



**FACULTAD DE POSTGRADO
TRABAJO FINAL DE GRADUACIÓN**

**DISEÑO E IMPLEMENTACIÓN DE UN MODELO DE
APRENDIZAJE AUTOMÁTICO PARA PREDECIR LA
CARTERA EN MORA DE TARJETAS DE CRÉDITO EN
EL SECTOR BANCARIO COMERCIAL DE HONDURAS
DIC-2018 A SEPT-2023**

SUSTENTADO POR:

**ALVARO LUIS MEDINA ALMENDARES
OSCAR EDUARDO GODOY SALGADO**

PREVIA INVESTIDURA AL TÍTULO DE

**MÁSTER EN
ANÁLITICA DE NEGOCIOS**

**TEGUCIGALPA M.D.C.,
FRANCISCO MORAZÁN,
HONDURAS,
C.A.
FEBRERO, 2024**

**UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA
UNITEC**

FACULTAD DE POSTGRADO

AUTORIDADES UNIVERSITARIAS

RECTORA

ROSALPINA RODRÍGUEZ

VICERRECTOR ACADÉMICO NACIONAL

JAVIER ABRAHAM SALGADO LEZAMA

SECRETARIO GENERAL

ROGER MARTÍNEZ MIRALDA

DIRECTORA NACIONAL DE POSTGRADO

ANA DEL CARMEN RETTALLY VARGAS

**DISEÑO E IMPLEMENTACIÓN DE UN MODELO DE
APRENDIZAJE AUTOMÁTICO PARA PREDECIR LA
CARTERA EN MORA DE TARJETAS DE CRÉDITO EN
EL SECTOR BANCARIO COMERCIAL DE HONDURAS
DIC-2018 A SEPT-2023**

**TRABAJO PRESENTADO EN CUMPLIMIENTO DE LOS
REQUISITOS EXIGIDOS PARA OPTAR AL TÍTULO DE
MÁSTER EN
ANÁLITICA DE NEGOCIOS**

ASESOR METODOLÓGICO

ALBA GABRIELA GARAY ROMERO

MIEMBROS DE LA TERNA:

**KEVIN EDUARDO FUNEZ FUNEZ
ERICK LEONARDO IZAGUIRRE ARDON
DANIEL ANTONIO LUNA RODRIGUEZ**

DERECHOS DE AUTOR

© Copyright 2024
Alvaro Luis Medina Almendares
Oscar Eduardo Godoy Salgado

Todos los derechos son reservados.



FACULTAD DE POSTGRADO

DISEÑO E IMPLEMENTACIÓN DE UN MODELO DE APRENDIZAJE AUTOMÁTICO PARA PREDECIR LA CARTERA EN MORA DE TARJETAS DE CRÉDITO EN EL SECTOR BANCARIO COMERCIAL DE HONDURAS DIC-2018 A SEPT-2023

**ALVARO LUIS MEDINA ALMENDARES
OSCAR EDUARDO GODOY SALGADO**

Resumen

El presente estudio se enfoca en el desarrollo y aplicación de un modelo predictivo de aprendizaje automático para examinar la relación entre las siguientes variables: La colocación de las tarjetas de crédito, la cantidad de tarjetas de crédito activas por periodo y el comportamiento de la cartera de tarjetas de crédito en mora en el sector bancario comercial de Honduras, así mismo el modelo a desarrollar realizará predicciones en data set para ofrecer una herramienta funcional en la toma de decisiones basada en datos para periodos posteriores. Esto se realizará a partir del análisis de conjuntos de datos detallados y de carácter “open source” obtenidos de la página web de la Comisión Nacional de Bancos y Seguros, (en adelante CNBS); se implementarán técnicas avanzadas de Machine Learning en busca de proporcionar una comprensión de cómo las variables anteriormente mencionadas pueden influir en la dinámica financiera de las entidades bancarias en este contexto específico. Este enfoque investigativo se alinea con la necesidad de anticipar y comprender las tendencias del mercado financiero para tomar decisiones estratégicas informadas.

Palabras claves: (CNBS, data set, Machine Learning, Mercado Financiero, Open Source).



GRADUATE SCHOOL

DESIGN AND IMPLEMENTATION OF AN AUTOMATIC LEARNING MODEL TO PREDICT CREDIT CARD DEFAULT PORTFOLIO IN THE HONDURAN COMMERCIAL BANKING SECTOR FROM DEC-2018 TO SEPT-2023

**ALVARO LUIS MEDINA ALMENDARES
OSCAR EDUARDO GODOY SALGADO**

Abstract

The present study focuses on the development and application of a predictive machine learning model to examine the relationship between the following variables: credit card placement, the quantity of active credit cards per period, and the delinquency behavior of the credit card portfolio in the commercial banking sector of Honduras. Additionally, the model to be developed will make predictions on a dataset to provide a functional tool for data-driven decision-making in subsequent periods. This will be done through the analysis of detailed and "open source" datasets obtained from the website of the National Banking and Insurance Commission (hereinafter CNBS). Advanced Machine Learning techniques will be implemented to gain an understanding of how the aforementioned variables may influence the financial dynamics of banking entities in this specific context. This research approach aligns with the need to anticipate and understand trends in the financial market for making informed strategic decisions.

Key words: (CNBS, data set, Machine Learning, Mercado Financiero, Open Source).

DEDICATORIA

Yo Alvaro Luis Medina Almendares dedico este proyecto a Dios por brindarme sabiduría y fortaleza en los momentos más difíciles de la vida, A mi esposa; por darme su apoyo, su amor y su entrega en todo momento y animarme a dar el máximo esfuerzo en esta maestría desde el primer día. A mis padres por siempre tener palabras de ánimo y alentarme a salir adelante y dar lo mejor de mí en cada proyecto que decido tomar.

Yo, Oscar Eduardo Godoy Salgado, dedico este proyecto a Dios por su guía y su gracia y a mi madre extraordinaria, quien ha sido mi pilar fundamental a lo largo de este arduo camino académico. Expresándole agradecimiento por ser mi fuente constante de inspiración, por alentarme a perseguir mis metas y por ser mi apoyo en cada paso que he dado.

AGRADECIMIENTO

Agradecemos a la institución/cuerpo docente por preparar un plan de estudios acertado y compartir sus conocimientos con la mejor disposición.

ÍNDICE DE CONTENIDO

DEDICATORIA	vii
AGRADECIMIENTO	viii
CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN	1
1.1 INTRODUCCIÓN	1
1.2 ANTECEDENTES DEL PROBLEMA	2
1.3 DEFINICIÓN DEL PROBLEMA	5
1.4 OBJETIVOS DEL PROYECTO.....	6
1.5 JUSTIFICACIÓN.....	7
CAPÍTULO II. MARCO TEÓRICO	9
2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL	9
2.1.1 Macroentorno.....	9
2.1.2 Microentorno	12
2.1.3 Entorno Nacional.....	13
2.2 CONCEPTUALIZACIÓN	16
2.2.1 MACHINE LEARNING	16
2.2.2 APRENDIZAJE SUPERVISADO	17
2.2.3 APRENDIZAJE SUPERVISADO POR REGRESIÓN.....	18
2.2.4 APRENDIZAJE SUPERVISADO POR CLASIFICACIÓN.....	19
2.2.5 DIFERENCIA ENTRE ALGORITMOS DE REGRESIÓN Y DE CLASIFICACIÓN	
20	
2.2.6 MODELO DE APRENDIZAJE POR REGRESIÓN LINEAL	21
2.2.7 MODELO DE APRENDIZAJE POR RANDOM FOREST.....	21
2.2.8 PRONÓSTICOS	23
2.2.9 CRÉDITO.....	23
2.2.10 TARJETA DE CRÉDITO	24
2.2.11 RIESGO CREDITICIO	25
2.2.12 GESTIÓN DEL RIESGO CREDITICIO	26
2.2.13 CARTERA.....	26
2.2.14 TIPOS DE CARTERA	26
2.2.15 INSTITUCIÓN BANCARIA COMERCIAL	27

2.3	TEORÍAS DE SUSTENTO	27
2.3.1	BASES TEÓRICAS	27
2.3.2	TEORÍA DE LA DEMANDA Y OFERTA DE CRÉDITO	27
2.3.3	TEORÍA DEL CICLO ECONÓMICO Y CICLO DE CRÉDITO	29
2.3.4	TEORÍA DEL RIESGO Y RENDIMIENTO.....	29
2.3.5	OTROS ESTUDIOS CONSULTADOS	31
2.3.5.1	Modelos predictores de la morosidad con variables macroeconómicas	31
2.3.5.2	Determinantes de la demanda de crédito. Una estimación con un modelo mensual de series de tiempo para Venezuela	31
2.4	METODOLOGÍAS DESARROLLADAS EN LOS ESTUDIOS CONSULTADOS ...	32
2.4.1	Determinantes de la demanda de crédito. Una estimación con un modelo mensual de series de tiempo para Venezuela	32
2.5	METODOLOGÍAS A DESARROLLAR EN EL PRESENTE ESTUDIO	32
2.5.1.1	Modelo Estadístico.....	32
2.5.1.2	Machine Learning	33
2.5.1.3	Segmentación de Datos.....	33
2.5.1.4	VARIABLES e Indicadores.....	33
2.5.1.5	Procesos de Limpieza y Preprocesamiento:.....	33
2.6	INSTRUMENTOS UTILIZADOS	34
2.7	MARCO LEGAL	34
2.7.1	Reglamento de la Ley de Tarjetas de Crédito.....	34
2.7.2	Ley de Protección al Consumidor	35
2.7.3	Ante Proyecto de Ley de Protección de Datos Personales	37
2.7.4	Ley de Transparencia y Acceso a la Información Pública	38
CAPÍTULO III. METODOLOGÍA		39
3.1	CONGRUENCIA METODOLÓGICA.....	39
3.2	MATRIZ DE CONGRUENCIA METODOLÓGICA.....	40
3.3	ESQUEMA DE VARIABLES DE ESTUDIO	43
3.4	OPERACIONALIZACIÓN DE LAS VARIABLES	44
3.5	ENFOQUE Y MÉTODOS	45
3.6	DISEÑO DE LA INVESTIGACIÓN	46

3.7	POBLACIÓN.....	47
3.8	MUESTRA.....	47
3.9	FUENTES DE INFORMACIÓN.....	49
3.9.1	FUENTES PRIMARIAS.....	49
3.9.2	FUENTES SECUNDARIAS.....	49
CAPÍTULO IV. RESULTADOS Y ANÁLISIS		51
4.1	INFORME DE PROCESO DE RECOLECCIÓN DE DATOS	51
4.2	ANÁLISIS EXPLORATORIO DE DATOS	53
4.3	RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS.....	58
4.3.1	RESULTADOS CUANTITATIVOS.....	58
4.3.1.1	MODELO DE REGRESIÓN LINEAL CON CANTIDAD DE TARJETAS ACTIVAS COMO VARIABLE INDEPENDIENTE.....	58
4.3.1.2	MODELO DE REGRESIÓN LINEAL CON TODAS LAS VARIABLES	60
4.3.1.3	MODELO RANDOM FOREST REGRESSOR POR BANCO.....	62
4.3.2	ANÁLISIS CUALITATIVO.....	66
4.4	ANÁLISIS INFERENCIAL Y MODELOS APLICADOS.....	67
4.4.1	REGRESION LINEAL	67
4.4.2	RANDOM FOREST REGRESSION PARA TODOS LOS BANCOS	68
4.4.3	RANDOM FOREST REGRESSOR PARA CALCULO DE PREDICCIONES DE CARTERA EN MORA PARA CADA BANCO EN PARTICULAR.	71
4.4.4	RANDOM FOREST REGRESSOR PARA CALCULO DE PRONOSTICO DE CARTERA EN MORA PARA CADA BANCO EN PARTICULAR.	76
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES.....		78
5.1	CONCLUSIONES	78
5.2	RECOMENDACIONES	80
CAPÍTULO VI. APLICABILIDAD.....		83
6.1	NOMBRE DE LA PROPUESTA	83
6.2	JUSTIFICACIÓN DE LA PROPUESTA.....	83
6.3	ALCANCE DE LA PROPUESTA	83
6.4	DESCRIPCIÓN Y DESARROLLO	84
6.4.1	DESCRIPCIÓN.....	84

6.4.2	DESARROLLO.....	85
6.4.3	EVALUACIÓN Y MÉTRICAS.....	85
6.5	MEDIDAS DE CONTROL	86
6.6	CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO	86
6.7	CONCORDANCIA DE LOS SEGMENTOS DE LA TESIS	92
	REFERENCIAS BIBLIOGRÁFICAS.....	95
	ANEXOS	102

ÍNDICE DE FIGURAS

Figura 1-	Deterioro en la cartera crediticia y su impacto en la rentabilidad bancaria	4
Figura 2.	Deterioro en la cartera crediticia y su impacto sobre la solvencia bancaria.....	4
Figura 3.	Penetración del pago a crédito en América Latina	10
Figura 4.	Penetración de mercado Latam por Marca de tarjeta.	11
Figura 5.	Cartera vía tarjeta de crédito en Millones de LPS a diciembre de cada año.	14
Figura 6.	Emisores de tarjetas de crédito en Honduras, Ranking de mayor crecimiento interanual de cuota de mercado.....	15
Figura 8.	Funcionamiento del aprendizaje supervisado.....	18
Figura 9.	Modelo de Regresión.....	19
Figura 10.	Modelo de clasificación.....	19
Figura 12.	Diferencia entre Modelos de Clasificación y Regresión.	20
Figura 14.	Estructura de un Random Forest.	22
Figura 16.	Representación gráfica de la teoría de oferta y demanda de crédito.	28
Figura 17.	Matriz de congruencia metodológica.	41
Figura 18.	Variación mensual de las variables de estudio.....	53
Figura 19.	Tarjetas en circulación y cartera en mora por mes.	53
Figura 20.	Promedio de cartera en mora por banco	54
Figura 21.	Promedio de tarjetas en circulación por banco.....	55
Figura 22.	Promedio de cartera total por banco. Fuente: (Elaboración Propia)	56
Figura 23.	Comportamiento de cartera en mora por año.	56
Figura 24.	Comportamiento de cartera total por año.	57
Figura 25.	Comportamiento cantidad de tarjetas activas en circulación por año.	58
Figura 26.	Cronograma de ejecución del proyecto.	87

ÍNDICE DE TABLAS

Tabla 1-	Diferencia entre banca comercial y banca de desarrollo.	27
Tabla 3-	Diferencias entre tasa activa y tasa pasiva.	30
Tabla 4.	Esquema de variables de estudio.	44
Tabla 5-	Operacionalización de las variables (Elaboración propia)	45
Tabla 6.	Población para el estudio.	47

Tabla 8. Resumen de métricas, modelo de regresión lineal con cantidad de tarjetas activas como variable independiente.	58
Tabla 7. Resumen de métricas, modelo de regresión lineal con todas las variables de estudio. ...	60
Tabla 9. Resumen de métricas, modelo Random forest regressor banco 1 al 3	62
Tabla 10. Resumen de métricas, modelo Random forest regressor banco 4 al 6	62
Tabla 11. Resumen de métricas, modelo Random forest regressor banco 7 al 9.	62
Tabla 12. Tabla 12- Resumen de métricas, modelo Random forest regressor banco 10 y 11.	63
Tabla 13. Alcance de la propuesta de aplicabilidad.	83
Tabla 15- Desglose de presupuesto para implementación del proyecto.	90
Tabla 17- Concordancia de los segmentos de la tesis con la propuesta.	93

ÍNDICE DE ANEXOS

ANEXO 1 : PROGRAMACIÓN KNIME, PREPARACIÓN DE DATA PARA MODELOS..	102
ANEXO 2 : DASHBOARD EDA EN MS POWERBI.	103
ANEXO 3: CORRELACIÓN ENTRE VARIABLES USANDO PYTHON MATPLOTLIB ..	104
ANEXO 4: FLUJO MODELO DE REGRESIÓN LINEAL	105
ANEXO 5: FLUJO MODELO PREDICCIÓN CARTERA EN MORA CON RANDOM FOREST REGRESSOR.	106
ANEXO 6: FLUJO MODELO PREDICCIÓN CARTERA EN MORA CON RANDOM FOREST REGRESSOR.	107
ANEXO 7 : FLUJO PREDICCIÓN DE CARTERA EN MORA CON EL MODELO ENTRENADO.	108

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1 INTRODUCCIÓN

Las primeras tarjetas de crédito fueron introducidas en los Estados Unidos de América. Este concepto se originó en la década de 1950; fue la compañía Diners Club emitió la primera tarjeta de crédito que permitía a los titulares realizar compras en varios establecimientos sin necesidad de portar dinero en efectivo. Originalmente se centraba en su utilización para amenidades como ser restaurantes y hoteles. (History of Credit Cards, s.f.)

Posteriormente en el año 1958, el Bank of America lanzó al mercado la “BankAmericard” que eventualmente se convertiría en la mundialmente conocida tarjeta “VISA”, la idea detrás de dicha tarjeta era la de ofrecer una forma de pago universalmente aceptada y permitir a los consumidores realizar compras en una amplia gama de comercios. (The History of Visa, s.f.)

A lo largo de los años, el concepto se ha ido expandiendo y evolucionando, dando lugar a la industria de tarjetas de crédito.

La tarjeta de crédito es un instrumento que permite adquirir bienes, servicios y efectuar retiros de dinero en el momento que el titular lo desee, hasta el margen o límite de crédito pre acordado con la empresa emisora de la tarjeta. El emisor de tarjetas de crédito facilita un plástico con el nombre de la persona que ha contratado una línea de crédito con un límite de compra y un límite de crédito. “El usuario puede utilizar la tarjeta de crédito en la realización de diversas compras de bienes o servicios, obligándose a restituir al emisor de la tarjeta de crédito, el valor de dichas compras. Si el usuario paga de contado el valor total de las compras realizadas en el mes, no pagará al emisor ningún interés. Por el contrario, si el usuario paga solamente una parte del importe de las compras realizadas, obtiene un crédito por la diferencia y debe pagar el interés fijado por el emisor”. (2020 CNBS - GPUF - Educación Financiera., s.f.)

La evolución de las tarjetas de crédito ha sido notable desde su invención. “Desde su encarnación original como una tarjeta de cartón del Diner’s Club hasta las variedades de metal con chip incrustado que encontramos hoy en día, la premisa fundamental detrás de una tarjeta de crédito ha permanecido igual: una tarjeta de crédito es un método para comprar ahora y pagar después”.(When Were Credit Cards Invented: The History of Credit Cards – Forbes Advisor, s. f.)

Las tarjetas de crédito han evolucionado rápidamente en las últimas décadas, pasando de tomar una impresión física de la tarjeta a deslizar, introducir, tocar o agitar su información en un terminal de pago (When Were Credit Cards Invented: The History of Credit Cards – Forbes Advisor, s. f.). En 2023, el uso de tarjetas de crédito representó más de un tercio (36.3%) de las transacciones minoristas en tiendas.(Credit Card Industry in 2023: Analysis and Trends in Payment Processing, s. f.).

El uso de tarjetas de crédito por parte de los consumidores crea un ciclo económico virtuoso en el que niveles más altos de consumo conducen a niveles más altos de producción, a su vez, creando más empleos e ingresos más altos, permitiendo que la economía prospere (Outlook Money s. f.2021). Sin embargo, también es importante tener en cuenta que el endeudamiento puede sostener una economía, pero el crecimiento eventualmente se detiene cuando los hogares operan con pérdidas. Cuando las personas no pueden permitirse pagar sus tarjetas de crédito, necesitan reducir su nivel de vida. Esto es malo para la economía y puede llevar a períodos de recesión (Credit Card Debt & its Effect on the Economy | Pocketsense, s. f.).

Las tarjetas de crédito también han jugado un papel sorprendentemente grande en la Gran Recesión. “Los economistas y los responsables de las políticas deberían considerar el mercado de tarjetas de crédito cuando intentan entender cómo una crisis de crédito dentro del sector financiero afecta a los consumidores” (Federal Reserve Bank of Philadelphia 2021).

1.2 ANTECEDENTES DEL PROBLEMA

De acuerdo con una investigación conducida por diario La Prensa de Honduras la cartera crediticia bancaria de Honduras vía tarjetas de crédito ascendió a 34,738.3 millones de lempiras para junio de 2022 de acuerdo con el “Informe de estabilidad financiera” del Banco Central de Honduras. (BCH, 2022)

“La expansión interanual fue de un 25.3% o 7,007.4 millones de lempiras respecto a junio de 2021 cuando la cartera se encontraba en 27,730.0 millones de lempiras.” (BCH, 2022)

De acuerdo con el informe del Banco Central dicho monto de expansión interanual se compone de 5,304.7 millones para consumo y 1,702.7 millones para comercio.

El BCH analiza el comportamiento de la cartera de tarjetas de crédito desde junio 2019 y denota que en el bienio 2021-2022 se demarca un dinamismo mayor en el uso de estas por parte de los tarjetahabientes.

A junio de 2022 el BCH y la CNBS reportan 802,375 tarjetas en circulación y 440.248 deudores. “De acuerdo con la tendencia a partir del 2022 se ha ampliado la cantidad de tarjetas circulando en un número más grande de deudores. La relación de tarjetas circulando entre tarjetahabientes es de aproximadamente dos plásticos por cliente” indica el informe de estabilidad financiera (BCH, 2022).

Además, es importante destacar que el uso de tarjetas de crédito ha tenido un impacto significativo en la economía hondureña. Según un informe del Banco Interamericano de Desarrollo (BID), el uso de tarjetas de crédito puede impulsar el crecimiento económico al facilitar las transacciones y proporcionar a los consumidores una mayor flexibilidad en su gasto. Sin embargo, el BID también advierte que un uso excesivo de las tarjetas de crédito puede llevar a niveles insostenibles de deuda, lo que genera efectos negativos en la economía a largo plazo (BID, 2021).

De acuerdo al IEF El coeficiente de cobertura de los préstamos en mora se situó en 214.7% al cierre del primer semestre de 2023, manteniéndose en valores por encima del porcentaje mínimo requerido por la Autoridad Supervisora que es del 110.0% (IEF junio 2023.pdf, s. f.)

En términos del Banco Central de Honduras “La principal forma en la que el riesgo de crédito se materializa en las instituciones bancarias es por el gasto que efectivamente se incurre por créditos que ya no pueden ser recuperados”. (IEF junio 2023.pdf, s. f.)

Trasladar al gasto la irrecuperabilidad de los préstamos significa reducción en las utilidades de las entidades bancarias, esta merma en la rentabilidad afecta a su vez la capitalización de las entidades. Por lo anterior, se efectúan pruebas de sensibilidad que permiten evaluar la resiliencia del sistema bancario ante diversos grados de deterioro en la cartera crediticia.

Si se toma la cartera de préstamos que ha sido refinanciada y readecuada por el sector privado no financiero (empresas y hogares). Las cifras a junio de 2023 indican que, la rentabilidad medida por el ROE de las instituciones bancarias en conjunto fue de 15.7%. Los resultados manifestaron que, en los dos últimos escenarios 7 y 8 bancos estarían reportando pérdidas, respectivamente, incluyendo 2 y 3 bancos sistémicos, en su orden.

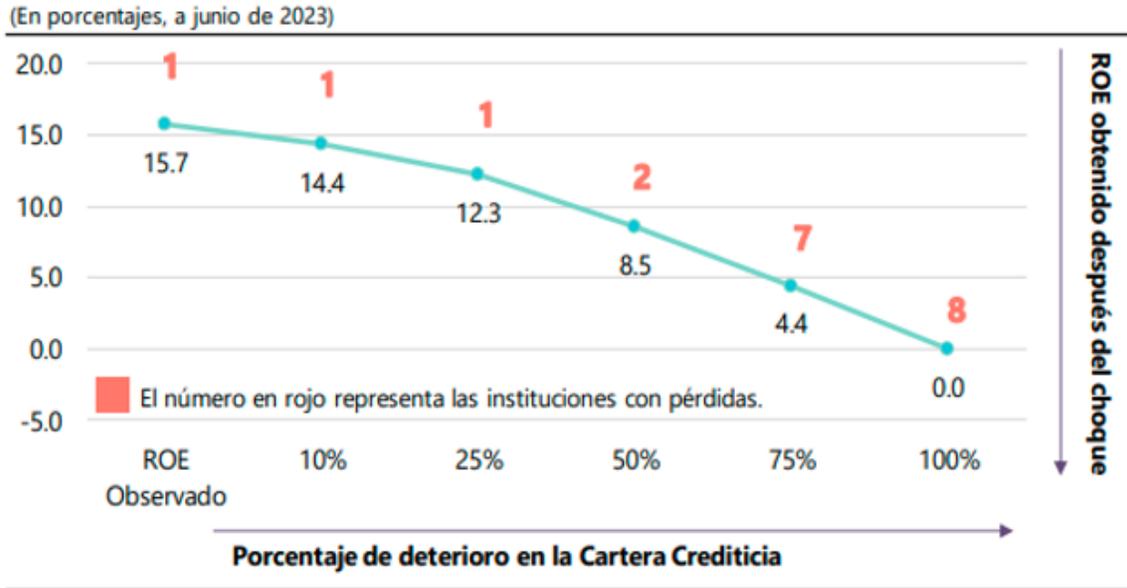


Figura 1-Deterioro en la cartera crediticia y su impacto en la rentabilidad bancaria

Fuente: (IEF 2022 BCH con cifras de CNBS.)

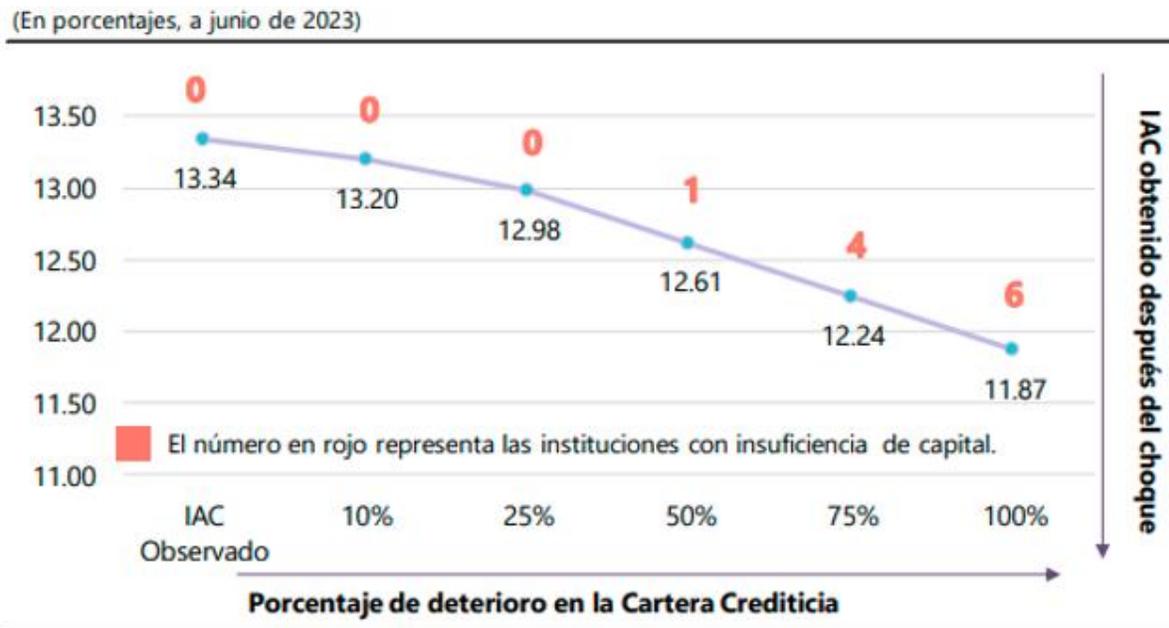


Figura 2. Deterioro en la cartera crediticia y su impacto sobre la solvencia bancaria.

Fuente: (IEF 2022 BCH con cifras de CNBS.)

En cuanto a la solidez bancaria, el Índice de Adecuación de Capital (IAC) se posicionó en 13.3% a junio de 2023. Del ejercicio de sensibilidad, se infiere que los bancos como conjunto no ubicarían su solvencia por debajo del mínimo regulatorio (11.50%) en ninguno de los escenarios de deterioro en el crédito. (IEF junio 2023.pdf, s. f.).

Si bien los escenarios en los que las instituciones comienzan a fallar de forma masiva son los más extremos y no se ven afectados solamente por la morosidad de tarjetas de crédito hoy por hoy, se puede concluir de estas pruebas que los bancos tienen cierta sensibilidad ante la cartera refinanciada y reestructurada (lo que incluye aquella que proviene de tarjetas de crédito), a pesar de que ésta representa una pequeña proporción de la cartera total en gran parte de los casos. “Por ello, es conveniente que mantengan la debida vigilancia sobre la calidad del portafolio de préstamos y que continúen fortaleciendo sus técnicas de gestión del riesgo crediticio” (IEF junio 2023.pdf, s. f.).

1.3 DEFINICIÓN DEL PROBLEMA

Planteamiento del problema:

El marcado incremento en la emisión de tarjetas de crédito y la creciente cantidad de tarjetas activas en el mercado crediticio de la banca comercial del sector privado hondureño han suscitado un escenario donde las instituciones financieras presentan una amplia gama de opciones al público. Este fenómeno va acompañado de una clara preferencia por parte de los consumidores hacia el uso de este método de pago. En este contexto, se hace imperativo analizar de manera detallada cómo se comporta la cartera crediticia, sobre todo la cartera en mora de tarjetas de crédito en relación con estas variables.

Problema de investigación:

El presente estudio se propone examinar el comportamiento de la cartera crediticia en mora de tarjetas de crédito de la banca comercial del sector privado hondureño. Específicamente, se busca comprender la influencia de una variable clave: la cantidad de tarjetas activas en circulación por institución bancaria sobre el comportamiento de la cartera en mora para 11 de los bancos que

conforman la banca comercial de Honduras; a su vez, se desea construir un modelo de machine learning que permita predecir el valor de la cartera en mora de tarjetas de crédito a partir de diversas variables que permitan verificar si se pueden obtener resultados satisfactorios para las instituciones objeto de este estudio.

