

**UNIVERSIDAD TÉCNICA FEDERICO SANTA MARIA
DEPARTAMENTO DE INFORMÁTICA
SANTIAGO - CHILE**



**“PROTOTIPO DE SISTEMA INTELIGENTE PARA LA GESTIÓN DE
INVENTARIO MEDIANTE MODELOS DE MACHINE LEARNING E
IOT”**

DANIEL ELIAS HERRERA JIMENEZ

**MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA**

**PROFESOR GUÍA: MAURICIO OLIVARES FAÚNDEZ
PROFESOR CORREFERENTE: JOSE LUIS MARTÍ LARA**

DICIEMBRE - 2025



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: **PROTOTIPO DE SISTEMA INTELIGENTE PARA LA GESTIÓN DE INVENTARIO MEDIANTE MODELOS DE MACHINE LEARNING E IOT**

Nombre del candidato(a): **DANIEL ELIAS HERRERA JIMENEZ**

Carrera / Grado: **INGENIERÍA CIVIL INFORMÁTICA**

Campus: **SAN JOAQUÍN** Departamento: **INFORMÁTICA**

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, MAURICIO OLIVARES FAÚNDEZ, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 28 diciembre 2025

Firma:

Estudiante o Candidato(a):

Fecha: 28 de diciembre de 2025

Firma:

DEDICATORIA

A mi Padre, que si bien no fue perfecto siempre supo que decir para mantenerme en el camino correcto, sin sus palabras y su apoyo es probable que este trabajo nunca hubiera llegado siquiera a ser escrito. Gracias por todo papá, esta victoria también es tuya.

“Quiero que me prometas que lucharás hasta tu último aliento para ser feliz”

Agradecimientos

Incluso a la hora de escribir estas palabras no puedo creer realmente que esté tan cerca de por fin terminar esta carrera, la misma que me dio años maravillosos y momentos de profunda introspección cuyo valor emocional trasciende lo económico.

Un sueño de niñez será cumplido y por eso mismo creo que es necesario darle las gracias a todas las personas que siempre estuvieron ahí para mí, incluso sin ellos saber que confiaba en ellos con todo mi corazón.

Quiero agradecer a mi familia, sin su apoyo esto no habría sido posible, Juan, mi difunto padre, Cynthia mi madre, mis hermanas Catalina y Matilda, y Matías, mi padrastro. Todo este proceso ha sido muy difícil para todos, pero su cariño y apoyo incondicional impulsa mi búsqueda de superación constante.

A mi pareja, Yazmin, la cual desde el momento en el que la conocí ha sido mi refugio en la adversidad, la parte de mi día que me ilumina y que me hace creer que un futuro mejor para ambos no solo es posible, sino que es alcanzable.

A mis amigos y compañeros, que incluso en los peores momentos siempre pude confiar en que los tendría a mi lado, tanto los que conocí en este camino como son Martín, Mauricio, Sarubbi, Octavio, Nicolás, Andrés, Joaquín, Javiera, como los que estuvieron desde antes, Odilo, Lucas, Eric, Ignacio y a quienes, por brevedad, no menciono explícitamente, pero cuyo afecto atesoro, gracias a todos, su compañía fue determinante para culminar esta etapa.

A los que me dieron esta oportunidad como lo son Franco Astorga y Nicolás Rodríguez, que han sido parte esencial de la escritura de esta memoria y son la razón por la cual está terminada.

Y a los que deje sin nombrar, muchísimas gracias a ustedes también.

Los errores u omisiones presentes en este documento son de mi exclusiva responsabilidad.

Resumen

En el área de Mantenimiento de Operaciones de la empresa, la gestión de inventario realizada mediante el sistema SAP y controles manuales ha generado problemáticas críticas, como descuadres de stock y demoras en la reposición. Para abordar estas ineficiencias, esta memoria presenta el desarrollo y validación de un prototipo de sistema inteligente que integra trazabilidad mediante códigos QR como solución IoT, un motor de predicción de demanda y dashboards interactivos. La metodología incluyó una evaluación experimental de tipo "todos contra todos" utilizando cuatro modelos predictivos: ARIMA, Croston/TSB, Random Forest y XGBoost, aplicados sobre un conjunto de datos históricos de 230 materiales. A nivel funcional, el prototipo desarrollado en Python logró automatizar la captura de datos y redujo el tiempo de registro en aproximadamente un 80%, disminuyendo la operación de minutos a unos 7 segundos por ítem. En cuanto al desempeño predictivo, los resultados contradijeron la expectativa inicial sobre la superioridad de los algoritmos de Machine Learning complejos. El modelo estadístico ARIMA resultó ser el más robusto y preciso, pues superó significativamente a las alternativas en el 46.5% de los casos y demostró ser la opción idónea para gestionar la demanda intermitente que caracteriza al 97% del inventario MRO analizado. El trabajo concluye demostrando la viabilidad técnica de la solución y su potencial para disminuir los quiebres de stock y optimizar el capital inmovilizado mediante decisiones basadas en datos.

Keywords– Mantenimiento, Sistema inteligente, Machine Learning, IoT, ARIMA, Demanda Intermitente.

Abstract

In the company's Operations Maintenance area, inventory management carried out through the SAP system and manual controls has led to critical issues, such as stock discrepancies and replenishment delays. To address these inefficiencies, this thesis presents the development and validation of an intelligent system prototype that integrates traceability via QR codes as an IoT solution, a demand prediction engine, and interactive dashboards. The methodology included an "all-against-all" experimental evaluation of four predictive models, specifically ARIMA, Croston/TSB, Random Forest, and XGBoost, on a historical dataset of 230 materials. Functionally, the Python-developed prototype successfully automated data capture and reduced registration time by approximately 80%, dropping from minutes to about 7 seconds per item. Regarding predictive performance, the results contradicted the initial expectation regarding the superiority of complex Machine Learning algorithms. The statistical model ARIMA proved to be the most robust and accurate, as it significantly outperformed alternatives in 46.5% of cases and proved to be the ideal choice for managing the intermittent demand that characterizes 97% of the analyzed MRO inventory. The work concludes by demonstrating the technical feasibility of the solution and its potential to reduce stockouts and optimize immobilized capital through data-driven decisions.

Keywords– Maintenance, Intelligent System, Machine Learning, IoT, ARIMA, Intermittent Demand.

Glosario de Abreviaturas

ABC: Análisis ABC. Metodología de clasificación de inventarios que categoriza los ítems según su valor de consumo anual (los ítems “A” son los más valiosos, “B” intermedios, y “C” los de menor valor).

ACF: Autocorrelation Function (Función de Autocorrelación). Herramienta estadística utilizada para identificar patrones de estacionalidad o ciclicidad en una serie de tiempo.

ARIMA: AutoRegressive Integrated Moving Average (Modelo Autorregresivo Integrado de Media Móvil). Modelo estadístico clásico para el análisis y predicción de series temporales.

BAPI: Business Application Programming Interface. Interfaz de programación estándar de SAP que permite a aplicaciones externas interactuar con los procesos de negocio del sistema ERP.

CEP: Complex Event Processing (Procesamiento de Eventos Complejos). Tecnología que consiste en combinar datos de múltiples fuentes para inferir eventos o patrones complejos en tiempo real.

CPS: Cyber-Physical System (Sistema Ciberfísico). Sistema que integra componentes computacionales con procesos físicos, donde los algoritmos monitorean y controlan mecanismos del mundo real.

CV: Coeficiente de Variación. Métrica estadística (desviación estándar dividida por la media) utilizada para medir la variabilidad relativa de una serie de datos.

DM: Test de Diebold-Mariano. Contraste estadístico utilizado para comparar la precisión predictiva de dos modelos de pronóstico.

EOQ: Economic Order Quantity (Cantidad Económica de Pedido). Modelo de optimización de inventario que determina la cantidad de pedido que minimiza los costos totales.

ERP: Enterprise Resource Planning (Planificación de Recursos Empresariales). Sistema de software de gestión que integra los principales procesos de negocio de una organización. SAP es el ejemplo más prominente en el documento.

HH: Horas-Hombre. Unidad de medida de trabajo que representa el trabajo realizado por una persona en una hora.

IoT: Internet of Things (Internet de las Cosas). Red de objetos físicos ("cosas") que incorporan sensores, software y otras tecnologías para conectarse e intercambiar datos con otros dispositivos y sistemas a través de internet.

KPI: Key Performance Indicator (Indicador Clave de Desempeño). Métrica cuantificable utilizada para evaluar el éxito de una organización o de una actividad particular.

LSTM: Long Short-Term Memory (Memoria a Corto y Largo Plazo). Un tipo de arquitectura de Red Neuronal Recurrente (RNN) avanzada, utilizada en aprendizaje profundo para el análisis de secuencias de datos.

MAE: Mean Absolute Error (Error Medio Absoluto). Métrica que mide el promedio de los errores absolutos entre las predicciones y los valores reales.

MAPE: Mean Absolute Percentage Error (Error Medio Absoluto Porcentual). Métrica que mide el error de predicción como un porcentaje del valor real.

ML: Machine Learning (Aprendizaje Automático). Campo de la inteligencia artificial que utiliza algoritmos para permitir que los sistemas aprendan y mejoren a partir de la experiencia (datos).

MRO: Maintenance, Repair and Operations (Mantenimiento, Reparación y Operaciones). Se refiere al inventario de repuestos, insumos y materiales necesarios para mantener el funcionamiento de los activos de una empresa, pero que no forman parte del producto final.

PDF: Portable Document Format. Formato de archivo universal para presentar documentos.

PWA: Progressive Web App (Aplicación Web Progresiva). Tipo de aplicación de software entregada a través de la web, construida usando tecnologías web comunes.

QR: Quick Response Code (Código de Respuesta Rápida). Código de barras bidimensional que almacena información y puede ser leído rápidamente por un dispositivo de imagen.

R²: Coeficiente de Determinación. Métrica estadística que indica la proporción de la varianza en la variable dependiente que es predecible a partir de las variables independientes en un modelo.

RFC: Remote Function Call. Protocolo de SAP que permite a un sistema externo ejecutar una función o módulo en un sistema SAP.

RFID: Radio Frequency Identification (Identificación por Radiofrecuencia). Tecnología que utiliza campos electromagnéticos para identificar y rastrear automáticamente etiquetas adheridas a objetos.

RMSE: Root Mean Square Error (Raíz del Error Cuadrático Medio). Métrica que mide la raíz cuadrada del promedio de los errores al cuadrado, penalizando más los errores grandes.

RNN: Recurrent Neural Network (Red Neuronal Recurrente). Clase de red neuronal artificial donde las conexiones entre nodos forman un grafo dirigido a lo largo de una secuencia temporal.

ROP: Reorder Point (Punto de Reorden). Nivel de inventario que activa una acción para reponer ese ítem en particular.

SAP: Sistemas, Aplicaciones y Productos en Procesamiento de Datos. Nombre de la empresa y de su software ERP, ampliamente utilizado en grandes corporaciones y lo que usamos en este documento para conseguir la información.

SBA: Syntetos-Boylan Approximation. Una variante refinada del método de Croston para la predicción de demanda intermitente.

SVM: Support Vector Machine (Máquina de Vectores de Soporte). Algoritmo de aprendizaje supervisado utilizado para clasificación y regresión.

TSB: Método de Teunter-Syntetos-Babai. Una variante del método de Croston que también modela la probabilidad de que haya demanda en un período.

UI: User Interface (Interfaz de Usuario). El medio a través del cual un usuario interactúa con una máquina, software o aplicación.

XYZ: Análisis XYZ. Metodología de clasificación de inventarios que categoriza los ítems según la regularidad o variabilidad de su demanda.

ÍNDICE

Agradecimientos	II
Resumen	III
Abstract	IV
Glosario de Abreviaturas.....	V
Introducción	1
Objetivos	2
Capítulo 1. Marco Teórico y Estado del Arte	3
1.1. Fundamentos de la gestión de inventario	3
1.2. Evolución de los modelos de predicción en gestión de inventario.....	3
1.3. Machine Learning aplicada al inventario MRO.....	4
1.4. Modelos especializados para demanda intermitente.....	5
1.5. Análisis de criticidad mediante clustering.....	6
1.6. IoT y sensorización para trazabilidad automática	6
1.7. Procesamiento de eventos complejos (CEP).....	7
1.8. Brechas entre disponibilidad tecnológica y adopción efectiva.....	8
1.9. Situación actual de la empresa.....	8
1.10. Síntesis crítica y aporte del proyecto	9
Capítulo 2. Metodología.....	12
2.1. Indicadores Clave de Desempeño (KPI).....	12
2.1.1. KPI 1 - Tiempo total de gestión de inventario.....	12
2.1.2. KPI 2 - Horas Hombre (HH) utilizadas en cuadraturas y ajustes	12
2.1.3. KPI 3 -Días sin servicio por falta de stock	13
2.1.4. KPI 4 - Disponibilidad de stock crítico	13
2.1.5. KPI 5 -Desviación del stock promedio respecto del stock óptimo.....	13
2.1.6. Metas Esperadas	14
2.2. Trazabilidad con códigos QR como solución IoT accesible.....	15
2.3. Modelos predictivos utilizados	16
2.3.1. ARIMA (AutoRegressive Integrated Moving Average).....	18
2.3.2. Croston y variantes (TSB, SBA).....	18
2.3.3. Random Forest.....	19
2.3.4. XGBoost (Extreme Gradient Boosting).....	19
2.3.5. Metodología de comparación de modelos.....	20
2.4. Métricas de evaluación y funciones de pérdida.....	21
2.4.1. Error Medio Absoluto (MAE)	22

2.4.2.	Error Medio Absoluto Porcentual (MAPE)	22
2.4.3.	Raíz del Error Cuadrático Medio (RMSE)	23
2.4.4.	Coeficiente de Determinación (R²)	23
2.4.5.	Estrategia de uso de las métricas en la metodología	23
2.5.	Segmentación analítica de ítems	24
2.5.1.	Justificación	24
2.5.2.	Criterios de segmentación	24
2.5.3.	Procedimiento	25
2.5.4.	Valor esperado del análisis	25
2.6.	Estrategia de comparación global y selección del modelo único	27
2.6.1.	Diseño de evaluación “todos contra todos” (una sola elección final)	27
2.6.2.	Métricas de desempeño y criterios de evaluación	27
2.6.3.	Contrastes estadísticos entre algoritmos	28
2.6.4.	Robustez práctica	30
2.6.5.	Procedimiento paso a paso	31
2.6.6.	Plantilla de resultados (resumen global)	33
Capítulo 3.	Implementación	34
3.1.	Módulo de Digitalización mediante Códigos QR	34
3.1.1.	Arquitectura general del módulo	34
3.1.2.	Funcionalidades principales	35
3.1.3.	Estructura de datos y Flujo Operativo	35
3.1.4.	Desarrollo e implementación técnica	38
3.2.	Módulo de algoritmos de Machine Learning	39
3.2.1.	Arquitectura general del módulo	39
3.2.2.	Flujo de Ejecución del módulo	41
3.2.3.	Tecnologías utilizadas	42
3.2.4.	Implementación funcional	43
3.2.5.	Modelos disponibles e implementación	44
3.2.6.	Evaluación comparativa y registro de resultados	45
3.3.	Módulo de Visualizaciones Dinámicas y Dashboards	47
3.3.1.	Arquitectura general y componentes	47
3.3.2.	Paneles Principales de Análisis	48
3.4.	Integración de módulos	50
3.4.1.	Arquitectura de integración y flujo de datos	50
3.4.2.	Flujo integrado de operación	51
3.4.3.	Sincronización y consistencia de datos	53

Capítulo 4.	Resultados.....	55
4.1.	Evaluación funcional y de integración.....	55
4.2.	Evaluación del Desempeño predictivo	56
4.2.1.	Comparativa Global de Modelos	56
4.2.2.	Análisis de Rendimiento por Segmento	57
4.2.3.	Contrastes Estadísticos	58
4.3.	Proyección del Impacto sobre los KPIs Operacionales	59
Capítulo 5.	Conclusiones	61
5.1.	Cumplimiento de los Objetivos del Proyecto.....	61
5.2.	Hallazgo Principal: Idoneidad del Modelo ARIMA para Demanda Intermitente	61
5.3.	Validación del Impacto Operacional y Aporte del Prototipo.....	62
5.4.	Líneas de Trabajo Futuro	63
Bibliografía		64

ÍNDICE DE FIGURAS

Figura 1: Mapa de Procesos actual.....	9
Figura 2: Patrones de demanda MRO.....	26
Figura 3: Flujo del módulo de digitalización.....	37
Figura 4: Pantalla de Inicio Entrada de Materiales	37
Figura 5: Pantalla de Inicio Salida de Materiales.....	38
Figura 6: Diagrama de la arquitectura del Módulo.....	40
Figura 7: Diagrama de Flujo Módulo de Machine Learning	42
Figura 8: Pantalla de Inicio Machine Learning.....	46
Figura 9: Pantalla de Machine Learning, Evaluando Buna Love Joy 095	46
Figura 10: Pantalla de Machine Learning en masa, Comparando Resultados	47
Figura 11: Pantalla de Inicio Dashboard y Métricas de la aplicación	49
Figura 12: Pantalla de Análisis de Valores, Censurando Datos Sensibles	50
Figura 13: Diagrama Flujo 1.....	52
Figura 14: Diagrama Flujo 2.....	53
Figura 15: Nuevo Mapa de procesos	56

ÍNDICE DE TABLAS

Tabla 1: Comparación de KPI.	15
Tabla 2: Clasificación de ítems según patrón de consumo.....	26
Tabla 3: Plantilla de resultados comparativos para modelo óptimo de pronóstico	33
Tabla 4: Funcionalidades del módulo de digitalización.....	35
Tabla 5: Tecnologías utilizadas por componente en Módulo de Digitalización	38
Tabla 6: Librerías utilizadas en módulo ML	43
Tabla 7: Resultados Evaluación Funcional.....	55
Tabla 8: Resumen global de modelos ganadores sobre 230 ítems evaluados	57
Tabla 9: Análisis de rendimiento por segmento de demanda	58
Tabla 10: Resultados de los contrastes estadísticos pareados	58

Introducción

La gestión de inventario es un área fundamental para el correcto funcionamiento de cualquier empresa. Sin embargo, a pesar de su relevancia, suele quedar relegada frente a otros procesos considerados más estratégicos o visibles en la adopción de tecnologías de vanguardia. Esto resulta particularmente crítico en áreas como mantenimiento, bodega y abastecimiento, donde mantener la continuidad operativa depende en gran medida de contar con insumos disponibles en el momento adecuado.

La falta de automatización en los procesos de inventario puede generar consecuencias graves, desde la falta temporal de stock hasta la detención total de operaciones por varios días. Estas interrupciones pueden derivar en pérdidas económicas significativas, que fácilmente pueden alcanzar miles o incluso millones de pesos si no se gestionan con la seriedad que ameritan.

Este trabajo propone el desarrollo de un prototipo que incorpore tecnologías emergentes como el uso de códigos QR (en adelante, etiquetas IoT), algoritmos de Machine Learning, entre ellos Extreme Gradient Boosting, Random Forest, ARIMA y Croston, y dashboards interactivos que permitan visualizar datos relevantes para la toma de decisiones. El objetivo es contribuir a una gestión de inventario más eficiente, precisa y adaptable a contextos dinámicos como el de las operaciones industriales.

La estructura del documento es la siguiente:

- El **Capítulo 1** presenta el marco teórico del problema y el estado del arte.
- El **Capítulo 2** describe la metodología empleada en el desarrollo del sistema propuesto, las tecnologías y los subsistemas a utilizar.
- El **Capítulo 3** detalla la implementación de los sistemas planteados.
- El **Capítulo 4** expone los resultados obtenidos.
- El **Capítulo 5** incluye las conclusiones del trabajo y posibles trabajos a futuro.

Se espera que esta memoria no solo aporte valor a la gestión de inventario de la empresa en cuestión, sino también al desarrollo de soluciones escalables que reconozcan la importancia estratégica del inventario en los procesos de mantenimiento industrial.

Objetivos

El objetivo principal de este proyecto es desarrollar un prototipo de sistema de inventario inteligente orientado al área de Mantenimiento de Operaciones de la empresa, con el fin de optimizar y automatizar los procesos actualmente utilizados. Para ello, se propone la incorporación de tecnologías emergentes como etiquetas IoT (basadas en códigos QR), modelos de Machine Learning (incluyendo ARIMA, regresión lineal, redes neuronales, entre otros) y dashboards interactivos que permitan una visualización dinámica y accesible de la información crítica del inventario.

