

Tipo de artículo: Artículos originales
Temática: Inteligencia Artificial
Recibido: 23/11/2024 | Aceptado: 10/01/2025 | Publicado: 30/09/2025

Identificadores persistentes:
DOI: [10.48168/innosoft.s24.a311](https://doi.org/10.48168/innosoft.s24.a311)
ARK: [ark:/42411/s24.a311](https://nbn-resolving.org/ark:/42411/s24.a311)
PURL: [42411/s24.a311](https://nbn-resolving.org/urn:nbn:pe:ulasalle:innosoft-24-a311)

Aplicación del algoritmo KNN para la predicción de enfermedad celíaca utilizando variables clínicas y serológicas

Application of KNN algorithm for predicting celiac disease using clinical and serological variables

Daniel Lévano Rodríguez¹[[0000-0001-5652-0601](#)]*, Flor Elizabeth Cerdán León²[[0000-0001-6747-6335](#)], Cesar Rolando Salazar Giraldo³[[0009-0002-0685-8035](#)], Jadira Dina Vasquez Castro⁴[[0009-0008-3813-9346](#)], Marita Abigail Carbajal Bazán⁵[[0009-0002-5519-2495](#)], Aldana Camila Zea Mendoza⁶[[0009-0005-9917-8937](#)]

¹Universidad Nacional Tecnológica de Lima Sur. Lima, Perú. dlevano@untels.edu.pe

²Universidad Nacional Tecnológica de Lima Sur. Lima, Perú. fcerdan@untels.edu.pe

³Universidad Nacional Tecnológica de Lima Sur. Lima, Perú. 2213110208@untels.edu.pe

⁴Universidad Nacional Tecnológica de Lima Sur. Lima, Perú. 2213100026@untels.edu.pe

⁵Universidad Peruana Unión. Lima, Perú. marita.carbajal@upeu.edu.pe

⁶Universidad Peruana Unión. Lima, Perú. aldana.zea@upeu.edu.pe

*Autor para correspondencia: dlevano@untels.edu.pe

Resumen

La enfermedad celíaca corresponde a una condición autoinmune con una prevalencia cercana al 1 % a nivel global, frecuentemente subdiagnosticada debido a la escasa sospecha clínica, lo que incrementa su morbilidad y mortalidad. En este contexto, la aplicación del algoritmo K-Nearest Neighbors (KNN) surgió como un modelo predictivo para contribuir a la detección de esta enfermedad mediante variables clínicas y serológicas. Se diseñó un modelo supervisado con el algoritmo KNN utilizando variables clínicas y serológicas extraídas de una base de datos académica de 2,206 registros. Para balancear las clases, se aplicó la técnica de sobremuestreo sintético (SMOTE). Los datos fueron segmentados para entrenamiento y validación, optimizando el parámetro de clasificación mediante validación cruzada. Además, se desarrolló una plataforma web diseñada para admitir el ingreso, análisis y emisión que permite la carga, procesamiento y generación de reportes médicos con acceso por roles y estimación de probabilidad diagnóstica. Este modelo alcanzó una exactitud del 94 %, una precisión del 97 % y una sensibilidad del 91 %. El algoritmo demostró ser útil para la predicción de la enfermedad celíaca a partir de datos clínicos y serológicos, y su implementación en la web permite su integración práctica en entornos clínicos.

Palabras claves: autoinmune, Django, enfermedad, KNN, predicción

Abstract

Celiac disease is an autoimmune condition with a global prevalence close to 1 %, often underdiagnosed due to low clinical suspicion, which increases both morbidity and mortality. In this context, the application of the K-Nearest Neighbors (KNN) algorithm emerged as a predictive model to support the detection of this disease using clinical and serological variables. A supervised model was developed using the KNN algorithm and clinical and serological data extracted from an academic dataset containing 2,206 records. To address class imbalance,

the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The data were split for training and validation, optimizing the classification parameter through cross-validation. In addition, a web platform was developed to support data input, analysis, and output, allowing the uploading, processing, and generation of medical reports with role-based access and diagnostic probability estimation. The model achieved 94 % accuracy, 97 % precision, and 91 % sensitivity. The algorithm proved to be effective for predicting celiac disease based on clinical and serological data, and its web-based implementation enables practical integration in clinical environments.

Keywords: *autoimmune, Django, disease, KNN, prediction*

Introducción

La enfermedad celíaca (EC) es un trastorno autoinmune sistémico con una prevalencia global cercana al 1 % [1]. Se asocia con factores genéticos como HLA-DQ2/DQ8 (antígeno leucocitario humano), dietéticos (gluten presente en trigo y centeno) e inmunológicos (diabetes tipo 1 y tiroiditis autoinmune) [2–6]. Su diagnóstico es complejo debido a la falta de consenso en su definición y a la baja sospecha clínica, lo que conduce a una elevada tasa de casos no diagnosticados (85 %), con aumento de morbilidad y mortalidad [7–9]. Este panorama ha impulsado la implementación de estrategias de tamizaje en poblaciones de riesgo, considerando síntomas digestivos, extradigestivos, antecedentes genéticos, dieta y enfermedades autoinmunes asociadas [7, 10].

