Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2025,
Volumen 9, Número 4.

https://doi.org/10.37811/cl_rcm.v9i2

DETERMINACIÓN DE UN MODELO MEDIANTE APRENDIZAJE AUTOMÁTICO PARA OBTENER LA IDONEIDAD DEL PERSONAL NAVAL PROFESIONAL

DETERMINING A MODEL USING MACHINE LEARNING TO DETERMINE THE SUITABILITY OF PROFESSIONAL NAVAL PERSONNEL

Danny Marcelo Vivanco Toala
Universidad Americana de Europa, México

Rosa Gabriela Camero Berrones
Universidad Americana de Europa, México

DOI: https://doi.org/10.37811/cl_rcm.v9i4.19442

Determinación de un Modelo mediante Aprendizaje Automático para obtener la Idoneidad del Personal Naval Profesional

Danny Marcelo Vivanco Toala[1]
vivanco.espol101@gmail.com
https://orcid.org/0009-0005-0785-971X
Universidad Americana de Europa
México

Rosa Gabriela Camero Berrones
gabriela.camero@gmail.com
https://orcid.org/0000-0003-4438-1645
Universidad Americana de Europa
México

RESUMEN

La falta de automatización del proceso para la selección del personal naval profesional para ocupar puestos sensibles en la institución militar ha contribuido a aumentar el grado de subjetividad en los evaluadores. El objetivo de este trabajo de investigación es entrenar y evaluar modelos de predicción utilizando datos obtenidos de los resultados de las evaluaciones realizadas al personal naval profesional desde el año 2017 hasta agosto del 2024. Se realizó un muestreo no probabilístico por conveniencia. El enfoque metodológico de esta investigación es cuantitativa con un diseño cuasiexperimental y un alcance de investigación correlacional. Este estudio emplea tres algoritmos de aprendizaje automático supervisado de clasificación para la predicción: Árboles de Decisión, Naives Bayes Bernoulli y K-Nearest Neighbors. Los resultados indican que la evaluación de la métrica exactitud (accuracy) para los tres algoritmos antes indicados dieron como resultado un valor superior al 90% y con la métrica AUC obtuvieron un valor superior al 0,9. Se puede concluir que, entre los tres modelos entrenados y evaluados, el algoritmo Árboles de Decisión es el mejor modelo a implementar, logrando la mejor exactitud del 99,81%, demostrando un buen rendimiento y la aplicabilidad para este trabajo de investigación.

Palabras clave: algoritmos, aprendizaje automático, clasificación, idoneidad, modelos

[1] Autor principal.
Correspondencia: vivanco.espol101@gmail.com

Determining a Model using Machine Learning to Determine the Suitability of Professional Naval Personnel

ABSTRACT

The lack of automation of the process for the selection of professional naval personnel to fill sensitive positions in the military institution has contributed to increase the degree of subjectivity in the evaluators. The objective of this research work is to train and evaluate prediction models using data obtained from the results of evaluations conducted on professional naval personnel from 2017 to August 2024. A non-probabilistic convenience sampling was performed. The methodological approach of this research is quantitative with a quasi-experimental design and a correlational research scope. This study employs three supervised machine learning classification algorithms for prediction: Decision Trees, Naives Bayes Bernoulli and K-Nearest Neighbors. The results indicate that the evaluation of the accuracy metric for the three algorithms indicated above resulted in a value higher than 90% and with the AUC metric they obtained a value higher than 0,9. It can be concluded that, among the three models trained and evaluated, the Decision Trees algorithm is the best model to implement, achieving the best accuracy of 99,81%, demonstrating good performance and applicability for this research work.

***Keywords*:** algorithms, classification, suitability, machine learning, models

Artículo recibido 21 julio 2025
Aceptado para publicación: 26 agosto 2025

INTRODUCCIÓN

En el campo de la inteligencia artificial, el empleo de algoritmos de aprendizaje automático se ha incrementado en las últimas dos décadas, ya que se ha convertido en una herramienta para la extracción de información de grandes conjuntos de datos y para la construcción de modelos utilizados en aplicaciones o sistemas, obteniendo el mejor resultado en base al análisis de datos predictivo realizado, esto contrasta en la actualidad con la falta de automatización de procesos críticos en las organizaciones, lo que provoca designar a través de procesos manuales a personas no calificadas para ocupar puestos catalogados como sensibles o estratégicos (Salamanca y Castro, 2020).

La aportación principal de este trabajo radica en la aplicación de técnicas de aprendizaje automático utilizando algoritmos supervisados de clasificación, para la selección rápida y oportuna de empleados de una organización que requieran ocupar puestos sensibles sin la necesidad de realizar procesos manuales que demanden demasiado recurso humano y tiempo a emplear.

Consecuentemente, este artículo de investigación se estructura en temas que abordan desde una revisión exhaustiva de la literatura de tres tipos de algoritmos de aprendizaje automático supervisado de clasificación, donde, se estableció como herramienta la matriz de confusión binaria, la cual permite evaluar el desempeño de un algoritmo supervisado de clasificación a través de sus métricas de confusión.

Asimismo, se describirá en detalle la metodología utilizada para el diseño y la determinación del mejor modelo obtenido de la evaluación de los algoritmos de aprendizaje supervisado de clasificación a través de sus métricas. Finalmente, se discutirá la aplicabilidad del mejor modelo seleccionado en base al análisis de los resultados de sus métricas. Este enfoque integral asegura que el estudio realizado no solo contribuye significativamente al campo de la inteligencia artificial, sino que también ofrece soluciones prácticas a problemas urgentes de aprendizaje automático aplicado a empresas.

Problemática y preguntas de investigación

Este artículo aborda la problemática, en cuanto a la falta de automatización del proceso para la selección del personal naval profesional para ocupar puestos sensibles en la institución militar, la cual ha contribuido a aumentar el grado de subjetividad en los evaluadores; es por ello, que al no contar con algoritmos de aprendizaje automático que coadyuven a predecir de manera correcta, rápida y oportuna

la idoneidad o no idoneidad de un postulante, pueda que algún candidato no idóneo sea considerado a un puesto sensible; y, en consecuencia a esto, ocurra algún incidente como fuga de información u otro suceso que ponga en riesgo la seguridad de las operaciones militares.

Este trabajo de investigación tiene la siguiente pregunta de investigación: ¿Qué tipo de algoritmo de aprendizaje automático y que métricas se deben emplear para analizar el desempeño y obtener los resultados deseados del algoritmo seleccionado?.

Justificación y relevancia

En cuando a la justificación de este artículo de investigación, es de carácter institucional y aportará a la institución militar en el ámbito tecnológico de la inteligencia artificial aplicando aprendizaje automático, utilizando como herramienta la matriz de confusión binaria para validar y generar los datos para obtener el mejor modelo a implementar.

Es así que, la importancia de realizar esta investigación para la institución militar, es porque permitirá mejorar el proceso de selección del personal y así determinar la idoneidad o no idoneidad del personal naval que requieren ostentar puestos sensibles, a fin de que no pongan en riesgo la seguridad de las operaciones militares, y así poder tomar las medidas correctivas antes de que ocurra algún incidente.