Preguntas de Investigación:

Para poder desarrollar el problema anteriormente planteado se proponen las siguientes preguntas que permitirán navegar a través del desarrollo del proyecto:

- 1) ¿Existe una vinculación estadísticamente significativa y lineal de al menos el 75% entre las variables: cantidad de tarjetas de crédito activas en circulación y el crecimiento de la cartera en mora de tarjeta de crédito?
- 2) 2) ¿Es posible el desarrollar un modelo de machine learning que permita predecir el valor de la cartera en mora con una precisión estadística mínima del 85% de certeza a partir del número de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes para cada institución bancaria estudiada?

1.4 OBJETIVOS DEL PROYECTO

OBJETIVO GENERAL

El objetivo principal de este estudio es diseñar e implementar un modelo de aprendizaje automático para predecir el valor de la cartera en mora de tarjetas de crédito a partir de la evaluación de las variables independientes (cantidad de tarjetas activas en circulación y cantidad de tarjetas colocadas por mes) y su relación, en el sector bancario comercial de Honduras.

OBJETIVOS ESPEFICOS

- A. Evaluar el impacto que tiene la cantidad de tarjetas en circulación activas en el comportamiento de la cartera crediticia en mora y validar si entre estas variables existe una relación lineal del 75% o más para la totalidad de las instituciones objetos de este estudio.

- B. Identificar patrones y tendencias específicas de la cartera en mora en el contexto de la banca comercial privada hondureña a través del análisis de los datos obtenidos del portal de transparencia de la CNBS.
- C. Utilizar los dataset obtenidos para construir un modelo de aprendizaje automático que permita realizar proyecciones de la cartera en mora para los 11 bancos objetos de este estudio y que conforman el 80% de la banca comercial privada, considerando las variables estudiadas y las variaciones temporales.
- D. Contribuir al conocimiento financiero local proporcionando referencia para futuras investigaciones sobre la dinámica de tarjetas de crédito en el sector bancario en Honduras.

1.5 JUSTIFICACIÓN

La relevancia de esta investigación radica en la necesidad de proporcionar una visión integral y actualizada sobre el comportamiento de la cartera crediticia de tarjetas de crédito en un entorno financiero dinámico. La comprensión de cómo estas variables impactan a nivel individual y en conjunto permitirá a las instituciones financieras ajustar estrategias y políticas de manera más precisa, considerando tanto su tamaño de mercado como las variaciones temporales.

Es por ello que se busca no solo comprender el comportamiento actual de la cartera crediticia en mora, sino también anticipar posibles escenarios futuros. Esto permitirá a las instituciones financieras tomar decisiones informadas y estratégicas para gestionar de manera eficaz sus carteras de tarjetas de crédito en el entorno dinámico del mercado hondureño.

El presente estudio reviste una importancia significativa en el contexto financiero de Honduras. Con el vertiginoso incremento en la cartera de tarjetas de crédito durante los últimos tres años indica un cambio sustancial en los patrones de consumo y la forma en que los hondureños gestionan sus transacciones financieras. Esta tendencia plantea interrogantes cruciales sobre como este fenómeno influye en la estabilidad, rendimiento y rentabilidad del sector bancario del país.

El desarrollo de un modelo predictivo se torna esencial en este escenario, ya que permitirá anticipar el impacto futuro con base en la cantidad de tarjetas en circulación y su impacto en la cartera en mora de tarjetas de crédito que se ajuste a cada una de las instituciones bancarias del sector comercial privado hondureño. Esto no solo proporciona una valiosa herramienta para

evaluar el comportamiento del sector, sino que también otorga una visión más precisa y fundamentada de las dinámicas financieras locales.

Adicional, el presente estudio puede servir para futuras investigaciones llevadas a cabo por o para los entes reguladores del sistema financiero, ya que la comprensión detallada de como el cambio en la cantidad de tarjetas de crédito en circulación puede influir en la estabilidad y desarrollo económico del país es esencial para garantizar un entorno financiero sostenible y robusto para nuestro país.

CAPÍTULO II. MARCO TEÓRICO

2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL

2.1.1 Macroentorno

El mercado de tarjetas de crédito en América Latina ha experimentado cambios significativos en los últimos años, especialmente en el período de 2021 a 2023. A continuación, se presenta un análisis del macroentorno de este mercado:

- *Factores Económicos:*

El crecimiento económico de América Latina ha sido uno de los principales factores que han impulsado el crecimiento del mercado de tarjetas de crédito. El aumento de los ingresos de los consumidores ha hecho que estos sean más propensos a utilizar tarjetas de crédito para realizar sus compras (BCIE, 2023).

Según datos del Banco Interamericano de Desarrollo (BID), “el crecimiento económico de América Latina se estima en un 3,5% en 2023”. Este crecimiento se debe a una serie de factores, como la recuperación de la demanda global, el aumento de los precios de las materias primas y la mejora de las condiciones financieras. (Banco de Integración Económica (BID), 2023).

- *Factores Demográficos:*

El crecimiento de la población en América Latina también ha contribuido al crecimiento del mercado de tarjetas de crédito. “La población de la región se estima en 650 millones de personas en 2023”, y se espera que alcance los 700 millones de personas en 2030. (BCIE, 2023)

El aumento de la población ha creado un mercado más grande para las tarjetas de crédito. Además, la creciente urbanización ha llevado a un aumento de la demanda de productos y servicios financieros, como las tarjetas de crédito.

- *Factores Tecnológicos:*

El desarrollo de las tecnologías financieras (fintech) también ha contribuido al crecimiento del mercado de tarjetas de crédito.

Las fintech están ofreciendo nuevos productos y servicios de tarjetas de crédito, como tarjetas de crédito sin cuotas de mantenimiento o con recompensas exclusivas (rankinglatam/shopify, 2023).

- *Crecimiento del mercado:*

“En 2022, el mercado de pagos de América Latina se situó en un valor promedio de USD 41,84 mil millones”. Se espera que este mercado se expanda a una tasa de crecimiento anual compuesta (CAGR) del 8,60% durante el período de pronóstico (2023-2028) (Penetración del pago de crédito en Latam, 2019). Según Stéphanie Chevalier Naranjo quien es periodista basada en datos de statista.com: “Este crecimiento se atribuye en gran medida al aumento en el uso de tarjetas de crédito como medio de pago electrónico líder en la región”.

En un informe realizado por Minsait, de todos los analizados, Brasil y la República Dominicana son las economías en las que los ciudadanos usan las tarjetas para pagar a plazos con mayor frecuencia, “con cerca del 82% y 75% de la población con tarjeta de crédito que dice usarla siempre o casi siempre para este fin. (X Edición del Informe de Tendencias en Medios de Pago, 2021).

De acuerdo con los datos que presenta el informe de Minsait “el porcentaje de personas que realiza pagos a crédito podría ser aún mayor en gran parte de los países latinoamericanos estudiados si su utilización estuviera vinculada a una tasa de interés preferencial o un tipo de interés más bajo” (Penetración del pago de crédito en Latam, 2019).

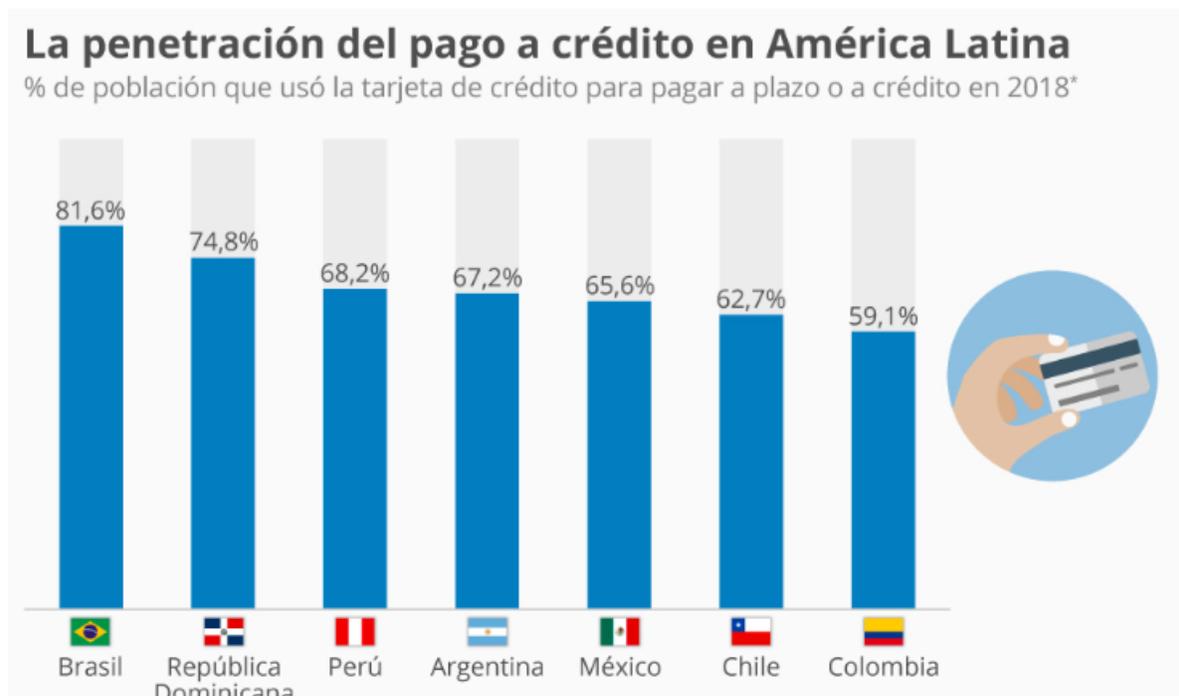


Figura 3. Penetración del pago a crédito en América Latina

Fuente: (Statista ,2020)

- *Conformación del Mercado:*

El mercado de tarjetas de crédito en América Latina está compuesto por más de 600 emisores y adquirentes (Mercado de Pagos de América Latina, Tamaño, Informe 2024-2032, s. f.).

Se ha visto el 33% aumento en las transacciones en línea a través de tarjetas de crédito nacionales, 13% en tarjetas de débito, 11% en carteras electrónicas, y 9% en vales en la región objetivo. Por lo tanto, un impresionante crecimiento en mayores oportunidades para los actores clave del mercado de pagos en América Latina es evidente a un ritmo rápido durante el periodo de pronóstico (Mercado de Pagos de América Latina, Tamaño, Informe 2024-2032, s. f.).

Las tarjetas de crédito representan el 50% del comercio electrónico y el 20% de los pagos presenciales. (Juárez, 2023).

A continuación, se ilustra la penetración de los principales emisores de tarjetas de crédito en América Latina:

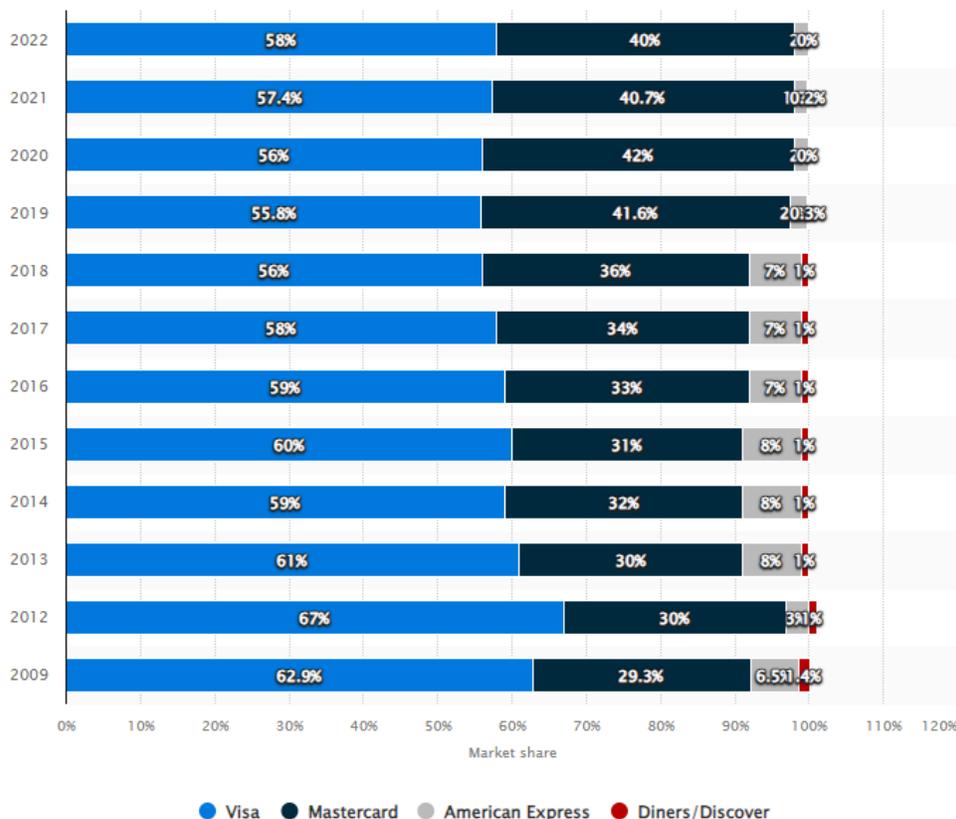


Figura 4. Penetración de mercado Latam por Marca de tarjeta.

Fuente: (Rankings Latam, 2023)

- *Valor de Cartera Total:*

El valor total de la cartera de crédito en América Latina ha experimentado un crecimiento constante. En 2022, el mercado de pagos de América Latina se situó en un valor promedio de USD 41,84 mil millones (Penetración del pago de crédito en Latam, 2019).

- *Valor de Cartera en Mora:*

La cartera en mora es un aspecto crucial en el análisis del mercado de tarjetas de crédito. Aunque no se dispone de datos específicos para América Latina en su conjunto, se puede observar que, en países como Colombia, la cartera en mora ha estado creciendo debido a la pandemia y sus años posteriores (Reporte FINLAC, 2015).

En resumen, el mercado de tarjetas de crédito en América Latina ha experimentado un crecimiento significativo en los últimos años, impulsado por el aumento en el uso de tarjetas de crédito como medio de pago electrónico. Sin embargo, también se enfrenta a desafíos, como el aumento de la cartera en mora sobre todo en el ámbito de los pagos en cuotas posteriores al periodo de la pandemia (año 2020 en adelante).

2.1.2 Microentorno

"En cuanto a la calidad de la cartera crediticia en la región, se reflejan interanualmente menores índices de morosidad, casos que son más pronunciados en Nicaragua, Honduras y Guatemala." (p.41 del informe de estabilidad financiera) (BCH, 2022); Esto nos indica una tendencia positiva de mejora en la calidad de la cartera crediticia en la región.

Según el informe de Estudio Económico de América Latina y el Caribe 2023 de la CEPAL, se espera que la economía de Centroamérica crezca un 3.8% en 2023, impulsada por la recuperación de la demanda interna y externa, así como por la inversión pública y privada (Cepal, Caribe, 2023). Además, el informe destaca que la región ha experimentado una mejora en la calidad de la cartera crediticia, lo que puede ser un indicador de una gestión más eficiente de riesgos crediticios y de un entorno económico más estable en estos países (Informes de Estabilidad Financiera, BCCR, s. f.).

La CEPAL destaca que la región debe seguir trabajando en mejorar la calidad de la cartera crediticia y en fortalecer la inversión pública y privada para mantener un crecimiento sostenible.

2.1.3 Entorno Nacional

El mercado de tarjetas de crédito en Honduras ha experimentado un crecimiento significativo en los últimos años. Este crecimiento se debe a diversos factores, como el aumento del poder adquisitivo de la población, la expansión de la banca comercial y el desarrollo del comercio electrónico.

Según datos de la Comisión Nacional de Bancos y Seguros en datos publicados por el diario El Heraldo: “En los últimos tres años en Honduras, se ha observado un notable crecimiento en la colocación de tarjetas de crédito.

En 2021, se emitieron un total de 11,968 tarjetas de crédito, cifra que aumentó a 16,686 en 2022 y alcanzó la cifra de 21,725 tarjetas de crédito en 2023” (El Heraldo, s.f.).

A diciembre de 2022, los once bancos emisores de tarjetas de crédito objetos de este estudio obtuvieron ingresos por L7,302.9 millones por concepto de intereses (representando el 14.5% de los ingresos financieros totales) (BCH, 2022); con esto podemos observar la significativa contribución de los intereses generados por las tarjetas de crédito al total de los ingresos financieros de los bancos emisores. Esta cifra muestra que las tarjetas de crédito son una fuente de ingresos clave para estas instituciones financieras.

Las tarjetas de crédito son de suma importancia como opción de financiamiento popular tanto para individuos como para empresas en el país Según el diario la prensa quien cita palabras redactadas por el Banco Central de Honduras en sus informes de estabilidad financiera: “La tarjeta de crédito es una de las fuentes de financiamiento preferidas por las personas y empresas en el país.” (BCH, 2022) esto lo podemos notar en como los clientes obtienen sus préstamos.

Las instituciones emisoras de tarjetas de crédito están experimentando un aumento en la demanda de crédito por parte de los clientes. “Los emisores de Tarjetas de Crédito han elevado sus préstamos, la variación interanual fue 30.1%, es decir, un aumento de L9,676.2 millones, asociado con el dinamismo del consumo de sus principales usuarios: los hogares. Del total de la cartera vía TC (L41,862.2 millones)” (p.42 del informe de estabilidad financiera) (BCH, 2022).

Las personas y las familias son los principales beneficiarios de estos préstamos y las tarjetas de crédito son un medio popular para financiar sus gastos, El total de la cartera vía tarjetas de crédito alcanza L41,862.2 millones. Esto muestra la magnitud de la cartera crediticia que está siendo gestionada a través de tarjetas de crédito en el mercado.

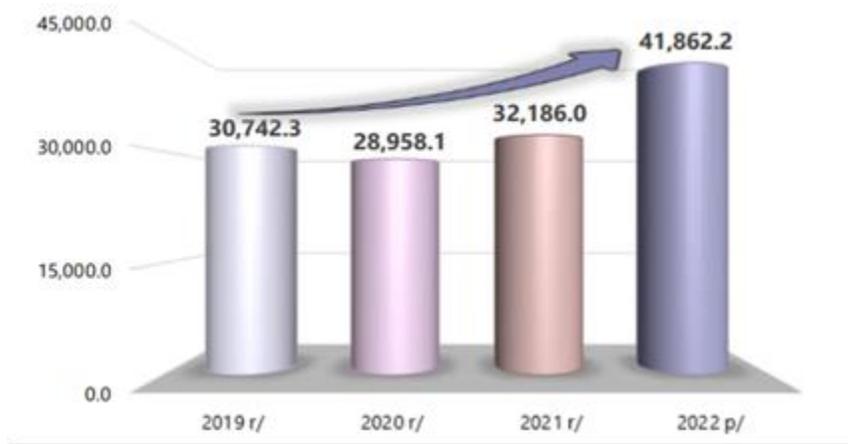


Figura 5. Cartera vía tarjeta de crédito en Millones de LPS a diciembre de cada año.

Fuente: (BCH, 2022)

- *Mercado de tarjetas de crédito para los años 2022, 2023 en adelante:*

De acuerdo con el reporte Rankings Latinoamérica “El mercado de tarjetas de crédito en Honduras ha experimentado un crecimiento significativo en los últimos años, lo que indica una tendencia positiva tanto para los consumidores como para las instituciones financieras”. (rankinglatam/shopify, 2023)

Según el último reporte publicado que data de marzo 2023, el tamaño del mercado alcanzó los USD 1.773 millones en 2022, mostrando un aumento desde los USD 1.493 millones en 2021. Además, se espera que el mercado continúe su trayectoria ascendente, con un tamaño proyectado de USD 2.300 millones para 2025 (rankinglatam/shopify, 2023).

En cuanto a participación de mercado considerando las tarjetas de crédito activas, los cinco principales emisores en Honduras dominan el panorama. A diciembre de 2022, BAC Honduras tiene la mayor participación de mercado con un 35,1 %, seguido por BANCO FINANCIERA COMERCIAL HONDUREÑA (FICOHSA) con un 21,1 %. BANCO ATLANTIDA, BANCO DEL PAIS y BANCO PROMERICA completan el top cinco con participaciones de mercado de 11,4%, 7,7% y 6,5%, respectivamente. (rankinglatam/shopify, 2023)

Número de Tarjetas de Crédito activas

rankingslatam	RK Δ		Δ Market Share	2023 03	2022 03
	1		↑ 0,59%	7,7%	7,2%
	2		↑ 0,42%	5,5%	5,0%
	3		↑ 0,34%	11,4%	11,0%
	4		↑ 0,30%	6,5%	6,2%
	5		↑ 0,15%	4,2%	4,1%

Figura 6. Emisores de tarjetas de crédito en Honduras, Ranking de mayor crecimiento interanual de cuota de mercado.

Fuente: (Rankings Latam ,2023)

Cabe destacar que uno de los indicadores clave del crecimiento del mercado es el número de tarjetas de crédito activas. Según investigaciones de diario El Heraldo en 2021, había 765 000 tarjetas de crédito activas en Honduras, que aumentaron a 863 000 en 2022 (El Heraldo, s.f.). Esta tendencia al alza sugiere una mayor adopción de las tarjetas de crédito como método de pago entre los consumidores hondureños.

En términos de comportamiento del consumidor, las compras promedio realizadas por los tarjetahabientes también han experimentado un aumento constante. (rankinglatam/shopify, 2023) afirma que, en 2021, la compra promedio por tarjeta fue de USD 44, que aumentó a USD 47 en 2022. Los consumidores se sienten más cómodos y seguros al usar tarjetas de crédito para sus transacciones diarias ya sea en terminales de venta o mediante eWallets o Apps. El número de transacciones realizadas con tarjetas de crédito ha experimentado un aumento sustancial. “En 2021, hubo 34 millones de transacciones de compra con tarjetas de crédito, y esta cifra aumentó a 40 millones en 2022”.

Los instrumentos financieros otorgados durante el mes de enero 2023 fueron 19,919, es decir, 6,718 tarjetas más que el mismo período del año pasado. “Las instituciones bancarias tienen a disposición para sus clientes 244 tipos tarjetas de crédito” (El Heraldo, s.f.).

Al menos dos instrumentos de crédito por persona es la relación de estos medios de pago que circulan entre los tarjetahabientes (El Heraldo, s.f.).

En la actualidad las cooperativas de ahorro y crédito también participan en la emisión de tarjetas de crédito a nivel nacional.

2.2 CONCEPTUALIZACIÓN

2.2.1 MACHINE LEARNING

Machine Learning también llamado aprendizaje automático o ML (por sus siglas en inglés), es una rama de la inteligencia artificial que tiene por objetivo generalizar comportamientos a partir de información en forma de ejemplos.

Es un subcampo de la inteligencia artificial que utiliza algoritmos entrenados en conjuntos de datos para crear modelos que permiten a las máquinas realizar tareas que de otro modo solo serían posibles para los humanos, como categorizar imágenes, analizar datos o predecir fluctuaciones de precios. “ML se refiere al uso general de algoritmos y datos para crear máquinas autónomas o semiautónomas” (What Is Machine Learning?, 2023).

ML es una de las formas más comunes de inteligencia artificial y a menudo alimenta muchos de los bienes y servicios digitales que utilizamos todos los días desde aplicaciones móviles hasta programas computacionales.

Machine learning es una rama de la inteligencia artificial (IA) y la informática que se centra en el uso de datos y algoritmos para imitar la forma en la que aprenden los seres humanos, con una mejora gradual de su precisión.

Machine learning es un componente importante del creciente campo de la ciencia de datos. Mediante el uso de métodos estadísticos, los algoritmos se entrenan para hacer clasificaciones o predicciones, y descubrir información clave dentro de los proyectos de minería de datos (¿Qué es machine learning? | IBM, s. f.).

Dicha información clave facilita posteriormente la toma de decisiones dentro de las aplicaciones y las empresas, lo que impacta de forma eficiente a las métricas de crecimiento clave.

2.2.2 APRENDIZAJE SUPERVISADO

El aprendizaje supervisado es un método de aprendizaje en cual se utilizan una serie de datos etiquetados o modelos que queremos que la máquina o el sistema aprenda. Para luego dar una respuesta acertada. Es una parte del campo de aplicación de la inteligencia artificial, Es decir, que en el aprendizaje supervisado se etiqueta un conjunto de datos de entrada que sirven para que la máquina obtenga claves y las aprenda para dar una respuesta correcta (Quiroa, s. f.).

El aprendizaje supervisado es una subcategoría del ML. Se define por su uso de conjuntos de datos etiquetados para entrenar algoritmos que clasifican datos o predicen resultados con precisión. Los algoritmos de aprendizaje supervisado utilizan un conjunto de entrenamiento para enseñar a los modelos a producir el resultado deseado. Este conjunto de entrenamiento incluye entradas y salidas correctas, lo que permite que el modelo aprenda con el tiempo (What Is Machine Learning?, 2023).

Para comenzar, el aprendizaje supervisado sigue un proceso de entrenamiento, que va enseñando al modelo a que logre producir el resultado esperado.

La serie de datos contiene datos de entrada y datos de salida con los que se obtienen los resultados que son correctos. Por esa razón, es que el modelo aprende con el paso del tiempo.

Efectivamente, “el algoritmo logra establecer el nivel de precisión por medio de la función de pérdida. Lo que le permite ajustarse hasta que el error vaya disminuyendo suficientemente” (Quiroa, s. f.).

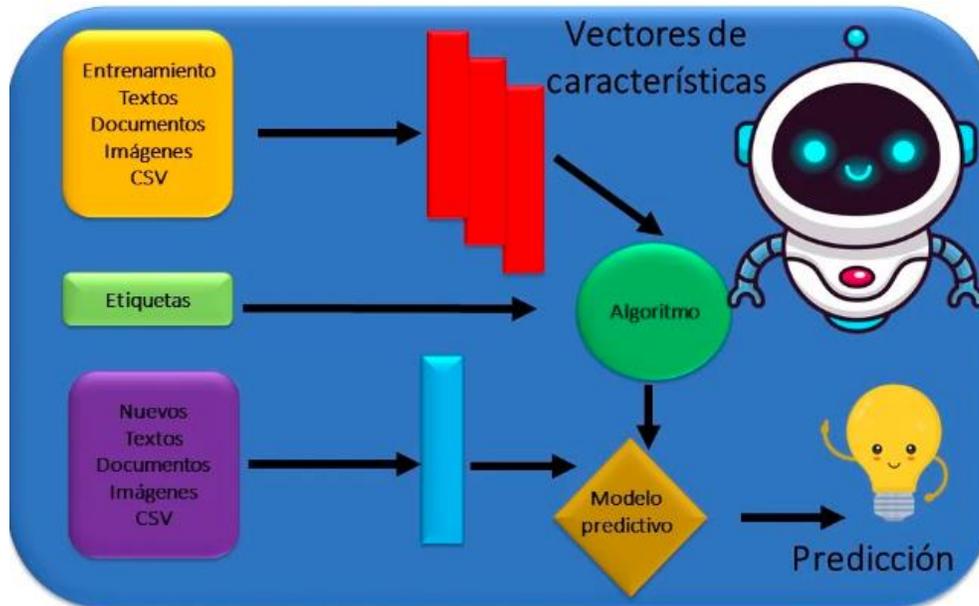


Figura 7. Funcionamiento del aprendizaje supervisado.

Fuente: (Economipedia, 2023)

Existen diferentes tipos de aprendizaje supervisado, aunque los más comunes son aquellos por regresión y por clasificación.

2.2.3 APRENDIZAJE SUPERVISADO POR REGRESIÓN

La regresión se utiliza para entender la relación entre las variables dependientes e independientes. Se utiliza comúnmente para hacer proyecciones, como por ejemplo para las ventas de ingresos para un negocio dado. Los algoritmos de regresión producen un solo valor de salida probabilístico que se determina en función del grado de correlación entre las variables de entrada (¿Qué es machine learning? | IBM, s. f.) (Naveen, 2020)

En los problemas de regresión se deben identificar las variables independientes con los cuales se entrenará el modelo solo que en este caso la variable dependiente (predicción) es discreta (Buitrago, 2020).

En palabras sencillas un modelo de regresión intenta explicar el comportamiento de una variable dependiente en relación con otras variables independientes (Buitrago, 2020).

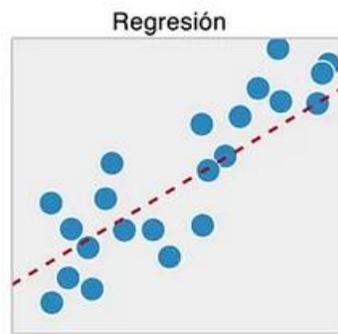


Figura 8. Modelo de Regresión.

Fuente: (Medium, Buitrago, 2020)

En los problemas de regresión los valores que se predicen son valores continuos. Se debe identificar cuáles son las variables independientes, para establecer una fórmula matemática, que se encargará de asociar el valor de la variable dependiente con las variables independientes. Esta variable dependiente es el valor a predecir.

2.2.4 APRENDIZAJE SUPERVISADO POR CLASIFICACIÓN

En los problemas de clasificación el resultado es una categoría o una clase. Lo más habitual es que se clasifique solamente entre dos posibles clases, comúnmente: Verdadero o Falso. Se deben identificar las variables independientes con los cuales se entrenará el modelo (Buitrago, 2020).

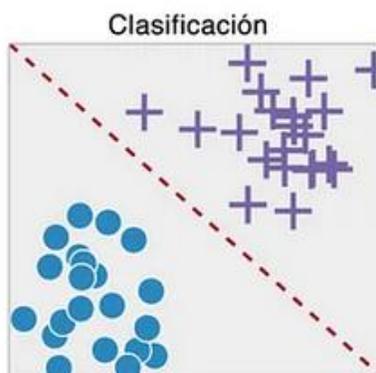


Figura 9. Modelo de clasificación.

