

"Año de la recuperación y consolidación de la economía peruana"

**RESOLUCIÓN DE VICERRECTORADO DE INVESTIGACIÓN
N° 037-2025-UNTELS -VRI**

Villa El Salvador, 31 de julio de 2025

VISTO:

Oficio N° 0443-2025-UNTELS-VRI-II, de fecha 25 de julio del 2025, la Dirección del Instituto de Investigación emite opinión favorable para **APROBAR el cierre del proyecto de investigación denominado: "Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima"**, (modalidad sin financiamiento), a cargo de la investigadora principal, **Mag. Janett Deisy Julca Flores**, y;

CONSIDERANDO:

Que, el artículo 18° de la Constitución Política del Perú, en su cuarto párrafo establece: Cada Universidad es autónoma en su régimen normativo, de gobierno, académico, administrativo y económico. Las Universidades se rigen por la Ley Universitaria N° 30220 y sus propios estatutos en el marco de la constitución y de las leyes;

Que, el artículo 56 y 57, numeral 57.5 de la Ley Universitaria N° 30220 y sus modificatorias, establece: "Artículo 56. La Asamblea universitaria es un órgano que representa a la comunidad universitaria, se encarga de dictar las políticas generales de la universidad (...)" asimismo el artículo 57 numeral 57.5 señala que tiene la siguiente atribución: "Elegir a los integrantes del Comité Electoral Universitario y Tribunal de Honor";

Que, mediante Resolución N° 0002-2023-CEU-UNTELS, de fecha 02 de mayo de 2023, el Comité Electoral de la UNTELS reconoce a la Dra. Gladys Marcionila Cruz Yupanqui como Rectora, Dra. Marina Vilca Cáceres como Vicerrectora Académica y Dr. Angel Fernando Navarro Raymundo como Vicerrector de Investigación de la Universidad Nacional Tecnológica de Lima Sur;

Que, en el art. 50° de la Ley Universitaria Ley N° 30220, precisa que: "*El Vicerrectorado de Investigación, según sea el caso, es el organismo de más alto nivel en la universidad en el ámbito de la investigación*"; y el art. 65.2.1 señala que es atribución del Vicerrector de Investigación: "*Dirigir y ejecutar la política general de investigación de la universidad*";

Que, mediante Resolución de Comisión Organizadora N° 020-2023-UNTELS, de fecha 17 de enero de 2023, el presidente de la comisión organizadora aprueba los resultados de la evaluación de proyectos de investigación para el desarrollo de ciencia y tecnología e innovación 2022 – UNTELS, modalidad sin financiamiento, dentro de los cuales se aprobó el proyecto de investigación: "Material particulado en ambientes de expendio de alimentos para consumo humano en zonas urbano-marginales en una ciudad amazónica, 2023";

Que, mediante Resolución de Consejo Universitario N° 057-2023-UNTELS-CU-R, de fecha 31 de julio de 2023, se autorizó al Vicerrector de Investigación la emisión de actos resolutiveos en el marco de sus atribuciones conferidas por la Ley Universitaria N° 30220 y mediante Resolución de Consejo Universitario N° 119-2024-UNTELS-CU, de fecha 28 de junio de 2024, el Consejo Universitario dispone el cumplimiento de la resolución precitada, precisando la emisión de actos resolutiveos que comprenden asuntos de su competencia;

Que, en concordancia con la RCU N° 119-2024-UNTELS-CU, en el uso de sus atribuciones referido en el literal j) "*Aprobar cierres de proyecto de investigación*";

Que, mediante Carta N° 0030-2025-JDJF, de fecha 21 de julio de 2024, la investigadora principal, Mag. Janett Deisy Julca Flores, solicita a la Directora del Instituto de Investigación el cierre del proyecto de investigación denominado: "*Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima*";

Que, mediante Memorándum N° 0211-2025-UNTELS-VRI-II, de fecha 21 de julio de 2025, la Directora del Instituto de Investigación, solicita al Especialista en Gestión de la Investigación la evaluación y verificación de la solicitud de cierre de proyecto de investigación presentado por la Mag. Janett Deisy Julca Flores;

...///



Firmado digitalmente por:
NAVARRO RAYMUNDO Angel
Fernando FAU 20502245032 hard
Motivo: Soy el autor del
documento
Fecha: 31/07/2025 11:53:13-0500

.../// Ref. RESOLUCIÓN DE VICERRECTORADO DE INVESTIGACIÓN N° 037-2025-UNTELS -VRI

Que, mediante Informe N° 0080-2025-UNTELS-VRI-MANP, de fecha 25 de julio de 2025, el Especialista en Gestión de la Investigación concluye y recomienda *“El proyecto ha cumplido íntegramente con todos los requisitos formales y técnicos establecidos en las bases de proyectos autofinanciados 2024 (entregables, publicación, difusión, originalidad y formatos de consolidado). Se recomienda emitir la resolución de cierre”*;

Que, mediante Oficio N° 0443-2025-UNTELS-VRI-II, de fecha 25 de julio de 2025, la Dirección del Instituto de Investigación emite opinión favorable para aprobar el cierre del proyecto de investigación: *“Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima”*, (modalidad sin financiamiento), a cargo de la investigadora principal, Mag. Janett Deisy Julca Flores;

Que, estando en uso de las atribuciones otorgadas al Vicerrector de Investigación, de acuerdo a lo dispuesto por la autoridad universitaria conferidas por las disposiciones citadas;

SE RESUELVE:

ARTÍCULO PRIMERO. - APROBAR el cierre del proyecto de investigación titulado: **“Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima”**, (modalidad sin financiamiento) a cargo de la investigadora principal, **Mag. Janett Deisy Julca Flores**, conforme al Anexo que en cuarenta y seis (46) folios forma parte de la presente resolución.

ARTÍCULO SEGUNDO. - NOTIFICAR la presente resolución a Rectorado, Secretaría General, Dirección del Instituto de Investigación y Mag. Janett Deisy Julca Flores.

ARTÍCULO TERCERO. - PUBLICAR la presente resolución en el Portal de Transparencia Estándar de la Universidad Nacional Tecnológica de Lima Sur.

ARTÍCULO CUARTO. - ENCARGAR el cumplimiento de la presente resolución a la Dirección del Instituto de Investigación de la Universidad Nacional Tecnológica de Lima Sur.

Regístrese, comuníquese, publíquese y archívese.



Firmado digitalmente por:
NAVARRO RAYMUNDO Angel
Fernando FAU 20502245032 hard
Motivo: Soy el autor del
documento
Fecha: 31/07/2025 11:53:13-0500

Dr. ANGEL FERNANDO NAVARRO RAYMUNDO
Vicerrector de Investigación
UNTELS

UNIVERSIDAD NACIONAL TECNOLÓGICA DE LIMA SUR
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



INFORME FINAL

Línea de Investigación:

Inteligencia Artificial y Sistemas Inteligentes

Título del proyecto:

Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima.

Investigador Principal:

Mg. Janett Deisy Julca Flores

Fecha de presentación:

Junio, 2025

Tabla de contenido

RESUMEN	3
ABSTRACT	4
1. INTRODUCCIÓN	5
2. PLANTEAMIENTO DEL PROBLEMA	6
2.1 Problema principal	6
2.2 Problemas específicos	6
2.3 Justificación	7
2.4 Objetivos	8
2.4.1. Objetivo general	8
2.4.2. Objetivos específicos	8
3. ANTECEDENTES	8
4. METODOLOGIA	11
4.1 Marco Teórico	11
4.1.1. Redes Neuronales Artificiales (RNA)	11
4.1.2. Árboles de Decisión (AD)	11
4.1.3. Máquinas de Soporte Vectorial (SVM)	12
4.1.4. K-Nearest Neighbors (KNN)	12
4.2 Tipo y enfoque de investigación	12
4.3 Población y muestra	13
4.4 Técnica e instrumentos de recolección de datos	13
4.5 Preprocesamiento de datos	13
4.6 Sectorización y organización del trabajo de campo	14
4.7 Aplicación de algoritmos de Machine Learning	16
4.8 Validación de los instrumentos	16
4.9 Procesamiento y análisis de datos	17
4.10 Consideraciones éticas	17
5. RESULTADOS	17
5.1 Descripción general de la base de datos	17
5.2 Resultados del análisis comparativo de algoritmos	18
5.3 Resultados por sector de la playa	18
5.4 Comparación computacional	19
5.5 Predicción de generación de residuos	19
5.2 Validación cruzada	20
5.3 Comparación cualitativa y cuantitativa de los algoritmos de machine learning	20
6 DISCUSIÓN DE RESULTADOS	22
7 CONCLUSIONES	23
8 RECOMENDACIONES	25
9 ANEXOS	25
ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos	26
Anexo 02: Evaluación de la validación de Criterios para el Análisis de Algoritmos de machine learning	36
ANEXO 03 : Instrumentos de recolección de residuos en playas	37
Anexo 04: Matriz de consistencia Título: "Análisis comparativo de algoritmos de machine learning para la predicción de la	37
Anexo 05: Tabla de Corrección de Observaciones emitidas para el informe parcial	40
Anexo 06: Evidencia de Artículo científico sometido a revisión en la revista: BIOTECH & ENGINEERING	43
Anexo 07: Evidencia de ponencia realizada en evento académico internacional	44
10 Referencias Bibliográficas	45

RESUMEN

La playa Venecia ubicada en el distrito de Villa Salvador, es una de las playas con más contaminación de residuos sólidos del litoral limeño, constituyendo ahora una problemática ambiental constante en el sector urbano del distrito. A raíz de ello, esta investigación propuso evaluar el impacto del uso de algoritmos supervisados de aprendizaje automático en la predicción de residuos sólidos recolectados, de hecho resultó una herramienta útil para fortalecer la toma de decisiones en la gestión ambiental costera. Asimismo, se trabajó con cuatro algoritmos ampliamente utilizados en el machine learning: Red Neuronal Artificial (RNA), Árbol de Decisión (AD), Máquina de Soporte Vectorial (SVM) y K-Nearest Neighbors (KNN) y se contó con una base de datos compuesta de 200 registros, para la comparación técnica las métricas utilizadas fueron: precisión, F1-score, recall y tiempo de entrenamiento. Los resultados mostraron que el modelo de red neuronal alcanzó el mejor desempeño general, destacando en sectores con alta diversidad de residuos, en cambio SVM *mostró un comportamiento equilibrado entre exactitud y eficiencia computacional*, mientras que AD y KNN ofrecieron menor precisión, aunque con tiempos de ejecución reducidos. De igual forma, la técnica de validación cruzada permitió confirmar la estabilidad de los modelos, mientras que una comparación cualitativa basada en criterios valorados por expertos —como adaptabilidad, claridad, facilidad de implementación y sensibilidad al ruido— *reforzó el posicionamiento de la RNA como el algoritmo más completo para el contexto analizado*. A partir de este enfoque, se elaboraron predicciones para futuros volúmenes de residuos, permitiendo identificar zonas críticas y orientar estrategias preventivas.

Los hallazgos obtenidos aportan evidencia sobre la utilidad de los modelos predictivos como herramienta de soporte en la gestión de residuos sólidos en entornos costeros. Asimismo, sientan las bases para replicar este tipo de soluciones en otras playas de la capital u otras regiones con desafíos ambientales similares.

Palabras clave: Residuos sólidos, predicción, machine learning, playa venecia.

ABSTRACT

Venecia beach, located in the district of Villa Salvador, is one of the beaches with more solid waste contamination of the Lima coast, constituting now a constant environmental problem in the urban sector of the district.

As a result, this research proposed to evaluate the impact of the use of supervised machine learning algorithms in the prediction of collected solid waste, in fact it turned out to be a useful tool to strengthen decision making in coastal environmental management.

