



Universidad
Popular del Cesar

PROYECTO DE GRADO



Universidad
Popular del Cesar

PROYECTO DE GRADO DE INGENIERÍA DE SISTEMAS

PREPARADO POR:

**Luis Fabian Contreras Maldonado
Daniel Ricardo Piña Corzo**

Febrero de 2026

Tabla de contenido

SECCIÓN I: Descripción de Proyecto Grado	3
1.1 Título del Proyecto de Grado	3
1.2 Dirección de Ejecución	3
1.3 Lapso de Ejecución	3
1.4 Organismo o Institución Responsable del Proyecto	3
1.5 Información de contacto de los estudiantes	3
1.6 Línea, sublínea y grupo de investigación del Proyecto	4
SECCIÓN II: Descripción Situacional	4
2.1 Identificación del Problema	4
2.2 Justificación del Proyecto	7
2.3 Objetivos del Proyecto	11
2.4 Bases Teóricas	12
2.4.1 Antecedentes	13
2.4.2 Marco Teórico	19
2.4.3 Marco Conceptual	22
2.5 Marco Metodológico	24
2.5.1 Tipo de investigación	24
2.5.2 Enfoque del proyecto	25
2.5.3 Población y muestra	26
2.5.4 Técnicas de recolección de datos	28
2.5.5 Técnicas de análisis de datos	29
Sección III. Desarrollo Científico Tecnológico	31
3.1 Desarrollo de las fases de la metodología propuesta	31
3.2 Análisis de resultados y discusión	33
3.2.1 Recolección de datos y preparación del sistema	33
3.2.2 Descripción del sistema desarrollado	34
3.2.3 Tecnologías utilizadas	35
3.2.4 Evolución del modelo por año	36
3.3 Conclusiones	41
3.4 Recomendaciones	42
2.9 Bibliografía	44

SECCIÓN I: Descripción de Proyecto Grado

1.1 Título del Proyecto de Grado

Aplicativo web para la predicción de ventas y optimización del abastecimiento en tiempo real de las MiPymes del sector retail en Valledupar.

1.2 Dirección de Ejecución

Calle 6 No. 19A-52, Barrio San Martín, Valledupar, Cesar, Colombia

1.3 Lapso de Ejecución

Seis meses (6 meses)

1.4 Organismo o Institución Responsable del Proyecto

Universidad Popular de Cesar Sede Sabanas

1.5 Información de contacto de los estudiantes

Nombres	Apellidos	Cédula	Teléfono	Correo
Luis Fabian	Contreras Maldonado	1003235269	+57 3164990870	lfabiancontreras@unicesar.edu.co
Daniel Ricardo	Piña Corzo	1193040752	+57 3014881717	drpina@unicesar.edu.co

1.6 Línea, sublínea y grupo de investigación del Proyecto

Línea de Investigación	Transformación Digital
Sub-línea de Investigación	Analytics
Área Temática	Business Intelligence
Grupo de Investigación	GISICO

SECCIÓN II: Descripción Situacional

2.1 Identificación del Problema

Las Micro, Pequeñas y Medianas Empresas (MiPymes) constituyen un pilar esencial de la economía colombiana, ya que representan el 99,5% de las empresas formales del país, generan aproximadamente el 79% del empleo y aportan cerca del 40% del Producto Interno Bruto (PIB).¹ No obstante, pese a su relevancia económica y social, enfrentan importantes desafíos en los procesos de transformación digital y adopción de tecnologías avanzadas como los softwares de análisis de datos. Según el informe *Digitalización y Desarrollo Sostenible de la MIPYME en Colombia* elaborado por ACOPI, una de las principales barreras para la implementación tecnológica es la falta de recursos financieros, la cual presenta una alta valoración en todos los tamaños empresariales: microempresas (3,64), pequeñas empresas (3,58) y medianas

¹ **BBVA Research (2024).** *Colombia: Una mirada a las mipymes en Colombia.* Banco Bilbao Vizcaya Argentaria – BBVA Research. Recuperado de <https://www.bbva.com/publicaciones/colombia-una-mirada-a-las-mipymes-en-colombia/>

empresas (3,51), en una escala de 1 (poco importante) a 5 (muy importante) (ACOPI, p. 31). Estos resultados evidencian que las restricciones económicas constituyen un obstáculo significativo para el avance en la digitalización de las MiPymes, afectando especialmente a las unidades empresariales de menor tamaño [8].

En el entorno actual, caracterizado por una alta competitividad y una transformación constante de los hábitos de consumo, las pequeñas y medianas empresas enfrentan importantes retos operativos y estratégicos. En particular, el sector *retail* —también conocido como comercio minorista— comprende aquellas empresas dedicadas a la venta directa de bienes al consumidor final a través de canales físicos o digitales. Dentro de este sector, las MiPymes deben gestionar eficientemente procesos clave como ventas, abastecimiento e inventarios para mantenerse competitivas. Sin embargo, muchas de estas organizaciones carecen de herramientas tecnológicas que les permitan prever la demanda de productos y gestionar de forma eficiente sus niveles de inventario, lo cual genera problemas de sobrestock y desabastecimiento que afectan su rentabilidad [9].

La gran mayoría de estas empresas toman decisiones clave, como compras, reposiciones y promociones a partir de la experiencia empírica de sus propietarios o registros manuales. Aunque puede ser útil en algunos casos, generalmente genera estimaciones inexactas que terminan resultando en falta de stock de productos altamente demandados o exceso de mercadería de artículos de baja rotación, afectando directamente su rentabilidad y sostenibilidad [9] [10].

Estas ineficaces formas de operar generan pérdidas económicas importantes y hacen que las MiPymes sean menos competitivas frente a las empresas que han implementado soluciones tecnológicas, al mismo tiempo que frenan su capacidad de reacción a la demanda, limitando su crecimiento. Por otro lado, la falta de datos precisos impide la toma de decisiones adecuadas, lo que impacta en la planificación de compras y la gestión del inventario de productos, afectando también la satisfacción del cliente, dado que no se puede garantizar en la oferta de productos ni la disponibilidad ni la oportunidad [3].

Frente a esta situación, la incorporación de tecnologías como el Machine Learning se presenta como una alternativa viable para estas organizaciones. La implementación de un sistema predictivo que utilice estas herramientas permitiría a las MiPymes anticipar la demanda de productos, ajustar los niveles de inventario en función de datos reales, permitiendo una toma de decisiones más acertada para evitar pérdidas económicas y mejorar su rendimiento general. [1] [5]

Por lo tanto, las problemáticas a investigar son:

¿Cómo podría la introducción de un sistema predictivo de ventas y una gestión optimizada del inventario contribuir a mejorar la eficiencia operativa y la rentabilidad en las MiPymes del sector retail en Valledupar?

¿Qué herramientas tecnológicas pueden contribuir a mejorar la gestión del inventario y la toma de decisiones de compra en las MiPymes del sector retail?

¿Cuál sería el impacto de la implementación de este sistema en la reducción de costos operativos, sobrestock y desabastecimiento?

2.2 Justificación del Proyecto

Las MiPymes del sector retail enfrentan la necesidad de fortalecer su gestión operativa en un entorno caracterizado por alta competitividad, incertidumbre en la demanda y cambios constantes en los hábitos de consumo. Diversos estudios señalan que factores como la variabilidad de la demanda, la amplia variedad de productos y los errores en la planificación del suministro inciden directamente en problemas de sobrestock o desabastecimiento, afectando la rentabilidad empresarial [6]. En este contexto, desafíos como la planificación de la demanda, el manejo eficiente del inventario y la toma de decisiones estratégicas requieren el apoyo de herramientas tecnológicas que permitan analizar información de manera oportuna; sin embargo, muchas MiPymes no cuentan con los recursos necesarios para su implementación [6].

