



# *Innovación y Software*

*VOL 4 N° 2 Septiembre - Febrero 2023*

*ISSN: 2708-0935*

Revista de la Facultad de Ingeniería de ULASALLE



## Vol. 4 N° 2 2023 Septiembre - Febrero

ISSN: 2708-0935

DOI: 10.48168/innosoft.s12

ARK: ark:/42411/s12

PURL: 42411/s12

Depósito Legal: 2023-08884

Periodicidad: Semestral

Publicado: 30/09/2023

### Editado por:

Universidad La Salle

RUC: 20456344004

Ave. Alfonso Ugarte No. 517, Cercado, Arequipa, Perú

### COMITÉ EDITORIAL

Editor jefe:

Dr. Yasiel Pérez Vera

Editores asociados:

MSc. Anié Bermudez Peña

MSc. Percy Oscar Huertas Niquén

Miembros del Consejo Editorial

Dr. José Manuel Patricio Quintanilla Paulet

Hno. Jacobo Meza Rodríguez

Dr.C José Javier Zavala Fernández

Dr. Glenn Roberto Arce Larrea

Dr.C Álvaro Rodolfo Fernández del Carpio

MSc. Paul Mauricio Mendoza del Carpio

Corrección de estilos

MSc. Orlando Alonso Mazeyra Guillén

Maquetación

Audrey Aramburú Huacac, Madeleine Bedoya Gomez, Fabricio Centeno



## EDITORIAL

Prólogo Editorial

**p. 5**

## ARTÍCULOS ORIGINALES

Modelo de clasificación de depresión en Tweets usando BERT

**Autores:** Guillermo José Aleman Zambrano, Marvik Irzovic Del Carpio-Lazo, Daniel Gustavo Mendiguri-Chávez, Daniela Carolina Vilchez-Silva, Franco Tejada-Toledo.

**p. 6 - 24**

Cómo integrar Design Thinking en el Ciclo de Vida de Verificación y Validación de Software para Simuladores de Vuelo

**Autores:** Valeria Raquel Filoniuk, María Soledad Martínez, Ana Claudia Diz, Silvia Edith Arias.

**p. 25 - 35**

Clasificación de categorías de noticias usando BERT

**Autores:** Bradlhy Luis Machado Medina, César Alonso Santillana Quirita, Sharmelyn Violeta Bautista Luque.

**p. 36 - 51**

Data leakage analysis in wireless networks using Super-vised and Unsupervised Testing

**Autores:** Shahzad Ashraf, Zeeshan Aslam.

**p. 52 - 62**

El impacto del Design Thinking en la tecnología

**Autores:** Melvin Kevin López Asto, Jair Joel Vásquez Cerna, Alberto Carlos Mendoza de los Santos.

**p. 63 -78**

Reconocimiento y Clasificación de Mensajes de Odio

**Autores:** Patrick Leopoldo Paredes Neira, Gary Jamil Vilca Tapia, Kristhyan Andree Kurt Lazarte Zubia.

**p. 79 - 95**

Garantía de seguridad de la información empresarial a través de la gestión de servicios

**Autores:** Luz Maria Solano Quincho, Miriam Maryori Horna Maguiña, Alberto Carlos Mendoza de los Santos.

**p. 96 - 106**

Detección de Ideas Suicidas en Twitter

**Autores:** Juliana Berrios Butron, Daria López Franco, Dante Rodolfo Tárraga.

**p. 107 -120**

Aplicación de modelo de regresión lineal para predecir el índice de popularidad en la plataforma Spotify

**Autores:** Cesar Vasquez Alvarez, Edith M. Coaquira Cuevas, Emerson D. Mendoza Hilasaca, Jeffrey J. Pinto Ñaupá.

**p. 121 - 135**



Predicción de precios de autos usados

**Autores:** Yoset Cozco Mauri, Fiorella Pilar Cayo Cayo, Jordy Pedro Valencia Jara, Henry Uriel Bustinza Torres.

**p. 136 - 152**

Aplicación de árboles de decisión para un pronóstico salarial

**Autores:** Corina Hanco Vargás, David Flores Silva, Jessica Hanco Velásquez, Alejandra Fernandez Ninahuaman.

**p. 153 - 165**

Aplicación de árboles de decisión para la identificación de adaptabilidad de estudiantes en educación online

**Autores:** Luis Emanuel Araoz Valencia, Walter Huaracha Condori, Víctor Raúl Quispe Quicaña, Alex Ronaldo Turpo Coila.

**p. 166 - 181**

Seguridad de la información en la prevención de pérdida de datos: una revisión sistemática

**Autores:** Mishael Jeanpier Rojas Valiente, Jose Maria Hermes Castillo Sarmiento, Alberto Carlos Mendoza De Los Santos.

**p. 182 - 200**

El impacto de la Inteligencia Artificial en la mejora de la atención al cliente: Una revisión sistémica

**Autores:** López Zavaleta Verónica, Rojas Ahumada Kenner, Mendoza de los Santos, Alberto.

**p. 201 - 222**





Nos complace dar la bienvenida a este segundo número del cuarto Volumen de la Revista Innovación y Software de la Facultad de Ingeniería en la Universidad La Salle. En esta ocasión, hemos preparado una edición que destaca la excelencia en la investigación y la innovación en el campo de la tecnología y la ingeniería de software, presentando un conjunto diverso de 14 artículos que abarcan una amplia gama de temas de vanguardia en el ámbito tecnológico. Desde la aplicación de modelos avanzados de procesamiento de lenguaje natural para abordar cuestiones de salud mental, pasando por la exploración de la inteligencia artificial en la mejora de la atención al cliente y transformando así la toma de decisiones en diversos campos empresariales, hasta el estudio del diseño centrado en el usuario como influencia en el ciclo de vida del software.

La diversidad de temas abordados en estos artículos refleja la naturaleza en constante evolución de la tecnología; y nos propone como un foro para la difusión de investigaciones innovadoras que tienen un impacto significativo en la sociedad y en la forma en que interactuamos con la tecnología en nuestro día a día, pues a medida que avanzamos hacia un futuro donde esta misma continuará desempeñando un papel central en nuestras vidas, es crucial mantenernos al tanto de los desarrollos más recientes y prometedores en el campo.

Esperamos que esta edición sea una fuente de inspiración y conocimiento para nuestros lectores, y que les ayude a comprender mejor cómo la tecnología está dando forma a nuestro mundo. Así mismo agradecemos a todos nuestros autores por su dedicación y esfuerzo en la creación de estas valiosas contribuciones y al mismo tiempo agradecemos a nuestros revisores y al comité editorial por su arduo trabajo en la producción de esta edición.

Por último, nos complace invitarlos a explorar estos artículos y a unirse a nosotros en el emocionante viaje de la innovación tecnológica, pues estamos seguros de que lo encontrarán estimulante y enriquecedor.

**Comité Editorial**



# Modelo de clasificación de depresión en Tweets usando BERT

*Depression classification model on Twitter using BERT*

6

## **Guillermo José Aleman Zambrano**

Universidad La Salle. Arequipa, Perú.

@ galemanz@ulasalle.edu.pe

<https://orcid.org/0000-0001-5471-4226>

## **Marvik Irzovic Del Carpio Lazo**

Universidad La Salle. Arequipa, Perú.

@ mdelcarpiol@ulasalle.edu.pe

<https://orcid.org/0000-0002-0019-2458>

## **Daniel Gustavo Mendiguri Chávez**

Universidad La Salle. Arequipa, Perú.

@ dmendiguric@ulasalle.edu.pe

<https://orcid.org/0000-0002-0588-6520>

## **Daniela Carolina Vílchez Silva**

Universidad La Salle. Arequipa, Perú.

@ dvilchezs@ulasalle.edu.pe


<https://orcid.org/0000-0002-7896-8228>


## **Franco Tejada Toledo**


Universidad La Salle. Arequipa, Perú.

@ ftejadat@ulasalle.edu.pe

<https://orcid.org/0000-0002-1675-7097>

 **ARK:** <ark:/42411/s12/a89>

 **DOI:** [10.48168/innosoft.s12.a89](https://doi.org/10.48168/innosoft.s12.a89)

 **PURL:** [42411/s12/a89](https://purl.org/42411/s12/a89)

RECIBIDO 08/03/2023 • ACEPTADO 15/06/2023 • PUBLICADO 30/09/2023



## RESUMEN

Hoy en día existen muchos indicios de depresión, así como muchos intentos de suicidio causados por este trastorno emocional, esto se ve reflejado mayormente en redes sociales principalmente en Twitter. Por ello, es importante que los especialistas y organizaciones que busquen salvaguardar la vida de las personas, utilicen herramientas de software que permitan abordar este problema. Para ello, en este trabajo se propone una herramienta web llamada "UBDevs-Depression-Classifier" que permite clasificar y obtener tweets de forma automática por algún tema específico. Se puso un mayor énfasis a tweets relacionados con el COVID-19 debido a que en los años 2020-2021 en el mundo se vivió una pandemia que incrementó los casos de depresión



en muchos lugares. Esta propuesta de investigación se centra en la utilización en un modelo basado en NLP (Natural Language Processing) para la clasificación de Tweets con el fin de encontrar aquellos que inciten a la depresión o den a entender que los usuarios se encuentren en un mal estado de ánimo, todo ello con el fin de mantener la salud mental y física de los usuarios de esta plataforma. Existen varios modelos usados como base para proyectos de NLP, sin embargo, en la actualidad BERT ha demostrado ser uno de los más eficientes por ello lo seleccionamos para el desarrollo de nuestra propuesta. Para evaluar la eficiencia del proyecto aplicamos la métrica F1 obteniendo un valor de 0.8806, resultado bastante aceptable respecto a una clasificación textual.

**Palabras claves:** Clasificación de depresión, clasificación de texto, procesamiento de lenguaje natural, BERT, redes sociales.

## ABSTRACT

*Today there are many signs of depression, as well as many suicide attempts caused by this emotional disorder, and this is reflected mostly on social networks, mainly on Twitter. For this reason, it is important for specialists and organizations seeking to safeguard people's lives to use software tools to address this problem. For this, in this work a web tool called "UBDevs-Depression-Classifier" is proposed, that allows you to automatically obtain and classify tweets for a specific topic. A greater emphasis was placed on tweets related to COVID-19 in the years 2020-2021 the world experienced a pandemic that increased cases of depression in many places. This research proposal focuses on the use of a model based on NLP (Natural Language Processing) for the classification of Tweets in order to find those that incite depression or imply that users are in a bad mood, all this in order to maintain the mental and physical health of the users of this platform. There are several models that are used as a basis for NLP projects, however, at present BERT has proven to be one of the most efficient, so we selected it for the development of our proposal. To evaluate the efficiency of the project we applied the F1 metric obtaining a value of 0.8806, a quite acceptable result with respect to a textual classification.*

**Keywords:** Depression classification, text classification, natural language processing, BERT, social networks.

---

## INTRODUCCIÓN

En la actualidad, debido al contexto social aislado, el estrés laboral y académico, la falta de recursos y estabilidad, entre otros factores han contribuido al aumento de depresión y sentimientos de soledad [1][2][3][4], forzando a muchos al suicidio y con mayor incidencia en jóvenes alrededor de los veinte años [5].



Por ello, es importante prevenir este tipo de casos a través de la detección temprana de personas con problemas de depresión, para que puedan ser asistidas por las autoridades competentes y salvar su vida, antes de formar parte del 60% al 70% de suicidios causados por depresión [6].

Analizar los sentimientos de la población a través del entendimiento de sus expresiones hace que este tema no sólo sea interesante, sino necesario en la lucha contra el problema global que significa una población depresiva, y el riesgo latente sobre la vida de aquellos que no encuentran ayuda a tiempo [7].

Bajo el contexto de la pandemia, el aislamiento social ha contribuido a que el uso de redes sociales incremente, donde se sabe que los jóvenes son quienes más lo utilizan, sobre todo con el fin de expresar pensamientos, comentarios, sentimientos y contenido [8].

Twitter es una de las redes sociales con mayor número de usuarios a nivel mundial [9], estos mayormente son jóvenes, los cuales encuentran en esta red social un lugar donde expresar ideas y sentimientos, que pueden ser tanto de alegría, enojo, tristeza e incluso depresión. Con tal fin se desea contribuir con la detección de personas emocionalmente vulnerables por problemas emocionales, especialmente con tendencias depresivas quienes se expresan por Twitter. Precisamente esta red social será utilizada para extraer tweets con su respectiva información y posteriormente clasificarlos.

Existen muchos modelos para realizar el procesamiento de lenguaje natural, uno de los más prometedores actualmente es el modelo propuesto por Google llamado BERT el cual analiza el texto de forma bidireccional lo que permite que sea muy eficiente [19]. Una característica importante es que para lograr que este modelo logre un objetivo específico primero tiene que pasar por una fase de entrenamiento con un dataset. Así mismo para que BERT pueda entender el texto primero debemos realizar una tokenización y codificación de las palabras [14].

Con el fin de que la propuesta de este trabajo tenga buenos resultados se realizó una limpieza previa de datos irrelevantes en los tweets antes del proceso de tokenización y codificación.

Los resultados obtenidos podrán ser visualizados mediante una herramienta web la cual presentará gráficos estadísticos de tipo Pie Chart, mostrando que cantidad de tweets han sido clasificados como depresivos y no depresivos. Estos gráficos lograrán que los usuarios comprendan mejor los valores estadísticos obtenidos.

La información generada en internet día a día es inconmensurable por la cantidad de usuarios que registran datos de diferentes temas, en distintas plataformas alrededor del mundo. Kauffman habla de cómo la analítica del big data se está convirtiendo de forma acelerada en una práctica de tendencia adoptada por muchas organizaciones con el objetivo de obtener información valiosa [11].





En tal sentido, hablamos del poder de la data existente en internet y continuamente renovada por los usuarios, para que organizaciones privadas o de interés público puedan tomar decisiones sobre la población que le compete o sobre la que es responsable.

Un gran inconveniente que tuvo Kauffman es que al trabajar con distintos tipos de comentarios y calificaciones enfocados a analizar algo en concreto, obtuvo también otros que están fuera de lo que se desea analizar. Por ello Kauffman determinó que a la calificación se pueden sumar varios aspectos que pueden llegar a ser importantes para su investigación.

La propuesta de Back puede extraer información sobre los sentimientos en grandes cantidades de datos generados por usuarios (big data no estructurado de SNS) utilizando Naive Bayes y NLP, en su estudio realizan múltiples pruebas donde cada uno de estos métodos muestra su desempeño; llegan a la conclusión que NLP tiene un menor rango de error pero tiende a demorar 5 veces más que el algoritmo de Bayes pero la precisión de Bayes es 63.5% , comparando con la precisión de NLP que es 72.28% entonces concluyen que si el objetivo es un sistema que procese una gran cantidad de datos la mejor opción es Bayes, pero si nuestro objetivo es tener un análisis confiable lo mejor es utilizar NLP [12].

Ya que este documento intenta enfocarse en la clasificación de texto mediante el uso de NLP es la mejor opción para poder obtener la mayor fiabilidad sobre la data estudiada.

Sobrino explica dos clases de métodos para la clasificación de textos enfocados en NLP:

#### A. Métodos basados en diccionarios

Para este método se hace uso de listados de palabras y frases ya etiquetadas con la polaridad del sentimiento que expresan, así como la intensidad o fuerza de dicho sentimiento. Luego los textos a ser clasificados se dividen y se busca en los diccionarios para obtener su respectivo resultado.

Ejemplo de este se da en el estudio de Taboada, donde haciendo uso de un diccionario previamente clasificado con polaridades y sus respectivas puntuaciones, luego el sistema extrae las palabras que tienen información de sentimiento mediante un análisis de adverbios, verbos y adjetivos, después se busca sus respectivas puntuaciones para así tener calcular el sentimiento global [10].

Algunas desventajas de este método es que una misma palabra puede contener diferentes polaridades lo cual puede llevar a resultados incorrectos ya que el contexto de la palabra dentro de la oración no está presente en este tipo de calificaciones. Otra desventaja es que se debe tener un gran diccionario con un inmenso contenido de palabras con calificaciones, el cual puede ser muchas veces difícil de construir y consultar. [10].



## B. Métodos basados en relaciones lingüísticas

En este método se busca opiniones y sentimientos con mayor probabilidad para ser usados luego en la categorización del texto global, obteniendo también la categoría gramatical de las palabras, siendo esta la característica más importante [10]. El estudio de Hatzivassiloglou [23] quién hace uso de este método, afirma que dependiendo de los conectores usados entre dos palabras, dice si habrá polaridades opuestas o iguales.

Por ejemplo, si dos palabras tienen el conector “y”, esta tendrá la misma polaridad, por otro lado, con “pero” ocurrirá lo contrario, para ahorrar así análisis extra, luego con estos resultados se calculará el sentimiento global [10].

Luego de presentar los métodos explicados por Sobrino, se buscó enfocar el documento hacia el análisis de sentimientos. El estudio de Leis habla sobre la identificación de características lingüísticas en tweets en español y patrones de comportamiento de los usuarios de Twitter que podría sugerir signos de depresión, lo cual firma un precedente sobre el análisis de tweets en español [13].

En [13] indican que los trastornos mentales son una preocupación importante en la salud pública y son una de las principales causas de morbilidad en el mundo. Las plataformas de redes sociales nos permiten observar las interacciones, pensamientos y sentimientos de la vida diaria de las personas, incluidas las de los pacientes que padecen trastornos mentales. Existen estudios que han analizado la influencia de los trastornos mentales, incluida la depresión, en el comportamiento de los usuarios de las redes sociales, básicamente se han centrado en mensajes en inglés.

El estudio se centró en detectar características especiales acerca de los tweets depresivos en español. Esto se logró gracias al etiquetado de parte del discurso, Part-of-Speech (POS) realizado mediante la herramienta Freeling Natural Language Processing para analizar los patrones de uso de categorías gramaticales (p. Ej., Adjetivos, sustantivos o pronombres) en el texto de los tweets. Así mismo se contó de la confirmación y evaluación de especialistas de la salud para un análisis más detallado de los resultados de las frecuencias de parte del habla (POS) en tweets de control, usuarios depresivos y conjuntos de datos de tweets depresivos.

El estudio llegó a la conclusión que los usuarios de Twitter que sufren depresión modifican las características generales de su idioma y la forma en que interactúan en las redes sociales. En base a estos cambios, estos usuarios pueden ser monitoreados y apoyados, lo que presenta nuevas oportunidades para estudiar la depresión y brindar servicios de atención médica a las personas con este trastorno mental.

Adicionalmente a la intención de analizar el texto de las publicaciones, existe la necesidad de limpiar la información receptada. Puesto que todo lo que está en internet no es precisamente



concreto llegando a tornarse muchas veces ambiguo y puede afectar la data recolectada y estudiada con falsas alarmas Kauffmann plantea seis clases de problemas al momento de validar la información que se encuentra en las redes sociales [11]:

- En dominios particulares, una palabra u oración puede tener un significado opuesto.
- Una oración interrogativa o una oración condicional puede no tener un sentimiento positivo o negativo, pero una palabra en particular puede tenerlo.
- Las frases sarcásticas pueden tener el sentimiento opuesto.
- Algunas oraciones pueden tener información de sentimiento, pero no usan palabras de sentimiento.
- Una sola palabra puede cambiar la polarización del sentimiento en dos oraciones similares, así como el hecho de que, para una persona diferente, una oración puede tener un sentimiento diferente.
- Los problemas del lenguaje natural cambian de lugar a lugar.

Existe un modelo llamado BERT mayormente utilizado para tareas como la inferencia del lenguaje y la respuesta de preguntas y ha demostrado tener muy buenos resultados gracias a la forma como procesa el texto. Por ello es muy popular y reconocido como técnica revolucionaria que da inicio a una nueva era en el procesamiento de lenguaje natural (NLP). Este es propuesto por Google como un modelo pre entrenado con un conjunto de información gigantesca para la comprensión del lenguaje natural.

El acrónimo BERT significa "Bidirectional Encoder Representations from Transformers" por ende la forma en la que analiza las oraciones es bidireccional para una mejor comprensión de texto, hace uso de una arquitectura basada en "Transformers" que consta de una codificación y una decodificación [19].

En Pandey, S. hacen el análisis de sentimientos sobre la red social de Twitter para poder hacer un sondeo de cuál fue el impacto en la vida social por el efecto de coronavirus, para el proyecto se utilizó BERT para el análisis que se realiza en la sección 4 de dicho documento, este análisis también toma en consideración los likes que son recibidos junto con los retweets, se puede presumir que entre una mayor de cantidad de likes los tweets tienen una mayor acogida [14].

Potena, D. tiene un trabajo parecido al de Pandey, S. con la diferencia de que a pesar de trabajar también con tweets estos aún conservan emoticones, los cuales son utilizados para poder tener una mayor polarización en el texto y ser mejor clasificado [14][15].

El trabajo de Pota es muy parecido al creado por Potena, D. pero Pota logró llegar un poco más allá logrando poder trabajar también no sólo el lenguaje del inglés sino que logró también el trabajar con italiano [15][16].



Así podemos darnos cuenta que BERT es una poderosa herramienta que no solo nos permite el entendimiento de texto en inglés, sino que también puede ser utilizado para varios lenguajes.

## **Materiales y métodos o Metodología computacional**

Para la propuesta de este trabajo, se busca desarrollar una herramienta que permita clasificar una serie de tweets como depresivos o no depresivos. Todos los resultados obtenidos podrán ser visualizados a través de una interfaz web por medio de un gráfico Pie Chart.

Se hace uso del modelo BERT, tomándolo como base para el procesamiento de texto. Se añade una red neuronal lineal, la cual le permite que se especialice en una tarea en particular que en nuestro caso, es el análisis de tweets para reconocer cuando es depresivo o no lo es. Luego este modelo junto con la red neuronal deben pasar a un entrenamiento para alcanzar el objetivo de clasificación específica.

Para lograr el objetivo, primero es necesario conseguir los recursos para este entrenamiento. Se utiliza tablas procedentes del sitio web de Kaggle [21], [22]. Este conjunto de datos está estructurado de la siguiente manera: Tweet y su respectiva clasificación (depresivo o no depresivo).

Los tweets incluyen emoticones, lo cuales dificultan la interpretación del texto, elevando la ambigüedad, por ello se suprimieron antes de pasar al tratamiento del texto de los tweets procedentes de las tablas. Este tratamiento consiste en dividir el texto en tokens (tokenización) que luego son codificados por BERT Tokenizer. Estas etapas previas son necesarias para que el modelo BERT pueda entender el texto.

Una vez terminado el entrenamiento, la herramienta estará lista para clasificar tweets. El uso de esta herramienta web será por medio de una interfaz gráfica en el que el usuario podrá buscar tweets por un tema específico para luego ser clasificados. Los resultados serán entregados finalmente en el gráfico Pie Chart el cual indica los porcentajes de cada clase de tweets (depresivo y no depresivo).





## A. Pipeline

El diseño de la propuesta puede verse reflejado en el siguiente pipeline:

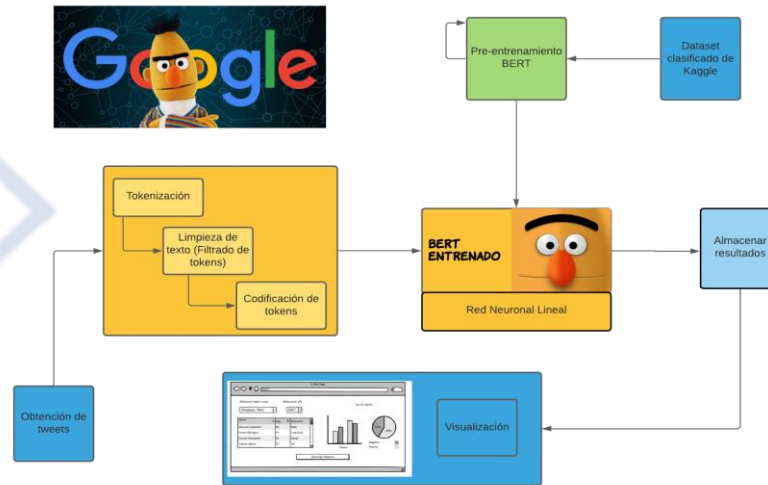


Figura 1. Pipeline descriptivo de la propuesta.

La entrada principal es la data de Tweets que se extrajeron de la API Twitter, la cual será analizada. Luego pasa por un proceso de tokenización tarea que es realizada el mismo modelo de BERT, segundo una limpieza de texto la cual consiste en eliminar los tokens que no aportan en el análisis como símbolos, menciones (no propias al lenguaje común), por último una codificación de tokens.

En el modelo se hará uso de una red neuronal lineal para el proceso de pre entrenamiento y entrenamiento. Como entrada para el modelo se tiene un dataset ya clasificado por dos clases: depresivo no depresivo, la cual servirá como input para el entrenamiento.

Como salida del modelo de BERT se obtendrán los resultados de clasificación de tweets los cuales pasarán a ser procesados con el fin que tenga una correcta estructura que sirva para una futura visualización de los datos. Luego pasarán a ser almacenados y por último los resultados serán presentados mediante una interfaz gráfica para representar de manera visual, así como estadísticas para tener un mejor entendimiento de la información.

## B. Datos

La data a analizar será extraída del API Tweepy, y tendrá la siguiente estructura:

Tabla 1. Data a analizar

Atributos	Descripción
tweet_id	Identificador del tweet



Atributos	Descripción
tweet	Cuerpo del mensaje del tweet
classification	Valoración: depresivo, no depresivo
twitter_username	Nombre de usuario
twitter_user_location	Ubicación del usuario
hashtags	Son los hashtags que se encuentran en el Tweet

En cuanto al dataset para el entrenamiento del modelo base se tendrá un archivo CSV con 116933 tweets y se tendrá los siguientes parámetros: Texto, Sentimiento.

### C. Principios de diseño

Con la intención de desarrollar un software de calidad, se planteó el diseño basado en dos clases de principios:

- Kiss ( "Keep it simple, Stupid!")

Se propone hacer el diseño lo más simple posible, de forma tal que la complejidad no sea tanta como para evitar que el proyecto no logre reflejar sus objetivos por perderse en detalles innecesarios; ni sea tan simple como para desmerecer la calidad del servicio planteado [17].

- Solid  
Representa cinco principios básicos de la programación orientada a objetos:

- S – Single Responsibility Principle (SRP)
- O – Open/Closed Principle (OCP)
- L – Liskov Substitution Principle (LSP)
- I – Interface Segregation Principle (ISP)
- D – Dependency Inversion Principle (DIP)

En cuanto a la etapa de desarrollo, se aplican los principios SOLID. Se busca mantener responsabilidades únicas sobre el código planteado, de modo tal que no sean necesarias



de modificar, a menos que estrictamente deban de hacerlo debido a que la función real (no la abstraída) cambie su concepto.

Las clases programadas, estarán dispuestas a extenderse para lograr mejorar, y no ser modificadas con otro fin que no sea la mejora de las mismas.

Aunque el proyecto es corto, en caso sea necesario extenderlo a través de subclases, estas no deberán evitar la funcionalidad haciéndose exclusivas, sino que se deberá poder seguir operando bajo las clases base, primeramente, planteadas para dicho fin.

El diseño se está planteando con secciones que realizan una tarea específica, cada parte del programa se encarga de un contexto único, y las funcionalidades diferentes son parte de otra clase. Por ejemplo: tokenización, limpieza y análisis [18].

## D. Propuesta UI

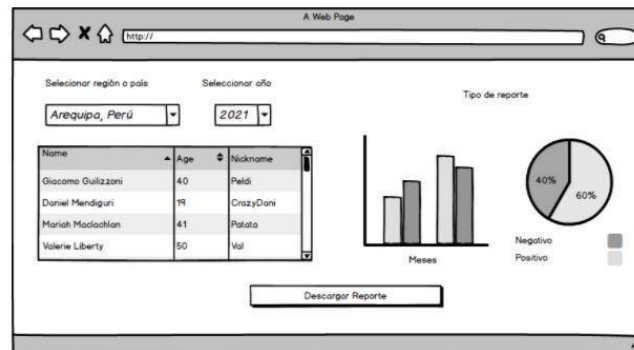


Figura 2. Prototipo de la interfaz de usuario.

Para la propuesta UI se desarrollará una aplicación web, en la cual se podrán visualizar los resultados del análisis de tweets por medio de gráficos estadísticos (Pie Chart). La extracción de tweets se hará de forma automática cuando el usuario busque un tema en concreto, además se mostrará una leyenda para reconocer las etiquetas de los tweets clasificados en depresivo o no depresivo.

## E. Implementación

Para la propuesta se siguieron los pasos descritos en el algoritmo de la siguiente subsección, que estará implementado en Python. Se dió inicio a obtener el conjunto de datos de Tweets relacionados con un contexto especificado por un especialista interesado en obtener información y alertas sobre casos de depresión. Esto se logrará gracias al API Tweepy la que se encargará de proporcionar todos los tweets con su respectiva información.



La implementación se realizó en un entorno virtual de Google Colab con las siguientes características: GPU del backend de Google Compute Engine en python 3, RAM 12.69 GB, Disco 68.35 GB. Para el entrenamiento de BERT para el reconocimiento de sentimiento se tardó 2 horas 21 minutos y 23 segundos utilizando la GPU del tipo CUDA proporcionado por Google Colab.

Con respecto a la evaluación del tweet, en primer lugar, es necesario un preprocesamiento de los tweets para convertirlo en tokens entendibles para el modelo entrenado BERT, así mismo una limpieza y estandarización del texto. Luego el algoritmo procede a clasificar cada tweet analizado si es depresivo o no depresivo.

Para el entrenamiento del modelo base de BERT junto a la red neuronal lineal, se hará uso de un dataset con ejemplos de tweets, frases y citas ya clasificados. BERT se encargará de la compresión de texto mientras que la red neuronal lineal se encargará de determinar si un tweet es o no depresivo.

## Algoritmo 1: Obtención de Tweets

**Input:** TOPIC => Contexto, N\_TWEETS\_REQUESTED  
=> número de tweets solicitados  
**Output:** TWEETS => Tweets de la API tweepy  
TWEETS = []  
**for** i = 1 ... N\_TWEETS\_REQUESTED  
=> número de tweets solicitados  
**return** TWEETS

## Algoritmo 2: Entrenamiento BERT

**Input:** DATA\_SET => Grupo de datos clasificados, MODEL => Modelo pre entrenado BERT, DATA\_TEST => Grupo de datos clasificados.  
**Output:** PORCENTAJE\_DE\_EFICIENCIA => Porcentaje que indica que tan efectivo es el modelo clasificando tweets en el entrenamiento.  
DATA\_LOADER = []  
**for** i = 1 DATA\_SET.tamaño  
DATA\_LOADER.añadir (tokenización (DATA\_SET[1].texto), DATASET[1].clasificación)  
MODEL.train (DATA\_LOADER)  
DATA\_LOADER\_TEST = []  
**for** i = 1 DATA\_TEST.tamaño





```
DATA_LOADER_TEST.añadir (tokenización (DATA_SET[1].texto),  
DATA_TEST[1].clasificación)  
PORCENTAJE_DE_EFICACIA = MODEL.evaluación  
return PORCENTAJE_DE_EFICACIA
```

### Algoritmo 3: Clasificación de tweets

**Input:** TWEETS => Tweets, Tweets de la API tweepy

**Output:** CLASIFICACIONES => Si es o no depresivo.

```
TWEETS_TOKENIZADOS = []
```

```
TOKENIZADO = []
```

```
LIMPIO = []
```

```
CODIFICADO = []
```

```
for i = 1 TWEETS.tamaño
```

```
TOKENIZADO = tokenización (TWEETS [i][Text])
```

```
LIMPIO = limpieza (tokenizado)
```

```
CODIFICADO = codificación (LIMPIO)
```

```
TWEETS_TOKENIZADOS.añadir (CODIFICADO)
```

```
CLASIFICACIONES = []
```

```
for i = 1 TWEETS.tamaño
```

```
CLASIFICACIONES.añadir(MODEL.clasificartweet((TWEETS_TOKENIZADOS [ i ]))
```

```
return CLASIFICACIONES
```

## Resultados y discusión

### A. Resultados de la implementación del modelo de BERT

Se obtuvieron los siguientes resultados de efectividad durante el entrenamiento de BERT. El entrenamiento duró 2 horas con 21 minutos y 23 segundos con una cantidad de 57732 tweets clasificados. Para ambas clases en el dataset (depresivo y no depresivo) se ha considerado la misma cantidad de tweets.

Tabla 2. Resultados del entrenamiento de BERT

Época	Entrenamiento		Validación	
	Loss	Accuracy	Loss	Accuracy
1	0.02922	0.99214	0.01582	0.99740
2	0.00869	0.99840	0.01578	0.99801
3	0.00544	0.99911	0.01257	0.99827
4	0.00268	0.99959	0.01311	0.99835
5	0.00104	0.99983	0.01594	0.99809



Como se puede observar en la tabla 2, se obtuvo luego de 5 épocas del entrenamiento de BERT un Accuracy de 0.99983 y un 0.99809 en la validación.

Estos son algunos ejemplos que se utilizaron para la prueba de clasificación de tweets.

Tabla 3. Ejemplos de tweets y sus resultados de clasificación.

<b>Tweet</b>	<b>Clasificación</b>
Up early to watch the GP today. Dodgy hamstring is still dodgy, so I wouldn't have been able to run this morning anyway	No depresivo.
is very sad because she can't be with her Mom during this sad time.	Depresivo
Suicide is the best way to get out of any situation...	Depresivo
@shaundiviney I want pizzaaaa.	No depresivo

## B. Métricas de exactitud

Se evaluó el rendimiento de los algoritmos Machine Learning de clases múltiples mediante el uso de the 10-Fold Cross-Validation strategy [20].

Para ello se utilizó la evaluación F1 por medio de la librería de python llamada sklearn.

La medida de evaluación F1 corresponde a la media armónica de Precisión (Ecuación 1) y Recuperación (Ecuación 2), donde T P (Verdadero Positivo) se refiere al número de documentos de una clase en la que el algoritmo clasificó correctamente; FP (Falso Positivo) se refiere al número de documentos que no pertenecen a una clase en la que el algoritmo clasificó erróneamente como pertenecientes; y F N (Falso Negativo) se refiere al número de documentos



de una clase en los que el algoritmo clasificó erróneamente como otra clase. La ecuación 3 define la medida F1.

$$P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN}, F_1 = \frac{TP}{TP+FN} \quad (1)$$

Aplicando la evaluación F1 con el dataset de tweets se obtuvo como resultado un valor de 0.8806.

### C. Interfaz gráfica

Para una mejor visualización de datos y usabilidad para el usuario, se plantea el uso de herramientas gráficas fáciles de comprender.

Se desarrolló una herramienta web gracias a los frameworks Vue.js y Flask para poder obtener de manera automática los tweets de un tema específico usando la librería Tweepy.

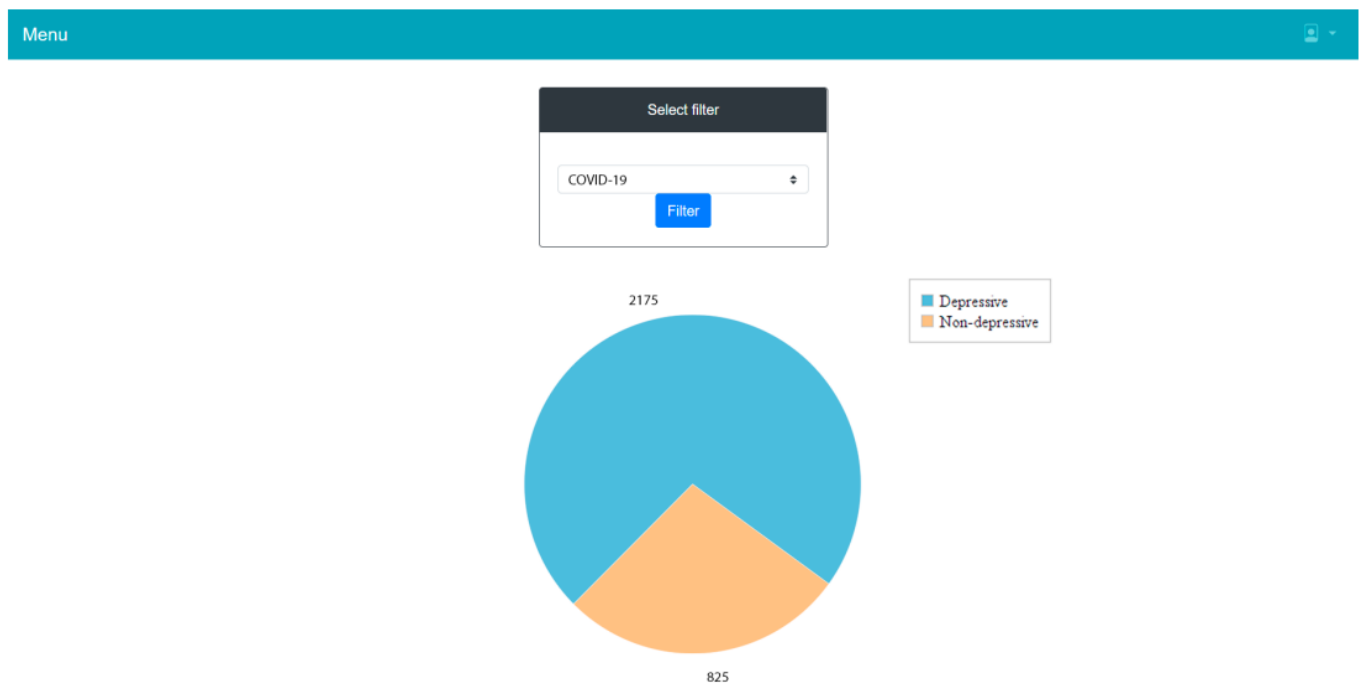


Figura 3. Interfaz de búsqueda de tweets por tema.

Así mismo mediante la herramienta web construida es permisible obtener un gráfico del tipo Pie chart para la visualización estadística por clases.



Esta herramienta permite filtrar los resultados de la búsqueda antes realizada. Así mismo muestra una leyenda por colores para identificar las clases de tweets en el gráfico, siendo estas: depresivo en color celeste y no depresivo en color anaranjado.

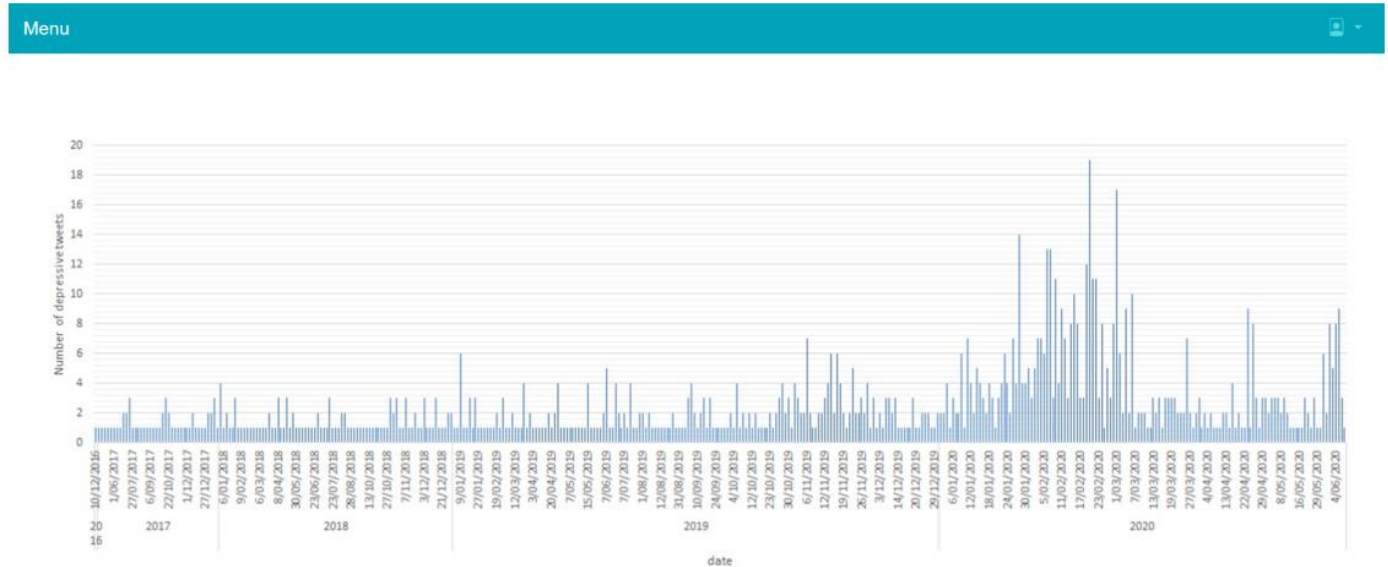


Figura 4. Histograma generado por la herramienta.

## Conclusiones

A pesar de existir varios modelos para el análisis, es importante considerar la eficiencia de los resultados y la calidad que se busca ofrecer, así que para el proyecto se utilizó una técnica de clasificación de textos bastante eficiente como es en la actualidad BERT, logrando un valor en la métrica F1 de 0.8806 porcentaje superior a otras técnicas que existen en el estado del arte.

La aplicación del Procesamiento del Lenguaje Natural por medio del uso de distintas técnicas a lo largo de los años se ha ido refinando, siendo la inteligencia artificial la herramienta más valorada para NLP en la actualidad, la cual se refleja en el modelo BERT, que demuestra ser mucho más eficiente que técnicas usadas previamente con simple reconocimiento de texto sin considerar el contexto.

Se debe recordar que la evaluación de sentimientos de esta propuesta es solo para tweets en inglés.

En el presente trabajo, fue desarrollada una herramienta web llamada "UBDevs-Depression-Classifer" la cual permite a especialistas, organizaciones gubernamentales o privadas con





implicancias o relaciones sobre la salud de la población, identificar problemas de depresión en la red social Twitter. Así mismo, obtener de forma temprana alertas sobre problemas de depresión, muchas veces desencadenantes y/o agravantes de otros problemas de salud física.

Se realizó mayor énfasis en los tweets de COVID-19 debido a que los casos de depresión se incrementaron en la pandemia que atraviesa el mundo en los años 2020-2021.

Como trabajo futuro se espera mejorar aún más el porcentaje de la métrica F1 ampliando el dataset de entrenamiento y de prueba. Además, se espera mejorar esta herramienta web considerando una línea de tiempo que permita a los especialistas observar la variación de índices de depresión respecto a una fecha o periodo determinado.

Así mismo, se podría incluir una mayor cantidad de filtros para la vista de gráficos, como por ejemplo un filtro por localización. Gracias a ello será más útil en el estudio de los tweets y podrán orientar su uso a labores de prevención de la depresión sobre determinadas regiones, ciudades o países. Como trabajo futuro, se espera desarrollar o ampliar la funcionalidad de la herramienta para incluir la identificación de palabras más importantes y frecuentes en los tweets depresivos, si es posible también aumentar el idioma español para la herramienta.

## Contribución de Autoría

**Guillermo José Aleman-Zambrano:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Marvik Irzovic Del Carpio-Lazo:** [Conceptualización](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#). **Daniel Gustavo Mendiguri-Chávez:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Daniela Carolina Vilchez-Silva:** [Visualización](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#). **Franco Tejada-Toledo:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Supervisión](#).

## Referencias

- [1] Chen, F., Zheng, D., Liu, J., Gong, Y., Guan, Z., & Lou, D. (2020). Depression and anxiety among adolescents during COVID-19: A cross-sectional study. *Brain, behavior, and immunity*, 88, 36.



- [2] Islam, M. A., Barna, S. D., Raihan, H., Khan, M. N. A., & Hossain, M. T. (2020). Depression and anxiety among university students during the COVID-19 pandemic in Bangladesh: A web-based cross-sectional survey. *PloS one*, 15(8), e0238162.
- [3] Lee, S. A., Jobe, M. C., Mathis, A. A., & Gibbons, J. A. (2020). Incremental validity of coronaphobia: Coronavirus anxiety explains depression, generalized anxiety, and death anxiety. *Journal of anxiety disorders*, 74, 102268.
- [4] Santini, Z. I., Jose, P. E., Cornwell, E. Y., Koyanagi, A., Nielsen, L., Hinrichsen, C., ... & Koushede, V. (2020). Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *The Lancet Public Health*, 5(1), e62-e70.
- [5] Bhuiyan, A. I., Sakib, N., Pakpour, A. H., Griffiths, M. D., & Mamun, M. A. (2020). COVID-19-related suicides in Bangladesh due to lockdown and economic factors: case study evidence from media reports. *International Journal of Mental Health and Addiction*, 1-6.
- [6] Reddy MS. Depression - the global crisis. *Indian J Psychol Med* 2012;34:201-3.
- [7] World Health Organization. (2014). Preventing suicide: A global imperative. World Health Organization.
- [8] Charoensukmongkol, P. (2018). The impact of social media on social comparison and envy in teenagers: The moderating role of the parent comparing children and in-group competition among friends. *Journal of Child and Family Studies*, 27(1), 69-79.
- [9] Anger, I., & Kittl, C. (2011, September). Measuring influence on Twitter. In *Proceedings of the 11th international conference on knowledge management and knowledge technologies* (pp. 1-4).
- [10] Sobrino Sande, J. C. (2018) Análisis de sentimientos en Twitter.
- [11] Kauffmann, E., Peral, J., Gil, D., Ferrández, A., Sellers, R., & Mora, H. (2020). A framework for big data analytics in commercial social networks: A case study on



- sentiment analysis and fake review detection for marketing decision-making. *Industrial Marketing Management*, 90, 523-537.
- [12] Back, B. H., & Ha, I. K. (2019). Comparison of sentiment analysis from large Twitter datasets by Naïve Bayes and natural language processing methods. *Journal of information and communication convergence engineering*, 17(4), 239-245.
- [13] Leis, A., Ronzano, F., Mayer, M. A., Furlong, L. I., & Sanz, F. (2019). Detecting signs of depression in tweets in Spanish: behavioral and linguistic analysis. *Journal of medical Internet research*, 21(6), e14199.
- [14] Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Social Network Analysis and Mining*, 11(1), 1-11.
- [15] Chiorrini, A., Diamantini, C., Mircoli, A., & Potena, D. (2021). Emotion and sentiment analysis of tweets using BERT. In *EDBT/ICDT Workshops*.
- [16] Pota, M., Ventura, M., Fujita, H., & Esposito, M. (2021). Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets. *Expert Systems with Applications*, 181, 115119.
- [17] Qadeer, S., & Wu, D. (2004). KISS: keep it simple and sequential. *ACM sigplan notices*, 39(6), 14-24.
- [18] Rotge, J. F. (2000, September). SGDL-Scheme: a high level algorithmic language for projective solid modeling programming. In *Proceedings of the Scheme and Functional Programming 2000 Workshop (Montreal, Canada (pp. 31-34))*.
- [19] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [20] Tan, P., Steinbach, M., and Kumar, V. (2013). *Introduction to Data Mining: Pearson NewInternational Edition*. Pearson Education Limited.
- [21] Virahonda, S. (2021). Depressive and Anxious Tweets. Disponible en: <https://www.kaggle.com/datasets/sergiovirahonda/depression-anxiety-tweets>



- [22] Kazanova, M. (2018). Sentiment140 dataset with 1.6 million tweets. Disponible en: <https://www.kaggle.com/datasets/kazanov/sentiment140>
- [23] Hatzivassiloglou, V., & McKeown, K. R. (1997, July). Predicting the semantic orientation of adjectives. In Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics (pp. 174-181). Association for Computational Linguistics



# Cómo integrar Design Thinking en el Ciclo de Vida de Verificación y Validación de Software para Simuladores de Vuelo

25

*How to integrate Design Thinking in the Verification and Validation of Software's Life Cycle for Flight Simulators*

## Valeria Raquel Filoniuk

Dirección de Análisis Operativo. Fuerza Aérea Argentina, Córdoba, Argentina.

@vfiloniuk@gmail.com

<https://orcid.org/0000-0003-0614-3814>

## María Soledad Martínez

Dirección de Análisis Operativo. Fuerza Aérea Argentina, Córdoba, Argentina.

@mariasolemartinez81@gmail.com

<https://orcid.org/0000-0003-2346-9859>

## Ana Claudia Diz

Dirección de Análisis Operativo. Fuerza Aérea Argentina, Córdoba, Argentina.

@anaclaudiadiz@gmail.com


<https://orcid.org/0000-0002-0585-860X>


## Silvia Edith Arias


Universidad Nacional de Córdoba, Universidad Tecnológica Nacional, Córdoba, Argentina.

@edith.edit@gmail.com

<https://orcid.org/0000-0001-9695-2812>

 **ARK:** <ark:/42411/s12/a93>

 **DOI:** [10.48168/innosoft.s12.a93](https://doi.org/10.48168/innosoft.s12.a93)

 **PURL:** [42411/s12/a93](https://purl.org/42411/s12/a93)

RECIBIDO 15/03/2023 • ACEPTADO 20/06/2023 • PUBLICADO 30/09/2023



## RESUMEN

Este artículo presenta la aplicación de la Metodología Design Thinking para dar una solución creativa e innovadora en el Ciclo de Vida de Verificación y Validación de Software en Simuladores de Vuelo, de tal manera de optimizar su funcionalidad. La utilización de este método en las primeras etapas de dicho proceso permite una interacción activa y temprana con el usuario, aumentando la generación de opciones de solución y la identificación de necesidades, consiguiendo contribuir a la calidad del producto final. Es importante asegurar la calidad en el desarrollo de software de simuladores de vuelos, debido a su funcionamiento crucial y a la esencialidad que aportan en la defensa aérea de nuestro país.

Para mejorar el proceso mencionado se trabaja principalmente con tres escenarios de Design Thinking: Ideación, Inspiración e Implementación. El objetivo principal de este artículo es integrar estos escenarios en el Ciclo de Vida de Verificación y Validación de manera que ayuden a desarrollar una nueva forma de testear software, incluyendo al usuario como fuente estratégica. Se espera que los resultados contribuyan a crear productos de calidad, ya que los mismos





representan un activo esencial para la formación de pilotos que asumen el desafío de custodiar el cielo argentino.

**Palabras claves:** Calidad, Design Thinking, Simulador de Vuelo, Usuario, Verificación y Validación de Software.

## ABSTRACT

*In order to ensure the Flight Simulator's functionality, this article exposes the application of the "Design Thinking" Methodology. This will lead to provide more creative and innovative solutions during the Verification and Validation process in the Flight Simulator Software's Life Cycle. The application of this method in the first stages of the process allows an active and early interaction with user's needs, increasing the building of solutions and the identification of needs, heighten the final product's standards. It is important to ensure quality in the software development of flight's simulators, due to its application is crucial and essential to the air defense of our country.*

*So as to improve the aforementioned process, Design Thinking works with three main scenarios: Ideation, Inspiration and Implementation. The main objective of this article is to integrate these scenarios with the Verification and Validation of the Software's Life Cycle and develop a new way to test software, including the user as a strategic source. It is expected that the results will contribute to create high quality products being these vital assets for the pilot's training process who daily assume the challenge of guarding the Argentine sky.*

**Keywords:** Design Thinking, Flight Simulator, Quality, User, Verification and Validation of Software.

## INTRODUCCIÓN

Existen una gran variedad de investigaciones que han sugerido que Design Thinking puede tener un impacto positivo en el desarrollo de software, ya que facilita una profunda comprensión de las necesidades del usuario y aumenta la colaboración del equipo [1].

Los proyectos de desarrollo de software se ven afectados debido a factores que impactan en el resultado de los mismos. Estos factores incluyen requerimientos incompletos, demoras, incapacidad para empatizar e incorporar al usuario final y diseñar soluciones que respondan a sus necesidades reales. Todo ello contribuye a incrementar el costo, tiempo, y a disminuir la calidad del producto software [2].

Los pilotos de Fuerza Aérea Argentina utilizan simuladores de vuelo para su formación. Estas herramientas son empleadas para que tengan un entrenamiento previo por medio de software antes de realizar prácticas presenciales y de esta forma tener un mayor conocimiento del funcionamiento de la aeronave [3]. El software utilizado para entrenadores de vuelo representa



un activo esencial para la defensa aérea de nuestro país, por lo que se considera de gran criticidad en materia de calidad. En consecuencia, es inevitable incluir a dichos usuarios desde las primeras etapas del proceso de Verificación y Validación de Software, conociendo sus necesidades reales y priorizando sus requerimientos, lo que permitirá mantenerlos satisfechos. Adicionalmente se obtendrá un software de mejor calidad y más usable, obteniendo una disminución del tiempo invertido por el usuario en su capacitación, reducción de soporte y mantenimiento.

Una alternativa para impulsar el involucramiento del usuario es la Metodología Design Thinking [4], [5], un proceso para incentivar la creatividad e innovación, integrando factores humanos y tecnológicos para la resolución de problemas. Su enfoque centrado en lo humano permite descubrir perspectivas de las personas y usuarios finales [2].

Design Thinking posibilita generar soluciones que se basan en la idea de ponerse en el lugar del usuario final [6], es decir, para quien va dirigido el producto o servicio [7].

Este artículo pretende dar solución a problemas existentes en el proceso de Verificación y Validación de Software para Simuladores de Vuelo. El propósito principal es integrar el enfoque de Design Thinking en las fases de dicho proceso incluyendo al usuario como fuente estratégica y luego analizar los resultados obtenidos.

## Materiales y Métodos o Metodología Computacional

Para la presente investigación se revisaron artículos relacionados a técnicas existentes destinadas a integrar Design Thinking (DT) en etapas tempranas del Ciclo de Vida de Verificación y Validación de Software. En las siguientes secciones se muestran conceptos claves relacionados a DT y Verificación y Validación de Software (V&V), y luego se menciona la metodología a aplicar basada en la unificación de ambos conceptos, adaptándola a las necesidades específicas en la organización objeto de estudio.

### Design Thinking

*Design Thinking* permite concebir proyectos informáticos para dar soluciones a los problemas del mundo real. Se centra en las personas, considerando multidisciplinariedad, colaboración y concreción de pensamientos y procesos, los cuales son caminos que llevan a soluciones innovadoras como base para la resolución de problemas, desde el punto de vista del usuario y con el apoyo de prototipos [8].

El proceso de DT se basa en tres pilares importantes [9], [10]:

*Inspiración*: consiste en identificar la necesidad y el problema para analizar soluciones. Incluye las siguientes etapas:

- Empatizar: comprender lo que el usuario necesita, en términos de lo que siente el usuario.



- Definir: establecer las necesidades del usuario en términos de un problema que se puede resolver.

*Ideación:* se basa en proponer una solución que satisfaga las necesidades del usuario.

*Implementación:* desarrollar la solución final, en la que intervienen las siguientes fases:

- Prototipo: crear una solución que cubra una selección de las características requeridas por el usuario, para que pueda experimentarlo.
- Prueba: utilizar la solución real, de manera que el usuario pueda proporcionar comentarios oportunos y el proceso pueda iterar [1].

### *Verificación y Validación de Software*

El *Testing de software* pertenece a una actividad o etapa del proceso de producción de software denominada *Verificación y Validación*. V&V es el nombre genérico dado a las actividades de comprobación que aseguran que el software respeta su especificación y satisface las necesidades de sus usuarios [11]. Si bien estos términos en su uso cotidiano pueden llegar a ser sinónimos, en Ingeniería de Software tienen significados diferentes y cada uno tiene una definición más o menos precisa. Por lo tanto, verificar consiste en comprobar que el software cumple los requerimientos funcionales y no funcionales que se le han especificado; y validar se basa en probar que el software hace lo que el usuario espera [12].

Para llevar a cabo el proceso de V&V, teniendo en cuenta la funcionalidad y complejidad con que operan los simuladores de vuelo, los especialistas los ejecutan en fases que se presentan a continuación:

Tabla 1. Fases del Ciclo de Vida de Verificación y Validación de Software

<i>Primer Fase: Análisis de Requerimientos</i>	<ul style="list-style-type: none"><li>✓ Analizar los requerimientos documentados.</li><li>✓ Comprender lo que los usuarios esperan que haga el software en base a los requerimientos documentados y priorizarlos.</li></ul>
<i>Segunda Fase: Planificación de Pruebas</i>	<ul style="list-style-type: none"><li>✓ Definir el alcance de las pruebas detallando niveles, tipos, estrategias y técnicas de prueba tomando como referencia los requerimientos documentados.</li></ul>



<i>Tercer Fase: Diseño y Desarrollo de Pruebas</i>	<ul style="list-style-type: none"><li>✓ Comprobar que los requerimientos documentados estén completos.</li><li>✓ Desarrollar casos de prueba en base a los requerimientos documentados.</li><li>✓ Probar todas las combinaciones posibles y priorizar cuál de ellas afecta más al producto software.</li></ul>
<i>Cuarta Fase: Ejecución de Pruebas</i>	<ul style="list-style-type: none"><li>✓ Documentar novedades y desviaciones del comportamiento del software sin participación del usuario.</li><li>✓ Transmitir novedades para que se solucionen.</li></ul>
<i>Quinta Fase: Evaluación de resultados</i>	<ul style="list-style-type: none"><li>✓ Validar si los requerimientos implementados cubren las expectativas de los usuarios.</li></ul>
<i>Sexta Fase: Cierre del Proceso</i>	<ul style="list-style-type: none"><li>✓ Reportar las novedades del proceso y la aceptación/rechazo del software testeado, sin intervención del usuario final.</li></ul>

La Tabla 1 muestra que los requerimientos son comprendidos y analizados una sola vez y para todo el ciclo y que solo se interactúa con los usuarios en la fase de Evaluación de Resultados. Durante el proceso, los usuarios son ajenos al ciclo de vida de V&V, hasta que el producto se prueba y se determina que hace lo que ellos esperan. Por consiguiente, se llega al final del proceso trabajando en una propuesta que, en determinadas circunstancias, no cumple las expectativas de los usuarios, repercutiendo en la calidad del producto final.

En tal sentido se infiere que podría usarse DT para mejorar el ciclo mencionado. Se ofrece en este marco una metodología ágil, centrada en herramientas innovadoras que se utilizarán a lo largo del proceso y en la interacción continua con los usuarios [13]. Por consiguiente, se implementarán técnicas basadas en la integración de factores humanos y tecnológicos para cada una de las fases enunciadas.

Design Thinking, por lo tanto, puede ser lo suficientemente flexible como para ser tomado como un marco general de desarrollo de soluciones que lo hace adecuado para ser utilizado en los procesos de desarrollo de software, que de hecho es un campo en el que se requiere un enfoque abierto, creativo y flexible para resolver problemas y encontrar oportunidades de mejora [1].