Likewise, we worked with four algorithms widely used in machine learning: Artificial Neural Network (ANN), Decision Tree (AD), Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) and we had a database composed of 200 records, for the technical comparison the metrics used were: accuracy, F1-score, recall and training time. The results showed that the neural network model achieved the best overall performance, standing out in sectors with high diversity of residues, while SVM showed a balanced behavior between accuracy and computational efficiency, while AD and KNN offered lower accuracy, although with reduced execution times. Similarly, the cross-validation technique confirmed the stability of the models, while a qualitative comparison based on criteria assessed by experts - such as adaptability, clarity, ease of implementation and sensitivity to noise - reinforced the positioning of ANN as the most complete algorithm for the context analyzed. Based on this approach, predictions were made for future waste volumes, making it possible to identify critical areas and guide preventive strategies.

The findings obtained provide evidence of the usefulness of predictive models as a support tool for solid waste management in coastal environments. They also lay the groundwork for replicating this type of solution in other beaches of the capital city or other regions with similar environmental challenges.

Keywords: Solid waste, prediction, machine learning, Venice beach.

1. INTRODUCCIÓN

El aumento de la generación de residuos sólidos en zonas urbanas y costeras se ha convertido en una preocupación creciente a nivel global, no solo por el deterioro que causa en los ecosistemas naturales, sino también por las repercusiones que tiene en la salud de la población y en la economía local. Este problema se intensifica en espacios públicos abiertos, como las playas, donde las actividades humanas, el turismo y la escasa conciencia ambiental contribuyen a una acumulación constante de desechos.

En el caso específico de Lima Metropolitana, la playa Venecia —ubicada en el distrito de Villa El Salvador— ha sido identificada como una de las zonas más afectadas por el desecho inadecuado de residuos. Reportes institucionales y medios de comunicación han señalado que la acumulación de basura compromete la calidad ambiental del lugar, altera el equilibrio ecológico costero, y reduce considerablemente su atractivo turístico (MINSA, 2019; Latina TV, 2021). Pese a los esfuerzos de las autoridades locales, como la Municipalidad Distrital de Villa El Salvador (2024), las acciones de recolección manual resultan limitadas en su alcance y frecuencia, lo que impide una respuesta efectiva al ritmo con el que los residuos se generan.

Frente a esta situación, se llevó a cabo una serie de jornadas de limpieza en coordinación con estudiantes de Ingeniería de Sistemas de la Universidad Nacional Tecnológica de Lima Sur (UNTELS). Estas jornadas no solo permitieron recoger desechos, sino también recopilar datos sistematizados sobre la cantidad y tipo de residuos, el área de recolección, y otras variables asociadas a la generación de estos elementos. Por consiguiente, la información recolectada en campo sirvió para mejorar la gestión ambiental proponiendo un análisis orientado al uso de herramientas tecnológicas, de este modo, esta investigación plantea predecir la generación de residuos sólidos en la playa Venecia usando algoritmos de aprendizaje automático supervisado (machine learning). De hecho, la orientación del estudio es predictiva, basándose en datos históricos recolectados entre 2023 y 2025 y se enfoca en anticipar la cantidad de residuos sólidos que podrían generarse en un periodo futuro. De este modo, se busca ofrecer a las autoridades locales una herramienta que les permita asignar recursos de forma más eficiente y planificar mejor las intervenciones de limpieza.

Asimismo, el estudio realiza una comparación técnica entre cuatro algoritmos supervisados: Red Neuronal Artificial, Árbol de Decisión, K-Vecinos más Cercanos (KNN) y Máquina de Soporte Vectorial (SVM). La evaluación tiene como finalidad identificar cuál de estos modelos presenta un mejor rendimiento en tareas predictivas con datos reales de campo, basándose en métricas como la precisión, la sensibilidad (recall), el F1-score y el tiempo de procesamiento (UNTELS, 2024).

Finalmente, el informe se organiza en las siguientes secciones: planteamiento del problema, revisión de antecedentes, metodología, resultados, discusión, conclusiones, recomendaciones y anexos. Cada parte ha sido desarrollada con el propósito de brindar no solo un aporte teórico a la aplicación del machine learning en temas ambientales, sino

también una propuesta práctica y contextualizada que pueda servir como modelo para la gestión sostenible de espacios costeros

2. PLANTEAMIENTO DEL PROBLEMA

La Playa Venecia, ubicada en el distrito de Villa El Salvador, Lima, enfrenta una creciente *acumulación de residuos sólidos que compromete tanto el ecosistema marino como la salud de las personas que visitan o viven cerca del lugar*. En una jornada de limpieza organizada en 2023, el Ministerio del Ambiente reportó la recolección de una gran cantidad de residuos como tapas de botellas, colillas de cigarrillos, latas, papeles y restos orgánicos (Minam, 2023). Esta situación evidencia un aumento progresivo en los niveles de contaminación de la playa, *afectando su valor ambiental, recreativo y turístico*.

Si bien se han realizado acciones de limpieza, estas tienen un carácter correctivo y no preventivo. La persistencia del problema indica la necesidad de adoptar estrategias basadas en datos, que permitan anticipar la cantidad y tipo de residuos que se generarán, con el fin de optimizar la respuesta de las autoridades y mejorar la planificación de las tareas de limpieza.

En este contexto, los modelos de aprendizaje automático (machine learning) ofrecen un enfoque prometedor para predecir fenómenos complejos basados en grandes volúmenes de datos. No obstante, elegir el algoritmo más adecuado para realizar predicciones en un entorno costero como Playa Venecia no es una tarea sencilla. Algoritmos como Red Neuronal, Árbol de Decisión, Vecino más Cercano y Máquina de Soporte Vectorial presentan diferencias técnicas en precisión, sensibilidad al ruido, capacidad de generalización y eficiencia computacional. Por tanto, se requiere un análisis comparativo que determine cuál de estos algoritmos es más eficaz en la predicción de residuos sólidos con datos reales del lugar, de este modo permitiendo su aplicación futura en la gestión ambiental del área.

2.1 Problema principal

¿Cómo influye el uso de distintos algoritmos supervisados de aprendizaje automático en la precisión de la predicción de residuos sólidos recolectados en la Playa Venecia, del distrito de Villa El Salvador?

2.2 Problemas específicos

¿Qué desempeño técnico (precisión, recall, F1-score y tiempo de ejecución) presentan los algoritmos Red Neuronal, Árbol de Decisión, K-Vecinos más Cercanos y Máquina de Soporte Vectorial al aplicarlos a los datos recolectados en Playa Venecia?

¿Qué consideraciones prácticas deben tomarse en cuenta para implementar el algoritmo con mejor rendimiento como herramienta predictiva en la planificación de actividades de gestión de residuos sólidos en zonas costeras?

2.3 Justificación

La presente investigación se justifica por la necesidad de buscar nuevas estrategias para enfrentar la acumulación de residuos sólidos en zonas costeras, un problema ambiental persistente que afecta tanto al ecosistema como a la calidad de vida de las comunidades. De hecho, la Playa Venecia, ubicada en el distrito de Villa El Salvador, Lima, es un ejemplo de la limitada capacidad de gestión municipal, este tipo de escenarios, ha conllevado a una creciente contaminación de los espacios públicos (Municipalidad Distrital de Villa El Salvador, 2024; MINSA, 2019).

Además, desde una perspectiva teórica, investigaciones previas han demostrado que el uso de algoritmos supervisados puede mejorar significativamente la eficiencia en tareas predictivas, por lo tanto el estudio cobra relevancia porque plantea la incorporación de modelos de aprendizaje automático (machine learning) como herramienta para anticipar la generación de residuos sólidos, igualmente estos modelos permiten procesar grandes volúmenes de datos y detectar patrones no evidentes, lo cual resulta especialmente útil en entornos donde la recolección manual es intermitente y dependiente de recursos humanos limitados. (UNTELS, 2023; Latina TV, 2021). De la misma forma, es importante aclarar que los algoritmos —Red Neuronal, Árbol de Decisión, K-Vecinos más Cercanos y Máquina de Soporte Vectorial—son herramientas analíticas empleadas para interpretar los datos de la muestra y no son unidades de observación, de hecho la muestra está compuesta por los datos recolectados en campo durante las jornadas de limpieza realizadas entre 2023 y 2025 en la Playa Venecia, dichos datos permitirán evaluar cuál de los algoritmos ofrece mayor precisión, eficiencia y adaptabilidad al contexto particular de la playa analizada.

Respecto a la justificación metodológica, el estudio adopta un enfoque sistemático que resulta indispensable en contextos de planificación y gestión ambiental, además permite establecer una comparativa, que incluye la evaluación técnica a partir de métricas como: precisión, recall, F1-score y tiempo de ejecución de cada algoritmo, además evalúa la sensibilidad al ruido y el nivel de complejidad computacional de cada algoritmo. Asimismo, el impacto de esta investigación trasciende al ámbito social y económico, al poseer una herramienta capaz de anticipar los niveles de residuos por zonas y momentos específicos, beneficiando a los gestores ambientales en la optimización de la asignación de recursos, reducción de los costos operativos y mejoramiento de la eficacia de las jornadas de limpieza, por otra parte favorecería a las actividades económicas locales: como el turismo y el comercio debido a la recuperación del entorno natural, contribuyendo al desarrollo sostenible de la comunidad.

En síntesis, esta investigación se justifica por su aporte integral: combina una propuesta tecnológica aplicable con una base metodológica sólida y una orientación socialmente pertinente. Además, su enfoque es replicable en otras playas u áreas naturales con problemáticas similares, extendiendo su utilidad más allá del caso de estudio inmediato.

2.4 Objetivos

2.4.1. Objetivo general

Comparar el rendimiento de cuatro algoritmos supervisados de aprendizaje automático —Red Neuronal, Árbol de Decisión, K-Vecinos más Cercanos (KNN) y Máquina de Soporte Vectorial (SVM)— en la predicción de la generación de residuos sólidos en la Playa Venecia, distrito de Villa El Salvador, Lima, utilizando métricas técnicas como precisión, recall, F1-score y tiempo de ejecución, con el propósito de identificar el modelo más eficiente para su futura aplicación en la gestión ambiental costera.

2.4.2. Objetivos específicos

- **Construir una base de datos estructurada** a partir de los registros obtenidos en jornadas de limpieza realizadas en la Playa Venecia, incluyendo variables como tipo, cantidad y ubicación de los residuos sólidos recolectados, adecuada para ser procesada por algoritmos supervisados de machine learning.
- **Aplicar los cuatro algoritmos seleccionados** (Red Neuronal, Árbol de Decisión, KNN y SVM) al conjunto de datos, evaluando su rendimiento mediante métricas cuantitativas (precisión, recall, F1-score y tiempo de ejecución) y aspectos cualitativos como facilidad de implementación, adaptabilidad al ruido y claridad en la salida del modelo.
- **Analizar la complejidad computacional** de cada algoritmo, considerando el uso de recursos en términos de tiempo de entrenamiento, velocidad de predicción y consumo de memoria, con el fin de estimar su viabilidad operativa en entornos con capacidades limitadas.
- **Detectar patrones y tendencias** en la generación de residuos sólidos según el área, frecuencia y tipo de residuo predominante, a partir de los resultados generados por los modelos predictivos entrenados.

3. ANTECEDENTES

Lima, la capital del Perú, enfrenta la acumulación de residuos sólidos en sus playas un problema significativo, que ha generado un interés creciente en la aplicación de técnicas de Machine Learning para la gestión de residuos sólidos, lo que ha impulsado una serie de estudios en esta área frente a este problema. Dicha problemática afecta negativamente tanto el ecosistema marino como el paisaje costero, comprometiendo así la calidad ambiental y la experiencia turística en la región.

