

UNIVERSIDAD SAN PEDRO

FACULTAD DE INGENIERIA

SECCIÓN DE POSGRADO DE LA FACULTAD DE INGENIERIA



**Modelo de aprendizaje automático para clasificación del
desempeño docente en una universidad, Chimbote, 2024**

Tesis para obtener el grado de Maestra en Ingeniería Informática y
de Sistemas con mención en Gestión de Tecnologías de
Información y Comunicaciones

Autora:

Aquino Gonzales, Aracelli

Asesor:

Villarreal Torres, Henry

ORCID: 0000-0002-5989-4534

CHIMBOTE – PERÚ

2024

Índice general

Índice general.....	i
Índice de tablas	ii
Palabras claves.....	iii
Constancia de originalidad.....	iv
Título.....	v
Resumen.....	vi
Abstract.....	vii
Introducción	1
Metodología.....	21
Resultados.....	24
Análisis y Discusión	38
Conclusiones.....	47
Recomendaciones	48
Referencias bibliográficas.....	49
Anexo y apéndices	58

Índice de tablas

Tabla 1 Métricas de modelos de clasificación mediante validación cruzada	24
Tabla 2 Métricas de modelos de clasificación mediante validación cruzada	27
Tabla 3 Métricas de modelos de clasificación por muestreo aleatorio	28
Tabla 4 Métricas de modelos de clasificación mediante validación cruzada	30
Tabla 5 Métricas de modelos de clasificación por muestreo aleatorio	31
Tabla 6 Métricas de modelos de clasificación mediante validación cruzada	33
Tabla 7 Métricas de modelos de clasificación por muestreo aleatorio	34
Tabla 8 Comparación de los modelos de clasificación con mejores métricas de rendimiento	35

Palabras claves

Tema	:	Aprendizaje automático
------	---	------------------------

Especialidad	:	Ingeniería Informática y de Sistemas
--------------	---	--------------------------------------

Línea de Investigación:

Línea de Investigación	:	Sistemas de la Información
------------------------	---	----------------------------

Área	:	Ingeniería y tecnología
------	---	-------------------------

Subárea	:	Ingeniería eléctrica, electrónica e informática
---------	---	---

Disciplina	:	Ingeniería de sistemas y comunicaciones
------------	---	---

Constancia de originalidad



VICERRECTORADO DE INVESTIGACIÓN

CONSTANCIA DE ORIGINALIDAD

El que suscribe, Vicerrector de Investigación de la Universidad San Pedro:

HACE CONSTAR

Que, de la revisión del trabajo titulado "**Modelo de aprendizaje automático para clasificación del desempeño docente en una universidad, Chimbote, 2024**" del (a) estudiante: **AQUINO GONZALES CORINA ARACELLI**, identificado(a) con Código N° **0200123084**, se ha verificado un porcentaje de similitud del **15%**, el cual se encuentra dentro del parámetro establecido por la Universidad San Pedro mediante resolución de Consejo Universitario N° **5037-2019-USP/CU** para la obtención de grados y títulos académicos de pre y posgrado, así como proyectos de investigación anual Docente.

Se expide la presente constancia para los fines pertinentes.

Chimbote, 05 de marzo de 2025

UNIVERSIDAD SAN PEDRO
VICERRECTORADO DE INVESTIGACIÓN

Dr. JAVIER MARTÍNEZ CARRIÓN
VICERRECTOR



NOTA: Este documento carece de valor si no tiene adjunta el reporte del Software TURNITIN.

Título

Modelo de aprendizaje automático para clasificación del desempeño docente en una
universidad, Chimbote, 2024

Machine learning model for classifying teaching performance at a university,
Chimbote, 2024

Resumen

El proyecto tuvo como objetivo desarrollar un modelo innovador basado en aprendizaje automático para clasificar el desempeño docente. Este modelo está diseñado para optimizar los procesos evaluativos mediante técnicas avanzadas de minería de datos. La investigación es de tipo aplicada, orientada a la resolución de problemas prácticos mediante el uso de herramientas tecnológicas avanzadas. Con un enfoque predictivo, se busca identificar patrones y tendencias mediante el análisis de datos. El diseño no experimental transversal permite observar y analizar las variables sin manipulación directa. La población consta de 800 registros de desempeño docente recopilados durante el año 2024. Debido a la disponibilidad completa de los datos, no se seleccionará muestra, la recolección de datos se realizará mediante la técnica de análisis documental, utilizando una ficha estructurada que será previamente validada por un panel de expertos y sometida a pruebas piloto para medir su confiabilidad y consistencia. Concluyendo que se determinó que la realización de actividades de investigación emergió como el factor más influyente, evidenciado por el valor más alto de Information Gain (0.067) y un destacado puntaje ReliefF (0.146). La adscripción a la Facultad de Ciencias de la Salud se posicionó como el segundo factor más relevante, con un Information Gain de 0.040 y un valor chi-cuadrado de 561.130, sugiriendo que las exigencias específicas de las carreras de salud podrían generar un entorno más propicio para la excelencia docente.

Abstract

The project aimed to develop an innovative model based on machine learning to classify teaching performance. This model is designed to optimize evaluation processes through advanced data mining techniques. The research is of an applied type, oriented to the resolution of practical problems through the use of advanced technological tools. With a predictive approach, it seeks to identify patterns and trends through data analysis. The non-experimental cross-sectional design allows observing and analyzing the variables without direct manipulation. The population consists of 800 teaching performance records collected during the year 2024. Due to the complete availability of the data, no sample will be selected. Data collection will be carried out using the documentary analysis technique, using a structured form that will be previously validated by a panel of experts and subjected to pilot tests to measure its reliability and consistency. Concluding that it was determined that carrying out research activities emerged as the most influential factor, evidenced by the highest Information Gain value (0.067) and a prominent ReliefF score (0.146). Affiliation with the Faculty of Health Sciences was ranked as the second most relevant factor, with an Information Gain of 0.040 and a chi-square value of 561.130, suggesting that the specific demands of health careers could generate a more conducive environment for teaching excellence.

Introducción

La investigación analizó estudios desarrollados que formaron parte de los antecedentes, seleccionando los siguientes:

En primer lugar, Taiwo et al. (2024) desarrollaron un estudio con el objetivo de diseñar un modelo predictivo para evaluar el éxito académico de estudiantes universitarios, enfocado en identificar estudiantes en riesgo. La metodología empleada fue de tipo cuantitativa, utilizando un diseño experimental y una muestra de 120 estudiantes seleccionados de manera intencional. Para el análisis, se implementaron algoritmos de redes neuronales que generaron un área bajo la curva (AUC) de 0.89, indicando un alto nivel de precisión en las predicciones. En conclusión, el estudio demostró que el modelo predictivo basado en redes neuronales es una herramienta efectiva para identificar a estudiantes en riesgo, contribuyendo al diseño de estrategias educativas preventivas y personalizadas.

Asimismo, Amer (2024) desarrolló un estudio con el objetivo de predecir el rendimiento académico en estudiantes de ingeniería mediante el uso de tres algoritmos de aprendizaje supervisado: regresión lineal, árboles de decisión y redes neuronales. La metodología empleada fue de tipo cuantitativa, utilizando una muestra de 200 estudiantes seleccionados de forma estratificada. Los resultados indicaron que el algoritmo de regresión lineal logró un RMSE de 0.35, destacándose como el más preciso entre los tres modelos evaluados. Las conclusiones del estudio subrayaron que la regresión lineal es una herramienta confiable para predecir el rendimiento académico, sugiriendo su aplicación en sistemas educativos para la identificación temprana de estudiantes en riesgo.

Igualmente, Pacol (2024) realizó un estudio con el objetivo de analizar la opinión de los estudiantes sobre el desempeño docente mediante técnicas de análisis de sentimientos. La metodología utilizada fue de tipo descriptivo, empleando un enfoque cuantitativo y una muestra de 300 estudiantes seleccionados de manera

aleatoria. El análisis se llevó a cabo utilizando un modelo de conjunto que alcanzó una precisión del 87% en la clasificación de los sentimientos expresados en las encuestas. Los resultados revelaron patrones claros de retroalimentación positiva y áreas de mejora en el desempeño docente. En conclusión, el estudio destacó que el análisis de sentimientos es una herramienta efectiva para recopilar y procesar retroalimentación estudiantil, contribuyendo a mejorar la calidad de la enseñanza.

Además, Al-Alawi et al. (2023) llevaron a cabo un estudio con el objetivo de predecir factores que afectan el rendimiento académico en estudiantes universitarios mediante el uso de algoritmos de minería de datos. La metodología empleada fue de tipo correlacional, utilizando metodologías de Knowledge Discovery en una muestra representativa de 500 estudiantes de diversas disciplinas académicas. Los resultados destacaron que el algoritmo de regresión logística alcanzó un 89% de precisión en las predicciones, lo que demuestra su efectividad en la identificación de patrones relevantes. La investigación concluyó que la implementación de algoritmos avanzados no solo mejora la comprensión de los factores clave en el éxito académico, sino que también proporciona herramientas prácticas para el diseño de intervenciones educativas más efectivas.

Por otro lado, Almufarreh et al. (2023) desarrollaron un marco de enseñanza y evaluación en educación superior mediante aprendizaje automático, con el objetivo de optimizar los procesos educativos y mejorar la calidad de las evaluaciones. La metodología utilizada fue descriptiva, basada en técnicas cualitativas, aunque no se especificó información detallada sobre la población y la muestra. Los resultados obtenidos destacaron que el marco propuesto logró una precisión del 92% en la clasificación de datos educacionales, lo que valida su efectividad en entornos simulados. Las conclusiones resaltan que la aplicación de este marco contribuye significativamente a la mejora de la evaluación educativa, recomendándose su implementación en contextos reales para validar su aplicabilidad.

Cabe destacar que Baffa et al. (2023) llevaron a cabo un estudio con el objetivo de predecir la empleabilidad de egresados utilizando aprendizaje automático, con énfasis en identificar factores clave que influyen en la inserción laboral. La

metodología utilizada fue cuantitativa, con un diseño predictivo basado en una muestra de 150 graduados seleccionados de diversas disciplinas académicas. El modelo Random Forest aplicado alcanzó un índice de precisión del 91%, destacándose como una herramienta efectiva para predecir empleabilidad. Las conclusiones subrayaron que el aprendizaje automático permite optimizar los procesos de inserción laboral en mercados competitivos, ofreciendo una base sólida para diseñar estrategias de desarrollo profesional.

Además, Ogundele et al. (2024) realizaron un estudio con el objetivo de evaluar el desempeño académico en estudiantes universitarios mediante el uso de regresión lineal. La metodología adoptada fue cuantitativa, con un diseño predictivo y una muestra de 100 estudiantes seleccionados de manera aleatoria. Los resultados del análisis arrojaron un error cuadrático medio (RMSE) de 0.28, lo que indica un alto grado de precisión en las predicciones realizadas por el modelo. En conclusión, el estudio demostró que la regresión lineal es un enfoque confiable y efectivo para predecir resultados académicos, recomendándose su uso en entornos educativos para identificar a estudiantes en riesgo de bajo rendimiento.

De igual forma, Tarhini et al. (2023) llevaron a cabo un estudio con el objetivo de identificar los factores determinantes en el aprendizaje efectivo mediante análisis factorial exploratorio. La metodología utilizada fue de enfoque cuantitativo, con un diseño correlacional y una muestra de 400 estudiantes seleccionados aleatoriamente. Los resultados mostraron una correlación positiva significativa ($r = 0.76$, $p < 0.05$) entre el uso de tecnología educativa y el rendimiento académico, indicando que la integración tecnológica favorece el aprendizaje. En conclusión, el estudio destacó la importancia de incorporar tecnología en los entornos educativos como una estrategia clave para mejorar el rendimiento académico y la experiencia de aprendizaje de los estudiantes.

Igualmente, Noaman et al. (2023) llevaron a cabo un estudio con el objetivo de implementar un sistema basado en inteligencia artificial (IA) para optimizar la gestión de cursos en universidades. La metodología utilizada fue de tipo cuantitativa, con un diseño experimental aplicado a una muestra de 50 profesores y 300 estudiantes

seleccionados de manera intencional. Los resultados del estudio indicaron que el sistema logró un 93% de satisfacción del usuario, evidenciando su efectividad en la gestión de recursos académicos. En conclusión, el sistema basado en IA se destacó como una herramienta eficiente para optimizar la organización y administración de cursos, con un impacto positivo en la experiencia de los usuarios involucrados.

Finalmente, Zhong (2023) investigó la calidad de enseñanza mediante el uso del algoritmo GA-SVM en instituciones de educación superior, con el objetivo de optimizar los modelos de enseñanza y mejorar los procesos educativos. La metodología utilizada fue de enfoque cuantitativo, con un diseño experimental aplicado en un entorno simulado. Aunque la investigación no especificó detalles sobre la población y la muestra, los resultados obtenidos fueron significativos, logrando un rendimiento del 95% en precisión al evaluar la efectividad del algoritmo en escenarios simulados. Además, los hallazgos destacaron el potencial del GA-SVM para identificar puntos críticos en los procesos de enseñanza, lo que permite ajustar los modelos en tiempo real. Finalmente, se concluyó que esta herramienta es una alternativa prometedora para mejorar la calidad educativa, subrayando la importancia de su validación en contextos educativos reales y diversos.

En el contexto nacional, se tiene la investigación de Polo (2024) desarrolló una investigación con el objetivo de diseñar un modelo predictivo que supere en confiabilidad a los algoritmos tradicionales bayesianos, de regresión, soporte vectorial y bosques aleatorios, para estimar de manera más eficiente la deserción estudiantil en centros de educación superior tecnológicos públicos de la región La Libertad. La metodología empleada fue de tipo aplicada con un diseño descriptivo, utilizando la metodología CRISP-DM y algoritmos de aprendizaje supervisado de la librería Scikit Learn de Python. La población estuvo conformada por 500 registros de estudiantes del programa de Computación e Informática del Instituto Superior Tecnológico Trujillo, recopilados entre 2012 y 2021. Se entrenaron cuatro algoritmos en un conjunto de datos correspondiente al 80% del total, reservando el 20% restante para la prueba del modelo final. Los resultados mostraron que el modelo ensamblado, obtenido mediante un algoritmo de votación, alcanzó una confiabilidad del 93%, con una tasa de error del

7%. Se concluyó que el modelo ensamblado propuesto garantiza una alta confiabilidad durante todo su ciclo de vida.

Para el caso de Quintana (2023) llevó a cabo un estudio con el propósito de analizar y validar el modelo de machine learning más adecuado para predecir la deserción estudiantil en la Universidad Nacional de Moquegua entre 2009 y 2019, además de identificar las características influyentes en este fenómeno. La investigación fue de tipo descriptivo correlacional con un diseño observacional, utilizando la metodología CRISP-DM y métricas como la matriz de confusión y validación cruzada K-Folds. La población incluyó a 329 estudiantes, de los cuales 109 eran desertores y 220 no desertores, seleccionados mediante muestreo probabilístico aleatorio estratificado. Se emplearon algoritmos como Regresión Logística, Árboles de Decisión, Máquinas de Vectores de Soporte y Naive Bayes, implementados en Python. Los resultados indicaron que el Árbol de Decisión logró un 76% de precisión, seguido de la Regresión Logística con 73%, Máquinas de Vectores de Soporte con 71% y Naive Bayes con 62%. Se concluyó que las características influyentes en la deserción incluyen datos generales, económicos y de vivienda.

Asimismo, Naranjo (2023) realizó una investigación con el objetivo de establecer la relación entre el desempeño docente y la calidad educativa en una universidad pública de Ecuador. La metodología fue de tipo correlacional descriptiva exploratoria, aplicando cuestionarios a una muestra representativa de 60 estudiantes. Los instrumentos utilizados presentaron altos niveles de confiabilidad, con coeficientes de 0.927 para desempeño docente y 0.877 para calidad educativa. La prueba de Kolmogorov arrojó un valor p de 0.000, indicando una correlación significativa. El coeficiente de Spearman fue de 0.815, evidenciando una fuerte relación positiva entre las variables. Se concluyó que mejoras en el desempeño docente impactan notablemente en la calidad educativa.

Reyes (2023) llevó a cabo un estudio con el fin de determinar las propiedades psicométricas de un cuestionario destinado a evaluar la competencia docente en estudiantes de una escuela técnico profesional policial. La muestra estuvo compuesta por 319 estudiantes. Para la validez de contenido, se utilizó el juicio de cinco expertos,

aplicando la V de Aiken, obteniendo puntuaciones superiores a 0.80. El análisis factorial exploratorio y confirmatorio arrojó índices de RMSEA = 0.061, SRMR = 0.052, CFI = 0.997 y TLI = 0.997, considerados aceptables. La confiabilidad, medida mediante el coeficiente Omega de McDonald, fue de 0.91 para el factor "Planeación y gestión de los procesos de enseñanza-aprendizaje" y de 0.96 para "Conducción y valoración de los procesos de enseñanza-aprendizaje". Se concluyó que el cuestionario posee validez y confiabilidad adecuadas.