Desde una perspectiva teórica, el aprovechamiento de herramientas como Machine Learning ofrece un beneficio importante para las empresas, permitiéndoles procesar grandes volúmenes de datos históricos para identificar patrones de comportamiento y pronosticar la demanda de sus productos con mayor capacidad de estimación, haciendo que la toma de decisiones sobre su inventario sea basada en datos, evitando el desabastecimiento y sobrestock. En definitiva, el uso de este tipo de herramientas representa una ventaja competitiva para las empresas debido a la capacidad de anticipar las necesidades del usuario. Además, el avance más significativo en estas tecnologías es la capacidad de aprender patrones complejos a partir de datos, lo que representa una ventaja frente a modelos tradicionales basados solo en datos históricos. Esta actualización basada en los datos de ventas más recientes permite detectar patrones de consumo

emergentes, responder a cambios en la demanda y optimizar los recursos de forma continua.[1] [2]

En términos prácticos, el desarrollo de este aplicativo representará un aporte significativo, ya que permitirá a las MiPymes contar con una herramienta tecnológica asequible, adaptada a su realidad, que mejore su capacidad de respuesta y facilite una gestión más eficiente y estratégica de sus operaciones, reduciendo pérdidas por desabastecimiento o sobrestock, lo que permitirá detectar tendencias de demanda y apoyar la toma de decisiones de forma oportuna. Se ha demostrado que los sistemas web automatizan procesos manuales lo que reduce el error humano en el control de inventario y venta. Asimismo, el uso de algoritmos de Machine Learning permitirá una planificación del abastecimiento evitando el sobrestock y una anticipación a la demanda con mayor precisión, contribuyendo a la rotación del inventario, su duración y evitará compras innecesarias [4] [6] [7].

Adicionalmente, el sistema propuesto incorporará un enfoque híbrido de modelado predictivo, en el cual los productos serán segmentados según su frecuencia de venta en diferentes categorías de comportamiento (alta, media y baja rotación). Esta segmentación permitirá aplicar modelos diferenciados, como algoritmos de Machine Learning para productos de alta demanda y métodos estadísticos especializados para productos de baja frecuencia, optimizando así la precisión de las predicciones y el uso eficiente de recursos computacionales.

Desde la perspectiva metodológica, esta investigación adopta un enfoque aplicado y tecnológico, en el cual se diseñará, desarrollará y evaluará un

aplicativo web para la predicción de ventas y la optimización del abastecimiento mediante análisis predictivo de datos. Los beneficios esperados son una mejora en la toma de decisiones operativas con base en análisis predictivo y datos actualizados de manera continua.

Fase 1 – Revisión bibliográfica especializada: Se recopilará información sobre técnicas de predicción de demanda, optimización del abastecimiento, modelos de Machine Learning aplicados a las ventas y la logística, analítica predictiva y aplicaciones web orientadas al sector comercial. Esta fase permitirá establecer el fundamento teórico y tecnológico del sistema.

Fase 2 – Identificación de requerimientos funcionales y tecnológicos: Se determinarán las necesidades operativas de las MiPymes locales mediante el análisis de estudios de casos relevantes y revisión de documentación técnica sobre sistemas similares implementados en el sector retail. A partir de esta información se establecerán funcionalidades claves y las tecnologías a implementar.

Fase 3 – Diseño y desarrollo del aplicativo: Con base en los requerimientos obtenidos, se desarrollará una plataforma web que integre modelos predictivos y procesamiento asincrónico de datos. El aplicativo contará con una interfaz amigable que permita la correcta visualización de las proyecciones de ventas y abastecimiento.

Fase 4 – Validación funcional y pruebas del sistema: Se realizarán pruebas técnicas y funcionales las cuales ayudarán a verificar la precisión de las predicciones, el correcto funcionamiento de los procesos de análisis de datos

y la estabilidad del sistema.

Fase 5 – Implementación: Una vez validado el aplicativo, se hará la implementación en una empresa del sector retail donde se probará el desempeño con datos reales, lo que nos permitirá observar el comportamiento en condiciones de uso real, para identificar posibles ajustes y recopilar la retroalimentación sobre la experiencia del usuario usando el sistema.

Fase 6 – Análisis de resultados: Se analizarán los resultados obtenidos para valorar el impacto de la solución en la eficiencia operativa de las MiPymes.

Desde una perspectiva social, este proyecto cobra especial relevancia debido al papel fundamental que desempeñan las MiPymes en la generación de empleo, dinamización económica y sostenimiento de numerosas familias en la región. En ciudades intermedias como Valledupar, el sector retail constituye una fuente importante de ingresos para pequeños empresarios y trabajadores locales. La mejora en la gestión operativa de estas empresas no solo impacta su rentabilidad, sino que contribuye directamente a la estabilidad laboral, la reducción de la informalidad y el fortalecimiento del tejido empresarial regional. Al optimizar procesos como la planificación de la demanda y el abastecimiento, se disminuyen pérdidas económicas que pueden comprometer la continuidad del negocio, favoreciendo así su sostenibilidad en el tiempo.

En el ámbito institucional, la investigación se alinea con los objetivos de transformación digital, innovación empresarial y fortalecimiento de las MiPymes promovidos a nivel nacional y regional. El desarrollo de soluciones tecnológicas aplicadas al sector productivo responde a la necesidad de cerrar

brechas digitales existentes en pequeñas empresas que, por limitaciones financieras o técnicas, no acceden fácilmente a herramientas avanzadas de análisis de datos. Asimismo, este proyecto aporta al desarrollo académico y tecnológico al integrar conocimientos de analítica de datos, Machine Learning y desarrollo web en un contexto empresarial real, generando evidencia aplicada que puede servir como referencia para futuras investigaciones o implementaciones en otras organizaciones del sector.

De igual manera, la propuesta fortalece la relación universidad–empresa al ofrecer una solución basada en investigación aplicada que atiende problemáticas concretas del entorno productivo. Esto permite que el conocimiento académico trascienda el ámbito teórico y tenga un impacto directo en la competitividad y modernización del sector comercial local.

2.3 Objetivos del Proyecto

Objetivo General:

Desarrollar un aplicativo web para la predicción de ventas y optimización del abastecimiento en tiempo real de las MiPymes del sector retail en Valledupar.

Objetivos Específicos:

1. Analizar y seleccionar algoritmos de Machine Learning adecuados para la predicción de ventas en MiPymes del sector retail.
2. Recolectar datos históricos y contextuales relevantes para el sistema de predicción de ventas y abastecimiento basado en un caso de uso representativo.
3. Diseñar un prototipo funcional de un sistema predictivo de ventas y

abastecimiento, que utilice Machine Learning para apoyar la toma de decisiones de compra.

4. Desarrollar los módulos esenciales de ingreso de datos, análisis predictivo y gestión de inventarios para el sistema web de predicción de ventas y optimización del abastecimiento.

5. Implementar el aplicativo web en una empresa del sector retail como caso de uso representativo para medir su impacto en la eficiencia operativa.

2.4 Bases Teóricas

Las MiPymes han sido históricamente un eje fundamental en la economía colombiana, especialmente en ciudades intermedias como Valledupar, donde constituyen una de las principales fuentes de empleo y dinamización del comercio local. Tradicionalmente, la gestión de inventarios en estas empresas se ha basado en métodos empíricos, registros manuales y decisiones sustentadas en la experiencia del propietario, lo que generaba limitaciones en el control del stock, errores en la reposición y baja capacidad de anticipación frente a variaciones en la demanda.

Desde una perspectiva histórica del estado del arte, la gestión de inventarios comenzó a estructurarse formalmente a inicios del siglo XX con modelos clásicos como el **Modelo de Cantidad Económica de Pedido (EOQ)** desarrollado por Harris (1913), el cual buscaba minimizar los costos de almacenamiento y pedido. Posteriormente, durante las décadas de 1950 y 1960, se incorporaron modelos probabilísticos para manejar la incertidumbre en la demanda, permitiendo una planificación más técnica del abastecimiento. En los años 80 y 90, con la expansión de los sistemas MRP (Material Requirements Planning) y ERP (Enterprise Resource Planning), las grandes

empresas comenzaron a integrar herramientas digitales para optimizar la cadena de suministro.

Sin embargo, estas tecnologías tardaron en llegar a las pequeñas y medianas empresas debido a barreras económicas y tecnológicas. A partir de los años 2000, con la masificación de internet y la reducción de costos en software empresarial, se inició una transición progresiva hacia sistemas digitalizados en MiPymes. Más recientemente, en la última década, el avance del **Machine Learning** ha transformado el enfoque tradicional de gestión, pasando de modelos descriptivos e históricos a modelos predictivos capaces de analizar grandes volúmenes de datos y anticipar patrones de consumo con mayor precisión [1].

