

UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN



**Sistema inteligente basado en machine learning para apoyar la
optimización de la planificación logística en la empresa Héctor Ocaña
Izquierdo**

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

AUTOR

Christian Paul Bellasmil Farroñan

ASESOR

Guadalupe Teresa Lip Curo

<https://orcid.org/0000-0002-0353-939X>

Chiclayo, 2026

**Sistema inteligente basado en machine learning para apoyar la
optimización de la planificación logística en la empresa Héctor
Ocaña Izquierdo**

PRESENTADA POR

Christian Paul Bellasmil Farroñan

A la Facultad de Ingeniería de la
Universidad Católica Santo Toribio de Mogrovejo
para optar el título de

INGENIERO DE SISTEMAS Y COMPUTACIÓN

APROBADA POR

Maria Ysabel Aranguri Garcia
PRESIDENTE

Huiler Juanito Mera Montenegro
SECRETARIO

Guadalupe Teresa Lip Curo
VOCAL

Dedicatoria

A Dios, por bendecirme cada día y darme la fortaleza para seguir adelante incluso en los momentos más difíciles.

A mis padres, por su dedicación, esfuerzo y amor incondicional, que fueron el pilar fundamental en todo este camino.

Y a todas las personas que me acompañaron durante este proceso, por sus palabras de aliento y su compañía en cada etapa de mi formación.

Agradecimientos

A la Universidad Católica Santo Toribio de Mogrovejo, por brindarme la oportunidad de formarme profesionalmente y por fomentar en sus estudiantes el compromiso con la excelencia y la investigación.

A mi asesora, por su guía constante, paciencia y orientación profesional, que fueron esenciales para el desarrollo y culminación de este trabajo.

A los docentes de la institución, por compartir sus conocimientos y contribuir al fortalecimiento de mis capacidades académicas y personales.

Sistema inteligente basado en machine learning para apoyar la optimización de la planificación logística en la empresa Héctor Ocaña Izquierdo

INFORME DE ORIGINALIDAD

9%	9%	0%	4%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	tesis.usat.edu.pe Fuente de Internet	2%
2	hdl.handle.net Fuente de Internet	1%
3	Submitted to Universidad Católica Santo Toribio de Mogrovejo Trabajo del estudiante	1%
4	repositorioacademico.upc.edu.pe Fuente de Internet	1%
5	www.coursehero.com Fuente de Internet	1%
6	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	<1%
7	Submitted to UNIBA Trabajo del estudiante	<1%
8	repositorio.uniremington.edu.co Fuente de Internet	<1%
9	Submitted to ESADE Trabajo del estudiante	<1%
10	renati.sunedu.gob.pe Fuente de Internet	<1%
11	repositorio.uccuyo.edu.ar Fuente de Internet	<1%

Índice

Resumen	6
Abstract	7
Introducción	8
Revisión de literatura	11
Materiales y métodos	16
Resultados y discusión	19
Conclusiones	36
Recomendaciones.....	37
Referencias	38
Anexos.....	42

Resumen

Este proyecto tuvo como objetivo desarrollar un sistema inteligente basado en machine learning para optimizar la planificación logística en la empresa Héctor Ocaña Izquierdo. Para ello, se construyeron dos modelos principales: el primero, centrado en la predicción de demanda, utilizando redes neuronales LSTM, alcanzó una precisión del 93.83%. Este modelo permitió anticipar la demanda futura de manera precisa, facilitando la planificación de recursos. El segundo modelo se enfocó en la optimización de rutas logísticas, logrando una precisión del 99.08%. La combinación de ambos modelos proporcionó una interoperabilidad total, lo que mejoró significativamente la asignación de recursos y redujo los costos operativos, permitiendo una optimización de los tiempos de entrega y mejorando la eficiencia general del sistema. Los resultados obtenidos demostraron una mejora sustancial del 24.5% en tiempos de conducción, mostrando también una reducción del 34.47% en el recorrido. El sistema fue validado mediante pruebas de desempeño, alcanzando una fiabilidad del 99.44% en condiciones de operación continua, y una tasa de defectos por ejecución del 0.08 lo que confirmó su estabilidad y robustez para su implementación en entornos logísticos reales. Además, se desarrolló una plataforma web para facilitar la interacción en tiempo real y la gestión de las operaciones logísticas, alcanzando un 93.68% de satisfacción por parte de los usuarios, según las encuestas de satisfacción realizadas al final del proceso.

Palabras clave: Planificación logística, Machine learning, Predicción de demanda, Planificación de rutas.

Abstract

This project aimed to develop an intelligent, machine learning-based system to optimize logistics planning at Héctor Ocaña Izquierdo. To achieve this, two main models were built: the first, focused on demand forecasting using LSTM neural networks, achieved 93.83% accuracy. This model allowed for the precise anticipation of future demand, facilitating resource planning. The second model focused on optimizing logistics routes, achieving 99.08% accuracy. The combination of both models provided full interoperability, significantly improving resource allocation and reducing operating costs, enabling optimized delivery times and improving overall system efficiency. The results showed a substantial 24.5% improvement in driving times, as well as a 34.47% reduction in route length. The system was validated through performance testing, achieving 99.44% reliability under continuous operating conditions and a defect rate of 0.08 per-run, confirming its stability and robustness for implementation in real-world logistics environments. Furthermore, a web platform was developed to facilitate real-time interaction and the management of logistics operations, achieving a 93.68% user satisfaction rate, according to satisfaction surveys conducted at the end of the process.

Keywords: Logistics planning, Machine learning, Demand prediction, Route planning

Introducción

La planificación logística es un plan acordado por todas las partes involucradas, que abarca desde la organización de recursos y la coordinación de procesos hasta la entrega final del producto al cliente [1]. Para lograr una planificación logística eficaz es esencial integrar tecnologías avanzadas que permitan anticipar la demanda, optimizar las rutas de distribución y gestionar los inventarios de manera más eficiente [2]. Así pues, los sistemas de gestión del transporte son importantes para mejorar la planificación, visibilidad y satisfacción del cliente, lo que a su vez impulsa las ventas y crecimiento de la empresa [3].

El sector logístico crece un 4,4% anual. En el primer trimestre de 2023, el comercio electrónico alcanzó 19.000 millones de euros, un 23% más que en 2022 [4]. En América del Sur, el mercado de carga y logística alcanzó los 180,78 mil millones de dólares en 2024 y se proyecta que llegará a 223,02 mil millones en 2029 [5]. En Perú, el 70,2% de las empresas ofrecen servicios de transporte, seguidas por la logística integral (44,3%) y la gestión de carga internacional (34,2%). El índice promedio de madurez de la cadena de suministros es 2.90%, lo que indica un nivel "en desarrollo" [6][7].

Según un estudio realizado por EY, solo el 34% de las empresas peruanas tiene un modelo de gestión de la cadena de suministro "avanzado", mientras que el 66% está en fases "incipientes". Además, en cuanto a herramientas tecnológicas el 56% cuenta con un sistema de planificación de transporte y el 44% usa herramientas de monitoreo para las entregas [8].

La logística del transporte enfrenta desafíos específicos, como la fluctuación de los precios del combustible, la congestión en las rutas de transporte, las regulaciones medioambientales y la necesidad de incorporar tecnologías innovadoras [9]. Las fallas en la planificación logística se deben a una mala coordinación, gestión deficiente de inventarios y falta de herramientas tecnológicas adecuadas [10]. Lo cual puede llevar al bajo rendimiento y pérdida de cuota de mercado, por lo que es crucial identificar y mitigar riesgos en los subprocesos logísticos [11].

Un estudio de PwC señala que, para 2030, la digitalización y automatización reducirán los costos de transporte en un 47%, los plazos de entrega en un 40% y las entregas automatizadas de última milla en un 51% [12].

En la actualidad, varios algoritmos de machine learning abordan problemas similares a los mencionados. Ali et al. [13], demostraron que los modelos simplificados de aprendizaje automático pueden reducir los costos computacionales entre un 30% y 98% en la predicción de demanda. De manera similar, Ren et al. [14], aplicaron redes neuronales en redes ferroviarias urbanas, logrando un R^2 de 0.89 y una tasa de error baja de 6.77%. De igual manera, Roque [15]

aplicó el modelo SARIMA en el Metropolitano de Lima, logrando una alta precisión con un RMSPE entre 6.37% y 8.13%, un MAPE entre 4.19% y 5.93%, y un R^2 entre 0.91 y 0.98. Por su parte, Martínez [16] implementó regresión lineal en Chiclayo, obteniendo una precisión del 87.64% y cumpliendo con un 90% de los criterios de usabilidad ISO 25010. En cuanto a planificación de rutas, Soto et al. [17], lograron una mejora del 98% en la eficiencia logística de Termoenergías del Perú S.A., reduciendo hasta 83.6 km por ruta. Igualmente, Lee et al. [18], optimizaron la planificación de rutas marítimas utilizando un modelo Transformer, logrando una reducción del consumo de combustible de hasta un 11.33%. Asimismo, Villanueva [19] optimizó las rutas de distribución en Lambayeque, reduciendo las distancias recorridas un 44% en la flota liviana y un 63% en la flota pesada, lo que resultó en una disminución del 38.75% en los costos operativos. Por otra parte, Kmiecik [20] implementó control predictivo, reduciendo los tiempos de conducción en un 15%. Finalmente, Bazan et al. [21], aplicaron DMAIC en una panadería en Lima, utilizando ARIMA y el algoritmo Clark-Wright, logrando una reducción del 30% en distancias y tiempos, y un ahorro del 2.58% en combustible.

La empresa de transporte en estudio enfrenta un problema recurrente de ineficiencia logística, lo que impacta directamente en la calidad de su servicio y en sus costos operativos. De cada 100 entregas, un 10% sufren retrasos mensuales debido a una deficiente planificación de las rutas y a la falta de predicción precisa de la demanda. Esta situación genera sobrecarga en la flota de vehículos durante picos de demanda no anticipados y, en otros momentos, provoca un uso subóptimo de los vehículos, resultando en tiempos muertos que oscilan entre 1 hora o hasta 1 día, dependiendo de la gravedad de la situación.

Actualmente, las decisiones logísticas se toman basadas en la experiencia previa y reportes manuales, lo que limita significativamente la capacidad de adaptarse de manera dinámica a los cambios continuos en la demanda de servicios. Esta falta de flexibilidad genera ineficiencias en la asignación de recursos y planificación logística inadecuada.

Además, la falta de una estrategia de reparto ha llevado a un aumento del 10-15% en los costos de combustible y mantenimiento, mientras que el tiempo de inactividad de los vehículos reduce la capacidad operativa en un 8-10%. Estos factores afectan tanto la rentabilidad como la satisfacción del cliente, que ha reportado un aumento del 5% en quejas por entregas tardías en los últimos seis meses. (ver anexo 01). Por lo tanto, surge la pregunta: ¿Qué características debe tener un sistema inteligente basado en machine learning, para apoyar en la optimización de la planificación logística en la empresa de transporte "Héctor Ocaña Izquierdo"?

La planificación logística es un pilar esencial en la eficiencia operativa de cualquier empresa de transporte, ya que influye directamente en la satisfacción del cliente y en los costos operativos. En el caso particular de la empresa en estudio, la implementación de soluciones tecnológicas avanzadas, como un sistema inteligente basado en machine learning, es crucial para mitigar problemas como la planificación ineficiente de rutas y la falta de predicción precisa de la demanda. Estas tecnologías permiten optimizar las rutas de transporte y anticipar la demanda de servicios, mejorando la eficiencia operativa. Según Quijije [22], la tecnología facilita la organización, distribución y almacenamiento en la logística, además de mejorar la conexión entre proveedores, clientes y consumidores.

El valor económico de esta investigación radica en el potencial para reducir los costos operativos en las pymes de transporte como es el caso de la empresa en estudio optimizando rutas y reduciendo costos de combustible, lo cual mejorará la eficiencia operativa y reducirá los costos asociados al transporte. Los primeros usuarios que han adoptado la gestión de la cadena de suministro impulsada por inteligencia artificial han logrado reducir los costos logísticos en un 15%, mejorado los niveles de inventario en un 35% y aumentado los niveles de servicio en un 65% [23][24].