También se tiene a Delgado (2022) quien desarrolló una investigación con el propósito de diseñar un modelo de gestión académica que permita disminuir el riesgo de deserción estudiantil en una universidad estatal del cantón Jipijapa, Ecuador. La metodología fue de tipo básica y proyectiva, estructurada en etapas descriptiva, analítica y predictivo-prospectiva. Se aplicó un análisis de regresión logística ordinal para validar el modelo teórico propuesto. La población estuvo conformada por 268 estudiantes, quienes respondieron a dos cuestionarios. Los hallazgos indicaron que el 94.8% de los estudiantes percibían la gestión académica en niveles bajo-medio, y el 92.1% presentaban un riesgo de deserción en niveles bajo-medio. El modelo funcional teórico obtenido fue: Riesgo de deserción 2 moderado = $-3.688 - 6.009$ estrategias metodológicas = 2 mediano, evidenciando una influencia del 84.9% de la gestión académica sobre el riesgo de deserción. Se concluyó que la implementación de un plan de acción con actividades innovadoras es esencial para promover la permanencia estudiantil.

En tal sentido, Velásquez (2022) realizó un estudio con el objetivo de determinar la asociación entre la evaluación del desempeño docente y la satisfacción académica en estudiantes de una universidad privada de Lima. La investigación fue de enfoque cuantitativo, con diseño no experimental y alcance correlacional. La población y muestra fueron censales, conformadas por 61 docentes y 1,703 estudiantes. Se emplearon dos instrumentos: una ficha de evaluación del desempeño docente en aula, con una confiabilidad de 0.76, y una encuesta de satisfacción académica del estudiante, con una confiabilidad de 0.97. Los resultados mostraron una asociación significativa, aunque débil, entre la evaluación del desempeño docente y la

satisfacción académica del estudiante ($\text{Chi-Cuadrado} = 10.34, p < 0.05$), indicando que a medida que aumenta la calificación del desempeño docente, también lo hace el nivel de satisfacción académica. Se concluyó que, aunque la correlación es positiva, su magnitud es débil, posiblemente debido a factores como la edad, madurez y ciclo de estudio de los estudiantes.

Respecto a la fundamentación científica de la investigación, se seleccionaron los siguientes referentes teóricos:

El aprendizaje automático (AA) ha emergido como una disciplina clave dentro de la inteligencia artificial (IA), destacándose por su capacidad de desarrollar modelos que permiten a las máquinas aprender y mejorar su rendimiento en tareas específicas sin requerir programación explícita para cada caso. Este campo se sustenta en algoritmos avanzados que identifican patrones en grandes volúmenes de datos, generando modelos predictivos y clasificatorios aplicables en ámbitos como la salud, la educación y la industria, con un impacto transformador en diversos sectores. El aprendizaje automático integra un conjunto amplio de teorías y aplicaciones que están revolucionando diversos sectores. Desde la creación de herramientas predictivas hasta la automatización de procesos, el AA ha demostrado ser esencial para tomar decisiones basadas en datos. Sin embargo, para maximizar su impacto y garantizar su equidad, es crucial abordar los desafíos relacionados con la calidad de los datos y los sesgos en los modelos, promoviendo así su implementación responsable y efectiva.

Una de las bases teóricas más relevantes en el AA son los algoritmos de clasificación, que asignan etiquetas a datos en función de sus características observadas. En el sector de la salud, por ejemplo, se ha demostrado la eficacia de algoritmos como K-vecinos más cercanos (KNN) y los árboles de decisión para detectar condiciones como la obesidad en adolescentes, logrando predicciones precisas basadas en datos complejos (López et al., 2023). En el ámbito financiero, los modelos de AA han sido utilizados para la detección de fraudes en transacciones, aprovechando patrones históricos para identificar actividades sospechosas con alta precisión (Dávila et al., 2023; Rodríguez-Tovar et al., 2023).

Los modelos de regresión, orientados a la predicción de valores continuos, son otra teoría fundamental en el AA. En el contexto industrial, se han empleado para proyectar la demanda de potencia eléctrica, utilizando algoritmos como Random Forest, que manejan eficazmente datos complejos y no lineales (Gallo et al., 2021). Estos modelos se basan en la premisa de que los datos históricos contienen información clave para identificar tendencias futuras, crucial para la planificación estratégica y la gestión de recursos.

El aprendizaje profundo, una subcategoría del AA, ha ganado relevancia gracias a su capacidad para procesar datos no estructurados como imágenes y texto. Las redes neuronales convolucionales (CNN) son un ejemplo destacado, con aplicaciones en visión por computadora y procesamiento de lenguaje natural. En el ámbito médico, estas redes han sido utilizadas para analizar imágenes diagnósticas, detectando características complejas que escapan al análisis humano convencional (Álvarez et al., 2020; Quiroz et al., 2021).

La teoría del aprendizaje supervisado y no supervisado subyace a muchas aplicaciones del AA. Mientras que el aprendizaje supervisado utiliza datos etiquetados para entrenar modelos con respuestas conocidas, el aprendizaje no supervisado busca identificar patrones en datos no etiquetados. Este último ha sido esencial en la segmentación de mercados, ayudando a las empresas a personalizar estrategias de marketing al identificar grupos de consumidores con características comunes (Pasquinelli et al., 2022; Polo-Triana et al., 2022).

Unos de los aspectos de mayor importancia son los datos de entrenamiento; es decir, el éxito de cualquier aplicación de AA depende de la calidad y cantidad de los datos de entrenamiento. En tareas como el análisis de sentimiento en redes sociales, un corpus de datos bien etiquetado es esencial para entrenar modelos que clasifiquen mensajes como positivos, negativos o neutros (Sánchez-Holgado et al., 2020). Además, la recolección y el preprocesamiento de datos son pasos críticos para garantizar resultados fiables y robustos. El sesgo en los modelos de AA representa un desafío significativo, con implicaciones para la precisión y equidad de los resultados. Los sesgos pueden originarse en los datos, los algoritmos o la interpretación de los

resultados, lo que hace crucial abordarlos, especialmente en áreas sensibles como la medicina y la justicia, donde la representatividad es esencial (Pasquinelli et al., 2022).

En la educación, el AA ha sido implementado para desarrollar modelos predictivos que evalúan el rendimiento académico y detectan estudiantes en riesgo de deserción, permitiendo a las instituciones adoptar estrategias de intervención temprana (Rico y Gaytán, 2022; Urteaga et al., 2020). En la industria, el AA ha optimizado procesos y mejorado la eficiencia. Por ejemplo, en el sector agroindustrial, se han utilizado algoritmos para prever la demanda y ajustar la producción, minimizando costos y desperdicios, y promoviendo la sostenibilidad (García-López et al., 2022).

La evaluación del desempeño docente universitario ha adquirido una importancia creciente en la educación superior, posicionándose como una herramienta estratégica para la mejora continua de la calidad educativa. Más allá de medir la efectividad de los docentes en su labor, este proceso tiene como objetivo proporcionar insumos valiosos para el desarrollo profesional y la consolidación de prácticas pedagógicas innovadoras. Diversas teorías y enfoques han enriquecido este campo, aportando perspectivas únicas sobre los criterios y metodologías más adecuados para su implementación.

Uno de los enfoques más destacados es el modelo de retroalimentación, que integra la autoevaluación y la evaluación por pares como elementos centrales. Este modelo promueve la participación activa de los docentes, no solo como sujetos evaluados, sino también como evaluadores críticos de sus colegas. La autoevaluación fomenta la reflexión sobre la práctica educativa, permitiendo identificar fortalezas y áreas de mejora, así como establecer metas de desarrollo profesional (Revilla-Mendoza & Palacios-Jiménez, 2020; Pacheco et al., 2018). La evaluación por pares, por su parte, favorece un entorno colaborativo, enriqueciendo las estrategias pedagógicas a través del intercambio de experiencias y buenas prácticas (Cipagauta-Moyano, 2019; Gómez y Valdés, 2019).

La pandemia de COVID-19 transformó profundamente las dinámicas educativas, forzando la adopción de modalidades de enseñanza a distancia y evidenciando la necesidad de adaptar los instrumentos de evaluación a estos nuevos

entornos. Durante este periodo, la investigación destacó la pertinencia de evaluar dimensiones específicas de la enseñanza virtual, como la capacidad de los docentes para emplear herramientas tecnológicas y mantener la interacción en entornos remotos. Sin embargo, también surgieron debates sobre la validez y fiabilidad de las herramientas tradicionales, como las encuestas de opinión estudiantil, para evaluar el desempeño docente en este contexto (De Los Heros y Solana-Villanueva, 2021; Vera, 2021). Este cambio ha impulsado una reevaluación de los métodos de evaluación, subrayando la importancia de enfoques más integrales y adaptativos (Chanta y Clavo, 2023; Zamora, 2021).

El clima organizacional y la motivación docente han demostrado ser factores determinantes en el desempeño y la calidad de la enseñanza. Un ambiente laboral positivo que fomente el bienestar y el reconocimiento profesional influye directamente en el compromiso y la dedicación de los docentes. La motivación, entendida como un motor interno que impulsa la excelencia en la práctica docente, es esencial para el desarrollo de experiencias de aprendizaje significativas (Rodríguez-Horna y Herrera-Vásquez, 2023; Sánchez-Huete et al., 2022). Por ello, la evaluación del desempeño debe ir acompañada de políticas institucionales que promuevan un entorno laboral favorable y motivador.

En la educación superior, el desempeño docente debe evaluarse considerando un conjunto integral de competencias específicas, que incluyen el dominio disciplinar, habilidades pedagógicas, capacidades comunicativas y la gestión efectiva del aula. Una evaluación centrada en estas competencias proporciona una visión más holística del impacto docente, permitiendo identificar áreas críticas para la formación continua y la mejora de la enseñanza (Achata y Quispe, 2018; Maussa y Hernández, 2018). Además, los criterios de evaluación deben estar alineados con los objetivos institucionales y las expectativas de los estudiantes, asegurando su relevancia y claridad (Pacheco et al., 2018; Gómez y Valdés, 2019).

La implementación de un sistema efectivo de evaluación del desempeño docente requiere un compromiso institucional sostenido y una cultura de evaluación centrada en la mejora continua. Esto implica no solo el diseño de instrumentos

pertinentes, sino también la formación de los docentes en su uso y en la interpretación constructiva de los resultados (Sánchez y Chávarry, 2022; Aranda y Llontop, 2023). La evaluación debe ser entendida como un proceso formativo, orientado a identificar oportunidades de desarrollo y a generar estrategias para optimizar las prácticas docentes (Arroyo, 2023). Además, un sistema de evaluación eficaz debe incluir mecanismos de seguimiento que aseguren la implementación de las recomendaciones derivadas del proceso evaluativo (Beltrán et al., 2021).

La investigación sobre el desempeño docente ha evolucionado hacia la implementación de modelos predictivos que permiten evaluar y mejorar la calidad educativa. En este contexto, el autorreporte de los estudiantes se ha identificado como una herramienta valiosa para obtener indicadores de desempeño didáctico, como se evidencia en el estudio de (Bazán y Velarde, 2021; Revilla-Mendoza y Palacios-Jimenez, 2020). Este enfoque no solo proporciona información sobre la efectividad de la enseñanza, sino que también permite a los docentes reflexionar sobre su práctica y realizar ajustes necesarios para mejorar el aprendizaje de los estudiantes.

Además, la relación entre el desempeño pedagógico y el aprendizaje de los estudiantes es fundamental. Soria et al. (2020) y Pacheco et al. (2018) destacan que un desempeño docente adecuado está directamente correlacionado con el rendimiento académico, lo que subraya la importancia de la preparación constante y el uso de técnicas pedagógicas efectivas. Por otro lado, Peceros (2014) como Cipagauta-Moyano (2019) enfatiza que el estilo motivacional autodeterminado de los docentes no solo impacta su desempeño, sino que también fomenta la autonomía de los estudiantes, creando un entorno de aprendizaje más dinámico y participativo.

Finalmente, la evaluación del desempeño docente debe ser un proceso integral que contemple la reflexión y la mejora continua. Gálvez y Milla (2018) como Gómez y Valdés (2019) proponen un modelo que considera tanto las competencias pedagógicas como los resultados de aprendizaje, mientras que Pacheco et al. (2018) y De Los Heros y Solana-Villanueva (2021) sugieren que una evaluación bien estructurada puede contribuir significativamente a la calidad educativa. En conjunto,

estos estudios resaltan la necesidad de adoptar enfoques innovadores y basados en datos para evaluar y mejorar el desempeño docente en el ámbito educativo.

La integración de la inteligencia artificial (IA) en la educación plantea tanto oportunidades como desafíos que demandan una reflexión crítica y una planificación estratégica. Este campo emergente tiene el potencial de transformar profundamente las prácticas pedagógicas, la personalización del aprendizaje y los procesos administrativos, pero su implementación efectiva requiere abordar problemas estructurales, éticos y de formación docente.

La implementación de la IA enfrenta varios obstáculos significativos. Guanga et al. (2024) destaca la falta de capacitación docente, la resistencia al cambio y la necesidad de infraestructura tecnológica adecuada como los principales retos para integrar esta tecnología en los entornos educativos. Asimismo, Guzmán et al. (2024) enfatiza la importancia de garantizar un uso ético de los datos y la formación continua de los docentes para abordar las complejidades que trae consigo esta tecnología. A pesar de estos desafíos, ambos autores coinciden en que una planificación equilibrada puede maximizar los beneficios de la IA, garantizando su integración efectiva y sostenible.

Entre las principales ventajas de la IA se encuentra su capacidad para personalizar el aprendizaje y mejorar la eficiencia en la enseñanza y la administración. Núñez-Michuy et al. (2023) argumenta que estas herramientas pueden transformar la interacción entre estudiantes y docentes, permitiendo una enseñanza adaptada a las necesidades individuales de los estudiantes. De manera complementaria, Guanga et al. (2024) resalta cómo la IA puede mejorar los procesos de evaluación, ofreciendo retroalimentación más precisa y oportuna. Sin embargo, el éxito de estas iniciativas depende de una infraestructura robusta y de una integración adecuada en los currículos existentes.

La innovación pedagógica es un aspecto clave en la preparación para el aprendizaje automático y la adopción de estrategias educativas modernas. Buendía et al. (2023) sugiere que herramientas visuales como las imágenes fractales pueden estimular la cognición en niños de 3 a 7 años, mejorando su atención y sentando las

bases para la comprensión de conceptos complejos. Este enfoque evidencia cómo las tecnologías innovadoras pueden enriquecer las prácticas pedagógicas, incluso desde la educación temprana.

Por otro lado, Salazar (2019) explora las lecciones aprendidas en la implementación de estrategias didácticas en la educación virtual. Su estudio resalta la importancia de la flexibilidad y la mentalidad abierta para enfrentar los desafíos de la enseñanza en línea, a la vez que señala las oportunidades para innovar en la interacción educativa y la metodología docente. La combinación de enfoques innovadores y tecnologías avanzadas como la IA podría amplificar aún más el alcance y la efectividad de estas estrategias.

La implementación de la IA en la educación es una tarea compleja que exige un enfoque equilibrado entre sus posibilidades y limitaciones. Guzmán et al. (2024) sugiere que los riesgos asociados, como el uso inadecuado de datos y las desigualdades tecnológicas, deben ser abordados con estrategias éticas y responsables. Además, Núñez-Michuy et al. (2023) y Buendía et al. (2023) enfatizan que la preparación pedagógica y la innovación son elementos fundamentales para aprovechar al máximo el potencial de la IA, especialmente en contextos educativos diversos y en evolución.

La investigación se justifica considerando que tuvo como propósito desarrollar un modelo de aprendizaje automático para la clasificación del desempeño docente en una universidad de Chimbote, con el objetivo de contribuir al análisis y mejora de la calidad educativa. Esta investigación se llevará a cabo debido a la necesidad de contar con herramientas tecnológicas avanzadas que permitan evaluar el rendimiento docente de manera precisa y objetiva, beneficiando a la institución y a los estudiantes a través de procesos de evaluación más eficientes.