Con el propósito de abordar de manera integral esta problemática, se han definido los siguientes objetivos específicos:

En primer lugar, se busca identificar los procesos críticos dentro del flujo operativo que conecta las áreas de Mantenimiento, Bodega y Abastecimiento, con el fin de comprender sus interdependencias y establecer metas concretas de mejora. Esta etapa permitirá enfocar los esfuerzos del prototipo en puntos de alto impacto operacional.

Posteriormente, se procederá a definir indicadores clave de desempeño (KPIs) que permitan medir objetivamente el impacto de las mejoras implementadas. Estos indicadores se relacionarán directamente con variables de costo, eficiencia operativa y continuidad del servicio.

Capítulo 1. Marco Teórico y Estado del Arte

1.1. Fundamentos de la gestión de inventario

La gestión de inventario constituye uno de los pilares esenciales de la operación en organizaciones productivas. Su función no se limita a garantizar la disponibilidad de insumos y repuestos necesarios para la continuidad de los procesos, sino que también influye directamente en la eficiencia económica, la sostenibilidad logística y la satisfacción de los usuarios internos. En sectores industriales de alta criticidad, una interrupción no planificada puede implicar pérdidas millonarias por hora, lo que hace indispensable contar con un sistema de inventario preciso, trazable y proactivo (Chen et al., 2019).

A pesar de su carácter estratégico, la gestión de inventario ha sido históricamente relegada frente a áreas percibidas como más críticas o visibles. En consecuencia, muchas organizaciones todavía dependen de métodos manuales, sistemas desconectados y prácticas reactivas que dificultan la trazabilidad, incrementan la carga administrativa y generan errores sistemáticos en los registros (Siponen et al., 2019). Incluso en entornos donde se utilizan sistemas ERP como SAP, persisten brechas operativas: por ejemplo, el ingreso de materiales sin código o la omisión del registro de salida en situaciones de urgencia. Estas omisiones producen vacíos de información que afectan tanto la toma de decisiones como la planificación del abastecimiento (Wamba et al., 2020).

Las consecuencias de una gestión ineficiente se reflejan en fenómenos como el sobre stock, la obsolescencia de materiales y, más críticamente, los quiebres de inventario que pueden comprometer operaciones completas (Mittal et al., 2018). Asimismo, las cuadraturas manuales necesarias para corregir inconsistencias representan una pérdida considerable de horas hombre, recursos que podrían destinarse a actividades de mayor valor agregado. Según Ghobakhloo y Fathi (2019), esta situación suele estar asociada a una limitada priorización estratégica y a la baja digitalización de procesos de soporte como mantenimiento y logística interna.

En síntesis, la gestión de inventario no es un proceso administrativo secundario, sino un componente estratégico que impacta en la resiliencia, la competitividad y la capacidad de adaptación de las organizaciones industriales. La literatura reciente enfatiza la necesidad de migrar desde esquemas manuales y reactivos hacia modelos más integrados, apoyados en digitalización y analítica avanzada (Praveen, 2020; Chen et al., 2019).

1.2. Evolución de los modelos de predicción en gestión de inventario

La predicción de demanda ha sido históricamente un desafío central en la gestión de inventario. Los enfoques iniciales se apoyaron en métodos estadísticos clásicos como el promedio móvil, el suavizamiento exponencial y los modelos autorregresivos integrados de media móvil (ARIMA). Estas metodologías parten del supuesto de patrones temporales relativamente estables, lo que resulta adecuado para productos con consumo regular y predecible (Pinçe et al., 2021). No obstante, en entornos de mantenimiento, reparación y operaciones (MRO), la demanda de repuestos suele ser intermitente, altamente variable y condicionada por factores exógenos como fallas técnicas, condiciones ambientales y dinámicas operativas (Siponen et al., 2019).

En escenarios industriales reales, estas limitaciones generan desviaciones significativas en las proyecciones, lo cual compromete la planificación del abastecimiento y eleva los costos asociados al sobre stock o a quiebres de inventario. Aktepe et al. (2023) evidencian que los modelos tradicionales presentan errores considerables al predecir la demanda de repuestos de flotas de transporte público, especialmente en situaciones de alta irregularidad. De manera similar, Chen et al. (2019) destacan que, si bien los métodos estadísticos siguen siendo relevantes por su simplicidad y bajo costo computacional, resultan insuficientes frente a la complejidad de los sistemas productivos actuales.

La incorporación de algoritmos de Machine Learning (ML) ha permitido superar estas limitaciones, gracias a su capacidad de capturar relaciones multivariantes y patrones no lineales. Técnicas como Random Forest, XGBoost y Support Vector Machines (SVM), así como arquitecturas de aprendizaje profundo como DeepAR, han demostrado un desempeño superior al de los modelos clásicos, reduciendo significativamente el error de predicción en demandas complejas e intermitentes (Salinas et al., 2017; Vidal et al., 2022; Ahmadov & Helo, 2023). Estos enfoques no solo mejoran la precisión, sino que también permiten incorporar variables contextuales como condiciones operativas, criticidad de los equipos y estacionalidad de fallas, aspectos claves en entornos MRO (Ifraz et al., 2023).

La transición desde modelos estadísticos tradicionales hacia algoritmos de ML refleja, por tanto, una evolución necesaria para enfrentar la incertidumbre y complejidad de la gestión de inventario industrial. Sin embargo, esta migración no debe entenderse como un reemplazo total, sino como una ampliación del repertorio metodológico: los modelos clásicos aún conservan valor en escenarios con baja variabilidad, mientras que los enfoques basados en ML resultan más adecuados para contextos dinámicos y críticos (Fiorentina et al., 2021).

1.3. Machine Learning aplicada al inventario MRO

La naturaleza de la demanda en inventarios de mantenimiento, reparación y operaciones (MRO) se caracteriza por su intermitencia, irregularidad y dependencia de factores contextuales, lo que dificulta su predicción mediante métodos estadísticos tradicionales. Frente a esta limitación, los algoritmos de *Machine Learning* han surgido como una alternativa prometedora, al permitir modelar relaciones no lineales y aprovechar múltiples variables explicativas (Chen et al., 2019; Pinçe et al., 2021).

Diversos estudios han demostrado la efectividad de estas técnicas en entornos industriales. Aktepe et al. (2023), por ejemplo, aplicaron modelos de ML al consumo de repuestos en flotas de transporte urbano, incorporando variables como el tiempo medio entre fallas (MTBF), kilometraje acumulado y frecuencia de mantenimientos. Sus resultados evidenciaron que algoritmos basados en árboles de decisión, como Random Forest y XGBoost, superan significativamente a los enfoques clásicos (ARIMA, suavizamiento exponencial) en términos de precisión. De manera similar, Vidal et al. (2022) evaluaron Support Vector Machines (SVM) y Random Forest en contextos de alta intermitencia, destacando que los modelos de ML alcanzaron reducciones notables en el error de predicción tras aplicar procesos de limpieza y normalización de datos.

El aprendizaje profundo también ha mostrado avances en esta área. Salinas et al. (2017) introdujeron el modelo DeepAR, capaz de generar pronósticos probabilísticos en múltiples series temporales relacionadas, lo cual resulta útil para ítems con baja frecuencia de

consumo, pero con correlaciones entre categorías. Ahmadov y Helo (2023) confirmaron el potencial de estas arquitecturas al aplicarlas en demandas intermitentes de comercio electrónico, encontrando mejoras sustanciales respecto a los modelos convencionales.

En términos aplicados, los enfoques de ML no solo incrementan la exactitud en la estimación del consumo de repuestos, sino que además habilitan estrategias de gestión más avanzadas, como la reposición automática, la priorización de ítems críticos y la planificación dinámica de compras. Sin embargo, la literatura coincide en señalar que su adopción depende de factores organizacionales clave: la disponibilidad de datos históricos de calidad, la interoperabilidad con sistemas ERP como SAP, y la capacidad de los usuarios para interpretar y confiar en los resultados (Ifraz et al., 2023; Wamba et al., 2020).

A modo de conclusión, la incorporación de Machine Learning en inventarios MRO representa un avance significativo respecto de los modelos estadísticos tradicionales, al ofrecer predicciones más robustas en escenarios de alta complejidad y variabilidad. No obstante, su efectividad está condicionada tanto por la madurez digital de las organizaciones como por la integración tecnológica que permita transformar los resultados analíticos en decisiones operativas concretas.

1.4. Modelos especializados para demanda intermitente

Uno de los principales retos en la gestión de inventarios MRO es la predicción de ítems cuya demanda se manifiesta de manera intermitente o esporádica. A diferencia de productos de consumo masivo con patrones regulares, muchos repuestos se solicitan solo en caso de fallas o mantenimientos específicos, lo que genera series temporales con largos períodos de inactividad seguidos de picos de consumo (Şahin et al., 2013). En este contexto, los modelos tradicionales como ARIMA o suavizamiento exponencial tienden a mostrar bajo desempeño, pues asumen continuidad y estabilidad en las series (Pinçe et al., 2021).

Para abordar este problema, Croston (1972) propuso un modelo específico que separa la estimación de dos componentes: la frecuencia entre demandas y la magnitud de la demanda cuando ocurre. Esta metodología ha demostrado ser más adecuada para ítems con consumos esporádicos que los métodos convencionales de series temporales (Şahin et al., 2013). Posteriormente, se han desarrollado variantes como el Syntetos–Boylan Approximation (SBA) y el Teunter–Syntetos–Babai (TSB), que buscan corregir sesgos en la estimación y mejorar la capacidad de adaptación frente a cambios en los patrones de demanda (Hasani, 2023).

La literatura reciente ha comparado estos enfoques con modelos de *Machine Learning*. Ifraz et al. (2023) evidenciaron que, al incorporar variables contextuales mediante aprendizaje supervisado, es posible superar las limitaciones de los métodos basados únicamente en series históricas. De manera similar, Vidal et al. (2022) aplicaron Random Forest y SVM a series con alta intermitencia, logrando mejoras notables tras procesos de limpieza de datos y normalización. Estos resultados sugieren que los modelos de Croston y sus variantes continúan siendo una referencia válida como línea base, pero su rendimiento puede ser superado por algoritmos más flexibles cuando se dispone de información contextual suficiente.

En conclusión, los modelos especializados como Croston, SBA y TSB representan un aporte fundamental para el tratamiento de la demanda intermitente en inventarios industriales, al capturar características que los métodos tradicionales pasan por alto. No obstante, la tendencia actual apunta hacia la combinación de enfoques, en la que los métodos clásicos sirven como referencia inicial y los algoritmos de ML permiten explotar variables adicionales, mejorando la capacidad predictiva en entornos de alta variabilidad y complejidad.

1.5. Análisis de criticidad mediante clustering

En la gestión de inventarios industriales no solo resulta relevante predecir cuánto se va a consumir, sino también identificar qué ítems requieren mayor atención estratégica. En este sentido, el análisis de criticidad permite priorizar los esfuerzos de planificación y control, enfocándose en los materiales que generan mayor impacto en la continuidad operativa y en los costos logísticos. Tradicionalmente, este proceso se ha abordado mediante técnicas como el análisis ABC (clasificación según valor de consumo anual) o el XYZ (clasificación según regularidad de la demanda). Sin embargo, estos métodos tienden a ser unidimensionales y no capturan la complejidad multivariable de los contextos industriales (Silver et al., 1998).

Con el desarrollo de nuevas herramientas analíticas, enfoques basados en clustering no supervisado se han posicionado como alternativas más flexibles. Fiorentia et al. (2021) aplicaron algoritmos como *K-Means* y *DBSCAN* para clasificar repuestos navales según su frecuencia de uso y costo asociado, identificando que un reducido grupo de ítems (aproximadamente un 4%) concentraba más del 50% del costo anual. Este hallazgo evidencia la capacidad del *clustering* para descubrir patrones ocultos que van más allá de las heurísticas clásicas de clasificación.

El análisis de criticidad mediante *clustering* también permite incorporar criterios múltiples, como riesgo de quiebre, tiempo de reposición, impacto económico y dependencia funcional con equipos críticos. Al integrar estas dimensiones, las organizaciones pueden construir perfiles de inventario más realistas, facilitando la toma de decisiones sobre niveles de stock, políticas de abastecimiento y planes de contingencia (Chen et al., 2019).

Por consiguiente, el uso de técnicas de clustering aporta un enfoque más robusto y adaptativo para el análisis de criticidad en inventarios MRO. Si bien las metodologías tradicionales como ABC/XYZ siguen siendo útiles por su simplicidad, las técnicas basadas en aprendizaje no supervisado ofrecen una mayor granularidad y precisión, lo que resulta clave para entornos donde los recursos son limitados y la continuidad operacional constituye una prioridad estratégica.

1.6. IoT y sensorización para trazabilidad automática

La trazabilidad del inventario es un aspecto crítico para asegurar la disponibilidad de repuestos y la eficiencia en los procesos logísticos. En este ámbito, las tecnologías de Internet of Things (IoT) han adquirido un papel fundamental al posibilitar la captura de datos en tiempo real sobre ubicación, estado y movimiento de materiales (Taj, 2023). La incorporación de sensores y dispositivos de identificación automática reduce la dependencia de registros manuales, disminuye errores y proporciona mayor visibilidad a lo largo de la cadena de suministro.

Entre estas tecnologías, las etiquetas RFID (Radio Frequency Identification) se han consolidado como una solución robusta en entornos industriales de alta criticidad. Zhang et al. (2012) demostraron la efectividad de un sistema de gestión de mantenimiento basado en RFID combinado con *Complex Event Processing (CEP)*, el cual permitió detectar eventos anómalos y reducir los errores derivados de la intervención manual. Sin embargo, su adopción enfrenta barreras relacionadas con los costos de infraestructura y las limitaciones técnicas en ambientes hostiles, como aquellos con polvo, humedad o interferencias electromagnéticas.

En respuesta a estas dificultades, han surgido alternativas más accesibles como el uso de códigos QR, que ofrecen una solución de bajo costo y fácil implementación. Estudios recientes destacan su aplicación en sistemas móviles y plataformas web progresivas (PWA), lo que permite la trazabilidad en tiempo real sin necesidad de infraestructura especializada (Desai, 2025; Noguerra, 2023). Estos enfoques, en conjunto con dispositivos de bajo costo como microcontroladores y cámaras integradas, han mostrado ser particularmente adecuados en organizaciones con menor madurez digital, al equilibrar simplicidad de uso y efectividad operativa.

En síntesis, la incorporación de tecnologías IoT en la gestión de inventarios contribuye a la creación de sistemas más confiables, transparentes y adaptativos. Mientras que el RFID se mantiene como un estándar en operaciones de gran escala, las soluciones basadas en códigos QR representan un camino de democratización tecnológica, facilitando la trazabilidad en organizaciones con recursos limitados. La literatura coincide en que el verdadero potencial de estas herramientas radica en su integración con sistemas ERP y plataformas analíticas, habilitando la transición hacia una gestión de inventarios alineada con los principios de la Industria 4.0 (Chen et al., 2019; Taj, 2023).

1.7. Procesamiento de eventos complejos (CEP)

Para elevar el nivel de inteligencia operativa en la gestión de stock, el Procesamiento de Eventos Complejos o CEP actúa como una herramienta esencial. Su mecánica consiste en correlacionar flujos de eventos individuales para identificar comportamientos significativos y detectar anomalías en tiempo real. Este proceso permite trascender el mero registro de datos para generar información de valor estratégico (Chen et al., 2019).

En el ámbito industrial, la integración de CEP con tecnologías de sensorización ha mostrado resultados prometedores. Zhang et al. (2012) propusieron un marco de gestión de mantenimiento basado en RFID y CEP, el cual permitió identificar anomalías en el uso de repuestos y detectar desviaciones en los patrones históricos de consumo. Este enfoque demostró que el análisis de secuencias de eventos no solo mejora la capacidad de reacción ante contingencias, sino que también habilita sistemas de alerta temprana y control automatizado.

La aplicación de CEP en inventarios MRO es especialmente relevante para la prevención de quiebres críticos y la detección de usos no autorizados, fenómenos que suelen pasar desapercibidos en sistemas manuales o desconectados. Asimismo, su integración con plataformas analíticas y modelos predictivos de *Machine Learning* amplifica el valor de los datos, al permitir retroalimentar en tiempo real los algoritmos de predicción y ajustar dinámicamente los parámetros de reposición (Wamba et al., 2020).

Concluyendo, el CEP constituye una capa de inteligencia adicional que potencia la trazabilidad y el control del inventario. Si bien su adopción implica desafíos en términos de interoperabilidad tecnológica y capacitación organizacional, la literatura evidencia su potencial para fortalecer la resiliencia operativa y avanzar hacia una gestión de inventarios más predictiva y automatizada, alineada con los principios de la Industria 4.0 (Chen et al., 2019; Zhang et al., 2012).

1.8. Brechas entre disponibilidad tecnológica y adopción efectiva

Existe una notable distancia entre la oferta tecnológica actual y la realidad operativa de muchas organizaciones que, pese a tener acceso a soluciones como IoT y predicción avanzada, persisten en el uso de métodos manuales. La literatura señala que esta falta de adopción efectiva constituye uno de los impedimentos más significativos para lograr la digitalización en los procesos de soporte (Ghobakhloo & Fathi, 2019; Mittal et al., 2018).

Entre los factores que explican este rezago destacan la resistencia al cambio organizacional, la falta de capacitación del personal y la baja priorización estratégica de funciones consideradas “no productivas”, como mantenimiento, logística interna o gestión de repuestos. Según Wamba et al. (2020), muchas empresas concentran sus esfuerzos de digitalización en áreas visibles de la cadena de valor, como ventas o producción, relegando la modernización de los procesos de soporte. Esta desarticulación genera un aprovechamiento parcial de las capacidades tecnológicas disponibles.

La literatura también señala la existencia de limitaciones técnicas y de interoperabilidad, especialmente en contextos donde coexisten sistemas ERP como SAP con registros manuales o aplicaciones desconectadas. Esta falta de integración no solo incrementa los costos operativos, sino que también impide capturar beneficios estratégicos derivados de la gestión basada en datos, como la reducción de quiebres de inventario o la optimización de niveles de stock (Chen et al., 2019; Siponen et al., 2019).

Por consiguiente, la brecha entre capacidad tecnológica instalada y su uso efectivo refleja que la transformación digital en la gestión de inventarios no depende únicamente de la adopción de herramientas, sino de su alineación con los flujos operativos, la cultura organizacional y la capacitación de los actores involucrados. Superar este rezago es un paso fundamental para avanzar hacia sistemas de inventario predictivos, trazables y resilientes, en línea con las exigencias de la Industria 4.0 (Mittal et al., 2018; Wamba et al., 2020).

1.9. Situación actual de la empresa

El análisis realizado en la empresa objeto de estudio evidencia que la gestión de inventarios se encuentra en un estado de baja sistematización y fuerte dependencia de procesos manuales, tal como se esquematiza en el flujo operativo de la Figura 1. Las operaciones de ingreso, salida, solicitudes de reposición y generación de alertas se realizan actualmente sin apoyo de sistemas automatizados, lo que implica un uso intensivo de recursos humanos y tiempos operativos. Como consecuencia, las cuadraturas periódicas entre el inventario físico y los registros administrativos requieren jornadas completas de trabajo, generando una carga adicional para el personal y reduciendo la trazabilidad en tiempo real.

La situación actual de la empresa ejemplifica claramente la brecha tecnológica y organizacional que describe la literatura. De acuerdo con Ghobakhloo y Fathi (2019), los esfuerzos de digitalización suelen privilegiar las áreas centrales del negocio, lo que relega a un segundo plano actividades de soporte como el mantenimiento y la logística interna. Esta realidad se confirma en el caso de estudio, pues a pesar de disponer de infraestructura como el sistema SAP, redes y lectores de códigos, su aplicación en el control de inventarios es prácticamente inexistente, un patrón de subutilización coincidente con lo reportado por Wamba et al. (2020) y Siponen et al. (2019).

El impacto de esta situación no se limita a una mayor carga administrativa, sino que se traduce en ineficiencias críticas: aumento de riesgos de quiebre de stock, generación de sobre inventarios y menor capacidad de reacción frente a contingencias operativas. Además, la ausencia de mecanismos automáticos de trazabilidad restringe la capacidad de incorporar estrategias modernas como la reposición dinámica, el análisis predictivo de consumo o la priorización de ítems críticos (Chen et al., 2019).

Bajo este contexto, la realidad de la empresa ilustra la distancia entre la disponibilidad tecnológica y su aplicación práctica en la gestión de inventarios. Este diagnóstico justifica la necesidad de implementar soluciones adaptadas al contexto organizacional, que integren herramientas accesibles como códigos QR, modelos de predicción y dashboards interactivos, con el objetivo de cerrar la brecha existente y avanzar hacia un sistema de inventario más eficiente, trazable y alineado con los principios de la Industria 4.0.

