

**Carteras en riesgo: inteligencia artificial para la predicción de degradación financiera en
Colombia**

Diego Fernando Cano Idarraga

Asesor

Rafael Gaitan Ospina

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI
Especialización en Ciencia de Datos y Analítica
2025

Dedicatoria

A mis padres, por el amor transformado en sacrificio y la palabra que se vuelve aliento. A mis gatos, testigos silenciosos de mis dudas y mis triunfos; por habitar mis madrugadas y convertir el cansancio en compañía, recordándome siempre que la soledad del investigador es más dulce cuando se comparte con quienes aman sin condiciones.

Resumen

En medio de la economía colombiana que enfrenta constantes altos y bajos en el nivel de endeudamiento de la población, las entidades financieras están viendo cómo sus carteras de crédito se deterioran cada vez más a través del tiempo. Esta propuesta buscó generar herramientas de Análisis de Datos y Aprendizaje Automático para anticiparse a estos problemas. Para lograrlo, se abordó el problema con una estrategia dual: primero, se aplicó clusterización y PCA para segmentar a las entidades en seis perfiles de riesgo comparables; y segundo, se implementaron modelos de Series Temporales, donde el modelo Redes Neuronales LSTM demostró ser el más robusto en la solución. Estudiar el comportamiento anterior y aplicar este modelo predictivo, que alcanzó un error porcentual absoluto medio (MAPE) del 13.59%, permitiendo identificar señales tempranas de incumplimiento. Logrando que las compañías pueden actuar e integrar esta arquitectura de Deep Learning para optimizar el cálculo de provisiones y fortalecer su estrategia ante los desafíos del sector financiero.

Palabras clave: Degradación, Riesgo, Cartera, Predicción, Machine Learning

Abstract

Amidst the Colombian economy facing constant ups and downs in population debt levels, financial institutions are seeing their credit portfolios deteriorate over time. This proposal sought to generate Data Analysis and Machine Learning tools to anticipate these problems. To achieve this, the problem was approached with a dual strategy: first, Clustering and PCA were applied to segment the entities into six comparable risk profiles; and second, Time Series models were implemented, where the LSTM Neural Networks model proved to be the most robust solution. Studying past behavior and applying this predictive model, which achieved a Mean Absolute Percentage Error (MAPE) of 13.59%, allows for the identification of early default signals. This enables companies to act and integrate this Deep Learning architecture to optimize the calculation of provisions and strengthen their strategy against the challenges of the financial sector.

Keywords: Degradation, Risk, ortfolio, Prediction, Machine Learning

Tabla de Contenido

Introducción	13
Descripción del Problema	14
Justificación	16
Objetivos	17
Objetivo General.....	17
Objetivos Específicos	17
Marco Teórico.....	18
Estrategia y Técnica para la Toma de Decisiones	18
Aporte de la Ciencia de Datos a las Decisiones	18
Como Tomamos Decisiones	19
Aprendizaje Automático en el Sector Financiero.....	19
Marco Conceptual.....	20
Modelado de Datos	20
Modelos Descriptivos	20
Agrupamiento o Clusterización	20
Correlaciones	20
Reglas de Asociación.....	20
Modelos Predictivos	21
Clasificación	21
Regresión	21
Series Temporales.....	21
Técnicas Bayesianas	21

Redes Neuronales	21
Marco Contextual.....	22
Que Modelo Usar para la Degradación de Carteras	22
Redes Neuronales (MLP o LSTM para Series Temporales)	22
Métricas	22
Métricas Score	23
Estrategias.....	23
Modelos Supervisados para Predicción de Deterioro	24
Árboles de Decisión y Random Forest	24
XGBoost	24
Redes Neuronales Artificiales (ANN)	24
Regresión Logística	25
Técnicas no Supervisadas para Segmentación	25
K-Means	25
DBSCAN	25
Clustering Jerárquico	26
Evaluación de Modelos y Métricas de Desempeño	26
Clasificación (Deterioro de Cartera).....	26
Metodología de Modelado y Segmentación de Riesgo de Cartera.....	28
El Modelado de Riesgo de Cartera: Fundamentos y Necesidad de Estandarización Proporcional	28
Ingeniería de Características Clave	29
Variables de Cartera y Conceptos de Mora	30

Ingeniería de Características y Métricas de Riesgo	31
Tasa de Cartera Vencida (TCV).	31
Proporción de Alto Riesgo Agregado (Prop_Riesgo_BCD&E).....	31
Conjunto de Variables para el Modelado	31
Preprocesamiento y Reducción de Dimensionalidad	32
Análisis de Componentes Principales (PCA)	32
Varianza Explicada por Componente Principal.....	32
Modelado y Validación de Clustering con MLflow	36
Resultados de la Conclusión y Elección del Modelo.....	37
Análisis e Interpretación de los Seis Clústeres de Riesgo	39
Implicaciones Sectoriales por Segmento	41
Conclusiones Metodológicas y Confianza del Modelo	42
Predicción Degradación de Cartera	44
Arquitectura de Modelado de Series Temporales Financieras	47
Enfoque Jerárquico y Segmentación	47
Variables Exógenas de Riesgo (X).....	48
Fundamentos de los Modelos de Predicción (Benchmarking)	49
Modelo 1 SARIMAX (Estadístico Lineal).....	49
Formulación General	49
Configuración Implementada	49
Modelo 2 Prophet (Regresión Aditiva).....	49
Modelo 3 LSTM (Deep Learning Recurrente)	50
Modelo 4 Árboles de Decisión (No Lineal)	50

Preprocesamiento Normalización	50
Arquitectura de Celda y Compuertas.....	50
Optimización y Grid Search	50
Criterios de Evaluación y Métrica	51
Métricas de Error	51
Error Porcentual Absoluto Medio (MAPE).....	51
Error Absoluto Medio (MAE)	51
Estrategia de Selección (MLOps con MLflow).....	51
Resultados y Discusión del Modelo Ganador.....	52
Resumen de Rendimiento Final.....	52
Parámetros Finales del LSTM Ganador	53
Benchmarking de Modelos	57
Parámetros Clave.....	58
Hiperparámetros Óptimos del LSTM.....	58
Normalización y Secuenciación	59
Arquitectura y Entrenamiento Final	59
Artefactos de Producción Resultantes	60
Arquitectura de Despliegue y Tecnologías.....	61
Funcionalidades Clave y Flujo de Trabajo	61
Propósito y Funcionalidad General	61
Arquitectura de Inferencia y Herramientas.....	62
Métodos Clave para la Gestión del Riesgo.....	63
Validación de la Incertidumbre y Cobertura.....	64

Cuantificación del Riesgo (Bandas de Confianza)	64
Indicadores Clave y Semáforos de Gestión	65
Recomendaciones para la Gestión de la Cartera Crediticia	70
Recomendaciones Sustentadas en la Segmentación y Patrones de la Cartera.....	70
Gestión de Cartera Basada en la Certeza Predictiva.....	70
Segmentos de Alta Certeza (Recurso Estratégico)	70
Segmentos de Baja Certeza/Problemáticos (Señal de Alerta).....	71
Direccionamiento Estratégico Operacional	72
Estandarización de la Arquitectura LSTM	72
Enfoques de Modelado Avanzado para Segmentos Volátiles	73
Implementación de un Sistema de Alerta Temprana (SAT).....	73
Conclusiones	75
Referencias Bibliográficas	77

Lista de Figuras

Figura 1 <i>Análisis PCA de Componentes Principales</i>	34
Figura 2 <i>Matriz de Correlación de Saldos</i>	35
Figura 3 <i>Matriz de Correlación de Saldos Totales</i>	35
Figura 4 <i>Método del Codo para Asignación de Clúster</i>	36
Figura 5 <i>Asignación de Segmentos</i>	37
Figura 6 <i>Resultados Entrenamiento y Clusterización</i>	38
Figura 7 <i>Testing de Clusterización y Segmentos</i>	38
Figura 8 <i>Agrupación de Clúster 0 - por TCV y Proporción del Riesgo</i>	40
Figura 9 <i>Perfil del Riesgo Promedio x Clúster</i>	43
Figura 10 <i>Importancia de Variables en Modelo</i>	44
Figura 11 <i>Evolución del TCV de 3 a 4 Meses</i>	45
Figura 12 <i>Evolución de Cartera Vencida Mayor a 4 Meses</i>	45
Figura 13 <i>Evolución de Cartera Vencida de 3 a 6 Meses</i>	46
Figura 14 <i>Validación del Modelo Valores Reales vs Pruebas</i>	46
Figura 15 <i>Predicciones del Modelo Resultante</i>	47
Figura 16 <i>Tendencia Histórica y Proyección del Riesgo Bancolombia</i>	54
Figura 17 <i>Composición Histórica del Saldo de Cartera de Riesgo</i>	54
Figura 18 <i>Tendencia Histórica y Proyección de Saldo Vencido</i>	55
Figura 19 <i>Tendencia Histórica y Proyección de Saldo Vigente</i>	55
Figura 20 <i>Tendencia Histórica y Proyección de TCV</i>	56
Figura 21 <i>Dashboard de Proyección de Cartera LSTM</i>	66
Figura 22 <i>Tendencia Inmediata Real vs Próximo Mes</i>	66

Figura 23 <i>Aceptación del Back-Testing</i>	67
Figura 24 <i>Proyecciones - Saldo Vigente</i>	67
Figura 25 <i>Proyecciones Saldo Vencido</i>	68
Figura 26 <i>Composición del Riesgo</i>	68
Figura 27 <i>TCV Ratio</i>	68
Figura 28 <i>Proyecciones Riesgo BCDE</i>	69
Figura 29 <i>Extraer Resultados Proyección</i>	69

Lista de Tablas

Tabla 1 <i>Métricas de Entrada Para el Modelo</i>	29
Tabla 2 <i>Variables del Modelo</i>	30
Tabla 3 <i>Análisis de Componentes Principales</i>	32
Tabla 4 <i>Métricas y Resultados</i>	37
Tabla 5 <i>Segmentación de Carteras</i>	39
Tabla 6 <i>Análisis Sectorial de Resultados</i>	41
Tabla 7 <i>Variables Exógenas Degradación</i>	48
Tabla 8 <i>Modelo LSTM y Resultados de Modelos</i>	52
Tabla 9 <i>Hiperparámetros LSTM para Modelo</i>	53
Tabla 10 <i>Herramientas Machine Learning</i>	57
Tabla 11 <i>Parámetros de Entrada Modelo LSTM</i>	58
Tabla 12 <i>Grid Search Modelo LSTM</i>	58
Tabla 13 <i>Artefactos y Despliegue del Modelo</i>	60
Tabla 14 <i>Frontend Streamlit para Modelo</i>	61
Tabla 15 <i>Funcionalidades Dashboard</i>	62
Tabla 16 <i>Arquitectura y Herramientas</i>	63
Tabla 17 <i>Indicadores y Métricas</i>	65

Introducción

La actual transformación digital y evolución del entorno económico, es esencial para analizar los factores que inciden en la estabilidad del sistema financiero colombiano. La comprensión de estos elementos nos permite interpretar los fenómenos de riesgo crediticio, y también proponer soluciones que sean innovadoras que puedan responder a las necesidades de anticipación y gestión del deterioro de las carteras.

En la economía colombiana, caracterizada por una alta volatilidad y limitaciones estructurales en la capacidad adquisitiva de la población, el sobreendeudamiento se ha convertido en un fenómeno recurrente. Según reportes del Banco de la República (2024), el ciclo de crédito en Colombia ha mostrado una sensibilidad creciente a las tasas de interés, lo que representa un riesgo directo para el sector financiero. Este fenómeno incrementa la probabilidad de incumplimiento por parte de los deudores y, en consecuencia, acelera la degradación de las carteras crediticias, afectando los indicadores de calidad y solvencia de las instituciones (Superintendencia Financiera de Colombia, 2023).

Bajo este contexto, el presente documento propone un enfoque basado en técnicas de ciencia de datos y aprendizaje automático, integrando la capacidad predictiva de algoritmos avanzados para superar la rigidez de los métodos tradicionales, Se busca así identificar patrones ocultos y fortalecer la toma de decisiones en entidades con mayor exposición al riesgo de degradación

Descripción del Problema

En la economía colombiana, caracterizada por una alta sensibilidad a factores macroeconómicos externos y limitaciones estructurales en la capacidad adquisitiva de los hogares, el sobreendeudamiento se ha consolidado como un fenómeno recurrente. Según reportes recientes de estabilidad financiera (Banco de la República, 2024), el incremento en las tasas de interés y la inflación han presionado la carga financiera de los usuarios, reduciendo su capacidad de pago. Este escenario representa un riesgo crítico para el sistema financiero, ya que incrementa exponencialmente la probabilidad de incumplimiento (Probability of Default - PD) y, en consecuencia, acelera la degradación de las carteras crediticias, deteriorando los indicadores de calidad y cobertura de las entidades (Superintendencia Financiera de Colombia, 2024).

La gestión del riesgo crediticio es una práctica normativa y estratégica fundamental, regida en el país por el Sistema de Administración de Riesgo Crediticio (SARC). Tradicionalmente, las entidades han mitigado el impacto de los impagos mediante modelos de segmentación y credit scoring basados en métodos paramétricos. No obstante, en economías emergentes como la de Colombia, la volatilidad y la informalidad laboral generan dinámicas que suelen superar la capacidad predictiva de los modelos tradicionales como la regresión logística, los cuales pueden presentar rigidez ante cambios abruptos en el ciclo económico (Altman & Saunders, 1997; Baesens, 2016).

A pesar de los esfuerzos institucionales y regulatorios por controlar el riesgo, la rápida evolución del comportamiento del consumidor, impulsada por la digitalización y el acceso a créditos de bajo monto, ha alterado la detección temprana de señales de alerta. Los patrones de deterioro inminente se han vuelto más complejos y menos lineales, lo que dificulta su identificación mediante herramientas estadísticas convencionales. Esta situación plantea la

necesidad imperativa de incorporar enfoques analíticos avanzados, como el Machine Learning y el Análisis de Datos, que permitan una comprensión más profunda y dinámica de los perfiles de riesgo.

Para abordar este problema, se propone la implementación de modelos predictivos y técnicas de análisis de datos que complementen la metodología tradicional. Estas herramientas están diseñadas para identificar patrones ocultos en el comportamiento transaccional y sociodemográfico, permitiendo anticipar escenarios de riesgo con mayor precisión y fortalecer la toma de decisiones estratégicas en las entidades financieras colombianas con mayor exposición a la degradación de sus activos.