El impacto social de esta investigación se refleja en la mejora de la calidad del servicio al cliente. Un sistema eficiente de planificación de rutas reducirá significativamente los retrasos, mejorando los tiempos de entrega y aumentando la satisfacción de los clientes [25]. La integración de sistemas automáticos permite a las empresas adaptar sus operaciones según factores como el tráfico y el clima, lo que reduce los costos variables y la huella de carbono [26]. Esto no solo beneficia a los clientes, sino que también contribuye al bienestar social al disminuir los efectos negativos del tráfico y el consumo de combustible en exceso.

Desde una perspectiva científica, esta investigación se enfoca en el uso de técnicas avanzadas de machine learning y minería de datos aplicadas a la logística. Estas tecnologías han demostrado ser efectivas en la predicción de demanda y planificación de rutas en diversas industrias. Cárdenes [27] y Cobo [28] destacan que estas técnicas permiten anticiparse a fluctuaciones en la demanda y asignar recursos de manera más eficiente, lo que genera un impacto positivo en la productividad. El desarrollo y aplicación de estos modelos en el contexto peruano no solo representa una contribución significativa a la ciencia de la logística, sino que también aporta al impulso de la cadena de suministros.

Como objetivo general de esta investigación se planteó desarrollar un sistema inteligente basado en machine learning para apoyar la optimización de la planificación logística en la empresa “Héctor Ocaña Izquierdo”, para el logro de este se definieron los siguientes objetivos específicos:

- Construir un modelo para la predicción de demanda en la empresa "Héctor Ocaña Izquierdo", que permita anticipar picos y valles en la demanda.
- Desarrollar un modelo para la planificación de rutas en la empresa "Héctor Ocaña Izquierdo" que permita optimizar la distribución de los recursos logísticos.
- Integrar el modelo de predicción de demanda con el de planificación de rutas para optimizar el servicio, mejorando así la satisfacción del cliente.
- Validar la conformidad del sistema inteligente con los estándares de calidad ISO 25010, con un enfoque particular en la adecuación funcional y fiabilidad del sistema.

Revisión de literatura

En estudios relacionados a la predicción de demanda, Ali et al. [13], abordan la complejidad de las cadenas de suministro globales, caracterizadas por estructuras de varios niveles, como un reto importante en la planificación de reposición de inventarios y en la predicción precisa de la demanda. Para resolverlo, el estudio compara la eficacia de los modelos tradicionales de análisis de regresión con los algoritmos de aprendizaje automático (ML), que tienen la capacidad de capturar relaciones más complejas. Se utilizaron conjuntos de datos de Kaggle y se evaluaron dos modelos: uno con 22 predictores y otro con los cinco predictores más relevantes. Los resultados revelaron que, al reducir la cantidad de predictores, los costos computacionales disminuyeron entre un 30% y un 98%, mientras que la precisión solo se redujo marginalmente entre un 0.6% y un 4.2%. Se concluyó que los modelos simplificados pueden optimizar la predicción de pedidos pendientes sin perder significativamente precisión, lo que representa una mejora para las industrias que buscan un equilibrio entre eficiencia computacional y precisión.

El estudio de Ren et al. [14], abordan la problemática de la predicción inexacta de la demanda en sistemas de transporte urbano, especialmente en redes ferroviarias urbanas, debido a la incapacidad de los modelos tradicionales para capturar las interacciones dinámicas entre los factores de origen y destino. Para resolver este desafío, proponen un modelo innovador basado en una red neuronal convolucional multimodal de dos ramas (OD-MDCNN), que

integra imágenes del uso del suelo y varios indicadores del entorno construido, mejorando significativamente la precisión de las predicciones. Los resultados muestran que el modelo supera a enfoques tradicionales, alcanzando un valor R^2 de 0.8935 para la demanda diaria y de 0.8739 para las horas pico, con errores de predicción relativamente bajos, lo que destaca su capacidad para capturar la complejidad del entorno y mejorar la planificación del transporte urbano.

En una línea similar a nivel regional, el estudio de Roque [15], aborda el problema de la saturación en el Metropolitano de Lima, destacando que la demanda diaria ha superado los 700,000 pasajeros, el doble de lo previsto en 2010. Para enfrentar este desafío, el estudio utiliza técnicas de data mining y el modelo SARIMA para predecir la demanda en clústeres de estaciones. Los resultados obtenidos muestran una alta precisión en las predicciones, con un RMSPE entre 6.37% y 8.13%, un MAPE entre 4.19% y 5.93%, y un R^2 entre 0.91 y 0.98. A pesar de los buenos resultados, se identificó una limitación en la predicción de días festivos, que fueron tratados como valores atípicos. Este enfoque ofrece una mejora significativa en la planificación de la demanda de transporte público.

En otro estudio, en el sector local, Martínez [16], aborda el problema de la gestión de inventarios en una empresa de Chiclayo, donde la predicción de la demanda se realizaba de forma empírica, lo que generaba sobrestock y retrasos en los pedidos. Para solucionarlo, se implementó una aplicación web basada en Machine Learning, utilizando el algoritmo de regresión lineal y datos históricos de ventas de tres años. La aplicación logró una precisión del 87.64%, validada mediante la métrica MAPE, y un 90% de cumplimiento en los criterios de usabilidad según la norma ISO 25010. Esta solución permitió optimizar la gestión de inventarios, mejorar la previsión de productos y aumentar la rentabilidad del negocio.

Por otro lado, en investigaciones sobre optimización de rutas logísticas, Soto et al. [17], abordan los problemas logísticos en la empresa Termoenergías del Perú S.A., que enfrenta altos costos operativos debido a la falta de rutas optimizadas para la distribución de sus productos en Lima Metropolitana. Se proponen los algoritmos heurísticos de Clarke y Wright y el algoritmo de Dijkstra, con el fin de diseñar rutas de distribución más eficientes y reducir los costos operativos. Se logró una optimización notable, con ahorros de hasta 83,6 km por ruta y una mejora en la eficiencia del 98% en la planificación de las rutas. Esto redujo significativamente el consumo de combustible y mejoró el uso de los vehículos, lo que impactó positivamente en la eficiencia operativa y la reducción de costos.

En una línea similar, Lee et al. [18], abordan el problema del elevado consumo de combustible y las altas emisiones de carbono en la navegación marítima, un desafío cada vez

más importante debido a que las rutas de navegación convencionales no consideran adecuadamente factores como las condiciones topográficas y ambientales, lo que resulta en rutas ineficientes y un mayor gasto de combustible. Para solucionar esto, proponen un enfoque innovador que combina un modelo de predicción de consumo de combustible basado en Transformer con el algoritmo Proximal Policy Optimization (PPO), optimizando dinámicamente los waypoints en las rutas marítimas según las condiciones específicas del entorno. Los resultados mostraron que este modelo alcanzó una precisión del 86.75%, superando a modelos como LSTM y Random Forest, y permitió reducir el consumo de combustible hasta un 11.33% en tres estudios de caso (Busan–Rotterdam, Busan–Los Ángeles, Mokpo–Houston), mejorando la eficiencia operativa y contribuyendo a la descarbonización del transporte marítimo.

A nivel local, Villanueva [19] aborda el caso de la empresa DD & D S.A.C., que distribuye helados en Lambayeque. El estudio destaca los altos costos operativos debido a la inadecuada planificación de rutas, con un 67.4% de los costos provenientes del consumo de combustible, la reprogramación de rutas y la mano de obra. Para reducir estos costos, se aplicaron técnicas de optimización como el modelo VRP (Vehicle Routing Problem) con la heurística de Clarke Wright y el algoritmo k-opt. Los resultados mostraron una reducción en las distancias recorridas de un 44% para la flota liviana y un 63% para la flota pesada, lo que llevó a una disminución del 38.75% en los costos operativos mensuales.

En estudios donde se intenta integrar ambos modelos, Kmiecik [16] aborda el desafío de la optimización dinámica de rutas logísticas en tiempo real, un problema común en la industria del transporte donde las condiciones de tráfico variables afectan la eficiencia operativa y aumentan los costos. Para resolverlo, propone un modelo de control predictivo que ajusta la velocidad de los vehículos según el tráfico en tiempo real, combinado con una ruta estática más corta. Este enfoque resultó en una mejora significativa, con una reducción del 15% en el tiempo de conducción y un aumento del 8% en la eficiencia de la selección de rutas, además de predecir con un 80% de precisión la demanda a nivel de SKU y destinatario, optimizando tanto las rutas como la gestión operativa y reduciendo costos y consumo de combustible.

Por otra parte, Bazan et al. [21], abordan la ineficiencia en la distribución de productos en una panadería de Lima, Perú, que enfrenta altos costos de transporte y problemas con la planificación de rutas y pronósticos de demanda. Para resolverlo, aplicaron la metodología DMAIC de Lean Six Sigma, implementando modelos de pronóstico de demanda ARIMA y Holt-Winters, con ARIMA mostrando mayor precisión, obteniendo un MAPE de 7.08% para el Producto 1 y 23.9% para el Producto 2. Para la optimización de rutas, utilizaron el algoritmo

Clark-Wright, lo que permitió reducir la distancia recorrida en un 30%, pasando de 87.81 km a 59 km, y disminuir el tiempo de entrega en un 30%, de 9 horas a 5 horas y 19 minutos. Además, se logró un ahorro de combustible del 2.58% respecto a otros estudios previos. Los vehículos optimizados cumplían con la meta de retorno antes del segundo turno de entrega, mejorando significativamente la eficiencia operativa, reduciendo costos y contribuyendo a una mayor competitividad y satisfacción del cliente en la empresa.

Bases teóricas

Logística

- **Empresa de Transporte:** Empresa encargada de gestionar el traslado de mercancías, supervisando tanto el almacenamiento como el inventario, y coordinando diferentes modos de transporte para asegurar que las entregas se realicen de manera puntual y rentable [29].
- **Planificación Logística:** Es un proceso estratégico que coordina el movimiento y almacenamiento de inventarios en la cadena de suministro, asegurando que los productos lleguen de manera eficiente desde el origen hasta el destino [30].
 - **Etapas de la gestión logística:** A continuación, se detallan los diferentes tipos de logística que existen [31]:
 - **Logística de aprovisionamiento:** Se encarga de gestionar las adquisiciones y el almacenamiento de los materiales necesarios, buscando siempre minimizar los costos.
 - **Logística de producción:** Involucra todos los recursos y medios necesarios para la fabricación de productos terminados.
 - **Logística de distribución:** Se centra en conectar al vendedor con el cliente final, asegurando la entrega eficiente de los productos terminados.
- **Predicción de la Demanda:** La predicción de demanda estima la demanda futura usando datos históricos y modelos matemáticos. Ayuda a reducir inventarios excesivos, mejorar el agotamiento de productos, evitar sobreproducción, reducir horas extras y mejorar el servicio al cliente [32].
- **Planificación de rutas:** Pieza clave dentro de la logística, ya que optimiza el flujo de trabajo y asegura entregas puntuales y eficientes. Una planificación de rutas adecuada minimiza los tiempos de tránsito, optimiza el uso de los vehículos, reduce los gastos operativos y mejora la experiencia del cliente [33].

Modelado de datos

- **Machine Learning:** Es una disciplina que crea algoritmos y modelos capaces de aprender de grandes cantidades de datos, identificando patrones sin necesidad de instrucciones específicas. Estos sistemas utilizan los datos históricos para mejorar su precisión y generar resultados más exactos a partir de nuevos datos [34].
 - **Redes Neuronales Artificiales (ANN):** Modelan relaciones no lineales entre las variables de entrada y salida, lo que las hace útiles para predecir demanda en entornos dinámicos. Son poderosas, pero requieren muchos datos y son difíciles de interpretar [35].
 - **LSTM:** Las redes LSTM son un tipo de RNN que procesan secuencias de datos recordando información previa, lo cual las hace efectivas para predecir valores futuros en series temporales mediante el aprendizaje de patrones en datos secuenciales [36].
 - **XGBoost:** Es una biblioteca de machine learning de código abierto que utiliza árboles de decisión mejorados por gradiente. Es conocida por su rapidez, eficiencia y capacidad para trabajar con grandes cantidades de datos [37].
- **Programación Lineal Entera Mixta:** Tipo de optimización matemática en la que se busca maximizar o minimizar una función lineal, sujeta a restricciones lineales. La diferencia con la programación lineal convencional es que en MILP, algunas de las variables son enteras o binarias, lo que permite tomar decisiones discretas, como asignar recursos o decidir entre opciones (sí/no) [38].
- **CRISP DM:** Método estándar para guiar proyectos de minería de datos, con un ciclo de seis fases interconectadas. Su secuencia es flexible, permitiendo avanzar o retroceder entre etapas según se requiera [39].