Teóricamente, este estudio se justificará al aportar conocimientos relacionados con la aplicación de algoritmos de aprendizaje automático en el ámbito educativo, proponiendo nuevas perspectivas sobre cómo evaluar el desempeño docente. Los resultados permitirán validar teorías existentes sobre el impacto de la tecnología en la

evaluación educativa o incluso establecer nuevos enfoques que podrían contribuir al desarrollo del conocimiento en esta área.

De manera práctica, la investigación se justificará por su utilidad en el diseño de una herramienta tecnológica que la universidad podrá aplicar para evaluar y clasificar el desempeño de su personal docente. Este modelo podrá ser adaptado a diferentes contextos académicos, contribuyendo a optimizar procesos institucionales y garantizar la mejora continua de la calidad educativa.

Socialmente, la investigación beneficiará a los estudiantes, quienes serán los principales receptores de una educación de mayor calidad al contar con docentes evaluados objetivamente. Además, contribuirá al fortalecimiento de la institución universitaria, permitiéndole responder mejor a las demandas de la sociedad en términos de excelencia académica y formación profesional de alto nivel.

Metodológicamente, la investigación será relevante porque propondrá el uso de técnicas de aprendizaje automático como un enfoque innovador para abordar problemas educativos. Esto incluirá la implementación de nuevos métodos y herramientas que podrán replicarse en otros contextos, promoviendo avances en el diseño metodológico de investigaciones futuras relacionadas con la evaluación educativa.

Finalmente, desde un punto de vista científico, este proyecto aportará datos y resultados que podrán ser utilizados como base para investigaciones futuras sobre la evaluación docente y el uso de tecnologías avanzadas en la educación. Los hallazgos contribuirán al desarrollo de estudios posteriores en los que se analicen nuevas variables, escenarios o métodos para optimizar los sistemas educativos.

El estudio se realizó basándose en el planteamiento del problema que el desarrollo de un modelo de aprendizaje automático para clasificar el desempeño docente en una universidad representa un desafío multidimensional que implica una serie de etapas críticas. En primer lugar, resulta fundamental definir claramente los objetivos de la clasificación, tales como garantizar la calidad de la enseñanza y

fomentar el desarrollo profesional del cuerpo docente. Estos objetivos pueden lograrse mediante la evaluación del desempeño basada en una variedad de indicadores y fuentes de datos confiables (Tan, 2022; Salem et al., 2021).

La selección de algoritmos de aprendizaje automático adecuados constituye un paso esencial en este proceso. Entre las opciones destacadas se encuentran la regresión logística, las máquinas de soporte vectorial, los árboles de decisión, los bosques aleatorios y los métodos de conjunto, como AdaBoost, los cuales han demostrado ser eficaces en contextos educativos (Zaidan, 2022; Guo et al., 2021). Una vez seleccionados los algoritmos, la optimización de sus hiperparámetros resulta clave para mejorar la precisión y la fiabilidad del modelo. Estudios previos han mostrado que enfoques como el análisis discriminante lineal y las redes neuronales artificiales pueden alcanzar un desempeño superior cuando se ajustan de manera adecuada (Zaidan, 2022; Zhang & Kang, 2021).

La recopilación y el preprocesamiento de datos también son pilares fundamentales en el desarrollo del modelo. La adquisición de datos precisa y completa, facilitada por tecnologías de macrodatos, asegura una base sólida para la construcción del modelo (Tan, 2022). Asimismo, los métodos de selección de características, como Relief-F y la optimización basada en colonias de hormigas, permiten identificar las variables más influyentes en el rendimiento docente, tales como indicadores demográficos, académicos y conductuales (Alam et al., 2021; Meenakshi et al., 2022).

Por otro lado, la integración de tecnologías avanzadas de inteligencia artificial, como la visión por computadora y el reconocimiento inteligente de voz, puede enriquecer el proceso de evaluación docente al proporcionar información adicional sobre la dinámica del aula y los estilos de enseñanza (Guo et al., 2021). Estas tecnologías ofrecen una perspectiva más completa y detallada del entorno educativo. Finalmente, la validación del modelo es indispensable para garantizar su eficacia. Métricas como la exactitud, la precisión, el recall y la puntuación F1 deben emplearse para evaluar su rendimiento y asegurar que cumpla con los estándares exigidos en la evaluación educativa (Xiongjun & Lv, 2022).

Mejorar el rendimiento docente en el ámbito universitario a través de algoritmos de aprendizaje automático requiere la implementación de diversas estrategias basadas en evidencias empíricas. Un enfoque destacado es la utilización de modelos de aprendizaje automático para evaluar y optimizar la calidad de la enseñanza. Por ejemplo, se ha demostrado que el algoritmo GA-SVM supera significativamente a los modelos SVM tradicionales, logrando una precisión del 98,36% en las evaluaciones de calidad docente, frente al 86,72% de la SVM convencional. Este avance evidencia la capacidad de los algoritmos avanzados para proporcionar evaluaciones más precisas y eficientes.

Además, el aprendizaje automático puede ser empleado para predecir el rendimiento académico de los estudiantes, lo que, de manera indirecta, permite evaluar la eficacia de la enseñanza. Métodos como KNN, Naive Bayes y los árboles de decisión han sido utilizados para clasificar datos de rendimiento estudiantil, destacándose el KNN por su rendimiento superior en términos de precisión, memoria y puntuación F1 (Bravo et al., 2022; Meenakshi et al., 2022). Estos enfoques no solo facilitan la identificación de patrones de éxito académico, sino que también proporcionan información valiosa para guiar la mejora docente.

Asimismo, el desarrollo de marcos integrados como el Marco de Enseñanza y Evaluación de Calidad (QTEF) ha demostrado ser efectivo al incorporar técnicas de aprendizaje automático para evaluar la competencia docente, particularmente en entornos de aprendizaje electrónico. Este marco garantiza que los educadores estén mejor preparados para fomentar el éxito académico de los estudiantes, adaptándose a las exigencias de los entornos educativos modernos (Almufarreh et al., 2023).

Por otro lado, el aprendizaje automático también desempeña un papel crucial en la predicción temprana del rendimiento estudiantil. Esta capacidad permite a los docentes implementar intervenciones personalizadas y estrategias que mejoren la participación y los resultados de aprendizaje (Kalaiselvi et al., 2022). Estas aplicaciones promueven un enfoque proactivo en la mejora de la enseñanza y el aprendizaje, asegurando un impacto positivo en la experiencia educativa.

Siguiendo estos pasos, es posible desarrollar un modelo robusto de aprendizaje automático que no solo clasifique el desempeño docente, sino que también contribuya a la mejora continua de la calidad educativa en un entorno universitario.

Para lo cual, se formuló la interrogante del problema general siguiente: ¿Cómo desarrollar un modelo de aprendizaje automático para la clasificación del desempeño docente en una Universidad de Chimbote en el año 2024?

Referente a la conceptualización y operacionalización de variables, tenemos:

Sobre la variable de Desempeño docente, se formuló:

Definición conceptual: El desempeño docente universitario se refiere a la calidad y eficacia con la que los profesores llevan a cabo su labor educativa en el ámbito superior, abarcando aspectos como la planificación de clases, la interacción con los estudiantes y la evaluación del aprendizaje. Este desempeño es fundamental para garantizar una educación de calidad que fomente el desarrollo integral de los estudiantes y prepare a las futuras generaciones para enfrentar los desafíos del mundo laboral (Arroyo, 2023).

Definición operacional: El desempeño docente universitario puede ser operacionalizado a través de indicadores específicos que midan la efectividad en la enseñanza, tales como el nivel de satisfacción estudiantil, los resultados académicos obtenidos y la implementación de metodologías innovadoras en el aula. Estos indicadores permiten evaluar no solo la calidad de la enseñanza, sino también el impacto que tiene en la formación profesional y personal de los estudiantes, asegurando así una mejora continua en los procesos educativos.

Estas dimensiones incluyen planificación, que hace referencia a la organización estructurada de objetivos, contenidos y estrategias educativas alineadas con los fines académicos. A continuación, se destaca la enseñanza, entendida como la ejecución efectiva de las estrategias pedagógicas para facilitar el aprendizaje significativo de los estudiantes.

Por otro lado, se considera el cumplimiento, el cual abarca la responsabilidad del docente para cumplir con las obligaciones académicas, como entrega de materiales, evaluaciones y asistencia. Relacionado con este aspecto, la responsabilidad constituye un valor esencial que refleja el compromiso y la ética profesional del docente en sus actividades diarias.

Asimismo, se resalta la formación académica y profesional, que comprende la preparación constante y el desarrollo de competencias actualizadas para fortalecer el desempeño docente. En relación con ello, la investigación juega un papel clave, pues promueve la generación de conocimiento y la aplicación de resultados en el ámbito educativo.

Finalmente, se incluye la participación, la cual implica la integración del docente en actividades institucionales, académicas y sociales, fomentando un ambiente colaborativo y de desarrollo continuo. Estas dimensiones en conjunto permiten evaluar de manera integral la labor docente en las instituciones educativas.

Sobre la variable de Factores asociados, se consideró:

Definición conceptual: Los factores asociados al desempeño docente son diversos y pueden incluir la experiencia laboral, la formación académica, el apoyo institucional y la motivación personal. Estos elementos influyen en cómo los docentes se relacionan con sus estudiantes y en su capacidad para adaptar las estrategias de enseñanza a las necesidades del grupo (Llovio et al., 2023).

Factores personales que influyen en el desempeño docente incluyen la motivación, la experiencia previa y las habilidades interpersonales del profesor. Estos aspectos pueden afectar directamente su capacidad para conectar con los estudiantes y fomentar un ambiente de aprendizaje positivo (Aindra et al., 2022)

Factores académicos que afectan el desempeño docente abarcan la preparación curricular, los recursos disponibles y el apoyo institucional. La calidad del material didáctico y la formación continua también son cruciales para garantizar que los educadores estén equipados con las herramientas necesarias para abordar las diversas necesidades de sus alumnos (Werang et al., 2022).

Factores de formación profesional que impactan el desempeño docente incluyen la formación inicial, los programas de desarrollo profesional y las oportunidades de capacitación. La actualización constante en metodologías pedagógicas y el acceso a talleres especializados son fundamentales para mejorar la práctica educativa y adaptarse a los cambios en el contexto educativo (Bastian et al., 2022).

Definición operacional: Los factores asociados al desempeño docente son esenciales para comprender cómo influyen en la efectividad de los educadores. Estos factores pueden incluir variables como el nivel de experiencia, las habilidades interpersonales y la capacidad de trabajar en equipo, así como el acceso a recursos tecnológicos que faciliten la enseñanza.

La investigación se realizó en base a la hipótesis general siguiente: El desarrollo de un modelo de aprendizaje automático permite clasificar de manera objetiva y precisa el desempeño docente en una universidad de Chimbote, optimizando los procesos evaluativos y promoviendo la mejora continua de la calidad educativa.

Asimismo, se consideró las hipótesis específicas siguiente:

H1: Los factores personales, académicos y de formación profesional de los docentes en una universidad de Chimbote durante el período 2024 se asocian significativamente con su desempeño docente.

H2: La implementación de técnicas de aprendizaje automático en el análisis de datos de desempeño docente en una universidad de Chimbote durante el período 2024 permite seleccionar algoritmos que clasifican con alta precisión el desempeño docente.

H3: La evaluación comparativa de diferentes modelos de clasificación basados en aprendizaje automático identifica aquellos con mayor precisión, interpretabilidad y capacidad de generalización en la clasificación del desempeño docente en una universidad de Chimbote durante el período 2024.

Respecto a la ejecución, se sigue el objetivo general siguiente: Desarrollar un modelo de aprendizaje automático para clasificar el desempeño docente en una universidad de Chimbote durante el año 2024, evaluando su precisión, interpretabilidad y capacidad de generalización.

En la misma línea, se plasmaron los objetivos específicos siguientes:

Identificar los factores pedagógicos, metodológicos y de formación profesional que influyen significativamente en el desempeño docente en una universidad de Chimbote durante el año 2024.

Implementar técnicas de aprendizaje automático para analizar los datos de desempeño docente, seleccionando los algoritmos más adecuados que permitan una clasificación precisa y objetiva.

Evaluar comparativamente diferentes modelos de clasificación basados en aprendizaje automático para determinar aquellos con mayor precisión, interpretabilidad y capacidad de generalización en la clasificación del desempeño docente.

Metodología

Según su finalidad, la investigación fue de tipo aplicada, ya que busco utilizar conocimientos existentes para resolver un problema específico relacionado con la clasificación del desempeño docente mediante modelos de aprendizaje automático. La investigación aplicada se centra en la aplicación práctica de teorías y conocimientos para abordar problemas concretos (Hernández-Sampieri & Mendoza, 2018).

Según su alcance, el estudio fue de tipo predictiva, puesto que pretendió identificar y analizar la relación entre variables como factores personales, académicos y de formación profesional de los docentes, y su desempeño. Las investigaciones correlacionales buscan determinar el grado de asociación entre dos o más variables (Hernández-Sampieri & Mendoza, 2018).

Según el tipo de ocurrencia, la investigación fue de tipo transversal, ya que recolectará datos en un único momento del tiempo, específicamente durante el año 2024, para analizar las variables de interés. Los estudios transversales se caracterizan por observar variables en un punto específico, proporcionando una medición del fenómeno en estudio (Hernández-Sampieri & Mendoza, 2018).

El diseño adoptado es no experimental, debido a que no se manipularán deliberadamente las variables independientes; en su lugar, se observarán y analizarán tal como se presentan en su contexto natural. Dentro de los diseños no experimentales, se empleará un diseño transversal correlacional, que permite examinar la relación entre variables en un momento determinado sin intervención del investigador (Hernández-Sampieri & Mendoza, 2018).

La población del estudio estuvo constituida por los registros de desempeño docente de una universidad en Chimbote durante el año 2024, totalizando 800 registros. Estos registros incluyen información relevante sobre factores personales, académicos y de formación profesional de los docentes. No se utilizará muestra, en tal sentido, no se hará uso de ninguna técnica de muestreo.

Se utilizó la técnica de análisis documental, enfocada en la revisión y evaluación de los registros del sistema informático de la universidad que contienen datos sobre el desempeño docente. El instrumento principal fue una ficha de recolección de datos, diseñada para extraer información pertinente de los registros digitales. El análisis documental es una técnica que permite obtener datos relevantes de documentos existentes, facilitando la comprensión de fenómenos específicos (Hernández-Sampieri & Mendoza, 2018).

Para asegurar la validez del instrumento, se realizó una validación de contenido mediante la evaluación de expertos en el área educativa y tecnológica, quienes revisarán la ficha de recolección de datos para garantizar que mida adecuadamente las variables de interés. La validez de contenido se refiere al grado en que un instrumento representa adecuadamente el constructo que se desea medir (Hernández-Sampieri & Mendoza, 2018).

La confiabilidad del instrumento se evaluó a través de una prueba piloto aplicada a una pequeña muestra de registros, permitiendo calcular el coeficiente de consistencia interna (por ejemplo, el alfa de Cronbach) para determinar la estabilidad y consistencia de las mediciones. La confiabilidad se refiere a la precisión y estabilidad de las mediciones obtenidas con un instrumento (Hernández-Sampieri & Mendoza, 2018).

El procesamiento y análisis de la información en esta investigación se estructuró siguiendo la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), reconocida por su eficacia en proyectos de minería de datos y aprendizaje automático. Esta metodología comprende seis fases interrelacionadas:

- **Comprensión del Negocio:** En esta etapa inicial, se identificarán los objetivos institucionales relacionados con la evaluación del desempeño docente, estableciendo criterios y métricas que el modelo de aprendizaje automático debe considerar para alinearse con las necesidades de la universidad. Una comprensión clara del problema del negocio y la definición de los objetivos del proyecto son fundamentales en esta fase.

- **Comprensión de los Datos:** Se procederá a la recopilación exhaustiva de los registros del sistema informático de la universidad, que contienen información sobre factores personales, académicos y de formación profesional de los docentes. Este análisis preliminar permitirá identificar la calidad, completitud y relevancia de los datos disponibles, sentando las bases para todo el proceso de minería de datos.
- **Preparación de los Datos:** En esta fase, se llevará a cabo la limpieza y transformación de los datos para asegurar su calidad y adecuación al análisis. Esto incluye la imputación de datos faltantes, normalización de variables y codificación de datos categóricos, garantizando que el conjunto de datos esté en el formato adecuado y sea apto para el análisis posterior.
- **Modelado:** Se seleccionarán y aplicarán diversos algoritmos de clasificación supervisada, como árboles de decisión, máquinas de soporte vectorial y redes neuronales, para construir modelos que clasifiquen el desempeño docente. Cada modelo será entrenado y ajustado utilizando técnicas de validación cruzada para optimizar su rendimiento. La fase de modelado implica la construcción y evaluación de modelos predictivos utilizando técnicas estadísticas y de aprendizaje automático.
- **Evaluación:** Los modelos desarrollados serán evaluados mediante métricas como precisión, sensibilidad, especificidad y el área bajo la curva ROC, para determinar su eficacia en la clasificación del desempeño docente. Además, se considerará la interpretabilidad y capacidad de generalización de cada modelo para seleccionar el más adecuado. La evaluación de los modelos creados es crucial para garantizar su precisión y relevancia.