Esta evolución evidencia una transición clara en la línea de investigación: desde modelos matemáticos clásicos de control de inventarios, hacia sistemas inteligentes basados en analítica avanzada y modelos predictivos apoyados en el análisis de datos. No obstante, la adopción de estas herramientas en MiPymes continúa siendo desigual, especialmente en contextos regionales donde las limitaciones de infraestructura tecnológica y recursos financieros ralentizan el proceso de transformación digital.

2.4.1 Antecedentes

La sección expone los antecedentes que fundamentan el desarrollo del presente proyecto, organizados en tres dimensiones: antecedentes históricos, que describen la evolución de la gestión de inventarios y predicción de ventas en la pequeñas y medianas empresas; antecedentes investigativos, que recopilan aportes recientes en el uso de técnicas de análisis de datos, Machine Learning y sistema web aplicados al sector retail; y antecedentes legales, que enmarcan el contexto normativo colombiano en materia de transformación

digital y fortalecimiento de MiPymes.

2.4.1.1 Antecedentes históricos

Las MiPymes han sido históricamente un eje fundamental en la economía colombiana, especialmente en ciudades intermedias como Valledupar, donde constituyen la principal fuente de empleo y dinamización del comercio local. Durante décadas, la gestión de inventarios y ventas en estas empresas se ha basado en métodos empíricos y registros manuales, lo que ha generado ineficiencias en el control de los niveles de stock y dificultades para responder a las fluctuaciones de la demanda. Con el auge de la digitalización en los últimos años, han surgido herramientas tecnológicas que permiten pasar de un enfoque meramente reactivo a uno predictivo, apoyado en el análisis de datos y en el uso de algoritmos inteligentes [1].

2.4.1.2 Antecedentes investigativos

En el ámbito internacional, **Peña Sanz (2024)** tuvo como objetivo analizar el impacto de la integración de Inteligencia Artificial en los procesos de ventas en grandes organizaciones comerciales. La metodología empleada fue cuantitativa, utilizando modelos predictivos como regresión lineal, K-Nearest Neighbors (KNN) y Random Forest aplicados a datos históricos de ventas de Walmart. Los resultados evidenciaron una mejora significativa en la precisión de los pronósticos y en la toma de decisiones estratégicas relacionadas con inventarios y abastecimiento. El autor concluye que la implementación de analítica avanzada permite pasar de modelos reactivos a modelos predictivos con alto nivel de confiabilidad [1]. Este estudio aporta a la presente investigación la validación técnica de modelos predictivos aplicados al sector retail.

En Latinoamérica, **Flores y Montalvo (2024)** desarrollaron un modelo de Machine Learning orientado a la predicción en la gestión de inventarios en PYMES de productos tecnológicos. El objetivo fue mejorar los indicadores de rotación y duración del inventario mediante algoritmos de aprendizaje supervisado. La metodología incluyó análisis exploratorio de datos, selección de variables relevantes y validación del modelo con métricas de error como MAPE y RMSE. Los resultados mostraron una reducción en el sobrestock y una mejora en la planificación del abastecimiento [5]. Este antecedente demuestra la aplicabilidad del Machine Learning en pequeñas empresas, lo cual respalda la viabilidad del presente proyecto.

En el sector farmacéutico se han desarrollado herramientas de pronóstico de demanda orientadas a reducir pérdidas por vencimiento de productos mediante modelos predictivos. Estas propuestas contemplan procesos ETL (Extracción, Transformación y Carga), análisis exploratorio de datos e implementación de modelos de series temporales, evidenciando una disminución en mermas y una mejora en la eficiencia del inventario. Se concluye que el uso de análisis predictivo mejora la sostenibilidad financiera de las empresas [2]. Este enfoque aporta un referente metodológico claro para la estructuración técnica del aplicativo propuesto.

En el ámbito empresarial, **Bernaola y Varillas (2022)** propusieron un sistema predictivo con Machine Learning para la gestión de inventarios en una empresa ferretera. Su objetivo fue evaluar el impacto del modelo en la optimización del abastecimiento. La metodología incluyó la implementación de algoritmos supervisados y la validación estadística mediante la prueba de Wilcoxon para medir diferencias significativas antes y después de la aplicación

del sistema. Los resultados demostraron mejoras estadísticamente significativas en la reducción del sobrestock y quiebres de inventario [7]. Este antecedente fortalece el enfoque cuantitativo de la presente investigación.

En el contexto del retail, **Andrade y Viacava (2024)** desarrollaron una propuesta de mejora en la planificación del abastecimiento mediante gestión de inventarios y técnicas de Business Process Management (BPM). El objetivo fue disminuir el sobrestock en una MYPE del sector retail. La metodología se basó en análisis de procesos, diagnóstico situacional y aplicación de herramientas de mejora continua. Los resultados evidenciaron una reducción en los niveles de inventario excedente y una mejora en la eficiencia operativa [6]. Este estudio aporta un enfoque complementario al integrar procesos organizacionales con herramientas analíticas.

En Colombia, **Gutiérrez (2024)** analizó la conciliación de inventarios en MiPymes de operaciones logísticas. El objetivo fue diseñar un modelo metodológico que mejorara la coherencia entre registros físicos y digitales. La metodología incluyó diagnóstico documental, análisis comparativo y propuesta de estandarización de procesos. Los resultados demostraron que la correcta conciliación mejora la confiabilidad de los datos, elemento fundamental para implementar modelos predictivos efectivos [3]. Este antecedente resalta la importancia de la calidad de los datos en sistemas de predicción.

Complementariamente, **Díaz (2024)** investigó el impacto de la gestión contable en la sostenibilidad de microempresas colombianas. El objetivo fue determinar la relación entre el uso de información financiera y la permanencia empresarial. La metodología fue de enfoque cuantitativo, utilizando análisis estadístico de

variables contables y su incidencia en indicadores de sostenibilidad. Los resultados confirmaron que la gestión basada en datos fortalece la toma de decisiones estratégicas [10]. Este estudio aporta evidencia sobre la importancia de integrar sistemas tecnológicos con información financiera estructurada.

Finalmente, a nivel gremial, **ACOPI (2022)** analizó el nivel de digitalización de las MiPymes en Colombia. El objetivo fue identificar barreras y avances en la adopción tecnológica. La metodología consistió en encuestas estructuradas a empresas clasificadas por tamaño. Los resultados evidenciaron que solo el 35,7% de las microempresas han incorporado tecnologías digitales avanzadas, reflejando una brecha tecnológica significativa [8]. Este hallazgo justifica la necesidad de soluciones tecnológicas accesibles y adaptadas al contexto de las MiPymes, como la propuesta en esta investigación.

2.4.1.3 Antecedentes legales

El marco normativo colombiano ha impulsado diversas políticas orientadas al fortalecimiento, formalización y digitalización de las MiPymes. La **Ley 905 de 2004**, que modifica la Ley 590 de 2000, estableció disposiciones para el desarrollo y promoción de la micro, pequeña y mediana empresa, reconociendo su importancia estratégica dentro del tejido productivo nacional. Posteriormente, el **Decreto 1008 de 2018**, mediante el cual se establece la Política de Gobierno Digital, promovió el uso estratégico de las tecnologías de la información, la gestión de datos y la innovación como herramientas para mejorar la competitividad y eficiencia organizacional. Estas disposiciones respaldan normativamente la implementación de soluciones basadas en analítica de datos y sistemas predictivos en el sector empresarial.

En relación con el desarrollo de la investigación y el tratamiento de información, resulta fundamental considerar la **Ley 1581 de 2012**, por la cual se dictan disposiciones generales para la protección de datos personales en Colombia. Esta norma regula la recolección, almacenamiento, uso y circulación de datos, estableciendo principios como legalidad, finalidad, libertad, veracidad y seguridad. En el contexto del presente proyecto, cualquier información comercial o empresarial utilizada deberá cumplir con estos lineamientos, garantizando la confidencialidad y uso adecuado de los datos.

Asimismo, el desarrollo del aplicativo web y los modelos predictivos debe enmarcarse dentro de la normativa sobre propiedad intelectual. La **Ley 23 de 1982** sobre derechos de autor y la **Decisión Andina 351 de 1993** protegen las obras literarias, científicas y el software como creación intelectual. En este sentido, el código fuente, la arquitectura del sistema y los algoritmos desarrollados estarán protegidos como obra de carácter informático. De igual manera, si se utilizan librerías, frameworks o herramientas de terceros, estas deberán cumplir con sus respectivas licencias de uso (por ejemplo, licencias MIT, GPL o Apache), respetando las condiciones establecidas por sus desarrolladores.

