Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

Tipo de artículo: Artículos originales
Temática: Inteligencia artificial
Recibido: 20/06/2023 | Aceptado: 08/09/2023 | Publicado: 30/09/2023

Identificadores persistentes:
DOI: 10.48168/innosoft.s12.a113
ARK: ark:/42411/s12/a113
PURL: 42411/s12/a113

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

Application of decision trees for the identification of adaptability of students in online education

Luis Emanuel Araoz Valencia [1][0000-0003-4274-6380], Walter Huaracha Condori [2][0000-0002-4155-235X], Víctor Raúl Quispe Quicaña [3][0000-0003-2294-0215], Alex Ronaldo Turpo Coila [4][0000-0002-6306-0137]

[1] Universidad Nacional de San Agustín. laraozv@unsa.edu.pe
[2] Universidad Nacional de San Agustín. whuaracha@unsa.edu.pe
[3] Universidad Nacional de San Agustín. vquispequic@unsa.edu.pe
[4] Universidad Nacional de San Agustín. aturpoco@unsa.edu.pe

[*] Autor para correspondencia: aturpoco@unsa.edu.pe

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 clave: 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.

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

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

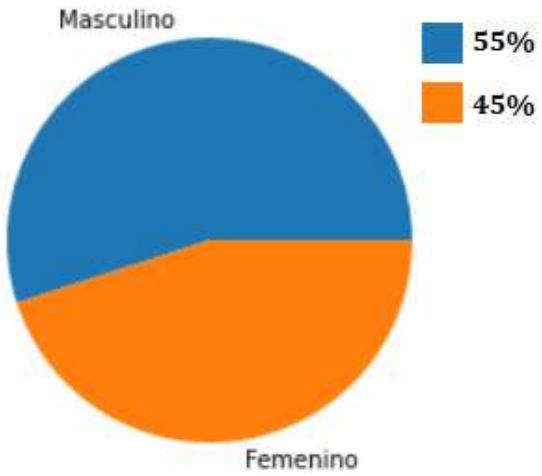

Figura 1. Distribución por género

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

Tipos de Conectividad

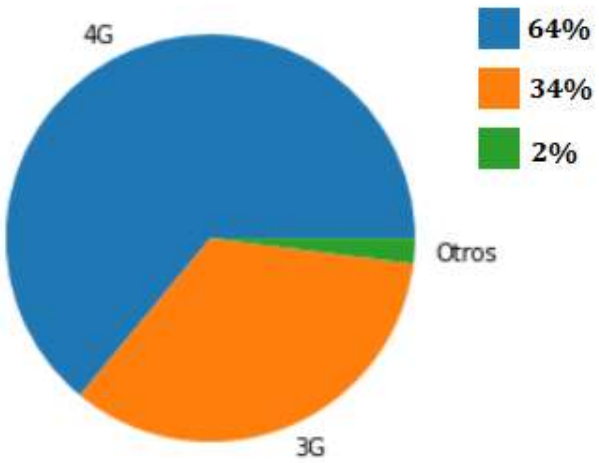

Figura 2. Distribución por tipo de conectividad

Nivel Educativo

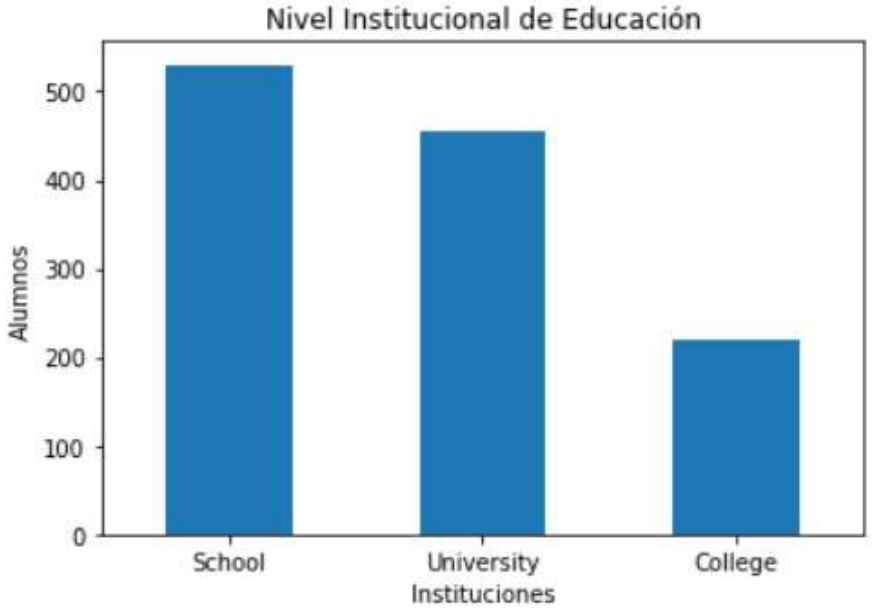

Figura 3. Distribución por nivel educativo

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

Edad

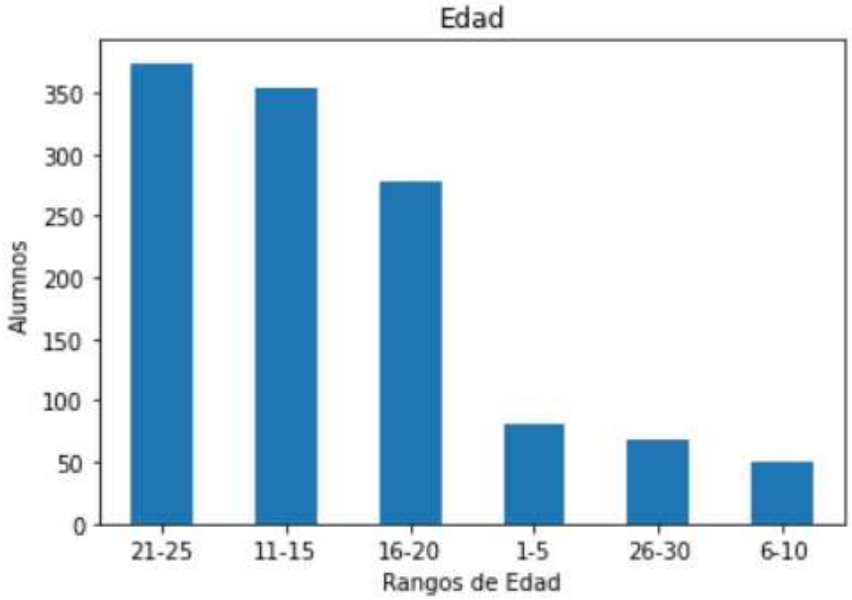

Figura 4. Distribución por edad

Tipo de Red

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

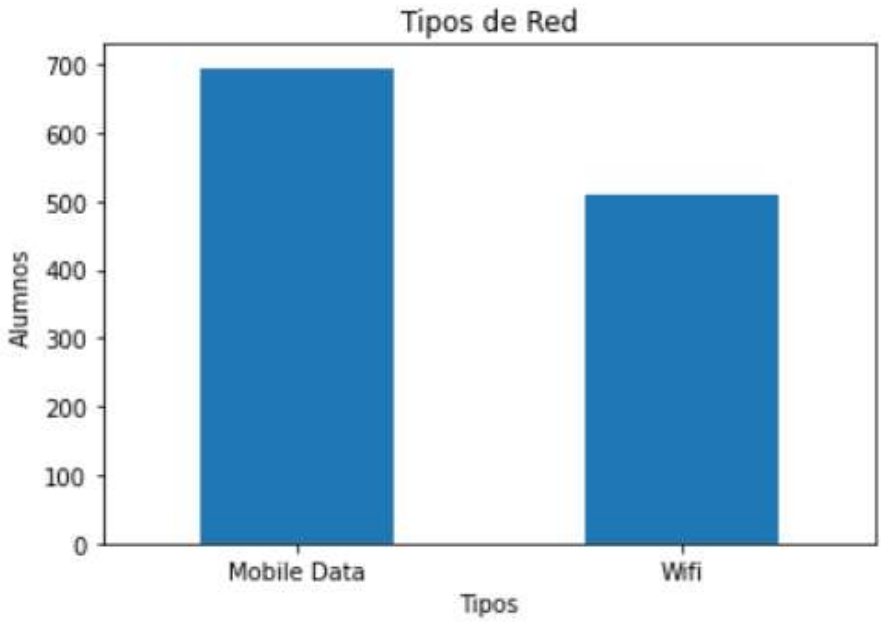

Figura 5. Distribución por tipo de red

Condición Financiera

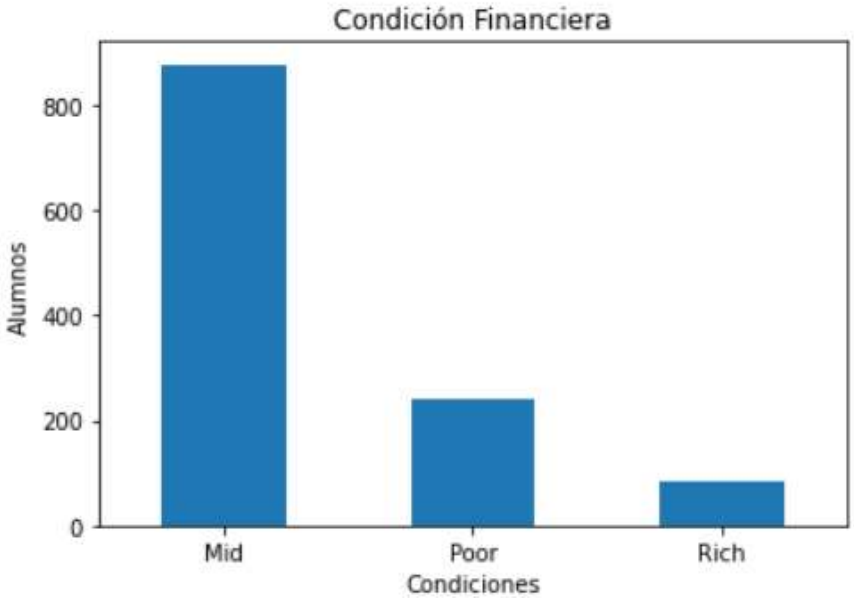

Figura 6. Distribución por condición financiera

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

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 basado 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.

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