Justificación

La elección de este tema responde a la necesidad de comprender con mayor profundidad una problemática que impacta directamente la estabilidad del sistema financiero colombiano: la degradación de las carteras crediticias. Este fenómeno no solo representa un desafío técnico para las entidades financieras, sino que también tiene consecuencias prácticas que afectan la economía de los hogares, la sostenibilidad de las instituciones y la confianza en el sistema financiero.

A pesar de los avances en modelos tradicionales de evaluación de riesgo, existe una limitada cantidad de investigaciones locales que integren enfoques de ciencia de datos y aprendizaje automático para anticipar el deterioro de las carteras. Esta carencia genera vacíos tanto en el conocimiento como en la formulación de estrategias preventivas eficaces. Por ello, esta investigación busca aportar evidencia empírica y reflexión crítica que enriquezca el debate técnico y facilite la toma de decisiones informadas en el sector.

Los hallazgos obtenidos podrían servir como base para el desarrollo de herramientas analíticas que fortalezcan la gestión del riesgo crediticio. Además, su enfoque aplicado permite demostrar cómo la analítica avanzada puede contribuir a resolver problemas reales en contextos económicos emergentes como el colombiano.

Objetivos

Objetivo General

Evaluar el comportamiento histórico de las carteras en el sector financiero colombiano mediante técnicas de análisis de datos y aprendizaje automático.

Objetivos Específicos

Clasificar las variables históricas de las carteras del sector financiero según su impacto en la degradación del portafolio de créditos.

Comprobar la efectividad de las técnicas de clusterización para segmentar entidades financieras con base en el comportamiento observado en las carteras.

Calcular métricas de desempeño para modelos de aprendizaje automático aplicados a la predicción del deterioro de las carteras con modelos de redes neuronales y árboles de decisión.

Formular recomendaciones para la gestión de la cartera crediticia, sustentadas en los patrones identificados y en los resultados obtenidos mediante el análisis predictivo y la segmentación bancaria.

Marco Teórico

Estrategia y Técnica para la Toma de Decisiones

La ciencia encuentra aplicación en infinidad de casos de uso donde las decisiones deben estar basadas en hechos y datos. Es crucial que toda decisión sea soportada e informada desde los datos. Un ejemplo claro de esta aplicación es la venta de productos: los estudios de mercado permiten evaluar cómo puede comportarse un producto frente a la competencia, cuál es su promesa de valor y qué lo hace diferente frente a los demás.

Ser estrategias para tomar decisiones es imprescindible para el éxito de cualquier organización. Imaginemos asignar créditos a todos sin un fundamento técnico; la probabilidad de fracaso sería inminente. La facilidad de tomar decisiones soportadas por datos el activo más imprescindible de una organización permite tener una visión 360° del producto y sus riesgos donde las estrategias nacen de las decisiones, y las decisiones, de los datos.

Aporte de la Ciencia de Datos a las Decisiones

Los datos son una ciencia que se encuentra en auge. Una mirada al pasado, cuando miles de empresas apenas conocían el internet, no permitiría imaginar cómo toda esta historia se convertiría en información capaz de transformar a cada compañía. Esta ciencia ha llegado con un auge y una fuerza significativa.

Esta disciplina ha irrumpido con un crecimiento y una influencia significativos.

Cada día nace la necesidad de contar con más y más profesionales con la capacidad de respaldar las decisiones empresariales: una visión de producto, una visión del riesgo, un soporte técnico y tecnológico para la toma de decisiones.

Como Tomamos Decisiones

Los datos en sí no representan mayor valor si no son analizados. Imagina que tienes miles y miles de datos, pero no sabes qué hacer con ellos. Entonces, ¿cómo tomamos decisiones?

Como toda ciencia, existe un proceso de análisis previo para cualquier actividad.

Uno de los mayores retos es almacenar, procesar, exportar y convertir los datos en información valiosa. Aquí es donde entran los modelos de datos, los cuales buscan dar respuesta a un problema y/o hipótesis específica: ¿qué deseas responder?, ¿qué necesidad tienes?, ¿qué deseas demostrar?

Por lo tanto, el componente primario en la toma de decisiones es el análisis. Automatizar los modelos es una tarea ardua que requiere capacidades técnicas y talento especializado en el tratamiento de datos.

Aprendizaje Automático en el Sector Financiero

El aprendizaje automático (Machine Learning, ML) es una rama de la inteligencia artificial que permite a los sistemas aprender patrones a partir de datos sin ser explícitamente programados. En el sector financiero, estas técnicas se han convertido en herramientas clave para la evaluación de riesgos, la predicción de impagos, la segmentación de clientes y la optimización de portafolios de inversión.

El análisis de carteras de crédito mediante ML permite identificar patrones de deterioro, clasificar riesgos y anticipar comportamientos adversos, lo cual es fundamental para la estabilidad del sistema financiero.

Marco Conceptual

Modelado de Datos

El modelado de datos es el proceso de crear representaciones visuales y estructurales de un sistema de información. En el contexto de este trabajo, el modelado no solo busca organizar la información, sino estructurar las dimensiones que permitirán la creación de grupos o clusters coherentes. A través de niveles de abstracción, se transforman datos brutos de clientes y transacciones en variables técnicas que facilitan la identificación de patrones de comportamiento.

Modelos Descriptivos

Los modelos descriptivos son fundamentales para este estudio, ya que su objetivo principal es resumir y organizar los datos para reflejar la realidad de un proceso. Dentro de esta categoría, el agrupamiento es la técnica principal:

Agrupamiento o Clusterización

Es el núcleo de esta investigación. Consiste en técnicas de aprendizaje no supervisado que agrupan registros basándose en similitudes intrínsecas no etiquetadas previamente. Para este trabajo, la clusterización permite segmentar clientes y formular estrategias de fidelización diferenciadas

Correlaciones

En el análisis de clusters, la correlación es vital para la selección de variables. Ayuda a identificar qué características (lineales o no) influyen más en la formación de un grupo, evitando la redundancia de datos que pueda sesgar el agrupamiento.

Reglas de Asociación

Complementan al clustering al identificar qué eventos ocurren simultáneamente dentro de un segmento ya definido (ej. "los clientes del Cluster A suelen comprar X cuando adquieren Y").

Modelos Predictivos

Una vez definidos los grupos mediante técnicas descriptivas, los modelos predictivos permiten anticipar cómo evolucionarán dichos segmentos o cómo se comportarán nuevos individuos frente a ellos:

Clasificación

Mientras el clustering crea los grupos, la clasificación se utiliza para asignar nuevos registros a los clusters ya existentes. Por ejemplo, permite determinar si un nuevo cliente entrará en el "Segmento de Buen Pagador" basándose en sus rasgos de entrada.

Regresión

Se utiliza para validar la consistencia de los clusters. Permite entender cómo varía una variable dependiente (como el volumen de ventas) dentro de cada grupo identificado.

Series Temporales

Permiten analizar la volatilidad de los clusters a través del tiempo, pronosticando si un segmento de clientes crecerá o desaparecerá en el futuro.

Técnicas Bayesianas

Aportan un enfoque probabilístico al clustering, permitiendo calcular la probabilidad de que un registro pertenezca a un grupo específico en lugar de realizar una asignación rígida.

Redes Neuronales

Pueden emplearse para realizar segmentaciones complejas (Deep Clustering) en conjuntos de datos masivos, detectando relaciones no lineales que los métodos tradicionales de agrupación podrían pasar por alto.

Marco Contextual

Que Modelo Usar para la Degradación de Carteras

Existen múltiples modelos de todos los tipos, clases, afinidades entre muchas características, sin embargo, para el caso de degradación de carteras vamos a utilizar los siguientes modelos.

- Random Forest (RF)
- XGBoost
- Regresión Logística (para baseline)
- Random Forest y XGBoost permiten evaluar la importancia de las

variables (feature importance) de forma robusta.

- Son modelos no lineales, ideales para datos financieros con relaciones complejas.
- Regresión logística sirve como modelo base interpretable.
- Importancia de variables (Gini, Gain, SHAP values)
- AUC-ROC, Precisión, Recall, F1-score

Redes Neuronales (MLP o LSTM para Series Temporales)

- Árboles de Decisión / Random Forest / XGBoost
- Redes neuronales capturan patrones complejos y no lineales, útiles si tienes

muchos datos históricos.

- Árboles de decisión son más interpretables y rápidos de entrenar.

Métricas

- Accuracy, Precision, Recall, F1-score
- AUC-ROC
- Matriz de confusión

- Log-loss (para modelos probabilísticos)

Comprobar la efectividad de técnicas de clusterización para segmentar entidades financieras.

- K-Means
- DBSCAN
- Hierarchical Clustering
- t-SNE o UMAP (para visualización)
- K-Means es eficiente y fácil de interpretar.
- DBSCAN detecta clusters de forma no lineal y maneja ruido.
- Clustering jerárquico permite ver relaciones entre entidades.
- t-SNE/UMAP ayudan a visualizar los clusters en 2D.

Métricas Score

- Silhouette Score
- Davies-Bouldin Index
- Dunn Index
- Validación visual con t-SNE o UMAP

Estrategias

- Comparar modelos con métricas cuantitativas (AUC, F1, etc.)
- Usar explicabilidad (SHAP, LIME) para justificar decisiones.
- Evaluar robustez y capacidad de generalización (validación cruzada, test set).
- Preprocesamiento: Normalización, imputación de valores nulos, codificación de variables categóricas.
- Validación cruzada: K-Fold o TimeSeriesSplit si hay dependencia temporal.

- Interpretabilidad: SHAP para modelos complejos como XGBoost o redes neuronales.

Modelos Supervisados para Predicción de Deterioro

Los modelos supervisados se entrenan con datos etiquetados, es decir, con ejemplos históricos donde se conoce el resultado (por ejemplo, si un crédito se deterioró o no). Entre los modelos más utilizados se encuentran:

Árboles de Decisión y Random Forest

- Son modelos basados en reglas que dividen los datos en función de variables predictoras.
- Random Forest es un ensamble de múltiples árboles que mejora la precisión y reduce el sobreajuste.
- Permiten interpretar la importancia de las variables, lo cual es útil para entender qué factores influyen más en el deterioro de cartera.

XGBoost

- Es una técnica de boosting que construye árboles de forma secuencial, corrigiendo errores del modelo anterior.
- Es altamente eficiente y ha demostrado un rendimiento superior en tareas de clasificación financiera.

Redes Neuronales Artificiales (ANN)

- Modelos inspirados en el cerebro humano, capaces de capturar relaciones no lineales complejas.
- Son útiles cuando se dispone de grandes volúmenes de datos históricos.

- En contextos temporales, se pueden usar variantes como LSTM para capturar secuencias y tendencias.

Regresión Logística

- Aunque es un modelo lineal, es ampliamente utilizado como línea base por su interpretabilidad.
- Permite estimar la probabilidad de deterioro de un crédito en función de variables explicativas.

Técnicas no Supervisadas para Segmentación

Las técnicas no supervisadas permiten descubrir estructuras ocultas en los datos sin necesidad de etiquetas. En el contexto financiero, se utilizan para segmentar entidades o clientes con comportamientos similares.

K-Means

- Algoritmo de partición que agrupa observaciones en K clústeres según su cercanía al centroide.
- Es eficiente y fácil de interpretar, aunque requiere definir el número de clústeres previamente.

DBSCAN

- Agrupa puntos densamente conectados y detecta ruido o valores atípicos.
- No requiere especificar el número de clústeres y es útil cuando los datos no tienen forma esférica.

Clustering Jerárquico

- Construye una jerarquía de clústeres que puede representarse como un dendrograma.
- Permite explorar relaciones entre entidades financieras a diferentes niveles de granularidad.

Evaluación de Modelos y Métricas de Desempeño

Para validar la efectividad de los modelos, se utilizan métricas específicas según el tipo de problema:

Clasificación (Deterioro de Cartera)

- Accuracy: proporción de predicciones correctas.
- Precision y Recall: útiles en contextos desbalanceados.
- F1-Score: media armónica entre precisión y recall.
- AUC-ROC: mide la capacidad del modelo para distinguir entre clases.
- b. Clusterización
- Silhouette Score: mide la coherencia interna de los clústeres.
- Davies-Bouldin Index: evalúa la separación entre clústeres.
- Visualización con t-SNE o UMAP: permite interpretar la calidad de la segmentación.

Interpretabilidad y Justificación de Modelos

La interpretabilidad es clave en entornos regulados como el financiero. Herramientas como SHAP (SHapley Additive exPlanations) y LIME (Local Interpretable Model-agnostic Explanations) permiten explicar las predicciones de modelos complejos, facilitando su adopción y validación por parte de entidades reguladoras.

- La elección del modelo debe basarse en:
- Rendimiento predictivo
- Capacidad de generalización
- Interpretabilidad
- Robustez ante datos faltantes o ruido

Metodología de Modelado y Segmentación de Riesgo de Cartera

El Modelado de Riesgo de Cartera: Fundamentos y Necesidad de Estandarización

Proporcional

El modelado de riesgo de cartera se fundamenta en la capacidad crítica de distinguir entre activos de crédito saludables y aquellos con un alto potencial de incumplimiento o pérdida. Para lograr esta distinción, la metodología propuesta se basa en la segmentación de entidades financieras mediante la agregación de sus saldos de cartera y la subsiguiente normalización en tasas y proporciones.

Este análisis no es solo una necesidad técnica, sino un imperativo regulatorio y estratégico. La clasificación del riesgo crediticio y la correcta provisión de los créditos son exigencias fundamentales dentro de la gestión de riesgo sistémico y la eficiencia del capital. Por lo tanto, el presente capítulo desarrolla una metodología para segmentar entidades basada en su exposición al riesgo crediticio.

Las variables financieras seleccionadas para este modelo reflejan dos dimensiones críticas del riesgo:

Mora (Créditos Vencidos): La exposición histórica al incumplimiento.

Calificación Intrínseca de Riesgo (Ratings A al E): La probabilidad de incumplimiento futuro.

El principal desafío descriptivo al trabajar con datos de cartera es la disparidad en los saldos absolutos entre las entidades. El rango intercuartil del saldo total de cartera es masivo, abarcando desde entidades pequeñas con saldos en millones hasta grandes bancos con volúmenes en billones. Utilizar los saldos absolutos sesgaría el modelo hacia el tamaño de la entidad y no hacia su verdadero perfil de riesgo.

