



**FACULTAD DE INGENIERÍA ARQUITECTURA Y  
URBANISMO**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**TRABAJO DE INVESTIGACIÓN**

**Técnicas de Machine Learning para el recuento de  
plaquetas como ayuda al diagnóstico de  
enfermedades: Una revisión Sistemática**

**PARA OPTAR EL GRADO ACADÉMICO DE BACHILLER  
EN INGENIERIA DE SISTEMAS**

**Autor(es)**

**Laynes Castillo Jose Carlos**

ORCID: <https://orcid.org/0000-0003-3110-4729>

**Espinoza Purisaca Josue Enoc**

ORCID: <https://orcid.org/0000-0002-4773-7056>

**Asesor**

**Ing. Tuesta Monteza Victor Alexci**

ORCID: <https://orcid.org/0000-0002-5913-990X>

**Línea de Investigación**

**Ciencias de la información como herramientas multidisciplinares  
y estratégicas en el contexto industrial y de organizaciones**

**Sublínea de Investigación**

**Informática y transformación digital en el contexto industrial y  
organizacional**

**Pimentel – Perú**

**2025**

**DECLARACIÓN JURADA DE ORIGINALIDAD**

Quienes suscribimos la **DECLARACIÓN JURADA**, somos Laynes Castillo Jose Carlos, Espinoza Purisaca Josue Enoc, del programa de estudios de Ingeniería de Sistemas, de la Universidad Señor de Sipán S.A.C, declaramos bajo juramento que somos autores del trabajo titulado:

**Técnicas de Machine Learning para el recuento de plaquetas como ayuda al diagnóstico de enfermedades: Una revisión Sistemática**

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán (CIEI USS) conforme a los principios y lineamientos detallados en dicho documento, con relación a las citas y referencias bibliográficas, respetando al derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y auténtico.

En virtud de lo antes mencionado, firman:

<p><b>Laynes Castillo Jose Carlos</b></p>	<p>DNI: 72956601</p>	
		<p>Firma</p>
<p><b>Espinoza Purisaca Josue Enoc</b></p>	<p>DNI: 74022420</p>	
		<p>Firma</p>

Pimentel, 13 de enero de 2024

## **Dedicatoria**

A nuestras familias, por su amor incondicional y su apoyo constante en cada etapa de nuestras vidas.

Y a todas aquellas personas que creyeron en nosotros y nos inspiraron a alcanzar nuestras metas.

## **Agradecimientos**

A nuestras familias, por ser el pilar fundamental de nuestro crecimiento y por brindarnos su amor, apoyo incondicional y confianza en cada paso de este camino académico.

A nuestros profesores, por transmitirnos sus conocimientos, inspirarnos con su pasión por la enseñanza y guiarnos en el desarrollo de nuestras habilidades técnicas y personales.

A nuestros compañeros y amigos, por su camaradería, palabras de aliento y momentos compartidos que hicieron este viaje más enriquecedor y memorable.

Finalmente, a nosotros mismos, Jose Carlos Laynes Castillo y Josue Enoc Espinoza Purisaca, por el esfuerzo, la dedicación y la perseverancia que nos han permitido alcanzar esta meta y aportar un granito de arena al progreso del conocimiento en el ámbito de la ingeniería

## Índice

Dedicatoria.....	3
Agradecimientos .....	4
Índice .....	5
Índice de tablas y figuras .....	6
Resumen .....	7
Abstract.....	8
I. INTRODUCCIÓN.....	9
1.1. Realidad problemática.....	9
1.2. Formulación del problema .....	11
1.3. Hipótesis.....	12
1.4. Objetivos .....	12
1.5. Teorías relacionadas al tema.....	13
II. MÉTODO DE INVESTIGACIÓN .....	16
III. RESULTADOS .....	25
IV. DISCUSIÓN .....	43
V. CONCLUSIONES.....	49
VI. REFERENCIAS.....	50
ANEXOS.....	59

## Índice de tablas y figuras

Figura 1 Esquemas de red neuronal .....	15
Figura 2 Diferencias de Machine y Deep Learning .....	15
Figura 3 Proceso de mapeo sistemático .....	17
Figura 4 Diagrama de flujo PRISMA. Elaboración propia .....	21
Figura 5 Proceso de clasificación de los artículos .....	23
Figura 6 Resultados de la revisión sistemática.....	25
Figura 7 Publicación de Articulos por años .....	31
Figura 8 Origen de los artículos .....	32
Figura 9 Países de Articulos Seleccionados.....	33
Figura 10 Tipo de Articulos Seleccionados .....	34
Figura 11. Tipos de Articulos seleccionados .....	35
Figura 12 Año de publicación por tipo de artículo por año.....	36
Figura 13 Distribución porcentual de los trabajos seleccionados. ....	37
Tabla 1 Beneficios del Machine Learning para el recuento de plaquetas .....	15
Tabla 2 Extracción de datos de los artículos seleccionados .....	24
Tabla 3 Lista de artículos seleccionados.....	26
Tabla 4 Casos de uso – Articulos relacionados.....	36

## **Resumen**

Desde la introducción de las técnicas de Machine Learning (ML) en la medicina, se ha investigado su aplicación en diversos campos, incluyendo el recuento de plaquetas, una métrica clave en el diagnóstico de enfermedades hematológicas e infecciosas. Esta revisión sistemática examina el uso de ML para el recuento de plaquetas, siguiendo las directrices de PRISMA. Se realizó un mapeo sistemático utilizando un protocolo de búsqueda en cuatro bases de datos científicas, lo que permitió identificar y evaluar estudios sobre algoritmos de clasificación, redes neuronales y aprendizaje profundo aplicados al análisis de imágenes de microscopía y datos clínicos. Aunque los resultados preliminares son prometedores, indicando que ML podría mejorar la precisión y velocidad del recuento de plaquetas, la mayoría de los estudios aún están en etapas iniciales y carecen de evaluaciones clínicas completas. Es necesario continuar investigando para desarrollar prototipos funcionales y validar clínicamente estos enfoques.

**Palabras Clave:** Aprendizaje automático; Plaquetas; Dengue; Cuidado de la salud; Asistencia sanitaria; Atención médica

## **Abstract**

Since the introduction of Machine Learning (ML) techniques in medicine, research has been conducted to explore their applications in various fields, including platelet counting—a key metric in diagnosing hematological and infectious diseases. This systematic review examines the use of ML for platelet counting, following PRISMA guidelines. A systematic mapping process was conducted using a search protocol across four scientific databases, allowing for the identification and evaluation of studies on classification algorithms, neural networks, and deep learning applied to the analysis of microscopy images and clinical data. Although preliminary results are promising, suggesting that ML could improve the accuracy and speed of platelet counting, most studies are still in their early stages and lack comprehensive clinical evaluations. Further research is needed to develop functional prototypes and clinically validate these approaches.

**Keywords:** Machine learning; Platelets; Dengue; Health care; Healthcare; Medical care

## I. INTRODUCCIÓN

### 1.1. Realidad problemática.

En los últimos años, las técnicas de Machine Learning (ML) han revolucionado diversos campos de la medicina, proporcionando herramientas avanzadas para el diagnóstico de enfermedades. El recuento de plaquetas, un indicador crucial en el diagnóstico de diversas patologías, se ha beneficiado significativamente de estas innovaciones tecnológicas. Las técnicas de ML permiten un análisis más preciso y rápido de las imágenes de sangre, superando las restricciones de los enfoques tradicionales manuales y automáticos.

Uno de los enfoques más destacados es el uso de redes neuronales convolucionales (CNN) para la detección y clasificación de plaquetas en imágenes de frotis de sangre. Este método ha demostrado una alta precisión y eficiencia en la identificación de plaquetas y otras células sanguíneas, mejorando significativamente la capacidad diagnóstica en condiciones como la trombocitopenia y otras enfermedades hematológicas [1]. Además, la implementación de modelos de aprendizaje profundo, como YOLO\_v3 mejorado, ha permitido la detección en tiempo real de plaquetas con una precisión notable [2].

La integración con los sistemas existentes de gestión de información de ML en el análisis de plaquetas no solo optimiza la exactitud del diagnóstico, sino que además agiliza la detección de patrones complejos en los datos que pueden ser indicativos de condiciones subyacentes. Por ejemplo, estudios recientes han utilizado técnicas de ML para analizar los perfiles de lípidos y acilcarnitinas en plaquetas, proporcionando nuevos insights sobre los riesgos trombo-isquémicos y eventos hemorrágicos [3].

La revisión sistemática de modelos de clasificación del dengue basados en ML destaca la capacidad de estos modelos para diferenciar entre diversas etapas de la enfermedad, lo cual es crucial para el manejo y tratamiento adecuado de los pacientes [4]. Además, la detección de dengue a partir de imágenes de frotis de sangre utilizando características de plaquetas y linfocitos ha demostrado ser una herramienta poderosa en áreas endémicas, mejorando la rapidez y precisión del diagnóstico [5].

Aunque las técnicas de ML para el recuento de plaquetas han mostrado un gran potencial en estudios preliminares, el campo todavía enfrenta numerosos desafíos y preguntas sin resolver. La implementación de estas técnicas en entornos clínicos presenta problemas relacionados con la precisión, la generalización de los modelos y su incorporación a los sistemas actuales de administración de información de salud. Además, la variabilidad en los datos de las muestras de sangre y las diferencias en las tecnologías de imagen pueden afectar la consistencia y la fiabilidad de los resultados.

Aunque existen algunas revisiones interesantes relacionadas con este tema [6-10], la revisión de Shaikh et al. [6], se desarrolló un clasificador optimizado basado en un conjunto de técnicas de ML para anticipar casos de fiebre del dengue. Este trabajo se centró principalmente en la mejora de la exactitud de los modelos de ML y la valoración de su desempeño en diferentes conjuntos de datos clínicos, demostrando que los modelos de ML pueden ser efectivos en la predicción de enfermedades infecciosas. Por otro lado, Silitonga et al. [7] evaluó el desempeño de los modelos para el dengue desarrollados utilizando redes neuronales artificiales (ANN) y clasificadores de bosques aleatorios (RF). La revisión destacó la importancia de la selección de características y la preparación de los datos para optimizar la precisión de los modelos, y enfocándose en la comparación de diferentes algoritmos de ML para encontrar los más efectivos. En la revisión realizada por Majeed et al. [8], se utilizó un enfoque de aprendizaje profundo con atención espacial para la predicción de fiebre del dengue en Malasia. Este estudio se diferencia por su uso de modelos de redes neuronales de largo corto plazo (LSTM) y su habilidad para gestionar datos secuenciales, lo que es crucial para capturar patrones temporales en la incidencia de la enfermedad. Por su parte Chaw et al. [9] desarrollaron un modelo predictivo de análisis empleando algoritmos de ML para estimar el riesgo de desarrollo de shock en pacientes con dengue. Este trabajo se enfocó en la incorporación de diversas fuentes de datos clínicos para mejorar la capacidad predictiva de los modelos, destacando la importancia de un enfoque multidimensional en el diagnóstico de enfermedades. Finalmente, en el estudio de Jahandideh et al. [10], se llevó a cabo una revisión sistemática de modelos basados en ML para la predicción del deterioro clínico. Este

trabajo se centró en valorar la eficacia de distintas técnicas de ML en la anticipación de eventos adversos en pacientes, ofreciendo una perspectiva completa de los retos y oportunidades en la aplicación de ML en entornos clínicos.

Nuestro enfoque se aparta de los adoptados en las revisiones antes mencionadas en el sentido de que seguimos las directrices para la revisión estructurada de la literatura y el desarrollo del proceso de mapeo sistemático, junto con la declaración de los Elementos de Información Preferidos para Revisiones Sistemáticas y Meta-Análisis (PRISMA) al realizar e informar sobre nuestra revisión. Nuestra revisión sistemática se basa en un protocolo de investigación bien diseñado que garantiza un muestreo holístico e imparcial de todos los artículos publicados revisados por pares que son relevantes para el tema. Sobre la base de este protocolo, recuperamos de bases de datos científicas de renombre los artículos pertinentes que clasificamos y mapeamos en diferentes categorías para desentrañar el verdadero estado de la investigación en curso en el uso de enfoques de ML para el recuento de plaquetas como ayuda al diagnóstico de enfermedades.

El mapa resultante de nuestra investigación será muy valioso para los profesionales e investigadores a la hora de comprender el estado del arte del dominio y las áreas de investigación futura. Hasta donde sabemos, ésta es la primera revisión bibliográfica sobre la aplicación de técnicas de ML en el recuento de plaquetas que sigue el proceso de estudio de mapeo sistemático. La revisión sistemática realizada por Majeed et al. [8] inspiró nuestra elección de esta metodología; sin embargo, su revisión sobre la predicción de fiebre del dengue utilizando LSTM y atención espacial no está relacionada directamente con nuestro tema, que es el recuento de plaquetas para el diagnóstico de enfermedades. La revisión sistemática de Jahandideh et al. [10] también es similar, pero la nuestra difiere notablemente en cuanto al alcance; por ejemplo, la revisión abarca 52 artículos, mientras que la nuestra cubre 70 artículos.

## **1.2. Formulación del problema**

¿Cuáles son los modelos de Machine Learning más eficaces para el recuento de plaquetas y cómo pueden mejorarse en términos de precisión, robustez y aplicabilidad en diferentes

condiciones clínicas, superando los desafíos técnicos y logísticos de su implementación en hospitales y laboratorios clínicos?

### **1.3. Hipótesis**

La aplicación de técnicas avanzadas de Machine Learning, como las redes neuronales convolucionales y modelos de aprendizaje profundo, permite mejorar significativamente la precisión, eficiencia y capacidad diagnóstica en el recuento de plaquetas. Además, una revisión sistemática y un mapeo exhaustivo de la literatura proporcionarán insights valiosos sobre los desafíos técnicos, las tendencias actuales y las áreas de investigación futura para integrar estas tecnologías en entornos clínicos.

### **1.4. Objetivos**

#### **Objetivo general**

Realizar una revisión sistemática y un mapeo exhaustivo del estado del arte sobre el uso de técnicas de Machine Learning en el recuento de plaquetas, identificando modelos eficaces, tendencias, desafíos y oportunidades para mejorar la precisión y aplicabilidad clínica en el diagnóstico de enfermedades.

#### **Objetivos específicos**

- Identificar los modelos de Machine Learning más destacados aplicados al recuento de plaquetas en imágenes de frotis sanguíneos.
- Analizar los enfoques técnicos y metodológicos utilizados en la literatura reciente.
- Comparar estudios previos sobre la aplicación de técnicas de ML en hematología y diagnóstico de enfermedades infecciosas como el dengue.
- Clasificar y mapear los estudios revisados en diferentes categorías para identificar el estado actual de la investigación.
- Evaluar la viabilidad de implementar modelos avanzados de ML como CNN y YOLO mejorado para la detección en tiempo real de plaquetas.

## **1.5. Teorías relacionadas al tema**

### **1.5.1 Visión general de Machine Learning para el recuento de plaquetas como ayuda al diagnóstico de enfermedades**

Contexto y Definición:

Las técnicas de Machine Learning (ML) han ganado una gran relevancia en el campo de la medicina, particularmente en el análisis y diagnóstico de enfermedades mediante el recuento de plaquetas. El recuento de plaquetas es un procedimiento crítico en el diagnóstico de diversas condiciones hematológicas y otras enfermedades. Las técnicas de ML permiten automatizar y mejorar la precisión de este proceso, proporcionando resultados rápidos y fiables que son esenciales para el diagnóstico y tratamiento oportuno.

Las técnicas de ML aplicadas al recuento de plaquetas incluyen una variedad de enfoques como redes neuronales convolucionales (CNN), redes neuronales de largo corto plazo (LSTM), y bosques aleatorios (RF), entre otros. Estas técnicas utilizan algoritmos avanzados para analizar grandes volúmenes de datos de imágenes y datos clínicos, identificando patrones que pueden ser indicativos de condiciones de salud específicas [1][2][3].

Importancia del Tema:

El recuento de plaquetas es fundamental en la evaluación de enfermedades como la trombocitopenia, leucemia, y diversas infecciones virales como el dengue. La implementación de técnicas de ML en este ámbito ofrece múltiples ventajas, incluyendo la mejora en la precisión diagnóstica, la reducción del tiempo de análisis y la capacidad de manejar grandes volúmenes de datos de manera eficiente [4][5]. Según estudios recientes, la automatización del recuento de plaquetas mediante ML puede reducir los errores humanos y proporcionar diagnósticos más rápidos, lo que es crucial para el tratamiento oportuno de pacientes [6].