Así también, este artículo contribuirá a la comunidad científica en cuanto al planteamiento de una nueva forma de utilizar la inteligencia artificial para el análisis de la información y el método para determinar el mejor modelo e identificar si los empleados de una empresa son idóneos o no idóneos para la ocupación de puestos sensibles en la organización.

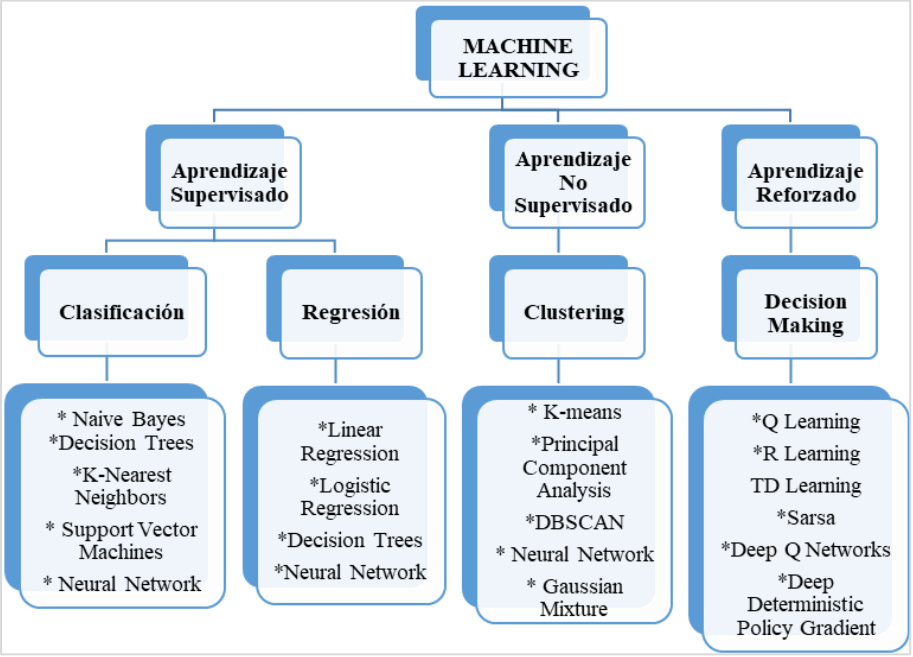
Fundamentación Teórica

La determinación de un modelo predictivo para obtener la idoneidad del personal naval profesional mediante aprendizaje automático, se fundamenta en la teoria de capacidad de aprendizaje de Valiant (1984), el cual indica que el aprendizaje computacional consiste en un modelo teórico llamado PAC (Probably Approximately Correct), que evalua si un algoritmo puede aprender correctamente funciones o conceptos utilizando un número finito de ejemplos y en un tiempo computacional razonable, sin tener la necesidad de una programación explícita, lo cual es clave para la automatización de procesos de selección a partir de datos históricos, es así, que esta teoría se articula con los principios de computabilidad interpuestos por Turing (1936), quien demostró que los números computables pueden

ser ejecutados por máquinas automáticas o llamadas máquinas de Turing bajo reestricciones de operaciones finitas para producir soluciones válidas, los mismos que son considerados como un modelo teórico para describir procesos algorítmicos, estableciendo así el marco de los algoritmos que sustentan los modelos de aprendizaje automático a emplearse.

En consecuencia, el aprendizaje automático se define como el campo de estudio que da a las computadoras la habilidad para aprender sin estar explícitamente programadas; es decir, el aprendizaje automático se usa para enseñar a las máquinas cómo manejar los datos más eficientemente, de igual forma se aplica, cuando después de ver los datos, no se pueden interpretar de forma correcta, dicho de otro modo, se usa el aprendizaje automático cuando se quiere aprender más de los datos para obtener un mayor conjunto de información de gran relevancia. Del mismo modo, el aprendizaje automático se puede clasificar en tres tipos de aprendizaje los cuales son: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje reforzado (Ali *et al.*, 2023; Sindayigaya y Dey, 2022) , donde el enfoque es en el algoritmo supervisado de tipo clasificación, según como se detallan en la figura 1 a continuación.

Figura 1 Tipos de algoritmos de aprendizaje de automático

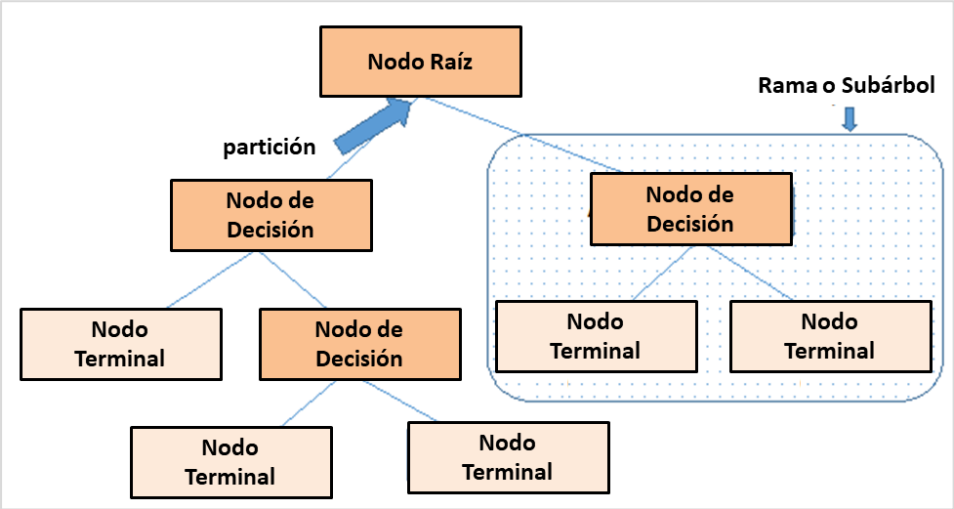


Fuente: Elaboración propia adaptado de Ali *et al.* (2023); Sindayigaya y Dey (2022)

Para este artículo se van a contextualizar tres algoritmos de tipo clasificación, los cuales son: Árboles de Decisión, Naive Bayes y K-Nearest Neighbors. Primeramente se va a enunciar, el algoritmo de árboles de decisión, el cual es un modelo predictivo, que consiste en estructuras en forma de árbol,

donde el algoritmo que utiliza para armar los árboles se denomina “partición binaria recursiva” en la cual cada nodo interno representa una pregunta sobre una característica de los datos y cada rama representa una posible respuesta a la pregunta realizada del nodo, y es así como se denomina partición binaria, de tal manera que se sigue con este proceso en forma recursiva hasta un cierto punto previamente estipulado en el que el proceso de bifurcación se detiene, teniendo a los nodos hojas del árbol con la predicción o clasificación final (Arana, 2021; Contreras *et al.*, 2024), de acuerdo a la figura 2 detallada a continuación.