Fuente: (Medium, Buitrago ,2020)

La clasificación en el aprendizaje supervisado se utiliza para identificar etiquetas o grupos. Esta técnica se utiliza cuando los datos de entrada pueden ser segregados en categorías o pueden ser etiquetados. Si tenemos un algoritmo que se supone que debe etiquetar ‘masculino’ o ‘femenino’, ‘gatos’ o ‘perros’, etc., podemos utilizar la técnica de clasificación (Naveen, 2020).

Se usa una matriz de confusión para entender y evaluar el resultado del modelo de clasificación. Es una forma de comparar los resultados predichos con los resultados esperados.

2.2.5 DIFERENCIA ENTRE ALGORITMOS DE REGRESIÓN Y DE CLASIFICACIÓN

La diferencia entre algoritmos de regresión y clasificación recae en los resultados o las predicciones que cada modelo genera, mientras que la regresión obtiene valores continuos es decir porcentajes o valores numéricos, la clasificación por otro lado obtiene una determinada respuesta entre dos variables (combinaciones binarias). (Difference Between Classification and Regression In Machine Learning - Dataaspirant, 2014)

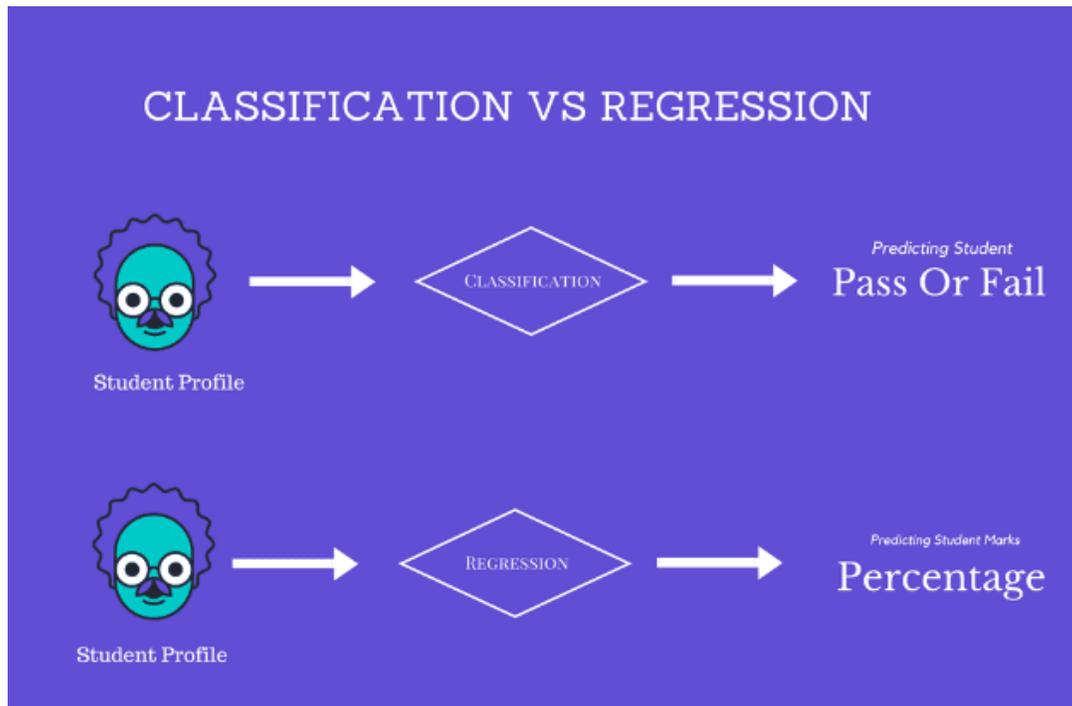


Figura 10. Diferencia entre Modelos de Clasificación y Regresión.

Fuente: (Data aspirant, 2014)

2.2.6 MODELO DE APRENDIZAJE POR REGRESIÓN LINEAL

“Un algoritmo de regresión lineal o regresión logística se utiliza para realizar estimaciones y tratar de comprender las relaciones entre diferentes variables”. Mide la relación entre la variable dependiente y una o varias variables independientes estimando las probabilidades utilizando una función logística, que sigue una distribución logística acumulativa.

Así, una de las principales funciones de este tipo de algoritmos consiste en poder realizar pronósticos o predicciones eficientes mediante Machine Learning (Ramon, 2014).

2.2.7 MODELO DE APRENDIZAJE POR RANDOM FOREST

“La regresión de bosque aleatorio es un algoritmo de aprendizaje supervisado y técnica de ensacado que utiliza un método de aprendizaje de conjunto para la regresión en el aprendizaje automático. Los árboles en los bosques aleatorios se ejecutan en paralelo, lo que significa que no hay interacción entre estos árboles mientras se construyen”.(Random Forest Regression in Python Explained | Built In, s. f.)

Los bosques aleatorios, o Random Forest, son algoritmos de aprendizaje supervisado que utilizan un método de aprendizaje compuesto por una multitud de árboles de decisión (Random Forest, s. f.). Cada árbol en el bosque aleatorio depende de los valores de un vector aleatorio muestreado de forma independiente y con la misma distribución para todos los árboles en el bosque (Cutler et al., 2011).

El error de generalización para los bosques converge a un límite a medida que el número de árboles en el bosque se vuelve grande. El error de generalización de un bosque de clasificadores de árboles depende de la fuerza de los árboles individuales en el bosque y de la correlación entre ellos (Cutler et al., 2011). Esto significa que los bosques aleatorios se vuelven más precisos a medida que se agregan más árboles

En palabras sencillas, los bosques aleatorios funcionan de la siguiente manera:

- Se selecciona un conjunto de datos de entrenamiento.
- Se construye un árbol de decisión utilizando ese conjunto de datos.
- Se repite el proceso anterior varias veces, creando un bosque de árboles de decisión.

- Para hacer una predicción, el bosque aleatorio consulta a todos los árboles de decisión y luego toma la decisión que la mayoría de los árboles apoyan.

Los bosques aleatorios tienen una serie de ventajas, entre ellas:

- Son relativamente fáciles de implementar y entrenar.
- Son muy precisos, incluso con conjuntos de datos pequeños o ruidosos.
- Son robustos al sesgo y al sobreajuste.

Los bosques aleatorios se pueden utilizar para una variedad de tareas de aprendizaje automático, incluyendo:

- Clasificación
- Regresión
- Clasificadores conjuntos
- Análisis de conglomerados

Los bosques aleatorios se basan en un método de aprendizaje compuesto. Esto significa que los árboles de decisión individuales se construyen de forma independiente, pero luego se combinan para crear una predicción final. (Cutler et al., 2011).

El bosque aleatorio es una técnica utilizada en la modelización de predicciones y análisis de comportamiento y se construye sobre árboles de decisión (What Is a Random Forest?, s. f.)

En resumen, los bosques aleatorios son una técnica de aprendizaje automático poderosa y versátil que puede utilizarse para una variedad de tareas.

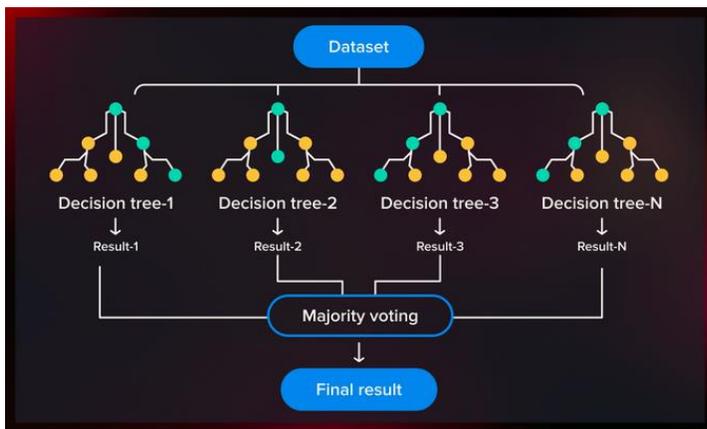


Figura 11. Estructura de un Random Forest.

Fuente: (Medium, Bexapin Khakim ,2022)

2.2.8 PRONÓSTICOS

Los pronósticos son predicciones de algo que puede suceder en un futuro y este puede ser realizado de distintas maneras de manera general los métodos de pronósticos se basan en análisis de datos históricos o por juicio.

“los pronósticos de corto plazo son necesarios para prácticamente todos los departamentos de una organización, sin embargo, se usan más frecuentemente en las áreas de mercadotecnia, producción y administración de la cadena de suministro” (Santana, 2014)

“Pueden implicar tomar datos históricos y proyectarlos hacia el futuro con alguna clase de modelo matemático, también puede ser una predicción subjetiva o intuitiva, o puede implicar una combinación de estas, es decir, un modelo matemático ajustado por el buen juicio de un administrador.” (Santana, 2014)

El pronóstico es una técnica que utiliza datos históricos como entradas para hacer estimaciones informadas que son predictivas en la determinación de la dirección de las tendencias futuras.

“Las empresas utilizan el pronóstico para determinar cómo asignar sus presupuestos o planificar los gastos anticipados para un período de tiempo futuro”. Esto se basa generalmente en la demanda proyectada de los bienes y servicios ofrecidos (What Is Machine Learning?, 2023).

2.2.9 CRÉDITO

El crédito es la capacidad de pedir prestado dinero con la promesa de devolverlo, a menudo con intereses (The Definition of Credit and Why You Need It - NerdWallet, s. f.). En términos más generales, el crédito se refiere a cualquier combinación de recursos financieros, como acciones, bonos y efectivo que se ponen a la disposición de una persona natural o jurídica para que los adquiera con la promesa de pagarlos más tarde pasado un periodo de tiempo convenido entre las partes y generalmente con la adición de intereses. Los créditos pueden ser mantenidos por inversores individuales o gestionados por profesionales financieros, fondos de cobertura, bancos y otras instituciones financieras.

Tipos de crédito

Existen muchos tipos de crédito, pero los dos más comunes son el crédito revolving y el crédito a plazos.

- Crédito revolving:

El crédito revolving es un tipo de crédito, generalmente emitido en forma de tarjeta de crédito, donde los usuarios reciben un límite de crédito, pero pueden gastar tanto o tan poco hasta ese monto como lo deseen. Los saldos se pagan en su totalidad o en parte cada mes, y cualquier saldo restante se transfiere (o se renueva) al mes siguiente. “Las tarjetas de crédito son diferentes de las tarjetas de cargo, otro tipo de crédito, donde el saldo debe pagarse en su totalidad cada mes” (The Definition of Credit and Why You Need It - NerdWallet, s. f.).

- Crédito a plazos:

El crédito a plazos es un tipo de crédito, generalmente emitido en forma de préstamo, que los prestatarios pagan en incrementos constantes a lo largo del tiempo (The Definition of Credit and Why You Need It - NerdWallet, s. f.). Ejemplos de crédito a plazos incluyen préstamos estudiantiles, préstamos para automóviles e hipotecas.

- Crédito por servicios:

El crédito por servicios es un tipo de crédito que describe los contratos que celebra con muchos proveedores de servicios, como empresas de servicios públicos y servicios de membresía. Estas empresas brindan el servicio y usted firma un contrato para pagarles después del hecho (The Definition of Credit and Why You Need It - NerdWallet, s. f.).

2.2.10 TARJETA DE CRÉDITO

Una tarjeta de crédito es esencialmente un medio para solicitar dinero prestado que viene acompañado de intereses y, a veces, tarifas adicionales por el uso del producto financiero.

También es una línea de crédito renovable, lo que significa que puede pedir prestado dinero repetidamente en una cuenta hasta un límite establecido. Antes de solicitar una tarjeta de crédito, primero debe considerar las ventajas y desventajas de usar una (Understanding Credit - Financial Aid & Scholarships, s. f.).

- Ventajas:
 - Uso para emergencias: Las tarjetas de crédito pueden ser útiles para cubrir gastos inesperados, como reparaciones en el hogar o facturas médicas.
 - Compra ahora, paga después: Las tarjetas de crédito le permiten realizar compras y pagarlas más tarde, puede ser útil para presupuestar y financiar compras grandes.
 - Protección de compras: Muchas tarjetas de crédito ofrecen protección contra compras fraudulentas y daños o robo de artículos comprados con la tarjeta.
 - Ayuda a establecer un buen crédito: Hacer pagos a tiempo y mantener un saldo bajo en su tarjeta de crédito puede ayudar a establecer un buen historial crediticio.

- Desventajas
 - Uso excesivo: Es fácil gastar de más con una tarjeta de crédito, lo que puede llevar a deudas y problemas financieros.
 - Intereses altos/tarifas anuales: Las tarjetas de crédito generalmente cobran altas tasas de interés sobre los saldos no pagados, y también pueden cobrar tarifas anuales y otras tarifas.
 - Aumentar su deuda: Usar una tarjeta de crédito para financiar compras puede aumentar su deuda general, lo que puede ser difícil de pagar.
 - Establecer un mal crédito si no se usa sabiamente: Hacer pagos atrasados o exceder su límite de crédito puede dañar su historial crediticio.

2.2.11 RIESGO CREDITICIO

El riesgo crediticio se define como “la posibilidad de que un prestatario no devuelva una deuda según lo acordado” (Credit Risk Definition - Risk.Net, s. f.). Este riesgo es inherente a la actividad de préstamo y puede resultar en pérdidas para el prestamista. Las pérdidas pueden incluir el principal y los intereses perdidos, la interrupción de los flujos de efectivo y el aumento de los costos de cobro (Burrows, 2023).

El riesgo crediticio es uno de los tres principales tipos de riesgo que enfrentan los bancos, junto con el riesgo de mercado y el riesgo operacional (Zamore et al., 2018). En la mayoría de los bancos, el riesgo crediticio representa la mayor parte de los activos ponderados por riesgo.

En general, cuanto mayor es el riesgo crediticio, mayor será la tasa de interés que se le pedirá al deudor que pague sobre la deuda (Burrows, 2023).

2.2.12 GESTIÓN DEL RIESGO CREDITICIO

Las instituciones financieras pueden utilizar una variedad de estrategias para gestionar el riesgo crediticio, incluyendo la realización de controles de crédito en los prestatarios potenciales, la exigencia de que los prestatarios tomen un seguro apropiado, la búsqueda de garantías sobre algunos activos del prestatario o una garantía de un tercero (Burrows, 2023).

Sin embargo, es importante tener en cuenta que los modelos de riesgo crediticio no son perfectos y pueden tener limitaciones. Por ejemplo, pueden no ser capaces de capturar completamente la complejidad del comportamiento humano, y pueden ser sensibles a la calidad y cantidad de los datos de entrenamiento disponibles (ResearchGate Link, s. f.).

2.2.13 CARTERA

En el ámbito financiero, la cartera (o “financiamiento” como se conoce internacionalmente) “se refiere al conjunto de activos y pasivos que posee una persona o entidad. La cartera de crédito es el conjunto de créditos que una institución financiera ha otorgado a sus clientes” (Financiamiento, s. f.).

2.2.14 TIPOS DE CARTERA

- Cartera de crédito:

La cartera de crédito es un activo importante para las instituciones financieras, ya que genera ingresos a través de los intereses que cobra a sus clientes. Sin embargo, también representa un riesgo, ya que los clientes pueden incumplir con sus pagos (Financiamiento, s. f.).

- Cartera en mora

La cartera en mora es el conjunto de créditos que los clientes de una institución financiera no han pagado a tiempo. La cartera en mora representa un riesgo para las instituciones financieras, ya que puede generar pérdidas (Financiamiento, s. f.).

“Las instituciones financieras utilizan una serie de medidas para reducir la cartera en mora, como la cobranza temprana y la refinanciación de los créditos” (BCH, 2022).

2.2.15 INSTITUCIÓN BANCARIA COMERCIAL

Una institución bancaria comercial “es una empresa que se dedica a la intermediación financiera, es decir, que se encarga de transferir recursos financieros entre personas, empresas y gobiernos” (Financial Portfolio, s. f.).

Las instituciones bancarias comerciales ofrecen una variedad de productos y servicios financieros, como cuentas de ahorro, cuentas corrientes, préstamos, tarjetas de crédito y seguros (Financial Portfolio, s. f.) (Understanding Credit - Financial Aid & Scholarships, s. f.).

Tabla 1-Diferencia entre banca comercial y banca de desarrollo.

Característica	Banco comercial	Banco de desarrollo
Propietario	Privado o público	Público
Objetivo	Generar ganancias	Promover el desarrollo económico
Clientes	Personas y empresas en general	Sectores específicos de la economía
Productos y servicios	Cuentas de ahorro, cuentas corrientes, préstamos, tarjetas de crédito, seguros, etc.	Préstamos a largo plazo, financiamiento de proyectos, garantías, etc.
Regulación	Regulada por la Comisión Nacional de Bancos y Seguros (CNBS)	Regulada por Ministerios nacionales e internacionales y regulación de la materia

Fuente:(Elaboración Propia)

2.3 TEORÍAS DE SUSTENTO

2.3.1 BASES TEÓRICAS

El mercado de tarjetas de crédito está influenciado por una serie de factores intrínsecos del mismo, por lo cual existen algunas teorías que pretenden explicar el comportamiento de dicho mercado. Entre estas teorías se encuentran las teorías de la demanda y oferta de crédito, la teoría del riesgo y rendimiento, y la teoría del ciclo económico y ciclo de crédito.

2.3.2 TEORÍA DE LA DEMANDA Y OFERTA DE CRÉDITO

Esta teoría proporciona un marco para comprender cómo los cambios en la disponibilidad de crédito pueden influir en el comportamiento de los consumidores y las decisiones de endeudamiento. En el contexto del mercado de tarjetas de crédito, se espera que variaciones en la colocación de tarjetas de crédito por parte de los bancos tengan un impacto en la demanda de préstamos revolventes y, por ende, en la composición de la cartera crediticia (Benavides, 2000).

Según esta teoría, la demanda de crédito está determinada por factores como los ingresos de los consumidores, el tipo de interés y la disponibilidad de crédito. La oferta de crédito, por su parte, está determinada por factores como la capacidad de los bancos para otorgar préstamos y las políticas regulatorias de los entes reguladores (Mankiw, 2017).

De acuerdo con información tomada de una publicación del banco BBVA (BBVA, 2015), Esta teoría es la conjunción de dos leyes económicas:

La ley de la oferta, que indica que la oferta es directamente proporcional al precio; cuanto más alto sea el precio del producto, más unidades se ofrecerán a la venta.

La ley de la demanda, que indica que la demanda es inversamente proporcional al precio; cuanto más alto sea el precio, menos demandarán los consumidores.

En el crédito se aplica igual, entre mayor sean los intereses que genere un crédito, menor será la demanda de los consumidores hacia dicho producto.

En conclusión, esta teoría sugiere que la oferta y la demanda de crédito están determinadas por factores directamente relacionados con los consumidores como lo son sus ingresos, y la institución que otorga el crédito como lo es el tipo de interés y la disponibilidad de crédito.

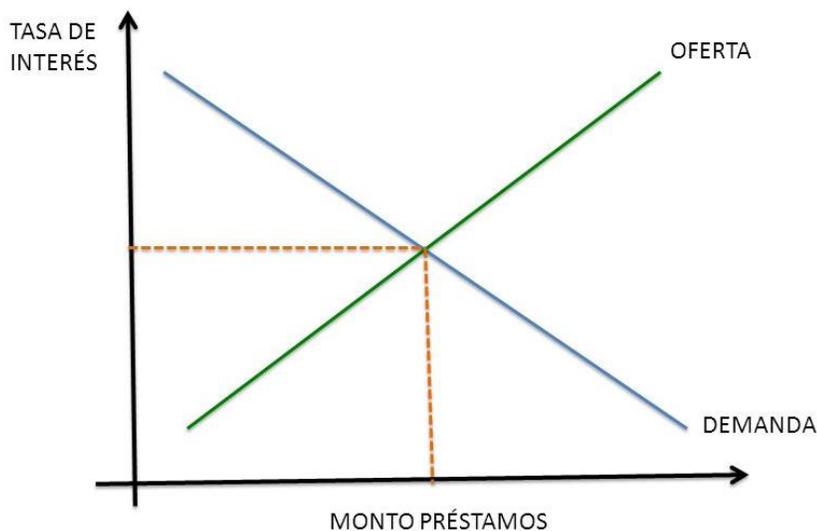


Figura 12. Representación gráfica de la teoría de oferta y demanda de crédito.

Fuente: (Adriana Chacón,2020)

2.3.3 TEORÍA DEL CICLO ECONÓMICO Y CICLO DE CRÉDITO

Según esta teoría, los ciclos económicos pueden influir en la demanda de crédito y la calidad de la cartera crediticia. “Durante períodos de expansión económica, es probable que la demanda de tarjetas de crédito aumente, lo que podría impactar la composición de la cartera crediticia y su variación” (Economic Research, s. f.)

Durante períodos de recesión económica, es probable que la demanda de tarjetas de crédito disminuya, lo que podría conducir a una cartera crediticia más deteriorada.

La evidencia empírica respalda las teorías mencionadas anteriormente. Por ejemplo, un estudio de la Reserva Federal de San Francisco encontró que “un aumento en la oferta de crédito está asociado con un aumento en la demanda de crédito”. (Economic Research, s. f.)

Por último, un estudio del Fondo Monetario Internacional encontró que “los ciclos económicos pueden influir en la demanda de crédito y la calidad de la cartera crediticia” (International Monetary Fund, s. f.).

2.3.4 TEORÍA DEL RIESGO Y RENDIMIENTO

Esta teoría sugiere que los bancos buscan un equilibrio entre la búsqueda de rendimiento y la gestión del riesgo al decidir a quién otorgar préstamos y en qué condiciones. Se espera que esta dinámica influya en la distribución de tarjetas de crédito y, por lo tanto, en la variación positiva o negativa de la cartera crediticia (Mayer, 2007)

Según esta teoría, los bancos otorgan préstamos a los clientes que tienen un mayor riesgo de incumplimiento, pero que también tienen un mayor potencial de rendimiento. Los bancos también pueden otorgar préstamos a clientes con un menor riesgo de incumplimiento, pero que ofrecen un menor rendimiento (Surti - Riesgo y rendimiento La búsqueda de rentabilidad.pdf, s. f.)

El riesgo se puede definir como “la probabilidad de que ocurra algo que impacte de forma negativa en la inversión realizada, mientras que el rendimiento es la capacidad que posee la inversión de generar beneficios económicos” (Guanajuato, 2021). Ambos términos están relacionados con la tasa de interés, que es el precio del dinero.

Las tasas de interés pueden ser activas o pasivas, dependiendo de si se refieren al precio del dinero prestado o invertido. Las tasas de interés influyen en el rendimiento de las inversiones, ya que determinan el costo de oportunidad y el nivel de riesgo (Guanajuato, 2021).

- La tasa activa es el precio del dinero que sale del banco.
- Por otra parte, la tasa pasiva es el precio del dinero que entra al banco.

La tasa activa suele ser mayor que la tasa pasiva, y la diferencia entre ambas tasas se conoce como margen de intermediación.

Tabla 2-Diferencias entre tasa activa y tasa pasiva.

Característica	Tasa activa	Tasa pasiva
Definición	Tasa de interés que cobran los bancos a los clientes por los préstamos que les otorgan.	Tasa de interés que pagan los bancos a los clientes por los depósitos que les hacen.
Expresión	Porcentaje del monto del préstamo.	Porcentaje del monto del depósito.
Aplicación	Saldo no pagado a final de mes.	Saldo del depósito durante el período de tiempo especificado.
Fuente	Ingresos.	Costos.
Relación	Mayor que la tasa pasiva.	Menor que la tasa activa.
Margen de intermediación	Diferencia entre ambas tasas.	Fuente de ingresos principal de los bancos.
Efectos en la economía	Pueden influir en el comportamiento de los consumidores y las empresas, así como en el ahorro y la inversión.	

Fuente: (Elaboración propia)

Se podría concluir que las teorías de la demanda y oferta de crédito, la teoría del riesgo y rendimiento, y la teoría del ciclo económico y ciclo de crédito proporcionan un marco útil para comprender el comportamiento del mercado de los créditos, incluyendo el tipo de crédito revolvente que son las tarjetas de crédito; como se evidencia en algunos estudios realizados por instituciones de prestigio y citados en esta investigación la evidencia empírica respalda estas teorías, lo que sugiere que son útiles para explicar los cambios en el mercado de tarjetas de crédito.

2.3.5 OTROS ESTUDIOS CONSULTADOS

2.3.5.1 Modelos predictores de la morosidad con variables macroeconómicas

El artículo presenta una propuesta que analiza la anticipación del riesgo de morosidad en los préstamos con problemas utilizando modelos ARIMA que identifican los indicadores macroeconómicos asociados a la morosidad de la cartera en cada segmento de destino del crédito ecuatoriano.

El objetivo es contribuir a la construcción de sistemas capaces de anticipar los pagos por defecto y evitar las crisis financieras que afectan a la estabilidad y el desarrollo económico.

El método consiste en aplicar modelos ARIMA estacionales y no estacionales a las series de tiempo de la morosidad de cada segmento de crédito, utilizando variables macroeconómicas como posibles factores explicativos.

Los resultados muestran que cinco de los nueve modelos generados fueron validados con una anticipación de al menos doce meses en el período de estudio, mientras que los otros cuatro modelos fueron descartados por tener bajos coeficientes de determinación o significancia estadística (Guillén-Franco & Peñafiel-Chang, 2018).

2.3.5.2 Determinantes de la demanda de crédito. Una estimación con un modelo mensual de series de tiempo para Venezuela

El artículo trata sobre los determinantes de la demanda de crédito en Venezuela, usando un modelo de series de tiempo mensuales para el periodo 1980-2000.

El autor analiza el impacto de variables como el ingreso real, la tasa de interés, la inflación, el tipo de cambio y el déficit fiscal sobre la demanda de crédito de los sectores privado y público.

El autor encuentra que la demanda de crédito es positivamente sensible al ingreso real y al tipo de cambio, y negativamente sensible a la tasa de interés y a la inflación. Además, el autor detecta una sustitución entre el crédito público y el privado, lo que implica que el déficit fiscal desplaza la demanda de crédito del sector privado.

El autor concluye que la política monetaria tiene un efecto limitado sobre la demanda de crédito, debido a la rigidez de la oferta de dinero y a la inestabilidad de la demanda de dinero.

Por tanto, el autor sugiere que se requieren reformas estructurales para mejorar el funcionamiento del mercado de crédito y estimular el crecimiento económico (Vera - Determinantes de la demanda de crédito. Una estima.pdf, s. f.).

En algunas economías es ahí donde entran en juego productos como las tarjetas de crédito y explica el porque es un medio financiero tan demandado en América Latina como lo hemos visto en el estudio Macroeconómico de esta investigación, así como por qué la cartera de crédito en Mora tiene una tendencia alcista en la región.

2.4 METODOLOGÍAS DESARROLLADAS EN LOS ESTUDIOS CONSULTADOS

Los estudios consultados utilizan diferentes metodologías para analizar el comportamiento del mercado de tarjetas de crédito y su relación con la economía.

A continuación, se presenta un resumen de las metodologías más destacadas:

2.4.1 Determinantes de la demanda de crédito. Una estimación con un modelo mensual de series de tiempo para Venezuela

Este estudio utiliza una metodología cuantitativa, basada en el análisis de series de tiempo. El autor utiliza un modelo de series de tiempo mensuales para estimar el impacto de variables económicas sobre la demanda de crédito en Venezuela.

2.5 METODOLOGÍAS A DESARROLLAR EN EL PRESENTE ESTUDIO

2.5.1.1 Modelo Estadístico

Se implementarán modelos que permitan evaluar la influencia de la Cantidad de tarjetas de crédito en circulación en relación con la Cartera en Mora de los bancos en Honduras. En este modelo, la variable independiente será la cantidad de tarjetas en circulación de forma mensual, la cantidad de tarjetas colocadas por banco, el periodo y la institución bancaria; La variable dependiente sería el comportamiento de la Cartera en Mora de tarjetas de crédito.

2.5.1.2 Machine Learning

Se emplearán técnicas de machine learning para desarrollar un modelo de aprendizaje automático que permita predecir el Valor de la Cartera en Mora en función de la cantidad de Tarjetas en Circulación y la cantidad de tarjetas colocadas por periodo. Los algoritmos a utilizar serán regresión lineal y Random Forest Regression.

2.5.1.3 Segmentación de Datos

Se segmentarán los datos por bancos. Esto permitirá analizar la influencia de la colocación de tarjetas en diferentes periodos de tiempo sin tener un sesgo tomando en cuenta el tamaño de la cartera de cada institución y su posicionamiento en el mercado.

Se realizarán análisis específicos para cada segmento, lo que te permitirá obtener información detallada sobre cómo la colocación de tarjetas de crédito afecta a cada institución.

2.5.1.4 Variables e Indicadores

- a. Cartera en mora real de tarjetas de crédito (Valor monetario en HNL)
- b. Cantidad de tarjetas de crédito en circulación por institución (número entero)
- c. Cantidad de colocación de tarjetas de crédito por periodo (número entero).
- d. Predicción de la cartera en mora por institución (Valor monetario en HNL).

Se utilizarán datos mensuales para evaluar el comportamiento de la Cartera en Mora y la Cantidad de tarjetas en circulación en un contexto temporal.

2.5.1.5 Procesos de Limpieza y Preprocesamiento:

- a. Eliminación de datos faltantes.
- b. Estandarización de variables
- c. Detección y tratamiento de valores atípicos.