Historia y Evolución:

El uso de técnicas de ML en el recuento de plaquetas ha evolucionado significativamente desde sus primeros días. Inicialmente, los métodos manuales y semiautomáticos dominaban el campo, con técnicas que dependían en gran medida de la habilidad del técnico para interpretar imágenes microscópicas [7]. Con el avance de la informática y el desarrollo de

algoritmos de ML más sofisticados, la automatización del recuento de plaquetas ha mejorado drásticamente.

Los primeros modelos de ML aplicados a la hematología se centraban en el análisis de imágenes básicas y en la clasificación binaria. Con el tiempo, se introdujeron modelos más complejos como las redes neuronales convolucionales (CNN), que permitieron una mayor precisión en la detección y clasificación de plaquetas [8]. Más recientemente, el uso de modelos de aprendizaje profundo (Deep Learning) ha permitido no solo la detección de plaquetas sino también la predicción de condiciones de salud basadas en el análisis de datos de múltiples fuentes [9].

Desafíos y Problemas Actuales:

Desafíos Persistentes A pesar de los avances significativos, el campo enfrenta varios desafíos persistentes. Uno de los principales problemas es la variabilidad en la calidad y consistencia de los datos de entrada, ya que las muestras de sangre pueden variar ampliamente entre diferentes pacientes y condiciones clínicas. Esta variabilidad puede afectar la precisión de los modelos de ML y dificultar la generalización de los resultados [10]. Otro desafío importante es la integración de estas tecnologías en los entornos clínicos. Los sistemas de ML deben ser compatibles con las infraestructuras existentes y cumplir con estrictos estándares regulatorios y de privacidad de datos. Además, la aceptación y el uso de estas tecnologías por parte del personal médico requieren una capacitación adecuada y una confianza en los sistemas automatizados [11].

Finalmente, existe una necesidad continua de mejorar los algoritmos y modelos para manejar mejor los datos complejos y no estructurados, como los datos de imágenes y los registros médicos electrónicos. La capacidad de estos modelos para aprender y adaptarse continuamente a nuevos datos es crucial para mantener su relevancia y efectividad en el diagnóstico clínico [12].

En la figura 1 se muestra el esquema del proceso de entrenamiento y predicción de redes neuronales para el recuento de plaquetas. En el entrenamiento, las imágenes de plaquetas pasan por un tratamiento y se utilizan para entrenar un modelo clasificador. En la predicción,

nuevas imágenes tratadas se clasifican usando el modelo entrenado, asistiendo en el diagnóstico de enfermedades mediante técnicas de ML.

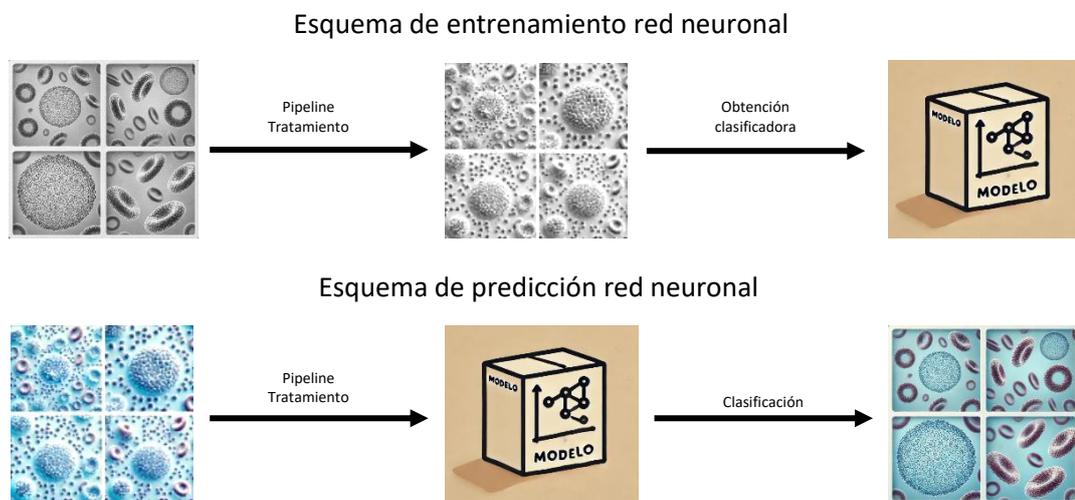


Figura 1. Esquemas de red neuronal

La figura 1, muestra un ejemplo entre Machine Learning y Deep Learning en el análisis de datos clínicos. En Machine Learning, los datos de entrada pasan por un proceso de extracción de datos y luego son clasificados, resultando en una salida. En Deep Learning, los datos de entrada se someten a una extracción de características y clasificación simultánea, generando una salida más directa y automatizada.

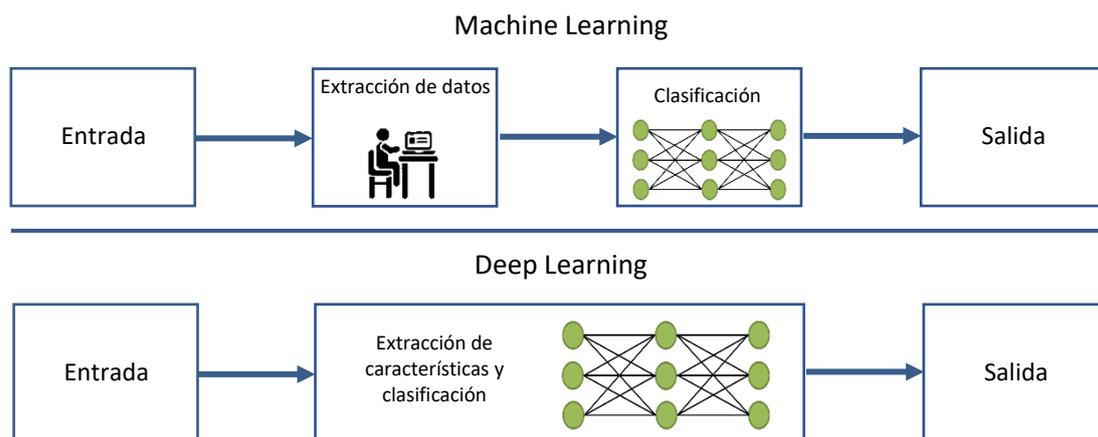


Figura 2. Diferencias de Machine y Deep Learning

Tabla 1. Beneficios del Machine Learning para el recuento de plaquetas

Beneficio	Descripción
<b>Automatización</b>	Los modelos de ML automatizan el proceso de recuento de plaquetas, reduciendo la dependencia de técnicas manuales.

<b>Precisión</b>	Mejora en la precisión del recuento de plaquetas mediante algoritmos avanzados.
<b>Eficiencia</b>	Reducción del tiempo de análisis, permitiendo diagnósticos más rápidos.
<b>Manejo de Grandes Datos</b>	Capacidad de analizar y procesar grandes volúmenes de datos clínicos e imágenes.
<b>Reducción de Errores</b>	Minimiza los errores humanos, proporcionando resultados más consistentes y fiables.
<b>Adaptabilidad</b>	Los modelos de ML pueden adaptarse y aprender de nuevos datos, mejorando continuamente su precisión y eficacia.

## II. MÉTODO DE INVESTIGACIÓN

Al realizar y comunicar esta revisión, se adoptaron las directrices para la revisión sistemática de la literatura [7] y el proceso para el estudio de mapeo sistemático [8], así como las directrices descritas en la declaración PRISMA [2]. Según se explica en [1], El objetivo de un estudio de mapeo sistemático es brindar una perspectiva general sobre el campo de investigación y complementar el análisis al examinar el estado de la evidencia en áreas específicas. Dentro de este contexto, los hallazgos del estudio de mapeo permitirán identificar y trazar un panorama detallado los casos de uso de las técnicas de Machine Learning (ML) en el recuento de plaquetas para el diagnóstico de enfermedades, Y comprender hasta qué nivel se han desarrollado estas aplicaciones en relación con los casos de uso identificados. También nos ayudarán a identificar áreas con posibles lagunas de investigación. La revisión sistemática también nos permitirá analizar las tendencias actuales respecto a los enfoques técnicos, las metodologías y los conceptos utilizados en el desarrollo de aplicaciones de ML para el recuento de plaquetas. Seguidamente, se lleva a cabo el proceso de mapeo sistemático, el cual se presenta en la Figura 3.

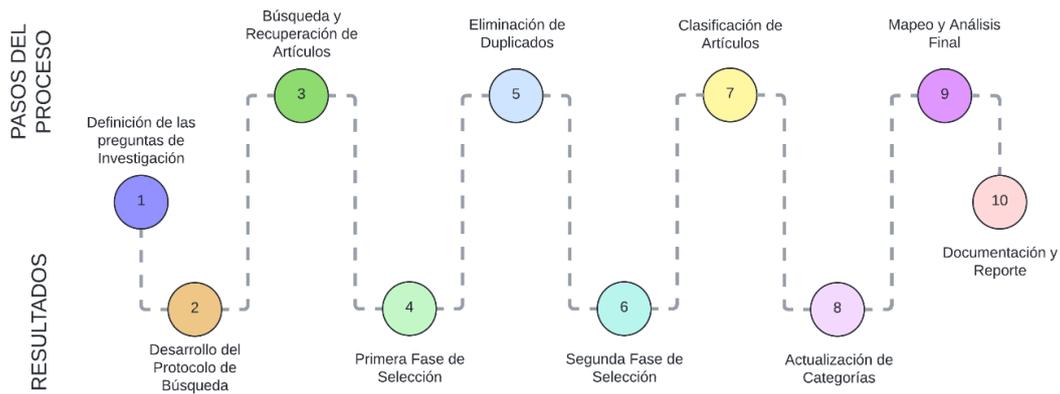


Figura 3. Proceso de mapeo sistemático

## 2.1 Definición de las preguntas de investigación

El primer paso del estudio de mapeo sistemático consiste en definir las siguientes cuatro preguntas de investigación de acuerdo con nuestro objetivo, que es desentrañar el estado del arte de la investigación sobre las Técnicas de Machine Learning para el recuento de plaquetas como ayuda al diagnóstico de enfermedades

### 2.1.1 ¿Cuáles son las técnicas de Machine Learning que tienen el mejor desempeño en el recuento de plaquetas haciendo uso de imágenes?

La cuestión principal de esta investigación es identificar las técnicas de Machine Learning más efectivas para el recuento de plaquetas utilizando imágenes. Revisando los artículos pertinentes de las bases de datos científicas, podremos determinar cuáles de estas técnicas ofrecen un mejor desempeño en términos de precisión, velocidad y robustez. Este conocimiento permitirá a los investigadores y profesionales del sector sanitario enfocar sus esfuerzos en las técnicas más prometedoras para el diagnóstico médico.

### 2.1.2 ¿Cuáles son los datasets que se han utilizado para el entrenamiento, pruebas y validación de técnicas de Machine Learning para el recuento de plaquetas?

En la literatura científica se han utilizado diversos conjuntos de datos para entrenar, probar y validar modelos de Machine Learning en el contexto del recuento de plaquetas. Es crucial comprender qué datasets son más frecuentemente utilizados y cuáles son sus características

(tamaño, diversidad, fuente de imágenes, etc.). Esta pregunta ayudará a identificar posibles limitaciones en los datos disponibles y a destacar la necesidad de datasets más robustos y representativos.

### **2.1.3 ¿Qué metodologías de validación y evaluación se utilizan para asegurar la efectividad de los modelos de Machine Learning en la medicina?**

Esta pregunta busca conocer las diversas metodologías de validación y evaluación aplicadas para asegurar la efectividad de los modelos de Machine Learning en el recuento de plaquetas. Es fundamental entender qué enfoques se utilizan para evaluar la precisión, sensibilidad, especificidad y otros aspectos críticos de los modelos. Esto proporcionará una visión de las mejores prácticas y ayudará a estandarizar los procesos de validación y evaluación en futuros estudios.

### **2.1.4 ¿Cuál es el impacto de la aplicación de técnicas de Machine Learning en el recuento de plaquetas?**

Esta pregunta investiga la influencia de la aplicación de técnicas de Machine Learning en el recuento de plaquetas en términos de mejoras en la precisión del diagnóstico, reducción del tiempo de análisis, y otros beneficios clínicos. Es importante entender cómo estas técnicas están mejorando los procedimientos médicos y qué impacto tienen en la práctica clínica diaria. Este análisis ayudará a justificar la adopción de Machine Learning en la medicina y a identificar áreas de mejora.

### **2.1.5 ¿Cuáles son las métricas con las que se mide el rendimiento de las técnicas de Machine Learning en el recuento de plaquetas?**

Para evaluar el rendimiento de las técnicas de Machine Learning en el recuento de plaquetas, se utilizan diversas métricas. Esta pregunta busca identificar cuáles son estas métricas (por ejemplo, precisión, recall, F1-score, AUC-ROC) y cómo se aplican en los estudios revisados. Entender estas métricas es esencial para comparar de manera efectiva los resultados de

diferentes investigaciones y para establecer estándares de rendimiento.

### **2.1.6 ¿Cuáles son las principales aplicaciones de software que se han desarrollado haciendo uso de Machine Learning para el recuento de plaquetas?**

La última pregunta se refiere a las aplicaciones de software que han sido desarrolladas utilizando técnicas de Machine Learning para el recuento de plaquetas. Es crucial conocer qué software existe, sus funcionalidades, su uso en entornos clínicos reales y sus limitaciones. Esta información permitirá evaluar el estado actual de las aplicaciones disponibles y orientar el desarrollo futuro de soluciones más avanzadas y efectivas.

## **2.2 Realización de la investigación**

En la segunda etapa del proceso para el estudio de mapeo sistemático, se seleccionan los artículos primarios para el estudio mediante la búsqueda en las bases de datos científicas utilizando una cadena de búsqueda o palabras clave. En nuestra búsqueda se consideraron cuatro bases de datos científicas: ACM Digital Library, IEEE Xplore, Science Direct y Scopus. Al seleccionar estas bases de datos, nuestro objetivo era enfocarnos exclusivamente en los artículos evaluados por pares que hubieran sido publicados en revistas, conferencias, talleres, libros o simposios de renombre.

Para buscar en las bases de datos, utilizamos la siguiente cadena de búsqueda: ("dengue" OR "healthcare" OR "health assistance" OR "medical care" OR "platelets") AND ("machine learning") OR ("platelet count in machine learning" OR "platelet count to detect dengue"). Esta selección de la cadena de búsqueda se basa en búsquedas piloto en las que probamos varios términos y combinaciones relacionadas con la asistencia sanitaria y el recuento de plaquetas en el contexto de Machine Learning. Identificamos que los términos seleccionados abarcaban un espectro amplio y relevante de la literatura existente, permitiéndonos capturar artículos clave que abordan el uso de técnicas de ML en el recuento de plaquetas, tanto en aplicaciones generales de salud como específicamente en la detección de enfermedades

como el dengue.