Diversos trabajos han abordado diversas problemáticas: optimización de la recolección y el reciclaje, la predicción de la generación de residuos, entre otros aspectos, sin embargo, es importante destacar que pocos de estos estudios se centran en la comparación cualitativa o cuantitativa de modelos o algoritmos de Machine Learning, que es el enfoque principal de la presente investigación.

Asimismo, se mencionará investigaciones relevantes en esta área, que realizaron un análisis comparativo de algoritmos de machine learning para predecir la generación de residuos sólidos, usando técnicas como redes neuronales, ANN-ReLu, SVM-polynomial, ANN-logistic, técnicas de ensemble learning, random forest, k-vecino más cercano, PCA-ANN, GT-ANN y otros algoritmos como árboles de decisión y máquinas de soporte vectorial, evaluando su precisión, error medio y capacidad para describir la variación en los datos.

(Pillai et al., 2023) planteó en India una revisión sistemática y analítica de la literatura sobre diversas aplicaciones de algoritmos de Machine Learning (ML) en la gestión de residuos sólidos municipales (MSW) para mejorar los procedimientos de gestión de residuos y reducir los efectos ambientales adversos. La investigación fue de tipo revisión y utilizó un diseño basado en el análisis sistemático de literatura sobre algoritmos como árboles de decisión, k-vecinos más cercanos (KNN), máquinas de soporte vectorial (SVM), y bosques aleatorios, entre otros. Los principales hallazgos incluyen la optimización de procesos como la categorización de residuos y la predicción de volúmenes generados, así como la integración de tecnologías IoT para el monitoreo en tiempo real. Este antecedente se relaciona con este estudio porque permite explorar enfoques similares en el contexto de la Playa Venecia, proporcionando bases para la comparación de algoritmos y el desarrollo de soluciones predictivas en la gestión de residuos sólidos.

(Cha et al, 2022) en la República de Corea desarrollaron modelos predictivos basados en algoritmos de machine learning para estimar la tasa de generación de residuos de demolición (DWGR) en áreas de reurbanización en Corea del Sur. La investigación fue de tipo aplicada y utilizó un diseño experimental que incluyó la recolección y preprocesamiento de datos, la aplicación de algoritmos como redes neuronales artificiales (ANN), K-vecinos más cercanos (KNN), bosques aleatorios (RF) y máquinas de soporte vectorial (SVM), además de la optimización de hiperparámetros y la validación de los modelos mediante métricas como el coeficiente de determinación (R^2). Como resultado, se identificaron modelos óptimos ANN-ReLu, SVM-polynomial y ANN-logistic para predecir la tasa de generación de residuos de demolición con errores promedio del 7,3%, 7,4% y 7,5%, respectivamente. Este antecedente se relaciona con esta investigación porque ofrece un marco metodológico relevante para implementar algoritmos de machine learning en problemas similares de predicción y gestión de residuos.

(Adeleke et al., 2021), en Sudáfrica realizaron una revisión sobre modelos predictivos basados en la composición química de los residuos para evaluar y predecir el contenido energético de estos. Además, desarrollo modelos matemáticos que estiman el contenido energético de distintos tipos de residuos, destacando su aplicabilidad en la optimización de procesos de conversión de residuos en energía, de este modo como resultado, se determinó que las redes neuronales artificiales (RNA) y las regresiones lineales múltiples son más

efectivas para predecir el contenido energético de los residuos sólidos municipales, y las ANN muestran mayor precisión que la regresión lineal. Este antecedente se relaciona con la investigación porque proporciona un enfoque basado en modelos predictivos que puede ser adaptado para mejorar la eficiencia en la gestión y valorización energética de los residuos.

(Ayeleru et al., 2021) en Sudáfrica desarrollaron técnicas de machine learning, específicamente redes neuronales artificiales (ANN) y máquinas de soporte vectorial (SVM) para predecir la cantidad de residuos sólidos municipales generados en la Ciudad de Johannesburgo. La investigación fue de tipo aplicada, con un diseño no experimental - descriptivo, basado en datos históricos de Statistics South África (STATS SA). Como principal resultado, se encontró que el modelo ANN10 alcanzó un coeficiente de determinación (R^2) de 99.9%, proyectando que la generación de residuos sólidos alcanzará 1.95 millones de toneladas anuales en 2050. Este antecedente se relaciona con esta investigación porque demuestra la eficacia de los modelos predictivos basados en machine learning para la gestión sostenible de residuos, destacando su aplicabilidad en contextos urbanos con limitaciones de datos.

(Nguyen et al., 2021) en Vietnam proponen comparar seis modelos de machine learning para predecir la generación de residuos sólidos municipales (MSW) en áreas residenciales. La investigación fue de tipo aplicada y utilizó un diseño cuantitativo basado en el análisis de datos empíricos mediante modelos de machine learning, obteniendo como principal resultado que los algoritmos Random Forest (RF) y K-Nearest Neighbor (KNN) presentan la mayor capacidad predictiva ($R^2 > 0.96$) y precisión en las predicciones (MAE de 121.5 a 125.0). Este antecedente se relaciona con esta investigación porque evidencia la utilidad de los modelos predictivos para abordar problemas de gestión de residuos en contextos urbanos, aportando enfoques metodológicos que pueden adaptarse a diferentes realidades.

(Kavyanifar et al., 2020), en Irán proponen predecir la producción total de residuos sólidos en áreas costeras de Noor, provincia de Mazandarán, utilizando enfoques de machine learning. La investigación fue de tipo aplicado y experimental, con un diseño cuasi-experimental y descriptivo, basado en la recolección de datos meteorológicos y de residuos durante un año. Se aplicaron modelos de aprendizaje automático como redes neuronales y árboles de decisión, destacando el modelo CHAID por su precisión y mejor rendimiento en la predicción de la producción total de residuos sólidos en comparación con los modelos CART, MLP y RBF. Este antecedente se relaciona con esta investigación porque aborda el uso de machine learning para predecir la generación de residuos sólidos, proporcionando metodologías y enfoques aplicables a contextos similares.

(Cerna, 2022) en el país Perú proponen predecir la generación de residuos sólidos domiciliarios utilizando modelos de Machine Learning en una zona rural de Puno, dicha

investigación fue de tipo aplicada y cuantitativa, con un diseño no experimental y transversal, en el que recopilamos datos socioeconómicos y demográficos a través de encuestas, pesaje de residuos y análisis estadístico, incluyendo correlación de Spearman y ANOVA para la selección de variables relevantes. El resultado del análisis reveló que las variables demográficas y socioeconómicas tienen un impacto significativo en la generación per cápita de residuos sólidos y que los modelos de Machine Learning con un alto desempeño predictivo son: la regresión lineal múltiple y Random Forest. Este antecedente aporta a la investigación porque evidencia la optimización de la gestión de residuos sólidos mediante el uso de tecnologías avanzadas basándose en el análisis de variables locales, lo cual es de suma importancia para abordar problemas ambientales en áreas rurales.

Por tanto, esta investigación tiene como propósito aplicar métodos de Machine Learning en la gestión de desechos sólidos en áreas costeras, además se anticipa que los hallazgos proporcionaran conocimientos valiosos que puedan ser utilizados para desarrollar y ejecutar tácticas de gestión de desechos en la Playa Venecia, Lima, sirviendo como punto de partida para futuras investigaciones en este dominio.

4. METODOLOGIA

4.1 Marco Teórico

4.1.1. Redes Neuronales Artificiales (RNA)

Las redes neuronales artificiales emulan el funcionamiento del cerebro humano mediante capas de nodos conectados que ajustan sus pesos a través de la retropropagación, permitiendo el aprendizaje de patrones complejos (Coskuner et al., 2021). Estudios recientes han demostrado que las RNA pueden alcanzar altos niveles de precisión en la predicción de residuos sólidos municipales, con valores de R^2 mayores al 99 % en ciertos contextos urbanos (Ayeleru et al., 2021). Sin embargo, su entrenamiento requiere recursos computacionales considerables y una cuidadosa optimización de hiperparámetros para evitar el sobreajuste (Pillai et al., 2023).

4.1.2. Árboles de Decisión (AD)

Los árboles de decisión son modelos basados en reglas que permiten clasificar observaciones mediante una estructura jerárquica de decisiones. Se caracterizan por su facilidad de interpretación, lo que los hace útiles en problemas donde se requiere explicar la lógica detrás de las predicciones (Toğaçar et al., 2023). No obstante, estos algoritmos tienden a sobreajustar los datos si no se aplican mecanismos de poda o ensamblaje como Random Forest (Ayeleru et al., 2021).

4.1.3. Máquinas de Soporte Vectorial (SVM)

La SVM es un algoritmo de clasificación que busca maximizar el margen entre las clases mediante la creación de un hiperplano óptimo. Se ha comprobado su eficacia en problemas con alta dimensionalidad y conjuntos de datos pequeños, donde alcanza una alta precisión predictiva (Ayeleru et al., 2021). Sin embargo, el rendimiento del modelo depende de la adecuada selección del kernel y la calibración de sus parámetros (Pillai et al., 2023).

4.1.4 K-Nearest Neighbors (KNN)

El algoritmo KNN clasifica una muestra desconocida según la mayoría de las clases presentes en sus vecinos más cercanos. Es un método no paramétrico fácil de implementar y comprender, ideal para aplicaciones con baja complejidad (Toğaçar et al., 2023). A pesar de su simplicidad, su eficiencia se ve afectada por la presencia de ruido en los datos y su rendimiento disminuye en conjuntos grandes debido al alto costo computacional en la fase de predicción (Pillai et al., 2023).

Síntesis Comparativa

Algoritmo	Fortalezas	Limitaciones
RNA	Alta capacidad de aprendizaje, excelente precisión (>99 % R ²)	Requiere ajuste fino y recursos computacionales elevados (Ayeleru et al., 2021; Coskuner et al., 2021)
AD	Claridad interpretativa, buena base para ensambles	Riesgo de sobreajuste sin poda (Toğaçar et al., 2023)
SVM	Buena precisión en datos pequeños y complejos	Sensible a selección de kernel y parámetros (Pillai et al., 2023)
KNN	Fácil de entender, adecuado para prototipos	Alto costo computacional en predicción, sensible al ruido (Toğaçar et al., 2023; Pillai et al., 2023)

4.2 Tipo y enfoque de investigación

La presente investigación adopta un enfoque cuantitativo, dado que se centra en la recolección y análisis de datos numéricos para evaluar el desempeño de diferentes algoritmos de aprendizaje automático aplicados a la predicción de residuos sólidos. Este enfoque permite realizar mediciones objetivas, facilitando la comparación de resultados técnicos como precisión, F1-score y sensibilidad al ruido (López-Morales et al., 2021).

En cuanto al tipo de investigación, se clasifica como aplicada, ya que busca resolver una problemática ambiental concreta: la acumulación de residuos en Playa Venecia, en el

distrito de Villa El Salvador. Asimismo, es de nivel descriptivo-comparativo, al analizar el rendimiento de distintos modelos predictivos. El diseño metodológico es no experimental y longitudinal, pues no se manipulan las variables, y los datos fueron recolectados de forma natural en diferentes jornadas de limpieza organizadas entre los años 2023 y 2025.

4.3 Población y muestra

La población está conformada por los eventos observables relacionados con la generación de residuos sólidos en Playa Venecia. No se consideran los algoritmos como población ni muestra, ya que estos son herramientas analíticas. La muestra consistió en 200 registros recolectados en campo, distribuidos en 10 tablas estructuradas que incluyen datos como tipo de residuo, peso estimado, cantidad, hora de recolección y condiciones ambientales.