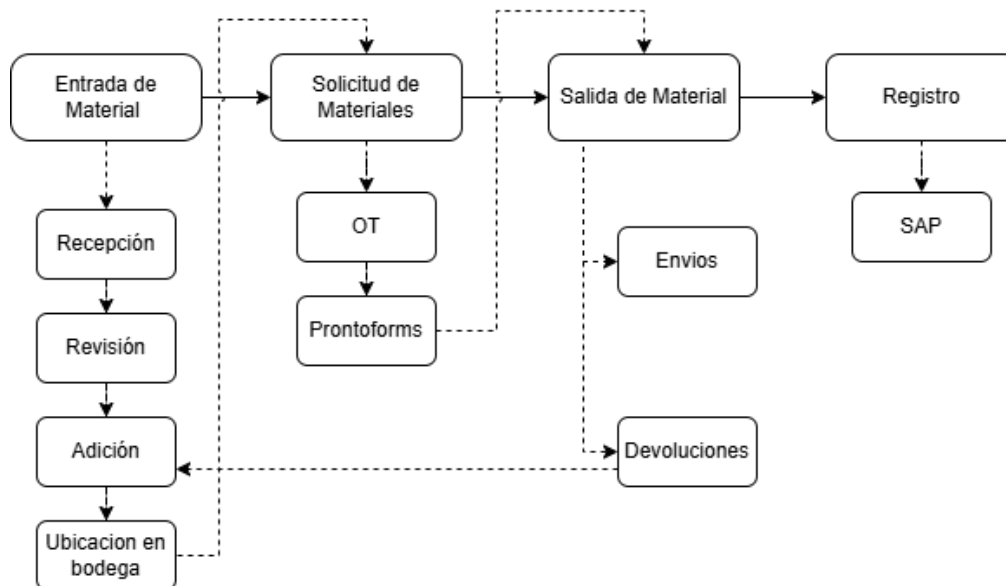


Figura 1: Mapa de Procesos actual
Fuente: Elaboración Propia

1.10. Síntesis crítica y aporte del proyecto

La revisión de la literatura evidencia una evolución progresiva en la gestión de inventarios industriales, pasando de enfoques manuales y estadísticos tradicionales hacia sistemas inteligentes, conectados y predictivos. Este cambio se sustenta en tres pilares fundamentales: la sensorización del inventario mediante tecnologías IoT, el análisis avanzado de datos a

través de modelos de Machine Learning y la visualización interactiva mediante dashboards (Chen et al., 2019; Taj, 2023; Salinas et al., 2017).

Los estudios revisados muestran que, si bien los modelos estadísticos clásicos (como ARIMA o suavizamiento exponencial) siguen teniendo valor en contextos de demanda regular, resultan insuficientes en escenarios con consumos intermitentes o altamente variables. En este ámbito, metodologías especializadas como Croston, SBA y TSB ofrecen un punto de partida adecuado, aunque su desempeño puede ser superado por algoritmos de *Machine Learning* (Random Forest, XGBoost, SVM, DeepAR) cuando se dispone de datos contextuales y registros de mayor calidad (Aktepe et al., 2023; Vidal et al., 2022; Ifraz et al., 2023). Asimismo, investigaciones recientes destacan que técnicas de *clustering* permiten refinar el análisis de criticidad y focalizar recursos en los ítems de mayor impacto económico y operativo (Fiorentia et al., 2021).

Sin embargo, la literatura también subraya la existencia de brechas relevantes entre la disponibilidad tecnológica y su adopción efectiva, asociadas a factores organizacionales, culturales y de interoperabilidad de sistemas (Ghobakhloo & Fathi, 2019; Wamba et al., 2020; Siponen et al., 2019). En el caso de la empresa analizada, esta brecha se manifiesta en una gestión de inventario altamente manual, donde el potencial de tecnologías ya disponibles permanece desaprovechado.

El presente proyecto busca aportar en este espacio, desarrollando un prototipo adaptado al contexto organizacional que combine:

- Códigos QR como solución IoT accesible, capaces de habilitar trazabilidad sin grandes inversiones.
- Modelos predictivos de inventario, que incluyan desde variantes de Croston hasta algoritmos de ML como XGBoost y Random Forest, para estimar demanda en escenarios intermitentes.
- Dashboards interactivos vinculados a SAP, orientados a mejorar la interpretación de datos y la toma de decisiones en tiempo real.
- Mecanismos de alerta temprana y priorización de ítems críticos, que permitan reducir riesgos de quiebre de stock y optimizar el uso de recursos.

Fundamentalmente, el valor de este trabajo reside en su capacidad para vincular la teoría con la práctica operativa. La propuesta no se limita a incorporar tecnologías validadas, sino que las ajusta a las restricciones y requerimientos reales de la organización. Así, se plantea una solución escalable y orientada por datos que, bajo los lineamientos de la Industria 4.0, refuerza tanto la resiliencia como la sostenibilidad del inventario MRO.

Otro de los objetivos es comparar distintos modelos de predicción basados en Machine Learning, evaluando su aplicabilidad en la estimación del consumo y la reposición de inventario. Esta comparación se basará en los datos históricos disponibles en el sistema SAP, considerando tanto su precisión como su capacidad de generalización frente a nuevas condiciones.

Además, se busca contrastar el desempeño de los modelos implementados con los procesos actuales de gestión de inventario, a fin de cuantificar su efectividad y determinar el valor agregado que pueden ofrecer al sistema global.

Adicionalmente, se pretende identificar los procesos que pueden ser automatizados dentro del ciclo de manejo de stock, incorporando mecanismos de detección temprana de quiebres, alertas inteligentes y recomendaciones proactivas de compra basadas en la predicción de la demanda.

Por último, se plantea diseñar y establecer dashboards interactivos que expongan de forma clara y funcional la información clave del inventario, facilitando la toma de decisiones por parte de los distintos usuarios del sistema, ya sea personal técnico, administrativos o jefaturas operativas.

Capítulo 2. Metodología

2.1. Indicadores Clave de Desempeño (KPI)

Con el fin de evaluar el impacto del prototipo propuesto sobre la gestión de inventarios MRO, se definieron indicadores que permiten cuantificar mejoras tanto en la eficiencia operativa como en la continuidad del servicio. La literatura señala que el diseño de KPIs debe responder simultáneamente a criterios de relevancia, medibilidad y alineación estratégica (Silver, Pyke, & Peterson, 1998; Chen et al., 2019). A continuación, se presentan los cinco principales indicadores seleccionados, cuyas definiciones, fórmulas y metas esperadas se encuentran resumidas en la Tabla 1

2.1.1. KPI 1 - Tiempo total de gestión de inventario

Este indicador mide la duración de un ciclo completo de gestión, que incluye la identificación de insumos, el registro de ingresos y salidas, así como la generación de solicitudes de reposición. Se formaliza en la Ecuación (1):

$$T_{ciclo} = \underline{t}_{tarea} \times N_{tareas} \quad (1)$$

Donde T_{ciclo} representa el tiempo que tarda un ciclo completo de gestión, $\underline{t}_{tarea}$ corresponde al tiempo promedio por tarea y N_{tareas} al número de tareas de cada ciclo

La importancia de este KPI radica en que una reducción significativa de T_{ciclo} refleja la capacidad del sistema para digitalizar flujos logísticos y disminuir la dependencia de registros manuales, en línea con lo reportado por Zhang et al. (2012) y Taj (2023). En entornos industriales, donde los costos de oportunidad del tiempo son elevados, este indicador actúa como una medida directa de eficiencia administrativa.

2.1.2. KPI 2 - Horas Hombre (HH) utilizadas en cuadraturas y ajustes

Este KPI contabiliza las horas-hombre (HH) destinadas a tareas correctivas como conciliaciones físico-digitales, ajustes de stock y búsquedas de materiales no trazables. La expresión formal se indica en la Ecuación (2):

$$HH_{corr} = \sum_{i=1}^m h_i \quad (2)$$

Donde h_i representa las horas invertidas en la tarea correctiva i .

La relevancia de este indicador radica en que un exceso de HH_{corr} denota falta de calidad de datos y procesos poco confiables. La incorporación de tecnologías de trazabilidad, como los códigos QR, apunta a disminuir estas pérdidas, permitiendo que los recursos humanos se reasignen a actividades de mayor valor agregado (Siponen et al., 2019). En términos organizacionales, constituye una métrica clara de eficiencia en el uso del recurso humano.

2.1.3. KPI 3 -Días sin servicio por falta de stock

Este indicador captura el impacto operativo de los quiebres de inventario sobre la continuidad de los procesos. Se expresa a través de dos dimensiones: el tiempo total de indisponibilidad (Ecuación 3a) y el costo monetario asociado (Ecuación 3b).

$$DSS = \sum_{j=1}^J d_j \quad (3a)$$

$$C_{inactividad} = \sum_{j=i}^J d_j \cdot c_j \quad (3b)$$

donde d_j corresponde a la duración del evento de quiebre j y c_j al costo por día de inactividad.

Este KPI es crucial en contextos industriales, ya que los quiebres de stock pueden generar pérdidas millonarias por hora de detención (Chen et al., 2019). Asimismo, conecta directamente la precisión de los modelos predictivos con el nivel de continuidad operativa (Aktepe et al., 2023).

2.1.4. KPI 4 - Disponibilidad de stock crítico

Este KPI mide la proporción de días en que los ítems catalogados como críticos permanecen sobre el nivel mínimo definido por políticas de inventario. Su formulación se presenta en la Ecuación (4):

$$Disponibilidad(\%) = 100 \times \frac{Días\ con\ stock \geq S_{min}}{Días\ del\ periodo} \quad (4)$$

donde S_{min} representa el nivel mínimo de seguridad definido para cada ítem.

Su importancia reside en que actúa como un indicador inverso al riesgo de quiebre, reflejando la capacidad del sistema para sostener la operación sin interrupciones (Silver et al., 1998). Además, al estar vinculado al rendimiento de los modelos de predicción, constituye una métrica de eficacia de las señales de reabastecimiento (Vidal et al., 2022).

2.1.5. KPI 5 -Desviación del stock promedio respecto del stock óptimo

Este KPI evalúa el grado en que el inventario real se aproxima al nivel óptimo definido por políticas de gestión, como (s, S) o ROP. Se expresa mediante la Ecuación (5):

$$\Delta(\%) = 100 \times \frac{\bar{S} - S^*}{S^*} \quad (5)$$

donde $\bar{S}$ es el stock promedio observado y S^* el stock óptimo teórico.

La importancia de este KPI radica en su capacidad para capturar tanto el sobre stock (capital inmovilizado y costos de almacenamiento) como el substock (riesgo de quiebre). Según Ifraz et al. (2023), los modelos predictivos basados en aprendizaje automático permiten reducir la desviación de $\Delta(\%)$, acercando el inventario observado al nivel teóricamente deseable.

2.1.6. Metas Esperadas

Los indicadores clave de desempeño (KPIs) definidos en este estudio se complementan con metas de referencia basadas en la literatura, lo que permite contextualizar el impacto esperado del prototipo. La digitalización de procesos logísticos puede reducir entre 20% y 40% los tiempos de gestión (Wamba et al., 2020) y en más de 30% las horas-hombre destinadas a cuadraturas manuales (Siponen et al., 2019). Asimismo, el uso de modelos predictivos incrementa la disponibilidad de stock crítico en 10% a 20% (Vidal et al., 2022) y favorece un mayor equilibrio entre stock promedio y stock óptimo (Chen et al., 2019).

En el ámbito de la predicción de demanda, métodos como Croston y algoritmos de Machine Learning avanzados, como XGBoost, permiten disminuir el error de pronóstico en rangos de 15% a 25% frente a enfoques tradicionales (Aktepe et al., 2023; Fiorentia et al., 2021). Estas metas deben entenderse como parámetros de referencia académicos, cuya validación final dependerá de la implementación práctica en la empresa.

Los KPIs seleccionados no sólo permiten medir el impacto de la digitalización del inventario en términos de eficiencia administrativa (KPI 1 y KPI 2), sino también en la continuidad operativa (KPI 3 y KPI 4) y en la optimización de recursos (KPI 5). Además, su formulación matemática permite comparaciones objetivas *antes y después* de la implementación del prototipo, en línea con las mejores prácticas de evaluación en entornos industriales digitalizados (Wamba et al., 2020; Ghobakhloo & Fathi, 2019).

Tabla 1: Comparación de KPI.
Fuente: Elaboración propia

KPI	Definición	Formula	Unidad	Meta esperada
Tiempo total de gestión de inventario	Duración de un ciclo operativo	$T_{ciclo} = \underline{t}_{tarea} \times N_{tareas}$	horas/ciclo	Reducción $\geq 20\%$
HH en cuadraturas y ajustes	Suma de HH destinadas a tareas correctivas	$HH_{corr} = \sum_{i=1}^m h_i$	HH/mes	Reducción $\geq 30\%$
Días sin servicio por falta de stock	Tiempo de indisponibilidad por quiebres y costo monetario asociado	$DSS = \sum_{j=1}^J d_j$ $C_{inactividad} = \sum_{j=i}^J d_j \cdot c_j$	Días, CLP/USD	Reducción $\geq 40\%$
Disponibilidad de stock crítico	Proporción de días con inventario $\geq$ stock mínimo operacional	$Disp(\%) = 100 \times \frac{DCS \geq S_{min}}{DdP}$	%	Aumento $\geq 10\%$
Desviación del stock promedio vs. óptimo	Diferencia relativa entre stock observado y stock óptimo definido	$\Delta(\%) = 100 \times \frac{\bar{S} - S^*}{S^*}$	%	Reducción $\geq 10\%$

2.2. Trazabilidad con códigos QR como solución IoT accesible

La trazabilidad constituye un eje crítico en la gestión de inventarios industriales, al permitir identificar en tiempo real la ubicación, el estado y el historial de cada ítem dentro del flujo logístico. Sin embargo, su implementación suele enfrentar barreras asociadas al costo de tecnologías avanzadas, como RFID o sensores especializados, lo que limita su adopción en organizaciones con presupuestos restringidos o con baja madurez digital (Taj, 2023; Ghobakhloo & Fathi, 2019).

En este contexto, los códigos QR emergen como una solución IoT accesible, al ofrecer un mecanismo de identificación digital de bajo costo, fácil implementación y alta compatibilidad con dispositivos móviles. A diferencia del RFID, que requiere infraestructura dedicada, los QR pueden ser generados e impresos con recursos básicos, siendo legibles tanto por lectores especializados como por smartphones y tablets, lo que democratiza su aplicación en distintos entornos industriales (Noguerra, 2023; Desai, 2025).

La información contenida en un código QR puede incluir atributos esenciales del inventario, tales como:

- Identificación única del ítem (código SAP o interno).
- Descripción del material y ubicación física dentro de la bodega.
- Estado operativo (nuevo, en uso, en reparación, crítico).
- Datos de trazabilidad como fecha de ingreso, proveedor o lote.

Al integrarse con un sistema ERP o con dashboards analíticos, estos registros permiten reducir la dependencia de entradas manuales y habilitar la actualización en tiempo real de los niveles de stock. En términos de impacto, la literatura destaca que el uso de QR puede reducir hasta un 30% las horas-hombre dedicadas a cuadraturas y búsquedas de materiales no trazables, al aumentar la confiabilidad de los registros digitales (Siponen et al., 2019; Wamba et al., 2020).

Adicionalmente, implementaciones recientes muestran que el uso combinado de QR con microcontroladores de bajo costo, como el ESP32-CAM, y plataformas web progresivas (PWA), permite construir sistemas de trazabilidad multiplataforma y escalables, adecuados incluso para operaciones en terreno con conectividad limitada (Desai, 2025). Esto se traduce en una mayor resiliencia operativa, ya que el personal de mantenimiento puede consultar y actualizar el estado de los materiales desde cualquier punto de la faena.

Por consiguiente, el uso de códigos QR como herramienta IoT se perfila como una táctica efectiva y realista para la transformación digital logística. Aunque esta tecnología no posee la complejidad técnica del RFID, su eficiencia económica y su sencilla integración con la infraestructura existente la posicionan como la opción idónea para resolver problemas de trazabilidad en industrias que, pese a tener restricciones presupuestarias, requieren asegurar la continuidad operativa (Noguerra, 2023; Chen et al., 2019).

2.3. Modelos predictivos utilizados

En la gestión de inventarios con demandas irregulares, ningún modelo es universalmente superior. La literatura especializada demuestra que tanto los métodos estadísticos clásicos como los algoritmos de Machine Learning (ML) presentan fortalezas y debilidades según el contexto (Pinçe et al., 2021; Aktepe et al., 2023; Ifraz et al., 2023). Por esta razón, en el presente proyecto se consideran diversas alternativas que serán aplicadas de manera independiente y posteriormente evaluadas comparativamente bajo criterios homogéneos.

La elección de los modelos considerados en este estudio responde a criterios de adecuación metodológica frente a los patrones de demanda característicos del inventario MRO. Existen

múltiples enfoques en la literatura para la predicción de inventarios, que abarcan desde modelos estadísticos básicos hasta arquitecturas avanzadas de *deep learning* como LSTM o DeepAR (Salinas et al., 2017). Sin embargo, la selección realizada se justifica en cuatro ejes principales:

1. Adecuación al tipo de demanda:

- ARIMA representa una línea base consolidada para ítems de consumo regular y estacional, proporcionando un referente estadístico de comparación (Box et al., 2016).
- Croston y sus variantes (SBA, TSB) fueron diseñados específicamente para demandas intermitentes, frecuentes en entornos de mantenimiento (Croston, 1972; Hasani, 2023).
- Random Forest y XGBoost permiten capturar relaciones multivariantes y no lineales, comunes en ítems cuya demanda está condicionada por factores externos como planes de mantenimiento, horas de operación o condiciones ambientales (Vidal et al., 2022; Aktepe et al., 2023).

2. Balance entre simplicidad e interpretabilidad:

Se optó por un conjunto que combina modelos de baja complejidad (ARIMA, Croston) con otros de mayor capacidad predictiva pero menos interpretables (Random Forest, XGBoost). Este balance permite contrastar la efectividad de métodos clásicos frente a algoritmos de Machine Learning sin introducir complejidad innecesaria en etapas iniciales.

3. Relevancia en la literatura aplicada:

Estos modelos han sido ampliamente utilizados en investigaciones previas sobre inventarios industriales y predicción de demanda MRO, lo que permite establecer comparaciones con evidencia empírica existente (Siponen et al., 2019; Pinçe et al., 2021; Ifraz et al., 2023).

4. Viabilidad computacional y de datos:

El alcance del proyecto contempla pruebas con datos históricos extraídos de SAP. En este contexto, los modelos seleccionados equilibran precisión y factibilidad: los estadísticos requieren menor poder de cómputo y los algoritmos de ensamble (RF, XGBoost) pueden ejecutarse en infraestructura estándar, sin necesidad de entornos especializados en deep learning.

En consecuencia, se descartaron metodologías como redes neuronales recurrentes (RNN, LSTM) o modelos probabilísticos bayesianos complejos, los cuales, aunque de gran potencial, demandan volúmenes de datos más extensos, mayor capacidad de cómputo y un grado de madurez digital que excede los alcances del prototipo actual. Por ello, los cuatro modelos seleccionados representan un rango metodológico realista y comparativamente robusto para evaluar la factibilidad de un sistema de inventario predictivo en el contexto de MRO industrial.

2.3.1. ARIMA (AutoRegressive Integrated Moving Average)

El modelo ARIMA combina autorregresión, diferenciación e integración de media móvil, siendo uno de los métodos más utilizados en series temporales. Su formulación general se expresa en la ecuación (6)

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \quad (6)$$

Donde $\phi(B)$ y $\theta(B)$ representan los polinomios de rezagos, d es el orden de diferenciación y ε_t corresponde a un ruido blanco (Box et al., 2016; Hyndman & Athanasopoulos, 2018).

Respecto a su aplicabilidad, este modelo presenta ventajas significativas como su alta interpretabilidad, una sólida base estadística y un bajo costo computacional, lo que facilita su implementación. Sin embargo, su efectividad se ve restringida en ciertos escenarios, mostrando un pobre desempeño frente a series intermitentes o altamente irregulares, tal como lo evidencia la literatura reciente (Pinçe et al., 2021).