La presentación clínica clásica de EC es diarrea crónica, dolor abdominal, abdomen distendido, vómitos, y reducción de peso corporal [7, 11, 12]. En cuanto a las formas atípicas predominan síntomas como anemia ferropénica, baja estatura, infertilidad, retraso puberal, y alteraciones hepáticas [10, 13]. El diagnóstico serológico se prioriza la medición de anticuerpos IgA específicos contra la enzima transglutaminasa (tTG-IgA), por su alta sensibilidad, y anti-endomisio IgA (EMA-IgA), de elevada especificidad (97 %) [14]. En casos con deficiencia de Inmunoglobulina A (IgA), se emplean como alternativas los anticuerpos IgG anti-tTG y anti-DGP (anti-péptidos de gliadina) [15–17].

La EC presenta una mayor prevalencia en Europa y Asia (0,8 %) y una menor en Sudamérica (0,4 %) [15]. Asimismo, la aparente baja prevalencia de esta patología se atribuye, en gran medida, a la subestimación diagnóstica. Esto conlleva un aumento en las tasas de morbilidad y mortalidad [7, 9]. En este contexto, resulta necesario explorar herramientas innovadoras que permitan optimizar la predicción diagnóstica utilizando datos accesibles, clínicos y serológicos. Numerosas investigaciones han explorado la aplicación de aprendizaje automático (AA) en la identificación de EC, destacando el algoritmo KNN (K - vecinos más cercanos) como medio de predicción de EC mediante datos de genes asociados [18]. También, ha sido utilizado en conjuntos de datos con variables clínicas [16].

El aprendizaje automático, que se clasifica en aprendizaje supervisado, no supervisado, semisupervisado y el

aprendizaje por refuerzo, permite que las computadoras adquieran conocimiento a partir de datos e identifiquen patrones para ser utilizados y analizados mediante métodos estadísticos y así dotar a las computadoras de la capacidad de aprender y tomar decisiones de forma autónoma [17]. En este ámbito, el algoritmo KNN, de tipo no paramétrico, se utiliza ampliamente para clasificar variables categóricas, considerando la clase predominante entre los K vecinos más cercanos, utilizando métricas de proximidad como la distancia euclidiana para determinar la similitud, en esto se basa su simplicidad, tolerancia al ruido y capacidad de adaptación a conjuntos de datos complejos [18-21].

El desequilibrio en los conjuntos de datos puede tener un impacto en el rendimiento del algoritmo KNN. [17]. Por lo tanto, se utiliza SMOTE (Técnica de Sobremuestreo Sintético de la Clase Minoritaria). SMOTE crea nuevos datos sintéticos a partir de combinaciones de instancias próximas dentro de la clase minoritaria, favoreciendo la igualdad entre categorías y aumentar la capacidad del modelo para generalizar nuevos datos [19,20]. La implementación de KNN en aplicaciones web puede realizarse con Django, un framework en Python y la API, los cuales facilitan la comunicación entre componentes y permiten ejecutar predicciones en tiempo real, favoreciendo su integración en entornos clínicos [21-23]. Django, que abarca el Localizador Uniforme de Recursos (URL), el Marco de Caché para aplicaciones web, el modelo Modelo-Vista-Plantilla (MVT) de Django y un sistema de almacenamiento que permite gestionar tanto los datos operativos como los modelos estructurados definidos en Django [24]. Una API (Interfaz de Programación de Aplicaciones) es un servicio web que utiliza REST para proporcionar respuestas a las solicitudes del cliente mediante transferencia de representaciones del estado del recurso solicitado, usando el protocolo HTTP y es enviado en formato JSON y HTML [25,26].

Este trabajo tuvo como propósito obtener un rendimiento aceptable en términos de exactitud, precisión y sensibilidad, utilizando el algoritmo KNN para la predicción de la enfermedad celíaca.

Materiales y métodos

Con base en la metodología CRISP-DM [27] se desarrolló el sistema predictivo como se ilustra en la Figura 1. Esta incluyó la comprensión del problema clínico, el análisis y limpieza del conjunto de datos, el balanceo mediante SMOTE, la construcción del modelo KNN en Python, y su posterior despliegue en una plataforma web usando Django.