2.6 INSTRUMENTOS UTILIZADOS

- Fuentes secundarias: artículos de revistas académicas, informes de organismos internacionales y estadísticas oficiales.
- Modelos de regresión Lineal: Para realizar la comprobación estadística de la relación que existe entre la cantidad de tarjetas de crédito activas en circulación y la cartera en mora de tarjetas de crédito.
- Modelos de Random forest: Al ser un algoritmo robusto ante la aparición de valores atípicos o la utilización de dataset con registros limitados se utilizará el mismo para desarrollar el modelo de aprendizaje automático para la predicción de la cartera en mora. equilibrio general: para analizar la relación entre la desigualdad del ingreso y el acceso al crédito.
- Herramientas de limpieza y visualización de datos.

Es importante señalar que los instrumentos utilizados en cada estudio dependen de los objetivos y el enfoque metodológico del estudio. Por ejemplo, los estudios que se centran en el análisis empírico utilizan principalmente modelos cuantitativos, mientras que los estudios que se centran en el análisis cualitativo utilizan principalmente fuentes secundarias y análisis de textos.

El presente estudio continuará la línea de análisis empírico por lo que se utilizará modelos de aprendizaje automático que proporcionen resultados con valores numéricos medibles estadísticamente.

2.7 MARCO LEGAL

2.7.1 Reglamento de la Ley de Tarjetas de Crédito

La Ley de Tarjetas de Crédito de Honduras, Decreto Legislativo No. 106-2006 («Leyes y Normativas», s. f.), establece las normas y regulaciones para la emisión y uso de tarjetas de crédito en el país.

La ley define una tarjeta de crédito como "un instrumento que, de conformidad con el contrato suscrito con el emisor, podrá ser utilizada dentro o fuera del país. De igual forma, las tarjetas emitidas, comercializadas y adquiridas en el extranjero podrán ser utilizadas en el territorio nacional".

La ley establece que la emisión de tarjetas de crédito está sujeta a la autorización del Banco Central de Honduras.

Los emisores de tarjetas de crédito deben cumplir con los requisitos establecidos en la ley, incluyendo la constitución de un fondo de garantía para cubrir los pagos de los tarjetahabientes.

La ley también establece los derechos y obligaciones de los emisores y los tarjetahabientes:

Los emisores están obligados a proporcionar información clara y precisa a los tarjetahabientes sobre los términos y condiciones de los contratos de tarjetas de crédito. Los tarjetahabientes, por su parte, están obligados a pagar sus saldos a tiempo y en su totalidad.

El Reglamento de la Ley de Tarjetas de Crédito, Decreto Ejecutivo No. 39-2018, establece las disposiciones complementarias a la ley. El reglamento establece, entre otras cosas, los requisitos para la autorización de los emisores de tarjetas de crédito, los términos y condiciones de los contratos de tarjetas de crédito, y los procedimientos de resolución de controversias.

La Ley de Tarjetas de Crédito de Honduras fue aprobada en 2006 y ha sido objeto de varias reformas desde entonces. La última reforma se publicó en La Gaceta en 2017 y establece que las tasas de interés en moneda nacional no deben ser mayores al 54%, mientras que en moneda extranjera no deben ser mayores al 30%.

Además, la reforma permite que los hondureños endeudados con tarjetas de crédito puedan hacer arreglos de pago con plazos de hasta sesenta meses.

2.7.2 Ley de Protección al Consumidor

La Ley de Protección al Consumidor en Honduras (ccichonduras.org); tiene como objetivo proteger los derechos de los consumidores y regular las relaciones entre consumidores y proveedores.

Fue aprobada inicialmente mediante el decreto legislativo 41-89 del 7 de abril de 1989. Desde entonces ha experimentado varias reformas que se detallan a continuación:

- Aprobación inicial: La Ley de Protección al Consumidor fue aprobada el 7 de abril de 1989.
- Reforma de 2008: Se promulgó el decreto 24-2008, que actualizó la ley para adaptarla a las necesidades modernas de protección al consumidor.

- **Publicación de 2021:** La ley se publicó el 27 de diciembre de 2021 en el Diario Oficial La Gaceta, con el objetivo de actualizar e incluir reformas que mejoren la protección del consumidor (*Ley de Protección al Consumidor Honduras, ¿para qué sirve y cuándo se aplica? / Tu Nota, s. f.*).
- **Reglamento actualizado:** El Acuerdo No. 84-2021 establece 106 artículos y deroga el acuerdo ejecutivo N. 15-2009, que contenía el antiguo Reglamento de la Ley de Protección al Consumidor (*Ley de Protección al Consumidor Honduras, ¿para qué sirve y cuándo se aplica? / Tu Nota, s. f.*).

La ley actual de protección al consumidor hace alusión directa a diversos tópicos como:

Aplicabilidad: Se aplica a todas las relaciones de consumo dentro del territorio de Honduras, ya sea entre personas naturales o jurídicas, y entre consumidores y proveedores.

Autoridad encargada: La Secretaría de Estado en el Despacho del Desarrollo Económico, a través de la Dirección General de Protección al Consumidor, es la autoridad encargada de la aplicación de la ley.

Derechos del consumidor: La ley actualizada desarrolla con más detalle los derechos del consumidor, como el acceso a productos y servicios, un trato justo, información exacta sobre lo que se consume, y un espacio para expresar quejas.

Prácticas abusivas: Define prácticas consideradas abusivas y violatorias de los derechos del consumidor, como bienes o servicios defectuosos, prácticas comerciales desleales y publicidad engañosa.

Sanciones: Establece sanciones para los proveedores que cometan infracciones, clasificadas como muy graves o graves

Estas reformas reflejan el compromiso de Honduras con la protección de los derechos de los consumidores y la adaptación a los cambios en las prácticas comerciales y las expectativas de los consumidores.

2.7.3 Ante Proyecto de Ley de Protección de Datos Personales

El anteproyecto de ley de protección de datos personales de Honduras (Ley de Protección de Datos Personales.pdf, s. f.) es un documento que establece las bases para la protección de datos personales en el país. Fue elaborado con el apoyo de la Agencia Española de Cooperación Internacional para el Desarrollo (AECID) y la Unión Europea.

El anteproyecto busca garantizar que Honduras cuente con una legislación actualizada y moderna que esté alineada con los estándares nacionales e internacionales en materia de protección de datos personales.

El anteproyecto incluye principios fundamentales como los derechos de acceso, rectificación, cancelación y oposición (derechos ARCO), así como la acción de Habeas Data, que permite a los individuos controlar su información personal en posesión de terceros. Se enfoca en la importancia de proteger los derechos de personalidad como el nombre, la intimidad, la imagen, el honor y los derechos de autor frente a los avances tecnológicos que pueden amenazarlos.

Este anteproyecto es un paso importante hacia la creación de un marco legal que proteja los datos personales de los ciudadanos hondureños y promueva su derecho a la autodeterminación informativa. El documento incluye y hace hincapié en los siguientes tópicos:

Derechos de Personalidad:

El documento discute la importancia de los derechos personales, como el derecho al nombre propio, la privacidad, la imagen, el honor y los derechos de autor, especialmente en el contexto de los avances tecnológicos que pueden amenazar estos derechos.

Legislación Hondureña:

Señala el reconocimiento y la protección de los derechos personales en la legislación hondureña, incluido el derecho al Habeas Data, que se propuso por primera vez en 2003 pero no se ratificó hasta 2013.

Protección de Datos:

El documento destaca la necesidad de “una norma legal que reconozca expresamente los derechos de los sujetos de datos, incluidos el acceso, la rectificación, la cancelación y la oposición, así como el recurso a una autoridad independiente y especializada con potestad sancionadora” (Ley

de Proteccion de Datos Personales.pdf, s. f.).

El documento presenta un proyecto de ley de Protección de Datos Personales, cuyo objetivo es “dar a las personas control sobre su información personal en manos de terceros y alinearse con los estándares internacionales de protección de datos”.

Este documento aborda la importancia de proteger los derechos personales en la era digital. Destaca la necesidad de una legislación sólida que reconozca y proteja estos derechos, especialmente en un contexto en el que las nuevas tecnologías pueden representar una amenaza. El documento también propone un anteproyecto de ley que busca brindar a las personas control sobre su información personal y garantizar su adecuada protección.

La adopción de una ley integral de protección de datos personales en Honduras sería un paso importante para garantizar la privacidad y la seguridad de la información personal de los ciudadanos. Esta ley podría tener un impacto positivo en la confianza de los consumidores en las empresas y el gobierno, así como en el desarrollo de la economía digital.

2.7.4 Ley de Transparencia y Acceso a la Información Pública

La Ley de Transparencia y Acceso a la Información Pública de Honduras (Ley de Transparencia y Acceso a la Información Pública y su Reglamento, s. f.), establecida por el Decreto Legislativo No. 170 – 2006, es una normativa de gran importancia para el fortalecimiento del Estado de Derecho y la consolidación de la democracia en el país. Esta ley fue publicada en el Diario Oficial La Gaceta el 30 de diciembre de 2006.

La ley tiene como finalidad el desarrollo y ejecución de la política nacional de transparencia, así como garantizar el ejercicio del derecho de toda persona al acceso a la información pública. Esta ley tiene como objetivos los siguientes:

- Garantizar el ejercicio del derecho que tienen los ciudadanos a participar en la gestión de los asuntos públicos.
- Hacer efectiva la transparencia en el ejercicio de las funciones públicas y en las relaciones del Estado con los particulares.
- Garantizar la protección, clasificación y seguridad de la información.

CAPÍTULO III. METODOLOGÍA

3.1 CONGRUENCIA METODOLÓGICA

En palabras de la Universidad Nacional Autónoma de Hidalgo en su espacio online de aplicación del conocimiento, la congruencia metodológica se define como “La presentación en forma de resumen de la investigación, debe ser realizada de forma que permita visualizar los elementos básicos del proyecto de investigación, para comprender y evaluar la coherencia y conexión lógica entre el problema, los objetivos, la hipótesis, las variables y metodología en la investigación a realizar” (Boletín Científico :: UAEH, s. f.).

En el presente capítulo se presentan los elementos que conforman la congruencia metodológica, esto implica que los elementos del proceso de investigación estén alineados y se complementen entre sí de manera coherente y lógica.

También implica que los métodos y técnicas utilizados sean apropiados para responder a los objetivos de la investigación, que los datos recopilados sean relevantes y adecuados para responder a las preguntas de investigación, y que el análisis de los datos esté en línea con los objetivos y la metodología utilizada.

La congruencia metodológica es importante para garantizar la validez y la fiabilidad de los resultados de la investigación. Cuando existe congruencia, se fortalece la credibilidad de los hallazgos y se evita la generación de conclusiones erróneas o sesgadas. Por el contrario, la falta de congruencia puede llevar a resultados confusos, contradictorios o poco confiables.

Como concepto se puede determinar que la congruencia metodológica “es una herramienta que permite analizar e interpretar la operatividad teórica del proyecto, sistematizando el conjunto de elementos como el problema, los objetivos, las variables y la operacionalización de las variables” (Comprendiendo la Matriz de Congruencia, 2023)

En términos más específicos, la congruencia metodológica asegura que haya una alineación lógica entre:

- Las preguntas de investigación planteadas.
- Los objetivos que se buscan alcanzar.

- Las hipótesis formuladas.
- Las variables identificadas y cómo estas se medirán a través de indicadores.

3.2 MATRIZ DE CONGRUENCIA METODOLÓGICA

“Una Matriz de Consistencia Metodológica (MCM): Es un cuadro horizontal, conformado por columnas y filas, que consiste en presentar y resumir en forma adecuada, panorámica y sucinta los elementos básicos del proyecto de investigación, para comprender y evaluar la coherencia y conexión lógica entre el problema, los objetivos, la hipótesis, las variables y metodología en la investigación a realizar” (Boletín Científico :: UAEH, s. f.).

Una matriz metodológica es una herramienta utilizada en la investigación y el desarrollo de proyectos para organizar y estructurar los métodos, técnicas y procedimientos que se utilizarán durante el proceso. Esta matriz proporciona una visión general de las diferentes etapas del proyecto, las actividades a realizar, los recursos necesarios y los plazos correspondientes (Hernández Sampieri et al., 2014).

Es por ello que la matriz metodológica tiene como función el brindar un panorama macro de cómo se relacionan el problema de investigación, los objetivos: general y específicos, así como las preguntas de investigación. También, se detallan con precisión las variables de estudio que han sido elegidas.

A continuación, se presenta la matriz metodológica correspondiente a este proyecto de investigación, con la que, de acuerdo con los conceptos mencionados, se pretende proporcionar una visión generalizada de las diferentes etapas que se abordarán a lo largo de este proyecto:

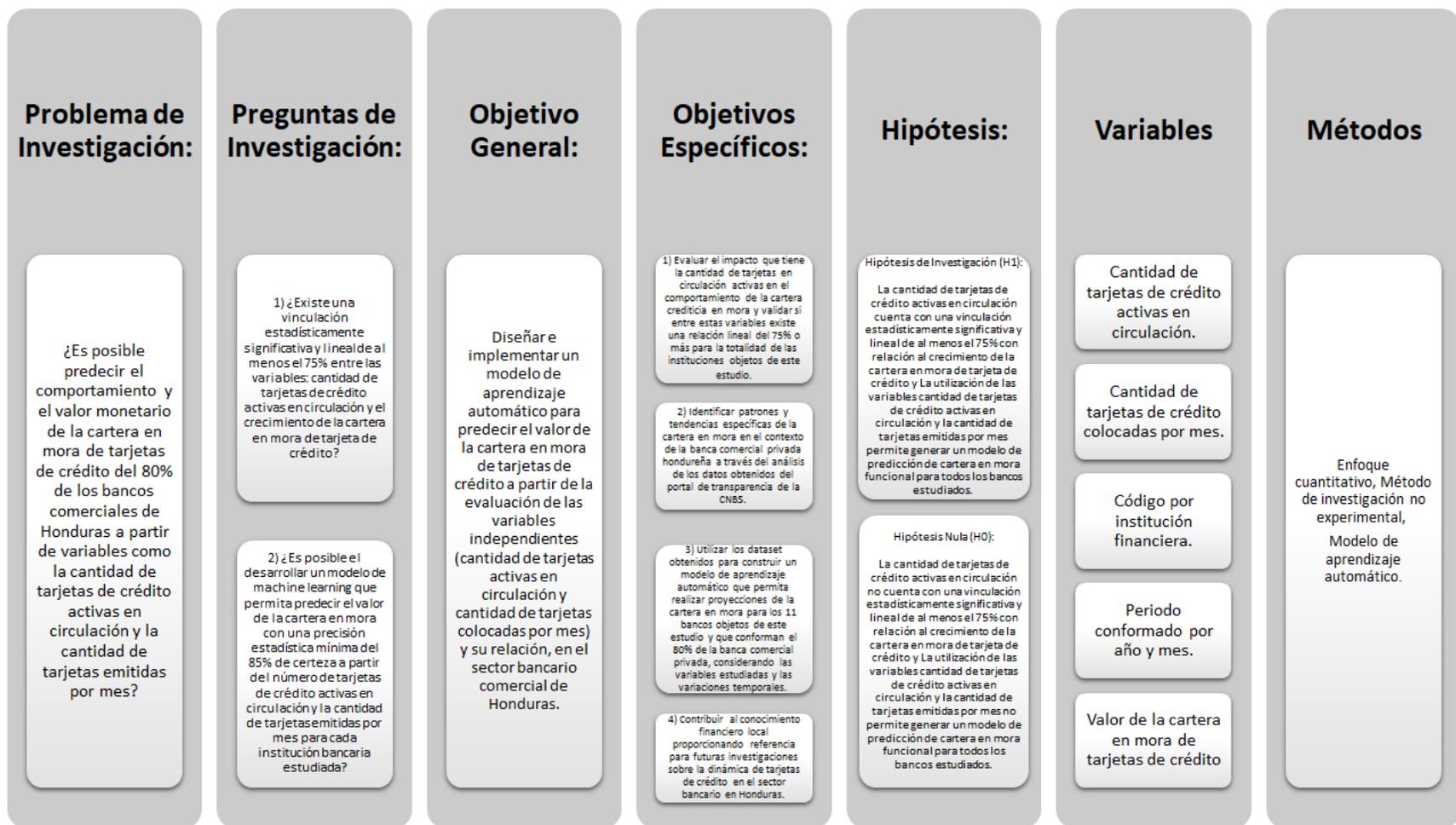


Figura 13. Matriz de congruencia metodológica.

Fuente:(Elaboración Propia)

El problema de investigación establece el tema general de la investigación. En este caso, el problema de investigación es el siguiente: ¿Es posible predecir el comportamiento y el valor monetario de la cartera en mora de tarjetas de crédito de todos los bancos comerciales de Honduras a partir de variables como la cantidad de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes?

Las preguntas de investigación son más específicas y ayudan a delimitar el problema de investigación. En este caso, las preguntas de investigación son las siguientes:

¿Existe una vinculación estadísticamente significativa y lineal de al menos el 75% entre las variables: cantidad de tarjetas de crédito activas en circulación y el crecimiento de la cartera en mora de tarjeta de crédito?

¿Es posible el desarrollar un modelo de machine learning que permita predecir el valor de la cartera en mora con una precisión estadística mínima del 85% de certeza a partir del número de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes para cada institución bancaria estudiada?

El objetivo general de la investigación establece el propósito general del estudio. En este caso, el objetivo general es el siguiente:

Diseñar e implementar un modelo de aprendizaje automático para predecir el valor de la cartera en mora de tarjetas de crédito a partir de la evaluación de las variables independientes (cantidad de tarjetas activas en circulación y cantidad de tarjetas colocadas por mes) y su relación, en el sector bancario comercial de Honduras.

Los objetivos específicos son objetivos más detallados que ayudan a alcanzar el objetivo general.

Las hipótesis son afirmaciones que se quieren comprobar con la investigación. En esta investigación se pretende desarrollar una de las siguientes hipótesis:

Hipótesis de Investigación (H1):

La cantidad de tarjetas de crédito activas en circulación cuenta con una vinculación estadísticamente significativa y lineal de al menos el 75% con relación al crecimiento de la cartera en mora de tarjeta de crédito y La utilización de las variables cantidad de tarjetas de crédito activas

en circulación y la cantidad de tarjetas emitidas por mes permite generar un modelo de predicción de cartera en mora funcional para todos los bancos estudiados.

Hipótesis Nula (H0):

La cantidad de tarjetas de crédito activas en circulación no cuenta con una vinculación estadísticamente significativa y lineal de al menos el 75% con relación al crecimiento de la cartera en mora de tarjeta de crédito y La utilización de las variables cantidad de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes no permite generar un modelo de predicción de cartera en mora funcional para todos los bancos estudiados

Las variables son los conceptos que se van a medir en la investigación. En este caso, las variables son las siguientes:

- Cantidad de tarjetas de crédito activas en circulación.
- Cantidad de tarjetas de crédito colocadas por mes (periodo).
- Valor de cartera en mora de tarjetas de crédito.
- Código por institución financiera.
- Periodo temporal (Mes y año).

3.3 ESQUEMA DE VARIABLES DE ESTUDIO

El Esquema de Variables de Estudio según Sampieri se refiere “a la estructura organizada que presenta las variables que serán examinadas en una investigación”. Este esquema es parte del proceso de diseño de la investigación y ayuda a definir cómo se medirán y analizarán las variables para responder a las preguntas de investigación y probar las hipótesis (Hernandez-Sampieri.-Metodologia-de-la-investigacion.pdf, s. f.).

Podemos concluir que este esquema es crucial para garantizar que la investigación sea sistemática y que los datos recopilados sean relevantes para los objetivos del estudio. “Además, proporciona una guía clara para el análisis estadístico y la interpretación de los resultados” (Hernandez-Sampieri.-Metodologia-de-la-investigacion.pdf, s. f.).

Variable dependiente	Variables independientes	Definición	Medida	Unidad de medida
Valor de la cartera en mora de tarjetas de crédito En moneda LPS, escala real.	Periodo	Fecha de corte del análisis	Fecha	Mes y año
	Institución bancaria	Entidad emisora de la tarjeta de crédito	Texto	Nombre de la institución bancaria
	Cantidad de tarjetas de crédito colocadas	Número total de tarjetas de crédito emitidas por la institución bancaria en el periodo de análisis	Conteo	Número de tarjetas (valor)
	Cantidad de tarjetas en circulación	Número total de tarjetas de crédito activas en el periodo de análisis	Conteo	Número de tarjetas (valor)

Tabla 3. Esquema de variables de estudio.

Fuente: (Elaboración Propia)

3.4 OPERACIONALIZACIÓN DE LAS VARIABLES

La operacionalización de las variables según Sampieri “es el proceso mediante el cual se lleva una variable del nivel abstracto a un plano concreto, haciendo que la variable sea medible y observable en la realidad empírica” (Hernandez-Sampieri.-Metodologia-de-la-investigacion.pdf, s. f.). Este proceso implica definir las variables de manera conceptual y operacional para asegurar que puedan ser evaluadas adecuadamente en el contexto de la investigación

La operacionalización de las variables se refiere a cómo se medirán o cuantificarán las mismas en el proyecto.

Tabla 4-Operacionalización de las variables (Elaboración propia)

Variable	Concepto	Escala	Unidad de medida	Operacionalización
Cantidad promedio de tarjetas de crédito activas en circulación	Es la cantidad promedio de tarjetas de crédito activas en circulación en el periodo de análisis.	Escala de conteo	Número promedio	Se recopilará la información sobre el número total de tarjetas de crédito activas en el periodo de análisis.
Cantidad promedio de tarjetas de crédito colocadas en un mes por institución bancaria	Es la cantidad promedio de tarjetas de crédito colocadas en un mes por institución bancaria.	Escala de conteo	Número	Se recopilará la información sobre el número total de tarjetas de crédito emitidas por institución bancaria en un mes.
Periodo (año mes)	Es el periodo de tiempo en el que se recopilaron los datos.	Escala de tiempo	Año mes	Se recopilará la información sobre el año y mes en el que se recopilaron los datos.
Predicción del valor de la cartera en mora de tarjetas de crédito en un periodo de tiempo determinado	Es el valor estimado del saldo de la cartera en mora de tarjetas de crédito en un periodo de tiempo determinado.	Escala de razón	Número de lempiras	Se calculará mediante un modelo de aprendizaje automático, a partir de las variables anteriores.

Fuente: (Elaboración Propia)

3.5 ENFOQUE Y MÉTODOS

Enfoque: es el nivel en el que se especifican los supuestos y las creencias.

Método: es el nivel en que se pone en práctica la teoría.

Dado que la investigación implica la recopilación y el análisis de datos numéricos, tiene un enfoque cuantitativo. Este enfoque nos permite medir la relación entre las tarjetas de crédito en circulación y el comportamiento de la cartera en mora. Así mismo se utilizará un método no experimental ya que, si bien se pretenden desarrollar varios modelos de aprendizaje automático para explicar tanto el nivel y tipo de relación de las variables, como también, para predecir el valor de la cartera en mora no se llevará al nivel de implementar dichos modelos en un ambiente de producción real por lo que la fase experimental de manipulación de las variables no se llevaría a cabo.

3.6 DISEÑO DE LA INVESTIGACIÓN

El diseño señala al investigador lo que debe hacer para alcanzar sus objetivos de estudio, contestar los interrogantes que se ha planteado y analizar la certeza de la(s) hipótesis formuladas en un contexto en particular (Hernández Sampieri et al., 2014)

El diseño desempeña un rol esencial en toda investigación, por la función que tiene de pautar el camino que nos permitirá cumplir con los propósitos del estudio. Para algunos autores como Kerlinger (1979, citado por Hernández Sampieri et al., 2003) si el diseño está bien concebido, el producto último de un estudio (sus resultados) tendrá mayores posibilidades de ser válido. Asimismo, la precisión de la información obtenida puede variar en función del diseño o estrategia elegida.

En el presente estudio el diseño de la investigación será cuantitativo, con un enfoque no experimental.

- Cuantitativo significa que la investigación se basará en datos numéricos.
- No experimental significa que la investigación no implicará la manipulación de una o más variables para observar el efecto en otra variable.

En este caso, la investigación se centrará en predecir el valor de la cartera en mora de tarjetas de crédito a partir de la cantidad de tarjetas de crédito activas en circulación y la cantidad promedio de tarjetas de crédito colocadas en un mes por institución.

Para ello, se utilizará un modelo de aprendizaje automático para entrenar un modelo que pueda predecir el valor de la cartera en mora de tarjetas de crédito a partir de los datos de las variables mencionadas.

El modelo de aprendizaje automático se entrenará utilizando datos históricos de los bancos comerciales de Honduras a excepción de los bancos Popular y Cuscatlán para los cuales no se obtuvieron los datos suficientes como para poderlos evaluar dentro del grupo de instituciones a estudiar.

Una vez entrenado el modelo, se utilizará para predecir el valor de la cartera en mora de tarjetas de crédito para un periodo de tiempo futuro.

El modelo se evaluará utilizando una muestra de datos de prueba, es decir, el segmento de datos que no se utilizó para entrenar el modelo (en este caso el 20% de los registros obtenidos).

Posteriormente se utilizará el total de los registros para hacer una comparativa entre los valores reales y los valores predichos por el modelo, así como la utilización del mismo en registros aleatorios basados en los registros reales para realizar pruebas del modelo entrenado cargado en un ambiente de pruebas en productivo.

La investigación se espera que contribuya a mejorar la capacidad de los bancos comerciales de Honduras para predecir el valor de la cartera en mora de tarjetas de crédito.

3.7 POBLACIÓN

El universo está conformado por toda la población o conjunto de unidades que se quiere estudiar y que podrían ser observadas individualmente en el estudio (Bravo, 1998, p. 179).

Para Hernández Sampieri, "una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones" (p. 65) (Hernández Sampieri et al., 2014). Es la totalidad del fenómeno a estudiar, donde las entidades de la población poseen una característica común la cual se estudia y da origen a los datos de la investigación.

Para este estudio la población se compone de los datos de tarjeta de crédito de 11 bancos comerciales del sector de la banca comercial de Honduras:

Tabla 5. Población para el estudio.

Datos de Cartera en mora y cantidad de tarjetas de crédito de los siguientes Bancos Comerciales:	
Banco Atlántida (Bancatlan)	Banpaís
Banco Azteca	Banco Ficohsa
Banco Davivienda	Banhcafé
Bac Credomatic	Banco Lafise
Banrural	Banco Promerica
Banco de Occidente	

Fuente: (Elaboración Propia)

3.8 MUESTRA

Andrea Lina Lavalle, Elda Beatriz Micheli y Natalia Rubio en su publicación "Análisis didáctico de regresión y correlación para la enseñanza media" mencionan que la proporción para

realizar el entrenamiento y prueba de un modelo de aprendizaje automático puede oscilar en proporciones 80/20 y 70/30 respectivamente” (Lavallo et al. - Análisis didáctico de regresión y correlación para la enseñanza, s. f.).

Se utilizarán dos sets de datos que contienen el conjunto de registros de todas las tarjetas de crédito colocadas en Honduras para los periodos comprendidos entre diciembre de 2018 a septiembre de 2023, para 11 de los bancos que conforman la banca comercial del país, se deja fuera de este estudio la data correspondiente a los bancos: Banco Cuscatlán (Anteriormente Banco de los Trabajadores) y Banco popular pues no se obtuvo data consistente para estas dos instituciones. En los sets de datos se tienen registros detallados de forma mensual, por tipo de institución, junto con el valor de la cartera crediticia de tarjeta de crédito, cartera en mora, márgenes por utilizar y cantidad de tarjetas en circulación de cada institución.

Para entrenamiento y prueba del algoritmo optaremos por utilizar una proporción del 80/20. 80% de los datos para entrenamiento de los modelos y 20% para la prueba de los mismos.

(Heidegger y la técnica en el mundo moderno, 2021) sostiene que: “la técnica es un modo de comprensión del mundo (como lo son también, por ejemplo, el arte o la política). En ese ámbito abierto que es la técnica, los entes resultan técnicamente descubiertos”.

Selección de datos para entrenamiento y prueba del algoritmo

La selección aleatoria es un método de selección de datos en el que cada elemento de la población tiene la misma probabilidad de ser seleccionado (Arias, s. f.). En esta investigación, se utilizará dicho método para dividir el conjunto de datos en dos muestras, una para entrenamiento y otra para prueba.

El procedimiento para realizar la selección aleatoria 80/20 será ejecutado al momento de realizar la limpieza y transformación de datos utilizando un software especializado en dicho tipo de tareas utilizando un generador de aleatoriedad que tomará una porción de la data (80% en este caso) para ser utilizado como muestra de entrenamiento del modelo y el 20% restante para ser utilizado como muestra de prueba del modelo.

Dicho procedimiento se utilizará para garantizar que las dos muestras sean representativas de la población y que la muestra de entrenamiento sea lo suficientemente grande para que el modelo de aprendizaje automático pueda aprender las relaciones entre las variables.

3.9 FUENTES DE INFORMACIÓN

Las fuentes de datos para esta investigación se obtuvieron mediante los data sets de acceso público que se pueden encontrar en el portal web de la Comisión Nacional de Bancos y Seguros (CNBS) de Honduras, así mismo para el sustento teórico de la misma se realizó revisión de literatura académica, informes de la CNBS, BCH, BCIE, BID, CEPAL, leyes y proyectos de ley aplicables a la temática del estudio, publicaciones financieras, investigaciones de temas similares o de ramos paralelos, sitios web confiables que citan fuentes bibliográficas consultadas y documentos emitidos por organizaciones internacionales, entre otros.

Se consulto literatura académica relacionada con tarjetas de crédito y la cartera crediticia para obtener antecedentes y contexto para la investigación.

También se recurrió a informes del BCH, revistas financieras y económicas que cubren la industria bancaria en Honduras y la región, base de datos de investigación en línea (Google Scholar, revista Scielo), para acceder a estudios y artículos académicos, periódicos locales (El Herald, La Prensa) ,y a informes de organizaciones internacionales como el Banco Mundial, el Fondo Monetario Internacional (FMI) el Banco Interamericano de Desarrollo (BID) ,el Banco Central de Honduras (BCH) que publican análisis y perspectivas sobre la economía y el sector financiero en América Latina, incluyendo Honduras.