Adicionalmente, el **Código de Comercio colombiano** establece obligaciones relacionadas con la conservación de registros contables y documentación empresarial, lo cual se vincula directamente con la digitalización de inventarios y sistemas de información financiera. Esto refuerza la importancia de que

cualquier solución tecnológica implementada garantice integridad, trazabilidad y respaldo adecuado de la información.

2.4.2 Marco Teórico

El marco teórico del presente proyecto se orienta a fundamentar los conceptos que respaldan la implementación de un aplicativo web con capacidades de aprendizaje automático, cuyo propósito es predecir ventas y optimizar el abastecimiento en MiPymes del sector retail de Valledupar, Cesar. La revisión incluye enfoques sobre Machine Learning, algoritmos de predicción y la transformación digital aplicada al entorno empresarial.

2.4.2.1 Machine Learning

El aprendizaje automático (Machine Learning) constituye una rama de la inteligencia artificial que permite a los sistemas mejorar su desempeño en tareas específicas a partir de la experiencia y los datos. En el ámbito de las MiPymes, su aplicación en la predicción de ventas permite desarrollar modelos capaces de reconocer patrones de comportamiento de los consumidores, anticipar necesidades de abastecimiento y reducir ineficiencias derivadas del manejo empírico de inventarios [5].

2.4.2.2 Algoritmos de predicción y clasificación

Los algoritmos de clasificación y predicción representan la base de los modelos de Machine Learning en contextos empresariales. Entre los más utilizados en la literatura destacan Random Forest, Regresión Logística, árboles de decisión, redes neuronales y k-vecinos más cercanos (kNN). Estos métodos permiten generar proyecciones confiables a partir de datos históricos, lo cual se traduce en la posibilidad de generar alertas tempranas frente a

riesgos de sobrestock o desabastecimiento [7].

2.4.2.3 Transformación digital en MiPymes

La transformación digital en las MiPymes va más allá de la simple digitalización de registros contables o de ventas. Implica la adopción estratégica de tecnologías de información que se integren a los procesos de negocio, mejoren la eficiencia operativa y fortalezcan la competitividad. En Colombia, el gremio de MiPymes ha enfrentado una brecha tecnológica que limita su capacidad de competir frente a grandes superficies; sin embargo, estudios recientes muestran que la incorporación de aplicaciones web predictivas representa una oportunidad para reducir esta brecha y garantizar mayor sostenibilidad [8].

2.4.2.3.1 Línea de investigación: Transformación digital empresarial

La transformación digital empresarial se centra en la integración de tecnologías de información y comunicación (TIC) en los procesos productivos, administrativos y comerciales. Esta línea aborda:

- Digitalización de procesos.
- Automatización operativa.
- Analítica de datos.
- Innovación basada en tecnología.
- Modelos de negocio digitales.

En el caso de las MiPymes, esta línea busca reducir brechas tecnológicas y mejorar la competitividad frente a grandes empresas.

2.4.2.3.2 Sublínea: Analítica de datos aplicada a la gestión empresarial

Dentro de la transformación digital, surge la sublínea de **analítica de datos aplicada a la gestión empresarial**, la cual se enfoca en el uso estratégico de

información para la toma de decisiones.

La analítica empresarial se clasifica en:

1. Analítica descriptiva → ¿Qué ocurrió?
2. Analítica diagnóstica → ¿Por qué ocurrió?
3. Analítica predictiva → ¿Qué podría ocurrir?
4. Analítica prescriptiva → ¿Qué se debe hacer?

El presente proyecto se ubica principalmente en la **analítica predictiva**, ya que busca anticipar comportamientos de demanda para optimizar inventarios.

2.4.2.3.3 Tema específico: Predicción de ventas y optimización del abastecimiento

En el sector retail, uno de los principales retos es la gestión eficiente del inventario. Una planificación inadecuada genera:

- Sobrestock → aumento de costos de almacenamiento.
- Desabastecimiento → pérdida de ventas y clientes.
- Baja rotación → inmovilización de capital.

El uso de modelos predictivos permite anticipar la demanda futura a partir de datos históricos, comportamiento de clientes y patrones estacionales.

2.4.2.4 Segmentación de la demanda en sistemas predictivos

En sistemas de predicción de demanda en entornos retail, es común segmentar los productos según su comportamiento de ventas. Esta segmentación permite adaptar el modelo predictivo a las características específicas de cada producto. Por ejemplo, los productos de alta rotación pueden ser modelados mediante algoritmos de Machine Learning supervisado, mientras que los productos de baja frecuencia requieren métodos especializados como el modelo de Croston, diseñado para series de demanda intermitente.

Este enfoque mejora la precisión de los pronósticos y evita el uso innecesario de modelos complejos en escenarios donde no aportan valor. Para efectos del presente sistema, la segmentación se determina a partir del historial de ventas de cada producto. Los productos de alta rotación (Tier A) son aquellos con alta frecuencia de días con venta y volumen suficiente para entrenar modelos supervisados. Los de rotación media (Tier B) presentan venta intermitente pero con historia suficiente para el clasificador. Los de baja rotación (Tier C) son productos con demanda esporádica o historia reducida, para los cuales se aplica el método de Croston y bootstrap histórico mensual.

2.4.3 Marco Conceptual

El marco conceptual define los principales términos y categorías analíticas que orientan el desarrollo del proyecto. Estos conceptos permiten establecer una visión clara de los elementos que intervienen en la predicción de ventas y la optimización de inventarios en MiPymes de Valledupar.

Tiempo real en analítica empresarial:

En el contexto de los sistemas de análisis de datos empresariales, el término tiempo real hace referencia a la capacidad de un sistema para actualizar sus análisis y recomendaciones utilizando la información más reciente disponible en la base de datos. Esto no implica necesariamente un procesamiento instantáneo de cada evento, sino la posibilidad de recalcular resultados de manera dinámica a medida que se incorporan nuevos datos al sistema.

El procesamiento en tiempo real puede implementarse mediante un enfoque basado en la acumulación de datos, donde las predicciones se ejecutan

cuando se alcanza una cantidad significativa de nuevos registros. Este mecanismo permite optimizar el uso de recursos computacionales y evitar cálculos innecesarios, garantizando que los resultados se generen a partir de información suficientemente representativa. De esta manera, el sistema mantiene un comportamiento dinámico y actualizado sin requerir procesamiento continuo por cada evento individual.

Datos empresariales

Los datos empresariales corresponden a la información generada en el desarrollo de las actividades comerciales de las MiPymes, incluyendo registros de ventas, inventario, proveedores y clientes. Su calidad, precisión y disponibilidad oportuna determinan la fiabilidad de los modelos predictivos que se construyan a partir de ellos [2].

Transformación de datos

La transformación de datos hace referencia al proceso mediante el cual la información bruta es depurada, organizada y adaptada para ser procesada por algoritmos de Machine Learning. Este proceso es esencial en los sistemas de predicción, ya que convierte datos no estructurados en insumos de valor para el entrenamiento de modelos [6].

Modelos predictivos

Los modelos predictivos son sistemas analíticos basados en técnicas estadísticas y de minería de datos, cuyo objetivo es estimar o anticipar resultados futuros a partir de información histórica y actual. En el contexto de este proyecto, los modelos predictivos permiten anticipar niveles de ventas,

identificar tendencias de consumo y mejorar la toma de decisiones en el abastecimiento de inventario [5].

Sistemas predictivos en MiPymes

Un sistema predictivo es un software que integra algoritmos de Machine Learning con datos de la organización, generando información procesable para la toma de decisiones. Su implementación en MiPymes del sector retail en Valledupar busca mejorar la planificación de compras, reducir pérdidas por sobrestock o desabastecimiento y fortalecer la sostenibilidad empresarial [3].

2.5 Marco Metodológico

El marco metodológico describe el conjunto de procedimientos, técnicas y estrategias que orientan el desarrollo de la investigación. En esta sección se especifica el tipo de estudio, el enfoque metodológico, las fases del proceso investigativo, las técnicas de recolección y análisis de datos, así como los instrumentos y herramientas utilizadas para cumplir los objetivos planteados. Su finalidad es garantizar rigor científico, coherencia entre el problema y la solución propuesta, y validez en los resultados obtenidos.