Figura 2 Estructura de un árbol de decisión para clasificación

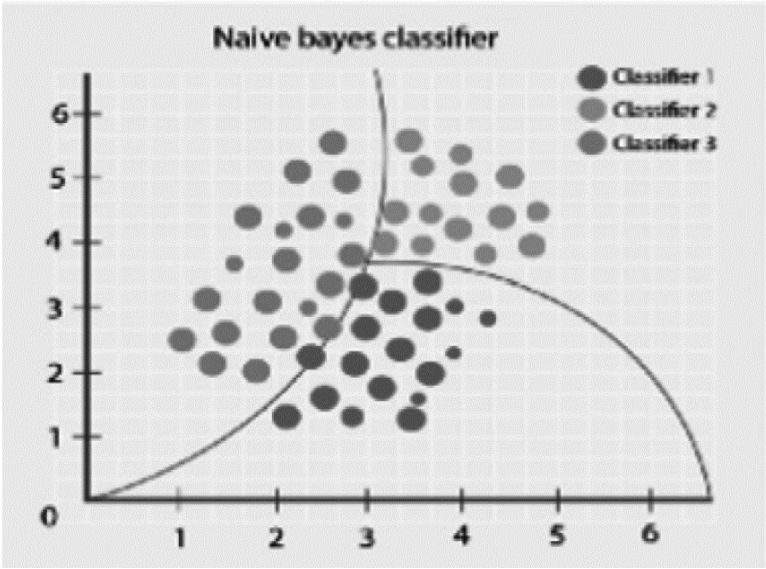


Fuente: Adaptado de Arana (2021)

En efecto, los algoritmos más representativo para árboles de decisión son el algoritmo CART por sus siglas en inglés (Classification and Regression Trees) que es para clasificación y regresión, y el algoritmo ID3 por sus siglas en inglés (Iterative Dichotomiser 3) que es un decotomizador iteractivo, donde la diferencia de estos dos algoritmos es que ID3 no aplica ningún tipo de poda en el árbol y solo funciona para variables de clasificación, mientras que CART ademas de trabajar con variables de clasificación también trabaja con variables continuas para los árboles de regresión, es así que, ID3 emplea la métrica de Information Gain o Ganancia de Información y CART emplea la métrica llamada Gini Index o Índice de Gini, donde al aplicar dichas métricas se puede evaluar y seleccionar las mejores divisiones o atributos al construir un árbol de decisión, ya que estas métricas te permiten encontrar el atributo que mejor separe las clases, construyendo un árbol mas eficiente y preciso (Lu *et al.*, 2022; Somvanshi *et al.*, 2016).

Por otro lado, el algoritmo Naive Bayes pertenece a la familia de clasificadores probabilísticos simples, soportado en el teorema de Bayes, el cual es una herramienta de aprendizaje automático empleada para la clasificación, el cual funciona calculando la probabilidad de que un dato pertenezca a una clase determinada, dada ciertas caracteristicas, donde se asume que estas caracteristicas son independientes entre sí, tomando como ejemplo la detección de "spam" o "no spam" en los correos electrónicos, basándose en las palabras claves que orientan cuando un correo puede ser "spam", en el cual Naive Bayes calcula la probabilidad de que un correo sea "spam" o "no spam", según las palabras que contenga el correo (Contreras *et al.*, 2024; Gupta *et al.*, 2022).

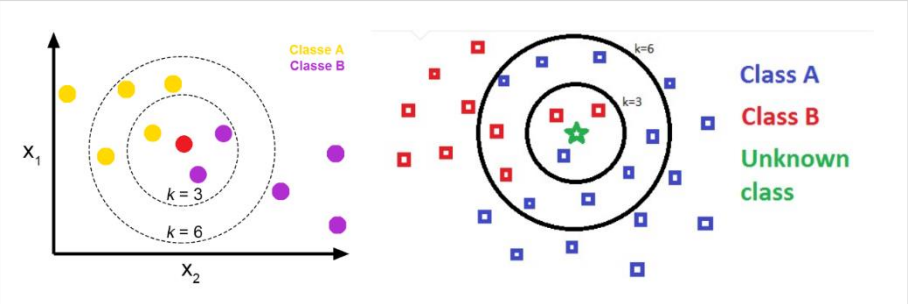
Figura 3 Estructura de un árbol de decisión para clasificación



Fuente: Adaptado de Gupta *et al.* (2022)

Del mismo modo, el algoritmo k-Vecinos Más Cercanos (KNN), es un algoritmo de clasificación no paramétrico, utilizado tanto para la clasificación como regresión, donde un punto de datos se clasifica según la mayoría de los votos de sus vecinos más cercanos, es decir asigna a un punto de datos no etiquetado la clase del conjunto de datos más cercano previamente eqtiquetados. El valor apropiado de K y la mejor opción de distancia se determina mediante validación cruzada para que el algoritmo sea mas eficiente, donde la caracteristica de este algoritmo es buscar a partir de un conjunto de datos de entrenamiento las K instancias más similares o los vecinos a un valor de salida determinado, asimismo, predice a que grupo pertenece el dato no etiquetado dependiendo de la forma de los datos más cercanos (Contreras *et al.*, 2024; Singh *et al.*, 2016), de acuerdo a la figura 3 presentada a continuación.

Figura 4 Proceso del algoritmo KNN



Fuente: Adaptado de Contreras *et al.* (2024)

Por otra parte, la Matriz de Confusión Binaria, es una herramienta que permite visualizar el desempeño de un algoritmo de aprendizaje supervisado, la cual se trata de una tabla de frecuencias donde las filas pertenecen a la clase predicha y las columnas a la clase real, representando así el número de predicción de cada clase mutuamente excluyentes; es decir, se puede observar que tipo de aciertos o errores está teniendo nuestro modelo a la hora de pasar por el proceso de aprendizaje con los datos. Es así que, a estos cuatro resultados posibles se los conoce como matriz de confusión: Verdadero Positivo, Verdadero Negativo, Falso Positivo y Falso Negativo, en la cual se ejemplifica el caso de una clasificación binaria, donde los verdaderos positivo y negativo (f11 y f00) corresponden al número de instancias correctamente clasificadas en cada clase, mientras que los totales F y f representa las frecuencias marginales correspondientes a: F1T – F0T es igual a la cantidad total de elementos en cada clase, f1T – f0T es el total de instancias clasificadas por el algoritmo como positivo y negativo en el caso binario y finalmente N representa el tamaño muestral utilizado para la clasificación (Borja-Robalino *et al.*, 2020; Chamorro, 2020), de acuerdo a la tabla 1 expuesta.

Tabla 1 Matriz de confusión binaria para variables dicotómicas

		Clasificador 1 "Real"		
Clasificador 2 "Predicción"		**Clase 1: Positivo**	**Clase 0: Negativo**	**TOTAL**
	Clase 1: Positivo	f11= Verdadero Positivo	f10= Falso Negativo	F1T
	Clase 0: Negativo	f01= Falso Positivo	f00= Verdadero Negativo	F0T
		f1T	f0T	N

Fuente: Elaboración propia adaptado de Borja-Robalino *et al.* (2020)

Del mismo modo, a partir de lo explicado en la matriz de confusión binaria se presentan las siguientes métricas de evaluación que permiten cuantificar la calidad del modelo: Exactitud (en su significado en inglés Accuracy), Precisión (en su significado en iglés Precision), Sensibilidad (en su significado en inglés Recall), Especificidad (en su significado en inglés Specificity) o también llamada tasa negativa verdadera, el Puntaje F1 (en su significado en inglés F1-Score) y la curva ROC–AUC (Borja-Robalino *et al.*, 2020; Contreras *et al.*, 2024; Romero, 2025).