## Resultados y discusión

Utilizar la técnica de Design Thinking en el Ciclo de Vida de V&V, favorece una mejora considerable en la identificación de problemas y necesidades de usuarios, recopilación de nuevas perspectivas de solución, generación de opciones de solución y prototipos que permitan representar dichas ideas [14].





La Tabla 2 expone cómo la metodología citada [14], [15], se adecúa perfectamente a las fases del Ciclo de Vida del Software:

Tabla 2. Integración del Design Thinking en el Ciclo de Vida de Verificación y Validación de Software

<i>Fases del Ciclo de Vida del V&amp;V</i>	<i>Design Thinking</i>		
<b>Análisis de Requerimientos</b>	Empatizar	Observación	<ul style="list-style-type: none"><li>• Observar qué actividades llevan a cabo los usuarios y cuáles son las más prioritarias.</li></ul>
		Mirar y Escuchar	<ul style="list-style-type: none"><li>• Pedir a los usuarios que expliquen cómo realizan algunas tareas y que vayan vocalizando lo que pasa por su mente cuando estén trabajando.</li></ul>
		Encuestas	<ul style="list-style-type: none"><li>• Confeccionar encuestas para que los usuarios expresen sus principales necesidades.</li></ul>
	Definir	Reuniones	<ul style="list-style-type: none"><li>• Establecer los requerimientos a probar en base a los resultados obtenidos anteriormente.</li></ul>
<b>Planificación de Pruebas</b>	Empatizar	Inspección del Software	<ul style="list-style-type: none"><li>• Examinar el software a testear con la participación del usuario con el objetivo de identificar qué tipos, estrategias y técnicas de pruebas se ensayarán.</li></ul>
	Definir	Plan de Pruebas	<ul style="list-style-type: none"><li>• Concretar qué tipos, estrategias y técnicas de pruebas se ensayarán tomando en cuenta las necesidades planteadas por el usuario, a través de un documento de Plan de Pruebas.</li></ul>





<b>Diseño y Desarrollo de Pruebas</b>	Idear	Escenarios	<ul style="list-style-type: none"><li>• Demostrar que los requerimientos a probar reflejen lo que el usuario espera por medio de la representación de Casos de Uso.</li></ul>
	Prototipar	Casos de Pruebas	<ul style="list-style-type: none"><li>• Diseñar casos de pruebas, priorizando aquellos que afecten en mayor medida los requisitos del usuario final.</li><li>• Establecer feedback con el usuario en la resolución de casos de pruebas.</li></ul>
<b>Ejecución de Pruebas</b>	Idear	Planillas de Reportes	<ul style="list-style-type: none"><li>• Documentar y validar con los usuarios la criticidad de las novedades que contribuyan a solucionar sus necesidades, detallando las desviaciones obtenidas en la ejecución de las pruebas.</li><li>• Comparar los resultados obtenidos con las necesidades del usuario.</li></ul>
<b>Evaluación de resultados:</b>	Prototipar	Pruebas de Usabilidad	<ul style="list-style-type: none"><li>• Verificar la facilidad de uso y aprendizaje del software en el entorno real del usuario, a través de la ejecución de un ejercicio práctico de simulación.</li></ul>
	Probar	Pruebas de Aceptación	<ul style="list-style-type: none"><li>• Validar con el usuario los resultados del ejercicio anterior.</li></ul>
	Probar	Encuestas de Satisfacción	<ul style="list-style-type: none"><li>• Solicitar comentarios a los usuarios a cerca de la valoración del software testeado.</li></ul>
<b>Cierre del Proceso:</b>	Probar	Informe Final	<ul style="list-style-type: none"><li>• Dar a conocer al usuario las novedades del proceso y determinar conjuntamente con el mismo, si es necesario llevar a cabo un nuevo ciclo.</li></ul>



Los resultados obtenidos en la aplicación de Design Thinking durante el Ciclo de Vida de V&V se detallan en la Tabla 3, donde podemos observar que es importante tener una comunicación constante con los usuarios para que el software de entrenamiento que ellos operan sea un producto usable y de calidad [8].

Tabla 3. Resultados aplicando Design Thinking

<i>Fases del Ciclo de Vida del V&amp;V</i>	<i>Resultados obtenidos en su ejecución habitual</i>	<i>Resultados obtenidos con la aplicación de Design Thinking</i>
<b>Análisis de Requerimientos</b>	Los requerimientos documentados son analizados, comprendidos y clasificados una sola vez, al comienzo del ciclo, y son utilizados a lo largo de todo el proceso sin involucrar a los usuarios.	Los requerimientos son definidos y priorizados a través del trato directo con los usuarios.
<b>Planificación de Pruebas</b>	Los tipos, estrategias y técnicas de prueba son definidas en base a los requerimientos documentados.	Los tipos, estrategias y técnicas de prueba son identificadas y definidas tomando en cuenta las necesidades planteadas por los usuarios.
<b>Diseño y Desarrollo de Pruebas</b>	Los Casos de Prueba son diseñados y desarrollados en base a requerimientos documentados.	Los Casos de Prueba son diseñados y desarrollados en base a escenarios.
<b>Ejecución de Pruebas</b>	Las desviaciones del comportamiento del software son documentadas y transmitidas sin tener en cuenta la opinión de los usuarios.	Las novedades son documentadas y validadas con el usuario en base a su criticidad, detallando las desviaciones obtenidas en la ejecución de las pruebas.
<b>Evaluación de Resultados</b>	Los requerimientos implementados son validados con el usuario final.	Los requerimientos implementados son validados con el usuario final.
<b>Cierre del Proceso</b>	Se reporta el fin del proceso sin intervención del usuario final.	Se reporta el fin del proceso con intervención del usuario final, determinando si es necesario iterar.



## Conclusiones

Los resultados obtenidos en este artículo demuestran que las fases que intervienen en el proceso de Ciclo de Vida de V&V, pueden ser optimizadas mediante la aplicación de Design Thinking. La metodología investigada pretende ser la base para conseguir una mejora sustancial en el desarrollo de software correctamente verificado y validado.

Esta debe ser conocida y asimilada por todos los especialistas involucrados en un proyecto software, para impulsar el cambio hacia formas de trabajo más eficientes incluyendo al usuario como fuente estratégica.

Design Thinking constituye una propuesta que propicia la participación del usuario. Su carácter iterativo permite adaptarse a los cambios, estableciendo un mecanismo que permite asegurar la calidad del software y mantener al usuario satisfecho.

## Contribución de Autoría

**Valeria Raquel Filoniuk:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#), [Escritura, revisión y edición](#). **María Soledad Martínez:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#), [Escritura, revisión y edición](#). **Ana Claudia Diz:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#), [Escritura, revisión y edición](#). **Silvia Edith Arias:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#), [Escritura, revisión y edición](#).



## Referencias

- [1] V. F. C. Osis, D. Q. Soto, A. C. Huarca and J. C. Suyo, "Casos de Estudio de Design Thinking en las etapas de Análisis y Diseño del Desarrollo de Software," *Revista Innovación y Software*, no.1, vol. 3, pp. 17-29, 2022.
- [2] J. C. Espinoza Vásquez and E. E. Espinoza Zapata, "Marco de trabajo en base a Design Thinking y metodologías ágiles de desarrollo de software", thesis, Universidad Peruana de Ciencias Aplicadas, Lima, 2017.
- [3] L. A. Baquero Hernández and C. C. León Suarez, "Integración e implementación de realidad virtual para un Simulador de entrenamiento de vuelo de uso libre", thesis, Fundación Universitaria Los Libertadores, Bogotá, 2022.
- [4] S. I. Pedraza Gutiérrez, J. F. Romero-González, J. C. Güiza Rodríguez and E. W. Giraldo Henao, "Diseño centrado en el usuario y experiencia de usuario en el sistema de control de acceso de la Universidad Libre," *Revista Científica de Sistemas e Informática*, no.1, vol.3, pp. 426, 2023.
- [5] M. L. F. Rojas, "Pensamiento de diseño y marcos éticos para la Inteligencia Artificial: una mirada a la participación de las múltiples partes interesadas," *Revista Desafíos*, no.1, vol.35, pp.1, 2023.
- [6] D. M. Ortega Ávila, "Aplicación de la Metodología Design Thinking para el Diseño y Comercialización de Agendas Personalizadas en la Ciudad de Santa Marta", thesis, Universidad Antonio Nariño, Colombia, 2023.
- [7] E. A. S. Gonzales, C. L. S. Bobadilla, J. C. A. Uñapilco and D. H. Quispe, "Design Thinking para resolver problemas con la selección de métricas en la Calidad del Software," *Revista Innovación y Software*, no.1, vol.3, p.67-80, 2022.
- [8] J. J. G. Valdivia, S. M. Z. Quentasi, D. M. C Yana, R. E. C. Apaza and Y. P. Vera, "Design Thinking en la Planificación de Pruebas de Software," *Revista Innovación y Software*, no.2, vol.1, pp. 40-51, 2020.



- [9] E. R. A. Serrato, "Propuesta de Alfabetización Digital para las mujeres beneficiarias del programa "Mujer Es", de la Fundación Misión y Vida, en el Departamento del Huila, República de Colombia", Tesis Doctoral, Corporación Universitaria Minuto de Dios, Colombia, 2021.
- [10] A. M. Ospina Rivera, and D. M. González Arévalo "Validación por medio de la aplicación del Design Thinking de la oferta de un servicio de medición de emisiones de gases efecto invernadero para el sector pecuario", Tesis de Maestría, Universidad EAN, Colombia, 2021.
- [11] L. Merino García, "Desarrollo de un sistema automático de validación y verificación de software," Thesis, Universidad Pública de Navarra, Pamplona, 2021.
- [12] C. Maximiliano, "Introducción al Testing de Software", Thesis, Universidad Nacional de Rosario, Santa Fe, 2021.
- [13] M. S. Martínez, D. I. Martínez, V. R. Filoniuk, G. G. Chiappori, A. C. Diz and S. E. Arias, "Aplicación de Norma ISO 9241-11 para la Evaluación de la Usabilidad en Simuladores de Vuelo", Revista Innovación y Software, no.2, vol.3, pp. 70-80, 2022.
- [14] G. I. H. Zapana, N. H. Castro, M. A. S. Tico, E. D. C Choquehuanca and A. D. T. Bejarano, "Aplicación del método Design Thinking en el área de requerimientos de software," Revista Innovación y Software, no.1, vol.2, p. 43-52, 2021.
- [15] B. L. V. Márquez, L. A. I. Hanampa and M. G. M. Portilla, "Design Thinking aplicado al Diseño de Experiencia de usuario", Revista Innovación y Software, no.1, vol. 2, pp.6-19, 2021.





# Clasificación de categorías de noticias usando BERT

## *Classification of news categories using BERT*

36

### **Bradlhy Luis Machado Medina**

Universidad La Salle. Arequipa, Perú.

@ bmachadom@ulasalle.edu.pe

### **César Alonso Santillana Quirita**


Universidad La Salle. Arequipa, Perú.


@ csantillanaq@ulasalle.edu.pe


### **Sharmelyn Violeta Bautista Luque**

Universidad La Salle. Arequipa, Perú.

@ sbautistal@ulasalle.edu.pe

 **ARK:** [ark:/42411/s12/a98](https://nbn-resolving.org/urn:nbn:org:ark:ark:/42411/s12/a98)

 **DOI:** [10.48168/innosoft.s12.a98](https://doi.org/10.48168/innosoft.s12.a98)

 **PURL:** [42411/s12/a98](https://purl.org/urn:nbn:org:ark:ark:/42411/s12/a98)

RECIBIDO 18/03/2023 • ACEPTADO 28/06/2023 • PUBLICADO 30/09/2023



## RESUMEN

El presente proyecto consiste en desarrollar un modelo de Procesamiento del Lenguaje Natural para clasificar noticias utilizando un conjunto de datos o DataSets ya evaluados. El objetivo principal es crear un sistema que pueda identificar y asignar automáticamente las noticias a una de las categorías predefinidas: negocios, entretenimiento, política, deportes o tecnología. Esto implica el preprocesamiento de datos, extracción de características, entrenamiento de un modelo de machine learning y posteriormente su evaluación de rendimiento utilizando métricas como "precisión", "recall" y "F1 – score". Esto permitirá determinar que tan bien el modelo puede predecir la categoría correcta para una noticia nueva o no etiquetada. Si el rendimiento del modelo es satisfactorio, se puede utilizar para clasificar noticias no etiquetadas en tiempo real. En resumen, se busca proporcionar una solución eficiente y precisa para organizar y etiquetar el contenido informativo de una noticia con ayuda de la Inteligencia Artificial.

**Palabras claves:** Clasificación de noticias, procesamiento del lenguaje natural, BERT, machine learning, inteligencia artificial.

## ABSTRACT

*The present project consists of developing a Natural Language Processing model to classify news using a set of data or DataSets already evaluated. The main objective is to create a system that can automatically identify and assign news to one of the predefined categories: business, entertainment, politics, sports or technology. This involves data preprocessing, feature extraction,*



*training a machine learning model and then evaluating its performance using metrics such as "accuracy", "recall" "F1 - score". This will allow to determine how well the model can predict the correct category for a new or unlabeled news item. If the performance of the model is satisfactory, it can be used to classify unlabeled news in real time. In summary, it seeks to provide an efficient and accurate solution for organizing and labeling the informative content of a news item with the help of Artificial Intelligence.*

**Keywords:** News classification, natural language processing, BERT, machine learning, artificial intelligence.

---

## INTRODUCCIÓN

En la actualidad, con la globalización, era digital y avance tecnológico, hay una enorme cantidad de información disponible en la nube o físicamente. Esto representa un desafío en términos de organización, clasificación, evaluación y etiquetado. La clasificación de noticias es una tarea fundamental para facilitar el acceso a la información relevante y brindar una experiencia más eficiente a los usuarios, en cuanto a sus preferencias o el tema que desee leer. En este contexto, el presente trabajo de investigación se enfoca en el desarrollo de un modelo de Procesamiento del Lenguaje Natural (*NLP*) para la clasificación automática de noticias.

Se planea diseñar y realizar un sistema capaz de identificar y asignar automáticamente las noticias a una de las siguientes categorías establecidas: negocios, entretenimiento, política, deportes o tecnología. Para lograr dicha clasificación, se empleará un conjunto de datos ya evaluados (*Dataste*) obtenidos de la plataforma [Tagle](#) que servirán como base para el entrenamiento del modelo de *machine Learning*.

Para llevar a cabo la clasificación automatizada de noticias, se realizarán varias etapas. En primera instancia, se lleva a cabo el procesamiento de los datos para limpiar, normalizar el texto y eliminar información innecesaria.

Luego, se utilizarán técnicas avanzadas de *NLP* y representaciones de texto, como el modelo *BERT*, para extraer las características relevantes de cada noticia. Posteriormente, se entrenará un modelo de *machine Learning* que aprenderá a asignar las noticias en sus respectivas categorías. Finalmente, se medirá el rendimiento del modelo mediante el uso de métricas como "precisión", "recall" "F1-score", para determinar la capacidad del modelo de predecir la categoría correcta para noticias no etiquetadas.



## Motivación

Este trabajo planea desarrollar una herramienta de clasificación de tema de textos, específicamente en las categorías ya definidas anteriormente de las noticias.

### ¿En qué dominio del conocimiento está trabajando?

En cuanto al tema de clasificación de temas de noticias, el dominio del conocimiento en el que se está trabajando es el campo de Procesamiento del Lenguaje Natural (*NLP*), Inteligencia Artificial (*IA*) y clasificación de textos.

### ¿Quiénes son los usuarios objetivo?

Los usuarios objetivo de este proyecto puede ser público en general, pero principalmente se enfocan en especialistas en análisis de contenido informativo, es decir periodistas, editores de noticias o investigadores en medios de comunicación. Estos pueden beneficiarse de un sistema automatizado que les permita clasificar y etiquetar rápidamente las noticias en diferentes categorías ten áticas ya preestablecidas.

### ¿Porque es interesante el tema que proponen?

El tema propuesto es interesante puesto que la cantidad de noticias generadas diariamente en el mundo y cada país es enorme, lo que dificulta su procesamiento manual y clasificación. Un sistema automatizado puede ayudar a gestionar esta gran cantidad de información y ahorrar tiempo y esfuerzo a los profesionales encargados de analizar, organizar y clasificar las noticias.

### ¿Cuáles son las preguntas que su proyecto de NLP intenta responder?

- Q1: ¿Que categoría de noticia es?
- Q2: ¿Como se clasificó la noticia?
- Q3: ¿La clasificación fue correcta?



## Problema

El problema en el que se centra el proyecto planteado es encontrar una solución eficiente y precisa que permita la clasificación automática de temas de noticias, superando las limitaciones del procesamiento manual y proporcionando una herramienta útil para los profesionales involucrados en el análisis y la gestión de noticias, reduciendo así el tiempo, esfuerzo y costo de hacerlo manualmente.

## Objetivos

Proporcionar una solución automatizada basada en *NLP* que resuelva el problema de la clasificación de temas de noticias, permitiendo una clasificación eficiente y precisa en tiempo real. Esto beneficiará a profesionales de medios de comunicación, analistas de contenido y cualquier persona involucrada en el procesamiento y análisis de grandes volúmenes de noticias.

## Datos

### ¿Qué datos necesitará?

Para el proyecto se necesitarán dos conjuntos principales de datos o *Dataste*, por un lado, los datos de entrenamiento que serán las noticias con su etiqueta ya clasificada en los temas mencionados, y por otro los del *testing*, que serán solamente noticias sin etiqueta alguna. Estos serán usados tanto para definir y entrenar el modelo, como también para probarlo y ver su funcionamiento.

### ¿Como recolectaran los datos?

Se descargará desde un repositorio en donde se encuentra almacenado.

### ¿Dónde planea obtenerlos?

En [Tagle](#), que es una plataforma donde se alojan y almacenan diversos *Dataste* de entrenamiento y *testing* de clasificación de *NLP*, específicamente el *Dataset* denominado [News Category Datase](#).



## ¿Como planea almacenarlos?

En archivos "csv" que son de texto plano para guardar datos tabulares. Cada línea del archivo representa una fila de datos, y los valores de cada fila están separados por comas. Estos organizaran los datos de manera organizada y estructural, además que tendrá poco peso de almacenamiento por el formato que se está usando, para que posteriormente sean consumidos desde el modelo que realizaremos.

## ¿Como accederá a ellos para utilizarlos en su proyecto?

Mediante la lectura de ficheros, abriendo y organizándolos con la librería "Pandas" (*pd*).

## Diseño

Para utilizar la herramienta propuesta y ayudar a los usuarios a realizar las tareas que responden a las preguntas mencionadas en la motivación, se deber `a tener en cuenta las siguientes funcionalidades:

- Categoría de noticia (Q1): Se debe permitir a los usuarios ingresar una noticia y utilizar el modelo de procesamiento del lenguaje natural desarrollado para clasificarla en una de las categorías predefinidas.
- Clasificación de la noticia (Q2): Después de clasificar la noticia en una categoría específica, la herramienta debe mostrar al usuario la etiqueta o categoría asignada.
- Evaluación de la clasificación (Q3): Se debe proporcionar una función de evaluación de la precisión de la clasificación realizada por el modelo. Esto incluye las métricas ya mencionadas, que permitirán al usuario determinar la calidad de la clasificación realizada y evaluar si fue correcta, además que, si tiene buena precisión en los resultados, generaremos confianza en nuestro diseño.

El código, debería ser una implementación eficiente del modelo de procesamiento del lenguaje natural y de las funciones de clasificación. Se usar `a *BERT* con la librería *Pythorch* y *transformers* para el procesamiento en una plataforma como [Colab](#). Además también se controlar `a el versionamiento con ayuda de dicha herramienta y tendrá comentarios en los códigos, además de definición importante en el archivo si es necesario para una mayor comprensión

## Trabajos relacionados





En cuanto a los trabajos relacionados, consultamos diversos sitios o páginas web que realizaban la implementación de modelos similares, como *ScalerTopic*<sup>1</sup>, *NewsClasificationusingBERT*<sup>2</sup> o *AnalyticsVidhya*<sup>3</sup>, que sirvieron como guía para elegir y trabajar el tema, además que nos pareció interesante porque es un tema de actualidad y con el constante crecimiento de día a día de información se va haciendo ineficiente hacerlo de forma manual.

## Revisión Literaria

En esta sección, presentaremos algunos de los trabajos relacionados que ya han clasificado diferentes noticias de diferentes idiomas utilizando diferentes modelos. Hay muchos trabajos donde la clasificación se centra en patrones lingüísticos, estructura de las oraciones, significado semántico de las oraciones, y se extraen algunas palabras clave para detectar alguna categoría específica.

Las Redes Neuronales Recurrentes y las Redes Neuronales Convolucionales son los enfoques más utilizados para realizar tareas de procesamiento del lenguaje natural en la clasificación de noticias.

Hayes *et al* [2], categorizaron noticias en amplias categorías temáticas utilizando métodos de NLP con técnicas de coincidencias de patrones, basados en algoritmos de *machine Learning* para identificar palabras y frases clave en un contexto adecuado.

Su estudio no realiza un análisis semántico o sintáctico completo de las historias, sino que depende de reconocimiento fragmentario. Se divide en dos fases: sintáctico y confirmación. La sintáctico busca identificar todas las categorías posibles en función de las palabras y frases presentes en la historia, mientras que la confirmación busca evidencia adicional para respaldar una hipótesis o determinar si el lenguaje utilizado puede haber llevado a una hipótesis incorrecta. Cabe destacar que el sistema logró una precisión promedio del 93%.

---

1 <https://www.scaler.com/topics/machine-learning/bbc-news-classification/>

2 <https://www.kaggle.com/code/foolofatook/news-classification-using-bert/notebook>

3 <https://www.analyticsvidhya.com/blog/2021/12/text-classification-of-news-articles/>



Wu *et al* [4], Realizan un webcrawler para obtener noticias de Internet y se procesa la segmentación de oraciones mediante el algoritmo Jieba. Luego, se marca la clasificación para cada noticia procesada y se ordenan las palabras por número de ocurrencias.

Chy *et al* [1], propusieron un modelo para clasificar noticias en bengalí utilizando el clasificador Naive Bayes para categorizar 34 tipos de noticias. Utilizaron su propio *web crawler* para recopilar un conjunto de datos de varias páginas noticias en línea y aplicaron el enfoque estadístico de *Naive Bayes* para clasificar diferentes noticias.

Aplican técnicas de estadístico, stemming y eliminación de palabras para mejorar la precisión de noticias y evalúa el sistema utilizando medidas de precisión y recall. Sin embargo, su grafico de Recall-Precision mostró una precisión inferior al 80% y no proporciono correctamente los nombres de las categorías de noticias.

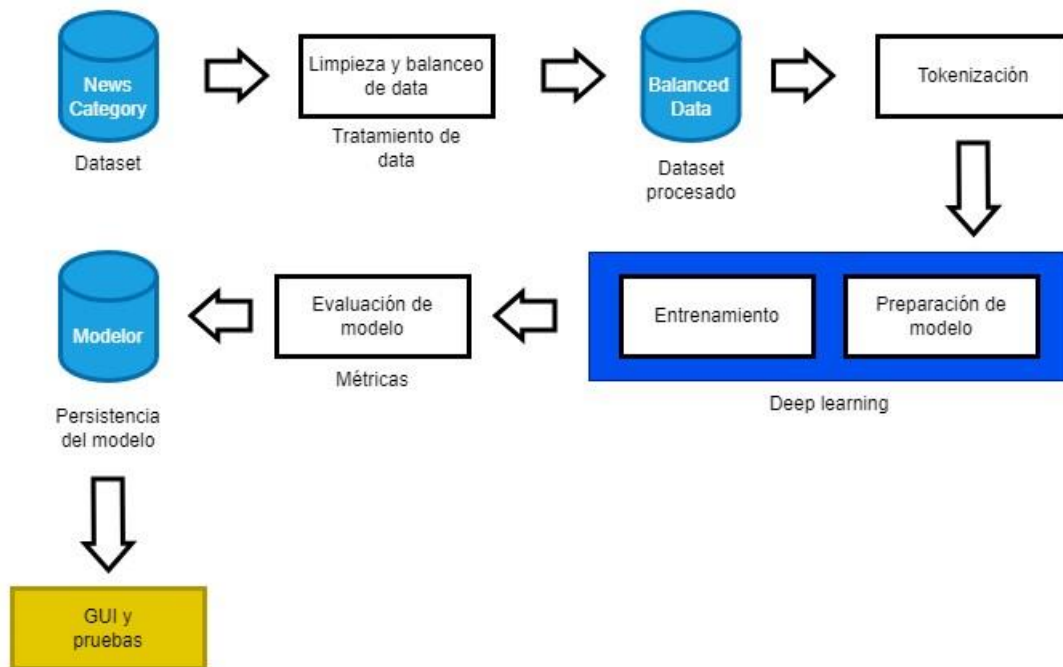
Rahman *et al* [3], implementan un enfoque a nivel de caracteres que categoriza noticias bengalís. Utilizan dos modelos de *deep Learning* CNN y LSTM, respectivamente, para la clasificación, usando la ley de Pareto. La mayor precisión y puntuación F1 se obtienen para el conjunto de las bases de datos utilizando el modelo LSTM del 94% aproximadamente.

Wang *et al* [5], presentan un paper académico que alberga un nuevo conjunto de datos, N24News, que contiene información tanto textual como visual de cada noticia y que tiene 24 categorías. El paper también muestra cómo usar un método multimodal para mejorar la clasificación de noticias.

En base a dicha información implementaremos un modelo usando *BERT* para categorizar diferentes tipos de noticias.



## Diseño



Para la clasificación de noticias seguiremos los pasos mostrados en la Fig. 1.

- *DataSet News Category Dataset v3*: Se obtienen los datos a partir de un *dataset* de noticias clasificado en formato *csv*.
- *Tratamiento de datos*: En esta fase vamos a limpiar y balancear el conjunto de datos.
- *Tokenización*: Hace referencia al proceso de dividir un texto en unidades más pequeñas para la mejor comprensión por parte del modelo.
- *Deep Learning*: En esta fase prepararemos el modelo a usar "bert-base-uncased" para el posterior entrenamiento de nuestro modelo con *BERT*.
- *Evaluación del modelo*: En esta etapa se realiza la evaluación del modelo usando las métricas que se estableció.
- *Persistencia*: En esta fase se almacena el modelo ya entrenado para su posterior consumo.
- *GUI y pruebas*: Finalmente desplegaremos nuestro modelo mediante una interfaz que permita ingresar *inputs* de noticias y que se nos dé a que categoría pertenece.



## Bibliotecas utilizadas:

- *Transformers*: Proporciona métodos para descargar y entrenar fácilmente modelos pre entrenados.
- *Torch*: Es un *framework* para diseñar y entrenar redes neuronales.
- *Pandas*: Pandas es una librería de Python especializada en la manipulación y el análisis de datos. *Sklearn*: Para evaluar el modelo con las métricas establecidas

## Resultados:

Se decidió trabajar colaborativamente en [Colab](#) para utilizar los recursos de la GPU. Además, utilizamos transformers. Al utilizar Transformers en nuestro clasificador de noticias, podemos aprovechar su potencial para procesar y comprender el lenguaje humano. Los modelos reentrenados disponibles en la biblioteca nos permitieron extraer características relevantes de los textos, capturar la semántica y contextualizar la información, lo que resultó fundamental para la clasificación de las noticias en diferentes categorías o temas. Las categorías y el tamaño del conjunto de datos se muestran a continuación:

- Las categorías de las noticias son las siguientes: *ENTERTAINMENT, POLITICS, STYLE&BEAUTY, TRAVEL, WELLNESS*
- El tamaño del conjunto de datos a usar es de 49070
- La data se encuentra balanceada, donde cada categoría tiene un total de 9814 datos.

## Proceso:

1. **Configuración de GPU:** Dado que entrenaremos una red neuronal usamos Google Colab porque proporciona acceso a GPUs gratuitas. Al asociar Colab con **CUDA**, se aprovechó mejor el poder de la GPU para acelerar las operaciones computacionales necesarias durante el entrenamiento del clasificador de noticias. Para verificar la disponibilidad de una GPU se utilizó la biblioteca PyTorch y se configuro el dispositivo de entrenamiento para usar la GPU disponible. Más tarde, en nuestro ciclo de entrenamiento, cargaremos datos en el dispositivo.
2. **Selección de categorías y balanceo de data:** Utilizaremos la biblioteca *panda* para leer un archivo JSON que contiene el conjunto de datos de noticias. Luego, filtramos las categorías deseadas y combinamos las columnas relevantes en un nuevo DataFrame. A continuación, se realizará el balanceo de datos utilizando *RandomUnderSampler* para igualar el número de muestras en todas las categorías,



¿como se muestra en la Fig.? Finalmente, el *DataFrame* balanceado se guarda en un archivo JSON en Google Drive. De esta manera realizamos el proceso de preparación de datos para el entrenamiento de un clasificador de noticias.

```
//Imprimiendo la cantidad total de categorías seleccionadas print ("Total data to be used:"), df.groupby("category").size().sum()
```

```
ENTERTAINMENT9814
POLITICS      9814
STYLE & BEAUTY 9814
TRAVEL        9814
HELLNESS      9814
```

```
Total data be used: 49070
```

3. **Tratamiento de datos:** Se emplea *sklearn* para dividir el conjunto de datos de noticias en entrenamiento y prueba. Donde el 20% de los datos se destinará a las pruebas, mientras que el 80% se utilizará para el entrenamiento, lo que es fundamental para evaluar la precisión y el rendimiento del clasificador.
4. **Tokenización:** Se crea un tokenizador *BERT* utilizando el modelo pre-entrenado de *Transformers* de *Hugging Face* para cargar el tokenizador *BERT* pre-entrenado. De esta manera se preparan los datos para el entrenamiento y la evaluación del modelo de clasificación utilizando *BERT*. Convirtiendo los textos en secuencias numéricas para obtener las etiquetas correspondientes y definir los parámetros necesarios para el modelo de clasificación de noticias.
5. **Modelo de clasificación:** En el siguiente pseudocódigo se puede ver que se crea un modelo de clasificación de secuencias utilizando el modelo pre-entrenado de *BERT*, además de un optimizador utilizando el algoritmo de optimización *AdamW* y se pasan los parámetros del modelo para ser optimizados. Además, se definen los parámetros para el entrenamiento, como el número de épocas, que fueron diez épocas, y la tasa de aprendizaje.

```
model = BertForSequenceClassification.from_pretrained(
'bert-base-uncased', num_labels=num_classes,
output_attentions=False,
output_hidden_states=False

batch_size = 32 # Tamaño del lote epochs = 10
# Número de épocas learning_rate = 2e-5 #
Tasa de aprendizaje
```





```
optimizer = AdamW(model.parameters(),  
lr=learning_rate, eps=1e-8)  
)
```

- 6. Preparación del modelo:** En la Fig.?? se prepara los conjuntos de datos y los cargadores de datos para el entrenamiento y la evaluación del modelo. Además, se mueve el modelo a la GPU si está disponible y se verifica la ubicación de los parámetros del modelo.

*#Crear objetos DataLoader para los conjuntos de entrenamiento y prueba*

```
train_dataset = NewsDataset(train_input_ids, train_attention_masks, train_labels) test_dataset =  
NewsDataset(test_input_ids, test_attention_masks, test_labels)
```

```
train_dataloader = DataLoader(train_dataset, batch_size=batch_size)  
test_dataloader = DataLoader(test_dataset, batch_size=batch_size)
```

- 7. Métricas:** Para calcular diferentes métricas de evaluación del modelo se importaron algunas funciones de la biblioteca *sklearn.metrics* y de *torch.nn.functional*. Las métricas que se usarán serán: precisión, Recall y F1-score. Estas funciones de métricas serán utilizadas después de realizar predicciones con el modelo entrenado y las etiquetas de prueba. De esta manera se proporciona una comprensión más completa del desempeño del clasificador de noticias.

- 8. Entrenamiento y prueba:** NO encontramos en el proceso de entrenamiento y evaluación del modelo de clasificación de noticias utilizando el algoritmo *BERT*. En cada época, se itera sobre los lotes de datos de entrenamiento y se calcula la pérdida, las predicciones y las métricas de desempeño: **precisión, F1, recall y precisión**, para el conjunto de entrenamiento. Luego, se realiza una evaluación similar en el conjunto de prueba. Al final de cada época, se imprimen y muestran las métricas promedio de entrenamiento y evaluación. Esto permite monitorear el desempeño del modelo a lo largo del entrenamiento y evaluar su capacidad para clasificar noticias en las 5 categorías indicadas.

```
For epoch in rango(epochs)  
#Inicializar variables de entrenamiento a 0.0  
#para el entrenamiento y evaluación  
Model.train()  
For batch in train_dataloader  
#get inputs y labels del batch  
#Realizar pasos de entrenamiento
```



```
#Actualizar métricas de entrenamiento  
Calcular promedios de métricas de entrenamiento  
Print métricas de entrenamiento  
Model.eval()  
Print métricas de validación
```

9. **Persistencia del modelo:** En la función guardar modelo se toma como entrada el modelo. Luego, se guarda el *state dict* del modelo en el archivo especificado utilizando la función *torch.save*. El *state dict* contiene los parámetros entrenados del modelo. Finalmente, se imprime un mensaje para indicar la ubicación donde se ha almacenado el modelo.

```
def guardar_modelo(modelo, nombre_archivo):  
    # Guardar el state_dict del modelo en un archivo path =  
    '/content/drive/MyDrive/NLP_NewsClas/' + nombre_archivo  
    torch.save(modelo.state_dict(), path)  
    # Imprimir un mensaje  
    print('Model stored in: ', nombre_archivo)  
La función cargar_modelo se encarga de cargar el estado de un modelo previamente guardado.  
  
def cargar_modelo(nombre_archivo):  
    path = '/content/drive/MyDrive/NLP_NewsClas/' + nombre_archivo # Crear una  
    configuración de Bert usando los parámetros por defecto config =  
    BertConfig(num_labels=5)  
    # Crear una instancia del modelo de clasificación de secuencias  
    # usando la configuración  
    modelo = BertForSequenceClassification(config) # Cargar el estado del modelo desde el archivo  
    modelo.load_state_dict(torch.load(path)) # Devolver el modelo cargado return modelo
```

10. **Prueba del modelo:** Finalmente se prueba el modelo entrenado utilizando una instancia del tokenizador BERT, el modelo previamente cargado y un texto de entrada. El texto se codifica utilizando el tokenizador y se pasa al modelo para realizar la inferencia. Se aplica una función *softmax* a la salida del modelo para obtener las probabilidades de cada clase, así como se muestra en la Fig. ???. Estas probabilidades se imprimen para mostrar la clasificación del texto de entrada. Además, se muestran las categorías y los códigos asignados a esas categorías en el conjunto de entrenamiento, lo que ayuda a verificar las etiquetas utilizadas durante el entrenamiento.



```
# Crear una instancia del tokenizador de Bert
tokenizer = BertTokenizer.from_pretrained('bert-base-uncased')

# Definir una entrada de texto para probar el modelo texto = "cuerpo
noticia"

# Codificar el texto usando el tokenizador entrada =
tokenizer(texto, return_tensors='pt')

# Obtener la salida del modelo usando el método __call__
salida = modelo(**entrada)

# Aplicar una función softmax para obtener las probabilidades probabilidades =
F.softmax(salida.logits, dim=1)

# Mostrar las probabilidades del modelo print(probabilidades)
```

## Comparación:

Comparando el desempeño que tuvo BERT con Naive Bayes y el antiguo modelo de Bert en la clasificación de categorías de noticias, se tomaron en cuenta algunos aspectos importantes como lo son:

**Preprocesamiento de datos:** Se aplicaron técnicas de estadístico utilizando medidas de precisión y recall.

**Tamaño del conjunto de datos:** Para un gran conjunto de datos, por ejemplo, más de 45k Naive Bayes es menos efectivo. Para un conjunto un conjunto de datos grande, fue mejor entrenar con un modelo más complejo como lo es BERT. No obstante, si el conjunto de datos es relativamente pequeño, Naive Bayes puede ser más efectivo debido a su menor requerimiento computacional.

En Naive Bayes

- **Representación de características:** Naive Bayes utiliza una representación de características basada en la frecuencia de palabras o n-gramas. BERT captura mejor el contexto y las relaciones semánticas entre las palabras, lo que condujo a un mejor rendimiento en tareas de clasificación de texto.
- **Eficiencia computacional:** Naive Bayes es un modelo más rápido en términos de entrenamiento y predicción en comparación con BERT, que es un modelo más complejo y requiere más recursos computacionales.
- **Evaluación del rendimiento:** En cuanto a las métricas de evaluación adecuadas para la clasificación de texto se usaron: precisión, recall y F1-score.



En la **Tabla 1** se muestra la cantidad de datos, número de categorías y las métricas de precisión para comparar el rendimiento de los modelos predecesores. Además, se agregó una comparación con un modelo de Bert, denominado Bert antiguo que usa la misma Base de Datos. Para el entrenamiento se usó la ley de Pareto.

Métricas	BERT mejorado	Bert antiguo	Naive Bayes	Roberta	VIT
Categorías	5	40	12	24	32
accuracy	94%	70%	85%	91%	92%
F1	94%	no se ubica	75%	87%	94%
Recall	94%	no se ubica	78%	no se ubica	no se ubica

Cuadro 1: Cuadro comparativo

En comparación con el modelo Bert antiguo y Naive Bayes, el modelo BERT mejorado muestra una mejora significativa en todas las métricas de evaluación. El accuracy aumentó del 0.7% al 94%, lo que indica un rendimiento mucho mejor en la clasificación de procesos. Además, el modelo BERT mejorado supera al Naive Bayes en términos de precisión, con un 94% frente al 85%. El F1 score también muestra una mejora considerable, alcanzando un valor de 94% para BERT mejorado y 75% para Naive Bayes.

En cuanto a la comparación con RoBERTa esta vendría a ser la versión optimizada de BERT que fue entrenada con textos mucho más grande y durante más tiempo, lo que le permite capturar aún más información contextual y semántica de una noticia, además de imágenes. Sin embargo, la calidad de las noticias de entrenamiento es menores cuando se le pasa poco texto, en este caso habría que enviar una foto para mejorar el F1 y el recall. Tal como se muestra en la Tabla 1. EL accuracy y el F1 son menores a los obtenidos por el entrenamiento realizado con Bert.

Se usa ViT, junto con el algoritmo Jieba para tokenizar el texto de noticias. Uso de machine Learning para entrenar a la computadora y registrar las palabras clave para cada categoría de noticias. Esto implica que se utiliza el procesamiento de lenguaje natural y el machine Learning en el proceso de clasificación de noticias.

Cabe destacar que la cantidad de datos y el número de categorías son diferentes en todas las comparaciones.



Es importante tener en cuenta que el rendimiento de los modelos puede variar dependiendo de diversos factores, como el tamaño y calidad de los datos de entrenamiento, la selección de características y la selección utilizados en el entrenamiento. También se destaca que el uso de la ley de Pareto, el balanceo de datos, el uso de GPU con la librería PyTorch y la limpieza de datos contribuyeron a la mejora considerable de las evaluaciones tomadas en el modelo BERT mejorado.

## Conclusiones:

En conclusión, la calidad y la preparación adecuada de los conjuntos de datos utilizados para entrenar modelos, como BERT, son factores críticos que influyen en los resultados de la evaluación y las métricas obtenidas. El uso de un dataset balanceado y una limpieza exhaustiva de los datos, centrándose en los campos relevantes para el entrenamiento, se revela como un aspecto crucial para obtener resultados óptimos.

Un dataset balanceado garantiza que el modelo se entrene con una representación equitativa de las distintas clases o categorías presentes en los datos. Esto evita sesgos y permite que el modelo aprenda de manera equilibrada, mejorando así su capacidad para clasificar correctamente las muestras de prueba.

La limpieza de datos también desempeña papel fundamental. Al eliminar ruido, datos irrelevantes o redundantes, y garantizar la integridad de los campos necesarios para el entrenamiento, se mejora la calidad y la coherencia del conjunto de datos. Esto se traduce en un aprendizaje más preciso por parte del modelo, lo que se reflejará en una evaluación más confiable y en métricas más sólidas.

## Contribución de Autoría

**Bradhy Luis Machado Medina:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **César Alonso Santillana Quirita:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Sharmelyn Violeta Bautista Luque:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#).





## Referencias

- [1] Abu Nowshed Chy, Md Hanif Seddiqui, and Sowmitra Das. Bangla news classification using naive bayes classifier. In *16th Int'l Conf. Computer and Information Technology*, pages 366–371. IEEE, 2014.
- [2] Philip J Hayes, Laura E Knecht, and Monica J Cellio. A news story categorization system. In *Second Conference on Applied Natural Language Processing*, pages 9–17, 2000.
- [3] Md Mahbubur Rahman, Rifat Sadik, and Al Amin Biswas. Bangla document classification using character level deep Lear Ning. In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pages 1–6. IEEE, 2020.
- [4] Meng-Jin Wu, Tzu-Yuan Fu, Yao-Chung Chang, and Chia-Wei Lee. A study on natural language processing classified news. In *2020 Indo-Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)*, pages 244–247. IEEE, 2020.
- [5] Zhen Wang, Xu Shan, Xiangxie Zhang, and Jie Yang. N24news: A new dataset for multimodal news classification, 2022.



# Análisis de fugas de datos en redes inalámbricas mediante pruebas supervisadas y no supervisadas

52

## *Data leakage analysis in wireless networks using Super-vised and Unsupervised Testing*

### **Shahzad Ashraf**


NFC Institute of Engineering and Technology Multan Pakistan. Multan, Pakistan.


 nfc.ie@hotmail.com


### **Zeeshan Aslam**

Alfanar Global Development Saudi Arabia. Damman, Arabia Saudita.

 zeeshan.aslam@alfanar.com

 **ARK:** [ark:/42411/s12/a108](https://nfdi.org/ark:/42411/s12/a108)

 **DOI:** [10.48168/innosoft.s12.a108](https://doi.org/10.48168/innosoft.s12.a108)

 **PURL:** [42411/s12/a108](https://nfdi.org/42411/s12/a108)

RECIBIDO 25/03/2023 • ACEPTADO 08/07/2023 • PUBLICADO 30/09/2023



## RESUMEN

Debido al creciente número de espectros inalámbricos, las múltiples frecuencias están enredando el proceso de gestión de recursos, lo que dificulta el funcionamiento. Además, los datos anteriores se vuelven vulnerables cuando se reciben informes de enigma de fuga de datos. En esta situación, es indispensable asegurar los datos en el conjunto de datos y detectar la cantidad real de datos durante el mecanismo de transformación de recursos en redes inalámbricas. Se ha desarrollado un sistema para detectar la fuga de datos utilizando técnicas de prueba supervisadas y no supervisadas mediante simulación en Python. Se obtienen los resultados previstos y reales, que se reducen mediante pruebas supervisadas y no supervisadas, el resultado sigue siendo del 96,03% y 94,53% respectivamente.

**Palabras claves:** Pruebas supervisadas, pruebas no supervisadas, redes neuronales, redes inalámbricas.

## ABSTRACT

*Due to an increasing number of wireless spectrums, the multiple frequencies are tangling resource management process that results hindrance in operation. In addition, the previous data become vulnerable when reports are received for data leakage enigma. In this situation, it is indispensable to secure the data in the dataset and detect the actual amount of data during resource transformation mechanism in wireless networks. A system as been developed to detect the leaked*



*data using supervised and unsupervised testing technique by conducting simulation in Python. The targeted and actual outcome is obtained which deduced through supervised and undersized testing, the outcome remained 96.03%, and 94.53% respectively.*

**Keywords:** : *Supervised testing, unsupervised testing, neural network, wireless networks.*

---

## INTRODUCTION

With the exponential increase in the number of wireless devices in recent years and due to the rapid growth of wireless services, it has become increasingly important to allocate and manage them accordingly. There is, however, generally a challenge to establish suitable strategies due to the fact that the non-convex objective average rate maximization problem is NP-hard. It is no secret that there has been a growing interest in numerical optimization as it relates to wireless resource allocation in recent years.

There are still significant challenges in implementing numerical optimization-based algorithms on practical systems, e.g., high computation costs, despite their ability to solve specific resource management problems with tremendous results [1]

As neural networks (NNs) [2], memorize features of example data during training and unintentionally reveal them during prediction, it is a major concern for machine learning applications to prevent models from revealing sensitive input data details. There is, however, no easy way to accomplish this. The literature is lacking studies that address deep learning-based wireless resource allocation systems with privacy protection.

These radio resource management algorithms cannot always provide an adequate degree of performance due to the dynamic nature of wireless networks. An algorithm that learns from interactions with the environment may be able to better handle such dynamics.

Numerous resource management schemes have been proposed in order to address heterogeneously [3], high service demands, fairness, and starvation, as well as transmission errors due to channel congestion. In spite of this, there are many schemes out there that prioritize voice services while allocating the remaining bandwidth to non-real-time applications without any guarantee that the delay will not affect the voice service.

In wireless communication, when multiple resources exchanges between heterogeneous network environment there is a great chance of distribution of data without the notice of the relevant agencies. This lost of data sometimes create a big hassle and situation becomes aggravate [4]. In this situation, It is imperative to develop a system that should detect the leaked data proactively. After detection of the leaked data the prevention measures can be taken for future. After going through different studies, no specific study is found that can detect the wastage of



data during resource transformation mechanism in wireless communication therefore, an intelligent data detection mechanism has been developed using supervised and unsupervised testing mechanism. This technique utilized hidden layers of NN and then generates the outcome which further processed by conducting simulation in Python.

The following are the main contributions of this work.

- Using intelligent supervised and unsupervised testing, the amount of data leakage would be identified.
- The tested data would be simulated in python to identify the amount of accurate detected data and the amount of lost data.
- The targeted and actual outcome is analyzed to deduce the impact of proposed method.

Rest of the manuscript is arranged as: Section 2 presents a comprehensive overview of previous work. Section 3, is enriched with proposed methodology using supervised and unsupervised testing and Section 4 assesses viability of performance of the proposed method. Finally, Section 5 concludes about findings and future research outlook.

## Literature review

It is necessary to advance wireless communication technologies both in terms of scale and complexity to support emerging applications. Machine learning and predictive regression methods [5], have been investigated in order to solve wireless resource allocation problems. Supervised learning involves training neural networks to approximate a function or algorithm so that computation time is minimized. Additionally, their use requires strategies that protect privacy while fulfilling the application requirements.

According to Jorge Cortés [6], privacy preserving data analysis must take into account dynamic data as well as data exchanged across networks, as well as systems and control perspectives. In order to protect their data against adversaries with arbitrary side information, they adopted differential privacy mechanisms that were initially used to analyze large, static datasets.

Under differential privacy constraints, they reviewed how multiple agents can perform signal estimation, consensus, and distributed optimization tasks. Several factors were ignored in this study, including how to deal with signals when multiple spectrums exist and how to avoid tangling them. Further, it is crucial to investigate the appropriate scales for privacy parameters based on specific application domains as well.

Li Ming [7], work was to identify the current situation of small- and medium-sized organizations' human resources, using deep learning data. Through the deep learning approach, human resources can be more productive, and business volumes can be reduced, thereby improving





human resource efficiency. His proposed model was improved by implementing a deep neural network. In an analysis of experimental data, several types of decent gradient processes were considered as well as a number of neurons in the hidden layer.

Because of the low calculation complexity and fast training speed, the model fails if there are non-linear relationships between variables, which is common for multiple frequencies.

According to Haoran Sun [8], wireless resource management can be improved by utilizing a learning-based approach. An unknown non-linear map is treated as a resource allocation algorithm and approximated using deep neural networks. As long as a DNN of moderate size is capable of learning accurate and effective non-linear mapping, such a DNN can be used to allocate resources in almost real time, since the input to get the output only involves a few simple operations. In order to approximate some of the algorithms of interest in wireless communications, they developed DNNs to approximate a class of 'learnable algorithms'. Using numerical simulations, the authors demonstrated that DNNs are superior to conventional algorithms for estimating two relatively complex algorithms for power allocation in wireless transmissions. This model only represents a very preliminary step towards understanding the capability of DNN and how to deal with challenging problems such as beamforming for IC/IMAC is still unknown. It also could not answer how to further reduce the computational complexity of DNN?

The application of differential privacy is well documented. However, wireless resource allocation schemes with differential privacy based on Deep Neural Networks (DNNs) [9], have never been studied. This study investigates the impact of Differential Privacy DP) [10], on model convergence and network performance using neural network resource allocation schemes.

These research findings show differentially private schemes can produce high-performance models, especially when Convolutional Neural Networks (CNNs) [11], are used.

## **Proposed method using supervised and unsupervised testing**

Differential privacy is achieved by adding noise to the data used in a machine learning model in a way that does not significantly affect the accuracy of the model. This noise ensures that individual data points cannot be easily identified in the model's output. In neural networks, DP can be incorporated into the training process by adding noise to the weights or gradients of the model during training.

There are many functionally interconnected neurons in a neural network that follow a topological structure [12]. The hierarchical structure of neuron is used to categorize neural network models into hierarchical and interconnection models.





According to a hierarchical model, neurons are classified into different layers in accordance with their functionality and are interconnected every year. An input, middle, and output layer structure is used in a hierarchical model, while any two neurons are linked in an interconnection model. Hierarchical models are widely used because of their easy analysis and good structure.

### 3.1 Selection Process

Regression problems include prediction of leaked data during resource allocation process. A regression model is a way of predicting an outcome in machine learning [13]. A regression model may be linear, elastic, neural, polynomial, or ridge. There is a wide range of models which use linear variables, among the most common of which is linear regression. The relationship between the dependent variable and the independent variable is necessary in a linear regression model. This makes linear regression applicable only to problems whose solutions are linearly separable. The model estimates linear relationships between variables because it's simple to calculate and fast to train. Linear regression models are sensitive to outliers, which is a disadvantage.

### 3.2 Identification of hidden layers

There are several mathematical iterations to calculate amount of actual hidden layers [14]. Two of the most commonly implemented iterations are represented by equations (1) and (2).

$$N_h = \sqrt{n_i \times n_o} \quad (1)$$

$$N_h = \sqrt{n_i \times n_o} + K \quad (2)$$

In these equations, the numbers of input and output layer nodes are represented by  $n_i$  and  $n_o$  as shown in figure 1. During training, a slight expansion in space is observed, i.e., in the range and the deep neural network should train continuously based on equation (2). 16, 17, and 1 neurons in total were assessed in the input, hidden, and output layers, respectively. The neuron activation function represents sigmoid functions, and the error function is seen to be the quadratic mean square value.

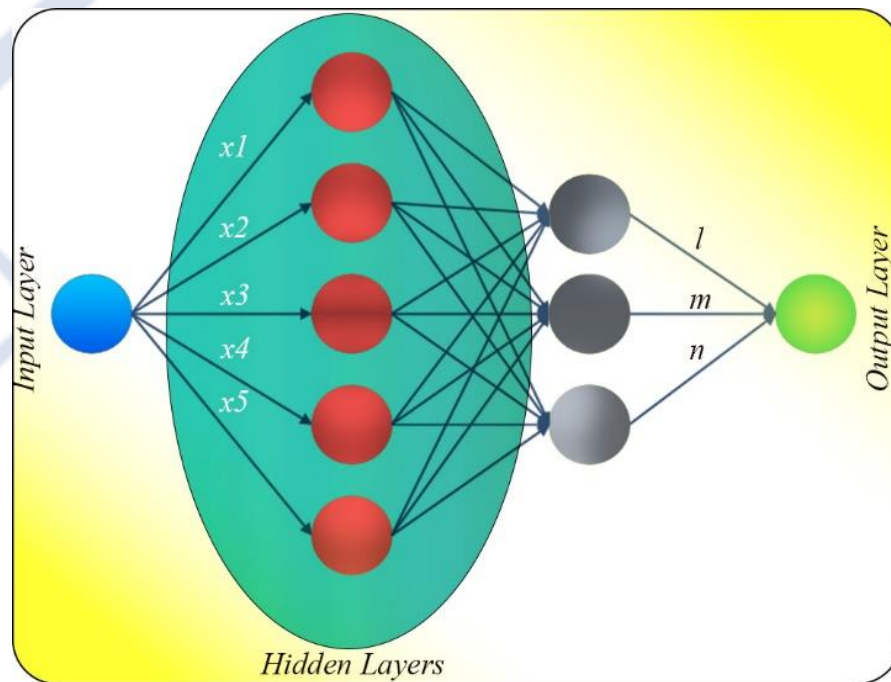


Figure 1. Data leakage detection through Supervised testing mechanism

In supervised testing,  $n_1$  as input layer is applied it generates the sublayers as  $x_1, x_2, \dots, x_5$ . These layers further undergo for analyze the actual amount of the hidden layers. Iteration by iteration, the hidden layers are identified. These hidden layers actually carries the data that has been separated during resource sharing process. These hidden layers further match the carrying data with sample data and if it matches, the layers change to  $l, m,$  and  $n$  in this case. The  $l, m,$  and  $n$  layers check the authenticity of the data and save off the ambient data in the form of the noises. Consequently, final data is transferred to the output layer  $n_o$ .

In next stage the unsupervised testing is carried out. In unsupervised testing, the data is not labeled and it appears as a mixture of heterogeneous raw table. Figure 2, illustrate the case where at stage one different data has been mixed and ready to transfer to another network. In stage 2, from the mixture of data, some data is labeled as  $a_1, a_2, a_3$  and  $a_4$  while other data is transferred. The labeled data is screen out from the hidden layers and the statistics of accuracy and the lost, from both supervised and unsupervised technique is placed in table 1 and table 2 respectively. The experiment is performed repeatedly to get all detected data. Based on 750 supervised samples and 250 unsupervised samples, Table 1 illustrates the accuracy values. Observe the final average value of accuracy outcomes after training each hidden layer neuron 25 times consecutively. The execution is terminated once it has run for 500 iterations.

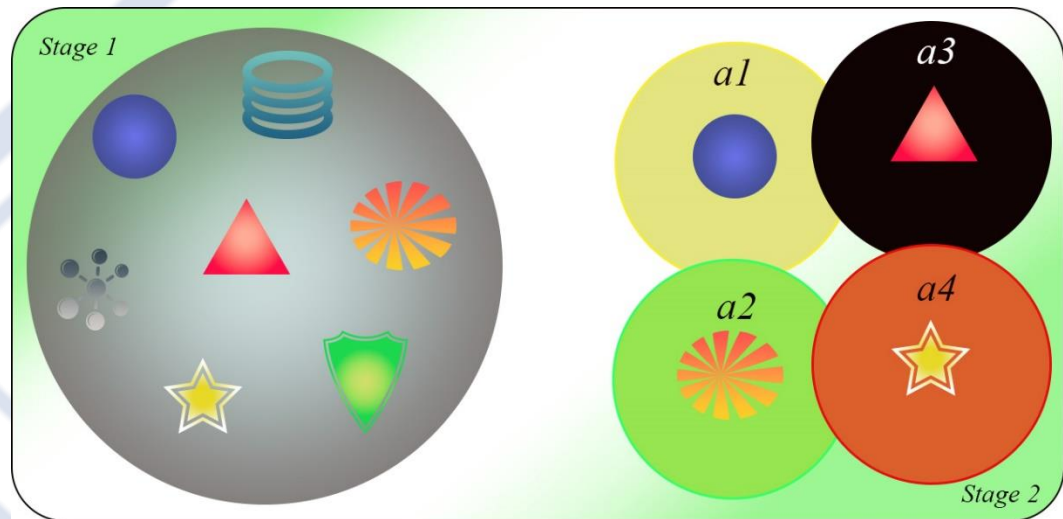


Figure 2. Data leakage detection through Unsupervised testing mechanism

As the number of neurons in the hidden layer increases, both supervised and unsupervised accuracy increase. As shown in table 2, both supervised and unsupervised neural networks are capable of predicting lost data. There is a reduction in loss values with an increase in the number of hidden neurons.

Table 1. Accuracy of leaked data for supervised and unsupervised testing

Hidden layers	Number of Iterations	Accuracy for supervised testing	Accuracy for unsupervised testing
1	30	86.76%	80.67%
2	60	87.62%	80.14%
3	90	88.54%	81.45%
4	120	89.43%	82.65%
5	150	90.67%	83.67%
6	180	90.87%	84.87%
7	210	91.69%	85.95%
8	240	92.78%	86.74%
9	270	92.54%	87.65%
10	300	93.65%	88.62%
11	330	94.75%	89.84%
12	360	94.86%	90.78%
13	390	94.97%	91.62%
14	420	95.46%	92.78%
15	450	95.54%	93.67%



Table 2. Loss of leaked data for supervised and unsupervised testing

Hidden layers	Number of Iterations	Loss supervised testing	Loss for unsupervised testing
1	30	0.0099	0.01017
2	60	0.0098	0.01016
3	90	0.0097	0.01015
4	120	0.0096	0.01014
5	150	0.0095	0.01013
6	180	0.0094	0.01012
7	210	0.0093	0.01011
8	240	0.0092	0.01010
9	270	0.0091	0.01009
10	300	0.0090	0.01008
11	330	0.0089	0.01007
12	360	0.0088	0.01006
13	390	0.0087	0.01005
14	420	0.0086	0.01004
15	450	0.0085	0.01003
16	500	0.0084	0.01002

## Performance discourse

After conducting experiments in Python [15], the targeted and the actual generated outcomes are analyzed. The amount of privacy leakage is directly proportional to the batch size and number of hidden layers. To improve the rate, one must raise the batch size and number of hidden layers.

Figure 3, depicts the curve for targeted and actual outcomes. Starting from 20 samples, the target and actual generated outcome is the same, however, both outcomes fluctuate up and down. At sample 70, the targeted detection was aimed at 38, and the same was obtained whereas at samples were reached 120, there was a difference between targeted and actual output. Here the target was 37 but received nearly 41%. Similarly, in sample 220, the target was 70% but the actual outcome was reached to 72%. In the end, the actual outcome left behind the targeted values which was 79% but achieved 92%.