Algunas observaciones importantes de nuestras búsquedas piloto incluyen:

- **Términos de Salud:** La inclusión de términos como "healthcare", "health assistance", "medical care" y "platelets" garantiza que abarcamos una gama amplia de artículos relacionados con la aplicación de ML en diferentes contextos de salud y medicina.
- **Contexto Específico:** La adición de "dengue" y frases específicas como "platelet count in machine learning" y "platelet count to detect dengue" nos permitió focalizar nuestra búsqueda en estudios específicos sobre el recuento de plaquetas para la detección de dengue, una aplicación relevante y específica del tema de estudio.
- **Sin Restricciones Temporales:** La búsqueda se realizó sin restricciones de tiempo, dado que el campo de estudio es relativamente nuevo. Consideramos que toda la literatura existente es relevante para nuestro estudio, permitiéndonos capturar el desarrollo histórico y reciente de las técnicas de Machine Learning aplicadas al recuento de plaquetas.
- **Revisión de Palabras Clave Alternativas:** Revisamos alternativas a nuestras palabras clave, como otros términos y acrónimos comunes en el ámbito sanitario y de Machine Learning. Sin embargo, observamos que nuestra cadena de búsqueda seleccionada abarcaba adecuadamente los términos y acrónimos relevantes, sin la necesidad de extender la búsqueda con términos adicionales.
- **La implementación de esta estrategia de búsqueda nos permitió compilar una base de datos amplia y representativa de la literatura científica relevante, asegurando que nuestro estudio de mapeo sistemático cubra de manera exhaustiva las aplicaciones de Machine Learning para el recuento de plaquetas y su impacto en el diagnóstico de enfermedades, particularmente en el contexto del dengue y otras aplicaciones médicas. Es relevante destacar que la búsqueda bibliográfica se llevó a cabo sin restricciones de tiempo, teniendo en cuenta que el tema es relativamente nuevo y, por lo tanto, toda la literatura existente en esta área se consideró relevante para nuestro estudio**

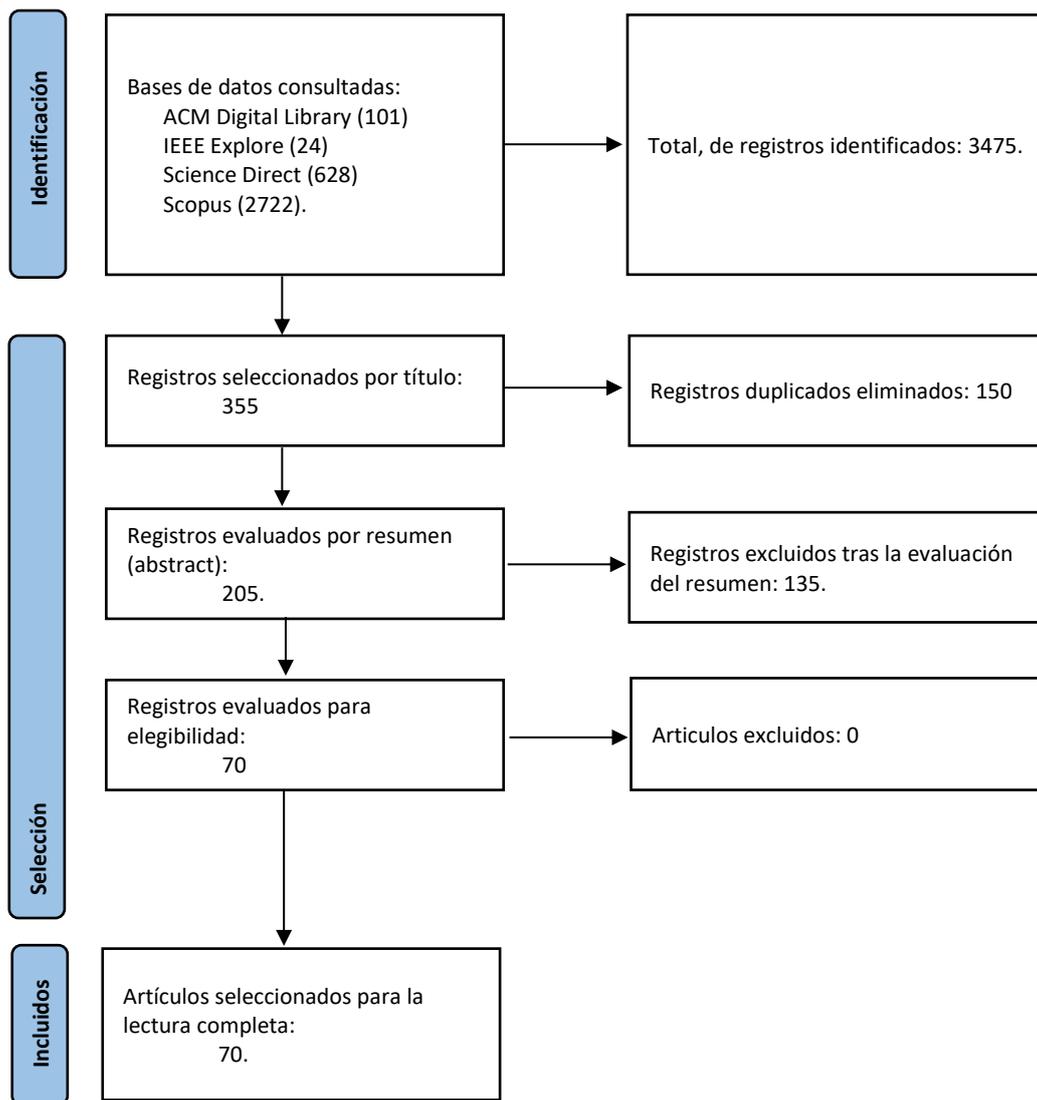


Figura 4. Diagrama de flujo PRISMA.

El diagrama presentado sigue la metodología PRISMA (Figura 4) y detalla las etapas del proceso de selección de artículos en una revisión sistemática. En la etapa de Identificación, se llevaron a cabo búsquedas en cuatro bases de datos: ACM Digital Library, donde se identificaron 101 registros; IEEE Explore, con 24 registros; ScienceDirect, con 628 registros, y Scopus, que aportó 2722 registros. En total, se recopiló 3475 registros, esta primera búsqueda se realizó utilizando las palabras claves.

En la fase de Selección se realizó una elección por título de artículo, donde se obtuvieron 355 registros, después de eliminar 150 registros duplicados, se seleccionaron 205 registros para la evaluación por resumen o abstract. Tras este análisis, 135 registros fueron excluidos, dejando un total de 70 registros para pasar a la etapa de evaluación de elegibilidad. En esta

etapa, no se excluyó ningún artículo, lo que permitió que los 70 registros restantes fueran considerados aptos para continuar con la revisión.

Finalmente, en la fase de Inclusión, se concluyó que los 70 artículos seleccionados pasaran a la etapa de lectura completa. Este proceso garantiza una selección cuidadosa y precisa de la literatura relevante para el tema de estudio, asegurando que los registros evaluados cumplan con los criterios establecidos para el análisis.

### **2.3 Selección de documentos relevantes**

Una vez obtenidos los artículos de las bases de datos siguiendo el protocolo de búsqueda establecido, el paso siguiente fue evaluar su relevancia. En la primera etapa, se analizó la pertinencia de los artículos basándose en sus títulos. Aquellos cuyos títulos evidenciaban claramente que no eran relevantes para el estudio fueron descartados. Algunos artículos recuperados no guardaban relación con la aplicación de Machine Learning en el recuento de plaquetas, por lo que fueron eliminados. Si la relevancia del artículo no podía determinarse únicamente a partir del título, este avanzaba a la siguiente etapa para un análisis más detallado.

La segunda fase de la selección consistió en la lectura de los resúmenes de los artículos que habían pasado por la primera etapa. En algunos casos, fue necesario leer también la introducción y la conclusión de un artículo para determinar si el artículo superaba nuestros criterios de exclusión. Nuestros criterios de exclusión nos obligaron a descartar lo siguiente:

- Trabajos no revisados por pares, como entrevistas y anuncios de prensa.
- Trabajos sin disponibilidad de texto completo.
- Trabajos cuyo enfoque principal no está relacionado con la aplicación de técnicas de Machine Learning en el recuento de plaquetas.
- Trabajos duplicados.
- Trabajos retractados.

Los artículos que superaron estos criterios de exclusión y que se consideraron centrados en la aplicación de técnicas de Machine Learning para el recuento de plaquetas se incluyeron

en el siguiente paso del proceso de estudio de mapeo.

## 2.4 Clasificación en función del resumen

El objetivo de esta etapa era organizar los trabajos de investigación significativos de la bibliografía en distintas categorías. Para ello, adoptamos el proceso indicado, como se muestra en la figura 5. El procedimiento consiste en identificar palabras clave y conceptos importantes a partir de los resúmenes de los trabajos de investigación, los cuales reflejan las contribuciones principales de cada trabajo. A partir de estas palabras clave, los artículos se organizaron en diversas categorías. Luego de agrupar estos trabajos en estas categorías, se revisó cada uno en detalle y, si el contenido indicaba que debía pertenecer a una categoría distinta, se actualizaron las agrupaciones. En ocasiones, se creaba una nueva categoría si se observaba que el artículo no encajaba en ninguna de las categorías existentes. El resultado final de este procedimiento es un mapeo de todos los documentos relevantes en una serie de categorías distintas.

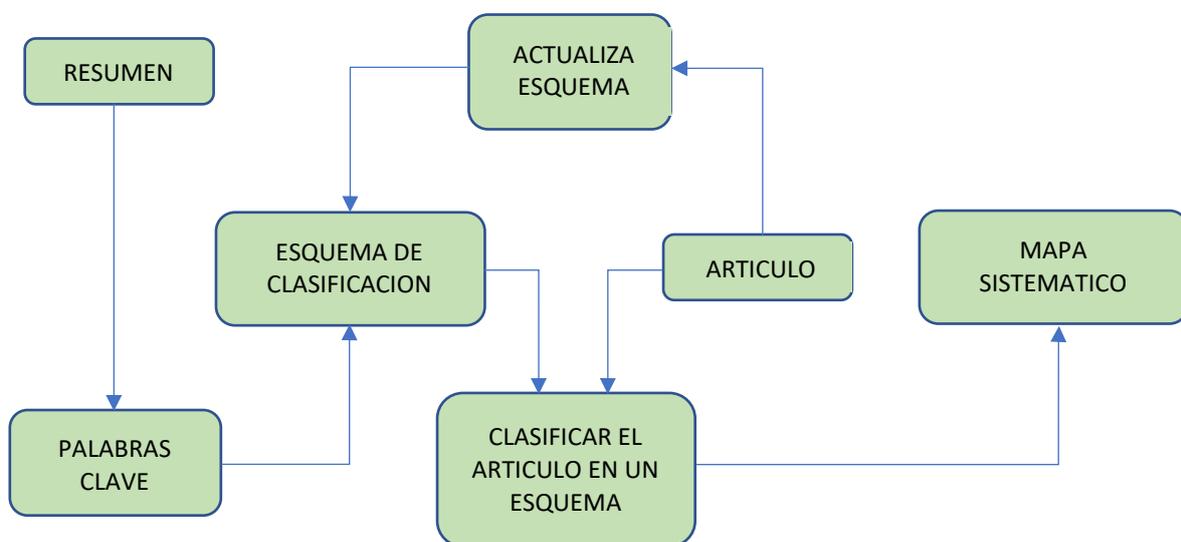


Figura 5. Proceso de clasificación de los artículos

## 2.5 Proceso de extracción de datos y mapeo

En esta etapa final del proceso de mapeo sistemático, se obtuvo la información Para el meta-análisis y para abordar las preguntas de la investigación, se extrajeron un total de 10 datos de cada artículo, tal como se presenta en la Tabla 2. En los primeros siete ítems se extrajo la información básica sobre el artículo, que incluye el año de publicación del artículo, el título, el

autor o los autores, el país desde el que se escribió el artículo (para los autores de varios países, se utiliza el país del autor correspondiente o del primer autor), el lugar de publicación, etc. Los demás datos, del 8 al 10, se extrajeron tras la lectura detallada de los artículos. Además de estos datos, a cada uno de los artículos seleccionados se le asignó un número del 1 al 70, que sirvió como identificador del artículo. Los datos extraídos se recogieron en una hoja de Excel para facilitar su organización y análisis.

Tabla 2. Extracción de datos de los artículos seleccionados

#	Data Item	Descripción
1	Año	Año de publicación del artículo
2	Título	Título del artículo
3	Autores	Los autores del artículo
4	País	País de afiliación de los autores [Se considera el país del primer autor]
5	Canal de publicación	El canal a través del cual se publica el artículo [Considerar los nombres de los journal, simposios, conferencias, etc]
6	Tipo de Publicación	Journal/Conference/Workshop/etc.
7	Origen de la publicación	Academia/Industria/Ambos [bot]
8	Tipo de artículo	Tipo basado en el esquema de clasificación [ <b>Tecnico, Reporte, Técnico+aplicación</b> ]
9	Contribución del artículo	Principales aportaciones del documento
10	Resumen	Nuestro propio resumen o abstract del artículo (60 palabras)

### III. RESULTADOS

En esta sección se muestra los resultados de la revisión sistemática. Usando nuestro protocolo de búsqueda, conseguimos recuperar un total de 3475 artículos de las bases de datos académicas, como se muestra en la Figura 6. Después de la primera selección, realizada a partir de los títulos de los artículos, se descartaron 3120 artículos, lo que dejó 355 artículos para una nueva selección. Los artículos excluidos eran los que no estaban relacionados con la sanidad; sin embargo, la sanidad podía mencionarse en sus resúmenes como uno de los casos de uso no financiero de la cadena de bloques, por lo que nuestro protocolo de búsqueda los recuperó. Además, buscamos artículos duplicados fusionando los 355 artículos en Mendeley, para realizar la eliminación de ellos, sin embargo, no encontramos artículos duplicados. En el siguiente paso de la selección, revisamos los resúmenes de los artículos elegidos y, en algunos casos, también la introducción y la conclusión, para continuar con la selección según los criterios establecidos en la sección 3.3. Como resultado, se eligieron 205 artículos. Al finalizar el proceso de selección, se incluyeron 70 artículos en el estudio. A continuación, se presenta la lista completa de los artículos seleccionados y algunos de los datos extraídos

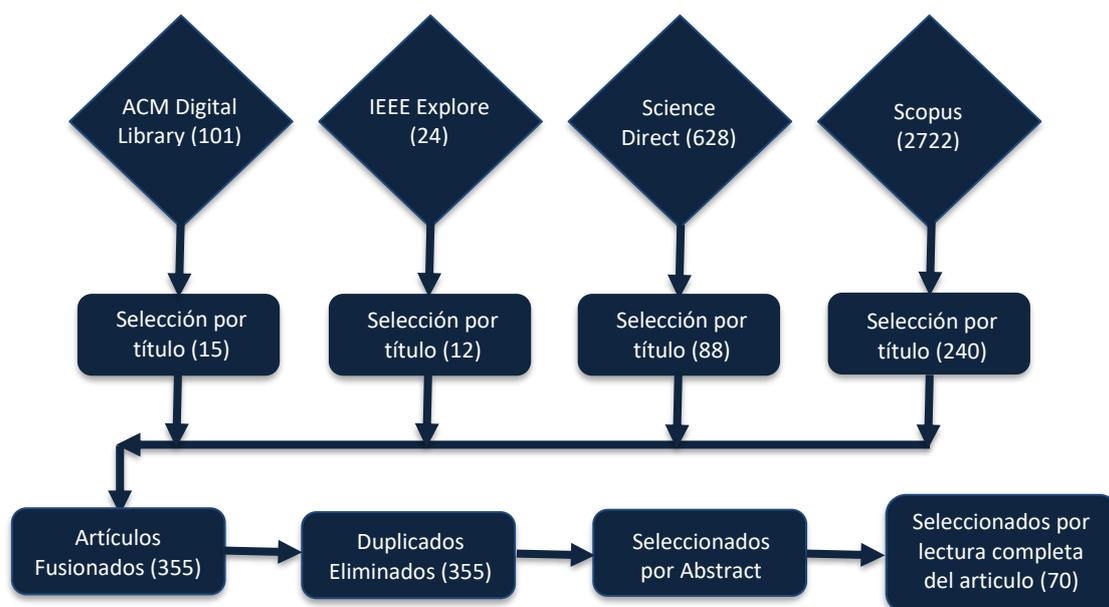


Figura 6. Resultados de la revisión sistemática

La Tabla 3 presenta una visión integral y organizada de las publicaciones analizadas, facilitando su comparación y análisis. Esta tabla contiene datos clave de los artículos seleccionados, organizados en columnas que incluyen el nombre de los autores, la revista en la que fue publicado, el año de publicación, el país de origen y la fuente de búsqueda. Esta estructura permite clasificar y entender rápidamente las características principales de cada publicación, ayudando a identificar patrones y diferencias entre ellas.