Exactitud, es la métrica más utilizada por investigadores en variables dicotomicas y multiclase, donde se define como la cantidad de predicciones positivas que son correctas sobre el número total de predicciones realizadas. Esta métrica suele expresarse como un porcentaje, oscilando entre 0 % (rendimiento nulo) y 100 % (máxima precisión posible). Un modelo alcanza una exactitud de 1.0 cuando el número de predicciones acertadas coincide completamente con el número de etiquetas reales. Esta medida resulta especialmente útil cuando se trabaja con conjuntos de datos balanceados, es decir, cuando todas las clases tienen una representación equitativa en el problema de clasificación. Sin embargo, su utilidad disminuye en contextos donde existe un fuerte desequilibrio de clases, como en casos en los que una categoría predomina ampliamente sobre las demás. En tales escenarios, es recomendable complementar el análisis con otras métricas como la sensibilidad, la especificidad, la precisión o el área bajo la curva ROC-AUC, las cuales ofrecen una evaluación más integral del desempeño del modelo, la cual es representada por la siguiente formula.

$$\boldsymbol{Exactitud} = \frac{\mathrm{VP} + \mathrm{VN}}{\mathrm{VP} + \mathrm{FP} + \mathrm{FN} + \mathrm{VN}}$$

Precisión, es una métrica fundamental en la evaluación de modelos de clasificación, ya que representa la proporción de predicciones verdaderamente positivas entre todos los casos identificados como positivos por el modelo. Esta medida refleja la capacidad del sistema para evitar clasificar erróneamente casos negativos como positivos. Su utilidad es especialmente relevante en contextos donde los falsos positivos tienen consecuencias significativas, como en la detección de spam o en diagnósticos médicos, donde se espera que los positivos predichos sean realmente correctos. Por ejemplo, en un sistema antispam, la precisión indica qué porcentaje de correos marcados como spam lo son efectivamente; en medicina, señala qué parte de los pacientes diagnosticados con una enfermedad realmente la padecen.

No obstante, la precisión por sí sola puede ser insuficiente, particularmente cuando los datos están desbalanceados. En estos casos, es necesario complementarla con métricas como la sensibilidad, la especificidad o la puntuación F1, lo que permite una evaluación más integral del desempeño del modelo, y se la calcula de la siguiente manera y se la representa por la siguiente fórmula.

$$\boldsymbol{Precisión} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FP}}$$

Sensibilidad, representa la proporción de verdaderos positivos correctamente identificados por un modelo con respecto al total de casos positivos reales. Esta medida resulta óptima cuando su valor se aproxima a 1, equivalente al 100 %, lo cual indica una identificación completa de las instancias positivas. Su aplicación es recomendada especialmente en contextos donde es prioritario minimizar los falsos negativos, como sucede en entornos médicos. En pruebas de detección de enfermedades como el cáncer, por ejemplo, esta métrica permite estimar qué porcentaje de pacientes enfermos fueron identificados correctamente. La sensibilidad es esencial cuando no detectar un caso positivo podría derivar en consecuencias críticas, ya que un valor bajo podría implicar que el modelo omite diagnósticos importantes, lo que pone en riesgo la salud del paciente, y se la calcula de la siguiente manera.

$$\boldsymbol{Sensibilidad} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FN}}$$

Especificidad, también conocida como tasa verdadera negativa, se define como el cociente entre las predicciones negativas correctas y el total de verdaderos negativos en un conjunto de datos. Esta métrica es especialmente útil en contextos donde se busca reducir la incidencia de falsos positivos, es decir, aquellos casos en los que el modelo clasifica erróneamente una instancia negativa como positiva. Su aplicación es frecuente en ámbitos médicos y de diagnóstico, donde es fundamental asegurar que los individuos sanos sean correctamente identificados como tales. Una baja especificidad podría derivar en diagnósticos erróneos y, en consecuencia, en la administración de tratamientos innecesarios, lo que puede implicar riesgos éticos, económicos y de salud para los pacientes, y es presentada por la siguiente formula.

$$\boldsymbol{Especificidad} = \frac{\text{VN}}{\text{VP} + \text{FN}}$$

Índice F1 Score, es el promedio ponderado entre precisión y la sensibilidad, ofreciendo una medida balanceada del rendimiento del modelo cuando ambas métricas son relevantes. Esta puntuación integra tanto los falsos positivos como los falsos negativos en su cálculo, lo que la convierte en una herramienta especialmente adecuada para escenarios de clasificación con clases desbalanceadas, donde la distribución de ejemplos entre las categorías no es uniforme. Por ello, el F1-score es ampliamente empleado para evaluar modelos en contextos donde el equilibrio entre exactitud y cobertura es crítico, y se la representa de la siguiente manera.

$$\boldsymbol{F1} = 2 * \frac{Precisión * Sesibilidad}{Precisión + Sensibilidad}$$

Curva ROC-AUC, es el área bajo la curva ROC (Receiver Operating Characteristic), también denominada AUC, constituye una de las métricas más relevantes para evaluar el rendimiento de un modelo de clasificación. Esta métrica cuantifica la capacidad del modelo para diferenciar correctamente entre clases positivas y negativas, reflejando su habilidad para establecer distinciones claras entre distintas categorías. Por ejemplo, permite evaluar si el sistema puede discriminar con precisión entre pacientes con una condición crítica y aquellos que no la presentan. Un valor de AUC igual a 1.0 indica un desempeño perfecto del modelo, mientras que un valor de 0.5 sugiere un rendimiento equivalente al azar, sin capacidad real de clasificación

Estado del arte: investigaciones recientes

Las investigaciones recientes sobre el empleo de los algoritmos de aprendizaje automático supervisado tanto para la toma de decisiones en los procesos de contratación en una organización como para la detección de ataques distribuidos de denegación de servicio, se presentan a través de los siguientes autores.

El artículo de investigación de Adillah *et al.* (2025) se basa en que los empleados son los activos más importantes de la organización, para lo cual la toma de decisiones en los procesos de contratación son cruciales para el éxito a largo plazo de la organización, donde decisiones incorrectas pueden generar altos costos en los procesos de recontratación y disminución en la productividad, lo que mantiene una

similitud con este trabajo de investigación, en cuanto al proceso de evaluación para determinar la idoneidad del personal naval profesional para ocupar puestos sensibles, es así que, en indicada investigación se desarrollo un modelo de predicción de decisiones de reclutamiento utilizando datos de los resultados finales del reclutamiento 2024 en el ministerio de finanzas, con los siguientes atributos: nivel educativo, la edad, el promedio de calificaciones (GPA), la puntuación SKD y la puntuación SKB, donde se emplearon cinco algoritmos de aprendizaje automático para la predicción: Máquina de Vectores de Soporte Lineal, Árbol de Decisión (C5.0), Bosque Aleatorio, k-Vecino Más Cercano (k-NN) y Clasificador Naive Bayes, donde entre los cinco modelos probados, Naive Bayes demostró ser el más efectivo, logrando una exactitud del 88% y un AUC de 0.97, demostrando su sólido rendimiento en la distinción de clases positivas y negativas; donde se puede tomar en consideración algunos de estos algoritmos con dichas métricas para la obtención del mejor modelo para la evaluación de la idoneidad del personal naval. Así también, se puede decir que, los factores claves que contribuyen al éxito del modelo de dicho estudio son la selección de características relevantes, la calidad de los datos y las técnicas adecuadas para el entrenamiento y validación de los algoritmos.