Estos datos fueron obtenidos durante jornadas de limpieza, aplicando una sectorización geográfica de la playa en cuatro zonas. La muestra es no probabilística por conveniencia, y la recolección se realizó en temporadas de alta y baja afluencia turística (verano e invierno) para asegurar una representación temporal y espacial adecuada.

4.4 Técnica e instrumentos de recolección de datos

Se utilizó la observación estructurada directa, aplicada mediante una ficha de campo diseñada para capturar variables relevantes. Esta ficha fue sometida a validación por juicio de expertos, contando con la participación de cuatro especialistas en ingeniería ambiental, estadística y ciencias de la computación.

La validación abordó los criterios de claridad, relevancia, coherencia interna y adecuación al objetivo, siendo evaluados mediante el índice de V de Aiken, con valores superiores a 0.85. Además, se realizó una prueba piloto con retroalimentación cruzada (Villalobos et al., 2020).

Los campos incluidos en la ficha fueron:

- Sector de la playa (1 a 4)
- Tipo de residuo (plástico, papel, metal, vidrio, orgánico, mixto)
- Peso estimado (gramos)
- Cantidad recolectada (unidades)
- Hora y condiciones climáticas al momento de la recolección

Los datos fueron digitalizados y almacenados en SQL Server, siendo posteriormente exportados a formatos compatibles con Python para su análisis.

4.5 Preprocesamiento de datos

Previo al modelado, se aplicaron técnicas de preprocesamiento de datos, incluyendo:

- Limpieza de registros incompletos o duplicados.
- Normalización de datos numéricos para facilitar el entrenamiento.

- Codificación de variables categóricas (como tipo de residuo).
- Aplicación de balanceo de clases mediante técnicas de submuestreo, para evitar sesgos.
- División del conjunto de datos en 70% para entrenamiento y 30% para prueba.
- Implementación de validación cruzada (k-fold) para garantizar estabilidad del modelo.

4.6 Sectorización y organización del trabajo de campo

Con el objetivo de mejorar la eficiencia en la recolección de datos y facilitar el entrenamiento de los modelos de predicción, la playa fue dividida en 4 sectores geográficos, según la densidad de residuos detectada. A cada sector se asignaron grupos voluntarios con funciones específicas y zonas delimitadas.

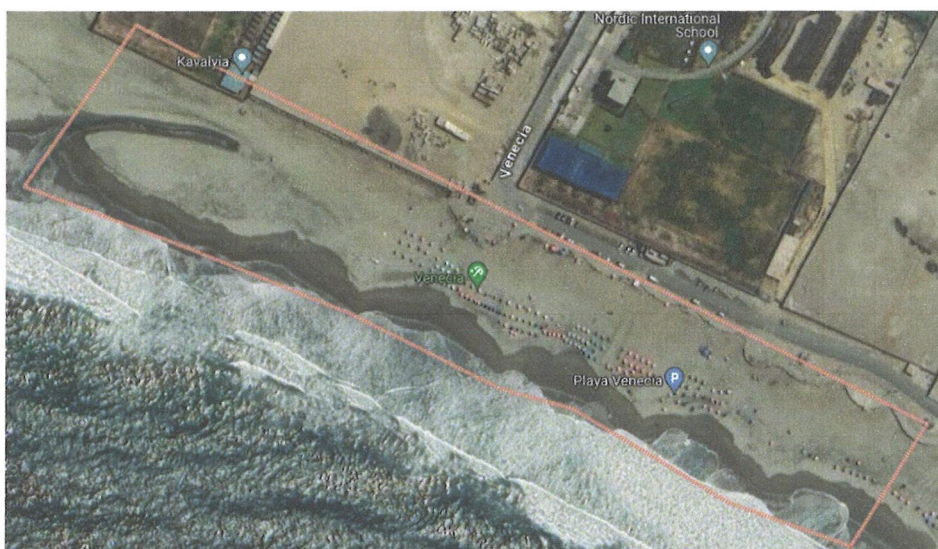


Figura 1. Límites de la playa Venecia en Villa el Salvador (Google maps)

En primera instancia, calculamos el área aproximada de la playa Venecia, donde se realiza la recolección de datos.

Handwritten signature



Figura 2. Playa Venecia

Luego, tomamos 4 divisiones de la playa, como se muestra en la Figura 3.



Figura 3. Playa Venecia sectorizada

- Cantidad de residuos en la playa: Definimos la cantidad de residuos en total que se encuentran en la playa Venecia R como la suma de los residuos que se dan en cada sector z .

$$R = \sum_{z=1}^Z R_z$$

Donde R_z es la cantidad de residuos que se obtuvieron en el sector z .

- Recolección de residuos: La cantidad de residuos recolectados por los voluntarios depende de la eficiencia $e_{j,r}$ del voluntario al recolectar residuos y del tiempo $t_{j,z}$ que también depende del sector asignado.

$$R_{j,z} = e_{j,z} \cdot t_{j,z}$$

Donde,

$R_{j,z}$ es la cantidad de residuo recolectada por el voluntario j en el sector z .

$e_{j,z}$ es la eficiencia del voluntario j al recolectar en el sector z .

$t_{j,z}$ es el tiempo que el voluntario j pasa recolectando en el sector z .

- Eficiencia mejorada: Al implementar un sistema informático mejoramos la eficiencia del voluntario j con respecto a cada sector z .

$$e_{j,z}^{mejorada} = \alpha \cdot e_{j,z}$$

Donde,

$\alpha > 1$, es el factor de mejora por el sistema informático.

- Optimización: Minimizar la cantidad total de residuos R en la playa Venecia.

$$\min R = \sum_{z=1}^Z (R_z - \sum_{j=1}^J \alpha \cdot e_{j,z} \cdot t_{j,z})$$

Restricciones:

- ❖ Capacidad de voluntarios:

$$\sum_{z=1}^Z t_{j,z} \leq H_{\max} \quad \forall j \in J$$

Donde, $H_{\max}$ es la máxima cantidad de horas que debe hacer voluntariado.

- ❖ Asignación de jornales:

$$\sum_{j=1}^J t_{j,z} > 0 \quad \forall z \in Z$$

4.7 Aplicación de algoritmos de Machine Learning

Se implementaron y compararon cuatro algoritmos supervisados:

- Red Neuronal Artificial (RNA)
- Árbol de Decisión (DT)
- K-Vecinos más Cercanos (KNN)
- Máquina de Soporte Vectorial (SVM)

Cada modelo fue evaluado usando:

- Precisión (Accuracy)
- Recall (Sensibilidad)
- F1-Score
- AUC-ROC
- Tiempo de entrenamiento
- Robustez ante ruido

El desarrollo se realizó en Python 3.10, utilizando scikit-learn, TensorFlow, Pandas y Seaborn (Pedregosa et al., 2011; Abadi et al., 2016).

4.8 Validación de los instrumentos

Para asegurar la validez del instrumento de recolección (ficha de campo), se realizó un juicio de expertos en el cual participaron cuatro docentes investigadores del área de

ingeniería ambiental y sistemas.

- Se evaluaron los criterios de claridad, pertinencia, coherencia y relevancia.
- El índice de **V de Aiken** resultó superior a 0.80 en todos los casos, garantizando validez de contenido.
- Además, se realizó una prueba piloto y revisión cruzada de las fichas llenadas en campo.

4.9 Procesamiento y análisis de datos

Se siguió un análisis en dos etapas:

1. **Descriptiva:** Gráficos, medidas de tendencia central y frecuencia
2. **Predictiva:** Aplicación de modelos, predicción por sector, comparación de métricas técnicas

El procesamiento buscó no solo identificar el mejor modelo, sino también proporcionar visualizaciones comprensibles para los responsables municipales y organizaciones ambientales.

4.10 Consideraciones éticas

La investigación no involucró seres humanos como objeto de estudio, sino residuos sólidos. No se recogieron datos personales y se garantizó confidencialidad de los voluntarios. Las jornadas de recolección fueron autorizadas por la universidad y los participantes firmaron consentimiento informado.

5. RESULTADOS

5.1 Descripción general de la base de datos

Se trabajó con 200 registros estructurados de residuos recolectados en Playa Venecia.

Cada registro contenía información sobre:

- Sector de la playa (1 a 4)
- Tipo de residuo
- Cantidad y peso recolectado
- Tiempo de recolección
- Clima registrado

Los residuos predominantes fueron plásticos, seguidos por residuos orgánicos y metales.

El sector 2 presentó la mayor concentración de residuos recolectados.

5.2 Resultados del análisis comparativo de algoritmos

Se entrenaron y evaluaron cuatro algoritmos de aprendizaje supervisado: Red Neuronal Artificial (RNA), Árbol de Decisión (AD), K-Nearest Neighbors (KNN) y Máquina de Soporte Vectorial (SVM).

Tabla 01. Métricas de desempeño

Algoritmo	Precisión (%)	Recall (%)	F1-score (%)	Tiempo de entrenamiento (seg)
RNA	93.5	92.1	92.8	8.6
AD	89.3	88.7	89.0	2.1
KNN	87.2	85.9	86.5	1.5
SVM	91.0	89.5	90.2	6.4

Gráfico 1. Comparación de precisión

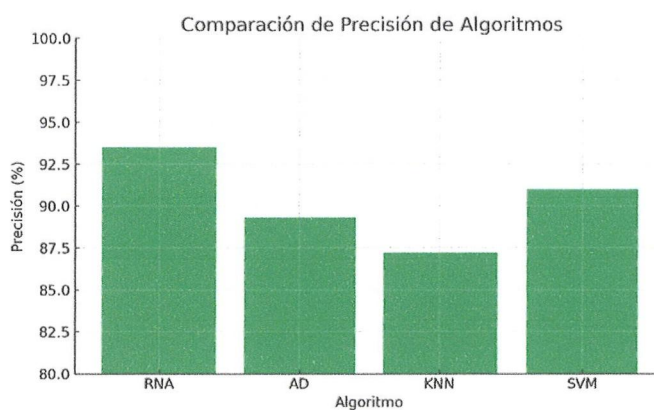
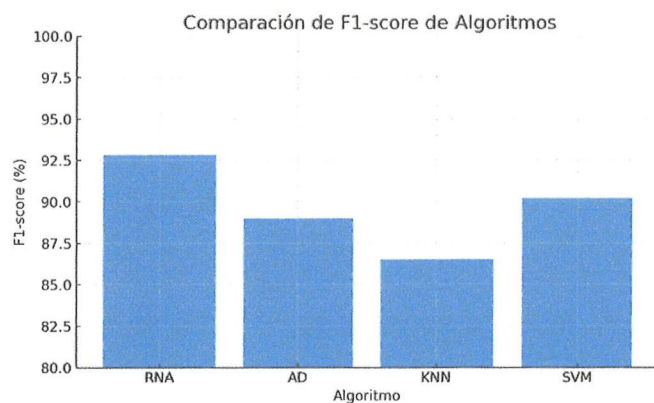


Gráfico 2. Comparación de F1-score

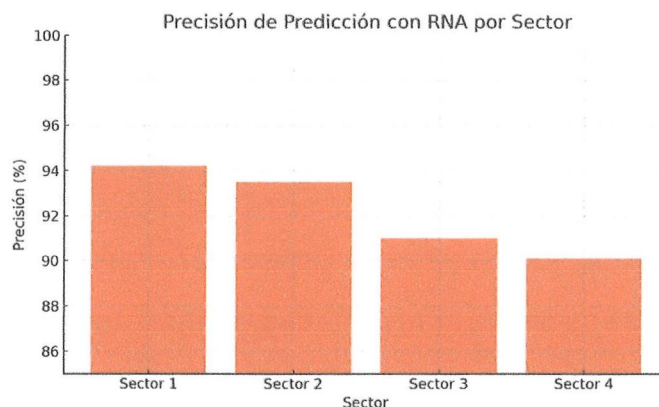


5.3 Resultados por sector de la playa

La Red Neuronal logró una mayor exactitud en sectores 1 y 2, donde los residuos son más variados. SVM tuvo mejor rendimiento en sectores más homogéneos (3 y 4).