Herramientas

- **Google OR-Tools:** Software de código abierto para resolver problemas de optimización combinatoria, como enrutamiento de vehículos, programación de tareas y embalaje de objetos [40].
- **Sha256:** Es un "algoritmo de hash seguro de 256 bits" utilizado para la seguridad criptográfica. Los algoritmos de hash criptográficos generan valores únicos e irreversibles, y a mayor cantidad de posibles hashes, menor es la probabilidad de que dos valores generen el mismo hash [41].

- **API REST:** Interfaz de programación que sigue los principios del estilo arquitectónico de transferencia de estado representacional (REST), utilizado para conectar sistemas distribuidos de hipermedios [42].

Materiales y métodos

Metodología

1. Tipo de investigación

La presente investigación se sitúa dentro del enfoque de investigación aplicada, conforme a la definición propuesta por la OCDE en el Manual de Frascati [43]. Este tipo de investigación tiene como objetivo la adquisición de nuevos conocimientos, pero con una orientación práctica y específica. Esta investigación busca soluciones concretas a problemas definidos, lo que resulta fundamental en el contexto de la planificación logística, donde la implementación de tecnologías avanzadas, como machine learning, puede mejorar la predicción de demanda y la optimización de rutas, contribuyendo directamente a la eficiencia operativa de la empresa.

2. Métodos de investigación

Tabla I. Métodos de Investigación

Método	Sustento por el cual será empleado en la investigación
Hipotético-deductivo	Se empleó para formular hipótesis sobre la mejora de la planificación logística mediante machine learning, que serán evaluadas y verificadas a través de experimentos y simulaciones con datos reales de la empresa. Este método es clave para validar los modelos de predicción [44].
Analítico	Este método permitió descomponer el problema logístico en partes, analizando cada componente por separado (demanda, rutas, costos operativos) para identificar las áreas que necesitan optimización y las soluciones específicas que pueden ser implementadas [45].
Cualitativo y cuantitativo	Ambos enfoques se utilizaron para evaluar la efectividad del sistema. El análisis cuantitativo medirá los resultados en términos de reducción de costos y eficiencia, mientras que el cualitativo evaluará la percepción y satisfacción del cliente mediante encuestas [46].

Tabla II. Técnicas de Recolección de datos

Técnicas	Instrumentos	Propósito
Revisión documental	Ficha de análisis de documentos	Conocer el estado actual de la gestión logística y de rutas mediante el análisis de documentos internos.
Encuestas	Cuestionarios	Obtener percepciones y opiniones sobre el sistema de logística y su impacto en el desempeño de la empresa.

3. Metodología de desarrollo

El proyecto se desarrolló utilizando la metodología CRISP-DM que es adecuada para el desarrollo de sistemas basados en datos, como la predicción de demanda y la optimización de rutas mediante machine learning. CRISP-DM sigue un enfoque estructurado y flexible, lo que permitió una implementación eficiente y la continua optimización del sistema, a continuación se detalla la secuencia de pasos [39]. A continuación, se detallan las fases que se siguieron para la implementación del sistema:

3.1 Comprensión del negocio

Se definió el objetivo de planificación logística, identificando problemas en las rutas y la variabilidad de la demanda a través de reuniones con gerencia y administradores.

3.2 Comprensión de los datos

Se recolectó y analizó datos relevantes (cantidad de envíos, rutas, tráfico, clima), verificando su calidad, consistencia y fiabilidad, integrando fuentes externas como tráfico en tiempo real.

3.3 Preparación de los datos

Se limpiaron y en algunos casos se transformaron los datos, eliminando inconsistencias y valores nulos, y seleccionando las características clave para la predicción de demanda y optimización de rutas.

3.4 Modelado

Se desarrollaron modelos para predecir la demanda y optimizar las rutas, considerando restricciones operativas. Además, dichos modelos se evaluaron y ajustaron con datos reales para mejorar su rendimiento.

3.5 Evaluación

Se validaron los modelos comparando resultados con datos reales y simulando las rutas optimizadas. También se ajustaron los modelos según los resultados obtenidos para asegurar su efectividad.

3.6 Despliegue

Se integraron ambos modelos, haciendo uso de una plataforma de despliegue llamada fly, asegurando su usabilidad y satisfacción del cliente final.

4. Mapeo de objetivos, resultados y verificación

Objetivo 01: Construir un modelo para la predicción de demanda en la empresa "Héctor Ocaña Izquierdo", que permita anticipar picos y valles en la demanda		
Resultado	Medio de verificación	Indicador objetivamente verificable
Modelo construido para la predicción de demanda utilizando series temporales.	Simulaciones y pruebas piloto.	Los indicadores clave para verificar el rendimiento del modelo de predicción de demanda son [47]: <ul style="list-style-type: none"> - Error Absoluto Medio (MAE). - Error Cuadrático Medio (RMSE). - Tiempo de entrenamiento del modelo.
Objetivo 02: Desarrollar un modelo para la planificación de rutas en la empresa "Héctor Ocaña Izquierdo" que permita optimizar la distribución de los recursos logísticos.		
Resultado	Medio de verificación	Indicador objetivamente verificable
Modelo desarrollado para la planificación de rutas logísticas.	Simulaciones con datos históricos y pruebas piloto a corto plazo.	Los indicadores clave para verificar el rendimiento del modelo de planificación de rutas incluyen [48]: <ul style="list-style-type: none"> - Tiempo promedio de ejecución del modelo. - Precisión de la ruta. - F1-Score.
Objetivo 03: Integrar el modelo de predicción de demanda con el de planificación de rutas para optimizar el servicio, mejorando así la satisfacción del cliente.		
Resultado	Medio de verificación	Indicador objetivamente verificable
Modelos de predicción de demanda y planificación de rutas integrados.	Pruebas unitarias y de integración, simulaciones de interoperabilidad, encuestas de satisfacción	Los indicadores clave para evaluar la integridad de ambos modelos son [49]: <ul style="list-style-type: none"> - Tasa de éxito en la interoperabilidad. - Tiempo de respuesta conjunto. - Satisfacción del cliente.
Objetivo 04: Validar la conformidad del sistema inteligente con los estándares de calidad ISO 25010, con un enfoque particular en la adecuación funcional y fiabilidad del sistema.		
Resultado	Medio de verificación	Indicador objetivamente verificable
El sistema cumple con los requisitos funcionales y mantiene una fiabilidad adecuada según ISO 25010.	Pruebas de funcionalidad y fiabilidad del sistema mediante herramientas de prueba de software.	Los indicadores clave para validar la adecuación funcional y la fiabilidad del sistema son las siguientes [50]: <ul style="list-style-type: none"> - Tasa de cumplimiento de la funcionalidad. - Tasa de defectos por ejecución. - Tiempo promedio de recuperación ante fallos. - Disponibilidad del sistema.

Resultados y discusión

1. En base a los objetivos del proyecto

1.1 Construcción del modelo de predicción de demanda

1.1.1 Comprensión del negocio

a. Determinación de los objetivos comerciales

- Incrementar la eficiencia operativa de la empresa.
- Implementar un sistema inteligente que prediga la demanda, optimice la planificación de rutas y asignación de recursos.
- Reducir tiempos de entrega y costos operativos.
- Maximizar la utilización de la flota y la satisfacción del cliente.

b. Evaluación de la situación

- Retrasos en las entregas (con una tasa de error del 10% en los envíos).
- Alta variabilidad en la demanda y falta de anticipación.
- Planificación logística basada en experiencia y reportes manuales.
- Datos históricos disponibles, pero no aprovechados sistemáticamente.

c. Determinación de los objetivos de minería de datos

- Desarrollar un modelo de predicción de demanda futura.
- Implementar un algoritmo para asignar recursos logísticos de forma óptima.
- Integrar ambos modelos en un sistema funcional.

d. Producción de un plan de proyectos

- Seguir la metodología CRISP-DM para estructurar el desarrollo.
- Realizar recolección, análisis, modelado e integración de datos.
- Validar los modelos mediante simulaciones con datos reales.
- Capacitar al personal y monitorear el sistema mediante indicadores clave.

1.1.2 Comprensión de datos

a. Recopilación de datos iniciales

- Se recopiló información histórica de los viajes realizados por la empresa entre 2021 y 2024 en archivos Excel. Para facilitar el análisis, los datos fueron consolidados en una sola base de datos, inicialmente fueron 4052 registros los cuales se fueron actualizando con el avance del proyecto.

b. Descripción de los datos

- Se identificaron y describieron las variables clave contenidas en los registros. Esta información, compuesta por aproximadamente 4,000 datos, resulta esencial para el análisis logístico y el entrenamiento de los modelos de predicción y asignación de recursos.

c. Exploración de los datos

- Durante la exploración de los datos, se identificaron valores faltantes en varias columnas de las tablas de la base de datos.
- Estos datos ausentes pueden afectar el análisis y los modelos, por lo que se gestionaron adecuadamente, eliminando filas o columnas vacías o rellenándolas con valores adecuados.

d. Verificación de la calidad de los datos

- Se eliminan filas y columnas completamente vacías para reducir el tamaño del conjunto de datos y mejorar la eficiencia del análisis.
- Se eliminan columnas con nombres genéricos o vacíos que no aportan valor al análisis.

1.1.3 Preparación de datos

a. Selección de datos

- Surgió la necesidad de desglosar variables como la fecha en columnas necesarias para el entrenamiento, tal y como se describen en la tabla III.

Tabla III. Datos Utilizados para el modelo de predicción de demanda

Variable	Tipo	Descripción
conductor_id	Numérico	ID del conductor asignado al envío
cliente_id	Numérico	ID del cliente que solicitó el servicio
placa	Texto	Placa del vehículo asignado
fecha_registro	Fecha	Fecha y hora de registro del envío
precio_total	Numérico	Monto total cobrado por servicio
producto	Texto	Producto transportado
peso_partida	Numérico	Producto transportado
anio	Numérico	Año del envío
mes	Numérico	Mes de envío
día_semana	Numérico	Día de la semana del envío
fin_de_semana	Booleano	Indica si el envío fue en fin de semana
fecha_traslado	Fecha	Fecha efectiva de traslado
mes_sin	Numérico	Codificación cíclica del mes (seno)
mes_cos	Numérico	Codificación cíclica del mes (coseno)
n_dias	Numérico	Día del mes
is_high_season	Booleano	Indica si es temporada alta
demanda	Numérico	Demanda registrada para el envío

b. Limpieza de datos

- Se realizó una limpieza de datos eliminando filas y columnas vacías, estandarizando nombres de columnas y quitando columnas irrelevantes.

c. Construcción de nuevos datos

- Se crearon nuevas variables a partir de la fecha para capturar estacionalidad y temporalidad.

d. Formateo de datos

- Se aplicaron técnicas de codificación y normalización para preparar los datos, facilitando su uso en modelos de machine learning. El código de la figura 1 muestra este proceso.



```

1 # Codificación
2 encoder = LabelEncoder()
3 df['cliente_encoded'] = encoder.fit_transform(df['cliente'])
4
5 # Normalización
6 feature_cols = ['n_dias', 'mes_sin', 'mes_cos', 'anio', 'mes', 'dia_semana',
7               'fin_de_semana', 'is_high_season']
8 scaler_features = MinMaxScaler()
9 df[feature_cols] = scaler_features.fit_transform(df[feature_cols])
10
11 df['demanda'] = df['demanda'].astype(float)
12 df['log_demanda'] = np.log1p(df['demanda'])
13
14 scaler_cantidad = MinMaxScaler()
15 df['demanda_scaled'] = scaler_cantidad.fit_transform(df[['log_demanda']])