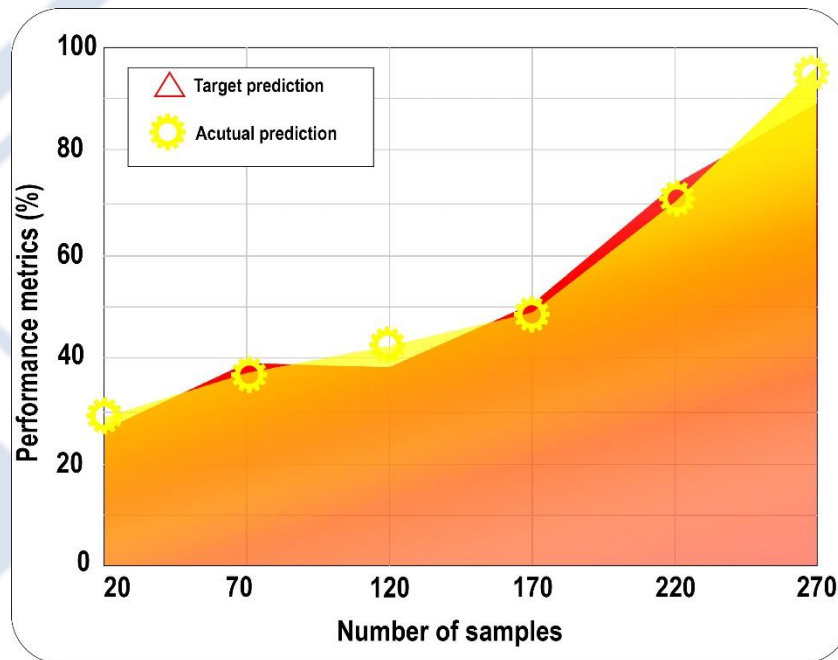


Figure 3. Target versus Actual leaked data analysis

Although the non-normalized dataset [16], may introduce some imperfections, these minute errors are minimized and do not significantly affect the effectiveness of the proposed prediction model. In addition, the proposed method achieved higher accuracy and prediction efficiency as well as a faster convergence rate. Consequently, the prediction model is very accurate. A supervised accuracy of 96.03% was achieved using the proposed method, while an unsupervised accuracy of 94.53% was achieved.

## Conclusion

The outcome of this study after analyzing the targeted and actual obtained data showed that using supervised and unsupervised technique to detected the amount of leaked data during data transfer in wireless network is a hallmark of shrewdness. Models proposed in this study have been shown to perform better in experiments. The proposed method achieved a supervised accuracy of 96.03%, while unsupervised accuracy is remained at 94.53%. Data leakage can be identified using this method with considerable ease.





## Contributor Roles

**Shahzad Ashraf:** [Conceptualización](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Escritura, revisión y edición](#). **Zeeshan Aslam:** [Análisis formal](#).

## Referencias

- [1] R. Zeeshan, and A. Muhammad, "Adopting proactive results by developing the Shrewd model of pandemic COVID-19," Arch. Community Med. Public Health, vol. 8, no. 2, pp. 062–067, Apr. 2022, doi: 10.17352/2455-5479.000175.
- [2] Z. Rasheed, S. Ashraf, N. A. Ibutoto, P. K. Butt, and E. H. Sadiq, "SDS: Scrumptious Dataflow Strategy for IoT Devices in Heterogeneous Network Environment," Smart Cities, vol. 5, no. 3, pp. 1115–1128, Sep. 2022, doi: 10.3390/smartcities5030056.
- [3] S. Ashraf, "Avoiding Vulnerabilities and Attacks with a Proactive Strategy for Web Applications," vol. 3, no. 2, p. 9.
- [4] Z. A. Arfeen, T. Ahmed, S. Ashraf, and S. Saleem, "Succulent link selection strategy for underwater sensor network," Int. J. Comput. Sci. Math., vol. 15, no. 3, p. 224, 2022, doi: 10.1504/IJCSM.2022.10049407.
- [5] M. Gao, Z. Chen, H. Naeem, and T. Ahmed, "CED-OR Based Opportunistic Routing Mechanism for Underwater Wireless Sensor Networks," Wirel. Pers. Commun., vol. 125, no. 1, pp. 487–511, Jul. 2022, doi: 10.1007/s11277-022-09561-w.
- [6] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential privacy in control and network systems," in 2016 IEEE 55th Conference on Decision and Control (CDC), Dec. 2016, pp. 4252–4272. doi: 10.1109/CDC.2016.7798915.
- [7] L. Ming, "A Deep Learning-Based Framework for Human Resource Recommendation," Wirel. Commun. Mob. Comput., vol. 2022, pp. 1–12, Jul. 2022, doi: 10.1155/2022/2377143.
- [8] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management," in



- 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Sapporo: IEEE, Jul. 2017, pp. 1–6. doi: 10.1109/SPAWC.2017.8227766.
- [9] A. Shahzad, "Towards Shrewd Object Visualization Mechanism," *Trends Comput. Sci. Inf. Technol.*, pp. 097–102, Nov. 2020, doi: 10.17352/tcsit.000030.
- [10] D. Muhammad, M. A. Khan, and T. Ahmed, "Fuzzy based efficient Cosmetology Paradigm," vol. 8, pp. 513–520, doi: 10.14741/ijmcr/v.8.4.3.
- [11] S. Saleem, and S. Afnan, "FTMCP: Fuzzy based Test Metrics for Cosmetology Paradigm," *Adv. Comput. Intell. Int. J. ACII*, vol. 4, no. 7, pp. 1–13, 2020, doi: 10.5121/acii.2020.7401.
- [12] S. Saleem, S. Ashraf, and M. K. Basit, "CMBA - A Candid Multi-Purpose Biometric Approach," *ICTACT J. Image Video Process.*, vol. 11, no. 01, p. 6, 2020, doi: 10.21917/ijivp.2020.0317.
- [13] T. Ahmed, "Sagacious Intrusion Detection Strategy in Sensor Network," in 2020 International Conference on UK-China Emerging Technologies (UCET), Glasgow, United Kingdom: IEEE, Aug. 2020, pp. 1–4. doi: 10.1109/UCET51115.2020.9205412.
- [14] S. Saleem, T. Ahmed, Z. Aslam, and M. Shuaeeb, "Iris and Foot based Sustainable Biometric Identification Approach," in 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Hvar, Croatia: IEEE, Sep. 2020, pp. 1–6. doi: 10.23919/SoftCOM50211.2020.9238333.
- [15] T. Ahmed, Z. Aslam, D. Muhammad, A. Yahya, and M. Shuaeeb, "Depuration based Efficient Coverage Mechanism for Wireless Sensor Network," *J. Electr. Comput. Eng. Innov. JECEI*, vol. 8, no. 2, pp. 145–160, 2020, doi: 10.22061/jecei.2020.6874.344.
- [16] A. Yahya et al., "Underwater routing protocols: Analysis of link selection challenges," *AIMS Electron. Electr. Eng.*, vol. 4, no. 3, pp. 234–248, 2020, doi: 10.3934/ElectrEng.2020.3.234.



# El impacto del Design Thinking en la tecnología

*The impact of Design Thinking on technology*

63

## Melvin Kevin López Asto

Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ t053300420@unitru.edu.pe

id <https://orcid.org/0009-0000-2532-3816>

## Jair Joel Vásquez Cerna

Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ t053300120@unitru.edu.pe


id <https://orcid.org/0000-0002-6182-5554>


## Alberto Mendoza de los Santos


Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ amendozad@unitru.edu.pe

id <https://orcid.org/0000-0002-0469-915X>

 **ARK:** <ark:/42411/s12/a94>

 **DOI:** [10.48168/innosoft.s12.a94](https://doi.org/10.48168/innosoft.s12.a94)

 **PURL:** [42411/s12/a94](https://purl.org/42411/s12/a94)

RECIBIDO 05/04/2023 • ACEPTADO 10/07/2023 • PUBLICADO 30/09/2023



## RESUMEN

En este artículo, exploramos el impacto del design thinking en la tecnología, tanto en el desarrollo de software como en el diseño de hardware. Los profesionales de la tecnología han adoptado este enfoque al diseñar interfaces de usuario, aplicaciones móviles y sitios web o plataformas web. Mediante la investigación y la empatía con los usuarios, se pueden identificar las fricciones y los puntos débiles en la experiencia del usuario y se pueden proponer soluciones centradas en el usuario que mejoren la usabilidad y la satisfacción general. En el desarrollo de software, el design thinking mejora la experiencia del usuario al crear aplicaciones y plataformas más intuitivas. Los resultados demostraron que la aplicación de este enfoque condujo a mejoras significativas en la usabilidad, la funcionalidad y la satisfacción general del usuario. Asimismo, en el diseño de hardware, el design thinking garantiza que los productos tecnológicos físicos sean estéticamente atractivos y funcionales. Los productos diseñados con este enfoque mostraron una combinación efectiva de estética y usabilidad, lo que resultó en una mayor aceptación por parte de los usuarios y una mejor experiencia general. La sinergia entre el design thinking y la tecnología impulsa la innovación centrada en el usuario, combinando la creatividad y el enfoque humano con las capacidades tecnológicas. Con iteración, experimentación y resolución de problemas complejos como elementos clave, el design thinking y la tecnología se complementan para crear soluciones tecnológicas significativas.

**Palabras claves:** Design Thinking, Innovación, Software, Hardware, Diseñador.



## ABSTRACT

*In this article, we explore the impact of design thinking in technology, both in software development and hardware design. Technology professionals have adopted this approach when designing user interfaces, mobile applications, and websites or web platforms. By researching and empathizing with users, frictions and pain points in the user experience can be identified and user-centric solutions can be proposed that improve usability and overall satisfaction. In software development, design thinking improves the user experience by creating more intuitive applications and platforms. Results showed that the application of this approach led to significant improvements in usability, functionality and overall user satisfaction. Similarly, in hardware design, design thinking ensures that physical technology products are aesthetically appealing and functional. Products designed with this approach showed an effective combination of aesthetics and usability, resulting in greater user acceptance and a better overall experience. The synergy between design thinking and technology drives user-centered innovation, combining creativity and human focus with technological capabilities. With iteration, experimentation and complex problem-solving as key elements, design thinking and technology complement each other to create meaningful technology solutions.*

**Keywords:** Design Thinking, Innovation, Software, Hardware, Designer.

## INTRODUCCIÓN

El avance tecnológico ha transformado nuestra sociedad en muchos aspectos, y el diseño de productos y servicios tecnológicos desempeña un papel fundamental en esta evolución. En este contexto, el design thinking ha surgido como un enfoque innovador que ha revolucionado la forma en que se concibe y desarrolla la tecnología, ha evolucionado a ser una metodología ampliamente utilizada en el desarrollo de productos y servicios innovadores. Su enfoque centrado en el usuario y su énfasis en la empatía, la colaboración y la experimentación han demostrado ser efectivos para abordar desafíos complejos. En el contexto de la tecnología, según [6, 13], el design thinking desempeña un papel crucial en el desarrollo de software y el diseño de hardware, generando un impacto significativo en la calidad y la experiencia del usuario.

### Design thinking

El design thinking es una metodología altamente centrada en el usuario que se enfoca en resolver problemas complejos y promover la innovación. Se trata de diseñar productos, servicios o experiencias que no solo resuelvan un problema, sino que también conecten emocionalmente con los usuarios y generen un impacto positivo en sus vidas. Según [3], se fundamenta en la empatía, la colaboración y la experimentación con el fin de comprender los deseos y las necesidades de los usuarios, generar ideas creativas y prototipos, y desarrollar soluciones prácticas y efectivas.





De manera general, podemos decir que el design thinking se presenta como un enfoque que combina la empatía, la colaboración y la experimentación para comprender y abordar las necesidades de los usuarios, fomentar la creatividad y crear soluciones prácticas y efectivas.

Las etapas del proceso del Design Thinking son:

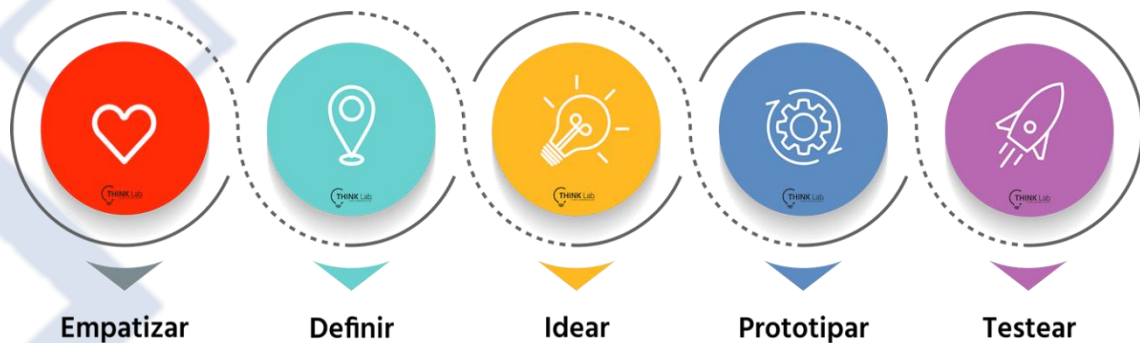


Figura 1: Procesos de Design Thinking. Fuente: think.ci

## **Empatizar:**

Centrarse en las condiciones y/o necesidades que presentan los usuarios para brindar soluciones reales. Según [6, 13], en esta etapa, el objetivo principal es sumergirse en la perspectiva del beneficiario que está directamente involucrado con la solución en desarrollo, a fin de generar resultados que se acerquen de manera única a su realidad.

## **Definir:**

Consiste en evaluar minuciosamente todos los problemas detectados en el anterior punto, con el propósito de plantear medidas para llegar a una solución definitiva. Según [6, 13], durante esta etapa, utilizando la información recopilada mediante la fase anterior, se busca de manera clara identificar el problema, lo cual es fundamental para poder identificar las soluciones que serán clave en el desarrollo de un proyecto innovador.

## **Idear:**

Se centra en la creatividad y el pensamiento libre, con el objetivo de realizar una lluvia de ideas que permita solucionar aquellos objetivos específicos previos. Según [6, 13], en esta etapa, el enfoque no se limita a una única idea, sino a cultivar un vasto océano de ideas que impida que la empresa se vuelva obsoleta. Se busca nutrir un amplio espectro de propuestas creativas y disruptivas que impulsen la innovación y aseguren que la organización se mantenga a la vanguardia.





## **Prototipar:**

Significa plasmar las ideas seleccionadas, ya sea de forma digital o físico. Según [6, 13], en esta etapa, cada idea obtenida se convierte en una realidad tangible. Se materializan modelos y prototipos de dichas ideas, permitiendo así su continua mejora y refinamiento antes de llegar al producto final.

## **Testear:**

Los usuarios se encargan de evaluar los errores y aciertos del prototipo elaborado. Según [6, 13], en esta etapa, se persigue activamente la capacidad de colocarse en la posición del beneficiario que está directamente involucrado con la solución en desarrollo. El objetivo primordial es generar resultados que reflejen una comprensión íntima y auténtica de su realidad, estableciendo así una conexión más profunda y significativa con sus necesidades y deseos.

## **Design thinking aplicado al desarrollo de software**

El design thinking representa el aspecto más desafiante del diseño de software, ya que se centra en la experiencia y sensaciones que el producto genera, en lugar de su apariencia visual [3]. En otras palabras, no es necesario enfocarse tanto en la parte estética del producto, sino hacer que el producto sea eficiente en cuanto a la función o funciones que debe desempeñar.

En el ámbito del diseño de soluciones de software, el enfoque de design thinking se presenta como una invaluable herramienta para comprender y satisfacer de manera efectiva los deseos y necesidades de los usuarios finales. Un ejemplo concreto de su aplicación radica en el diseño de la experiencia de usuario (UX), un elemento fundamental en el desarrollo de software. El enfoque de design thinking en lugar de basarse únicamente en las funcionalidades técnicas, este enfoque coloca al usuario como el epicentro del proceso de diseño. Esto permite crear interfaces de usuario que sean más que meros instrumentos, convirtiéndose en herramientas que potencian la productividad y brindan una experiencia satisfactoria.

El desarrollo de software implica un esfuerzo conjunto, complicado y creativo, donde la calidad del producto está estrechamente ligada a las personas, organizaciones y procesos utilizados para su creación y entrega [3]. Design thinking ha sido aplicado en el desarrollo de una gran variedad de software, desde aplicaciones móviles hasta soluciones empresariales. Aquí hay algunos ejemplos de software que se han mejorado a partir de la aplicación de design thinking:

## **Airbnb:**

La plataforma de alojamiento compartido utilizó el enfoque de design thinking para mejorar la experiencia de los usuarios en su sitio web y aplicación móvil. El resultado fue un diseño más



intuitivo y fácil de usar, que permitió a los usuarios encontrar y reservar alojamientos de manera más eficiente. Según [17], Airbnb es una plataforma en línea de economía colaborativa que actúa como intermediario entre personas que buscan alojamiento y anfitriones que ofrecen sus espacios a precios asequibles, brindando experiencias únicas y memorables para los huéspedes.



Figura 2: Logo de la plataforma web Airbnb. Fuente: stickpng.com

## **IBM:**

La empresa de tecnología utilizó el design thinking para desarrollar soluciones empresariales más centradas en el usuario. El resultado fue una mejor integración entre diferentes sistemas y una mayor eficiencia en los procesos empresariales.

Según [12], en el corazón de IBM Design Thinking hay un modelo de comportamiento para comprender de una mejor manera los deseos y necesidades de los usuarios y visualizar un futuro mejor: un ciclo continuo de observar, reflexionar y hacer.



Figura 3: Logo de la empresa tecnológica IBM. Fuente: freebiesupply.com

## **SAP:**

La empresa de software empresarial utilizó el enfoque de design thinking para desarrollar soluciones más intuitivas y fáciles de usar para sus clientes. El resultado fue una mejor comprensión de los procesos empresariales y una mejora en la eficiencia en el manejo de la información empresarial.

Según [7], SAP ha llevado este enfoque un paso más allá, aprovechando su profundo conocimiento del mercado y las necesidades cambiantes de sus clientes. Con una visión vanguardista, SAP ha diseñado y desarrollado productos innovadores que van más allá de las



expectativas convencionales. Su compromiso no se limita solo a ofrecer un producto, sino a proporcionar una experiencia completa y única para sus clientes.



Figura 4: Logo de la empresa de software empresarial SAP. Fuente: conectasoftware.com

## Design thinking en el diseño de hardware

El design thinking desempeña un papel fundamental en el diseño de hardware y productos tecnológicos físicos. Al aplicar los principios del design thinking, los diseñadores de hardware pueden crear productos que sean intuitivos, funcionales y satisfagan las necesidades de los usuarios.

Según [11], el Design Thinking aplicado al diseño de hardware innovador ha demostrado ser una metodología efectiva para mejorar la experiencia del usuario y la funcionalidad del producto. En su capítulo titulado 'Design Thinking for Innovative Hardware' en el libro Design Thinking for Innovation, los autores exploran cómo el enfoque centrado en el usuario y orientado a la innovación puede impulsar el desarrollo de hardware innovador. Los casos de estudio y ejemplos proporcionados destacan la importancia de identificar las necesidades del usuario, la generación de ideas creativas, la iteración y el prototipado, y la colaboración interdisciplinaria.

El enfoque de design thinking puede ayudar a los diseñadores a crear productos que ofrezcan una experiencia de usuario satisfactoria y que sean intuitivos y fáciles de usar. Aquí hay algunos ejemplos de hardware que se han mejorado a partir de la aplicación de design thinking:

### Nest Thermostat:

La empresa de termostatos inteligentes utilizó el enfoque de design thinking para crear un dispositivo fácil de usar y con una interfaz intuitiva para los usuarios. El resultado fue un termostato inteligente que se adapta automáticamente a las preferencias del usuario y reduce el consumo de energía. El Nest Thermostat es un ejemplo destacado de cómo el design thinking ha influido en su diseño y funcionalidad [16]. Al aplicar los principios del design thinking, los creadores del termostato Nest lograron crear un producto que se adapta a las necesidades y preferencias individuales de los usuarios, ofreciendo una experiencia intuitiva y personalizada.



Figura 5: Producto - Nest Thermostat. Fuente: nytimes.com

## Tesla Model S:

El Tesla Model S es un ejemplo destacado de cómo el design thinking ha influido en su diseño y desarrollo [5]. La empresa de automóviles eléctricos utilizó el enfoque de design thinking para diseñar un coche eléctrico que ofrecía una experiencia de conducción intuitiva y agradable para el usuario. El enfoque centrado en el usuario, la innovación tecnológica y el diseño aerodinámico del Model S son ejemplos claros de cómo el design thinking ha influido en su éxito.



Figura 6: Producto - Tesla Model S. Fuente: voitureblog.com

## OXO Good Grips:

OXO Good Grips es un ejemplo destacado de cómo el design thinking ha influido en el diseño de productos centrados en la ergonomía y la usabilidad [2]. Al aplicar los principios del design thinking, OXO ha logrado crear una línea de utensilios de cocina y herramientas que se adaptan de manera intuitiva a las necesidades de los usuarios, ofreciendo una experiencia cómoda y accesible, además permitían a las personas con discapacidades físicas disfrutar de la cocina.



Figura 7: Producto - OXO Good Grips. Fuente: amazon.com

## Sinergia entre el design thinking y la tecnología

Según [4], las técnicas de pensamiento de diseño se pueden utilizar para desarrollar y rediseñar no solo productos (en el sentido habitual, como producto o servicio), sino también, en principio, la tecnología de la actividad como tal (procesos). Después de todo, todos los participantes en el equipo del proyecto trabajan principalmente con la creación y reestructuración de procesos, diseñando la estructura organizativa óptima, etc. En el contexto de este estudio, el producto es un proceso comercial mejorado, el usuario del proceso comercial es el consumidor y cliente. El design thinking y la tecnología se complementan entre sí, creando una sinergia poderosa para impulsar la innovación y generar soluciones tecnológicas exitosas. Aquí están algunos aspectos clave de esta sinergia:

### Innovación centrada en el usuario:

El design thinking fomenta la generación de ideas innovadoras centradas en las necesidades de los usuarios. Al combinar este enfoque con la tecnología, se pueden desarrollar productos y servicios tecnológicos que realmente resuelvan problemas y mejoren la vida de las personas.

### Iteración y experimentación:

Tanto el design thinking como la tecnología valoran la iteración y la experimentación. La tecnología permite crear prototipos rápidos y realizar pruebas para validar ideas, mientras que el design thinking proporciona el enfoque humano y el marco para aprender de los resultados y mejorar continuamente.

### Diseño centrado en la experiencia del usuario:

La tecnología es más que solo funcionalidad; también se trata de brindar una experiencia satisfactoria al usuario. El design thinking se enfoca en comprender y abordar las necesidades y





emociones de los usuarios, lo que se traduce en productos tecnológicos más intuitivos, accesibles y agradables de usar.

### Resolución de problemas complejos:

La tecnología a menudo enfrenta desafíos complejos que requieren soluciones creativas. El design thinking proporciona un enfoque estructurado y colaborativo para abordar estos problemas, alentando la generación de ideas no convencionales y la exploración de diferentes perspectivas.

### Tabla de Resultados

Autor	Año	País	Título	Impacto generado de Design Thinking	Puntos de vista
E. V. Vasilieva and T. E. Tochilkina	2020	Rusia	Design thinking and process transformation: synergy of these approach (La sinergia entre el design thinking y la transformación de procesos)	El impacto tecnológico en relación Design Thinking y el enfoque centrado en el ser humano se manifiesta a través de las herramientas tecnológicas que permiten recopilar datos, analizar información, visualizar conceptos y mejorar la experiencia del cliente y del empleado.	La centralización del design thinking en el ser humano es una metodología eficaz para mejorar los procesos comerciales al priorizar la experiencia de los clientes y empleados, adaptarse a los cambios tecnológicos y comerciales, y enfocarse en el conocimiento intensivo. Al utilizar técnicas de design thinking, las empresas pueden desarrollar y rediseñar productos y procesos comerciales de manera efectiva.
Contreras Bello, Johan Sebastian	2021	Colombia	Design Thinking: metodología para el diseño de productos y servicios multimedia innovadores	El Design Thinking ha tenido un impacto inmejorable en la industria tecnológica, mejorando los procesos de innovación. Es una herramienta indispensable para generar productos y servicios multimedia innovadores, reorientar estrategias y fomentar la creatividad como impulsor de la innovación. Su enfoque centrado en el ser humano y la colaboración ha llevado a avances significativos	Al emplear el pensamiento visual y seguir la metodología del design thinking, es posible estructurar ideas y generar soluciones creativas orientadas a mejorar la experiencia del usuario en productos multimedia. Este enfoque se basa en identificar las necesidades de los usuarios, generar ideas y prototipos iterativos. En otras palabras, el design thinking ofrece un marco efectivo para fomentar



				en el desarrollo tecnológico y la experiencia del usuario.	la innovación en el ámbito de los productos y servicios multimedia.
Quistan Adams, George Alexander	2022	Perú	Propuesta de Sistema Web en la tienda Fierce Man usando la Metodología Design Thinking para la Gestión de Información y Toma de Decisiones	El impacto del Design Thinking en la Tienda Fierce Man se traduce en mejoras en la eficiencia operativa, competitividad e innovación. Al utilizar esta metodología centrada en el usuario, se logró optimizar los métodos de trabajo y procedimientos de la empresa, gracias a la interacción constante con los trabajadores. El enfoque creativo del Design Thinking permitió encontrar soluciones innovadoras adaptadas a las necesidades reales de los trabajadores y clientes. En otras palabras, se infiere que tuvo un impacto positivo en la empresa.	El design thinking es una metodología altamente efectiva para reconocer las necesidades de los usuarios en relación con el acceso, la organización y la presentación de información. Al utilizar este enfoque, se logra mejorar significativamente la eficiencia en la gestión de la información. Además, esta metodología ofrece un enfoque estructurado y creativo para abordar problemas complejos, lo cual resulta fundamental para desarrollar soluciones que satisfagan plenamente sus requerimientos, lo que a su vez eleva la calidad de las decisiones tomadas.
T. A. De Medeiros, J. A. T. M. Leite, P. S. Fernandes, V. Y. Suzuki y D. Nicodemo	2023	Brasil	Design Thinking e Técnica Delphi na elaboração de um livro digital educativo para unidades de alimentação e nutrição	El Design Thinking tuvo un impacto relevante en la creación del libro digital al poner enfoque en los usuarios, emplear prototipos iterativos, fomentar la colaboración y promover la generación de ideas creativas. Estos elementos fueron fundamentales para comprender a fondo las necesidades de los gestores de UAN, mejorar continuamente el contenido del libro a través de la retroalimentación de los usuarios, involucrar a un equipo multidisciplinario para aprovechar diferentes perspectivas y conocimientos, y desarrollar un libro digital que se adaptara efectivamente a las necesidades	El enfoque del design thinking ha sido fundamental en el desarrollo de este libro digital sobre prevención de quemaduras en UAN. Gracias a este enfoque, se pudo utilizar una metodología centrada en el usuario, incorporar ilustraciones educativas y utilizar un lenguaje accesible para los gestores de UAN, lo que facilita la comprensión y aplicación de las medidas de seguridad por parte del personal.



				de prevención de quemaduras en los entornos laborales de UAN.	
C. D. Corte et al.	2022	Brasil	Design Thinking: solução para o projeto de uma lixeira com o uso da fabricação digital	El Design Thinking tuvo un impacto significativo en el desarrollo del artículo, se promovió la colaboración de diferentes personas con perspectivas y habilidades diversas en la creación de una papelería. Además de fomentar un enfoque colaborativo y centrado en el usuario, se utilizaron técnicas de prototipado e iteraciones basadas en la retroalimentación de los usuarios para mejorar el diseño y la funcionalidad del producto. Además, se realizó una validación de ideas a través de pruebas y comentarios de los usuarios, lo que permitió ajustar y refinar la solución propuesta.	El design thinking juega un papel fundamental en este trabajo al proporcionar métodos y enfoques que permiten abordar de manera efectiva los procesos de diseño, desde la identificación de necesidades hasta la materialización de soluciones.
M. Vico C., S. De Francisco Vela y M. Pachón Buitrago	2021	Chile	Cuando el pensamiento en Diseño y el mundo legal se encuentran: Innovación para mejorar la prestación de servicios públicos	El enfoque del Design Thinking ha demostrado ser invaluable para el desarrollo de este artículo, al destacar la importancia de la originalidad y la autenticidad en áreas como el diseño para la justicia, diseño legal y diseño para políticas públicas. En este artículo, se presentan ejemplos concretos en los cuales el diseño ha generado un impacto positivo al mejorar la interacción y comprensión de los tribunales y cortes, al comunicar conceptos legales, al explorar temas relacionados con la salud y la educación, y al consolidar el diseño como una disciplina relevante y esencial.	El design thinking tiene un impacto significativo en el ámbito legal al permitir una aproximación centrada en el usuario y la generación de soluciones innovadoras y efectivas. Al fomentar la colaboración entre profesionales del diseño y del derecho, se pueden crear servicios públicos más accesibles, eficientes y adaptados a las necesidades de los usuarios.
R. M. Gómez Caballero	2021	Colombia	Innovación en Servicios de Salud usando herramientas de	El Design Thinking influyó en el desarrollo del artículo al proporcionar un enfoque metodológico para abordar la innovación en servicios de salud.	El design thinking impulsa la generación de soluciones innovadoras en el diseño de servicios de salud al fomentar la colaboración, la empatía y la



			<p>Pensamiento de Diseño</p>	<p>Mediante las etapas de descubrir, definir, desarrollar y entregar, se aplicaron técnicas de investigación en diseño y se realizó un análisis en profundidad de los intereses de los actores clave en el servicio de ortopedia y traumatología del Hospital Universitario de Santander ESE-HUS. Este enfoque permitió identificar los puntos de contacto clave y co-crear el concepto del servicio, así como definir las oportunidades y requerimientos para mejorarlo.</p>	<p>comprensión profunda de las necesidades de los actores involucrados. Esto permite mejorar la comunicación, la gestión de la innovación y la entrega de servicios que satisfacen las necesidades reales de los pacientes y generan valor en el sistema de salud.</p>
<p>Rodríguez Macías, Víctor Manuel</p>	<p>2019</p>	<p>México</p>	<p>El Design Thinking en la construcción de experiencia de marca: caso específico Airbnb México</p>	<p>El enfoque del Design Thinking ha desempeñado un papel fundamental en el éxito de Airbnb México al proporcionar experiencias completas a través de canales digitales, como su plataforma en línea y aplicación móvil. Al poner énfasis en comprender las necesidades y emociones de los usuarios, Airbnb México ha logrado ofrecer soluciones innovadoras y personalizadas, lo que les ha otorgado una posición ventajosa en el mercado. Al adoptar un enfoque centrado en las personas y aprovechar técnicas colaborativas, la empresa ha experimentado impactos positivos en la toma de decisiones estratégicas.</p>	<p>El uso del Design Thinking ha sido un elemento clave en el éxito y la diferenciación de Airbnb México en el entorno digital. Su enfoque centrado en el usuario y la capacidad para generar soluciones innovadoras les ha permitido adaptarse eficazmente a las demandas cambiantes del mercado y ofrecer una experiencia superior a sus clientes.</p>
<p>M. Villius Zetterholm y P. Jokela</p>	<p>2023</p>	<p>Suiza</p>	<p>Addressing Complexity in the Pandemic Context: How Systems Thinking Can Facilitate Understanding of Design Aspects for</p>	<p>El Design Thinking generó un impacto significativo al proporcionar un enfoque metodológico estructurado para abordar el diseño de tecnologías preventivas en el contexto de la pandemia de COVID-19. Al utilizar el pensamiento sistémico y el modelo DSRP, se logró comprender las complejidades y</p>	<p>El design thinking juega un papel crucial al organizar nuestra comprensión de situaciones complejas, como una pandemia, permitiendo identificar puntos de influencia para el diseño de soluciones preventivas efectivas y adaptadas a las necesidades de los usuarios. Al adoptar un</p>



			Preventive Technologies	conexiones subyacentes del dominio de la pandemia, lo cual permitió explorar diferentes enfoques de diseño y analizar cómo se relacionan con la complejidad del sistema, los desafíos de implementación y la comprensión del usuario. El Design Thinking fue fundamental para identificar puntos de apalancamiento y considerar los requisitos contextuales de las tecnologías preventivas, brindando una guía clara para el diseño efectivo en el contexto de las epidemias y pandemias.	enfoque centrado en el usuario y utilizar herramientas de pensamiento sistémico, se pueden encontrar soluciones innovadoras y adaptadas a las necesidades de la sociedad en tiempos de crisis sanitaria.
--	--	--	-------------------------	---	--

## Discusión de Resultados

El Design Thinking ha demostrado ser una metodología efectiva y versátil que ha generado impacto en varios ámbitos:

- En la industria tecnológica, el design thinking ha tenido un impacto inmejorable al mejorar los procesos de innovación, generando productos y servicios multimedia innovadores, reorientando estrategias y fomentando la creatividad como motor de la innovación. Su enfoque centrado en el ser humano y la colaboración han impulsado avances significativos en el desarrollo tecnológico y la experiencia del usuario, generando mejoras tangibles en la industria.
- En la tienda Fierce Man y en la compañía Airbnb México, el design thinking ha generado un impacto significativo tanto en Airbnb México como en la Tienda Fierce Man. En el caso de Airbnb México, ha sido fundamental para su éxito y diferenciación en el entorno digital, permitiéndoles adaptarse a las demandas cambiantes del mercado y proporcionar una experiencia superior a sus clientes. Por otro lado, en la Tienda Fierce Man, ha generado mejoras en la eficiencia operativa, competitividad e innovación al optimizar los procedimientos de la empresa y métodos de trabajo. La metodología ha permitido implementar soluciones innovadoras adaptadas a las necesidades reales de los trabajadores y clientes, generando un impacto positivo en ambas organizaciones.
- En la creación de un libro digital y en el desarrollo de una papelería, el design thinking desempeñó un papel relevante al utilizar prototipos iterativos, fomentar la colaboración y estimular la generación de ideas creativas. Estos elementos resultaron fundamentales





para comprender las necesidades de los gestores de una manera profunda y permitieron mejorar de manera continua la funcionalidad y el diseño del producto a través de la retroalimentación de los usuarios. Gracias a este enfoque se logró desarrollar un libro digital y una papelería innovadora y funcional que se adaptó eficazmente a las necesidades reales de los usuarios.

- En el diseño legal y políticas públicas, el design thinking ha tenido un impacto significativo al resaltar la importancia de la colaboración y la interdisciplinariedad en estas áreas. Se han mostrado casos en los que el diseño ha mejorado la interacción y comprensión en cortes y tribunales, la comunicación de temas legales, la exploración de asuntos de salud y educación, y ha fortalecido su relevancia y necesidad como disciplina.
- En la innovación en servicios de salud, el design thinking ha sido una herramienta necesaria en el campo de la innovación en servicios de salud, al brindar un enfoque sistemático y creativo para identificar y abordar oportunidades de mejora. Al utilizar técnicas de investigación y análisis en profundidad, se logra una comprensión de los desafíos más profunda y se generan soluciones más efectivas para satisfacer las necesidades de los usuarios en el ámbito de la salud.

## Conclusiones

El design thinking tiene un impacto importante en el ámbito tecnológico, abarcando tanto el desarrollo de software como el diseño de hardware. Al adoptar un enfoque centrado en el usuario, esta metodología permite a los equipos de desarrollo comprender a fondo las necesidades y deseos de los usuarios. Esto, a su vez, conduce a la generación de soluciones tecnológicas más efectivas y satisfactorias.

Algunas conclusiones clave son:

- En el desarrollo de software, el design thinking ayuda a crear aplicaciones y plataformas más intuitivas, con una mejor experiencia de usuario. El enfoque en la comprensión, observación, ideación y prototipado permite a los equipos ajustar el software a las necesidades y requerimientos reales de los usuarios.
- En el diseño de hardware, el design thinking asegura que los productos tecnológicos físicos sean estéticamente atractivos y funcionales. La empatía y la iteración desempeñan un papel importante para crear productos que se ajusten a los deseos y necesidades de los usuarios.
- La sinergia entre el design thinking y la tecnología fomenta la innovación centrada en el usuario. Al combinar la creatividad y el enfoque humano del design thinking con las capacidades tecnológicas, se pueden desarrollar soluciones tecnológicas que realmente aborden los problemas y mejoren la vida de las personas.



## Contribución de Autoría

**Jair Joel Vásquez Cerna:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Redacción - borrador original](#). **Melvin Kevin López Asto:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Redacción - borrador original](#). **Alberto Carlos Mendoza de los Santos:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Redacción - borrador original](#).

## Referencias

- [1] C. D. Corte et al., "Design Thinking: Solução para o projeto de uma lixeira com o uso da fabricação digital", Revista de Arquitetura IMED, vol. 11, n.º 1, p. 24, octubre de 2022.
- [2] D. A. Norman and C. Röcker, "Design for everyday things: Revised and expanded edition." Basic Books, 2016.
- [3] E. Vásquez, J. Carlo, E. Zapata, and E. Enrique, "Marco de trabajo en base a Design Thinking y metodologías ágiles de desarrollo de software," Universidad Peruana de Ciencias Aplicadas (UPC), 2017.
- [4] E. V. Vasilieva and T. E. Tochilkina, "Design thinking and process transformation: synergy of these approaches", 2020.
- [5] F. Maurer and S. Carvalho, "Design thinking as a strategy for innovation: The case of Tesla Motors," International Journal of Innovation Management, vol. 20, no. 7, art. no. 1650044, 2016.
- [6] G. A. Quistan Adams, "Propuesta de Sistema Web en la tienda Fierce Man usando la Metodología Design Thinking para la Gestión de Información y Toma de Decisiones," Universidad Tecnológica del Perú (UTP), 2022. [En línea]. Disponible en: <https://hdl.handle.net/20.500.12867/5482>
- [7] J. S. Contreras Bello, "Design Thinking: metodología para el diseño de productos y servicios multimedia innovadores", 2020. [En línea]. Disponible en: <http://hdl.handle.net/10654/38632>



- [8] M. Vico C., S. De Francisco Vela y M. Pachón Buitrago, "Cuando el pensamiento en Diseño y el mundo legal se encuentran: Innovación para mejorar la prestación de servicios públicos", *RChD: Creación y pensamiento*, vol. 6, n.º 11, p. 1, noviembre de 2021.
- [9] M. Villius Zetterholm y P. Jokela, "Addressing complexity in the pandemic context: How systems thinking can facilitate understanding of design aspects for preventive technologies", *Informatics*, vol. 10, n.º 1, p. 7, enero de 2023.
- [10] R. M. Gómez Caballero, "Innovación en servicios de salud usando herramientas de pensamiento de diseño", *Kepes*, vol. 18, n.º 23, pp. 15-48, enero de 2021.
- [11] R. Roth and R. Roy, "Design Thinking for Innovative Hardware," in *Design Thinking for Innovation*. Springer, 2018, pp. 235-256.
- [12] S. Hamm, "IBM Design Thinking. Field Guide", 2016.
- [13] T. A. De Medeiros, J. A. T. M. Leite, P. S. Fernandes, V. Y. Suzuki y D. Nicodemo, "Design Thinking e Técnica Delphi na elaboração de um livro digital educativo para unidades de alimentação e nutrição", *Brazilian J. Develop.*, vol. 9, n.º 05, pp. 16407-16423, mayo de 2023.
- [14] T. Brown, "Change by Design. How design thinking transforms organizations and inspires innovation. *The Power of Design Thinking*", 6.
- [15] T. Brown, "Design Thinking," *Harvard Business Review*, vol. 86, no. 6, pp. 84-92, Jun. 2008.
- [16] T. Fadell and R. Erickson, "Reinventing the thermostat: Designing Nest," *ACM Interactions*, vol. 19, no. 1, pp. 52-57, 2012.
- [17] V. M. Rodríguez Macias, "El Design Thinking en la construcción de experiencia de marca: caso específico Airbnb México", 2019. [En línea]. Disponible en: [https://ru.dgb.unam.mx/handle/DGB\\_UNAM/TES01000790712](https://ru.dgb.unam.mx/handle/DGB_UNAM/TES01000790712)



# Reconocimiento y Clasificación de Mensajes de Odio

## *Recognition and Classification of Hate Messages*

79

### **Patrick Leopoldo Paredes Neira**

Universidad La Salle. Arequipa, Perú.

@ pparedesn@ulasalle.edu.pe

### **Gary Jamil Vilca Tapia**


Universidad La Salle. Arequipa, Perú.


@ gvilcat@ulasalle.edu.pe


### **Kristhyan Andree Kurt Lazarte Zubia**

Universidad La Salle. Arequipa, Perú.

@ klazartez@ulasalle.edu.pe

 **ARK:** [ark:/42411/s12/a99](https://nbn-resolving.org/ark:/42411/s12/a99)

 **DOI:** [10.48168/innosoft.s12.a99](https://doi.org/10.48168/innosoft.s12.a99)

 **PURL:** [42411/s12/a99](https://nbn-resolving.org/ark:/42411/s12/a99)

RECIBIDO 16/04/2023 • ACEPTADO 15/07/2023 • PUBLICADO 30/09/2023



## RESUMEN

El uso masivo de las redes sociales y el anonimato que este brinda ha posibilitado no solamente la comunicación inmediata entre los usuarios, sino también que acrezca la difusión del discurso de odio contra ciertos grupos de nuestra sociedad en forma de mensajes ofensivos para ellos, esto ha desembocado en un grave problema social; el cual sigue siendo tema de investigación actual junto con NLP. El propósito del presente trabajo es hacer una comparación de nuestro modelo de reconocimiento de "HateCheck" contra los resultados del autor, utilizando la misma base de datos que ellos. Para ello haremos uso de las principales métricas como son: precisión, recall y F1.

**Palabras claves:** Odio, HateCheck, Modelos de Detección del Discurso de Odio, NLP, Procesamiento del Lenguaje Natural, Discurso de Odio, Procesamiento del Español.

## ABSTRACT

*The massive use of social networks and the anonymity that this provides has made possible not only the immediate communication between users, but also the spread of hate speech against certain groups of our society in the form of offensive messages to them, this has led to a serious social problem, which remains a topic of current research along with NLP. The purpose of the present work is to make a comparison of our "HateCheck" recognition model against the author's*



*results, using the same database as them. To do so, we will make use of the main metrics such as: precision, recall and F1.*

**Keywords:** *Hate, HateCheck, Hate Speech Detection Models, NLP, Natural Language Processing, Hate Speech, Spanish Processing.*

---

## MOTIVACIÓN

### A. ¿En qué dominio del conocimiento está trabajando?

Basándonos en nuestro problema propuesto y nuestro objetivo principal, podemos concluir que el dominio que estamos abarcando es el de la Sociología, debido a que abordaremos el estudio de una muestra de la sociedad como un conjunto de relaciones culturales y los mensajes que estos escriben para su posterior clasificación.

### B. ¿Quiénes son los usuarios objetivo?

En general, cualquier persona que pueda acceder a un medio de comunicación a través de la internet, por ejemplo: foros, redes sociales, etc. Debido a que a través de toda la información recopilada podremos clasificar de manera más eficiente los mensajes que inciten al odio. Para ello haremos una amplia búsqueda de diferentes bases de datos que se detallarán posteriormente.

### C. ¿Por qué es interesante el tema que proponen?

Tenemos que tener en cuenta que la libertad de expresión es uno de los derechos que todos poseemos por el simple hecho de que somos seres humanos. Pero nuestros derechos acaban cuando se atenta contra los de otro individuo, y de todas las formas de comunicación que existen en la actualidad, la más popular son los mensajes de texto, los cuales los usamos en todo momento (documentos, foros, redes sociales, correos electrónicos, etc.), es por ello que resulta agotador pensar en toda la información que se tiene que procesar por una persona, en lugar de un ordenador de forma sistemática.

Además, luego de clasificar estos mensajes se podrá ver quiénes fueron los usuarios que tienen más inclinación a esos pensamientos para su posterior tratamiento. Incluso, saber cuáles son los temas que provocan una sensación de odio en las personas y su reducción en los medios de comunicación.

### D. ¿Cuáles son las preguntas que su proyecto de NLP intenta responder?

Nuestro proyecto intenta responder sobre cuán viable sería implementar un modelo de reconocimiento de frases de odio en español para su posterior clasificación, así como obtener de





forma cuantitativa la precisión de este, teniendo en cuenta que normalmente los modelos son entrenados para el reconocimiento del inglés y que los investigadores de NLP tanto en inglés como en otros idiomas suelen tener como lengua materna el inglés.

- ¿Para qué me serviría usar esta herramienta de HateCheck?
- ¿Se puede mejorar el sesgo de rendimiento con respecto a otras herramientas?

## Problema

Definimos a los mensajes de odio o a la incitación al odio como un abuso dirigido contra los miembros que conforman un determinado grupo, estos están basados en la edad, discapacidad, identidad de género, la raza, los orígenes nacionales o étnicos, la religión, la orientación sexual (haciendo foco a estos últimos), lo que puede generar una serie de discrepancias de ideología con personas que no comparten esas idas o no pertenecen a dicho grupo. Este problema está íntimamente relacionado a la discriminación racial.

Es por tales motivos y basándonos en estas definiciones, enfocamos la detección de la incitación al odio como la clasificación binaria de contenidos como odiosos o no odiosos.

## Objetivo

El presente trabajo tiene como finalidad el poder utilizar una Inteligencia Artificial y entrenarla con un dataset de diferentes frases de odio en español, categorizados según el grupo social al cual va dirigido para que esta pueda reconocer, discernir y clasificar si un texto es de odio. Para ello tendremos 7 categorías: personas negras, homosexuales, indígenas, judías, discapacitadas, transexuales y mujeres.

Asimismo, nuestra implementación será comparada con los resultados del equipo de investigación del artículo "HateCheck: Functional Test for Hate Speech Detection Models" bajo las 3 principales métricas conocidas: precisión, recall y F1 [1].

## Datos

### A. ¿Qué datos necesitará?

Para la realización del presente proyecto se ha visto necesario tener un dataset el cual no solamente contenga los mensajes de odio, sino que también debe estar clasificado en 7 categorías: personas negras, homosexuales, indígenas, judías, discapacitadas, transexuales y mujeres.



## B. ¿Cómo recolectarán los datos?

Debido a las limitaciones de nuestros recursos, optaremos por usar una data en español que sea de uso libre cuyos autores pertenezcan a una organización o universidad de prestigio.

## C. ¿Dónde planea obtenerlos?

La búsqueda lo haremos en los principales repositorios de bases de datos como: Kaggle, Hugging Face, Google, UCI Machine Learning Repository y GitHub. Se referenciará apropiadamente a los autores.

## D. ¿Cómo planea almacenarlos?

Los datos serán llevados a un archivo de Excel con extensión .xls para que puedan ser fácilmente visualizados y posteriormente manipulados.

## E. ¿Cómo accederá a ellos para utilizarlos en su proyecto?

Ya que trabajaremos con Python y el notebook Jupyter (Google Colab), usaremos la librería Pandas y Numpy que nos permitirá la lectura de la data y su obtención en una estructura de datos como diccionarios y matrices, para su posterior procesamiento.

## Diseño

### A. ¿Cómo sugiere usar la herramienta para ayudar a los usuarios a realizar las tareas que respondan a las preguntas que enumeró en la motivación?

El usuario deberá cargar un archivo Excel con extensión .xls o en su defecto deberá tipear los mensajes que desea clasificar, luego nuestro "HateCheck" le asignará su respectiva categoría para finalmente devolver al usuario otro archivo .xls.

Al usar esta herramienta poder comprender los abusos más comunes descritos en la industria de la tipografía también poder tener diagnósticos más específicos con "HateCheck" ya que cuenta con pruebas funcionales, y también es ampliamente usada en la detección de discursos del odio [3], a diferencia de Bert que clasifica erróneamente los comentarios relacionados a mujeres más que en otras categorías.

El sesgo de rendimiento del "HateCheck" podemos mejorarlo de manera significativa evitando el uso de palabras soeces en expresiones exclamativas. También podemos mejorarla evitando usar negaciones antes de ingresar cada mensaje [2].



## Trabajos relacionados

- Strong negative emotions explicitly expressed about a protected group or its members: Resembles "expressed hatred" [5]
- Sobre el lenguaje ofensivo Palmer et al. (2020) [7] recopilan tres datasets para evaluar el rendimiento del modelo en lo que denominan lenguaje ofensivo complejo, donde se utilizan los insultos, adjetivos y distanciamiento lingüístico.

## Revisión Literaria

En este presente trabajo se centrará específicamente en los problemas de clasificación de texto que contienen lenguaje ofensivo. Durante el transcurso del proyecto encontramos problemas que afectan la precisión y la eficacia del modelo que estamos desarrollando, dentro de los cuales nos encontramos con: el dominio del idioma, las métricas comparativas de modelos y la clasificación que reciben algunas de las palabras que solo son soeces mas no ofensivas.

Para poder identificar los enfoques que abordan algunos de estos problemas se requirió de la revisión de artículos que se encuentran disponibles en la página de la IEEE, además de ello daremos una breve descripción de cada enfoque realizado por diferentes trabajos.

Para nuestro primer enfoque dará solución al problema de la eficiencia del modelo por ello usaremos un enfoque basado en pruebas, la necesidad de este enfoque es porque los modelos de detección tienen debilidades que no se pueden observar por medio de las métricas. El beneficio de usar este modelo para nuestra investigación es que cuenta con bastantes pruebas funcionales que nos ayudarán a clasificar de mejor manera la información, este enfoque se basa principalmente en el reconocimiento de características que presentan gran dependencia a muchas palabras específicas. [9].

En segundo lugar, tenemos el enfoque que nos ayuda a tener un contexto claro al momento de hacer la clasificación de mensajes que inciten al odio. Para abordar este desafío se ha propuesto agregar tareas adicionales al corpus de entrenamiento, como la predicción binaria y la multitiqueta. La necesidad de este enfoque surgió ya que dentro de las redes sociales se usa el sarcasmo. Para nuestro estudio esto representa un factor clave al momento de procesar información [14].

En tercer lugar, el idioma representa un factor importante al momento de hacer la clasificación de información, esto se debe a que tenemos limitaciones al momento de hacer un análisis de comentarios en una red social famosa como por ejemplo twitter. La necesidad de este enfoque es debido a que nuestro modelo solamente está basado en el idioma inglés, en vista a esta necesidad se desarrolla el multilingual hatecheck que cuenta con idiomas variados y mejoras en sus pruebas funcionales. Pero esta herramienta cuenta con errores que serán cubiertos en su



mayoría, errores como el corpus del modelo al contener tantos idiomas disminuye la certeza de sus resultados [15].

Finalmente encontramos un último enfoque que se basa en mezclar varias técnicas para mejorar la precisión de los datos de salida del modelo. La necesidad de este enfoque es cuando tenemos como un conjunto de entrada datos que provienen de redes sociales entonces lo que se espera es contar con un corpus grande. Las técnicas encontradas son el bigrama que usa de representación de características BOW, después encontramos la técnica de los diccionarios que están basados en N-gramas y finalmente nos encontramos con diccionarios que nos sirvan para generar características maestras. Para concluir hemos podido encontrar a través de la literatura un enfoque de aprendizaje por transferencia basados en modelos preentrenados de tipo BERT que utilizan las técnicas anteriormente mencionadas [16].

La elección de los enfoques depende del contexto en el cuál vamos a querer usar el modelo. Si los datos de entrada fueran de redes sociales usaríamos la herramienta multilingual, en caso quisiéramos contar con una mayor precisión usaríamos el enfoque basado en técnicas.

Para finalizar dirémos que en todo modelo hemos encontrado muchos puntos débiles es por ello que los enfoques nos ayudan a fortalecer esos puntos débiles. Como se comentó al inicio, existen problemas que no podemos darles solución como es el caso de comparar dos modelos, ya que las métricas no nos ayudan a compararlos. Se puede concluir que para realizar el trabajo debemos contar con un corpus lo suficientemente grande para realizar una clasificación más precisa.

## Diseño

### A. Propuesta

Para este trabajo se utilizó un dataset de frases relacionadas con mensajes de odio que están previamente categorizadas en 7 tipos dependiendo al grupo social al que se refiere: homosexuales, personas de color, transexuales, indígenas, judíos, discapacitados, mujeres y también está la categoría de que no es un mensaje de odio. Como ventaja de este dataset es que ya se le realizó preprocesamiento y limpieza de emojis, pero mantiene las formas alternativas de palabras de odio que reemplazan ciertas letras por números o que las palabras están separadas por espacios.

Con respecto a los pasos de tokenización, tags y lemas se utilizará SpaCy junto con el modelo preentrenado `es_dep_news_trf`

Sobre el proceso de reconocimiento de entidades con nombre (Named Entity Recognition, NER) también se utilizará SpaCy junto con su modelo preentrenado con esta tarea, en caso de que no se obtenga buenos resultados se optará por el uso de otras alternativas como BERT.



Finalmente, para la clasificación usaremos BERT y afinaremos sus parámetros y usaremos etiquetas para que realice un entrenamiento supervisado a fin de obtener buenos resultados.

## B. Pipeline

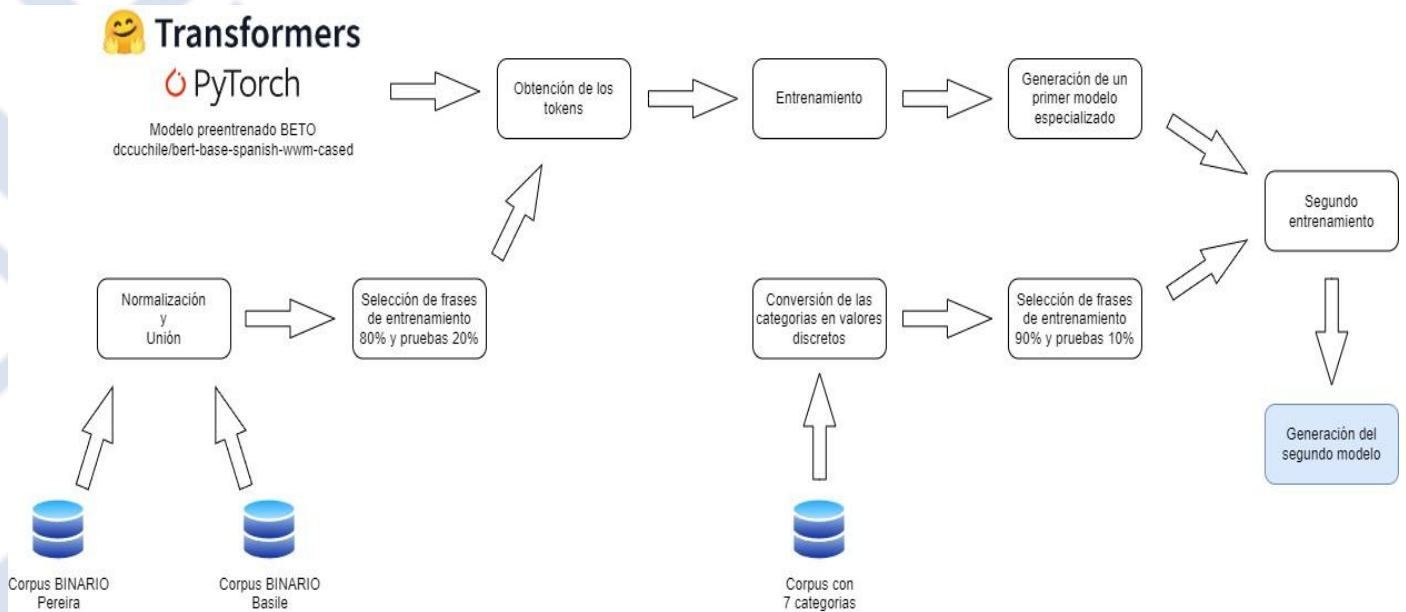


Fig. 1: Pipeline del proceso de clasificación de mensajes de odio

## C. Principios básicos de diseño que nos guió para crearlo

Los principios de procesamiento que están presentes son los principios estándares para el reconocimiento de texto antes de ser procesados, de forma que generamos las características del modelo BERT. Las reglas aplicadas al corpus son las siguientes:

- Pasar todo el texto a minúsculas.
- Remover URLs contenidas en el texto.
- Remover emojis que contenga el dataset.
- Remover puntuación.
- Remover saltos de línea innecesarios.

La selección de las características y la parametrización de los modelos se realizó de acuerdo a la revisión de la literatura. Se procedió a realizar una comparativa de la información de los estudios revisado. En ello se identificó la utilización del modelo SVM con características de tipo unigrama





para establecer el punto de referencia con los demás modelos en comparación. La siguiente tabla contempla los modelos utilizados y los mejores resultados obtenidos según la literatura.

TABLE I: Comparación de métodos

Estudio	(Paul, et al, 2020)	(Esmá, et al, 2022)
Modelo	Modelo Propuesto%	CNN static y word2vec
Precisión	0.6443	Sin información
Recall	0.9628	Sin información
F1 score	- -	89.6
Resultados	Basados en unigramas	CNN simple de una capa

## D. Fortalezas y debilidades del diseño

**1) Fortalezas:** Nuestro diseño ofrece varias ventajas significativas para el procesamiento de lenguaje natural (NLP) en el reconocimiento de mensajes de odio. Aquí se mencionan algunas:

- Contextualización bidireccional:

Nuestro diseño captura el contexto y la relación entre las palabras en una oración de manera bidireccional, gracias a BERT. Esto significa que comprende tanto las palabras anteriores como las posteriores en un texto, lo que mejora la comprensión del significado y la intención detrás de las palabras clave relacionadas con el odio. Esta contextualización mejorada ayuda a identificar mensajes de odio más precisamente.

- Representaciones lingüísticas variadas:

Nuestro diseño será entrenado en un dataset ya clasificado, lo que le permite aprender patrones y características del lenguaje de manera objetiva. Como resultado, el diseño desarrolla representaciones lingüísticas ricas que capturan la semántica y el significado de las palabras, con un previo análisis morfológico. Esto es beneficioso para el reconocimiento de mensajes de odio, ya que puede captar sutilezas en el lenguaje que indican contenido ofensivo.

- Adaptabilidad a diferentes idiomas:

El diseño será entrenado y aplicado al idioma español, pero tenemos datasets de diferentes idiomas lo que lo hace adaptable a diferentes lenguajes para su reconocimiento y clasificación. Esto es especialmente valioso para el reconocimiento de mensajes de odio, ya que el contenido ofensivo puede estar presente en diferentes idiomas y culturas.

Estas ventajas hacen que nuestra propuesta sea una opción versátil para el procesamiento de lenguaje natural en el reconocimiento de mensajes de odio, aunque es importante considerar que ningún modelo es perfecto y siempre se deben tener en cuenta las limitaciones y los posibles sesgos inherentes en los datos y el entrenamiento.