Este enfoque sistemático garantiza que el procesamiento y análisis de la información se realicen de manera estructurada y alineada con los objetivos de la investigación, asegurando la validez y confiabilidad de los resultados obtenidos.

Resultados

Se evidencia los hallazgos para el primer objetivo específico, identificar los factores asociados que influyen significativamente en el desempeño docente en una universidad de Chimbote durante el año 2024. Se tienen los resultados en la Tabla 1.

Tabla 1

Métricas de modelos de clasificación mediante validación cruzada

Atributo	Gain	Chi Cuadrado	ReliefF
Realiza investigación	0.067	988.965	0.146
Adscrito a Ciencias de la Salud	0.040	561.130	0.080
Grado de Doctor	0.027	429.477	0.050
Adscrito a Derecho y Ciencia Política	0.024	424.354	0.068
Trabajan en Huacho	0.024	354.877	0.064
Sexo	0.017	134.960	0.104
Trabajan en Huaraz	0.016	275.859	0.028
Adscritos a Ingeniería	0.013	211.953	0.048
Trabajan en Piura	0.012	202.779	0.078
No tienen horas carga no lectiva	0.011	72.781	0.114

El análisis de relevancia de atributos reveló patrones significativos en su capacidad predictiva y discriminativa. El atributo relacionado a los docentes que realizan investigación emergió como el factor más influyente en todas las métricas evaluadas, exhibiendo un Information Gain de 0.067 y un valor chi-cuadrado (χ^2) de 988.965, lo que evidencia una asociación estadística robusta con la variable objetivo. Esta predominancia se confirma además por su elevado puntaje ReliefF de 0.146, que destaca su capacidad excepcional para discriminar entre clases en el espacio multidimensional de características.

En segundo orden de importancia, corresponde a los docentes que se encuentran adscrito a la Facultad de Ciencias Económicas, demostró una capacidad predictiva sustancial, reflejada en un Information Gain de 0.040 y un valor χ^2 de 561.130. Su puntaje ReliefF de 0.080 corrobora su posición como segundo atributo más relevante, sugiriendo una contribución fundamental en el proceso de clasificación.

Respecto al grado académico de doctor, presentó una influencia moderada pero significativa, con un Information Gain de 0.027 y un valor χ^2 de 429.477. A pesar de mostrar un puntaje ReliefF más modesto (0.050), mantiene una relevancia notable en el contexto global del modelo predictivo.

Que el atributo referido a los docentes que labora en la Facultad de Derecho, exhibió una contribución significativa, evidenciada por un valor χ^2 de 424.354 y un Information Gain de 0.024. Su puntaje ReliefF intermedio de 0.068 confirma su importancia en la estructura predictiva del modelo.

La característica sobre los docentes que trabajan en la Filial Huacho se posicionó como un atributo de considerable relevancia, con un valor χ^2 de 354.877 y un Information Gain de 0.024. Su puntaje ReliefF de 0.064, aunque menor, sugiere una capacidad significativa para capturar patrones subyacentes en los datos. El atributo se considera el sexo de un docente mostró una relevancia estadística moderada, con un valor χ^2 de 134.960 y un Information Gain de 0.017. Notablemente, su elevado puntaje ReliefF de 0.104 sugiere una capacidad discriminativa particularmente efectiva en determinados subgrupos de la población estudiada.

Con respecto a los docentes que trabajan en la Filial Huaraz se presentó niveles moderados de relevancia, con un valor χ^2 de 275.859 y un Information Gain de 0.016. Su puntaje ReliefF de 0.028 indica una influencia secundaria pero potencialmente significativa en el contexto predictivo.

Los atributos respecto a los docentes de Facultad de Ingeniería, que laboran en la Filial Piura y que no tiene horas no lectivas, mostraron niveles más bajos de relevancia estadística, con valores de Information Gain que oscilan entre 0.013 y 0.011, y valores χ^2 entre 211.953 y 72.781. Sin embargo, es notable que los docentes

que no cuentan con horas no lectivas, a pesar de presentar los valores más bajos en las métricas tradicionales, exhibió un puntaje ReliefF relativamente alto de 0.114, sugiriendo una potencial importancia en las interacciones con otros atributos que podría no ser capturada por las métricas univariadas.

A continuación, el segundo objetivo específico, implementar técnicas de aprendizaje automático para analizar los datos de desempeño docente, seleccionando los algoritmos más adecuados que permitan una clasificación precisa y objetiva, se ha desarrollado tres experimentos con diversos parámetros de configuración, se describen: La configuración inicial se estableció considerando tres aspectos fundamentales: el esquema de muestreo, la replicabilidad y el protocolo de validación. En primer lugar, se implementó una estrategia de muestreo estratificado, donde se destinó el 70% de los datos al conjunto de entrenamiento, garantizando así la preservación de las proporciones originales de las clases presentes en el conjunto de datos completo. Esta decisión metodológica resulta crucial para mantener la representatividad de la distribución de clases en la fase de entrenamiento.

Para asegurar la reproducibilidad del experimento, un aspecto fundamental en la investigación científica, se activó específicamente la opción "Replicable (deterministic) sampling". Esta configuración permite que los resultados sean verificables y reproducibles por otros investigadores, fortaleciendo así la validez científica del estudio.

En cuanto al proceso de validación, se implementaron dos esquemas de validación cruzada diferentes. Por un lado, se ejecutó una validación cruzada de 5 particiones (5-fold cross-validation), donde el conjunto de datos se dividió en cinco subconjuntos de igual tamaño. En este esquema, cuatro subconjuntos se utilizaron para el entrenamiento del modelo, mientras que el subconjunto restante se empleó para la evaluación de este. Este proceso se iteró cinco veces, rotando sistemáticamente el subconjunto de evaluación.

Paralelamente, se implementó una validación cruzada de 10 particiones (10-fold cross-validation), siguiendo el mismo principio, pero con una granularidad más fina. En este caso, el conjunto de datos se dividió en diez subconjuntos iguales, donde

nueve se destinaron al entrenamiento y uno a la evaluación, repitiendo el proceso diez veces. Esta doble aproximación en la validación permite obtener una evaluación más robusta y comprensiva del rendimiento de los modelos bajo diferentes condiciones de particionamiento de datos.

Tabla 2

Métricas de modelos de clasificación mediante validación cruzada

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	0.949	0.000	0.626	0.591	0.592	0.600	0.591	0.184	6.586
Random Forest	0.212	0.134	0.652	0.609	0.608	0.609	0.611	0.308	0.792
Gradient Boosting	0.910	0.066	0.695	0.637	0.637	0.637	0.637	0.267	0.647
Neural Network	2.199	0.095	0.708	0.656	0.654	0.656	0.654	0.294	0.626
Logistic Regression	0.155	0.056	0.717	0.663	0.663	0.661	0.663	0.312	0.616
kNN	0.127	0.244	0.622	0.594	0.594	0.594	0.577	0.173	1.590
SVM	0.382	0.111	0.585	0.589	0.589	0.588	0.589	0.162	0.679
Naive Bayes	0.047	0.008	0.712	0.665	0.665	0.664	0.665	0.316	0.635

En este experimento, el modelo Gradient Boosting destacó con un AUC de 0.695 y un LogLoss de 0.647, mostrando un buen equilibrio entre capacidad discriminativa y calibración de las probabilidades. Este modelo presentó valores aceptables de F1-Score (0.637), precisión y recall, lo que lo posiciona como una de las mejores opciones.

El modelo de Regresión Logística, por otro lado, obtuvo el AUC más alto (0.717) y un LogLoss de 0.616, evidenciando un rendimiento muy competitivo. Las métricas de precisión y recall bien equilibradas (0.663 y 0.661, respectivamente) lo consolidan como otra alternativa sólida.

Los modelos de Árbol de Decisión y Random Forest tuvieron un rendimiento inferior en comparación con Gradient Boosting y Regresión Logística. El Árbol de Decisión, aunque simple, mostró un AUC de 0.626, mientras que Random Forest logró un AUC de 0.652. Sin embargo, ambos modelos presentaron valores altos de LogLoss, lo que indica limitaciones en la estimación de probabilidades.

Por otro lado, Naive Bayes destacó por su simplicidad y por un AUC competitivo (0.712), aunque sus métricas de precisión y recall estuvieron por debajo de otros modelos. Los modelos SVM y kNN, en contraste, tuvieron un rendimiento menor, reflejado en valores bajos de AUC (0.585 y 0.622, respectivamente).

Tabla 3

Métricas de modelos de clasificación por muestreo aleatorio

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	0.779	0.001	0.639	0.598	0.599	0.61	0.598	0.202	6.667
Random Forest	0.208	0.071	0.651	0.615	0.615	0.614	0.617	0.213	0.8
Gradient Boosting	0.823	0.069	0.692	0.647	0.647	0.647	0.647	0.279	0.655
Neural Network	4.164	0.105	0.717	0.66	0.66	0.659	0.66	0.306	0.625
Logistic Regression	0.163	0.057	0.719	0.661	0.661	0.661	0.661	0.306	0.616
kNN	0.126	0.266	0.6	0.646	0.646	0.646	0.582	0.143	0.671
SVM	0.344	0.132	0.6	0.582	0.58	0.58	0.582	0.143	0.671
Naive Bayes	0.048	0.014	0.708	0.659	0.659	0.659	0.659	0.306	0.643

En el segundo experimento, el modelo Gradient Boosting se mantuvo consistente con un AUC de 0.692 y un LogLoss de 0.655. Este modelo nuevamente mostró valores equilibrados en F1-Score, precisión y recall, confirmando su capacidad como una de las opciones más confiables.

El modelo de Regresión Logística sobresalió con un AUC de 0.719, manteniendo la mejor capacidad discriminativa. Además, el LogLoss se mantuvo bajo (0.616), lo que refuerza su estabilidad en la estimación de probabilidades.

El desempeño de los modelos Árbol de Decisión y Random Forest mostró mejoras leves respecto al primer experimento. El Árbol de Decisión incrementó su AUC a 0.639, pero el LogLoss continuó siendo elevado (6.667). Random Forest mantuvo un rendimiento moderado con un AUC de 0.651.

Por otro lado, Naive Bayes mejoró su LogLoss (0.643) mientras mantenía un AUC competitivo de 0.708. El modelo de Red Neuronal, aunque mostró un AUC elevado (0.717), reflejó un posible sobreajuste debido a los valores desproporcionados en el conjunto de entrenamiento.

A partir del primer experimento, se tiene que los modelos Gradient Boosting y Regresión Logística son los más prometedores. Ambos presentan un equilibrio destacado entre discriminación y calibración, siendo consistentes en las métricas evaluadas. La Regresión Logística, además, destaca por su estabilidad en el AUC y LogLoss, mientras que Gradient Boosting se posiciona como una alternativa robusta en escenarios de clasificación más complejos. Aunque otros modelos como Random Forest y Naive Bayes mostraron desempeños moderados, estos son menos competitivos frente a los anteriores. El Árbol de Decisión, aunque simple, tiene limitaciones significativas debido a su alto LogLoss, y los modelos SVM y kNN tuvieron un desempeño inferior en términos generales.

En el segundo experimento, la configuración metodológica se centró en tres aspectos fundamentales: el esquema de validación, la replicabilidad y la estratificación de los datos. En primer lugar, se implementó como técnica principal de muestreo una validación cruzada de 10 particiones (10-fold cross-validation). Sin embargo, es importante destacar que se introdujo una variación significativa en el protocolo estándar: uno de los subconjuntos se reservó, excluyéndolo tanto del proceso de entrenamiento como de la validación. Esta decisión metodológica proporciona un conjunto de datos independiente que puede ser utilizado para evaluaciones posteriores o pruebas finales, fortaleciendo así la robustez de la validación del modelo.

Con el objetivo de mantener el rigor científico y la reproducibilidad del experimento, se conservó activa la opción "Replicable (deterministic) sampling". Esta configuración resulta crucial para garantizar que los resultados puedan ser verificados y reproducidos por otros investigadores, contribuyendo así a la validez científica del estudio.

Asimismo, se mantuvo la estrategia de estratificación de los datos mediante la opción "Stratify sample (when possible)". Esta decisión metodológica asegura que la distribución de las clases en los diferentes subconjuntos generados durante la validación cruzada sea representativa de la distribución original del conjunto de datos completo, minimizando así posibles sesgos en el proceso de evaluación del modelo.

Tabla 4

Métricas de modelos de clasificación mediante validación cruzada

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	1.156	0.000	0.634	0.604	0.611	0.611	0.604	0.206	6.209
Random Forest	0.218	0.127	0.680	0.655	0.657	0.657	0.655	0.297	0.866
Gradient Boosting	1.045	0.063	0.719	0.665	0.665	0.666	0.665	0.317	0.618
Neural Network	6.775	0.100	0.731	0.662	0.662	0.665	0.661	0.317	0.607
Logistic Regression	0.162	0.055	0.734	0.666	0.666	0.666	0.666	0.317	0.601
kNN	0.128	0.283	0.650	0.646	0.646	0.646	0.597	0.294	1.829
SVM	0.448	0.116	0.607	0.582	0.580	0.579	0.559	0.143	0.671
Naive Bayes	0.060	0.003	0.718	0.654	0.654	0.656	0.294	0.294	0.627

En el segundo experimento, los resultados revelaron patrones significativos en el desempeño de los diversos modelos implementados. El modelo de Gradient Boosting demostró una notable capacidad discriminativa, evidenciada por un AUC de 0.719, acompañado de un LogLoss de 0.618, lo cual indica una calibración precisa de las probabilidades estimadas. Asimismo, las métricas de evaluación complementarias, incluyendo la precisión, F1-Score y Recall, convergieron en 0.666, lo que denota una consistencia remarcable en su capacidad predictiva.

Por su parte, la Regresión Logística emergió como el modelo más sobresaliente en esta primera fase experimental, alcanzando un AUC de 0.734, el valor más elevado entre todos los modelos evaluados. Este desempeño excepcional se complementó con un LogLoss de 0.601, el más bajo registrado, lo que confirma su robustez en términos

de calibración probabilística. Adicionalmente, las métricas de precisión y recall (0.666) evidenciaron un equilibrio óptimo en su capacidad predictiva.

En cuanto a la Red Neuronal, esta exhibió un comportamiento particularmente interesante, alcanzando un AUC de 0.731, aunque con un LogLoss de 0.607, ligeramente superior al de la Regresión Logística. Este patrón, junto con un F1-Score de 0.662, sugiere un modelo altamente competitivo, si bien con posibles indicios de sobreajuste que merecen consideración. El modelo Random Forest, por otro lado, mostró limitaciones significativas en su desempeño, con un AUC de 0.634 y un LogLoss de 0.866, valores que indican restricciones en su capacidad de calibración probabilística. No obstante, mantiene una posición moderadamente aceptable en el conjunto de modelos evaluados.

Tabla 5

Métricas de modelos de clasificación por muestreo aleatorio

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	0.943	0.000	0.630	0.608	0.618	0.608	0.221	0.615	6.839
Random Forest	0.210	0.072	0.669	0.631	0.631	0.632	0.240	0.616	0.851
Gradient Boosting	0.911	0.070	0.711	0.666	0.666	0.666	0.320	0.655	0.631
Neural Network	2.680	0.103	0.727	0.666	0.666	0.669	0.318	0.646	0.611
Logistic Regression	0.159	0.055	0.735	0.667	0.666	0.667	0.310	0.649	0.603
kNN	0.131	0.368	0.652	0.622	0.622	0.622	0.229	0.606	1.824
SVM	0.460	0.152	0.591	0.591	0.588	0.587	0.160	0.567	0.669
Naive Bayes	0.059	0.009	0.709	0.652	0.653	0.652	0.294	0.643	0.641

En el segundo caso, los resultados mostraron una notable consistencia con respecto a la primera iteración. El modelo Gradient Boosting mantuvo su robustez con un AUC de 0.711 y un LogLoss de 0.631, confirmando su estabilidad y confiabilidad. De manera similar, la Regresión Logística reafirmó su superioridad al alcanzar un AUC de 0.735 y un LogLoss de 0.603, consolidando así su posición como el modelo más eficaz en ambas fases experimentales.