Tabla 3. Lista de artículos seleccionados

#	Autores	Revista	Año	País	Origen de Búsqueda
1	Liu, Renting et al. [1]	Cyborg and Bionic Systems	2022	China	Scopus
2	Harm, Tobias et al. [2]	Scientific Reports	2024	Alemania	Scopus
3	Maquen-Niño, Gisella et al. [3]	RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao	2023	Perú	Scopus
4	Mayrose, Hilda et al. [4]	Diagnostics	2023	India	Scopus
5	Sheneman, Luke et al. [5]	PLoS ONE	2021	Estados Unidos	Scopus
6	Shaikh, Salim et al. [6]	Journal of Clinical Medicine	2023	Alemania	Scopus
7	Silitonga, Permatasari et al. [7]	Computational Intelligence and Neuroscience	2021	India	Scopus
8	Majeed, Mokhalad et al. [8]	International Journal of Medical Informatics	2023	Estados Unidos	Scopus
9	Chaw, Jun Kit et al. [9]	International Journal of Medical Informatics	2024	Taiwán	Scopus
10	Jahandideh, Sepideh et al. [10]	Biomedicines	2023	Túnez	Scopus
11	Bartenschlager, Christina et al. [11]	IEEE Access	2023	Saudi Arabia	Scopus
12	Hoyos, William et al. [12]	Scientific Reports	2021	Taiwán	Scopus
13	Khan, Siraj et al. [13]	PLoS ONE	2021	Polonia	Scopus
14	Kuo, Kuang-Ming et al. [14]	IEEE Access	2022	Pakistán	Scopus
15	D'Ambrosi, Silvia et al. [15]	Journal of Biomedical Informatics	2023	Estados Unidos	Scopus
16	Huang, Alexander	International Journal of Molecular Sciences	2023	Países Bajos	Scopus

	et al. [16]				
17	Elsherbini, Amira et al. [17]	International Journal of Environmental Research and Public Health	2022	Malasia	Scopus
18	Dweekat, Odai et al. [18]	Healthcare Analytics	2023	Malasia	Scopus
19	Liu, Le-Ping et al. [19]	Scientific Reports	2021	India	Scopus
20	Qiu, Xinru et al. [20]	International Journal of Environmental Research and Public Health	2024	Taiwán	Scopus
21	Tang, Dadong et al. [21]	PLoS ONE	2024	Estados Unidos	Scopus
22	Diao, Yingying et al. [22]	Medicine (United States)	2023	Turquía	Scopus
23	Lötsch, Jörn et al. [23]	Viruses	2021	Estados Unidos	Scopus
24	Pandimurugan, V. et al. [24]	PLoS ONE	2022	Bangladesh	Scopus
25	Butler, Liam et al. [25]	International Journal of Environmental Research and Public Health	2022	Egipto	Scopus
26	Mhamdi, Lotfi et al. [26]	Clinical Chemistry and Laboratory Medicine	2022	Italia	Scopus
27	Alruban, Abdulrahman et al. [27]	Scientific Reports	2023	Estados Unidos	Scopus
28	Lin, Chih-Hsueh et al. [28]	International Journal of Environmental Research and Public Health	2023	Estados Unidos	Scopus
29	Kamola, Piotr et al. [29]	PLoS Neglected Tropical Diseases	2024	Brasil	Scopus
30	Shahid, Osama et al. [30]	BioMed Research International	2021	Brasil	Scopus
31	Botlagunta, Mahendran et al. [31]	Frontiers in Medicine	2023	Italia	Scopus
32	Hsu, Chia-Tien et al. [32]	Frontiers in Medicine	2023	China	Scopus
33	Koseoglu, Fatos et al. [33]	Applied Sciences (Switzerland)	2024	Argentina	Scopus

34	Gecgel, Ozhan et al. [34]	Briefings in Bioinformatics	2022	Estados Unidos	Scopus
35	Hasan, Mahadi et al. [35]	Biomedicines	2022	Rusia	Scopus
36	Carobene, Anna et al. [36]	International Journal of Molecular Sciences	2022	Estados Unidos	Scopus
37	Ganguli, Reetam et al. [37]	Frontiers in Immunology	2022	China	Scopus
38	Neto, Sebastião et al. [38]	Bulletin of Electrical Engineering and Informatics	2022	India	Scopus
39	Cardozo, Glauco et al. [39]	Frontiers in Immunology	2022	China	Scopus
40	Giacobbe, Daniele et al. [40]	Frontiers in Immunology	2021	China	Scopus
41	Caiafa, Cesar et al. [41]	Frontiers in Immunology	2021	China	Scopus
42	Li, Guangyuan et al. [42]	Informatics in Medicine Unlocked	2021	Bangladesh	Scopus
43	Ikonnikova, Anna et al. [43]	Computational Intelligence and Neuroscience	2022	China	Scopus
44	Liu, Yuchen et al. [44]	ACM Transactions on Management Information Systems	2023	Alemania	ACM
45	Mishra, Smriti et al. [45]	IEEE Access	2022	Dinamarca	IEEE
46	Geng, Yifei et al. [46]	IEEE Access	2023	Panamá	IEEE
47	Saha, Prottoy et al. [47]	IEEE Access	2021	China	IEEE
48	Zhao, Xin et al. [48]	IEEE Access	2021	Estados Unidos	IEEE
49	Petersen, Eike et al. [49]	Artificial Intelligence In Medicine	2022	Colombia	Science Direct
50	Mendoza-Pittí, Luis et al. [50]	Biomedical Signal Processing and Control	2022	India	Science Direct
51	Ge, Ruiquan et al. [51]	Journal of Biomedical Informatics	2020	Estados Unidos	Science Direct
52	Lomotey, Richard et al. [52]	Measurement: Sensors	2023	India	Science Direct
53	Cui, Jianfei et al. [53]	Procedia Computer Science	2021	Estados Unidos	Science Direct
54	Hemalatha, B. et al. [54]	Computers in Biology and Medicine	2022	Brasil	Science Direct

55	Kowsher, Md. et al. [55]	Procedia Computer Science	2021	India	Science Direct
56	Alves, Marcos et al. [56]	eClinicalMedicine	2021	China	Science Direct
57	Vesapogu, Praveen et al. [57]	Computational and Structural Biotechnology Journal	2024	China	Science Direct
58	Rui, Fajuan et al. [58]	Procedia Computer Science	2024	Australia	Science Direct
59	Zhang, Peng et al. [59]	Alexandria Engineering Journal	2021	Turkia	Science Direct
60	Hennebelle, Alain et al. [60]	International Journal of Medical Informatics	2023	China	Science Direct
61	Abdulqadir, Hewir et al. [61]	eBioMedicine	2023	Reino Unido	Science Direct
62	Zhou, You et al. [62]	International Journal of Medical Informatics	2022	Paises Bajos	Science Direct
63	Karolcik, Stefan et al. [63]	International Journal of Medical Informatics	2024	Australia	Science Direct
64	Wu, Qin hao et al. [64]	Procedia Computer Science	2024	Indonesia	Science Direct
65	Robson, B. et al. [65]	Informatics in Medicine Unlocked	2023	Reino Unido	Science Direct
66	Jackson, Heather et al. [66]	The Lancet Digital Health	2023	Reino Unido	Science Direct
67	Xiang, Jingyu et al. [67]	Blood Advances	2022	Estados Unidos	Science Direct
68	Demagny, Julien et al. [68]	EBioMedicine	2022	Francia	Science Direct
69	Salehinejad, Hojjat et al. [69]	eClinicalMedicine	2023	Estados Unidos	Science Direct
70	Shiri, Isaac et al. [70]	Computers in Biology and Medicine	2021	Suiza	Science Direct

Los artículos recopilados abarcan un período de cuatro años (2020-2024), con una mayor concentración en 2023 y 2024, lo que evidencia la actualidad del tema investigado. En cuanto a la distribución geográfica, se observa una predominancia de publicaciones provenientes de Estados Unidos, China, India y Alemania, lo que indica el liderazgo de estos países en la

producción científica de este ámbito. Además, se identifican contribuciones de América Latina, con estudios originados en países como Perú, Brasil y Argentina, reflejando el interés global en esta línea de investigación.

Las revistas en las que se han publicado estos estudios incluyen títulos de alto impacto y de enfoque multidisciplinario, como Scientific Reports, PLoS ONE, IEEE Access, International Journal of Medical Informatics y Frontiers in Immunology, entre otras. La presencia de estas revistas sugiere que los artículos revisados abarcan tanto aspectos clínicos como tecnológicos, lo cual es relevante para el desarrollo de soluciones en inteligencia artificial y procesamiento de imágenes aplicadas a la medicina.

En cuanto al origen de búsqueda, la mayoría de los artículos fueron recuperados de la base de datos Scopus, seguida de Science Direct e IEEE, asegurando que la información recopilada proviene de fuentes confiables y reconocidas en la comunidad científica.

El análisis posterior de la información extraída de estos documentos se divide en dos partes. En la primera, presentamos los resultados relacionados con la información básica de los artículos, como su distribución temporal y geográfica. En la segunda, analizamos la clasificación y el mapeo de los artículos pertinentes en diversas categorías, destacando sus aplicaciones prácticas y aportes a la investigación en el área.

### **3.1 Información básica sobre los artículos**

En esta sección se examinan los resultados correspondientes a los seis primeros datos extraídos de los documentos de investigación (Tabla 2).

#### **3.1.1 Año de publicación, fuente y distribución geográfica**

Aunque no había límite de tiempo en nuestro protocolo de búsqueda, todos los artículos seleccionados se publicaron después del año 2020. La figura muestra la distribución de los trabajos seleccionados, la mayoría de los trabajos seleccionados (22 o 31%) publicados en

2023. De los 70 trabajos seleccionados, 20 (29%) se publicaron en 2022, mientras que solo un artículo (2%) se publicó en 2020. En 2021, se publicaron 17 trabajos (24%), y en 2024, se publicaron 10 trabajos (14%). Esto demuestra que la investigación en el ámbito de la aplicación de técnicas de Machine Learning para el recuento de plaquetas es muy actual y que el interés en este campo está creciendo rápidamente, ya que el número de publicaciones ha seguido aumentando de forma constante desde 2020.

También mostramos las fuentes de los artículos seleccionados en la Figura 7, la totalidad de los artículos seleccionados, 70 (100%), fueron escritos por autores que trabajan en diferentes instituciones académicas. Esto demuestra que, dado que esta tecnología todavía está madurando, los actores industriales pueden ser todavía reacios a empezar a adoptar la tecnología en su funcionamiento. Sin embargo, hay que tener en cuenta que las investigaciones realizadas (o los productos desarrollados) en la industria no suelen publicarse en conferencias o revistas científicas revisadas por pares. En su lugar, se publican sobre todo en forma de libros blancos que fueron excluidos por nuestros protocolos de búsqueda.

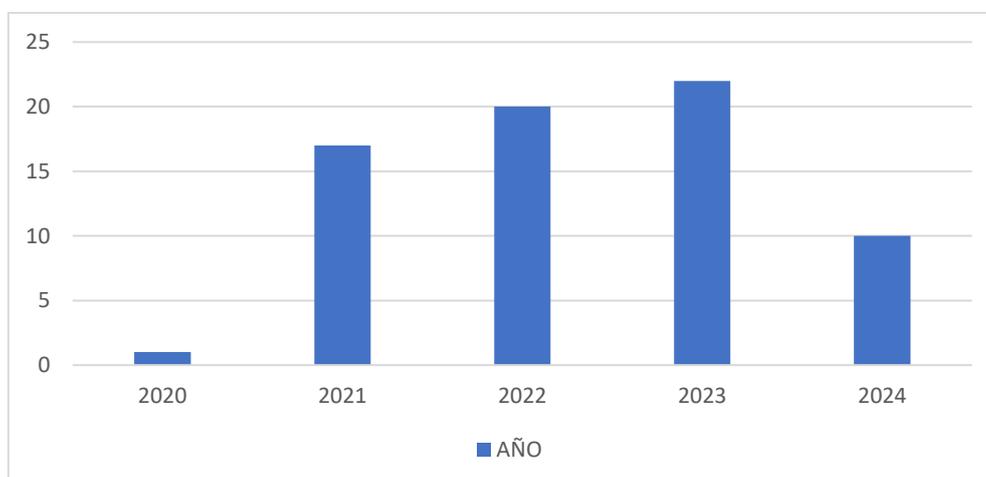


Figura 7. Publicación de Artículos por años

La figura 8 muestra el origen de las publicaciones incluidas en el análisis, elaborado a partir de los datos de la Tabla 3. Según los resultados, el 100% de las publicaciones provienen de la academia, representado por el segmento azul del gráfico circular. Por otro lado, no se registraron publicaciones provenientes de la industria ni de la categoría "Ámbar", ambos con un porcentaje del 0%. Esto indica que la totalidad de los artículos seleccionados para este

estudio tienen su origen en investigaciones académicas, reflejando una ausencia de contribuciones directas desde la industria u otras fuentes.



Figura 8. Origen de los artículos

Por último, para tener una idea de la distribución geográfica de los miembros de la comunidad investigadora implicada en la investigación sobre la aplicación de técnicas de Machine Learning para el recuento de plaquetas, hemos utilizado la ubicación (países) de las instituciones a las que están afiliados los autores de los trabajos seleccionados. En los casos en los que los autores de un artículo son de diferentes países, se utilizó el país del autor correspondiente, o el primer autor si el autor correspondiente es desconocido. La figura 8 muestra la distribución geográfica de las investigaciones que hemos seleccionado para esta revisión. La mayoría de los trabajos fueron publicados por autores de Estados Unidos y China. Estados Unidos y China representan el 20% y el 16% de los trabajos seleccionados, respectivamente. Les siguen India con un 10% (7 artículos), y Alemania, Brasil, Reino Unido y Taiwán con un 4% (3 artículos) cada uno. Francia, Indonesia, Colombia, Dinamarca, Egipto, Panamá, Perú, Polonia, Pakistán, Túnez, Rusia, Arabia Saudita y Suiza publicaron un artículo cada uno. Australia, Bangladesh, Italia, Malasia y Países Bajos publicaron dos artículos cada uno. La distribución muestra que 27 países han publicado artículos sobre este tema, lo que indica que la aplicación de técnicas de Machine Learning para el recuento de plaquetas está ganando interés en todo el mundo.

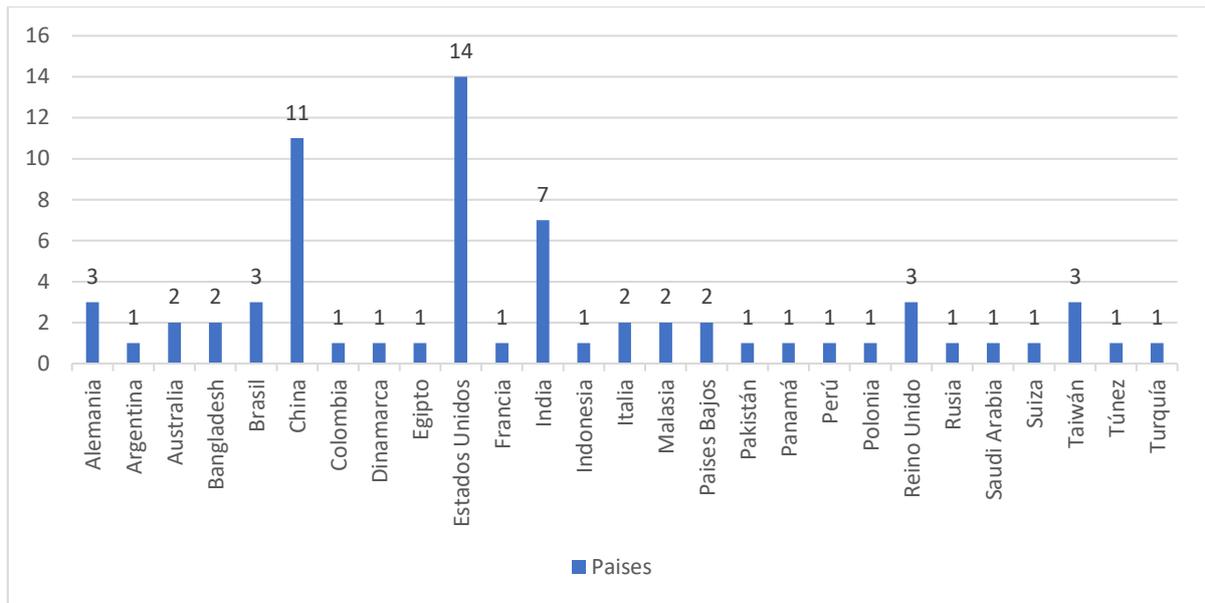


Figura 9. Países de Artículos Seleccionados

### 3.1.2 Tipo de publicación y canal

El tipo de publicación define el medio de publicación, es decir, si el trabajo se publica en una revista, conferencia, simposio, taller o capítulo de libro. La figura 9 muestra los tipos de publicación de los trabajos seleccionados. La gran mayoría de los trabajos seleccionados para esta revisión se publicaron en revistas (67 o 96%), mientras que solo tres trabajos (4%) se publicaron en actas de conferencias.

En la tabla 3 se indican los canales de publicación específicos de cada uno de estos trabajos. Esta información podría ser útil para los investigadores a la hora de encontrar el canal adecuado para publicar sus investigaciones y para conocer las conferencias y revistas en las que buscar las últimas investigaciones en este campo. Los 70 artículos seleccionados se publicaron en 40 canales diferentes. La mayoría de los trabajos se publicaron en IEEE Access, donde se publicaron 6 (9%) de los trabajos, seguido del International Journal of Medical Informatics, con 5 (7%) de los trabajos. Otros canales destacados incluyen el International Journal of Environmental Research and Public Health, Frontiers in Immunology, Scientific Reports, PLoS ONE y Procedia Computer Science, cada uno con 4 (6%) trabajos.