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 expresarse 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 del 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

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

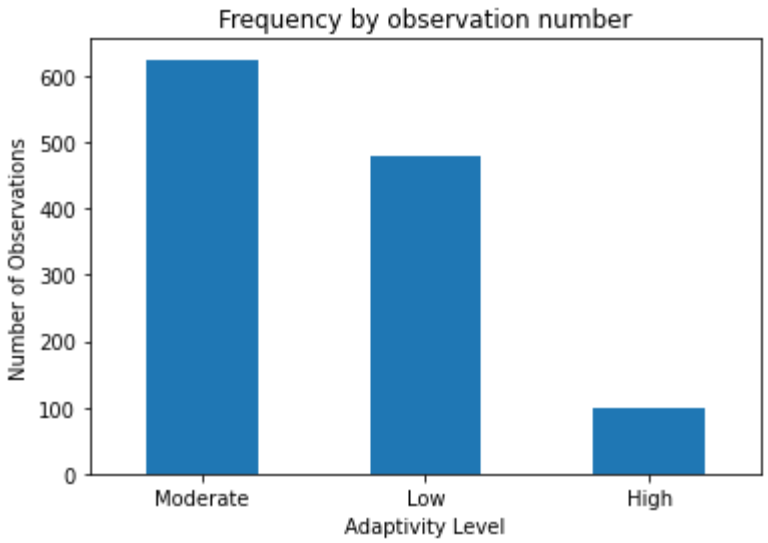

Figura 8. Verificación de desbalance de datos de salida

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.

Analisis y Discusion

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.

Revista Innovación y Software
Vol. 4, No. 2, Mes Septiembre-Febrero, 2023
ISSN: 2708-0935
Pág. 166-181
https://revistas.ulasalle.edu.pe/innosoft

```
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,» KurzweilAl.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-Rodriguez, "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 LInea] https://academiccommons.columbia.edu/doi/10.7916/D82N59NB

[13]M. L Garg , ”Predictive Analytics: A Review of Trends and Techniques” July 2018 [En Linea] 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 Linea] 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