2.3.2. Croston y variantes (TSB, SBA)

El método de Croston (1972) fue diseñado específicamente para abordar demandas intermitentes, operando mediante la separación de la estimación en dos componentes: la magnitud de la demanda cuando esta ocurre y el intervalo temporal entre dichas ocurrencias. Para ello, se calculan el tamaño medio de la demanda mediante la Ecuación (7) y el intervalo medio entre demandas a través de la Ecuación (8):

1. Tamaño medio de la demanda ($\hat{z}_t$)

$$\hat{z}_t = \hat{z}_{t-1} + \alpha(z_t - \hat{z}_{t-1}) \quad (7)$$

2. Intervalo medio entre demandas ($\hat{p}_t$)

$$\hat{p}_t = \hat{p}_{t-1} + \alpha(p_t - \hat{p}_{t-1}) \quad (8)$$

Cada componente se estima por separado utilizando suavizamiento exponencial con un parámetro α . Finalmente, estos valores se combinan para obtener la predicción de tasa de demanda, tal como se expresa en la Ecuación (9):

Predicción final

$$\hat{y}_{t+h} = \frac{\hat{z}_t}{\hat{p}_t} \quad (9)$$

Adicionalmente, se han desarrollado variantes como TSB (Teunter–Syntetos–Babai) y SBA (Syntetos–Boylan Approximation), las cuales refinan el modelo original para corregir sesgos y mejorar la adaptación en escenarios de demanda muy baja.

En cuanto a su desempeño, este enfoque destaca por ser adecuado para series con una gran cantidad de ceros y patrones esporádicos, logrando corregir los sesgos inherentes a los modelos tradicionales. No obstante, presenta limitaciones importantes, ya que es un método univariante que no considera variables externas y asume rígidamente que la naturaleza intermitente de la demanda mantendrá su patrón histórico en el futuro.

2.3.3. Random Forest

Random Forest es un algoritmo de ensamble fundamentado en la generación de múltiples árboles de decisión, los cuales son entrenados sobre subconjuntos aleatorios de datos y variables. Para obtener el resultado final, el modelo combina las predicciones individuales de cada árbol mediante el cálculo del promedio, en el caso de regresión, o por voto mayoritario en tareas de clasificación. Este proceso de agregación se formaliza matemáticamente en la Ecuación (10):

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (10)$$

Donde B es el número de árboles y $T_b(x)$ es la predicción del árbol b para las características x.

Este algoritmo es ampliamente valorado por su robustez frente al ruido y a los valores atípicos (outliers), así como por su capacidad para capturar relaciones no lineales y multivariantes complejas sin requerir supuestos estrictos. No obstante, su implementación conlleva ciertas limitaciones, entre las que destacan una menor interpretabilidad global comparada con modelos más simples y la necesidad de realizar una ingeniería de características (feature engineering) exhaustiva para poder capturar la estacionalidad y las dependencias temporales de la serie (Vidal et al., 2022).

2.3.4. XGBoost (Extreme Gradient Boosting)

XGBoost es un algoritmo de gradient boosting optimizado que construye modelos secuenciales, corrigiendo en cada paso los errores del anterior. Su principal ventaja radica en la incorporación de términos de regularización en la función objetivo global L , definida en la Ecuación (11), lo que previene el sobreajuste:

$$L = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (11)$$

El término de regularización $\Omega(f)$, detallado en la Ecuación (12), controla la complejidad del modelo penalizando tanto el número de hojas (T) como la magnitud de los pesos (mediante regularización L2 o Ridge, representada por el parámetro λ):

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (12)$$

Finalmente, la actualización de la predicción en el paso t se realiza según la Ecuación (13):

$$\hat{y}^{(t)} = \hat{y}^{(t-1)}_i + \eta f_t(x_i) \quad (13)$$

Donde l es la función de pérdida, w_j es el peso asignado a la hoja j , γ el parámetro de penalización por hoja y η es la tasa de aprendizaje.

Este algoritmo destaca por su alto rendimiento y precisión, manejando eficientemente grandes volúmenes de datos y valores faltantes. Sin embargo, presenta como limitaciones una menor interpretabilidad en comparación con modelos lineales y una curva de aprendizaje más compleja para el ajuste de sus múltiples hiper parámetros.

2.3.5. Metodología de comparación de modelos

Con el propósito de evaluar de manera objetiva el rendimiento de los modelos predictivos propuestos, se estableció una metodología experimental que garantiza la comparabilidad, reproducibilidad y consistencia de los resultados.

Esta metodología considera las etapas de preprocesamiento de datos, entrenamiento y validación, y evaluación de desempeño, conforme a las mejores prácticas reportadas en la literatura sobre predicción de demanda intermitente (Vidal et al., 2022; Aktepe et al., 2023).

1. Preprocesamiento de datos

En la primera etapa, los registros históricos provenientes del sistema de inventario se someten a un proceso de depuración y estandarización.

Esto incluye la limpieza de valores faltantes o atípicos, la normalización temporal y la homogeneización de unidades de medida.

Posteriormente, se generan variables derivadas destinadas a capturar la dinámica temporal de la demanda, tales como retardos (*lags*), tendencias lineales y factores de estacionalidad.

Una vez completado este proceso, los datos quedan preparados para la etapa de modelamiento y validación.

2. Entrenamiento y validación

En esta fase, los conjuntos de datos se dividen en entrenamiento (80%) y prueba (20%), manteniendo la secuencia cronológica para evitar fugas de información (*data leakage*).

Se implementa además una validación cruzada temporal (*rolling time-series cross-validation*), con el fin de evaluar la estabilidad del modelo ante distintos periodos y garantizar que el desempeño no dependa de un único intervalo de prueba (Hyndman & Athanasopoulos, 2018).

Cada modelo se entrena y evalúa de manera independiente bajo las mismas condiciones experimentales y con idéntico horizonte de predicción.

3. Evaluación de desempeño

El rendimiento de los modelos se cuantifica mediante métricas estándar de error de predicción ampliamente utilizadas en la literatura, entre ellas el Error Medio Absoluto (MAE), el Error Medio Absoluto Porcentual (MAPE), la Raíz del Error Cuadrático Medio (RMSE) y el coeficiente de determinación (R^2).

Estas métricas, que serán desarrolladas en detalle en el apartado siguiente, permiten analizar de forma complementaria la precisión, estabilidad y capacidad explicativa de cada modelo.

Adicionalmente, se realiza un análisis cualitativo del equilibrio entre precisión, interpretabilidad y costo computacional, con el fin de determinar el modelo que mejor balance logra entre exactitud y aplicabilidad práctica en contextos operativos reales (Wamba et al., 2020).

4. Selección del modelo óptimo

Una vez calculadas las métricas de desempeño, los resultados obtenidos por cada modelo se comparan de manera global sobre el mismo conjunto de prueba. El modelo con mejor rendimiento promedio general, considerando tanto las métricas de error como los criterios de robustez y eficiencia, se selecciona como la referencia para la etapa de simulación operativa dentro del prototipo. Este enfoque permite una evaluación objetiva y cuantitativa de la efectividad relativa de los modelos implementados, asegurando que la decisión final se base en evidencia empírica y no en criterios subjetivos.

De este modo, el diseño experimental propuesto habilita la comparación de los modelos bajo parámetros controlados y uniformes, asegurando una validación objetiva y reproducible de su rendimiento previo a su despliegue en el sistema.

2.4. Métricas de evaluación y funciones de pérdida

Para evaluar y comparar el rendimiento de los modelos predictivos implementados, resulta necesario establecer métricas que permitan medir de forma objetiva la precisión y la fiabilidad de las predicciones generadas. En este proyecto se emplean cuatro métricas ampliamente reconocidas en la literatura sobre pronóstico de demanda: el Error Medio Absoluto (MAE), el

Error Medio Absoluto Porcentual (MAPE), la Raíz del Error Cuadrático Medio (RMSE) y el Coeficiente de Determinación (R^2).

Cada una de estas métricas proporciona una perspectiva complementaria del desempeño del modelo, lo que permite una evaluación integral del mismo. Esta combinación posibilita identificar no solo el modelo más preciso en promedio, sino también aquel con mayor robustez frente a patrones de consumo irregulares y menor sensibilidad a errores extremos (Hyndman & Athanasopoulos, 2018; Vidal et al., 2022).

2.4.1. Error Medio Absoluto (MAE)

El MAE (*Mean Absolute Error*) mide la magnitud promedio de los errores cometidos por el modelo sin considerar su dirección, expresándose en las mismas unidades que la variable de demanda. Su formulación se presenta en la Ecuación (14):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (14)$$

El MAE permite interpretar directamente cuánto se desvía en promedio la predicción del valor real, por ejemplo, un MAE de 5 indica que el modelo se equivoca, en promedio, en ± 5 unidades del ítem.

Su principal ventaja radica en su facilidad de interpretación y su baja sensibilidad a valores atípicos en comparación con otras métricas cuadráticas. No obstante, su limitación radica en que todos los errores tienen el mismo peso, sin distinguir entre desviaciones menores o extremas.

2.4.2. Error Medio Absoluto Porcentual (MAPE)

El MAPE (*Mean Absolute Percentage Error*) expresa el error como un porcentaje relativo al valor real, permitiendo comparar el rendimiento del modelo entre ítems con distintas magnitudes de consumo. Su formulación se presenta en la Ecuación (15):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (15)$$

Un MAPE de 10% implica que, en promedio, el error de predicción equivale al 10% del valor real. Esta métrica es especialmente útil para la comparación Inter productos o entre líneas de inventario de diferentes escalas. Sus principales ventajas son su intuitividad y fácil interpretación gerencial. Sin embargo, no puede calcularse cuando $y_t = 0$ y puede generar distorsiones en ítems de bajo consumo, donde errores pequeños se traducen en porcentajes altos.

2.4.3. Raíz del Error Cuadrático Medio (RMSE)

El RMSE (*Root Mean Square Error*) penaliza de forma más severa los errores grandes, al elevarlos al cuadrado antes de promediarlos, lo que lo convierte en una métrica sensible a desviaciones extremas. Su formulación se muestra en la Ecuación (16):

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (16)$$

Un RMSE alto sugiere la presencia de errores grandes o valores atípicos frecuentes. Esta métrica es útil cuando los errores significativos implican costos operativos elevados, como en la falta de un repuesto crítico. Su ventaja es que penaliza fuertemente las predicciones inestables, favoreciendo modelos más consistentes. Como desventaja, puede verse afectado por unos pocos valores atípicos y no es tan intuitivo para la interpretación de usuarios no técnicos.

2.4.4. Coeficiente de Determinación (R^2)

El R^2 mide la proporción de la varianza observada en la demanda que es explicada por el modelo. Su formulación se expresa en la Ecuación (17):

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (17)$$

Donde $\bar{y}$ es el valor medio de la demanda observada.

Un R^2 de 0.9 indica que el 90% de la variabilidad de la demanda es explicada por el modelo. Su principal ventaja es su interpretación intuitiva como porcentaje de varianza explicada, siendo especialmente útil para comparar modelos entrenados sobre los mismos datos. No obstante, un valor alto de R^2 no siempre implica alta precisión predictiva, ya que puede verse afectado por sobreajuste o por series no estacionarias.

2.4.5. Estrategia de uso de las métricas en la metodología

Durante el proceso de validación, las métricas descritas se aplican de manera complementaria para obtener una evaluación equilibrada del rendimiento de los modelos:

- MAPE: permite comparar la precisión relativa entre ítems y seleccionar el modelo con mejor rendimiento global.
- MAE: traduce los errores a unidades físicas, facilitando su interpretación económica.
- RMSE: destaca modelos que minimizan errores críticos o extremos.

- R^2 : mide la capacidad explicativa global del modelo respecto a la variabilidad de los datos.

Este conjunto de métricas proporciona una visión integral del desempeño, permitiendo identificar el modelo más preciso, robusto y operacionalmente aplicable en el contexto del sistema de inventario inteligente desarrollado.

2.5. Segmentación analítica de ítems

Aunque el propósito principal de este estudio es identificar un único algoritmo con el mejor desempeño global en la predicción de demanda del inventario, resulta metodológicamente pertinente realizar una segmentación analítica de los ítems. Este procedimiento no busca modificar la selección del modelo, sino evaluar su robustez y consistencia de desempeño bajo distintos patrones de consumo. De esta forma, se obtiene una visión más granular y validada empíricamente del comportamiento del modelo frente a la heterogeneidad propia de los inventarios industriales MRO.

2.5.1. Justificación

La literatura especializada en predicción de inventarios MRO demuestra que los algoritmos de pronóstico no mantienen un rendimiento uniforme en todos los contextos (Croston, 1972; Aktepe et al., 2023; Vidal et al., 2022). Los métodos estadísticos tienden a destacar en series estables y regulares, mientras que los algoritmos de Machine Learning suelen obtener mejores resultados en patrones complejos o no lineales. En consecuencia, la segmentación analítica permite comprobar si el modelo seleccionado como ganador global conserva un rendimiento competitivo dentro de cada categoría de ítem, reforzando así la validez externa y la capacidad generalizadora del prototipo desarrollado.

Adicionalmente, este análisis responde a la naturaleza heterogénea de las series temporales de inventarios MRO, donde conviven ítems de alta rotación con otros de consumo intermitente o estacional. Evaluar el modelo bajo esta diversidad permite comprender mejor sus límites de aplicación y su sensibilidad frente a diferentes comportamientos de demanda.

2.5.2. Criterios de segmentación

La clasificación de los ítems se fundamenta en los patrones teóricos de demanda ilustrados en la Figura 2, los cuales caracterizan el comportamiento del consumo según su variabilidad y la frecuencia de los intervalos entre demandas (Chen et al., 2019). Para operacionalizar estos conceptos, se definieron umbrales cuantitativos basados en series históricas extraídas del sistema SAP, cuyos detalles y métricas específicas se presentan en la Tabla 2.

- **Alta rotación:** ítems con consumo frecuente y estable.
 - Criterio: ≥ 1 salida mensual en promedio.
- **Estacionales:** ítems con demanda concentrada en periodos específicos, vinculada generalmente a mantenimientos programados o ventanas operativas.
 - Criterio: detección de picos recurrentes mediante análisis de autocorrelación (ACF) o patrones periódicos definidos.

- **Intermitentes:** ítems con largos periodos sin consumo, seguidos de eventos aislados de alta demanda.
 - Criterio: $\geq 50\%$ de periodos sin consumo y alta variabilidad en los intervalos entre demandas (inter-arrival times).
- **Irregulares:** ítems con consumo continuo, pero altamente volátil, sin tendencia ni estacionalidad definida.
 - Criterio: coeficiente de variación ($CV = \sigma/\mu$) $> 1,5$.

2.5.3. Procedimiento

El proceso metodológico de segmentación se desarrolla en las siguientes etapas:

1. **Extracción de datos históricos:** obtención de registros de consumo desde SAP, correspondientes al último año de operación.
2. **Cálculo de métricas estadísticas:** determinación de frecuencia de salidas, intervalos medios, coeficiente de variación y análisis de autocorrelación para cada ítem.
3. **Asignación de segmentos:** clasificación de los ítems conforme a los criterios definidos en la sección anterior.
4. **Evaluación comparativa:** aplicación del modelo seleccionado (y de los modelos alternativos, con fines comparativos) a los ítems de cada segmento.
5. **Cálculo de métricas de desempeño:** estimación de errores (MAE, RMSE, R^2 , MASE) para cada segmento de manera independiente.
6. **Análisis integrador de resultados:** comparación de los valores medios de error entre segmentos y síntesis global para verificar la estabilidad y adaptabilidad del modelo en contextos diversos.

2.5.4. Valor esperado del análisis

El análisis de segmentación permitirá:

- Verificar si el modelo con mejor desempeño global mantiene una precisión aceptable en todos los segmentos.
- Detectar posibles debilidades locales (por ejemplo, bajo desempeño en ítems estacionales o altamente intermitentes).
- Identificar oportunidades para el futuro uso de enfoques híbridos o jerárquicos.
- Aportar evidencia adicional sobre la validez externa del prototipo y su aplicabilidad operativa en distintos escenarios de demanda.

De esta forma, la segmentación contribuye a consolidar la evaluación empírica del sistema propuesto, asegurando que su rendimiento no dependa exclusivamente de un tipo particular de comportamiento de consumo.

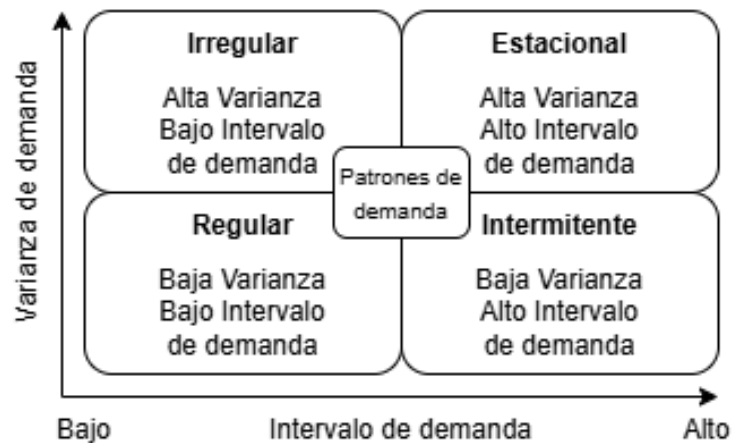


Figura 2: Patrones de demanda MRO
Fuente: Adaptado de Chen et al. (2019)

Tabla 2: Clasificación de ítems según patrón de consumo
Fuente: Elaboración Propia

Segmento	Definición	Métrica utilizada	Criterio de clasificación
Alta rotación	Ítems con consumo frecuente y regular.	Frecuencia de salidas mensuales.	≥ 1 salida mensual promedio durante el periodo analizado.
Estacionales	Ítems con consumo concentrado en periodos específicos (por ejemplo, mantenimientos programados).	Análisis de autocorrelación (ACF) y estacionalidad.	Picos de consumo recurrentes en intervalos definidos (ej. trimestral, semestral).
Intermitentes	Ítems con largos periodos sin consumo, seguidos de demandas puntuales.	Proporción de periodos sin consumo y variabilidad entre demandas.	$\geq 50\%$ de periodos sin consumo + alta variabilidad en intervalos de demanda.
Irregulares	Ítems con consumo continuo, pero altamente volátil.	Coeficiente de variación ($CV = \sigma/\mu$).	$CV > 1,5$

2.6. Estrategia de comparación global y selección del modelo único

El propósito de esta sección es establecer una metodología sistemática para la selección de un único algoritmo de pronóstico que sea aplicable de forma transversal a todos los ítems del inventario MRO. A diferencia de enfoques que asignan un modelo distinto a cada serie, este estudio busca determinar el modelo con mejor desempeño global, equilibrando precisión, estabilidad temporal y viabilidad operativa.

En este contexto, se evaluarán los modelos ARIMA, Croston (y sus variantes TSB y SBA), Random Forest y XGBoost, aplicados sobre el conjunto completo de series de demanda extraídas del sistema SAP. La comparación se realizará bajo un marco de evaluación uniforme, asegurando condiciones idénticas de entrenamiento, horizonte de pronóstico y métricas de validación.

2.6.1. Diseño de evaluación “todos contra todos” (una sola elección final)

La evaluación seguirá un diseño de tipo *all-against-all*, en el cual todos los algoritmos serán entrenados y validados sobre los mismos datos y horizontes de predicción.

Parámetros de evaluación:

- Fuente de datos: registros históricos extraídos desde SAP.
- Periodo de análisis: último año completo de operación.
- Frecuencia de modelado: semanal o mensual, manteniendo consistencia entre ítems.
- Horizonte de predicción: Establecido como h fijado en 12 semanas según este análisis

Se aplicará una división de datos cronológica (hold-out), que separa el conjunto de datos históricos en un segmento de entrenamiento (80%) y un segmento de prueba (20%). Esta estrategia simula un escenario realista donde el modelo, una vez entrenado con el pasado, debe pronosticar el futuro. Todos los algoritmos se entrenarán y evaluarán sobre estas mismas divisiones para garantizar una comparabilidad directa.

El preprocesamiento será homogéneo para todas las series: limpieza, imputación de valores faltantes, agregación temporal, generación de *lags* y variables estacionales, y normalización si corresponde. Esta uniformidad asegura que las diferencias observadas se atribuyan al desempeño del modelo y no a la manipulación de datos.

2.6.2. Métricas de desempeño y criterios de evaluación

La evaluación del rendimiento de los modelos predictivos implementados se basa en cuatro métricas ampliamente reconocidas en la literatura sobre pronóstico de demanda: Error Medio Absoluto (MAE), Error Medio Absoluto Porcentual (MAPE), Raíz del Error Cuadrático Medio (RMSE) y Coeficiente de Determinación (R^2), cuyas formulaciones se presentan en las Ecuaciones (14) a (17).

Estas métricas fueron seleccionadas por su complementariedad: mientras MAE permite interpretar el error en unidades reales y cuantificar su impacto económico, MAPE posibilita la

comparación relativa entre ítems con diferentes escalas de consumo. Por su parte, RMSE enfatiza la penalización de errores grandes, siendo útil para ítems críticos donde las desviaciones tienen mayor costo, y R^2 mide la capacidad explicativa global del modelo respecto de la variabilidad observada en la demanda (Hyndman & Athanasopoulos, 2018; Vidal et al., 2022).