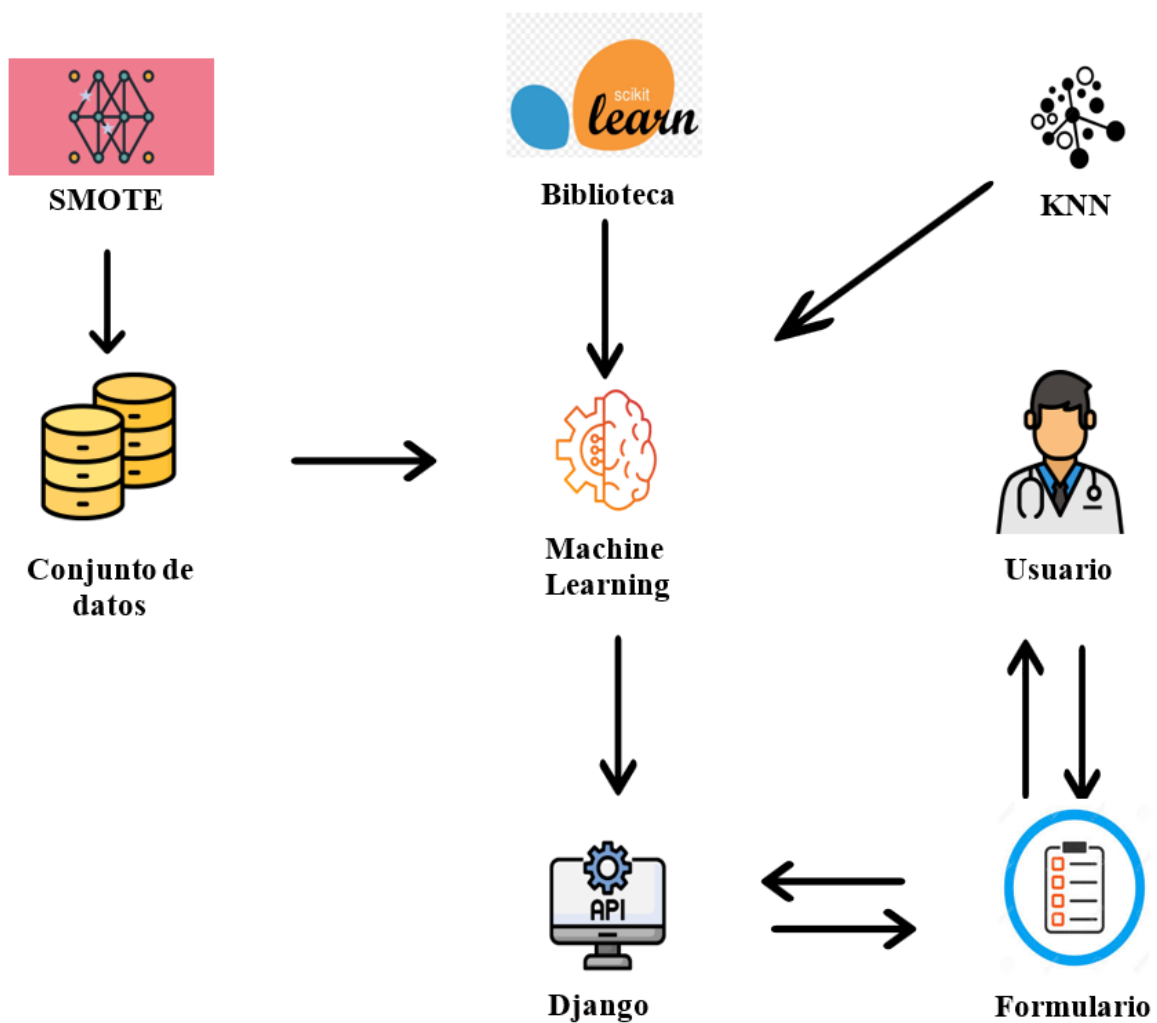


Figura 1. Arquitectura del sistema predictivo con KNN

Conjunto de datos

El conjunto de datos denominado `celiac_disease_lab_data` fue recopilado por el laboratorio de Biotecnología de la Universidad de Wageningen y se encuentra disponible públicamente a través de la plataforma Kaggle, donde fue seleccionado como fuente de datos para el presente estudio. Este diseño tiene 14 atributos y 2206 instancias con valores numéricos, de ellos 1843 pertenecen a personas con enfermedad celíaca y 363 a personas sin enfermedad celíaca.

Limpieza de datos

En el conjunto de datos no se hallaron datos vacíos, quedando 2206 registros; sin embargo, se excluyó las variables “Abdomen” y “Short.Stature” debido a que no se establecieron criterios explícitos para su definición en el conjunto de datos. Las variables “Marsh”, “Cd_type” se excluyeron por estar directamente relacionadas con el diagnóstico definitivo de la enfermedad, resultando 10 atributos para análisis. El propósito central de este estudio fue analizar clínicos y serológicos con el fin de aplicar el algoritmo KNN en la predicción de enfermedad celíaca.

La Tabla 1 presenta el análisis del conjunto de datos, detallando tanto el tipo de atributo como sus respectivos rangos de valores.

Tabla 1. Información de los atributos del conjunto de datos

ATRIBUTO	TIPO	RANGO DE VALORES
Edad	Discreto	[1,35]
Género	Nominal	Femenino $\rightarrow$ 1, Masculino $\rightarrow$ 0
Diabetes	Nominal	Si $\rightarrow$ 1, No $\rightarrow$ 0
Tipo de diabetes	Nominal	Tipo 1 $\rightarrow$ 1, Tipo 2 $\rightarrow$ 2, Ninguno $\rightarrow$ 0
Diarrea	Nominal	Esteatorrea $\rightarrow$ 0, Acuosa $\rightarrow$ 1, Inflamatorio $\rightarrow$ 2
Pérdida de peso	Nominal	Si $\rightarrow$ 1, No $\rightarrow$ 2
Heces pastosas	Nominal	Si $\rightarrow$ 1, No $\rightarrow$ 0
IgA	Continuo	[0.34, 9]
IgG	Continuo	[5,15.3]
IgM	Continuo	[0.5,2.7]