La Red Neuronal, en esta segunda fase, demostró una ligera mejora en su comportamiento, registrando un AUC de 0.727 y un LogLoss de 0.611, lo que sugiere un ajuste más controlado de sus parámetros. Por su parte, el Árbol de Decisión, aunque mostró una leve mejoría en su AUC (0.630), mantuvo un LogLoss significativamente elevado (6.839), lo que compromete su aplicabilidad en contextos que requieren alta precisión.

Es particularmente notable el desempeño del modelo Naive Bayes, que a pesar de su relativa simplicidad, mantuvo un rendimiento competitivo en ambas fases experimentales, alcanzando en la segunda iteración un AUC de 0.709 y un LogLoss de 0.641. Estos resultados sugieren su viabilidad como alternativa eficiente en escenarios donde la complejidad computacional debe equilibrarse con la precisión predictiva.

Este análisis exhaustivo de los resultados experimentales proporciona evidencia substancial para la selección informada de modelos en función de los requisitos específicos de implementación, considerando tanto la capacidad discriminativa como la calibración probabilística como criterios fundamentales de evaluación.

El tercer experimento se estructuró mediante una configuración metodológica que incorporó técnicas avanzadas de muestreo y validación, implementadas en dos fases diferenciadas. En la fase inicial, se adoptó el método de Bootstrap como estrategia principal de muestreo, una técnica de remuestreo que permite generar múltiples conjuntos de datos mediante el muestreo con reemplazo del conjunto original. Esta aproximación resulta particularmente valiosa para obtener estimaciones más robustas del rendimiento del modelo y evaluar su variabilidad.

La primera parte del protocolo de validación se fundamentó en una validación cruzada estratificada de 10 particiones (10-fold stratified cross-validation). Esta metodología asegura que la distribución de las clases en cada una de las particiones refleje la distribución del conjunto de datos completo, proporcionando así una evaluación más precisa y representativa del rendimiento del modelo.

En la segunda fase del experimento, se implementó un enfoque basado en muestreo aleatorio con dos niveles de iteración distintos. Inicialmente, se ejecutó el proceso con 10 repeticiones, seguido de una intensificación a 50 repeticiones. En ambos casos, se mantuvo constante la proporción del conjunto de entrenamiento en un 70% del total de los datos, aplicando estratificación para preservar la distribución de las clases. Esta progresión en el número de repeticiones permite evaluar la estabilidad y consistencia de los resultados bajo diferentes intensidades de muestreo, proporcionando una comprensión más profunda de la robustez del modelo.

Esta configuración experimental, que combina técnicas de Bootstrap, validación cruzada estratificada y muestreo aleatorio repetido, proporciona un marco comprehensivo para la evaluación del rendimiento y la estabilidad de los modelos implementados.

Tabla 6

Métricas de modelos de clasificación mediante validación cruzada

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	1.085	0.001	0.736	0.737	0.745	0.736	0.548	0.477	3.089
Random Forest	0.234	0.070	0.849	0.777	0.777	0.778	0.548	0.477	0.680
Gradient Boosting	1.150	0.069	0.820	0.760	0.760	0.759	0.515	0.753	0.515
Neural Network	11.066	0.112	0.735	0.734	0.734	0.735	0.460	0.724	0.538
Logistic Regression	0.184	0.063	0.792	0.722	0.722	0.722	0.433	0.733	0.548
kNN	0.139	0.295	0.778	0.720	0.720	0.720	0.720	0.700	1.716
SVM	0.576	0.123	0.646	0.603	0.607	0.607	0.143	0.560	0.671
Naive Bayes	0.063	0.005	0.772	0.703	0.704	0.704	0.400	0.700	0.583

El modelo Gradient Boosting sobresalió como una de las mejores opciones, con un AUC de 0.820 y un LogLoss bajo de 0.515. Esto demuestra su capacidad para equilibrar la discriminación entre clases y la calibración de probabilidades. Además, sus métricas de precisión, F1-Score y Recall (todas alrededor de 0.760) consolidan su desempeño consistente.

Por otro lado, Random Forest obtuvo el AUC más alto en este conjunto de resultados (0.849), lo que indica una destacada capacidad discriminativa. Sin embargo, el LogLoss de 0.680 sugiere cierta imprecisión en la calibración de las probabilidades, lo que limita su rendimiento en escenarios donde la confianza en las predicciones es crítica.

El modelo de Regresión Logística mostró un desempeño estable, con un AUC de 0.792 y un LogLoss bajo de 0.548. Las métricas de precisión y Recall (ambas en 0.722) estuvieron equilibradas, lo que refuerza su utilidad como un modelo robusto y confiable para problemas de clasificación.

La Red Neuronal, aunque competitiva en términos de AUC (0.735) y LogLoss (0.538), presentó valores de entrenamiento significativamente altos (11.066), lo que sugiere un posible sobreajuste. Este modelo requeriría ajustes adicionales para mejorar su generalización.

También se tienen a los modelos más simples, como Naive Bayes, demostraron ser opciones viables para problemas menos complejos. Con un AUC de 0.772 y un LogLoss de 0.583, Naive Bayes mostró un buen equilibrio entre simplicidad y capacidad predictiva. Sin embargo, los modelos SVM y kNN tuvieron un rendimiento más limitado, con métricas inferiores en comparación con los modelos previamente destacados.

Tabla 7

Métricas de modelos de clasificación por muestreo aleatorio

Modelo	Train	Test	AUC	F1	Precisión	Recall	MCC	Especificidad	LogLoss
Tree	4.105	0.004	0.796	0.739	0.744	0.738	0.478	0.743	3.651
Random Forest	0.960	0.356	0.833	0.765	0.765	0.765	0.523	0.758	0.736
Gradient Boosting	4.826	0.366	0.818	0.761	0.761	0.760	0.515	0.755	0.519
Neural Network	41.806	0.702	0.794	0.733	0.733	0.734	0.458	0.743	0.553
Logistic Regression	0.765	0.305	0.790	0.725	0.725	0.725	0.442	0.649	0.552
kNN	0.600	2.061	0.777	0.715	0.715	0.715	0.422	0.708	1.798
SVM	2.031	0.814	0.716	0.662	0.663	0.662	0.318	0.657	0.615
Naive Bayes	0.212	0.023	0.771	0.709	0.709	0.711	0.413	0.706	0.587

El desempeño del modelo Gradient Boosting se mantuvo consistente en este conjunto, con un AUC de 0.818 y un LogLoss bajo de 0.519. Sus métricas de precisión y Recall (ambas de 0.760) refuerzan su capacidad para realizar predicciones confiables y bien calibradas.

Por su parte, Random Forest obtuvo un AUC sobresaliente de 0.833, consolidándose como el modelo con la mejor capacidad discriminativa en este experimento. No obstante, su LogLoss de 0.736 indica que podría beneficiarse de ajustes adicionales en la calibración. La Regresión Logística también mostró un desempeño sólido, con un AUC de 0.790 y un LogLoss de 0.552. Sus métricas de precisión y Recall (ambas de 0.725) fueron equilibradas, lo que reafirma su estabilidad y confiabilidad en aplicaciones prácticas.

La Red Neuronal, con un AUC de 0.794 y un LogLoss de 0.553, se mantuvo competitiva, aunque los valores excesivamente altos en el entrenamiento (41.806) sugieren un sobreajuste significativo. Este modelo, aunque prometedor, requiere ajustes adicionales para garantizar su generalización.

Por otro lado, Naive Bayes destacó como un modelo simple pero efectivo, con un AUC de 0.771 y un LogLoss de 0.587, lo que lo convierte en una alternativa viable para problemas menos exigentes. Los modelos SVM y kNN, en cambio, mostraron un desempeño inferior, siendo menos recomendados para problemas complejos.

Para el tercer objetivo específico para evaluar comparativamente diferentes modelos de clasificación basados en aprendizaje automático para determinar aquellos con mayor precisión, interpretabilidad y capacidad de generalización en la clasificación del desempeño docente.

Tabla 8
Comparación de los modelos de clasificación con mejores métricas de rendimiento

Descripción	Modelo	AUC	F1	MCC	LogLoss	Train	Test
Experimento 1	Gradient Boosting	0.692	0.647	0.279	0.655	0.823	0.069
Experimento 2	Logistic Regression	0.734	0.669	0.317	0.601	0.162	0.055
Experimento 3	Random Forest	0.833	0.765	0.523	0.519	4.826	0.366

El análisis comparativo de los resultados experimentales revela patrones significativos en el rendimiento de los tres modelos principales evaluados. A continuación, se presenta un análisis detallado de cada modelo: En el primer experimento, el modelo Gradient Boosting exhibió un rendimiento moderado pero consistente. El valor de AUC de 0.692 indica una capacidad discriminativa aceptable, aunque inferior a los otros modelos evaluados. Es particularmente notable el F1-Score de 0.647, que evidencia un equilibrio satisfactorio entre precisión y exhaustividad. La calibración del modelo, reflejada en un LogLoss de 0.655, sugiere una estimación confiable de las probabilidades. Adicionalmente, la diferencia entre los valores de entrenamiento (0.823) y prueba (0.069) indica un ajuste apropiado del modelo, sin manifestar signos significativos de sobreajuste.

En el segundo experimento, la Regresión Logística demostró un desempeño sobresaliente en varios aspectos. El modelo alcanzó un AUC de 0.734, posicionándose como una opción robusta para tareas de clasificación. Este rendimiento se complementa con un F1-Score de 0.669, el más elevado entre los modelos evaluados en su categoría, lo que denota un equilibrio óptimo entre precisión y exhaustividad. Es especialmente destacable el LogLoss de 0.601, el más bajo registrado, que evidencia una calibración excepcional de las probabilidades predichas. La diferencia contenida entre los valores de entrenamiento (0.162) y prueba (0.055) confirma la capacidad del modelo para generalizar efectivamente.

El tercer experimento, centrado en Random Forest, reveló resultados particularmente notables. El modelo logró el AUC más alto de todos los experimentos con 0.833, demostrando una capacidad discriminativa superior. Este rendimiento excepcional se refleja también en su F1-Score de 0.765, que supera significativamente a los otros modelos evaluados. El LogLoss de 0.519, el más bajo registrado en todos los experimentos, indica una calibración óptima de las probabilidades predichas. Sin embargo, es importante señalar que los valores de entrenamiento (4.826) y prueba (0.366), aunque más elevados que en los otros modelos, sugieren un potencial riesgo de sobreajuste que podría mitigarse mediante una optimización cuidadosa de los hiperparámetros.

En términos comparativos, cada modelo presenta características distintivas que los hacen apropiados para diferentes contextos de aplicación. Random Forest destaca por su superior capacidad discriminativa y equilibrio en las métricas de rendimiento, aunque requiere una atención especial para controlar el riesgo de sobreajuste. La Regresión Logística sobresale por su estabilidad y excelente calibración de probabilidades, características particularmente valiosas en aplicaciones que requieren interpretaciones probabilísticas precisas. Por su parte, Gradient Boosting, aunque con un rendimiento más modesto en términos absolutos, demuestra un comportamiento equilibrado y robusto, con un riesgo minimizado de sobreajuste.

Análisis y Discusión

En el marco del análisis predictivo del desempeño docente en educación superior, donde la evaluación objetiva y precisa del profesorado constituye un elemento fundamental para la mejora continua de la calidad educativa, la presente investigación implementó técnicas de aprendizaje automático para identificar factores determinantes y desarrollar modelos predictivos robustos. El estudio, realizado en una universidad de Chimbote durante el año 2024, analizó múltiples variables relacionadas con el desempeño docente, incluyendo aspectos académicos, profesionales y contextuales, mediante la aplicación de diversos algoritmos de clasificación. A continuación, se presenta el análisis y discusión de los resultados obtenidos, contrastándolos con investigaciones previas y examinando sus implicaciones para la gestión educativa y el desarrollo docente. La discusión se estructura en torno a los tres objetivos específicos planteados: la identificación de factores influyentes, la implementación de técnicas de aprendizaje automático y la evaluación comparativa de los modelos de clasificación.

Se presenta el análisis y discusión en función de los resultados mostrados en la Tabla 1; referentes al primer objetivo específico de la investigación identificar los factores asociados que influyen significativamente en el desempeño docente en una universidad de Chimbote durante el año 2024. El análisis de los resultados obtenidos revela patrones significativos en la relevancia de diversos atributos para la predicción del desempeño docente. En primera instancia, la variable relacionada a los docentes que realizan investigación emerge como el factor más determinante, evidenciado por el valor más elevado de Gain (0.067) y un destacado índice ReliefF (0.146). Asimismo, su significativo estadístico Chi Cuadrado (988.965) corrobora la existencia de diferencias estadísticamente significativas en el desempeño docente entre aquellos que desarrollan actividades investigativas y quienes no lo hacen.

En segunda instancia, la adscripción a la facultad de Ciencias de la Salud se posiciona como el segundo factor de mayor relevancia, con un Gain de 0.040 y un Chi Cuadrado de 561.130. Este resultado podría atribuirse a las exigencias inherentes a las carreras del área de la salud. Por otra parte, los docentes con grado de doctor

demuestran una influencia considerable en el desempeño docente, evidenciada por un Gain de 0.027 y un Chi Cuadrado de 429.477. Esta correlación sugiere que el nivel de formación académica del docente incide significativamente en su desempeño, presumiblemente debido a que la formación avanzada o especializada se relaciona con prácticas pedagógicas más efectivas y un dominio más profundo de la materia.

En cuanto a las variables relacionadas con la adscripción disciplinar y ubicación geográfica, tanto los docentes adscritos a la facultad de Derecho y Ciencia Política (Gain = 0.024) como los docentes adscritos a la facultad de Ingeniería (Gain = 0.013) muestran una influencia moderada. De manera análoga, las variables geográficas Trabajan en Huacho (Gain = 0.024), Trabajan en Huaraz (Gain = 0.016) y Trabajan en Piura (Gain = 0.012) sugieren que la ubicación del docente puede incidir en su desempeño, posiblemente debido a factores contextuales como la disponibilidad de recursos e infraestructura académica.

Con respecto a la variable Sexo, si bien presenta valores moderados de Gain (0.017) y Chi Cuadrado (134.960), su índice ReliefF (0.104) indica una influencia no despreciable en el desempeño docente. No obstante, este efecto resulta menos pronunciado en comparación con factores como la actividad investigadora o la adscripción a Ciencias de la Salud, lo cual sugiere que, aunque podría existir una brecha de género, su impacto es menos determinante que otros factores analizados. Finalmente, la variable "No tienen horas de carga no lectiva" presenta el valor más bajo de Gain (0.011), aunque mantiene un Chi Cuadrado de 72.781 y un ReliefF de 0.114. Esta configuración de métricas podría sugerir que la ausencia de horas no lectivas destinadas a actividades de gestión o asesoría podría limitar las oportunidades de planificación y mejora continua de la docencia, aunque su impacto global resulta menos significativo en comparación con otros factores analizados en el estudio.

En contraste con el estudio de Quintana (2023), quien encontró que el Árbol de Decisión fue el modelo con mejor rendimiento (77%), nuestro análisis de relevancia de atributos utilizó múltiples métricas (Information Gain, Chi-cuadrado y ReliefF) para identificar factores influyentes, proporcionando así una perspectiva más

comprehensiva de las variables predictoras. Esta diferencia metodológica podría explicar las variaciones en los hallazgos principales.

Nuestros resultados sobre la importancia del grado académico (Gain = 0.027, $\chi^2 = 429.477$) encuentran cierta resonancia con el estudio de Velásquez (2022), quien identificó una asociación positiva entre las competencias docentes y la satisfacción académica. Sin embargo, nuestro estudio proporciona una cuantificación más precisa de esta relación mediante métricas específicas.

La investigación de Naranjo (2023) se centró en aspectos específicos como los materiales didácticos, encontrando una correlación alta (0.910) con la calidad educativa. Si bien nuestro estudio no evaluó directamente esta variable, nuestros hallazgos sobre la importancia de la carga no lectiva (Gain = 0.011, ReliefF = 0.114) podrían estar indirectamente relacionados con el tiempo disponible para la preparación de materiales didácticos.

Es particularmente relevante contrastar nuestros hallazgos con las observaciones de Almufarreh et al. (2023) sobre las limitaciones del aprendizaje automático en la evaluación docente. Nuestro estudio aborda parcialmente estas preocupaciones al utilizar múltiples métricas de evaluación y considerar variables contextuales como la ubicación geográfica y la adscripción facultativa, proporcionando así un marco más robusto para la interpretación de los resultados.