El uso conjunto de estas métricas permite una evaluación integral, combinando precisión, estabilidad y capacidad de generalización. En esta sección, MAE y MAPE se utilizarán como indicadores principales de desempeño, mientras que RMSE y R^2 se considerarán métricas complementarias orientadas a identificar la robustez y consistencia de los modelos ante escenarios de alta variabilidad en la demanda.

2.6.3. Contrastes estadísticos entre algoritmos

Con el propósito de asegurar que la elección del modelo ganador no se deba al azar ni a una ventaja marginal en un subconjunto reducido de ítems, se aplicarán contrastes estadísticos de significancia sobre los resultados de error obtenidos. Este procedimiento busca determinar si las diferencias observadas en las métricas de desempeño son estadísticamente significativas y no atribuibles a la variabilidad inherente de los datos.

De acuerdo con la literatura en validación de pronósticos (Diebold & Mariano, 1995; Demšar, 2006; Hyndman & Athanasopoulos, 2018), se emplearán dos pruebas complementarias: el contraste de Diebold–Mariano (DM) y el análisis de rangos con signo de Wilcoxon.

2.6.3.1. Test de Diebold–Mariano (DM)

Este método evalúa si dos modelos de pronóstico presentan una diferencia estadísticamente significativa en su precisión media. Para ello, compara las pérdidas de predicción entre dos modelos A y B , definidas según la ecuación (18)

$$d_t = L(e_{A,t}) - L(e_{B,t}) \quad (18)$$

Donde $L(\cdot)$ es la función de pérdida y el error del modelo A en el tiempo t esta planteado como la ecuación (19):

$$e_{A,t} = y_t - \hat{y}_{A,t} \quad (19)$$

Gracias a estos valores, planteamos el estadístico DM en la ecuación (20):

$$DM = \frac{\bar{d}}{\sqrt{\widehat{Var}(\underline{d})}} \quad (20)$$

Bajo la hipótesis nula H_0 , ambos modelos tienen la misma precisión en promedio. Si el valor-p asociado al test < 0.05 , se rechaza H_0 , concluyendo que uno de los modelos es significativamente mejor.

En este estudio, la evaluación se aplicará de forma pareada entre algoritmos, comparando sus series de errores generadas sobre el conjunto de prueba.

Los resultados se consolidarán mediante el conteo de victorias significativas y, en caso necesario, con la agregación de valores-p mediante el método de Fisher, siguiendo las recomendaciones de Demšar (2006).

2.6.3.2. Test de rangos con signo de Wilcoxon

Se trata de una alternativa no paramétrica para comparar métodos emparejados sin asumir normalidad ni homogeneidad de varianzas. En el contexto del proyecto, permite identificar si las diferencias de error son sistemáticamente positivas o negativas a lo largo de los ítems del inventario.

Para cada ítem i , se calcula la diferencia entre los errores de dos modelos A y B según la ecuación (21):

$$d_i = e_{A,i} - e_{B,i} \quad (21)$$

A continuación, se toman los valores absolutos $|d_i|$, se ordenan de menor a mayor y se asignan rangos R_i . Luego, se restauran los signos originales para distinguir si las diferencias favorecen a un modelo u otro. Las sumas de rangos positivos y negativos se definen como W^+ y W^- , respectivamente, permitiendo calcular el estadístico final mediante la ecuación (22):

$$W = \min\left(\sum_{i=1}^n R_i^+, \sum_{i=1}^n R_i^-\right) \quad (22)$$

Bajo la hipótesis nula H_0 , las diferencias de desempeño se distribuyen alrededor de cero, es decir, ningún modelo es sistemáticamente mejor. Si el valor-p obtenido es menor a 0.05, se rechaza H_0 , indicando que uno de los modelos supera significativamente al otro en la mayoría de los casos.

La principal fortaleza de este enfoque radica es su robustez frente a distribuciones asimétricas y su capacidad de detectar superioridades consistentes incluso en contextos con alta variabilidad de la demanda, como ocurre en los inventarios MRO.

2.6.3.3. Criterios de decisión con significancia estadística

La selección del modelo único final se basará en una combinación de criterios estadísticos y prácticos:

1. Desempeño promedio y consistencia:

El modelo debe presentar el menor error medio global en las métricas principales (MAE, MAPE, RMSE) calculado como un promedio por ítem (enfoque macro).

Complementariamente, se evaluará la tasa de éxito del modelo, seleccionando aquel que logre minimizar el error en la mayor cantidad de ítems individuales (enfoque de frecuencia o conteo de victorias).

2. Consistencia por ítem:

El modelo ganador debe mantener un desempeño competitivo en la mayoría de los ítems, y no solo en el promedio agregado.

3. Significancia estadística:

- Contraste DM: las diferencias respecto al segundo mejor modelo deben ser estadísticamente significativas ($p < 0.05$).
- Test Wilcoxon: la distribución de diferencias de error entre ítems debe mostrar una tendencia sistemática en favor del modelo ganador.

4. Robustez práctica:

Además de la precisión estadística, el modelo debe demostrar un desempeño estable y generalizable, eficiencia computacional y facilidad de mantenimiento operativo.

2.6.4. Robustez práctica

Además de las métricas de precisión y los contrastes estadísticos, la selección del modelo final debe considerar su robustez práctica, entendida como la capacidad del algoritmo para mantener un desempeño estable, eficiente y operacionalmente viable en contextos reales de implementación.

En el ámbito de los sistemas MRO, un modelo estadísticamente preciso pero inestable o costoso de mantener puede resultar inaplicable en la práctica, comprometiendo la continuidad operativa del sistema predictivo (Chen et al., 2019; Vidal et al., 2022).

La robustez práctica se evaluará a partir de tres dimensiones complementarias:

1. Consistencia de Desempeño:

Esta dimensión verifica que el modelo no solo sea preciso en promedio, sino también confiable. Se evalúa de dos formas:

- Sensibilidad a errores grandes: Se analiza la métrica RMSE (Raíz del Error Cuadrático Medio). Un modelo con un RMSE bajo con relación a su MAE es más robusto, ya que indica que no comete errores extremos de gran magnitud, los cuales suelen tener un alto costo operativo.
- Consistencia entre segmentos de ítems: Se verifica que el modelo seleccionado como ganador global no deteriore significativamente su desempeño en categorías específicas de demanda. Esto se analiza en la tabla de resumen por segmentos, asegurando que el modelo dominante sea consistente o que el MAE promedio no se dispare en ningún grupo.

2. Costo Computacional y Latencia:

Dado que la planificación de inventarios MRO obedece a ciclos de reposición periódicos y no a demandas de tiempo real estricto, no se definieron métricas formales de tiempo de ejecución como criterio primario de selección. La estrategia priorizó la calidad analítica y la precisión del pronóstico por sobre la velocidad de inferencia, bajo la premisa de que el costo operativo de un error de predicción supera ampliamente el costo de unos segundos o minutos adicionales de cómputo en un proceso por lotes. No obstante, se realiza una evaluación cualitativa para descartar modelos cuya complejidad impida su ejecución en tiempos razonables dentro de la infraestructura disponible.

3. Operatividad e Interpretabilidad:

Se considera la facilidad de mantenimiento y el grado de interpretabilidad de los resultados. Modelos como ARIMA y Croston ofrecen alta transparencia, mientras que para los algoritmos de Machine Learning (Random Forest, XGBoost), la interpretabilidad se evalúa a nivel de su capacidad para manejar patrones complejos sin requerir una parametrización excesivamente especializada.

En caso de que dos algoritmos presenten rendimientos estadísticamente equivalentes, se priorizará aquel que muestre mayor consistencia de desempeño (menor RMSE relativo), menor complejidad operativa y mejor eficiencia de cómputo, privilegiando la sostenibilidad técnica del prototipo desarrollado.

De esta manera, la robustez práctica se integra como un criterio esencial que garantiza que el modelo elegido no solo sea el más preciso, sino también el más confiable y viable para su implementación dentro del sistema de inventario inteligente propuesto.

2.6.5. Procedimiento paso a paso

El proceso experimental para evaluar y seleccionar el modelo de pronóstico se desarrolla siguiendo ocho etapas secuenciales, diseñadas para asegurar la comparabilidad, reproducibilidad y validez de los resultados.

1. Preprocesamiento Único.

Se aplicará un pipeline común a todas las series: depuración de registros, agregación temporal a frecuencia semanal (de forma consistente para todos los ítems) y construcción de variables derivadas (rezagos, ventanas móviles). Este preprocesamiento homogéneo es idéntico para todos los algoritmos con el fin de asegurar que las diferencias de desempeño se deban a los modelos mismos y no a la preparación de los datos.

2. División de Datos para Validación (Hold-out).

Se definirá una única división cronológica de los datos en un conjunto de entrenamiento (80%) y uno de prueba (20%) para todos los algoritmos e ítems. Este

esquema simula un escenario operativo real y garantiza un marco de comparación estandarizado.

3. Entrenamiento y Predicción de Modelos Candidatos.

Sobre el conjunto de entrenamiento, se entrenarán por separado los modelos ARIMA, Croston/TSB, Random Forest y XGBoost. Cada modelo generará pronósticos para el periodo correspondiente al conjunto de prueba. No se combinan modelos; cada uno compite individualmente bajo idénticas condiciones experimentales.

4. Cálculo de Métricas por Ítem.

El desempeño de cada modelo sobre cada ítem se evaluará con las métricas MAE (Ecuación 14), MAPE (Ecuación 15), RMSE (Ecuación 16) y R^2 (Ecuación 17). Estas métricas permiten capturar la precisión promedio, la sensibilidad a errores grandes y la capacidad explicativa.

5. Segmentación Analítica de Ítems.

Cada ítem del inventario se clasificará en un segmento de demanda (Alta Rotación, Intermitente, Estacional, Irregular o Regular) según los criterios estadísticos definidos en la Sección 2.5.2. Esta clasificación se utilizará en la evaluación de robustez.

6. Contrastes Estadísticos Pareados.

Para cada par de algoritmos, se aplicarán:

- Test de Rangos con Signo de Wilcoxon (Ecuaciones 21-22), para verificar si la superioridad de un modelo sobre otro es sistemática a lo largo de los ítems.
- Test de Diebold-Mariano (DM) (Ecuaciones 18-20), para contrastar si las diferencias en la precisión media son estadísticamente significativas. Se considerará una significancia con $p < 0,05$.

7. Evaluación de Robustez Práctica.

Además del error promedio, se analizarán:

- Consistencia de Desempeño, evaluando el RMSE para detectar sensibilidad a errores extremos.
- Consistencia por Segmentos, verificando el rendimiento de los modelos en cada una de las categorías de demanda identificadas (alta rotación, intermitente, etc.).
- Operatividad, considerando la interpretabilidad y la simplicidad de cada modelo.

8. Selección Final del Modelo Único.

Se seleccionará un solo algoritmo para todo el inventario, priorizando una combinación de criterios:

- (i) Menor error global (MAE).
- (ii) Superioridad estadísticamente significativa frente a las alternativas (p-valores de los tests).
- (iii) Mayor robustez práctica (bajo RMSE relativo y consistencia entre segmentos).

En caso de empate, se favorecerá el modelo más estable y de menor complejidad operativa.

2.6.6. Plantilla de resultados (resumen global)

Los resultados consolidados de la comparación entre algoritmos, obtenidos a partir de las métricas macro y micro, los contrastes de significancia y los indicadores de robustez práctica, se presentarán en la Tabla 3

Tabla 3: Plantilla de resultados comparativos para modelo óptimo de pronóstico
Fuente: *Elaboración Propia.*

Algoritmo	MAE	MAPE	RMSE	R ²
ARIMA				
Croston / TSB				
Random Forest				
XGBoost				

Capítulo 3. Implementación

La fase de implementación constituye la materialización práctica de los conceptos teóricos y metodológicos desarrollados en los capítulos anteriores. En este apartado se detalla el proceso de desarrollo del prototipo de sistema inteligente de gestión de inventario, su arquitectura modular, las tecnologías empleadas y la integración funcional entre sus componentes.

El sistema fue concebido como una solución integral de carácter predictivo, automatizado y trazable, orientada a la gestión de inventarios industriales del tipo MRO (Maintenance, Repair and Operations). Su propósito es optimizar los procesos de control de stock, reducir la incertidumbre en la planificación y mejorar la disponibilidad de materiales críticos.

La arquitectura del prototipo integra en una plataforma de escritorio tres pilares tecnológicos fundamentales:

- Digitalización mediante códigos QR para la captura automatizada de datos,
- Modelos predictivos de Machine Learning para la estimación de demanda,
- Dashboards interactivos para la visualización y toma de decisiones.

De esta manera, el sistema busca demostrar la factibilidad de un ecosistema ciberfísico local alineado con los principios de la Industria 4.0.

El prototipo se estructura en tres módulos principales:

1. Módulo de Digitalización mediante Códigos QR.
2. Módulo de Predicción mediante Algoritmos de Machine Learning.
3. Módulo de Visualización y Soporte a la Decisión.

Cada uno aborda una dimensión específica del proceso de gestión de inventario, pero todos operan dentro de una misma arquitectura integrada.

3.1. Módulo de Digitalización mediante Códigos QR

Este módulo constituye la base del sistema y reemplaza los registros manuales tradicionales por un proceso automatizado de captura y trazabilidad. Su objetivo es garantizar la confiabilidad de los datos, reducir errores de registro y generar información estructurada para alimentar el módulo analítico.

3.1.1. Arquitectura general del módulo

El módulo fue desarrollado en Python bajo una interfaz gráfica Tkinter, actuando como núcleo operativo del sistema. Los usuarios registran transacciones mediante:

- Escaneo de códigos QR que contienen los atributos del material (código SAP, ubicación, descripción, etc.).

- Búsqueda manual mediante formularios integrados cuando el código no está disponible.

Los registros se gestionan localmente a través de dos archivos principales:

- *MATERIALES.xlsx*: actúa como la base maestra de inventario.
- *movimientos.csv*: funciona como registro cronológico (log) de todas las transacciones.

3.1.2. Funcionalidades principales

El diseño funcional del módulo se orienta a la automatización de la captura de datos y la simplificación de la interacción con el usuario. Para ello, integra capacidades de visión artificial para la lectura de códigos, gestión de archivos locales y una interfaz gráfica unificada. A continuación, la Tabla 4 detalla las principales funcionalidades implementadas y su descripción técnica dentro del sistema.

Tabla 4: Funcionalidades del módulo de digitalización
Fuente: Elaboración Propia

Funcionalidad	Descripción
Escaneo de códigos QR	Utiliza PyAV y OpenCV para capturar video desde una cámara y decodificar códigos QR en tiempo real.
Registro de movimientos	Permite registrar ingresos y salidas, asociando salidas a una Orden de Trabajo (OT) y adjuntando respaldos PDF.
Gestión de inventario local	Actualiza automáticamente el stock en <i>MATERIALES.xlsx</i> y registra los movimientos en <i>movimientos.csv</i> .
Interfaz gráfica unificada	Centraliza las operaciones mediante formularios, tablas y visualizaciones integradas en Tkinter.

3.1.3. Estructura de datos y Flujo Operativo

Para garantizar la integridad de la información procesada por el prototipo, se definió una estructura de datos basada en la exportación estándar del sistema ERP de la empresa (SAP) y un registro incremental de eventos locales.

- Origen de los datos (Input)
La información base se obtiene mediante una exportación de movimientos de mercancía desde SAP en formato Excel (.xlsx). Este archivo actúa como la "fuente de verdad" inicial del inventario. De las múltiples columnas que genera el reporte de SAP, el prototipo extrae y procesa las siguientes para construir el maestro de materiales:
 - Material: Código único identificador (ID SAP) del repuesto.

- Texto breve de material: Descripción técnica del ítem.
- Ctd.en UM entrada: Cantidad movida. El sistema interpreta el signo de este valor según la Clase de movimiento (ej. 101 para ingresos positivos, 261 para salidas/consumos negativos).
- Fe.contabilización: Fecha del registro original para establecer la línea de tiempo histórica.
- Registro de transacciones (Output)

La captura de datos se realiza a través de interfaces específicas según el tipo de operación. Para la recepción de materiales, se utiliza el formulario de ingresos detallado en la Figura 4, el cual permite validar las cantidades físicas contra las guías de despacho. Por su parte, el retiro de insumos se gestiona mediante la interfaz de salidas mostrada en la Figura 5, que obliga a vincular el consumo con una Orden de Trabajo (OT).

Una vez confirmada la operación en estas pantallas, el sistema no sobrescribe el historial pasado, sino que genera un nuevo registro en un archivo de bitácora llamado movimientos.csv. Este archivo estructurado por comas permite una lectura rápida para los algoritmos de Machine Learning y contiene:

 - timestamp: Fecha y hora exacta de la operación local.
 - accion: Tipo de movimiento (ej. "AGREGAR", "SALIDA").
 - ot: Número de Orden de Trabajo asociada (para trazabilidad de salidas).
 - codigo_sap / material: Identificación del ítem.
 - cantidad_movida: Flujo del movimiento (positivo o negativo).
 - cantidad_final: Saldo de stock resultante tras la operación.
 - archivo: Ruta del respaldo digital (PDF/Foto) adjunto a la transacción.
- Diagrama de Flujo del Proceso

El funcionamiento lógico del módulo, desde la carga de datos hasta la actualización de la visualización, se detalla en la Figura 3

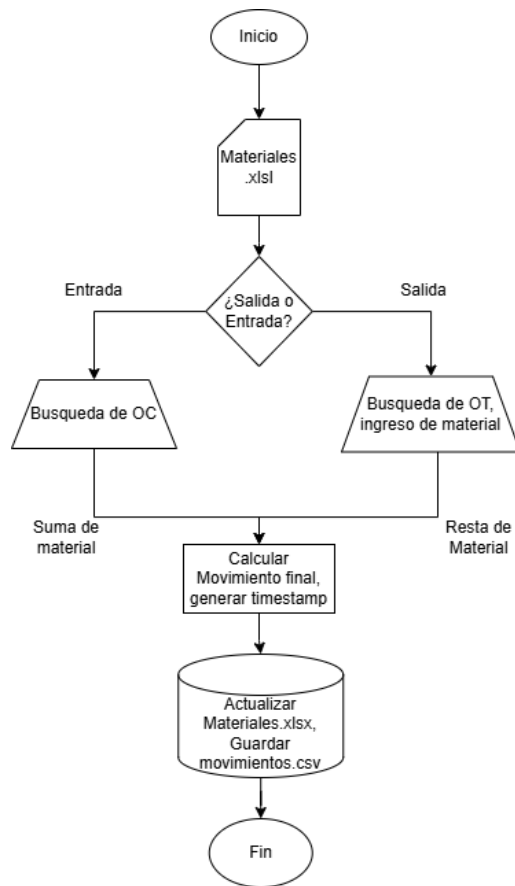


Figura 3: Flujo del módulo de digitalización
Fuente: Elaboración Propia

Excel inventario: c:\Users\daniel.herrera.aep\... Escritorio\Memoria\MATERIALES.xlsx

Abrir... Recargar Volver al inicio

Guía de despacho (Ingreso)

Archivo de guía: c:\Users\daniel.herrera.aep\... Escritorio\Memoria\Guías de despacho.xlsx Cargar Guía...

Guía cargada (edita 'Cant. recibida' si faltó material)

Código	Material	Cant. guía	Cant. recibida	Cant. actual	Cant. final
12345	MANGUEROTE	5	5	78	83
9876543	VALVULVA	5	5	27	32
12347	ACEITE	5	5	26	31

Editar recibido Confirmar ingresos Historial de Guías

Guía 1234567 cargada. Edita 'Cant. recibida' con doble clic o botón.

Figura 4: Pantalla de Inicio Entrada de Materiales
Fuente: Elaboración Propia

Excel inventario: c:\Users\daniel.herrera.aep\...escritorio\Memoria\MATERIALES.xlsx

Abrir... Recargar Volver al inicio Cámara: Integrated Camera Refrescar Iniciar Detener

Datos del material (SALIDA)

Buscar código: CODIGO Buscar Re-escanear (imagen)

OT (obligatoria): 1000046841 Adjuntar archivo...