En el estudio de Salau *et al.* (2023), se desarrolló un modelo predictivo basado en el algoritmo Bernoulli Naive Bayes para la detección de ataques distribuidos de denegación de servicio (DDoS), logrando una exactitud del 85.99 % al ser evaluado con métricas como la exactitud y curva ROC-AUC. Aunque el dominio de aplicación fue la ciberseguridad, los hallazgos destacan la viabilidad del uso de técnicas supervisadas de clasificación binaria en contextos de alta sensibilidad operativa. Esta línea de trabajo puede extrapolarse al campo de la defensa, particularmente en el análisis automatizado de la idoneidad del personal naval profesional, donde la exactitud y la curva ROC-AUC del modelo son críticas para evitar falsos positivos o negativos. Por tanto, la implementación de modelos de clasificación similares permitiría no solo discriminar entre perfiles idóneos y no idóneos, sino también optimizar el proceso de toma de decisiones en instituciones militares, contribuyendo a la eficiencia operativa mediante sistemas inteligentes y adaptativos de evaluación de recursos humanos.

Hipótesis

La determinación de un modelo mediante la evaluación de algoritmos de aprendizaje automático supervisados de clasificación, determinará la idoneidad del personal naval profesional para ocupar

puestos sensibles con una métrica exactitud (accuracy) mayor al 90% y una métrica AUC-ROC mayor al 0,9.

Objetivos de la investigación

Objetivo general

- Determinar un modelo que permita obtener la idoneidad del personal naval para el cumplimiento de funciones en puestos sensibles asignados mediante aprendizaje automático.

Objetivos específicos

- Estructurar y depurar los datos a través de las variables que determinan la idoneidad y no idoneidad del personal naval para cumplir funciones o tareas en puestos sensibles.
- Entrenar los algoritmos de aprendizaje automático con los nuevos datos de las variables que determinan la idoneidad del personal naval.
- Evaluar los modelos que se obtengan del entrenamiento de los algoritmos de machine learning.

METODOLOGÍA

El enfoque metodológico a aplicarse en esta investigación es cuantitativa, basada en la metodología de Hernández-Sampieri *et al.* (2014), ya que permite la medición objetiva de variables relacionadas con la idoneidad del personal naval profesional y la aplicación de algoritmos de aprendizaje automático. Así también, el diseño de investigación a emplear es cuasiexperimental, debido a que responde a la determinación de un modelo predictivo en un contexto institucional real, sin la asignación aleatoria de los postulantes, pero comparando los resultados antes y después de la intervención tecnológica. El alcance de investigación de este articulo es correlacional, ya que busca identificar relaciones causa efecto entre las variables predictoras seleccionadas y verificar los niveles de idoneidad del personal evaluado, sin establecer inicialmente causalidad directa, asimismo, responde a un diseño de intervención institucional tecnológica, dado que la investigación no se limitó al análisis teórico, sino que propone un modelo automatizado basado en inteligencia artificial para automatizar los procesos para la evaluación del persona naval profesional para poder ostentar puestos sensibles, permitiendo observar sus efectos en la toma de decisiones estratégicas (Creswell y Creswell, 2017).

La muestra tomada para esta investigación es de 2.651 registros desde el año 2017 hasta agosto del 2024, la cual es la data oficial del personal naval que fue evaluado para obtener la idoneidad o no

idoneidad para ocupar puestos sensibles en sus repartos o unidades militares, emitidos por la institución militar, con la cual se realizó un tipo de muestreo no probabilístico por conveniencia.

El instrumento de investigación que se utilizó es una hoja electrónica de datos en Excel con un total de 22 variables de estudio, las cuales tienen que ver con los ámbitos de: información militar publicada en redes sociales, deudas con entidades financieras y la institución militar, faltas disciplinariarias, información judicial por narcotráfico y asociación ilícita, antecedentes penales, impedimento de salida del país, responsabilidad en procesos de investigación dentro de la institución, impedimento legal para ejercer cargos públicos, examen de conocimiento para tratamiento de documentación militar, autorización para verificación de datos personales recopilados de los ámbitos de las variables antes descritas y autorización para realizar prueba de confianza.

Es así que, se seleccionaron 05 variables, las cuales son: no alcanzar nota mínima en examen de conocimiento por dos ocasiones, registra que no autoriza verificar datos personales, registra que no autoriza realizarse prueba de confianza, registra falta de pérdida de documentación militar por negligencia y registra falta por responsabilidad en procesos de investigación dentro de la institución, las mismas que fueron evaluadas como de mayor importancia por los algoritmos seleccionados y se explicará su evaluación en los resultados obtenidos.

En cuanto al método de validación del instrumento se firmó un “Acuerdo de confidencialidad y no divulgación de información manejada por los servidores públicos del Ministerio de Defensa Nacional, Fuerzas Armadas y entidades adscritas”, con la Dirección de Inteligencia de la Institución Militar para entrega de los datos, ya que es información oficial y calificada.

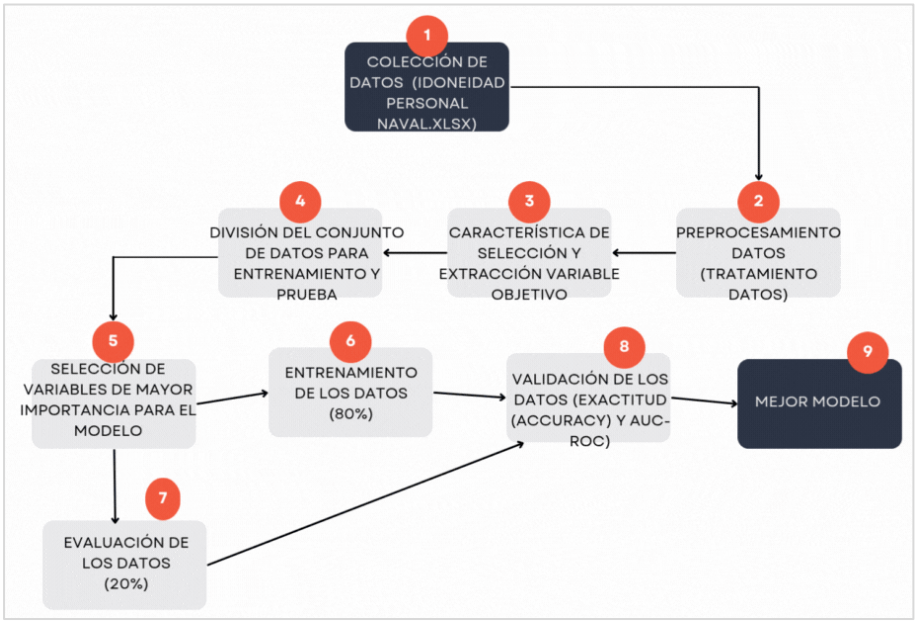
Del mismo modo, los algoritmos de aprendizaje automático necesitan aprender de los datos y para su aplicabilidad, se empleó la herramienta Google Colab, la cual es un servicio gratuito alojado en la nube, que permite escribir y ejecutar código abierto “Python” directamente del navegador, sin la necesidad de instalar o realizar alguna configuración en un computador, asimismo, dicha herramienta permite integrarse con aplicaciones web e incorporar las estadísticas de análisis de los datos en una base de datos de producción, así también, se emplearon las librerías Scikit-Learn, Statsmodels, PyMC y NTLK, Pandas, Numpy, Scipy y Matplotlib para aplicación de los algoritmos de aprendizaje automático supervisados de clasificación los cuales son: Árboles de Decisión, Naive Bayes y K-Nearest Neighbors.