3.9.1 FUENTES PRIMARIAS

No se utilizaron fuentes primarias de información puesto que la data se obtuvo de la API de uso público de la CNBS y no directamente de las instituciones bancarias del sector comercial de Honduras.

3.9.2 FUENTES SECUNDARIAS

Datos de uso público abiertos obtenidos de la página de la Comisión Nacional de Bancos y Seguros (CNBS), que incluyen información detallada sobre la cartera de mora de tarjeta de crédito, colocación de tarjetas de crédito de forma mensual y por Institución financiera.

La CNBS es una institución no gubernamental que supervisa el sistema financiero de Honduras. Los datos de la CNBS se consideran una fuente confiable de información financiera.

Sin embargo, el uso de datos secundarios también tiene algunas limitaciones:

Una limitación es que los datos secundarios pueden no ser tan precisos como los datos obtenidos directamente de las fuentes originales. Esto se debe a que los datos secundarios pueden haber sido recopilados, procesados o interpretados por terceros.

Otra limitación es que los datos secundarios pueden no ser tan actualizados como los datos obtenidos directamente de las fuentes originales. Esto se debe a que los datos secundarios pueden tardar un tiempo en recopilarse, procesarse e interpretarse.

Por lo tanto, los resultados de esta investigación deben interpretarse teniendo en cuenta las limitaciones del uso de datos secundarios.

CAPÍTULO IV. RESULTADOS Y ANÁLISIS

4.1 INFORME DE PROCESO DE RECOLECCIÓN DE DATOS

Identificación de los datos necesarios

El primer paso en el proceso de recolección de datos fue identificar los datos necesarios para el proyecto. La identificación de los datos necesarios es un paso fundamental en cualquier proyecto de investigación.

En este caso, los datos necesarios se identificaron teniendo en cuenta los objetivos de la investigación. Los objetivos de la investigación eran predecir el valor de la cartera en mora de tarjetas de crédito. Por lo tanto, los datos necesarios eran los que podían proporcionar información sobre las variables que podrían estar relacionadas con el valor de la cartera en mora de tarjetas de crédito.

Los datos identificados fueron los siguientes:

- Cantidad de tarjetas de crédito activas en circulación: Esta variable puede ser un indicador de la demanda de tarjetas de crédito.
- Cantidad de tarjetas de crédito colocadas en un mes por institución: Esta variable puede ser un indicador de la oferta de tarjetas de crédito.
- Valor de la cartera en mora de tarjetas de crédito: Esta variable es el objetivo de la investigación.
- Valor de la cartera crediticia de tarjetas de crédito por institución: Esta variable puede ser un indicador del riesgo de la cartera de tarjetas de crédito.

Estos datos se obtuvieron de la Comisión Nacional de Bancos y Seguros (CNBS) de Honduras, a través de su portal de datos abiertos.

Acceso a los datos

Los datos de la CNBS se obtuvieron a través de su portal de datos abiertos. Este portal proporciona acceso a una gran cantidad de datos sobre el sistema financiero de Honduras.

Para acceder a los datos de la CNBS, se utilizó la API de datos abiertos. Esta API proporciona una interfaz fácil de usar para descargar los datos.

Se obtuvo información completa necesaria para la realización de este estudio para 11 de los bancos que conforman la banca comercial de Honduras, los bancos: Cuscatlán y Popular se han omitido del estudio a falta de la obtención de data suficiente para ambas instituciones.

Descarga de los datos

Los datos se descargaron en formato CSV, que es un formato de texto estándar que es fácil de leer y analizar.

Preparación de los datos

Una vez descargados los datos, se prepararon para su análisis. En este caso, se realizaron las siguientes acciones:

- Eliminación de registros duplicados: Se eliminaron los registros duplicados para evitar sesgos en el análisis.
- Corrección de errores de formato: Se corrigieron los errores de formato para garantizar que los datos fueran precisos y consistentes.
- Transformación de fecha a periodo: luego de revisar la data confirmamos que la fecha de corte el último día de cada mes, decidimos remover el día y trabajar solo con el periodo (AÑO/MES) para que de este modo el modelo pudiera leer una columna estandarizada de 6 dígitos en lugar de tener que validar el valor numérico que representaría el día en el entendido que no todos los meses del año tienen la misma cantidad de días.

Estas acciones de preparación de los datos son necesarias para garantizar que los datos sean adecuados para el análisis.

4.2 ANÁLISIS EXPLORATORIO DE DATOS

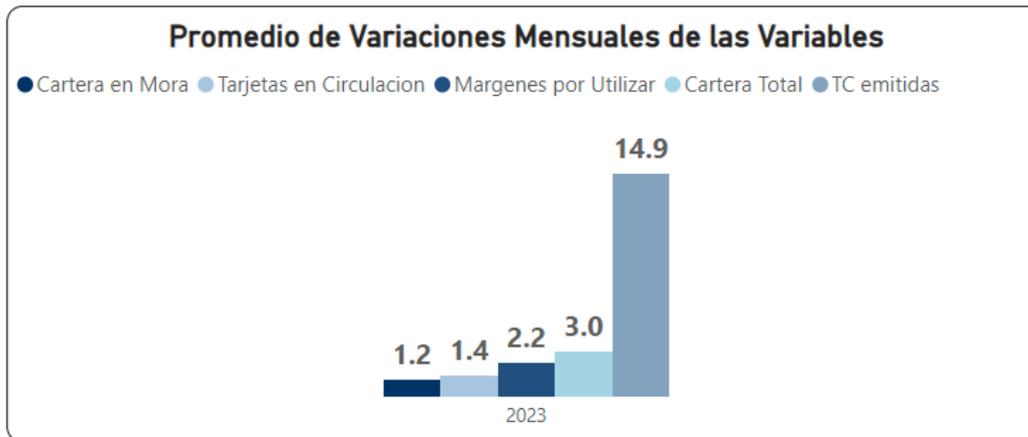


Figura 14. Variación mensual de las variables de estudio.

Fuente: (Elaboración Propia)

Como se puede observar el crecimiento de la cartera en mora parece ser proporcional al crecimiento en la cantidad de tarjetas en circulación, no así con a cantidad de tarjetas emitidas que se posiciona fuera de la normalidad de las demás variables.

La gráfica siguiente muestra la cantidad de tarjetas en circulación y la cartera en mora en Honduras, desde enero hasta septiembre de 2023.

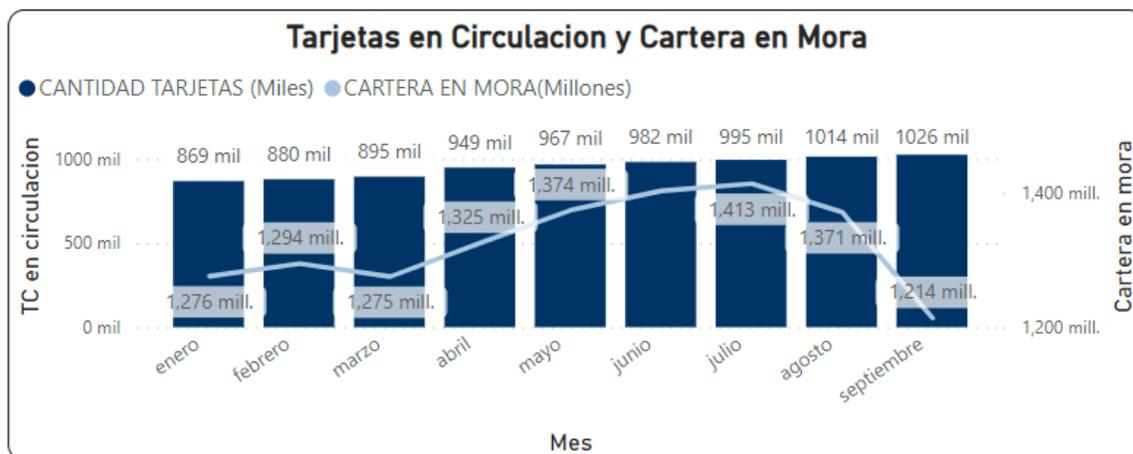


Figura 15. Tarjetas en circulación y cartera en mora por mes.

Fuente: (Elaboración Propia)

La cantidad de tarjetas en circulación ha ido aumentando progresivamente durante el período analizado, pasando de 868 mil en enero a 102 mil en septiembre.

La cartera en mora, por su parte, disminuyó ligeramente durante el período analizado, pasando de 1275 millones de lempiras en enero a 1213 millones de lempiras en septiembre.

En general, la gráfica muestra una tendencia positiva en el sector de las tarjetas de crédito en Honduras. La cantidad de tarjetas en circulación está aumentando, lo que indica que hay un mayor uso de este producto financiero. Además, la cartera en mora está disminuyendo, lo que refleja una mejor gestión de los riesgos crediticios.

El siguiente gráfico muestra el promedio de cartera en mora de los 11 bancos comerciales objetos de este estudio, expresado en millones de lempiras. El gráfico abarca el período de enero de 2023 a septiembre de 2023.

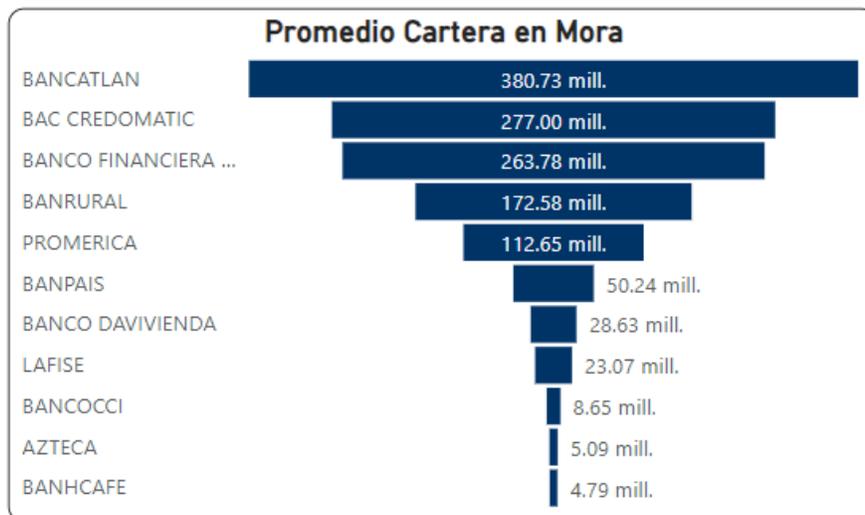


Figura 16. Promedio de cartera en mora por banco

Fuente: (Elaboración Propia)

Como se puede observar, la cartera en mora promedio de los bancos comerciales en Honduras varía significativamente entre los diferentes bancos.

BANCO ATLÁNTIDA, BAC CREDOMATIC y FICOHSA son los bancos con la cartera en mora promedio más alta, mientras que BANCO AZTECA, BANCO DE OCCIDENTE y BANHCAFÉ son los bancos con la cartera en mora promedio más baja. Es importante señalar que

la cartera en mora es una variable según el tamaño del banco, la cartera crediticia total, cantidad de tarjetas en circulación entre otras.

La siguiente gráfica muestra el promedio de tarjetas en circulación de los 11 principales bancos comerciales en Honduras, expresado en miles de tarjetas. El gráfico abarca el período de enero de 2023 a septiembre de 2023.

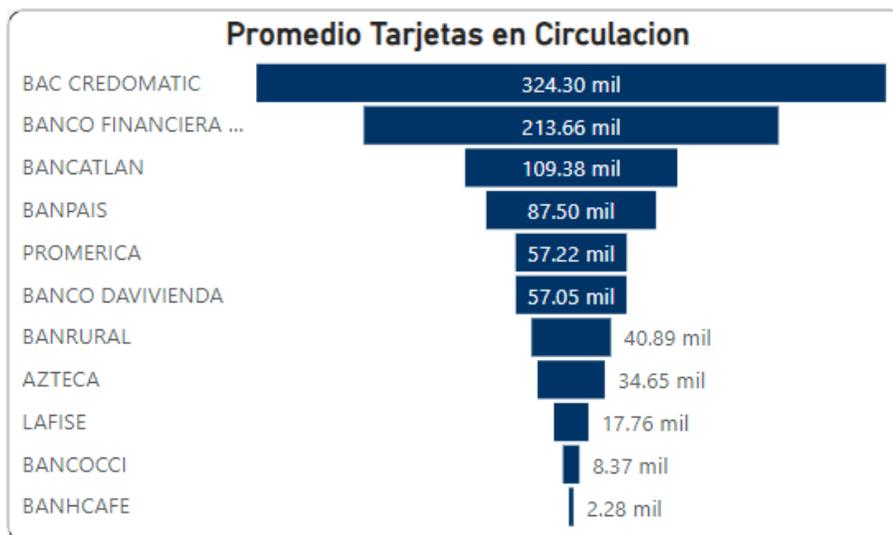


Figura 17. Promedio de tarjetas en circulación por banco.

Fuente: (Elaboración Propia)

Como se puede observar, BAC CREDOMATIC es el banco con el mayor promedio de tarjetas en circulación, seguido de FICOHSA y BANCO ATLÁNTIDA. Estos tres bancos representan el 99% del total promedio de tarjetas en circulación en 2023 de los 11 bancos comerciales objetos de este estudio.

La gráfica muestra el promedio de cartera total de los 11 bancos comerciales estudiados, expresado en millones de lempiras. Al igual que los gráficos anteriores el gráfico abarca el período de enero de 2023 a septiembre de 2023

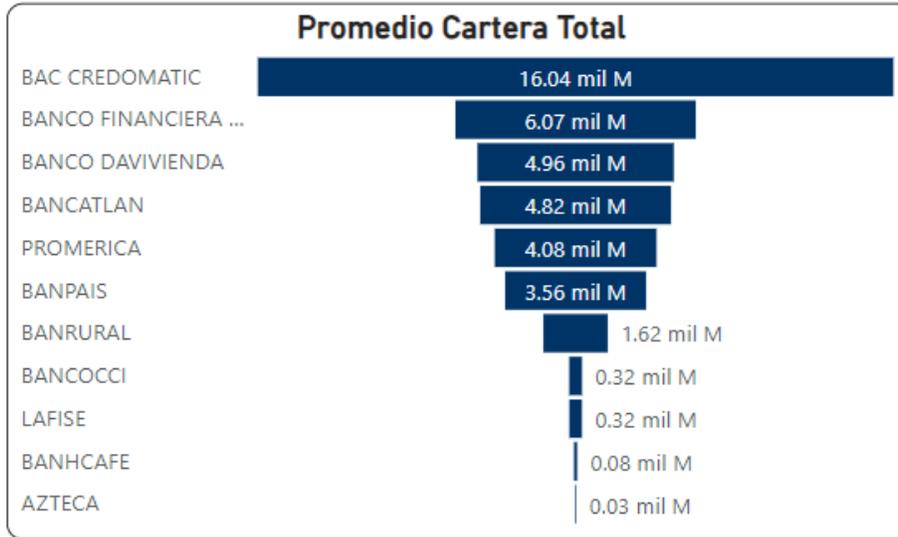


Figura 18. Promedio de cartera total por banco.

Fuente: (Elaboración Propia)

Como se puede observar, BAC Credomatic es el banco con la mayor cartera total de Tarjeta de Crédito con 16041 millones de lempiras. Esto representa el 40% del total de la cartera total de los bancos comerciales estudiados, seguido de Banco FICOHSA y BANCO ATLÁNTIDA, estos tres bancos representan el 67% del total de la cartera total de dichos Bancos.

El siguiente gráfico muestra la evolución de la cartera en mora de los bancos comerciales en Honduras, expresado en millones de lempiras. El gráfico abarca el período de enero de 2022 a septiembre de 2023.

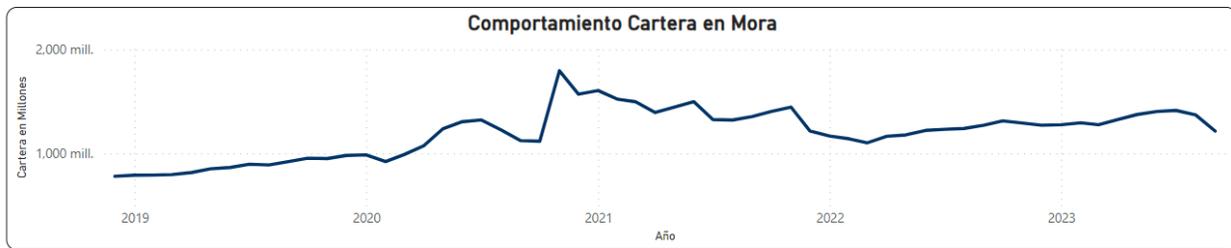


Figura 19. Comportamiento de cartera en mora por año.

Fuente: (Elaboración Propia)

Como se puede observar en el gráfico, la cartera en mora ha presentado una tendencia a la baja en los últimos meses del 2023. En enero de 2022, la cartera en mora era de 1166 millones de lempiras. A pesa de esta disminución en los últimos meses para septiembre de 2023 la cartera en mora ha aumentado a 1213 millones de lempiras.

La siguiente gráfica muestra el comportamiento de la cartera total de tarjetas de crédito de los principales bancos de Honduras del año 2019 al 2023.

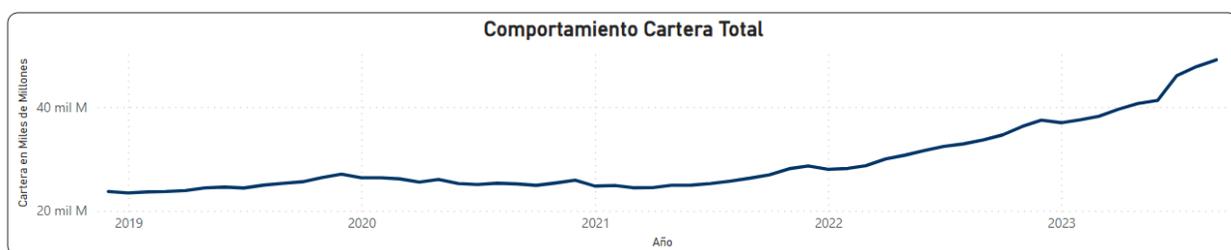


Figura 20. Comportamiento de cartera total por año.

Fuente: (Elaboración Propia)

La gráfica muestra que la cartera total de tarjetas de crédito ha crecido de manera constante durante los últimos cinco años. En enero 2019, la cartera total era de 23,704 millones de lempiras. En enero 2020, la cartera total creció a 26,350 mil millones de lempiras.

En enero 2021, la cartera total disminuyó a 24,758 millones de lempiras influenciado en gran parte por la pandemia. En enero 2022, la cartera total creció a 27,980 mil millones de lempiras. Y en enero 2023, la cartera total creció a 36,960 millones de lempiras.

La cartera total de tarjetas de crédito se ha duplicado en los últimos cinco años. Esto es un crecimiento significativo que demuestra la decantación del consumidor hondureño hacia este medio de pago y financiamiento en Honduras.



Figura 21. Comportamiento cantidad de tarjetas activas en circulación por año.

Fuente: (Elaboración Propia)

La gráfica muestra el comportamiento de la cantidad de tarjetas de crédito en circulación en Honduras de 2019 a 2023.

La tendencia al alza de la cantidad de tarjetas de crédito en circulación es positiva, se observa una disminución en cantidad de tarjeta de crédito en los años que estuvo en su auge la pandemia causada por el COVID-19, pero esta cantidad ha subido considerablemente desde el 2022 a septiembre 2023.

El crecimiento de la cantidad de tarjetas de crédito en circulación en Honduras tiene implicaciones importantes para el sistema financiero hondureño. El aumento de la penetración de las tarjetas de crédito representa una oportunidad para los bancos y otras instituciones financieras para expandir sus negocios. Sin embargo, también representa un desafío, ya que los bancos deben estar preparados para manejar el aumento del riesgo crediticio.

4.3 RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS

4.3.1 RESULTADOS CUANTITATIVOS

4.3.1.1 MODELO DE REGRESIÓN LINEAL CON CANTIDAD DE TARJETAS ACTIVAS COMO VARIABLE INDEPENDIENTE

Tabla 6. Resumen de métricas, modelo de regresión lineal con cantidad de tarjetas activas como variable independiente.

Identificador	Predicción Cartera en Mora
R^2	0.5919
mean absolute error	36.5194
mean squared error	6,373,421,957,366,880
root mean squared error	79,833,714

mean signed difference	652155.094
mean absolute percentage error	37%
adjusted R²	0.5919

Fuente: (Elaboración propia)

- R² (Coeficiente de Determinación):

El valor del R² es 0.5919, indicando que el modelo explica aproximadamente el 59.19% de la variabilidad en la cartera en mora.

Aunque el modelo tiene cierta capacidad para explicar las variaciones en la cartera en mora, todavía queda un porcentaje considerable de variabilidad sin explicar (alrededor del 40.81%).

- MAE (Mean Absolute Error):

El valor de 36.5194 representa la magnitud promedio de los errores en las predicciones del modelo en unidades de cartera en mora. Un MAE bajo indica que las predicciones tienden a desviarse poco de los valores reales.

El MAE proporciona una medida de la magnitud promedio de los errores absolutos. En este caso, el valor sugiere que las predicciones pueden tener desviaciones sustanciales en términos monetarios con respecto a los valores reales.

- Mean Signed Difference:

La dirección promedio es positiva, lo que indica que las predicciones tienden a ser, en promedio, mayores que los valores reales.

Al igual que en el caso anterior, hay una tendencia sistemática en el modelo para sobreestimar la cartera en mora.

Esta observación podría ser clave para realizar ajustes y mejorar la precisión del modelo.

- RMSE (Root Mean Squared Error):

El RMSE, es una medida que nos dice cuánto varían las predicciones del modelo respecto a los valores reales de la cartera en mora. En este caso, el RMSE es de 79,833,714 lempiras, lo que equivale a aproximadamente 79.8 millones de lempiras, dado que la cartera en mora está expresada en millones de lempiras.

Un RMSE más bajo indica que las predicciones del modelo tienden a estar más cercanas a los valores reales, lo que sugiere una mayor precisión en las predicciones. En este contexto, un RMSE de 79.8 millones de lempiras indica que, en promedio, las predicciones pueden desviarse de los valores reales de la cartera en mora por esa cantidad.

- MAPE (Mean Absolute Percentage Error):

El MAPE es del 37%, indicando el porcentaje promedio de error en las predicciones en relación con los valores reales.

Un MAPE del 37% sugiere que las predicciones tienen, en promedio, un error del 37% en relación con los valores reales. Este valor indica que hay un margen de mejora en la precisión del modelo.

En base a estos resultados podemos decir que este modelo tiene cierta capacidad para explicar la variabilidad en la cartera en mora, pero hay oportunidades para mejorar la precisión, sin embargo, la relación entre las variables cartera en mora y cantidad de tarjetas de crédito en circulación por sí sola no podrían ser utilizadas para diseñar un modelo de predicción ya que la precisión del 59% es un valor muy bajo para considerar confiables sus resultados.

4.3.1.2 MODELO DE REGRESIÓN LINEAL CON TODAS LAS VARIABLES

Tabla 7. Resumen de métricas, modelo de regresión lineal con todas las variables de estudio.

Identificador	Predicción Cartera en Mora
R²	0.7682
mean absolute error	46,332,521
mean squared error	3,620,190,719,843,500
root mean squared error	60,168,020
mean signed difference	305445.7066
mean absolute percentage error	33%
adjusted R²	0.7682

Fuente: (Elaboración propia)

- R² (Coeficiente de Determinación):

El valor del R² es 0.7682, lo que significa que el modelo explica aproximadamente el 76.82% de la variabilidad en la cartera en mora.

Un R^2 alto indica que las variables incluidas en el modelo son relevantes y tienen un impacto sustancial en la capacidad del modelo para explicar las variaciones en la cartera en mora. Sin embargo, aún hay un 23.18% de variabilidad que no se explica.

- MAE (Mean Absolute Error):

Se ha calculado un MAE de 46,332,521. Este valor representa la magnitud promedio de los errores en las predicciones del modelo en unidades de cartera en mora. Un MAE bajo indica que las predicciones tienden a desviarse poco de los valores reales.

- Mean Signed Difference:

La dirección promedio es positiva, lo que significa que las predicciones tienden a ser ligeramente mayores que los valores reales.

Este resultado indica una tendencia sistemática en el modelo para sobreestimar la cartera en mora. Es importante investigar por qué esto está ocurriendo y si hay ajustes necesarios en el modelo.

- RMSE (Root Mean Squared Error):

Dado que la cartera en mora está expresada en millones de lempiras, este RMSE equivale aproximadamente a 60.2 millones de lempiras. Un RMSE más bajo sugiere una mayor precisión en las predicciones, indicando que estas tienden a estar más cercanas a los valores reales.

Así, con un RMSE de 60.2 millones de lempiras, las predicciones del modelo pueden desviarse, en promedio, de los valores reales de la cartera en mora por esa cantidad.

- MAPE (Mean Absolute Percentage Error):

El MAPE es del 33%, sugiriendo que, en promedio, las predicciones pueden desviarse en un 33.37% de los valores reales.

El MAPE proporciona una medida de la precisión relativa en términos porcentuales. Un MAPE del 33% sugiere una moderada precisión, pero aún hay espacio para mejorar.

Podemos concluir que el modelo parece tener un buen rendimiento general al explicar la variabilidad en la cartera en mora, pero hay áreas específicas, como la dirección de las predicciones y la precisión porcentual, que podrían beneficiarse de un análisis más detenido y ajustes en el modelo.

Sin embargo, al poseer una precisión de más del 75% se pueden utilizar estas variables para construir un modelo de predicción de cartera en mora.

4.3.1.3 MODELO RANDOM FOREST REGRESSOR POR BANCO

Tabla 8. Resumen de métricas, modelo Random forest regressor banco 1 al 3

Institución	Bancatlan (cod. 1)	Banco Azteca (cod. 2)	Banco Davivienda (cod. 3)
MAPE	0.06%	0.07%	0.48%
MAE	14,276,273.84	322,598.84	8,793,344.40
MSE	354,952,127,682,441.0	210,332,814,632.71	121,516,601,388,901.23
RMSE	18,840,173.23	458,620.55	11,023,456.87
R2	0.925	-0.23	-0.093

Fuente: (Elaboración propia)

Tabla 9. Resumen de métricas, modelo Random forest regressor banco 4 al 6

Institución	Bac Credomatic (cod. 4)	Banrural (cod. 5)	Banco de Occidente (cod.6)
MAPE	0.09%	0.21%	0.12%
MAE	34,349,830.76	12,749,258.94	1,033,197.77
MSE	5,253,023,498,578,076.0	309,642,396,082,741.44	1,748,013,030,038.17
RMSE	72,477,744.85	17,596,658.66	1,322,124.43
R2	0.69	0.92	0.91

Fuente: (Elaboración propia)

Tabla 10. Resumen de métricas, modelo Random forest regressor banco 7 al 9.

Institución	Banpaís (cod. 7)	Banco Ficohsa (cod. 8)	Banhcafé (cod.9)
MAPE	0.10%	0.02%	0.02
MAE	4,571,489.28	5,644,041.02	442,110.55
MSE	39,117,348,150,660.91	67,815,918,343,922.41	789,030,419,633.79
RMSE	6,254,386.31	8,235,042.09	888,273.84
R2	0.88	0.98	0.97

Fuente: (Elaboración propia)

Tabla 11.Tabla 12-Resumen de métricas, modelo Random forest regressor banco 10 y 11.

Institución	Banco Lafise (cod. 10)	Banco Promerica (cod. 11)
MAPE	0.12%	0.04%
MAE	3,408,598.48	3,031,798.29
MSE	26,431,862,913,785.56	15,331,043,518,451.78
RMSE	5,141,192.75	3,915,487.64
R2	0.68	0.98

Fuente: (Elaboración propia)

Interpretación de resultados por banco:

Banco Atlántida (Código 1):

En resumen, las métricas indican que el modelo tiene un rendimiento bastante bueno para el Banco 1, con predicciones que se ajustan bien a los valores reales, según estas medidas específicas, un R2 de 0.92 es un valor extremadamente bueno ya que entre más cercano a 1 sea este resultado, más preciso es el modelo y sus respuestas.

Banco Azteca (Código 2):

Estos resultados indican que el modelo no se ajusta tan bien a los datos del Banco 2, y podría ser necesario explorar otras técnicas o ajustes en el modelo para mejorar su rendimiento en este conjunto de datos específico. También puede ser indicio de la necesidad de contar con un data set que contenga más registros para este banco en particular.

Banco Davivienda (Código 3):

Estos resultados indican que el modelo no se ajusta tan bien a los datos del Banco 3, y podría ser necesario explorar otras técnicas o ajustes en el modelo para mejorar su rendimiento en este conjunto de datos específico.

También puede ser indicio de la necesidad de contar con un data set que contenga más registros para este banco en particular.

Bac Credomatic (Código 4):

En este caso, el R² es 0.6924, lo que indica que el modelo explica aproximadamente el 69.24% de la varianza en la variable objetivo. Este valor es relativamente alto y sugiere que el modelo tiene un buen poder explicativo en comparación con la varianza total. El modelo parece

tener un buen rendimiento en términos de porcentaje de error promedio y capacidad explicativa (R^2).

Banrural (Código 5):

El R^2 es 0.9224, lo que indica que el modelo explica aproximadamente el 92.24% de la varianza en la variable objetivo. Este valor es muy alto y sugiere que el modelo tiene un excelente poder explicativo en comparación con la varianza total.

En general, el modelo parece tener un buen rendimiento para el "Banco 5", con un bajo error porcentual promedio y una capacidad explicativa significativa.