CODIGO SAP: CODIGO

MATERIAL: ACOPIE

CANTIDAD: 0

Cantidad: 1 Añadir a lista (+) Añadir a lista (-)

UBICACIÓN X (p): 43.0

UBICACIÓN Y (p): 40.0

Lista de movimientos (pendiente)

Código	Material	Δ (cambio)	Cant. actual	Cant. final
12345	MANGUERO	5	5	78
9876543	VALVULA	5	5	27
12347	ACEITE	5	5	26

Quitar seleccionados Vaciar lista Editar cantidad Confirmar lista

Figura 5: Pantalla de Inicio Salida de Materiales
Fuente: Elaboración Propia

3.1.4. Desarrollo e implementación técnica

La implementación técnica del módulo se fundamenta en el uso del lenguaje Python, seleccionado por su versatilidad y su robusto ecosistema de bibliotecas de código abierto. La arquitectura de software integra componentes específicos para resolver las distintas capas de la aplicación: desde la gestión de la interfaz gráfica y el procesamiento de imágenes en tiempo real, hasta la manipulación estructurada de datos. En la Tabla 5 se presenta el desglose de las tecnologías y librerías empleadas, asociando cada una con su función crítica dentro del flujo de digitalización.

Tabla 5: Tecnologías utilizadas por componente en Módulo de Digitalización
Fuente: Elaboración Propia

Componente	Librería / Tecnología	Función
Interfaz principal	Tkinter	Control general del inventario y registro de transacciones.
Lector QR	PyAV, OpenCV	Captura y decodificación de códigos QR.
Gestión de datos	Pandas	Lectura, escritura y manipulación de archivos Excel/CSV.
Almacenamiento local	Archivos Excel / CSV	Persistencia del stock y registro histórico.

En conjunto, este módulo permite la captura estructurada de información de terreno, asegurando continuidad de datos hacia el módulo de predicción.

3.2. Módulo de algoritmos de Machine Learning

El Módulo de Machine Learning (ML) constituye el núcleo analítico del sistema inteligente de inventario, siendo responsable de transformar los datos históricos de consumo en pronósticos de demanda precisos y adaptativos. Su objetivo principal es reducir la incertidumbre en la planificación del stock, optimizar los niveles de inventario y evitar quiebres o sobreabastecimientos en materiales de tipo MRO.

El módulo se encuentra embebido en la misma aplicación de escritorio, lo que permite ejecutar modelos predictivos de forma local, autónoma y reproducible, sin depender de conexión a servidores o servicios externos. Esto refuerza su aplicabilidad en entornos industriales aislados o con infraestructura limitada.

3.2.1. Arquitectura general del módulo

El diseño del módulo sigue una arquitectura modular y jerárquica, tal como se esquematiza en la Figura 6. Esta estructura se organiza en tres capas funcionales que interactúan de manera secuencial para transformar los datos crudos en información visual:

1. Capa de Entrada de Datos

Controla la lectura, validación y limpieza del archivo Excel de movimientos históricos.

- Se ejecuta mediante la función *load_movimientos_from_excel()*, que filtra los datos relevantes (principalmente salidas de inventario) y los estructura en una serie temporal.

2. Capa de Procesamiento Predictivo

Representa el corazón del motor analítico, encapsulado en la clase *MLEngine*.

- Esta capa ejecuta el ciclo de modelado completo: preprocesamiento, entrenamiento, validación y comparación de resultados entre los cuatro algoritmos implementados.

3. Capa de Salida e Integración Visual

Los resultados se devuelven como estructuras en memoria (diccionarios y *DataFrames* de Pandas), las cuales son consumidas directamente por el módulo de visualización.

- Esto elimina la necesidad de archivos intermedios o procesos externos y permite una retroalimentación inmediata al usuario.

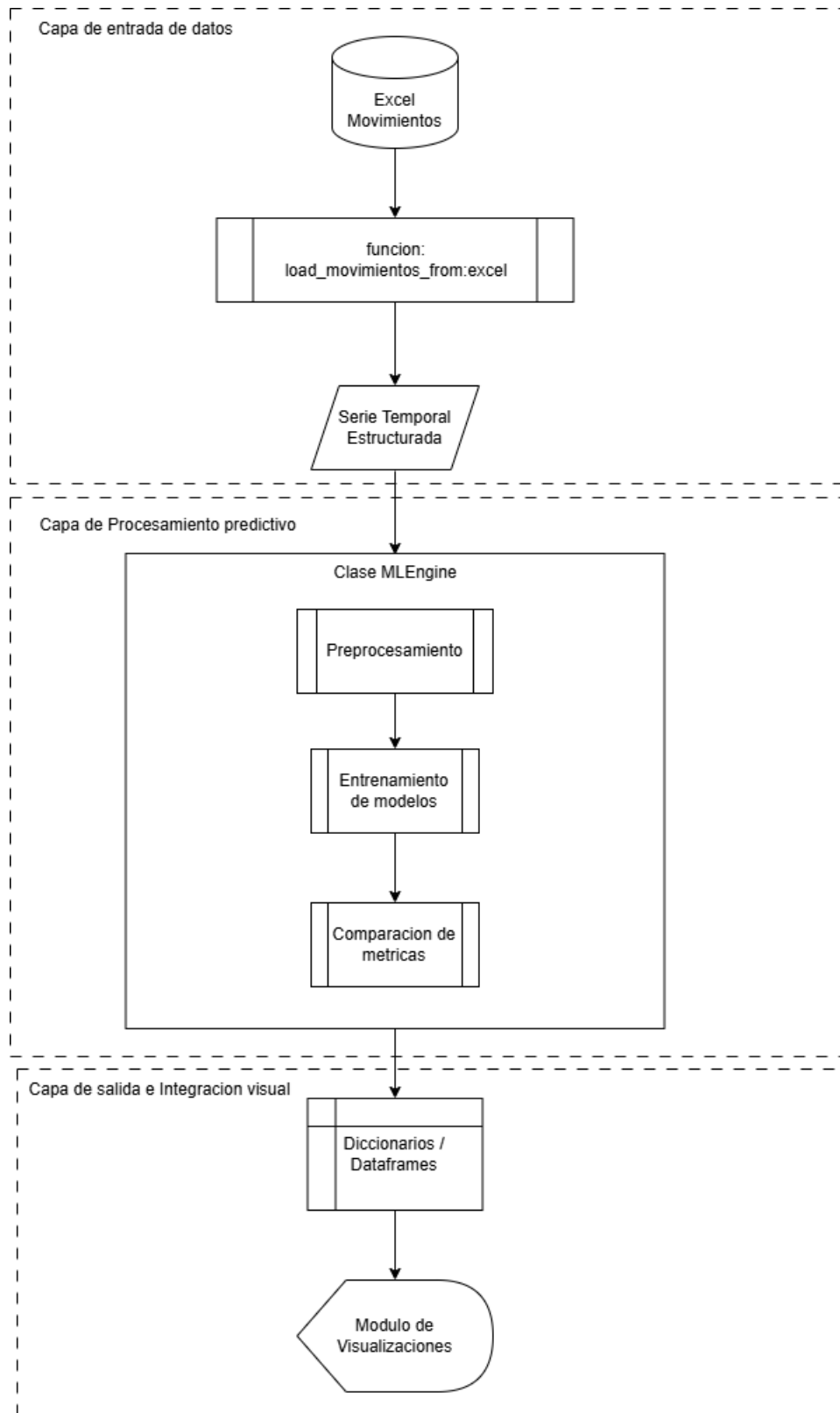


Figura 6: Diagrama de la arquitectura del Módulo
Fuente: elaboración propia

3.2.2. Flujo de Ejecución del módulo

El proceso analítico sigue un flujo lógico de cinco etapas secuenciales, tal como se detalla en el diagrama de la Figura 7. Este procedimiento se ejecuta automáticamente una vez que el usuario inicia la simulación desde la interfaz:

1. Extracción y carga de datos:

El usuario selecciona un archivo Excel (.xlsx) con el historial de movimientos de materiales. El sistema lo carga, verifica su integridad y extrae las transacciones relevantes (especialmente las salidas).

2. Preprocesamiento:

Los datos se transforman en series temporales mediante agregación semanal, eliminación de valores atípicos y generación de variables derivadas. La función `_create_features()` produce indicadores como rezagos (lag), promedios móviles y estacionalidades para enriquecer los modelos predictivos.

3. Entrenamiento de modelos:

Cada algoritmo se entrena sobre el 80% más antiguo de la serie temporal (conjunto de entrenamiento), aplicando la división temporal definida en el Capítulo 2.

4. Predicción y validación:

Los modelos generan pronósticos sobre el 20% más reciente (conjunto de prueba) y se evalúan según métricas estandarizadas:

- a. MAE (Mean Absolute Error)
- b. RMSE (Root Mean Square Error)
- c. MAPE (Mean Absolute Percentage Error)
- d. R^2 (Coeficiente de determinación)

5. Visualización y análisis:

Los resultados se devuelven a la interfaz en tiempo real, actualizando las curvas de predicción y la tabla comparativa de desempeño.

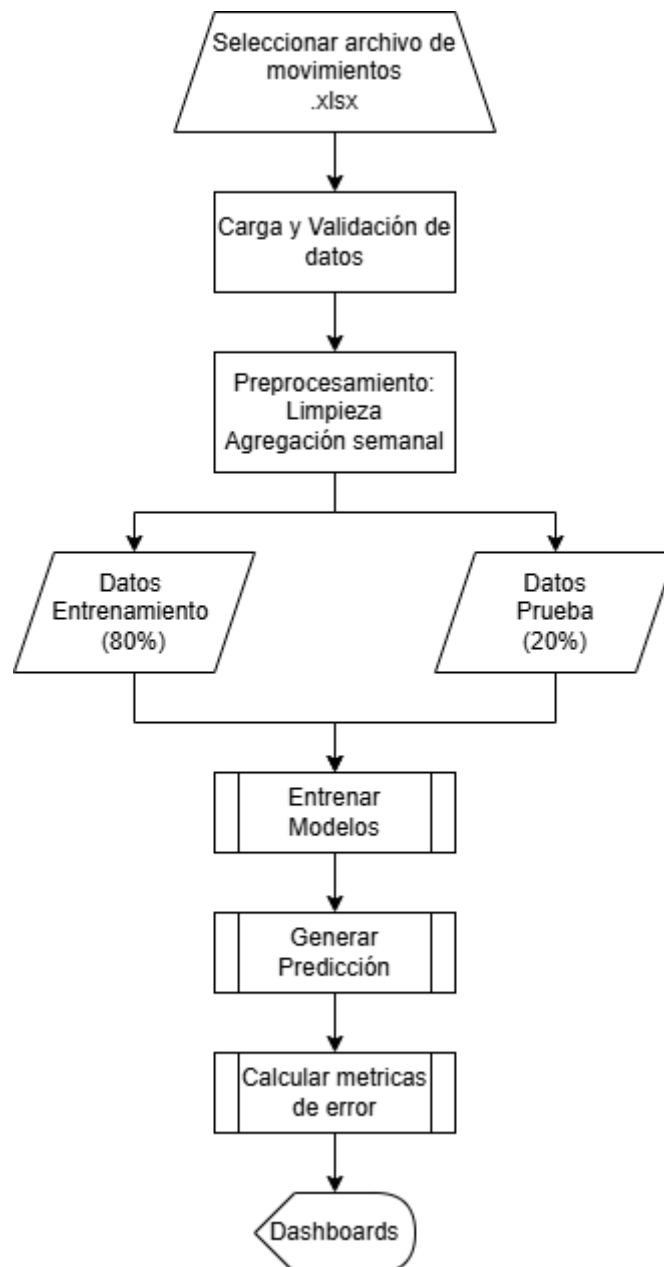


*Figura 7: Diagrama de Flujo Módulo de Machine Learning
Fuente: Elaboración Propia*

3.2.3. Tecnologías utilizadas

El módulo se desarrolló íntegramente en Python, empleando las librerías estándar del ecosistema científico de datos. A continuación, la Tabla 6: Librerías utilizadas en módulo ML detalla el conjunto de herramientas seleccionadas y describe la función específica que cumple cada una dentro de la arquitectura del sistema.

Tabla 6: Librerías utilizadas en módulo ML
Fuente: Elaboración Propia

Librería Framework	Función principal
pandas, numpy	Manipulación de estructuras de datos, operaciones vectoriales y cálculo matricial.
Statsmodels	Implementación del modelo ARIMA y herramientas estadísticas complementarias (ACF, PACF).
scikit-learn	Implementación del modelo Random Forest y cálculo de métricas de evaluación (MAE, RMSE, R ² , MAPE).
Xgboost	Implementación del modelo de Gradient Boosting optimizado (XGBoost).
Scipy	Aplicación de contrastes estadísticos no paramétricos, como el Test de Wilcoxon, para validación comparativa.

Esta combinación de librerías permite garantizar reproducibilidad, precisión estadística y eficiencia computacional, factores esenciales para la evaluación comparativa entre modelos.

3.2.4. Implementación funcional

El motor de predicción se encapsula en la clase *MLEngine*, compuesta por las siguientes funciones principales:

- `get_timeseries_for_item()`

Extrae los datos históricos correspondientes a un ítem específico, ordenándolos cronológicamente y manejando valores nulos o inconsistentes.

- `_create_features()`

Genera variables de entrada adicionales (lags, rolling means, desviaciones) para enriquecer los modelos basados en aprendizaje supervisado.

- `run_predictions()`

Ejecuta el ciclo completo de entrenamiento y predicción para los cuatro modelos en paralelo, aplicando el mismo esquema de evaluación.

- `calculate_metrics()`

Evalúa los resultados obtenidos mediante las métricas MAE, RMSE, MAPE y R^2 , registrando la precisión y robustez de cada algoritmo.

Este diseño modular permite integrar nuevos algoritmos sin alterar la estructura base, favoreciendo la extensibilidad del sistema y su adecuación a distintos tipos de demanda.

Las principales funciones de la clase *MLEngine* son:

- `get_timeseries_for_item()` y `_create_features()`: Pre procesamiento y generación de variables.
- `run_predictions()`: Entrenamiento y predicción simultánea de los modelos.
- `calculate_metrics()`: Evaluación comparativa mediante las métricas definidas.

Este diseño modular permite la comparación estandarizada entre algoritmos y la futura inclusión de nuevos modelos (como Prophet o LSTM) sin alterar la estructura principal.

3.2.5. Modelos disponibles e implementación

El módulo incorpora cuatro algoritmos principales, seleccionados para cubrir diferentes comportamientos de demanda observados en inventarios MRO. Su selección se fundamenta en el análisis metodológico descrito en el Capítulo 2.

1. ARIMA (AutoRegressive Integrated Moving Average):

Modelo estadístico clásico implementado con *statsmodels.tsa.arima.model.ARIMA*. Es idóneo para series con tendencia o estacionalidad débil, actuando como línea base para evaluar el desempeño de modelos más complejos.

2. Croston y variantes (TSB):

Implementado mediante una función personalizada (*croston_forecast()*), diseñada para series con demanda intermitente o esporádica. Este método descompone la serie en dos componentes (tamaño de demanda y tiempo entre demandas), permitiendo predicciones más estables para materiales de baja rotación.

3. Random Forest:

Algoritmo de ensamble implementado con *RandomForestRegressor* de *scikit-learn*. Su fortaleza radica en capturar relaciones no lineales entre las variables y adaptarse a variaciones abruptas de consumo sin requerir supuestos de distribución estadística.

4. XGBoost (Extreme Gradient Boosting):

Implementado con *XGBRegressor*, representa la versión más avanzada del módulo, capaz de manejar grandes volúmenes de datos y patrones altamente no lineales. Este modelo utiliza boosting secuencial para minimizar errores residuales, mostrando mayor desempeño en ítems con alta variabilidad y estacionalidad mixta.

3.2.6. Evaluación comparativa y registro de resultados

La interacción con el módulo se realiza desde la interfaz principal mostrada en la Figura 8, la cual permite al usuario seleccionar entre un análisis individual por material o una ejecución masiva (batch). Al finalizar la ejecución de los modelos, el sistema realiza una evaluación comparativa interna siguiendo los criterios definidos en la metodología:

1. Cálculo de métricas:

Se obtienen automáticamente los valores de MAE, RMSE, MAPE y R^2 para cada algoritmo.

2. Selección del modelo ganador:

Se aplica la metodología multicriterio definida en el Capítulo 2, priorizando el equilibrio entre precisión y estabilidad.

3. Estructura de resultados:

Los resultados se almacenan en memoria y se presentan al usuario mediante paneles interactivos.

- Análisis Individual: Como se observa en la Figura 9, el sistema grafica la serie histórica original junto con las predicciones de cada modelo, despliega la tabla de métricas específica para ese ítem y muestra un cuadro de texto con la recomendación del modelo ganador y su justificación analítica.
- Análisis Masivo (Batch): Para la toma de decisiones gerenciales, el sistema genera un reporte consolidado presentado en la Figura 10. Este panel resume el rendimiento global, agrupa los resultados por segmentos de demanda y muestra los tests de significancia estadística (Wilcoxon y Diebold-Mariano) para validar la superioridad del modelo seleccionado.

Esta salida es consumida por el Módulo de Visualización, logrando una integración total entre los procesos de análisis, decisión y monitoreo en tiempo real dentro de la misma aplicación.

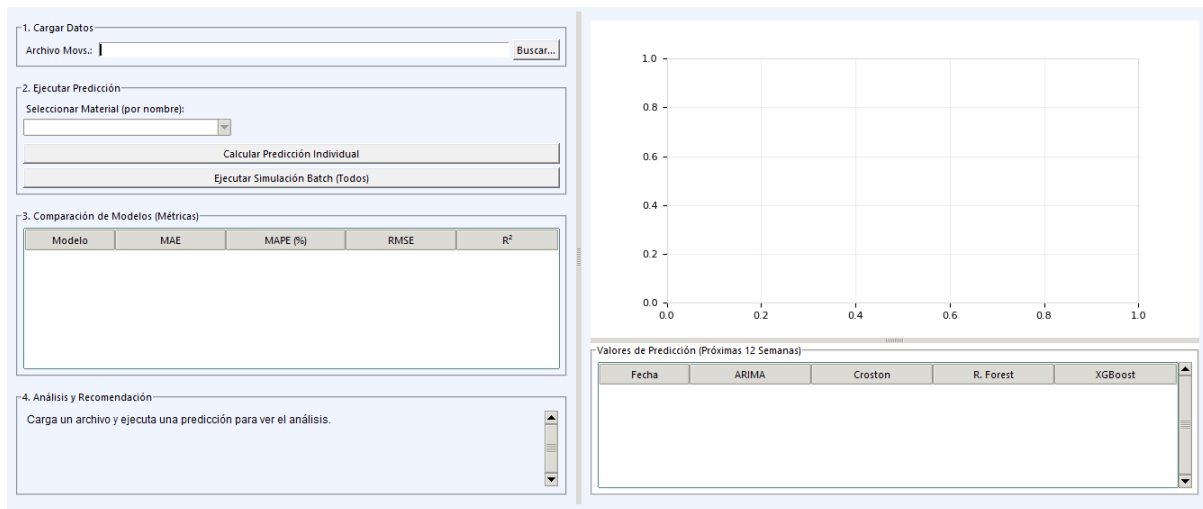


Figura 8: Pantalla de Inicio Machine Learning
Fuente: Elaboración Propia

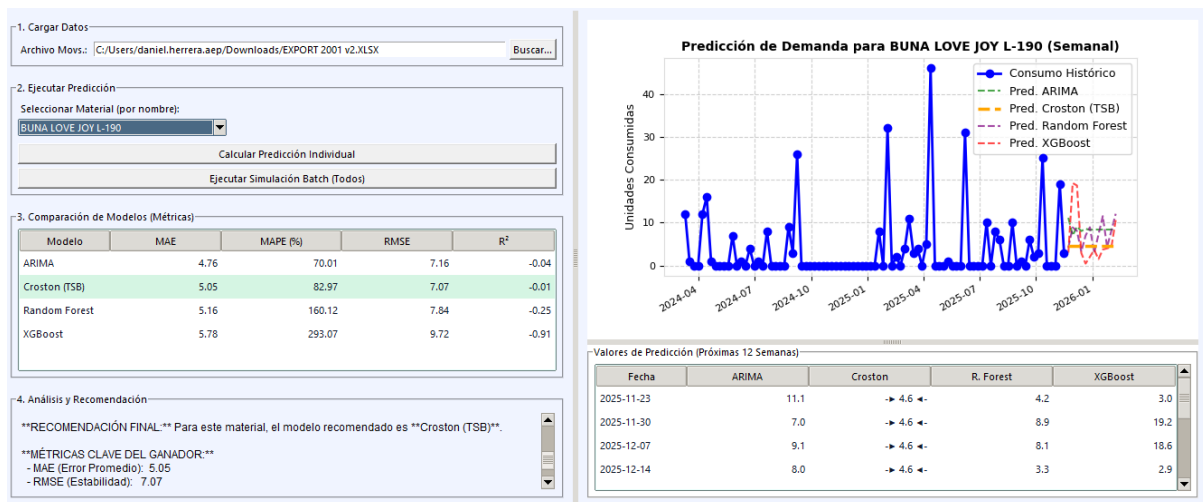


Figura 9: Pantalla de Machine Learning, Evaluando Buna Love Joy 095
Fuente: Elaboración Propia



Figura 10: Pantalla de Machine Learning en masa, Comparando Resultados
Fuente: Elaboración Propia

3.3. Módulo de Visualizaciones Dinámicas y Dashboards

El módulo de visualización constituye la capa de presentación e interacción del sistema inteligente de inventario. Su función principal es transformar los resultados analíticos del módulo de Machine Learning en información comprensible, dinámica y visualmente estructurada, permitiendo una interpretación inmediata de los datos.