## 2) Debilidades:

- **Manejo de texto estructurado:** Los mensajes de odio a menudo se presentan en forma de texto no estructurado, como comentarios en redes sociales o publicaciones en foros. Nuestro planteamiento no es efectivo en el procesamiento de este tipo de texto, ya que a pesar de capturar el contexto y el significado gramatical más informal, esto se ve afectada por su capacidad en relación al testing y entrenamiento ya que será a partir de un dataset estructurado y previamente clasificado.
- **Requerimientos computacionales y de recursos:** Es un modelo de lenguaje profundo y complejo que requiere una cantidad significativa de recursos computacionales y memoria para su entrenamiento y aplicación (BERT). Esto dificulta su implementación en entornos con recursos limitados, lo que limita su accesibilidad y uso en algunos casos.
- **Tamaño del modelo:** Dependiendo del tamaño del modelo y del data set que usamos, principalmente con una relación de 10 000 palabras puede presentar un desafío para el modelo de reconocimiento y clasificación de mensajes de odio.
- **Dependencia de datos etiquetados:** Aunque BERT se puede adaptar a tareas específicas con datos etiquetados limitados, aún requiere una cantidad significativa de datos anotados para un mejor desempeño. La recopilación y el etiquetado de grandes conjuntos de datos pueden ser costosos y consume tiempo, especialmente cuando se trata de mensajes de odio, que a menudo requieren una revisión manual exhaustiva.

Es importante tener en cuenta estas desventajas y considerarlas en el desarrollo y la implementación de modelos de procesamiento de lenguaje natural para el reconocimiento de mensajes de odio. Además, es crucial realizar una evaluación constante.

## Código

Ahora, se mostrará el paso a paso para implementar este sistema de reconocimiento y clasificación de mensajes de odio en español:

### A. Importamos las bibliotecas y habilitamos el uso de CUDA

Para empezar con la implementación tenemos que importar `tensorflow` y como usamos la GPU para acelerar el proceso también debemos de habilitar el uso de CUDA, se obtiene el nombre de la GPU. Limitamos el uso de la VRAM, debemos de tener en cuenta de que la GPU sea compatible con CUDA y si no tenemos alguna se utilizará la CPU.



```
import tensorflow as tf

# Obtener el nombre del dispositivo GPU. device_name =
tf.test.gpu_device_name()

# El nombre del dispositivo de GPU debe ser similar al
, → siguiente:
if device_name == '/device:GPU:0':
    print('Found GPU at: ', device_name)
else: raise SystemError('GPU device not found')
```

Fig. 2: Código Python

```
import torch if
torch.cuda.is_available():
    # Apuntamos a la primera GPU CUDA
    #torch.cuda.set_device(torch.cuda.device_count() - 1) device =
    torch.device("cuda") print('# GPUs: ',
    torch.cuda.device_count()) print('Nombre de la GPU:',
    , → torch.cuda.get_device_name(0))

else:
    print('Se utilizará la CPU') device =
    torch.device("CPU")

model_str = "dccuchile/bert-base-spanish-wwm-cased"
```

Fig. 3: Código Python

## B. Leemos el corpus

Para leer el corpus de datos usamos pandas y se lee el siguiente archivo. CORPUS DE DATOS Posteriormente se realiza la tokenización; es decir, se le agrega un token específico a cada palabra para poder analizarla.

```
import pandas as pd df=
pd.read_excel('corpus_bin.xlsx')
df.groupby("HS").size() #1 -> Hate speech
```

Fig. 4: Código Python

```
train_df = df.sample(frac=0.8, random_state=2023) test_df =
df.drop(train_df.index) train_df.head()
```

Fig. 5: Código Python



```
from transformers import BertTokenizer, \
    BertForSequenceClassification, AdamW
tokenizer = BertTokenizer.from_pretrained(model_str)
```

Fig. 6: Código Python

## C. Tokenizamos el texto y obtenemos las etiquetas

En este paso se convierte el texto en secuencias numéricas usando el tokenizador BERT estableciendo un máximo de 512 por texto. Esto puede variar y lo podemos modificar para disminuirlo y aumentaría el tiempo de ejecución.

```
max_length = 512

def encode_text(texts):
    input_ids = []
    attention_masks = []
    for text in texts:
        encoded_text = tokenizer.encode_plus(
            text, add_special_tokens=True,
            max_length=max_length,
            truncation=True,
            padding='max_length',
            return_attention_mask=True,
            return_tensors='pt'
        )
        input_ids.append(encoded_text['input_ids'])
        attention_masks.append(encoded_text['attention_mask'])
    input_ids = torch.cat(input_ids, dim=0)
```

Fig. 7: Código Python

## D. Cargamos el modelo y los detalles para entrenar

En este paso se carga el modelo pre entrenado se le da como parámetro el modelo y el número de CAT del modelo pre entrenado

```
model = BertForSequenceClassification.from_pretrained(
    model_str, num_labels=2,
)

```

Fig. 8: Código Python

## E. Realizamos el entrenamiento



Para el entrenamiento por épocas, se establece el número que se va a ciclar su entrenamiento, luego se ingresa la data completa y luego se transforma las etiquetas en el tipo de dato necesario en nuestro caso convertimos a tipo entero, en vectores binarios, a tipo flotante. Este mismo proceso se repite para convertir las predicciones a tipo entero, luego se obtiene la clase con mayor probabilidad, esto para ejemplo. También debemos usar las métricas de sklearn para calcular la precisión, el f1, el recall y la precisión para cada lote o época.

```
for n_epoca in range(epocas): print(f"### Epocas: {n_epoca + 1}
de {epocas}
-> ###\n")

train_loss = 0.0 train_accuracy =
0.0 train_f1 = 0.0 train_recall =
0.0 train_precision = 0.0
model.train()

for batch in train_dataloader: batch_input_ids
= batch['input_ids'].to(device)
batch_attention_mask =
batch['attention_mask'].to(device) batch_labels =
batch['labels'].to(device)

# Convertir las etiquetas a tipo entero batch_labels =
batch_labels.long()

# Convertir las etiquetas en vectores binarios batch_labels
= F.one_hot(batch_labels, num_classes=N_CAT)
model.zero_grad()

# Convertir las etiquetas a tipo flotante batch_labels =
batch_labels.float()

outputs = model(
    batch_input_ids, token_type_ids=None,
    attention_mask=batch_attention_mask,
    labels=batch_labels
)

loss = outputs[0] logits =
outputs[1] train_loss +=
loss.item()

labels = batch_labels.to('cpu').numpy() preds =
logits.detach().cpu().numpy()

# Convertir las etiquetas y las predicciones a
-> tipo entero
labels = labels.astype(int) preds =
preds.astype(int)

# Obtener la clase con mayor probabilidad para
-> cada ejemplo
labels = np.argmax(labels, axis=1) preds =
np.argmax(preds, axis=1)

# Usar las métricas de sklearn para calcular
-> la precisión, el f1, el recall y la
-> precisión para cada lote
train_accuracy += accuracy_score(labels,
-> preds) train_f1 += f1_score(labels, preds,
-> average='macro', zero_division=0) train_recall +=
recall_score(labels, preds,
-> average='macro', zero_division=0) train_precision +=
precision_score(labels, preds, average='macro',
zero_division=0) loss.backward()

adam.step() scheduler.step()
```

Fig. 9: Código Python





```
avg_train_loss = train_loss /
,→ len(train_dataloader) avg_train_accuracy =
train_accuracy /
,→ len(train_dataloader) avg_train_f1 = train_f1 /
len(train_dataloader) avg_train_recall = train_recall /
,→ len(train_dataloader) avg_train_precision =
train_precision / ,→ len(train_dataloader)

# Imprimir y dar formato a las métricas con dos
,→ decimales
print(f'Train loss: {avg_train_loss:.2f}') print(f'Train accuracy:
,→ {avg_train_accuracy:.2f}') print(f'Train f1: {avg_train_f1:.2f}')
print(f'Train recall: {avg_train_recall:.2f}') print(f'Train precision:
,→ {avg_train_precision:.2f}') print('\n' * 30)

val_loss = 0.0 val_accuracy = 0.0
val_f1 = 0.0 val_recall = 0.0
val_precision = 0.0 model.eval()
```

```
for batch in test_dataloader: batch_input_ids = ,→
batch['input_ids'].to(device)
batch_attention_mask =
,→ batch['attention_mask'].to(device) batch_labels =
batch['labels'].to(device)

# Convertir las etiquetas a tipo entero batch_labels =
batch_labels.long()

# Convertir las etiquetas en vectores binarios batch_labels =
F.one_hot(batch_labels, ,→ num_classes=N_CAT)

# Convertir las etiquetas a tipo flotante batch_labels =
batch_labels.float()

with torch.no_grad(): outputs = model(
    batch_input_ids, token_type_ids=None,
    attention_mask=batch_attention_mask,
    labels=batch_labels
)

loss = outputs[0] logits =
outputs[1] val_loss += loss.item()

preds = logits.detach().cpu().numpy() labels =
batch_labels.to('cpu').numpy()
```

Fig. 10: Código Python

```
# Convertir las etiquetas y las
,→ predicciones a tipo entero
labels = labels.astype(int) preds =
preds.astype(int)

# Obtener la clase con mayor probabilidad
,→ para cada ejemplo
labels = np.argmax(labels, axis=1) preds =
np.argmax(preds, axis=1)

# Usar las métricas de sklearn para
,→ calcular la precisión, el f1, el
,→ recall y la precisión para cada lote
val_accuracy += accuracy_score(labels,
,→ preds) val_f1 += f1_score(labels, preds, ,→
average='macro', zero_division=0) val_recall +=
recall_score(labels, preds,
,→ average='macro', zero_division=0) val_precision +=
precision_score(labels,
,→ preds, average='macro', ,→
zero_division=0)
```

```
avg_val_loss = val_loss / len(test_dataloader) avg_val_accuracy
= val_accuracy /
,→ len(test_dataloader) avg_val_f1 = val_f1 /
len(test_dataloader) avg_val_recall = val_recall /
len(test_dataloader) avg_val_precision = val_precision / ,→
len(test_dataloader)

# Imprimir y dar formato a las métricas con dos
,→ decimales
print(f'Validation loss: {avg_val_loss:.2f}') print(f'Validation
accuracy:
,→ {avg_val_accuracy:.2f}') print(f'Validation f1: {avg_val_f1:.2f}')
print(f'Validation recall: {avg_val_recall:.2f}') print(f'Validation
precision: ,→ {avg_val_precision:.2f}') print('\n')
```

Fig. 11: Código Python



## X. Resultados Obtenidos

Luego de realizar el entrenamiento por época del sistema de reconocimiento de mensajes de odio, obtenemos resultados que sorprendentemente superaron a la implementación de [16], la cual hace uso de uno de los clasificadores más conocidos, ya que ha demostrado ser muy eficaz y precisión en la clasificación de textos.

Una ventaja de ese clasificador es que suele funcionar bien incluso con una pequeña de datos de entrenamiento, además su dataset que ellos usaron para sus pruebas (testing y training) se pueden encontrar en las siguientes enlaces: HatEval y HateNet no dió referencia de cuál escogió.

Pero en [16] se menciona que en el caso del conjunto de datos HaterNet, los tres modelos más precisos (mBERT, XLM y BETO) predijeron las mismas etiquetas erróneas 54 veces de 600 y 207 de 1600 en el caso del conjunto de datos HatEval. En cuanto al mejor sistema BETO, las instancias mal etiquetadas están se inclinan hacia FP en el caso del conjunto de datos HatEval (71,03 en el conjunto de datos HaterNet (61,32 Respecto a nuestro resultado obtenido, hemos aumentado y superado el valor obtenido por [16], a pesar que estamos usando un dataset con menos registros.

Por otro lado, el aumento de lo obtenido se debe a qué antes de la clasificación estamos realizando una clasificación binaria del dataset, extrayendo todos los mensajes que contienen odio de los que no la tienen, eliminando los falsos positivos y aumentando la precisión del modelo LaSalleBHS. Luego entrenamos el modelo con el dataset en español y lo entrenamos en un total de 7 épocas, luego de ello subimos el modo al Drive para su uso en una interfaz más usable.

### A. Comparación de Resultados

Para la comparación de los resultados, estamos tomando como referencia [16] y [10] de tal forma que tomamos dos modelos, uno más antiguo que el otro. El primero es BETO, un modelo basado en BERT pero netamente entrenado para el idioma español. Y el segundo se trata de una Red Neuronal Convolutiva (CNN en inglés), clasifica mensajes en español según 5 categorías: Homosexuales, personas de color, indígenas, judíos, mujeres y también está la categoría de que no es un mensaje de odio.



TABLE II: Comparación de métodos: BETO

Estudio	(Paul, et al, 2020)	LaSalleBHS
Modelo	BETO	LaSalle Beto hatespeech spanish
Precisión	78.87	1.00
Recall	76.51	1.00
F1 score	77.23	1.00
Pérdida	NA	0.00
Dataset	HaterNet	aCORPUS DE DATOS BINARIO

TABLE III: Comparación de métodos: CNN

Estudio	(Sun, et al, 2019)	LaSalleBHS
Modelo	GloVe + CNN	LaSalle Beto hatespeech spanish
Precisión	74.20	1.00
Recall	65.27	1.00
F1 score	67.20	1.00
Pérdida	NA	0.00
Dataset	HaterNet	aCORPUS DE DATOS BINARIO

## XI. Conclusiones

Inicialmente, utilizamos un modelo BERT multilingüe y luego aplicamos un modelo de detección de discurso de odio creado en BETO. Sin embargo, el modelo BERT presenta errores en la clasificación. Al etiquetar una frase, BERT asigna tres etiquetas a una palabra que puede ser considerada como una mala palabra o malsonante. Además, cuando el corpus de entrenamiento no contiene nuevas palabras, como en nuestro caso, por ejemplo, la palabra "muerdealmohadas", el clasificador no podrá encontrar su significado y clasificará incorrectamente la frase. Por otro lado, BETO realiza un etiquetado más preciso, asignando una única etiqueta cuando encuentra una palabra ofensiva o malsonante. Esto evita confusiones durante la clasificación y mejora las métricas, como la precisión, F1, entre otras [4].

## XII. Trabajos Futuros

Se espera mejorar la precisión de nuestro modelo de tal modo que pueda hacer la clasificación a través de un corpus más grande y permitir hacer una predicción más exacta haciendo uso de una variación de LSTM bidireccional [4].



## Contribución de Autoría

**Patrick Leopoldo Paredes Neira:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Gary Jamil Vilca Tapia:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Kristhyan Andree Kurt Lazarte Zubia:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
- [2] VAN AKEN, Betty, et al., "Challenges for toxic comment classification: An in-depth error analysis", 2018.
- [3] BANKO, Michele; MACKEEN, Brendon, "RAY, Laurie. A unified taxonomy of harmful content. En Proceedings of the fourth workshop on online abuse and harms", 2020, p. 125-137.
- [4] PLAZA-DEL-ARCO, Flor Miriam, et al. "Comparing pretrained language models for Spanish hate speech detection. *Expert Systems with Applications*", 2021, vol. 166, p. 114120.
- [5] Thomas Davidson, Dana Warmesley, Michael Macy and Ingmar Weber "Automated hate speech detection and the problem of offensive lenguaje". In Proceedings of the 11th International AAI Conference on Web and Social Media, 2017, p. 512-515
- [6] Alexis Palmer, Christine Carr, Melissa Robinson, Jordan Sanders. "COLD: Annotation scheme and evaluation data set for complex offensive language in English", *Journal for Language Technology and Computational Linguistics*, p. 1-28.



- [7] Alexis Palmer, Christine Carr, Melissa Robinson, Jordan Sanders. "COLL: Annotation scheme and evaluation data set for complex offensive language in English", *Journal for Language Technology and Computational Linguistics*, p. 1-28.
- [8] B. Lohitha, M. V and J. J. Amarnath, "A Comparison of Different Models for the Detection of Hate Speech," 2022 1st International Conference on Computational Science and Technology (ICCST), CHENNAI, India, 2022, pp. 492-496, doi: 10.1109/ICCST55948.2022.10040400.
- [9] RÖTTGER, Paul, et al. HateCheck: Functional tests for hate speech detection models. arXiv preprint arXiv:2012.15606, 2020.
- [10] BALKIR, Esmá, et al. Necessity and sufficiency for explaining text classifiers: A case study in hate speech detection. arXiv preprint arXiv:2205.03302, 2022.
- [11] OTTER, Daniel W.; MEDINA, Julian R.; KALITA, Jugal K. A survey of the usages of deep learning for natural language processing. *IEEE transactions on neural networks and learning systems*, 2020, vol. 32, no 2, p. 604-624.
- [12] WU, Stephen, et al. Deep learning in clinical natural language processing: a methodical review. *Journal of the American Medical Informatics Association*, 2020, vol. 27, no 3, p. 457-470.
- [13] SHISHAH, Wesam; FAJRI, Ricky Maulana. Large Comparative Study of Recent Computational Approach in Automatic Hate Speech Detection. *TEM Journal*, 2022, vol. 11, no 1, p. 82.
- [14] PÉREZ, Juan Manuel, et al. Assessing the impact of contextual information in hate speech detection. *IEEE Access*, 2023, vol. 11, p. 30575-30590.
- [15] RÖTTGER, Paul, et al. MULTILINGUAL HATECHECK: Functional Tests for Multilingual Hate Speech Detection Models. arXiv preprint arXiv:2206.09917, 2022.
- [16] ABRO, Sindhu, et al. Automatic hate speech detection using machine learning: A comparative study. *International Journal of Advanced Computer Science and Applications*, 2020, vol. 11, no 8.





# Garantía de seguridad de la información empresarial a través de la gestión de servicios

*Ensuring the security of business information through service management*

96

## Luz Maria Solano Quincho

Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ t023300520@unitru.edu.pe

id <https://orcid.org/0009-0002-2243-7388>

## Miriam Maryori Horna Maguiña

Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ t513300720@unitru.edu.pe


id <https://orcid.org/0009-0009-4199-5765>


## Alberto Mendoza de los Santos


Universidad Nacional de Trujillo,  
Trujillo, Perú.

@ amendozad@unitru.edu.pe

id <https://orcid.org/0000-0002-0469-915X>

 **ARK:** <ark:/42411/s12/a95>

 **DOI:** [10.48168/innosoft.s12.a95](https://doi.org/10.48168/innosoft.s12.a95)

 **PURL:** [42411/s12/a95](https://purl.org/42411/s12/a95)

RECIBIDO 28/04/2023 • ACEPTADO 31/07/2023 • PUBLICADO 30/09/2023



## RESUMEN

En esta revisión sistemática, se examinaron los enfoques y técnicas utilizados por las empresas para proteger su información y cumplir con las expectativas de los clientes. Siguiendo una metodología rigurosa que involucró una investigación sistemática en bases de datos y un análisis exhaustivo de los artículos seleccionados, se identificaron seis estudios relevantes que proporcionaron información valiosa. Los hallazgos resaltaron la relevancia de incorporar la seguridad informática en la estrategia general de gestión de servicios y la necesidad de involucrar a todas las partes interesadas. Además, se enfatiza el papel crucial del elemento humano en la protección de la información y se recomienda invertir en programas de capacitación y concientización, asimismo se resalta la importancia de un enfoque holístico dentro de la gestión de servicios. El documento ofrece una visión integral y recomendaciones efectivas para fortalecer las prácticas de seguridad y protección de datos empresariales, siendo relevante para las organizaciones que buscan mejorar su enfoque en la protección de los datos en un entorno empresarial dinámico y en continua transformación.

**Palabras claves:** Gestión de seguridad de la información, gestión de servicios de TIC, seguridad de la información, sector empresarial..



## ABSTRACT

*In this systematic review, we examined the approaches and techniques used by companies to protect their information and meet customer expectations. Following a rigorous methodology involving systematic database research and a thorough analysis of selected articles, six relevant studies were identified that provided valuable information. The findings highlighted the relevance of incorporating IT security into the overall service management strategy and the need to involve all stakeholders. In addition, the crucial role of the human element in information protection is emphasised and investment in training and awareness programmes is recommended, as well as the importance of a holistic approach to service management. The document provides a comprehensive view and effective recommendations for strengthening enterprise data protection and security practices, and is relevant for organisations seeking to improve their approach to data protection in a dynamic and ever-changing business environment.*

**Keywords:** *Information security management, ICT service management, information security, business sector.*

---

## INTRODUCCIÓN

En la era actual, las empresas se respaldan en las soluciones informáticas como un componente esencial para garantizar la integridad y confianza de los servicios que brindan a sus usuarios según [2]. El uso de esta herramienta se ha vuelto indispensable para las empresas, ya que permite planificar, diseñar, implementar y mejorar sus servicios de manera sistemática y controlada. Esto resulta en un aumento significativo de la productividad y efectividad en la entrega de los servicios, lo que se traduce en beneficios tangibles como mayor agilidad, rendimiento y satisfacción para los usuarios.

Los servicios de informática y seguridad trabajan en conjunto como una alianza indispensable para preservar los datos sensibles [3]. Así mismo, la gestión de servicios ha evolucionado con nuevos enfoques, prácticas y herramientas que permiten una mayor agilidad y adaptabilidad a los cambios en el entorno empresarial. Algunos de estos enfoques incluyen la gestión basada en la nube, la gestión ágil y la gestión de seguridad informática.

Dado que los datos empresariales son un activo de gran valor, resulta necesario salvaguardar los procesos que la gestionan de cualquier vulnerabilidad que pueda amenazar la continuidad y estabilidad de la organización según [7], es por eso que es imprescindible tener un equipo dedicado a la seguridad informática de la empresa. Este equipo coordina y supervisa las acciones de los departamentos; priorizando, transmitiendo información, mejorando la eficacia y tomando decisiones dentro de la organización.



Esta revisión sistemática examina la gestión de servicios como garantía de la seguridad de la informática en las empresas, abordando dos elementos esenciales: la gestión de servicios y la seguridad de la información en el sector empresarial.

## Materiales y métodos

### Tipo de Estudio

El presente estudio busca examinar la conexión entre la gestión de servicios y la seguridad de la información en el ámbito empresarial mediante una revisión sistemática. En este estudio, se empleó la lista de verificación de la metodología PRISMA para asegurar que se elabore de manera correcta la revisión sistemática.

La pregunta de investigación que se planteó para guiar el proceso de investigación fue la siguiente: ¿Cuál es el impacto que tiene la gestión de servicios en la seguridad de la información en el entorno empresarial?

### Fundamentación de la Metodología

Las revisiones sistemáticas son una forma estructurada, explícita y sistemática de resumir y analizar la evidencia relacionada con una pregunta específica. En general, se describe el método utilizado para buscar, seleccionar, analizar y sintetizar la evidencia presentada (según Letelier y otros). En base a la definición de las revisiones sistemáticas se resalta la importancia de sintetizar y sistematizar la información, con el objetivo de identificar y enfocarse en lo más relevante de toda la documentación disponible en diferentes contextos de investigación, pero que comparten un objetivo común.

Continuamos con el detalle del procedimiento en cada etapa de este estudio.

### Búsqueda inicial

En la búsqueda inicial se hizo combinando los siguientes términos 'seguridad de la información', 'gestión de servicios', 'empresas' en las bases de datos ScienceDirect, Scielo, Scopus y PubMed. Luego, se agregaron los términos 'seguridad informática', 'gestión de la seguridad de la información', 'negocios', 'Business environment' y a la búsqueda se unieron los operadores AND y OR. Debido a que los artículos encontrados en "PubMed" tenían poca relación con el artículo de revisión, se decidió retirarlos de la búsqueda.

### Búsqueda Sistemática

Para llevar a cabo la búsqueda sistemática reciente, se efectuaron búsquedas en las diferentes bases de datos, como Scopus, Scielo y ScienceDirect. Se pudo reducir los resultados solo



considerando las publicaciones realizadas en los últimos 6 años, es decir a partir del año 2018 hasta el año 2023. La búsqueda avanzada que se realizó en Scopus fue:

- (KEY (security) AND KEY (information) AND KEY (company) AND KEY (service) AND KEY (management)) AND (LIMIT-TO (PUBYEAR,2022) OR LIMIT-TO (PUBYEAR,2020) OR LIMIT-TO (PUBYEAR,2019) OR LIMIT-TO (PUBYEAR,2018)) AND (LIMIT-TO (LANGUAGE,"English"))

En base de datos Scielo se usó la siguiente búsqueda avanzada:

- (gestión de servicios) OR (seguridad de la información) OR (empresas) AND la:("pt") AND year\_cluster:("2020" OR "2021" OR "2018" OR "2019") AND wok\_subject\_categories:("business" OR "services")

Mientras que en el buscador de ScienceDirect se buscó lo siguiente:

- Service management tools - information security - companies Year: 2019-2023 Title: information security

En resumen, se obtuvieron 6 resultados en la búsqueda realizada en Scopus, 108 resultados en Scielo y 93 resultados en ScienceDirect. Antes de continuar con el análisis de los artículos se establecieron criterios para determinar qué artículos serían incluidos y cuáles serían excluidos.

## **Criterios de inclusión:**

- Durante la elaboración de la revisión sistemática, se tuvieron en cuenta los artículos publicados en un periodo de 6 años, abarcando desde el año 2018 hasta la actualidad (2023).
- Los artículos encontrados deben abordar la pregunta de investigación.
- Se seleccionan artículos publicados en los siguientes idiomas: español, inglés y portugués.

## **Criterios de exclusión:**

- No se deben considerar los artículos que engloban sectores irrelevantes para la revisión sistemática.
- Los documentos que no son artículos.

Siguiendo los criterios establecidos, se eliminaron 184 artículos tras leer el título y 5 artículos que eran duplicados quedando así 19. Luego continuamos con la lectura del resumen y, a partir de



eso, se descartaron 6 pues estaban enfocados en el aprendizaje automático, 2 no eran considerados artículos y 5 están centrados en el área de la salud pública.

Finalmente, solo 6 artículos cumplían con los criterios de inclusión y exclusión, por eso se eligieron para realizar la presente investigación como se muestra en la Figura 1. Los artículos elegidos abordan el tema de la gestión de la seguridad informática en el sector empresarial. Además, algunos de ellos presentan ejemplos de cómo la seguridad informática beneficia a las empresas y cómo el factor humano desempeña un papel influyente en este aspecto.

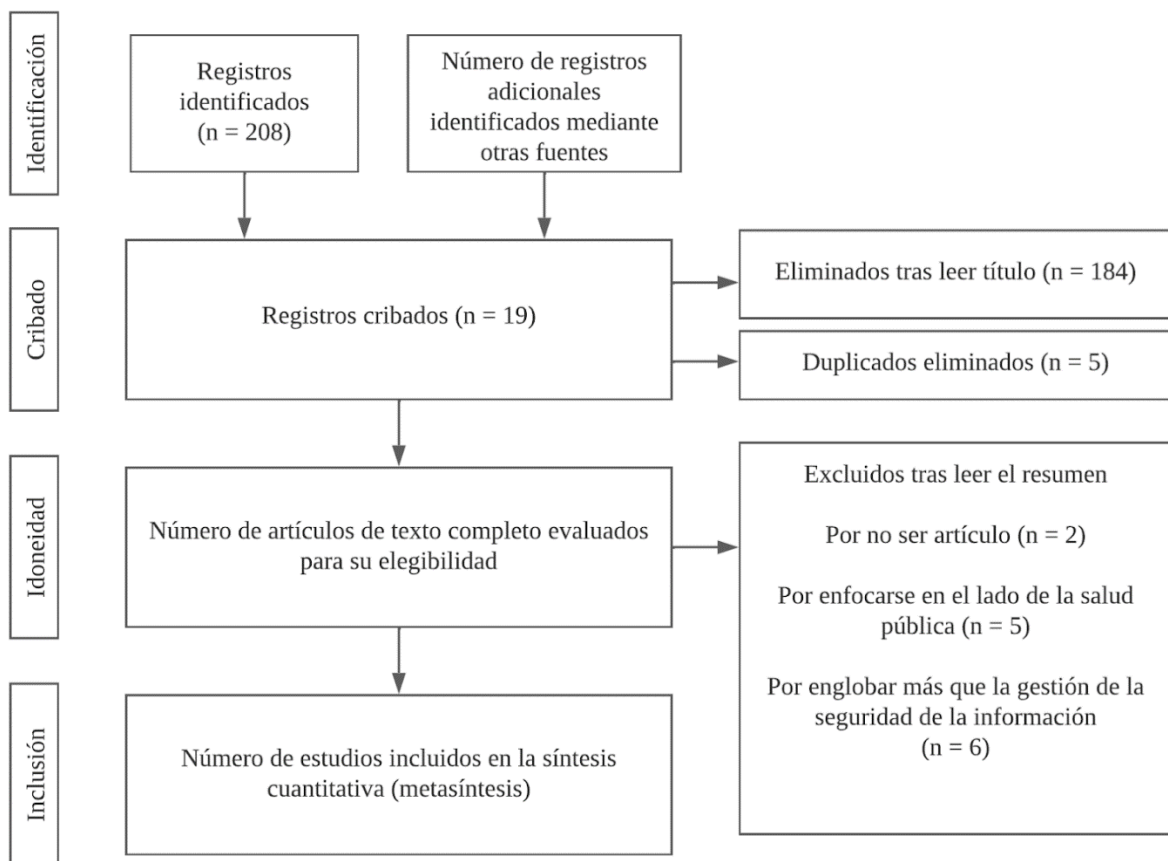


Figura 1. Diagrama de flujo de la metodología PRISMA en cuatro niveles

## Resultados y discusión

Primero identificamos los datos de los artículos que resultaron de la búsqueda sistemática y se les aplicaron los criterios de exclusión e inclusión antes mencionados; para ver en qué país se desarrolló, en qué revista fueron publicados, como abordan el tema, todo esto lo expresamos en la Tabla 1.





En la Figura 2 exhibe un gráfico que muestra el número de colaboradores involucrados en la elaboración de las investigaciones seleccionadas para la revisión, clasificados según su nacionalidad. Esta representación nos brinda datos sobre la ubicación geográfica de los investigadores centrados en el tema de investigación de la revisión sistemática realizada.

Tabla 1. Investigaciones seleccionadas para la elaboración de la revisión sistemática

N°	Artículo	País	Impacto de la gestión de servicios en la seguridad informática
1	A. F. Cadena, 2022. Aplicación de la Seguridad Informática como Herramienta de Gestión Empresarial en el Contexto del Gobierno TIC.	Colombia	La gestión de servicios forma parte integral de un sistema completo, el cual se centra en abordar los riesgos globales de un negocio, logrando implementar importantes normas para la empresa, tales como la norma ISO 27000 y la norma NIST SP 800, proveen en las empresas enfoques y prácticas que respaldan su fundamental labor de salvaguardar la seguridad de la información.
2	A. Da Veiga, V. Liudmila, A. Botha, M. Herselman, 2020. Defining organisational information security culture- Perspectives from academia and industry.	Sudáfrica, Rusia	La gestión de servicios influye al proporcionar los medios y las prácticas necesarias para establecer y fomentar un enfoque de seguridad de la información sólido en las empresas. Se ocupa de diseñar, entregar, operar y mejorar los servicios de TI, lo que implica la implementación de políticas, controles de seguridad y procedimientos.
3	D.M. Altamirano, 2019. Modelo para la gestión de la seguridad de la información y los riesgos asociados a su uso.	Cuba	La gestión de servicios de TIC se manifiesta al proponer un modelo que brinda la automatización de las acciones de operación, monitorización y revisión de los controles de seguridad informática, lo cual contribuye a la protección de la seguridad de la información y detección de eventos de seguridad en redes de computadoras
4	K. Hughes-Lartey , M. Li, F. E. Botchey y Z. Qin, 2021. Human factor, a critical weak point in the information security of an	China, Ghana	La gestión de servicios puede mejorar la seguridad de la información al asegurar que los sistemas y procesos estén diseñados y operados de manera segura y eficiente. Además, una buena gestión de servicios



	organization's Internet of things. <i>Heliyon</i> , 7(3), e06522.		puede ayudar a detectar y responder rápidamente a posibles amenazas o incidentes de seguridad.
5	E. K. Szczepaniuk, H. Szczepaniuk, T. Rokicki, y B. Klepacki, 2020. Information security assessment in public administration.	Polonia	La gestión de servicios puede garantizar que las aplicaciones utilizadas por las entidades de la administración pública estén actualizadas y protegidas contra vulnerabilidades conocidas, que se realicen copias de seguridad y recuperación de desastres, y que se implementen medidas adecuadas para proteger la información confidencial. Entonces la gestión de servicios puede contribuir a mejorar la seguridad de la información en las organizaciones de la administración pública.
6	I. I. Livshitz, P. A. Lontsikh, N. P. Lontsikh, E. A. Golovina y O. Safonova, 2020. <i>The Effects of Cyber-security Risks on Added Value of Consulting Services for IT-security Management Systems in Holding Companies.</i>	Rusia	La gestión de servicios garantiza una mejor seguridad de la información al identificar y evaluar áreas problemáticas donde pueden ocurrir pérdidas. Además, el uso de herramientas y prácticas robustas de gestión de riesgos cibernéticos puede mejorar aún más la seguridad de la información en las empresas.

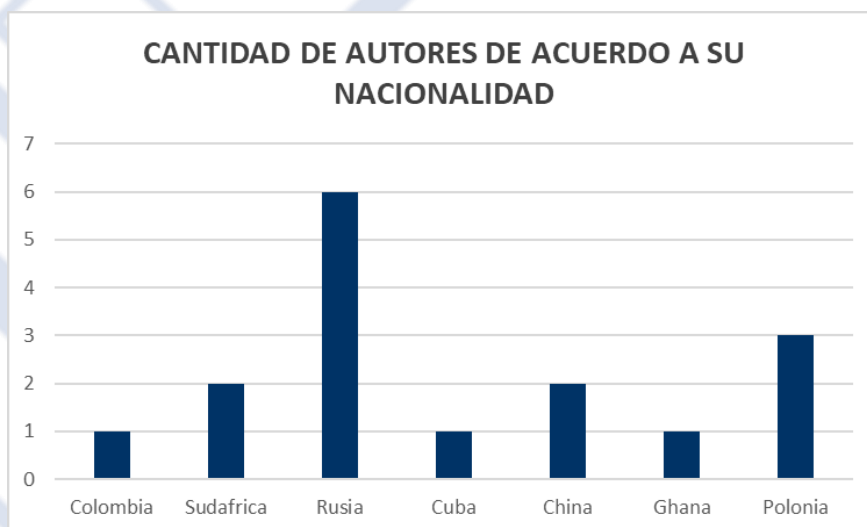


Figura 2. Cantidad de autores de acuerdo a su nacionalidad



En la figura siguiente se representa el porcentaje de artículos por año, según se puede observar en la Tabla 2. Podemos decir que en el año 2018 no se encontraron artículos significativos que cumplieran con los criterios de selección que se han mencionado anteriormente, lo que representa un porcentaje de 0%. En el año 2019 se identificó una publicación, mientras que en el año 2020 se seleccionaron la mitad de las publicaciones incluidas en esta revisión sistemática, lo que equivale al 50% de los artículos. Por otro lado, en los años 2021 y 2022 solo se encontró una publicación en cada año, lo que representa un porcentaje de 16.67%.

Tabla 2. Porcentaje de artículos según el año de publicación.

<b>Año</b>	<b>Nº de Publicaciones</b>	<b>Porcentaje</b>
2022	1	16.67%
2021	1	16.67%
2020	3	50.00%
2019	1	16.67%
2018	0	0.00%

## Concepto Gestión de servicios de TIC

Es esencial enfocarse en la excelencia de los servicios de TI y su alcance con los objetivos de la organización al analizar las prioridades del sector. Para servicios críticos, se requiere seguir una secuencia específica en cada paso para garantizar un aporte de valor constante por parte del equipo de TI.

Según [1]: La gestión de servicios abarca diversas disciplinas especializadas que generan valor para los clientes a través de la entrega de servicios. Estas disciplinas se materializan a través de funciones y procesos que abarcan las diferentes etapas dentro del ciclo de vida del servicio.

Según [4]: "La Gestión de Servicios es una práctica que se basa en la colaboración de procesos buscando aumentar la calidad de servicios interconectados y que se encuentran en constante evolución, pero esto depende del servicio acordado con el cliente. Incluye diferentes áreas de gestión, como la gestión de redes, la gestión de sistemas y el desarrollo de sistemas, así como procesos clave como la gestión de activos, la gestión de cambios, y la gestión de conflictos. En estas citas se puede ver el concepto del autor de que la gestión de servicios en un entorno TIC es fundamental para entregar valor a los clientes y garantizar la excelencia en los servicios ofrecidos. Porque es una disciplina que se basa en procesos que incluye administración de sistemas, desarrollo web y de sistemas, y muchas otras áreas de proceso.



## Concepto Gestión de la seguridad de la información

Según [6], proteger la información implica asegurar de manera eficaz la integridad de los datos mediante la implementación de un enfoque que gestione los riesgos. Esto, a su vez, fomenta la generación de confianza entre las partes o entidades relacionadas.

Según [3], el propósito principal de la seguridad informática radica en proteger los datos y sistemas de información, asegurando su integridad, confidencialidad y disponibilidad, al evitar cualquier tipo de acceso o divulgación no autorizada. Este concepto está vinculado a la seguridad, la cual implica la inexistencia de riesgos o la preparación para enfrentarlos, y se refiere a un estado en el que no hay presencia de peligro, daño o amenaza. Se implementan medidas preventivas y reactivas en los sistemas organizativos y técnicos con el fin de proteger y asegurar la información, manteniendo intacta su integridad, confidencialidad y disponibilidad.

Entonces resumiendo, garantizar la seguridad de la información permite la puesta en práctica de una variedad de medidas destinadas a proteger el correcto desarrollo de la información. Las citas mencionadas enfatizan que el principal objetivo de la seguridad informática es garantizar el resguardo de los datos y sistemas de información contra posibles riesgos o amenazas.

## Discusión

Según [4]: La Gestión de Servicios se sustenta en la cooperación entre procesos, con la finalidad de fortalecer la eficacia de los servicios que están en operación y en constante desarrollo, de acuerdo a los niveles de servicio establecidos con los clientes en una empresa. En esta revisión sistemática, se realiza un análisis detallado sobre el estado actual de la gestión de servicios como una herramienta que asegura la información dentro del sector empresarial, en la que se está apreciando varios puntos a argumentar; en primer lugar, se puede investigar cómo en una organización, la gestión de servicios juega un papel esencial al colaborar en identificar y evaluar los riesgos relacionados con la seguridad informática.

Según [2]: Con el fin de implementar un sistema de control de la seguridad de la información de manera exclusiva, es necesario seguir una serie de etapas esenciales, que incluya una gestión de riesgos sistemática y personalizada según los requerimientos particulares de la entidad, con el objetivo de obtener un enfoque eficaz y oportuno. Esto implica la ejecución de estrategias y controles de seguridad adecuados para resguardar los sistemas y los datos.

Por lo tanto, se puede discutir cómo la gestión de servicios facilita la detección temprana y la respuesta eficiente a incidentes de seguridad. Esto conlleva a establecer procedimientos y protocolos claros para la gestión de incidentes, así como la capacitación del personal para responder ante eventos críticos. De acuerdo a nuestros resultados obtenidos de la investigación realizada, observamos que la gestión de servicios se encarga de fortalecer la seguridad de la información, potencia la detección y manejo de incidentes, y fomenta un enfoque de seguridad en cada uno de los niveles de la organización.



## Conclusiones

La gestión de servicios desempeña un papel crucial dentro de las empresas para proteger sus datos, garantizando que estos se mantengan completos, estén accesibles cuando se necesiten y sean confidenciales. Al adoptar prácticas recomendadas, las organizaciones pueden administrar sus servicios de tecnología de la información de forma efectiva y asegurarse de proteger sus datos de manera adecuada.

Se destaca la importancia de integrar aspectos organizativos, humanos y estratégicos en la gestión de servicios, no limitándose únicamente a las medidas de seguridad tecnológicas. La colaboración entre diferentes departamentos de la organización es fundamental para establecer prioridades, tomar decisiones informadas y mantener la eficiencia institucional.

Un aspecto crucial que aún no ha sido abordado en la investigación es el conocimiento y las capacidades del personal en cuanto a la seguridad de la información, ya que son quienes manejan la información de la organización. Hasta ahora, la implementación de la gestión de servicios ha sido esencial para respaldar eficazmente los datos relevantes de una empresa.

## Contribución de Autoría

**Luz Maria Solano Quincho:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Redacción - borrador original](#). **Miriam Maryori Horna Maguiña:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Redacción - borrador original](#). **Alberto Carlos Mendoza de los Santos:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Validación](#), [Redacción - borrador original](#), [Supervisión](#).

## Referencias

- [1] J.L. Baud, 2016. ITIL V3 Entender el enfoque y adoptar las buenas prácticas. Barcelona, España: Editorial Angel María Sanchez Conejo.
- [2] A. Infante, s.f. Aplicación de la Seguridad Informática como Herramienta de Gestión Empresarial en el Contexto del Gobierno TIC. Universidad Piloto de Colombia. Gobierno TIC y Seguridad Informática. Recuperado de: <http://repository.unipiloto.edu.co/bitstream/handle/20.500.12277/12277/Aplicaci%C3%B3n%20de%20la%20Seguridad%20Inform%C3%A1tica%20como%20Herramienta%20de%20Gesti%C3%B3n%20Empresarial%20en%20el%20Contexto%20del%20Gobierno%20TIC.pdf?sequence=1&isAllowed=y>





- [3] R. Godoy, 2014. Seguridad de Información. Guatemala: Revista de la Segunda Cohorte del Doctorado en Seguridad Estratégica. Disponible en: [https://books.google.com.pe/books?id=xKkYBgAAQBAJ&printsec=frontcover&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](https://books.google.com.pe/books?id=xKkYBgAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false)
- [4] D. A. Lobos Anfuso, M. Baquinzay, y M. S. Bustos Aguiar, n.d. GESTIÓN DE SERVICIOS TIC (Tecnología de la información y las comunicaciones) - ITIL. Biblioteca de Infraestructura de Tecnologías de Información. Disponible en: [//www.editorial.unca.edu.ar/Publicacione%20on%20line/CIENCIA%20Y%20TECNOLOGIA/Revista%2013%20ONLINE/2.Gestion%20de%20servicios.pdf](http://www.editorial.unca.edu.ar/Publicacione%20on%20line/CIENCIA%20Y%20TECNOLOGIA/Revista%2013%20ONLINE/2.Gestion%20de%20servicios.pdf)
- [5] L. Letelier, J. Manríquez, y G. Rada, 2005. Revisiones sistemáticas y metaanálisis ¿son la mejor opción? Disponible en: [https://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0034-98872005000200015](https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0034-98872005000200015)
- [6] NormaISO27001. (s.f.). Referencias normativas ISO 27000. Recuperado de: <https://normaISO27001.es/referencias-normativas-iso-27000/#:~:text=ISO%20%2F%20IEC%2027000%3A%202018%20nos,sobre%20seguridad%20de%20la%20informaci%C3%B3n>
- [7] P. Aguilera, 2010. Seguridad informática, Ciclos Formativos. Madrid: Editex.



# Detección de Ideas Suicidas en Twitter

## *Suicidal Idea Detection on Twitter*

107

### **Juliana Berrios Butron**

Universidad La Salle. Arequipa, Perú.

@jberriosb@ulasalle.edu.pe

### **Daria López Franco**


Universidad La Salle. Arequipa, Perú.


@dlopezf@ulasalle.edu.pe


### **Dante Rodolfo Tárraga**

Universidad La Salle. Arequipa, Perú.

@dtarragau@ulasalle.edu.pe

 **ARK:** [ark:/42411/s12/a97](https://nbn-resolving.org/ark:/42411/s12/a97)

 **DOI:** [10.48168/innosoft.s12.a97](https://doi.org/10.48168/innosoft.s12.a97)

 **PURL:** [42411/s12/a97](https://nbn-resolving.org/ark:/42411/s12/a97)

RECIBIDO 02/05/2023 • ACEPTADO 02/08/2023 • PUBLICADO 30/09/2023



## RESUMEN

El análisis de sentimientos es una nueva tendencia hoy en día para entender cómo se sienten las personas en diferentes situaciones de su vida cotidiana. Los datos de las redes sociales se utilizan durante todo el proceso de análisis y clasificación, que consiste en datos de texto. Utilizando las redes sociales, se puede monitorizar o analizar el nivel emocional. En este trabajo de investigación clasificaremos los datos de redes sociales como twitter relativos al suicidio y los clasificaremos como: pensamiento suicida activo, pensamiento suicida pasivo, sarcasmo relacionado con el pensamiento suicida, tweets relacionados con el suicidio (concienciación sobre el suicidio, noticias, charla sobre el suicidio) y otros.

**Palabras claves:** Suicidio, pensamientos suicidas, procesamiento de lenguaje natural.

## ABSTRACT

*Sentiment analysis is a new trend nowadays to understand how people feel in different situations of their daily life. Social network data is used during the whole process of analysis and classification, which consists of text data. Using social networks, the emotional level can be monitored or analyzed. In this research work we will classify data from social networks such as twitter regarding suicide and classify it as: active suicidal thinking, passive suicidal thinking, sarcasm related to suicidal thinking, tweets related to suicide (suicide awareness, news, suicide talk) and others.*



**Keywords:** *Suicide, suicidal thoughts, natural language processing.*

---

## INTRODUCCIÓN

Las tasas de suicidio han sido un tema preocupante en muchos países, entre ellos Perú. Según una revista de Neuro-Psiquiatría, la cantidad de personas que se suicidaron en un lapso del 2017-2021 incluye un total de 2579 suicidios [8]. Esto resalta la urgencia de implementar estrategias efectivas de prevención del suicidio. Estudios han demostrado que el suicidio es prevenible a través de intervenciones oportunas y específicas. En este contexto, las aplicaciones de NLP pueden desempeñar un papel importante a la hora de señalar e identificar a las personas en situación de riesgo mediante el análisis de patrones lingüísticos y sentimientos en las publicaciones de las redes sociales.

El objetivo de este proyecto de aplicación NLP es detectar pensamientos suicidas expresados en Twitter. Sabemos que en la actualidad las redes sociales se han convertido en una parte integral de la vida diaria de muchas personas y pueden tener un impacto significativo en la salud mental. Muchas personas utilizan las redes sociales para conectarse con otros, pero también pueden exponerse a una gran cantidad de información negativa, comparación social y ciberacoso. Twitter es una de las plataformas de redes sociales más populares y la gente la utiliza a menudo para expresar sus emociones y pensamientos. El proyecto desarrolla un modelo de aprendizaje automático que pueda analizar tweets para identificar patrones lingüísticos asociados a los pensamientos suicidas y marcar tweets que puedan indicar un riesgo de suicidio entre las categorías de activo o pasivo, usando modelos pre entrenados como Roberta y Bert. Identificar los pensamientos suicidas en los tweets podría ayudar a intervenir precozmente, lo que a su vez puede mejorar la calidad de vida de las personas afectadas y salvar vidas.

## Preguntas de Investigación

El objetivo del proyecto es desarrollar un algoritmo basado en procesamiento de lenguaje natural (NLP) para clasificar el contenido de texto subido a las redes sociales, para identificar tweets respecto al suicidio.

- RQ1: ¿Que palabras indican posibles pensamientos suicidas? (activa, pasiva, sarcástica, relacionado al tema u otros)
- RQ2: ¿Qué modelo será más efectivo para este proyecto? Ya sea de Bert o Roberta



## DATOS

Los datos clasificados se obtuvieron de Kaggle [6], hay tres columnas principales, los ids de los tweets, la clasificación y especificar si el dato se usara para entrenar al modelo, validarlo y otros. Estos datos cuentan con cerca de 81519 tweets. Este conjunto de datos se caracteriza por ser el único que toma en cuenta una clasificación para los tweets relacionados al suicidio considerando las categorías de activo, cuando el texto denota alguna intención suicida, pasiva las cuales son potencialmente activas en un futuro, sarcasmo dado que es propio de la plataforma, relacionado al suicidio como noticias o charlas y finalmente otros. La importancia de una clasificación es que los tipos de pensamientos suicidas pasivos o activos requieren un apoyo y ayuda diferente. Para obtener los datos se tuvo de base el código proporcionado por raza8899 (usuario de Github). El código toma el id del tweet, que lo obtenemos del dataset, y se devolvemos la sentencia que es lo que se va a analizar. Finalmente se crea un nuevo dataset con las sentencias, su respectiva etiqueta y el dataset al que pertenece. Se mostrará el pseudocódigo.

## Funciones

- `Get_specific_tweet`: Recibe el id del tweet para retornar solo el contenido de este.
- `Snscraper_get_tweet`: Esta función usa `TwitterTweetScrapper` de la librería `snscraper` tomando como parametros el id del tweet y el modo `single`.
- `Get_items`: Devuelve los objetos (tweets) obtenidos de un conjunto de datos.

## Pseudocódigo

```
importar libreria snsrape
importar libreria json
importar libreria pandas
importar libreria numpy

función get_specific_tweet(id_tweet):
for tweet in enumerate(snscraper_get_tweet()
.get_items())
dato = tweet transformado a json

# Verificamos que se obtenga el contenido
if "rawContent" in dato:
retornamos el rawContent
else:
return '0'

df = leemos el dataset usando pandas
```



```
ids = lista de la columna de ids del dataset
lab = lista de la columna de labels del
dataset
dat = lista de la columna de dataset del
dataset
test_list = convertimos los ids a enteros
tweet = array vacio para agregar los tweets
con el contenido
```

```
for i in range (longitud del dataset o
longitud deseada):
```

```
try:
```

```
tweet = get_specific_tweet(test_list[i])
cargamos el tweet al arreglo de tweets
junto con su respectivo label y dataset
```

```
except:
```

```
pass #pasamos los tweets a los que no se
tuvo acceso
```

```
arr = convertir el arreglo de tweets a un
array de numpy
con el nuevo arreglo de datos lo pasamos a
formato csv usando de cabecera sentence,
label y dataset
```

## REVISION LITERARIA

Para identificar otras soluciones propuestas se realizó una revisión literaria respecto al tema utilizando palabras claves como: suicide, suicide ideation, nlp, clasification.

### A. Natural language processing applied to mental illness detection: a narrative review

Este artículo [4] propone la revisión de investigaciones relacionadas al procesamiento de lenguaje natural aplicado en la detección de enfermedades mentales. Principalmente identifica varias fuentes de datos como son posts en redes sociales, entrevistas, repositorio clínico y escritura narrativa. Detalla los métodos de machine learning y deep learning que son en su mayoría utilizados. Se concluye que ambos métodos son eficientes pero que dependen mucho de los datos utilizados, en especial el volumen y la fuente, debido a que no existen muchos conjuntos de datos. Los temas más recurrentes en este ámbito de investigación son la depresión y el suicidio,





las bases de datos más usadas son Twitter y Reddit debido a su facilidad de obtención de los datos, el lenguaje más usado para la entrenar a los modelos es el inglés.

## **B. Detection of suicide-related posts in Twitter data streams. IBM Journal of Research and Development**

La detección temprana de señales de advertencia de suicidio en las redes sociales, específicamente en Twitter, es necesaria para prevenir el suicidio y mejorar la salud mental de las personas. Sin embargo, la identificación manual de publicaciones relacionadas con el suicidio es un proceso muy lento y costoso. Por ello se necesita un enfoque innovador que pueda detectar automáticamente estas publicaciones y proporcionar una alerta temprana a los profesionales de la salud mental. Este artículo [2] presenta un nuevo enfoque que utiliza técnicas avanzadas de NLP para detectar señales de advertencia de suicidio en Twitter.

Para resolver el problema de la detección de pensamientos suicidas en las redes sociales en línea, los autores proponen un nuevo enfoque que utiliza Twitter para cuantificar las señales de advertencia de suicidio y detectar los mensajes con contenido relacionado con el suicidio. Para ello se utilizan dos métodos diferentes: el procesamiento del lenguaje natural (NLP) y el aprendizaje automático. El método de NLP combina características generadas a partir del texto basadas en un conjunto de léxicos, que se componen de temas lingüísticos comúnmente exhibidos por las personas en situación de riesgo. El segundo método, denominado clasificador de angustia, se basa en el aprendizaje automático.

En el estudio, se concluye que el enfoque propuesto resulta efectivo en la detección de contenido relacionado con el suicidio en Twitter. Los autores se enfocaron en detectar contenido relacionado con la angustia y el suicidio, utilizando los dos enfoques mencionados anteriormente. Aunque el estudio tiene algunas limitaciones, como la prueba con solo dos usuarios de Twitter, se sugiere que futuras investigaciones podrían ampliar las pruebas con más usuarios y líneas temporales, especialmente con aquellos que se identifiquen como personas con enfermedades mentales. En general, este trabajo presenta un enfoque innovador con el potencial de salvar vidas y prevenir el suicidio al detectar las ideas suicidas en las redes sociales en línea.

## **C. Exploring and Learning Suicidal Ideation Connotations on Social Media with Deep Learning**

Debido a esta creciente disponibilidad de contenidos en las redes sociales (como Twitter, Facebook y Reddit, etc.), urge identificar a las personas afectadas y ofrecerles ayuda [3].

El artículo [3] busca realizar la comparación entre 3 arquitecturas de deep learning. Primero hizo la recolección de los datos, los clasifico manualmente y luego se utilizaron en tres modelos para determinar cuál era el mejor.

Los datos fueron recolectados de 3 fuentes principales, sitios de micro-blogueo, Tumblr y Reddit. Donde 300 datos fueron de forums de suicidio y 2000 datos fueron de posts seleccionados aleatoriamente en Tumblr, entre otras plataformas el conjunto de datos total consistía en 5217 sentencias. Para la clasificación utilizaron a tres observadores quienes manualmente identificaron



si el dato se consideraba relacionado al suicidio o no. En caso de ambigüedad se optó por colocarlo como suicida. Para la categoría de suicidio se considera que este explicito el intento de suicidio del usuario. Para la categoría de no suicidio se considera sarcasmo, reportes, o textos relacionados a la sensibilización del tema.

Para el preprocesamiento se quitaron las urls, la mención de usuarios, hashtags y los signos como el guion y el apóstrofe (-, '), se consideraron las palabras que repitieran las letras. Luego esta información paso por tres arquitecturas. Recurrent Neural Network (RNN) modelo que se destaca por ser eficiente al clasificar información secuencial de longitud arbitraria. Long Short Term Memory (LSTM), método adecuado para clasificar, procesar y predecir series temporales y captar dependencias a largo plazo en las frases. Finalmente, C-LSTM, este modelo combina la eficiencia de codificación espacial y la extracción automática de características de la red neuronal convolucional (Convolutional neural network, CNN), combinada con la eficaz clasificación de textos mediante LSTM.

El procesamiento de clasificación se basó en tres etapas. Integración de palabras, integración de frases y clasificación. Para la primera etapa se usaron las N palabras más frecuentes en un tweet para formar una capa, se utilizó word2vec. La segunda etapa se usó en el modelo C-LSTM, donde se añade un CNN unidimensional y luego una capa de integración. La integración de sentencias alimenta la capa de LSTM. Finalmente, la clasificación donde el modelo alimenta las sentencias aprendidas. En comparación a los tres modelos el que dio mejores resultados fue el modelo híbrido C-LSTM.

## D. Detection of Suicidal Ideation on Twitter using Machine Learning and Ensemble Approaches

El suicidio es uno de los principales problemas de salud pública que consume muchas vidas. Según las estadísticas de la Organización Mundial de la Salud (OMS), alrededor de un millón de personas mueren por suicidio cada año.

La OMS menciona que el suicidio es la principal causa de muerte entre los adolescentes y la sexta entre los adultos. Los factores incluyen problemas personales como desesperanza, abuso de sustancias, ansiedad, esquizofrenia; factores sociales como el aislamiento de la sociedad, la pérdida de seres queridos, el desempleo, etc. A pesar del creciente número de casos de suicidio, se puede prevenir hasta cierto punto mediante la comprensión de los factores de riesgo relacionados con el comportamiento suicida en las primeras etapas del proceso suicida. El proceso de suicidio comienza con pensamientos o ideas suicidas. Luego madura hasta el intento de suicidio y finalmente hasta el suicidio consumado [1].

Hay una creciente inclinación de las personas a formar comunidades en línea e interactuar entre sí. Las investigaciones han revelado que las personas se sienten cómodas para hablar sobre tendencias suicidas en línea en lugar de hablar de ello cara a cara; la razón es una sensación de autocontrol y la característica anónima de las redes sociales. Existen varios casos en los que las víctimas de suicidio también revelaron sus sentimientos finales antes de su muerte en las redes sociales.



El trabajo de investigación hace contribuciones novedosas a la literatura ya existente de las siguientes maneras.

- El trabajo extrae una serie de características relevantes para diferenciar entre tweets suicidas y no suicidas con la ayuda de una nueva técnica de ingeniería de características.
- El documento construye un nuevo conjunto de datos extrayendo los tweets relacionados con el suicidio y el no suicidio de Twitter. El conjunto de datos se anota y luego se usa para entrenar su modelo para la diferenciación binaria (suicida y no suicida) de los tweets. El trabajo revela el conocimiento oculto desde la perspectiva de la minería de datos.
- La evaluación comparativa se realiza en varios métodos de aprendizaje automático y conjuntos para distinguir su eficacia.

En el documento, se implementan varios algoritmos de aprendizaje automático y enfoques de conjuntos como arboles de decisión, Naive Bayes, Multinomial Naive Bayes, Support vector machine (SMO), Regresion, Bagging, Random Forest, AdaBoost, votacion y Stacking para clasificar los tweets relacionados con el suicidio. en dos grupos usando los tweets reales extraídos a través de la API de Twitter. Random Forest resulto ser el m as efectivo y funciona mejor con una precisión del 98,5%, una precisión del 98,7% y un valor de recuperación del 98,2%.

## DISEÑO

Dada la comparativa entre los dos modelos se hicieron dos pipelines, uno para Roberta sin un preprocesamiento y otro para Bert con los datos limpios, de esta forma medir la eficiencia de cada modelo.