Banco de Occidente (Código 6):

El R^2 es 0.9120, lo que indica que el modelo explica aproximadamente el 91.20% de la varianza en la variable objetivo. Este valor es alto y sugiere que el modelo tiene un buen poder explicativo en comparación con la varianza total.

En general, el modelo parece tener un buen rendimiento para el "Banco 6", con un bajo error porcentual promedio y una capacidad explicativa significativa.

Banpaís (Código 7):

El R^2 es 0.8843, lo que indica que el modelo explica aproximadamente el 88.43% de la varianza en la variable objetivo. Este valor es alto y sugiere que el modelo tiene un buen poder explicativo en comparación con la varianza total.

En general, el modelo parece tener un buen rendimiento para el "Banco 7", con un bajo error porcentual promedio y una capacidad explicativa significativa.

Banco Ficohsa (Código 8):

El R^2 es 0.9813, lo que sugiere que el modelo explica aproximadamente el 98.13% de la varianza en la variable objetivo. Este valor es excepcionalmente alto y sugiere que el modelo tiene un poder explicativo muy fuerte en comparación con la varianza total.

En general, el modelo parece tener un rendimiento excepcionalmente preciso para el "Banco 8", con un error porcentual extremadamente bajo y una capacidad explicativa muy fuerte.

Banhcafé (Código 9):

El R^2 es 0.9790, lo que sugiere que el modelo explica aproximadamente el 97.90% de la varianza en la variable objetivo. Este valor es excepcionalmente alto y sugiere que el modelo tiene un poder explicativo muy fuerte en comparación con la varianza total.

En general, el modelo parece tener un rendimiento excepcionalmente preciso para el "Banco 9", con un error porcentual extremadamente bajo y una capacidad explicativa muy fuerte.

Banco Lafise (Código 10):

El R^2 es 0.6873, lo que sugiere que el modelo explica aproximadamente el 68.73% de la varianza en la variable objetivo. Este valor es moderado y sugiere que el modelo tiene un poder explicativo razonable, pero puede haber espacio para mejoras.

En resumen, el modelo para el "Banco 10" muestra un rendimiento aceptable en términos de error porcentual y métricas de precisión absoluta.

Sin embargo, el poder explicativo del modelo (R^2) es moderado, lo que sugiere que puede haber factores no capturados que afectan la variable objetivo.

Banco Promerica (Código 11):

El R^2 es 0.9839, lo que sugiere que el modelo explica aproximadamente el 98.39% de la varianza en la variable objetivo. Este valor es excepcionalmente alto y sugiere que el modelo tiene un poder explicativo muy fuerte en este caso.

En resumen, el modelo para el "Banco 11" muestra un rendimiento excepcionalmente preciso en términos de error porcentual y métricas de precisión absoluta. Además, el modelo tiene un poder explicativo muy fuerte, indicando una buena capacidad para explicar la variabilidad en la variable objetivo.

Interpretación General de Resultados:

Bancatlan muestra un excelente rendimiento predictivo con un bajo MAPE, MAE, MSE, y un alto R^2 .

Banco Azteca presenta un MAE y un RMSE relativamente bajos, aunque su R^2 es negativo, lo que sugiere un ajuste deficiente del modelo.

Banco Davivienda tiene un RMSE considerablemente alto y un R^2 negativo, lo que indica

un mal ajuste del modelo y predicciones poco precisas.

Bac Credomatic muestra un buen rendimiento predictivo con un bajo MAPE y RMSE, aunque su R^2 es moderado.

Banrural presenta un rendimiento sobresaliente con un bajo MAE, MSE, RMSE, y un alto R^2 , lo que sugiere un modelo muy ajustado y preciso.

Banco de Occidente tiene un rendimiento destacado con un bajo MAE, MSE, y RMSE, y un alto R^2 , lo que indica un modelo bien ajustado y predicciones precisas.

Banpaís muestra un buen rendimiento predictivo con un MAPE y un RMSE moderados y un alto R^2 .

Banco Ficohsa tiene un rendimiento excelente con un MAPE muy bajo, un RMSE bajo y un R^2 alto, lo que sugiere un modelo altamente preciso y bien ajustado.

Banhcafé también presenta un rendimiento sobresaliente con un MAPE muy bajo, un RMSE bajo y un R^2 alto, indicando un modelo de alta calidad y precisión.

Banco Lafise muestra un rendimiento moderado con un MAPE relativamente bajo y un RMSE aceptable, aunque su R^2 es moderado.

Banco Promerica tiene un rendimiento excelente con un MAPE muy bajo, un RMSE bajo y un R^2 alto, lo que sugiere un modelo altamente preciso y bien ajustado.

4.3.2 ANÁLISIS CUALITATIVO

En este estudio no existe un resultado con enfoque cualitativo por las siguientes razones:

El enfoque del presente estudio es cuantitativo. En los estudios cuantitativos, se recolectan datos numéricos que se analizan utilizando métodos estadísticos. En este caso, el objetivo de este estudio es desarrollar un modelo de aprendizaje automático para predecir el valor de la cartera en mora de tarjetas de crédito. Para ello, se recopilieron datos numéricos sobre la cantidad promedio de tarjetas de crédito activas en circulación, la cantidad promedio de tarjetas de crédito colocadas en un mes por institución y el período de tiempo. Estos datos se analizaron utilizando métodos estadísticos para determinar la relación entre las variables y la predicción del valor de la cartera en mora.

Otra razón por la que no existe un análisis cualitativo es debido a que los datos necesarios para realizar un análisis cualitativo no están disponibles. En los estudios cualitativos, se recolectan datos no numéricos, como entrevistas, observaciones y documentos. Estos datos se analizan utilizando métodos cualitativos, como el análisis de contenido y la teoría fundamentada. En este caso, no se dispone de datos cualitativos sobre los factores que influyen en el valor de la cartera en mora de tarjetas de crédito.

4.4 ANÁLISIS INFERENCIAL Y MODELOS APLICADOS

4.4.1 REGRESION LINEAL

Preprocesamiento de Datos:

- a. Se utilizaron dos conjuntos de datos: uno sobre el comportamiento de tarjetas de crédito y otro sobre la cantidad de tarjetas emitidas y en circulación.
- b. Los datos se limpiaron de valores faltantes, garantizando la integridad de la información.
- c. Se realizó un pivote para transformar datos de cartera en mora, valor de cartera crediticia y márgenes por utilizar en columnas, facilitando el análisis y el modelado.
- d. Se estandarizo el formato de fecha para garantizar la consistencia en el manejo temporal de los datos.

Transformación de Variables:

- a. Se realizaron conversiones de variables categóricas a numéricas mejora la capacidad de los modelos para interpretar y aprender patrones.

Modelado y Evaluación:

- a. Se realizo la división del conjunto de datos en 80% para entrenamiento y 20% para prueba para garantizar la validez y generalización del modelo.
- b. La elección de modelos de regresión lineal (Linear Regression Learner y Linear Regression Predictor) es adecuada para comprender las relaciones lineales entre las variables. La evaluación de correlaciones destaca relaciones fuertes y positivas

entre emisiones, tarjetas activas, cartera en mora, márgenes por utilizar y total de cartera, proporcionando insights valiosos. (Definir la relación)

- c. Se realizó la evaluación con el Numeric Scorer proporciona métricas sólidas para medir el desempeño del modelo.

4.4.2 RANDOM FOREST REGRESSION PARA TODOS LOS BANCOS

Se aplicó un modelo de machine learning para generar una herramienta de predicción de cartera en mora, basado en la data disponible, en este caso específico se utilizó Random Forest Regressor de la siguiente manera:

Paso N°1: Importación de librerías.

```
python

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_absolute_percentage_error
from datetime import datetime
import seaborn as sns
```

Aquí, estamos importando las bibliotecas necesarias para el análisis de datos, modelado y visualización.

Paso N°2: Cargar el DataFrame desde un archivo CSV.

```
python

df_1 = pd.read_csv(r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PRO
```

Se utiliza la librería “pandas” para cargar un conjunto de datos desde un archivo CSV, en este caso haciendo uso de una ruta absoluta, el archivo CSV se encuentra alojado dentro del ordenador en una carpeta física.

Paso N°3: Iterar sobre los códigos de banco únicos.

```
python
```

```
for codigo_banco in df_1['Institucion (to number)'].unique():
```

Se utiliza un comando tipo bucle en este caso el “for” para iterar sobre los códigos de banco únicos en la columna 'Institucion (to number)' del DataFrame.

Paso N°4: Preparar datos para regresión.

```
python
```

```
DF_BANCO_231123 = df_1[df_1['Institucion (to number)'] == codigo_banco]  
X_reg = DF_BANCO_231123[['Fecha Reporte', 'Emitidas en el Mes', 'Total Ta  
y_reg = DF_BANCO_231123['CARTERA EN MORA']
```

Se filtró el DataFrame para obtener datos específicos del banco actual y dividiendo los datos en características (X_reg) y la variable objetivo (y_reg). En este caso las características son: Fecha Reporte, Emitidas en el mes y Total Tarjetas Activas, la variable objetivo es Cartera en Mora.

Paso N°5: Preprocesamiento y entrenamiento del modelo de regresión.

```
python
```

```
# (código para la normalización y división de datos)  
regressor = RandomForestRegressor(random_state=42)  
regressor.fit(X_train_reg, y_train_reg)
```

En este paso se está utilizando un modelo de regresión de bosque aleatorio para entrenar el modelo con datos normalizados y divididos. Se utiliza un random state fijo de 42 para crear el objeto random forest regressor. esto con el fin de introducir reproducibilidad en experimentos de machine learning que involucran elementos aleatorios.

Paso N°6: Realizar predicciones y calcular el error.

```
python

y_pred_reg = regressor.predict(X_test_reg)
mape_reg = mean_absolute_percentage_error(y_test_reg, y_pred_reg)
print(f'Mean Absolute Percentage Error (regresión) para Banco {codigo_ban
```

Se están haciendo predicciones en el conjunto de prueba y calculando el error de regresión.

Paso N°7: Agregar resultados al DataFrame original.

```
python

DF_BANCO_231123['CARTERA EN MORA (PREDICTION RESULTS)'] = y_pred_full
DF_BANCO_231123['DIFERENCIA PORCENTUAL'] = ((DF_BANCO_231123['CARTERA EN MO
```

Se agregan los resultados obtenidos en el paso anterior al dataframe original para poder verificar la diferencia entre la cartera en mora real, la cartera en mora predicha por el modelo y la variación en porcentaje calculado entre ambos valores.

Paso N°8: Guardar resultados en archivos CSV.

```
python

results_path = r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO
DF_BANCO_231123.to_csv(results_path.format(codigo_banco=codigo_banco), ir
print(f'Resultados para Banco {codigo_banco} guardados en {results_path.:
```

Finalmente, se guardan los resultados (con predicciones y diferencia porcentual) en un archivo CSV específico para cada banco. Este código se repite para cada banco único en el conjunto de datos, generando modelos y guardando los resultados para cada uno.

4.4.3 RANDOM FOREST REGRESSOR PARA CALCULO DE PREDICCIONES DE CARTERA EN MORA PARA CADA BANCO EN PARTICULAR.

Se aplicó un modelo de machine learning basado en Random Forest Regressor para generar una herramienta de predicción de cartera en mora, utilizando la data disponible; esta vez de cada banco; con la finalidad de que el modelo se entrenara y aprendiera tomando en cuenta las particularidades de comportamiento por cada institución ya que sabemos que la construcción de la cartera total, cartera en mora y la manera en que la cantidad de tarjetas en circulación afecta el crecimiento o decrecimiento varia para cada institución. Así mismo decidimos obtener y validar; para cada modelo; las métricas comúnmente utilizadas para evaluar la calidad de un modelo de regresión Aquí hay una breve descripción de cada una:

- Mean Absolute Percentage Error (MAPE):

Es una métrica que mide el porcentaje promedio de error entre los valores observados y los valores predichos; Se calcula como la media del valor absoluto de los errores porcentuales individuales.

- Mean Absolute Error (MAE):

Mide la magnitud promedio de los errores entre los valores observados y los valores predichos; Es la media del valor absoluto de los errores individuales.

- Mean Squared Error (MSE):

Mide el promedio de los cuadrados de los errores entre los valores observados y los valores predichos; Penaliza más fuertemente los errores grandes debido al cuadrado.

- R-squared (R²) Score:

También conocido como el coeficiente de determinación. Proporciona una medida de la proporción de la varianza en la variable dependiente que es predecible a partir de las variables independientes.

R² puede variar entre 0 y 1, donde 1 indica una predicción perfecta y 0 indica que el modelo no explica la variabilidad de la variable dependiente.

Paso N°1: Importación de librerías.

```
# Importar librerías necesarias
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_absolute_percentage_error, mean_absolute_error, mean_squared_error
import seaborn as sns
import pickle
```

Se importan las librerías necesarias, incluyendo Pandas para manipulación de datos, NumPy para operaciones numéricas, scikit-learn para modelos y métricas, Seaborn para visualización y Pickle para guardar el modelo.

Paso N°2: Cargar el DataFrame desde el Archivo CSV.

```
# Cargar el DataFrame desde el archivo CSV
df_1 = pd.read_csv(r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO TESIS\BD_TC_DEPURADA')
Carga un conjunto de datos desde un archivo CSV utilizando la librería “pandas”.
```

Paso N°3: Procesamiento de Bancos.

```
# Preguntar si quieres procesar un banco específico
opcion_banco = input("¿Quieres procesar un banco específico? (sí/no): ").lower()
```

Se pregunta al usuario si se quiere procesar un banco específico.

Al escribir el comando “Sí” se avanza a la siguiente etapa del proceso, si se escribe el comando “no” se ejecutan todos los bancos.

```
while opcion_banco != 'no':
    if opcion_banco == 'sí' or opcion_banco == 'si':
        # Pedir el código del banco
        codigo_banco = input("Ingresa el código del banco: ")
        bancos_a_procesar = [codigo_banco]
    else:
        # Procesar todos los bancos
        bancos_a_procesar = df_1['Institucion (to number)'].unique()
```

Al hacerlo por banco específico (comando “Sí”) Utiliza un bucle para procesar bancos; de lo contrario, al utilizar el comando “End” se culmina la ejecución del código.

Paso N°4: Modelo de Regresión para Cada Banco.

```

# Modelos por banco para evaluar la relación
for codigo_banco in bancos_a_procesar:
    DF_BANCO_231123 = df_1[df_1['Institucion (to number)'] == int(codigo_banco)] # Convertir

# Verificar si hay datos para el banco actual
if len(DF_BANCO_231123) == 0:
    print(f"No hay datos para el banco {codigo_banco}.")
    continue

X_reg = DF_BANCO_231123[['Fecha Reporte', 'Emitidas en el Mes', 'Total Tarjetas Activas']]
y_reg = DF_BANCO_231123['CARTERA EN MORA']

```

Primero se toma el código de banco ingresado por el usuario y se compara contra los códigos de banco del dataframe haciendo la conversión del código ingresado en formato integer (int) contra los códigos en la columna Institucion (to number). Se realiza una comprobación de los datos que existen para el código de banco ingresado por el usuario, en caso que ingrese un código no existente en el dataframe original se devuelve el mensaje “No hay datos para el banco”, caso contrario. el código, continua su ejecución y Selecciona características (`X_reg`) y la variable objetivo (`y_reg`) para el modelo de regresión.

```

# Convertir la columna 'Fecha Reporte' a un formato numérico (año * 100 + mes)
X_reg['Fecha Reporte'] = X_reg['Fecha Reporte'] // 100 * 12 + X_reg['Fecha Reporte'] % 100

```

Convierte la columna 'Fecha Reporte' a un formato numérico (año * 100 + mes). Esto se realiza para que la columna que originalmente contiene formato (AAAAMM) se pueda convertir a un número que el código interpretará como valor continuo y de esta forma comparar los valores reales y predichos de la cartera en mora para cada periodo subsecuente.

```

# Normalización de datos
scaler_reg = StandardScaler()
X_reg_scaled = scaler_reg.fit_transform(X_reg)

```

Se realiza la normalización de los datos utilizando StandardScaler.

```

X_train_reg, X_test_reg, y_train_reg, y_test_reg = train_test_split(X_reg_scaled, y_reg, test_siz

```

Se dividen los datos en conjuntos de entrenamiento y prueba.

```
test_size=0.2, random_state=42)
```

Test_size = 0.2 indica que se están dividiendo los datos del dataframe original en una proporción 80/20. 80% para entrenar el modelo y 20% para la prueba del mismo. El dato siguiente Random_state = 42 es un número fijo de utilizado para crear el objeto random forest regressor. esto con el fin de introducir reproducibilidad en experimentos de machine learning que involucran elementos aleatorios. Por convencionalidad se utiliza el número 42 pero puede ser cambiado por el número que el ejecutor del código decida.

```
regressor = RandomForestRegressor(random_state=42)
regressor.fit(X_train_reg, y_train_reg)

y_pred_reg = regressor.predict(X_test_reg)
```

Se inicializa y entrena un modelo de regresión de bosque aleatorio, luego realiza predicciones en el conjunto de prueba.

```
mape_reg = mean_absolute_percentage_error(y_test_reg, y_pred_reg)
print(f'Mean Absolute Percentage Error (regresión) para Banco {codigo_banco}: {mape_reg}%')

# Calcular otras métricas
mae_reg = mean_absolute_error(y_test_reg, y_pred_reg)
print(f'Mean Absolute Error (regresión) para Banco {codigo_banco}: {mae_reg}')

mse_reg = mean_squared_error(y_test_reg, y_pred_reg)
print(f'Mean Squared Error (regresión) para Banco {codigo_banco}: {mse_reg}')

rmse_reg = np.sqrt(mse_reg)
print(f'Root Mean Squared Error (regresión) para Banco {codigo_banco}: {rmse_reg}')

r2_reg = r2_score(y_test_reg, y_pred_reg)
print(f'R2 (regresión) para Banco {codigo_banco}: {r2_reg}')
```

Se calculan las métricas mencionadas anteriormente para evaluación del modelo.

```
# Realizar predicciones en todo el conjunto de datos
y_pred_full = regressor.predict(X_reg_scaled)
```

Al tener un dataset con datos limitados para cada banco, se utiliza el modelo ya entrenado

y probado para realizar predicciones en todo el conjunto de datos.

```
# Agregar predicciones al DataFrame original
DF_BANCO_231123['CARTERA EN MORA (PREDICTION RESULTS)'] = y_pred_full

# Calcular la diferencia porcentual
DF_BANCO_231123['DIFERENCIA PORCENTUAL'] = ((DF_BANCO_231123['CARTERA EN MORA'] - DF_BANCO_231123['CARTERA EN MORA REAL'] ) / DF_BANCO_231123['CARTERA EN MORA REAL'] ) * 100
```

Se agregan las predicciones al DataFrame original y posteriormente se calcula la diferencia porcentual entre las predicciones realizadas por el modelo y los valores reales de cartera en mora, esta diferencia porcentual se almacena en una nueva columna dentro del dataframe.

```
# Imprimir los resultados y guardar en un archivo CSV
results_path = rf'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO TESIS\results_regresion_DF_BANCO_231123.to_csv(results_path, index=False)

print(f'Resultados para Banco {codigo_banco} guardados en {results_path}')
```

Imprime los resultados y guarda el DataFrame resultante en un archivo CSV para su posterior revisión y análisis.

```
# Guardar el modelo en un archivo
with open(f'modelo_regresion_{codigo_banco}.pkl', 'wb') as model_file:
    pickle.dump(regressor, model_file)

# Guardar el scaler en un archivo
with open(f'scaler_reg_{codigo_banco}.pkl', 'wb') as scaler_file:
    pickle.dump(scaler_reg, scaler_file)
```

Se almacena el modelo y el scaler en un archivo pkl utilizando la librería pickle.

```
# Preguntar si quieres procesar otro banco
opcion_banco = input("¿Quieres procesar otro banco? (Sí/No/End): ").lower()

# Terminar la ejecución si se ingresa "end"
if opcion_banco == 'end':
    break

print("Fin de la ejecución.")
```

El código pregunta si deseamos procesar otro banco, en caso de colocar SÍ, nos preguntará el código de banco a procesar y se ejecutará todo el proceso nuevamente hasta que escribamos el comando End para que se termine la ejecución y nos devuelva el mensaje “Fin de la ejecución.

4.4.4 RANDOM FOREST REGRESSOR PARA CALCULO DE PRONOSTICO DE CARTERA EN MORA PARA CADA BANCO EN PARTICULAR.

Este código realiza la carga de un modelo de regresión guardado, así como de su correspondiente escalador, luego aplica este modelo a un nuevo conjunto de datos, realiza predicciones/pronósticos y añade los resultados al DataFrame guardando finalmente el nuevo archivo en formato csv. A continuación, se muestra el desglose paso a paso:

Paso N°1: Importación de Librerías.

```
import pandas as pd
import pickle
```

Se importan las librerías necesarias, incluyendo Pandas para manipulación de datos y Pickle para cargar modelos y escaladores.

Paso N°2: Cargar el Modelo y el Escalador.

```
with open(r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO TESIS\result pr
regressor = pickle.load(model_file)
```

Se abre el archivo que contiene el modelo entrenado (modelo_regresion_5.pkl) y se carga en la variable regressor haciendo uso de la librería pickle.

```
with open(r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO TESIS\result pr
scaler_reg = pickle.load(scaler_file)
```

Se abre el archivo que contiene el escalador entrenado (scaler_reg_5.pkl) y lo carga en la variable scaler_reg. haciendo uso de la librería pickle.

Paso N°3: Cargar el Nuevo Conjunto de Datos.

```
nuevo_df = pd.read_csv(r'C:\Users\Lenovo PC\Documents\Python Scripts\BD PROYECTO TES
print(nuevo_df.head())
print(nuevo_df.columns)
```

Se debe contar con un nuevo conjunto de datos que contenga las mismas columnas que el data set que se utilizó para generar el modelo y escalador entrenado. En este caso un data set con

las columnas Fecha Reporte, Institución, Emitidas en el Mes y Total Tarjetas Activas. También se indica mostrar las primeras filas y las columnas para asegurar la carga correcta del archivo.

Paso N°4: Preprocesamiento del Nuevo Conjunto de Datos.

```
nuevo_df['Fecha Reporte'] = nuevo_df['Fecha Reporte'] // 100 * 12 + nuevo_df['Fecha
```

Ajusta la columna 'Fecha Reporte' para que tenga el mismo formato que en el conjunto de entrenamiento.

```
X_nuevo = nuevo_df[['Fecha Reporte', 'Emitidas en el Mes', 'Total Tarjetas Activas']]
```

Selecciona las características relevantes (X_nuevo) del nuevo conjunto de datos.

```
X_nuevo_scaled = scaler_reg.transform(X_nuevo)
```

Normaliza el conjunto de datos nuevo utilizando el escalador previamente cargado.

```
y_pred_nuevo = regressor.predict(X_nuevo_scaled)
```

Aplica el modelo cargado para hacer predicciones sobre el nuevo dataset normalizado.

```
nuevo_df['CARTERA EN MORA (PREDICTION RESULTS)'] = y_pred_nuevo
```

Añade las predicciones al DataFrame original bajo una nueva columna llamada 'CARTERA EN MORA (PREDICTION RESULTS)'.

Paso N°5: Guardar el DataFrame Resultante.

```
nuevo_df.to_csv('resultados_nuevo_conjunto.csv', index=False)
```

Se guarda el DataFrame resultante, que ahora incluye las predicciones, en un archivo CSV llamado 'resultados_nuevo_conjunto.csv'. El parámetro index=False evita que se incluyan los índices en el archivo CSV.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

Objetivo Específico A:

La relación entre la cantidad de tarjetas activas en circulación y el comportamiento de la cartera crediticia en mora varía significativamente entre los bancos objeto de estudio. Si bien algunos bancos muestran una relación lineal del 75% o más entre estas variables, otros pueden presentar relaciones más complejas o menos directas. Esta variabilidad sugiere la importancia de considerar las características específicas de cada institución bancaria al diseñar estrategias de gestión de riesgos y modelos predictivos.

B. Objetivo Específico B:

El análisis de patrones y tendencias específicas de la cartera en mora en el contexto de la banca comercial privada hondureña revela la heterogeneidad en el comportamiento de este indicador entre los bancos estudiados. Esto sugiere la existencia de factores internos y externos que influyen en la dinámica de la cartera en mora y subraya la importancia de un enfoque detallado y específico para comprender y gestionar el riesgo crediticio en cada institución financiera.

C. Objetivo Específico C:

La construcción de un modelo de aprendizaje automático para proyectar la cartera en mora de los bancos objeto de estudio permite identificar patrones y relaciones complejas entre las variables estudiadas. Los resultados indican la viabilidad de utilizar enfoques de machine learning para mejorar la precisión de las proyecciones de la cartera en mora y optimizar la gestión de riesgos en el sector bancario hondureño.

D. Objetivo Específico D:

Este estudio contribuye al conocimiento financiero local al proporcionar una visión detallada y actualizada sobre la dinámica de las tarjetas de crédito en el sector bancario hondureño. Los hallazgos y los modelos desarrollados pueden servir como referencia para futuras investigaciones y prácticas en el ámbito de la gestión de riesgos y la toma de decisiones financieras en Honduras.

Conclusión General:

El diseño e implementación de un modelo de aprendizaje automático para predecir la cartera en mora de tarjetas de crédito en el sector bancario hondureño es un paso crucial hacia una gestión de riesgos más eficiente y una toma de decisiones más informada. Los resultados obtenidos subrayan la importancia de considerar la heterogeneidad entre las instituciones financieras y la utilidad de enfoques avanzados de análisis de datos para comprender y gestionar el riesgo crediticio en un entorno bancario dinámico y cambiante lleno de particularidades tanto propias del sector como de cada institución financiera.

En general, los resultados de la interpretación por banco son alentadores, ya que indican que el modelo de aprendizaje automático desarrollado tiene el potencial de ser una herramienta útil para las instituciones financieras para predecir el valor de la cartera en mora de tarjetas de crédito, a pesar de esto, pero en vista que dichas variables no son suficientes para obtener predicciones con una precisión aceptable (más del 80%) para todos los bancos que fueron objetivo de este estudio se concluye como aceptada la hipótesis nula.

5.2 RECOMENDACIONES

Con base en los resultados de la investigación se presenta la siguiente serie de recomendaciones fundamentadas para mejorar el rendimiento del modelo de predicción de la cartera en mora de tarjetas de crédito en el sector bancario hondureño:

Personalización del Modelo por Banco:

Dado que se observó una variabilidad significativa en el desempeño del modelo entre los distintos bancos, se recomienda personalizar y ajustar el modelo de predicción según las características y el comportamiento específico de cada institución bancaria. Esto podría implicar la inclusión de variables adicionales o la adaptación de los algoritmos de aprendizaje automático para cada banco en particular.

Exploración de Modelos Alternativos:

Aunque se obtuvieron resultados positivos con el modelo de Random forest, se sugiere explorar y comparar otros modelos de aprendizaje más complejos, como redes neuronales, gradient boosting, o modelos de series temporales. La exploración de modelos alternativos puede ayudar a capturar patrones más complejos y mejorar la capacidad predictiva del modelo.

Refinamiento de Variables Predictivas:

Se recomienda realizar un análisis detallado de las variables independientes utilizadas en el modelo y considerar la inclusión de nuevas variables que puedan proporcionar una mejor explicación del comportamiento de la cartera en mora. Esto podría incluir variables macroeconómicas, tendencias del mercado financiero y datos sociodemográficos que puedan influir en el comportamiento de los clientes y la cartera en mora.

Validación Cruzada y Evaluación Continua del Modelo:

Es fundamental realizar validaciones cruzadas y evaluaciones continuas del modelo a medida que se recopilan nuevos datos. Esto ayudará a identificar posibles desviaciones en el rendimiento del modelo y permitirá ajustar y mejorar el modelo de manera proactiva.

Incorporación de Métricas Adicionales:

Además de las métricas tradicionales como el MAE, el MSE y el RMSE, se recomienda incorporar métricas adicionales que puedan proporcionar una evaluación más completa del rendimiento del modelo. Esto incluye métricas de precisión, sensibilidad, especificidad y áreas bajo la curva ROC (Receiver Operating Characteristic) que pueden ser especialmente útiles para modelos de clasificación.

Colaboración Interinstitucional:

Se sugiere promover la colaboración y el intercambio de información entre las instituciones bancarias para mejorar la calidad de los datos y desarrollar modelos predictivos más robustos y generalizables. El establecimiento de estándares comunes y la compartición de mejores prácticas pueden contribuir significativamente a la mejora del rendimiento del modelo en el sector bancario hondureño.

Sostenibilidad y Evaluación Continua:

Es fundamental reconocer que el desarrollo y la implementación de un modelo de predicción de la cartera en mora no constituyen un proyecto a corto plazo, sino un esfuerzo continuo que requiere dedicación y compromiso a largo plazo. Para alcanzar un nivel de confianza en los resultados del modelo de al menos el 90%, es necesario mantener una evaluación constante y alimentar el modelo con nueva información y datos actualizados.

El trabajo constante y la retroalimentación continua son clave para mejorar la precisión y la confiabilidad del modelo a lo largo del tiempo. Esto implica la incorporación de nuevos datos, la revisión periódica de las variables y el ajuste del modelo según sea necesario para reflejar los cambios en el entorno bancario y económico.

Solo a través de un compromiso sostenido y una evaluación continua se puede alcanzar un nivel de confianza suficiente en los resultados del modelo para que sea considerado una herramienta confiable y productiva para la gestión de riesgos y la toma de decisiones en el sector bancario hondureño.

Al mantener un enfoque de trabajo constante y una mentalidad orientada a la mejora continua, se podrá integrar efectivamente el modelo de predicción de la cartera en mora en las operaciones diarias de las instituciones financieras, lo que contribuirá significativamente a la estabilidad y la eficiencia del sector bancario en Honduras.