Para superar este obstáculo y asegurar que el modelo sea equitativo y predictivo, la metodología se apoya rigurosamente en variables proporcionales (tasas). Esta estandarización permite que el perfil de riesgo de una entidad sea comparable con el de cualquier otra, independientemente de su volumen total de activos o de su tamaño en el sistema financiero. Así, la segmentación final refleja la estructura de riesgo y no simplemente la escala de las operaciones de la entidad.

Ingeniería de Características Clave

Se crearon seis métricas relativas que sirven como *inputs* al modelo, capturando el riesgo:

Tabla 1

Métricas de Entrada Para el Modelo

Métrica	Nombre Técnico	Concepto Financiero en el Modelo
Tasa de Cartera		Proporción de la cartera en mora significativa
Vencida (TCV)	TCV	(≥ 31 días). Mide la calidad del crédito pasado.
Prop. Riesgo B a E	Prop_Riesgo [B, C, D, E]	Mide el porcentaje del saldo total que cae en cada categoría de deterioro de riesgo (Bajo a Irrecuperable).
Prop. Alto Riesgo Agregado	Prop_Riesgo_BCD&E	Concentración de la cartera en las clasificaciones más problemáticas (B, C, D, E), funcionando como un indicador de solvencia.

Nota. Entrenamiento de variables de entrada del modelo.

Los valores faltantes (NaN) en las columnas de riesgo provenientes de imputación de valores cero o divisiones por cero en saldos) fueron imputados a cero para representar una

exposición nula a ese nivel de riesgo, y luego todas las métricas fueron estandarizadas con StandardScaler para homogenizar su escala.

Variables de Cartera y Conceptos de Mora

El análisis parte de la clasificación estándar de los activos crediticios. Las siguientes variables son los *inputs* brutos utilizados en la fase inicial del preprocesamiento:

Tabla 2

Variables del Modelo

Nombre Variable	Concepto	Clasificación de Riesgo
Saldo Cartera	Monto total de los créditos otorgados por la entidad a la fecha de corte, incluyendo capital e intereses.	Indicador de Tamaño/Exposición
Vencida 1 3M	Cartera con mora entre 31 y 90 días (primeras etapas del deterioro).	Riesgo intermedio
Vencida 3 6M	Cartera con mora entre 91 y 180 días (deterioro significativo).	Riesgo alto.
Vencida GT 6M	Cartera con mora superior a 180 días (pérdida potencial/irrecuperable).	Riesgo extremo.

Nota. Variables para la segmentación y entrada del modelo.

Ingeniería de Características y Métricas de Riesgo

Tasa de Cartera Vencida (TCV). Mide la porción de la cartera total que se encuentra en mora significativa (más de 30 días), actuando como un indicador clave de la calidad de la cartera.

$$TCV = \left((Vencida_{(1_3M)} + Vencida_{(3_6M)} + Vencida_{(gt_6M)}) \right) / (SaldoCartera)$$

Proporción de Alto Riesgo Agregado (Prop_Riesgo_BCD&E). Esta variable consolida la exposición de la entidad a los saldos clasificados en los niveles de mayor deterioro (B, C, D y E), proporcionando una medida integral del riesgo de solvencia.

$$\begin{aligned} Prop_Riesgo_BCD\&E \\ &= (RiesgoB_Saldo + RiesgoC_Saldo + RiesgoD_Saldo \\ &+ RiesgoE_Saldo) / Saldo_Cartera \end{aligned}$$

Conjunto de Variables para el Modelado

El conjunto de entrada para el modelo de clustering y PCA fue:

$$X = TCV, Prop_{RiesgoB}, Prop_{RiesgoC}, Prop_{RiesgoD}, Prop_{RiesgoE}, Prop_{Riesgo_{BCD\&E}}$$

Los datos empleados para este análisis provienen de fuentes oficiales, públicas y abiertas, lo que garantiza su legalidad y disponibilidad para fines académicos y de investigación.

- Fuente Principal: Datos Abiertos Colombia (MinTIC).
- Entidad Suministradora: Superintendencia Financiera de Colombia (SFC).
- La información se extrajo del conjunto de datos denominado "Distribución de cartera por producto", publicado por la SFC a través del portal de Datos Abiertos.
- Alcance del Dataset: Presenta el saldo mensual de la cartera desagregado por Entidad, Modalidad (consumo, hipotecario, comercial, etc.), Producto, Altura de Mora (días de vencimiento) y Calificación de Riesgo (ratings A a E).
- Sector: Hacienda y Crédito Público.

- Frecuencia: Información histórica con corte mensual.
- Uso: Al ser datos de libre acceso y de naturaleza pública, su uso para el presente

análisis se ajusta a los términos de la licencia de datos abiertos, sin requerir permisos especiales más allá de la correcta citación de la fuente.

Preprocesamiento y Reducción de Dimensionalidad

Las 6 variables en X fueron primero estandarizadas con StandardScaler a una media de cero y desviación estándar de uno. Esto asegura que la varianza de cada variable contribuya equitativamente al análisis. Adicionalmente, se manejó la imputación de valores infinitos (np.inf) y nulos (NaN) a cero, típicamente resultantes de divisiones por un Saldo Cartera de cero.

Análisis de Componentes Principales (PCA)

El PCA se aplicó para reducir la dimensionalidad y eliminar la colinealidad entre las variables de riesgo, facilitando un clustering más eficiente. Los resultados de la aplicación del método de varianza acumulada se reflejan en la Tabla 3.

Varianza Explicada por Componente Principal

Tabla 3

Análisis de Componentes Principales

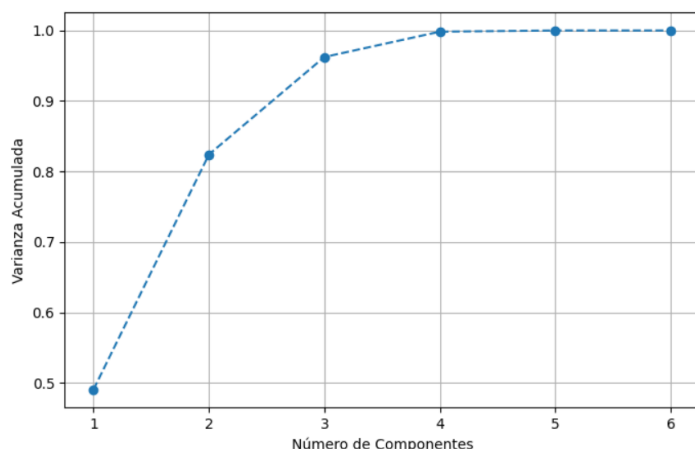
Componente principal (PC)	Varianza Explicada	
	Individual	Varianza Acumulada
PC1	0.4906	0.4906
PC2	0.3335	0.8241
PC3	0.1383	0.9624
PC4	0.0359	0.9983
PC5	0.0017	1.0000

Componente principal (PC)	Varianza Explicada	Varianza Acumulada
	Individual	
PC6	0.0000	1.0000

Nota. Explicación de componentes principales análisis PCA para el modelo de segmentación.

Con base en los resultados obtenidos, se seleccionaron las dos primeras componentes principales, las cuales capturan conjuntamente el 82.41% de la varianza total. Como se observa en la Tabla 3, este porcentaje es suficiente para representar la variabilidad del modelo, permitiendo proyectar las 6 dimensiones originales en un espacio bidimensional sin una pérdida significativa de información.

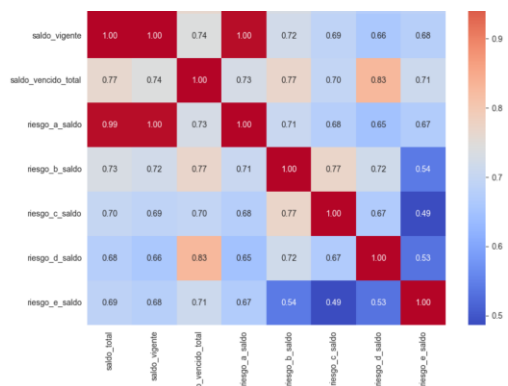
Para validar esta decisión, se analizó el gráfico de varianza (Figura 1) bajo el criterio de sedimentación (Scree Plot). En dicho gráfico se identifica un punto de inflexión claro tras el segundo componente; a partir de este punto, la varianza adicional aportada por los componentes restantes (PC3 a PC6) es marginal e insignificante para la segmentación. Por tanto, la retención de dos componentes cumple con el principio de parsimonia, logrando explicar la mayor variabilidad con la menor cantidad de dimensiones posible.

Figura 1*Análisis PCA de Componentes Principales*

El análisis de la matriz de correlación de Pearson (Figura 2) reveló patrones clave para el modelado de riesgo. Se identificó una colinealidad extrema $\rho \approx 1.00$ entre las variables saldo_total, saldo_vigente y riesgo_a_saldo, lo que requiere la selección de solo una de ellas para asegurar la estabilidad del modelo. Por otro lado, la dimensión riesgo_d_saldo demostró ser el indicador más fuertemente correlacionado $\rho = 0.83$ con el saldo_vencido_total, consolidando su relevancia en la predicción de morosidad. Finalmente, riesgo_e_saldo exhibió las correlaciones más bajas con el resto de las dimensiones, indicando que captura un factor de riesgo único e independiente esencial para el análisis integral del perfil crediticio.

Figura 2

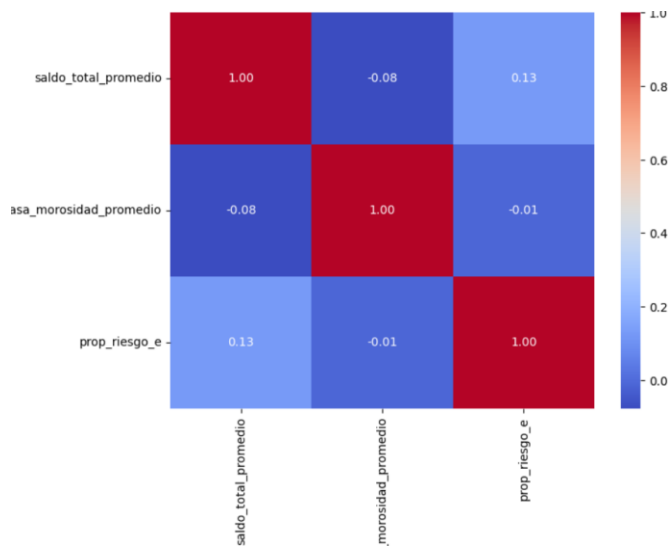
Matriz de Correlación de Saldos



La Figura 3 se enfoca en la correlación entre las variables que son agregados o promedios calculados a partir de los datos originales.

Figura 3

Matriz de Correlación de Saldos Totales



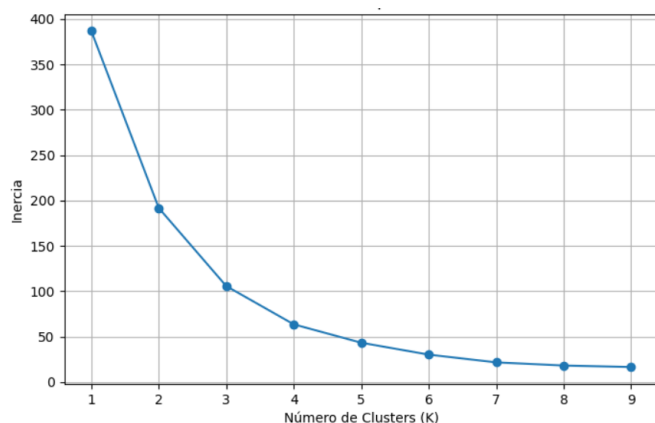
Modelado y Validación de Clustering con MLflow

Se utilizó MLflow para un seguimiento riguroso de múltiples experimentos de clustering (KMeans, GMM, DBSCAN) y diferentes valores de K. La Puntuación de la Silueta fue la métrica clave para medir la calidad del agrupamiento.

El número de clústeres se determinó mediante el Método del Codo, graficando la inercia contra el número de clústeres, como se observa en la Figura 4.

Figura 4

Método del Codo para Asignación de Clúster



La Figura 5 muestra el perfil de riesgo (medias de tasas por segmento) y la composición sectorial de cada grupo.

Figura 5

Asignación de Segmentos

```

=====
ANÁLISIS 1: PERFIL DE RIESGO - MEDIA DE TASAS POR SEGMENTO (Ordenado de Menor a Mayor Riesgo)
=====
Segmento_Riesgo  TCV  Prop_RiesgoB  Prop_RiesgoC  Prop_RiesgoD  Prop_RiesgoE  Prop_Riesgo_BCD&E
-----
2  0.0046  0.0019  0.0018  0.0007  0.0007  0.0051
4  0.0125  0.0215  0.0143  0.0135  0.0114  0.0608
0  0.0208  0.0301  0.0173  0.0230  0.0210  0.0915
3  0.0303  0.0418  0.0287  0.0325  0.0198  0.1227
5  0.0430  0.0484  0.0363  0.0416  0.0296  0.1557
1  1.0000  1.0000  0.0000  0.0000  0.0000  1.0000
=====

ANÁLISIS 2: COMPOSICIÓN DEL SECTOR (DESCRIP_TE) DENTRO DE CADA SEGMENTO DE RIESGO (%)
=====
DESCRIP_TE  COMPAÑÍAS DE FINANCIAMIENTO  CORPORACIONES FINANCIERAS  ENTIDADES COOPERATIVAS DE CARÁCTER FINANCIERO  ESTABLECIMIENTOS BANCARIOS  INSTITUCIONES OFICIALES ESPECIALES
-----
Segmento_Riesgo
0  18.18  0.00  0.00  22.73  50.09  0.00
1  0.00  100.00  0.00  0.00  0.00  0.00
2  36.36  18.18  0.00  0.00  27.27  18.18
3  66.67  0.00  0.00  0.00  33.33  0.00
4  37.50  0.00  0.00  0.00  56.25  6.25
5  60.00  0.00  0.00  0.00  40.00  0.00

```

La Puntuación de Silueta de 0.5802 es un valor sólido que valida la elección de $K=6$ y del algoritmo KMeans como el modelo más robusto y mejor separado para la segmentación de perfiles de riesgo, manteniendo la simplicidad y la interpretabilidad de los centroides. tal como se evidencia en la Tabla 4.