### A. Pipeline Bert

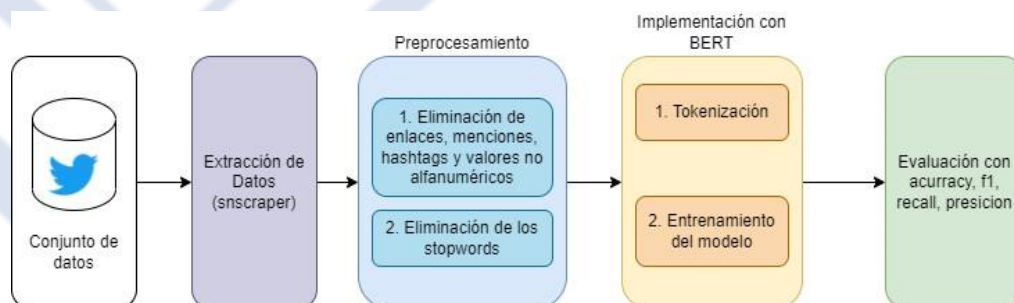


Figura 1. Marco de Trabajo para la Detección de suicidio en tweets con Bert

### B. Pipeline Roberta



## IMPLEMENTACION Y RESULTADOS

Se utilizó Google Colab para entrenar los modelos, ah  
utilizaron los códigos de (referencias a los códigos). En ambos casos se utilizaron los recursos de la GPU para el entrenamiento del modelo.

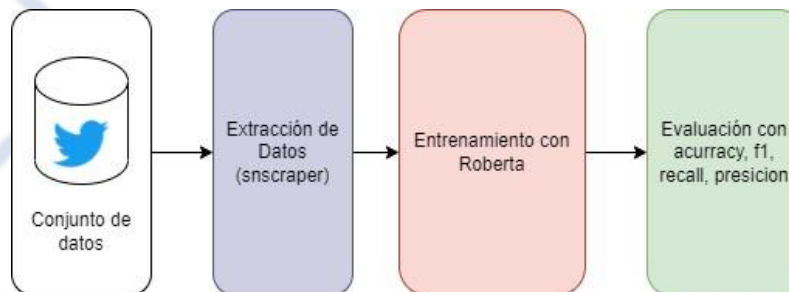


Figura 2. Marco de Trabajo para la Detección de suicidio en tweets con Roberta

### A. Bert

BERT, que significa Bidirectional Encoder Representations from Transformers, es un modelo de lenguaje basado en Transformers que fue presentado por Google en 2018. Representa un avance significativo en el procesamiento del lenguaje natural y ha demostrado un rendimiento excepcional en una amplia gama de tareas relacionadas con el lenguaje. Por eso es ampliamente utilizado y reconocido por su capacidad para comprender y generar texto en un contexto más profundo.

#### Ventajas:

- **Contexto bidireccional:** Una de las principales ventajas de BERT es su capacidad para comprender el contexto en ambos sentidos, es decir, tanto las palabras que lo preceden como las que lo siguen. Esto le permite capturar las sutilezas de las relaciones entre palabras y generar representaciones de palabras más ricas y precisas.
- **Representaciones profundas:** BERT utiliza una arquitectura basada en Transformers con múltiples capas, lo que le permite capturar relaciones semánticas y sintácticas complejas en el texto. A través del entrenamiento en grandes cantidades de datos, BERT desarrolla una comprensión profunda del lenguaje y es capaz de inferir significados incluso en oraciones con estructuras gramaticales complejas.
- **Preentrenamiento y adaptabilidad:** BERT se preentrena en grandes corpus de texto sin etiquetas, lo que le permite aprender representaciones generales del lenguaje. Una vez preentrenado, se puede adaptar a tareas específicas de NLP mediante un proceso de afinación, lo que permite su uso en una variedad de aplicaciones y dominios.





- **Rendimiento sobresaliente:** BERT ha demostrado un rendimiento excepcional en una amplia gama de tareas de procesamiento del lenguaje natural, como la clasificación de texto, el etiquetado de entidades, la respuesta a preguntas y la traducción automática. Su capacidad para comprender el contexto y las relaciones complejas entre palabras lo convierte en una opción destacada en comparación con modelos previos.

## Desventajas:

- **Alta demanda computacional:** Debido a su arquitectura basada en Transformers y su gran tamaño, BERT requiere una cantidad significativa de recursos computacionales para entrenar y utilizar. Esto puede ser una limitación en entornos con recursos limitados o en aplicaciones que requieren un procesamiento rápido en tiempo real.
- **Dependencia de grandes conjuntos de datos:** Para alcanzar su máximo rendimiento, BERT necesita ser entrenado en grandes cantidades de datos. Esto puede ser un desafío en dominios de nicho o en idiomas con recursos limitados, donde puede ser difícil reunir suficiente información para entrenar el modelo de manera efectiva.
- **Dificultad para manejar ambigüedades:** Aunque BERT tiene un rendimiento sobresaliente, aun puede tener dificultades para manejar casos de ambigüedad en el lenguaje. En situaciones donde el contexto no es suficiente para determinar el significado exacto de una palabra o frase, BERT puede producir resultados imprecisos o inconsistentes.
- **Falta de aplicabilidad:** Dado que BERT es un modelo de aprendizaje automático de caja negra, puede resultar difícil comprender como toma decisiones o llega a ciertas conclusiones. Esto puede plantear desafíos en escenarios donde se requiere una explicación clara y transparente del proceso de toma de decisiones.

## B. RoBERTa

RoBERTa es un modelo de lenguaje basado en Transformers que se basa en el éxito de BERT y mejora a un mas su rendimiento. Fue presentado por Facebook AI en 2019 y se enfoca en el preentrenamiento sin supervisión para aprender representaciones de alta calidad del lenguaje.

## Ventajas:

- **Mejor preentrenamiento:** RoBERTa utiliza una estrategia de preentrenamiento más extensa en comparación con BERT. Se entrena en grandes conjuntos de datos sin etiquetas, utilizando técnicas como el entrenamiento de máscara de palabras (masking) y la predicción de oraciones siguientes. Esto resulta en representaciones de palabras más sólidas y una mejor comprensión del lenguaje.
- **Mayor capacidad de generalización:** Al tener un enfoque de preentrenamiento más exhaustivo, RoBERTa logra una mayor capacidad de generalización. Puede capturar mejor las relaciones semánticas y sintácticas complejas en el texto, lo que le permite





adaptarse a diferentes tareas de procesamiento del lenguaje natural con un rendimiento mejorado.

- **Afinación mas efectiva:** RoBERTa se beneficia de técnicas de afinación mejoradas. Después del preentrenamiento, el modelo se ajusta a tareas específicas de NLP mediante un proceso de afinación en conjuntos de datos etiquetados más grandes y diversificados. Esto permite una adaptación más efectiva y un mejor rendimiento en comparación con BERT.
- **Resultados lideres en el estado del arte:** RoBERTa ha logrado resultados sobresalientes en una amplia gama de tareas de procesamiento del lenguaje natural, incluyendo clasificación de texto, extracción de información on, traducción automática y mas. Su capacidad para capturar el contexto y las relaciones complejas en el lenguaje lo coloca en la vanguardia del rendimiento en NLP.

## Desventajas:

- **Requiere recursos computacionales significativos:** Al igual que BERT, RoBERTa es computacionalmente intensivo y requiere una gran cantidad de recursos para su entrenamiento y uso eficiente. Esto puede limitar su aplicabilidad en entornos con recursos limitados o en aplicaciones que necesitan un procesamiento rápido en tiempo real.
- **Dificultad para la interpretación y explicación:** Debido a su complejidad y enfoque de aprendizaje automático de caja negra, RoBERTa puede ser difícil de interpretar y explicar. Comprender como el modelo toma decisiones o llega a conclusiones específicas puede resultar desafiante, lo que puede ser problemático en aplicaciones que requieren transparencia y aplicabilidad.
- **Dependencia de grandes conjuntos de datos:** Al igual que BERT, RoBERTa también se beneficia del entrenamiento en grandes conjuntos de datos. Esto puede ser un desafío en dominios o idiomas con recursos limitados, donde puede ser difícil reunir suficiente información para el entrenamiento efectivo del modelo.
- **Sensibilidad al ruido en los datos de entrenamiento:** RoBERTa puede ser sensible al ruido presente en los datos de entrenamiento. La presencia de datos incorrectos o mal etiquetados puede afectar negativamente su rendimiento y conducir a resultados imprecisos o inconsistentes en ciertos casos.

## Resultados

Se utilizo sklearn para evaluar ambos modelos. Los resultados de ambos modelos fue el siguiente, Bert obtuvo 0.96684 su mejor valor en precisión lo que significa que el modelo tiene una baja tasa de falsos positivos; mientras que Roberta obtuvo 0.95800 en recall, indica que el modelo es efectivo al capturar la mayoría de las instancias positivas y tiene una baja tasa de falsos negativos.



Tabla I Resultados de los modelos.

• Modelo	• Accuracy	• F1	• Recall	• Precision
• Bert	• 0.95802	• 0.95756	• 0.95802	• 0.96684
• Roberta	• 0.93798	• 0.94447	• 0.95800	• 0.93187
• Roberta • DataSet Limpio	• 0.954039	• 0.95213	• 0.95486	• 0.951526

## CONCLUSIONES

El modelo Bert muestra una mayor precisión (0.96684) y exactitud (0.95802), lo que lo hace adecuado para aplicaciones donde minimizar los falsos positivos y lograr una precisión global en la clasificación son cruciales. Sin embargo, muestra un nivel ligeramente inferior de recall (0.95802) en comparación con Roberta, lo que indica que podría perder algunos casos positivos. Además, su puntuación F1 (0.95756) sugiere un ligero equilibrio entre precisión y recall.

Por otro lado, el modelo Roberta demuestra un mayor recall (0.95800) y puntuación F1 (0.94447), lo que indica su fortaleza al capturar casos positivos y mantener un buen equilibrio entre precisión y recall. Esto lo convierte en una opción confiable cuando reducir los falsos negativos es importante. Sin embargo, tiene una precisión más baja (0.93187) y una exactitud menor (0.93798) en comparación con Bert, lo que sugiere una posible mayor tasa de falsos positivos y un mayor índice general de clasificación incorrecta.

Es importante tener en cuenta que la elección entre Bert y Roberta debe basarse en los requisitos específicos de la tarea en cuestión, así como en la importancia relativa de la precisión, el recall y la exactitud en el contexto de la aplicación.

En términos de precisión y exactitud, tanto Bert como Roberta muestran resultados inferiores en comparación con el modelo Random Forest mencionado en el artículo, que logra una precisión del 98,7%. Esto sugiere que el modelo Random Forest es más preciso y capaz de clasificar los casos correctamente con mayor certeza que los modelos de lenguaje natural como Bert y Roberta. Sin embargo, se debe tener en cuenta que los modelos de lenguaje natural como Bert y Roberta pueden tener otras ventajas, como la capacidad de comprender y generar texto en lenguaje natural de manera más general, lo que los hace más versátiles en diferentes aplicaciones. Respecto a las palabras más recurrentes se recopilaron en una imagen.



## Pensamientos Suicidas



Figura 3. Palabras Recurrentes

### TRABAJOS FUTUROS

Para los trabajos futuros hay varias consideraciones a tener en cuenta al trabajar con datos relacionados con la ideación suicida en Twitter, especialmente teniendo en cuenta los desafíos encontrados durante la recopilación de datos.

En primer lugar, sería beneficioso expandir el conjunto de datos para mejorar el rendimiento de los modelos. Dado la dificultad para obtener datos relacionados con los pensamientos suicidas, se deben realizar esfuerzos para recopilar un conjunto de datos más amplio para las categorías de activo, pasivo, sarcasmo y relaciones dado que la suma de estos datos no se compara con la categoría de otros que son la mayor parte. Esto podría implicar colaborar con organizaciones de salud mental, investigadores o utilizar técnicas de búsqueda avanzadas para capturar una gama más amplia de tweets relacionados con la ideación suicida.

En segundo lugar, considerando la naturaleza sensible del tema, sería importante incorporar medidas adicionales para garantizar el manejo ético de los datos. Esto podría incluir la implementación de técnicas robustas de anonimización para proteger la privacidad y confidencialidad de las personas involucradas en los tweets. Colaborar con profesionales de salud mental y expertos en ética podría proporcionar orientación valiosa para desarrollar protocolos adecuados para la recopilación y el uso de datos.



Además, explorar la información contextual y los matices del lenguaje específicos de Twitter podría ser beneficioso para mejorar la comprensión de los modelos sobre los diferentes tipos de ideación suicida. Con ello obtener más información sobre que características tienen las personas relacionado a estos tipos de pensamientos. Esto podría mejorar la capacidad de los modelos para clasificar y diferenciar con precisión entre la ideación suicida activa y pasiva, el sarcasmo, las discusiones relacionadas con el suicidio y otras categorías relevantes.

## Contribución de Autoría

**Juliana Berrios Butron:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Daria López Franco:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Dante Rodolfo Tárraga:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1] Rabani, S. T., Khan, Q. R., & Khanday, A. M. U. D. (2020). Detection of suicidal ideation on Twitter using machine learning & ensemble approaches. *Baghdad Science Journal*, 17(4), 1328-1339.
- [2] M. J. Vioules, B. Moulahi, J. Az` e and S. Bringay, (2018). Detection of suicide-related posts in Twitter data streams. *IBM Journal of Research and Development*, vol. 62, no. 1, pp. 7:1-7:12, 1 Jan.-Feb. 2018, doi: 10.1147/JRD.2017.2768678.
- [3] Sawhney, R., Manchanda, P., Mathur, P., Shah, R., & Singh, R. (2018, October). Exploring and learning suicidal ideation connotations on social media with deep learning. In *Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis* (pp. 167-175).
- [4] Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: a narrative review. *NPJ digital medicine*, 5(1), 46.



- [5] Otter, D. W., Medina, J. R., & Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *IEEE transactions on neural networks and learning systems*, 32(2), 604-624.
- [6] Lech, N. (2008). Suicidal ideation on Twitter. [Database]. <https://www.kaggle.com/datasets/natalialech/suicidal-ideation-on-twitter><https://www.kaggle.com/datasets/natalialech/suicidal-ideation-on-twitter>
- [7] Rabani, S. T., Khan, Q. R., & Khanday, A. M. U. D. (2020). Detection of suicidal ideation on Twitter using machine learning & ensemble approaches. *Baghdad Science Journal*, 17(4), 1328-1339.
- [8] Coco C. C., Joel A. P., Carlos S. Francklin C. H., Waldir P. H. (2022). Suicidios en el Perú: Descripción epidemiológica a través del Sistema Informático Nacional de Defunciones (SINADEF) en el periodo 2017- 2021.





# Aplicación de modelo de regresión lineal para predecir el índice de popularidad en la plataforma Spotify

121

*Linear Regression application to predict the popularity index in Spotify*

## Cesar Vasquez Alvarez

Universidad Nacional de San Agustín.

@cvasqueza@unsa.edu.pe

## Edith M. Coaquira Cuevas

Universidad Nacional de San Agustín.

@ecoquiracu@unsa.edu.pe

## Emerson D. Mendoza Hilasaca


Universidad Nacional de San Agustín.


@emendozahi@unsa.edu.pe


## Jeffrey J. Pinto Ñaupá

Universidad Nacional de San Agustín.

@jpinton@unsa.edu.pe

 **ARK:** [ark:/42411/s12/a110](https://nbn-resolving.org/urn:nbn:pe:unsa:42411-s12-a110)

 **DOI:** [10.48168/innosoft.s12.a110](https://doi.org/10.48168/innosoft.s12.a110)

 **PURL:** [42411/s12/a110](https://nbn-resolving.org/urn:nbn:pe:unsa:42411-s12-a110)

RECIBIDO 07/05/2023 • ACEPTADO 15/08/2023 • PUBLICADO 30/09/2023



## RESUMEN

En la actualidad los servicios de música en streaming se han convertido en uno de los principales medios de consumo de música alrededor del mundo. Spotify ofrece servicios de transmisión de música y abarca más de treinta millones de canciones. Cada año hay un incremento en la producción de canciones por lo cual es más difícil que una canción se establezca como un hit en el mercado. El presente trabajo tuvo como objetivo aplicar la técnica de modelado de Regresión Lineal para encontrar una tendencia del conjunto de datos sobre el índice de popularidad de las canciones en la plataforma Spotify, de esta manera predecir un resultado con nuevos datos que ingresen. Se aplicó una metodología cuantitativa basada en datos medibles que se tomaron como datasets. Como resultado se obtuvo un error cuadrático medio de 94.79 y una varianza de 0.20. La conclusión del trabajo es que el dataset utilizado no fue el ideal acorde a nuestro objetivo.

**Palabras claves:** Python, Regresión Lineal, Predicción.



## ABSTRACT

*Currently, streaming music services have become one of the main means of music consumption around the world. Spotify offers music streaming services and covers more than thirty million songs. Every year there is an increase in the production of songs so it is more difficult for a song to establish itself as a hit in the market. The objective of this work was to apply the Linear Regression modeling technique to find a trend of the data set on the popularity index of songs on the Spotify platform, in this way predict a result with new data that enters. A quantitative methodology was applied based on measurable data that were taken as datasets. As a result, a mean square error of 94.79 and a variance of 0.20 were obtained. The conclusion of the work is that the dataset used was not the ideal according to our objective.*

**Keywords:** Python, Linear Regression, Predict.

---

## INTRODUCCIÓN

A lo largo del tiempo, la industria musical ha pasado por revoluciones tecnológicas y comunicativas. La industria musical está experimentando un fuerte crecimiento gracias al avance de Internet y el surgimiento de nuevas tecnologías de la comunicación, como son las redes sociales o diversas plataformas de contenido, como Spotify o YouTube [1]. La forma de consumo de música ha cambiado y en la actualidad, los servicios de música de streaming se han convertido en uno de los principales medios en los que se consume música alrededor de todo el mundo y su crecimiento no muestra señales de estancamiento. En la investigación [2] menciona los servicios más populares de plataformas de música, entre los cuales Spotify lidera el ranking seguido de sus principales competidores Apple Music y Amazon Music.

Smith et al. [3] mencionan a Spotify como una empresa de software el cual ofrece servicios de transmisión de música, Spotify es lanzado en el 2008 y en los siguientes 10 años tuvo un crecimiento continuo convirtiéndose en un icono para una nueva generación de organizaciones ágiles. Sus oficinas de desarrollo e investigación cuentan con una forma de trabajo y estructuras organizativas con un diseño que promueve la colaboración, innovación y autonomía.

El éxito de las canciones es regido en una serie de factores tal como lo presenta Interiano et al [4], hicieron el análisis de más 500000 canciones lanzadas en el Reino Unido entre los años 1985 y 2015 para ello usaron la técnica de Machine Learning "bosques aleatorios" para predecir el éxito de las canciones que le permitió cuantificar la contribución de las características puramente musicales en el éxito de las canciones. Entonces sugiriendo una escala temporal de la dinámica de la moda en la música exitosa. También mencionan el descubrimiento de varias tendencias multi décadas que tuvieron relevancia para comprender la dinámica del éxito el cual define como



ocupar puestos en las listas principales, correlacionar el éxito con las características acústicas y explorar la previsibilidad del éxito.

En el campo laboral de empresas de servicios multimedia descritas por Braga [5] tales como Spotify, Apple, Amazon o youtube music han tenido un crecimiento drástico en los últimos años principalmente debido a que el mundo se ha vuelto cada vez más digital con el tiempo, desarrollando así tecnologías de rápido desarrollo y fácil acceso, a lo que el autor también nos brinda un panorama de la constante competencia de estas empresas por liderar el mercado pero entre estas empresas está Spotify ya que destaca por ser una compañía pionera del rubro y de momento ninguna plataforma hace temblar su dominio.

En el caso de Lopes y Simoes [6] están de acuerdo que la industria de la música ha experimentado enormes cambios en relación con sus hábitos de producción, distribución y consumo debido al desarrollo exponencial de las plataformas de streaming, ellos a través de estas buscaron analizar qué factores influyen en la intención de adquirir un servicio de streaming de música y, en consecuencia, en su recomendación pues debido a la variedad de competencia viene a ser un algo complejo y multidimensional a esto también se cuenta el factor de las estrategias que emplea cada empresa para captar más clientes además de la implementación de algoritmos que buscan la satisfacción del cliente pues cada plataforma usa varios métodos que resuelvan estas dudas.

Para Twomey y Kroll [7], el método de Regresión Lineal, juega un papel importante en estudios que involucren datos cuantitativos, si bien el método no es ideal, el uso de varias técnicas pueden ayudar a superar sus debilidades. En adición Sravani y Bola [8] indican que la regresión lineal es un método perteneciente a algoritmos de Machine Learning, se basa en aprendizaje supervisado, este método define la relación entre 2 variables (una dependiendo de la otra) ajustando la línea de regresión a los datos, por otro lado Regresión lineal Múltiple utiliza múltiples variables independientes unas de otras.

Este trabajo tiene como objetivo usar el algoritmo de Regresión Lineal para encontrar una tendencia del conjunto de datos sobre el éxito de canciones en Spotify y de esta manera poder predecir un resultado con nuevos datos que puedan entrar.

## Materiales y métodos

### Estado del Arte

Existen trabajos como los de Interiano et al [4] en la cual realizaron un análisis de un gran repertorio de músicas lanzadas los años 1985 y 2015 en el Reino Unido para comprender la dinámica del éxito, correlacionar el éxito con las características acústicas y explorar la previsibilidad del éxito. Haciendo uso de la técnica de Machine Learning "bosques aleatorios"



teniendo en cuenta la función de sus características acústicas y agregando variables permitió cuantificar la contribución de las características puramente musicales en el éxito de las canciones y sugirió la escala temporal de la dinámica de la moda en la música popular.

Por otra parte Hernández y Beltrán [9] desarrollaron el modelo de dos redes neuronales que identifican las transiciones de las diferentes partes de la estructura de las piezas musicales y las diferencias entre las transiciones para etiquetarlas, hicieron uso de técnicas de aprendizaje profundo y aprendizaje automático con Pytorch sus resultados obtenidos fueron similares al estado del arte que tomaron como caso similar.

Asimismo Quin et al.[10] realizan una predicción de tendencias de géneros musicales según la influencia y la similitud de la música a través del análisis de datos, haciendo uso de métodos como Deepwalk, Cosine Similarity, modelo de similitud musical basado en el Análisis de Componentes Principales y la distancia Euclidiana. Los resultados que obtuvieron fueron que las tendencias del género musical se resumen y predicen a partir de la influencia y similitud de la música.

Se tomaron como referencia las investigaciones anteriores mencionadas porque se enfocan en un análisis de diferentes puntos de vista de la música, también aporta experiencia a nuestro trabajo de investigación la cual se centra en el análisis y predicción del índice de popularidad a partir de características o atributos de un tema musical.

## Marco teórico

### Inteligencia Artificial

La Inteligencia Artificial (IA) es una rama de las ciencias computacionales encargada de estudiar modelos de cómputo capaces de realizar actividades propias de los seres humanos en base a dos de sus características primordiales: el razonamiento y la conducta [11].

### Data reduction

Es la transformación de información digital numérica o alfabética derivada empírica o experimentalmente en una forma corregida, ordenada y simplificada [12]. El propósito de la reducción de datos puede ser doble: reducir la cantidad de registros de datos mediante la eliminación de datos no válidos o producir datos de resumen y estadísticas en diferentes niveles de agregación para varias aplicaciones.

### Data cleaning

Es el proceso de detectar, corregir o eliminar registros inexactos de un conjunto de registros, tabla o base de datos y se refiere a identificar partes incompletas, incorrectas, inexactas o





irrelevantes de los datos y luego reemplazar, modificar, o eliminar los datos sucios o gruesos. Data cleaning se puede realizar de forma interactiva con herramientas de gestión de datos, o como procesamiento por lotes a través de secuencias de comandos o un firewall de calidad de datos [13].

## Variables dummy

Permiten medir el efecto de una característica de determinados individuos en determinada muestra [14]. También son definidas como una variable utilizada para explicar valores cualitativos en un modelo de regresión, son variables las cuales suelen tomar valores binarios [15].

## Herramientas

### Regresión Lineal

Según la página MathWorks [16] la define como una técnica de modelado estadística la cual es empleada para describir una variable de respuesta continua como una función de una o varias variables predictoras. Asimismo ayudan a comprender y predecir el comportamiento de sistemas complejos, como también analizar datos experimentales, financieros y biológicos.

Por otra parte Gupta, García y Chin [17] mencionan sobre el amplio uso que tiene en la estimación estadística también sobre los beneficios que tiene su modelo lineal como son la simplicidad y facilidad de uso del método, asimismo recalcan que su principal inconveniente es su alta desviación del modelo lo cual hace que pueda producir resultados deficientes.

### Python

Es un lenguaje de alto nivel de programación interpretado cuya filosofía hace hincapié en la legibilidad de su código, se utiliza para desarrollar aplicaciones de todo tipo [18]. Es un lenguaje de programación multiparadigma, ya que soporta parcialmente la orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, dinámico y multiplataforma.

### Google Colab

Es una herramienta que te permite ejecutar scripts de Python a través de los servidores de Google. Esto te permite ejecutar celdas de código como si se tratara de un cuaderno de Jupyter Notebook. Pero no sólo eso, Google Colab es perfecto para implementar algoritmos de aprendizaje máquina, ya que no te limitas a los recursos de tu computadora. Esto quiere decir que tienes a tu disposición el GPU y las TPU's de Google para potencializar el cómputo de tu proyecto [19].





## **Numpy**

Es una biblioteca para el lenguaje de programación Python que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas [20].

## **Pandas**

Pandas es un paquete de Python para la ciencia de datos y Machine Learning, ofrece estructuras de datos poderosas, expresivas y flexibles que facilitan la manipulación y análisis de datos [21]. Es una biblioteca de código abierto que proporciona herramientas de análisis y manipulación de datos de alto rendimiento utilizando sus potentes estructuras de datos.

## **Matplotlib**

Matplotlib es una biblioteca Python open source que permite crear visualizaciones de datos [22]. Esta biblioteca sirve para generar gráficas a partir de datos contenidos en listas, vectores, en el lenguaje de programación Python y en su extensión matemática NumPy.

## **Skelearn**

La librería scikit-learn, también llamada sklearn, es un conjunto de rutinas escritas en Python para hacer análisis predictivo, que incluyen clasificadores, algoritmos de clusterización, etc. Está basada en NumPy, SciPy y matplotlib, de forma que es fácil aprovechar el código que usan estas librerías [23].

## **Seaborn**

Es una librería de visualización de datos para Python desarrollada sobre matplotlib. Ofrece una interfaz de alto nivel para la creación de atractivas gráficas. Además, está íntimamente integrada con las estructuras de datos de pandas, lo que permite utilizar el nombre de los Data Frames y campos directamente como argumentos de las funciones de visualización [24].

## **Resultados y discusión**

### **Tratamiento de datos**

En procesos de aprendizaje automático, el tratamiento de datos o también llamado pre procesamiento tiene como objetivo manipular y/o transformar los datos sin procesar para evitar alterar el aprendizaje y futuras predicciones [25].



El dataset utilizado para este trabajo contiene inicialmente 2000 datos y 18 variables, entre los cuales, al momento de graficarlas se pudo apreciar que contenían valores anómalos los cuales claramente tenían que ser tratados.

## Transformación de datos

Primero, importamos el dataset y las librerías necesarias.

```
import numpy as np
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt
%matplotlib inline
from mpl_toolkits.mplot3d import Axes3D
from matplotlib import cm
plt.rcParams['figure.figsize'] = (16, 9)
plt.style.use('ggplot')
from sklearn import linear_model
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score

from google.colab import drive
drive.mount('/content/drive')

data = pd.read_csv("/content/drive/My Drive/songs_normalize_new(2).csv")
```

Figura 1. Importación de librerías y dataset

Segundo, identificamos las variables adecuadas y necesarias para aplicar el modelo, las cuales fueron todas aquellas que contienen valores continuos, eliminando así las variables Nombre del Artista, Nombre de la canción, Explícito y Género.

```
data.drop(['artist', 'song', 'explicit', 'genre'],1)
```

Figura 2. Eliminación de variables irrelevantes

Tercero, identificamos las variables dependientes e independientes.

### Variables Independientes:

- x<sub>1</sub>: Duración
- x<sub>2</sub>: Año
- x<sub>3</sub>: Bailabilidad



- x<sub>4</sub>: Energía
- x<sub>5</sub>: Tonalidad
- x<sub>6</sub>: Volumen
- x<sub>7</sub>: Modalidad
- x<sub>8</sub>: Pronunciación
- x<sub>9</sub>: Acústica
- x<sub>10</sub>: Instrumentalidad
- x<sub>11</sub>: En vivo
- x<sub>12</sub>: Valencia
- x<sub>13</sub>: Tempo

### Variables Dependientes:

y<sub>1</sub>: Popularidad

Cuarto, graficamos las variables por medio de histogramas para buscar posibles valores anómalos.

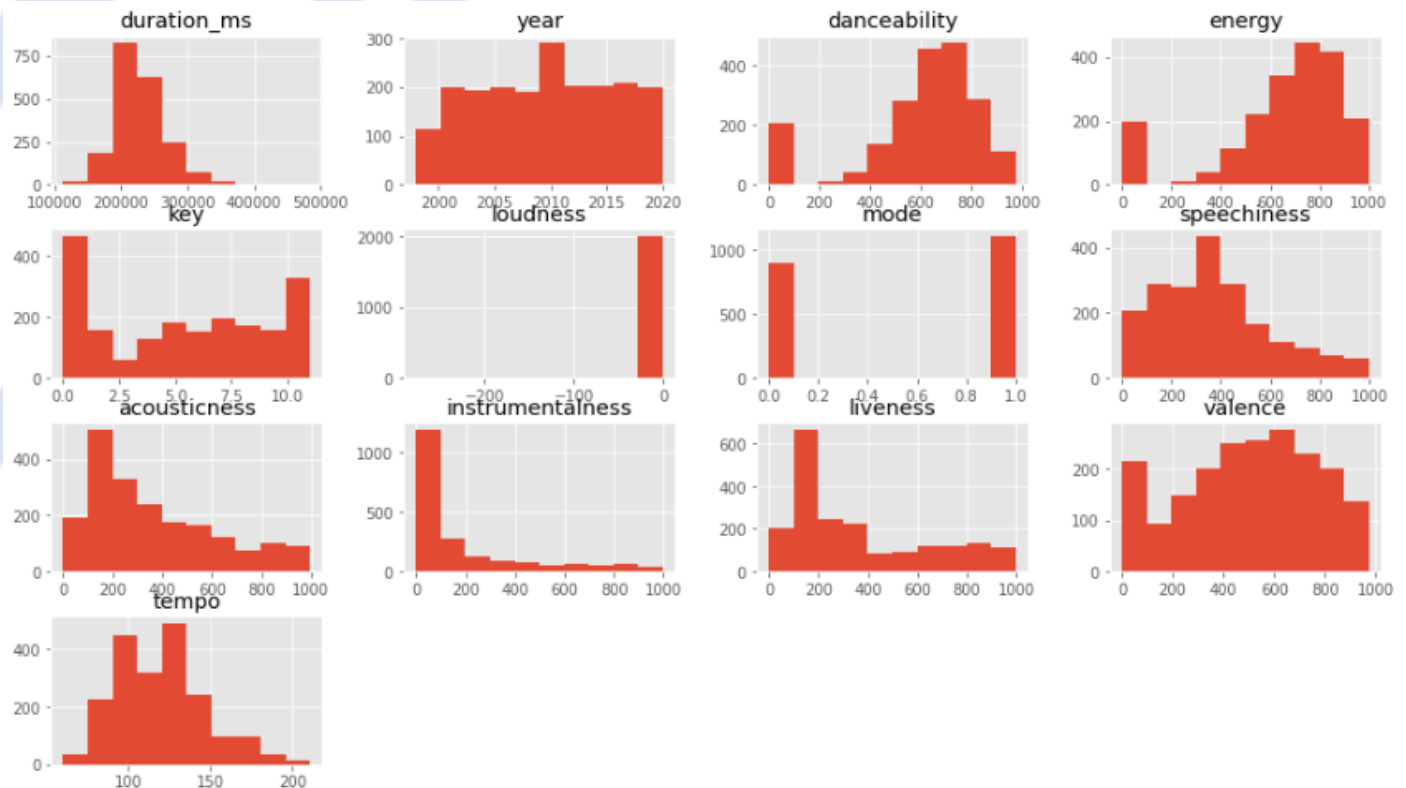


Figura 3. Histograma de variables independientes



Quinto, procedemos a eliminar conjuntos de valores que estén alejados del grupo o iguales a cero, como en las variables Bailabilidad (danceability), Energía (energy), Tonalidad (key), Instrumentalidad (instrumentalness), como también a disminuir el rango de las variables Duración (duration\_ms), Volumen (loudness) para tener una mejor distribución.

```
filtered_data = data[(data['duration_ms'] < 390000) & (data['loudness'] > -15) & (data['instrumentalness'] >= 100) & (data['danceability'] >= 180) & (data['energy'] >=180)& (data['key'] >=1.5)]
```

Figura 4. Data Cleaning y Data Reduction

364.56 ecm 0.04 varianza

Sexto, procedemos a crear y entrenar el modelo de Regresión Lineal.

```
#CREAMOS EL MODELO
model=linear_model.LinearRegression()
#ENTRENAMOS (predictoras, salida)
model.fit(XY_train,z_train)
#PREPARAMOS PARA LA PREDICCIÓN
z_pred=model.predict(XY_train)
```

Figura 5. Creación y entrenamiento del modelo

Séptimo, calculamos métricas para analizar el estado del modelo, Error Cuadrático Medio (MSE, por sus siglas en inglés) y la varianza.

```
# Error cuadrático medio cerca o =0
print("Error cuadrático medio: %.2f" % mean_squared_error(z_train, z_pred))
# Evaluamos el puntaje de varianza
print('Varianza: %.2f' % r2_score(z_train, z_pred))
```

```
Error cuadrático medio: 364.56
Varianza: 0.04
```

Figura 6. Métricas



Octavo, observamos que nuestro MSE es demasiado elevado por lo que decidimos agregar la variable género al modelo. Esta variable al ser de carácter nominal tuvo que ser tratada con Dummy Coding, generando así 11 columnas adicionales a las 13 ya tratadas.

```
# La columna genre es nominal (aplicamos dummy coding)
genre_dummy=pd.get_dummies(filtered_data["genre"],prefix="Género")
filtered_data=pd.concat([filtered_data,genre_dummy],axis=1)
```

Figura 7. Dummy coding

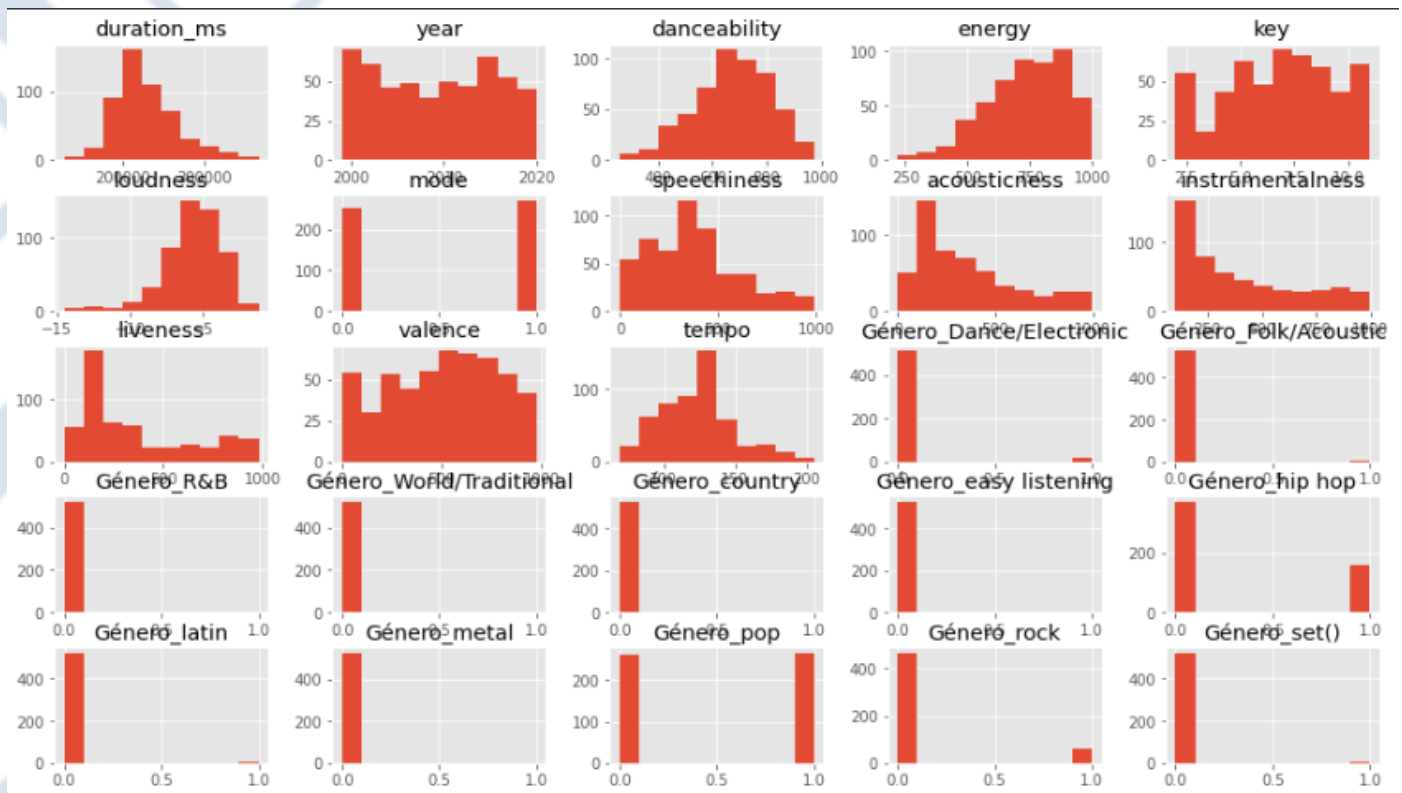


Figura 8. Histogramas del total columnas

Noveno, calculamos y revisamos las métricas nuevamente.

```
Error cuadrático medio: 347.80
Varianza: 0.08
```

Figura 9. Métricas con Dummy Coding





Observamos que se pudo disminuir el MSE, pero aún así sigue siendo demasiado alto. Décimo, tras una serie de pruebas de datos se incluyó la variable Explícito aplicando un mapeo para transformar sus valores de Verdadero y Falso como string a valores booleanos. Además, en la variable tonalidad, sus valores iguales a 0 se igualaron a su media para evitar eliminarlos y disminuir aún más el número de datos.

```
data2["Explícito"]=list(map(lambda ele:ele=="VERDADERO",filtered_data["explicit"])))
```

Figura 10. Mapeo de la variable Explícito

```
data2["Tonalidad"]=list(map(lambda x: 5.378000 if x==2 else x, filtered_data["key"])))
```

Figura 11. Cambio de valores de 0 a media

Obteniendo con estas nuevas variante, nuevas métricas

```
Error cuadrático medio: 94.79  
Varianza: 0.20
```

Figura 11. Métricas finales

Consideramos estos valores aún muy altos, pero si hacemos una comparación con las métricas iniciales se logró una reducción del 75% del MSE.

Finalmente, procedemos a realizar una predicción, que si bien no será exacta (por el valor del MSE) será mejor que la predicción que se pudo obtener con el modelo inicial.

```
z_Dosmil = model.predict([[300000, True, 2020, 0.7, 0.1, 8, -10, 1, 0.5, 0, 0.5, 150, 0, 0, 0, 0, 0, 0, 0, 1, 0]])  
print(int(z_Dosmil))  
-466
```

Figura 12. Predicción



## Conclusiones

1. Regresión lineal es un modelo que se usa para predecir la dependencia entre dos o más variables continuas.
2. Para desarrollar un modelo de regresión lineal ideal se debe procurar tener un dataset con valores continuos y no tan dispersos. De no ser así, procurar tener más de 2000 datos para que al realizar data cleaning y/o data reduction no se tenga al final, por ejemplo, un aproximado de 400 datos como en este trabajo.
3. En este trabajo, al aplicar dummy coding a variables nominales se pudo lograr una mejora en las métricas del modelo.
4. Consideramos que si se hubiera tratado la variable "Nombre del Artista" con algún método nuestro modelo mejoraría considerablemente.
5. Este conjunto de datos no es ideal para usar un modelo de Regresión Lineal.

## Contribución de Autoría

**Cesar Vasquez Alvarez:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Edith Coaquira Cuevas:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Emerson Mendoza Hilasaca:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#). **Jeffrey Pinto Ñaupa:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura](#), [revisión y edición](#).

## Referencias

- [1] Gómez Herrero, R. (2021). Evolución de la Industria Musical. Siglo XX-Siglo XXI. UVaDOC Principal. <https://uvadoc.uva.es/handle/10324/48012> [Accessed: June 22, 2022].
- [2] García Pizarro, A. (2021). El auge de la música en streaming. UVaDOC Principal. <https://uvadoc.uva.es/handle/10324/51809> [Accessed: June 22, 2022].



- [3] D. Smite, N. B. Moe, G. Levinta and M. Floryan, "Spotify Guilds: How to Succeed With Knowledge Sharing in Large-Scale Agile Organizations," in *IEEE Software*, vol. 36, no. 2, pp. 51-57, March-April 2019, doi: 10.1109/MS.2018.2886178.
- [4] Interiano, M., Kazemi, K., Wang, L., Yang, J., Yu, Z., & Komarova, N. L. (2018). Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society Open Science*, 5(5), 171274. doi:10.1098/rsos.171274
- [5] . M. Braga, "Spotify vs. Apple : a battle of titans", doctoral thesis, Universidade de Catolica Portuguesa, 2021.
- [6] M. Lopes Barata y P. Simões Coelho, "Music streaming services: understanding the drivers of customer purchase and intention to recommend", *ScienceDirect*, Volume 7, Issue 8, agosto de 2021, art. n.º e07783.
- [7] Golbaz, S., Nabizadeh, R., & Sajadi, H. S. (2019). Comparative study of predicting hospital solid waste generation using multiple linear regression and artificial intelligence. *Journal of Environmental Health Science and Engineering*, 17(1), 41-51.
- [8] Sravani, B., & Bala, M. M. (2020, June). Prediction of student performance using linear regression. In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.
- [9] Hernández Oliván, C., & Beltrán Blázquez, J. R. Análisis musical mediante inteligencia artificial.
- [10] C. Qin, H. Yang, W. Liu, S. Ding and Y. Geng, "Music Genre Trend Prediction Based on Spatial-Temporal Music Influence and Euclidean Similarity," *2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2021, pp. 406-411, doi: 10.1109/YAC53711.2021.9486510.
- [11] López Takeyas, B. (2007). *Introducción a la inteligencia artificial*. Instituto Tecnológico de Nuevo Laredo.



<http://itnuevolaredo.edu.mx/takeyas/Articulos/Inteligencia%20Artificial/ARTICULO%20Introduccion%20a%20la%20Inteligencia%20Artificial.pdf>

- [12] Contributors to Wikimedia projects. (2009, 29 de julio). Data reduction - Wikipedia. Wikipedia, the free encyclopedia. [https://en.wikipedia.org/wiki/Data\\_reduction](https://en.wikipedia.org/wiki/Data_reduction)
- [13] Contributors to Wikimedia projects. (2005, 31 de diciembre). Data cleansing - Wikipedia. Wikipedia, the free encyclopedia. [https://en.wikipedia.org/wiki/Data\\_cleansing](https://en.wikipedia.org/wiki/Data_cleansing)
- [14] Alonso, J. C., & Muñoz, A. (2014). Interpretacion de variables Dummy en modelos log-lin. Cali, Colombia: Departamento de Economía, Universidad Icesi.
- [15] Variable ficticia - Definición, qué es y concepto | Economipedia. Economipedia. <https://economipedia.com/definiciones/variable-ficticia.html>
- [16] ¿Qué es la regresión lineal? MathWorks - Creadores de MATLAB y Simulink - MATLAB y Simulink - MATLAB & Simulink. <https://la.mathworks.com/discovery/linear-regression.html> (accedido el 13 de agosto de 2022).
- [17] M. R. Gupta, E. K. Garcia and E. Chin, "Adaptive Local Linear Regression With Application to Printer Color Management," in IEEE Transactions on Image Processing, vol. 17, no. 6, pp. 936-945, June 2008, doi: 10.1109/TIP.2008.922429.
- [18] Colaboradores de los proyectos Wikimedia. (2002, 13 de febrero). Python - Wikipedia, la enciclopedia libre. Wikipedia, la enciclopedia libre. <https://es.wikipedia.org/wiki/Python>
- [19] ¿Qué es Google Colaboratory? (s. f.). 330ohms. <https://blog.330ohms.com/2021/08/10/que-es-google-colaboratory/>
- [20] Colaboradores de los proyectos Wikimedia. (2012, 4 de febrero). NumPy - Wikipedia, la enciclopedia libre. Wikipedia, la enciclopedia libre. <https://es.wikipedia.org/wiki/NumPy>



- [21] Introducción a la Librería Pandas de Python. (s. f.). Aprende IA. <https://aprendeia.com/introduccion-a-la-libreria-pandas-de-python-parte-1/>
- [22] Matplotlib: Funciones principales. Cursos de Programación de 0 a Experto © Garantizados. <https://unipython.com/matplotlib-funciones-principales/>
- [23] Espacio de recursos de ciencia de datos. (s. f.). Espai de recursos de ciencia de dades. <http://datascience.recursos.uoc.edu/es/preprocesamiento-de-datos-con-sklearn/>
- [24] Seaborn presentación. (s. f.). Interactive Chaos. <https://interactivechaos.com/es/manual/tutorial-de-seaborn/presentacion>
- [25] Gao, J. (2012). Data preprocessing.





# Predicción de precios de autos usados

## Used Car Price Prediction

136

### Yoset Cozco Mauri

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@ ycozco@unsa.edu.pe

<https://orcid.org/0000-0003-2027-962X>

### Fiorella Pilar Cayo Cayo

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@ fcayoc@unsa.edu.pe

<https://orcid.org/0000-0001-6068-6225>

### Jordy Pedro Valencia Jara

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@ jquispejar@unsa.edu.pe


<https://orcid.org/0000-0001-6336-5419>


### Henry Uriel Bustinza Torres


Universidad Nacional de San Agustín.  
Arequipa, Perú.

@ hbustinzat@unsa.edu.pe

<https://orcid.org/0000-0001-7092-1516>

 **ARK:** <ark:/42411/s12/a111>

 **DOI:** [10.48168/innosoft.s12.a111](https://doi.org/10.48168/innosoft.s12.a111)

 **PURL:** [42411/s12/a111](https://purl.org/ark:/42411/s12/a111)

RECIBIDO 19/05/2023 • ACEPTADO 21/08/2023 • PUBLICADO 30/09/2023



## RESUMEN

En varias ocasiones cuando se vende un automóvil de segunda mano no se sabe el precio exacto por varios factores, con lo cual puede ser vendido a menos o más valor del que en realidad vale, generando pérdidas tanto a los compradores como a los vendedores que intentan calcular el valor por intuición o con su conocimiento en autos, pero está a veces falla.

Entonces con ayuda de organizar los datos en una base de datos, y compararlos con ayuda de un software, así pudimos conseguir un precio mucho más preciso, esto tomando en cuenta diferentes factores que podrían modificar el precio del auto, como el tiempo de uso calidad de las distintas piezas, entre otros, logrando un precio mucho más preciso que al solo usar intuición.

**Palabras claves:** Automóvil, inteligencia artificial, precio de venta.

## ABSTRACT

*On several occasions when a second-hand car is sold, the exact price is not known due to various factors, with which it can be sold at less or more value than it is actually worth, you will lose both buyers and sellers who try to calculate the value by intuition or with his knowledge in cars, but this sometimes fails.*



*Then, with the help of organizing the data in a database, and comparing them with the help of software, we were able to get a much more accurate price, taking into account different factors that could modify the price of the car, such as time of use and quality. of the different parts, among others, achieving a much more accurate price than just using intuition.*

**Keywords:** Car, artificial intelligence, selling Price.

---

## INTRODUCCIÓN

Considerando la demanda actual de autos personales o de uso privado, la demanda en el sector automovilístico ha aumentado, y en el mercado de segunda mano la realidad es que ha ido creciendo, creando así una oportunidad de negocio tanto para compradores como vendedores. Comprar un auto usado en muchos países es una buena opción por el precio asequible. Dado que después de algunos años de uso, se puede obtener una ganancia de vender de nuevo.

Es por este motivo que es necesario saber con certeza qué factores influyen en el precio final. En los mercados de segunda mano, los precios de los autos usados no son constantes, es por este problema que la evaluación del auto para posteriormente calcular el precio se hace una práctica constante.

Además por la reciente emergencia sanitaria que sufrió todo el mundo, tener un automóvil para uso personal a paso de ser un privilegio a una necesidad, que al menos por familia se cuente con un automóvil para el transporte de esta misma, tanto por salud como por seguridad, siendo un mercado muy activo últimamente, más que en años anteriores, y en la zona en que vivimos Sudamérica es muy normal comprar autos de segunda mano siendo una aplicación a nuestro entorno o un problema que prácticamente podemos ver a diario.

Debido al incremento del valor de los vehículos, se aprecian cada vez más a las ferias de autos que muestran gran variedad en modelos de autos y precios siendo muchas veces más convenientes que una casa automotriz. Cada zona geográfica (países) tiene sus características que ofrecerán más o menos oportunidades a la industria automotriz dependiendo principalmente de su condición socio-económica.

El precio depende de factores como modelo del automóvil, año de fabricación, marca, opiniones de consumidores, ciudad, versión, seguridad, color, si es concesionario/particular, kilometraje, tipo de combustible (GLP, GNV, Gasolina, Diesel), llantas de aleación, sistema de frenos, aceleración, dimensiones, seguridad, aire acondicionado, su estado físico, el número de dueños anteriores, interior y dirección asistida, tales variables se utilizan para predecir el precio de dichos vehículos para el comercio.



También considerando que algunos vehículos pueden ser de uso particular o laboral pueden tener mayor o menor desgaste según la función que pudieron desempeñar antes del momento de estar a la venta. Tener en cuenta que tanto fueron usados sus diferentes componentes considerando el precio actual de un modelo completamente nuevo del mismo ayudará bastante a la hora de calcular un precio aproximado aparte de brindar un precio justo a nuestro usuario.

Ya que la mayoría de compradores en lo primero que se fijan es el precio, sin saber el valor de cada pieza y calcular el valor por uso, desgaste y/o pieza no original, como ya mencionamos anteriormente hay muchas características que alteran el precio final, pero en la mayoría de casos el comprador no cuenta con el tiempo o la información suficiente para calcular, y poder ver si el precio es razonable, de igual forma un vendedor, mayormente vende a un valor echado a suerte e intuición, y en varias ocasiones este puede ser más bajo de lo que realmente vale, o demasiado alto, y termina perdiendo dinero o tiempo, por no saber calcular el precio adecuado.

Ahora se debe considerar que los diferentes objetos ya de por si tienen un tiempo de vida útil y esto podría afectar al momento de calcular el valor que pueden tener ya que si consideramos un primer dueño que lo uso todo el tiempo donde se tenía la seguridad que dicho vehículo funciona de forma correcta, al momento de vender el precio bajo de forma considerable.

Teniendo en cuenta todo lo anterior el objetivo principal de este trabajo es encontrar el mejor modelo predictivo para predecir el precio de un auto de segunda mano, teniendo en cuenta todos los factores que hacen que varíe el precio entre un vehículo u otro.

## **Materiales y métodos o Metodología computacional**

La metodología utilizada para este proyecto se inicia con la recopilación de datos en bruto en un conjunto de datos lógico o conjunto de datos disponible en línea. El siguiente paso es crear un entorno para Pre-procesamiento de datos mediante limpieza y reducción de datos. Después obteniendo información útil del conjunto de datos, entrenamos y probamos para los análisis de datos.

Por último, en función de este resultado, se selecciona el modelo de aprendizaje automático para la predicción de precio y clasificación de precio basada en Inteligencia Artificial.



**Materiales:** revisar Dataset

**Metodología computacional:**

- Recolección de datos
- Creación de un entorno
- Preprocesamiento de Datos
- Entrenamiento
- Pruebas y análisis
- Selección del modelo
- Predicción del valor

## DataSet

Esta investigación se basó en el estudio de un dataset de vehículos usados y su precio. Este dataset de vehículos se encuentra en el sitio web Zenodo[2].

Las características disponibles en nuestro dataset son:

- Title: String. Vehicle name, a combination of brand and model. In some cases it includes model variation and other information.
- Price: Integer. Selling price.
- ref\_num: Integer. ID.
- location: String. Location of the vehicle.
- seller: String. Type of seller: professional or private.
- color: String. Vehicle color.
- door\_num: Integer. Number of doors.
- fuel\_type: String. Type of fuel.
- hp: Integer. Horse power.
- mileage: Integer. Kilometers.
- transmisión: String Either manual or automatic transmission.
- year: Integer. First registration.

## Pre - Procesamiento

En este paso revisamos que en nuestro dataset existen valores atípicos, puede ser debido a la gran cantidad de datos, estos valores han sido filtrados, dado que existían datos como cilindraje en la columna de kilometraje o tipo, valores que no corresponden a esta columna.



	A	B	C	D	E	F	G	H	I	J	K
1	color	door_num	fuel_type	hp	location	mileage	price	ref_num	seller	title	transmission
2	Blanco		5 gasolina	60	Málaga	170000	1450	426875509	Profesional	Renault - Clio	manual
3	Negro		5 híbrido	108	Madrid	14950	23390	429436719	Profesional	RENAULT - Zoe	automático
4	Blanco		5 diesel	165	Cuenca	260000	8990	426571125	Profesional	OPEL - Zafira Tc	manual
5	Marrón		5 diesel	313	Alicante	235000	21900	418805767	Particular	Bmw - Serie 7	automático
6	Rojo		5 gasolina	68	Jaén	74000	1800	428978271	Particular	Peugeot - 107	manual
7	Beige		5 gasolina	60	Valencia	148200	5200	429155464	Profesional	SKODA - Fabia	manual
8	Blanco		5 gasolina	90	Valencia	10958	13900	429155465	Profesional	RENAULT - Clio	manual
9	Azul		5 gasolina	75	Cádiz	41338	11990	426955686	Profesional	SKODA - Fabia	manual
10	Blanco		5 gasolina	82	Almería	115000	8500	423852498	Profesional	PEUGEOT - 200	manual
11	Azul		5 diesel	120	Salamanca	127000	9500	384357096	Profesional	OPEL - Insignia	manual
12	Blanco		5 diesel	130	Badajoz	151348	13499	416515268	Profesional	NISSAN - QASH	manual
13	Azul (Misano Blau)		5 diesel	116	Madrid	21500	29900	429446085	Profesional	BMW - Serie 1 1	manual
14	Gris / Plata (*RIV*)		5 gasolina	120	Murcia	102666	13200	429454609	Profesional	RENAULT - Cap	automático
15	Negro		5 diesel	120	Madrid	236000	1499	426420195	Profesional	Renault - Scénic	manual
16	Negro		5 gasolina	265	Barcelona	202000	12500	429382406	Profesional	Audi - S3	manual
17	Gris		5 diesel	231	Madrid	194000	8000	428695384	Particular	BMW Serie 3	automático
18	Blanco (T9T9 Blanco		5 diesel	150	Madrid	100082	21100	413132116	Profesional	AUDI - A4 Avant	automático
19	BlueMetallic		5 diesel	260	Madrid	180000	16500	306916699	Profesional	BMW - 330/325	manual
20	Gris / Plata		5 híbrido	416	Madrid	157000	42890	399596154	Profesional	PORSCHE - Par	automático
21	Gris / Plata		5 diesel	204	La Coruña	210000	17990	427831866	Profesional	AUDI - A6 Avant	manual
22	Gris / Plata		5 diesel	116	Madrid	96800	15900	429445783	Profesional	BMW - Serie 1 1	manual
23	Rojo		5 gasolina	82	Tenerife	65000	7990	427055980	Profesional	CITROEN - C1 F	manual
24	Gris / Plata		3 gasolina	160	Málaga	12000	18200	429200774	Profesional	ABARTH - 500 1	manual
25	Azul		5 diesel	258	Badajoz	68000	55900	372478528	Profesional	MERCEDES-BE	automático
26	5 puertas	diesel	140cv	340	Pontevedra	manual	6500	425410432	Particular	Volkswagen - Gc	año 2011
27	Gris / Plata (A90 So		4 diesel	265	Barcelona	111659	39500	426453621	Profesional	BMW - Serie 7 7	automático
28	Gris / Plata		3 diesel	125	Madrid	192000	1190	429386417	Profesional	SAAB - 93 2.2 T	manual
29	Negro		5 diesel	125	León	200000	2900	429164059	Profesional	SAAB - 93 SE 2.	manual
30	Gris / Plata		2 gasolina	61	Albacete	129000	4500	427018803	Profesional	SMART - fortwo	automático
31	Negro (*MUR*)		5 diesel	165	Albacete	213715	4900	429453596	Profesional	SSANGYONG -	manual
32	Gris		5 diesel	105	Zamora	348000	1800	298389356	Particular	Alfa romeo - 156	manual
33	Blanco (369 Blanco G		5 diesel	95	Madrid	124713	9100	427556441	Profesional	RENAULT - Meg	manual
34	Blanco		3 gasolina	265	Ourense	105000	20400	389199950	Profesional	VOLKSWAGEN	automático

Figura 1. Formato CSV usado para procesar el dataset

En la Figura 1 observamos valor atípico en la fila 26 columna C, cilindraje en fuel\_type. Usamos un filtro para evitar este tipo de valores atípicos, este dataset fue trasladado a otro documento libre de valores atípicos, la única condición fue eliminar estos valores atípicos.

Los valores no numéricos durante el procesamiento de datos en código serán suprimidos o modificados dado que no nos resultan útiles.

```
Splitting the data and target

[ ] X = car_dataset.drop(['title', 'price', 'location', 'ref_num'], axis=1)
    Y = car_dataset['price']
```

Figura 2. Descartando datos como título, locación, precio y ref num para variable X.





## Entrenamiento

Usamos el método linear Regresion para el entrenamiento con nuestro DataSet, por el tipo de datos y los resultados que queremos obtener resulta un método eficiente. Este resultado será contrastado con otro método Lasso para tener un análisis más exhaustivo de los resultados. Para este entrenamiento hemos usado la plataforma de google colab y el gestor de archivos en la nube google drive.