A diferencia de soluciones que dependen de plataformas externas (como Power BI o Tableau), este módulo fue desarrollado dentro de la misma aplicación de escritorio, garantizando independencia operativa y actualización en tiempo real sin requerir conectividad externa.

3.3.1. Arquitectura general y componentes

El diseño del módulo se organiza en tres componentes funcionales principales, integrados directamente en la interfaz Tkinter:

1. Gestor de datos en memoria:

Se encarga de recibir y consolidar la información proveniente tanto del módulo de digitalización como del módulo analítico. Los datos se mantienen temporalmente en estructuras de pandas DataFrame, lo que permite actualizaciones instantáneas sin necesidad de archivos intermedios.

2. Motor gráfico:

Desarrollado con Matplotlib y embebido en Tkinter mediante el componente *FigureCanvasTkAgg*, este motor genera los gráficos directamente dentro de la interfaz. Su diseño garantiza que las visualizaciones se actualicen en tiempo real dentro del ciclo principal de la aplicación, manteniendo una interacción fluida.

3. Interfaz de usuario modular por pestañas:

La interfaz organiza sus funcionalidades en tres pestañas principales:

- Dashboard General (KPIs): vista consolidada de indicadores clave.
- Predicción de Demanda (ML): panel interactivo para simulaciones predictivas.
- Análisis de Valor de Inventario: herramienta de evaluación económica del stock actual.

Esta arquitectura integrada permite mantener una sincronización directa entre captura, análisis y visualización, alineada con los principios de sistemas ciberfísicos de la Industria 4.0.

3.3.2. Paneles Principales de Análisis

El módulo presenta tres paneles visuales diseñados para cubrir las dimensiones operativa, analítica y económica del sistema.

1. Dashboard General (KPIs)

Constituye el panel central del sistema, tal como se aprecia en la Figura 11. Este tablero resume los Indicadores Clave de Desempeño (KPIs) definidos en el Capítulo 2, ofreciendo una visión global del estado del inventario. Cada indicador se representa mediante tarjetas de resumen en la parte superior y gráficos dinámicos (barras, líneas o dona) en la sección inferior, los cuales se recalculan automáticamente al modificarse los datos subyacentes. Entre los KPIs más relevantes se incluyen:

- Tiempo promedio de gestión de inventario,
- Nivel de disponibilidad de stock crítico,
- Días sin servicio por quiebre de material,
- Variación del stock promedio respecto al óptimo.

2. Panel de Predicción de Demanda (ML)

Es la herramienta principal para el análisis predictivo dentro de la aplicación. Permite al usuario:

- Seleccionar un ítem específico o ejecutar simulaciones en lote (batch) sobre múltiples materiales.

- Visualizar las curvas de consumo histórico superpuestas con las predicciones de los modelos ARIMA, Croston, Random Forest y XGBoost.
- Comparar los resultados mediante una tabla de métricas que incluye MAE, RMSE, MAPE y R^2 .
- Revisar un panel de texto interpretativo que destaca el modelo con mejor desempeño y su justificación analítica.

3. Panel de Análisis de Valor de Inventario

Proporciona una perspectiva económica del sistema, ilustrada en la Figura 12. Este módulo calcula y muestra el valor monetario del inventario actual a partir de:

- Cantidad de unidades en stock.
- Costo unitario promedio (derivado de los datos de ingreso).
- Valor total por material.

Esta información se presenta tanto en tablas de detalle como en visualizaciones comparativas, facilitando la identificación de los materiales de mayor impacto financiero.



Figura 11: Pantalla de Inicio Dashboard y Métricas de la aplicación
Fuente: Elaboración Propia

Dashboard de Análisis de Inventario

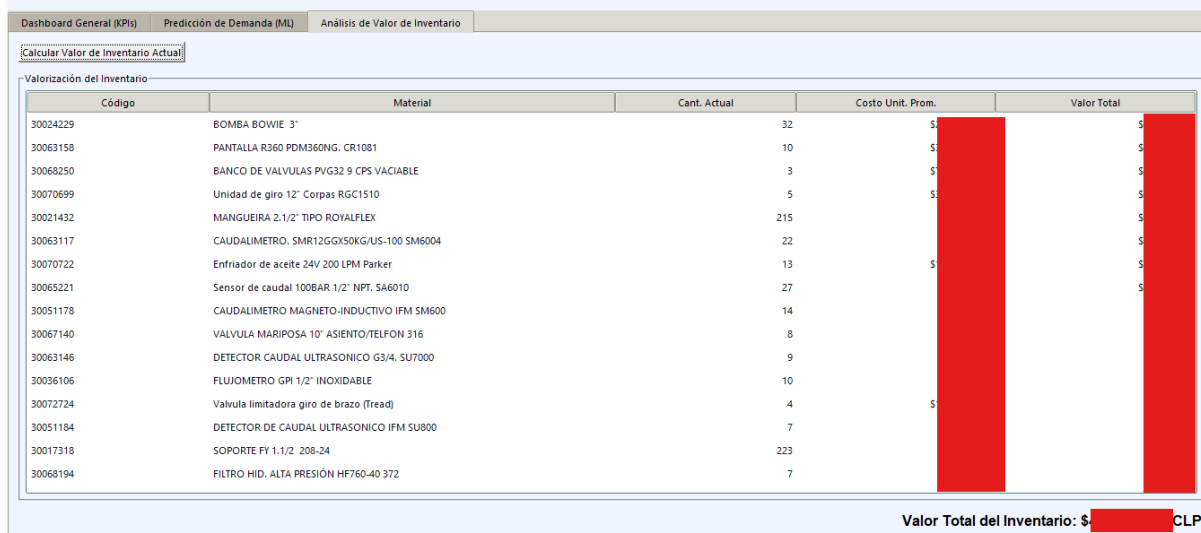


Figura 12: Pantalla de Análisis de Valores, Censurando Datos Sensibles
Fuente: Elaboración Propia

3.4. Integración de módulos

La integración de los módulos desarrollados constituye el núcleo funcional del sistema. Permite que los procesos de captura de datos, análisis predictivo y visualización operen de forma coordinada dentro de una única aplicación de escritorio, consolidando una plataforma confiable y de bajo acoplamiento.

A diferencia de sistemas fragmentados que dependen de múltiples plataformas o conexiones externas, el prototipo fue diseñado como un ecosistema autocontenido. Este enfoque reduce la latencia y elimina la duplicidad de registros, ya que los módulos comparten estructuras de datos en memoria y se comunican mediante procesos controlados por la interfaz principal.

3.4.1. Arquitectura de integración y flujo de datos

La aplicación se organiza bajo una arquitectura de tres capas funcionales que operan dentro del mismo proceso, conectadas mediante el intercambio de estructuras de datos de Pandas y callbacks de la interfaz gráfica.

1. Capa de Adquisición (Módulo de Digitalización):

Esta capa es la interfaz de usuario principal para el registro de movimientos. Un evento, como el escaneo de un código QR o el ingreso manual de un movimiento, es capturado por la interfaz de Tkinter. La información (código de ítem, cantidad, OT) es procesada y registrada en los archivos locales *movimientos.csv* (log de transacciones) y *MATERIALES.xlsx* (actualización de stock).

2. Capa de Procesamiento (Módulo de Machine Learning):

Esta capa se activa cuando el usuario interactúa con la pestaña de "Predicción de Demanda". No recibe eventos en tiempo real del Módulo QR, sino que opera a demanda del usuario:

- Entrada: Lee los datos históricos consolidados desde el archivo Excel de movimientos.
- Procesamiento: Ejecuta el ciclo de entrenamiento y evaluación de los modelos predictivos.
- Salida: Devuelve los resultados (predicciones, métricas, etc.) como un objeto en memoria (diccionario de Python).

3. Capa de Visualización (Dashboards Internos):

Esta capa es responsable de renderizar la información. Recibe los resultados de la capa de procesamiento y actualiza los componentes de la interfaz de Tkinter, como los gráficos de Matplotlib y las tablas *Treeview*. La actualización es inmediata porque ocurre dentro del mismo *mainloop* de la aplicación.

Tal como se esquematiza en la Figura 13, la comunicación entre capas se gestiona mediante hilos concurrentes (threads), especialmente para las tareas de larga duración como la simulación batch y la operación de la cámara. Esto mantiene la interfaz fluida y receptiva, evitando que se 'congele' durante operaciones de alto costo computacional.

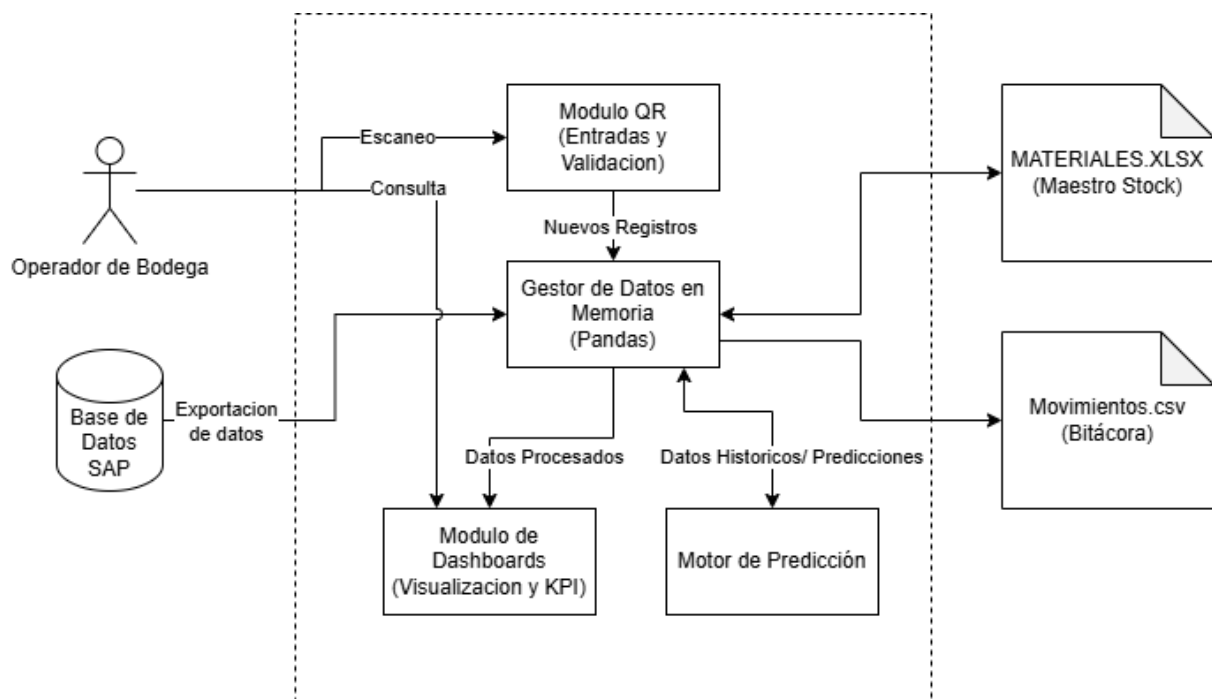


*Figura 13: Arquitectura general del sistema
Fuente: Elaboración propia*

3.4.2. Flujo integrado de operación

El ciclo completo de funcionamiento del sistema puede resumirse en dos flujos de trabajo principales, representados por la Figura 14 y la Figura 15:

Flujo 1: Registro de un Movimiento de Inventario

1. Entrada de Datos: El usuario escanea un código QR o busca un material manualmente en las pantallas de "Ingreso" o "Salida".
2. Procesamiento y Persistencia: El sistema valida la información y actualiza el stock en el archivo *MATERIALES.xlsx*. La transacción se registra con un sello temporal en el archivo *movimientos.csv*.
3. Retroalimentación Visual: El movimiento aparece inmediatamente en las tablas de historial de la aplicación, y los KPIs del dashboard se actualizan si es relevante.

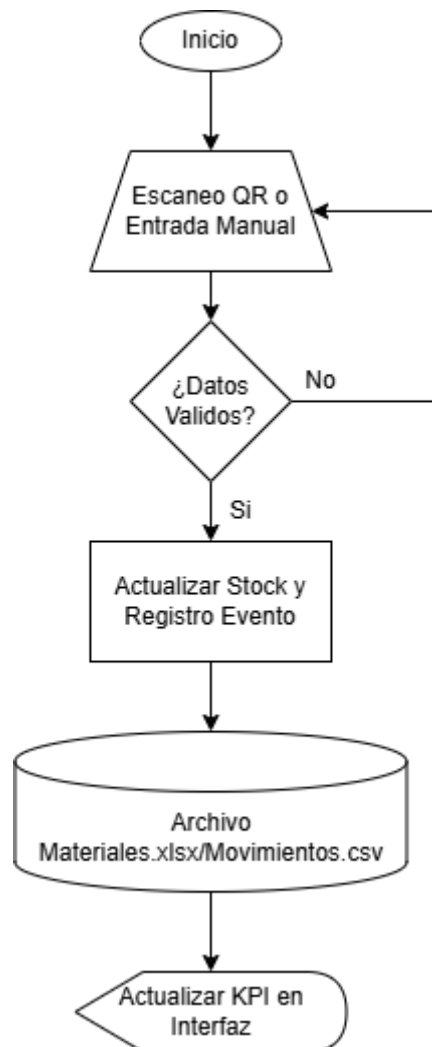


Figura 14: Diagrama Flujo 1
Fuente: Elaboración Propia

Flujo 2: Ejecución de una Simulación Predictiva

1. Entrada de Datos: El usuario carga un archivo Excel con el historial completo de movimientos en la pestaña "Predicción de Demanda (ML)".
2. Procesamiento Predictivo: Al ejecutar una simulación, el Módulo de Machine Learning lee los datos, los procesa y ejecuta los cuatro algoritmos de pronóstico.

3. Visualización de Resultados: Los resultados (curvas, métricas, recomendaciones) se envían como un objeto al Módulo Visual, que actualiza de inmediato el dashboard de predicción.

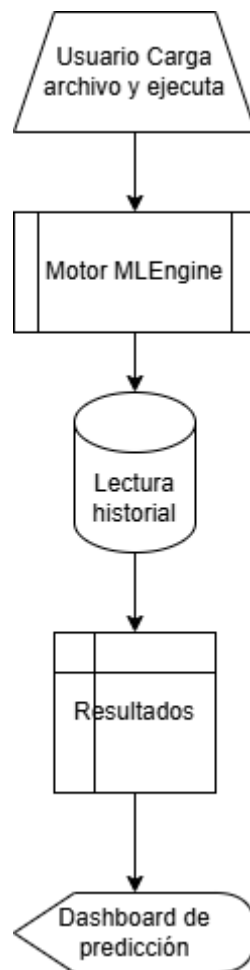


Figura 15: Diagrama Flujo 2
Fuente: Elaboración propia

Este esquema integrado consolida el sistema como una plataforma de análisis local, capaz de unificar trazabilidad física (mediante archivos) y modelado predictivo dentro de un único entorno operativo.

3.4.3. Sincronización y consistencia de datos

Para garantizar la coherencia dentro de este sistema basado en archivos, el prototipo utiliza un esquema de sincronización simple pero efectivo:

- Sellos Temporales (Timestamps): Cada registro en movimientos.csv incluye la fecha y hora exacta del evento, lo que permite reconstruir el historial cronológico del inventario.
- Fuente Única de Verdad: El archivo *MATERIALES.xlsx* actúa como la base de datos central del estado del stock, mientras que movimientos.csv es el registro inmutable de todas las transacciones.

- Callbacks de Interfaz: Funciones internas (*_batch_refresh_view*, *_update_prediction_display*) se encargan de refrescar las visualizaciones automáticamente cada vez que los datos subyacentes son modificados por una acción del usuario.

No existen mecanismos de rollback automático ni sincronización con bases de datos externas como SQLite o SAP, ya que la operación del prototipo está contenida en su entorno local de archivos.

La integración de los módulos en una única arquitectura monolítica no solo optimiza el flujo de información, sino que también garantiza la trazabilidad completa de los datos desde su captura hasta su interpretación final. Este acoplamiento controlado reduce errores humanos, elimina redundancias y mejora la coherencia temporal entre los procesos de digitalización, predicción y visualización. En consecuencia, la integración contribuye directamente a los KPI 3 y KPI 5, al disminuir los días sin servicio por falta de stock y al estabilizar la variación del inventario frente a los niveles óptimos. Con ello, el sistema alcanza una sinergia funcional que cierra el ciclo de la gestión de inventario inteligente, integrando observación, análisis y acción dentro de un mismo entorno operativo.

Capítulo 4. Resultados

El prototipo desarrollado fue sometido a un proceso de evaluación funcional y de desempeño predictivo. Las pruebas se realizaron utilizando un conjunto de datos históricos de consumo, extraídos de SAP, correspondientes a 230 materiales con suficiente historial de demanda para un análisis robusto.

4.1. Evaluación funcional y de integración

Se verificó la correcta operatividad de los módulos del sistema en un entorno de pruebas controlado. Los resultados obtenidos se presentan en la Tabla 7, los cuales confirman la coherencia de la arquitectura integrada y la estabilidad operativa del prototipo.

*Tabla 7: Resultados Evaluación Funcional
Fuente: Elaboración Propia*

Parámetro Evaluado	Resultado Observado
Tiempo promedio de registro (vía QR)	~7 segundos por ítem
Integración de datos (carga de archivos)	Correcta e instantánea
Estabilidad de la interfaz (UI)	Rendimiento estable, sin bloqueos
Latencia en renderizado de gráficos	~0.8 segundos promedio
Incidencia de errores críticos	0%

Estos resultados demuestran la viabilidad técnica del prototipo, su capacidad para reducir tiempos operativos y su estabilidad para el manejo de los flujos de datos y visualizaciones.

Además de las métricas cuantitativas, la implementación del prototipo generó una reingeniería cualitativa del flujo de trabajo. Al contrastar el mapa de procesos original (Figura 1) con el flujo optimizado tras la implementación del sistema planteado en la Figura 16, se evidencian tres mejoras estructurales clave:

- **Digitalización del Punto de Control:**
Como se observa en el diagrama, tanto el flujo de Entrada (tras la recepción y revisión física) como el de Salida (vinculado a la Solicitud y OT) convergen obligatoriamente en la etapa de "Escaneo QR". Este paso actúa como una compuerta digital que valida la transacción en tiempo real, abarcando incluso subprocesos complejos como envíos y devoluciones, asegurando que ningún movimiento físico ocurra sin su contraparte digital.
- **Automatización del Registro:**
A diferencia del proceso anterior donde el "Registro" dependía de la digitación manual posterior, en el nuevo flujo el bloque de "Registro" es una consecuencia automática del escaneo. Esto alimenta simultáneamente al sistema base (SAP/Archivos locales) y garantiza la integridad de los datos, eliminando los errores de transcripción humana.
- **Extensión hacia la Inteligencia Analítica:**
La diferencia más significativa es la adición de una nueva rama operativa a la derecha del diagrama. El proceso no termina en el almacenamiento del dato; el registro activa la etapa de "Predicción". Esto permite pasar de un esquema meramente transaccional a uno analítico, donde la información fluye hacia el "Análisis" y culmina en una "Toma de decisiones" informada, cerrando el ciclo de valor del sistema inteligente.