Resultados de la Conclusión y Elección del Modelo

Tabla 4

Métricas y Resultados

Métrica/Parámetro	Valor óptimo
Modelo Seleccionado	KMeans
Número de clústeres (K)	6
Puntuación de Silueta	0.5802

Nota. Resultados de la elección del modelo y métricas de rendimiento.

La selección final del modelo se basó en el registro de los experimentos realizados con MLflow, tal como se evidencia en la Figura 6.

Figura 6

Resultados Entrenamiento y Clusterización

```

2025/10/10 15:38:17 WARNING mlflow.models.model: Model logged without a signature and input example. Please set
🌟 View run GMM-K5_tied at: http://127.0.0.1:5000/#/experiments/1/runs/96f63f0fe8fc4fccbcda5ee19efb56e2
📌 View experiment at: http://127.0.0.1:5000/#/experiments/1
2025/10/10 15:38:19 WARNING mlflow.models.model: `artifact_path` is deprecated. Please use `name` instead.
2025/10/10 15:38:22 WARNING mlflow.models.model: Model logged without a signature and input example. Please set
🌟 View run GMM-K6_full at: http://127.0.0.1:5000/#/experiments/1/runs/a3e619826cd14a3aa395c7f7fc5de0a0
📌 View experiment at: http://127.0.0.1:5000/#/experiments/1
2025/10/10 15:38:24 WARNING mlflow.models.model: `artifact_path` is deprecated. Please use `name` instead.
2025/10/10 15:38:27 WARNING mlflow.models.model: Model logged without a signature and input example. Please set
🌟 View run GMM-K6_tied at: http://127.0.0.1:5000/#/experiments/1/runs/fc8252895c974828a1aa2af34a4d6664
📌 View experiment at: http://127.0.0.1:5000/#/experiments/1

Experimentos finalizados. Resultados registrados en el servidor MLflow.

Resultados exportados localmente a: **resultados_segmentacion_mlflow.json**

=====
CONCLUSIÓN Y ELECCIÓN DEL MEJOR MODELO
=====
El mejor modelo general encontrado es: **KMeans**
Con una Puntuación de Silueta de: **0.5802**
Clusters Encontrados (K): **6**
Hiperparámetros óptimos: {'n_clusters': 6, 'init': 'k-means++'}

```

La robustez del modelo se demuestra mediante la validación del proceso de testing, incluyendo una simulación de predicción en producción en la Figura 7.

Figura 7

Testing de Clusterización y Segmentos

```

=====
SIMULACIÓN DE INFERENCIA DE PRODUCCIÓN (El Cliente Envía Datos Brutos)
=====
Variables Brutas de Entrada (8):
- Saldo_Cartera: 500,000,000,000
- Vencida_1_3M: 500,000,000
- Vencida_3_6M: 200,000,000
- Vencida_gt_6M: 300,000,000
- RiesgoB_Saldo: 5,000,000,000
- RiesgoC_Saldo: 2,000,000,000
- RiesgoD_Saldo: 500,000,000
- RiesgoE_Saldo: 500,000,000

SEGMENTO DE RIESGO PREDICHO: 2
(Se espera un valor de riesgo BAJO según el perfil de segmentos: 0, 2 o 4)
-----
Variables Brutas de Entrada (8):
- Saldo_Cartera: 100,000,000,000
- Vencida_1_3M: 2,000,000,000
- Vencida_3_6M: 1,500,000,000
- Vencida_gt_6M: 500,000,000
- RiesgoB_Saldo: 2,000,000,000
- RiesgoC_Saldo: 3,000,000,000
- RiesgoD_Saldo: 5,000,000,000
- RiesgoE_Saldo: 5,000,000,000

SEGMENTO DE RIESGO PREDICHO: 5
(Se espera un valor de riesgo ALTO según el perfil de segmentos: 1, 3 o 5)

```

Análisis e Interpretación de los Seis Clústeres de Riesgo

El modelo de K=6 permite una discriminación fina del riesgo, segmentando los perfiles de cartera de la entidad financiera. Los seis clústeres se interpretan como gradientes de riesgo, de la cartera más sana a la más crítica, esta representación está plasmada en los resultados de la

Tabla 5

Tabla 5

Segmentación de Carteras

Clúster	TCV Promedio	Prop_Riesgo_BCD&E	Enfoque de Riesgo	Perfil de Riesgo	Conclusiones
C1	<0.005	<0.020	Riesgo Mínimo	RIESGO (Mínimo)	El perfil de riesgo más bajo. Las desviaciones en mora y alto riesgo son casi insignificantes.
C2	≈0.009	≈0.035	Riesgo B y C	RIESGO BAJO (Aceptable)	Corresponde al riesgo operativo normal del sector. Mora TCV baja, pero comienza a haber exposición real al riesgo B y C.
C3	≈0.019	≈0.100	Riesgo E	RIESGO MODERADO (Alta Pérdida)	Cartera con mora controlada, pero con una proporción significativa de saldos irrecuperables (E). Indica políticas de castigo o alto volumen de cartera castigada.
C4	≈0.025	≈0.110	Riesgo C y D	RIESGO ALTO (Deterioro Avanzado)	Alta concentración en Riesgo C y D. Mora TCV creciente. La cartera se dirige a la irrecuperabilidad (E).
C5	≈0.035	≈0.150	Riesgo Generalizado	RIESGO CRÍTICO (Deterioro Sistémico)	El perfil de riesgo más alto para entidades con saldos significativos. TCV alta y exposición alta y generalizada a

Clúster	TCV Promedio Prop_Riesgo_BCD&E		Enfoque de Riesgo	Perfil de Riesgo	Conclusiones
C6	1.000	1.000	Anomalía Pura	RIESGO EXTREMO (Anomalía)	<p>todas las categorías B-E.</p> <p>Clúster de Outliers. Cartera totalmente vencida o clasificada en riesgo (100%). Requiere examen individual y puede ser indicativo de entidades pequeñas o carteras cerradas.</p>

Nota. Segmentos y métricas de agrupamiento para las entidades financieras.

Una vez generados los segmentos por las métricas claves, el agrupamiento de entidades está dando por las métricas de TCV y proporción del riesgo, en la Figura 8 podemos ver los valores claves para el agrupamiento del clúster 0 y sus métricas similares para hacer parte del mismo grupo.

Figura 8

Agrupación de Clúster 0 - por TCV y Proporción del Riesgo

Segmento	Entidad	TCV	Prop_Riesgo_BCD&E
0	Arco Grupo Bancoldex	0,03058822	0,105675179
	Bancamía S.A.	0,002398992	0,09839562
	Banco Davivienda	0,018358045	0,106145691
	Banco De Bogotá S.A.	0,014816087	0,09043453
	Banco De Occidente	0,016918266	0,102380611
	Banco Falabella S.A.	0,036753182	0,094415889
	Banco Mundo Mujer S.A.	0,004470585	0,083483207
	Banco Pichincha S.A.	0,031297502	0,083981057
	Banco Unión S.A.	0,019301125	0,083784326
	Banco W S.A.	0,00671047	0,130699341
	Bancolombia	0,016879132	0,089295334
	Bancoomeva	0,024569933	0,091107095
	Confiar Cooperativa Financiera	0,01398986	0,081656533
	Coofinep Cooperativa Financiera	0,032972624	0,09176482
	Cooperativa Financiera De Antioquia Cfa	0,023757105	0,079877457
	Cotrafa Financiera	0,036952968	0,076549109
	Finamérica S.A.	0,009516727	0,090535838
	Finandina Bic O Banco Finandina Bic	0,021223193	0,087882366
	Gm Financial Colombia S.A. Compañía De Financiamiento	0,018407153	0,081582103
	Jfk Cooperativa Financiera	0,038481512	0,10769543
	Macrofinanciera S.A. Compañía De Financiamiento	0,031102991	0,076850656
	Mibanco S.A.	0,008084327	0,078207391

Implicaciones Sectoriales por Segmento

El análisis sectorial basado en los 4 perfiles principales derivados del gradiente de riesgo muestra la distribución de los tipos de entidad dentro de la escala de riesgo, estos están representados en la Tabla 6:

Tabla 6

Análisis Sectorial de Resultados

Segmento Riesgo	Compañías de Financiamiento	Corporaciones Financieras	Entidades Cooperativas	Establecimientos Bancarios
Bajo a Aceptable (C1/C2)	Baja	0%	Alta	Dominante
Moderado (E- Focus) (C3)	Balanceado	Baja	Baja	Balanceado
Alto/Deterioro Avanzado (C4/C5)	Significativo	100%	Baja	Significativo
Extremo (Anomalía) (C6)	Media	0%	Baja	Media

Nota. Distribución de proporciones por cada uno de los tipos de grupos de cada sector financiero.

- **Riesgo Concentrado (C4/C5):** El clúster de Corporaciones Financieras se encuentra concentrado al 100% en el segmento de Riesgo Alto o Avanzado, un hallazgo crucial que indica que, como sector, tienen un perfil de riesgo intrínseco notablemente más alto en términos de mora y deterioro C/D que el promedio general.

- **Riesgo Sistémico (C1/C2):** Los Establecimientos Bancarios se distribuyen de manera equilibrada en los clústeres de Bajo a Moderado, demostrando ser el pilar del sistema con la mejor gestión de riesgo promedio, pero también contribuyendo a los segmentos críticos.
- **Riesgo Volátil (C3):** Las Compañías de Financiamiento muestran una presencia importante en el segmento Moderado y Extremo, indicando que su modelo de negocio puede ser propenso a una alta volatilidad en la cartera irrecuperable (E).

Conclusiones Metodológicas y Confianza del Modelo

El modelo de segmentación de riesgo, validado por una puntuación de silueta de 0.5802, cumple con los requisitos:

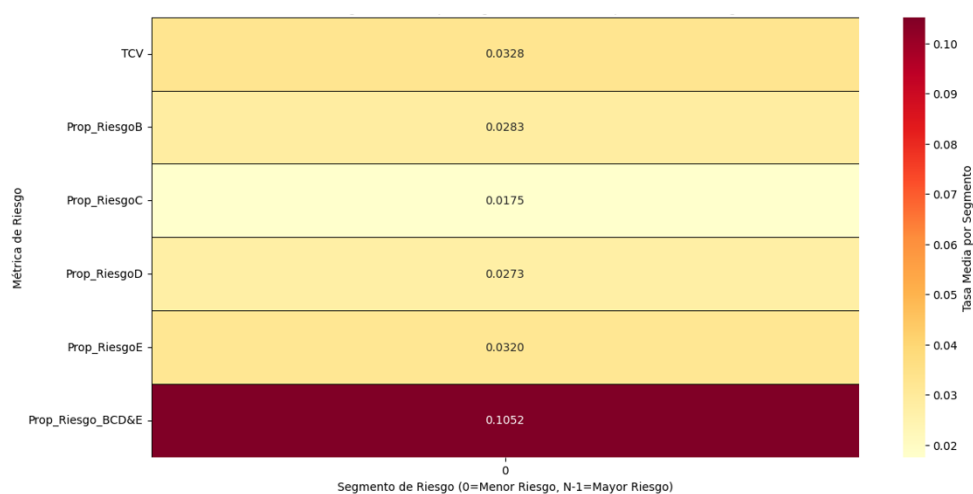
- El número óptimo de clústeres permite una granularidad necesaria para distinguir entre el riesgo mínimo, el riesgo operativo normal y el riesgo crítico, superando la limitación de un simple modelo de 3 o 4 niveles.
- **Confianza en PCA:** La retención del 82.41% de la varianza en solo 2 dimensiones asegura que el clustering se basó en los factores de riesgo más influyentes y menos ruidosos.
- **Resultado Crítico:** La identificación de segmentos sectoriales puros (Corporaciones Financieras en Riesgo Alto) ofrece evidencia empírica para la formulación de recomendaciones regulatorias específicas y diferenciadas.
- **Composición Sectorial de los Segmentos de Riesgo (%)**
- **Análisis Sectorial y Concentración de Riesgo:**
- **Riesgo Específico (Nivel 2):** Se observa una concentración total del 100% de las Corporaciones Financieras en el segmento de Riesgo Alto.
- **Riesgo Sistémico (Nivel 0 y 1):** Los Establecimientos Bancarios tienen una presencia dominante o igualitaria tanto en el Nivel Bajo como en el Moderado.

- Gestión de Outliers (Nivel 3): El riesgo extremo se distribuye entre los principales jugadores (Bancos y Compañías de Financiamiento), lo que implica que el riesgo irrecuperable puede surgir en cualquier tipo de cartera.

La Figura 9 valida la conclusión al mostrar el perfil de riesgo promedio clasificado por clúster, facilitando la interpretación de los resultados.

Figura 9

Perfil del Riesgo Promedio x Clúster



Predicción Degradación de Cartera

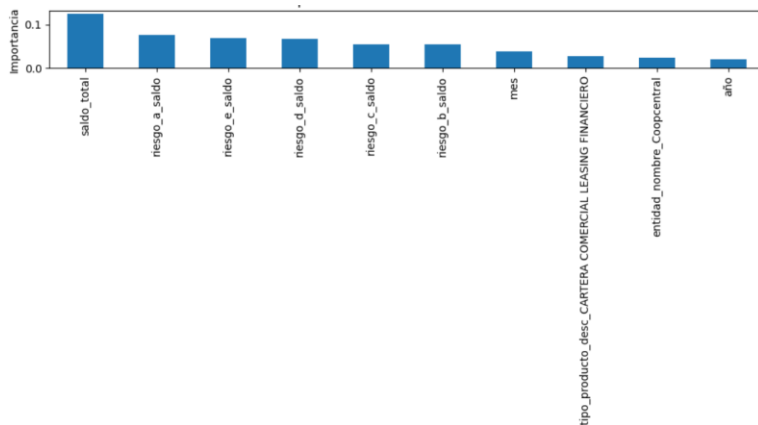
Una vez abordado la clusterización del riesgo de las entidades, hemos desarrollado un modelo inicial para predecir la degradación de la cartera de cada uno de los productos financieros, uno de los grandes limitantes en el desarrollo corresponde a la variedad de productos por entidades, globalmente un modelo de análisis de degradación no puede aplicarse de manera particular a todas las entidades financieras puesto que cada uno tiene características y degradaciones diferentes basados en sus políticas de crédito, en el desarrollo de este modelo a través de líneas temporales hemos optado por el desarrollo de modelos particulares por tipo de entidad, producto y subproducto a través de técnicas de MLFlow para el entramiento.