2.5.1 Tipo de investigación

Este trabajo se clasifica como una investigación **aplicada de carácter cuantitativo**, dado que busca resolver un problema práctico de las MiPymes del sector retail en Valledupar mediante el desarrollo de un aplicativo web basado en técnicas de **Machine Learning**. La investigación se centra en analizar datos históricos de ventas e inventario para generar modelos predictivos que apoyen la toma de decisiones estratégicas en la planificación del abastecimiento.

Desde el punto de vista metodológico, el estudio adopta un diseño tecnológico–experimental, incorporando un esquema de actualización incremental de predicciones, en el cual los modelos podrán recalcular resultados a partir de nuevos datos sin reprocesar completamente la información histórica. Asimismo, se integrará un enfoque de modelado híbrido basado en la segmentación de productos según su comportamiento de demanda, permitiendo la aplicación de distintos modelos predictivos (Machine Learning y métodos estadísticos) de acuerdo con la frecuencia de ventas.

En cuanto a la normatividad aplicable a la investigación, el proyecto se enmarca en los lineamientos establecidos por la **Resolución 8430 de 1993 del Ministerio de Salud de Colombia**, la cual regula los aspectos éticos de la investigación con datos, clasificando este estudio como una investigación sin riesgo, dado que no interviene directamente sobre personas ni manipula variables biológicas o psicológicas. Asimismo, se dará cumplimiento a la **Ley 1581 de 2012 sobre protección de datos personales**, garantizando el tratamiento adecuado, confidencialidad y uso exclusivo de la información empresarial con fines académicos y de investigación.

2.5.2 Enfoque del proyecto

Para el presente proyecto, el uso de Extreme Programming (XP) se considera adecuado, ya que se trata de un equipo reducido y esta metodología facilita el desarrollo colaborativo y flexible. XP promueve prácticas como la programación en parejas, las pruebas continuas y la integración frecuente, lo que ayuda a mantener un código confiable y de calidad. Además, al trabajar con ciclos cortos, permite entregar avances funcionales de manera temprana y obtener retroalimentación rápida de los usuarios. Esto resulta útil en un

entorno como el de las MiPymes, donde los cambios en la demanda pueden requerir ajustes constantes en el aplicativo web².

2.5.3 Población y muestra

Enfoque metodológico: estudio de caso único

Dado que el presente proyecto tiene como objetivo el desarrollo e implementación de un aplicativo web funcional en un contexto real, la investigación adopta un diseño de estudio de caso único de tipo instrumental (Stake, 1995; Yin, 2018). Este enfoque es ampliamente utilizado en proyectos aplicados de ingeniería de software cuando el interés científico no radica en la generalización estadística, sino en la comprensión profunda de un fenómeno en su contexto real, en este caso, la predicción de ventas y la optimización del abastecimiento en una MiPyme del sector retail. La validez del estudio no depende del número de unidades analizadas, sino de la riqueza y representatividad del caso seleccionado y del rigor metodológico aplicado en el diseño, implementación y evaluación del sistema.

Unidad de análisis (Población objetivo): La unidad de análisis está conformada por los procesos de gestión de ventas e inventarios de las MiPymes del sector retail en la ciudad de Valledupar. Este universo está integrado por los establecimientos de comercio al por menor registrados ante la Cámara de Comercio de Valledupar que operan en categorías como

² ESIC Business & Marketing School. (2024). *¿Qué es la programación extrema? Características*. Disponible en: <https://www.esic.edu/rethink/business/que-es-la-programacion-extrema-caracteristicas-c>

papelería, útiles escolares, productos de oficina y bienes de consumo frecuente. Dado que dichas empresas comparten estructuras operativas similares (gestión manual o semi-automatizada de inventarios, volumen de transacciones medianas y demanda con patrones estacionales), el caso seleccionado es representativo de las problemáticas comunes del segmento.

Caso de estudio seleccionado (Muestra): Para la implementación y prueba del aplicativo se seleccionará una MiPyme del sector retail ubicada en Valledupar, dedicada a la comercialización de productos de papelería y útiles escolares y de oficina. La elección se fundamenta en criterios de selección intencionada (muestreo no probabilístico por conveniencia teórica), considerando: (a) disponibilidad y accesibilidad real de los datos históricos de ventas e inventario; (b) voluntad de la empresa para participar y facilitar información; (c) representatividad de sus características operativas frente al perfil típico de las MiPymes retail de la ciudad; y (d) existencia de problemáticas concretas de predicción de demanda y desabastecimiento que justifican la intervención tecnológica propuesta.

Los datos que constituirán la muestra de entrenamiento y validación del modelo predictivo corresponden al conjunto completo de registros históricos de ventas e inventario suministrados por dicha empresa, que incluyen aproximadamente 250 transacciones mensuales (considerando productos repetidos), con variables como producto, fecha, cantidad vendida y nivel de inventario. Se aplicará una ventana temporal mínima de doce (12) meses de datos para garantizar la captura de ciclos estacionales relevantes para el entrenamiento del modelo.

Criterios de inclusión y exclusión: Se incluirán los registros que contengan información completa de las variables clave: producto, fecha de transacción,

cantidad vendida y nivel de inventario. Serán excluidos aquellos registros con más del 20 % de campos incompletos o con inconsistencias no subsanables mediante técnicas de imputación. Adicionalmente, se excluirán períodos atípicos no representativos (por ejemplo, cierres temporales o eventos excepcionales documentados) que puedan distorsionar el comportamiento histórico de la demanda. De igual forma, el sistema está diseñado exclusivamente para productos con precio unitario fijo que incluyen IVA en su valor final, quedando excluidos los productos exentos de IVA según la normativa tributaria colombiana, así como aquellos comercializados por peso o medida variable, cuya estructura de precio es incompatible con el modelo de predicción implementado.

Alcance y transferibilidad: Aunque el sistema será implementado y validado en una sola empresa, el diseño del aplicativo contempla una arquitectura modular y parametrizable que permitirá su adaptación a otras MiPymes del sector retail con características similares. Los resultados obtenidos durante la fase de prueba servirán como evidencia del funcionamiento del sistema en condiciones reales, lo que constituye la base para futuras escalabilidades o investigaciones comparativas con mayor número de unidades.

2.5.4 Técnicas de recolección de datos

La recolección se basará en fuentes primarias y secundarias, los datos serán exportados en formatos CSV/Excel y sometidos a un proceso de limpieza y validación para garantizar consistencia y calidad.

2.5.4.1. Revisión documental

Se realizará el análisis de registros históricos de ventas e inventario

proporcionados por la empresa objeto de estudio. Estos registros incluirán información como fecha de venta, producto, cantidad vendida, precios y niveles de inventario. Los datos serán exportados en formatos CSV o Excel para su posterior procesamiento.

2.5.4.2 Observación directa de procesos

Se analizará el procedimiento actual de gestión de inventarios y abastecimiento dentro de la empresa, con el fin de comprender el flujo de información y detectar posibles inconsistencias en el registro de datos.

2.5.4.3 Entrevista semiestructurada:

Se realizará una entrevista al encargado del área de inventarios o administración, con el propósito de identificar cómo se realiza actualmente la planificación de la demanda y qué criterios se utilizan para la toma de decisiones.

2.5.4.4 Extracción y análisis de bases de datos internas:

Se empleará la técnica de minería de datos para recopilar información almacenada en los sistemas digitales de la empresa. Posteriormente, los datos serán sometidos a un proceso de limpieza, depuración y validación para garantizar su consistencia y calidad antes de ser utilizados en los modelos predictivos.

2.5.5 Técnicas de análisis de datos

Limpieza de datos: Depuración de inconsistencias, eliminación de duplicados y tratamiento de valores faltantes.

Transformación: Organización de la información en un formato uniforme para facilitar el análisis.

Exploración (EDA): Identificación de tendencias, productos más vendidos, patrones de consumo y estacionalidad.

Modelado predictivo: Aplicación de un enfoque híbrido que combina algoritmos de Machine Learning (como Gradient Boosting) para productos de alta rotación y métodos estadísticos especializados (como Croston) para productos de demanda intermitente. Además, se implementará un proceso de predicción recursiva que permite generar proyecciones futuras utilizando valores previamente estimados.