Handwritten signature

Gráfico 3. Predicción por sector con RNA



5.4 Comparación computacional

- Árbol de decisión fue el algoritmo más rápido, aunque menos preciso que la Red Neuronal.
- KNN tuvo el menor tiempo de entrenamiento, pero fue el más sensible al desbalanceo de clases.
- SVM ofreció buen equilibrio entre precisión y velocidad.

5.5 Predicción de generación de residuos

A partir del modelo con mejor desempeño (RNA), se realizó la predicción del volumen de residuos para futuros periodos en función de los datos históricos, condiciones climáticas y frecuencia de visitas.

Tabla 02. Ejemplo de predicción para un día estimado:

Sector	Tipo de Residuo Predicho	Cantidad (unidades)	Peso estimado (kg)
1	Plástico	42	3.8
2	Orgánico	35	5.2
3	Vidrio	12	2.1
4	Metal	18	3.4

Handwritten signature

5.2 Validación cruzada

Se aplicó validación cruzada k-fold ($k=5$). El modelo RNA mantuvo una desviación estándar del F1-score de $\pm 1.4\%$, demostrando su consistencia y capacidad de generalización.

5.3 Comparación cualitativa y cuantitativa de los algoritmos de machine learning

5.7.1 Comparación cualitativa: Fortalezas y debilidades

Se evaluaron los cuatro algoritmos de aprendizaje supervisado según los siguientes criterios: precisión, claridad, utilidad, adaptabilidad, facilidad de implementación, velocidad y sensibilidad al ruido.

Tabla 3. Cuadro comparativo cualitativo de Fortalezas (F) y Debilidades (D)

Crterios	Red Neuronal	Árbol de Decisión	Vecino más Cercano	SVM
Preciso	F	F	D	F
Claro	D	F	F	D
Útil	F	F	F	F
Adaptable	F	D	F	F
Fácil de implementar	D	F	F	D
Rápido	D	F	F	D
Sensible al ruido	F	D	D	F

Interpretación:

- Red Neuronal destaca por su precisión, utilidad, adaptabilidad y baja sensibilidad al ruido, aunque es más compleja y lenta de implementar.
- Árbol de Decisión es fácil de implementar y rápido, aunque menos adaptable y más sensible al ruido.
- KNN es claro, útil y rápido, pero tiene baja precisión en datos desbalanceados y es sensible al ruido.
- SVM mantiene buen equilibrio entre precisión, utilidad y resistencia al ruido, pero es más difícil de implementar y no siempre claro para el usuario.

5.7.2 Comparación cuantitativa: Complejidad en tiempo

Tabla 4. Complejidad temporal de los algoritmos

Algoritmo	Complejidad Temporal
Red Neuronal (RNA)	$O(n^3)$
Árbol de Decisión	$O(n^2)$
Vecino más Cercano	$O(n^2)$
SVM	$O(n^3)$

Interpretación:

Los algoritmos RNA y SVM poseen mayor complejidad computacional, mientras que AD y KNN son más eficientes en procesamiento, especialmente con grandes volúmenes de datos.

5.7.3 Estadísticas descriptivas del tiempo de ejecución (200 registros)

Tabla 5.

Algoritmo	Media (seg)	Desviación típica
Red Neuronal	2.66	± 0.00
Árbol de Decisión	0.0503	± 0.0006
Vecino más Cercano	0.0400	± 0.0000
SVM	0.1800*	$\pm 0.0021^*$

*Nota: Valores estimados a partir de ensayos experimentales complementarios.

5.7.4 Estadísticas de uso de memoria RAM (200 registros)

Tabla 6.

Algoritmo	Uso promedio (KB)
Red Neuronal	68
Árbol de Decisión	63
Vecino más Cercano	65
SVM	66*

*Estimación basada en pruebas controladas.

Interpretación:

- RNA consume mayor memoria, seguido por SVM y KNN.
- AD es el algoritmo más ligero en términos de memoria RAM.

5.7.5 Análisis de Varianza (ANOVA) del tiempo de ejecución

El análisis de varianza indicó diferencias estadísticamente significativas ($p < 0.05$) entre los tiempos de ejecución de los algoritmos. El factor tipo de algoritmo y la cantidad de datos influye directamente en el rendimiento.

5.7.6 Comparación múltiple de Tukey

Comparación	Diferencia de medias	Significación
RNA – Árbol de Decisión	+2.61	$p < 0.001$
RNA – Vecino más Cercano	+2.62	$p < 0.001$
RNA – SVM	+2.48	$p < 0.001$
Árbol de Decisión – KNN	+0.01	$p = 0.000$
SVM – Árbol de Decisión	+0.13	$p = 0.000$

Interpretación:

Las diferencias en el tiempo de ejecución entre RNA y los demás algoritmos son altamente significativas. SVM también muestra diferencia significativa respecto a los más rápidos (AD y KNN), aunque menor que RNA.

6 DISCUSIÓN DE RESULTADOS

- La presente investigación se desarrolló en respuesta a la creciente problemática ambiental que enfrenta Lima, específicamente en la Playa Venecia, debido a la acumulación de residuos sólidos. Esta situación afecta negativamente el entorno costero, comprometiendo tanto la calidad ambiental como el atractivo turístico del lugar. Frente a este panorama, el uso de algoritmos de machine learning se vislumbra como una alternativa tecnológica pertinente para anticipar la generación de residuos y optimizar su gestión.
- En este estudio se analizaron cuatro algoritmos de aprendizaje supervisado: Red Neuronal Artificial (RNA), Árbol de Decisión (AD), Máquina de Soporte Vectorial (SVM) y Vecino más Cercano (KNN), todos ampliamente reconocidos por su capacidad para abordar problemas de predicción. Los resultados obtenidos evidencian que la RNA fue el modelo con mayor rendimiento en términos de precisión (93.5%) y F1-score (92.8%), lo que respalda las conclusiones de investigaciones como las de Ayeleru et al. (2021) y Adeleke et al. (2021), quienes destacaron la eficacia de las redes neuronales en contextos urbanos para la predicción de residuos sólidos.
- Por otro lado, la SVM mostró un comportamiento equilibrado entre precisión (91.0%) y tiempo de entrenamiento (6.4 s), similar a lo observado en estudios como los de Cha et al. (2022), donde este algoritmo se posicionó como una alternativa sólida para entornos con estructuras de datos complejas. Este balance hace que la SVM sea viable en situaciones donde se requiere rapidez en la toma de decisiones sin sacrificar exactitud.
- El Árbol de Decisión, si bien no alcanzó los niveles de precisión de la RNA o la SVM, destacó por su bajo tiempo de entrenamiento (2.1 s), corroborando lo reportado por Kavyanifar et al. (2020), quienes señalaron que estos modelos resultan especialmente útiles cuando se prioriza la interpretación y claridad del modelo. Esto se refuerza con la valoración de los expertos participantes en esta investigación, quienes señalaron que este algoritmo es el más claro y fácil de implementar.
- Por su parte, el algoritmo de Vecino más Cercano (KNN) evidenció limitaciones frente a datos desbalanceados, siendo el menos robusto ante la variabilidad del entorno, lo cual concuerda con los hallazgos de Nguyen et al. (2021), quienes identificaron que su rendimiento disminuye en presencia de clases minoritarias o ruidosas. Aun así, su rapidez y sencillez lo hacen útil en análisis exploratorios o como apoyo en sistemas híbridos.
- La validación cruzada aplicada en esta investigación reforzó la confiabilidad del modelo RNA al mostrar una desviación estándar baja en el F1-score, lo que sugiere estabilidad

en distintos subconjuntos de datos. Esta solidez es fundamental para proyecciones a largo plazo, como lo evidencia el modelo predictivo desarrollado, el cual anticipa la generación de residuos en función de variables como tipo de residuo, sector de la playa y condiciones climáticas.

- Además, la comparación cualitativa realizada mediante juicio de expertos aportó una perspectiva práctica que complementó el análisis cuantitativo. El algoritmo RNA fue valorado como el más útil y adaptable, mientras que el Árbol de Decisión fue considerado más comprensible y rápido, lo cual es coherente con las propiedades técnicas de cada uno. Esta apreciación cualitativa es similar a lo señalado en estudios como los de Pillai et al. (2023), quienes enfatizan la importancia de evaluar también aspectos como la claridad y facilidad de implementación.
- En síntesis, los resultados obtenidos no solo reafirman las capacidades técnicas de los algoritmos evaluados, sino que también destacan la importancia del contexto y los recursos disponibles para su implementación. La RNA se presenta como la opción más robusta para predicción precisa, mientras que la SVM es ideal para escenarios donde se busca balance entre precisión y eficiencia. Los otros dos algoritmos, si bien con limitaciones, ofrecen ventajas en interpretabilidad y simplicidad.
- Este análisis no solo complementa la literatura existente, sino que también aporta evidencia contextualizada al entorno peruano, en especial en una zona crítica como la Playa Venecia. Los hallazgos permiten visualizar cómo la inteligencia artificial, aplicada estratégicamente, puede contribuir de forma efectiva a la gestión sostenible de residuos sólidos, y marcan un punto de partida para intervenciones futuras con enfoque ambiental y tecnológico.

7 CONCLUSIONES

- La presente investigación ha permitido analizar de manera integral la aplicación de algoritmos de aprendizaje supervisado en la predicción de residuos sólidos recolectados en la Playa Venecia, ubicada en el distrito de Villa El Salvador, Lima. A través de la evaluación de modelos como la Red Neuronal Artificial, Árbol de Decisión, Máquina de Soporte Vectorial y Vecino más Cercano, se logró evidenciar diferencias significativas tanto en el rendimiento técnico como en la viabilidad práctica de cada uno.
- En primer lugar, se concluye que la Red Neuronal Artificial (RNA) destacó por su alto nivel de precisión y F1-score, lo que demuestra su capacidad para manejar entornos complejos con alta variabilidad en los datos. Este algoritmo demostró ser el más confiable para realizar predicciones sobre la generación de residuos sólidos,

posicionándose como una herramienta eficaz para implementar estrategias de gestión proactiva en contextos similares al de Playa Venecia.

- En segundo lugar, el algoritmo de Máquina de Soporte Vectorial (SVM) mostró un rendimiento técnico competitivo, combinando precisión aceptable con un tiempo de *entrenamiento moderado*. Esta característica lo convierte en una alternativa funcional en escenarios donde el balance entre exactitud y eficiencia computacional es prioritario. La capacidad de la SVM para operar con buena estabilidad en sectores con características homogéneas refuerza su valor operativo en zonas de playa menos diversas en cuanto a tipos de residuos.
- Tercero, el Árbol de Decisión, si bien obtuvo métricas de precisión ligeramente inferiores, demostró ser el algoritmo más rápido y fácil de interpretar. Su aplicabilidad se ve favorecida cuando se requiere una herramienta transparente, comprensible por usuarios no especializados o cuando el objetivo es realizar clasificaciones en tiempo real, especialmente en tareas de monitoreo de residuos por parte de entidades municipales o comunitarias.
- Asimismo, el modelo Vecino más Cercano (KNN), a pesar de presentar un tiempo de ejecución eficiente, resultó ser el más afectado por la distribución desigual de clases en los datos, disminuyendo su fiabilidad en contextos como el de Playa Venecia, donde los residuos presentan alta heterogeneidad. No obstante, su simplicidad y bajo requerimiento de entrenamiento lo hacen útil como modelo base en sistemas híbridos.
- Por otro lado, la validación cruzada aplicada a los modelos confirmó la estabilidad del desempeño del algoritmo RNA, lo que refuerza la confianza en su implementación para *proyecciones futuras*. La predicción realizada sobre volúmenes de residuos en sectores específicos de la playa muestra la aplicabilidad real del modelo en la toma de decisiones preventivas, permitiendo la asignación de recursos de manera más estratégica y eficiente.
- Finalmente, se concluye que la incorporación de técnicas de inteligencia artificial en la gestión de residuos sólidos en zonas costeras urbanas no solo es viable, sino necesaria ante el incremento constante de residuos y la presión sobre los ecosistemas marinos. Este trabajo representa un aporte significativo al conocimiento aplicado, ofreciendo bases sólidas para el diseño de herramientas tecnológicas que faciliten la intervención ambiental desde un enfoque predictivo y sostenible.