IgA: Inmunoglobulina A, IgG: Inmunoglobulina G, IgM: Inmunoglobulina M

Balanceo de datos

Se obtuvo una representación con un desequilibrio entre las clases “sí tiene” y “no tiene” con 1843 y 363 instancias respectivamente. Esta desproporción puede afectar negativamente al modelo predictivo; para mitigar esta problemática se decidió implementar la técnica SMOTE únicamente en el sobre el conjunto destinado al entrenamiento del modelo; así se crearon instancias sintéticas para la clase minoritaria basándose en las características de sus vecinos más cercanos, de esta manera se equilibró la clase “no tiene” con la clase “sí tiene” en la Figura 2. Con ello se obtuvo un equilibrio de datos de 1843 instancias para ambas clases.

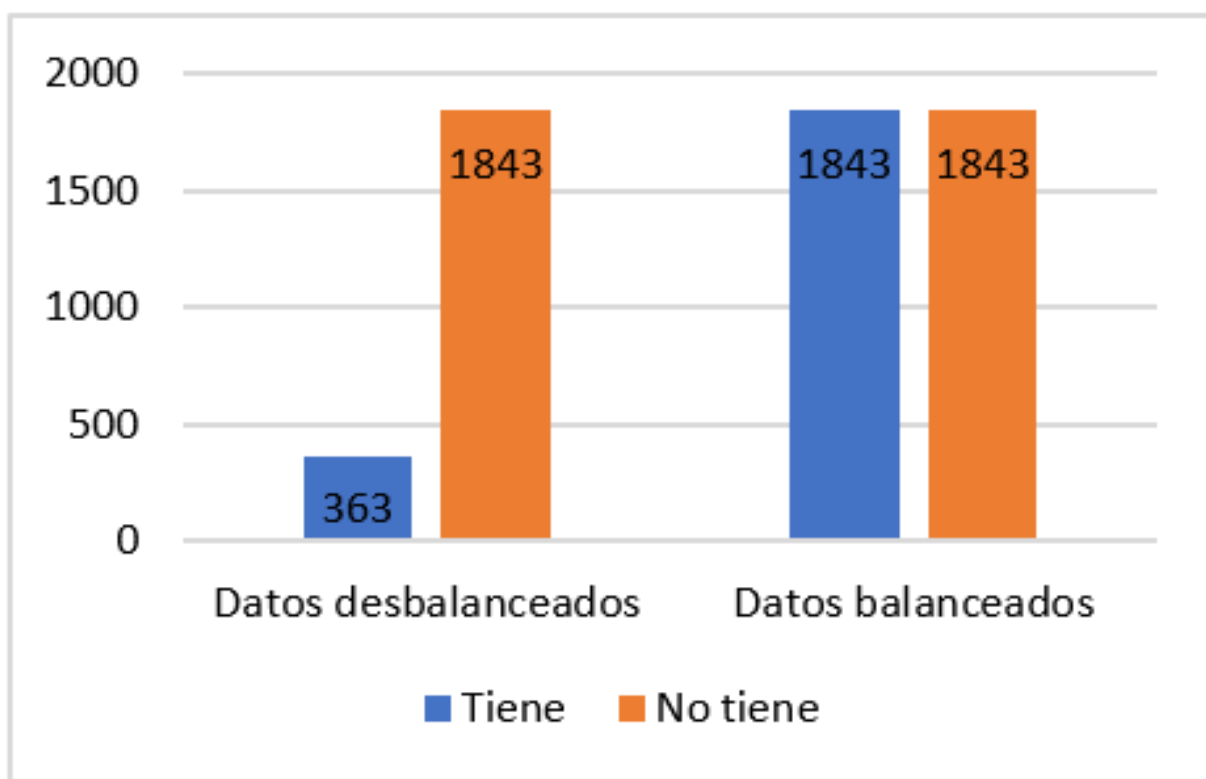


Figura 2. Balanceo de clases mediante la técnica de SMOTE

Modelado con KNN

Para el modelado predictivo se utilizó el algoritmo de vecinos más cercanos debido a su efectividad en la clasificación de datos clínicos, su facilidad de interpretación y porque no requiere suposiciones rigurosas acerca del comportamiento de las variables [17]. Se empleó la implementación del algoritmo KNN disponible en la biblioteca Scikit-learn de Python durante la etapa de entrenamiento. El conjunto de datos fue cargado en

una estructura tipo marco de datos (dataframe) y posteriormente dividido en dos subconjuntos: dividiéndose en un 70% de los registros para el proceso de entrenamiento y el 30% restante para evaluación, empleando `train_test_split` de Scikit-learn. El valor óptimo del parámetro k (número de vecinos) se determinó mediante una evaluación sistemática de la tasa de error para valores entre 1 y 25, el cual se presenta en la Figura 3. Para la selección del modelo, se optó por el valor de k que presentó la menor tasa de error en el conjunto de datos analizados.