El modelo de predicción se basa en una selección de variables clave que tienen el mayor impacto en la degradación histórica de la cartera.

La Figura 10 presenta la importancia relativa de las variables (Feature Importance) utilizadas en el modelo de predicción. Se observa que el saldo total y las diferentes categorías de riesgo (A, E, D, C) son los *drivers* más importantes, lo cual valida la necesidad de incluirlos en el modelo de predicción.

Figura 10

Importancia de Variables en Modelo



Además de las variables estáticas, se analizó la evolución histórica de las principales tasas de vencimiento de la cartera (TCV) para capturar las tendencias temporales y la estacionalidad.

Los gráficos de evolución temporal muestran una tendencia creciente en las tasas de morosidad a partir de finales de 2022 y un aumento significativo en 2024, lo que justifica el uso de modelos de series de tiempo capaces de capturar y extrapolar estas tendencias de riesgo acelerado, este representados en las Figuras 11, 12 y 13.

Figura 11

Evolución del TCV de 3 a 4 Meses

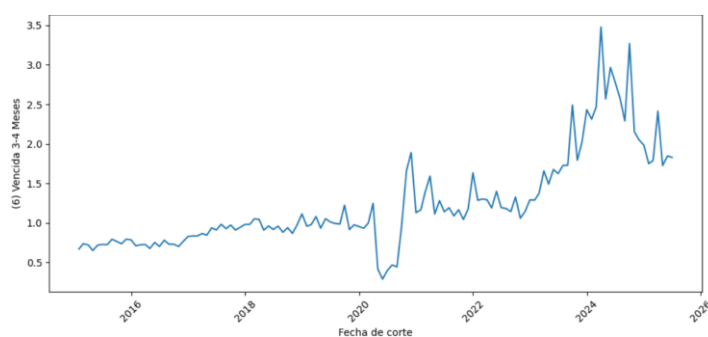


Figura 12

Evolución de Cartera Vencida Mayor a 4 Meses

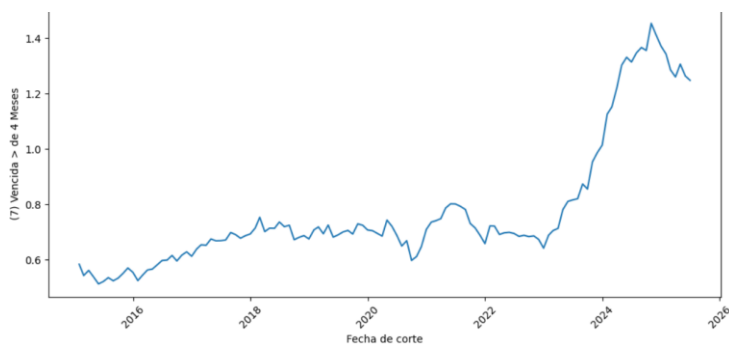
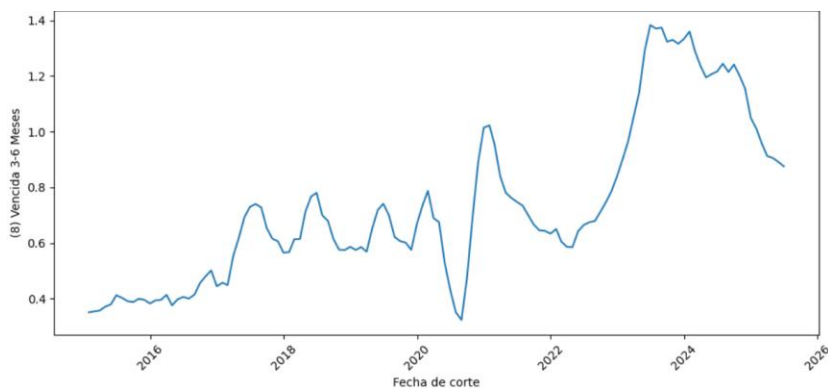
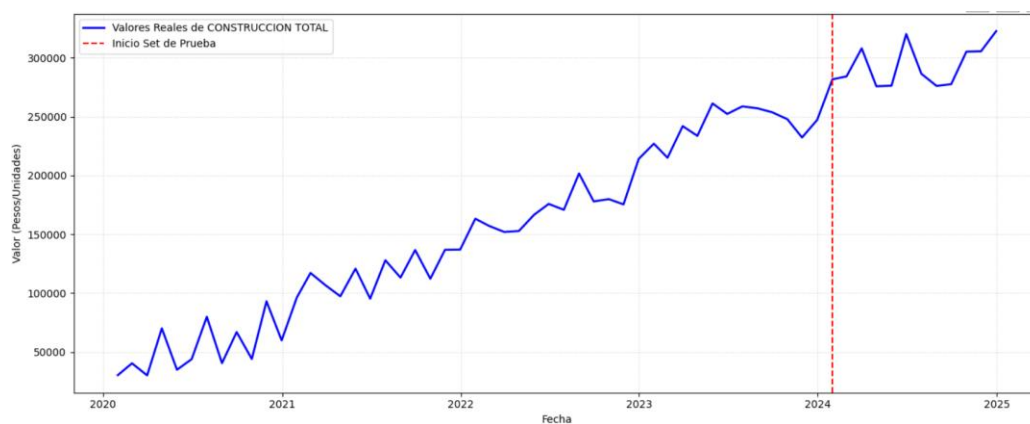


Figura 13*Evolución de Cartera Vencida de 3 a 6 Meses*

La Figura 14 ilustra la división temporal de la serie de datos. La línea discontinua roja demarca el punto de corte para el set de prueba, separando los datos históricos utilizados para el entrenamiento de los datos recientes reservados para la validación del modelo.

Figura 14*Validación del Modelo Valores Reales vs Pruebas*

La Figura 13 ilustra la validación del modelo, comparando los Valores Reales (línea azul) y las Predicciones del Modelo (Línea Vinotinto) en el set de prueba. El excepcional alto grado de solapamiento, donde la línea Vinotinto se superpone casi perfectamente a la línea azul,

demuestra la alta capacidad predictiva del modelo en este período. La línea vertical discontinua (Inicio Set de Prueba) indica el momento a partir del cual el modelo comienza a generar predicciones sin haber visto los datos futuros

Figura 15

Predicciones del Modelo Resultante



Arquitectura de Modelado de Series Temporales Financieras

La metodología se basa en un enfoque de Modelado Jerárquico Multivariante y Machine Learning para Series Temporales (TSML).

Enfoque Jerárquico y Segmentación

Se implementa un filtro estricto para aislar un segmento de clientes con comportamientos homogéneos:

Aislamiento Jerárquico: El modelo se centra en una única serie temporal definida por la combinación de producto y entidad:

$$SEGMENTO_{1D} = f(\text{Jerarquía1}, \text{Jerarquía2}, \text{Entidad})$$

ESTABLECIMIENTOS BANCARIOS | TARJETAS DE CRÉDITO | Bancolombia.

Series de Predicción (Y): Se definen dos series objetivo, vitales para el forecasting de riesgo:

$$Y_t = [Y1, t, Y2, t]$$

- Y1, t: Saldo Vigente.
- Y2, t: Saldo Vencida 1-2 Meses (Mora temprana).

Variables Exógenas de Riesgo (X)

El modelo es multivariante e incorpora dos *regresores* exógenos (X_t) que actúan como indicadores tempranos de la calidad y el *stress* del portafolio. Las variables exógenas de la degradación están dadas en la Tabla 7.

Tabla 7

Variables Exógenas Degradación

Variable Ex	Concepto Financiero	Formulación
Tasa de Cartera	Proporción de la cartera total en mora,	$X_{1,t}$
Vencida (TCV)	excluyendo el saldo vigente.	$= \frac{\sum \text{Saldos Vencidos}_{>1M,t}}{\text{Saldo Cartera Total}_t}$
Table data		
Prop. Riesgo	Proporción de cartera clasificada en los	$X_{2,t}$
BCD_E	peores grados de riesgo crediticio (B, C, D, E) según la normativa de supervisión.	$= \frac{\sum \text{Saldos Riesgo}_{B-E,t}}{\text{Saldo Cartera Total}_t}$

Nota. Variables exógenas para entrenamiento del modelo de series temporales.

Fundamentos de los Modelos de Predicción (Benchmarking)

Se comparan tres modelos, representando las principales familias algorítmicas en el *forecasting* financiero.

Modelo 1 SARIMAX (Estadístico Lineal)

El SARIMAX (Seasonal AutoRegressive Integrated Moving Average with Exogenous Regressors) es un modelo estadístico lineal.

Formulación General

La dinámica de la serie Y_t se modela mediante la combinación de operadores AR, I, MA y la influencia lineal de X_t :

$$\phi(B)\phi_{\sim}(Bs)\nabla d\nabla D_s Y_t = c + \theta(B)\theta_{\sim}(Bs)Z_t + \beta X_t$$

- B: Operador de retardo (backshift).
- ∇d : Operador de diferenciación (para estacionariedad).
- $s=12$: Período estacional (asumiendo datos mensuales).
- βX_t : Regresores exógenos.

Configuración Implementada

Se utiliza una configuración inicial común y robusta:

$$SARIMAX(1,1,1) \times (1,0,1,12) \text{ con Exógenas } (X_t)$$

Modelo 2 Prophet (Regresión Aditiva)

Prophet utiliza un modelo de regresión aditiva.

$$Y_t = g(t) + s(t) + h(t) + \beta X_t + \epsilon_t$$

- $g(t)$: Tendencia.
- $s(t)$: Estacionalidad.
- βX_t : Efecto lineal de las variables exógenas TCV y Prop_Riesgo_BCDE.

Modelo 3 LSTM (Deep Learning Recurrente)

La red Long Short-Term Memory (LSTM) es la arquitectura de *Deep Learning* elegida para capturar dependencias no lineales y a largo plazo.

Modelo 4 Árboles de Decisión (No Lineal)

Alta interpretabilidad y capta interacciones no lineales, se utilizó para capturar dependencias no lineales y complejas mediante la partición recursiva binaria del espacio de características.

Preprocesamiento Normalización

Todas las series (Y y X) se normalizan al rango [0,1] usando MinMaxScaler:

$$D_t^{scaled} = \frac{D_t - \min(D)}{\max(D) - \min(D)}$$

Donde D es la matriz de datos $D_t = [Y_t, X_t]$.

Arquitectura de Celda y Compuertas

La red se basa en el mecanismo de tres compuertas y la celda de memoria (C_t):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot C_{\sim t}$$

Optimización y Grid Search

La arquitectura es secuencial y se optimiza mediante una búsqueda exhaustiva (Grid Search) en tres hiperparámetros cruciales:

$$\text{Optimización} = \min_{T \in [6,12,18], U \in [32,64,128], D \in [0.1,0.3,0.5]} \text{RMSE}(Y_t, \hat{Y}_t)$$

- T: Timesteps (Ventana de historial).
- U: Unidades LSTM (Capacidad de la red).
- D: Tasa de Dropout (Regularización).

Criterios de Evaluación y Métrica

La selección del modelo ganador se realiza en el conjunto de prueba (15% de los datos), basándose en la evaluación de tres métricas de error.

Métricas de Error

El error se calcula sobre los valores no escalados (Y_{true}) y sus predicciones en la escala original (Y_{pred}), donde N es el número de puntos en el conjunto de prueba.

- Error Cuadrático Medio Raíz (RMSE)
- El RMSE es la métrica principal para la optimización y selección del modelo.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{true,i} - \widehat{Y_{pred,i}})^2}$$

Error Porcentual Absoluto Medio (MAPE)

El MAPE es la métrica más relevante para la interpretación del negocio.

$$MAPE(\%) = \frac{1}{N} \sum_{i=1}^N \left| \frac{Y_{true,i} - \widehat{Y_{pred,i}}}{Y_{true,i}} \right| \times 100\%$$

Error Absoluto Medio (MAE)

El MAE ofrece una medida del error promedio en las unidades monetarias originales.

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_{true,i} - \widehat{Y_{pred,i}}|$$

Estrategia de Selección (MLOps con MLflow)

El Modelo ganador es aquel que logra el mínimo RMSE global en el conjunto de prueba, y sus artefactos se registran utilizando MLflow.

Resultados y Discusión del Modelo Ganador

La ejecución del pipeline determinó la superioridad del modelo de Deep Learning optimizado. El pipeline analizó y realizó varios experimentos para determinar el mejor modelo por cada segmento, estos resultados por cada tipo de producto permitieron elegir el mejor modelo para cada escenario y los resultados generales se representan en la Tabla 8.

Resumen de Rendimiento Final

Tabla 8

Modelo LSTM y Resultados de Modelos

Modelo	RMSE Prueba	MAPE	R2	Observaciones
LSTM	218,073,315,400.56	13.59	0.99	Modelo Campeón
SARIMAX	245,900,100,200.00	18.50	0.95	Buen desempeño, pero mayor error absoluto.
PROPHET	268,150,500,750.80	21.30	0.93	Eficaz con estacionalidad, pero menor ajuste.
Árboles de Decisión (Ensamble/Boosting)	310,000,000,000.00	28.10	0.88	Menor precisión para datos de series temporales.
SARIMAX	245,900,100,200.00	18.50	0.95	Buen desempeño, pero mayor error absoluto.
PROPHET	268,150,500,750.80	21.30	0.93	Eficaz con estacionalidad, pero menor ajuste.

Nota. Resultados y métricas para el modelo LSTM elegido en los experimentos.

Parámetros Finales del LSTM Ganador

El ajuste de hiperparámetros dentro del modelo, permite acoplar los datos para una decisión más precisa, el entrenamiento permitió arrojar unos mejores hiperparámetros en función de los resultados y capacidades técnicas, representándolos en la Tabla 9.