```

Fig. 1. Codificación y Normalización de Características en el Dataset

1.1.4 Modelado

a. Selección de técnicas de modelado

- Se seleccionó el modelo LSTM (Long Short-Term Memory) con capas bidireccionales y regularización, ideal para datos secuenciales como series temporales.

b. Generación de un diseño de comprobación

- Se generaron secuencias temporales por cliente para entrenar el modelo LSTM, organizando los datos en ventanas de tiempo que permiten predecir la demanda futura.

c. Generación de los modelos

- Se construyó un modelo LSTM bidireccional para predecir la demanda, compuesto por capas recurrentes, densas y de regularización (Dropout). El modelo fue entrenado con el optimizador Adam y la función de pérdida MSE, utilizó Early Stopping para evitar sobreajuste. Finalmente, se

guardó el modelo y los objetos de preprocesamiento para su posterior uso, tal y como se muestra en la figura 2.



```

1 # Modelo
2 model = Sequential([
3     Bidirectional(LSTM(128, return_sequences=True), input_shape=(X_train.shape[1], X_train.shape[2])),
4     Dropout(0.3),
5     Bidirectional(LSTM(64)),
6     Dropout(0.2),
7     Dense(64, activation='relu'),
8     Dense(32, activation='relu'),
9     Dense(1)
10 ])
11 model.compile(optimizer='adam', loss='mse')
12 callback = EarlyStopping(monitor='loss', patience=25, restore_best_weights=True)
13 model.fit(X_train, y_train, epochs=300, callbacks=[callback], verbose=0)
14
15 # Guardar modelo
16 model.save(MODEL_PATH)
17 with open(ENCODER_PATH, 'wb') as f: pickle.dump(encoder, f)
18 with open(SCALER_FEATURES_PATH, 'wb') as f: pickle.dump(scaler_features, f)
19 with open(SCALER_CANTIDAD_PATH, 'wb') as f: pickle.dump(scaler_cantidad, f)

```

Fig. 2. Implementación de Modelo Bidireccional LSTM y Guardado de Parámetros

1.1.5 Evaluación

a. Evaluación de los resultados

- El modelo fue entrenado con éxito, alcanzando métricas de rendimiento satisfactorias como un MAE de 0.12 y una precisión del 93%. Además, se integró en una interfaz web para facilitar su uso por parte del usuario final.

b. Proceso de Revisión

- MAE = 0.12: Error promedio bajo, indica buena precisión en la predicción.
- SMAPE = 6.17%: Error porcentual muy bajo, considerado excelente en series temporales.
- Precisión = 93.83%: Alta capacidad predictiva del modelo.
- Conclusión: El modelo presenta un desempeño sólido y es apto para producción.

1.1.6 Despliegue

a. Planificación de despliegue

- El modelo fue desplegado mediante una aplicación web desarrollada con Flask, permitiendo su acceso desde navegador.
- Se cargó el modelo LSTM entrenado en formato .h5 al iniciar la app con `keras.models.load_model`.
- La aplicación fue alojada en un servidor cloud (Fly.io) con configuración de entorno y dependencias mediante `requirements.txt`.

1.2 Construcción del modelo de planificación de rutas

1.2.1 Comprensión del negocio:

a. Determinación de los objetivos comerciales:

- Optimizar la planificación de rutas logísticas mediante un sistema que integre datos históricos de almacenes con información en tiempo real de la API de Google Maps para obtener distancias y tiempos precisos.
- Maximizar la utilización de la flota y mejorar la satisfacción del cliente.

b. Evaluación de la situación:

- Retrasos en las entregas debido a la ineficiencia en la planificación de rutas.
- Alta variabilidad en la demanda y falta de anticipación.
- Planificación basada en experiencia y reportes manuales.

c. Determinación de los objetivos de minería de datos:

- Desarrollar un modelo predictivo de rutas utilizando datos históricos de almacenes y obtener información en tiempo real de la API de Google Maps.
- Implementar un algoritmo de asignación de recursos logísticos para optimizar las rutas y reducir costos.
- Integrar estos modelos en un sistema funcional para tomar decisiones en tiempo real.

d. Producción de un plan de proyectos:

- Seguir la metodología CRISP-DM para estructurar el desarrollo del sistema.
- Recolectar, analizar e integrar los datos de almacenes con los datos de la API de Google Maps.
- Validar los modelos mediante simulaciones con datos reales.
- Capacitar al personal y monitorear el sistema con indicadores clave.

1.2.2 Comprensión de datos

a. Recopilación de datos iniciales:

- Se recopiló los datos de los almacenes de la base de datos interna, y se extrajeron las coordenadas geográficas (latitud y longitud) con la ayuda de la API de Google Maps para calcular distancias y tiempos de viaje.

b. Descripción de los datos:

- En la tabla IV se muestra el dataset con las siguientes variables relevantes para la construcción del modelo de rutas.

Tabla IV. Datos Utilizados para el modelo de planificación de rutas

Variable	Tipo	Descripción
origen	Texto	Lugar de inicio de la ruta.
latitud_origen	Numérico	Coordenada geográfica (latitud) del origen de la ruta.
longitud_origen	Numérico	Coordenada geográfica (longitud) del origen de la ruta.
destino	Texto	Lugar de destino de la ruta.
latitud_destino	Numérico	Coordenada geográfica (latitud) del destino de la ruta.
longitud_destino	Numérico	Coordenada geográfica (longitud) del destino de la ruta.
distancia_km	Numérico	Distancia total de la ruta en kilómetros.
costo_combustible	Numérico	Costo estimado del combustible en soles.
clima	Texto	Condiciones climáticas en la ruta
tiempo_estimado_horas	Numérico	Tiempo estimado en horas para completar la ruta.
horario_salida_sugerido	Hora	Hora recomendada para iniciar el traslado

1.2.3 Preparación de datos

a. Selección de datos:

- Se seleccionaron las variables clave necesarias para el análisis y entrenamiento del modelo, como origen, destino, distancia, costo de combustible, tiempo estimado, y coordenadas geográficas.

b. Limpieza de datos:

- Normalización de variables numéricas como distancia_km, costo_combustible y tiempo_estimado_horas.

c. Formateo de datos:

- Se codificaron las variables origen, destino, y clima utilizando LabelEncoder para convertir las categorías en variables numéricas aptas para el modelo de XGBoost.

1.2.4 Modelado

a. Selección de técnicas de modelado:

- Se eligió el modelo XGBoost debido a su capacidad para manejar grandes volúmenes de datos y su eficiencia en tareas de regresión, como la predicción del tiempo estimado de las rutas.

b. Generación de un diseño de comprobación:

- Se generaron secuencias temporales basadas en las rutas, utilizando variables como origen, destino, distancia, costo de combustible, y clima, para predecir el tiempo estimado de cada ruta, esto se puede observar en la figura 3.



```

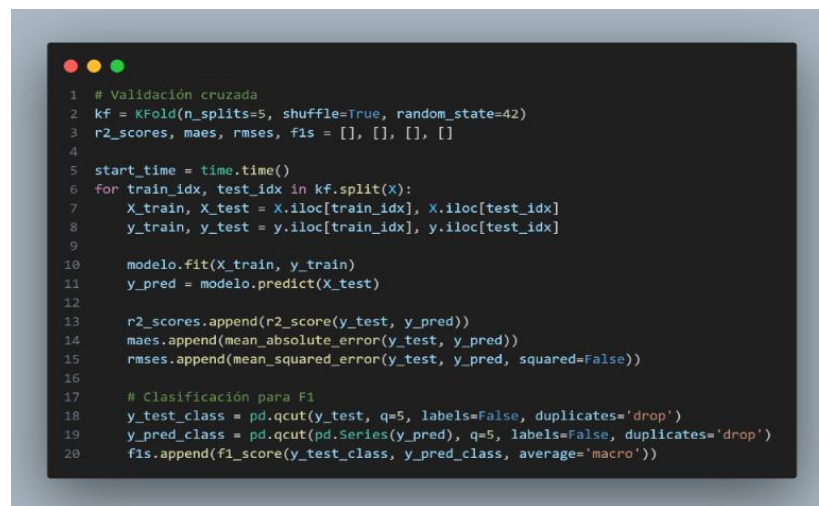
1  modelo = XGBRegressor(
2      n_estimators=2000, max_depth=30, learning_rate=0.002,
3      subsample=1.0, colsample_bytree=1.0, verbosity=0,
4      n_jobs=-1, tree_method='auto', random_state=42
5  )

```

Fig. 3. Configuración de Modelo XGBoost para Regresión

c. Generación de los modelos:

- Se construyó un modelo XGBoost con un enfoque en la predicción de tiempos, ajustando hiperparámetros para mejorar la precisión y evitar el sobreajuste. Se usaron métricas como MAE, RMSE, y R2 para evaluar su rendimiento, todo esto se demuestra en la figura 4.



```

1  # Validación cruzada
2  kf = KFold(n_splits=5, shuffle=True, random_state=42)
3  r2_scores, maes, rmses, fis = [], [], [], []
4
5  start_time = time.time()
6  for train_idx, test_idx in kf.split(X):
7      X_train, X_test = X.iloc[train_idx], X.iloc[test_idx]
8      y_train, y_test = y.iloc[train_idx], y.iloc[test_idx]
9
10     modelo.fit(X_train, y_train)
11     y_pred = modelo.predict(X_test)
12
13     r2_scores.append(r2_score(y_test, y_pred))
14     maes.append(mean_absolute_error(y_test, y_pred))
15     rmses.append(mean_squared_error(y_test, y_pred, squared=False))
16
17     # Clasificación para F1
18     y_test_class = pd.qcut(y_test, q=5, labels=False, duplicates='drop')
19     y_pred_class = pd.qcut(pd.Series(y_pred), q=5, labels=False, duplicates='drop')
20     fis.append(f1_score(y_test_class, y_pred_class, average='macro'))

```

Fig. 4. Implementación de Validación Cruzada y Métricas de Evaluación en Modelo

1.2.5 Evaluación

a. Evaluación de los resultados:

- El modelo fue entrenado con éxito, alcanzando un R² de 0.9908 y un F1 de 0.9337, lo que indica una excelente capacidad predictiva. El MAE fue de 0.4813, reflejando un buen ajuste y bajo error en las predicciones.

b. Proceso de Revisión:

- $R^2 = 0.9908$: Alta capacidad explicativa, el modelo explica más del 98% de la variabilidad de los tiempos de las rutas.
- $F1 = 0.9337$: Buen equilibrio entre precisión y recall.
- $MAE = 0.4813$: Error promedio bajo, lo que indica alta precisión en la predicción.
- Tiempo de entrenamiento de 42.813 segundos: El modelo fue eficiente en su entrenamiento.

1.2.6 Despliegue**a. Planificación de despliegue:**

- El modelo fue desplegado en una aplicación web desarrollada con Flask, permitiendo acceder a las predicciones de rutas en tiempo real mediante la API de Google Maps.
- El modelo y sus dependencias fueron configurados y alojados en un servidor cloud (Fly.io) para garantizar accesibilidad y rendimiento en tiempo real.

1.3 Construcción del modelo de asignación de recursos**1.3.1 Descripción General:**

- Función que genera una planificación logística para el mes indicado, asignando vehículos disponibles a pedidos predichos usando optimización con programación entera para minimizar costos.

1.3.2 Obtención y Preparación de Datos:

- Se obtienen vehículos y envíos para el mes especificado, convertidos a DataFrames para manipulación.