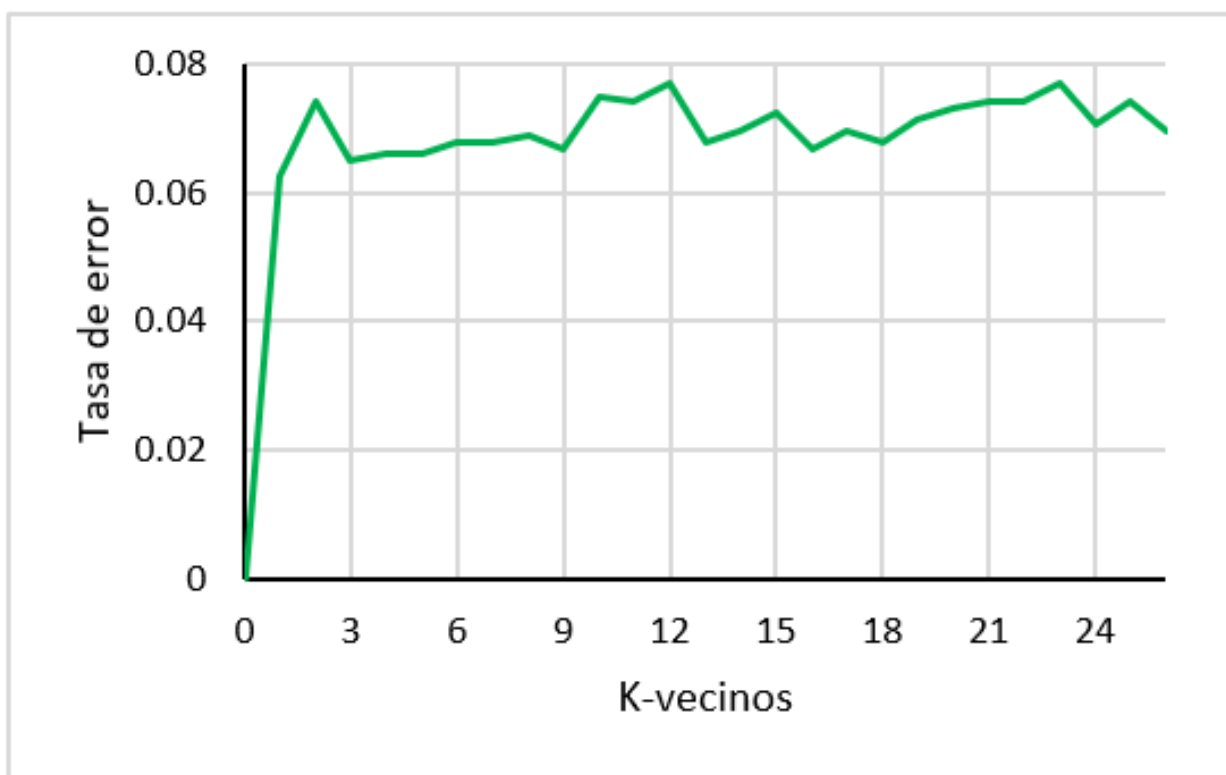


Figura 3. Reporte de tasa de error

Como prueba inicial, el modelo fue entrenado con un valor de K vecino igual a 2. Los resultados obtenidos para este número de vecinos se presentan en la Tabla 2, la cual resume las principales métricas de desempeño del clasificador.

Tabla 2. Reporte de clasificación

	Precisión	Sensibilidad	Puntaje -F1	Número de casos
Si tiene	0.92	0.93	0.93	569
No tiene	0.93	0.92	0.92	537
Exactitud			0.93	1106

Posteriormente, se evaluó el modelo con K igual a 1, obteniendo los valores obtenidos que se detallan en la Tabla 3, a partir de los cuales se identificó el valor óptimo del parámetro K vecino.

Tabla 3. Reporte de clasificación óptimo

	Precisión	Sensibilidad	Puntaje -F1	Número de casos
Si tiene	0.97	0.91	0.94	569
No tiene	0.91	0.97	0.94	537
Exactitud			0.94	1106

Despliegue del Sistema

El patrón arquitectónico Modelo–Vista–Template (MTV) pertenece al entorno de desarrollo Django. El modelo predictivo, entrenado previamente con el algoritmo KNN, fue serializado utilizando la biblioteca Joblib y almacenado en el archivo modelo_entrenado.pkl. Este archivo contiene una versión persistente del modelo que puede ser reutilizada en producción sin necesidad de reentrenar, lo cual permite optimizar el rendimiento y facilita la implementación del modelo en entornos web.

El formulario web (Template prediccion.html) permite al personal médico ingresar variables clínicas y serológicas del paciente. Una vez completado, los datos son enviados mediante una solicitud POST a una vista (views.py), donde se transforman al formato requerido por el modelo. En ese punto, el archivo modelo_entrenado.pkl se carga dinámicamente en memoria, se ejecuta la predicción con el modelo KNN y se genera un diagnóstico (positivo o negativo para enfermedad celíaca), acompañado de la probabilidad asociada.

El resultado es renderizado en una nueva plantilla HTML (resultado.html), con opciones adicionales como la exportación del informe en PDF y Excel. Esta estructura desacopla el entrenamiento del modelo de su ejecución en tiempo real, optimizando el rendimiento, reduciendo la latencia de respuesta y facilitando su integración en contextos clínicos digitales.