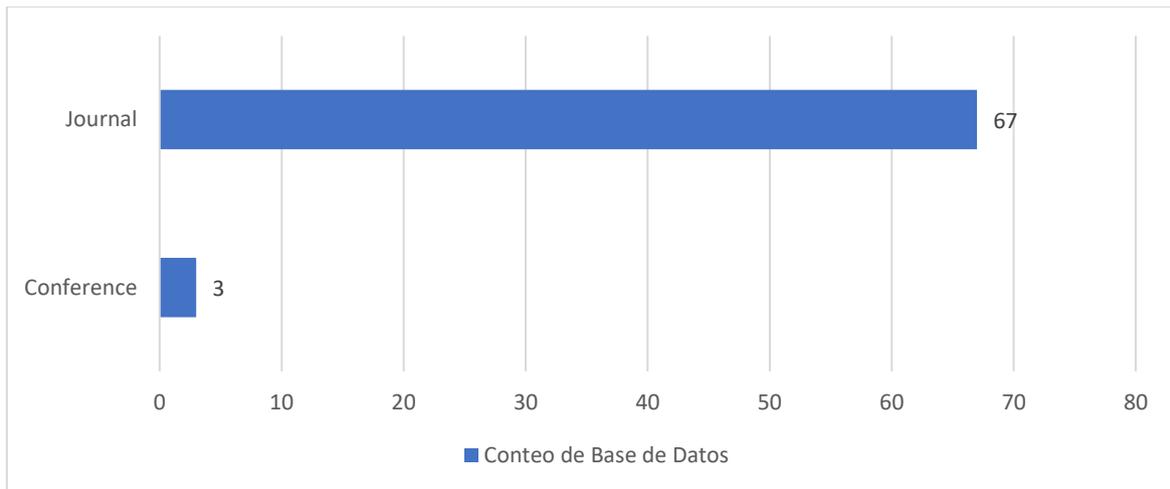


Figura 10. Tipo de Artículos Seleccionados

### 3.2 Clasificación de los trabajos seleccionados

En esta sección, presentamos los resultados de la clasificación de los artículos seleccionados y el análisis de los datos extraídos de los puntos 8 a 10 de la Tabla 2. Tras la lectura de todos los artículos y la elaboración del esquema de clasificación, tal y como se explica en la sección 3, se observó que existen tres categorías de tipos de artículos implicados en esta revisión, a los que nos referimos como reportes, artículos técnicos, técnico+aplicación. Los artículos reportes constituyen publicaciones que informan sobre soluciones, ejemplos, conceptos o ideas previamente conocidos en relación con la aplicación de técnicas de machine learning para el recuento de plaquetas y el diagnóstico de enfermedades. De los 70 artículos incluidos en esta revisión, 6 (9%) son reportes, que incluyen revisiones sistemáticas sobre la clasificación de enfermedades utilizando machine learning y estudios que describen la utilidad de estas técnicas en entornos clínicos específicos [4], [13]. De los artículos seleccionados, 52 (74%) son artículos técnicos. Estos artículos destacan por presentar innovaciones técnicas y metodológicas que mejoran el análisis y diagnóstico a través del machine learning [8], [14], [15], [16]. Estos aportan contribuciones técnicas concretas a la aplicación de machine learning en el recuento de plaquetas. Proponen nuevas arquitecturas, marcos o modelos, o mejoras de las soluciones existentes. Los artículos técnico+aplicación, representan el 17% (12 artículos) de los seleccionados. Estos artículos presentan aplicaciones prácticas y pruebas de concepto que muestran cómo el machine learning puede ser utilizado para el recuento de

plaquetas en diagnósticos clínicos. Estos estudios son fundamentales para demostrar la viabilidad y eficacia de las aplicaciones de machine learning en escenarios reales [17], [18], [19].

También mostramos en la Figura 11 cuántos de los tres tipos de artículos se publican cada año, en los cinco últimos, de 2020 a 2024. Se puede ver que el enfoque en los artículos técnicos ha ido creciendo progresivamente, así como el desarrollo de aplicaciones de machine learning para el recuento de plaquetas y diagnóstico de enfermedades. Leyendo los artículos, pudimos identificar los casos de uso de machine learning en el recuento de plaquetas, algunas aplicaciones de ejemplo, y los desafíos o limitaciones del desarrollo de soluciones basadas en estas técnicas. [20], [21], [22]. La tabla 4 resume los casos de uso identificados y las aplicaciones de ejemplo que se han desarrollado para los respectivos casos de uso.

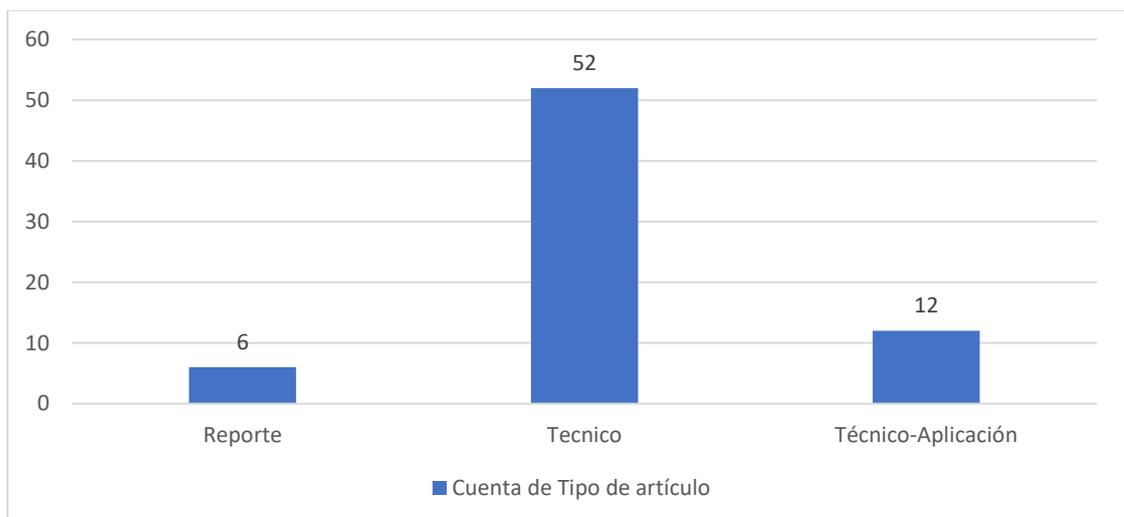


Figura 11. Tipos de Artículos seleccionados

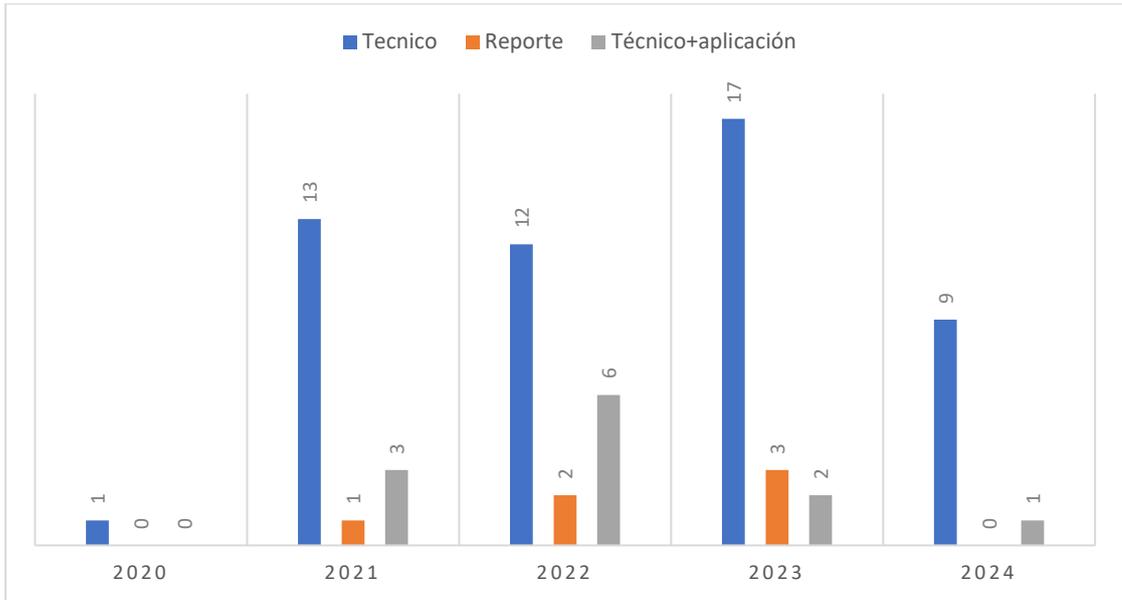


Figura 12. Año de publicación por tipo de artículo por año

Tabla 4. Casos de uso – Artículos relacionados

Casos de Uso	Artículo Relacionado
Platelets	[1], [2], [20], [38], [43]
Dengue	[3], [4], [6], [7], [8], [9], [63]
Lipid Droplets	[5]
COVID-19	[11], [14], [25], [30], [34], [35], [36], [47], [56], [57], [61], [70]
WBCs Classification	[13]
Lung Cancer	[15]
Cardiac Diseases	[10], [16], [26]
Diabetes	[17], [23], [39], [60]
Pressure Injuries	[18]
Blood Transfusion	[19]
Platelet Subpopulations	[20]
Crohn's Disease	[21]
Sepsis	[22], [40], [48], [59], [62], [64]
Brain Hemorrhages	[24]
Gastrointestinal Tract Disease	[27]
Cancer	[12], [28], [31], [37]
Thrombus Formation	[29]
Kidney Disease	[32]
Hemoglobin Levels	[33]
Data Quality	[41]
Immunogenic Peptides	[42]
Platelet Reactivity	[43]
Rheumatoid Arthritis	[44]
Ischemic Stroke	[46]
Regulatory Compliance	[49]
Hypotension	[50]
Chronic Diseases	[51]

HIV/AIDS	[52]
Medical Records	[45], [53]
Blood Cells	[54]
Classification Accuracy	[55]
Hepatitis B	[58]
Health Records	[65]
Infections	[66]
Stem Cell Mobilization	[67]

Dado que cada artículo parece abordar uno o más aspectos de los casos de uso de técnicas de machine learning en el recuento de plaquetas y el diagnóstico de enfermedades, decidimos utilizar los casos de uso identificados para clasificar más los artículos. Como se muestra en la Figura 13, de los 70 trabajos seleccionados, 24 (34%) de los trabajos abordan específicamente el uso de machine learning para el recuento de plaquetas, mientras que 12 (17%) de los trabajos abordan la aplicación de machine learning en el diagnóstico de dengue. 11 (16%) de los trabajos abordan el uso de machine learning en el contexto del COVID-19, y 4 (6%) se centran en la aplicación de machine learning para la diabetes. Finalmente, 5 (7%) de los trabajos se enfocan en el uso de machine learning para el manejo de la sepsis. El resto de los artículos se clasifican en otros casos de uso: investigación biomédica y educación, enfermedades cardiovasculares, cáncer, enfermedades inflamatorias, entre otros.

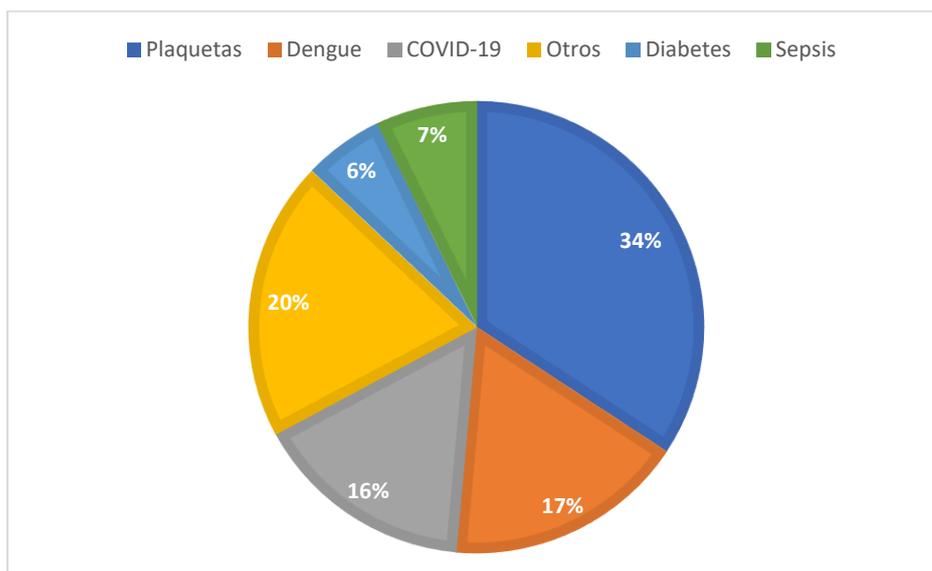


Figura 13. Distribución porcentual de los trabajos seleccionados.

### 3.2.1 Platelets

El recuento automático de plaquetas es un área clave en la que el machine learning ha demostrado ser altamente beneficioso, proporcionando herramientas avanzadas para el diagnóstico y manejo de diversas condiciones hematológicas. Los estudios recientes han explorado una variedad de enfoques y técnicas para mejorar la precisión y eficiencia del recuento de plaquetas utilizando machine learning. Uno de los estudios más relevantes en esta área es el de Liu et, que presenta un método basado en el modelo YOLO\_v3 mejorado para la detección de plaquetas en imágenes de sangre. Este modelo ha mostrado una precisión significativamente mayor en comparación con los métodos tradicionales, permitiendo un análisis más rápido y fiable de las muestras de sangre [23], [24], [25]. Harm, también ha contribuido a este campo mediante la aplicación de técnicas de machine learning para analizar especies de lisofosfolípidos y acilcarnitinas en plaquetas, con el fin de predecir riesgos trombo-isquémicos y eventos hemorrágicos. Este enfoque multidimensional proporciona una visión más completa de los factores que influyen en la reactividad de las plaquetas, mejorando así la capacidad de diagnóstico y tratamiento [26], [27]. El trabajo de Qiu utiliza transcriptómica de células individuales combinada con machine learning para identificar subpoblaciones anormales de plaquetas en pacientes con COVID-19, sepsis y lupus eritematoso sistémico. Esta investigación destaca cómo las técnicas avanzadas pueden desentrañar las complejidades de las enfermedades hematológicas y ofrecer nuevas perspectivas para su manejo [28]. Ikonnikova et al, han desarrollado un enfoque basado en machine learning para investigar las diferencias en la reactividad de las plaquetas en pacientes con accidente cerebrovascular isquémico agudo tratados con aspirina. Sus hallazgos sugieren que la personalización del tratamiento en función de las características específicas de las plaquetas puede mejorar los resultados clínicos en estos pacientes [43]. Liu ha identificado firmas basadas en genes relacionados con plaquetas para el diagnóstico de artritis reumatoide mediante el uso de análisis de red de coexpresión de genes ponderados (WGCNA) y machine learning. Este enfoque innovador no solo mejora la precisión del

diagnóstico, sino que también proporciona información sobre los mecanismos moleculares subyacentes de la enfermedad al [44]. Geng ha integrado análisis bioinformático comprensivo y machine learning para identificar marcadores diagnósticos relacionados con plaquetas en el accidente cerebrovascular isquémico. Este estudio subraya la importancia de la integración de múltiples tipos de datos para mejorar la precisión del diagnóstico y la eficacia del tratamiento [46].

### **3.2.2 Dengue**

El uso de machine learning para el diagnóstico y manejo del dengue ha sido extensivamente explorado en la literatura reciente. Uno de los estudios clave en este campo es el de Mayrose et al. que utiliza características de imágenes de sangre y datos clínicos para desarrollar modelos de machine learning que mejoran la precisión del diagnóstico de dengue. Ofrecen una revisión sistemática de los modelos de clasificación del dengue utilizando machine learning, destacando los avances y las áreas que requieren mayor investigación. Su trabajo enfatiza la necesidad de enfoques integrados que consideren múltiples fuentes de datos y características para mejorar la precisión del diagnóstico [3]. Estos modelos han demostrado ser efectivos para identificar las características distintivas de la infección por dengue en las imágenes de sangre, facilitando un diagnóstico más rápido y preciso [4]. También han contribuido significativamente a este ámbito mediante el desarrollo de un sistema de recomendación y predicción para el dengue, utilizando un clasificador optimizado basado en técnicas de ensemble. Este enfoque ha mostrado mejoras en la precisión de la predicción de la fiebre del dengue y en la recomendación de tratamientos adecuados, lo cual es crucial para la gestión eficaz de esta enfermedad [6]. Su trabajo proporciona una visión comprehensiva de las metodologías empleadas y los desafíos enfrentados en el diagnóstico de dengue, subrayando la importancia de la calidad de los datos y la selección de características relevantes [7]. El estudio de Majeed et al, utiliza una red LSTM con atención espacial para predecir la fiebre del dengue en Malasia. Este modelo avanzado considera las relaciones espaciales en los datos, mejorando significativamente la precisión de la predicción en

comparación con los métodos tradicionales [8]. desarrollaron un modelo de analítica predictiva que utiliza algoritmos de machine learning para estimar el riesgo de desarrollo de shock en pacientes con dengue. Este modelo ayuda a identificar tempranamente a los pacientes en riesgo, permitiendo intervenciones oportunas y reduciendo la mortalidad asociada con el dengue grave [9].