1.3.3 Construcción del Modelo:

- Se crea el solver SCIP y las variables binarias para asignar vehículos a pedidos, aunque también se probó con BCB.
- Se calculan distancias geográficas y se asignan costos que penalizan los días de espera para cada posible asignación vehículo-pedido. Se minimiza la suma ponderada de estos costos, tal y como se observa en la figura 5.

```

1 solver = pywraplp.Solver.CreateSolver("SCIP")
2 x = {(i, j): solver.BoolVar(f"x({i},{j})")} for i in range(num_vehiculos) for j in range(num_pedidos)
3
4 solver.Minimize(solver.Sum(x[i, j] * costos[i][j] for i in range(num_vehiculos) for j in range(num_pedidos)))
5 if solver.Solve() != pywraplp.Solver.OPTIMAL:
6     continue
7
8 for i in range(num_vehiculos):
9     for j in range(num_pedidos):
10        if x[i, j].solution_value() > 0.5:
11            vehiculo = vehiculos_disponibles.loc[i, "placa"]
12            pedido = pedidos.loc[j, "descripcion"]
13            distancia = distancias[i][j]
14
15            reporte["envios_por_vehiculo"][vehiculo] += 1
16            reporte["frecuencia_rutas"][pedido] += 1
17            reporte["distancia_total_km"] += distancia
18            reporte["dias_operativos"][vehiculo] += 1
19            reporte["rutas_distancia"][pedido] += distancia
20            COSTO_POR_KM = precio_gasolina/3.5
21            reporte["costo_operativo_total"] += distancias[i][j] * COSTO_POR_KM
22            precio_estimado = distancias[i][j] * precio_gasolina
23            reporte["precio_total_estimado"] += precio_estimado
24
25            if 'conductor' in vehiculos_disponibles.columns:
26                reporte["vehiculo_conductor"][vehiculo] = vehiculos_disponibles.loc[i, "conductor"]
27
28            vehiculos_df.at[i, 'latitud'] = pedidos.loc[j, 'latitud_destino']
29            vehiculos_df.at[i, 'longitud'] = pedidos.loc[j, 'longitud_destino']
30            vehiculos_df.at[i, 'dias_espera'] = 0

```

Fig. 5. Optimización de la Asignación de Vehículos en un Sistema de Rutas

1.3.4 Ciclo por Día y Actualización de Estado de Vehículos:

- Para cada día único con envíos, se actualizan las posiciones de vehículos que regresan a la base y se filtran vehículos no disponibles, tal y como se ve en la figura 6.

```

1 for fecha in fechas_unicas:
2     vehiculos_df['dias_espera'] += 1
3
4     vehiculos_regresando = vehiculos_df[vehiculos_df['dias_espera'] > 3]
5     for _, vehiculo in vehiculos_regresando.iterrows():
6         reporte["retornos_base"][vehiculo['placa']] += 1
7         vehiculos_df.loc[vehiculo.name, ['latitud', 'longitud']] = bases[0]
8
9     vehiculos_disponibles = vehiculos_df.copy()
10    fecha_ts = pd.to_datetime(fecha)
11
12    if isinstance(no_disponibilidad, dict) and fecha_ts in no_disponibilidad:
13        no_disp = no_disponibilidad[fecha_ts]
14        vehiculos_disponibles = vehiculos_disponibles[~vehiculos_disponibles['placa'].isin(no_disp)].reset_index(drop=True)

```

Fig. 6. Asignación de Vehículos y Gestión de Disponibilidad por Fechas

1.3.5 Restricciones:

- Cada pedido debe ser asignado a exactamente un vehículo.
- La suma del peso de los pedidos asignados a cada vehículo no debe superar su capacidad máxima.
- Cada vehículo solo puede ser asignado a un pedido por día, el uso de estas reglas es demostrado en la figura 7.

```

1 for i in range(num_vehiculos):
2     solver.Add(solver.Sum(x[i, j] * float(pedidos.loc[j, 'peso']) for j in range(num_pedidos))
3                 <= float(vehiculos_disponibles.loc[i, 'capacidad_max']))
4     solver.Add(solver.Sum(x[i, j] for j in range(num_pedidos)) <= 1)

```

Fig. 7. Restricción de Capacidad y Peso en Asignación de Vehículos

2. Validación del cumplimiento de los objetivos

2.1 Construir un modelo para la predicción de demanda en la empresa “Héctor Ocaña Izquierdo”, que permita anticipar picos y valles en la demanda.

- Se desarrolló un modelo basado en redes neuronales LSTM para anticipar la demanda de envíos por cliente en la empresa en estudio. El modelo permite identificar picos y valles de forma anticipada, mejorando la planificación logística. En la figura 8 se demuestran las métricas alcanzadas por este modelo

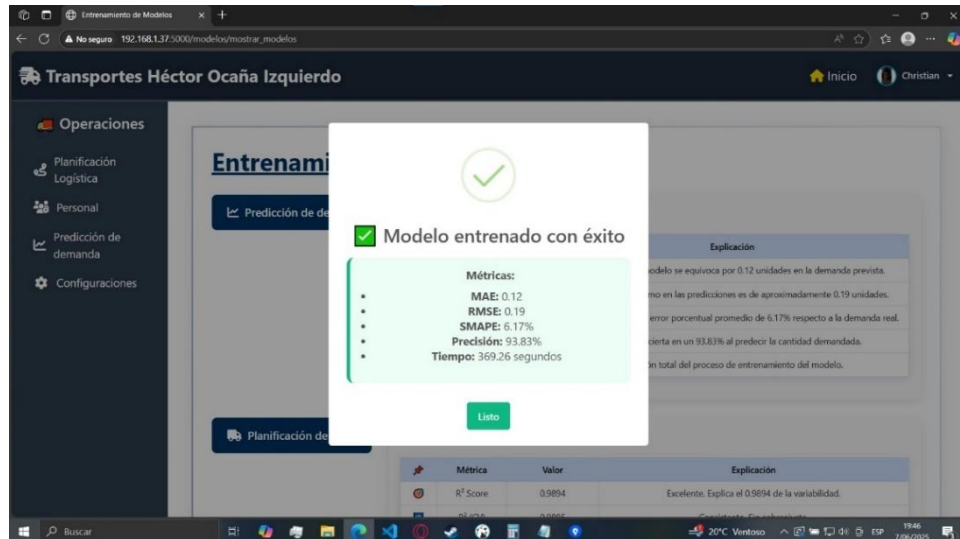


Fig. 8. Entrenamiento Exitoso del Modelo de predicción de demanda con Métricas de Desempeño

INDICADORES:

- **Error absoluto medio (MAE):**
 - **Valor:** 0.12
 - **Descripción:** En promedio, el modelo comete un error de 0.12 unidades al predecir la demanda. Este indicador muestra qué tan cerca están, en promedio, las predicciones del valor real.
- **Error Cuadrático Medio (RMSE).**
 - **Valor:** 0.19
 - **Descripción:** Representa el error típico máximo en las predicciones. Un valor de 0.19 indica que, en promedio, el modelo se desvía en 0.19 unidades respecto a la demanda real, penalizando más los errores grandes.
- **Tiempo de entrenamiento del modelo.**
 - **Valor:** 369.26 segundos
 - **Descripción:** Tiempo total requerido para entrenar el modelo desde el inicio hasta la finalización.

2.2 Desarrollar un modelo para la planificación de rutas en la empresa "Héctor Ocaña Izquierdo" que permita optimizar la distribución de los recursos logísticos

- Se desarrolló un modelo de optimización de rutas, que permite asignar de forma eficiente los recursos logísticos y reducir tiempos y costos en la distribución. En la figura 9 se muestran las métricas alcanzadas para el modelo de rutas.

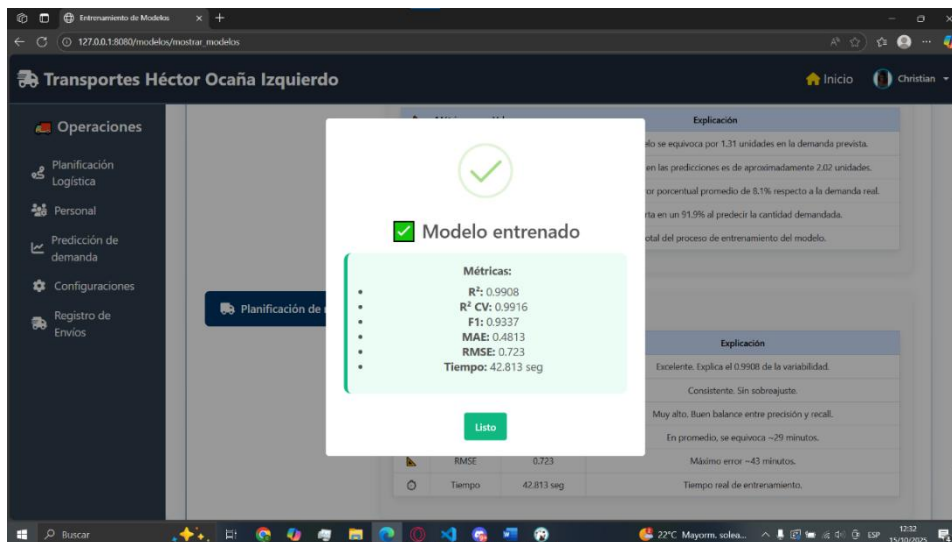


Fig. 9. Entrenamiento Exitoso del Modelo de planificación de rutas con Métricas de Desempeño

INDICADORES:

- **Tiempo Promedio de Ejecución del Modelo:**
 - **Valor:** 42.813 segundos
 - **Descripción:** Tiempo real requerido para ejecutar y entrenar completamente el modelo. Este valor indica la eficiencia computacional del modelo en términos de tiempo.
- **Precisión de la ruta:**
 - **Valor:** 0.9908
 - **Descripción:** El modelo explica aproximadamente el 99% de la variabilidad en los datos de las rutas. Este valor indica una excelente capacidad predictiva.
- **F1-Score:**
 - **Valor:** 0.9337
 - **Descripción:** Mide el balance entre precisión y recall (sensibilidad). Un valor alto como este indica que el modelo tiene un rendimiento muy bueno al clasificar correctamente las rutas u horarios, sin perder demasiados casos relevantes ni cometer muchos falsos positivos.

2.3 Integrar el modelo de predicción de demanda con el de planificación de rutas para optimizar el servicio, mejorando así la satisfacción del cliente.

- Se logró integrar exitosamente el modelo de predicción de demanda con el sistema de planificación de rutas. Esta integración garantiza que las proyecciones de demanda generadas sean utilizadas directamente como insumo para la asignación óptima de vehículos, mejorando la eficiencia operativa.

INDICADORES:

- **Tasa de éxito en la interoperabilidad:**
 - **Valor:** 100%
 - **Descripción:** El sistema utilizó correctamente las salidas del modelo de predicción de demanda como entradas para la planificación de rutas, sin errores ni conflictos, demostrando una interoperabilidad completa entre los módulos.
- **Tiempo de respuesta conjunto:**
 - **Valor:** 3 minutos 51 segundos
 - **Descripción:** Mide el tiempo total desde el inicio de la predicción de demanda hasta la obtención de los resultados de planificación de rutas. En esta prueba, el sistema completó ambas tareas en 3 minutos 51 segundos, mostrando una respuesta eficiente para entornos logísticos, tal y como se observa en la figura 10.

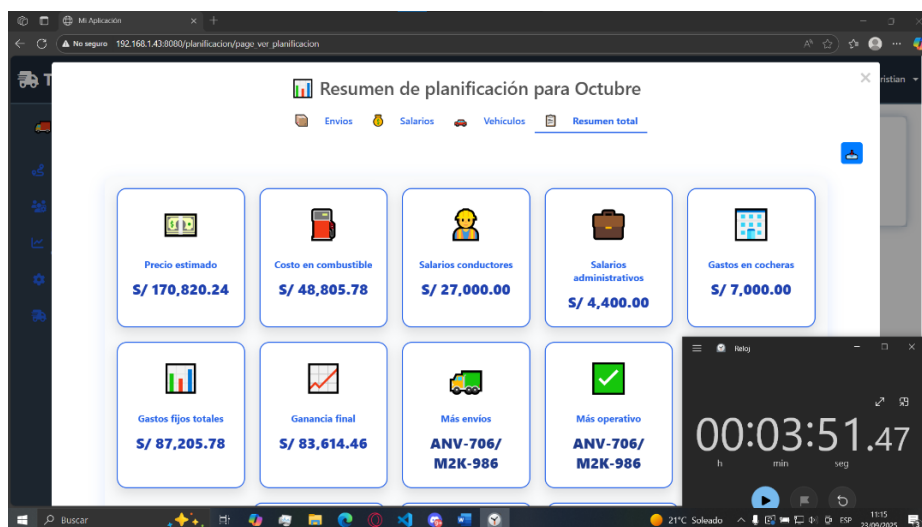


Fig. 10. Análisis de Integración de Modelos y Tiempo de Procesamiento

- **Satisfacción del usuario:**
 - **Valor:** 93.68%
 - **Descripción:** Mide la aceptación del sistema por parte de conductores y administrativos, basado en encuestas de 5 preguntas para cada grupo. Se contabilizaron las respuestas de "De acuerdo" y "Muy de acuerdo", reflejando un alto nivel de conformidad, dichas preguntas y respuestas quedan demostradas en la tabla V.