La lógica de predicción está encapsulada en una API interna desarrollada en Django, que gestiona las solicitudes HTTP, estructura las respuestas en formato JSON, y permite la interacción fluida con el frontend. El sistema incluye un esquema de autenticación basado en roles: el perfil “Admin” (dirigido al personal administrativo) tiene acceso a funciones como el registro de pacientes, acceso al formulario, la gestión de usuarios y el historial de predicciones; mientras que el perfil “Usuario” (dirigido a médicos autorizados) puede registrar pacientes y acceder al formulario mostrado en la Figura 4 para realizar las predicciones clínicas correspondientes.

Formulario de Predicción de Enfermedad Celíaca

Edad:

Género

☐ Masculino ☐ Femenino

¿Tiene diabetes?

☐ No ☐ Sí

¿Presenta antecedente de diabetes?

☒ Ninguno ☐ Tipo 1 ☐ Tipo 2

Presentación clínica de diarrea:

☐ Esteatorrea ☐ Acuosa ☐ Inflamatoria

¿Pérdida de peso?

☐ No ☐ Sí

¿Presenta heces pastosas?

☐ No ☐ Sí

Nivel de Inmunoglobulina tipo A (g/L):	Nivel de Inmunoglobulina tipo G (g/L):	Nivel de Inmunoglobulina tipo M (g/L):

Figura 4. Interfaz de formulario

Resultados y discusión

Se construyó la matriz correspondiente, cuyos resultados se encuentran en la Tabla 4. Esta permite evaluar la eficacia del modelo en la clasificación de pacientes con diagnóstico positivo y negativo de enfermedad celiaca, a partir del análisis de variables clínicas y serológicas. En dicha matriz se observa que el modelo identificó a 517 pacientes con enfermedad celíaca (Verdaderos positivos) y a 520 pacientes sanos (Verdaderos negativos). Sin embargo, clasificó erróneamente a 17 personas sanas como si tuvieran la enfermedad (falsos positivos) y no logró detectar la enfermedad en 52 casos, clasificándolos como sanos (falsos negativos).

Tabla 4. Matriz de confusión P

K_neighbors=1		Predicción	
		Si tiene	No tiene
Valores reales	Si tiene	VP=517	FN=52
	No tiene	FP=17	VN=520

VP: Verdaderos Positivos, FP: Falsos Positivos, FN: Falsos negativos, VN: Verdaderos Negativos.

En la primera fase del entrenamiento del modelo KNN, el valor k fue igual a 2. Este logró una exactitud superior al 93 %, con unas métricas balanceadas por ambas clases, con una precisión y sensibilidad del 92 %-93 %. Sin embargo, cuando se comenzó a optimizar el número de vecinos y se empezó a incrementar el número k hasta valores superiores, se detectó que el valor k=1 reportaba una mayor exactitud del 94 %. En este caso óptimo, la clase “sí tiene” logró una precisión del 97 % y una sensibilidad del 91 %. En la clase “no tiene” obtuvo una precisión del 91 % y una sensibilidad del 97 %. Como se observa en la Figura 5, las métricas correspondientes a los valores mencionados.