CAPÍTULO VI. APLICABILIDAD

6.1 NOMBRE DE LA PROPUESTA

Propuesta de Implementación de Modelo Predictivo para Proyección y Evaluación de Cartera en Mora en Instituciones del Sector Bancario Comercial en Honduras.

6.2 JUSTIFICACIÓN DE LA PROPUESTA

La presente propuesta de implementación de un modelo predictivo para la proyección y evaluación de la cartera en mora de tarjetas de crédito en instituciones bancarias, está enfocada para instituciones financieras pequeñas como ser OPDF's o cooperativas pequeñas que no cuentan con un departamento de analítica como tal, esta propuesta ofrece varias ventajas:

- **Optimización de Recursos:** Permite identificar de manera más eficiente las áreas de mayor riesgo, lo que facilita la asignación de recursos.
- **Mejora en la Toma de Decisiones:** Proporciona insights basados en datos para respaldar decisiones estratégicas relacionadas con la gestión de cartera.
- **Reducción de Pérdidas:** Al anticipar posibles problemas de morosidad, se pueden tomar medidas preventivas para reducir las pérdidas, activar medidas de prevención de mora alta para clientes seleccionados, entre otras.

6.3 ALCANCE DE LA PROPUESTA

El alcance de esta propuesta incluye la implementación y despliegue del modelo predictivo desarrollado utilizando la metodología Random Forest en 11 bancos específicos:

Tabla 12. Alcance de la propuesta de aplicabilidad.

Código asignado al banco	Nombre Comercial de la institución financiera
1	Banco Atlántida (Bancatlan)
2	Banco Azteca
3	Banco Davivienda
4	Bac Credomatic
5	Banrural
6	Banco de Occidente
7	Banpaís
8	Banco Ficohsa

9	Banhcafé
10	Banco Lafise
11	Banco Promerica

Fuente: (Elaboración propia)

La aplicación abordará la evaluación de la cartera en mora utilizando datos históricos para el entrenamiento del modelo y posteriormente se utilizará el modelo entrenado para realizar predicciones y proyecciones de cartera en mora para periodos posteriores.

6.4 DESCRIPCIÓN Y DESARROLLO

6.4.1 DESCRIPCIÓN

El proyecto se centra en el desarrollo e implementación de un modelo predictivo destinado a evaluar la cartera en mora de tarjetas de crédito en instituciones financieras, específicamente en el contexto de once bancos. El proceso se divide en dos partes fundamentales:

Código de Desarrollo del Modelo

- Carga y Pre procesamiento de Datos: Se inicia con la carga de datos desde un archivo CSV, que incluye información detallada de diferentes bancos. Luego, se realiza un proceso de preprocesamiento que incluye la conversión de fechas y la normalización de variables clave.
- Entrenamiento y Evaluación del Modelo: Utilizando el algoritmo Random Forest, se desarrolla un modelo predictivo individualizado para cada banco.
- Este modelo se entrena y evalúa utilizando métricas estándar, asegurando su capacidad para prever la cartera en mora con precisión.
- Generación y Almacenamiento de Predicciones: El modelo se utiliza para generar predicciones tanto para conjuntos de prueba como para el conjunto completo de datos de cada banco. Los resultados y métricas se almacenan para su análisis posterior.

Código de Aplicación del Modelo a Nuevos Datos

- **Carga del Modelo y Scaler Pre-entrenado:** Se inicia cargando el modelo de Random Forest y el objeto StandardScaler, ambos previamente entrenados durante el desarrollo del modelo.
- **Preprocesamiento y Normalización de Nuevos Datos:** Se carga un nuevo conjunto de datos, se ajusta la columna de fechas y se seleccionan las variables relevantes. Luego, se normalizan estos datos utilizando el mismo scaler que se empleó en el entrenamiento del modelo.
- **Aplicación del Modelo a Nuevos Datos:** El modelo se aplica a los nuevos datos normalizados, generando predicciones precisas sobre la cartera en mora para instituciones financieras en el nuevo dataframe que contiene registros no vistos durante el entrenamiento.

Este enfoque integral asegura la aplicabilidad del modelo a nuevas instancias, ofreciendo una herramienta valiosa para la evaluación de riesgos financieros en diferentes bancos.

6.4.2 DESARROLLO

- **Selección de Bancos:** El modelo se desarrolló y evaluó individualmente para cada uno de los 11 bancos seleccionados.
- **Metodología Random Forest:** Se utilizó Random Forest, un algoritmo de aprendizaje automático eficaz para problemas de regresión, para desarrollar el modelo.
- **Preprocesamiento de Datos:** Los datos se preprocesaron mediante normalización y división en conjuntos de entrenamiento y prueba.

6.4.3 EVALUACIÓN Y MÉTRICAS

- **Métricas de Evaluación:** Se emplearon métricas como Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), y R^2 para evaluar la precisión del modelo.

- **Análisis de Resultados:** Se interpretaron las métricas para cada banco, destacando el rendimiento del modelo en términos de predicción de cartera en mora.

6.5 MEDIDAS DE CONTROL

Validación Continua

- **Monitoreo Continuo:** Implementación de un sistema de monitoreo para evaluar continuamente el rendimiento del modelo en tiempo real.
- **Actualización del Modelo:** Se planifican actualizaciones periódicas del modelo para garantizar la adaptabilidad a cambios en los datos y en el comportamiento del usuario.

Implementación Gradual

- **Pilotos por Banco:** Se sugiere implementar el modelo en forma de pilotos para cada banco antes de una implementación a gran escala.
- **Capacitación del Personal:** Proporcionar capacitación al personal para comprender e interpretar los resultados del modelo.

6.6 CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO

Proyecto: Implementación de Modelo Predictivo para Proyección y Evaluación de Cartera en Mora en Instituciones Financieras del Sector Bancario Comercial en Honduras.

Producto Mínimo Viable (MVP)

Esta propuesta se centra en el desarrollo e implementación de un modelo predictivo destinado a evaluar la cartera en mora de tarjetas de crédito en instituciones financieras, específicamente en el contexto de 11 bancos. El modelo se desarrollará utilizando la metodología Random Forest y se implementará en la nube utilizando servicios de alojamiento de datos y procesamiento de datos.

Detalles de la propuesta:

Adquisición de Tecnología

Servidores y Equipos: Se contratarán los servicios de un proveedor de nube pública para alojar el modelo. Esto elimina la necesidad de adquirir y mantener servidores y equipos físicos.

Software Especializado: Se utilizarán servicios a plataformas de desarrollo y software de análisis de datos por suscripción, así como software de código abierto (opensource) para el desarrollo de los modelos.

Recursos Humanos

Personal de Desarrollo y Soporte: Se contratará un empleado nuevo plenamente calificado para la institución para desarrollar y mantener el modelo sin necesidad de invertir en capacitación adicional.

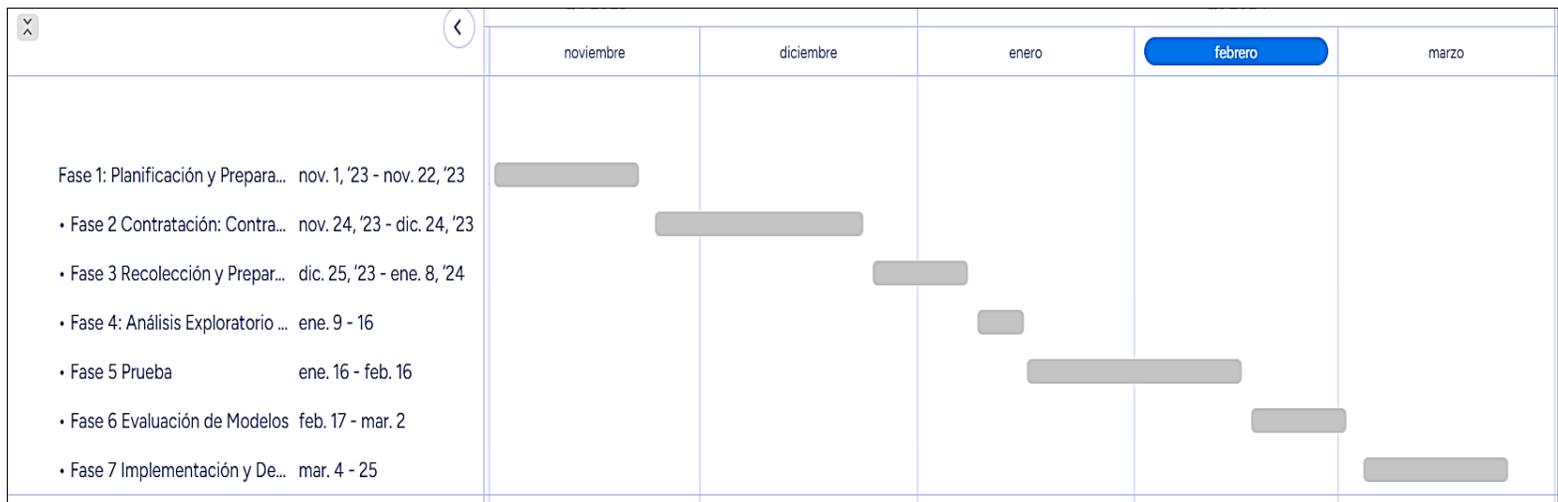
6.6.1 Cronograma

Fecha de inicio: 1 de noviembre de 2023

Fecha de finalización: 4 de marzo de 2024

Figura 22. Cronograma de ejecución del proyecto.

Fuente: (Elaboración propia)



6.6.2 Descripción de las fases:

- **Fase 1: Planificación y Preparación (3 Semanas):**

Semana 1-3:

- Definición de Objetivos del Proyecto-
- Identificar claramente el objetivo del proyecto, que es desarrollar un modelo de aprendizaje automático para predecir la cartera en mora de tarjetas de crédito en el sector bancario comercial de Honduras.
- - Adquisición de Tecnología y Recursos: Investigar y adquirir las herramientas y tecnologías necesarias para el desarrollo del modelo, como software de análisis de datos, plataformas de aprendizaje automático y hardware adecuado.

- **Fase 2 Contratación: Contratación de un Analista de BI con Conocimientos en Python e implementación de modelos de Machine Learning: (4 Semanas):**

Semana 1:

- Publicación de la plaza laboral.
- Revisión de los CV.
- Entrevista con los candidatos.

Semana 2:

- Evaluación de los candidatos.
- Selección del candidato.

Semana 3 y 4:

- Capacitación y adaptación del nuevo integrante.

- **Fase 3 Recolección y Preparación de Datos (2 semanas):**

Semana 1:

- Identificación de Datos Necesarios: Definir las variables clave necesarias para el modelo, como la cantidad de tarjetas en circulación, la cartera en mora y el valor de la cartera crediticia.
- Preparación de Datos: Limpiar y preprocesar los datos, eliminando registros duplicados, corrigiendo errores de formato y transformando la fecha a un formato estandarizado.

- **Fase 4: Análisis Exploratorio de Datos (1 Semana):**

Semana 1:

- Exploración y Visualización de Datos: Realizar análisis exploratorio de datos para comprender las relaciones y tendencias entre las variables. Generar gráficos y visualizaciones para identificar patrones y anomalías en los datos.

- **Fase 5 Prueba (4 semanas):**

Semana 1 y 2:

- Implementación del modelo en entornos de prueba.
- Recopilación de datos de prueba.
- Configuración de métricas de rendimiento.

Semana 3 y 4:

- Evaluación del rendimiento del modelo.
- Ajuste del modelo según sea necesario.

- **Fase 6 Evaluación de Modelos (2 semanas):**

Semana 1 y 2:

- Selección de Modelos: Evaluar diferentes modelos de aprendizaje automático, como

regresión lineal y Random Forest Regression, para determinar el más adecuado para el problema.

- Validar su rendimiento utilizando métricas específicas.

- **Fase 7 Implementación y Despliegue (3 semanas):**

Semana 1 y 2:

- Desarrollo de la Solución: Desarrollar la solución de software que incluya la integración del modelo de aprendizaje automático en una aplicación o plataforma.

Semana 3:

- Pruebas y Ajustes: Realizar pruebas exhaustivas para garantizar que el modelo funcione correctamente en diferentes escenarios. Realizar ajustes y optimizaciones según sea necesario para mejorar el rendimiento del modelo.

6.6.3 Presupuesto

A continuación, se proporciona una estimación general del presupuesto (expresado en lempira hondureño) necesario para un despliegue productivo de los modelos.

Cabe mencionar que puede ajustarse según las necesidades y condiciones específicas del proyecto de acuerdo al nivel de implementación o envergadura de la institución bancaria que desee embarcarse en el proyecto o los recursos disponibles para el mismo.

Tabla 13-Desglose de presupuesto para implementación del proyecto.

Categoría	Descripción	Monto (LPS)
Adquisición de tecnología y software especializado	Licencias de Power BI Premium, espacio en la nube y máquina virtual de Azure	12,000 por mes
Recursos humanos	Salario del empleado que desarrollará y mantendrá el modelo	35,000 por mes
Total		165,000.00 LPS

Fuente: (Elaboración propia)

6.6.4 Desglose a detalle del presupuesto planteado

6.6.4.1 Adquisición de Tecnología:

- Adquisición de Tecnología y Software Especializado (12,000 LPS):

Adquisición de tecnología: Se estima que el costo de las licencias de Power BI Premium, el espacio en nube y la máquina virtual de Azure será de (12,000 LPS) desglosados de la siguiente forma:

- *4 licencias de Power BI Premium:* El costo de cada licencia es de 20 USD, por lo que el costo total de las 4 licencias será de 80 USD.

Al tipo de cambio actual (1 USD = 25 LPS), el costo total de las licencias será de 2,000 LPS (admin, ¿Qué es power Bi?, evotic, 2023).

- *Espacio en nube:* El costo del espacio en nube dependerá del tamaño y la duración del almacenamiento de los datos. Para este proyecto, se estima que el costo será de 4,000 LPS.
- *Máquina virtual de Azure:* El costo de la máquina virtual dependerá del tipo de máquina, la cantidad de recursos y la duración del arrendamiento. Para este proyecto, se estima que el costo será de 6,000 LPS (*Calculadora de precios / Microsoft Azure, s. f.*).

Software Especializado (0 LPS):

- *Plataforma de Desarrollo:* Se utilizará Python como lenguaje de programación y Visual Studio Code como entorno de desarrollo integrado (IDE). Ambos son software libre y gratuito, por lo que no se incurren en costos en esta parte del proyecto.

Python “es un lenguaje de programación de propósito general de alto nivel, interpretado, multiplataforma y orientado a objetos” (*Welcome to Python.Org, 2023*). Es un lenguaje popular para el desarrollo de aplicaciones de aprendizaje automático, ya que es fácil de aprender y usar, y tiene una gran biblioteca de herramientas y módulos.

Visual Studio Code es un IDE gratuito y de código abierto desarrollado por Microsoft. Es una herramienta potente y flexible que se puede utilizar para desarrollar una amplia gama de aplicaciones, incluidos proyectos de aprendizaje automático (*Visual Studio Code - Code Editing*).

Redefined, s. f.)

El uso de Python y Visual Studio Code para el desarrollo del modelo predictivo permitirá al banco ahorrar costos y obtener una solución de alta calidad.

6.6.4.2 Recursos Humanos:

Salario del empleado que desarrollará y mantendrá el modelo (35,000 LPS)

Esto nos da un total de 105,000 LPS puesto que se está contemplando 3 meses desde su contratación.

El costo total del proyecto de implementación de un modelo predictivo para el banco se estima en 165,000 LPS.

La propuesta de MVP es totalmente viable, ya que los costos son relativamente bajos y los beneficios potenciales son significativos. El modelo predictivo puede ayudar a la institución a mejorar su eficiencia y rentabilidad, lo que podría generar un ahorro de millones de LPS.

Además, la propuesta utiliza software libre y gratuito, lo que reduce aún más los costos. El uso de Python y Visual Studio Code para el desarrollo del modelo también permite al banco contar con un equipo interno de desarrollo que puede mantener y mejorar el modelo en el futuro.

En conclusión, la propuesta de MVP es una solución viable y rentable que puede ayudar al banco a mejorar sus operaciones.

6.7 CONCORDANCIA DE LOS SEGMENTOS DE LA TESIS

La propuesta se alinea integralmente con los distintos segmentos de la tesis, consolidando la validación, extensión del alcance y contribución a la investigación:

Tabla 14-Concordancia de los segmentos de la tesis con la propuesta

Segmento de la Tesis	Propuesta	Concordancia
Problema de Investigación	¿Es posible predecir el comportamiento y el valor monetario de la cartera en mora de tarjetas de crédito del 80% de los bancos comerciales de Honduras a partir de variables como la cantidad de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes?	El problema de investigación se aborda directamente con la propuesta, ya que se busca desarrollar un modelo de aprendizaje automático que permita predecir el valor de la cartera en mora de tarjetas de crédito a partir de la evaluación de la cantidad de tarjetas de crédito activas y su relación.
Preguntas de Investigación	¿Existe una vinculación estadísticamente significativa y lineal de al menos el 75% entre las variables cantidad promedio de tarjetas de crédito activas en circulación, cantidad de tarjetas emitidas por mes y el crecimiento de la cartera en mora de tarjeta de crédito?) ¿Es factible el desarrollar un modelo de machine learning que permita predecir el valor de la cartera en mora con una precisión estadística mínima del 85% de certeza a partir del número de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes para cada institución bancaria estudiada?	Las dos preguntas de investigación se abordan directamente con la propuesta. La primera pregunta se aborda mediante la estimación de la relación entre la cantidad promedio de tarjetas de crédito activas en circulación y el crecimiento de la cartera en mora. La segunda pregunta se aborda mediante la evaluación de la precisión del modelo de aprendizaje automático desarrollado.
Objetivo General	Diseñar e implementar un modelo de aprendizaje automático para predecir el valor de la cartera en mora de tarjetas de crédito a partir de la evaluación de las variables independientes (cantidad de tarjetas activas en circulación y cantidad de tarjetas colocadas por mes) y su relación, en el sector bancario comercial de Honduras.	El objetivo general se aborda directamente con la propuesta, ya que se busca desarrollar un modelo de aprendizaje automático que cumpla con los requisitos establecidos en el objetivo general.
Hipótesis Nula	La cantidad promedio de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes no cuenta con una vinculación estadísticamente significativa y lineal de al menos el 75% con relación al crecimiento de la cartera en mora de tarjeta de crédito lo que permite generar un modelo de predicción de cartera en mora para todos los bancos estudiados.	La hipótesis nula se aborda directamente con la propuesta, ya que se busca probar su falsedad mediante la evaluación de la relación entre la cantidad promedio de tarjetas de crédito activas en circulación y el crecimiento de la cartera en mora.
Hipótesis de Investigación	La cantidad promedio de tarjetas de crédito activas en circulación y la cantidad de tarjetas emitidas por mes cuenta con una vinculación estadísticamente significativa y lineal de al menos el 75% con relación al crecimiento de la cartera en mora de tarjeta de crédito lo que permite generar un modelo de predicción de cartera en mora para todos los bancos estudiados	La hipótesis alternativa se aborda directamente con la propuesta, ya que se busca probar su veracidad mediante la evaluación de la relación entre la cantidad promedio de tarjetas de crédito activas en circulación y el crecimiento de la cartera en mora.
Variables	cantidad de tarjetas activas en circulación, cantidad de tarjetas colocadas, periodo, valor de cartera de tc y valor de cartera en mora de tc	Las variables identificadas en la tesis se utilizan en la propuesta para el desarrollo del modelo de aprendizaje automático.
Enfoque y Métodos	Enfoque cuantitativo, Método de investigación experimental, Modelo de aprendizaje automático.	El enfoque y los métodos identificados en la tesis se utilizan en la propuesta para el desarrollo y evaluación del modelo de aprendizaje automático.

Fuente: (Elaboración propia)

Como se puede observar en la matriz anterior la propuesta presentada es concordante con los segmentos de la tesis en un alto grado. La propuesta aborda directamente el problema de investigación, las preguntas de investigación, el objetivo general, las hipótesis y las variables identificadas en la tesis. Además, la propuesta utiliza el enfoque y los métodos identificados en la tesis para el desarrollo y evaluación del modelo de aprendizaje automático.

Por lo tanto, la propuesta es una solución viable para el problema de investigación planteado en la tesis y proporciona información relevante en los siguientes ámbitos:

Validación de Resultados:

La propuesta se fundamenta en la validación de los resultados obtenidos a través del modelo predictivo previamente desarrollado. Este proceso garantiza la confiabilidad y efectividad de las predicciones, respaldando la toma de decisiones basada en datos.

Extensión el Alcance:

La implementación propuesta va más allá de los confines teóricos, llevando los hallazgos de la investigación a la práctica mediante el uso de dos modelos, uno de aprendizaje automático para desarrollarlo basado en data histórica y otro para aplicar lo aprendido en datos nuevos para generar proyecciones y ayudar a la dirección del negocio; Al integrar el modelo en el entorno operativo de las instituciones financieras, se amplía significativamente el alcance de la investigación, proporcionando una aplicación tangible y beneficiosa.

Contribuir a la Investigación:

La propuesta destaca por su contribución sustancial al campo de investigación. Al aplicar los descubrimientos en un contexto operativo real, la propuesta agrega valor al conocimiento existente, ofreciendo una perspectiva práctica que enriquece el ámbito de estudio.

En conclusión, esta propuesta busca una integración efectiva del modelo predictivo en las operaciones cotidianas de las instituciones financieras. Con el propósito de mejorar la toma de decisiones y la gestión de la cartera en mora de tarjetas de crédito, por lo cual se presenta como una extensión práctica y valiosa de la investigación llevada a cabo durante la tesis de maestría.

REFERENCIAS BIBLIOGRÁFICAS

- 2020 CNBS - GPUF - Educación Financiera. (s.f.). Obtenido de cnbs.gob:
<https://www.cnbs.gob.hn/educacionfinanciera/prestamos-y-creditos/la-tarjeta-de-credito/>
- Banco de Integración Económica (BID). (2023). *Banco Interamericano de Desarrollo (BID). Perspectivas económicas de América Latina 2023. Washington, D.C.: BID, 2023.*
- Bard, G. (4 de Diciembre de 2023). *Google Bard*. Obtenido de <https://bard.google.com/>
- BCH. (2022). *Informe de Estabilidad Financiera*. Departamento de estabilidad Financiera.
- BCIE. (2023). *Banco Centroamericano de Integración Económica (BCIE). Informe de coyuntura económica de Centroamérica, Panamá y República Dominicana. Tegucigalpa, Honduras: BCIE, 2023.*
- Benavides, S. (2000). *Modelos predictor de la morosidad con variables macroeconómicas*.
- BID. (2021). *Banco Interamericano de Desarrollo*. Obtenido de <https://publications.iadb.org/publications/spanish/document/El-impacto-de-las-tarjetas-de-credito-en-la-economia-de-América-Latina.pdf>
- ccichonduras.org. (s.f.). Obtenido de https://www.ccichonduras.org/website/Descargas/LEYES/LEYES_ADMINISTRATIVAS_ESPECIALES/DECRETO_NO__24_2008__LEY_DE_PROTECCION_AL_CONSUMIDOR.PDF
- CEPAL. (2023). *Comisión Económica para América Latina y el Caribe (CEPAL). Panorama de la economía de América Latina y el Caribe 2023. Santiago de Chile: CEPAL, 2023.*
- COLADE. (2023). *FELEBAN.NET*. Obtenido de <https://www.felaban.net/colade/tarjetas>
- El Heraldo. (s.f.). *Publicaciones y Noticias S.A*. Obtenido de <https://www.elheraldo.hn/economia/aumento-cartera-tarjetas-credito-variacion-interanual-honduras-EG12500015#:~:text=Las%20instituciones%20bancarias%20tienen%20a,244%20tipos%20tarjetas%20de%20cr%C3%A9dito.>
- History of Credit Cards*. (s.f.). Obtenido de DinersClub: www.dinersclub.com/about-us/history
- Mankiw, N. G. (2017). *Principles of economics. Cengage Learning*.
- Mayer, D. S. (2007). *“The Impact of Homeownership on the Use of Payday Loans”*.
- Mitchell, T. (1997). *Machine Learning*. McGraw Hill.
- rankinglatam/shopify. (2023). *rankingslatam*. Obtenido de <https://rankingslatam.com/es->

la/blogs/industry-news/credit-card-market-in-honduras-2023-03-rankings
 (2015). *Reporte FINLAC*.

Santana, G. A. (2014). Metodología para el pronóstico de la demanda en ambientes multiproducto y de alta variabilidad. *researchgate.net*.

The History of Visa. (s.f.). Obtenido de Visa Inc.: www.usa.visa.com/about-visa/our-history

admin. (2023, septiembre 1). ¿Qué es Power BI?: Ventajas y tipos de licencias. *EVOTIC | Transformación Digital Inteligente*. <https://evotic.es/business-intelligence-bi/que-es-power-bi-ventajas-y-tipos-de-licencias/>

Anguren Martín—2011—*Credit Cycles Evidence Based on a Non Linear Mode.pdf*. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.bde.es/f/webbde/GAP/Secciones/Publicaciones/InformesBoletinesRevistas/RvistaEstabilidadFinanciera/12/May/Fic/ref2012226.pdf>

Arias, E. R. (s. f.). Muestreo aleatorio. *Economipedia*. Recuperado 7 de diciembre de 2023, de <https://economipedia.com/definiciones/muestreo-aleatorio.html>

BBVA. (2015, diciembre 11). *Teoría de la oferta y la demanda*. *BBVA NOTICIAS*. <https://www.bbva.com/es/teoria-la-oferta-la-demanda/>

Boletín Científico: UAEH. (s. f.). Recuperado 6 de diciembre de 2023, de <https://www.uaeh.edu.mx/scige/boletin/huejutla/n8/m1.html>

Buitrago, B. (2020, septiembre 15). *Machine Learning—Modelos de Regresión I. iWannaBeDataDriven*. <https://medium.com/iwannabedatadriven/machine-learning-modelos-de-regresi%C3%B3n-i-d293ae235e9a>

Burrows, L. (2023, julio 11). *What Is Credit Risk Management*. *Experian Insights*. <https://www.experian.com/blogs/insights/credit-risk-management/>

Calculadora de precios | Microsoft Azure. (s. f.). Recuperado 8 de diciembre de 2023, de <https://azure.microsoft.com/es-es/pricing/calculator/?service=virtual-machines>

Caribe, C. E. para A. L. y el. (2023). *Estudio económico de Centroamérica y la República Dominicana en 2023 y perspectivas para 2024*. *Comisión Económica para América Latina y el Caribe*. <https://www.cepal.org/es/publicaciones/68687-estudio-economico-centroamerica-la-republica-dominicana-2023-perspectivas-2024>

Classification in Machine Learning: A Guide for Beginners. (s. f.). Recuperado 4 de diciembre

de 2023, de <https://www.datacamp.com/blog/classification-machine-learning>

Comprendiendo la Matriz de Congruencia—Incorruptible. (2023, marzo 11).
<https://incorruptible.mx/que-es-una-matriz-de-congruencia/>

Credit Card Debt & its Effect on the Economy | Pocketsense. (s. f.). Recuperado 2 de diciembre de 2023, de <https://pocketsense.com/credit-card-debt-its-effect-economy-3111.html>

Credit Card Industry in 2023: Analysis and Trends in Payment Processing. (s. f.). Recuperado 2 de diciembre de 2023, de <https://www.insiderintelligence.com/insights/credit-card-industry/>

Credit risk definition—Risk.net. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.risk.net/definition/credit-risk>

Curso de implementación de servicios en Microsoft Azure | UdeMy. (s. f.). Recuperado 8 de diciembre de 2023, de <https://www.udemy.com/course/implementacion-de-soluciones-en-microsoft-azure/>

Cutler, A., Cutler, D., & Stevens, J. (2011). *Random Forests.* En *Machine Learning—ML* (Vol. 45, pp. 157-176). https://doi.org/10.1007/978-1-4419-9326-7_5

Difference Between Classification and Regression In Machine Learning—Dataaspirant. (2014, septiembre 27). <https://dataaspirant.com/classification-and-prediction/>

Drozd, L. (2021). *Why Credit Cards Played a Surprisingly Big Role in the Great Recession.* *Economic Research.* (s. f.). *San Francisco Fed.* Recuperado 5 de diciembre de 2023, de <https://www.frbsf.org/economic-research/>

Figuroa, A. (2011). *La desigualdad del ingreso y los mercados de crédito.* *Revista de la CEPAL,* 2011(105), 39-55. <https://doi.org/10.18356/ada2b51b-es>

Figuroa-Céspedes, I., Lambiasi Pérez, R., Cáceres Zapata, P., Figuroa-Céspedes, I., Lambiasi Pérez, R., & Cáceres Zapata, P. (2022). *Actitud lúdica y rol mediador de aprendizajes en educadoras de párvulos: Para aprender jugando se necesitan dos.* *Revista de estudios y experiencias en educación,* 21(47), 371-386. <https://doi.org/10.21703/0718-5162202202102147020>

Financial Portfolio: What It Is, and How to Create and Manage One. (s. f.). *Investopedia.* Recuperado 5 de diciembre de 2023, de <https://www.investopedia.com/terms/p/portfolio.asp>

Guanajuato, U. de. (2021, diciembre 7). *Clase digital 2. Riesgo y rendimiento.* *Licenciatura en*

Contador público. <https://blogs.ugto.mx/contador/clase-digital-2-riesgo-y-rendimiento/>

Guillén-Franco, E., & Peñafiel-Chang, L. (2018). Modelos predictor de la morosidad con variables macroeconómicas. *Revista Ciencia Unemi*, 11(26), 13-24.