Validación: Comprobación de la precisión del modelo mediante métricas como MAE y RMSE, garantizando que los resultados sean útiles para la toma de decisiones.

Simulación de demanda: Se aplicarán técnicas de simulación estocástica para generar múltiples escenarios futuros de demanda, permitiendo estimar percentiles (pesimista, esperado y optimista) que apoyen la toma de decisiones en la gestión de inventarios.

Simulación de precios e ingresos: A partir del historial de precios unitarios registrados por la empresa, se simulará la variabilidad futura del precio de cada producto mediante técnicas de bootstrap. Combinando los escenarios de demanda con los escenarios de precio, el sistema estimará rangos de ingresos mensuales proyectados bajo tres escenarios: pesimista (P20), esperado (P50) y optimista (P80), como apoyo a la planificación financiera de la empresa.

Sección III. Desarrollo Científico Tecnológico

3.1 Desarrollo de las fases de la metodología propuesta

El desarrollo de TargetBI siguió la metodología XP (Extreme Programming), que se adapta bien a equipos pequeños con entregas incrementales. El proyecto fue desarrollado por dos integrantes, donde cada uno asumió responsabilidades específicas sobre componentes del sistema para luego integrarlos en un producto final unificado, aplicando la práctica de integración continua propia de esta metodología.

Fase 1: Análisis y planificación

Se identificaron las necesidades operativas de la papelería caso de estudio y se definieron los requerimientos funcionales del sistema. Se revisó literatura sobre predicción de ventas, gestión de inventarios y aprendizaje automático aplicado a MiPymes, y se seleccionaron las tecnologías a utilizar.

Fase 2: Diseño del sistema

Se diseñó la arquitectura general del aplicativo, definiendo cómo se comunicarían entre sí la interfaz de usuario, la API desarrollada en Python y la base de datos en Supabase. Se estructuró el modelo de datos, los flujos de navegación de la interfaz y los puntos de integración entre las partes desarrolladas por cada integrante del equipo.

Fase 3: Desarrollo del backend y base de datos

Se construyó la API en Python usando FastAPI, que es el componente encargado de procesar los datos y ejecutar los modelos de predicción. Esta API fue publicada en Render, una plataforma en la nube que permite que el aplicativo funcione en internet y pueda ser usado desde cualquier dispositivo

con navegador. En esta fase también se configuró la base de datos en Supabase y se desarrollaron todos los módulos del pipeline: procesamiento de datos, extracción de características del historial de ventas, modelos predictivos, simulación de escenarios, cálculo de métricas de precisión y política de inventario.

Fase 4: Desarrollo del frontend

Se construyó la interfaz de usuario con Flutter Web, integrando los seis módulos funcionales del sistema: registro de ventas, registro de productos, categorización de productos, gestión de inventario, carga masiva de datos y dashboard de análisis predictivo. Cada módulo fue desarrollado de forma independiente y luego integrado al sistema completo.

Fase 5: Pruebas y depuración

Una vez integradas ambas partes del sistema, se realizaron pruebas funcionales de cada módulo y se verificó la correcta comunicación entre la interfaz de usuario y la API. Se corrigieron los errores identificados durante el proceso y se validó el correcto funcionamiento del pipeline de predicción con los datos históricos cargados.

Fase 6: Despliegue e implementación

Se publicó la API en Render y se dejó el sistema disponible para usarse desde el navegador. Se cargaron los datos históricos recopilados manualmente en la papelería mediante la funcionalidad de carga masiva por Excel, y se validó el funcionamiento completo del aplicativo en condiciones reales de operación, utilizando el año 2025 como periodo de validación de la precisión del modelo.

3.2 Análisis de resultados y discusión

El presente proyecto resultó en el desarrollo e implementación de TargetBI, un aplicativo web funcional para la predicción de ventas y optimización del abastecimiento en MiPymes del sector retail. Para su validación, el sistema fue implementado en una papelería ubicada en Valledupar que comercializa útiles escolares, productos de oficina y artículos de papelería en general. Aunque el caso de uso fue esta papelería específica, el sistema fue diseñado desde su base para poder ser adoptado por cualquier pequeña empresa del sector retail que cuente con registros históricos de ventas.

3.2.1 Recolección de datos y preparación del sistema

La recolección de datos se realizó directamente en el negocio de forma manual. Dado que la empresa no contaba con un sistema digital de registro de ventas, la información histórica fue recopilada a partir de los registros físicos disponibles, como facturas y anotaciones del propietario, y posteriormente digitalizados por los autores del proyecto en una hoja de Excel. Se recopilaron datos correspondientes a los años 2023, 2024 y 2025, cubriendo un total de 237 productos activos.

Una vez consolidada la información, se realizó un proceso de revisión y limpieza para garantizar su calidad: se corrigieron registros incompletos, se eliminaron entradas duplicadas o con valores inconsistentes, y se estandarizó el formato de fechas, productos y cantidades. Con los datos ya organizados, se utilizó la funcionalidad de carga masiva del sistema, que permite subir la información histórica mediante un archivo Excel con una plantilla predefinida, sin necesidad de conocimientos técnicos por parte del usuario.

El año 2025 fue utilizado como año de validación para medir qué tan preciso fue el sistema comparando sus predicciones contra lo que realmente ocurrió en el negocio. Con base en los tres años de historial cargados, el sistema generó las proyecciones para el año 2026. Es importante contextualizar que el sistema operó bajo las condiciones reales de una MiPyme: volumen de datos limitado, infraestructura de bajo costo basada en servicios gratuitos, y alta variabilidad en la demanda diaria, característica estructural de los negocios de retail de pequeña escala.

3.2.2 Descripción del sistema desarrollado

TargetBI fue construido como un sistema integral que cubre el ciclo operativo completo de una pequeña empresa comercial, no únicamente el análisis predictivo. El aplicativo cuenta con seis módulos funcionales.

El módulo de registro de ventas funciona como un punto de venta básico: el usuario busca el producto por nombre o escaneando su código de barras, indica la cantidad vendida y el sistema calcula automáticamente el subtotal, el IVA y el total a cobrar, registrando también el método de pago. Cada venta registrada alimenta el historial que usa el modelo predictivo, permitiendo que el sistema aprenda de forma continua con la operación diaria del negocio.

El módulo de registro de productos permite ingresar nuevos artículos al catálogo indicando nombre, costo de compra, precio de venta y stock inicial. El sistema genera automáticamente un código interno para cada producto, aunque también admite código de barras si el artículo lo tiene. El módulo de categorización permite organizar los productos en grupos personalizados, facilitando la gestión del catálogo cuando se trabaja con un volumen amplio de

referencias. El módulo de inventario presenta una vista consolidada de los 237 productos activos, mostrando el stock disponible, el estado de cada producto y el valor total del inventario, que al momento de la evaluación ascendía a \$12.015.138 COP, identificando automáticamente los productos con stock bajo. El módulo de carga de datos permite subir información histórica de forma masiva mediante un archivo Excel con plantilla descargable. Finalmente, el módulo de dashboard es el núcleo analítico del sistema, donde se visualizan las predicciones, los escenarios de demanda, el comportamiento histórico por mes y los indicadores de precisión del modelo.

3.2.3 Tecnologías utilizadas

- **Frontend:** Flutter Web — interfaz accesible desde el navegador sin necesidad de instalar nada en el equipo.
- **Backend:** API desarrollada en Python con FastAPI, publicada en Render para funcionar en internet.
- **Base de datos:** Supabase — almacenamiento de la información del negocio en la nube, basado en PostgreSQL.
- **Gradient Boosting** — modelo de aprendizaje automático para predecir la demanda de productos que se venden con frecuencia regular.
- **Croston** — método estadístico para productos de venta esporádica o irregular.
- **Simulación estocástica** — genera tres escenarios de demanda futura: conservador, esperado y optimista.
- **Bootstrap de precios** — simula la variabilidad futura del precio por producto para estimar rangos de ingresos más realistas.
- **Política de inventario (modelo s, S)** — calcula automáticamente los niveles mínimo y máximo de stock recomendados para cada producto.