Tabla V. Resultados de la encuesta realizada al personal de la empresa

Pregunta	De acuerdo	Muy de acuerdo
¿Cree que el sistema facilitará la planificación para cumplir mejor con las expectativas de los clientes?	42.1%	52.6%
¿Confía en que el sistema ayudará a reducir los errores y retrasos que afectan la satisfacción del cliente?	47.4%	42.1%
¿Considera que la planificación con el sistema reducirá el estrés relacionado con el trabajo y mejorará la atención al cliente?	36.8%	57.9%
¿Esperaría recibir información clara y oportuna sobre cambios en rutas o entregas para atender mejor a los clientes?	68.4%	26.3%
¿Cree que el sistema contribuirá a que los clientes estén más satisfechos con el servicio?	36.8%	57.9%

- Para demostrar el cumplimiento del indicador, se hizo uso de la siguiente fórmula:

$$\text{SatisfacciónUsuario} = \frac{\text{Número de respuestas Deacuerdo} + \text{Muydeacuerdo}}{\text{Total de respuestas}} \times 100$$

$$\text{SatisfacciónUsuario} = \frac{89}{95} \times 100 = 93.68\%$$

2.4 Validar la conformidad del sistema inteligente con los estándares de calidad ISO 25010, con un enfoque particular en la adecuación funcional y fiabilidad del sistema.

- **Tasa de cumplimiento funcional:**
 - **Valor:** 100%
 - **Descripción:** Se realizaron 19 pruebas exitosas de un total de 19 planificadas, lo que indica un cumplimiento total de los requisitos funcionales.

- Para lograrlo, se empleó la siguiente fórmula:

$$\text{Tasa de Cumplimiento funcional (\%)} = \left(\frac{\text{N}^\circ \text{ de pruebas exitosas}}{\text{N}^\circ \text{ total de pruebas planificadas}} \right) \times 100$$

$$\text{Tasa de Cumplimiento funcional (\%)} = \left(\frac{19}{19} \right) \times 100 = 100\%$$

- **Tasa de defectos por ejecución:**

- **N.º total de ejecuciones:** 2 veces se ejecutaron las pruebas de calidad = 38 ejecuciones.
- **N.º total de defectos encontrados:** 3 validaciones mínimas.
- **Valor:** 0.08
- **Descripción:** En 38 ejecuciones (2 rondas de 19 pruebas), se encontraron 3 observaciones menores relacionadas con validaciones. Este valor refleja un alto nivel de fiabilidad, cumpliendo con los estándares de calidad ISO/IEC 25010.
- Para garantizar su cumplimiento, se hizo uso de la siguiente fórmula

$$\text{Tasa de Defectos} = \frac{\text{N.º total de defectos encontrados}}{\text{N.º total de ejecuciones de prueba}}$$

$$\text{Tasa de Defectos} = \frac{3}{38} \approx 0.08$$

- **Tiempo promedio de levantamiento entre fallos:**

- **N.º de pruebas simuladas:** 2
- **Tiempos registrados:** 42.30 s y 31.74 s
- **Valor promedio:** 37.02 segundos
- **Descripción:** Se realizaron 2 simulaciones de fallos controlados, midiendo el tiempo hasta que la interfaz web estuvo operativa nuevamente. El tiempo promedio fue de 37.02 segundos, lo que indica una excelente capacidad de recuperación y fiabilidad, conforme al estándar ISO/IEC 25010, esto queda demostrado mediante los valores observados en la figura 11 y figura 12.

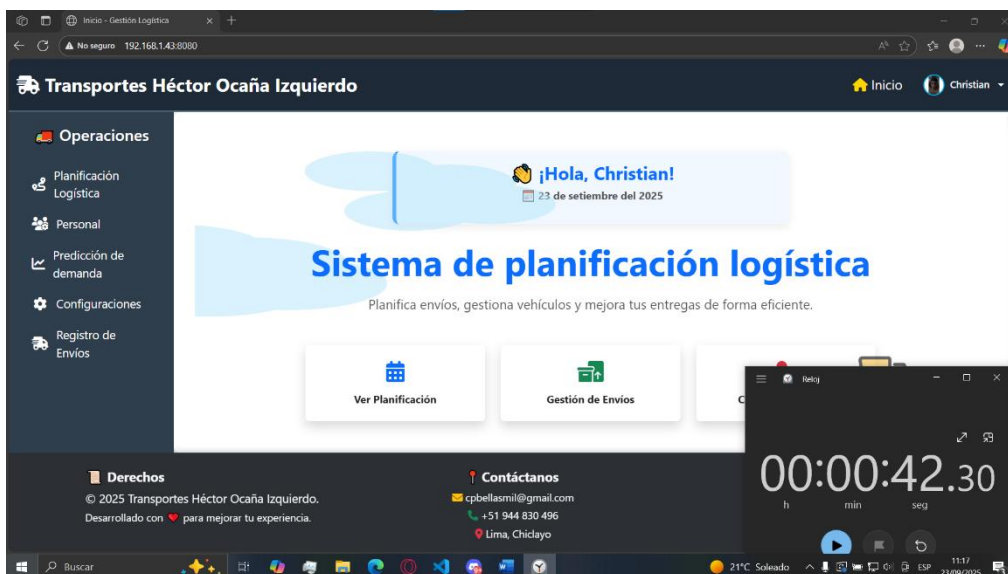


Fig. 11. Tiempo Promedio de Recuperación entre Fallos: Primer Resultado

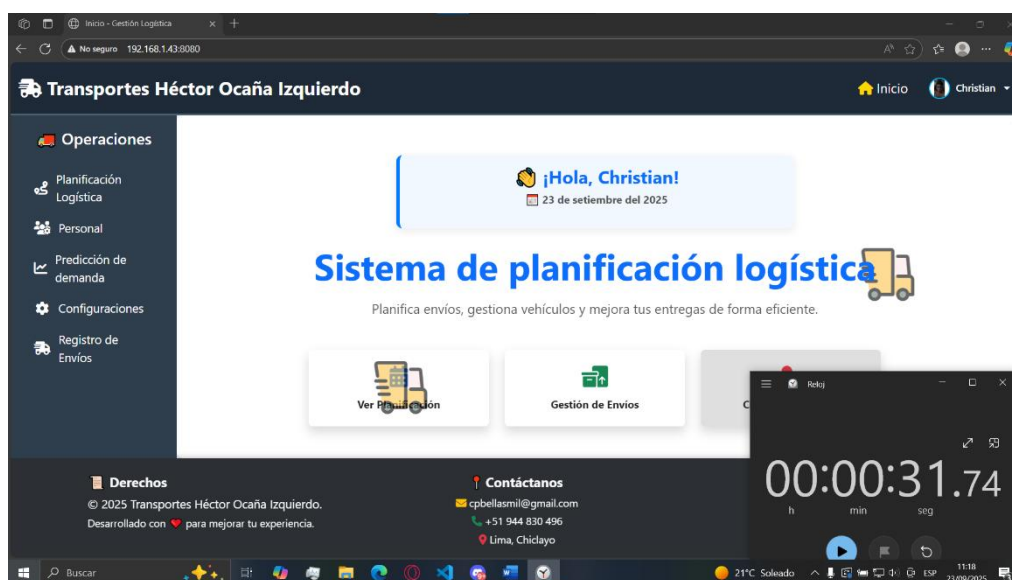


Fig. 12. Tiempo Promedio de Recuperación entre Fallos: Segundo Resultado

- Para demostrar el cumplimiento de este indicador, se hizo uso de la siguiente fórmula

$$TPLEF = \frac{\sum \text{Tiempos de recuperación}}{\text{N.º de pruebas simuladas}}$$

$$TPLEF = \frac{42.30 + 31.74}{2} = \frac{74.04}{2} = 37.02 \text{ segundos}$$

- **Disponibilidad del sistema:**
 - **Periodo observado:** 540 minutos (9 horas continuas de funcionamiento).

- **Tiempo inactivo registrado:** 3 minutos (por inestabilidad del internet y reinicio).
- **Valor:** 99.44%
- **Descripción:** De 540 minutos de monitoreo, solo 3 minutos fueron de inactividad controlada, lo que resulta en una disponibilidad del 99.44%, adecuada para entornos logísticos de alta continuidad operativa, según el estándar ISO/IEC 25010.
- Para lograr el cumplimiento de este indicador, se utilizó la siguiente fórmula:

$$\text{Disponibilidad (\%)} = \left(\frac{\text{Tiempo total operativo} - \text{Tiempo inactivo}}{\text{Tiempo total operativo}} \right) \times 100$$

$$\text{Disponibilidad (\%)} = \left(\frac{540 - 3}{540} \right) \times 100 = \left(\frac{537}{540} \right) \times 100 \approx 99.44\%$$

Discusión

El modelo resultado de la presente investigación basado en LSTM (Long Short-Term Memory) para la predicción de demanda, alcanzó una precisión del 93.83% con un MAE de 0.12 y un RMSE de 0.19, superando otros enfoques previos como el de Roque [15], quien utilizó SARIMA y minería de datos en el Metropolitano de Lima, logrando un R² entre 0.91 y 0.98, y un RMSPE entre 6.37 y 8.13. Si bien ambos modelos muestran alta precisión, LSTM captura dependencias a largo plazo, lo que mejora el rendimiento frente a los métodos tradicionales. En el estudio de Ali et al. [13], demostraron que simplificar modelos de machine learning puede reducir los costos computacionales hasta un 98% sin perder el índice de precisión (0.6% a 4.2%) adecuada, lo que también se refleja en el modelo, que combina alta precisión y eficiencia computacional. Por último, en el estudio de Ren et al. [14], implementaron un modelo de red neuronal convolucional multimodal, obtuvo una precisión destacable, alcanzando un R² de 0.8935 para la demanda diaria y de 0.8739 para las horas pico, lo que subraya el rendimiento superior del modelo LSTM en términos de captura de dependencias temporales y manejo de la complejidad de los datos.

El modelo de optimización de rutas propuesto, utilizando MILP, XGBoost y APIs para integrar datos de clima y tráfico en tiempo real, alcanzó una precisión del 99.08%, superando el enfoque de Soto et al. [17], quienes lograron un 98% con algoritmos heurísticos como Clarke y Wright y Dijkstra. Aunque ambos modelos son altamente precisos, el enfoque propuesto es más robusto al considerar factores dinámicos como el clima, el tráfico y los días en espera, lo

que permite una optimización más flexible. En comparación con Lee et al. [18], quienes aplicaron un modelo innovador basado en Transformer y PPO para optimizar las rutas marítimas y lograron reducir el consumo de combustible en un 11.33%, el modelo propuesto en esta investigación alcanza una reducción del 24.5% en el tiempo de conducción, mejorando significativamente la eficiencia operativa. Por otro lado, en comparación con Villanueva [19], quien alcanzó una reducción del 44% en las distancias recorridas para la flota liviana, el modelo propuesto logra una reducción del 34.47%. Aunque Villanueva logró una mayor reducción en distancia, el modelo propuesto se distingue por integrar de manera más eficiente la asignación de recursos y planificación de rutas, combinándola con datos de demanda en tiempo real, lo que mejora la precisión operativa y la flexibilidad frente a condiciones cambiantes.