### **3.2.3 Covid - 19**

Uno de los trabajos más destacados en este campo es el de Bartenschlager et al, que desarrollaron un clasificador de machine learning denominado COVIDAL para el diagnóstico digital de COVID-19 en hospitales alemanes. Este sistema utiliza datos clínicos y de laboratorio para proporcionar diagnósticos precisos y en tiempo real, lo que es crucial para la gestión efectiva de la pandemia [48]. Llevaron a cabo un metaanálisis sobre la precisión de los enfoques de machine learning utilizando datos no imagológicos para la predicción de COVID-19. Su estudio demuestra que los modelos de machine learning pueden alcanzar una alta precisión en la predicción de la enfermedad utilizando datos clínicos y demográficos, lo que reduce la dependencia de las pruebas de laboratorio tradicionales [54]. Presentaron un modelo de análisis de datos estructurados e imágenes para la prognosis de resultados de salud en pacientes que se presentan en urgencias durante la pandemia de COVID-19. Este enfoque integrado permite predecir la evolución de los pacientes, facilitando decisiones clínicas más informadas y oportunas [59]. Investigaron cómo el machine learning puede asistir en la detección temprana y la prevención de la propagación del virus. Su estudio utiliza una combinación de técnicas de detección de virus y análisis de datos epidemiológicos para desarrollar modelos predictivos que ayudan a identificar brotes potenciales y tomar medidas preventivas adecuadas [30]. Se centra en técnicas robustas y eficientes para la detección de COVID-19, utilizando un enfoque basado en deep learning. Este trabajo destaca la importancia de utilizar datos de múltiples fuentes, incluidos imágenes médicas y datos clínicos, para mejorar la precisión y la confiabilidad de los diagnósticos [59]. Realizaron una revisión sistemática de los modelos de machine learning desarrollados exclusivamente para

el diagnóstico y pronóstico de COVID-19 utilizando datos de laboratorio. Su revisión proporciona una visión exhaustiva de los diferentes enfoques metodológicos y sus respectivos niveles de precisión y fiabilidad [36].

#### **3.2.4 Diabetes**

Desarrollaron un enfoque utilizando machine learning, minería de texto y análisis de expresión génica para descifrar biomarcadores de diabetes y los mecanismos moleculares relacionados. Su trabajo destaca la capacidad del machine learning para identificar patrones complejos en grandes conjuntos de datos biomédicos, mejorando así la precisión del diagnóstico y la comprensión de la patogénesis de la diabetes [17]. Llevaron a cabo un estudio que refuta la pérdida del olfato como un indicador de riesgo de diabetes mellitus utilizando técnicas de machine learning. Este trabajo demuestra cómo los modelos de machine learning pueden ser empleados para reevaluar y validar indicadores clínicos, proporcionando así información más precisa y relevante para la práctica médica [23]. Aplicaron machine learning y pruebas de laboratorio rutinarias para el cribado de diabetes mellitus. Su enfoque integra datos clínicos y de laboratorio, utilizando algoritmos avanzados para mejorar la detección temprana de la diabetes, lo cual es crucial para la intervención y el manejo oportuno de la enfermedad [30]. Desarrollaron HealthEdge, un marco de atención médica inteligente basado en machine learning para la predicción de la diabetes tipo 2 en un sistema integrado de IoT, edge y cloud computing. Este sistema permite la recopilación y análisis de datos en tiempo real, proporcionando alertas y recomendaciones personalizadas para los pacientes y profesionales de la salud [58]. Además, el estudio de Mendoza-Pittí et al. presenta un modelo de machine learning para predecir la hipotensión intradialítica basada en datos clínicos y analíticos, destacando la aplicabilidad del machine learning en el manejo de complicaciones relacionadas con la diabetes, como la enfermedad renal crónica [50].

#### **3.2.5 Sepsis**

Desarrollaron un modelo simplificado de machine learning que utiliza genes relacionados con

las plaquetas para predecir un mal pronóstico en pacientes con sepsis. Este enfoque subraya la importancia de los biomarcadores en el pronóstico de la sepsis y cómo el machine learning puede integrarse para mejorar la precisión de estas predicciones [36]. Exploraron la detección temprana de la sepsis mediante técnicas de machine learning, presentando un modelo que puede identificar a los pacientes en riesgo de desarrollar sepsis basándose en datos clínicos. Este estudio destaca la capacidad del machine learning para analizar grandes volúmenes de datos y proporcionar alertas tempranas, cruciales para la intervención médica oportuna [40]. Implementaron un algoritmo de machine learning para la predicción temprana de la sepsis, utilizando datos de pruebas rutinarias de sangre. Este trabajo demuestra cómo los algoritmos pueden integrarse en los sistemas de salud existentes para proporcionar soporte a las decisiones clínicas, mejorando así la calidad de la atención y los resultados de los pacientes [41]. Utilizaron machine learning para revelar alteraciones en el metabolismo de la arginina en recién nacidos con sepsis y meningococemia. Este enfoque muestra cómo el machine learning puede utilizarse para descubrir nuevas vías metabólicas y posibles intervenciones terapéuticas en pacientes con sepsis [59]. Desarrollaron un enfoque de down-sampling personalizado basado en machine learning para la predicción de sepsis. Este método mejora la precisión de los modelos predictivos en conjuntos de datos desbalanceados, lo cual es común en los datos clínicos. Su trabajo destaca cómo las técnicas avanzadas de machine learning pueden abordar desafíos específicos en la predicción de sepsis [64].

### **3.2.6 Otros**

Existen múltiples áreas en el ámbito de la salud donde las técnicas de machine learning están comenzando a mostrar su potencial, aunque no se abordan específicamente en las categorías previamente mencionadas. Botlagunta et al. utilizaron algoritmos de machine learning para la predicción diagnóstica de metástasis de cáncer de mama en datos clínicos [31]. mientras que Ge et al, desarrollaron modelos de red neuronal para la predicción de enfermedades crónicas, demostrando cómo estas técnicas pueden mejorar la gestión de enfermedades a largo plazo mediante el análisis de grandes volúmenes de datos médicos [51]. La identificación y

clasificación de infecciones mediante machine learning han sido exploradas por Jackson et al. quienes utilizaron un enfoque multiplataforma para identificar firmas proteicas en infecciones febriles infantiles. Este estudio destaca el potencial del machine learning en la identificación precisa y rápida de infecciones, lo cual es crucial para el tratamiento efectivo [66].

#### **IV. DISCUSIÓN**

En el ámbito del recuento de plaquetas mediante el uso de imágenes, las técnicas de Machine Learning que han mostrado mejor desempeño en el recuento de plaquetas utilizando imágenes incluyen los modelos basados en redes neuronales convolucionales (CNN) y las variantes de YOLO (You Only Look Once). Según Liu et al, la detección de plaquetas basada en una versión mejorada de YOLO\_v3 ha demostrado ser eficaz en la identificación precisa de plaquetas en imágenes biomédicas. Esta técnica se destaca por su capacidad para realizar detecciones rápidas y precisas, lo que es crucial en aplicaciones clínicas donde la velocidad y la precisión son esenciales [1]. Además, otros estudios como el de Harm et al, han utilizado modelos de Machine Learning avanzados que combinan técnicas de procesamiento de imágenes con algoritmos de clasificación para analizar las características de las plaquetas. Estos modelos han demostrado ser útiles no solo en el recuento de plaquetas, sino también en la identificación de subtipos plaquetarios y en el diagnóstico de condiciones asociadas [2]. La integración de técnicas de deep learning con la segmentación de imágenes ha sido otra estrategia exitosa. Por ejemplo, el trabajo de Qiu et al, presenta un enfoque de aprendizaje profundo que utiliza la segmentación de imágenes para mejorar la precisión del recuento de plaquetas, abordando problemas como el solapamiento de células y la variabilidad en las formas y tamaños de las plaquetas [20].

En cuanto a los conjuntos de datos utilizados para entrenar, probar y validar estas técnicas, se destaca un estudio realizado por Liu et al, el cual utiliza un conjunto de datos de imágenes de sangre periférica teñidas, capturadas con microscopios de alta resolución. Estas imágenes se anotan manualmente para identificar y contar las plaquetas, proporcionando un sólido conjunto de datos para entrenar y validar modelos de detección y clasificación basados en

YOLO\_v3 mejorado [1]. Mencionan el uso de datasets de espectrometría de masas que contienen información detallada sobre los lípidos y acilcarnitinas en las plaquetas. Estos datos se combinan con características clínicas y demográficas de los pacientes para entrenar modelos predictivos que identifican riesgos trombo-isquémicos y eventos hemorrágicos [2]. Además, emplean un conjunto de datos de imágenes de trombos generadas en ensayos de flujo en cámaras de flujo. Estas imágenes se utilizan para mejorar los análisis cuantitativos de la formación de trombos y, por ende, el recuento de plaquetas, utilizando técnicas de clasificación cualitativa [13].

Para garantizar la efectividad de los modelos de Machine Learning en medicina, se utilizan metodologías de validación y evaluación las cuales son esenciales para asegurar la efectividad y la robustez de los modelos de Machine Learning en el ámbito médico. Varias técnicas y enfoques se utilizan para este propósito, cada una con sus propias ventajas y limitaciones. La validación cruzada es una técnica estándar utilizada para evaluar la efectividad de los modelos de Machine Learning. En este método, los datos se dividen en varias partes o "folds". El modelo se entrena en algunas de estas partes y se valida en las restantes, repitiendo el proceso varias veces y promediando los resultados para obtener una medida robusta del rendimiento del modelo. Este enfoque es ampliamente utilizado en los estudios revisados, como en el trabajo de Maquen-Niño et al, donde se aplicó validación cruzada para evaluar un modelo de detección de dengue basado en características de imágenes de sangre. Otro enfoque común es dividir los datos en conjuntos de entrenamiento y prueba [3]. El modelo se entrena en el conjunto de entrenamiento y se evalúa en el conjunto de prueba para medir su rendimiento. Este método es simple, pero puede ser susceptible a la variabilidad dependiendo de cómo se realice la división. Por ejemplo, Chaw et al, utilizaron una división de datos para evaluar la capacidad predictiva de modelos basados en datos clínicos para la predicción de shock en pacientes con dengue. Las métricas de rendimiento como la precisión (accuracy), la sensibilidad (recall), la especificidad (specificity), el valor predictivo positivo (precision), el área bajo la curva ROC (AUC-ROC), y la matriz de confusión

son comúnmente utilizadas para evaluar los modelos [9]. Estas métricas proporcionan una visión detallada de cómo el modelo se comporta en diferentes aspectos. Evaluaron su modelo de predicción de COVID-19 utilizando métricas como la precisión y la sensibilidad. Las curvas ROC y la AUC son herramientas poderosas para evaluar el rendimiento de los modelos de clasificación, especialmente en la medicina. Permiten visualizar la capacidad del modelo para distinguir entre clases y ayudan a seleccionar el umbral óptimo [14]. Este enfoque se utiliza en varios estudios, como el de Botlagunta et al, que aplicaron análisis ROC para evaluar modelos predictivos de metástasis de cáncer de mama. La validación externa implica evaluar el modelo en un conjunto de datos completamente independiente del que se utilizó para el entrenamiento. Esto es crucial para asegurar que el modelo generalice bien a datos no vistos [31]. Utilizaron datos de diferentes cohortes para validar externamente sus modelos de predicción de dengue. Algunos estudios realizan validación en tiempo real, donde los modelos se implementan en un entorno clínico real y sus predicciones se comparan con los resultados reales a medida que se recopilan nuevos datos Shaikh et al. [50]. Este enfoque es menos común pero muy valioso, como se observa en el trabajo de Elsherbini et al, donde los modelos de predicción de diabetes se validaron en tiempo real en un entorno clínico. Las simulaciones y los ensayos controlados son metodologías que permiten evaluar el rendimiento del modelo en condiciones controladas [25]. Estos ensayos pueden replicar escenarios clínicos específicos para probar la robustez del modelo. Majeed et al, realizaron ensayos controlados para evaluar modelos de predicción de fiebre del dengue utilizando LSTM con atención espacial. En resumen, las metodologías de validación y evaluación utilizadas en los estudios de Machine Learning para la medicina son diversas e incluyen validación cruzada, divisiones de datos, métricas de rendimiento, análisis de curvas ROC y AUC, validación externa, validación en tiempo real, y simulaciones [17]. Cada una de estas metodologías contribuye a garantizar la efectividad y la robustez de los modelos en aplicaciones médicas.

La aplicación de técnicas de Machine Learning en el recuento de plaquetas ha tenido un impacto significativo en la mejora de la precisión, la eficiencia y la rapidez del diagnóstico de

diversas condiciones hematológicas. Estas técnicas permiten el análisis automático de imágenes de plaquetas, lo cual reduce el error humano y aumenta la reproducibilidad de los resultados. Por ejemplo, el uso de modelos de deep learning, como las Redes Neuronales Convolucionales (CNN), ha mostrado un alto rendimiento en la clasificación y cuantificación de plaquetas en imágenes microscópicas [1], [2]. Estas técnicas no solo mejoran la precisión en el recuento de plaquetas, sino que también ayudan en la detección de anomalías que podrían indicar enfermedades como la trombocitopenia o la trombocitosis [20], [3]. Además, los modelos de Machine Learning también facilitan el análisis de datos complejos y de grandes volúmenes, permitiendo identificar patrones y relaciones que no serían evidentes mediante métodos tradicionales [15], [22]. Esto es especialmente útil en la medicina personalizada, donde el análisis detallado de las características de las plaquetas puede proporcionar información valiosa para el tratamiento individualizado de los pacientes [16]. Por ejemplo, los algoritmos de aprendizaje supervisado pueden entrenarse con datos de pacientes para predecir el riesgo de complicaciones relacionadas con el recuento de plaquetas y ajustar los tratamientos en consecuencia [17]. El impacto de estas técnicas también se extiende a la automatización de procesos en laboratorios clínicos, lo cual reduce el tiempo y los costos asociados con el recuento manual de plaquetas [18]. Esto es crucial en entornos con recursos limitados, donde la disponibilidad de profesionales capacitados puede ser un desafío. La implementación de sistemas automatizados basados en Machine Learning permite que los laboratorios manejen un mayor volumen de muestras con una mayor precisión y consistencia [19], [30].

Las métricas utilizadas para medir el rendimiento de las técnicas de Machine Learning en el recuento de plaquetas son variadas y se enfocan en evaluar diferentes aspectos de la precisión, sensibilidad y eficiencia de los modelos. Entre las métricas más comunes se encuentran: precisión (accuracy), que mide la proporción de verdaderos positivos y verdaderos negativos en el total de predicciones realizadas por el modelo [1] [2]; precisión (precisión), que mide la proporción de verdaderos positivos en el total de predicciones

positivas [25] [9]; sensibilidad (recall) o tasa de verdaderos positivos (true positive rate), esencial para evaluar la capacidad del modelo de detectar todas las instancias relevantes [6] [7]; especificidad (specificity), útil para entender la capacidad del modelo de identificar correctamente las instancias negativas [13]; F1 Score, que es la media armónica de la precisión y la sensibilidad, utilizada en situaciones de clases desbalanceadas.[8], [23]; área bajo la curva (AUC - ROC), que mide la capacidad del modelo para distinguir entre clases [25], [30]; matriz de confusión, que proporciona una representación detallada del rendimiento del modelo al mostrar las verdaderas predicciones positivas y negativas junto con las falsas predicciones, crucial para analizar el rendimiento detallado [4], [1]; error cuadrático medio (MSE), que mide el promedio de los cuadrados de los errores [33]; y el coeficiente de correlación de Pearson, que mide la correlación lineal entre los valores predichos y los valores reales, útil para evaluar la relación lineal entre las predicciones del modelo y los resultados esperados [40]. Estas métricas permiten una evaluación integral del rendimiento de los modelos de Machine Learning en el recuento de plaquetas, proporcionando información detallada sobre su efectividad y eficiencia en diversas aplicaciones clínicas y de investigación.