3.2.4 Evolución del modelo por año

Uno de los hallazgos más importantes del proyecto es la mejora progresiva del sistema a medida que dispone de más datos históricos. La siguiente tabla resume el comportamiento del modelo en cada año evaluado:

Año	Ingreso real acumulado	Error neto	Error real (WAPE)	sMAPE	Observación
2023 (abr–dic)	\$15.320.600	+46,8%	56,9%	43,6%	Primer año, datos muy limitados
2024 (ene–dic)	\$34.862.400	-7,7%	37,2%	41,7%	Mejora notable con historial completo
2025 (ene–nov)	\$32.740.000	-15,4%	37,0%	42,6%	Comportamiento estable respecto a 2024

Antes de analizar los resultados, es importante entender qué mide cada indicador. El **error neto** refleja la diferencia total entre lo predicho y lo real considerando la dirección de cada error: los meses donde el modelo sobreestimó compensan parcialmente los meses donde subestimó, por lo que los valores se cancelan entre sí. El **error real o WAPE** mide el error puro de cada mes de forma independiente, sin que unos errores cancelen a otros, lo que da una visión más honesta de qué tan lejos estuvo el modelo en promedio, y por eso siempre es mayor o igual al error neto en valor absoluto. El **sMAPE** o error porcentual absoluto medio simétrico expresa ese mismo error en porcentaje de forma simétrica, tratando igual un error por exceso que uno por defecto, lo que lo hace útil para comparar la precisión entre meses o productos con volúmenes de venta muy distintos. Un sMAPE de 0% indica predicción

perfecta y a menor valor mayor precisión.

En contextos de alta variabilidad como el de las MiPymes con poco historial de datos, valores por encima del 30% son completamente esperados en los primeros años de operación, como se evidencia en los resultados obtenidos, donde el SMAPE global mejoró progresivamente conforme el modelo acumuló más información del negocio.

En 2023 el sistema partió prácticamente desde cero con datos disponibles solo desde abril. Un error cercano al 47% era completamente esperado y no representa un fallo del modelo, sino una limitación natural de cualquier sistema predictivo sin historial suficiente para aprender los patrones del negocio. Para 2024, con un año completo de datos disponible, el error neto se redujo a -7,7%, lo que representa una mejora muy significativa.

En 2025 el comportamiento se mantuvo estable en niveles similares a 2024, confirmando que el sistema alcanzó un nivel de funcionamiento consistente. Esta evolución demuestra un principio clave: el modelo mejora con el tiempo a medida que acumula más información del negocio, y su potencial real se hace visible a partir del segundo año de operación.

Análisis de precisión — Año 2025

El año 2025 fue el año de validación principal. El sistema fue evaluado en dos dimensiones: unidades físicas vendidas e ingresos en dinero.

Métricas globales — Unidades:	
Métrica	Valor
Error neto	-16,3%
Error real (WAPE)	20,4%
MAE	63 unidades/mes
RMSE	84 unidades/mes
sMAPE	24,4%
Total sobreestimado	+48 UND
Total subestimado	-648 UND

Métricas globales — Ingresos:	
Métrica	Valor
Error neto	-15,4%
Error real (WAPE)	37,0%
MAE	\$1.100.000 COP/mes
RMSE	\$1.200.000 COP/mes
sMAPE	42,6%
Total sobreestimado	+\$2.500.000 COP
Total subestimado	-\$9.700.000 COP

La diferencia entre el error en unidades (24,4%) y en dinero (42,6%) tiene una explicación directa: el sistema predice con mayor precisión cuántos productos se venderán que el valor monetario total de esas ventas. Esto ocurre porque

en una papelería conviven productos de precios muy distintos, desde artículos de bajo costo como lapiceros y cuadernos hasta artículos de mayor valor como calculadoras o cartuchos. Cuando la mezcla de productos que se vende en un mes cambia respecto al historial, el total en dinero varía aunque la cantidad sea similar. Este comportamiento es propio del tipo de negocio y no indica un fallo del modelo.

Detalle mes a mes — Ingresos en dinero (2025):					
Periodo	Real	Predicción esperada	Error %	MAE	sMAPE
Ene-25	\$4.350.400	\$2.394.766	-45,0%	\$2,0M	88,0%
Feb-25	\$4.967.500	\$3.803.761	-23,4%	\$1,2M	26,5%
Mar-25	\$4.235.000	\$2.754.213	-35,0%	\$1,5M	42,4%
Abr-25	\$2.022.700	\$2.302.952	+13,9%	\$280K	13,0%
May-25	\$1.452.700	\$1.650.095	+13,6%	\$197K	12,7%
Jun-25	\$2.993.100	\$1.806.318	-39,7%	\$1,2M	41,5%
Jul-25	\$3.185.100	\$1.482.789	-53,4%	\$1,7M	72,9%
Ago-25	\$3.185.700	\$4.294.240	+34,8%	\$1,1M	29,6%
Sep-25	\$1.654.500	\$2.588.358	+56,4%	\$934K	44,0%
Oct-25	\$2.051.600	\$1.181.549	-42,4%	\$870K	53,8%
Nov-25	\$2.641.700	\$1.330.412	-49,6%	\$1,3M	66,0%

Los meses con mejor desempeño fueron abril (sMAPE 13,0%) y mayo (12,7%), donde las predicciones se acercaron mucho a la realidad. Los meses con mayor error fueron enero (88,0%), julio (72,9%) y noviembre (66,0%). Enero muestra el error más alto del año porque el modelo aún no tenía suficiente información sobre el comportamiento de inicio de año escolar, que

es uno de los picos más importantes para una papelería. Julio y noviembre corresponden a períodos de cambio de temporada escolar, patrones que el modelo no había capturado completamente en su primer ciclo anual. Con los datos de 2025 ya incorporados al historial, se espera que estas desviaciones disminuyan durante 2026.

Un aspecto positivo a destacar es que el sistema tiende a ser prudente en sus estimaciones: en dinero subestimó \$9,7 millones acumulados frente a solo \$2,5 millones de sobreestimación, y en unidades subestimó 648 frente a solo 48 de sobreestimación. Esto es operativamente favorable para el negocio, ya que una predicción conservadora lleva al dueño a reabastecer con cautela sin arriesgarse a comprar de más.

Proyecciones para 2026

Con base en los tres años de historial recopilados, el sistema generó las proyecciones para el año 2026. El ingreso esperado es de \$31.627.379 COP, con un escenario optimista de \$60.212.409 COP y uno conservador de \$5.954.207 COP. En cuanto a unidades, se proyectan 3.307 unidades, con un escenario optimista de 5.690 y uno conservador de 794. El promedio mensual estimado es de \$2.635.615 COP.

El sistema identifica octubre como el mejor mes proyectado, impulsado por productos como lapiceros Parker y calculadoras Cassio, y mayo como el mes de menor actividad esperada. Esta información le permite al negocio planificar sus compras con anticipación, reforzar el inventario antes de los meses de mayor demanda y ser más cuidadoso en los meses de baja rotación.

Limitaciones del sistema

Se identificaron tres limitaciones principales. Primero, el volumen de datos del primer año fue insuficiente para capturar completamente los patrones estacionales del negocio, lo que explica los errores más altos en 2023 y en ciertos meses de 2025. Segundo, la infraestructura gratuita utilizada impone restricciones de velocidad y capacidad que limitan la escala de datos procesables, aunque son suficientes para el caso de uso actual. Tercero, el sistema está diseñado para productos con precio fijo, quedando fuera del alcance artículos vendidos por peso o medida variable.

Estas limitaciones son propias del contexto en el que opera cualquier sistema predictivo implementado en una MiPyme en sus primeros años de uso y no representan fallas del sistema. A medida que el negocio continúe operando con TargetBI, el modelo irá acumulando más historial de forma automática y sus predicciones se volverán naturalmente más precisas.

3.3 Conclusiones

Se logró desarrollar e implementar TargetBI, un aplicativo web funcional que integra predicción de ventas, gestión de inventario y punto de venta en una sola herramienta accesible para MiPymes del sector retail, demostrando que es posible construir soluciones de inteligencia de negocios con infraestructura de bajo costo y datos recopilados manualmente.

La evolución del error del modelo a lo largo de los tres años evaluados confirma que los sistemas de aprendizaje automático aplicados a MiPymes requieren de un período de maduración. El error neto pasó de +46,8% en 2023 a -7,7% en 2024, evidenciando una mejora sustancial a partir de cada año de operación cuando el modelo dispuso de un historial completo.