Es así que, el procedimiento de aplicación metodológica, inicia con la data oficial obtenida, la misma que mantiene los registros iniciales que se cargarán en los algoritmos seleccionados para ser entrenados y evaluados, ya que al tratarse de datos reales, se aplicó un proceso de validación, limpieza y preprocesamiento para asegurar la calidad de los datos para su posterior análisis.

En cuanto, al procedimiento validación de los datos se incluyó la detección y tratamiento de valores nulos, la verificación de la coherencia lógica entre campos, la revisión de tipos de datos y la eliminación de registros duplicados. Adicionalmente, se identificaron y trataron valores atípicos que pudieran afectar el desempeño de los modelos predictivos.

Posteriormente, se realizó un procesamiento que consistió en la codificación de variables categóricas, la normalización de variables numéricas, transformación de fechas en variables temporales, dividir el conjunto de datos en características y la variable objetivo o resultado, con este tratamiento se preparó al conjunto de datos de entrenamiento (train) con el 80% de la data y de evaluación (test) con el 20% de la data para su aplicación en los algoritmos de aprendizaje automático seleccionados, y una vez evaluados los modelos con las variables de importancia seleccionadas, se procede a realizar una evaluación de los resultados obtenidos con las métricas Exactitud (Accuracy) y ROC-AUC, para la determinación del mejor modelo a emplear, de acuerdo a la figura 5 del proceso propuesto presentado a continuacion:

Figura 5 Proceso propuesto para obtener el mejor modelo de algoritmo supervisado de clasificación



Fuente: Elaboración propia

RESULTADOS Y DISCUSIÓN

Los resultados encontrados van en concordancia con los objetivos y metodologia planteada en este artículo de investigación, donde para dar respuesta con el objetivos específicos de entrenamiento y evaluación de los algoritmos de aprendizaje automático con los datos de las variables que determinan la idoneidad del personal naval, se analizó los tipos de algoritmos de aprendizaje automático supervisado de clasificación, donde de acuerdo a la problemática del tema planteado y al estado del arte investigado es pertinente aplicar los algoritmos Árboles de Decisión, Naive Bayes Bernoulli y K-Nearest Neighbors.

Para la selección de las variables que contribuyen al mejor modelo de los algoritmos, se seleccionó y se entrenó el algoritmo de árboles de decisión con las 05 variables que más registros mantienen con personal evaluado como No Idóneo de la matriz de datos del registro oficial, ya que es el algoritmo que mantiene la función para evaluar las características de importancia de las variables, pudiéndose determinar con el valor de ganancia, que no existieron variables menores al 1%, es decir las 05 variables son importantes, de acuerdo a las variables y figura 6 que se detallan a continuación.

Variable 1.- reg_nota_min_exam_conoc_dos_opor

Variable 2.- reg_no_autoriza_verificar_datos_DHP

Variable 3.- reg_no_autoriza_prue_confianza_DHP

Variable 4.- reg_falt_negli_talento_humano

Variable 5.- reg_resp_investigacion_inspectoria

Figura 6 Valor de ganancia de las 05 variables seleccionadas empleadas en los modelos

Fuente: Elaboración propia

Algoritmo de clasificación Árboles de Decisión

Se realizó el entrenamiento y evaluación del algoritmo, tomando las 05 variables anteriormente seleccionadas, teniendo como resultado la obtención de los valores de las siguientes métricas: Accuracy

Test con un valor del 99,81%, el cual ha aprendido bien de su conjunto de datos del Accuracy Train, asimismo, la métrica AUC mantiene un valor de 0,98; es decir, que el modelo clasifica perfectamente todas las instancias y tiene un rendimiento excelente y por último una diferencia del Gini Train que tiene un valor del 0,9608 y Gini Test con un valor de 0,9600 el cual da una diferencia del del 0,08%, obteniendo un buen porcentaje casi de 0% en la diferencia de Gini, porque cuando existe un diferencia de Gini en positivo y si se pasa del 10% se debe obligatoriamente realizar un sobreajuste al modelo para mejorar el Gini Test, el cual indica que es un buen modelo por los valores obtenidos en sus métricas, de acuerdo a la figura 7 y tabla 2 presentadas a continuación.

Figura 7 Obtención de las métricas del algoritmo Árboles de Decisión para entrenamieno y prueba

Fuente: Elaboración propia

Tabla 2 Valor de las métricas del algoritmo Árboles de Decisión

Componente	Accuracy	Precision	Recall	Specificity	F1-Score	AUC	Gini
Entrenamiento (Train)	0,9981	0,9980	1,0000	0,9608	0,9990	0,9804	0,9608
Evaluación (Test)	0,9981	0,9980	1,0000	0,9600	0,9990	0,9800	0,9600

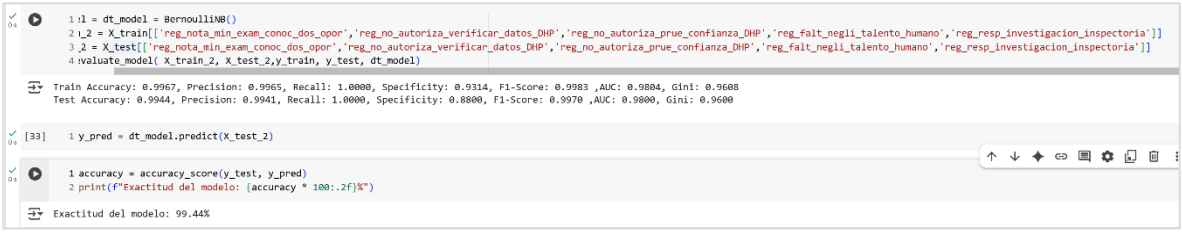
Fuente: Elaboración propia

Algoritmo de clasificación Naive Bayes Bernoulli

Se realizó el entrenamiento y evaluación del algoritmo, tomando las 05 variables anteriormente seleccionadas, teniendo como resultado la obtención de los valores de las siguientes métricas: Accuracy Test con un valor del 99,44%, el cual ha aprendido bien de su conjunto de datos del Accuracy Train, asimismo, la métrica AUC mantiene un valor de 0,98; es decir, que el modelo clasifica perfectamente todas las instancias y tiene un rendimiento excelente y por último una diferencia del Gini Train que tiene un valor del 0,9608 y Gini Test con un valor de 0,9600 el cual da una diferencia del del 0,08%, obteniendo un buen porcentaje casi del 0% en la diferencia de Gini, el cual indica que es un buen modelo

por los valores obtenidos en sus métricas, de acuerdo a la figura 8 y tabla 3 presentadas a continuación.