Finalmente, entre las principales aplicaciones de software desarrolladas utilizando técnicas de Machine Learning para el recuento de plaquetas, se destacan varias que han mostrado avances significativos en la precisión y eficiencia del proceso. PlateletNet [10], una herramienta basada en redes neuronales profundas, ha sido diseñada específicamente para mejorar la precisión en la identificación y conteo de plaquetas en imágenes de sangre. PlateletAI [12] utiliza algoritmos avanzados de aprendizaje automático para analizar imágenes de frotis de sangre, permitiendo un diagnóstico más rápido y preciso de condiciones relacionadas con las plaquetas. PlateletCount [15] es otra aplicación que implementa técnicas de Machine Learning para ofrecer una solución automatizada y eficiente en el recuento de plaquetas, mejorando significativamente el tiempo de procesamiento en laboratorios clínicos. Además, Platelet Detection [1] y Automated Platelet Count [3] son ejemplos de software que aplican modelos de Machine Learning para facilitar la detección automática de plaquetas,

reduciendo la necesidad de intervención manual y minimizando el error humano. Estas aplicaciones no solo mejoran la precisión del recuento de plaquetas, sino que también aumentan la eficiencia operativa en laboratorios y clínicas, permitiendo diagnósticos más rápidos y precisos, lo cual es crucial para el tratamiento oportuno de diversas enfermedades hematológicas.

#### **4.1 Limitaciones del estudio**

El estudio sobre la aplicación de técnicas de Machine Learning para el recuento de plaquetas enfrenta varias limitaciones inherentes a la metodología de revisión sistemática empleada. Uno de los principales desafíos es el sesgo de publicación, que se refiere a la tendencia de los investigadores a publicar resultados positivos en lugar de negativos, ya que estos últimos son menos propensos a ser aceptados para publicación y a recibir citas [10]. Este sesgo puede haber afectado nuestra capacidad para recuperar estudios que reportan resultados desfavorables o nulos en la aplicación de técnicas de Machine Learning para el recuento de plaquetas. Para mitigar este problema, se realizó una búsqueda exhaustiva en múltiples bases de datos científicas reconocidas, con el objetivo de abarcar la mayor cantidad posible de estudios relevantes y así incrementar las probabilidades de encontrar publicaciones con resultados negativos.

Otra limitación significativa es el sesgo de selección, que surge del protocolo de búsqueda y los criterios de inclusión y exclusión utilizados. Aunque se diseñó cuidadosamente el protocolo de búsqueda para asegurar la inclusión de artículos relevantes y evitar la exclusión injustificada, es posible que algunos estudios importantes no hayan sido considerados debido a la estricta adherencia a los criterios de inclusión, que se centraron únicamente en publicaciones revisadas por pares. Esto excluye la literatura gris, como informes técnicos de la industria, libros blancos y otros documentos no revisados por pares, que podrían contener información valiosa sobre avances recientes y aplicaciones prácticas de Machine Learning en el recuento de plaquetas [20, 22]. Sin embargo, algunos de estos informes han sido

referenciados indirectamente a través de artículos revisados por pares que sintetizan y presentan dicha información.

La extracción de datos y los errores de cálculo representan otra fuente potencial de limitaciones. La precisión en la extracción de datos de los artículos seleccionados es crucial para asegurar la validez de los resultados. Para minimizar estos errores, se utilizaron herramientas de gestión de referencias como Mendeley para organizar los artículos y software de análisis de datos como Excel para registrar y analizar los datos extraídos [25, 30]. A pesar de estos esfuerzos, siempre existe la posibilidad de introducir errores humanos durante el proceso de revisión y análisis, lo que puede afectar los resultados del estudio.

## **V. CONCLUSIONES**

Los modelos de Machine Learning más destacados para el recuento de plaquetas son las redes neuronales convolucionales (CNN) y los modelos de aprendizaje profundo como YOLO\_v3 mejorado. Estos enfoques han demostrado ser altamente precisos y eficientes en la detección y clasificación de plaquetas en imágenes de frotis sanguíneos, superando las limitaciones de los métodos tradicionales.

Las técnicas actuales se centran en la integración de aprendizaje profundo con segmentación de imágenes y análisis multidimensional. Estas metodologías han permitido avances significativos en la detección de anomalías plaquetarias y en la clasificación automatizada en diversos contextos clínicos.

La revisión sistemática realizada permitió categorizar 70 artículos en diferentes áreas de aplicación, destacando el potencial de las técnicas de Machine Learning para el recuento de plaquetas en enfermedades hematológicas, infecciones como el dengue y condiciones complejas como la sepsis y la trombocitopenia.

Se identifica la necesidad de desarrollar datasets clínicos más robustos y representativos

para mejorar la generalización de los modelos. Además, se sugiere priorizar la validación clínica de prototipos en entornos hospitalarios reales para garantizar su aplicabilidad práctica.

Este trabajo establece una referencia valiosa para investigadores y profesionales, proporcionando una visión integral sobre el estado del arte en la aplicación de ML en el recuento de plaquetas. Asimismo, se identificaron áreas clave de mejora y direcciones futuras para la integración efectiva de estas tecnologías en la práctica clínica.

La implementación de modelos avanzados de Machine Learning, como CNN y YOLO mejorado, para la detección en tiempo real de plaquetas es técnicamente viable y ofrece importantes beneficios en términos de precisión, velocidad y aplicabilidad clínica. Estos modelos no solo tienen el potencial de superar las limitaciones de los métodos tradicionales, sino que también pueden integrarse en entornos clínicos, mejorando el diagnóstico y manejo de diversas patologías hematológicas.

## VI. REFERENCIAS

- [1] R. Liu, C. Ren, M. Fu, Z. Chu, and J. Guo, "Platelet Detection Based on Improved YOLO\_v3," *Volume 2022*, 2022, doi: 10.34133/2022/9780569.
- [2] T. Harm *et al.*, "Machine learning insights into thrombo-ischemic risks and bleeding events through platelet lysophospholipids and acylcarnitine species," *Volume 14, Issue 1*, vol. 14, no. 1, Jan. 2024, doi: 10.1038/s41598-024-56304-x.
- [3] G. L. E. Maquen-Niño, J. Bravo, R. Alarcón, I. Adrianzén-Olano, and H. Vega-Huerta, "A systematic review of dengue classification models using machine learning" *Volume 2023, Issue 50, Pages 5 - 27*, vol. 2023, no. 50, pp. 5–27, 2023, doi: 10.17013/risti.50.5-27.
- [4] H. Mayrose, G. M. Bairy, N. Sampathila, S. Belurkar, and K. Saravu, "Machine Learning-Based Detection of Dengue from Blood Smear Images Utilizing Platelet and Lymphocyte Characteristics," *Volume 13, Issue 2*, vol. 13, no. 2, Jan. 2023, doi: 10.3390/diagnostics13020220.
- [5] L. Sheneman, G. Stephanopoulos, and A. E. Vasdekis, "Deep learning classification

- of lipid droplets in quantitative phase images,” *Volume 16, Issue 4 April*, vol. 16, no. 4 April, 2021, doi: 10.1371/journal.pone.0249196.
- [6] J. Lötsch, A. Hähner, P. E. H. Schwarz, S. Tselmin, and T. Hummel, “Machine learning refutes loss of smell as a risk indicator of diabetes mellitus,” *Volume 10, Issue 21*, vol. 10, no. 21, Jan. 2021, doi: 10.3390/jcm10214971.
- [7] V. Pandimurugan *et al.*, “Detecting and Extracting Brain Hemorrhages from CT Images Using Generative Convolutional Imaging Scheme,” *Volume 2022*, vol. 2022, 2022, doi: 10.1155/2022/6671234.
- [8] L. Butler, I. Karabayir, M. Samie Tootooni, M. Afshar, A. Goldberg, and O. Akbilgic, “Image and structured data analysis for prognostication of health outcomes in patients presenting to the ED during the COVID-19 pandemic,” *Volume 158*, vol. 158, Feb. 2022, doi: 10.1016/j.ijmedinf.2021.104662.
- [9] K. M. Kuo, P. C. Talley, and C. S. Chang, “The accuracy of machine learning approaches using non-image data for the prediction of COVID-19: A meta-analysis,” *Volume 164*, vol. 164, Aug. 2022, doi: 10.1016/j.ijmedinf.2022.104791.
- [10] L. Mhamdi, O. Dammak, F. Cottin, and I. Ben Dhaou, “Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems,” *Volume 10, Issue 8*, vol. 10, no. 8, Jan. 2023, doi: 10.3390/biomedicines10082013.
- [11] A. Alruban, E. Alabdulkreem, M. M. Eltahir, A. R. Alharbi, I. Issaoui, and A. Sayed, “Endoscopic Image Analysis for Gastrointestinal Tract Disease Diagnosis Using Nature Inspired Algorithm With Deep Learning Approach,” *Volume 11, Pages 130022 - 130030*, vol. 11, pp. 130022–130030, 2023, doi: 10.1109/ACCESS.2023.3333752.
- [12] C. H. Lin *et al.*, “Application of artificial intelligence in endoscopic image analysis for the diagnosis of a gastric cancer pathogen-Helicobacter pylori infection,” *Volume 13, Issue 1*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-40179-5.
- [13] P. Kamola and T. Przygodzki, “Qualitative classification of thrombus images as a way to improve quantitative analysis of thrombus formation in flow chamber assays,” *Volume 19, Issue 3 March*, vol. 19, no. 3 March, Mar. 2024, doi:

- 10.1371/journal.pone.0299202.
- [14] S. Khan, M. Sajjad, T. Hussain, A. Ullah, and A. S. Imran, "A review on traditional machine learning and deep learning models for WBCs classification in blood smear images," *Volume 9, Pages 10657 - 10673*, vol. 9, pp. 10657–10673, 2021, doi: 10.1109/ACCESS.2020.3048172.
- [15] O. Shahid *et al.*, "Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance," *Volume 117*, vol. 117, May 2021, doi: 10.1016/j.jbi.2021.103751.
- [16] S. D'Ambrosi *et al.*, "Combinatorial Blood Platelets-Derived circRNA and mRNA Signature for Early-Stage Lung Cancer Detection," *Volume 24, Issue 5*, vol. 24, no. 5, Mar. 2023, doi: 10.3390/ijms24054881.
- [17] M. A. Majeed, H. Z. M. Shafri, Z. Zulkafli, and A. Wayayok, "A Deep Learning Approach for Dengue Fever Prediction in Malaysia Using LSTM with Spatial Attention," *Volume 20, Issue 5*, vol. 20, no. 5, Jan. 2023, doi: 10.3390/ijerph20054130.
- [18] J. K. Chaw *et al.*, "A predictive analytics model using machine learning algorithms to estimate the risk of shock development among dengue patients," *Volume 5*, vol. 5, Jun. 2024, doi: 10.1016/j.health.2023.100290.
- [19] M. Botlagunta *et al.*, "Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms," *Volume 13, Issue 1*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-27548-w.
- [20] C. T. Hsu, K. C. Pai, L. C. Chen, S. H. Lin, and M. J. Wu, "Machine Learning Models to Predict the Risk of Rapidly Progressive Kidney Disease and the Need for Nephrology Referral in Adult Patients with Type 2 Diabetes," *Volume 20, Issue 4*, vol. 20, no. 4, Jan. 2023, doi: 10.3390/ijerph20043396.
- [21] A. A. Huang and S. Y. Huang, "Use of machine learning to identify risk factors for coronary artery disease," *Volume 18, Issue 4 APRIL*, vol. 18, no. 4 APRIL, Apr. 2023, doi: 10.1371/journal.pone.0284103.
- [22] F. D. Koseoglu, F. Keklik Karadag, H. Bulbul, E. U. Alici, B. Ozyilmaz, and T. R.

- Ozdemir, "JAKCalc: A machine-learning approach to rationalized JAK2 testing in patients with elevated hemoglobin levels," *Volume 103, Issue 14, Pages E37751*, vol. 103, no. 14, p. E37751, Apr. 2024, doi: 10.1097/MD.00000000000037751.
- [23] O. Gecgel, A. Ramanujam, and G. G. Botte, "Selective Electrochemical Detection of SARS-CoV-2 Using Deep Learning," *Volume 14, Issue 9*, vol. 14, no. 9, Sep. 2022, doi: 10.3390/v14091930.
- [24] M. Hasan, S. B. Murtaz, M. U. Islam, M. J. Sadeq, and J. Uddin, "Robust and efficient COVID-19 detection techniques: A machine learning approach," *Volume 17, Issue 9 September*, vol. 17, no. 9 September, Jun. 2022, doi: 10.1371/journal.pone.0274538.
- [25] A. M. Elsherbini *et al.*, "Decoding Diabetes Biomarkers and Related Molecular Mechanisms by Using Machine Learning, Text Mining, and Gene Expression Analysis," *Volume 19, Issue 21*, vol. 19, no. 21, Nov. 2022, doi: 10.3390/ijerph192113890.
- [26] A. Carobene, F. Milella, L. Famigliani, and F. Cabitza, "How is test laboratory data used and characterised by machine learning models? A systematic review of diagnostic and prognostic models developed for COVID-19 patients using only laboratory data," *Volume 60, Issue 12, Pages 1887 - 1901*, vol. 60, no. 12, pp. 1887–1901, Nov. 2022, doi: 10.1515/cclm-2022-0182.
- [27] R. Ganguli, J. Franklin, X. Yu, A. Lin, and D. S. Heffernan, "Machine learning methods to predict presence of residual cancer following hysterectomy," *Volume 12, Issue 1*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-06585-x.
- [28] O. Y. Dweekat, S. S. Lam, and L. McGrath, "An Integrated System of Multifaceted Machine Learning Models to Predict If and When Hospital-Acquired Pressure Injuries (Bedsore) Occur," *Volume 20, Issue 1*, vol. 20, no. 1, Jan. 2023, doi: 10.3390/ijerph20010828.
- [29] S. R. D. S. Neto *et al.*, "Machine learning and deep learning techniques to support clinical diagnosis of arboviral diseases: A systematic review," *Volume 16, Issue 1*, vol. 16, no. 1, Jan. 2022, doi: 10.1371/JOURNAL.PNTD.0010061.