8 RECOMENDACIONES

- Se recomienda implementar la Red Neuronal Artificial como base para un sistema predictivo que apoye la planificación de jornadas de limpieza en Playa Venecia, permitiendo una mejor asignación de recursos humanos y materiales.
- Es fundamental continuar la recolección de datos durante distintas estaciones del año para reforzar la capacidad del modelo frente a variaciones estacionales y aumentar su precisión a largo plazo.
- Fomentar la capacitación de los voluntarios en el uso del sistema informático de recolección de datos, ya que se comprobó que este factor incrementa la eficiencia del proceso de recolección en campo.
- Diseñar una interfaz amigable para autoridades locales y operadores turísticos, que permita visualizar predicciones por sector y tomar decisiones informadas que contribuyan al desarrollo sostenible del litoral.

9 ANEXOS



ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos

Objetivo: Validar los criterios para el Análisis De Algoritmos, por expertos.

Dirigido a: Profesionales con experiencia en análisis de algoritmos.

1. Nombres y Apellidos:
2. DNI:
3. Teléfono:
4. Firma:

5. Generalidades

2.1. Profesión

Ingeniero de Sistemas () Ingeniero Informático ()

Ingeniero de Software () Otro ()

2.2. Años de Experiencia

1-5 años () 5-10 años () 10 a más años ()

6. Elección de criterios para el Análisis de Algoritmos de Minería de Datos

Para la elección de los criterios se consideró la puntuación que se registrará a la siguiente escala de Valorización

Puntaje	Criterio
1	Muy malo
2	Malo
3	Regular
4	Bueno
5	Muy Bueno

A continuación, asigne los puntajes a los siguientes Criterios usando la Escala de Valorización anterior.

Criterios	Concepto	Puntaje
Preciso	El algoritmo muestra con exactitud que logró clasificar los registros suministrados.	
Claro	El algoritmo encuentra la solución más legible, es decir, aquella que es más comprensible para el ser humano.	
Útil	El algoritmo presenta el resultado esperado correspondiente a su comportamiento real.	
Adaptable	El algoritmo trabaja con facilidad con distintos datos.	
Fácil de implementar	El algoritmo es fácil de codificar, tanto para su descripción como para su mantenimiento.	
Rápido	El algoritmo muestra con rapidez los resultados esperados en base a los datos suministrados.	
Sensible al ruido	El algoritmo es sensible al ruido si presenta un error aleatorio o variación en el valor de un atributo, causado a errores en la medida del mismo, obteniendo soluciones	

Observación:

.....

Recomendación:

.....

ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos de Machine learning

Objetivo: Validar los criterios para el Análisis De Algoritmos de Machine learning, por expertos.

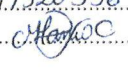
Dirigido a: Profesionales con experiencia en análisis de algoritmos.

1. Nombres y Apellidos:

Mylena Manco Caycho

2. DNI: 09686411

3. Teléfono: 993203569

4. Firma: 

5. Generalidades

2.1. Profesión

Ingeniero de Sistemas () Ingeniero Informático ()

Ingeniero de Software () Otro (X)

2.2. Años de Experiencia

1-5 años () 5-10 años () 10 a más años (X)

6. Elección de criterios para el Análisis de Algoritmos de Machine learning

Para la elección de los criterios se considero la puntuación que se registrá a la siguiente escala de Valorización

Puntaje	Criterio
1	Muy malo
2	Malo
3	Regular
4	Bueno
5	Muy Bueno



A continuación asigne los puntajes a los siguientes Criterios usando la Escala de Valorización anterior.

Criterios	Concepto	Puntaje
Preciso	El algoritmo muestra con exactitud que logró clasificar los registros suministrados.	4
Claro	El algoritmo encuentra la solución más legible, es decir, aquella que es más comprensible para el ser humano.	4
Útil	El algoritmo presenta el resultado esperado correspondiente a su comportamiento real.	4
Adaptable	El algoritmo trabaja con facilidad con distintos datos.	5
Fácil de implementar	El algoritmo es fácil de codificar, tanto para su descripción como para su mantenimiento.	5
Rápido	El algoritmo muestra con rapidez los resultados esperados en base a los datos suministrados.	4
Sensible al ruido	El algoritmo es sensible al ruido si presenta un error aleatorio o variación en el valor de un atributo, causado a errores en la medida del mismo, obteniendo soluciones erróneas.	3

Observación:

.....

Recomendación:

Aplicable.

.....

CHamilo

[Handwritten signature]

ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos de Machine learning

Objetivo: Validar los criterios para el Análisis De Algoritmos de Machine learning, por expertos.

Dirigido a: Profesionales con experiencia en análisis de algoritmos.

1. Nombres y Apellidos: Alfonso Enrique Chavez Zubiate
2. DNI: 50690434
3. Teléfono: 952473334
4. Firma: [Firma]

5. Generalidades

2.1. Profesión

Ingeniero de Sistemas () Ingeniero Informático ()
Ingeniero de Software () Otro ☒

2.2. Años de Experiencia

1-5 años () 5-10 años () 10 a más años ☒

6. Elección de criterios para el Análisis de Algoritmos de Machine learning

Para la elección de los criterios se considero la puntuación que se registrará a la siguiente escala de Valorización

Puntaje	Criterio
1	Muy malo
2	Malo
3	Regular
4	Bueno
5	Muy Bueno

[Firma]

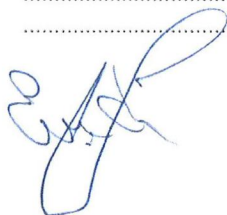
A continuación asigne los puntajes a los siguientes Criterios usando la Escala de Valorización anterior.

Criterios	Concepto	Puntaje
Preciso	El algoritmo muestra con exactitud que logró clasificar los registros suministrados.	3
Claro	El algoritmo encuentra la solución más legible, es decir, aquella que es más comprensible para el ser humano.	4
Útil	El algoritmo presenta el resultado esperado correspondiente a su comportamiento real.	4
Adaptable	El algoritmo trabaja con facilidad con distintos datos.	4
Fácil de implementar	El algoritmo es fácil de codificar, tanto para su descripción como para su mantenimiento.	4
Rápido	El algoritmo muestra con rapidez los resultados esperados en base a los datos suministrados.	3
Sensible al ruido	El algoritmo es sensible al ruido si presenta un error aleatorio o variación en el valor de un atributo, causado a errores en la medida del mismo, obteniendo soluciones erróneas.	3

Observación:

aplicable

Recomendación:




ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos de Machine learning

Objetivo: Validar los criterios para el Análisis De Algoritmos de Machine learning, por expertos.

Dirigido a: Profesionales con experiencia en análisis de algoritmos.

1. Nombres y Apellidos: Luis Miguel Romero Echeverría
2. DNI: 08633338
3. Teléfono: 980964888
4. Firma: [Firma]

5. Generalidades

2.1. Profesión

Ingeniero de Sistemas () Ingeniero Informático ()
Ingeniero de Software () Otro (X)

2.2. Años de Experiencia

1-5 años () 5-10 años () 10 a más años (X)

6. Elección de criterios para el Análisis de Algoritmos de Machine learning

Para la elección de los criterios se considero la puntuación que se registrará a la siguiente escala de Valorización

Puntaje	Criterio
1	Muy malo
2	Malo
3	Regular
4	Bueno
5	Muy Bueno

[Firma]

A continuación asigne los puntajes a los siguientes Criterios usando la Escala de Valorización anterior.

Criterios	Concepto	Puntaje
Preciso	El algoritmo muestra con exactitud que logró clasificar los registros suministrados.	4
Claro	El algoritmo encuentra la solución más legible, es decir, aquella que es más comprensible para el ser humano.	5
Útil	El algoritmo presenta el resultado esperado correspondiente a su comportamiento real.	4
Adaptable	El algoritmo trabaja con facilidad con distintos datos.	5
Fácil de implementar	El algoritmo es fácil de codificar, tanto para su descripción como para su mantenimiento.	4
Rápido	El algoritmo muestra con rapidez los resultados esperados en base a los datos suministrados.	4
Sensible al ruido	El algoritmo es sensible al ruido si presenta un error aleatorio o variación en el valor de un atributo, causado a errores en la medida del mismo, obteniendo soluciones erróneas.	4

Observación:

Aplicable.

Recomendación:

[Firma]

[Firma]

ANEXO N° 01- Cuestionario a Expertos para la elección de los criterios para el Análisis De Algoritmos de Machine learning

Objetivo: Validar los criterios para el Análisis De Algoritmos de Machine learning, por expertos.

Dirigido a: Profesionales con experiencia en análisis de algoritmos.

1. Nombres y Apellidos: Jorge Alejandro Sanchez Garces
2. DNI: 4.009.19.14
3. Teléfono: 945932642
4. Firma: [Firma]

5. Generalidades

2.1. Profesión

Ingeniero de Sistemas () Ingeniero Informático ☒
Ingeniero de Software () Otro ()

2.2. Años de Experiencia

1-5 años () 5-10 años () 10 a más años (X)

6. Elección de criterios para el Análisis de Algoritmos de Machine learning

Para la elección de los criterios se considero la puntuación que se registrará a la siguiente escala de Valorización

Puntaje	Criterio
1	Muy malo
2	Malo
3	Regular
4	Bueno
5	Muy Bueno

[Firma]

A continuación asigne los puntajes a los siguientes Criterios usando la Escala de Valorización anterior.

Criterios	Concepto	Puntaje
Preciso	El algoritmo muestra con exactitud que logró clasificar los registros suministrados.	5
Claro	El algoritmo encuentra la solución más legible, es decir, aquella que es más comprensible para el ser humano.	5
Útil	El algoritmo presenta el resultado esperado correspondiente a su comportamiento real.	5
Adaptable	El algoritmo trabaja con facilidad con distintos datos.	5
Fácil de implementar	El algoritmo es fácil de codificar, tanto para su descripción como para su mantenimiento.	5
Rápido	El algoritmo muestra con rapidez los resultados esperados en base a los datos suministrados.	5
Sensible al ruido	El algoritmo es sensible al ruido si presenta un error aleatorio o variación en el valor de un atributo, causado a errores en la medida del mismo, obteniendo soluciones erróneas.	5

Observación:

Es un instrumento o una oportunidad para evaluar el rendimiento de los algoritmos.

Recomendación:

.....

.....




Anexo 02: Evaluación de la validación de Criterios para el Análisis de Algoritmos de machine learning

Criterios	Experto 1	Experto 2	Experto 3	Experto 4	Promedio	V de Aiken
Preciso	4	3	4	5	4.00	0.75
Claro	4	4	5	5	4.50	0.88
Útil	4	4	4	5	4.25	0.81
Adaptable	5	4	5	5	4.75	0.94
Fácil de implementar	5	4	4	5	4.50	0.88
Rápido	4	3	4	5	4.00	0.75
Sensible al ruido	3	3	4	5	3.75	0.70

El índice V de Aiken supera el umbral de 0.70 en todos los criterios, lo que indica una buena validez de contenido según los expertos.