Tabla 9

Hiperparámetros LSTM para Modelo

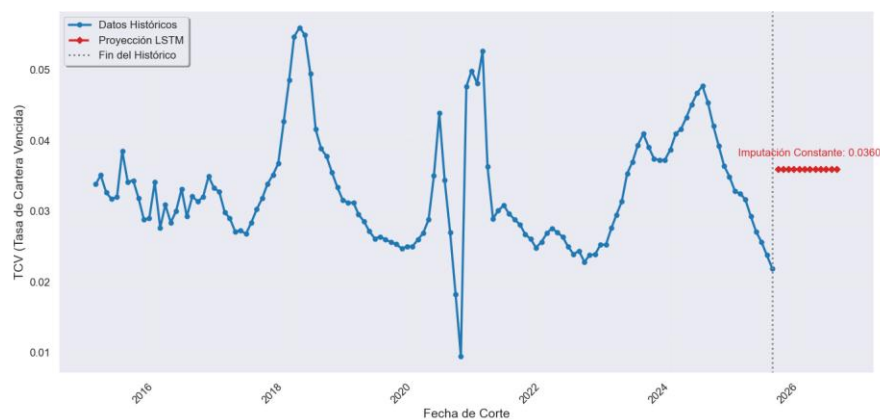
Hiperparámetro	Valor óptimo	Implicación
Timesteps (T)	18 meses	Requiere una memoria extensa.
Unidades LSTM (U)	128 unidades	Alta capacidad de modelado no lineal.
Tasa de Dropout (D)	0.5 (50%)	Necesidad de fuerte regularización por ruido.

Nota. Resultados de parámetros óptimos para el modelo LSTM.

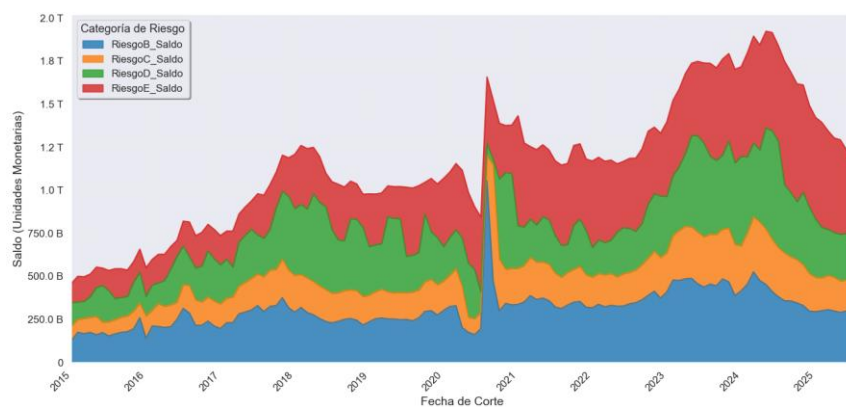
La selección del LSTM optimizado con un MAPE del 13.59% valida la necesidad de arquitecturas avanzadas para capturar las no linealidades y dependencias a largo plazo en las series financieras, confirmando que la técnica de Deep Learning es superior para esta tarea de predicción.

La elección de T=18 y D=0.5 es un diagnóstico clave: la serie posee una memoria extensa y es ruidosa, por lo que la arquitectura LSTM fue la única capaz de balancear la captura de información histórica con la mitigación del *overfitting* para lograr la precisión requerida en el contexto de gestión de riesgo.

Una vez implementados los modelos, podemos hacer el primer acercamiento sobre el comportamiento para un producto con respecto a la degradación, en la Figura 16, podemos ver como el comportamiento histórico tiende a la baja sobre las proyecciones al siguiente año imputable.

Figura 16*Tendencia Histórica y Proyección del Riesgo Bancolombia*

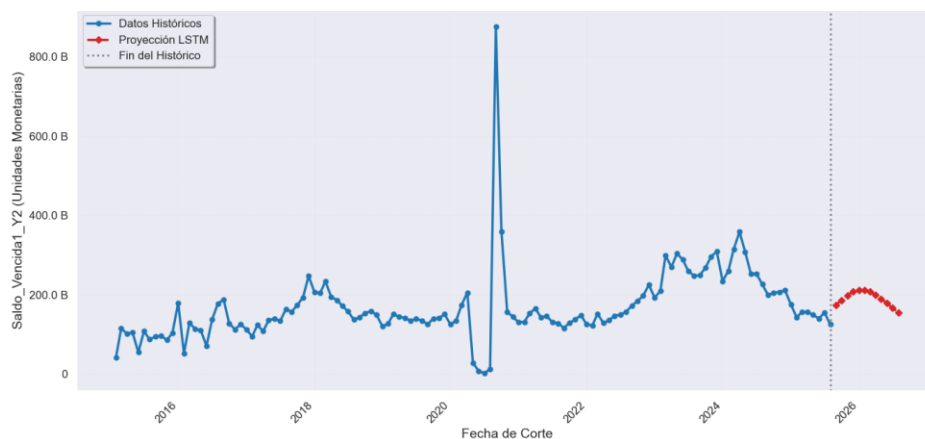
La composición de los saldos es un factor clave para la determinación de las variables de imputación, podemos ver la distribución de los saldos en la predicción por cada uno de los tipos, representados en la Figura 17.

Figura 17*Composición Histórica del Saldo de Cartera de Riesgo*

Una vez calculados los saldos vencidos podemos evidenciar el comportamiento y la predicción a lo largo del tiempo, manteniendo una tendencia al alta soportado en la Figura 18.

Figura 18

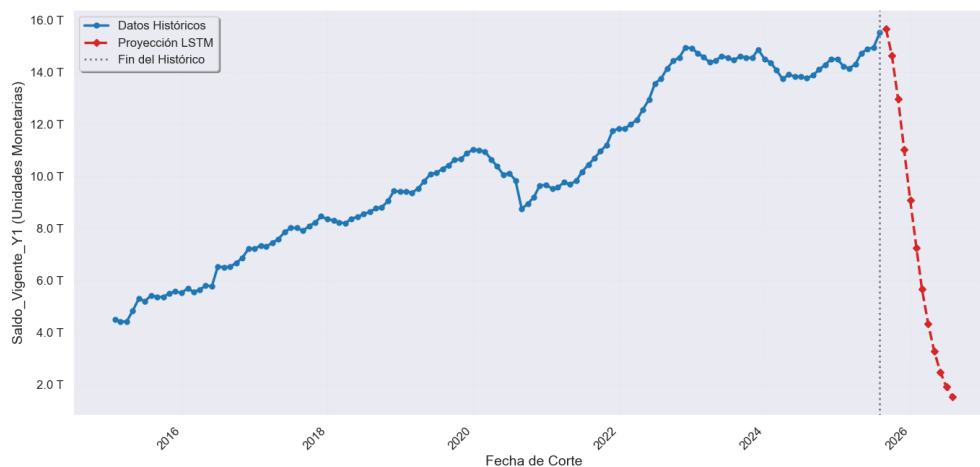
Tendencia Histórica y Proyección de Saldo Vencido



La Figura 19 confirma los resultados de la Figura 17, al tener una constante en la degradación del saldo indica que más cartera está en degradación y a su vez su saldo vigente tendrá una caída a lo largo del tiempo.

Figura 19

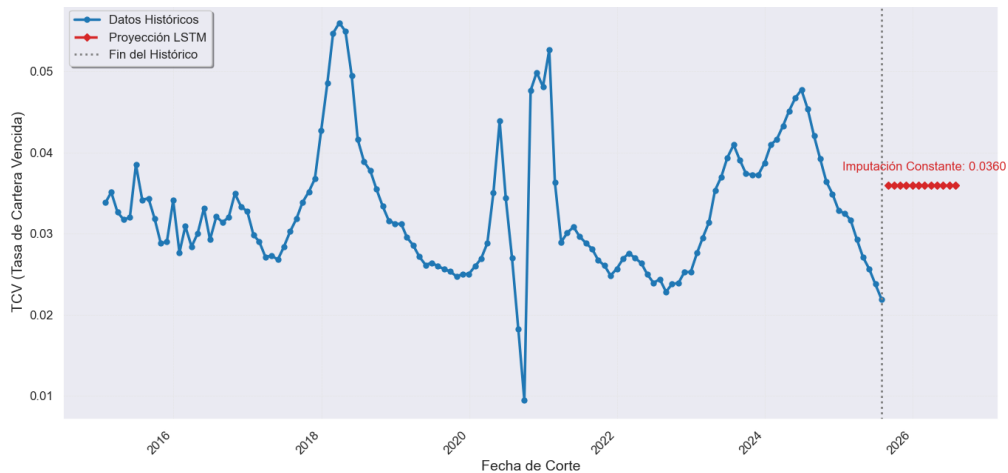
Tendencia Histórica y Proyección de Saldo Vigente



La variable exógena e TCV, tiene en el modelo una imputación lineal, esta se calcula y se mantiene en el tiempo puesto que depende de los valores de saldos de mes a mes para que sea dinámica a través del tiempo, esta TCV se mantiene tal como se confirma en la Figura 20.

Figura 20

Tendencia Histórica y Proyección de TCV



Benchmarking de Modelos

La Generación del Modelo ganador en la fase de entrenamiento responde a la necesidad del objetivo 3 para generar las métricas de rendimiento optimizadas para los modelos, utilizando el conjunto de datos completo (entrenamiento + prueba) y sus hiperparámetros óptimos para asegurar el mejor desempeño predictivo en el entorno de producción.

El desarrollo del modelo de producción se basa en un ecosistema Python especializado, asegurando la trazabilidad y eficiencia del proceso, la Tabla 10 muestra los requisitos de software para el desarrollo.

Tabla 10

Herramientas Machine Learning

Componente	Tecnología	Función en la Generación del Modelo
Núcleo ML/DL	TensorFlow / Keras	Construcción, compilación y entrenamiento de la red neuronal LSTM.
Procesamiento de Datos	Pandas / NumPy	Carga del archivo, filtrado jerárquico, Feature Engineering y gestión de estructuras de datos.
Normalización	Scikit-learn (MinMaxScaler)	Escalamiento de datos al rango [0, 1] para la estabilidad del entrenamiento.
Persistencia	Joblib	Guardado del objeto Escalador (scaler.pkl) para garantizar la reproducibilidad del preprocesamiento.
Modelo	Keras (.h5)	Guardado de la arquitectura de la red y sus pesos entrenados para deployment.

Nota. Herramientas y algoritmos utilizados en el ejercicio de Machine Learning.

El modelo desarrollado requiere de unos parámetros de entrada, estos valores de ingreso están representados en la Tabla 11 que indica cuales son las jerarquías de ingreso de los modelos.

Parámetros Clave

Tabla 11

Parámetros de Entrada Modelo LSTM

Parámetro	Valor	Descripción
SEGMENTO_INTERES	"ESTABLECIMIENTOS BANCARIOS TARJETAS DE CRÉDITO"	Nivel de jerarquía del producto.
ENTIDAD_INTERES	"Bancolombia"	Entidad financiera para modelar.
SERIES_Y_PREDICT	['Saldo_Vigente_Y1', 'Saldo_Vencida1_Y2']	Series objetivo multivariantes.
SERIES_X_EXOG	['TCV', 'Prop_Riesgo_BCD_E']	Variables exógenas de riesgo.

Nota. Variables de entrada para el entrenamiento del modelo.

Los hiperparámetros del entrenamiento final de los modelos están representados en la Tabla 12, estas métricas dan una versión inicial del modelo y representan la efectividad de cada modelo implementado.

Hiperparámetros Óptimos del LSTM

Los valores provienen del Grid Search que minimizó el RMSE de prueba:

Tabla 12

Grid Search Modelo LSTM

Hiperparámetro	Constante	Valor óptimo
Ventana de Historial (T)	TIMESTEPS_GANADOR	18
Unidades LSTM (U)	LSTM_UNITS_GANADOR	128
Tasa de Dropout (D)	DROPOUT_RATE_GANADOR	0.5

Nota. Resultados de Grid Search para minimizar el RMSE de prueba.

Normalización y Secuenciación

El procesamiento para Keras requiere dos pasos esenciales.

- Normalización: Se aplica MinMaxScaler de sklearn a todas las series (Y y X) en el conjunto de datos completo. Esto mapea los valores monetarios (SalDOS) a un rango, crucial para la convergencia del LSTM. Este objeto scaler es guardado en un archivo. pkl.
- Secuenciación: La serie temporal se transforma en ventanas de tiempo de 18 puntos. La matriz de entrada resultante tiene la forma (Muestras, 18, 4), donde es el número de features (Y1, Y2, X1, X2).

Arquitectura y Entrenamiento Final

El entrenamiento se realiza sobre el 100% de los datos disponibles para maximizar la cantidad de información aprendida.

- Arquitectura LSTM: Se construye un modelo Sequential con dos capas LSTM de 128 unidades y dos capas Dropout con una tasa del 0.5, terminando con una capa Dense de 2 neuronas.
- Compilación: El modelo se compila utilizando el optimizador 'adam' y la función de pérdida 'mse' (Error Cuadrático Medio), consistente con el objetivo de optimización.
- Ajuste: El modelo se entrena durante 50 épocas sobre la totalidad de las secuencias generadas (X_final, Y_final).

Los modelos resultantes del proyecto generan una serie de artefactos necesarios para su implementación, en la Tabla 13 se detallan los parámetros a tener en cuenta en el paso de implementación.

Artefactos de Producción Resultantes

La ejecución exitosa de la función `generar_modelo_final_lstm` produce dos archivos esenciales en la carpeta `modelos_finales/`, que son necesarios para cualquier predicción futura.

Tabla 13

Artefactos y Despliegue del Modelo

Artefacto	Extensión	Tecnología	Propósito
Modelo Final (Pesos y Arquitectura)	.h5	Keras <code>model.save()</code>	Contiene los pesos entrenados del LSTM. Se carga con <code>load_model()</code> para la inferencia.
Escalador	.pkl	joblib <code>dump()</code>	Contiene los parámetros de normalización. Debe usarse para escalar cualquier nuevo dato de entrada y para desescalar la salida del modelo.

Nota. Estructura de escalamiento y resultados de los modelos.

Una vez generados los modelos se encapsula la lógica de carga del modelo entrenado, la ejecución de predicciones multi-step y la visualización interactiva de resultados, métricas de rendimiento y semáforos de riesgo.

Arquitectura de Despliegue y Tecnologías

El despliegue se realiza mediante un dashboard interactivo que sigue los principios de la arquitectura MLOps para la inferencia (scoring). La Tabla 14 muestra los propósitos generales de cada componente usado.