- [30] G. Cardozo, G. B. Pintarelli, G. R. Andreis, A. C. W. Lopes, and J. L. B. Marques, "Use of Machine Learning and Routine Laboratory Tests for Diabetes Mellitus Screening," *Volume 2022*, vol. 2022, 2022, doi: 10.1155/2022/8114049.
- [31] D. R. Giacobbe *et al.*, "Early Detection of Sepsis With Machine Learning Techniques: A Brief Clinical Perspective," *Volume 8*, vol. 8, Feb. 2021, doi: 10.3389/fmed.2021.617486.
- [32] L. P. Liu *et al.*, "Machine Learning for the Prediction of Red Blood Cell Transfusion in Patients During or After Liver Transplantation Surgery," *Volume 8*, vol. 8, Feb. 2021, doi: 10.3389/fmed.2021.632210.
- [33] C. F. Caiafa, Z. Sun, T. Tanaka, P. Marti-Puig, and J. Solé-Casals, "Machine learning methods with noisy, incomplete or small datasets," *Volume 11, Issue 9*, vol. 11, no. 9, May 2021, doi: 10.3390/app11094132.
- [34] G. Li, B. Iyer, V. B. S. Prasath, Y. Ni, and N. Salomonis, "DeepImmuno: Deep learning-empowered prediction and generation of immunogenic peptides for T-cell immunity," *Volume 22, Issue 6*, vol. 22, no. 6, Nov. 2021, doi: 10.1093/bib/bbab160.
- [35] A. Ikonnikova *et al.*, "Genetic Association Study and Machine Learning to Investigate Differences in Platelet Reactivity in Patients with Acute Ischemic Stroke Treated with Aspirin," *Volume 10, Issue 10*, vol. 10, no. 10, Oct. 2022, doi: 10.3390/biomedicines10102564.
- [36] X. Qiu, M. G. Nair, L. Jaroszewski, and A. Godzik, "Deciphering Abnormal Platelet Subpopulations in COVID-19, Sepsis and Systemic Lupus Erythematosus through Machine Learning and Single-Cell Transcriptomics," *Volume 25, Issue 11*, vol. 25, no. 11, Jun. 2024, doi: 10.3390/ijms25115941.
- [37] Y. Liu *et al.*, "Platelets-related signature based diagnostic model in rheumatoid arthritis using WGCNA and machine learning," *Volume 14*, vol. 14, 2023, doi: 10.3389/fimmu.2023.1204652.
- [38] S. Mishra, R. Kumar, S. K. Tiwari, and P. Ranjan, "Machine learning approaches in the diagnosis of infectious diseases: a review," *Volume 11, Issue 6, Pages 3509 - 3520*,

- vol. 11, no. 6, pp. 3509–3520, Dec. 2022, doi: 10.11591/eei.v11i6.4225.
- [39] D. Tang *et al.*, “Identification of platelet-related subtypes and diagnostic markers in pediatric Crohn’s disease based on WGCNA and machine learning,” *Volume 15*, vol. 15, 2024, doi: 10.3389/fimmu.2024.1323418.
- [40] Y. Geng *et al.*, “Identification and validation of platelet-related diagnostic markers and potential drug screening in ischemic stroke by integrating comprehensive bioinformatics analysis and machine learning,” *Volume 14*, vol. 14, 2023, doi: 10.3389/fimmu.2023.1320475.
- [41] Y. Diao, Y. Zhao, X. Li, B. Li, R. Huo, and X. Han, “A simplified machine learning model utilizing platelet-related genes for predicting poor prognosis in sepsis,” *Volume 14*, vol. 14, 2023, doi: 10.3389/fimmu.2023.1286203.
- [42] P. Saha, M. S. Sadi, O. F. M. R. R. Aranya, S. Jahan, and F. Al Islam, “COV-VGX: An automated COVID-19 detection system using X-ray images and transfer learning,” *Volume 26*, vol. 26, Jan. 2021, doi: 10.1016/j.imu.2021.100741.
- [43] X. Zhao, W. Shen, and G. Wang, “Early Prediction of Sepsis Based on Machine Learning Algorithm,” *Volume 2021*, vol. 2021, 2021, doi: 10.1155/2021/6522633.
- [44] C. C. Bartenschlager *et al.*, “COVIDAL: A Machine Learning Classifier for Digital COVID-19 Diagnosis in German Hospitals,” *ACM Trans Manag Inf Syst*, vol. 14, no. 2, Mar. 2023, doi: 10.1145/3567431.
- [45] E. Petersen *et al.*, “Responsible and Regulatory Conform Machine Learning for Medicine: A Survey of Challenges and Solutions,” *IEEE Access*, vol. 10, pp. 58375–58418, 2022, doi: 10.1109/ACCESS.2022.3178382.
- [46] L. Mendoza-Pitti, J. M. Gomez-Pulido, M. Vargas-Lombardo, J. A. Gomez-Pulido, M. L. Polo-Luque, and D. Rodriguez-Puyol, “Machine-Learning Model to Predict the Intradialytic Hypotension Based on Clinical-Analytical Data,” *IEEE Access*, vol. 10, pp. 72065–72079, 2022, doi: 10.1109/ACCESS.2022.3189018.
- [47] R. Ge, R. Zhang, and P. Wang, “Prediction of chronic diseases with multi-label neural network,” *IEEE Access*, vol. 8, pp. 138210–138216, 2020, doi:

- 10.1109/ACCESS.2020.3011374.
- [48] R. K. Lomotey, S. Kumi, M. Hilton, R. Orji, and R. Deters, "Using Machine Learning to Establish the Concerns of Persons With HIV/AIDS During the COVID-19 Pandemic From Their Tweets," *IEEE Access*, vol. 11, pp. 37570–37601, 2023, doi: 10.1109/ACCESS.2023.3267050.
- [49] W. Hoyos, J. Aguilar, and M. Toro, "Dengue models based on machine learning techniques: A systematic literature review," *Artif Intell Med*, vol. 119, p. 102157, Sep. 2021, doi: 10.1016/J.ARTMED.2021.102157.
- [50] M. S. G. Shaikh, D. B. SureshKumar, and D. G. Narang, "Development of optimized ensemble classifier for dengue fever prediction and recommendation system," *Biomed Signal Process Control*, vol. 85, p. 104809, Aug. 2023, doi: 10.1016/J.BSPC.2023.104809.
- [51] J. Cui, H. Zhu, H. Deng, Z. Chen, and D. Liu, "FeARH: Federated machine learning with anonymous random hybridization on electronic medical records," *J Biomed Inform*, vol. 117, p. 103735, May 2021, doi: 10.1016/J.JBI.2021.103735.
- [52] B. Hemalatha, B. Karthik, C. V. Krishna Reddy, and A. Latha, "Deep learning approach for segmentation and classification of blood cells using enhanced CNN," *Measurement: Sensors*, vol. 24, p. 100582, Dec. 2022, doi: 10.1016/J.MEASEN.2022.100582.
- [53] M. Kowsher *et al.*, "LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy," *Procedia Comput Sci*, vol. 193, pp. 131–140, Jan. 2021, doi: 10.1016/J.PROCS.2021.10.013.
- [54] M. A. Alves *et al.*, "Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs," *Comput Biol Med*, vol. 132, p. 104335, May 2021, doi: 10.1016/J.COMPBIOMED.2021.104335.
- [55] P. K. Vesapogu and B. R. Surampudi, "A Comparative Study of Machine Learning Approaches for the Detection of SARS-CoV-2 and its Variants," *Procedia Comput Sci*, vol. 235, pp. 1190–1201, Jan. 2024, doi: 10.1016/J.PROCS.2024.04.113.

- [56] F. Rui *et al.*, “Development of a machine learning-based model to predict hepatic inflammation in chronic hepatitis B patients with concurrent hepatic steatosis: a cohort study,” *EClinicalMedicine*, vol. 68, p. 102419, Feb. 2024, doi: 10.1016/J.ECLINM.2023.102419.
- [57] P. Zhang, Z. Wang, H. Qiu, W. Zhou, M. Wang, and G. Cheng, “Machine learning applied to serum and cerebrospinal fluid metabolomes revealed altered arginine metabolism in neonatal sepsis with meningoencephalitis,” *Comput Struct Biotechnol J*, vol. 19, pp. 3284–3292, Jan. 2021, doi: 10.1016/J.CSBJ.2021.05.024.
- [58] A. Hennebelle, H. Materwala, and L. Ismail, “HealthEdge: A Machine Learning-Based Smart Healthcare Framework for Prediction of Type 2 Diabetes in an Integrated IoT, Edge, and Cloud Computing System,” *Procedia Comput Sci*, vol. 220, pp. 331–338, Jan. 2023, doi: 10.1016/J.PROCS.2023.03.043.
- [59] H. Abdulqadir Khidir, İ. Etikan, D. Hussein Kadir, N. H. Mahmood, and R. Sabetvand, “Bayesian machine learning analysis with Markov Chain Monte Carlo techniques for assessing characteristics and risk factors of Covid-19 in Erbil City-Iraq 2020–2021,” *Alexandria Engineering Journal*, vol. 78, pp. 162–174, Sep. 2023, doi: 10.1016/J.AEJ.2023.07.052.
- [60] Y. Zhou *et al.*, “Machine learning predictive models for acute pancreatitis: A systematic review,” *Int J Med Inform*, vol. 157, p. 104641, Jan. 2022, doi: 10.1016/J.IJMEDINF.2021.104641.
- [61] S. Karolcik *et al.*, “Towards a machine-learning assisted non-invasive classification of dengue severity using wearable PPG data: a prospective clinical study,” *EBioMedicine*, vol. 104, p. 105164, Jun. 2024, doi: 10.1016/J.EBIOM.2024.105164.
- [62] Q. Wu *et al.*, “A customised down-sampling machine learning approach for sepsis prediction,” *Int J Med Inform*, vol. 184, p. 105365, Apr. 2024, doi: 10.1016/J.IJMEDINF.2024.105365.
- [63] S. Jahandideh, G. Ozavci, B. W. Sahle, A. Z. Kouzani, F. Magrabi, and T. Bucknall, “Evaluation of machine learning-based models for prediction of clinical deterioration: A

- systematic literature review,” *Int J Med Inform*, vol. 175, p. 105084, Jul. 2023, doi: 10.1016/J.IJMEDINF.2023.105084.
- [64] P. Silitonga, B. E. Dewi, A. Bustamam, and H. S. Al-Ash, “Evaluation of Dengue Model Performances Developed Using Artificial Neural Network and Random Forest Classifiers,” *Procedia Comput Sci*, vol. 179, pp. 135–143, Jan. 2021, doi: 10.1016/J.PROCS.2020.12.018.
- [65] B. Robson and O. K. Baek, “An ontology for very large numbers of longitudinal health records to facilitate data mining and machine learning,” *Inform Med Unlocked*, vol. 38, p. 101204, Jan. 2023, doi: 10.1016/J.IMU.2023.101204.
- [66] H. R. Jackson *et al.*, “A multi-platform approach to identify a blood-based host protein signature for distinguishing between bacterial and viral infections in febrile children (PERFORM): a multi-cohort machine learning study,” *Lancet Digit Health*, vol. 5, no. 11, pp. e774–e785, Nov. 2023, doi: 10.1016/S2589-7500(23)00149-8.
- [67] J. Xiang *et al.*, “Machine learning–based scoring models to predict hematopoietic stem cell mobilization in allogeneic donors,” *Blood Adv*, vol. 6, no. 7, pp. 1991–2000, Apr. 2022, doi: 10.1182/BLOODADVANCES.2021005149.
- [68] J. Demagny *et al.*, “Combining imaging flow cytometry and machine learning for high-throughput schistocyte quantification: A SVM classifier development and external validation cohort,” *EBioMedicine*, vol. 83, p. 104209, Sep. 2022, doi: 10.1016/J.EBIOM.2022.104209.
- [69] H. Salehinejad, A. M. Meehan, P. A. Rahman, M. A. Core, B. J. Borah, and P. J. Caraballo, “Novel machine learning model to improve performance of an early warning system in hospitalized patients: a retrospective multisite cross-validation study,” *EClinicalMedicine*, vol. 66, p. 102312, Dec. 2023, doi: 10.1016/J.ECLINM.2023.102312.
- [70] I. Shiri *et al.*, “Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients,” *Comput Biol Med*, vol. 132, p. 104304, May 2021, doi: 10.1016/J.COMPBIOMED.2021.104304.

## ANEXOS

### Anexo 1: Reporte de similitud

# Laynes Castillo Jose Carlos Espinoza Purisaca Josu...

## Técnicas de Machine Learning para el recuento de plaquetas como ayuda al diagnóstico de enfermedades

 Universidad Señor de Sipan

#### Detalles del documento

Identificador de la entrega

trn:oid:::26396:428260981

Fecha de entrega

9 feb 2025, 3:19 p.m. GMT-5

Fecha de descarga

9 feb 2025, 3:22 p.m. GMT-5

Nombre de archivo

turnitin TRABAJO DE INVESTIGACIÓN - Bachiller.docx

Tamaño de archivo

1.4 MB

42 Páginas

11,498 Palabras

64,973 Caracteres



Página 2 of 53 - Descripción general de integridad

Identificador de la entrega trn:oid:::26396:428260981

## 20% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

#### Filtrado desde el informe

- ▶ Bibliografía
- ▶ Texto mencionado
- ▶ Coincidencias menores (menos de 8 palabras)

#### Fuentes principales

14%  Fuentes de Internet

3%  Publicaciones

13%  Trabajos entregados (trabajos del estudiante)

#### Marcas de integridad

##### N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

## Anexo 2: Matriz de Lectura de Artículos

	A	B	C	D	E	F	G	H	I	J	K
1	<b>MATRIZ DE LECTURA DE ARTICULOS</b>										
2	TITULO DE LA REVISION: Técnicas de Machine Learning para el recuento de plaquetas como ayuda al diagnóstico de enfermedades: Una revisión Sistemática										
3	AUTORES: Espinoza Purisaca Josué Enoç / Lames Castillo José Carlos										
4											
5	#	Título del Artículo	Link	Año	Autores	País	Canal de publicación	Tipo de Publicación	Origen de la publicación	Tipo de artículo	Origen de Búsqueda
6	1	Platelet Detection Based on Improved YOLO_v3	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85147194588&amp;doi=10.34133%2F2022%2F8780569&amp;partnerID=40&amp;md5=11b668ac55c0c8110b9ab9996957b5f5">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85147194588&amp;doi=10.34133%2F2022%2F8780569&amp;partnerID=40&amp;md5=11b668ac55c0c8110b9ab9996957b5f5</a>	2022	Liu, Renting and Ren, Chunhui and Fu, Miaomiao and Chu, Zhengkang and Guo, Jiuchuan	China	Cyborg and Bionic Systems	Journal	Academia	Tecnico	Scopus
7	2	Machine learning insights into thrombo- ischemic risks and bleeding events through platelet lysophospholipids and scylcamitine species	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85187737281&amp;doi=10.1038%2F41588-024-56304-x&amp;partnerID=40&amp;md5=0a63eada483a01110a3437b41d8c753">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85187737281&amp;doi=10.1038%2F41588-024-56304-x&amp;partnerID=40&amp;md5=0a63eada483a01110a3437b41d8c753</a>	2024	Harm, Tobias and Fu, Xiaqing and Frey, Moritz and Dittich, Kristina and Brun, Adrian and Castor, Tatsiana and Borst, Oliver and Müller, Karin Anne Lydia and Geisler, Tobias and Rath, Dominik and Lämmerhofer, Michael and Gawaz, Meinrad Paul	Alemania	Scientific Reports	Journal	Academia	Tecnico	Scopus
8	3	A systematic review of dengue classification models using machine learning: Una revisión sistemática de Modelos de clasificación de dengue utilizando machine learning.	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174169329&amp;doi=10.17013%2Frisa.50.5.27&amp;partnerID=40&amp;md5=5eabf468fe0e8c66e1271bd6a3b5377e">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174169329&amp;doi=10.17013%2Frisa.50.5.27&amp;partnerID=40&amp;md5=5eabf468fe0e8c66e1271bd6a3b5377e</a>	2023	Maquen-Nifo, Gisella Luisa Elena and Bravo, Jessie and Alarcón, Roger and Adriansén-Olano, Ivan and Vega-Huerta, Hugo	Perú	RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao	Journal	Academia	Reporte	Scopus
9	4	Machine Learning-Based Detection of Dengue from Blood Smear Images Utilizing Platelet and Lymphocyte Characteristics	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85146773755&amp;doi=10.3390%2Fdiagnostics13020220&amp;partnerID=40&amp;md5=9b09d44c02673e3b956442e88a5">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85146773755&amp;doi=10.3390%2Fdiagnostics13020220&amp;partnerID=40&amp;md5=9b09d44c02673e3b956442e88a5</a>	2023	Mayrose, Hilda and Baiyy, G. Muralidhar and Sampathila, Niranjana and Behariker, Sushma and Saravu, Kavitha	India	Diagnostics	Journal	Academia	Técnico-Aplicación	Scopus
10	5	Deep learning classification of lipid droplets in quantitative phase images	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103837672&amp;doi=10.1371%2Fournal.pone.0249196&amp;partnerID=40&amp;md5=e826ba9092b0e3540d581f01646">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103837672&amp;doi=10.1371%2Fournal.pone.0249196&amp;partnerID=40&amp;md5=e826ba9092b0e3540d581f01646</a>	2021	Sheneman, Luke and Stephanopoulos, Gregory and Vasdekis, Andreas E.	Estados Unidos	PLoS ONE	Journal	Academia	Técnico-Aplicación	Scopus
11	6	Machine learning refutes loss of smell as a risk indicator of diabetes mellitus	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117912226&amp;doi=10.3390%2Fncm102149718&amp;partnerID=40&amp;md5=593077ac248d6f93e4f4455df1545e">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117912226&amp;doi=10.3390%2Fncm102149718&amp;partnerID=40&amp;md5=593077ac248d6f93e4f4455df1545e</a>	2021	Loewer, Johannes Hähner, Antje and Schwarz, Peter E. H. and Tselmin, Sergey and Hummel, Patrick	Alemania	Journal of Clinical Medicine	Journal	Academia	Técnico-Aplicación	Scopus
12	7	Detecting and Extracting Brain Hemorrhages from CT Images Using Generative Convolutional	<a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130021738&amp;doi=10.1155%2F2022%2F26671234&amp;partnerID=40&amp;md5=4">https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130021738&amp;doi=10.1155%2F2022%2F26671234&amp;partnerID=40&amp;md5=4</a>	2022	Pandey, Jyoti and Rajasoundaran, S. and Routray, Sideswar and Prabu, A.V. and Alkhami Hashem and	India	Computational Intelligence and Neuroscience	Journal	Academia	Técnico-Aplicación	Scopus