El sistema predice con mayor precisión las unidades vendidas (sMAPE 24,4%)

que los ingresos en dinero (sMAPE 42,6%), diferencia explicada por la heterogeneidad de precios dentro del catálogo de una papelería. Este hallazgo es relevante para entender las capacidades y límites del modelo en negocios con alta variedad de productos y precios distintos.

La recolección manual de datos en un negocio sin sistema digital previo, aunque costosa en tiempo, demostró ser viable y suficiente para alimentar el modelo predictivo. Esto confirma que la falta de digitalización previa no es un obstáculo insuperable para que una MiPyme adopte herramientas de analítica de datos.

El uso de la metodología XP con un equipo de dos personas permitió desarrollar un sistema completo con división clara de responsabilidades e integración continua, validando que esta metodología es apropiada para proyectos de desarrollo de software en contextos académicos con recursos limitados.

3.4 Recomendaciones

Se recomienda que la papelería caso de estudio continúe operando con TargetBI de forma constante, registrando todas sus ventas diarias a través del módulo de punto de venta, ya que cada transacción registrada mejora la calidad de las predicciones futuras de forma automática.

Se recomienda que en futuras implementaciones del sistema en otras MiPymes se priorice la recolección de al menos tres años de datos históricos antes de tomar decisiones estratégicas basadas en las predicciones, dado que el modelo muestra potencial a partir de mas años de operación.

Se recomienda explorar la posibilidad de migrar hacia infraestructura de pago de bajo costo a medida que el negocio crezca, ya que los servicios gratuitos actuales, aunque suficientes para el caso de uso presente, tienen limitaciones

de capacidad que podrían volverse restrictivas con un mayor volumen de productos o transacciones.

Se recomienda extender el sistema para que soporte productos vendidos por peso o medida variable, lo que ampliaría el universo de MiPymes que podrían beneficiarse de la herramienta más allá de negocios con precios fijos.

Se recomienda que otros equipos que adopten esta solución documenten sus resultados de precisión año a año, ya que la generación de más casos de estudio en diferentes tipos de MiPymes permitirá seguir validando y mejorando los modelos utilizados.

2.9 Bibliografía

1. **Peña Sanz, I. (2024). *Big data e inteligencia artificial en la predicción y mejora de los procesos de venta*. España. Recuperado de <https://repositorio.comillas.edu/jspui/bitstream/11531/78819/2/TFG%20-%20Pena%20Sanz%2C%20Ignacio.pdf>**
Fecha de consulta: 1 de febrero de 2026.
2. **Alonso Antonio Miguel Rojas. (2024). *Desarrollo de modelos de pronóstico de demanda con variables predictoras para mejorar la gestión de inventarios en una empresa de alimentos*. Chile. Recuperado de <https://repositorio.uchile.cl/bitstream/handle/2250/204012/Desarrollo-de-modelos-de-pronostico-de-demanda-con-variables-predictoras-para-mejorar-la-gestion-de-inventarios-en-una-empresa-de-alimentos.pdf>**
Fecha de consulta: 1 de febrero de 2026.
3. **Enoc Gutiérrez Pallares. (2024). *Diseño metodológico para conciliación de inventarios en mi-pymes de operaciones logísticas*. México. Recuperado de <https://rinacional.tecnm.mx/jspui/handle/TecNM/8484>**
Fecha de consulta: 1 de febrero de 2026.
4. **Fernando Enrique & Oscar Darío. (2023). *Implementación de un sistema web y su influencia en la percepción en los procesos de ventas y control de inventario en la empresa Servicell e Imports C & R, Cajamarca*. Perú. Recuperado de <https://drive.google.com/file/d/1WYJhGHRCr-D5P4VyuwSumRxq79mfnjcj/view>**
Fecha de consulta: 1 de febrero de 2026.

5. Flores Taquiri & Montalvo Celis. (2024). *Machine learning para la predicción en la gestión de inventarios dirigida a pymes de venta de productos tecnológicos*. Perú. Recuperado de <https://repositorio.ucv.edu.pe/handle/20.500.12692/151376>

Fecha de consulta: 1 de febrero de 2026.

6. Andrade Maldonado & Viacava Malaga. (2024). *Propuesta de mejora de la planificación del abastecimiento a través de la gestión de inventarios para la disminución del sobrestock en una mype de retail*. Perú. Recuperado de <https://repositorioacademico.upc.edu.pe/handle/10757/683747>

Fecha de consulta: 1 de febrero de 2026.

7. Bernaola Velarde & Varillas Trujillo. (2022). *Sistema predictivo con Machine Learning para la gestión de inventario para la empresa Inversiones Ferreteras Mendoza S.A.C*. Perú. Recuperado de https://alicia.concytec.gob.pe/vufind/Record/UCVV_4bd6e3b6f2fa78892699d59081992189

Fecha de consulta: 1 de febrero de 2026.

8. Jorge Aníbal & Alejandro Quiñonez & Edgar Julian. (2022). *Acopi: digitalización y desarrollo sostenible de la mipymes en Colombia*. Colombia. Recuperado de <https://repositorio.unillanos.edu.co/server/api/core/bitstreams/bbd44333-9969-4de2-a055-5c0379f5db40/content>

Fecha de consulta: 1 de febrero de 2026.

9. **Juan Nicolás Rojas López. (2023). *Diseño de herramientas de apoyo para la gestión financiera y de costos de tiendas de la comuna cuatro de la ciudad de Popayán.* Colombia. Recuperado de <https://fupvirtual.edu.co/repositorio/files/original/300398c179b00bd3a0420ca26ef803d4c2c2abe6.pdf>**

Fecha de consulta: 1 de febrero de 2026.

10. **Ana María Díaz Hernández. (2024). *Análisis del impacto de la gestión contable en la efectiva sostenibilidad de las microempresas del sector comercial en Medellín.* Medellín <https://repository.uniminuto.edu/server/api/core/bitstreams/b56a36cd-a599-4928-873b-914606d7a49c/content>**

Fecha de consulta: 1 de febrero de 2026.

Anexo A: Carta de certificación de revisión del anteproyecto

Anexo B: Carta de aprobación de la propuesta.



FACULTAD DE INGENIERIA Y TECNOLOGICAS
Programa Ingeniería de Sistemas



Valledupar, 12 de junio de 2025

Señores:

DANIEL RICARDO PIÑA CORZO

LUIS FABIAN CONTRERAS MALDONADO

Estudiante de Ingeniería de Sistemas Ciudad

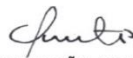
Cordial saludo,

El comité de investigación del programa de Ingeniería de Sistemas de la Universidad Popular del Cesar se permite informarle que, en sesión de la fecha arriba indicada, ha revisado la propuesta entregada por usted con la cual manifiesta haber atendido los requerimientos de este comité, con respecto a la propuesta "APLICATIVO WEB PARA LA PREDICCIÓN DE VENTAS Y OPTIMIZACIÓN DEL ABASTECIMIENTO EN TIEMPO REAL DE LAS MIPYMES DEL SECTOR RETAIL EN VALLEDUPAR.", hecho lo cual emite el siguiente concepto:

Se aprueba la pertinencia de la propuesta, en consecuencia, ajustar propuesta con el director si es necesario y proceder con la construcción del **ANTEPROYECTO**. Este deberá ser presentado ante este comité para su revisión y asignación de jurados, quienes evaluarán y calificarán el trabajo como fase previa al Proyecto Final; Es importante tener en cuenta que los jurados podrán realizar las observaciones que consideren necesarias sobre el anteproyecto.

Al momento de entregar su anteproyecto, tengan en cuenta lo siguiente:

- 1.- Acatar lo dispuesto en el protocolo de anteproyecto.
- 2.- Adjuntar el último documento de propuesta que revisó el comité.
- 3.- Adjuntar la presente carta
- 4.- Entregar el documento del Anteproyecto en Formato Word, al correo investigacionsistemas@unicesar.edu.co
- 5.- Entregar informe generado por la aplicación turnitin



ALVARO OÑATE BOWEN
Director Comité de Investigación
Programa de Ingeniería de Sistemas
Universidad Popular del Cesar

www.unicesar.edu.co

Teléfono conmutador PBX: (+57 605 588 5592)

Balneario Hurtado, Vía a Patillal

Valledupar - Cesar, Colombia

Anexo C: Resultado de revisión que evidencie una tolerancia máxima de 25% de no originalidad.

El informe de turnitin no se pudo realizar dado que la universidad no cuenta con una licencia activa.