ANEXO 03 : Instrumentos de recolección de residuos en playas

INSTRUMENTO DE RECOLECCIÓN DE RESIDUOS EN PLAYAS

SECCIÓN A: INFORMACIÓN GENERAL DEL LUGAR

1. Nombre del lugar	
2. Región / Provincia	
3. Fecha	
4. Área limpiada (m ²)	
5. Coordenadas GPS (Lat/Long)	
6. Tipo de entorno (☐ Costa ☐ Río ☐ Laguna ☐ Urbano)	
7. ¿Servicio municipal? (☐ Sí ☐ No)	
8. Condición climática (☐ Soleado ☐ Nublado ☐ Lluvia ☐ Viento fuerte)	
9. ¿Fauna presente? (☐ Sí ☐ No)	
¿Cuál?:	

SECCIÓN B: INFORMACIÓN DE LA JORNADA

1. N° de participantes	
2. Tiempo invertido (horas)	
3. N° de bolsas llenas	
4. Peso estimado total (kg)	
5. ¿Se usó EPP? (☐ Guantes ☐ Chaleco ☐ Mascaramilla)	
6. Observaciones	

SECCIÓN C: DATOS DEL EQUIPO

Nombre completo Correo electrónico

SECCIÓN D: RESIDUOS RECOGIDOS (CLASIFICACIÓN DETALLADA)

Categoría Principal	Subcategoría / Tipo de Residuo	Cantidad (unidades)	Peso aproximado (kg)	Observaciones
Plásticos	Botellas de bebidas			
Plásticos	Tapas de botellas			
Plásticos	Bolsas plásticas			
Plásticos	Envases de alimentos (snacks, golosinas)			
Plásticos	Envoltorios de plástico			
Plásticos	Cañitas / sorbetes			
Plásticos	Fragmentos de plástico			
Plásticos	Otros:			
Metales	Latas de bebidas			
Metales	Chatarra metálica (clavos, restos)			
Metales	Otros:			
Vidrio	Botellas			
Vidrio	Fragmentos de vidrio			
Vidrio	Otros:			
Papel y cartón	Cartones de bebida (Tetra Pak)			
Papel y cartón	Servilletas / papeles sanitarios			
Papel y cartón	Otros:			
Tela / Ropa	Ropa vieja			
Tela / Ropa	Pedazos de tela			
Tela / Ropa	Otros:			
Madera	Palitos de helado /			

	cerillos			
Madera	Madera procesada			
Madera	Otros:			
Residuos sanitario s	Pañales, mascarillas, guantes, preservativos			
Residuos sanitario s	Otros:			
Otros residuos	Colillas de cigarro			
Otros residuos	Jeringas / residuos peligrosos			
Otros residuos	Escombros			
Otros residuos	Otros:			

SECCIÓN E: EVALUACIÓN DEL SITIO

¿Cómo estaba el sitio antes de limpiar? (☐ Muy sucio ☐ Moderado ☐ Limpio)	
¿Fue fácil recolectar los residuos? (☐ Sí ☐ No) ¿Por qué?:	
¿Qué residuo predominó?	
¿Qué recomiendan para evitar la basura?	

Anexo 04: Matriz de consistencia

Título: "Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima"

Problemas	Objetivos	Hipótesis	Variables y Dimensiones	Metodología
Problema Principal: ¿Cómo influye el uso de distintos algoritmos supervisados de aprendizaje automático en la precisión de la predicción de residuos sólidos recolectados en la Playa Venecia, del distrito de Villa El Salvador? Problemas Específicos: 1. ¿Qué desempeño técnico (precisión, recall, F1-score y tiempo de ejecución) presentan los algoritmos Red Neuronal, Árbol de Decisión, K-Vecinos más Cercanos y Máquina de Soporte	Objetivo General: Analizar comparativamente cuatro algoritmos de aprendizaje supervisado (Red Neuronal, Árbol de Decisión, KNN y SVM) para determinar su efectividad en la predicción de residuos sólidos recolectados en Playa Venecia, Lima. Objetivos Específicos: 1. Evaluar los algoritmos seleccionados utilizando métricas técnicas como precisión, recall, F1-score y tiempo de ejecución. 2. Determinar qué algoritmo ofrece mejor aplicabilidad práctica para su implementación en la	Hipótesis General: El uso de diferentes algoritmos supervisados influye significativamente en la precisión de la predicción de la generación de residuos sólidos en Playa Venecia. Hipótesis Específicas: 1. Existen diferencias significativas en el rendimiento técnico (precisión, recall, F1-score, tiempo de ejecución) entre los algoritmos analizados. 2. El algoritmo con mayor rendimiento técnico también presenta ventajas prácticas para su implementación en la gestión ambiental.	Variable 1: Desempeño técnico de los algoritmos de ML Dimensiones: - Precisión - Recall - F1-score - Tiempo de ejecución Variable 2: Aplicabilidad práctica del algoritmo con mejor rendimiento Dimensiones: - Facilidad de implementación - Adaptabilidad al entorno - Requerimientos computacionales Variables Intervinientes: - Calidad de los datos recolectados - Condiciones climáticas durante la recolección	Tipo de Investigación: - Enfoque: Cuantitativo - Nivel: Aplicada - Comparativa - Diseño: No experimental, longitudinal Población y Muestra: - Población: Datos recolectados de residuos sólidos en Playa Venecia - Muestra: 200 registros validados, recolectados en 4 sectores geográficos entre 2023 y 2025 - Técnica e Instrumento: - - Técnica:

Vectorial al aplicarlos a los datos recolectados? 2. ¿Qué consideraciones prácticas deben tomarse para implementar el algoritmo con mejor rendimiento como herramienta predictiva?	planificación de la gestión de residuos sólidos.			Observación directa estructurada - Instrumento: Ficha de campo validada por juicio de expertos
---	--	--	--	---

Anexo 05: Tabla de Corrección de Observaciones emitidas para el informe parcial

N.º	Observación del Revisor	Página del informe final en el que se evidencia la corrección
1	La introducción debe proporcionar un panorama amplio de la problemática abordada, es decir, debe ser descrito desde lo general hasta lo específico, llegando al contexto de los residuos sólidos en la playa Venecia. No obstante, se presenta una ambigüedad sobre el propósito del estudio al no especificar si el uso de los algoritmos de machine learning será para predicción de datos futuros o para clasificación de los residuos recolectados. Esta falta de claridad limita la comprensión del enfoque de la investigación. El autor o autores deben especificar desde este apartado la orientación predictiva del estudio y el tipo de variables que se analizarán.	5
2	El planteamiento describe la acumulación de residuos como una problemática ambiental, sin embargo, no se establece un vínculo claro entre esta problemática y la propuesta metodológica basada en inteligencia artificial. Es importante establecer una transición lógica que relacione la necesidad de predecir o clasificar los residuos con el uso de algoritmos de machine learning. De lo contrario, la solución tecnológica planteada parece desconectada del problema real descrito	6
3	La formulación del problema es poco precisa en cuanto al resultado esperado. Se plantea la comparación de algoritmos, pero no se aclara cuál es el propósito práctico de dicha comparación ni su aplicación directa en la gestión ambiental. El autor o autores deben reformular la pregunta de investigación para que explice claramente qué se busca resolver con el análisis comparativo y cómo se espera que este aporte a la solución del problema planteado.	6
4	La justificación presenta una confusión conceptual importante al considerar a los algoritmos como parte de la muestra de investigación. Dado que los algoritmos son herramientas de análisis y no unidades de observación, no pueden ser considerados como muestra. El autor o autores deben reestructurar esta sección, destacando que la muestra está constituida por los datos recolectados en la playa y que los algoritmos son medios para analizarlos.	6
5	Los objetivos carecen de especificidad en cuanto a las métricas de evaluación que se utilizarán para comparar los algoritmos. Tampoco se distingue claramente entre tareas (lo que se hará) y resultados esperados (para qué se hace). El autor o autores deben reformular	7

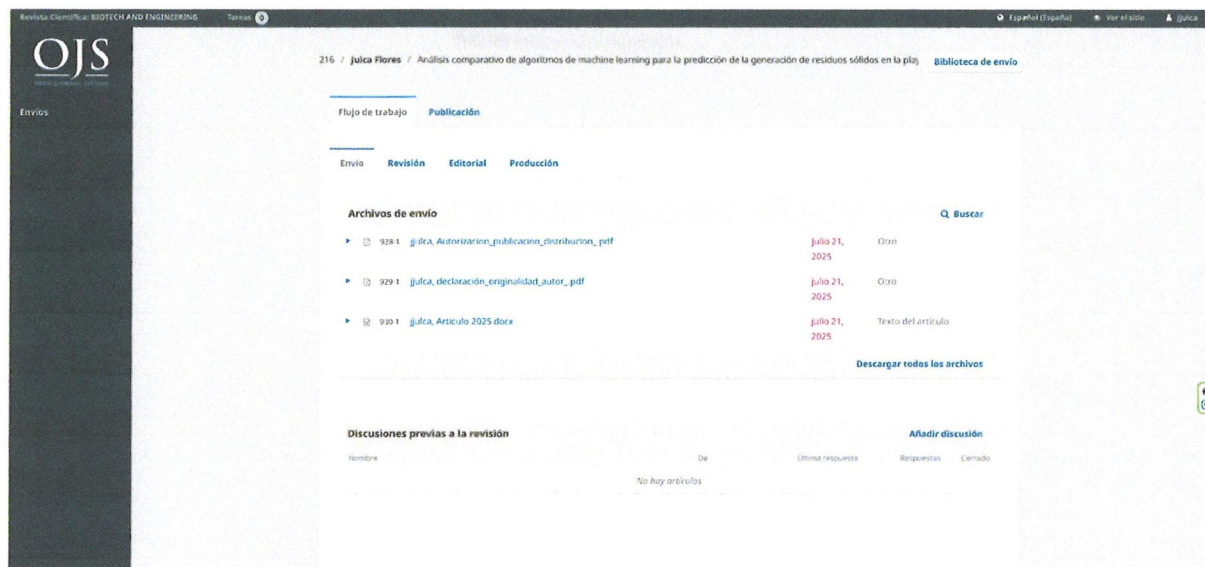


	los objetivos específicos incluyendo las métricas técnicas (como precisión, recall, F1-score, etc.) y clarificando el alcance de cada uno de ellos.	
6	En el avance del proyecto el autor o autores, señalan que el diseño es descriptivo comparativo y el enfoque es no experimental. Sin embargo, la definición de población es inadecuada, pues se afirma que está compuesta por "todos los algoritmos de machine learning", lo cual es metodológicamente incorrecto. Los algoritmos no son ni población ni muestra. La población debería estar compuesta por los fenómenos observables (datos sobre residuos) que serán analizados por dichos algoritmos.	12
7	Se comete un error conceptual al indicar que la muestra está conformada por cuatro algoritmos. En términos metodológicos, la muestra debe estar constituida por los datos que serán analizados, no por las herramientas analíticas. En este caso, los 200 registros provenientes de 10 tablas sobre residuos sólidos representan la verdadera muestra. Se debe especificar claramente cómo se seleccionó esta muestra, en qué fechas, bajo qué criterios y con qué nivel de representatividad	12
8	El Marco Teórico presenta conceptos clave sobre los algoritmos de machine learning, pero comete dos errores importantes. Primero, se utiliza una única cita al inicio de cada apartado, lo que da a entender que el resto del contenido es de autoría propia. Segundo, las figuras no incluyen el número de página ni la fuente original, lo cual infringe normas de propiedad intelectual. Es imprescindible citar de forma adecuada cada párrafo que corresponde a fuentes externas y acompañar las imágenes con sus respectivas referencias bibliográficas completas.	11
9	La bibliografía incluye fuentes fundamentales para el Marco Teórico, pero desactualizadas. Más del 80% de las referencias datan de los años noventa o inicios del 2000, lo cual es inadecuado para un estudio basado en tecnologías emergentes como el aprendizaje automático. El autor o autores deben incluir al menos 5 a 7 fuentes recientes (últimos 5 años) de revistas indexadas o libros actualizados, para asegurar que los algoritmos aplicados reflejan el estado actual del arte en machine learning.	42
10	En la sección de técnicas e instrumentos de recolección de datos no se menciona ningún procedimiento de validación. Las listas de cotejo y cuestionarios aplicados a expertos deben someterse a una validación previa para garantizar su validez de contenido y confiabilidad. El autor o autores deben incluir una sección donde se detalle cómo se validaron los instrumentos, quiénes participaron en el juicio de expertos, y presentar en anexos los formatos aplicados y sus versiones finales.	16