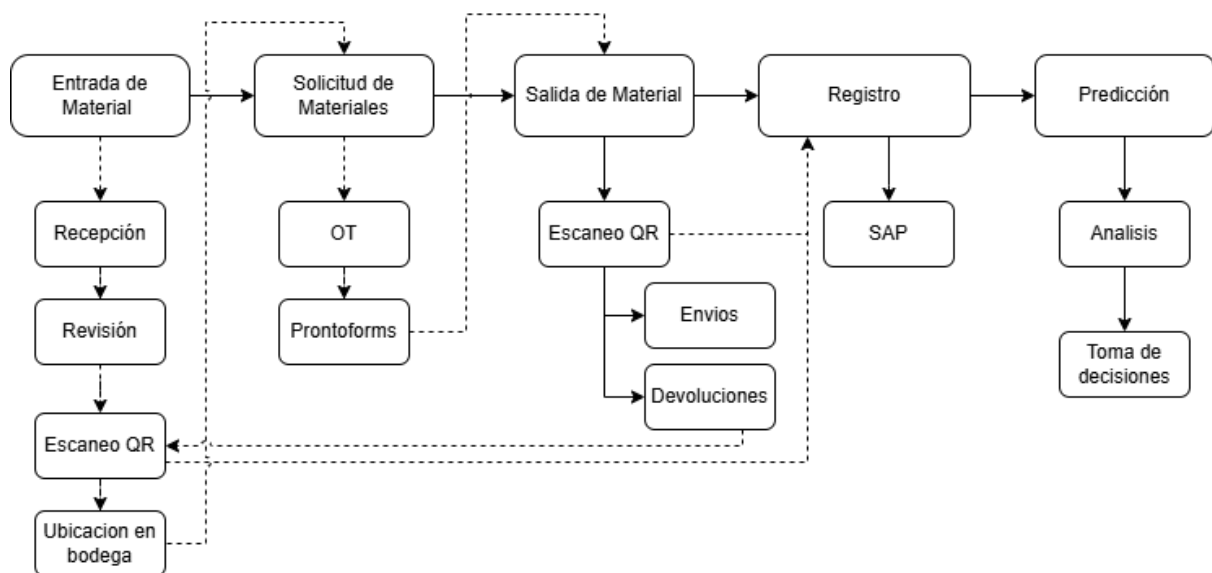


Figura 16: Nuevo Mapa de procesos
Fuente: Elaboración propia

4.2. Evaluación del Desempeño predictivo

El núcleo de la evaluación consistió en una comparación "todos contra todos" de los cuatro algoritmos implementados, siguiendo la metodología descrita en el Capítulo 2.

4.2.1. Comparativa Global de Modelos

Contrario a la expectativa inicial de que los modelos de Machine Learning dominarían las predicciones, el análisis global reveló que el modelo estadístico ARIMA fue el de mejor rendimiento general. Como se detalla en la Tabla 8, este algoritmo fue seleccionado como la

opción óptima para la mayor cantidad de materiales con 107 ítems, lo que representa un 46.5% del total y supera a las alternativas más complejas.

Tabla 8: Resumen global de modelos ganadores sobre 230 ítems evaluados
Fuente: Elaboración Propia

Modelo	N.º de Ítems (Mejor Modelo)	Porcentaje del Total
ARIMA	107	46.5%
Croston (TSB)	50	21.7%
XGBoost	44	19.1%
Random Forest	27	11.7%
Datos Insuficientes	2	0.9%

Este resultado subraya la importancia de no descartar los modelos clásicos, especialmente en contextos de inventarios MRO, y motivó un análisis más profundo a nivel de segmentos de demanda para comprender la causa de este desempeño.

4.2.2. Análisis de Rendimiento por Segmento

La segmentación analítica de los 230 ítems arrojó un resultado clave: la demanda del inventario es mayoritariamente intermitente.

Tabla 9: Análisis de rendimiento por segmento de demanda
Fuente: Elaboración Propia

Segmento de Demanda	N.º de Ítems	Porcentaje del Total	Modelo Más Frecuente	MAE Promedio
Intermitente	223	97.0%	ARIMA	2.70
Irregular	5	2.2%	Croston (TSB)	15.50
Indeterminado	2	0.9%	Datos Insuficientes	-1.00

La Tabla 9 revela que el dominio de ARIMA en los resultados globales se debe a su excelente desempeño en el segmento intermitente, que representa el 97% de los materiales analizados. Aunque Croston fue diseñado para este tipo de demanda, la implementación de ARIMA en la librería statsmodels demostró ser más robusta y precisa para este conjunto de datos específico. Este hallazgo es fundamental, ya que valida la selección de ARIMA como el modelo más confiable para el inventario de la empresa.

4.2.3. Contrastes Estadísticos

Para validar formalmente la superioridad del modelo ARIMA, se aplicaron los tests de Wilcoxon y Diebold-Mariano de forma pareada entre todos los algoritmos. Este procedimiento determina si las diferencias observadas en el error de predicción son estadísticamente significativas o si podrían deberse a la variabilidad inherente de los datos.

Tabla 10: Resultados de los contrastes estadísticos pareados
() indica una diferencia estadísticamente significativa ($p < 0.05$).*
Fuente: Elaboración Propia

Comparación	p-value Wilcoxon	p-value Diebold-Mariano
ARIMA vs. Croston (TSB)	0.0000 *	0.0095 *
ARIMA vs. Random Forest	0.0000 *	0.0010 *
ARIMA vs. XGBoost	0.0000 *	0.0000 *
Croston (TSB) vs. Random Forest	0.4616	0.5361
Croston (TSB) vs. XGBoost	0.1509	0.0586
Random Forest vs. XGBoost	0.0000 *	0.0002 *

Los resultados, presentados en la Tabla 10, son concluyentes.

El análisis de los p-values revela dos hallazgos clave:

1. Superioridad de ARIMA: Las tres primeras filas de la tabla muestran que los p-values para las comparaciones de ARIMA contra todos los demás modelos son muy inferiores a 0.05. Esto confirma con alta confianza estadística que la ventaja en precisión de ARIMA es real y no producto del azar.
2. Equivalencia entre los otros modelos: En las comparaciones entre Croston, Random Forest y XGBoost, los p-values son, en su mayoría, superiores a 0.05. Esto sugiere que, si bien hay diferencias en sus errores promedio (MAE), estas no son estadísticamente significativas. En la práctica, su desempeño predictivo es equivalente para este conjunto de datos.

Estos resultados proporcionan una validación empírica sólida para la selección de ARIMA como el algoritmo de pronóstico por defecto para el prototipo, al ser el único modelo que demostró una superioridad clara y estadísticamente significativa.

4.3. Proyección del Impacto sobre los KPIs Operacionales

Con el fin de vincular los hallazgos técnicos con los objetivos operativos del negocio, se proyecta el impacto de los resultados de la simulación sobre los Indicadores Clave de Desempeño (KPIs) definidos en el Capítulo 2.

- Reducción de Quiebres de Stock (KPI 3 - Días sin servicio):

El resultado más significativo del estudio es la identificación de ARIMA como el modelo predictivo con el menor error y la mayor consistencia para el 97% del inventario analizado. Un pronóstico más preciso se traduce directamente en señales de reposición más fiables. Esto permite anticipar las necesidades reales de consumo con mayor antelación, lo que proyecta una disminución directa en la cantidad de quiebres de stock y, por consiguiente, en los días de inactividad operacional.

- Optimización del Capital Inmovilizado (KPI 5 - Desviación del stock óptimo):

Al contar con un modelo de pronóstico validado y estadísticamente superior, la empresa puede definir niveles de stock óptimo (S^*) y de seguridad (S_{min}) con mayor confianza. El uso de ARIMA permite mantener niveles de inventario más ajustados a la demanda real, reduciendo los costos asociados al sobre stock (capital inmovilizado, almacenamiento) sin incrementar el riesgo de quiebres.

- Eficiencia en las Cuadraturas y Conciliaciones (KPI 2 - Horas-Hombre en ajustes):

Una predicción de consumo más precisa reduce las discrepancias entre el inventario físico y el registrado en el sistema a lo largo del tiempo. Esto implica que las cuadraturas periódicas serán más rápidas y requerirán menos ajustes, proyectando una reducción significativa en las horas-hombre destinadas a estas tareas correctivas.

- Mejora en la Disponibilidad de Stock Crítico (KPI 4 - Disponibilidad de stock crítico):

Al aplicar un modelo robusto como ARIMA, se mejora la estabilidad del sistema de gestión. Esto se traduce en una mayor capacidad para mantener los ítems críticos por encima de su nivel mínimo de seguridad, incrementando el indicador de disponibilidad y asegurando la continuidad operativa.

En conjunto, estos resultados confirman que el prototipo no solo es técnicamente viable, sino que, al identificar y validar el modelo predictivo más adecuado para el contexto específico de la empresa, aporta un valor operativo tangible y directamente medible a través de los KPIs del negocio.

Capítulo 5. Conclusiones

En este capítulo se presentan las conclusiones principales derivadas del desarrollo y evaluación del prototipo de sistema inteligente para la gestión de inventario. Se abordan los hallazgos clave en relación con los objetivos planteados, se sintetiza el aporte del trabajo y se proponen líneas de mejora para futuras iteraciones.

5.1. Cumplimiento de los Objetivos del Proyecto

El objetivo principal de desarrollar un prototipo funcional que integrara digitalización, predicción y visualización fue cumplido exitosamente. El sistema implementado demostró ser una herramienta coherente y operativa, validando la viabilidad de la arquitectura propuesta. Respecto a los objetivos específicos, su cumplimiento se detalla a continuación:

- Identificación de procesos críticos:

Se analizó el flujo operativo del inventario, determinando que el registro manual de movimientos y la falta de trazabilidad en tiempo real constituían los puntos de mayor impacto negativo (cuellos de botella). La digitalización de estos procesos mediante códigos QR resultó ser la intervención más crítica del proyecto, logrando reducir los tiempos de gestión en aproximadamente un 80% y eliminando los errores de transcripción humana.

- Evaluación de modelos predictivos:

Se implementó y evaluó sistemáticamente un conjunto de cuatro modelos. Tras la validación empírica, se seleccionó el modelo ARIMA como el motor predictivo principal del sistema, dado que demostró un desempeño superior en el 46.5% de los ítems evaluados. Esta elección se fundamenta en su capacidad para manejar la demanda intermitente del inventario MRO con mayor robustez y menor error promedio que los algoritmos de Machine Learning complejos.

- Implementación de dashboards:

Se desarrolló un panel de control interactivo dentro de la misma aplicación, que integra KPIs, resultados de predicción y análisis de segmentos, facilitando una toma de decisiones informada.

5.2. Hallazgo Principal: Idoneidad del Modelo ARIMA para Demanda Intermitente

La conclusión más relevante del estudio es la superioridad estadísticamente significativa del modelo ARIMA en el contexto del inventario MRO analizado. A pesar de la creciente popularidad de los algoritmos de Machine Learning, ARIMA fue seleccionado como el mejor modelo para el 46.5% de los ítems, superando a todos sus competidores.

Este resultado contradice la tendencia general de la literatura reciente, donde los algoritmos de Machine Learning suelen superar a los métodos clásicos. Sin embargo, un análisis crítico de los datos revela que esta “anomalía” responde a tres factores estructurales del inventario analizado:

1. Ausencia de Estacionalidad Compleja:

Modelos como XGBoost o Random Forest destacan principalmente cuando existen patrones estacionales no lineales o múltiples ciclos superpuestos. El inventario MRO

de la empresa, al ser predominantemente intermitente (97%) y reactivo a fallas, carece de esta estructura cíclica definida, lo que anula la principal ventaja comparativa de los algoritmos complejos.

2. Relación Señal/Ruido y Sobreajuste:

En series temporales cortas y con alta presencia de ceros, los algoritmos de ensamble tienden a confundir el “ruido” aleatorio de la demanda con patrones reales, incurriendo en sobreajuste. Por el contrario, la parsimonia del modelo ARIMA actuó como un filtro más eficiente, suavizando la variabilidad sin intentar modelar falsos patrones en la intermitencia.

3. Tamaño del Dataset:

Aunque 230 materiales son suficientes para una validación estadística, constituyen una muestra pequeña para el entrenamiento profundo de modelos de Inteligencia Artificial ("Small Data"). La evidencia sugiere que la complejidad de un modelo de Gradient Boosting no se justifica sin un volumen masivo de datos históricos que permita generalizar reglas de comportamiento más allá de la estadística lineal.

5.3. Limitaciones y Alcance del Estudio

Es importante establecer las delimitaciones técnicas y metodológicas que enmarcaron el desarrollo de este prototipo. En primer lugar, debido a las restricciones temporales propias del proyecto de memoria, no fue posible implementar una integración bidireccional en tiempo real con el sistema SAP (vía RFC/BAPI). Por consiguiente, la solución opera bajo un esquema de actualización por lotes (batch), lo cual valida la lógica de procesamiento, pero difiere de la inmediatez de un entorno productivo integrado.

En segundo lugar, el análisis histórico para el entrenamiento de los modelos predictivos se acotó a un ciclo operativo de dos años. Esta decisión responde a la inconsistencia estructural de los datos anteriores a dicho periodo, los cuales fueron generados bajo protocolos de ingreso manual distintos y heterogéneos que impedían su homologación confiable para algoritmos de Machine Learning.

Finalmente, si bien la validación se llevó a cabo en un entorno de pruebas y no en producción continua, este proceso se realizó bajo el control y supervisión directa de los actores reales del sistema (personal de mantenimiento y bodega). Esto garantizó que, a pesar de ser una simulación controlada, los flujos de trabajo y la usabilidad del prototipo fueran validados por quienes efectivamente utilizarían la herramienta, asegurando su alineación con la realidad operativa de la empresa.

5.4. Validación del Impacto Operacional y Aporte del Prototipo

El prototipo demostró ser una solución técnicamente viable con un claro potencial de impacto operativo. La evaluación funcional del módulo de digitalización validó una reducción del tiempo de registro superior al 80%.

El aporte principal del proyecto radica en haber aplicado una metodología rigurosa para seleccionar la herramienta analítica más adecuada para el problema real de la empresa. Al identificar y validar a ARIMA como el modelo más eficaz, el sistema proporciona una base sólida para mejorar directamente los KPIs de negocio, proyectando una reducción de quiebres

de stock (KPI 3), una optimización del capital inmovilizado (KPI 5) y una mejora en la disponibilidad de ítems críticos (KPI 4).

5.5. Líneas de Trabajo Futuro

A partir de los resultados y las limitaciones identificadas en el prototipo actual, se proponen las siguientes líneas de trabajo futuro para evolucionar el sistema hacia una solución de nivel productivo:

1. Integración Directa con SAP:

La principal limitación del prototipo es su dependencia de la carga manual de archivos. El siguiente paso crucial es incorporar un motor de integración directa con SAP mediante servicios RFC/BAPI. Esto permitiría una sincronización bidireccional y automática de los datos, eliminando la intervención manual y habilitando la retroalimentación de pronósticos al sistema ERP.

2. Evolución hacia una Arquitectura Cliente-Servidor:

Para facilitar el acceso multiusuario y la escalabilidad, se propone extender el sistema hacia una arquitectura distribuida. Esto implicaría desarrollar un backend (utilizando frameworks como Flask o FastAPI) que aloje el motor de ML y la lógica de negocio, y una interfaz web (frontend) que permita el acceso desde cualquier dispositivo.

3. Implementación de Metodologías de Validación Avanzadas:

Para profundizar en el análisis de robustez, se recomienda implementar backtesting con origen rodante. Esta técnica permitiría evaluar la estabilidad temporal de los modelos, es decir, cómo varía su desempeño a lo largo de distintas ventanas de tiempo, fortaleciendo la confiabilidad del modelo seleccionado.

4. Desarrollo de un Módulo de Optimización Prescriptiva:

Para cerrar el ciclo de gestión, se propone construir un nuevo módulo que utilice los pronósticos de demanda para generar recomendaciones de acción. Esto implicaría integrar modelos de optimización de inventarios (ej. Cálculo de Cantidad Económica de Pedido - EOQ y Punto de Reorden - ROP) para automatizar las sugerencias de compra y reposición.

Bibliografía

- Aktepe, A., Türe, M., & Tükenmez, I. (2023). *Demand Forecasting of Spare Parts in the Public Transport Sector Using Machine Learning Algorithms*. *Alexandria Engineering Journal*, 67, 173–184. <https://doi.org/10.1016/j.aej.2023.04.016>
- Chen, J., Gusikhin, O., Finkenstaedt, W., & Liu, Y. (2019). Maintenance, Repair, and Operations Parts Inventory Management in the Era of Industry 4.0. *IFAC-PapersOnLine*, 52(13), 171–176. <https://doi.org/10.1016/j.ifacol.2019.11.171>
- Fiorentia, C., Ponis, S., & Spanos, N. (2021). *Predicting the Performance of Forecasting Strategies for Naval Spare Parts Demand: A Machine Learning Approach*. *Procedia CIRP*, 99, 647–652. <https://doi.org/10.1016/j.procir.2021.03.108>
- Ghobakhloo, M., & Fathi, M. (2019). Corporate adoption of digitalization in supply chain management: Evidence from SMEs. *Computers in Industry*, 106, 70–83. <https://doi.org/10.1016/j.compind.2019.01.002>
- Ifraz, M., Gunawan, I., & Jiang, Z. (2023). *Forecasting Spare Parts Sporadic Demand Using Traditional Methods and Machine Learning: A Comparative Study*. *SMU Data Science Review*, 3(2). <http://scholar.smu.edu/datasciencereview/vol3/iss2/9>
- Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2018). Smart manufacturing: Characteristics, technologies and enabling factors. *Journal of Manufacturing Systems*, 48, 144–156. <https://doi.org/10.1016/j.jmsy.2018.01.003>
- Praveen, B. N. K. (2020). Inventory Management Using Machine Learning. *International Journal of Engineering Research and Technology (IJERT)*, 9(6). <https://doi.org/10.17577/ijertv9is060661>
- Siponen, M., Haapasalo, H., & Harkonen, J. (2019). Maintenance, Repair, and Operations Inventory Reduction and Operational Development. *International Journal of Industrial and Systems Engineering*, 32(1), 1–20. <https://doi.org/10.1504/IJISE.2019.099780>
- Vidal, G. H. P., Caiado, R. G. G., Scavarda, L. F., & Santos, R. S. (2022). MRO Inventory Demand Forecast Using Support Vector Machine: A Case Study. En *Springer Proceedings in Mathematics & Statistics* (pp. 221–233). https://doi.org/10.1007/978-3-031-14763-0_18
- Wamba, S. F., Queiroz, M. M., & Trinchera, L. (2020). Big data analytics-enabled digital transformation: A research agenda. *Information & Management*, 57(1), 103237. <https://doi.org/10.1016/j.im.2019.103237>
- Wits, W. W., Rijks, T. G. S. M., & van Es, J. (2016). Additive Manufacturing of Spare Parts for Inventory Reduction. *Procedia CIRP*, 55, 220–225. <https://doi.org/10.1016/j.procir.2016.07.010>

- Zhang, Y., Wang, Y., & Wang, L. (2012). A framework for maintenance management using RFID and complex event processing. *Journal of Intelligent Manufacturing*, 23, 2343–2356. <https://doi.org/10.1007/s10845-010-0482-2>
- Ahmadvov, Y., & Helo, P. (2023). Demand forecasting of intermittent e-commerce sales with deep learning methods. *AI and Ethics*, 4, 251–264. <https://doi.org/10.1007/s44163-023-00085-1>
- Ç. Pinçe, S., Basten, R. J., & Tan, T. (2021). A quantitative comparison of spare parts demand forecasting methods. *European Journal of Operational Research*, 293(1), 235–253. <https://doi.org/10.1016/j.ejor.2020.12.037>
- Desai, A. (2025). Inventory management with QR-based PWA and geolocation. *arXiv preprint arXiv:2506.11011*. <https://arxiv.org/abs/2506.11011>
- Hasani, Y. (2023). Comparative analysis of Croston-based methods and Machine Learning algorithms for intermittent demand. Master's Thesis, Erasmus University Rotterdam. <https://thesis.eur.nl/pub/70009>
- Noguerra, C. P., Jr. (2023). Design and Evaluation of an Innovative Mobile Solution: QR Code-based Inventory Monitoring System. *International Journal Of Advanced Engineering And Management Research*, 08(03), 231-242. <https://doi.org/10.51505/ijaemr.2023.8317>
- Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2017). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *arXiv preprint arXiv:1704.04110*. <https://arxiv.org/abs/1704.04110>
- Şahin, M., Çimen, M., & Eren, T. (2013). Forecasting aviation spare parts demand using Croston-based methods and artificial neural networks. *International Journal of Production Research*, 51(14), 4210–4224. <https://doi.org/10.1080/00207543.2012.754539>
- Taj, M. (2023). Internet of Things applications in supply chain management: A review. *Materials Today: Proceedings*, 81, 1393–1400. <https://doi.org/10.1016/j.matpr.2023.03.061>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley. <https://doi.org/10.1002/9781118619193>
- Gardner, E. S. (2006). Exponential smoothing: The state of the art—Part II. *International Journal of Forecasting*, 22(4), 637–666. <https://doi.org/10.1016/j.ijforecast.2006.03.005>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. <https://otexts.com/fpp2/>
- Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling* (3rd ed.). Wiley.
- Croston, J. D. (1972). Forecasting and Stock Control for Intermittent Demands. *Journal Of The Operational Research Society*, 23(3), 289-303. <https://doi.org/10.1057/jors.1972.50>
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal Of Business And Economic Statistics*, 13(3), 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. <http://jmlr.org/papers/v7/demsar06a.html>