Figura 8 Obtención de las métricas del algoritmo Naive Bayes Bernoulli para entrenamieno y prueba



Fuente: Elaboración propia

Tabla 3 Valor de las métricas del algoritmo Naive Bayes Bernoulli

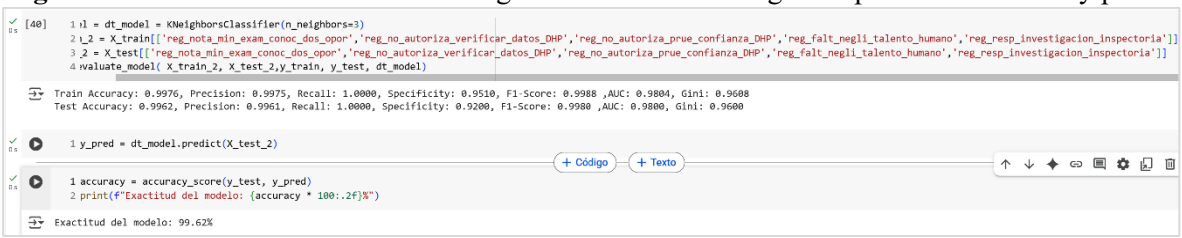
Componente	Accuracy	Precision	Recall	Specificity	F1-Score	AUC	Gini
Entrenamiento (Train)	0,9967	0,9965	1,0000	0,9314	0,9983	0,9804	0,9608
Evaluación (Test)	0,9944	0,9941	1,0000	0,8800	0,9970	0,9800	0,9600

Fuente: Elaboración propia

Algoritmo de clasificación K-Nearest Neighbors

Se realizó el entrenamiento y evaluación del algoritmo, tomando las 05 variables anteriormente seleccionadas, teniendo como resultado la obtención de los valores de las siguientes métricas: Accuracy Test con un valor del 99,62%, el cual ha aprendido bien de su conjunto de datos del Accuracy Train, asimismo, la métrica AUC mantiene un valor de 0,98; es decir, que el modelo clasifica perfectamente todas las instancias y tiene un rendimiento excelente y por último una diferencia del Gini Train que tiene un valor del 0,9608 y Gini Test con un valor de 0,9600 el cual da una diferencia del del 0,08%, obteniendo un buen porcentaje casi de 0% en la diferencia de Gini, el cual indica que es un buen modelo por los valores obtenidos en sus métricas, de acuerdo a la figura 9 y tabla 4 presentadas a continuación.

Figura 9 Obtención de las métricas del algoritmo K-Nearest Neighbors para entrenamieno y prueba



Fuente: Elaboración propia

Tabla 4 Valor de las métricas del algoritmo K-Nearest Neighbors

Componente	Accuracy	Precision	Recall	Specificity	F1-Score	AUC	Gini
Entrenamiento (Train)	0,9976	0,9975	1,0000	0,9510	0,9988	0,9804	0,9608
Evaluación (Test)	0,9962	0,9961	1,0000	0,9200	0,9988	0,9800	0,9600

Fuente: Elaboración propia

Evaluación de las métricas de los algoritmos de evaluación seleccionados

Una vez obtenidas las métricas de los 03 algoritmos de clasificación antes evaluados, se puede concluir que, los tres modelos son buenos ya que cumplen con los valores planteados en la hipótesis en cuanto a las métricas de Exactitud y AUC; sin embargo, el algoritmo Árboles de Decisión demostró ser el más efectivo, es decir el mejor modelo a implementar, logrando la mejor exactitud del 99,81% y un AUC de 0,98, exponiendo un buen rendimiento en la distinción de clases positivas y negativas, teniendo como factores claves que contribuyen al éxito del modelo la selección de características relevantes, la calidad de los datos y las técnicas adecuadas para el entrenamiento y validación de los algoritmos, de acuerdo a la tabla 5 expuesta a continuación.

Tabla 5 Selección del mejor modelo a implementar dados por las métricas Accuracy y AUC

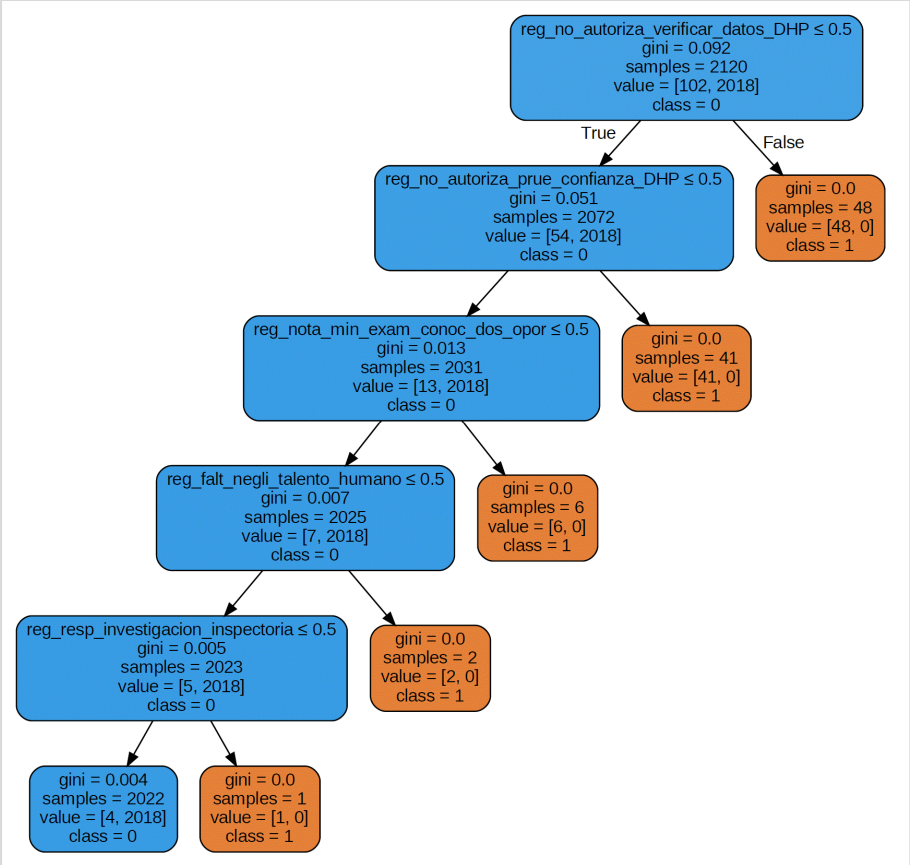
Algoritmos	Accuracy	AUC
Árboles de Decisión	0,9981	0,9804
Naive Bayes Bernoulli	0,9944	0,9800
K-Nearest Neighbors	0,9962	0,9800