Respecto al segundo objetivo específico, implementar técnicas de aprendizaje automático para analizar los datos de desempeño docente, seleccionando los algoritmos más adecuados que permitan una clasificación precisa y objetiva, se realizaron tres experimentos con diferentes parámetros con el propósito de obtener las métricas apropiadas en diversos contextos; se tiene el análisis para el primer experimento realizado; las métricas de clasificación obtenidas mediante validación cruzada revela patrones significativos en el rendimiento de los ocho algoritmos evaluados. En primera instancia, se observa una variabilidad notable en la relación entre los valores de entrenamiento (Train) y prueba (Test), lo cual proporciona indicios sobre la capacidad de generalización de cada modelo.

El Árbol de Decisión exhibe un rendimiento moderado, caracterizado por un AUC de 0.626 y un F1-Score de 0.591. No obstante, su baja especificidad (0.184) sugiere una limitada capacidad para identificar correctamente los casos negativos. Particularmente notable es su elevado LogLoss (6.586), lo cual indica deficiencias significativas en la calibración de las probabilidades predichas. Por su parte, Random Forest demuestra una mejora sustancial respecto al árbol simple, evidenciando la robustez característica de los métodos de conjunto. Este algoritmo alcanza valores superiores en métricas clave como AUC y F1-Score (0.65-0.61), además de presentar un Coeficiente de Correlación de Matthews (MCC) más favorable, lo que sugiere un mejor equilibrio en la clasificación binaria.

El algoritmo Gradient Boosting destaca por su rendimiento consistente, alcanzando un AUC aproximado de 0.69-0.70 y un F1-Score de 0.63-0.64. Su MCC intermedio ($\sim 0.26-0.30$) y un LogLoss comparativamente menor indican una sólida capacidad discriminativa con una penalización reducida en las estimaciones probabilísticas. La Red Neuronal presenta características particulares, con un valor de Train notablemente elevado (2.199), lo cual podría indicar cierta complejidad en la interpretación del modelo. Sin embargo, sus métricas de generalización son competitivas, con un AUC superior a 0.70 y un F1-Score que excede 0.65, aunque la disparidad entre Train y Test sugiere la presencia de un sobreajuste moderado.

La Regresión Logística, si bien no sobresale en ninguna métrica específica, mantiene un rendimiento consistente y ofrece la ventaja adicional de una mayor interpretabilidad. Sus resultados sirven como punto de referencia válido para la evaluación comparativa de los demás algoritmos. En cuanto a los algoritmos k-NN, SVM y Naive Bayes, cada uno presenta características distintivas. El k-NN alcanza un F1-Score cercano a 0.59-0.60 con un AUC moderado, mientras que SVM demuestra un rendimiento intermedio con un F1-Score de 0.58-0.59 y un AUC entre 0.63-0.65. Sorprendentemente, Naive Bayes exhibe un rendimiento competitivo con un F1-Score aproximado de 0.66 y un MCC relativamente alto (0.66-0.67), lo cual valida su aplicabilidad en este contexto específico.

El análisis de los resultados del segundo experimento obtenido mediante validación cruzada revela patrones distintivos en el comportamiento de los diversos modelos de clasificación evaluados. En particular, la Red Neuronal emerge como uno de los modelos más destacados, alcanzando un AUC de 0.731 y un F1-Score de 0.662, junto con un MCC sobresaliente de 0.665. No obstante, es importante señalar que el modelo exhibe un valor de Train inusualmente elevado (6.775), lo cual sugiere la presencia de un sobreajuste significativo, aunque sus métricas de predicción final y una prueba aproximado de 0.100 indican que el modelo mantiene una capacidad discriminativa robusta tras la validación cruzada.

El algoritmo Gradient Boosting demuestra un rendimiento sobresaliente, alcanzando un AUC de 0.719 y un F1-Score de 0.662, valores que lo posicionan entre los modelos más efectivos del experimento. Su MCC de 0.317 y un LogLoss relativamente bajo de 0.828 sugieren un equilibrio favorable entre la precisión predictiva y la calibración de probabilidades. Aunque presenta cierto grado de sobreajuste, evidenciado por un Train de 1.045, su prueba de 0.063 indica una capacidad de generalización satisfactoria. Random Forest exhibe mejoras significativas respecto al Árbol de Decisión simple, con un AUC de 0.680 y un F1-Score de 0.657, complementados por un MCC de 0.295 y un LogLoss moderado de 0.866. Particularmente notable es la similitud entre sus valores de Train (0.218) y Test (0.127), lo cual sugiere un nivel de sobreajuste controlado, consistente con la naturaleza del algoritmo que reduce la varianza mediante la combinación de múltiples árboles.

El modelo SVM alcanza un rendimiento competitivo con un AUC de 0.710 y un F1-Score de 0.660, aproximándose a los resultados obtenidos por Gradient Boosting y la Red Neuronal. Sin embargo, su MCC de 0.286 resulta comparativamente menor, sugiriendo ciertas limitaciones en su capacidad discriminativa global. Resulta particularmente interesante el desempeño del algoritmo Naive Bayes, que alcanza un AUC de 0.718 y un F1-Score de 0.654, con un LogLoss moderado de 0.627. Estos resultados son notables considerando la relativa simplicidad del modelo, sugiriendo

que las suposiciones de independencia condicional inherentes al algoritmo son razonablemente válidas para este conjunto de datos específico.

A continuación, el análisis de los resultados de tercer experimento mediante validación cruzada revela patrones significativos en el rendimiento de los algoritmos evaluados. Random Forest emerge como uno de los modelos más destacados, alcanzando un AUC aproximado de 0.84-0.85 y un F1-Score cercano a 0.77, posicionándose como uno de los algoritmos más efectivos del experimento. La diferencia contenida entre sus valores de Train (0.234) y Test (0.070) sugiere un nivel de sobreajuste controlado, mientras que su elevada Especificidad (0.680) indica una capacidad superior para clasificar correctamente los casos negativos.

El algoritmo Gradient Boosting demuestra un rendimiento sobresaliente, con un AUC aproximado de 0.82 y un F1-Score de 0.76, complementado por un LogLoss relativamente bajo (~ 0.515) que indica una calibración probabilística adecuada. Particularmente notable es su elevado MCC (~ 0.75), que evidencia un equilibrio excepcional en la clasificación de casos positivos y negativos. Sin embargo, la disparidad entre Train (1.150) y Test (0.069) sugiere cierta tendencia al sobreajuste, aunque mantiene una capacidad de generalización competitiva. La Red Neuronal exhibe un rendimiento robusto, alcanzando un AUC de 0.735 y un F1-Score de 0.734. Aunque estos valores son ligeramente inferiores a los obtenidos por Random Forest y Gradient Boosting, el modelo destaca por su elevado Recall (0.795), característica particularmente valiosa en contextos donde la minimización de falsos negativos es prioritaria. La diferencia entre Train (1.106) y Test (0.112) indica la presencia de un sobreajuste moderado.

La Regresión Logística mantiene un desempeño consistente, con un AUC aproximado de 0.79 y un F1-Score de 0.72. Sus valores de Train (0.184) y Test (0.063) reflejan un nivel de sobreajuste mínimo, mientras que su LogLoss (~ 0.548) y Especificidad (0.700) sugieren un equilibrio favorable entre la detección de verdaderos positivos y la clasificación correcta de casos negativos. El modelo Naive Bayes demuestra un rendimiento sorprendentemente competitivo, alcanzando valores de F1-Score entre 0.70-0.73 y AUC entre 0.70-0.72. Su LogLoss moderado (~ 0.583) indica

una calibración probabilística razonable, lo cual resulta notable considerando la simplicidad inherente del algoritmo.

En contraste con el estudio de Quintana (2023), quien encontró que el Árbol de Decisión fue el modelo con mejor rendimiento (77%), nuestro análisis de relevancia de atributos utilizó múltiples métricas (Information Gain, Chi-cuadrado y ReliefF) para identificar factores influyentes, proporcionando así una perspectiva más comprehensiva de las variables predictoras. Esta diferencia metodológica podría explicar las variaciones en los hallazgos principales.

Nuestros resultados sobre la importancia del grado académico (Gain = 0.027, $\chi^2 = 429.477$) encuentran cierta resonancia con el estudio de Velásquez (2022), quien identificó una asociación positiva entre las competencias docentes y la satisfacción académica. Sin embargo, nuestro estudio proporciona una cuantificación más precisa de esta relación mediante métricas específicas.

La investigación de Naranjo (2023) se centró en aspectos específicos como los materiales didácticos, encontrando una correlación alta (0.910) con la calidad educativa. Si bien nuestro estudio no evaluó directamente esta variable, nuestros hallazgos sobre la importancia de la carga no lectiva (Gain = 0.011, ReliefF = 0.114) podrían estar indirectamente relacionados con el tiempo disponible para la preparación de materiales didácticos.

Es particularmente relevante contrastar nuestros hallazgos con las observaciones de Almufarreh et al. (2023) sobre las limitaciones del aprendizaje automático en la evaluación docente. Nuestro estudio aborda parcialmente estas preocupaciones al utilizar múltiples métricas de evaluación y considerar variables contextuales como la ubicación geográfica y la adscripción facultativa, proporcionando así un marco más robusto para la interpretación de los resultados.

Finalmente, para el tercer objetivo específico, evaluar comparativamente diferentes modelos de clasificación basados en aprendizaje automático para determinar aquellos con mayor precisión, interpretabilidad y capacidad de generalización en la clasificación del desempeño docente; la comparación sistemática de los tres

experimentos revela patrones significativos en el rendimiento de los modelos seleccionados. El algoritmo Random Forest, implementado en el tercer experimento, emerge como el modelo con la capacidad predictiva más sobresaliente, alcanzando un AUC de 0.833 y un F1-Score de 0.765, valores que superan significativamente a los obtenidos por los otros modelos evaluados. Este rendimiento superior se complementa con un MCC de 0.523 y un LogLoss de 0.519, los cuales indican un equilibrio excepcional en la clasificación y una calibración probabilística óptima. Sin embargo, es importante señalar que los valores de Train (4.826) y Test (0.366) sugieren la presencia de cierto grado de sobreajuste, aspecto que requiere consideración en su implementación práctica.

En el segundo experimento, la Regresión Logística demuestra un rendimiento robusto y consistente, alcanzando un AUC de 0.734 y un F1-Score de 0.669. Particularmente notable es la estrecha diferencia entre sus valores de Train (0.162) y Test (0.055), lo cual sugiere una capacidad de generalización superior con un riesgo mínimo de sobreajuste. El modelo exhibe un MCC de 0.317 y un LogLoss de 0.601, métricas que, aunque inferiores a las de Random Forest, indican un equilibrio satisfactorio entre precisión y calibración probabilística.

El algoritmo Gradient Boosting, implementado en el primer experimento, aunque presenta métricas comparativamente más modestas con un AUC de 0.692 y un F1-Score de 0.647, mantiene un rendimiento competitivo en el contexto general de clasificación. Su MCC de 0.279 y LogLoss de 0.655 sugieren una capacidad discriminativa adecuada, aunque con oportunidades de mejora en términos de calibración probabilística. La diferencia entre Train (0.823) y Test (0.069) indica la presencia de un sobreajuste moderado.

En términos comparativos globales, los resultados evidencian una progresión clara en el rendimiento a través de los tres experimentos, con Random Forest destacándose significativamente en todas las métricas evaluadas. Sin embargo, es importante considerar que cada modelo presenta características distintivas que podrían resultar ventajosas en diferentes contextos de aplicación. La Regresión Logística, por ejemplo, ofrece un equilibrio atractivo entre rendimiento y estabilidad, mientras que

Gradient Boosting, a pesar de sus métricas más modestas, mantiene una capacidad predictiva competitiva con potencial de optimización mediante ajuste de hiperparámetros.

En contraste con el estudio de Amer (2024), que se limitó a evaluar tres modelos (regresión logística, árbol de decisiones y SVM), nuestro estudio proporcionó una evaluación más exhaustiva incluyendo ocho algoritmos diferentes y tres experimentos distintos. Esta aproximación más comprehensiva nos permitió identificar patrones de rendimiento que no fueron evidentes en el estudio de Amer, particularmente en relación con los métodos de conjunto.

Un hallazgo interesante en comparación con Quintana (2023) se relaciona con el impacto de la ingeniería de características. Mientras que Quintana encontró mejoras significativas con la introducción de variables dummy y la reducción de características, nuestro estudio se centró más en la optimización de los modelos y su capacidad de generalización. Esto se evidencia en nuestros resultados de la Regresión Logística, que mostró un rendimiento robusto ($AUC = 0.734$, $F1\text{-Score} = 0.669$) con una excelente capacidad de generalización ($\text{Train} = 0.162$, $\text{Test} = 0.055$), sin necesidad de una extensa ingeniería de características.

Conclusiones

En relación con el primer objetivo específico, orientado a identificar los atributos más relevantes en la predicción del desempeño docente, se determinó que la realización de actividades de investigación emergió como el factor más influyente, evidenciado por el valor más alto de Information Gain (0.067) y un destacado puntaje ReliefF (0.146). La adscripción a la Facultad de Ciencias de la Salud se posicionó como el segundo factor más relevante, con un Information Gain de 0.040 y un valor chi-cuadrado de 561.130, sugiriendo que las exigencias específicas de las carreras de salud podrían generar un entorno más propicio para la excelencia docente.

Respecto al segundo objetivo específico, enfocado en la evaluación de diferentes configuraciones experimentales, se encontró que la validación cruzada proporcionó resultados más robustos y consistentes en comparación con el muestreo aleatorio simple. La implementación de Bootstrap con validación cruzada estratificada de 10 particiones demostró ser particularmente efectiva, permitiendo una evaluación más comprehensiva del rendimiento de los modelos y minimizando el impacto de la variabilidad en los datos.

En cuanto al tercer objetivo específico, centrado en la comparación del rendimiento de los modelos de clasificación, Random Forest destacó significativamente, alcanzando el AUC más alto (0.833) y el mejor F1-Score (0.765) entre todos los modelos evaluados. Este rendimiento superior se complementó con un MCC de 0.523 y un LogLoss de 0.519, indicando un equilibrio óptimo en la clasificación y una calibración probabilística adecuada. La Regresión Logística también demostró un rendimiento notable, con un AUC de 0.734 y un F1-Score de 0.669, destacando particularmente por su estabilidad y menor riesgo de sobreajuste.

Recomendaciones

Se recomienda expandir la investigación mediante la incorporación de técnicas avanzadas de optimización de hiperparámetros y validación cruzada anidada. Específicamente, se sugiere implementar metodologías como Bayesian Optimization o Random Search en conjunto con validación cruzada estratificada multinivel, lo cual permitiría una exploración más exhaustiva del espacio de hiperparámetros mientras se mantiene la robustez en la evaluación del rendimiento.

Se sugiere ampliar el alcance del análisis de relevancia de atributos mediante la incorporación de técnicas de selección de características basadas en modelos (model-based feature selection) y métodos de interpretabilidad post-hoc como SHAP (SHapley Additive exPlanations) values.

Se propone desarrollar un framework metodológico estandarizado para la evaluación comparativa de modelos de predicción del desempeño docente, que incluya no solo métricas tradicionales de rendimiento como AUC y F1-Score, sino también indicadores de equidad algorítmica y robustez ante sesgos institucionales.