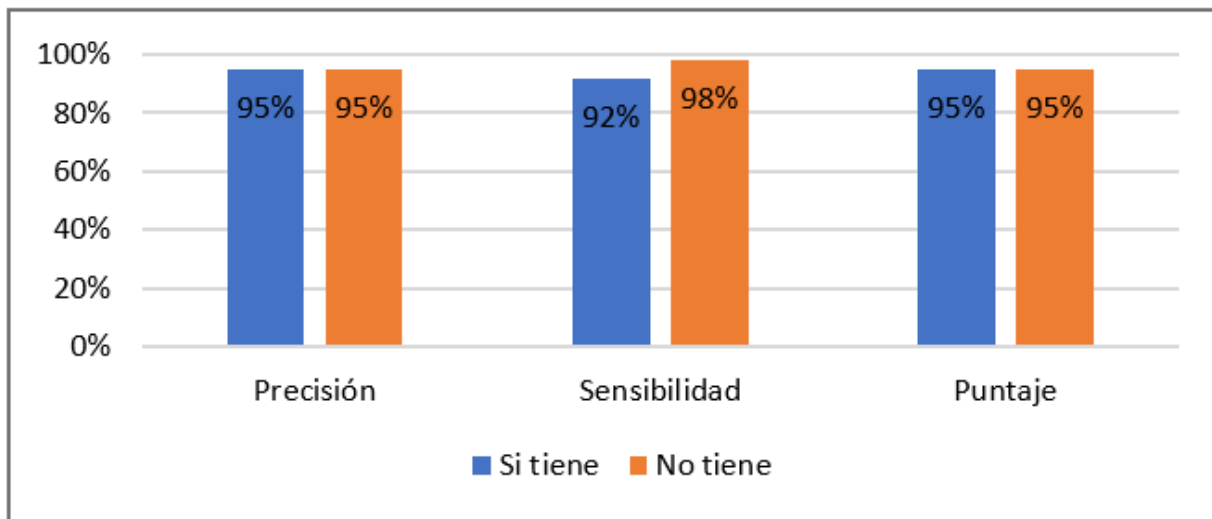


Figura 5. Métricas del modelo KNN por clase

En el desarrollo con Django, se elaboró una interfaz de programación de aplicaciones (API) destinada a recibir datos clínicos ingresados por el usuario, procesarlos y generar un diagnóstico codificado en formato. La aplicación en acceso se encuentra regida por un sistema de autenticación y administración de roles. El usuario con un personal médico autorizado puede acceder a la generación de informes, cuyas credenciales se encuentran generadas y administradas exclusivamente por el administrador. De esa forma el formulario de la interfaz de predicción fue usado por el personal médico para introducir los datos clínicos y serológicos del paciente, que fueron enviados al servidor a través de una vista en Django, que procesó la información, cargó el modelo (modelo_entrenado.pkl) y generó la predicción de enfermedad celíaca en tiempo real, mostrando el resultado en pantalla. Se hicieron pruebas con el conjunto de datos para determinar el nivel de confianza del modelo, los resultados se presentan en la Tabla 5.

Tabla 5. Predicción de enfermedad celíaca con datos clínicos

Atributo	Prueba 1	Prueba 2	Prueba 3
Edad	8	20	15
Género	Masculino	Femenino	Masculino
Diabetes	No	Si	Si
Tipo de Diabetes	Ninguno	2	1
Diarrea	Esteatorrea	Inflamatoria	Acuosa

Atributo	Prueba 1	Prueba 2	Prueba 3
Pérdida de peso	No	Si	No
Heces pastosas	No	Si	No
IgA	3	2.20	2.40
IgG	10	9	9.50
IgM	1.30	1.30	1.50
Predicción	No tiene EC.	Si tiene EC.	Si tiene EC.

Un estudio previo con una muestra de 50 personas reportó una exactitud del 96 % al aplicar el algoritmo KNN para predecir enfermedad celíaca a partir de variables clínicas [16]. El modelo desarrollado en este estudio obtuvo una exactitud general del 94 % lo que evidencia su solidez en el proceso de clasificación y resalta que la incorporación de variables serológicas contribuye significativamente a mejorar la eficacia. La eficacia del algoritmo KNN en distintos contextos ha sido reportada en diversos estudios. Se ha utilizado con éxito sobre datos de expresión génica; sin embargo, estas aproximaciones no son comparables directamente con la nuestra, dado que el enfoque de este trabajo se centra en variables clínicas y serológicas [28].

Una limitación relevante del presente estudio es el empleo de una base de datos pública proveniente de la plataforma Kaggle, la cual no proporciona detalles suficientes sobre las condiciones y criterios de recolección aplicados. Esto reduce la posibilidad de extrapolar los resultados a escenarios clínicos. Sin embargo, este estudio realizó un proceso de limpieza de datos, eliminando atributos sin definición clara o sin relevancia clínica. Además, se corrigió el desbalance entre clases mediante técnicas de sobremuestreo, lo que permitió mejorar la capacidad predictiva del modelo. Los resultados respaldan el uso del algoritmo de vecinos más cercanos como apoyo en el diagnóstico inicial de la enfermedad celíaca, al basarse en criterios clínicos y serológicos. Esta solución automatizada se alinea con la práctica actual. Además, tiene el potencial de reducir el infradiagnóstico, la morbilidad y la mortalidad asociadas.

Conclusiones

El presente estudio confirma que el uso del algoritmo KNN constituye una herramienta eficaz para predecir la enfermedad celíaca, utilizando parámetros clínicos y serológicos, respaldado por un alto nivel de exactitud en los resultados obtenidos. Esta aproximación ayuda a optimizar el diagnóstico temprano, especialmente en casos que podrían pasar desapercibidos bajo criterios clínicos convencionales.

Se desarrolló una interfaz web basada en Django que permitirá al personal médico ingresar datos clínicos y serológicos del paciente, generando de forma automática un informe de predicción, el cual puede ser enviado

al correo electrónico registrado. Esta implementación refuerza la aplicabilidad del modelo en contextos reales, favoreciendo su integración en entornos clínicos digitales. Como proyección futura, se destaca la necesidad de validar externamente el modelo mediante conjuntos de datos clínicamente verificados y estandarizados. Además, se sugiere ampliar su uso en cohortes poblacionales y explorar su aplicabilidad frente a otras enfermedades con perfiles clínicos similares, consolidando así su utilidad como herramienta diagnóstica complementaria en la práctica médica.

Contribución de Autoría

Daniel Lévano Rodriguez: [Conceptualización](#), [Investigación](#), [Metodología](#), [Redacción - borrador original](#). Flor Elizabeth Cerdán León: [Supervisión](#), [Administración de proyectos](#). Cesar Rolando Salazar Giraldo: [Software](#), [Validación](#). Jadira Dina Vasquez Castro: [Curación de datos](#). Marita Abigail Carbajal Bazán: [Análisis formal](#). Aldana Camila Zea Mendoza: [Visualización](#), [Escritura](#), [revisión y edición](#).