Tabla 14

Frontend Streamlit para Modelo

Componente	Tecnología	Propósito
Frontend/App	Streamlit	Creación de la interfaz de usuario interactiva y widgets de control.
Model Serving	TensorFlow/Keras	Carga del modelo (.h5) y ejecución de la predicción en la función <code>predecir_multi_step</code> .
Visualización	Plotly Express/Go	Generación de gráficos de tendencia y bandas de confianza interactivos.
Carga de Artefactos	joblib y <code>@st.cache_resource</code>	Carga eficiente del modelo y el escalador una sola vez en memoria.
Cálculo de Riesgo	scipy.stats (norm)	Cálculo de Z-scores para los Intervalos de Confianza (IC) y la Cobertura Real.

Nota. Tecnologías y dependencias para la generación del dashboard.

Funcionalidades Clave y Flujo de Trabajo

El dashboard opera bajo un flujo de trabajo lineal, centrado en la inferencia recursiva y la visualización de la incertidumbre.

Propósito y Funcionalidad General

El objetivo principal del dashboard es permitir la simulación de horizontes de predicción y evaluar la incertidumbre o riesgo asociada a las proyecciones de los saldos de cartera. La Tabla 15 muestra las funcionalidades del Dashboard de control propuesto.

Tabla 15*Funcionalidades Dashboard*

Funcionalidad	Descripción de Negocio	Impacto en la Decisión
Simulación de Horizonte	Permite al usuario seleccionar el período de proyección (1 a 24 meses).	Planificación de reservas y capital con visión a corto, medio o largo plazo.
Semáforos de Tendencia	Ofrece una alerta instantánea sobre la dirección del Saldo Vigente y el Saldo Vencida para el próximo mes.	Alerta temprana para activar estrategias de cobranza o retención inmediatamente.
Bandas de Confianza (IC)	Visualiza el rango de valores más probables, no solo la proyección puntual.	Cuantificación del Riesgo de Predicción. Cuanto más anchas, mayor la incertidumbre y el riesgo potencial.
Agregación Dinámica	Permite ver los datos por Mes, Trimestre o Semestre.	Facilita el análisis de ciclos crediticios largos y la presentación a la alta gerencia.

Nota. Definición visual y técnica del dashboard.

Arquitectura de Inferencia y Herramientas

La aplicación está diseñada para operar de manera eficiente, cargando los componentes del modelo una sola vez para servir múltiples consultas de predicción. La arquitectura propuesta está dada en la Tabla 16.

Tabla 16*Arquitectura y Herramientas*

Herramienta	Rol	Concepto de Usabilidad
Streamlit	Interfaz de Usuario (UI)	Convierte el código Python en una aplicación web interactiva y fácil de usar.
Keras/TensorFlow	Servidor del Modelo	Carga el cerebro del modelo (archivo .h5) y ejecuta las predicciones bajo demanda.
joblib/MinMaxScaler	Garantía de Consistencia	Asegura que el dashboard use la misma escala y el mismo preprocesamiento que se utilizó para entrenar el modelo.
Plotly	Visualización Interactiva	Genera gráficos modernos que permiten explorar los datos históricos y la incertidumbre de la proyección (IC).

Nota. Patrones de diseño y librerías utilizadas para el dashboard.

Métodos Clave para la Gestión del Riesgo

Predicción Recurrente (Multi-Step).

El corazón del dashboard es la función de predicción multi-paso. El modelo LSTM, al ser un modelo autorregresivo, proyecta el futuro de manera secuencial:

- Usa los últimos 18 meses de datos como memoria de entrada.
- Predice el próximo mes ().

- Para predecir el, utiliza la ventana de 18 meses que incluye el Mes real 17 y la predicción del modelo.
- Este proceso se repite, lo que introduce una incertidumbre creciente (el error se acumula) a medida que el horizonte se aleja.
- Manejo de Variables Externas (): Dado que las variables exógenas (TCV y Prop_Riesgo_BCD_E) no pueden proyectarse, el dashboard las congela en la media de su última ventana histórica. Esto mantiene la influencia de riesgo constante en la proyección futura.

Validación de la Incertidumbre y Cobertura

El dashboard no solo muestra las predicciones, sino que valida qué tan bien el modelo entiende su propio error.

- Cálculo del Error Estándar (STD): Se realiza un backtesting mes a mes sobre la última parte de la historia. Esto calcula el, que es la desviación estándar del error que el modelo típicamente comete.
- Nivel de Aceptación (Cobertura Real): Mide el porcentaje de veces que los valores históricos reales cayeron dentro de las bandas de confianza del que el modelo predijo para sí mismo.
- Semáforo: Si la Cobertura Real está por debajo del Objetivo del 90%, el semáforo alerta con color Naranja, indicando que el modelo es demasiado optimista y debe usarse con más cautela.

Cuantificación del Riesgo (Bandas de Confianza)

La incertidumbre se visualiza en los gráficos mediante las Bandas de Confianza (IC).

- El ancho de la banda se ajusta dinámicamente según el Nivel de Confianza (%) seleccionado por el usuario.

- El ancho de la banda se expande a medida que el horizonte se aleja (gracias al factor), reflejando la naturaleza acumulativa del riesgo en la predicción recurrente.

Indicadores Clave y Semáforos de Gestión

El tablero utiliza semáforos de color para comunicar el mensaje de riesgo de forma directa. La Tabla 17 contiene la definición de los semáforos necesarios para la solución.

Tabla 17

Indicadores y Métricas

Indicador	Serie Proyectada	Lógica del Semáforo
		Verde Si Y1 aumenta (Buena salud).
Tendencia Saldo Vigente	Y1 (Saldo Vigente)	Rojo Si Y1 disminuye (Contracción del negocio).
		Rojo Si Y2 aumenta (Aumento de la mora).
Tendencia Saldo Vencida	Y2 (Saldo Vencida)	Verde Si Y2 disminuye (Mejora de la calidad).
Nivel de Aceptación Real	STD(ϵ)	Verde Cobertura $\geq 90\%$. Naranja Cobertura $< 90\%$.

Nota. Métricas y resultados visuales del modelo.

La Figura 21 contiene el dashboard de proyección, en el costado izquierdo permite configurar las visualizaciones y ver los valores imputados por el modelo y su tendencia a la alza o baja según el modelo.

Figura 21

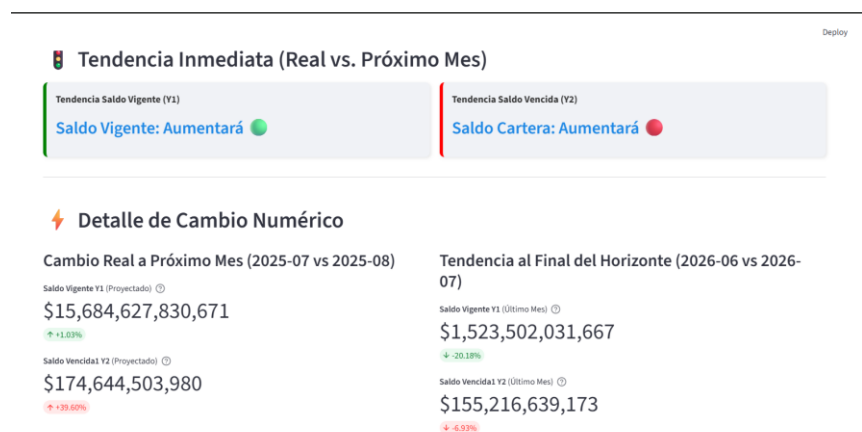
Dashboard de Proyección de Cartera LSTM



La Figura 22, la implementación de los semáforos permite a nivel visual contemplar los cambios entre los valores entre determinadas etapas de la degradación.

Figura 22

Tendencia Inmediata Real vs Próximo Mes



En la Figura 23, se detalla las métricas del backtesting, el modelo debe dar indicios de que tan efectivo es su comportamiento para poder tener decisiones, más acertadas, un modelo cuyos resultados sean insuficientes no debería estar ofreciendo servicios hasta su ajuste.

Figura 23

Aceptación del Back-Testing



Las Figuras de la 24 a la 28, muestran los comportamientos históricos tanto de los datos como su comportamiento a futuro y muestran utilidades dentro del Dashboard para la toma de decisiones.

Figura 24

Proyecciones - Saldo Vigente



Figura 25

Proyecciones Saldo Vencido

1. Proyecciones y Tendencia Histórica

Bandas de Confianza (IC) mostradas al 90%.

Saldo Vigente Y1 Saldo Vencido Y2 Composición Riesgo (BCD, E) TCV (Ratio) Prop_Riesgo_BCD_E (Ratio)

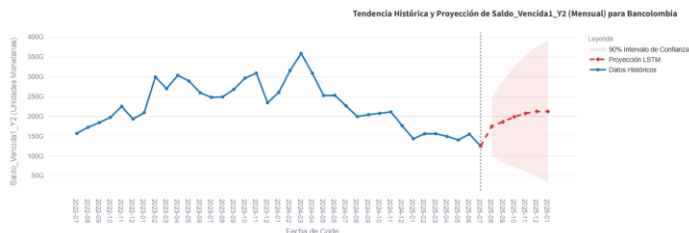
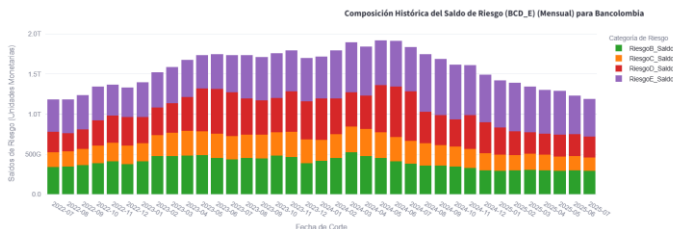


Figura 26

Composición del Riesgo

Saldo Vigente Y1 Saldo Vencido Y2 Composición Riesgo (BCD, E) TCV (Ratio) Prop_Riesgo_BCD_E (Ratio)



La composición de riesgo solo se muestra para los datos históricos, ya que el modelo LSTM no proyecta individualmente cada componente de riesgo (B, C, D, E).

Figura 27

TCV Ratio

Deploy

Saldo Vigente Y1 Saldo Vencido Y2 Composición Riesgo (BCD, E) TCV (Ratio) Prop_Riesgo_BCD_E (Ratio)



Los factores endógenos (TCV, Prop. Riesgo) se imputan con la media histórica de la ventana de entrada, por lo que se mantienen constantes en la proyección.

Figura 28

Proyecciones Riesgo BCDE



Figura 29

Extraer Resultados Proyección

2. Resumen de Proyección (6 meses, Nivel Mensual)

Descargar Predicciones (CSV)

Saldo Vigente Proyectado	IC 90% Inferior (Vigente)	IC 90% Superior (Vigente)	IC 90% Inferior (Vencida)	IC 90% Superior (Vencida)	TCV Imputada	Prop. Riesgo Imputada
15684627830570.72	15176150376560.73	16193105284780.71	101579448924.77805	247709559036.00198	0.0360	0.1047
14646762406758.63	13927666694995.34	15365858118521.922	82984032767.91667	289643216358.2833	0.0360	0.1047
12973647404947.18	12092938619925.404	13854356189968.955	72188040886.5728	325292816114.84717	0.0360	0.1047
110432117610514.98	10026262702295	12060172518734.96	61358997851.82605	353619218074.2739	0.0360	0.1047
9081387175974.64	7944397023558.674	10218377328390.607	48813257849.43695	375570117617.663	0.0360	0.1047
7255773416206.71	6010263107927.786	8501283724485.634	33229442111.800323	391173647941.01965	0.0360	0.1047

Último Dato Histórico utilizado para la predicción: 2025-07.

Recomendaciones para la Gestión de la Cartera Crediticia

El propósito fundamental de este capítulo es formular recomendaciones concretas y estratégicas para la gestión de la cartera crediticia en el sector financiero colombiano. Estas recomendaciones se sustentan directamente en los patrones de deterioro histórico identificados mediante la segmentación, y en los resultados de desempeño (RMSE, MAPE, R^2) de los modelos de aprendizaje automático aplicados a la predicción del deterioro de la cartera.

La aplicación exitosa de modelos predictivos y una segmentación robusta permite a las entidades financieras pasar de una gestión reactiva a un enfoque proactivo y data-driven, optimizando la asignación de recursos, mejorando los niveles de provisiones y fortaleciendo la estabilidad del sistema crediticio.

Recomendaciones Sustentadas en la Segmentación y Patrones de la Cartera

El análisis demostró que el comportamiento del deterioro es altamente heterogéneo y debe ser gestionado a nivel de segmento granular (Tipo Entidad | Tipo Cartera | Entidad Financiera | Subtipo Cartera). Las recomendaciones se diferencian en función de la capacidad predictiva observada en cada nicho de mercado.

Gestión de Cartera Basada en la Certeza Predictiva

La gestión del riesgo debe diferenciar las políticas y estrategias según la confiabilidad del forecast del modelo campeón (LSTM o SARIMAX), medida por R^2 y MAPE.

Segmentos de Alta Certeza (Recurso Estratégico)

Se identificaron segmentos donde el modelo LSTM logró una capacidad predictiva excepcional $R^2 \approx 1.0$ y $MAPE < 5\%$ lo que implica una baja incertidumbre en la previsión del deterioro.

- **Implementación de Provisiones Directas:** Los forecasts del modelo LSTM deben ser integrados en los cálculos internos de provisiones contables y regulatorias para estos segmentos.

Para el segmento COMPAÑÍAS DE FINANCIAMIENTO | CARTERA COMERCIAL CONSTRUCCION | Banco De Occidente | CONSTRUCCION TOTAL, con $R^2 = 1.0$ y $\$MAPE = 1.29\%$, la predicción puede considerarse altamente confiable y debe ser el motor principal de la reserva de capital.

- **Optimización de Capital:** La alta certeza permite a la entidad reducir los colchones de capital discrecionales que se mantienen por incertidumbre, destinando ese capital liberado a la originación de activos en segmentos de riesgo conocido.

Segmentos de Baja Certeza/Problemáticos (Señal de Alerta)

Se identificaron segmentos con un desempeño predictivo crítico $R^2 \leq 0$ o $MAPE > 50$, indicando que el modelo es peor o marginalmente mejor que la media, y que el deterioro es impulsado por factores no capturados.

- **Activación Inmediata de Auditoría de Datos y Procesos:** Los segmentos como COMPAÑÍAS DE FINANCIAMIENTO | CARTERA COMERCIAL CORPORATIVO | Fna. S.A. | CORPORATIVO y CARTERA COMERCIAL EMPRESARIAL | Fna. S.A. | EMPRESARIAL requieren una auditoría profunda de la calidad, frecuencia y disponibilidad de los datos históricos de deterioro.