Referencias bibliográficas

- Achata, C. y Quispe, Y. A. (2018). *El currículo nacional y el desempeño docente*. Revista de Investigaciones, 7(2), 598-606. <https://doi.org/10.26788/riepg.v7i2.321>
- Al-Alawi, L., Al Shaqsi, J., & Tarhini, A. (2023). *Using machine learning to predict factors affecting academic performance: the case of college students on academic probation*. Education and Information Technologies. 28, 12407-12432. <https://doi.org/10.1007/s10639-023-11700-0>
- Alam, T. M., Mushtaq, M., Shaukat, K., Hameed, I. A., Umer Sarwar, M., & Luo, S. (2021). *A Novel Method for Performance Measurement of Public Educational Institutions Using Machine Learning Models*. Applied Sciences, 11(19), 9296. <https://doi.org/10.3390/app11199296>
- Almufarreh, A., Noaman, K. M., & Saeed, M. N. (2023). *Academic Teaching Quality Framework and Performance Evaluation Using Machine Learning*. Applied Sciences, 13(5), 3121. <https://doi.org/10.3390/app13053121>
- Álvarez, M., Quirós, L. M., & Cortés, M. V. (2020). *Inteligencia artificial y aprendizaje automático en medicina*. Revista Médica Sinergia, 5(8), e557. <https://doi.org/10.31434/rms.v5i8.557>
- Amer, A. A. S. (2024). *Predictive Model for Student's Academic Performance Using Machine Learning Techniques*. International Journal For Research In Applied Science And Engineering Technology, 12(7), 966-975. <https://doi.org/10.22214/ijraset.2024.63659>
- Aranda, M. T. F., & Llontop, L. R. A. (2023). *Política de evaluación y su implicancia en la labor docente de una comunidad educativa peruana*. Ensaio: Avaliação E Políticas Públicas Em Educação, 31(121), e0234017. <https://doi.org/10.1590/S0104-40362023003104017>
- Arroyo, M. P. (2023). *Políticas públicas educativas sobre evaluación de desempeño: Tensión y realidad*. Revista Dialogus, (11), 107–118. <https://doi.org/10.37594/dialogus.v1i11.786>

- Baffa, N. M. H., Miyim, N. M. A., & Dauda, N. A. S. (2023). *Machine Learning for Predicting Students' Employability*. UMYU Scientifica, 2(1), 001-009. https://doi.org/10.56919/usci.2123_001
- Bazán, A., & Velarde, N. M. (2021). *Auto-reporte del estudiantado en criterios de desempeño didáctico en clases de Psicología*. Journal of Behavior, Health & Social Issues, 13(1), 22–35. <https://doi.org/10.22201/fesi.20070780e.2021.13.1.78071>
- Beltrán, M. P. H., Arias, C. L. E. V., Morales, S. E., Vázquez, N. T. L., & González, B. S. (2021). *Seguimiento a resultados de evaluación del desempeño docente*. Ciencia Latina Revista Científica Multidisciplinar, 5(6), 13508-13524. https://doi.org/10.37811/cl_rcm.v5i6.1339
- Bravo, L. E. C., Nieves-Pimiento, N., & Gonzalez-Guerrero, K. (2022). *Prediction of University-Level Academic Performance through Machine Learning Mechanisms and Supervised Methods*. Ingeniería, 28(1), e19514. <https://doi.org/10.14483/23448393.19514>
- Buendía, A. R., Pérez, C., & Chilet, S. E. (2023). *El uso de imágenes fractales y la Estrategia integral en la estimulación cognitiva de niños de 3 a 7 años: preparación para el aprendizaje automático*. FitoVida, 2(2), 14–24. <https://doi.org/10.56275/fitovida.v2i2.23>
- Chanta, L. F., & Clavo, P. J. (2023). *Desempeño docente en la educación básica: Una mirada sistemática*. Revista de Climatología, 23, 280-289. <https://doi.org/10.59427/rcli/2023/v23cs.280-289>
- Cipagauta-Moyano, M. E. (2019). *La evaluación docente en educación superior: características y desafíos*. Entorno, (68), 105–110. <https://doi.org/10.5377/entorno.v0i68.8460>
- Dávila, R. C., Castillo, R. A., Vargas, A. R., Velarde, L., García, E., García, C. F., Pasquel, R. F., & Guanilo, C. E. (2023). *Application of Machine Learning Models in Fraud Detection in Financial Transactions*. Data and Metadata, 2, 109. <https://doi.org/10.56294/dm2023109>

- Delgado, Z. M. (2022). *Modelo de gestión académica para disminuir el riesgo de deserción estudiantil en una universidad estatal, Cantón Jipijapa-Ecuador 2022*. (Tesis de doctorado, Universidad César Vallejo). <https://repositorio.ucv.edu.pe/handle/20.500.12692/94901>
- De Los Heros, M. G., & Solana-Villanueva, N. (2021). *La evaluación al docente de Educación Superior durante la pandemia de COVID-19*. *EDU REVIEW*. International Education and Learning Review Revista Internacional De Educación Y Aprendizaje, 9(3), 241–254. <https://doi.org/10.37467/gkarevedu.v9.2924>
- Gallo, A., Pérez, F., & Salinas, D. (2021). *Minería de datos y proyección a corto plazo de la demanda de potencia en el sistema eléctrico ecuatoriano*. *Revista Técnica Energía*, 18(1), 72-85. <https://doi.org/10.37116/revistaenergia.v18.n1.2021.461>
- Gálvez, E., y Milla, R. (2018). *Evaluación del desempeño docente: Preparación para el aprendizaje de los estudiantes en el Marco de Buen Desempeño Docente*. *Propósitos Y Representaciones*, 6(2), 407–452. <https://doi.org/10.20511/pyr2018.v6n2.236>
- García-López, Y., González-Sáez, L. Y., & Cabrera-Hernández, J. A. (2022). *Aplicaciones de aprendizaje automático para el análisis industrial de la provisión azucarera en Matanzas, Cuba*. *Revista U.D.C.A Actualidad & Divulgación Científica*, 25(2). <https://doi.org/10.31910/rudca.v25.n2.2022.2334>
- Gómez, L. F., y Valdés, M. G. (2019). *La evaluación del desempeño docente en la educación superior*. *Propósitos Y Representaciones*, 7(2), 479–515. <https://doi.org/10.20511/pyr2019.v7n2.255>
- Guo, J., Bai, L., Yu, Z., Zhao, Z., & Wan, B. (2021). *An AI-Application-Oriented In-Class Teaching Evaluation Model by Using Statistical Modeling and Ensemble Learning*. *Sensors*, 21(1), 241. <https://doi.org/10.3390/s21010241>

- Guanga, U. R., Carolina, A., Lozada, R. F., Reinoso, M. del C., & Paz, R. B. (2024). *Desafíos de la Educación para la Implementación de la Inteligencia Artificial*. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 3588-3602. https://doi.org/10.37811/cl_rcm.v8i3.11576
- Guzmán, N. del R., Álvarez, N. F., & Pacheco, X. O. (2024). *Alcances y limitaciones de la IA en educación*. *RECIMUNDO*, 8(1), 215–223. [https://doi.org/10.26820/recimundo/8.\(1\).ene.2024.215-223](https://doi.org/10.26820/recimundo/8.(1).ene.2024.215-223)
- Kalaiselvi, G., Akshata, Hema, S. M., & Iswarya, M. (2022). *Enhancing the Learning Experience of Students by Early Prediction of Student Academic Performance using Machine Learning*. *International Journal of Advanced Research in Science Communication and Technology*, 2(1), 935–940. <https://doi.org/10.48175/ijarsct-4755>
- López, S. E., Pérez, J. A., Flores, A., & Juárez, S. (2023). *Algoritmos de clasificación para la detección de obesidad en adolescentes: Un estudio comparativo entre KNN y árboles de decisión*. *Revista de Investigación en Tecnologías de la Información*, 11(23), 70–81. <https://doi.org/10.36825/RITI.11.23.007>
- Maussa, E. y Hernández, F. (2018). *Evaluación del desempeño del docente universitario y construcción de una cultura evaluativa integral*. *Enfoque Latinoamericano*, 1(2), 33-46. <https://core.ac.uk/reader/268097540>
- Meenakshi, Sanchez, D. T., Peconillo, L. B., Jr, De Vera, J. V., Mahajan, R., Kumar, T., & Bhosle, A. A. (2022). *Machine Learning Techniques for Quality Management in Teaching Learning Process in Higher Education by Predicting the Student's Academic Performance*. *International Journal of Next-Generation Computing*, 13(3). <https://doi.org/10.47164/ijngc.v13i3.837>
- Naranjo, L. Y. (2023). *Factores de desempeño docente que inciden en la calidad educativa de una Universidad Pública, Ecuador 2022*. (Tesis de maestría, Universidad César Vallejo). <https://repositorio.ucv.edu.pe/handle/20.500.12692/115903>

- Noaman, K. M., Almufarreh, A., & Saleh, M. (2023). *AI-Based Course Management Systems in Universities*. Applied Sciences. <https://doi.org/10.3390/app13053121>
- Núñez-Michuy, C. M., Agualongo-Chela, L. M., Vistin, J. M., & López, M. (2023). *La Inteligencia Artificial en la pedagogía como modelo de enseñanza*. Magazine de las Ciencias: Revista de Investigación e Innovación, 8(2), 120–135. <https://doi.org/10.33262/rmc.v8i1.2932>
- Ogundele, I. M., Taiwo, O., Babalola, A. E. & Ayeni, O. C. (2024). *Prediction of Student Academic Performance Based on Machine Learning Model*. International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2024, pp. 1-11. <https://doi.org/10.1109/seb4sdg60871.2024.10629703>
- Pacol, C. (2024). *Sentiment Analysis of Students' Feedback on Faculty Performance*. WSEAS Transactions on Information Science and Applications. <https://doi.org/10.37394/23209.2024.21.7>
- Pacheco, M., Ibarra, I., Iñiguez, M. E., Lee, H. & Sánchez, C. V. (2018). *La evaluación del desempeño docente en la educación superior*. Revista Digital Universitaria (RDU). 19(6). <http://doi.org/10.22201/codeic.16076079e.2018.v19n6.a2>
- Pasquinelli, M., Cafassi, E., Monti, C., Peckaitis, H., & Zarauza, G. (2022). *Cómo una máquina aprende y falla – una gramática del error para la inteligencia artificial*. Hipertextos, 10(17), 13-29. <https://doi.org/10.24215/23143924e054>
- Peceros, B. (2014). *Desempeño docente y su relación con autodeterminación, autoeficacia y orientación a la meta: Un estudio con docentes de Religión de Lima*. Studium Veritatis, 12(18), 327–382. <https://doi.org/10.35626/sv.18.2014.51>
- Polo, V. J. (2024). *Modelo predictivo basado en Machine Learning Supervised y la deserción estudiantil en centros de Educación Superior Tecnológicos Públicos de la región la Libertad*. (Tesis de doctorado, Universidad Nacional del Santa). <https://repositorio.uns.edu.pe/handle/20.500.14278/4737>

- Polo-Triana, S., Ramírez-Sierra, Y., Arias-Osorio, J., Martínez-Vega, R., & Lamos-Díaz, H. (2022). *Métodos de aprendizaje automático para predecir el comportamiento epidemiológico de enfermedades arbovirales: revisión estructurada de literatura*. *Revista De La Universidad Industrial De Santander Salud*, 55(1). <https://doi.org/10.18273/saluduis.55.e:23017>
- Quintana, J. O. (2023). *Análisis predictivo de la deserción estudiantil en los alumnos de la Universidad Nacional de Moquegua entre 2009 y 2019, usando Machine Learning*. (Tesis de doctorado, Universidad Nacional del Altiplano). <https://repositorio.unap.edu.pe/handle/20.500.14082/19742>
- Quiroz, N. H., Posadas-Martínez, M. L., Rossi, E., Giunta, D. H., & Risk, M. R. (2021). *Aprendizaje automático aplicado en área de la salud. Parte I*. *Revista Del Hospital Italiano De Buenos Aires*, 41(4), p. 206–209. <https://doi.org/10.51987/revhospitalbaire.v41i4.131>
- Revilla-Mendoza, J. & Palacios-Jimenez, A. (2020). *Ser docente en tiempos de pandemia por covid-19: evaluación del desempeño docente en una universidad pública de lima*. *Ágora Revista Científica*, 7(2), 58-62. <https://doi.org/10.21679/arc.v7i2.197>
- Reyes, M. E. (2023). *Propiedades psicométricas del cuestionario de evaluación de la competencia docente para estudiantes de escuelas técnico profesional policial*. (Tesis de doctorado, Universidad Cesar Vallejo). <https://repositorio.ucv.edu.pe/handle/20.500.12692/130624>
- Rico, A., & Gaytán, N. D. (2022). *Modelos predictivos del rendimiento académico a partir de características de estudiantes de ingeniería*. *IE Revista De Investigación Educativa De La REDIECH*, 13, e1426. https://doi.org/10.33010/ie_rie_rediech.v13i0.1426
- Rodríguez-Horna, E., & Herrera-Vásquez, C. (2023). *Relación entre clima organizacional, nivel de desempeño y motivación docente*. *Dékau Perú*, 1(1), 1–9. <https://doi.org/10.55996/dekape.v1i1.175>

- Rodríguez-Tovar, K. L., Gutiérrez-Portela, F., & Hernández-Aros, L. (2023). *Análisis del uso de técnicas supervisadas de aprendizaje automático y profundo en la detección de fraude financiero*. *Revista Tecnología En Marcha*, 36(8), Pág. 50–56. <https://doi.org/10.18845/tm.v36i8.6927>
- Salem S, Al-Habashneh O, & Lasassmeh O. (2021) *Data mining techniques for classifying and predicting Teachers' performance based on their evaluation reports*. *Indian Journal of Science and Technology* .14(2): 119-130. <https://doi.org/10.17485/IJST/v14i2.2149>
- Salazar, M. R. (2019). *Experiencias y aprendizajes de la implementación de estrategias didácticas en educación virtual*. *Revista Científica Internacional*, 2(1), 1–9. <https://doi.org/10.46734/revcientifica.v1i1.1>
- Sánchez, E. C., & Chávarry, P. (2022). *Plan de motivación herzbergiana y desempeño laboral del docente universitario*. *Horizonte Empresarial*, 9(1), 42-57. <https://doi.org/10.26495/rce.v9i1.2182>
- Sánchez-Holgado, P., Martín-Merino, M., & Blanco, D. (2020). *Del data-driven al data-feeling: análisis de sentimiento en tiempo real de mensajes en español sobre divulgación científica usando técnicas de aprendizaje automático*. *Anuario Electrónico De Estudios En Comunicación Social "Disertaciones"*, 13(1), 35–58. <https://doi.org/10.12804/revistas.urosario.edu.co/disertaciones/a.7691>
- Sánchez-Huete, J. C., Yaegashi, S. F. R., Marquès-Donoso, A., & Glatz, E. T. M. de M. (2022). *Parámetros del desarrollo profesional docente*. *Imagens Da Educação* , 12(4), 01-27. <https://doi.org/10.4025/imagenseduc.v12i4.62694>
- Soria, L., Ortega, W., & Ortega, A. (2020). *Desempeño pedagógico docente y aprendizaje de los estudiantes universitarios en la carrera de Educación*. *Praxis & Saber*, 11(27), e10329. <https://doi.org/10.19053/22160159.v11.n27.2020.10329>

- Taiwo, O., Ogundele, I. M., & et al. (2024). *Predictive Models for Academic Success in Higher Education*. IEEE Conference Proceedings. <https://doi.org/10.1109/seb4sdg60871.2024.10629703>
- Tan, C. (2022). *Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data*. Journal of Function Spaces, 2022, 1–8. <https://doi.org/10.1155/2022/7064287>
- Tarhini, A., Al-Alawi, L., & et al. (2023). *Determinants of Effective Learning in Higher Education*. Education and Information Technologies. <https://doi.org/10.1007/s10639-023-11700-0>
- Urteaga, I., Siri, L., & Garófalo, G. (2020). *Predicción temprana de deserción mediante aprendizaje automático en cursos profesionales en línea*. RIED-Revista Iberoamericana de Educación a Distancia, 23(2), 147–167. <https://doi.org/10.5944/ried.23.2.26356>
- Velásquez, M. A. (2022). *Evaluación del desempeño docente y satisfacción académica en estudiantes de una universidad privada de Lima*. (Tesis de doctorado, Universidad Femenina del Sagrado Corazón). <https://repositorio.unife.edu.pe/items/40f3c08b-b747-482f-a68e-491a16a30804>
- Vera, J. A. (2021). *Estrés laboral y desempeño en docentes de Psicología Clínica de la Universidad Técnica de Babahoyo*. Revista Ecuatoriana De Psicología, 4(9), 143–151. <https://doi.org/10.33996/repsi.v4i9.59>
- Xiongjun, X., & Lv, D. (2022). *The Evaluation of Music Teaching in Colleges and Universities Based on Machine Learning*. Journal of Mathematics, 2022, 1–7. <https://doi.org/10.1155/2022/2678303>
- Zaidan, A. O. (2022). *Development of a classification algorithm based on an academic data model using hyperparametric tuning for hiring teachers at iraqi universities*. T-Comm, 16(9), 33–38. <https://doi.org/10.36724/2072-8735-2022-16-9-33-38>

- Zamora, E. (2021). *La evaluación del desempeño docente mediante cuestionarios en la universidad: Su legitimidad según la literatura y los requerimientos para que sea efectiva*. *Actualidades Investigativas En Educación*, 21(3), 1–23. <https://doi.org/10.15517/aie.v21i3.46221>
- Zhang, X., & Kang, Y. (2021). *Examining and Predicting Teacher Professional Development by Machine Learning Methods*. In *Communications in computer and information science* (pp. 255–269). https://doi.org/10.1007/978-981-16-5188-5_19
- Zhong, Y. (2023). *Evaluation and analysis of teaching quality of university teachers using machine learning algorithms*. *Journal of Intelligent Systems*, 32(1), 20220204. <https://doi.org/10.1515/jisys-2022-0204>