Como primer paso se hace una carga de datos, con una inspección inicial se revisaron una parte de los datos, tamaño en filas columnas y la información que ofrece este dataset.

```
Data Collection and process

[212] #Cargar datos del archivo csv a pandas dataframe
car_dataset = pd.read_csv('/content/drive/MyDrive/ia_dataset/data_cars.csv')

# Inspeccion de los 5 primeras filas
car_dataset.head()

  door_num  fuel_type  hp  location  mileage  price  ref_num  seller  title  transmission  year
0         5   gasolina   60   Málaga  170000   1450  426875509  Profesional  Renault - Clio  manual  2000
1         5   hibrido  108   Madrid   14950  23390  429436719  Profesional  RENAULT - Zoe Intens 80 kW R110 Bateria 50kWh  automático  2020
2         5   diesel  165   Cuenca  260000   8990  426571125  Profesional  OPEL - Zafira Tourer 2.0 CDTI 165 CV Selective  manual  2012
3         5   diesel  313   Alicante  235000  21900  418805767  Particular  Bmw - Serie 7  automático  2013
4         5   gasolina   68   Jaén   74000   1800  428978271  Particular  Peugeot - 107  manual  2007

[214] #Revisamos el numero de filas , columnas
car_dataset.shape

(6489, 11)

car_dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6489 entries, 0 to 6488
Data columns (total 11 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   door_num    6489 non-null   int64
1   fuel_type   6489 non-null   object
2   hp          6489 non-null   int64
3   location    6489 non-null   object
4   mileage     6489 non-null   int64
5   price       6489 non-null   int64
6   ref_num     6489 non-null   int64
7   seller      6489 non-null   object
8   title       6486 non-null   object
9   transmission 6489 non-null   object
10  year        6489 non-null   int64
dtypes: int64(6), object(5)
memory usage: 557.8+ KB
```

Figura 3. Dataset Collect e inspección inicial



```
# Revisando la cantidad de datos faltantes
car_dataset.isnull().sum()

door_num      0
fuel_type     0
hp            0
location      0
mileage       0
price         0
ref_num       0
seller        0
title         3
transmission  0
year         0
dtype: int64

title missing = 4
```

**Figura 4.** Revisión de datos faltantes.

```
[217] #Revisando la distribución por categoría de nuestros datos
print(car_dataset.fuel_type.value_counts())
print(car_dataset.seller.value_counts())
print(car_dataset.transmission.value_counts())

diesel      3701
gasolina    2483
híbrido     225
Otro        65
eléctrico   9
GLP         6
Name: fuel_type, dtype: int64
Profesional 5266
Particular  1223
Name: seller, dtype: int64
manual      4386
automático  2103
Name: transmission, dtype: int64
```

**Figura 5.** Revisión de la distribución por categoría de los datos



Para poder usar ambos modelos propuestos (Regression, Lasso) es necesario presentar datos de forma numérica, se deben escoger los datos a usar, estos mismos deben ser codificados y el resto suprimidos(drop).

```
Encoding categorical data

# codificando "fuel_type" column
car_dataset.replace({'fuel_type':{'':0, 'diesel':1, 'eléctrico':2,'gasolina':3,'GLP':4,'híbrido':5,'Otro':6}},inplace=True)
# codificando "seller" column
car_dataset.replace({'seller':{'Profesional':0,'Particular':1}},inplace=True)
# codificando "transmission" column
car_dataset.replace({'transmission':{'manual':0,'automático':1}},inplace=True)
# codificando "location" column
car_dataset.replace({'location':{'Álava':0,'Albacete':1,'Alicante':2,'Almería':3,'Asturias':4,'Ávila':5,'Badajoz':6,'Balears':7,'Barcelona':8,
'Burgos':9,'Cáceres':10,'Cádiz':11,'Cantabria':12,'Castellón':13,'Ciudad Real':14,'Córdoba':15,'Cuenca':16,
'Girona':17,'Granada':18,'Guadalajara':19,'Guipúzcoa':20,'Huelva':21,'Huesca':22,'Jaén':23,'La Coruña':24,
'La Rioja':25,'Las Palmas':26,'León':27,'Lleida':28,'Lugo':29,'Madrid':30,'Málaga':31,'Melilla':32,'Murcia':33,
'Navarra':34,'Ourense':35,'Palencia':36,'Pontevedra':37,'Salamanca':38,'Segovia':39,'Sevilla':40,'Soria':41,
'Tarragona':42,'Tenerife':42,'Teruel':43,'Toledo':44,'Valencia':45,'Valladolid':46,'Vizcaya':47,'Zamora':48,'Zaragoza':49,'':50}},inplace=True)
```

**Figura 6.** Codificación de datos a usar: fuel\_type, seller, transmission, location

```
car_dataset.head()
```

	door_num	fuel_type	hp	location	mileage	price	ref_num	seller	title	transmission	year
0	5	3	60	31	170000	1450	426875509	0	Renault - Clio	0	2000
1	5	5	108	30	14950	23390	429436719	0	RENAULT - Zoe Intens 80 kW R110 Bateria 50kWh	1	2020
2	5	1	165	16	260000	8990	426571125	0	OPEL - Zafira Tourer 2.0 CDTi 165 CV Selective	0	2012
3	5	1	313	2	235000	21900	418805767	1	Bmw - Serie 7	1	2013
4	5	3	68	23	74000	1800	428978271	1	Peugeot - 107	0	2007

**Figura 7.** Verificación inicial de los datos codificados



```
[220] X = car_dataset.drop(['title', 'price', 'location', 'ref_num'], axis=1)
      Y = car_dataset['price']
```

**print(X)**

	door_num	fuel_type	hp	mileage	seller	transmission	year
0	5	3	60	170000	0	0	2000
1	5	5	108	14950	0	1	2020
2	5	1	165	260000	0	0	2012
3	5	1	313	235000	1	1	2013
4	5	3	68	74000	1	0	2007
...	...	...	...	...	...	...	...
6484	5	1	165	141608	0	1	2016
6485	3	1	90	110000	0	0	2011
6486	5	1	115	72997	0	0	2018
6487	5	1	190	136000	0	1	2007
6488	5	1	150	51338	0	0	2017

[6489 rows x 7 columns]

**Figura 8.** Asignación de valores X Y, verificación de contenido de X

**print(Y)**

0	1450
1	23390
2	8990
3	21900
4	1800
...	...
6484	9350
6485	5999
6486	19390
6487	13250
6488	24500

Name: price, Length: 6489, dtype: int64

**Figura 9.** Verificación del contenido de la variable Y



Se subdividen los datos para entrenamiento y pruebas con la función de sklearn.model train\_test\_split con los parámetros X Y, tamaño de prueba 0.1 ( 10% de los datos será para prueba y 90% serán de entrenamiento) , estado aleatorio (división de datos).

```
splitting training an test data  
[ ] X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.1, random_state = 2)
```

**Figura 10.** Separación de datos de prueba y entrenamiento.

## RESULTADOS

### Selección del modelo

#### Modelo de entrenamiento

#### Regresión Lineal

Modelo de entrenamiento regresión lineal fue escogido por el tipo de datos que manejamos y el objetivo principal que tenemos.

```
[ ] # Loading the linnear regression model  
lin_reg_model = LinearRegression()  
  
[225] lin_reg_model.fit(X_train,Y_train)  
  
LinearRegression()
```

**Figura 11.** Creación de variable lin\_reg\_model y entrenamiento con los datos de entrenamiento





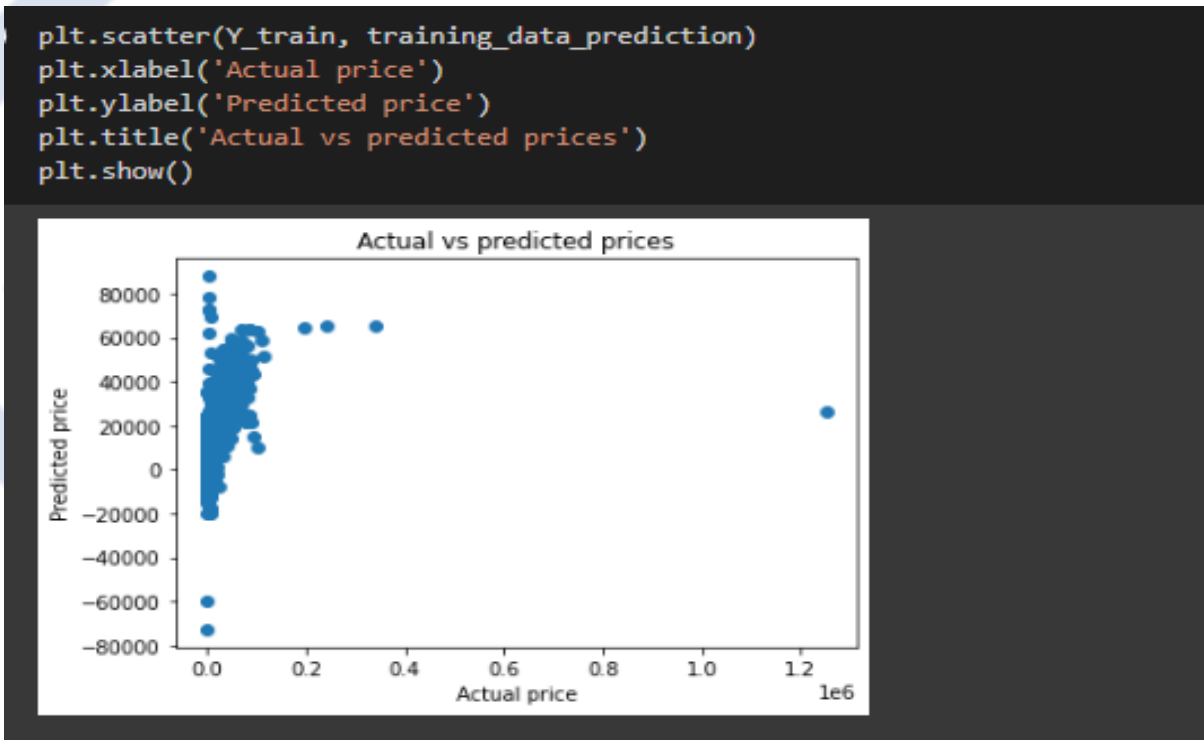
```
Model Evaluation

# predict on training data
training_data_prediction = lin_reg_model.predict(X_train)

[227] # R square error
error_score = metrics.r2_score(Y_train,training_data_prediction)
print("R score error:",error_score)

R score error: 0.2372542581549829
```

**Figura 12.** Creación de la variable training\_data\_prediction y predicción, cálculo de error



**Figura 13.** Visualización de resultados datos de entrenamiento.

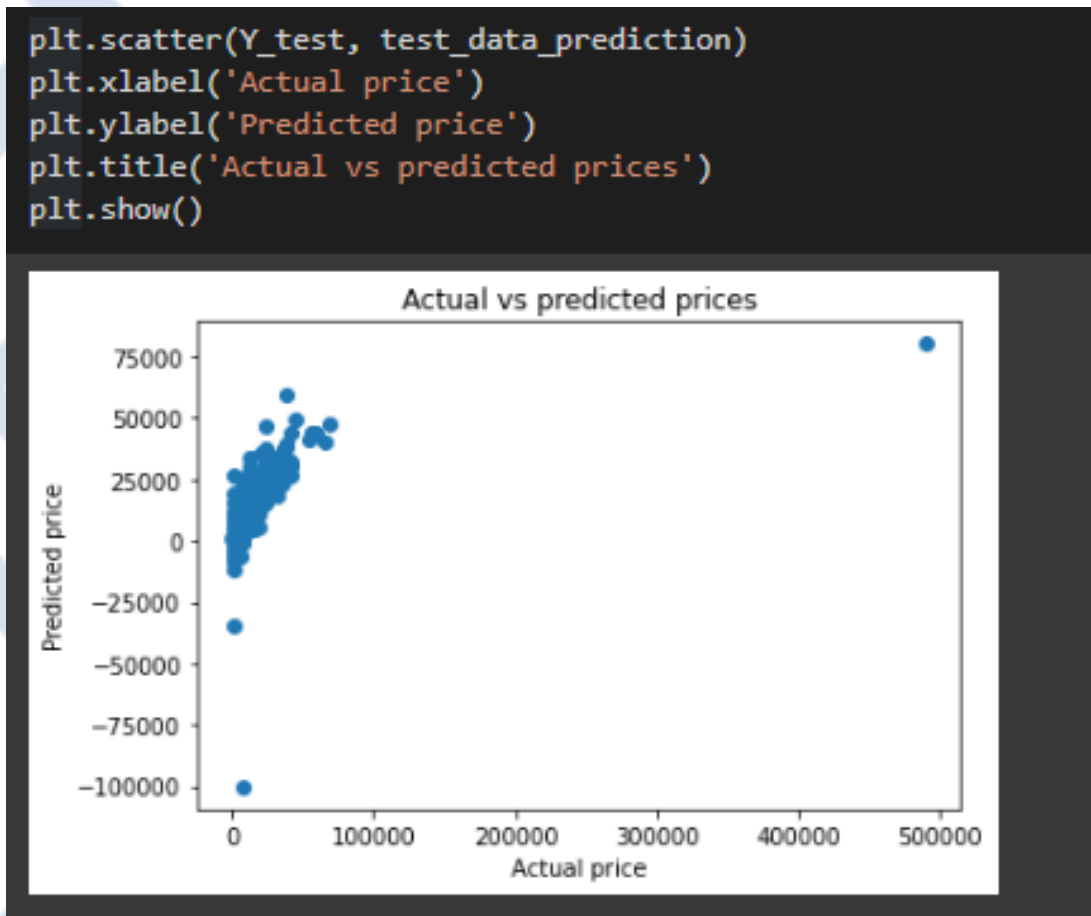


```
[235] # prediction on test data
      test_data_prediction = lin_reg_model.predict(X_test)

# R square error
error_score = metrics.r2_score(Y_test, test_data_prediction)
print("R score error:", error_score)

R score error: 0.31320393613194475
```

**Figura 14.** Predicción en los datos de prueba y calculo de error



**Figura 15.** Visualización de resultados en datos de prueba.



## Lasso

Como segunda opción hemos escogido el modelo Lasso que tiene similitud de comportamiento pero nos dio una perspectiva diferente en el entrenamiento.

```
[237] # Loading the Lasso model
lass_reg_model = Lasso()

[238] lass_reg_model.fit(X_train,Y_train)

Lasso()

Model Evaluation

[239] # predict on training data
training_data_prediction = lass_reg_model.predict(X_train)

[240] # R square error
error_score = metrics.r2_score(Y_train,training_data_prediction)
print("R score error:",error_score)

R score error: 0.23725421982942563
```

**Figura 16.** Carga del modelo, predicción en datos de entrenamiento, cálculo de error.

```
# prediction on test data
test_data_prediction = lass_reg_model.predict(X_test)

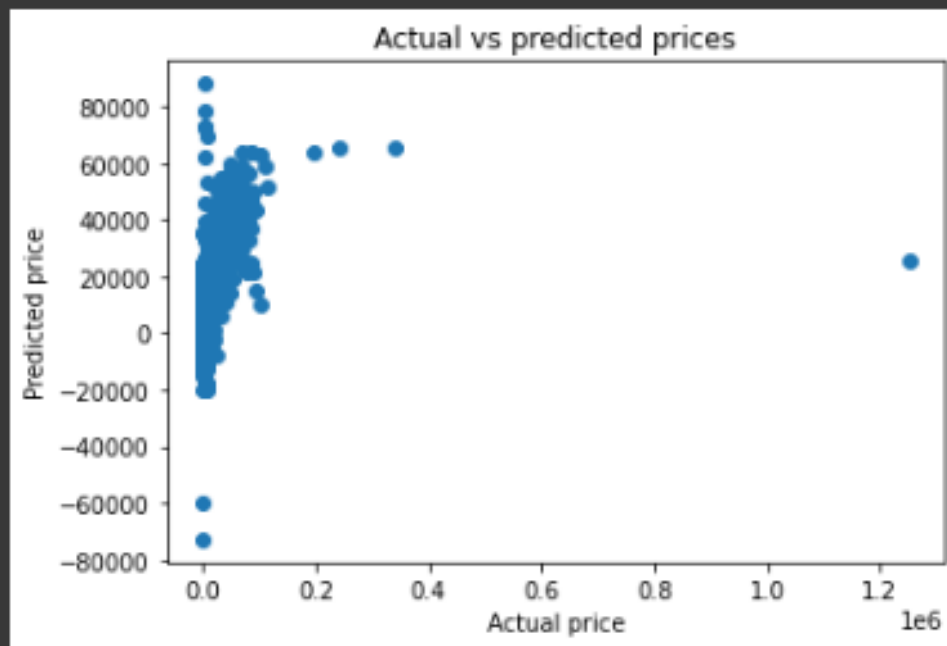
# R square error
error_score = metrics.r2_score(Y_test,test_data_prediction)
print("R score error:",error_score)

R score error: 0.3132182033298697
```

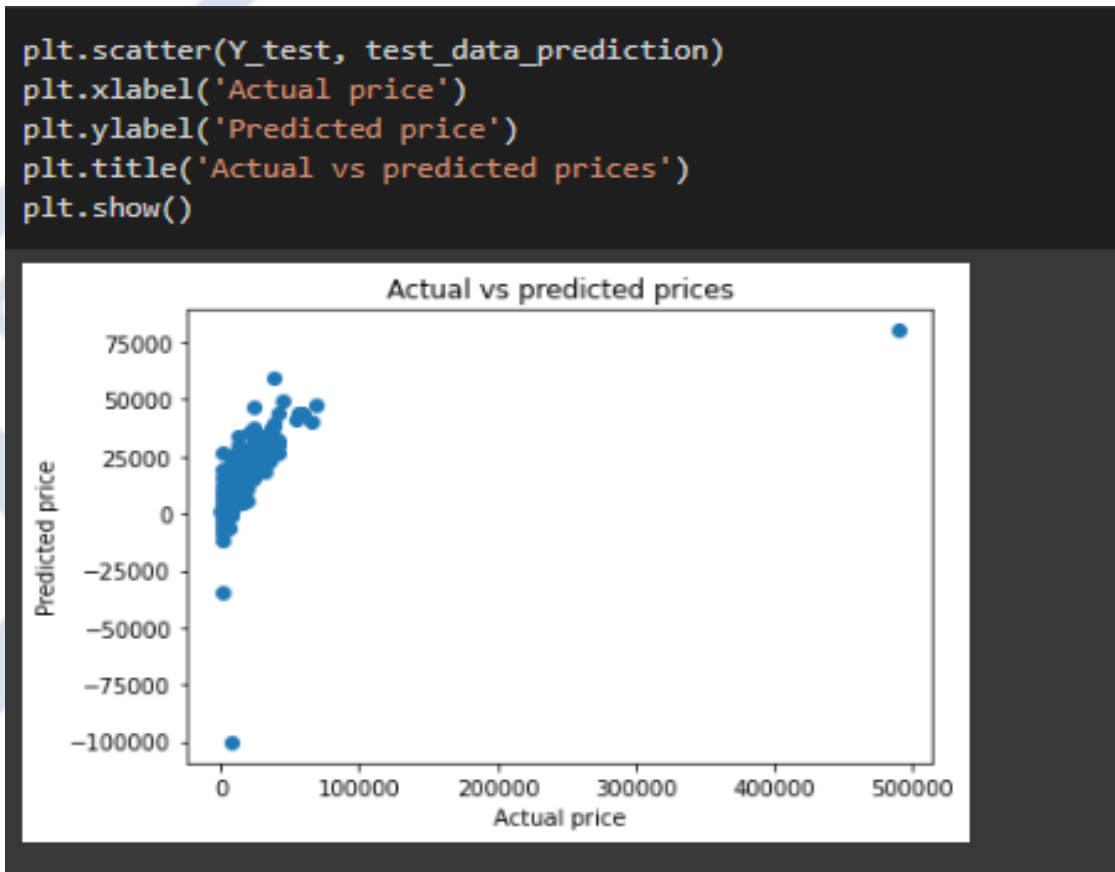
**Figura 17.** Entrenamiento en datos de prueba, cálculo de error.



```
plt.scatter(Y_train, training_data_prediction)  
plt.xlabel('Actual price')  
plt.ylabel('Predicted price')  
plt.title('Actual vs predicted prices')  
plt.show()
```



**Figura 18.** Visualización de resultados en datos de entrenamiento modelo Lasso



**Figura 19.** Visualización de resultados en datos de prueba modelo Lasso.

## Conclusión

Como pudimos apreciar la recopilación de datos puede llegar a ser fundamental para este trabajo ya que con ella logramos aproximar un precio justo para el mercado actual, considerando también diferentes factores ya sea el paso del tiempo las diferentes situaciones a las que podría haber sido expuesto el auto, valor de los diferentes materiales y desgaste de los mismos, precio normal al que estaría disponible en el tiempo actual. Lo cual sirve de mucho al comprador y al vendedor para llegar a un común acuerdo que beneficiaría a ambos.

La cantidad de variables que rodean este segmento que son los autos usados y su correcta predicción de precios, hacen que la incorrecta selección de parámetros afecte drásticamente el resultado. En este documento manualmente hemos normalizado y limpiado algunos datos que puedan generar ruido. Este preprocesamiento es necesario para aumentar el desempeño de nuestro modelo a usar.





Durante el análisis logramos corroborar la cohesión de los resultados entre la predicción y los datos proporcionados, motivo por el cual escogemos el modelo Lasso que mejor resultado nos ha proporcionado usando esta condición.

## Contribución de Autoría

**Yoset Cozco Mauri:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Fiorella Pilar Cayo Cayo:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Jordy Pedro Valencia Jara:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Henry Uriel Bustinza Torres:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1] A. Yadav, E. Kumar, and P. K. Yadav, "Object detection and used car price predicting analysis system (UCPAS) using machine learning technique," *Linguist. Cult. Rev.*, vol. 5, no. S2, pp. 1131–1147, 2021, doi: 10.21744/lingcure.v5ns2.1660.
- [2] Duran, D., & Martinez, P. (2021, November 5). Carads. Zenodo. Retrieved August 17, 2022, from <https://zenodo.org/record/5651148>
- [3] Thai-Nichi Institute of Technology, Institute of Electrical and Electronics Engineers, and IEEE Thailand Section., "Proceedings of 2018 5th International Conference on Business and Industrial Research (ICBIR): smart technology for next generation of information, engineering, business and social science: 17-18 May, 2018, Thai-Nichi Institute of Technology, Bangkok, Th," 2018 5th Int. Conf. Bus. Ind. Res., pp. 115–119, 2018.
- [4] G. Reiter Director and R. Blázquez, "Universidad Torcuato Di Tella Escuela de Negocios MBA Tesis de Maestría Valuación de autos usados utilizando la técnica de Regresión Lineal Múltiple," 2007.



- [5] I. Permatasari, "No Title על העיוורון," J. Geotech. Geoenvironmental Eng. ASCE, vol. 120, no. 11, p. 259, 2015.
- [6] A. Gavazza, A. Lizzeri, and N. Roketskiy, "A quantitative analysis of the used-car market," Am. Econ. Rev., vol. 104, no. 11, pp. 3668–3700, 2014, doi: 10.1257/aer.104.11.3668.
- [7] C. Chen, L. Hao, and C. Xu, "Comparative analysis of used car price evaluation models," AIP Conf. Proc., vol. 1839, no. May 2017, 2017, doi: 10.1063/1.4982530.
- [8] J. M. Ceferino and J. Ignacio, "LIII Reunión Anual," pp. 0–16, 2018.
- [9] E. Bustamante et al., "negocio de certificación de autos usados," 2022.
- [10] J. Berkovec, "New Car Sales and Used Car Stocks: A Model of the Automobile Market," RAND J. Econ., vol. 16, no. 2, p. 195, 1985, doi: 10.2307/2555410.



# Aplicación de árboles de decisión para un pronóstico salarial

153

*Application of decision trees for wage forecasting*

## **Corina Hanco Vargas**

Universidad Nacional de San Agustín.  
Perú

@ chancov@unsa.edu.pe

 <https://orcid.org/0000-0001-5471-4226>

## **David Flores Silva**

Universidad Nacional de San Agustín.  
Perú.

@ dfloressi@unsa.edu.pe

 <https://orcid.org/0000-0002-0019-2458>

## **Jessica Hanco Velásquez**

Universidad Nacional de San Agustín.  
Perú

@ jhancove@unsa.edu.pe


 <https://orcid.org/0000-0002-0588-6520>


## **Alejandra Fernandez N.**


Universidad Nacional de San Agustín.  
Perú.

@ mfernandezn@unsa.edu.pe

 <https://orcid.org/0000-0002-7896-8228>

 **ARK:** <ark:/42411/s12/a112>

 **DOI:** [10.48168/innosoft.s12.a112](https://doi.org/10.48168/innosoft.s12.a112)

 **PURL:** [42411/s12/a112](https://purl.org/42411/s12/a112)

RECIBIDO 05/06/2023 • ACEPTADO 01/09/2023 • PUBLICADO 30/09/2023



## RESUMEN

Este artículo detalla el proceso que se realizó para el pronóstico salarial en una base de datos de un censo dado en 1996, donde se utilizó el lenguaje de programación Python, para el análisis de los datos del dataset se utilizó el servidor Google Colab para ejecutar los algoritmos en la nube, ya que el equipo considero que la velocidad de análisis de datos en Google Colab es más rápido. También se hizo uso de una de las técnicas de minería de datos para clasificar las variables usando árboles de decisión que tienen la capacidad de representar gráficamente varias soluciones alternativas con el fin de determinar los cursos/rutas de acción más efectivos para la clasificación de la obtención del sueldo de una persona.

**Palabras claves:** Árboles de decisión, Minería de datos, Pronóstico salarial.

## ABSTRACT

*This article details the process that was carried out for the salary forecast in a database of a census given in 1996, where the Python programming language was used, for the analysis of the data of the dataset the Google Colab server was used to execute the algorithms in the cloud,*



*since the team considered that the speed of data analysis in Google Colab is faster. One of the data mining techniques was also used to classify the variables using decision trees that have the ability to graphically represent several alternative solutions in order to determine the most effective courses/routes of action for the classification of the obtainment. of a person's salary.*

**Keywords:** *Decision trees, Data mining, Salary forecast.*

---

## INTRODUCCIÓN

La situación laboral ha cambiado en todo el mundo a causa de la pandemia por el Covid-19, según noticias dadas por las Naciones Unidas [1], se predecía que para finales del 2021 se habrían perdido 125 millones de empleos y la recuperación desigual atribuiría al crecimiento de la brecha entre los mercados laborales entre países industrializados y las naciones en desarrollo, diferencias entre los países pobres y ricos. Sin embargo, desde antes de la pandemia ya había conocimiento de los problemas en el mercado laboral, tanto en la búsqueda como también en el tipo de selección que se daba para los empleados. Como ejemplo de ello, un informe de la Organización Internacional del Trabajo (OIT), analizó que el principal problema de los mercados laborales en el mundo era el empleo de mala calidad donde millones de personas, a pesar de tener un trabajo fijo, se veían obligadas a aceptar condiciones de trabajo deficientes. Y se predecía que para 2018, la mayoría de los 3300 millones de personas empleadas en el mundo no gozarían de un nivel suficiente de seguridad económica, bienestar material o igualdad de oportunidades. Es más, el avance de la reducción del desempleo a nivel mundial no se veía reflejado en la mejora de la calidad del trabajo. Lo que en general se preveía por diversos déficits de trabajo decente, y se advertía de que, a ese ritmo, la consecución del objetivo de trabajo decente para todos, establecido entre los Objetivos de Desarrollo Sostenible (ODES), sería inalcanzable para muchos países [2].

En el artículo de Madero [3] titulado "Factores relevantes del desarrollo profesional y de compensaciones en la carrera laboral del trabajador" se pudo concluir que destacaban diferencias en el apoyo, desarrollo y compensación profesional, por en un estudio realizado con población entre profesionales de México y Estados Unidos. Las diferencias de género influyen en lo que se espera de cada uno en su labor en el trabajo, sin embargo, las diferencias de ubicación influyen en los factores de apoyo y compensación profesional esperados entre mexicanos y estadounidenses, destacando como factores influyentes las diferencias de género como también las nacionalidades.

Además, en otro estudio por Heras et. al [4] que, aparte de resaltar las diferencias de salarios por géneros, define como determinantes económicos del salario: el nivel de capital humano y un indicador de éxito. Siendo algunas de sus variables *edad, antigüedad y nivel de estudios*, como parte del capital humano y haciendo comparaciones salariales para el indicador de éxito, ya que



este es definido como “la consecución salarial asociada al hecho de recibir más salario que aquellos con los que se comparte la misma dotación de capital humano”.

Para una mayor comprensión de las condiciones en las que está el mercado laboral en el mundo, la OIT (Organización Internacional del Trabajo) [5] ofrece un informe referencial del asunto, en el cuál se indican 4 principales preocupaciones del mundo laboral actualmente: Primero, las proyecciones de crecimiento económico no se verán reflejados en la disminución de pobreza en países de menores ingresos; segundo, la edad vendría a ser un filtro más en la contratación de empleadores, por la fuerza que se dedica en un trabajo; tercero, aún existen deficiencias en la calidad del empleo, que consiste en: un trabajo seguro, saludable, con el acceso a la protección social, para expresar opiniones propias, y que defienda los derechos fundamentales, como la no discriminación; esto también influenciado por el informalismo; y cuarto, se da la segmentación entre trabajadores por su ubicación geográfica, sexo y edad.

Se habla entonces de ciertos factores que determinan el conseguir un trabajo decente y por lo tanto un buen salario. Con el avance de la tecnología se ha desarrollado un campo que ayudaría a analizar y trabajar con datos recolectados para determinar predicciones, la inteligencia artificial, que cuenta con modelos de aprendizaje automático, machine learning, que servirían para la clasificación de grandes cantidades de datos, según Arana [6] los árboles de decisión son el mejor ejemplo de ello no sólo por su interpretabilidad sino también por ser la base de los modelos más potentes utilizados en la actualidad.

En el trabajo realizado por Ponce J. et. al [7] se aplicó la utilización de árboles de decisión para un análisis de la competitividad empresarial, tomando como indicadores recursos financieros, la mercadotecnia, recursos humanos y tecnología. Este se realizó con el fin de determinar el cambio de crecimiento o disminución de esos indicadores con respecto al tiempo, finalmente se concluyó en la mejora de técnicas en la toma de decisiones más convenientes para una empresa.

Por lo tanto, se propone realizar un análisis de una base de datos obtenida de Becker [8], quién tomó como variables la edad, género, nivel de educación, país de natalidad y datos adicionales, para determinar si una persona podría obtener un sueldo mayor o menor de 50 000 dólares al año. Se utilizará una de las técnicas de minería de datos para la clasificación de estos, usando árboles de decisión a fin de determinar qué variables si son las determinantes para la estimación de la obtención del sueldo de una persona pudiendo utilizarse esto como ayuda más precisa de estudios de factores determinantes en el mercado laboral.





## Fundamentación teórica

### Árboles de decisión

Representación gráfica de varias soluciones alternativas que están disponibles para resolver un problema determinado, con el fin de determinar los cursos/rutas de acción más efectivos, [9]. La elección de esta misma se debió a la guía realizada por el docente en una implementación.

### KDD

Técnica de minería de datos para mejorar la toma de decisiones con grandes cantidades de datos. Su procedimiento suele constar de los siguientes pasos, según la escuela ESAN BUSINESS [10]:

- Comprensión del área de estudio y fijación de objetivos.
- Implementación de un dataset objetivo.
- Limpieza y procesamiento de información.
- Minería de datos.
- Interpretación y análisis de patrones encontrados.
- Utilización del conocimiento obtenido para la toma de decisiones.

## Materiales y métodos o Metodología computacional

Este trabajo utilizó el método predictivo de Árboles de decisión KDD de Machine Learning para el pronóstico salarial en una base de datos de un censo dado en 1996. Este dataset consta originalmente de 15 atributos (columnas) y 32560 instancias (filas).

Nº	Variable	Descripción
1	age	Continuous
2	workclass	Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked
3	fnlwgt	Continuous
4	education	Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool
5	education-num	Continuous



6	marital-status	Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse
7	occupation	Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op- inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv- house-serv, Protective-serv, Armed-Forces
8	relationship	Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.
9	race	White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Blac
10	sex	Female, Male
11	capital-gain	Continuous
12	capital-loss	Continuous
13	hours-per-week	continuous
14	native-country	United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinidad&Tobago, Peru, Hong, Holand-Netherlands.
15	salary	$\leq 50K$ or $> 50K$

## Resultados y discusión

### Análisis del problema

Mediante un árbol de decisiones se predice si el salario de una persona es  $\leq 50K$  o  $> 50K$ , teniendo en cuenta 11 variables predictivas.



## Pre Procesamiento de data

### Data collecting

Del dataset descargado de la página Kaggle, la tabla cuenta con 32560 filas conocidas como registros y 15 columnas, las primeras 14 variables predictoras y la última la variable a predecir. De todas las variables 6 son del tipo int64 y 9 de ellas de tipo object como se muestra en la Figura 1, esta misma la cual será tratada para el estudio en procesamiento de datos antes de entrenar nuestro modelo predictivo.

```
➤ sueldo = pd.read_csv('/content/drive/My Drive/salary.csv', engine='python')
sueldo.head()
sueldo.info()

↳ <class 'pandas.core.frame.DataFrame'>
RangeIndex: 32561 entries, 0 to 32560
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  -
0   age                   32561 non-null  int64
1   workclass             32561 non-null  object
2   fnlwgt                32561 non-null  int64
3   education             32561 non-null  object
4   education-num        32561 non-null  int64
5   marital-status       32561 non-null  object
6   occupation            32561 non-null  object
7   relationship         32561 non-null  object
8   race                  32561 non-null  object
9   sex                   32561 non-null  object
10  capital-gain          32561 non-null  int64
11  capital-loss          32561 non-null  int64
12  hours-per-week        32561 non-null  int64
13  native-country       32561 non-null  object
14  salary                32561 non-null  object
dtypes: int64(6), object(9)
memory usage: 3.7+ MB
```

Figura 1. Información descriptiva del dataset descargado de internet

### Data transformation

Debido a que se va a realizar el entrenamiento de datos para un árbol de decisión se vio necesario transformar las variables de tipo object a variables numéricas, para eso haciendo uso de la librería LabelEncoder, Figura 2.



```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
salary.workclass = le.fit_transform(salary.workclass)
salary.education = le.fit_transform(salary.education)
salary.maritalStatus = le.fit_transform(salary.maritalStatus)
salary.occupation = le.fit_transform(salary.occupation)
salary.relationship = le.fit_transform(salary.relationship)
salary.race = le.fit_transform(salary.race)
salary.sex = le.fit_transform(salary.sex)
salary.nativeCountry = le.fit_transform(salary.nativeCountry)
#salary.salary = le.fit_transform(salary.salary)
salary.head(5)
```

Figura 2. Código de transformación de variables objeto a int64

	age	workclass	education	educationNum	maritalStatus	occupation	relationship	race	sex	hoursPerWeek	nativeCountry	salary
0	39	7	9	13	4	1	1	4	1	40	39	<=50K
1	50	6	9	13	2	4	0	4	1	13	39	<=50K
2	38	4	11	9	0	6	1	4	1	40	39	<=50K
3	53	4	1	7	2	6	0	2	1	40	39	<=50K
4	28	4	9	13	2	10	5	2	0	40	5	<=50K

Figura 3. Resultado del proceso de data transformation

## Data cleaning

### Eliminación de columnas:

Para el tratado de la data set se realizó la eliminación de las columnas `fnlwgt`, `capital-gain`, y `capital-loss` razones justificadas por información innecesaria para el caso se estudió, gran cantidad de valores desconocidos, valores desconocidos.

```
[71] salary = sueldo.drop(['fnlwgt', 'capital-gain', 'capital-loss'],1)
salary.head()
```

Figura 4. Código de eliminación de columnas del dataset



## Data reduction

### Reconocimiento y eliminación de ruido:

Se realizó el reconocimiento de valores ruidosos en el dataset en las columnas de cómo se puede ver en las Figura 5 y Figura 6.

```
[27] salary.age.hist()  
plt.show() # que sea hasta 75
```

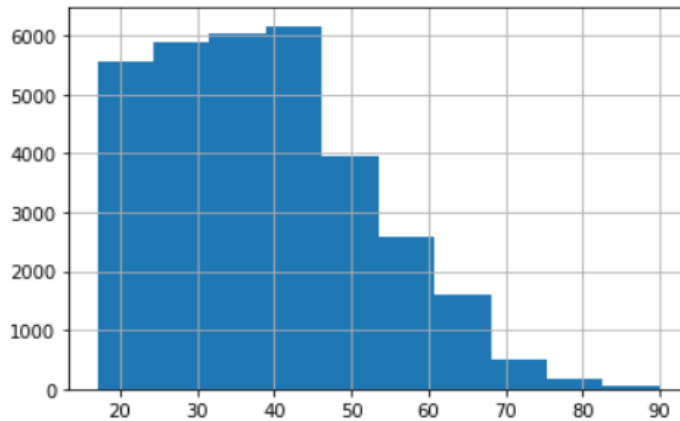


Figura 5. Reconocimiento de ruido mediante un histograma de la columna age

```
[36] salary.nativeCountry.hist()  
plt.show() # mayor o igual a 39
```

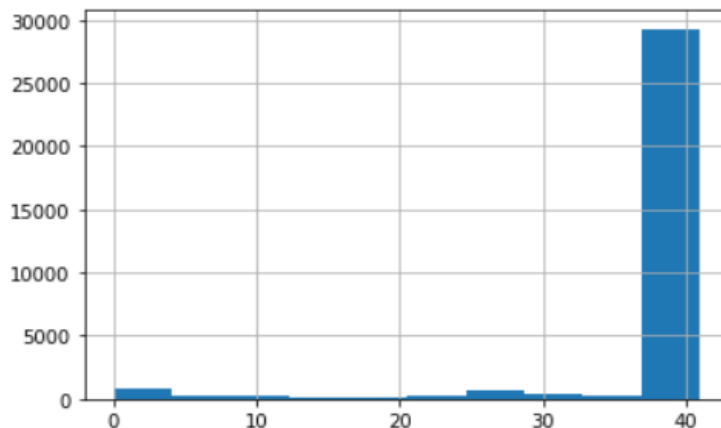


Figura 6. Reconocimiento de ruido mediante un histograma de la columna nativeCountry





## Eliminación de ruido

Los valores ruidos reconocidos en el dataset fueron eliminados como se muestra en la Figura 7 estableciendo valores limitados a las variables.

```
[74] filtered_salary = salary[(salary['age'] <= 75) & (salary['workclass'] >= 4) & (salary['education'] >= 6) & (salary['educationNum'] >= 9) & (salary['maritalStatus'] <= 4) & (salary['race'] >= 2) & (salary['hoursPerWeek'] <= 60) & (salary['nativeCountry'] >= 39)]
```

Figura 7. Código de la eliminación de datos que generan ruido

## Preparación de Datos

En esta sección se definieron las variables predictoras las cuales son age, workclass, education, educationNum, maritalStatus, occupation, relationship, race, sex, hoursPerWeek y la definicion de la variable a predecir que salary.

```
#Variables Predictoras  
x = filtered_salary .iloc[:,0:10] #que considere todas las filas, las variables predictoras  
  
#Variable a Predecir  
y = filtered_salary .iloc[:,11] #Se encuentra en la ultima columna
```

Figura 8. Definición de variables

Se realizó también la división de datos para testing & training, asignando un 80% de los datos para el entrenamiento del modelo y un 20% para pruebas.

```
#x_train y y_train para entrenamiento  
#x_test y y_test para prueba  
x_train, x_test, y_train, y_test = train_test_split(x,y,train_size=0.80)
```

Figura 9. Asignación de datos para testing y training

## Creación del modelo



Haciendo uso de la librería DecisionTreeClassifier se hace la construcción del árbol indicando que la evaluación será la entropía, además se aclara que se realizará un árbol con una profundidad de 4

```
[80] #llamamos al constructor del arbol de decision
#arbol = DecisionTreeClassifier() el arbol completo
arbol = DecisionTreeClassifier(criterion="entropy",max_depth=4)

#Entrenamos el modelo
arbol.fit(x_train, y_train)

#realizo una prediccion
y_pred = arbol.predict(x_test)
```

Figura 10. Código de creación del modelo de arbol de desicion

## Validación del entrenamiento

En la matriz de confusión se muestra que 2688 datos el algoritmo los clasificó como salario  $\leq 50K$  y si eran  $\leq 50K$ , 430 datos fueron clasificados por el algoritmo como  $> 50K$  y si eran  $> 50K$ , 672 datos el algoritmo los clasificó como  $\leq 50K$  y no eran  $\leq 50K$ , 107 datos fueron clasificados por el algoritmo como  $> 50K$  y no eran  $> 50K$ .

```
[81] matriz = confusion_matrix(y_test, y_pred)
print('Matriz de confucion')
print(matriz)
```

```
Matriz de confucion
[[2688  107]
 [ 672  435]]
```

Figura 11. Matriz de confusión

El modelo trabaja con una exactitud de 0.80 y tiene un error de 0.20

```
✓ [83] accuracy_score = accuracy_score(y_test, y_pred)
s print('Exactitud del modelo')
print(accuracy_score)
```

```
Exactitud del modelo
0.800358790363916
```

Figura 12. Código de medición de exactitud del modelo



## Resultado del árbol de decisión

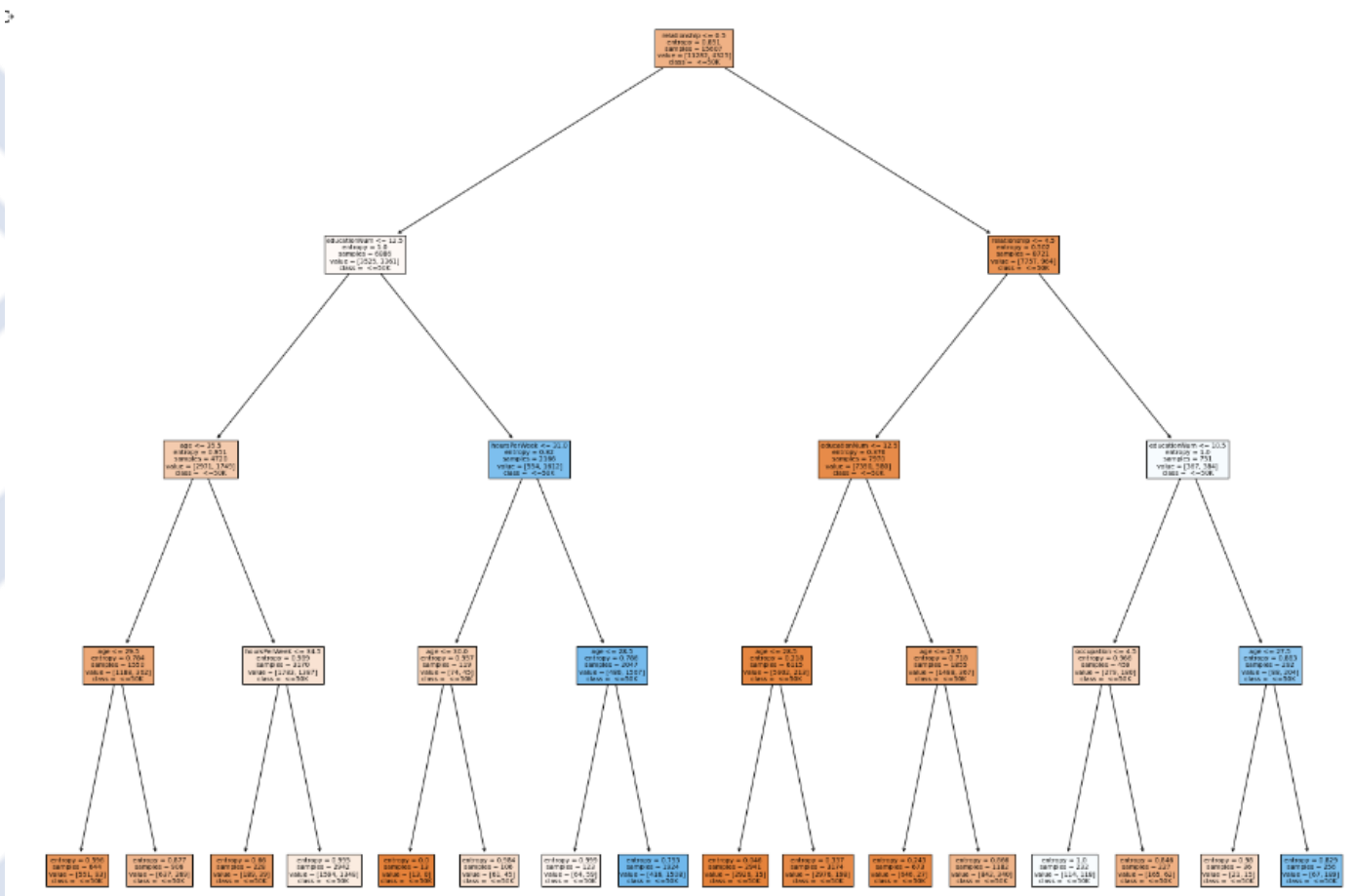


Figura 13. Resultado de árbol de decisiones con un profundidad de 4

## Conclusiones

Con la base de datos inicial se tenían en general 14 variables que definirían si una persona podría obtener un sueldo mayor a 50 000 dólares al año, que vendría a ser lo mínimo necesario para la satisfacción de las necesidades básicas en EEUU. Quedando como últimas variables predictoras, después de su procesamiento, en el árbol de decisión: "estado marital", "edad", "nivel de educación", "ocupación" y "horas trabajadas por semana".



Como parte de los trabajos futuros que se podrían implementar a raíz de este, pueden ser:

- Comparativas de distintos métodos predictivos aplicados a esta base de datos, aunque en este caso sólo se vió el algoritmo KDD.
- Para predicciones más exactas es necesaria una actualización de una base de datos.
- Consideración de las variables que sí determinan la obtención de un salario mayor a 50 000 dólares anuales por parte de quienes estén en búsqueda de empleo.

## Contribución de Autoría

**Corina Hanco Vargas:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **David Flores Silva:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Jessica Hanco Velásquez:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Alejandra Fernandez Ninahuaman:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1] "El panorama laboral después de la pandemia es peor de lo que se preveía | Noticias ONU." <https://news.un.org/es/story/2021/10/1499052>.
- [2] "El gran problema del empleo en el mundo: las malas condiciones de trabajo | Agora: Inteligencia Colectiva para la Sostenibilidad," Feb. 19, 2019. <https://www.agorarsc.org/el-gran-problema-del-empleo-en-el-mundo-las-malas-condiciones-de-trabajo/>.
- [3] S. Madero, "Factores relevantes del desarrollo profesional y de compensaciones en la carrera laboral del trabajador," p. 22, 2010.
- [4] R. L. Heras, A. Maroto Sánchez, Á. Martín-Román, and A. Moral De Blas, "Éxito salarial: indicadores por género en la distribución salarial." [Online]. Available: [www.iaes.es](http://www.iaes.es)



- [5] Organización Internacional del Trabajo, "RESUMEN EJECUTIVO," 2020.
- [6] C. Arana, "Modelos de Aprendizaje Automático Mediante Árboles," Buenos Aires, Feb. 2021. [Online]. Available: <http://hdl.handle.net/10419/238403>
- [7] J. Ponce, E. Vicente, R. Rodríguez, and S. Muñoz, "Análisis de la competitividad empresarial aplicando árboles de decisión," pp. 66–80, Aug. 2020. [Online]. Available: <https://orcid.org/0000-0002-2965-7263>
- [8] B. Becker and R. Kohavi, "UCI Machine Learning Repository: Census Income Data Set." <https://archive.ics.uci.edu/ml/datasets/Census+Income>.
- [9] "Árbol de decisión en Machine Learning (Parte 1) - sitiobigdata.com", sitiobigdata.com, 2019. [Online]. Available: <https://sitiobigdata.com/2019/12/14/arbol-de-decision-en-machine-learning-parte-1/#>.
- [10] "Minería de datos: ¿en qué consiste el knowledge discovery in databases?", ESAN Graduate School of Business, 2018. [Online]. Available: <https://www.esan.edu.pe/conexion-esan/mineria-de-datos-en-que-consiste-el-knowledge-discovery-in-databases#:~:text=El%20KDD%20es%20un%20proceso,recursos%20%C3%BAtiles%20para%20una%20compa%C3%B1a>.





# Aplicación de árboles de decisión para la identificación de adaptabilidad de estudiantes en educación online

166

*Application of decision trees for the identification of adaptability of students in online education*

## Luis Emanuel Araoz Valencia

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@laraozv@unsa.edu.pe

<https://orcid.org/0000-0003-4274-6380>

## Walter Huaracha Condori

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@whuaracha@unsa.edu.pe

<https://orcid.org/0000-0002-4155-235X>

## Víctor Raúl Quispe Quicaña

Universidad Nacional de San Agustín.  
Arequipa, Perú.

@vquispequic@unsa.edu.pe


<https://orcid.org/0000-0003-2294-0215>


## Alex Ronaldo Turpo Coila


Universidad Nacional de San Agustín.  
Arequipa, Perú.

@aturpoco@unsa.edu.pe

<https://orcid.org/0000-0002-6306-0137>

 **ARK:** [ark:/42411/s12/a113](https://nbn-resolving.org/ark:/42411/s12/a113)

 **DOI:** [10.48168/innosoft.s12.a113](https://doi.org/10.48168/innosoft.s12.a113)

 **PURL:** [42411/s12/a113](https://nbn-resolving.org/ark:/42411/s12/a113)

RECIBIDO 20/06/2023 • ACEPTADO 08/09/2023 • PUBLICADO 30/09/2023



## RESUMEN

Debido a la pandemia mundial por Covid-19, se instauró la educación online en el aprendizaje de los estudiantes. Sin embargo, la efectividad de esta modalidad, así como la adaptabilidad de los estudiantes es algo que puede depender de algunos factores. En ese sentido, el presente artículo de investigación presenta una descripción del uso de árboles de decisión para determinar la adaptabilidad de estudiantes en la educación online, usando para ello un dataset de 1205 registros con datos como el tipo de conexión e internet, dispositivo, condición financiera, entre otros datos importantes. Así mismo, se empleó herramientas como Google Colab, Python y librerías populares en trabajos similares de Inteligencia artificial y Machine Learning. El modelo del árbol de decisión elaborado tuvo una precisión y exactitud de 92%.

**Palabras claves:** Inteligencia artificial, aprendizaje automático, árboles de decisión, Python, clasificación, educación en línea.



## ABSTRACT

*Due to the global pandemic by Covid-19, online education was established in student learning. However, the effectiveness of this modality, as well as the adaptability of the students, is something that may depend on some factors. In this sense, this research article presents a description of the use of decision trees to determine the adaptability of students in online education, using a dataset of 1205 records with data such as the type of connection and internet, device, condition. financial, among other important data. Likewise, tools such as Google Colab, Python and popular libraries were used in similar works of Artificial Intelligence and Machine Learning.*

**Keywords:** *Artificial Intelligence, Machine Learning, decision trees, Python, classification, online education.*

---

## INTRODUCCIÓN

En la actualidad tras haber vivido una situación pandémica, todos hemos sido testigos de la virtualización en múltiples campos laborales, incluyendo el campo de la educación, por lo cual, se ha generado diferentes percepciones en los que se destacan calidad, aprendizaje, entre otras [1]. En ese sentido, existe la necesidad de conocer los niveles de adaptabilidad de los estudiantes al estudio virtual tanto en el nivel escolar como el universitario, sin embargo el definir cómo los estudiantes se adaptan a la virtualidad depende de muchos factores, y los que los se han llegado a considerar esta vez son: Tipo de Internet disponible, tipo de conexión, condición financiera, locación del estudiante, tipo de institución y nivel de educación [2].

La Inteligencia Artificial según Minsky [3] es el “estudio de cómo programar computadoras que posean la facultad de hacer aquello que la mente humana puede realizar”. En nuestros días la IA es un tema de gran importancia, ya que aborda muchos aspectos de las tendencias actuales. El uso de IAs es muy diverso y en la actualidad tiene muchas posibilidades de aplicación en múltiples áreas como la robótica, las ciencias sociales, como apoyo en ciencias empresariales, y también en el área de la educación [4].

Para realizar un análisis de los datos contenidos en el dataset [2] es necesaria la aplicación de minería de datos. En [5] hace mención a la minería de datos como una herramienta apropiada para el tratamiento de grandes cantidades de datos, por medio de la aplicación de la estadística y matemática determinar patrones y tendencias que nos permitan entender el comportamiento de estos datos y poder generar conocimiento.



La minería de datos reúne un conjunto de técnicas tales como: regresión logística, redes bayesianas, redes neuronales, árboles de decisión entre otros. Actualmente existe un gran interés por aplicar las técnicas de minería de datos al ámbito educativo, generando la creación de Minería de Datos Educativa, una comunidad de investigación educativa que busca analizar y explorar datos de entornos educativos con el fin de entender mejor el desempeño y las condiciones de aprendizaje de los estudiantes [6].

La clasificación que emplea árboles de decisión es una de las que más se usan como un modelo predictivo [5], además de que esta técnica de minería de datos es un método rápido y eficaz para la categorización de un conjunto de datos. Dicho de otra manera, esta técnica permite clasificar una población en un modelo de segmentos de tipo ramas que construyen un árbol invertido, el cual será utilizado para predecir una variable objetivo [7]. Un árbol de decisión contiene en su estructura nodos internos, nodos de probabilidad, nodos hojas y arcos, estos serán recorridos según se vaya evaluando las condiciones hasta llegar a un nodo hoja el cual devuelve una decisión.

Ante lo expuesto previamente, el presente trabajo tiene como objetivo el desarrollo de un sistema de clasificación del nivel de adaptabilidad de estudiantes frente a la educación online. En ese sentido, se hace uso de árboles de decisión, una de las diferentes técnicas de la Inteligencia Artificial, aplicada a un dataset de 1205 registros de estudiantes y que toma en consideración diferentes factores de la situación de los estudiantes en la educación en línea. De tal manera, al aplicar la técnica de árboles de decisión, se pueda determinar si el nivel de adaptabilidad del estudiante es bajo, moderado o alto.

## Revisión de la literatura

En el trabajo de Chiok[6] se hace uso de cuatro técnicas de minería de datos como son: regresión logística, árboles de decisión, redes neuronales y redes bayesianas a un conjunto de 914 datos de muestra. Estos datos académicos fueron tomados de estudiantes matriculados en el curso de Estadística General de la UNALM de los semestres 2013 II y 2014 I, a partir de estos datos poder predecir la clasificación final que puede obtener un estudiante (Aprobado o Desaprobado) cuando este tenga que matricularse en el curso. Realizó un análisis de los resultados obtenidos en cada una de las técnicas de minería de datos por medio de la aplicación de métricas a partir de un matriz de confusión. Del análisis de resultado concluyó que la técnica de clasificación red Naive de Bayes obtuvo un atasa de clasificación de 71.0%.

Suzan[8] aplica 6 técnicas de clasificación de Machine Learning siendo estas: árbol de decisión, bosque aleatorio, redes de Naive Bayes, support vector machine, K-Nearest Neighbors y redes neuronales a un dataset que contienen datos recolectados por medio de formularios de encuesta enviados a estudiantes de los diferentes niveles educativos, estos formularios. Como resultado final de la comparación y análisis de los resultados obtenidos independientemente de cada



técnica, con un 89.63% el algoritmo de Random Forest (Bosque aleatorio) fue el mejor de los algoritmos de clasificación.

Mediante el trabajo realizado por Dazhong Wu [14] se buscó realizar un sistema de aprendizaje automático en la nube mediante Análisis Predictivo, para esto se busco el uso como herramienta el algoritmo del Bosque Aleatorio el cual analiza el desgaste de procesos, Los resultados del uso de este algoritmo han demostrado que el algoritmo de bosque aleatorio puede generar predicciones muy precisas acelerando el proceso del aprendizaje dando razones al uso del Machine Learning mediante el Análisis Predictivo.

Finalmente el trabajo realizado por Chen Tan y Jianzhong Lin [15] en la que se presenta un modelo predictivo basado en QoE que detecta aspectos técnicos de la enseñanza y e-learning para evaluar el desempeño de sistemas de educación virtual en la pandemia de COVID-19 utilizando algoritmos de minería de datos. Sus resultados mostraron que el modelo de predicción sugerido cumple los factores de exactitud del 98.3%, precisión del 98.8% y recuerdo de 99.3% al predecir los aspectos conductuales de la enseñanza y el aprendizaje electrónico para los estudiantes en los sistemas de educación virtual.