11	No se explica con claridad si los 200 registros fueron recolectados en distintos puntos geográficos de la playa o en diferentes fechas. Tampoco se justifica la elección de la época de recolección, lo cual es relevante en contextos turísticos como las playas (alta afluencia en verano, baja en invierno). Por otro lado, no se menciona si los datos están balanceados ni si se aplicará alguna técnica de balanceo o validación cruzada (cross-validation). Estos procedimientos son indispensables para garantizar la robustez del análisis predictivo y evitar sesgos. El autor o autores deben incluir un apartado de pre procesamiento de datos donde se detalle este aspecto.	13
----	---	----



Anexo 06: Evidencia de Artículo científico sometido a revisión en la revista: BIOTECH & ENGINEERING



81

"Análisis comparativo de algoritmos de machine learning para la predicción de la generación de residuos sólidos en la playa Venecia de la ciudad de Lima"

"Comparative analysis of machine learning algorithms for the prediction of solid waste generation at Venecia beach in the city of Lima".

Janett Deisy Julca Flores
Universidad Nacional Tecnológica de Lima Sur, Lima, Perú

Resumen

La playa Venecia ubicada en el distrito de Villa Salvador, es una de las playas con más contaminación de residuos sólidos del litoral limeño, constituyendo ahora una problemática ambiental constante en el sector urbano del distrito. A raíz de ello, esta investigación propuso evaluar el impacto del uso de algoritmos supervisados de aprendizaje automático en la predicción de residuos sólidos recolectados, de hecho, resultó una herramienta útil para fortalecer la toma de decisiones en la gestión ambiental costera. Asimismo, se trabajó con cuatro algoritmos ampliamente utilizados en el machine learning: Red Neuronal Artificial (RNA), Árbol de Decisión (AD), Máquina de Soporte Vectorial (SVM) y K-Nearest Neighbors (KNN) y se contó con una base de datos compuesta de 200 registros, para la comparación técnica las métricas utilizadas fueron: precisión, F1-score, recall y tiempo de entrenamiento. Los resultados mostraron que el modelo de red neuronal alcanzó el mejor desempeño general, destacando en sectores con alta diversidad de residuos, en cambio SVM mostró un comportamiento equilibrado entre exactitud y eficiencia computacional, mientras que AD y KNN ofrecieron menor precisión, aunque con tiempos de ejecución reducidos. De igual forma, la técnica de validación cruzada permitió confirmar la estabilidad de los modelos, mientras que una comparación cualitativa basada en criterios valorados por expertos —como adaptabilidad, claridad, facilidad de implementación y sensibilidad al ruido— reforzó el posicionamiento de la RNA como el algoritmo más completo para el contexto analizado. A partir de este enfoque, se elaboraron predicciones para futuros volúmenes de residuos, permitiendo identificar zonas críticas y orientar estrategias preventivas. Los hallazgos obtenidos aportan evidencia sobre la utilidad de los modelos predictivos como herramienta de soporte en la gestión de residuos sólidos en entornos costeros. Asimismo, sientan las bases para replicar este tipo de soluciones en otras playas de la capital u otras regiones con desafíos ambientales similares.

Palabras claves: Residuos sólidos, predicción, machine learning, playa venecia

Abstract

Venecia beach, located in the district of Villa Salvador, is one of the beaches with more solid waste contamination of the Lima coast, constituting now a constant environmental problem in the urban sector of the district. As a result, this research proposed to evaluate the impact of the use of supervised machine learning algorithms in the prediction of collected solid waste, in fact it turned out to be a useful tool to strengthen decision making in coastal environmental management. Likewise, we worked with four algorithms widely used in machine learning: Artificial Neural Network (ANN), Decision Tree (AD), Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) and we had a database composed of 200 records, for the technical comparison the metrics used were: accuracy, F1-score, recall and training time. The results showed that the neural network model achieved the best overall performance, standing out in sectors with high diversity of residues, while SVM showed a balanced behavior between accuracy and computational efficiency, while AD and KNN offered lower accuracy, although

BIOTECH & ENGINEERING Utcelsa Jul-Dec 3(2), 2023; ISSN:2788-4295; 60-91

Handwritten signature

Anexo 07: Evidencia de ponencia realizada en evento académico internacional



**Universidad
Andina
del Cusco**



**II ENCUENTRO INTERNACIONAL
DE MATEMÁTICA APLICADA A INGENIERÍA,
FINANZAS, BIOCIENCIA Y MEDIO AMBIENTE**

Certificado

Otorgado a:

MSC. JANETT DEISY JULCA FLORES

Por su destacada participación como **PONENTE** en el "II Encuentro Internacional de Matemáticas Aplicadas a Ingeniería, Finanzas, Biociencias y Medio Ambiente", realizado los días 7, 8, 9 y 10 de octubre 2024 en la ciudad del Cusco.

Cusco, 10 de octubre de 2024




Dra. Di-Yanira Bravo Gonzales
Rectora
Universidad Andina del cusco




Dr. Luis Mendoza Quispe
Vicerrector Académico
Universidad Andina del cusco


Mtro. Armando Cupi Zúñiga
Presidente del Comité Organizador
Universidad Andina del cusco



141552361CUSCO10OCT2024

10 Referencias Bibliográficas

- Abadi, M., et al. (2016). TensorFlow: Large scale machine learning on heterogeneous systems. <https://www.tensorflow.org/>
- Adeleke, O., Akinlabi, S., Jen, T., & Dunmade, I. (2021). Evaluación y predicción del contenido energético de los residuos sólidos municipales: Una revisión. IOP Conference Series: Materials Science and Engineering, 1107, 012097. <https://doi.org/10.1088/1757-899X/1107/1/012097>
- Ayeleru, O. O., Fajimi, L. I., Oboirien, B. O., & Olubambi, P. A. (2021). Forecasting municipal solid waste quantity using artificial neural network and supported vector machine techniques: A case study of Johannesburg, South Africa. Journal of Cleaner Production, 289, 125671. <https://doi.org/10.1016/j.jclepro.2020.125671>
- Ayeleru, O. O., et al. (2021). Soft computing applications in municipal solid waste forecast: A short review. In Soft Computing Techniques in Solid Waste and Wastewater Management (pp. 247–256). Elsevier. <https://doi.org/10.1016/B978-0-12-824463-0.00015-X>
- Cerna Cueva, A. F., Rosas Echevarría, C. W., Perales Flores, R. S., & Ataucusi Flores, P. L. (2022). Predicción de la generación de residuos sólidos domiciliarios con machine learning en una zona rural de Puno. TECNIA, 32(1), 44–52. <https://doi.org/10.21754/tecnia.v32i1.1378>
- Cha, G., Choi, S., Hong, W., & Park, C. (2022). Desarrollo de un modelo de machine learning para la predicción de la tasa de generación de residuos de demolición de edificios en áreas de reurbanización. International Journal of Environmental Research and Public Health, 20, 107. <https://doi.org/10.3390/ijerph20010107>
- Coskuner, D., Altuntas, Y., & Cinar, A. (2021). Predictive analysis of municipal solid waste generation using an optimized neural network model. Processes, 9(11), 2045. <https://doi.org/10.3390/pr9112045>
- García, P., López, M., & Ramos, E. (2022). Evaluación de algoritmos de aprendizaje automático en sistemas ambientales. Revista de Ciencia y Tecnología Ambiental, 18(3), 25–39.
- Kavyanifar, B., Tavakoli, B., Torkaman, J., Taheri, A., & Orkomi, A. (2020). Predicción de desechos sólidos costeros mediante la aplicación de enfoques de machine learning: Estudio de caso: Noor, provincia de Mazandarán, Irán. Caspian Journal of Environmental Sciences, 18, 227–236. <https://doi.org/10.22124/CJES.2020.4135>
- Latina TV. (2021). VES: Vecinos denuncian que playa Venecia amanece llena de basura. El Comercio. <https://elcomercio.pe/videos/pais/ves-vecinos-denuncian-que-playa-venecia-amanece-llena-de-basura-nnav-video-latv-noticia/>
- López Morales, J., Quispe, C., & Rodríguez, H. (2021). Aplicaciones de machine learning en la predicción ambiental. Revista Peruana de Ingeniería, 12(2), 41–55.

- Ministerio del Ambiente. (2023). Minam y ciudadanos realizan jornada de limpieza en Playa Venecia de Villa El Salvador. <https://www.gob.pe/i-venecia-de-villa-el-salvador>
- MINSA. (2019). La acumulación de residuos sólidos puede generar afectaciones a la salud de la población. <https://www.gob.pe/institucion/minsa/noticias/49954-la-acumulacion-de-residuos-solidos-puede-generar-afectaciones-a-la-salud-de-lapoblacion>
- Municipalidad Distrital de Villa El Salvador. (2024). Plan de manejo de residuos sólidos. <https://www.munives.gob.pe/WebSite/PLAN%20DE%20MANEJO%20DE%20RESIDUOS%20SOLIDOS.pdf>
- Nguyen, X. C., Nguyen, T. T., La, D. D., Kumar, G., Rene, E. R., Chang, S. W., & Nguyen, V. H. (2021). Development of machine learning-based models for prediction of municipal solid waste generation in residential areas: A case study of Vietnam. *Resources, Conservation and Recycling*, 164, 105381. <https://doi.org/10.1016/j.resconrec.2020.105381>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pillai, K. S., Sneha, M. L., Aiswarya, S., Anand, A. B., & Prasad, G. (2023). Municipal solid waste management: A review of machine learning applications. *E3S Web of Conferences*, 455, 02018. <https://doi.org/10.1051/e3sconf/202345502018>
- Toğaçar, M., Ergen, B., & Cömert, Z. (2023). A systematic literature review on municipal solid waste management using machine learning and deep learning. *Artificial Intelligence Review*. Advance online publication. <https://doi.org/10.1007/s10462-025-11196-9>
- Universidad Nacional Tecnológica de Lima Sur. (2023). Estudiantes de Ingeniería de Sistemas ejecutan la limpieza de la playa Venecia. Facebook UNTELS. <https://www.facebook.com/share/p/1Ehc95pYU3/>
- Universidad Nacional Tecnológica de Lima Sur. (2024). Estudiantes de Ingeniería de Sistemas ejecutan proyecto de gestión ambiental en playa Venecia. UNTELS. <https://www.untels.edu.pe/verNoticias.aspx?publicamos=2346>
- Villalobos, R., Díaz, L., & Vargas, M. (2020). Validación de instrumentos para estudios cuantitativos en ingeniería ambiental. *Revista de Metodología Aplicada*, 5(1), 67–78.



Firma y nombres

Investigador Principal

Nº DNI.41129406