Anexo y apéndices

Anexo 01: Matriz de operacionalización de variables

Variable	Definición		Dimensión	Indicador	Ítems	Escala
	Conceptual	Operacional				
Desempeño docente	Calidad y eficacia con la que los profesores desarrollan su labor educativa, abarcando planificación, interacción y evaluación en el ámbito universitario (Morales, 2022).	Medida a través de registros documentales de desempeño docente, considerando factores pedagógicos, metodológicos y de formación profesional. Evaluado mediante una ficha estructurada validada por expertos.	Planificación	Sílabos, guías, material educativo	Sílabos, guías, material educativo	Escala ordinal
			Enseñanza	Metodologías, plataformas virtuales, instrumentos de evaluación	Metodologías, plataformas virtuales, instrumentos de evaluación	
			Cumplimiento	Asistencias, atenciones, entrega de documentos	Asistencias, atenciones, entrega de documentos	
			Responsabilidad	Tutoría, investigación, comisiones	Tutoría, investigación, comisiones	
			Formación académica y profesional	Cursos especialización, diplomados, programas especialización	Cursos especialización, diplomados, programas especialización	
			Investigación	Trabajos de investigación, asesorías y jurados	Trabajos de investigación, asesorías y jurados	
Factores asociados	Aspectos personales, académicos y de formación profesional que influyen en la eficacia del desempeño docente (Llovio et al., 2023).	Analizados mediante registros y encuestas que incluyen variables como experiencia docente, nivel académico, habilidades pedagógicas y uso de metodologías innovadoras.	Factores personales	Experiencia docente en años	Experiencia docente en años	Escala dicotómica y politómica
			Factores académicos	Nivel académico alcanzado	Nivel académico alcanzado	
			Factores metodológicos	Frecuencia en el uso de metodologías activas	Frecuencia en el uso de metodologías activas	

Anexo 02: Matriz de consistencia

Problema	Variables	Objetivos	Hipótesis	Metodología
¿Cómo desarrollar un modelo de aprendizaje automático para la clasificación del desempeño docente en una universidad de Chimbote en el año 2024?	Desempeño docente Factores asociados	<p>Objetivo General</p> <p>Desarrollar un modelo de aprendizaje automático para clasificar el desempeño docente en una universidad de Chimbote durante el año 2024, evaluando su precisión, interpretabilidad y capacidad de generalización.</p>	<p>Hipótesis General</p> <p>El desarrollo de un modelo de aprendizaje automático permite clasificar de manera objetiva y precisa el desempeño docente en una universidad de Chimbote, optimizando los procesos evaluativos y promoviendo la mejora continua de la calidad educativa.</p>	<p>Tipo de investigación: Aplicada, ya que busca resolver un problema práctico mediante el uso de herramientas tecnológicas avanzadas.</p> <p>Diseño: No experimental transversal, pues no se manipularán variables y los datos se recolectarán en un único momento temporal.</p> <p>Población y muestra: La población comprende 800 registros de desempeño docente en una universidad de Chimbote.</p> <p>Técnicas e instrumentos de recolección de datos: Técnica de análisis documental utilizando una ficha de recolección de datos</p>
		<p>Objetivos Específicos</p> <ol style="list-style-type: none"> 1. Identificar los factores pedagógicos, metodológicos y de formación profesional que influyen significativamente en el desempeño docente en una universidad de Chimbote durante el año 2024. 2. Implementar técnicas de aprendizaje automático para analizar los datos de desempeño docente, seleccionando los algoritmos más adecuados que permitan una clasificación precisa y objetiva. 3. Evaluar comparativamente diferentes modelos de clasificación basados en aprendizaje automático para determinar aquellos con mayor precisión, interpretabilidad y capacidad de generalización. 	<p>Hipótesis específicas</p> <ol style="list-style-type: none"> 1. Los factores personales, académicos y de formación profesional de los docentes en una universidad de Chimbote durante el período 2024 se asocian significativamente con su desempeño docente. 2. La implementación de técnicas de aprendizaje automático en el análisis de datos de desempeño docente en una universidad de Chimbote permite seleccionar algoritmos que clasifican con alta precisión el desempeño docente. 3. La evaluación comparativa de diferentes modelos de clasificación basados en aprendizaje automático identifica aquellos con mayor precisión, interpretabilidad y capacidad de generalización. 	

Anexo 3. Formato de publicación en repositorio



REPOSITORIO INSTITUCIONAL DIGITAL
FORMULARIO DE AUTORIZACIÓN PARA LA PUBLICACIÓN DE DOCUMENTOS DE INVESTIGACIÓN


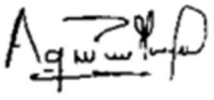
1. Información del Autor			
Aquino Gonzales, Aracelli Corina		41944062	corina.aquino@usapedro.edu.pe
Apellidos y Nombres		DNI	Correo Electrónico
2. Tipo de Documento de Investigación			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Testis	Trabajo de Suficiencia Profesional	Trabajo Académico	Trabajo de Investigación
3. Grado Académico o Título Profesional ¹			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Bachiller	Título Profesional	Título Segunda Especialidad	Maestría
<input type="checkbox"/> Doctorado			
4. Título del Documento de Investigación			
Modelo de aprendizaje automático para clasificación del desempeño docente en una universidad, Chimbote, 2024			
5. Programa Académico			
MAESTRIA EN INGENIERIA INFORMATICA Y DE SISTEMAS CON MENCIÓN EN GESTIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIONES			
6. Tipo de Acceso al Documento			
<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Abierto o Público ² (info-uso-repo/semantica/openAccess)	Acceso restringido ³ (info-uso-repo/semantica/restringidoAccess) (*)		(*)
(*) En caso de restringido sustentar motivo			

A. Originalidad del Archivo Digital

Por el presente dejo constancia que el archivo digital que entrego a la Universidad, es la versión final del trabajo de Investigación sustentado y aprobado por el Jurado Evaluador y forma parte del proceso que conduce a obtener el grado académico o título profesional.

B. Otorgamiento de una licencia CREATIVE COMMONS ⁴

C. I autor, por medio de este documento, autoriza a la Universidad, publicar su trabajo de Investigación en formato digital en el Repositorio Institucional Digital, al cual se podrá acceder, preservar y difundir de forma libre y gratuita, de manera íntegra a todo el documento. ⁵

Huella Digital		Firma	

Lugar	Día	Mes	Año
Chimbote	19	03	2024

1. Resolución de Consejo Directivo N°001-2018-UNZU-CO, Reglamento del Registro Nacional de Trabajos de Investigación para optar Grados Académicos y Títulos Profesionales, Art. 2, inciso 2.2.
2. Ley N°27444 que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto y D.S. 008-2018-PCM.
3. Si el autor elige el tipo de acceso abierto o público, otorga a la Universidad en Pedro una licencia no exclusiva, para que se pueda hacer amplia de forma en la obra y difundir en el Repositorio Institucional Digital. Respetando siempre los derechos de Autor y Propiedad Intelectual de acuerdo y en el Marco de la Ley N°11.
4. En caso de que el autor elija la segunda opción, únicamente se publicará los datos del autor y resumen de la obra, de acuerdo a la directiva M004-2018-CONCYTEC-0000 (Numerales 1.2 y 1.7) que norma el funcionamiento del Repositorio Nacional Digital.
5. Las licencias Creative Commons (CC) es una organización internacional sin fines de lucro que promueve la aplicación de los autores un conjunto de licencias flexibles y de herramientas tecnológicas que facilitan la difusión de información, recursos educativos, obras artísticas y científicas, entre otros. Estas licencias también garantizan que el autor obtenga crédito por su obra.
6. Según el inciso 1.2.2 del artículo 1.º del Reglamento del Registro Nacional de Trabajos de Investigación para optar grados académicos y títulos profesionales (RDNAT) las universidades, instituciones y escuelas de educación superior tienen como obligación registrar todos los trabajos de investigación y proyectos, incluyendo los realizados en sus repositorios institucionales prestando el servicio de acceso abierto o restringido, los cuales serán posteriormente indexados por el Repositorio Digital (RDNAT), a través del Repositorio ALICIA.

Nota: - En caso de falsedad en los datos, se procederá de acuerdo a ley (Ley 27444, art. 33, párr. 32.3).

Anexo 4. Reporte de similitud

Modelo de aprendizaje automático para clasificación del desempeño docente en una universidad, Chimbote, 2024

INFORME DE ORIGINALIDAD



FUENTES PRIMARIAS

1	hdl.handle.net Fuente de Internet	2%
2	repositorio.ucv.edu.pe Fuente de Internet	2%
3	www.coursehero.com Fuente de Internet	1%
4	doaj.org Fuente de Internet	<1%
5	repositorio.uns.edu.pe Fuente de Internet	<1%
6	repositorio.unife.edu.pe Fuente de Internet	<1%
7	www.researchgate.net Fuente de Internet	<1%
8	repositorio.uwiener.edu.pe Fuente de Internet	<1%
9	www.slideshare.net Fuente de Internet	

		<1 %
10	www.tdx.cat Fuente de Internet	<1 %
11	repositorio.unap.edu.pe Fuente de Internet	<1 %
12	es.scribd.com Fuente de Internet	<1 %
13	publicar.claec.org Fuente de Internet	<1 %
14	Submitted to Universidad Internacional de la Rioja Trabajo del estudiante	<1 %
15	repositorio.uct.edu.pe Fuente de Internet	<1 %
16	repositorio.unc.edu.pe Fuente de Internet	<1 %
17	repositorio.uta.edu.ec Fuente de Internet	<1 %
18	Submitted to Universidad Rey Juan Carlos Trabajo del estudiante	<1 %
19	repositorio.ug.edu.ec Fuente de Internet	<1 %
20	Submitted to Universidad Politécnica del Perú Trabajo del estudiante	

		<1 %
21	publicaciones.usanpedro.edu.pe Fuente de Internet	<1 %
22	Submitted to Corporación Universitaria Minuto de Dios, UNIMINUTO Trabajo del estudiante	<1 %
23	Submitted to ITESM: Instituto Tecnológico y de Estudios Superiores de Monterrey Trabajo del estudiante	<1 %
24	archive.org Fuente de Internet	<1 %
25	repositorio.ugto.mx Fuente de Internet	<1 %
26	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	<1 %
27	Submitted to Trabajo del estudiante	<1 %
28	repositorio.uss.edu.pe Fuente de Internet	<1 %
29	edoc.pub Fuente de Internet	<1 %
30	pesquisa.bvsalud.org Fuente de Internet	<1 %

		<1 %
21	publicaciones.usanpedro.edu.pe Fuente de Internet	<1 %
22	Submitted to Corporación Universitaria Minuto de Dios, UNIMINUTO Trabajo del estudiante	<1 %
23	Submitted to ITESM: Instituto Tecnológico y de Estudios Superiores de Monterrey Trabajo del estudiante	<1 %
24	archive.org Fuente de Internet	<1 %
25	repositorio.ugto.mx Fuente de Internet	<1 %
26	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	<1 %
27	Submitted to Trabajo del estudiante	<1 %
28	repositorio.uss.edu.pe Fuente de Internet	<1 %
29	edoc.pub Fuente de Internet	<1 %
30	pesquisa.bvsalud.org Fuente de Internet	<1 %

31	repositorio.cidecuador.org Fuente de Internet	<1 %
32	repositorio.unemi.edu.ec Fuente de Internet	<1 %
33	digitum.um.es Fuente de Internet	<1 %
34	worldwidescience.org Fuente de Internet	<1 %
35	www.dykinson.com Fuente de Internet	<1 %
36	www.ulv.edu.ec Fuente de Internet	<1 %
37	asele2020.unileon.es Fuente de Internet	<1 %
38	Submitted to Instituto Superior Tecnológico Espíritu Santo Trabajo del estudiante	<1 %
39	Submitted to Universidad Tecnológica del Peru Trabajo del estudiante	<1 %
40	Submitted to Escuela Superior Politécnica del Litoral Trabajo del estudiante	<1 %
41	scite.ai Fuente de Internet	<1 %

42	Submitted to Pontificia Universidad Javeriana Cali Trabajo del estudiante	<1 %
43	Submitted to Universidad Distrital FJDC Trabajo del estudiante	<1 %
44	Submitted to institutoeuropeodeposgrado Trabajo del estudiante	<1 %
45	ojs.tdea.edu.co Fuente de Internet	<1 %
46	repositorio.unal.edu.co Fuente de Internet	<1 %
47	repositorio.unheval.edu.pe Fuente de Internet	<1 %
48	repositorio.upsc.edu.pe Fuente de Internet	<1 %
49	www.ebizlatam.com Fuente de Internet	<1 %
50	ojs.docentes20.com Fuente de Internet	<1 %
51	www.openhr.cloud Fuente de Internet	<1 %
52	academica-e.unavarra.es Fuente de Internet	<1 %
53	ciencialatina.org Fuente de Internet	<1 %

		<1 %
54	docplayer.es Fuente de Internet	<1 %
55	gisweb.ciat.cgiar.org Fuente de Internet	<1 %
56	jalayo.blogspot.com Fuente de Internet	<1 %
57	revistasnicaragua.cnu.edu.ni Fuente de Internet	<1 %
58	terra.ar2.bumeran.com Fuente de Internet	<1 %
59	www.diveo.net Fuente de Internet	<1 %
60	www.netfilter.com.br Fuente de Internet	<1 %
61	www.revistasocialfronteriza.com Fuente de Internet	<1 %
62	www.scribd.com Fuente de Internet	<1 %
63	candidaturas.cespu.pt Fuente de Internet	<1 %
64	doi.org Fuente de Internet	<1 %

65	masters.franquicias-negocios.com Fuente de Internet	<1 %
66	memoriascimted.com Fuente de Internet	<1 %
67	prezi.com Fuente de Internet	<1 %
68	reliefweb.int Fuente de Internet	<1 %
69	repositorio.iberopuebla.mx Fuente de Internet	<1 %
70	repositorio.uladech.edu.pe Fuente de Internet	<1 %
71	repositorio.unprg.edu.pe Fuente de Internet	<1 %
72	repositorio.usmp.edu.pe Fuente de Internet	<1 %
73	revista.ister.edu.ec Fuente de Internet	<1 %
74	revistahorizontes.org Fuente de Internet	<1 %
75	revistas.upel.edu.ve Fuente de Internet	<1 %
76	scielo.senescyt.gob.ec Fuente de Internet	<1 %

77	www.cirugiaycirujanos.com Fuente de Internet	<1 %
78	www.colegiodeprofesores.cl Fuente de Internet	<1 %
79	www.cruce.iteso.mx Fuente de Internet	<1 %
80	www.fepafem.org.ve Fuente de Internet	<1 %
81	www.lareferencia.info Fuente de Internet	<1 %
82	www.memfod.edu.uy Fuente de Internet	<1 %
83	www.scielo.org.pe Fuente de Internet	<1 %
84	www.suneo.mx Fuente de Internet	<1 %
85	www2.eco.uva.es Fuente de Internet	<1 %
86	catalonica.bnc.cat Fuente de Internet	<1 %
87	digibug.ugr.es Fuente de Internet	<1 %
88	dspace.unl.edu.ec Fuente de Internet	<1 %

89	editorial.inudi.edu.pe Fuente de Internet	<1 %
90	edsociety.iberojournals.com Fuente de Internet	<1 %
91	medicalresearchjournal.org Fuente de Internet	<1 %
92	openworks.wooster.edu Fuente de Internet	<1 %
93	polodelconocimiento.com Fuente de Internet	<1 %
94	repositorio.unfv.edu.pe Fuente de Internet	<1 %
95	revistas.uam.es Fuente de Internet	<1 %
96	revistas.uis.edu.co Fuente de Internet	<1 %
97	revistas.unid.edu.pe Fuente de Internet	<1 %
98	ridum.umanizales.edu.co Fuente de Internet	<1 %
99	scielo.sld.cu Fuente de Internet	<1 %
100	upepe.sep.gob.mx Fuente de Internet	<1 %

101	www.ehu.eus Fuente de Internet	<1 %
102	www.grafiati.com Fuente de Internet	<1 %
103	www.icesse.ase.ro Fuente de Internet	<1 %
104	www.medicadepot.com Fuente de Internet	<1 %
105	www.uco.es Fuente de Internet	<1 %
106	www.dropbox.com Fuente de Internet	<1 %

Excluir citas Apagado Excluir coincidencias < 6 words
 Excluir bibliografía Activo