Heidegger y la técnica en el mundo moderno. (2021, diciembre 27). <https://filco.es/heidegger-tecnica/>

Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. McGraw Hill España.
<https://dialnet.unirioja.es/servlet/libro?codigo=775008>

Hernandez-Sampieri.-Metodologia-de-la-investigacion.pdf. (s. f.). Recuperado 7 de diciembre de 2023, de <https://perio.unlp.edu.ar/catedras/wp-content/uploads/sites/151/2021/08/Hernandez-Sampieri.-Metodologia-de-la-investigacion.pdf>

IEF junio 2023.pdf. (s. f.). Recuperado 2 de diciembre de 2023, de <https://www.bch.hn/estadisticos/EF/LIBINFORMEEF/IEF%20junio%202023.pdf>

Infografía: La penetración del pago a crédito en América Latina. (2019, junio 14). *Statista Daily Data*. <https://es.statista.com/grafico/18346/uso-de-tarjeta-de-credito-para-pagos-a-plazo-en-latinoamerica>

Informes de Estabilidad Financiera. (s. f.). Recuperado 14 de diciembre de 2023, de <https://www.bccr.fi.cr/publicaciones/informes-de-estabilidad-financiera>

Installments: What it is that and how it works in Latin America. (2019, diciembre 14). *LABS English*. <https://labsnews.com/en/articles/payments/installments-the-key-to-consumers-pockets-in-latin-america/>

International Monetary Fund. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.imf.org/external/error.htm?URL=https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Global-Financial-Stability-Report-2015-Update-Vulnerabilities-Legacies-and-Policy-Challenges-43020>

Juárez, E. (2023, junio 11). Tarjetas de crédito bancarias, con su mayor crecimiento en 15 años. *El Economista*. <https://www.economista.com.mx/sectorfinanciero/Tarjetas-de-credito-bancarias-con-su-mayor-crecimiento-en-15-anos-20230611-0041.html>

Lavalle et al. - Análisis didáctico de regresión y correlación para.pdf. (s. f.). Recuperado 7 de diciembre de 2023, de <https://www.scielo.org.mx/pdf/relime/v9n3/v9n3a4.pdf>

Lex Fridman (Director). (2019, enero 11). Deep Learning Basics: Introduction and Overview. <https://www.youtube.com/watch?v=O5xeyoRL95U>

Ley de Protección al Consumidor Honduras, ¿para qué sirve y cuándo se aplica? | Tu Nota. (s. f.). Recuperado 6 de diciembre de 2023, de <https://www.tunota.com/honduras-hoy/articulo/ley-de-proteccion-al-consumidor-honduras-para-que-sirve-y-cuando-se-aplica-2022-09-05>

Ley de Protección de Datos Personales.pdf. (s. f.). Recuperado 6 de diciembre de 2023, de <https://cei.iaip.gob.hn/doc/Ley%20de%20Proteccion%20de%20Datos%20Personales.pdf>

Ley de Transparencia y Acceso a la Información Pública y su Reglamento. (s. f.). Recuperado 6 de diciembre de 2023, de <https://www.tsc.gob.hn/biblioteca/index.php/leyes/17-ley-de-transparencia-y-acceso-a-la-informacion-publica>

Leyes y Normativas. (s. f.). Comisión Nacional de Bancos y Seguros. Recuperado 6 de diciembre de 2023, de <https://www.cnbs.gob.hn/leyes-y-normativas/>

Mercado de Pagos de América Latina, Tamaño, Informe 2024-2032. (s. f.). Recuperado 4 de diciembre de 2023, de <https://www.informesdeexpertos.com/informes/mercado-de-pagos-de-america-latina>

Morales—2020—El papel del crédito en el ciclo económico. Una in.pdf. (s. f.). Recuperado 5 de diciembre de 2023, de https://www.depfe.unam.mx/especializaciones/revista/2-1-2020/07_HPE_Ortega-Morales_2020.pdf

Naveen. (2020, julio 20). What is Supervised Learning? Intellipaat. <https://intellipaat.com/blog/what-is-supervised-learning/>

Payment Methods—Adyen. (s. f.). Recuperado 4 de diciembre de 2023, de <https://www.adyen.com/payment-methods?paymentMethodTypes=Cards>

(PDF) Estudio de oferta y demanda de la tarjeta de crédito como instrumento de pago, en la ciudad de Loja, año 2010 | Verónica Paredes—Academia.edu. (s. f.). Recuperado 5 de diciembre de 2023, de https://www.academia.edu/77179363/Estudio_de_oferta_y_demanda_de_la_tarjeta_de_credito_como_instrumento_de_pago_en_la_ciudad_de_Loja_a%C3%B1o_2010

Plans and Pricing—Choose a Package | DataCamp. (s. f.). Recuperado 8 de diciembre de 2023, de <https://www.datacamp.com/pricing>

¿Qué es el aprendizaje no supervisado? (s. f.). TIBCO Software. Recuperado 4 de diciembre de 2023, de <https://www.tibco.com/es/reference-center/what-is-unsupervised-learning>

¿Qué es machine learning? | IBM. (s. f.). Recuperado 4 de diciembre de 2023, de <https://www.ibm.com/es-es/topics/machine-learning>

Quiroa, M. (s. f.). Aprendizaje supervisado. Economipedia. Recuperado 4 de diciembre de 2023, de <https://economipedia.com/definiciones/aprendizaje-supervisado.html>

Ramírez—Económicos y del crédito en América Central y la R.pdf. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.cemla.org/PDF/ic/2018-pic-xix/2018-pic-xix3.pdf>

Ramon, A. (2014, octubre 28). Sistemas de predicción. Machine Learning. SPRI. <https://www.spri.eus/es/teics-comunicacion/sistemas-de-prediccion-machine-learning/>

Random Forest. (s. f.). Corporate Finance Institute. Recuperado 4 de diciembre de 2023, de <https://corporatefinanceinstitute.com/resources/data-science/random-forest/>

Random Forest Regression in Python Explained | Built In. (s. f.). Recuperado 24 de noviembre de 2023, de <https://builtin.com/data-science/random-forest-python>

ResearchGate Link. (s. f.). Recuperado 5 de diciembre de 2023, de https://www.researchgate.net/publication/323430569_Credit_Risk_Research_Review_and_Agenda

Surti—Riesgo y rendimiento La búsqueda de rentabilidad.pdf. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.imf.org/external/pubs/ft/fandd/spa/2021/06/pdf/risk-return-search-for-yield-surti-basics.pdf>

The Definition of Credit and Why You Need It—NerdWallet. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.nerdwallet.com/article/finance/what-is-credit>

The Impact Of Credit Card Usage On Consumer Spending And The Economy. (s. f.). <https://Outlookmoney.Com/>. Recuperado 2 de diciembre de 2023, de <https://outlookmoney.com/outlook-money-spotlight/the-impact-of-credit-card-usage-on-consumer-spending-and-the-economy-8846>

Understanding Credit—Financial Aid & Scholarships. (s. f.). Recuperado 5 de diciembre de 2023, de <https://financialaid.berkeley.edu/financial-literacy-and-resources/understanding-credit/>

Vera—Determinantes de la demanda de crédito. Una estima.pdf. (s. f.). Recuperado 5 de diciembre de 2023, de <https://www.scielo.org.mx/pdf/ineco/v62n245/0185-1667-ineco->

62-245-107.pdf

Visa, Mastercard penetration Latin America 2022. (s. f.). Statista. Recuperado 4 de diciembre de 2023, de <https://www.statista.com/statistics/734326/credit-cards-market-share-latin-america/>

Visual Studio Code—Code Editing. Redefined. (s. f.). Recuperado 8 de diciembre de 2023, de <https://code.visualstudio.com/>

Welcome to Python.org. (2023, diciembre 8). Python.Org. <https://www.python.org/about/>

What is a Random Forest? (s. f.). NVIDIA Data Science Glossary. Recuperado 4 de diciembre de 2023, de <https://www.nvidia.com/en-us/glossary/data-science/random-forest/>

What Is Machine Learning? Definition, Types, and Examples. (2023, noviembre 29). Coursera. <https://www.coursera.org/articles/what-is-machine-learning>

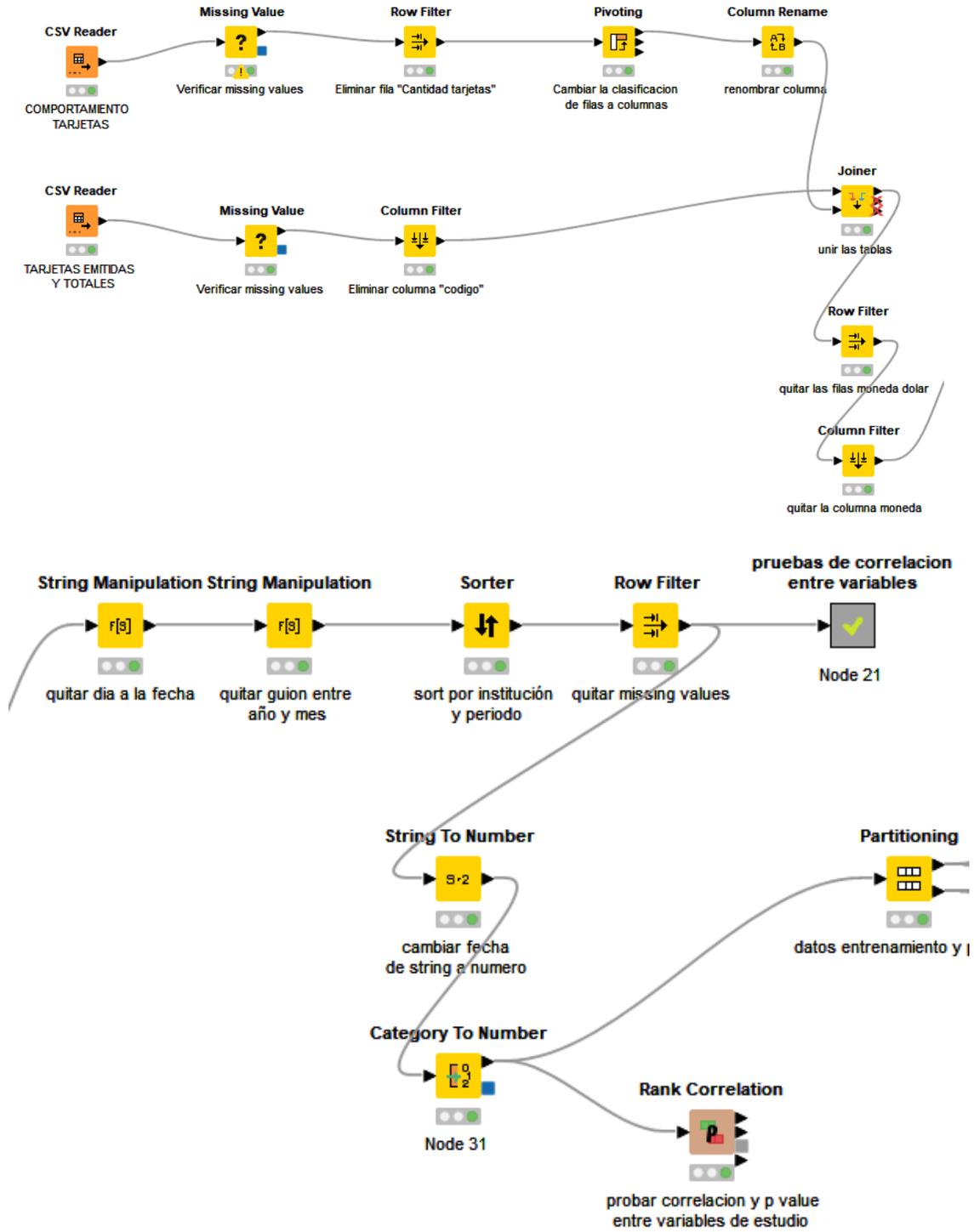
When Were Credit Cards Invented: The History of Credit Cards – Forbes Advisor. (s. f.). Recuperado 2 de diciembre de 2023, de <https://www.forbes.com/advisor/credit-cards/history-of-credit-cards/>

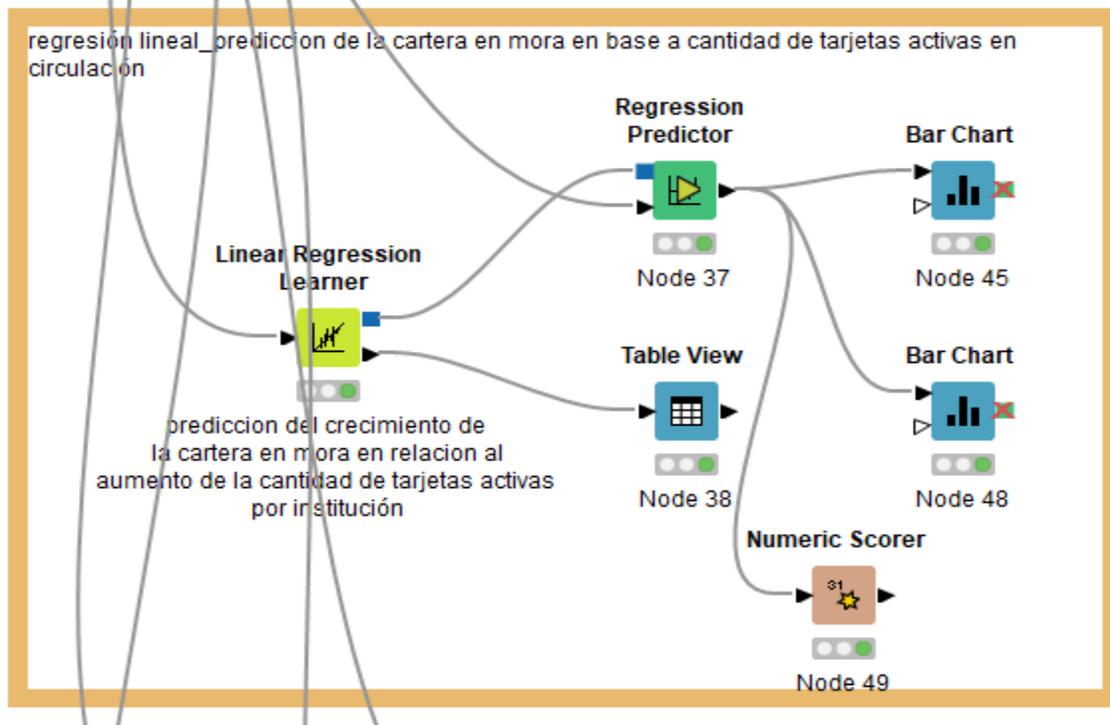
X Edición del Informe de Tendencias en Medios de Pago. (2021, enero 28). Minsait. <https://www.minsait.com/es/actualidad/insights/tendencias-en-medios-de-pago>

Zamore, S., Ohene Djan, K., Alon, I., & Hobdari, B. (2018). Credit Risk Research: Review and Agenda. Emerging Markets Finance and Trade, 54, 811-835.
<https://doi.org/10.1080/1540496X.2018.1433658>

ANEXOS

ANEXO 1 : PROGRAMACIÓN KNIME, PREPARACIÓN DE DATA PARA MODELOS.

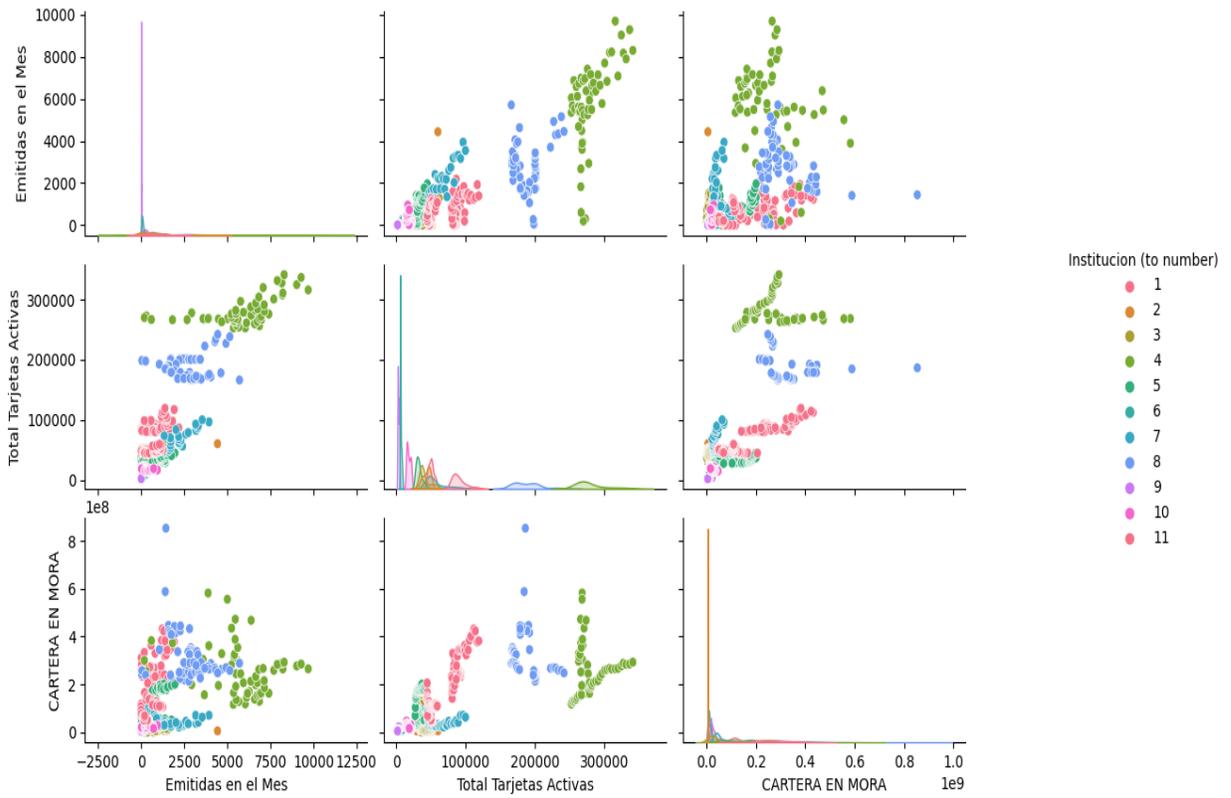




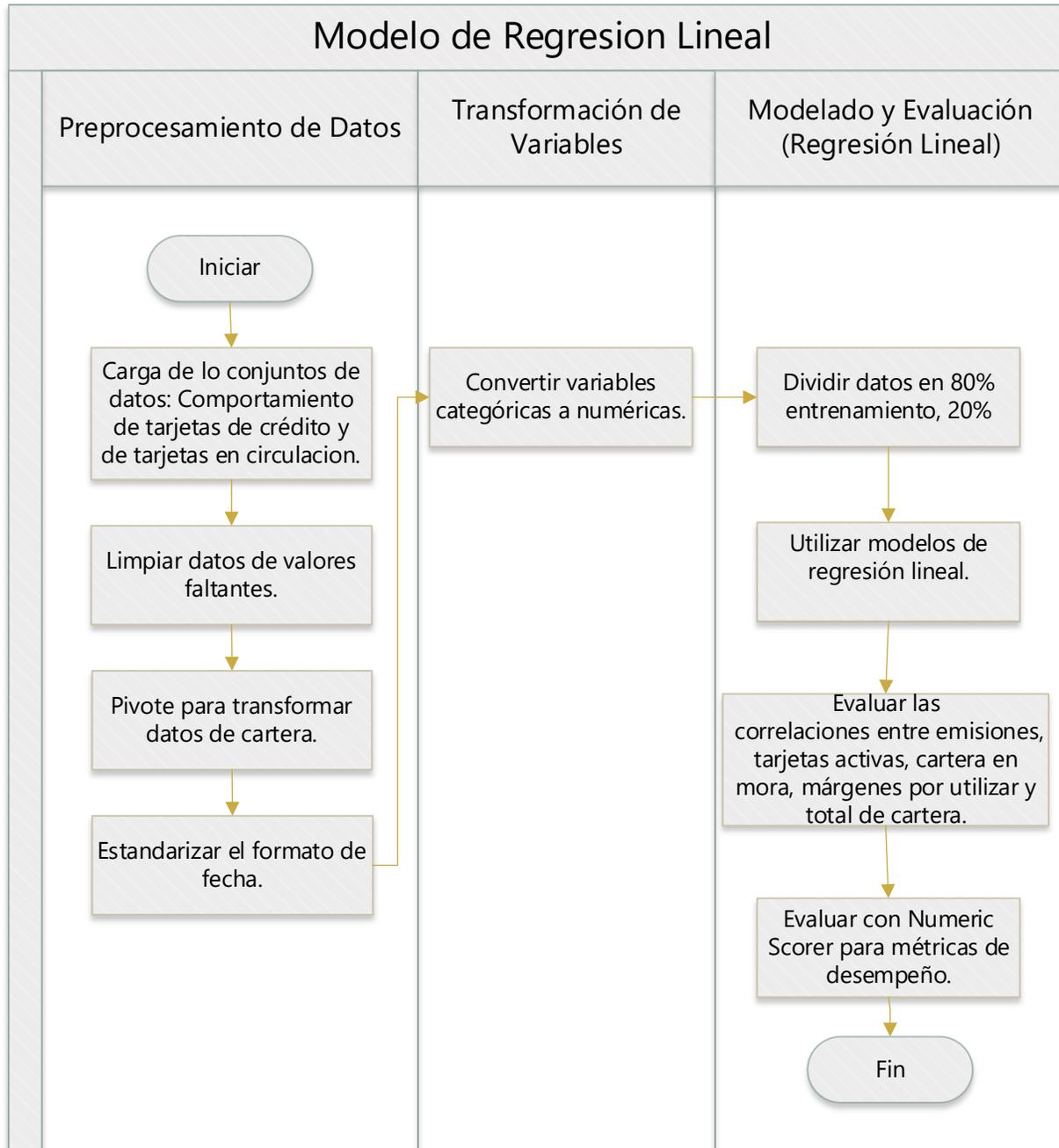
ANEXO 2 : DASHBOARD EDA EN MS POWERBI.



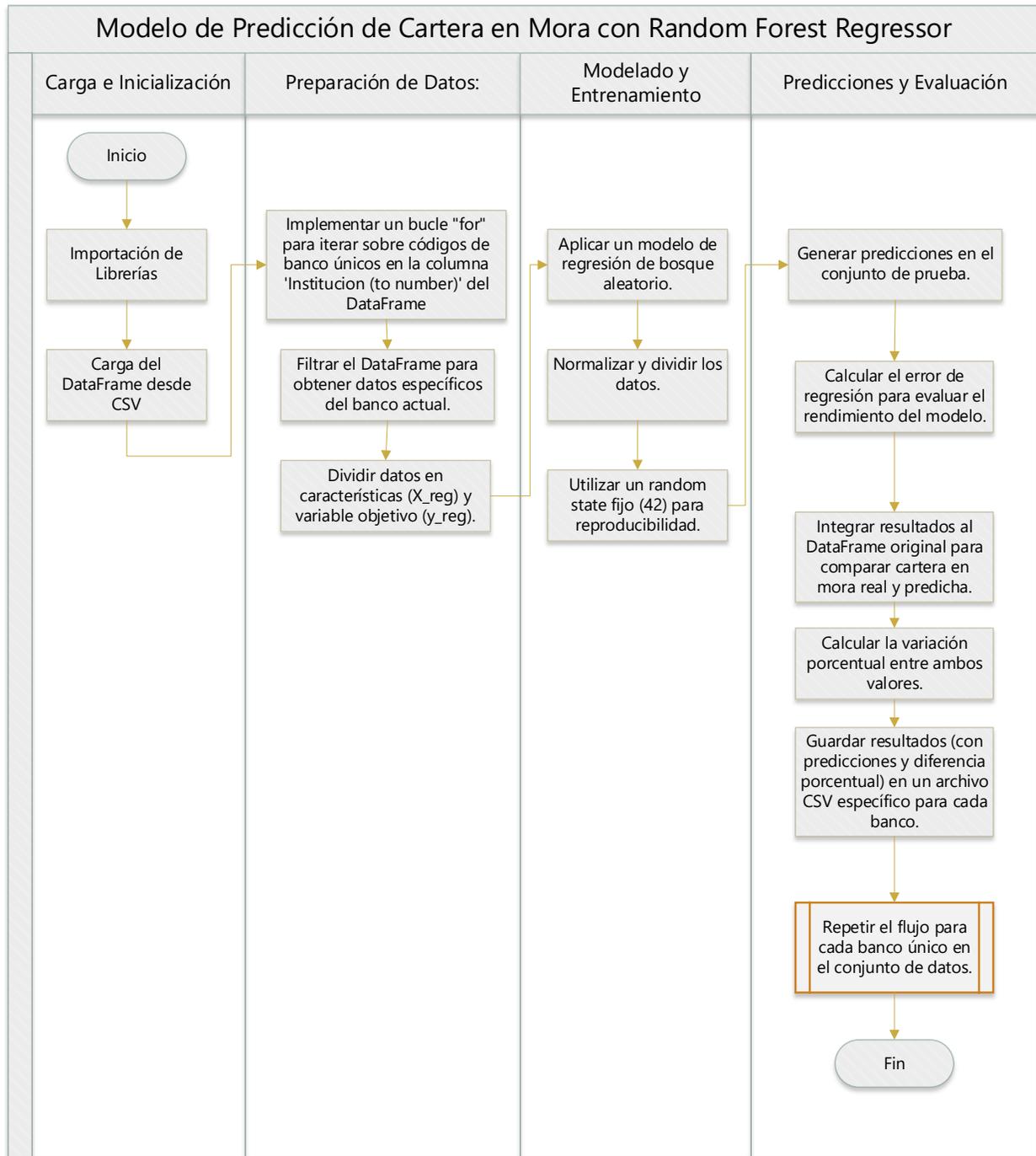
ANEXO 3: CORRELACIÓN ENTRE VARIABLES USANDO PYTHON MATPLOTLIB



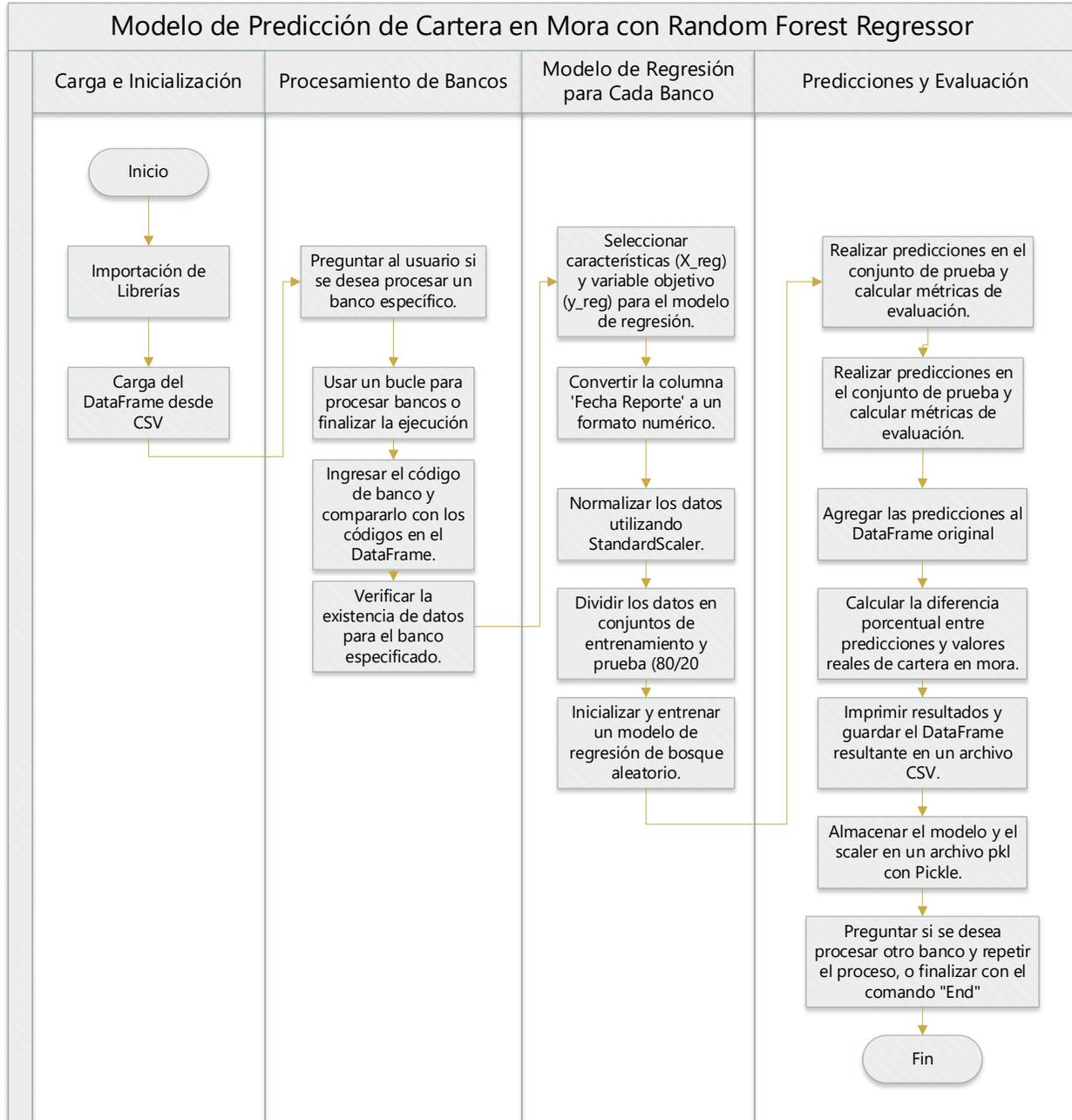
ANEXO 4:FLUJO MODELO DE REGRESIÓN LINEAL



ANEXO 5: FLUJO MODELO PREDICCIÓN CARTERA EN MORA CON RANDOM FOREST REGRESSOR



ANEXO 6: FLUJO MODELO PREDICCIÓN CARTERA EN MORA CON RANDOM FOREST REGRESSOR



ANEXO 7 :FLUJO PREDICCIÓN DE CARTERA EN MORA CON EL MODELO ENTRENADO.