Fuente: Elaboración propia

Dado que el algoritmo de árbol de decisión es el mejor modelo a implementar, el árbol generado en la figura 10 permite identificar con alta precisión al personal evaluado pertenecientes a la clase 0 (No Idóneo) en el nodo raíz del árbol, representando la gran mayoría de los casos 2.018 de 2.120 muestras,

mediante una secuencia de divisiones basadas en los registros de las variables seleccionadas de mayor importancia que guían el modelo, las cuales son: la no autorización para verificar datos del evaluado (reg_no_autoriza_verificar_datos_DHP), la no autorización para que se realice la prueba de confianza el evaluado (reg_no_autoriza_prue_confianza_DHP), el no cumplir con la nota mínima en el examen de conocimiento sobre el tratamiento de la documentación militar realizado por dos ocaciones (reg_nota_min_exam_conoc_dos_opor), la existencia de faltas disciplinarias por negligencia en pérdida de documentación militar (reg_falt_negli_talento_humano) y la responsabilidad disciplinaria en procesos dentro de la institución (reg_resp_investigacion_inspectoria). A lo largo de las divisiones, el índice Gini disminuye progresivamente, indicando una mayor pureza de clasificación. El árbol alcanza nodos terminales completamente puros es decir con Gini = 0, donde se identifica con total certeza a personal evaluado de la clase 1, es decir personal que obtuvo la calificación de Idóneo. Esto demuestra que el modelo no solo discrimina efizcamente entre clases, sino que también proporciona una lógica clara y trazable para fundamentar decisiones institucionales basadas en el historial del evaluado.

Figura 9 Generacion del árbol de decisión del mejor modelo a implementar



Fuente: Elaboración propia

DISCUSIÓN

Los resultados de esta investigación van en concordancia con lo aplicado por Adillah *et al.* (2025); Salau *et al.* (2023), los cuales proponen en sus estudios algoritmos supervisados de clasificación como Árboles de Decisión, Clasificador Naive Bayes , K-Nearest Neighbors y Máquina de Soporte Vectorial Lineal, donde para la evaluación del mejor algoritmo, utilizaron las métricas de Exactitud (Accuracy) y de AUC-ROC, para demostrar el modelo más efectivo a emplear, es por ello, que en esta investigación se tomaron las mismas métricas de Exactitud y AUC para evaluar los algoritmos seleccionados de Árboles de Decisión, Clasificador Naive Bayes , K-Nearest Neighbors; y asi, poder determinar el mejor modelo para la evaluación del personal naval para la idoneidad o no idoneidad para ocupar puestos sensibles.

CONCLUSIONES

La investigación realizada de los estudios del estado del arte antes expuestos y el análisis de los datos oficiales obtenidos de la institución militar de los años 2017 hasta agosto del 2024, con dos posibles resultados Idóneo o No Idóneo para ocupar puestos sensibles, permitió determinar que para la solución de este problema se deben aplicar algoritmos supervisados de clasificación, los cuales están indicados en la fundamentación teórica, para la selección del mejor algoritmo o modelo a implementar.

Lo resultados obtenidos del entrenamiento y evaluación de los algoritmos de clasificación Árboles de Decisión, Naive Bayes Bernoulli y el algoritmo K-Nearest Neighbors, permitió determinar que los tres modelos seleccionados son buenos ya que cumplen con los valores planteados en la hipótesis en cuanto a las métricas de Exactitud y AUC; sin embargo, el algoritmo Árboles de Decisión demostró ser el más efectivo, es decir el mejor modelo a implementar, logrando la mejor Exactitud del 99,81% y un AUC de 0,98, demostrando un buen rendimiento en la distinción de clases positivas y negativas, teniendo como factores claves que contribuyen al éxito del modelo la selección de características relevantes, la calidad de los datos y las técnicas adecuadas para el entrenamiento y validación de los algoritmos.

En cuanto a los desafíos, aunque no invalidan los hallazgos, se sugiere para estudios futuros se debería ampliar la muestra y seguir investigando estados del arte para analizar y corroborar los tipos métricas que permitan seleccionar el mejor algoritmo a implementar en base a los casos similares del estado de arte ya estudiado.

Se recomienda investigar los tipos de sistemas o aplicaciones que permitan implementar el modelo seleccionado.

REFERENCIAS BIBLIOGRAFICAS

Adillah, M. F. N., Suakanto, S., & Utama, N. I. (2025). Implementation of Machine Learning-Based Classification Model in Employee Recruitment Decision Prediction. *Journal La Multiapp*, *6*(2), 341-352.

Ali, A. N. F., Sulaima, M. F., Razak, I. A. W. A., Kadir, A. F. A., & Mokhlis, H. (2023). Artificial intelligence application in demand response: Advantages, issues, status, and challenges. *IEEE access*, *11*, 16907-16922.

Arana, C. (2021). *Modelos de aprendizaje automoático mediante árboles de decisión*.

Borja-Robalino, R., Monleón-Getino, A., & Rodellar, J. (2020). Estandarización de métricas de rendimiento para clasificadores Machine y Deep Learning. *Revista Ibérica de Sistemas e Tecnologias de Informação*(E30), 184-196.

Contreras, L., Fuentes, H., & Rodríguez, J. (2024). *Algoritmos supervisados y de ensamble con python: Implementación y estrategia de optimización*. Ediciones de la U.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Chamorro, A. (2020). Malware Detection with Machine Learning.

Gupta, V., Mishra, V. K., Singhal, P., & Kumar, A. (2022). An overview of supervised machine learning algorithm. 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART),

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). Metodología de la Investigación, Sexta Edición México. *DF, Editores, SA de CV*.

Lu, Y., Ye, T., & Zheng, J. (2022). Decision tree algorithm in machine learning. 2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA),

Romero, J. L. (2025). Análisis integral de algoritmos de clasificación en aprendizaje automático: perspectivas, comparaciones y aplicaciones. *Serie Científica de la Universidad de las Ciencias Informáticas*, *18*(1), 283-304.

Salamanca, I. N., & Castro, E. J. (2020). Técnicas de aprendizaje automático aplicadas en los sistemas de predicción. *Tecnología Investigación y Academia*, *8*(1), 37-53.

Salau, A. O., Assegie, T. A., Akindadelo, A. T., & Eneh, J. N. (2023). Evaluation of Bernoulli Naive Bayes model for detection of distributed denial of service attacks. *Bulletin of Electrical Engineering and Informatics*, *12*(2), 1203-1208.

Sindayigaya, L., & Dey, A. (2022). Machine learning algorithms: a review. *Inf Syst J*, *11*(8), 1127-1133.

Singh, A., Thakur, N., & Sharma, A. (2016). A review of supervised machine learning algorithms. 2016 3rd international conference on computing for sustainable global development (INDIACom),

Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. (2016). A review of machine learning techniques using decision tree and support vector machine. 2016 international conference on computing communication control and automation (ICCUBEA),

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math*, *58*(345-363), 5.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, *27*(11), 1134-1142.