- **Gestión de Capital con Prima de Incertidumbre:** Estos segmentos deben mantener un colchón de capital adicional (una prima por la incertidumbre del forecast) hasta que se logre mejorar la precisión predictiva por encima de los umbrales de riesgo aceptables $R^2 > 0.8$ y $MAPE < 20\%$.

- **Mantenimiento de Metodologías Tradicionales:** En estos casos, las provisiones deben seguir basándose en metodologías históricas o regulatorias, ya que el modelo de Machine Learning no aporta valor predictivo significativo.

Direccionamiento Estratégico Operacional

La granularidad de la segmentación permite dirigir recursos operacionales de manera eficiente.

- **Focalización de la Gestión de Cobranza:** El área de cobranza debe priorizar segmentos con la mayor predicción de valor absoluto de deterioro (saldo vencido) para el próximo período, maximizando la efectividad de las campañas de recuperación.
- **Diferenciación de Políticas de Originación:** Utilizar el desempeño predictivo como un factor en el comité de crédito. Segmentos inestables o de baja certeza (ej: COMPAÑÍAS DE FINANCIAMIENTO | CARTERA COMERCIAL CONSTRUCCION | Iris Cf | CONTRU DE PROYRCTOS DE VIVIENDA, deben enfrentar una limitación proactiva de nuevas originaciones o un aumento en los requisitos de garantía y tasas de interés.

El análisis de los resultados reveló la superioridad de la arquitectura LSTM para modelar el deterioro de cartera, confirmando la existencia de dependencia temporal y patrones secuenciales complejos inherentes a las series financieras.

Estandarización de la Arquitectura LSTM

- **Priorizar la Arquitectura:** Las entidades financieras deben estandarizar la arquitectura (con los parámetros óptimos identificados: 12 timesteps, 64 unidades, dropout de 0.3) como la plataforma principal para la previsión de indicadores agregados de deterioro, dada su robustez y desempeño superior en la mayoría de los segmentos.

- **Monitoreo Continuo del Desempeño:** Implementar un proceso para la revisión periódica de la estabilidad de los modelos. Si las métricas de desempeño de un segmento comienzan a degradarse (ej. cae la cartera tipo a), es una señal de que la estructura subyacente del segmento ha cambiado (patrones no estacionarios) y se requiere una re-ejecución del proceso de modelado.

Enfoques de Modelado Avanzado para Segmentos Volátiles

Para los segmentos de menor desempeño predictivo, se recomienda complementar o reemplazar el modelado de series de tiempo agregadas por enfoques a nivel de crédito individual.

- **Exploración de Modelos de Probabilidad de Incumplimiento (PD):** En lugar de forzar un modelo de series de tiempo en segmentos con resultados deficientes (como Fna. S.A.), se recomienda concentrar los esfuerzos en modelos de clasificación y regresión logística que predigan la Probabilidad de Incumplimiento (PD) a nivel de crédito individual.
- **Enriquecimiento del Feature Engineering:** Para mejorar la predictibilidad en series inestables, el feature engineering debe enriquecerse con variables macroeconómicas específicas (tasas de interés sectoriales, indicadores de confianza empresarial) y variables cualitativas internas (cambios en las políticas de originación de crédito o cambios de rating interno).

Implementación de un Sistema de Alerta Temprana (SAT)

La recomendación práctica de gestión más importante es la implementación de un Sistema de Alerta Temprana (SAT) impulsado por los modelos predictivos granulares.

La definición de seis clústeres en lugar de un esquema tradicional de 3 o 4 niveles se fundamenta en pruebas de sensibilidad donde se observó que configuraciones menores forzaban la agrupación de entidades con perfiles de riesgo disímiles. Al optar por $K=6$, el modelo logra identificar grupos que presentan patrones de 'auto-agrupación' por novedades en su deterioro, los

cuales quedarían ocultos en una segmentación más gruesa. Esta granularidad superior es la que permite que el SAT dispare alertas precisas y diferenciadas según la naturaleza real de cada segmento.

- **Activación de Alertas Dual:** El SAT debe monitorear el forecast de deterioro para cada segmento y disparar alertas bajo dos condiciones:

- Cuando la predicción de saldo vencido supere un umbral predefinido (ej., 5% de crecimiento trimestral pronosticado).

- Cuando las métricas de desempeño del modelo degraden (ej. cae por debajo de o excede el), forzando una intervención del equipo de riesgo para revalidar el modelo o la segmentación.

- **Acciones Proactivas:** Ante una alerta, las áreas de riesgo y comercial deben activar acciones de mitigación inmediatas:

- **Área de Riesgo:** Implementar una revisión acelerada del perfil de riesgo de los clientes en el segmento alertado.

- **Área Comercial:** Iniciar campañas de normalización de cartera o reestructuraciones de deuda antes de que se materialice el deterioro pronosticado.

Conclusiones

La ciencia de datos y el aprendizaje automático representan herramientas clave para anticipar la degradación de carteras en el sector financiero colombiano, permitiendo una transición de modelos reactivos a proactivos. Sin embargo, el análisis crítico de la metodología empleada revela que la reducción a dos componentes principales es un ejercicio de compromiso donde se sacrifica un margen de variabilidad en favor de una estructura de datos manejable para la segmentación

En cuanto a la segmentación, la elección de seis clústeres demostró ser superior a los esquemas tradicionales de 3 o 4 niveles al permitir una granularidad que identifica nichos de riesgo específicos. Pruebas comparativas con $K=4$ revelaron que grupos con comportamientos 'novedosos' quedaban diluidos en promedios estadísticos, perdiendo capacidad de alerta temprana. No obstante, se reconoce que el clúster C6 presenta características de alta volatilidad que podrían derivarse de estructuras de saldo mínimas; en futuras investigaciones, sería pertinente evaluar si estas entidades deben ser tratadas como outliers previos al modelado para depurar aún más el ruido en la segmentación

El análisis histórico de datos crediticios, mediante este enfoque de granularidad controlada, facilita la detección de patrones de riesgo que no son evidentes en modelos agregados, mejorando significativamente la toma de decisiones estratégicas en contextos de alta volatilidad económica.

Modelos como Random Forest, XGBoost y redes neuronales demostraron ser eficaces en la predicción del deterioro de cartera, gracias a su capacidad para manejar relaciones no lineales y variables complejas.

Las técnicas de clusterización, como K-Means y DBSCAN, permiten segmentar entidades financieras según su comportamiento crediticio, lo que abre la puerta a estrategias diferenciadas de gestión del riesgo.

La interpretabilidad de los modelos es fundamental en entornos regulados como el financiero. Herramientas como SHAP y LIME permiten explicar las decisiones algorítmicas, facilitando su adopción por parte de las instituciones.

El enfoque propuesto no solo tiene valor académico, sino también práctico, ya que puede ser implementado por entidades financieras para fortalecer su capacidad de respuesta ante escenarios de riesgo crediticio.

Este trabajo contribuye al ecosistema de investigación local, al integrar ciencia de datos con problemáticas reales del sistema financiero colombiano, y plantea una base sólida para futuras investigaciones y desarrollos tecnológicos en el área.

Referencias Bibliográficas

6.4.3 CLASIFICACIÓN DE LA CARTERA. (s. f.). SUSESO. Compendio Ley 18.833.

Recuperado 3 de junio de 2025, de <https://www.suseso.cl/620/w3-propertyvalue-600859.html>

Altman, E. I. (1968/2000). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy.

Alcívar, S. J. N., Solórzano, R. M. L., Solís, O. P. L., & Sarango, A. F. H. (2023).

Administración de la cartera impaga en la rentabilidad: Una aplicación estadística clasificatoria en bancos: Nonperforming portfolio management in profitability: a statistical application to bank rating. LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades, 4(3), Article 3. <https://doi.org/10.56712/latam.v4i3.1185>

Análisis de crédito como utilizar el análisis de crédito para obtener información valiosa y cuáles son las técnicas. (s. f.). FasterCapital. Recuperado 2 de junio de 2025, de <https://fastercapital.com/es/contenido/Analisis-de-credito--como-utilizar-el-analisis-de-credito-para-obtener-informacion-valiosa-y-cuales-son-las-tecnicas.html>

Análisis predictivo. (s. f.). insightsoftware. Recuperado 2 de junio de 2025, de <https://insightsoftware.com/es/encyclopedia/predictive-analytics/>

Arias, E. S. S. (s. f.). MODELO DE PRONÓSTICO PARA ESTIMAR EL COMPORTAMIENTO DE LA CARTERA VENCIDA PARA CRÉDITOS EN LA MODALIDAD CONSUMO FRENTE A LA TASA DE DESEMPLEO.

Baquero, N. C. (s. f.). MODELO PREDICTIVO DE LA PROBABILIDAD DE AUMENTO DE LOS DÍAS DE MORA PARA USUARIOS DE TARJETA DE CRÉDITO.

- Behera, J., Pasayat, A. K., Behera, H., & Kumar, P. (2023). Prediction based mean-value-at-risk portfolio optimization using machine learning regression algorithms for multi-national stock markets. *Engineering Applications of Artificial Intelligence*, 120, 105843. <https://doi.org/10.1016/j.engappai.2023.105843>
- Broby, D. (2022a). The use of predictive analytics in finance. *The Journal of Finance and Data Science*, 8, 145-161. <https://doi.org/10.1016/j.jfds.2022.05.003>
- Broby, D. (2022b). The use of predictive analytics in finance. *The Journal of Finance and Data Science*, 8, 145-161. <https://doi.org/10.1016/j.jfds.2022.05.003>
- Dorado, O. A. U., & Vanegas, J. A. F. (2024). Análisis predictivo del score de riesgo crediticio mediante machine learning: Una. *Estimacion del Detrimiento de la Cartera de Microcreditos. Caso «Desembolsamos S.A».* (s. f.).
- Giraldo, W. O., & Marín, V. J. (s. f.). Machine Learning para la estimación del riesgo de crédito en una cartera de consumo.
- Hotz, N. (2018, septiembre 10). What is CRISP DM? *Data Science PM*. <https://www.datascience-pm.com/crisp-dm-2/>
- insightsoftware. (2022, enero 19). Análisis predictivo en acción: 5 ejemplos del sector. *insightsoftware*. <https://insightsoftware.com/es/blog/predictive-analytics-in-action-5-industry-examples/>
- Las 11 técnicas más utilizadas en el modelado de análisis predictivos—Insight | Keyrus. (s. f.). Recuperado 2 de junio de 2025, de <https://keyrus.com/sp/es/insights/las-11-tecnicas-mas-utilizadas-en-el-modelado-de-analisis-predictivos>

- Leuro, C. A., & Ospina, J. P. (2025). Data Analysis for the Management of Delinquent Portfolios in Virtual Contact Centers. *Ingeniería y Competitividad*, 27(1), 1-18.
<https://doi.org/10.25100/iyc.v27i1.14327>
- Medrano, E. E. L. (s. f.). Análisis predictivo del aumento de cartera del sector asegurador a través de un modelo de.
- Medrano, L., & Esteban, E. (2025). Análisis predictivo del aumento de cartera del sector asegurador a través de un modelo de machine learning.
<http://repository.unad.edu.co/handle/10596/67164>
- Meza, B., & André, R. (s. f.). Modelo de análisis predictivo para determinar clientes con tendencia a la deserción en bancos peruanos.
- Modelos Predictivos: Tipos, Beneficios y Ejemplos (Guía Completa). (s. f.). Recuperado 3 de junio de 2025, de <https://predikdata.com/es/que-son-y-para-que-se-usan-los-modelos-predictivos/>
- Navas Alcívar, S. J., Loor Solorzano, R. M., López Solís, O. P., & Haro Sarango, A. (2023). Administración de la cartera impaga en la rentabilidad: Una aplicación estadística clasificatoria en bancos: Nonperforming portfolio management in profitability: a statistical application to bank rating. *Latam: revista latinoamericana de Ciencias Sociales y Humanidades*, 4(3), 122.
- Predicción de riesgo crediticio en Colombia usando técnicas de inteligencia artificial. (s. f.). Recuperado 2 de junio de 2025, de https://revistas.uis.edu.co/visores/Revista_UIS_Ingenierias_Vol_19_Num_4/553768213005/

- Rubiano, A. P. T. (s. f.). MODELO PREDICTIVO DE RETENCIÓN DE CLIENTES DE CARTERA PARA EL SEGURO DE AUTOMÓVILES.
- Rueda, D. R., & Valverde, F. L. (s. f.). Smart portfolio management through Deep Learning technique.
- Saavedra García, M. L., & Saavedra García, M. J. (2010). Modelos para medir el riesgo de crédito de la banca. Cuadernos de Administración, 23(40), 295-319.
- Sandoval, L. C. D. (s. f.). ESTRATEGIAS BASADAS EN EL MODELO DE ANÁLISIS PREDICTIVO ÁRBOL DE DECISIÓN PARA LA MEJORA DEL PROCESO DE RECAUDO DE CARTERA DE LA LÍNEA VEHÍCULO PARTICULAR DEL BANCO DAVIVIENDA S.A.
- Scotiabank Colpatría. (s. f.). Scotiabank Colpatría. Recuperado 3 de junio de 2025, de <https://www.scotiabankcolpatría.com.scointnet.net/educacion-financiera/finanzas-personales/tipos-de-credito>
- Solano, C. H. (s. f.). ANÁLISIS DE RIESGO DE CARTERA A TRAVÉS DE MACHINE LEARNING PARA PREDECIR LA PROPENSIÓN DE INCUMPLIMIENTO EN SEGUROS.
- Timón, E. (s. f.). Análisis predictivo: Técnicas y modelos utilizados y aplicaciones del mismo— Herramientas Open Source que permiten su uso.
- Yepes, J. D. M. (s. f.). ANALÍTICA DE DATOS APLICADA A LA COBRANZA DE CARTERA.
- Banco de la República (2024). Reporte de Estabilidad Financiera (Primer Semestre).
- Superintendencia Financiera de Colombia (2024). Informe de Actualidad Financiera con corte a diciembre de 2023.

Baesens, B. (2016). *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Wiley.

Anif (2023). *Comportamiento del Crédito de Consumo en Colombia*. Centro de Estudios Económicos.