## Materiales

### Descripción del Dataset

#### Género

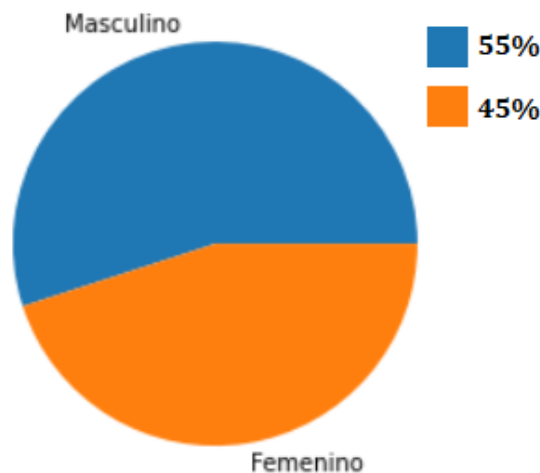


Figura 1. Distribución por género



## Tipos de Conectividad

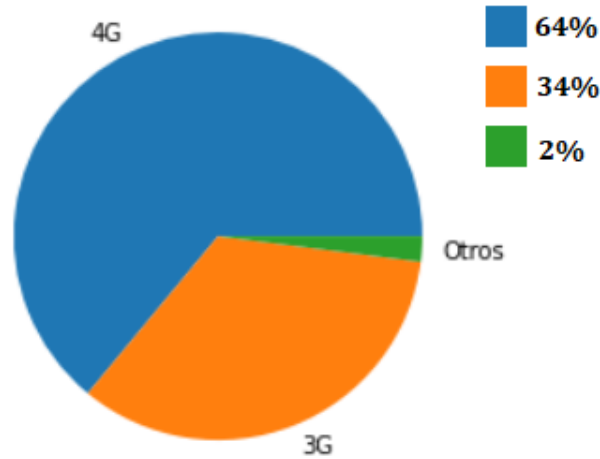


Figura 2. Distribución por tipo de conectividad

## Nivel Educativo

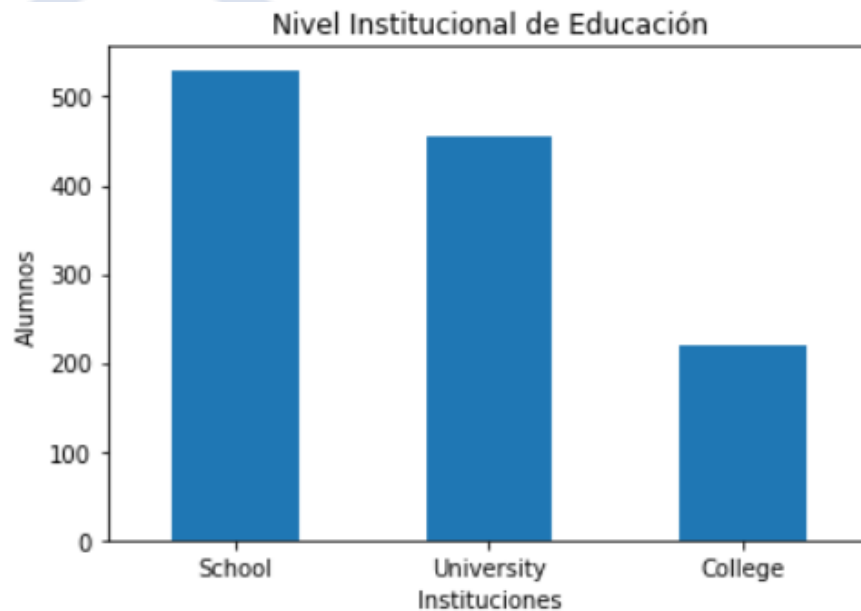


Figura 3. Distribución por nivel educativo





## Edad

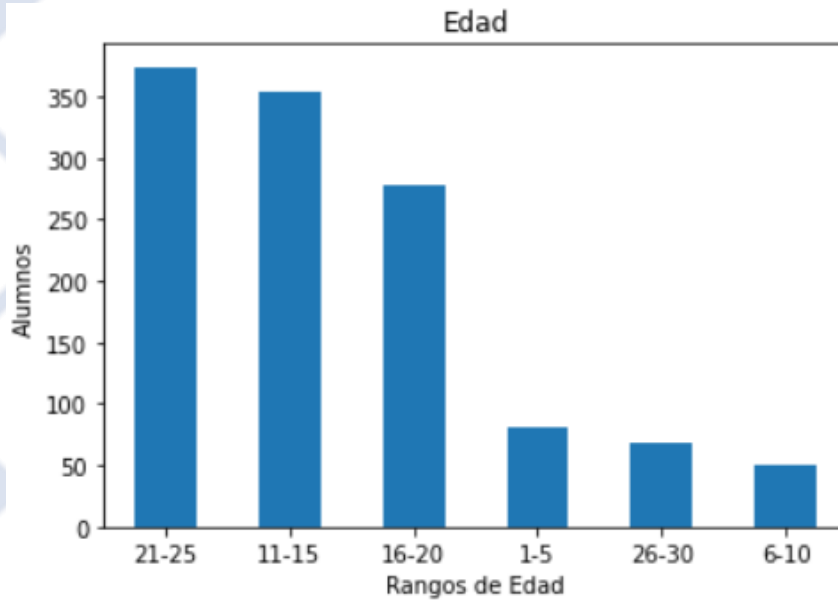


Figura 4. Distribución por edad

## Tipo de Red

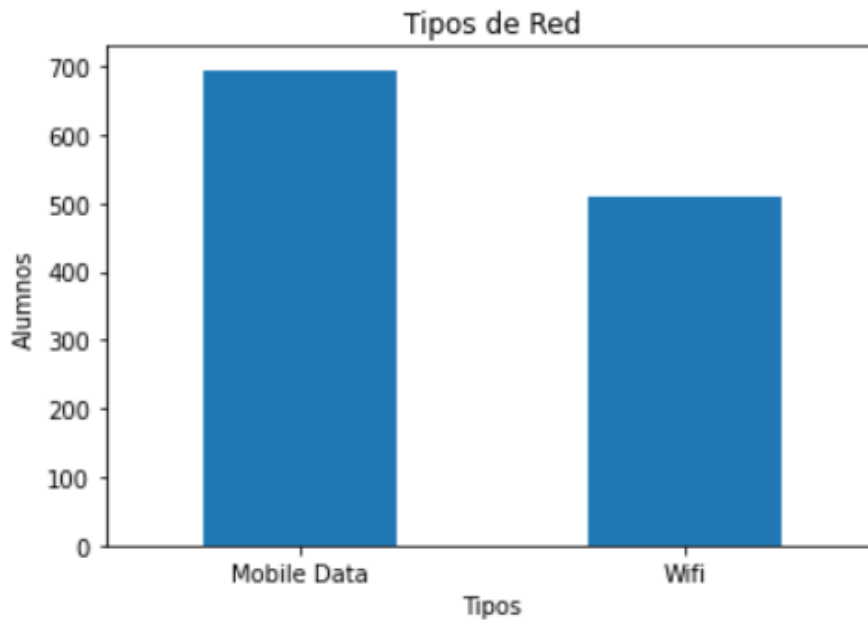


Figura 5. Distribución por tipo de red



## Condición Financiera

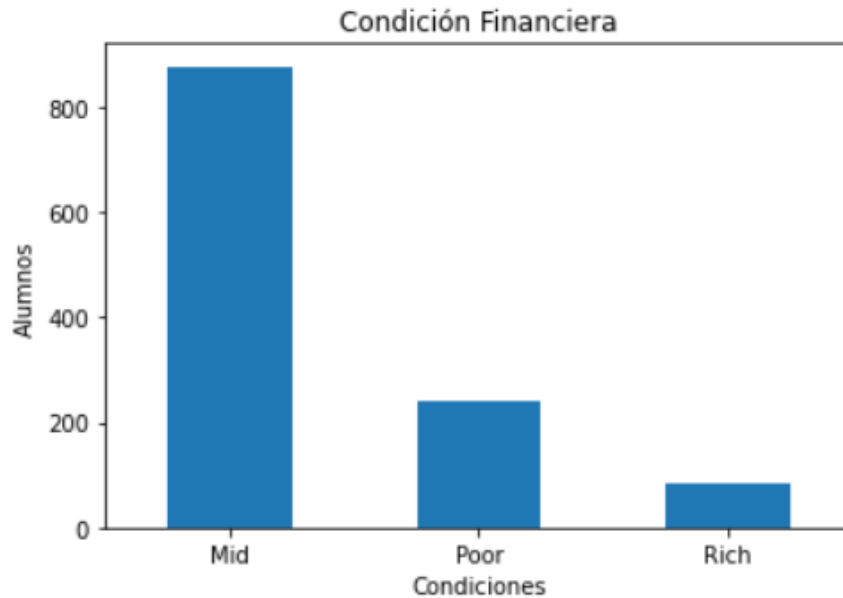


Figura 6. Distribución por condición financiera

## Herramientas

### Librerías

#### Pandas

Librería de Python que se especializa en el manejo y análisis de estructuras de datos. Permite leer y escribir fácilmente ficheros en formato CSV, Excel y SQL siendo perfecto para manejo y desarrollo de datasets y para el manejo de Machine Learning .

#### Numpy

Paquete utilizado para el uso de Machine Learning debido a que esta librería proporciona una estructura de datos de matriz que tiene algunos beneficios sobre las listas regulares de Python. los cuales son: ser más compacto, acceso rápido a los procesos de lectura y escritura, y más eficiente.

#### Sklearn

La librería scikit-learn o sklearn, es un conjunto de rutinas escritas en Python usada para realizar análisis predictivo, incluyendo clasificadores, algoritmos de clusterización, entre otros, los cuales



están basados en los paquetes NumPy, SciPy y matplotlib permitiendo normalizar, transformar y discretizar variables.

## Google Colab

Es una herramienta que nos proporciona Google permitiéndonos escribir y ejecutar código Python desde el navegador, siendo de gran ayuda y utilidad para trabajos de aprendizaje automático. Los recursos de este son limitados lo cual es necesario para que Colab pueda dar dichos recursos gratuitamente, siendo prohibidas las acciones asociadas a operaciones informáticas en bloques.

## Python

Es un lenguaje de programación interpretado de alto nivel que se utiliza para el desarrollo de aplicaciones de todo tipo, no es necesaria su compilación para ejecutarlo, sino que se ejecutan directamente por el ordenador usando un interpretador, por esto mismo no es necesario que sea traducido a lenguaje máquina.

## Método o Metodología Computacional

### Árbol de decisión (del inglés "Decision Tree", AD)

Es el algoritmo más utilizado porque proporciona un método rápido y eficaz de categorización de conjuntos de datos que es fácilmente comprensible e implementado en comparación con otros algoritmos de clasificación [7]. Un árbol de decisión se construye como una estructura de árbol de diagrama de flujo en la que cada nodo interno representa una prueba de una característica y los nodos de las hojas representan la salida final correspondiente [8]. La estructura del AD comienza con un nodo raíz y consiste en dividir los datos en subconjuntos cada vez más pequeños que contienen instancias de valores similares. Los árboles de decisión se construyen a partir de la estrategia de "Divide y Vencerás". donde cada uno de los atributos se van asignando a los nodos de manera recursiva y descendente [6].

La entropía [9] es una medida de desorden o incertidumbre, el objetivo de los modelos de aprendizaje automático y los científicos de datos en general es reducir la incertidumbre. La entropía puede calcularse por medio de la siguiente fórmula:

$$E(S) = - \sum_{i=1}^n p_i \log_2 p_i$$

Donde:

$p_i$ , define la probabilidad de que ocurra un evento.



La ganancia de información [10] se puede definir como una medida de cuánta información proporciona una característica sobre una clase. La obtención de información ayuda a determinar el orden de los atributos en los nodos de un árbol de decisión. Se puede expresar de la siguiente manera:

$$Gain(S, A) = E(s) \sum \frac{|Sv|}{|s|} H(Ev)$$

En [11] se mencionan algunas ventajas y desventajas de un AD, estos son:

## Ventajas de un AD

- Simple de entender, interpretar, visualizar.
- Los árboles de decisión llevan a cabo de forma implícita la filtración de variables o la selección de funciones.
- Puede manejar datos numéricos y categóricos.
- También puede manejar problemas de múltiples salidas.
- Los árboles de decisión requieren relativamente poco esfuerzo por parte de los usuarios para la preparación de datos.
- Las relaciones no lineales entre parámetros no afectan el rendimiento del árbol.

## Desventajas de un AD

- Los aprendices de árboles de decisión pueden crear árboles demasiado complejos que no generalizan bien los datos. Esto se llama sobreajuste.
- Los árboles de decisión pueden ser inestables porque pequeñas variaciones en los datos pueden generar un árbol completamente diferente. Esto se denomina varianza, que debe reducirse mediante métodos como bagging y boosting.
- Los algoritmos codiciosos no pueden garantizar la devolución del árbol de decisión globalmente óptimo. Esto se puede mitigar entrenando múltiples árboles, donde las características y las muestras se muestrean aleatoriamente con reemplazo.
- Los aprendices de árboles de decisión crean árboles sesgados si dominan algunas clases. Por lo tanto, se recomienda equilibrar el conjunto de datos antes de ajustarlo al árbol de decisión.



## Fuente de Información

La fuente de información fue basada en el dataset "Students Adaptability Level in Online Education", accedido en el sitio web Kaggle [2]. Este mismo dataset fue la base para otro trabajo de investigación que comparaba diversas técnicas de Machine Learning [8]. En ese sentido, el dataset cuenta con 1205 registros y 14 atributos, los cuales se detallan en la Tabla 1.

Tabla 1. Atributos del Dataset fuente utilizado

Variable	Descripción	Valores
Gender	Género del estudiante	Girl (0), Boy (1)
Age	Rango de edad del estudiante	1 - 5 (0), 6 - 10 (1), 11 - 15 (2), 16 - 20 (3), 21 - 25 (4), 26 - 30 (5), 30+ (6)
Education Level	Nivel de educación del estudiante	School (0), College (1), University (2)
Institution Type	Tipo de institución educativa	Non Government Ins (0), Government Ins (1)
IT Student	Es estudiante de Tecnologías de la Información	No (0), Yes (1)
Location in Town	Si el estudiante estudia en la ciudad	No (0), Yes (1)
Load-Shedding	Nivel de carga eléctrica	Low (0), High (1)
Financial Condition	Condición financiera de la familia del estudiante	Poor (0), Mid (1), Rich (2)
Internet Type	Tipo de internet usado mayormente en el dispositivo que utiliza el estudiante	2G (0), 3G (1), 4G (2)
Network Type	Tipo de conectividad de red usado en el equipo que utiliza el estudiante	Mobile Data (0), Wifi (1)
Class Duration	Duración diaria de las clases del estudiante	0 (0), 1 - 3 Hours (1), 3 - 6





		Hours (2)
Self LMS	La institución educativa cuenta con un Sistema para el Manejo del Aprendizaje (LMS) propio	No (0), Yes (1)
Device	Tipo de dispositivo utilizado por el estudiante para acceder a clases	Tab (0), Mobile (1), Computer (2)
Adaptability Level	Nivel de adaptabilidad de los estudiantes (variable de salida o dependiente)	Low (0), Moderate (1), High (2)

## Tratamiento de datos

El dataset mencionado anteriormente, fue cargado en la herramienta Colab. No obstante, antes de aplicar la técnica de Árboles de Decisión, es necesario el preprocesamiento de los datos para verificar y/o corregir valores vacíos (missing). En ese sentido, dicho preprocesamiento se realizó mediante código en la herramienta Colab.

```
Gender          0
Age             0
Education Level 0
Institution Type 0
IT Student      0
Location        0
Load-shedding   0
Financial Condition 0
Internet Type   0
Network Type    0
Class Duration  0
Self Lms        0
Device          0
Adaptivity Level 0
dtype: int64
```

Figura 7. Verificación de valores missing

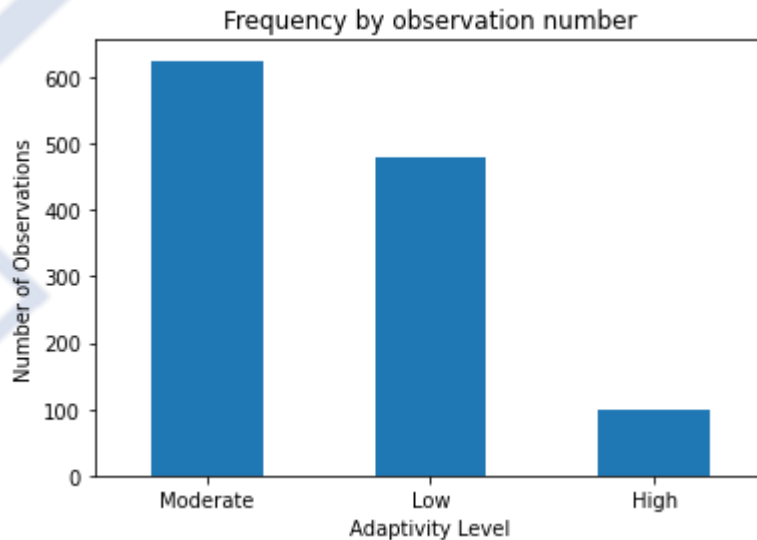


Figura 8. Verificación de desbalance de datos de salida

## Transformación de datos

Una vez que los datos han sido tratados, en el caso de los árboles de decisión, los datos respectivos de cada campo deben ser numéricos. En el caso del dataset utilizado, existían campos que eran categóricos, por lo que se procedió a transformarlos en valores numéricos.

## Implementación de modelo

Finalmente, se aplica el modelo empleando los árboles de decisión. Todo ello mediante código en Python y en el entorno de Google Colab. Mediante este paso, se obtendrá el árbol de decisión correspondiente al dataset utilizado, con el cual se podrá determinar, de acuerdo a los valores de los campos, el nivel de adaptabilidad del estudiante en la educación en línea.

## Análisis y Discusión

Para examinar la eficiencia del árbol de decisión propuesto fueron considerados los siguientes factores predictivos principales que incluyen precisión, exactitud, exhaustividad, F1 score.

En la tabla 2 observamos los valores de la matriz de confusión obtenida al realizar las predicciones de nuestro modelo y compararlas con los valores reales. A partir de estos valores podremos obtener las variables TP(True Positive), TN(True Negative), FP(False Positive), FN(False Negative) que serán necesarias para calcular los parámetros de predicción de la tabla 3.



```
Matriz de Confusión:  
[[ 14  0  4]  
 [  0 87  8]  
 [  1  6 121]]
```

Tabla 2. Matriz de confusión

	High	Low	Moderate
High	14	0	4
Low	0	87	8
Moderate	1	6	121

Tabla 3. Parámetros de predicción

Parámetros de predicción	Resultados
Precisión del modelo	0.921161825726141
Accuracy Score	0.921161825726141
Exhaustividad del modelo	0.921161825726141
F1 Score del modelo	0.921161825726141

Al ejecutar la función `classification_report()` en nuestro modelo la tabla 4 que nos genera la precisión, exhaustividad, y f1-score para cada clase objeto. Al mismo tiempo de esto, además tiene algunos valores extra: Exactitud, macro avg y weighted avg.

Tabla 4. Reporte de clasificación

	Precisión	Exhaustividad	F1-score	Support
High	0.93	0.78	0.85	18
Low	0.94	0.92	0.93	95
Moderate	0.91	0.95	0.93	128



Exactitud			0.92	241
Macro avg	0.93	0.88	0.90	241
Weighted avg	0.92	0.92	0.92	241

## Conclusión

El árbol de decisión es una de las técnicas de minería de datos que pueden ser aplicados para realizar un análisis predictivo, además de ser una técnica fácil de entender y aplicar. En el desarrollo del trabajo se realizó la limpieza y el preprocesamiento de los datos para poder realizar el entrenamiento del modelo de árbol de decisión el cual obtuvo una precisión de 0.92. El uso de las diferentes herramientas y librerías de python permitieron que las diferentes fases de análisis del dataset se desarrollen de manera óptima. Como trabajo futuro se deja la implementación de un modelo propiamente para cada nivel de educación (escuela, colegio, universidad).

## Contribución de Autoría

**Luis Emanuel Araoz Valencia:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Walter Huaracha Condori:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Víctor Raúl Quispe Quicaña:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Alex Ronaldo Turpo Coila:** [Conceptualización](#), [Análisis formal](#), [Investigación](#), [Visualización](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1] R. P. S. Elizabeth, «PERCEPCIONES DE LA EDUCACIÓN VIRTUAL EN CONFINAMIENTO,» 19 Agosto 2021. [En línea]. Available: <http://201.159.222.95/bitstream/123456789/2251/1/RODRIGUEZ%20PONCE%20SUSANA%20ELIZABETH.pdf>. [Último acceso: Junio 2022].



- [2] M. H. Suzan, «Students Adaptability Level in Online Education,» Abril 2022. [En línea]. Available: <https://www.kaggle.com/datasets/mdmahmudulhasansuzan/students-adaptability-level-in-online-education>. [Último acceso: Junio 2022].
- [3] M. Minsky, «The age of Intelligent Machines: Thoughts About Artificial Intelligence,» KurzweilAI.net., 1990.
- [4] Ocaña-Fernández, Yolvi, Luis Alex Valenzuela-Fernández, y Luzmila Lourdes Garro-Aburto. "Artificial Intelligence and Its Implications in Higher Education." *Propósitos Y Representaciones* 7.2 (2019).
- [5] B. Díaz-Landa, R. Meleán-Romero y W. Marín-Rodríguez, "Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión", *Telos Revista de Estudios Interdisciplinarios en Ciencias Sociales*, vol. 23, n.º 3, pp. 616–639, septiembre de 2021. Accedido el 19 de agosto de 2022. [En línea]. Disponible: <https://doi.org/10.36390/telos233.08>
- [6] C. H. Menacho Chiok, "Predicción del rendimiento académico aplicando técnicas de minería de datos", *Anales Científicos*, vol. 78, n.º 1, p. 26, junio de 2017. Accedido el 19 de agosto de 2022. [En línea]. Disponible: <https://doi.org/10.21704/ac.v78i1.811>
- [7] P. E. Ramírez, «Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados,» 26 Enero 2018. [En línea]. [Último acceso: Junio 2022].
- [8] M. Hasan Suzan, N. A. Samrin, A. A. Biswas y A. Pramanik, "Students' Adaptability Level Prediction in Online Education using Machine Learning Approaches", en 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 6–8 de julio de 2021. IEEE, 2021. Accedido el 20 de agosto de 2022. [En línea]. Disponible: <https://doi.org/10.1109/icccnt51525.2021.9579741>





- [9] S. T. "Entropy: How Decision Trees Make Decisions". Medium. <https://towardsdatascience.com/entropy-how-decision-trees-make-decisions-2946b9c18c8> (accedido el 14 de agosto de 2022).
- [10] C. Ayuya. "Entropy and Information Gain to Build Decision Trees in Machine Learning". Engineering Education (EngEd) Program | Section. <https://www.section.io/engineering-education/entropy-information-gain-machine-learning/> (accedido el 14 de agosto de 2022).
- [11] P. Gupta. "Decision Trees in Machine Learning". Medium. <https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052> (accedido el 14 de agosto de 2022).
- [12] Shanna S. Jaggars, «Adaptability to Online Learning: Differences Across Types of Students and Academic Subject Areas» February 2013 [En Línea] <https://academiccommons.columbia.edu/doi/10.7916/D82N59NB>
- [13] M. L Garg , "Predictive Analytics: A Review of Trends and Techniques" July 2018 [En Línea] [https://www.researchgate.net/profile/Vaibhav-Kumar-16/publication/326435728\\_Predictive\\_Analytics\\_A\\_Review\\_of\\_Trends\\_and\\_Techniques/links/5c484f6692851c22a38a6027/Predictive-Analytics-A-Review-of-Trends-and-Techniques.pdf](https://www.researchgate.net/profile/Vaibhav-Kumar-16/publication/326435728_Predictive_Analytics_A_Review_of_Trends_and_Techniques/links/5c484f6692851c22a38a6027/Predictive-Analytics-A-Review-of-Trends-and-Techniques.pdf)
- [14] Dazhong W. "Cloud-Based Machine Learning for Predictive Analytics: Tool Wear Prediction in Milling" 2016 [En Línea] [https://www.researchgate.net/profile/Soundar-Kumara/publication/313456166\\_Cloud-Based\\_Machine\\_Learning\\_for\\_Predictive\\_Analytics\\_Tool\\_Wear\\_Prediction\\_in\\_Milling/links/605fbecc92851cd8ce6fbc07/Cloud-Based-Machine-Learning-for-Predictive-Analytics-Tool-Wear-Prediction-in-Milling.pdf](https://www.researchgate.net/profile/Soundar-Kumara/publication/313456166_Cloud-Based_Machine_Learning_for_Predictive_Analytics_Tool_Wear_Prediction_in_Milling/links/605fbecc92851cd8ce6fbc07/Cloud-Based-Machine-Learning-for-Predictive-Analytics-Tool-Wear-Prediction-in-Milling.pdf)
- [15] Tan, C., Lin, J. A new QoE-based prediction model for evaluating virtual education systems with COVID-19 side effects using data mining. *Soft Comput* (2021). <https://doi.org/10.1007/s00500-021-05932-w>



# Seguridad de la información en la prevención de pérdida de datos: una revisión sistemática

182

*Information security in data loss prevention: a systematic review*

## **Michael Jeanpier Rojas Valiente**

Universidad Nacional de Trujillo. Perú.

@mjrojasv@unitru.edu.pe

<https://orcid.org/0009-0004-5917-714X>

## **Jose Maria HermesCastillo Sarmiento**

Universidad Nacional de Trujillo. Perú

@jmcastillos@unitru.edu.pe


<https://orcid.org/0009-0009-2668-2381>


## **Alberto Mendoza De Los Santos**


Universidad Nacional de Trujillo. Perú

@amendozad@unitru.edu.pe

<https://orcid.org/0000-0002-0469-915X>

 **ARK:** [ark:/42411/s12/a92](https://nbn-resolving.org/urn:nbn:org:ark:42411/s12/a92)

 **DOI:** [10.48168/innosoft.s12.a92](https://doi.org/10.48168/innosoft.s12.a92)

 **PURL:** [42411/s12/a92](https://nbn-resolving.org/urn:nbn:org:ark:42411/s12/a92)

RECIBIDO 25/06/2023 • ACEPTADO 15/09/2023 • PUBLICADO 30/09/2023



## RESUMEN

Actualmente, se refleja una crecida notable en los casos de pérdida de datos, lo que puede resultar en grandes pérdidas económicas para las organizaciones. La pérdida de datos es un problema frecuente en los diferentes sistemas de información y puede ocurrir debido a diversos motivos, como fallos en el hardware, errores humanos, ataques maliciosos y desastres naturales, entre otros. Esta revisión sistemática tiene como objeto identificar el impacto de la Gestión de la seguridad de la información en la prevención de la pérdida de datos. Para realizar esta investigación, se utilizó la metodología PRISMA y se realizaron búsquedas en 6 bases de datos: Scopus, Scielo, Alicia, Google Académico, Science Direct y Redalyc. Se encontraron un total de 59 documentos, y luego de aplicar los criterios de exclusión e inclusión, se escogieron 9 documentos para su análisis. De los documentos seleccionados, se encontró que el aumento masivo de información guardada en sistemas de bases de datos ha llevado a un incremento en los intentos de violación de seguridad por parte de personas externas. Esta investigación sostiene que las organizaciones deben implementar una adecuada gestión de la seguridad de la información para proteger sus datos y la privacidad de sus clientes. Se concluye que una buena Gestión de la Seguridad de la Información ayuda a disminuir y prevenir la pérdida de datos al



identificar y corregir vulnerabilidades, lo que tiene una relación de mejora en la protección de los datos almacenados por las organizaciones.

**Palabras claves:** Base de datos, gestión de la seguridad de la información, pérdida de datos, protección de datos.

## ABSTRACT

*Currently, there is a notable increase in data loss cases, which can result in great economic losses for organizations. Data loss is a frequent problem in different information systems and can occur due to various reasons, such as hardware failure, human error, malicious attack, and natural disaster, among others. This systematic review aims to identify the impact of Information Security Management on the prevention of data loss. To carry out this research, the PRISMA methodology was used and searches were carried out in 6 databases: Scopus, Scielo, Alicia, Google Scholar, Science Direct and Redalyc. A total of 59 documents were found, and applying the exclusion and inclusion criteria, 9 documents were chosen for analysis. Of the selected documents, it was found that the massive increase in information stored in database systems has led to an increase in security breach attempts by outsiders. This research argues that organizations must implement adequate information security management to protect their data and the privacy of their customers. It is concluded that a suitable Information Security Management helps to reduce and prevent data loss by identifying and correcting vulnerabilities, which has an improvement relationship in the protection of data stored by organizations.*

**Keywords:** : Database, data loss, data protection, Information security management.

## INTRODUCCIÓN

En la actualidad, ante los distintos cambios y necesidades que se van generando día a día, las organizaciones usan distintas Tecnologías de la Información y Comunicación (TIC) con la intención de cumplir sus objetivos. Según Consuelo, "las TIC son el conjunto de tecnologías que permiten el acceso, producción, tratamiento y comunicación de información presentada en diferentes códigos (texto, imagen, sonido, ...)" [1]. Para Antonio, "la información aparece como un recurso estratégico de primer orden, cuya adecuada administración puede aportar a las empresas nuevas fórmulas de competir" [2]. Teniendo en cuenta estas definiciones podemos notar la relación de las TIC con la información de una organización, por ende, información debe ser protegida ante cualquier vulnerabilidad, ataque cibernético, desastres naturales, accidentes, etc.

En este contexto aparece la seguridad de la información. Según [3] "la seguridad de la información se orienta a proteger los activos de información sin importar su forma o estado, valiéndose de metodologías, normas, técnicas, herramientas, estructuras organizacionales,



tecnología y otros elementos, para la aplicación y gestión de las medidas de seguridad apropiadas en cada caso". Teniendo en cuenta dicha definición debe ser desarrollada correctamente para asegurar el éxito de una organización, en este contexto aparece el concepto de gestión de seguridad de la información. El cual se entiende como un desarrollo estructurado y continuo que involucra a todas las partes de una organización, incluyendo personas, procesos y tecnologías, con el objetivo de proteger la información de la organización.

Por otro lado, la protección de los datos se define como proveer permisos para el acceso a los datos, asimismo la seguridad de los datos es garantizar que los datos se puedan restaurar ante cualquier incidente [4]. Según lo mencionado anteriormente se contempla que la Gestión de la Seguridad de la Información se relaciona con la protección de los datos manejados por una organización. Para Cristian, "en la actualidad el activo máspreciado para una organización es su información y los datos que maneja" [5]. Además, Dixon expresa que "es importante adoptar un enfoque proactivo en la prevención de pérdida de datos e información, con el objetivo de anticiparse a los posibles eventos adversos. No se debe limitar únicamente a una perspectiva técnica o reactiva" [6].

Contemplando todo lo mencionado anteriormente, se puede decir que la protección de los datos es imprescindible para las organizaciones actualmente, debido a la masiva cantidad de datos sensibles que manejan. Si bien la pérdida de datos puede tener consecuencias financieras y de reputación graves, también puede resultar en una pérdida de la confianza de los clientes. Es por ello que la Gestión de la Seguridad de la Información es primordial en las organizaciones de todos los tamaños y sectores.

Aunque se han realizado esfuerzos importantes para optimizar la Gestión de la Seguridad de la Información, la pérdida de datos es aún un problema significativo. Por consiguiente, es crucial investigar cómo la Gestión de Seguridad de la Información puede prevenir la pérdida de datos.

Por esta razón, se desarrolló una revisión sistemática, la cual se centra en determinar el impacto de una correcta gestión de la seguridad de la información en la prevención de pérdida de datos, siendo la pregunta de investigación la siguiente: ¿Qué impacto tiene la Gestión de la Seguridad de la Información en la prevención de pérdida de datos?

Para realizar esta revisión sistemática se utilizó la metodología PRISMA. Buscamos la información en 6 bases de datos: Scopus, Scielo, Alicia, Google Academic, Science Direct y Redalyc. Se encontró en total 59 documentos y aplicando los criterios se consideró 9 documentos.





## Métodos y Metodología computacional

### Tipo de Estudio

En el actual trabajo se realizó una revisión sistemática de la literatura enfocándose en la Gestión de la Seguridad de la información y el impacto en la prevención de pérdida de datos, para ello se hizo uso de la metodología PRISMA.

La pregunta de investigación planteada que conducirá el desarrollo metodológico fue la siguiente: ¿Qué impacto tiene la Gestión de la Seguridad de la Información en la prevención de pérdida de datos?

### Fundamentación de la metodología

“Las revisiones sistemáticas son una manera eficiente de presentar la evidencia disponible sobre un tópico determinado. En una publicación se puede sintetizar la evidencia de cientos de investigaciones discriminando aquellos estudios que están bien diseñados de los de mala calidad” [7].

Según [8], “la finalidad de un artículo de revisión sistemática es examinar, seleccionar y determinar cuál es la bibliografía publicada óptima para responder una pregunta de investigación claramente formulada mediante un proceso sistemático y explícito”.

De acuerdo con [9], “PRISMA se ha concebido como una herramienta para contribuir a mejorar la claridad y la transparencia en la publicación de revisiones sistemáticas”. Por ende, la metodología planteada nos ayudara y garantizara realizar una adecuada recopilación en la obtención de información relevante que nos ayude a responder nuestra pregunta planteada.

Para poder realizar de manera correcta la revisión sistemática seguiremos las recomendaciones y pasos que nos detallan en [10], quienes como pauta esencial aconsejan que se realice “una búsqueda por al menos dos autores de manera paralela en las diversas bases de datos”. Además, nos describen la siguiente secuencia (ver Figura 1) que se debe seguir: “planteamiento de la pregunta, búsqueda en base de datos, selección de artículos, extracción de datos y análisis estadísticos” [10].





**Figura 1.**Proceso de elaboración de una revisión sistemática

*Nota.* Proceso de elaboración de una revisión sistemática. Tomada de “Revisiones Sistemáticas: definición y nociones básicas”.

## Estrategia de búsqueda

Para empezar con la búsqueda se partió identificando y combinando términos referentes a la pregunta de investigación planteada: “Information security management”, “database”, “data protection”, “data loss”. Además, se procedió a delimitar los motores de base de datos tales como Scopus, Alicia, Redalyc, Science Direct, Scielo y Google Académico.

Posteriormente se determinó establecer relaciones entre los términos identificados mediante operadores booleanos, así como también se acortó la cantidad de publicaciones a un marco temporal menor a los 5 años (es decir, desde el año 2019 hasta la actualidad); aclarar también que el marco temporal mencionado fue tomado en cuenta como un criterio de inclusión, como se observa en la Tabla 1.

Las fórmulas usadas en cada motor de base de datos son las siguientes:

### Scopus: 17

(TITLE-ABS-KEY (“information security management”) AND TITLE-ABS-KEY (“database”) OR TITLE-ABS-KEY (“data protection”)) AND (LIMIT-TO (PUBYEAR, 2023) OR LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR,



2019)) AND (LIMIT-TO (SUBJAREA, "ENGI")) AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "Spanish")).

## **Scielo: 2**

("information security management") OR ("Data Protection") AND ("database")

## **Alicia: 19**

("information security management") AND ("data loss") OR ("database")

## **Google Academic: 2**

allintitle: ("information security management") OR ("data loss") AND ("database")

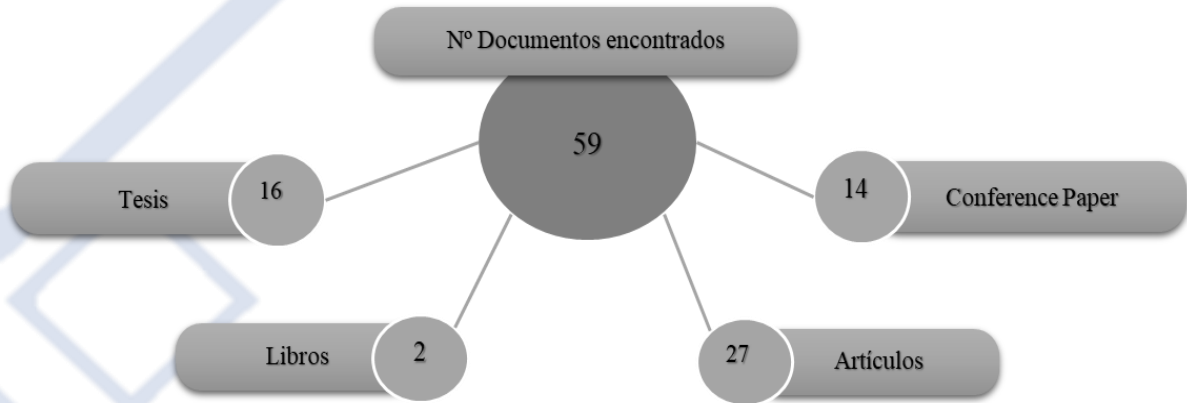
## **Science Direct: 16**

information security management AND (database OR "data loss" or "data protection")

## **Redalyc: 3**

"gestión de la seguridad de la información" + "protección de datos" site: redalyc.org

Al término de la búsqueda obtuvimos un total de 59 artículos (ver Figura 2) que se relacionan con la temática propuesta en la presente revisión sistemática los cuales se distribuyen según las bases de datos indicadas previamente: 17 en Scopus, 2 en Scielo, 19 en Alicia, 2 en Google académico, 16 en Science Direct y 3 en Redalyc. Para realizar la búsqueda en "Redalyc" se usó "Google académico" debido a que no cuenta con un apartado de búsqueda avanzada, por tal motivo en su fórmula especificamos el sitio de donde queremos obtener los artículos.



**Figura 2.** Clasificación por tipo de documento

## Criterios de inclusión y de exclusión

Como se observa en la tabla 1, se establecieron ciertos criterios los cuales fueron aplicados a los 59 artículos encontrados en nuestra primera búsqueda. Este procedimiento se realizó teniendo en cuenta a [11], que nos detalla que “los filtros son importantes especialmente cuando hay “ruido” documental y necesitamos discriminar entre muchas referencias”.

**Tabla 1.** Criterios de inclusión y exclusión

<b>Criterios</b>	
<b>Inclusión</b>	I01 Se incluyen artículos cuya fecha de publicación sea entre los años 2019 y 2023.
	I02 Se incluyen artículos cuyo idioma sea inglés y español.
	I03 Se incluyen publicaciones que tengan un enfoque en la gestión de seguridad de la información y protección de datos
<b>Exclusión</b>	E01 Se excluye publicaciones que poseen una temática diferente (salud, economía, lógica y/o matemáticas, etc.).
	E02 Se excluye publicaciones en las cuales no intervengan algún tipo de tecnologías de la información.
	E03 Se excluye publicaciones enfocadas en aspectos legales.



Para realizar la selección de artículos que nos ayudaran a darle respuesta a nuestra pregunta de investigación, empezamos comparando los criterios mencionados y el titulo quedándonos 38 artículos adecuados y a partir de ellos identificar si tenemos duplicados(n=0). Posteriormente se procedió a leer el resumen y compararlo con nuestros criterios, de esta manera se descartaron por incluir una temática irrelevante (n=17), por no incluir intervención de TI (n=0) y por enfocar aspectos legales(n=4).

Concretamente, se obtuvo 17 artículos seleccionados de los cuales 8 artículos eran inaccesibles, porque al entrar a la revista y registrarnos, de igual forma nos pedía pagar para adquirir los artículos. Por tal motivo quedaron 9 documentos seleccionados para el estudio de la revisión sistemática, tal como se aprecia en la Figura 3.

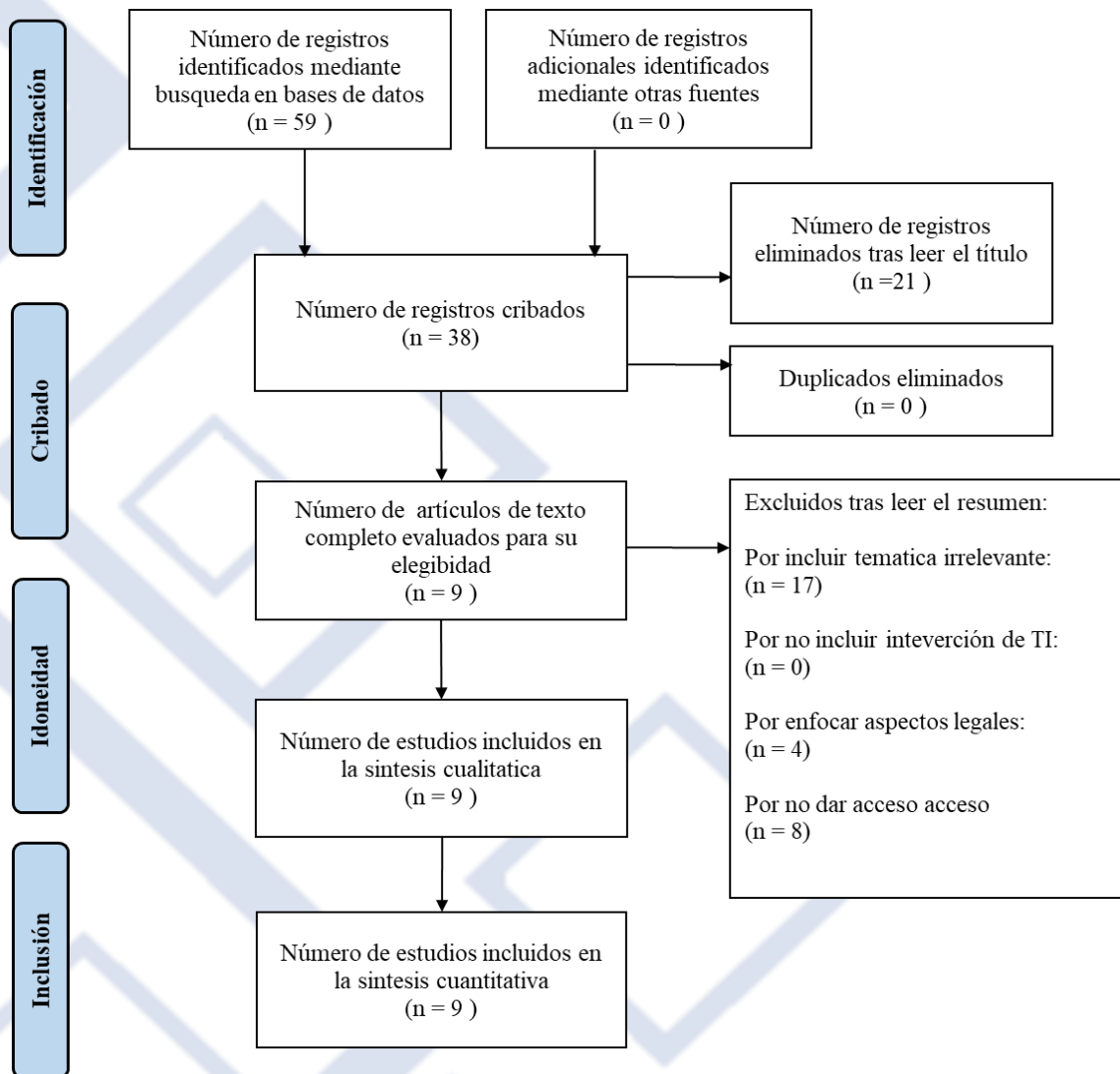


Figura 3. Diagrama de flujo PRISMA en cuatro niveles



## Resultados

A partir de los 9 documentos seleccionados se procedió a obtener los datos para conocer cuál es su relación con nuestro tema de estudio. En la Tabla 2 se muestra los datos de cada artículo tal cual se obtuvieron de las revistas donde fueron publicadas.

**Tabla 2.** Artículos seleccionados para la revisión sistemática

<b>Id</b>	<b>Autores</b>	<b>País</b>	<b>Universidad /Instituto</b>	<b>Fuente</b>	<b>Año</b>	<b>Título</b>
1	Weixiang Jiang	China	Changzhou University	Scopus	2022	Research on Machine Learning Algorithm for Internet of Things Information Security Management System Research and Implementation
2	Muñoz Hernández, Helmer; Zapata Cantero, Laura Giseth; Requena Vidal, Dina Marcela; Ricardo Villadiego, Leonela	Venezuela	Universidad del Zulia	Redalyc	2019	Riesgos informáticos y alternativas para la seguridad informática en sistemas contables en Colombia
3	Gerson De La Cruz Rodríguez, Ronny Adrián Méndez Fernández, Alberto Carlos Méndez Fernández	Perú	Universidad La Salle	Alicia	2023	Seguridad de la información en el comercio electrónico basado en ISO 27001: Una revisión sistemática



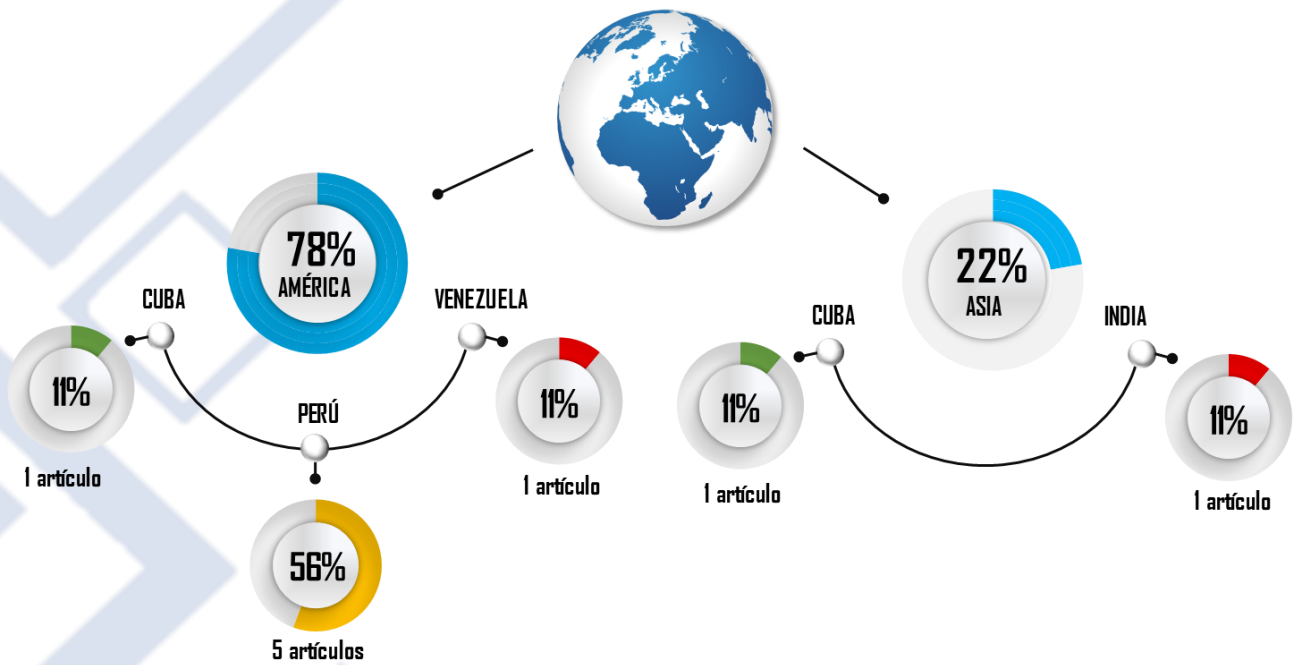


4	Camapaza Quispe, Abdon Anders	Perú	Universidad Andina del Cusco	Alicia	2019	Diseño del plan de seguridad informática basado en la ntp iso/iec 27001:2014 para la municipalidad del centro poblado de Salcedo - Puno
5	Poma Rosales, Luis Alejandro	Perú	Universidad Privada Antenor Orrego	Alicia	2019	Plan de mejora de la seguridad de la información del seguro social de salud – essalud aplicando estándar iso/IEC 27001
6	Chimoy Asto, Guillermo Enrique	Perú	Universidad Privada del Norte	Alicia	2020	Propuesta de optimización del sistema de gestión de seguridad de la información en una entidad financiera
7	Narro Mestanza, Sarita Morelia	Perú	Universidad Privada del Norte	Alicia	2022	El sistema de gestión de seguridad de la información y la gestión de riesgos en el área informática de una universidad pública, región



						Cajamarca 2020
8	Silvio Amable Machuca Vivar,Nelly Valeria Vinueza Ochoa,Carlos Roberto Sampetro Guamán,Alberto Leonel Santillán Molina	Cuba	Universidad de Cienfuegos	Scielo	2022	Habeas data y protección de datos personales en la gestión de las bases de datos
9	B.Lakshmi,K.Parish Venkata Kumar,A.Shahnaz Banu Asst, K.Anji Reddy received	India	Velagapudi Ramakrishna Siddhartha Engineering College	Google academi co	2023	Data Confidentiality and Loss Prevention using Virtual Private Database

A continuación, se realizó la búsqueda de las universidades a las que pertenecen las revistas en donde fueron publicados los 9 artículos (ver Tabla 2), para saber a qué país pertenecen. Pudimos rescatar que nuestro tema de estudio tuvo un mayor interés en los continente América y Asia, además de Perú ser el país con más artículos publicados sobre nuestro tema de investigación y el resto de países solo cuentan con un artículo publicado tal como se observa en la Figura 4.



**Figura 4.** Porcentaje de publicaciones según el país de las Universidades o Instituciones

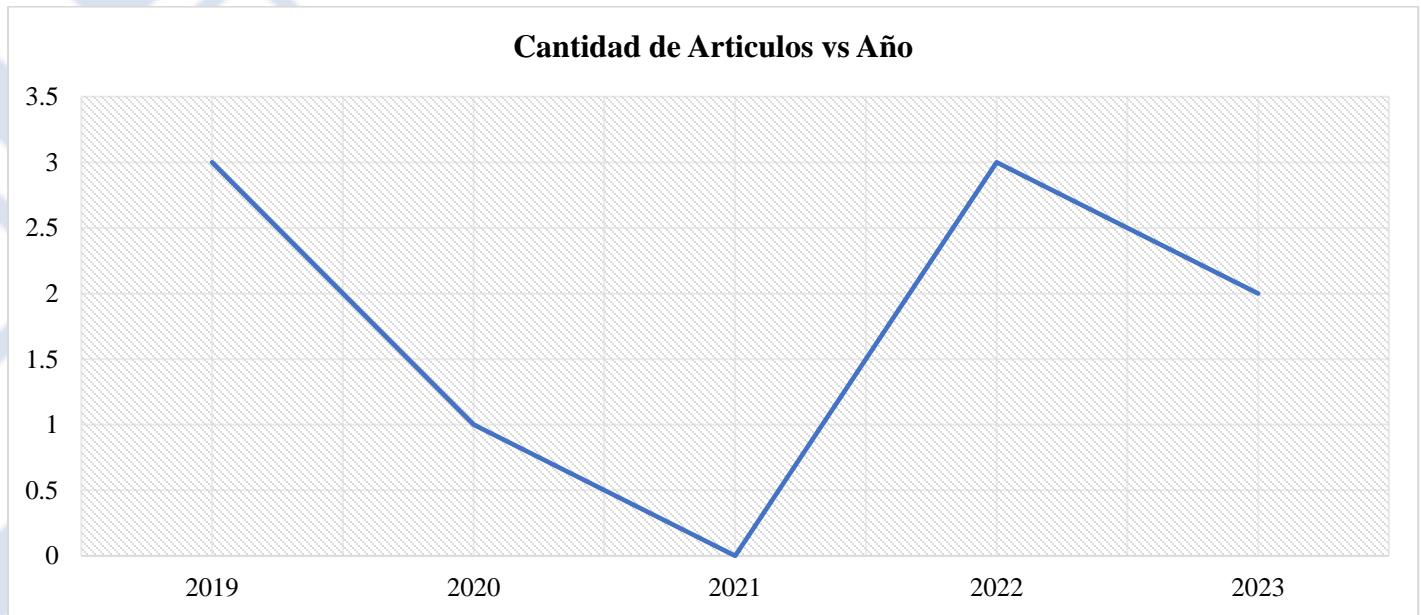
A continuación, se muestra el porcentaje de artículos publicados por año, donde concretamente podemos observar en la Tabla 3. En el año 2020 se tiene un porcentaje de 11.11%, en el año 2021 no hubo artículos publicados que hayan cumplido con nuestros criterios establecidos, en el 2023 se tiene un porcentaje de 22.22% y finalmente los años 2019 y 2022 se tiene un porcentaje de 33.33%; siendo estos los años que se publicaron una mayor cantidad de artículos referentes a nuestra investigación.

**Tabla 3.** Porcentaje de cantidad de artículos por año

Año	Cantidad de Artículos	Porcentaje
2019	3	33.33%
2020	1	11.11%
2021	0	0.00%
2022	3	33.33%
2023	2	22.22%
Total	9	100%



Además, también usamos un gráfico de líneas para mostrar la cantidad de artículos publicados a través del marco temporal establecido para esta revisión sistemática. Observando que en los años 2019 y 2022 fue donde se publicaron un mayor número de artículos referentes a nuestro tema de investigación, mientras que en el 2021 no hubo publicación alguna tal como se aprecia en la Figura 5.



**Figura 5.** Cantidad de Artículos por año

## Concepto de Sistemas de Base de datos

“Un marco de base de datos en ese punto se compone de una variedad de información relacionada y muchos, proyectos para llegar a esta información, el objetivo fundamental de una estructura de base de datos es dar una situación que sea útil y productiva para ser utilizada donde necesitamos concentrar y almacenar datos de la base” [21].

Considerando la definición citada podemos deducir la importancia de tener un espacio en donde almacenar la información, para posteriormente recuperarla y a su vez realizar diversos procesos que apoyen al desarrollo y mejora de la organización. Según [22], “los sistemas de base de datos son de gran uso, el cual va desde el uso de bases de datos ligeras, bases de datos en tiempo real (en algunas ocasiones obtenida a partir de la optimización de bases de datos relacionales) y bases de datos relacionales con potentes gestores como aplicación proveedora del servicio”.

## Concepto de Gestión de seguridad de la información



[23] define a la gestión de la seguridad de la información “como un conjunto de procesos para gestionar eficientemente la accesibilidad de la información, buscando asegurar la confidencialidad, integridad y disponibilidad de los activos de información minimizando a la vez los riesgos de seguridad de la misma”.

Teniendo en cuenta esa definición podemos centrar en la actualidad la necesidad de la gestión de la seguridad de la información en las organizaciones, debido a la cantidad masiva de datos que están en constante interacción, es decir almacenar tales datos trae consigo vulnerabilidades y filtraciones. Por ende, como comenta [24], “para gestionar la seguridad de la información de una entidad se debe partir de una premisa fundamental y es que la seguridad absoluta no existe”.

## Concepto de datos sensibles

Ana Gloria [25] “define a los datos sensibles como aquellos datos personales que afecten a la esfera más íntima de su titular, o cuya utilización indebida pueda dar origen a discriminación o conlleve un riesgo grave para éste”.

Mientras María Luisa [26] “define a los datos sensibles como aquellos datos que sirven para identificar a una persona y que en algunos casos la vulnerabilidad o exposición de estos significarían una amenaza, debido que estos datos solo pueden ser proporcionados por la persona”.

Tomando en cuenta las definiciones mencionadas los datos sensibles solo pueden ser revelados por la propia persona y solo ella puede decidir a quien se los proporcionara, es decir debe sentir la plena confianza para hacerlo y que quien las recepcione las mantenga de manera confidencial.

## Discusión

La presente revisión sistemática se centra en la importancia y el impacto que la Gestión de la Seguridad de la Información tiene en la actualidad dentro de muchas organizaciones que manejan cantidades de datos, las cuales son almacenadas en un sistema de base de datos que asegure un grado de confidencialidad y privacidad tanto para la organización como para los clientes, fortaleciendo de manera positiva la confianza de las personas para su interacción.

También asumir que es un problema latente, el cual no se puede erradicar, pero si disminuir y prevenir.

Una de las limitaciones fueron la poca cantidad de artículos en bases de datos como Scielo, Redalyc, Google Académico, además de algunos artículos de acceso restringido (es decir, se debe





pagar) por lo que nos limita a solo analizar cuáles fueron los resultados que tiene la Gestión de la Seguridad de la Información en la protección de datos.

Según los artículos seleccionados, nos detallan soluciones idóneas como la identificar los riesgos de seguridad, la integración de un sistema de gestión de seguridad informática, la capacitación constante del personal que lo usara y realizar un seguimiento para medir su efectividad dentro de la organización.

Teniendo en cuenta lo dicho anteriormente podemos afirmar que una correcta Gestión de la Seguridad de la información impacta positivamente ante la prevención de pérdida de datos.

Se recomienda que las organizaciones desarrollen una óptima y constante Gestión de la Seguridad de la información para poder minimizar la posibilidad de pérdida de datos. Además, este estudio proporcionara a los futuros investigadores tener una visión clara de la importancia y necesidad de la Gestión de la Seguridad de la Información en las organizaciones.

## Conclusiones

En la actualidad, las organizaciones y empresas han tenido una expansión muy notable alrededor del mundo, asimismo aumentaron los datos que las organizaciones deben almacenar, para poder llevar una gestión posterior ya sea administrativa, financiera, etc.

Este aumento masivo de datos trae consigo personas externas que buscan vulnerar o infringir la seguridad de los datos, y obtener información confidencial de la organización e incluso obtener información privada de sus clientes para fines ilegales o no adecuados sin el debido permiso del propietario.

Este problema está presente en todas las organizaciones, si bien es cierto casi todas cuentan con protección de seguridad de la información ninguna queda excluida de poder ser vulnerada.

Por ende, la presente investigación sostiene el impacto por el cual las organizaciones deben contar con una correcta gestión de la seguridad de la información. Y que el peligro de caer en una pérdida y/o robo de la información es constante, por lo tanto, las organizaciones deben tener una mayor concentración en la protección de sus datos asegurándole al cliente una mayor confianza.

Según la revisión de la literatura seleccionada queda demostrado que una correcta gestión de la seguridad de la información impacta positivamente en la protección de los sistemas de base de datos, ya que al ser un problema constante y no tener una solución que erradique totalmente el problema, ayudara a disminuir e identificar las vulnerabilidades, luego proponer mejoras, posteriormente corregirlas y prevenir de esta manera, la pérdida de datos.



## Contribución de Autoría

**Michael Jeanpier Rojas Valiente:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#). **Jose Maria Hermes Castillo Sarmiento:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#). **Alberto Carlos Mendoza De Los Santos:** [Conceptualización](#), [Análisis formal](#), [Recursos](#), [Supervisión](#).

## Referencias

- [1] C. B. Ortí, "LAS TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN (T.I.C.)", Universidad de Valencia, Valencia, España. [En línea]. Disponible: <http://pregrado.udg.mx/sites/default/files/formatosControlEscolar/pwtic1.pdf>.
- [2] A. Paños Álvarez, "REFLEXIONES SOBRE EL PAPEL DE LA INFORMACIÓN COMO RECURSO COMPETITIVO DE LA EMPRESA", Revistas UM, vol. 2, n.º 1999, pp. 21-38, 1. [En línea]. Disponible: <https://revistas.um.es/analesdoc/article/view/2701/2671>.
- [3] J. Figueroa Suárez, R. Rodríguez-Andrade, C. Bone Obando y J. Saltos-Gómez, "La seguridad informática y la seguridad de la información", Polo del Conocimiento, vol. 2, 2017. [En línea]. Disponible: <https://polodelconocimiento.com/ojs/index.php/es/article/view/420/pdf>.
- [4] Á. Pisco Gómez, J. J. Regalado Jalca, J. Gutiérrez García, O. Quimis Sánchez, K. Marcillo Parrales y J. Marcillo Merino, FUNDAMENTOS SOBRE LA GESTIÓN DE BASE DE DATOS. 3Ciencias, 2017.
- [5] Cristian Camilo Gantiva Rincon, "ANÁLISIS DE SOLUCIONES DPL (PREVENCIÓN DE PERDIDA DE DATOS) COMO ESTRATEGIA PARA LA SEGURIDAD DE LA INFORMACIÓN EN ORGANIZACIONES COLOMBIANAS", trabajo de grado, UNIVERSIDAD NACIONAL ABIERTA Y A DISTANCIA – UNAD, Acacías, 2021.
- [6] [6] "La prevención en la pérdida de datos e información, vital en la vigilancia privada", Universidad Militar Nueva Granada, Bogotá, Informe de la facultad de relaciones internacionales, estrategia y seguridad, 2021.