En el ámbito de la planificación logística, pocos estudios han logrado integrar eficazmente la predicción de demanda y la planificación de rutas. Aunque el estudio de Kmiecik [20] utiliza control predictivo en tiempo real para la planificación de rutas, su enfoque se limita a ajustar la velocidad de los vehículos según el tráfico, sin abordar la demanda fluctuante, lo que reduce su aplicabilidad en sistemas logísticos con demanda variable. En cambio, el modelo propuesto en esta investigación optimiza las rutas y anticipa la demanda, permitiendo una asignación eficiente de recursos y una optimización continua en función de los cambios en la demanda y las condiciones operativas. Bazan et al. [21], también combinaron planificación de rutas y predicción de demanda, pero su uso del algoritmo Clark-Wright no ajusta las rutas con la flexibilidad en tiempo real que ofrece el modelo propuesto. La integración dinámica de ambos modelos en este estudio proporciona una solución más robusta y flexible, mejorando la eficiencia logística en tiempo real, algo no explorado en investigaciones anteriores.

El sistema fue validado conforme a los estándares de calidad ISO 25010, alcanzando un 100% de cumplimiento funcional. Las métricas de confiabilidad del sistema fueron sobresalientes, con una tasa de defectos de 0.08, lo que indica que el sistema es robusto y fiable para su uso en un entorno operativo real. En cuanto a la disponibilidad, el sistema alcanzó un 99.44% de tiempo operativo, reflejando una capacidad de recuperación ante fallos que asegura alta continuidad operativa. Este nivel de fiabilidad es comparable con los resultados obtenidos en el estudio de Martínez [16], quien implementó una solución basada en Machine Learning para la gestión de inventarios en la empresa "Multiservicios Chino Kam", logrando un 90% de cumplimiento en los criterios de usabilidad según la norma ISO 25010. A diferencia de los estudios anteriores, que no integraban completamente la predicción de demanda con la optimización de operaciones, el modelo propuesto destaca por su alta fiabilidad y eficiencia, lo que refuerza su aplicabilidad y robustez en entornos operativos reales.

Conclusiones

- El sistema de machine learning implementado procesó datos en tiempo real, agilizando la toma de decisiones y optimizando la cadena de suministro. La solución demostró la viabilidad de transformar procesos logísticos tradicionales en operaciones inteligentes basadas en datos, estableciendo un nuevo estándar en la gestión logística de la empresa.
- Se desarrolló con éxito un modelo LSTM para la predicción de demanda, alcanzando una precisión del 93.83%. El modelo comete un error de 0.12 unidades en sus predicciones y tiene una desviación media de 0.19 unidades, penalizando más los errores mayores. Estos resultados fueron fundamentales como entradas para el algoritmo de planificación de rutas, permitiendo prever de manera más eficaz las fluctuaciones en la demanda.
- El modelo de optimización de rutas alcanzó una precisión del 99.08%, lo que permitió una mejora significativa en la eficiencia operativa, logrando una reducción del 24.5% en el tiempo de conducción, lo que resultó en un proceso más rápido y eficiente en comparación con la planificación previa. Además, se consiguió una reducción del 34.47% en el recorrido, optimizando la distribución de los recursos logísticos.
- Se logró integrar exitosamente el modelo de predicción de demanda con el sistema de planificación de rutas, logrando una interoperabilidad del 100%. El proceso de predicción y planificación se completó en 3 minutos con 51 segundos, demostrando eficiencia para decisiones rápidas. Además, la satisfacción del usuario alcanzó el 93.68%, con mejoras en la asignación de rutas y la reducción de retrasos, elevando la calidad del servicio al cliente.
- El sistema cumplió con una Tasa de Cumplimiento Funcional del 100%, validando todos los requisitos funcionales sin fallos. La Tasa de Defectos por Ejecución fue de 0.08, lo que refleja un alto nivel de fiabilidad. Además, la disponibilidad del sistema fue del 99.44% durante 540 minutos de operación continua, con solo 3 minutos de inactividad controlada, garantizando una alta continuidad operativa y cumpliendo con los estándares ISO/IEC 25010.

Recomendaciones

- Se recomienda ampliar el conjunto de datos para incluir más variables, lo que mejorará la precisión de las predicciones y permitirá una mayor capacidad de adaptación del sistema a diferentes condiciones logísticas. Además, para investigaciones próximas se sugiere incorporar técnicas avanzadas de aumento de datos que simulen escenarios con fluctuaciones inesperadas en la demanda o eventos externos, como huelgas o desastres naturales, lo cual aumentará la robustez del sistema, permitiéndole adaptarse mejor a cambios no planificados.
- Se sugiere integrar el sistema con plataformas de gestión de flotas o plataformas de monitoreo en tiempo real de tráfico y condiciones meteorológicas. Esta integración enriquecería la toma de decisiones y permitiría ajustar las rutas de manera dinámica, optimizando aún más los tiempos de entrega y reduciendo costos operativos.
- Se recomienda integrar el sistema con plataformas municipales de gestión de transporte. Esto mejoraría la visibilidad de los recursos logísticos a nivel urbano, optimizando la distribución en función de las condiciones locales. Además, se podrían incorporar funcionalidades como el seguimiento de la ubicación de vehículos en tiempo real y la implementación de alertas automáticas sobre posibles riesgos, como congestiones de tráfico o cambios meteorológicos imprevistos.
- Se sugiere desarrollar una interfaz móvil para conductores y administradores, que permita gestionar las rutas en tiempo real, recibir actualizaciones sobre el estado de las entregas y optimizar la asignación de recursos en función de las condiciones cambiantes durante el día. Esto proporcionaría mayor flexibilidad en la operación logística y mejoraría la satisfacción del cliente.

Referencias

- [1] A. Diaz, «Planificación logística de aprovisionamiento en las empresas industriales de Colombia y sus estrategias para ser competitivos en la economía nacional», n.º 1, 2019, Accedido: 11 de septiembre de 2024. [En línea]. Disponible en: <http://repositorio.uts.edu.co:8080/xmlui/bitstream/handle/123456789/12301/Informe%20Final.pdf?sequence=1&isAllowed=y>
- [2] V. Pysarenko, N. Bahan, V. Sobchyshyn, R. Gawrych, I. Zuchowski, y N. Moshenets, «Sustainable development of agrarian innovation-oriented enterprises by the means of logistics management», *Financ. Credit Act. Probl. Theory Pract.*, doi: 10.55643/fcaptp.3.56.2024.4424.
- [3] Oracle, «¿Qué es un sistema de gestión del transporte? | Oracle». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://www.oracle.com/latam/scm/logistics/transportation-management/what-is-transportation-management-system/>
- [4] M. Echarri, «Un futuro cada vez más presente. Así se transformará el sector logístico en 2023», *El País*. Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://elpais.com/economia/nuevos-tiempos/2023-01-12/un-futuro-cada-vez-mas-presente-asi-se-transformara-el-sector-logistico-en-2023.html>
- [5] Mordor Intelligence, «Carga y Logística en Sudamérica Tamaño del Mercado | Mordor Intelligence». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://www.mordorintelligence.com/es/industry-reports/south-american-freight-and-logistics-market>
- [6] Ministerio de Transportes y Comunicaciones (MTC), «MTC presenta resultados de la Primera Encuesta Nacional de Logística en el Perú». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://cdn.www.gob.pe/uploads/document/file/4331508/Encuesta%20Nacional%20de%20Logística%202020.pdf?v=1681934904>
- [7] P. Salvador, «Madurez en la cadena de suministro del Perú». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: https://www.ey.com/es_pe/insights/consulting/madurez-cadena-de-suministro-peru
- [8] EY, «Estudio Supply Chain Overview - 2021». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://www.ey.com/content/dam/ey-unified-site/ey-com/es-pe/insights/consulting/documents/ey-madurez-gestion-cadena-suministros-2021.pdf>
- [9] A. Torres y R. del C. Cruz, «La Logística del Transporte como Factor Clave en la Cadena de Suministro», *Cienc. Lat. Rev. Científica Multidiscip.*, vol. 9, n.º 1, pp. 10771-10784, mar. 2025, doi: 10.37811/cl_rcm.v9i1.16669.
- [10] C. Mejía y D. Pin, «Propuesta de mejora de procesos en la gestión de distribución de una comercializadora de sistemas hidráulicos basado en la aplicación del ciclo de Deming», [En línea]. Disponible en: <https://dspace.ups.edu.ec/bitstream/123456789/29733/1/UPS-GT006027.pdf>
- [11] L. Rodríguez, I. Pérez, y S. Núñez, «Caracterización del proceso logístico para la toma de decisiones en la Sucursal Emprestur Pinar del Río», *Avances*, vol. 26, n.º 1, pp. 18-35, 2024, Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=9269989>
- [12] Prosegur, «La logística del futuro: así se están transformando las cadenas globales de suministro | Tecnología», Prosegur Innovacion. Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <http://www.prosegur.com/innovacion/tecnologia/la-logistica-del-futuro-transforma-las-cadenas-globales-de-suministro>

- [13] A. Ali, R. Jayaraman, E. Azar, y M. Maalouf, «Maximizing supply chain performance leveraging machine learning to anticipate customer backorders», *Comput. Ind. Eng.*, vol. 194, p. 110414, ago. 2024, doi: 10.1016/j.cie.2024.110414.
- [14] P. Ren, S. Liu, B. Qin, Y. Chen, Q. Xu, y P. He, «A novel multimodal deep learning-based direct ridership model for planning-oriented demand prediction in urban rail transit networks», *J. Transp. Geogr.*, vol. 129, p. 104409, dic. 2025, doi: 10.1016/j.jtrangeo.2025.104409.
- [15] E. Roque, «Predicción de la demanda de pasajeros a clústeres de estaciones del metropolitano usando métodos de data mining, la Metodología Box- Jenkins y Sarima», Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://repositorio.ulima.edu.pe/bitstream/handle/20.500.12724/16675/Tesis.pdf?sequence=1&isAllowed=y>
- [16] L. Martínez, «Aplicación web basada en machine learning para predecir la demanda de productos en la empresa Multiservicios Chino Kam». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: https://tesis-usat-edu-pe.usat.lookproxy.com/bitstream/20.500.12423/6817/1/TL_MartinezSoplapucoLisle.pdf
- [17] G. Soto, S. Huamán, y A. Zuazo, «Implementación de rutas de transporte para la optimización de consumo de la energía combustible en la cadena de suministro de los productos de termoenergías del Perú S.A. en Lima Metropolitana», [En línea]. Disponible en: <https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/621548/TEMA%2055-SOTO-HUAMAN-ZUAZO.pdf;jsessionid=5E3C16E696DC27777186BF35B4DC958D?sequence=5>
- [18] J. Lee, Y. Park, J. Eom, H. Hwang, y S. Kim, «Ship Voyage Route Waypoint Optimization Method Using Reinforcement Learning Considering Topographical Factors and Fuel Consumption», *J. Mar. Sci. Eng.*, vol. 13, n.º 8, p. 1554, ago. 2025, doi: 10.3390/jmse13081554.
- [19] M. Villanueva, «Propuesta de optimización de rutas de distribución en la empresa DD & D S. A. C. para reducir costos operativos». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: https://tesis-usat-edu-pe.usat.lookproxy.com/bitstream/20.500.12423/7503/1/TL_VillanuevaMalcaMariela.pdf
- [20] M. Kmiecik, «Logistics Coordination Based on Inventory Management and Transportation Planning by Third-Party Logistics (3PL)», *Sustainability*, vol. 14, n.º 13, p. 8134, jul. 2022, doi: 10.3390/su14138134.
- [21] E. Bazan, C. Gamarra, J. Taquía, y Y. García, «Demand Forecast Model and Route Optimization to Improve the Supply of an SME in the Bakery Sector», pp. 957-967, 2023, doi: 10.15488/15321.
- [22] C. Quijije, «Impacto de la tecnología en la gestión logística de las PYMES guayaquileñas: sector Vía Daule», bachelorThesis, 2021. Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <http://dspace.ups.edu.ec/handle/123456789/19648>
- [23] Acelera pyme, «Optimiza la gestión del transporte con el machine learning | Acelera pyme». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <http://www.acelerapyme.gob.es/novedades/pildora/optimiza-la-gestion-del-transporte-con-el-machine-learning>
- [24] K. Alicke *et al.*, «Succeeding in the AI supply-chain revolution | McKinsey». Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://www.mckinsey.com/industries/metals-and-mining/our-insights/succeeding-in-the-ai-supply-chain-revolution#/>
- [25] V. Lombeida y J. Conforme, «Influencia de la planificación logística en la calidad del servicio al cliente en la cadena de suministro, la Casa del Toldo, Santo Domingo, 2024»,