Referencias

- [1] S. Jabeen *et al.*, “Disease specific symptoms indices in patients with celiac disease—a hardly recognised entity,” *Front Nutr*, vol. 9, Sep 2022.
- [2] B. Lebwohl and A. Rubio-Tapia, “Epidemiology, presentation, and diagnosis of celiac disease,” *Gastroenterology*, vol. 160, no. 1, pp. 63–75, Jan 2021.
- [3] A. F. Syam *et al.*, “Prevalence and factors associated with celiac disease in high-risk patients with functional gastrointestinal disorders,” *PLoS One*, vol. 19, no. 6, p. e0297605, Jun 2024.
- [4] Ö. Aydemir *et al.*, “Polymorphisms in intron 1 of hla-dra differentially associate with type 1 diabetes and celiac disease,” Jun 2023.
- [5] E. Crehuá-Gaudiza *et al.*, “Diagnóstico de enfermedad celiaca en la práctica clínica: presente y futuro,” *An Pediatr (Engl Ed)*, vol. 94, no. 4, pp. 223–229, Apr 2021.
- [6] T. B. Houmich and B. Admou, “Celiac disease: Understandings in diagnostic, nutritional, and medicinal aspects,” *Int J Immunopathol Pharmacol*, vol. 35, Jan 2021.
- [7] P. Singh *et al.*, “Who to screen and how to screen for celiac disease,” *World J Gastroenterol*, vol. 28, no. 32, pp. 4493–4507, Aug 2022.
- [8] A. H. Abend *et al.*, “Estimation of prevalence of autoimmune diseases in the united states,” *J Clin Invest*, vol. 135, no. 4, Dec 2024.
- [9] D. Andari *et al.*, “Clinical presentations and outcomes of celiac disease in children and adolescents,” *Front Pediatr*, vol. 13, Jan 2025.
- [10] C. M. Trovato *et al.*, “Clinical presentations of celiac disease,” *Nutrients*, vol. 17, no. 1, p. 129, Dec 2024.
- [11] A. A. Esmail *et al.*, “Celiac disease among outpatient attendees with gastrointestinal complaints,” *USTJMS*, vol. 3, Feb 2025.
- [12] H. Wieser *et al.*, “Gastrointestinal and hepatobiliary manifestations associated with untreated celiac disease,” *J Clin Med*, vol. 13, no. 15, p. 4579, Aug 2024.
- [13] M. Alfawaz *et al.*, “Clinical characteristics of celiac disease patients in qassim region,” *J Family Med Prim Care*, vol. 13, no. 3, pp. 827–832, Mar 2024.
- [14] G. Ortiz *et al.*, “New cut-off values of antitransglutaminase antibodies processed by chemiluminescence,” *JPGN Rep*, vol. 6, no. 2, pp. 107–112, May 2025.

- [15] N. Conrad *et al.*, “Incidence, prevalence, and co-occurrence of autoimmune disorders over time,” *The Lancet*, vol. 401, no. 10391, pp. 1878–1890, Jun 2023.
- [16] S. K. Dooraki, “Early detection of celiac disease through machine learning algorithms,” *JCMCR*, vol. 5, no. 3, Mar 2024.
- [17] A. F. A. H. Alnuaimi and T. H. K. Albaldawi, “An overview of machine learning classification techniques,” *BIO Web Conf*, vol. 97, Apr 2024.
- [18] F. Piccialli *et al.*, “Precision medicine and machine learning towards the prediction of potential celiac disease,” *Sci Rep*, vol. 11, p. 5683, Mar 2021.
- [19] M. Torres-Vásquez *et al.*, “Impact of oversampling algorithms in the classification of guillain-barré syndrome,” *Ingenius*, no. 25, pp. 20–31, 2021.
- [20] A. R. Vico *et al.*, “Predictors of the post-stroke status in the discharge from the hospital,” *Enfermería Global*, vol. 22, no. 1, Jan 2023.
- [21] C. L. Vidal-Silva *et al.*, “Experiencia académica en desarrollo rápido con python y django,” *Formación universitaria*, vol. 14, no. 5, Oct 2021.
- [22] C. E. Acosta and E. R. Vizcaino, “Implementación de un sistema para evaluar la cobertura de la red sigfox,” *Enfoque UTE*, Nov 2022.
- [23] M. L. Barbero *et al.*, “Analyzing the critical success factors in the online-game development,” *Dyna (Medellin)*, vol. 98, no. 5, Sep 2023.
- [24] D. N. and N. P. K. S., “Design and development of we-cdss using django framework,” *IEEE Access*, vol. 10, pp. 119 575–119 592, 2022.
- [25] C. M. C. Estrada *et al.*, “Diseño de un sistema web para el control de curriculum vitae electrónico,” *RITI*, vol. 10, no. 20, Dec 2022.
- [26] A. Ehsan *et al.*, “Restful api testing methodologies,” *Applied Sciences*, vol. 12, no. 9, Apr 2022.
- [27] C. Molina and V. Bonilla, “Aplicación de la metodología crisp-dm en el análisis de gases disueltos,” *Revista Técnica energía*, vol. 21, no. 1, Jul 2024.
- [28] J. Carreras, “Artificial intelligence analysis of celiac disease using transcriptomic panel,” *Healthcare*, vol. 10, no. 8, Aug 2022.