- [7] Z. Ortiz, "Revisiones sistemáticas", bioquímica y patología clínica, vol. 69, no. 2, pp. 34-36, May, 2005 [En línea]. Disponible en <https://www.redalyc.org/articulo.oa?id=65169204>.
- [8] A. Torres Fonseca, D. López Hernández, "Criterios para publicar artículos de revisión sistemática", Especialidades Medico Quirúrgicas, vol. 19, no. 3, pp. 393-399, Jul-Sep., 2014 [En línea]. Disponible en <https://www.redalyc.org/articulo.oa?id=47332498021>.
- [9] C. Pérez Rodrigo, "Las revisiones sistemáticas: declaración PRISMA", Revista Española de Nutrición Comunitaria, vol. 1, no. 18, pp. 57-58, 2012 [En línea]. Disponible en [https://renc.es/imagenes/auxiliar/files/Nutr\\_1-2012%20Taller%20escritura.pdf](https://renc.es/imagenes/auxiliar/files/Nutr_1-2012%20Taller%20escritura.pdf).
- [10] B. Moreno, M. Muñoz, J. Cuellar, S. Domancic, J. Villanueva, " Revisiones Sistemáticas: definición y nociones básicas", Revista Clínica de Periodoncia, Implantología y Rehabilitación Oral, vol. 11, no. 3, pp. 184-186, 2018, doi:10.4067/S0719-01072018000300184. Disponible en <https://www.scielo.cl/pdf/piro/v11n3/0719-0107-piro-11-03-184.pdf>.
- [11] L. Arnau Sabatés, J. Sala Roca, " La revisión de la literatura científica: Pautas, procedimientos y criterios de calidad", 2020. Disponible en <https://www.scielo.cl/pdf/piro/v11n3/0719-0107-piro-11-03-184.pdf>.
- [12] W. Jiang, " Research on Machine Learning Algorithm for Internet of Things Information Security Management System Research and Implementation", Wireless Communications and Mobile Computing, p. 6, Jun, 2022, doi: 10.1155/2022/8933468.
- [13] H. Muñoz Hernández, L. G Zapata Cantero, D. M. Requena Vidal, L. Ricardo
- [14] Villadiego, " Riesgos informáticos y alternativas para la seguridad informática en sistemas contables en Colombia", Revista Venezolana de Gerencia, vol. 2, pp. 528-537, 2019, [En línea]. Disponible en <https://www.redalyc.org/articulo.oa?id=29063446029>.



- [15] G. De La Cruz Rodríguez, R. A. Méndez Fernández, A. C. Mendoza De Los Santos, "Seguridad de la información en el comercio electrónico basado en ISO 27001: Una revisión sistemática", *Revista Innovación y Software*, vol. 4, no. 2, pp. 219-236, 2023, [En línea]. Disponible en <https://n2t.net/ark:/42411/s11/a79>.
- [16] A. A. Camapaza Quispe, "Diseño del plan de seguridad informática basado en la ntp iso/iec 27001:2014 para la municipalidad del centro poblado de salcedo - puno", Título, Ing. de Sistemas, Univ. Andina del Cuzco, Cuzco, Perú, 2019, p.116. Disponible en <https://hdl.handle.net/20.500.12557/3468>.
- [17] L. A. Poma Rosales, "Plan de mejora de la seguridad de la información del seguro social de salud – Essalud aplicando estándar iso/iec 27001", Título, Ing. de Sistemas, Univ. Privada Antenor Orrego, Trujillo, Perú, 2019, p.215. Disponible en <https://hdl.handle.net/20.500.12759/6307>.
- [18] G. E. Chimoy Asto, "Propuesta de optimización del sistema de gestión de seguridad de la información en una entidad financiera", Maestría, Ing. de Sistemas, Univ. Privada del Norte, Trujillo, La Libertad, Perú, 2020. Disponible en <https://hdl.handle.net/11537/24709>.
- [19] S. M. Narro Mestanza, "El sistema de gestión de seguridad de la información y la gestión de riesgos en el área informática de una universidad pública, región Cajamarca 2020", Maestría, Ing. de Sistemas, Univ. Privada del Norte, Cajamarca, Cajamarca, Perú, 2021. Disponible en <https://hdl.handle.net/11537/30041>.
- [20] S. A. Machuca Vivar, N. V. Vinuesa Ochoa, C. R. Sampedro Guamán, A. L. Santillán Molina, "Habeas data y protección de datos personales en la gestión de las bases de datos", *Revista Universidad y Sociedad*, vol. 14, no. 2, pp. 244-251, mar.-abr., 2022.
- [21] B. Lakshmi, K. Parish Venkata Kumar, A. Shahnaz Banu, K. Anji Reddy, "Data Confidentiality and Loss Prevention using Virtual Private Database ", *International*



Journal on Computer Science and Engineering, vol. 5, no. 3, pp. 143-149, mar., 2013.

- [22] R. E. Yarleque Ferrer, "Estructura De Datos Introducción, Conceptos, Tipos De Datos, Clasificación General, Arrays, Listas Enlazadas, Pilas, Colas, Inicialización Y Asignación De Valores". 2018.
- [23] A. M. Rubinos Carvajal, H. A. Nuevo León, "Seguridad en base de datos", Revista Cubana de Ciencias Informáticas, vol. 5, no. 1, p. 17, sep., 2011. Disponible en <https://www.redalyc.org/articulo.oa?id=378343671005>.
- [24] E. Carrasquero, L. Pérez, "SISTEMA DE GESTIÓN DE SEGURIDAD EN REDES LAN", Revista Electrónica de Estudios Telemáticos, vol. 15, no. 2, pp. 133-143, jul.- dic., 2016. Disponible en <https://www.redalyc.org/articulo.oa?id=78457627003>.
- [25] M. A. Villena Aguilar, "SISTEMA DE GESTION DE SEGURIDAD DE INFORMACION PARA UNA INSTITUCION FINANCIERA", Titulo, Ing. Informática, Pontificia Univ. Católica del Perú, Lima, Perú, 2006, p.72. Disponible en <http://hdl.handle.net/20.500.12404/362>.
- [26] A. G. Robles Osolio, "El derecho a la privacidad y la protección de datos personales transfronterizos", Revista Eurolatinoamericana de Derecho Administrativo, vol. 8, no. 1, pp. 35-60, ene-jun, 2021, Doi: <https://doi.org/10.14409/redoeda.v8i1.9543>. Disponible en <https://www.redalyc.org/journal/6559/655969720001/>.





# El impacto de la Inteligencia Artificial en la mejora de la atención al cliente: Una revisión sistémica

201

*The impact of Artificial Intelligence in improving Customer Service: A Systemic Review*

## López Zavaleta Verónica

Universidad Nacional de Trujillo.Perú.

@ t533300120@unitru.edu.pe

id <https://orcid.org/0000-0001-8656-0399>

## Rojas Ahumada Kenner

Universidad Nacional de Trujillo.Perú.

@ karojasa@unitru.edu.pe


id <https://orcid.org/0000-0001-6336-9321>


## Mendoza de los Santos, Alberto


Universidad Nacional de Trujillo.Perú.

@ amendozad@unitru.edu.pe

id <https://orcid.org/0000-0002-0469-915X>

 **ARK:** <ark:/42411/s12/a90>

 **DOI:** [10.48168/innosoft.s12.a90](https://doi.org/10.48168/innosoft.s12.a90)

 **PURL:** [42411/s12/a90](https://purl.org/42411/s12/a90)

RECIBIDO 22/05/2023 • ACEPTADO 23/09/2023 • PUBLICADO 30/09/2023



## RESUMEN

La Inteligencia Artificial (IA) está ganando una importancia creciente en el sector del servicio al cliente, permitiendo automatizar procesos y mejorar la eficacia de las interacciones con los clientes. Esta revisión sistemática tiene como objetivo explorar de cómo se está utilizando la IA en el servicio de atención al cliente, identificar las limitaciones y ventajas de la implementación de la IA y analizar cómo influyen en la satisfacción del cliente, la recolección de literatura y selección se utilizó las pautas de la metodología PRISMA. A medida que la IA se ha convertido en una tecnología cada vez más relevante en el ámbito del servicio al cliente, es esencial examinar de manera sistemática su influencia en este contexto específico. Se encontró en que sectores se está utilizando más la IA en la atención al cliente y se presentarán las perspectivas futuras de esta tecnología en el servicio al cliente. Se pudo concluir que los chatbots son una de las aplicaciones más utilizadas en la atención al cliente para brindar respuestas rápidas y efectivas a las consultas de los usuarios. Sin embargo, es importante considerar los beneficios y limitaciones de esta tecnología, así como la importancia de la interacción humana en la satisfacción del cliente.



**Palabras claves:** Atención al cliente, Gestión de servicios, Inteligencia artificial, Satisfacción del cliente, PRISMA.

## ABSTRACT

*Artificial Intelligence (AI) is gaining more and more importance in the customer service industry, making it possible to automate processes and improve the effectiveness of customer interactions. This systematic review aims to analyze the use of artificial intelligence in customer service, identify the limitations and advantages of the implementation of AI and analyze how they influence customer satisfaction, the collection of literature and selection was used. the guidelines of the PRISMA methodology. As AI has become an increasingly relevant technology in the field of customer service, it is essential to systematically examine its influence in this specific context. It was found in which sectors AI is being used more in customer service and the future perspectives of this technology in customer service will be presented. It was possible to conclude that chatbots are one of the most used applications in customer service to provide quick and effective responses to user queries. However, it is important to consider the benefits and limitations of this technology, as well as the importance of human interaction in customer satisfaction.*

**Keywords:** Customer service, Service management, Artificial intelligence, Customer satisfaction, PRISMA.

## INTRODUCCIÓN

La inteligencia artificial (IA) se puede definir como un modelo de procedimiento que permite a los sistemas informáticos y máquinas imitar la capacidad de la inteligencia humana, procesando y transformando datos en información [3], implicando la visión artificial, procesamiento de lenguaje natural y sistemas expertos, con el objetivo de tomar decisiones, resolver problemas y comprender la comunicación humana [11]. En los últimos años, la aplicación de la inteligencia artificial en la mejora del servicio al cliente ha experimentado un crecimiento significativo. Este hecho se debe a que los avances en IA tienen la capacidad de mejorar la experiencia del cliente al incrementar el conocimiento de las compañías acerca de las preferencias y patrones de compra de esos clientes (Evans, 2019) [3]. En este contexto, la realización de una revisión sistemática se convierte en una herramienta valiosa para detectar los últimos avances en esta área y analizar su impacto en la experiencia del cliente. Por ende, el objetivo principal de este estudio es generar un análisis de la aplicación de la inteligencia artificial en la mejora del servicio al cliente. A través de la revisión de los artículos más recientes se pretende responder a la pregunta central del trabajo: ¿Cuáles son los principales avances de la inteligencia artificial en la mejora de la atención al cliente? Asimismo, se plantearon otras preguntas de investigación, como ¿Cuáles son las aplicaciones más efectivas de la inteligencia artificial en la atención al cliente?, ¿En qué sectores se está implementando este tipo de tecnología?, ¿Cuáles son las limitaciones principales? y



¿Cuáles son los resultados con respecto a la satisfacción del cliente al ser atendidos por una tecnología que utiliza inteligencia artificial?

La aplicación de la inteligencia artificial en la atención al cliente proporciona ventajas tanto para las compañías como para los clientes. Las empresas pueden optimizar sus servicios y brindar una experiencia de usuario mejorada a través de la automatización del proceso y operaciones de alta frecuencia y eficiencia [11]. Además, la IA puede resolver problemas que el tradicional servicio al cliente tiene dificultades para resolver [5]. así como también puede personalizar los servicios y las recomendaciones de productos al procesar las compras y preferencias anteriores de los clientes, lo cual mejora notablemente la experiencia del usuario [3]. En el futuro, la investigación en la aplicación de servicio al cliente inteligente podría resultar en una nueva revolución para la industria en términos de soporte activo, servicio personalizado y alta confiabilidad [5].

Sin embargo, también es esencial considerar las posibles limitaciones de la aplicación de la inteligencia artificial en la atención al cliente, como la falta de capacidad para manejar situaciones complejas y para comprender el lenguaje natural en su totalidad [28]. Por tanto, el presente estudio no solo busca identificar los avances más recientes en la aplicación de la IA en la atención al cliente, sino también analizar los posibles retos y limitaciones asociados a su implementación, buscando así ser de gran utilidad para las empresas que desean mejorar su experiencia de atención al cliente y aprovechar las oportunidades que ofrece la IA en este ámbito.

## **Materiales y métodos o Metodología computacional**

La presente revisión sistemática, fue realizada gracias a las pautas establecidas de la Metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Urrútia & Bonfill, 2010) [20], para garantizar la elección de calidad de la literatura y que esté alineada a nuestro tema planteado.

La exploración de artículos para escogerlos se realizó en cinco bases de datos accesibles en la web, en este caso fueron SCOPUS, SCIENCEDIRECT, ACADEMIC GOOGLE, SCIELO, DIALNET y RESEARCHGATE. La pregunta que se planteó para orientar el proceso de búsqueda fue la siguiente: ¿Cuáles son los principales avances de la inteligencia artificial en la mejora de la atención al cliente? Durante el proceso de búsqueda se plantearon cuatro preguntas más, que resultaron ser interesantes para realizar esta revisión sistemática las preguntas son las siguientes: ¿Cuáles son las aplicaciones más efectivas de la inteligencia artificial en la atención al cliente? ¿En qué sectores se está implementando este tipo de tecnología? ¿Cuáles son las limitaciones principales? ¿Cuáles son los resultados con respecto a la satisfacción del cliente al ser atendidos por una tecnología que utiliza inteligencia artificial?



## Proceso de recolección de información

Para iniciar el proceso de búsqueda de literatura concerniente a este tema, se utilizó una combinación y términos relacionados con Innovación, servicio al cliente, mejora de la atención al cliente e inteligencia artificial.

Para la búsqueda de artículos en inglés se utilizaron como base de datos principales a SCOPUS Y SCIENCEDIRECT. En la primera base de datos mencionada se emplearon los siguientes términos con el operador booleano AND, de la siguiente manera: TITLE-ABS-KEY (artificial AND intelligence AND service AND management AND service AND management ) AND ( LIMIT-TO ( PUBYEAR , 2023 ) OR LIMIT-TO ( PUBYEAR , 2022 ) OR LIMIT-TO ( PUBYEAR , 2021 ) OR LIMIT-TO ( PUBYEAR , 2020 ) OR LIMIT-TO ( PUBYEAR , 2019 ) ).

En el caso de SCIENCEDIRECT se utilizó la siguiente combinación: artificial intelligence AND customer service AND Customer Satisfaction AND service AND management TITLE-ABS-KEY ( intelligence AND artificial ) AND ( LIMIT-TO ( PUBYEAR , 2023 ) OR LIMIT-TO ( PUBYEAR , 2022 ) OR LIMIT-TO ( PUBYEAR , 2021 ) OR LIMIT-TO ( PUBYEAR , 2020 ) OR LIMIT-TO ( PUBYEAR , 2019 ) )

Posteriormente se procedió a buscar artículos publicados en español, idioma nativo de los autores. En este caso, se realizó una búsqueda exhaustiva en ACADEMIC GOOGLE, utilizando la siguiente oración en el buscador : " El impacto de la inteligencia artificial en el servicio al cliente" y además limitando que los artículos mostrados sean desde el año 2019. De igual manera se realizó en SCIELO Y DIALNET.

## Criterios de exclusión y de inclusión

Después de la búsqueda inicial, se realizó una serie de criterios de inclusión para poder diferenciar a los artículos más relevantes para esta investigación. Estos criterios fueron: (1) Artículos publicados entre los años 2019 y 2023.(2)Estudios que investigaron el uso, impacto de la inteligencia artificial, en diferentes sectores con la banca, comercio electrónico, atención médica, turismo, hotelería.(3)Los estudios que utilizaron medidas objetivas y subjetivas para evaluar la efectividad de la inteligencia artificial para mejorar la atención al cliente.(4) Estudios que utilizaron métodos experimentales, cuasiexperimentales o de observación para evaluar el impacto de la aplicación de la inteligencia artificial en la mejora de la atención al cliente.

Por parte de los criterios de exclusión fueron las siguientes:(1) Artículos que no estaban disponibles en línea o no estaban completamente disponibles. (2) Los estudios no utilizaron objetivos o medidas para medir la eficacia de la inteligencia artificial para mejorar de la experiencia del cliente. (3) La investigación se centra en el desarrollo de herramientas de inteligencia artificial más que en su eficacia para mejorar la experiencia del cliente.





Se extrajeron documentos clave de los estudios incluidos en la revisión y se llevó a cabo una evaluación de la calidad metodológica de cada estudio. Se utilizaron técnicas de síntesis narrativa para resumir y analizar los resultados de los estudios incluidos en la revisión. En primera instancia se seleccionaron 85 artículos, de las cinco bases de datos mencionadas. En la Figura 1 se puede visualizar en porcentaje, que tanta información se recolectó en primera instancia de las bases de datos. En la Tabla 1 muestra la cantidad exacta de artículos escogidos por base de datos.

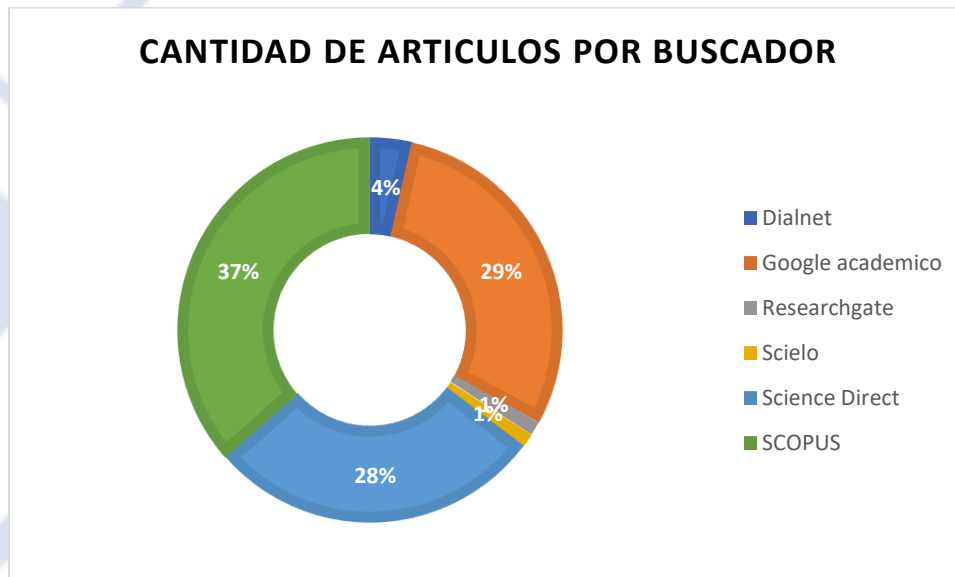


Figura 1. Porcentaje de artículos recolectados por base de datos. Elaboración propia

Tabla 1. Cantidad de artículos seleccionados por base de datos, sin criterios de inclusión y exclusión. Elaboración propia.

Buscadores	Cantidad
Dialnet	3
Google académico	25
Researchgate	1
Scielo	1
Science Direct	24
SCOPUS	31
<b>Total general</b>	<b>85</b>





En la Tabla 1 se puede observar que Scopus, ScienceDirect y Google Académico se encontró la mayoría de artículos que hablan del tema de esta revisión sistemática. En la Figura 2, se muestra como fue el flujo de identificar los artículos más relevantes para este trabajo. De un total de 85 artículos seleccionados inicialmente, se redujeron a 30. Durante el proceso de selección, se consideraron los criterios de inclusión y exclusión anteriormente mencionados.

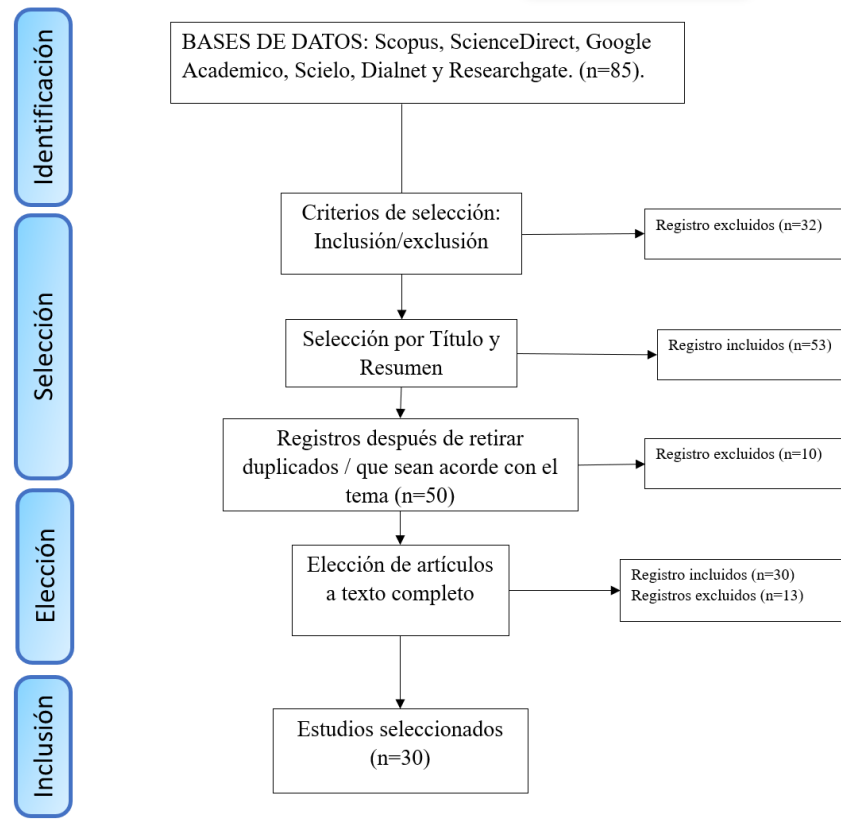


Figura 2. Flujograma de discriminación de artículos encontrados. PRISMA. Elaboración propia.

## Resultados y discusión

Algunos de los artículos más relevantes se muestran resumidos en la siguiente Tabla 2.



Tabla 2. Datos resaltantes de los estudios obtenidos para la revisión. Elaboración Propia.

Nº	Título	Base de datos	País	Año	Objetivos/Resultados
1	Impact of artificial intelligence assimilation on firm performance: The mediating effects of organizational agility and customer agility	Science Direct	Países Bajos	2022	Se descubrió que las familias en esta muestra se distinguen principalmente por sus habilidades de alfabetización en Internet, la frecuencia de uso de VA, la confianza en la tecnología y el grado preferido de mediación de medios de los padres.
2	The dark side of artificial intelligence in service: The "watching-eye" effect and privacy concerns	Science Direct	China	2023	Características de diseño de los dispositivos de IA, en lugar de la mera existencia de los dispositivos de IA, pueden causar problemas éticos como la invasión de la privacidad.
3	Key technologies of artificial intelligence in electric power customer service	Science Direct	China	2021	Aplicar el modelo al servicio al cliente inteligente de energía eléctrica, verificamos a través de una comparación experimental que el modelo de fusión propuesto mejora el rendimiento del reconocimiento de intenciones.
4	Artificial intelligence service recovery: The role of empathic response in hospitality customers' continuous usage intention	Science Direct	China	2022	Varias ideas sobre las implicaciones gerenciales para las empresas sobre cómo emplear la inteligencia artificial para facilitar la participación del cliente.
5	Reshaping the contexts of online customer engagement behavior via artificial intelligence: A conceptual framework	Science Direct	Finlandia	2021	El desarrollo futuro de la teoría del sistema de inteligencia artificial debe basarse en el concepto de sistema de información.
6	Voice artificial intelligence service failure and customer complaint behavior: The mediation effect of customer emotion	Science Direct	China	2023	Los hallazgos ofrecen información sobre la interacción entre la IA y la voz humana al revelar que el rendimiento de las IA de voz puede influir en los resultados del comportamiento del consumidor.



7	Chatbots in customer service: Their relevance and impact on service quality	Science Direct	Portugal	2022	Se conoce su efecto positivo en la calidad del servicio, que es el servicio efectivo de los bots de chat, así como la capacidad de los bots de chat en el servicio al cliente.
8	Artificial intelligence-driven music biometrics influencing customers' retail buying behavior	Science Direct	Finlandia	2021	Presenta de manera organizada el impacto de la música en las respuestas cognitivas y emocionales de los clientes y cómo éstas afectan las intenciones de comportamiento.
9	Exploring the influence of artificial intelligence technology on consumer repurchase intention: The mediation and moderation approach	Science Direct	Pakistan	2023	Los hallazgos de este estudio revelaron que la tecnología de inteligencia artificial influye positivamente en la participación del consumidor en las redes sociales y en la mejora de la tasa de conversión.
10	Robots, inteligencia artificial y realidad virtual: una aproximación en el sector del turismo	Google academico	España	2019	Cómo usar robots, big data, realidad virtual e inteligencia artificial en el ámbito del turismo en un amplio sentido.
11	Investigating the customer trust in artificial intelligence: The role of anthropomorphism, empathy response, and interaction	SCOPUS	Vietnam	2023	En general, los documentos sugieren que mejorar la calidad de la comunicación y la respuesta empática de la IA es esencial para ganar la confianza del cliente y fomentar la adopción de la tecnología.
12	Increasing customer service efficiency through artificial intelligence chatbot	SCOPUS	Brazil	2022	Los chatbots de IA reducen las colas de los centros de llamadas y permiten que los asistentes humanos se centren en problemas más complejos. Además, los chatbots de IA son ágiles, accesibles y están disponibles las 24 horas del día, los 7 días de la semana. El estudio también destaca la importancia de la innovación en procesos empresariales impulsada por la IA y cómo los chatbots de IA pueden reducir los costos del servicio al cliente.
13	Human-Computer Interaction in Customer Service: The Experience	SCOPUS	Alemania	2022	El artículo resume la investigación actual sobre la



	with AI Chatbots—A Systematic Literature Review				experiencia del cliente con chatbots de IA y brinda recomendaciones para las empresas que utilizan esta tecnología en su servicio al cliente.
14	Customer satisfaction in service delivery with artificial intelligence: A meta-analytic study	SCOPUS	Brazil	2022	En general, el estudio proporciona una visión general práctica que puede ser utilizada por investigadores y profesionales del servicio para comprender la complejidad tecnológica involucrada en la satisfacción del cliente que consume servicios con soporte de IA.
15	Conversational natural language processing for automated customer support services	SCOPUS	India	2023	El estudio presentó dos modelos de aprendizaje automático, recuperación de información y secuencia a secuencia, para mejorar los servicios de atención al cliente mediante el uso de chatbots. Los experimentos evaluaron los servicios al cliente a través de la traducción automática y el resumen de texto en términos de superposición de palabras y similitudes entre las respuestas de los chatbots y las respuestas humanas a las consultas de los clientes.
16	La usabilidad percibida de los Chatbots sobre la atención al cliente en las organizaciones	Dialnet	Perú	2021	Realiza un análisis actualizado del uso de chatbots para la atención al cliente en las organizaciones. Además, se resaltan puntos de relevancia y se examinan algunas implicaciones tanto para la investigación como para la práctica.

Con la ayuda de la Tabla 2 podemos sacar los siguientes datos para esta revisión sistemática, ya que en esta mejor especificado los artículos que se utilizaran, se podrá verificar de qué base datos se obtuvieron los artículos más relevantes, de que año son y de qué país provienen.



Figura 3. Cantidad de estudios incluidos por Base de Datos para la revisión sistemática. Elaboración propia.

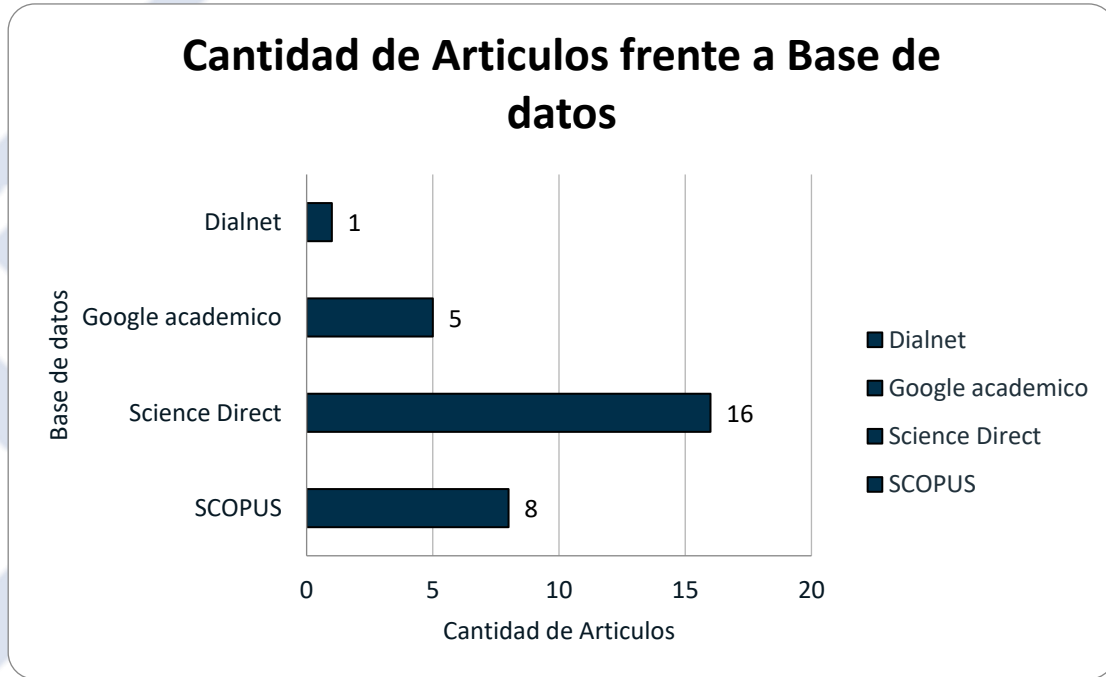


Tabla 3. Cantidad de estudios incluidos por Base de Datos para la revisión sistemática. Elaboración propia.

Base de datos	Cantidad de Artículos
Dialnet	1
Google académico	5
Science Direct	16
SCOPUS	8
Suma total	30

En la Figura 3 se observa que hay más artículos de la base de datos de ScienceDirect que pasaron por los criterios de inclusión, siendo 16, como se ve más detalladamente en la Tabla 3. Y siendo Scopus la base de datos con más artículos, con un total de 8. El que menos aportó es Dialnet con un solo artículo.



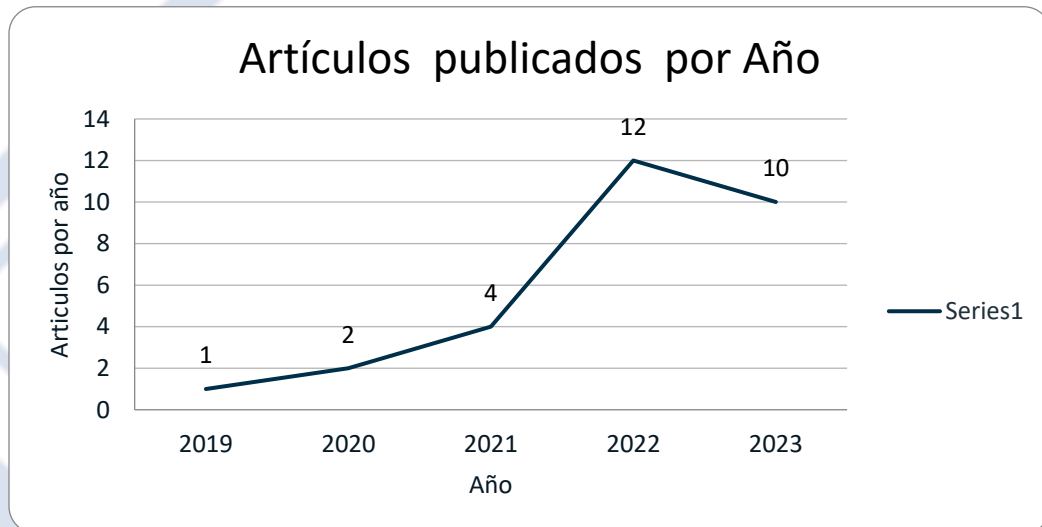


Figura 4. Cantidad de artículos incluidos por año para la revisión sistemática. Elaboración propia.

En la Figura 4 se observa que hay más artículos que los artículos del año 2022 son los que más ayudarán para esta investigación, hay en total 12 artículos de ese año, esto demuestra que las conclusiones que se mostrarán más adelante, estarán basadas en artículos publicados recientemente.

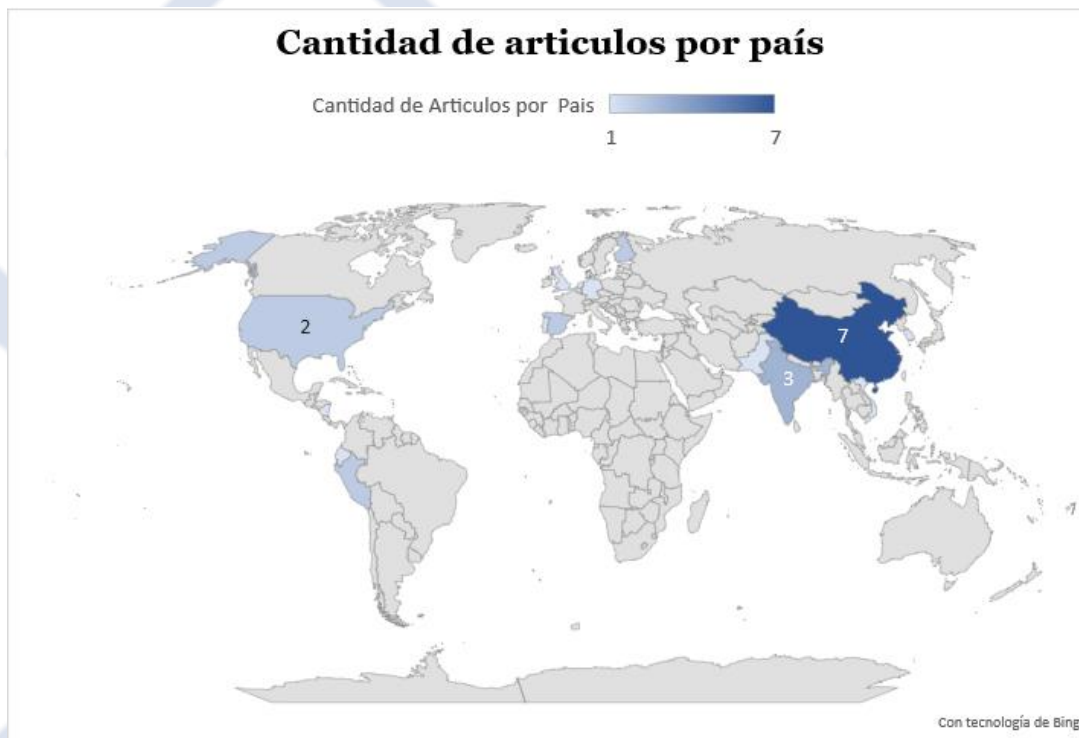


Figura 5. Mapa de la cantidad de estudios incluidos por país para la revisión sistemática. Elaboración propia.



En la Figura 5 se muestra de que país hay más artículos seleccionados, aquí se obtuvo que China tiene la mayor cantidad de artículos 7 en total (Azul oscuro), países como Alemania, Portugal, Ecuador (Celeste); tienen solo 1 artículo para nuestro interés.

## Inteligencia Artificial

La inteligencia artificial se puede definir como definirse generalmente como la capacidad de procesar y transformar datos en información para informar un comportamiento dirigido a objetivos ( Paschen, Kietzmann y Kietzmann, 2019 )[3]. Más específicamente, la IA se refiere a "agentes computacionales que actúan de manera inteligente" ( Poole y Mackworth, 2010 , p. 3)[3]. Está diseñado para simular el potencial de la energía humana en lugar de su energía real, mientras exceden su capacidad de precisión ( Dwivedi et al., 2019)[3]. Esto se logra mediante el modelado de la inteligencia biológica y natural utilizando un conjunto de modelos algorítmicos [5]. La Inteligencia Artificial en los últimos tiempos, ha cambiado el panorama empresarial. Rodrigo Perez-Vega, Valtteri Kaartemo, Cristiana R. Lages, Niloofar Borghei Razavi, Jaakko Männistö, (2021) nos dicen, que, con la capacidad de automatizar tareas, analizar e interpretar grandes cantidades de datos y tomar decisiones complejas con estos datos, la IA puede ayudar a las empresas a mejorar las operaciones, disminuir costos y elevar la satisfacción del cliente [3]. Las empresas están cada vez más adoptando las tecnologías de inteligencia artificial respaldadas por el análisis de datos como respuesta a las presiones de márgenes sostenidas, los ciclos de estrategia más cortos y las mayores expectativas de los clientes.

Estas definiciones nos llevan a decir que la inteligencia artificial es una tecnología cada vez más importante en el mundo empresarial y está cambiando la forma en que las empresas operan y compiten. La IA se utiliza para agilizar el cumplimiento de tareas, analizar grandes cantidades de datos y tomar decisiones complejas, lo que puede mejorar la eficiencia operativa, disminuir los costos y aumentar la satisfacción del cliente.

## Servicio al cliente

El servicio al cliente hace referencia al conjunto de actividades, técnicas y estrategias que las empresas utilizan para atender a las necesidades de sus demandantes, tratando de cumplir expectativas. Así, el objetivo principal del servicio al cliente es satisfacer las necesidades de estos, hacerlos fieles y mantenerlos a largo plazo.

Según Kotler y Armstrong (2018), "el servicio al cliente es la capacidad de la empresa para proporcionar el servicio esperado por el cliente, satisfacer sus necesidades y superar sus expectativas"[4]. Asimismo, Parasuraman, Zeithaml y Berry (1985) definieron el servicio al cliente como "el conjunto de actividades diseñadas para mejorar el nivel de satisfacción del



cliente, incluyendo la atención al cliente, la resolución de problemas y el seguimiento de los clientes después de la venta” [4].

## Asistentes virtuales

Los asistentes virtuales son definidos como aplicaciones capaces de comprender los comandos de voz y realizar tareas para los usuarios, son cada vez más accesibles en todo el mundo. (Rebeca Wald, Jessica T.Piotrowski,Theo Araújo,Johanna MF van Oosten, 2022 ) [1].

Además, podemos decir que los asistentes virtuales son programas informáticos diseñados para interactuar con los usuarios y ayudarles en diversas tareas, como responder preguntas, realizar reservas, proporcionar información, hacer recomendaciones y más [2]. Estos programas emplean tecnología de procesamiento de lenguaje natural y aprendizaje automático para entender el lenguaje humano y dar respuestas precisas y útiles.

Los asistentes virtuales pueden ser un asistente de voz, asistente de texto, Asistentes especializados (programas diseñados para ayudar en tareas específicas, como la programación de citas médicas o la gestión de proyectos), Asistentes inteligentes (Son programas que utilizan inteligencia artificial y aprendizaje automático para ofrecer una experiencia de usuario más personalizada y eficiente.).

## Mejora de la satisfacción del cliente a ser atendido por IA

Se sabe que el servicio al cliente tradicional suele tener muchas más dificultades que resolver, por lo que se plantea la aplicación de la inteligencia artificial con la finalidad de proporcionar una mejor experiencia de usuario para los clientes. Según [5] con la continua investigación y análisis de la tecnología de servicio al cliente ocasionará una revolución en la industria de servicio al cliente con mejores servicios personalizados, soporte activo y alta confiabilidad de apoyo.

La IA puede tener un efecto positivo en la satisfacción del cliente, pero su éxito de su implementación depende de factores como el tipo de servicio, el nivel de interacción humana y el valor percibido del servicio [26]. Esta ha sido implementada en varios sectores. Uno de ellos llega a ser el sector hotelero que según [14] permite una respuesta automática más personalizada para los clientes, por lo que muchos hoteles están optando por utilizar máquinas para brindar sus servicios, generando un impacto directo en la satisfacción de los huéspedes, ya que puede mejorar la velocidad y personalización del servicio. Por otro lado, se encuentra el sector financiero, [11] nos dice que con la aplicación y el poder de IA y ML, las empresas financieras pueden optimizar y brindar mejores servicios a clientes y usuarios. La tecnología está automatizando el proceso y ofreciendo operaciones de alta frecuencia y eficiencia.



## Calidad de servicios habilitados para IA

Hace referencia a la capacidad de los tipos de inteligencia artificial, como lo son los robots, y siendo el más conocido aplicado a la atención al cliente, el chatbot, el cual según [27] proporciona un servicio de alta calidad al usuario, teniendo un impacto significativo en la satisfacción y lealtad del usuario. Se destaca su influencia favorable en la calidad del servicio, convirtiéndose en un objetivo funcional importante para los chatbots y mostrando su potencial en el ámbito de la atención al cliente [13].

## Discusión

La Inteligencia Artificial se está implementando de manera más masiva en los últimos tiempos, para optimizar los diferentes procesos que tienen las empresas. En los últimos tiempos, se han experimentado progresos significativos en el ámbito de los sistemas que cuentan con la Inteligencia Artificial como su componente central, que les han permitido superar el rendimiento humano en tareas específicas (Silver et al., 2017, He et al. , 2015)[1].

Una de estas tareas que busca agilizar es el servicio de atención al cliente. La IA posee la capacidad de transformar radicalmente la manera en que las empresas se relacionan con sus clientes. ( [McLean & Osei-Frimpong, 2019](#) ) [3].

Además, se puede afirmar que la Inteligencia Artificial, como tecnología revolucionaria para la innovación de servicios, puede tomar decisiones y realizar acciones de forma independiente, proporcionar un mejor contacto e interacción con los clientes y se basa en tecnologías, como la computación cognitiva y el aprendizaje automático (Kuo et al., 2017, Murphy et al., 2017, Tung & Au, 2018) [6].

En este punto, para responder a las preguntas que planteamos al inicio, recopilamos una serie de datos gracias a nuestra recopilación de artículos. Donde para contestar a la pregunta: ¿Cuáles son las aplicaciones más efectivas de la inteligencia artificial en la atención al cliente ?, obtuvimos los siguientes datos.

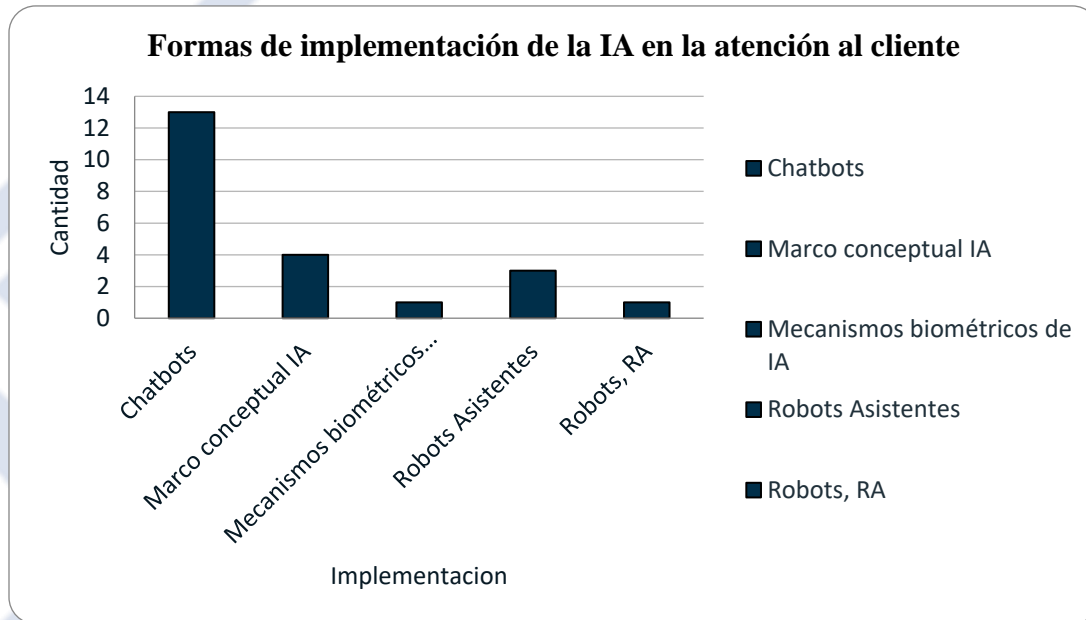


Figura 6. Tecnologías de Inteligencia artificial más usadas según bibliografía. Elaboración propia.

La Figura 6 , se puede observar que las tecnología más implementada en el servicio de atención al cliente son los Chatbots con un tota , ya que estos pueden agilizar la atención y ser más rápidos , a la par que pueden atender a varios clientes simultáneamente. La aplicación de chatbots de inteligencia artificial (IA) puede mejorar significativamente la eficiencia del servicio al cliente [23]. Los chatbots de IA reducen las colas de los centros de llamadas y permiten que los asistentes humanos se centren en problemas más complejos. Además, los chatbots de IA son ágiles, accesibles y están disponibles las 24 horas del día, los 7 días de la semana.[23].

Cabe destacar que los chatbots son un tipo de asistente virtual, pero lo consideramos aparte por ser la tecnología más usada, entonces los asistentes virtuales, solo los asistentes de voz, asistentes de texto, asistentes especializados: Son programas diseñados para ayudar en tareas específicas, como la programación de citas médicas o la gestión de proyectos, etc.

Un punto importante e interesante es en el caso de Mecanismos biométricos de IA, que se basa en la captura, por medio de cámaras, de imágenes mediante cámaras para analizar a los clientes a la hora de que ellos ven un producto, y además está relacionado con la música, ya que con esto hacía que la música que escuchaban los clientes dentro del establecimiento comercial, pueda cambiar su comportamiento indirectamente. La biometría basada en IA, como el reconocimiento facial y la música, sirve como trampolín para la innovación de creación de valor, que captura los estados cognitivos y emocionales de los clientes y genera respuestas conductuales.[14].

Nuestra segunda pregunta es: ¿En qué sectores se está implementando este tipo de tecnología?, para contestar esta pregunta obtuvimos los siguientes resultados.





La figura 7 nos muestra a los 3 principales sectores donde se está aplicando a la IA como aliado para la atención al cliente son, Financiero (5), Turismo/Hotelería (4), Servicios Públicos (4) que es como los clientes obtienen información como servicios de energía eléctrica o agua potable. Estos 3 sectores se caracterizan por ser los más demandados, ya que los clientes requieren continuamente información rápida.

La tercera pregunta fue ¿Cuáles son las limitaciones principales al integrar la IA en la atención al cliente?, respondiendo a esto según la literatura nos explica que en algunos casos los asistentes virtuales no están capacitados para responder preguntas, muy específicas de los clientes y estos quieren dejar de usar el servicio. En ese caso, estudios sugieren que la necesidad de interacción humana puede influir negativamente en la satisfacción del usuario con el uso de chatbots[15] Por ejemplo, se ha demostrado que el efecto del antropomorfismo del chatbot en la adopción aumenta a medida que aumenta la necesidad de interacción humana ( Sheehan et al., 2020)[15]. Esto sugiere que los clientes con una gran necesidad de interacción humana pueden necesitar chatbots más parecidos a una persona. Esto suele ser un reto más complejo para el desarrollo de estos.



**Figura 7.** Cantidad de Sectores donde se aplica más la IA en el servicio de atención al cliente según bibliografía. Elaboración propia.



Y la última pregunta es: ¿Cuáles son los resultados con respecto a la satisfacción del cliente al ser atendidos por una tecnología que utiliza inteligencia artificial? En general, los resultados muestran que la satisfacción del cliente se basa principalmente en la calidad del servicio y el nivel de personalización brindado. En la investigación (Xing'an Xu, Juan Liu, 2022), nos dice que para llegar a retener al cliente es necesario darle en algunas ocasiones respuestas humorísticas y hace que este sea más tolerante [12].

Otro estudio (Xingyang Lv, Yue Liub, Jingjing Luo, Yuqing Liuc, Chunxiao Li, 2021), se ha comprobado que el diseño atractivo del asistente de inteligencia artificial influye de manera positiva en la tolerancia del cliente hacia los fallos en el servicio. Se ha descubierto que hay dos factores que actúan como mediadores en este efecto: la ternura percibida y las expectativas de rendimiento. Sin embargo, la influencia de la ternura tiene un límite, que depende de la gravedad de la falla y la presión temporal a la que se enfrenta el cliente.[8]. En algunos casos los clientes se sienten desconfiados del uso de Inteligencia artificial para que los atiendan, como en la investigación de (Yaou Hu, Hyounae Kelly Min, 2023), que pone la premisa que la mera existencia de los dispositivos de IA, pueden causar problemas éticos como la invasión de la privacidad [2].

Es crucial resaltar que la satisfacción del cliente está ligada a la dificultad del problema que intenta resolver. En ciertos casos, los clientes pueden experimentar desilusión o insatisfacción si la tecnología de inteligencia artificial no puede proporcionar una solución efectiva a su situación. Por consiguiente, la satisfacción del cliente con un servicio que involucra asistencia de inteligencia artificial se basa en diversos elementos, como la calidad del servicio en sí, el nivel de experiencia del usuario, la complejidad de los problemas presentados y las expectativas previas del cliente en relación al servicio ofrecido.

## Conclusiones

La inteligencia artificial en el servicio al cliente muestra que los chatbots se están convirtiendo en una tecnología cada vez más utilizada para mejorar la eficiencia y reducir el costo del servicio al cliente. En particular, los sectores de turismo/hotelería, finanzas y servicios públicos son las aplicaciones más frecuentes de esta tecnología para la atención al cliente.

Sin embargo, se sabe que los chatbots tienen limitaciones. Por ejemplo, los asistentes virtuales pueden no estar preparados para responder preguntas muy específicas de los clientes, lo que puede llevar a la insatisfacción del cliente. Además, la experiencia del usuario es un factor clave en la satisfacción del servicio, por lo que la interacción humana es fundamental para la satisfacción del usuario. Además, descubrimos que la satisfacción del cliente con los servicios impulsados por IA depende de varios factores, que influyen: Expectativas de la calidad del servicio en sí, el nivel de experiencia del usuario, la complejidad del problema presentado e información previa del cliente, relacionada con el servicio prestado.



En síntesis, la IA tiene el potencial de ser una herramienta invaluable para elevar la calidad del servicio al cliente, su impacto es muy bueno y a medida que pasan los años se van actualizando. No obstante, resulta fundamental considerar que la tecnología no puede sustituir por completo la interacción humana, y el servicio al cliente debe ser diseñado desde la perspectiva del usuario con el fin de asegurar su satisfacción. Por lo tanto, al implementar la inteligencia artificial en el servicio al cliente, se deben considerar tanto las ventajas como las limitaciones de la inteligencia artificial para garantizar la satisfacción y lealtad del cliente.

## Contribución de Autoría

**Rojas Ahumada Kenner Alexander:** Conceptualización, Investigación, Metodología, Validación, Redacción - borrador original. **López Zavaleta Verónica:** Conceptualización, Investigación, Metodología, Visualización, Escritura, revisión y edición. **Alberto Carlos Mendoza De Los Santos:** Conceptualización, Metodología, Análisis formal, Supervisión, Curación de datos, Redacción - borrador original.

## Referencias

- [1] J. Walk, N. Kühn, M. Saidani y J. Schatte, "Artificial intelligence for sustainability: Facilitating sustainable smart product-service systems with computer vision", *J. Cleaner Prod.*, p. 136748, marzo de 2023. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jclepro.2023.136748>
- [2] Y. Hu y H. (. Min, "The dark side of artificial intelligence in service: The "watching-eye" effect and privacy concerns", *Int. J. Hospitality Manage.*, vol. 110, p. 103437, abril de 2023. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.ijhm.2023.103437>
- [3] N. Ameen, A. Tarhini, A. Reppel y A. Anand, "Customer experiences in the age of artificial intelligence", *Comput. Human Behav.*, vol. 114, p. 106548, enero de 2021. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.chb.2020.106548>
- [4] S. Z. Kamoopuri y A. Sengar, "Hi, May AI help you? An analysis of the barriers impeding the implementation and use of artificial intelligence-enabled virtual assistants in retail", *J. Retailing Consum. Services*, vol. 72, p. 103258, mayo de



2023. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jretconser.2023.103258>
- [5] X. Wu, X. Liu y Y. An, "Key technologies of artificial intelligence in electric power customer service", *Global Energy Interconnection*, vol. 4, n.º 6, pp. 631–640, diciembre de 2021. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.gloi.2022.01.005>
- [6] X. Lv, Y. Yang, D. Qin, X. Cao y H. Xu, "Artificial intelligence service recovery: The role of empathic response in hospitality customers' continuous usage intention", *Comput. Human Behav.*, vol. 126, p. 106993, enero de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.chb.2021.106993>
- [7] R. Perez-Vega, V. Kaartemo, C. R. Lages, N. Borghei Razavi y J. Männistö, "Reshaping the contexts of online customer engagement behavior via artificial intelligence: A conceptual framework", *J. Bus. Res.*, noviembre de 2020. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jbusres.2020.11.002>
- [8] X. Lv, Y. Liu, J. Luo, Y. Liu y C. Li, "Does a cute artificial intelligence assistant soften the blow? The impact of cuteness on customer tolerance of assistant service failure", *Ann. Tourism Res.*, vol. 87, p. 103114, marzo de 2021. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.annals.2020.103114>
- [9] H. Alboqami, "Trust me, I'm an influencer! - Causal recipes for customer trust in artificial intelligence influencers in the retail industry", *J. Retailing Consum. Services*, vol. 72, p. 103242, mayo de 2023. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jretconser.2022.103242>
- [10] C. Freeman y I. Beaver, "The effect of response complexity and media on user restatement with multimodal virtual assistants", *Int. J. Human-Comput. Stud.*,



- vol. 119, pp. 12–27, noviembre de 2018. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.ijhcs.2018.06.002>
- [11] V. Mahalakshmi, N. Kulkarni, K. V. Pradeep Kumar, K. Suresh Kumar, D. Nidhi Sree y S. Durga, "The Role of implementing Artificial Intelligence and Machine Learning Technologies in the financial services Industry for creating Competitive Intelligence", *Mater. Today: Proc.*, vol. 56, pp. 2252–2255, 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.matpr.2021.11.577>
- [12] X. Xu y J. Liu, "Artificial intelligence humor in service recovery", *Ann. Tourism Res.*, vol. 95, p. 103439, julio de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.annals.2022.103439>
- [13] C. V. Misischia, F. Poecze, y C. Strauss, "Chatbots in customer service: Their relevance and impact on service quality", *Procedia Comput. Sci.*, vol. 201, pp. 421–428, 2022.
- [14] W. Rodgers, F. Yeung, C. Odindo y W. Y. Degbey, "Artificial intelligence-driven music biometrics influencing customers' retail buying behavior", *J. Bus. Res.*, vol. 126, pp. 401–414, marzo de 2021. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jbusres.2020.12.039>
- [15] M. Song, X. Xing, Y. Duan, J. Cohen y J. Mou, "Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention", *J. Retailing Consum. Services*, vol. 66, p. 102900, mayo de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.jretconser.2021.102900>
- [16] S. Nazir, S. Khadim, M. Ali Asadullah y N. Syed, "Exploring the influence of artificial intelligence technology on consumer repurchase intention: The mediation and moderation approach", *Technol. Soc.*, p. 102190, diciembre de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.techsoc.2022.102190>





- [17] M. d. M. Alonso Almeida, "Robots, inteligencia artificial y realidad virtual: una aproximación en el sector del turismo", Cuadernos de Turismo, n.º 44, pp. 13–26, noviembre de 2019. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.6018/turismo.44.404711>
- [18] L. Jaime y S. Enrique, "El impacto tecnológico de los agentes virtuales en empresas integrados como herramientas de atención al cliente", 2022.
- [19] C. I. Hildebrand, "Inteligencia artificial y el aprendizaje automatizado en la industria bancaria: Mejoras en la eficiencia y en el servicio al cliente", 2020.
- [20] J. Radua, "PRISMA 2020 – An updated checklist for systematic reviews and meta-analyses", Neurosci. & Biobehav. Rev., vol. 124, pp. 324–325, mayo de 2021. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1016/j.neubiorev.2021.02.016>
- [21] N. T. K. Chi y N. Hoang Vu, "Investigating the customer trust in artificial intelligence: The role of anthropomorphism, empathy response, and interaction", CAAI Trans. Intell. Technol., agosto de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1049/cit2.12133>
- [22] Z. Yang, J. Zhou y H. Yang, "The Impact of AI's Response Method on Service Recovery Satisfaction in the Context of Service Failure", Sustainability, vol. 15, n.º 4, p. 3294, febrero de 2023. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.3390/su15043294>
- [23] I. M. D. Andrade y C. Tumelero, "Increasing customer service efficiency through artificial intelligence chatbot", Revista de Gestão, mayo de 2022. Accedido el 10 de mayo de 2023. [En línea]. Disponible: <https://doi.org/10.1108/rege-07-2021-0120>
- [24] L. Nicolescu y M. T. Tudorache, "Human-computer interaction in customer service: The experience with AI chatbots—A systematic literature review", Electronics (Basel), vol. 11, núm. 10, p. 1579, 2022.



- [25] Y.-S. (sandy) Huang y P. Dootson, "Chatbots and service failure: When does it lead to customer aggression", *J. Retail. Consum. Serv.*, vol. 68, núm. 103044, p. 103044, 2022.
- [26] L. M. Aguiar-Costa, C. A. X. C. Cunha, W. K. M. Silva, y N. R. Abreu, "Customer satisfaction in service delivery with artificial intelligence: A meta-analytic study", *RAM Rev. Adm. Mackenzie*, vol. 23, núm. 6, 2022.
- [27] C.-L. Hsu y J. C.-C. Lin, "Understanding the user satisfaction and loyalty of customer service chatbots", *J. Retail. Consum. Serv.*, vol. 71, núm. 103211, p. 103211, 2023.
- [28] A. Kaur y A. Singh, "Conversational natural language processing for automated customer support services", en *INSTRUMENTATION ENGINEERING, ELECTRONICS AND TELECOMMUNICATIONS – 2021 (IEET-2021): Proceedings of the VII International Forum*, 2023.
- [29] O. D. Casazola Cruz, G. Alfaro Mariño, J. Burgos Tejada, y O. A. Ramos More, "La usabilidad percibida de los chatbots sobre la atención al cliente en las organizaciones: una revisión de la literatura", *Interfases*, núm. 014, pp. 184–204, 2021.
- [30] Item Type y I.-R. Thesis, "La relación entre el uso de la inteligencia artificial y la satisfacción de los clientes del sector hotelero", *Edu.pe*. [En línea]. Disponible en: [https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/667149/Glasi novich\\_AG.pdf?sequence=3&isAllowed=y](https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/667149/Glasi novich_AG.pdf?sequence=3&isAllowed=y). [Consultado: 03-may-2023].



**LaSalle**  
Universidad - PERÚ