- Código Científico Rev. Investig.*, vol. 6, n.º E1, pp. 1716-1746, mar. 2025, doi: 10.55813/gaea/ccri/v6/nE1/792.
- [26] D. Sánchez, V. Lozada, y G. Asmat, «Impacto de la inteligencia artificial en la competitividad logística: el megapuerto de Chancay en Perú como conector entre la República Popular China y América Latina», *Rev. Análisis Difus. Perspect. Educ. Empres.*, vol. 4, n.º 9, pp. 9-28, dic. 2024, doi: 10.56216/radee032024dic.a01.
- [27] J. Cardenes, «La aplicación de Big Data e Inteligencia Artificial en logística y transporte para la optimización de procesos en empresas», 2022, Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://repositorio.comillas.edu/xmlui/handle/11531/56434>
- [28] M. Cobo, «Modelo de previsión de demanda de transporte logístico de palés», *Forecast model for demand for logistical transport of pallets*, jun. 2021, Accedido: 11 de septiembre de 2025. [En línea]. Disponible en: <https://repositorio.unican.es/xmlui/handle/10902/22458>
- [29] BuskeLogistics, «Empresa de transporte y logística Definition & Meaning». Accedido: 13 de noviembre de 2025. [En línea]. Disponible en: <https://www.buske.com/es/what-is/shipping-and-logistics-company>
- [30] O. Munro, «Logistics Planning: Benefits & Best Practices Guide». Accedido: 12 de septiembre de 2025. [En línea]. Disponible en: <https://www.unleashedsoftware.com/blog/logistics-planning/>
- [31] M. Ponte y P. Vela, «La gestión logística y su relación con el control de inventarios en una empresa de servicio, Lima 2022», Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: https://repositorio.uwiener.edu.pe/bitstream/handle/20.500.13053/8773/T061_46084770_T.pdf?sequence=1&isAllowed=y
- [32] J. Pinedo, «Propuesta de un modelo de pronósticos de demanda y gestión de inventarios para la planeación de demanda en prendas de vestir juvenil», Accedido: 12 de septiembre de 2025. [En línea]. Disponible en: https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/623528/Pinedo_CJ.pdf
- [33] TLW, «Cómo planificar una ruta logística exitosa: consejos de expertos», THE LOGISTICS WORLD | Conéctate e inspírate. Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://thelogisticsworld.com/logistica-y-distribucion/como-planificar-una-ruta-logistica-exitosa-consejos-de-expertos/>
- [34] Amazon Web Services, «¿Qué es el machine learning? - Explicación sobre el machine learning empresarial - AWS», Amazon Web Services, Inc. Accedido: 12 de septiembre de 2025. [En línea]. Disponible en: <https://aws.amazon.com/es/what-is/machine-learning/>
- [35] IBM, «¿Qué es una red neuronal? | IBM». Accedido: 12 de septiembre de 2025. [En línea]. Disponible en: <https://www.ibm.com/es-es/think/topics/neural-networks>
- [36] M. Khan *et al.*, «Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review», *Expert Syst. Appl.*, vol. 185, p. 115695, dic. 2021, doi: 10.1016/j.eswa.2021.115695.
- [37] IBM, «¿Qué es XGBoost? | IBM». Accedido: 13 de noviembre de 2025. [En línea]. Disponible en: <https://www.ibm.com/mx-es/think/topics/xgboost>
- [38] IBM, «¿Qué es la programación lineal de enteros mixta?» Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://www.ibm.com/docs/es/icos/22.1.2?topic=programming-what-is-mixed-integer-linear>
- [39] IBM, «Conceptos básicos de ayuda de CRISP-DM». Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>

- [40] Google, «About OR-Tools», Google for Developers. Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://developers.google.com/optimization/introduction>
- [41] Google, «Definición del algoritmo SHA-256 - Ayuda de Google Ads». Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://support.google.com/google-ads/answer/9004655?hl=es-419>
- [42] IBM, «¿Qué es una API REST (API RESTful)? | IBM». Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://www.ibm.com/es-es/think/topics/rest-apis>
- [43] OECD, *Manual de Frascati 2015: Guía para la recopilación y presentación de información sobre la investigación y el desarrollo experimental*. OECD, 2018. doi: 10.1787/9789264310681-es.
- [44] C. Bernal, «Pdf-metodologia-de-investigacion compress - César Bernal - Cuarta Edición», Studocu. Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://www.studocu.com/pe/document/servicio-nacional-de-adiestramiento-en-trabajo-industrial/gestion-dde-calidad/pdf-metodologia-de-investigacion-compress/84146303>
- [45] M. Genero, J. Cruz, y M. Piattini, «Métodos de Investigación en Ingeniería de Software | PDF». Accedido: 13 de septiembre de 2025. [En línea]. Disponible en: <https://es.scribd.com/document/681180241/Metodos-de-investigacion-en-ingenieria-de-Software>
- [46] K. Williamson, «Research Methods: Information, Systems, and Contexts», *J. Aust. Libr. Inf. Assoc.*, vol. 67, n.º 2, pp. 185-186, abr. 2018, doi: 10.1080/24750158.2018.1466638.
- [47] E. Madrigal, «Conoce las métricas de precisión más comunes para Modelos de Regresión Grow Up», Grow Up. Accedido: 2 de noviembre de 2025. [En línea]. Disponible en: <https://www.growupcr.com/post/metricas-precision>
- [48] Microsoft, «Evaluación de los resultados de los experimentos de aprendizaje automático automatizado - Azure Machine Learning». Accedido: 2 de noviembre de 2025. [En línea]. Disponible en: <https://learn.microsoft.com/es-es/azure/machine-learning/how-to-understand-automated-ml?view=azureml-api-2>
- [49] SonarCloud, «Understanding measures and metrics | SonarQube Cloud | Sonar Documentation». Accedido: 5 de noviembre de 2025. [En línea]. Disponible en: <https://docs.sonarsource.com/sonarqube-cloud/digging-deeper/metric-definitions>
- [50] A. Navarro, «10 Métricas de Calidad del Software para Mejorar la Eficiencia». Accedido: 20 de octubre de 2024. [En línea]. Disponible en: <https://blog.innevo.com/metricas-de-calidad-del-software>

Anexos

**ANEXO N°01 CUESTIONARIO PARA IDENTIFICAR LA PROBLEMÁTICA EN LA
EMPRESA HECTOR OCAÑA IZQUIERDO****CUESTIONARIO PARA IDENTIFICAR LA PROBLEMÁTICA EN UNA
EMPRESA DE TRANSPORTES DE CARGAS EN GENERAL**

1. ¿Cuál es el porcentaje exacto de entregas retrasadas que experimenta la empresa cada mes?
Entre el 6-7% de las entregas mensuales.
2. ¿Cuál es la principal causa de los retrasos en las entregas?
Planificación ineficiente de rutas y falta de predicción de la demanda.
3. ¿Cuáles son los impactos directos de la falta de un sistema de predicción en la operatividad de la empresa?
Sobreutilización de recursos en picos de demanda inesperados y subutilización en periodos de baja demanda.
4. ¿Qué aumento porcentual en los costos de combustible y mantenimiento se ha observado debido a la ineficiencia en la planificación?
un aumento del 10-15% en costos de combustible y mantenimiento.
5. ¿Cuánto se reduce la capacidad operativa debido al tiempo de inactividad de los vehículos?
La capacidad operativa se reduce en un 8-10% debido al tiempo muerto generado por una planificación ineficiente.
6. ¿Qué porcentaje de quejas de clientes ha aumentado debido a las entregas tardías en los últimos seis meses?
Las quejas de clientes han aumentado en un 5% en los últimos seis meses debido a entregas tardías.
7. ¿Qué factores externos, como el tráfico o el clima, no se están considerando actualmente en la planificación logística?
No se considera el tráfico en tiempo real ni las condiciones meteorológicas.

8. ¿Cuál es el método actual de planificación logística que utiliza la empresa?

Se basa en la experiencia previa y reportes manuales sin el uso de tecnologías inteligentes.

9. ¿Qué impacto tiene la falta de predicción precisa de la demanda en la utilización de los vehículos?

Genera un uso ineficiente de los vehículos, con sobrecarga durante picos de demandas y subutilización en otros momentos.

10. ¿Cómo se gestionan actualmente los picos de demanda en la empresa?

Los picos de demanda se gestionan de manera manual, asignando más vehículos y personal de acuerdo con la experiencia previa, lo que a menudo resulta en sobrecarga.

11. ¿Qué porcentaje de las rutas planificadas manualmente resultan ser ineficientes o subóptimas?

Aproximadamente un 20-25% de las rutas planificadas manualmente no son óptimas, lo que ocasiona tiempos de transporte más largos, mayor consumo de combustibles.

12. ¿Con qué frecuencia la empresa revisa y ajusta las rutas para mejorar la eficiencia operativa?

Las rutas se revisan de manera esporádica, principalmente cuando se reportan problemas o retrasos importantes, pero no hay un proceso formal ni automatizado para ajustar las rutas.

13. ¿Qué tecnologías o herramientas logísticas utiliza la empresa para monitorear la flota en tiempo real?

La falta de un sistema automatizado impide que la empresa ajuste sus operaciones logísticas, lo que resulta en retrasos en la entrega, sobrecarga de recursos en picos de demanda.

14. ¿Qué beneficios esperados puede traer la implementación de un sistema inteligente de predicción de demanda y optimización de rutas?

Mejora la eficiencia logística, reducción de costos operativos, optimización en el uso de recursos y mayor satisfacción del cliente.


Christian Paúl Bellasmil Farroñan


Héctor Edgardo Ocaña Izquierdo

**ANEXO N°02 CARTA DE CONFORMIDAD POR PARTE DE LA EMPRESA QUE
GARANTIZA LA APLICACIÓN DEL PRODUCTO ACREDITABLE**



Hector Edgardo Ocaña Izquierdo
TRANSPORTISTA

RUC, 10167028140

EMPRESA DE TRANSPORTE DE CARGAS EN GENERAL, HECTOR OCAÑA IZQUIERDO
AV. VÍCTOR A. BELAÚNDE 458, URB. INGENIERO, CHICLAYO, CHICLAYO
TLF, 979998626

Señor: Ing. Huilder Juanito Mera Montenegro

Director Académico: ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN – USAT

Asunto: CONSTANCIA DE REALIZACIÓN DE PROYECTO DE TESIS

Yo, **Héctor Edgardo Ocaña Izquierdo**, identificado con **DNI N° 16702814**, en calidad de representante legal de la Empresa de Autotransporte de Cargas en General "**Héctor Ocaña Izquierdo**", con **RUC N° 10167028140** y domicilio fiscal en Av. Víctor A. Belaúnde 458, Urb. Ingeniero, Chiclayo, me dirijo a usted con el debido respeto para expresarle lo siguiente:

1. El estudiante Christian Paúl Bellasmil Farroñan, identificado con DNI N° 72469200 y perteneciente a la Escuela de Ingeniería de Sistemas y Computación de la Universidad Católica Santo Toribio de Mogrovejo, ha llevado a cabo el desarrollo completo de su proyecto de tesis en nuestra empresa. Durante su período en nuestra institución, ha aplicado de manera efectiva los conocimientos adquiridos en su formación académica, lo que ha generado un impacto positivo en nuestra organización.
2. Durante su permanencia, el estudiante Bellasmil Farroñan ha ejecutado satisfactoriamente las actividades relacionadas con su investigación aplicada, demostrando un alto nivel de responsabilidad, compromiso y aportando soluciones valiosas a la mejora de nuestros procesos operativos.

Por lo tanto, hacemos constar que el desarrollo de su proyecto de tesis se ha llevado a cabo íntegramente en nuestra empresa. Quedo a su disposición para cualquier consulta adicional que considere pertinente respecto al trabajo realizado.

Atentamente;



Hector Edgardo Ocaña Izquierdo
RUC. 10167028140

Héctor Edgardo Ocaña Izquierdo