



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE CIENCIAS

TESIS

Aplicación y análisis del modelo de
Credit Scoring para mujeres jóvenes

Alumna:

Itzel Coba

Cuacuas

Actuaria

Asesor:

Dra. Nora

Gavira Durón

Ciudad de México

2020





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ÍNDICE

- INTRODUCCIÓN 4
 - I.1 Prólogo
 - I.2 Objetivo General
 - I.3 Objetivos Específicos
 - 1.4 Hipótesis
 - 1.5 Marco Temporal y Espacial
 - I.6 Marco metodológico

CAPÍTULO I

- MARCO TEÓRICO 12
 - I.1 Antecedentes del Riesgo de crédito
 - I.2 Definiciones
 - I.3 Basilea
 - I.4 Modelo Credit Risk
 - I.5 Aspectos a considerar en la evaluación de un crédito
 - I.6 Factores y causas que miden el Riesgo de Crédito.
 - I.7 Modelos de Credit Scoring

CAPÍTULO II

- SISTEMA CREDITICIO EN MÉXICO 27
 - II.1 Tarjeta de crédito
 - II.2 Situación actual de la cartera crediticia en México
 - II.3 Microfinancieras en México y su impacto
 - II.4 Requisitos para solicitar un microcrédito

CAPÍTULO III

- METODOLOGÍA PARA CALCULAR SCORE55
 - III.1 Técnicas para evaluar el riesgo crediticio
(Modelo de Regresión lineal, Modelo Proit y Modelo Logit)
 - III.2 Base de datos

CAPÍTULO IV

- SELECCIÓN DE VARIABLES78
 - IV.1 Descripción de las variables

CAPÍTULO V

- APLICACIÓN DE MODELO A CASO REAL84

CAPÍTULO VI

- ANÁLISIS DE RESULTADOS 99

CAPÍTULO VI

- RESULTADOS Y CONCLUSIONES 106
- BIBLIOGRAFIA..... 109

INTRODUCCIÓN

Estudiar y analizar la forma en que cada país, ciudad o entidad realiza la administración de riesgos financieros es interesante porque de acuerdo a cada situación se implementan distintos métodos y medidas. Actualmente las instituciones financieras están cada vez más presentes en el mundo actual, y relacionadas entre sí, esto a causa de que una entidad financiera puede ser cliente de otra, que a su vez es cliente de una tercera provocando su expansión geográfica; se hace cada vez más necesario reforzar la regulación de su actividad financiera. La regulación para este tipo de entidades debe ser muy estricta, ya que las instituciones financieras tienen cada vez más influencia en la sociedad, y su objetivo es mantener una estabilidad monetaria y financiera, así como la protección del cliente. Es por ello que una entidad financiera debe identificar los riesgos a los que está expuesta, debe saber cuantificarlos y ser capaz de mitigarlos o protegerse de ellos en caso de que sea necesario.

Entendiendo como riesgo la probabilidad de ocurrencia de un hecho que tuviera consecuencias adversas para la entidad. Los riesgos a los que se enfrentan se agrupan en cuatro grandes bloques: riesgo de mercado, riesgo de crédito, riesgo de liquidez y riesgo operacional. Todos ellos

afectan en mayor o menor medida a una entidad financiera pero para este texto nos enfocaremos en el riesgo de crédito.

El término conocido como 'microfinanza' es una herramienta que permite el acceso sostenible a servicios financieros a personas de bajos recursos y de esta manera les sirve como apoyo ante la posibilidad de aumentar sus ingresos, invertir en bienes o tener mayor capacidad para hacer frente a sus obligaciones. De igual manera permite que tengan un mejor manejo financiero, planificar su futuro e invertir en mejores condiciones de vida, educación y salud.

En la actualidad, las pequeñas y medianas empresas constituyen el mayor sector empresarial del país. Sin embargo, aunque estas resultan ser de gran importancia en la economía, no cuentan con el apoyo adecuado que estimule el desarrollo y crecimiento de estas empresas. Una de las principales limitaciones es el acceso a las fuentes de financiamiento formales para así permitirles un mayor desarrollo de sus operaciones. Esto se debe a la falta de información financiera provista por estas empresas, así como, por no contar con un historial financiero que permita evaluar su capacidad de pago.

Por lo tanto, con el conocimiento de esta situación, surge un gran interés en identificar y evaluar los riesgos financieros a los que se encuentran expuestas estas empresas, de esta manera se obtiene el conocimiento requerido para proponer y crear instrumentos que les permitan actuar

oportunamente y de esta forma, tomar buenas decisiones para su empresa.

Para estimar el riesgo del crédito y mitigar las posibles pérdidas, se han diseñado modelos estadísticos, en base a un conjunto de atributos, utilizados como variables en los modelos, que permiten calificar objetivamente la calidad de un cliente. Sin embargo, estos modelos se han desarrollado para las grandes instituciones relegando a empresas de menor tamaño.

Debido al incremento de la presencia de mujeres en el sector financiero, que a temprana edad han buscado la forma de mostrar su importancia en la economía del país, resulta atractivo estudiar en este proyecto como es el comportamiento de este sector de la población en relación con el riesgo de crédito en microfinanzas.

I.1 Prólogo

Administración de riesgos es el término que hace referencia al proceso aplicado por las instituciones financieras de forma sistemática para identificar, evaluar, analizar, monitorear y mitigar los riesgos a los que se encuentran expuestas, de una forma que permita a las instituciones minimizar pérdidas y maximizar oportunidades.

Las instituciones financieras se especializan en la acumulación de capital y se encargan de transferirlo por medio de inversiones ó préstamos a interés. Los bancos son las instituciones financieras más conocidas, y basan sus operaciones en la captación de ahorro y otorgamiento de crédito. Por otra parte, las microfinancieras son instituciones dirigidas en específico a un sector de bajos recursos, debido a que éstos difícilmente pueden cumplir con los requisitos que exige la banca tradicional, ofreciendo servicios de ahorro y crédito.

En 2001, en México, empezó el Programa Nacional de Financiamiento al Microempresario (Pronafim), el cual agrupa la prestación de servicios de aproximadamente 200 microfinancieras. Los créditos que éste tipo de institución puede otorgar tiene un máximo de 30,000 Udis¹ (190,000 pesos

1 UDIS – (Unidades de Inversión) Son unidades de valor que se basan en el incremento de los precios (la inflación) y son usadas para solventar las obligaciones de créditos.

aproximados en enero de 2020), a un plazo no mayor de tres años con pagos semanales y una tasa de interés que depende de cada caso.

El incumplimiento de créditos provoca considerables pérdidas a las instituciones financieras al no tener la capacidad de respuesta ante estas situaciones, y si estas instituciones no tienen una adecuada administración de riesgos se podría convertir en una posible causa de ruina. Para prevenir este tipo de problemas existen compañías aseguradoras que ofrecen a estas entidades un seguro para este tipo de riesgo.

Para que las instituciones financieras puedan ofrecer un crédito a sus clientes, primero deben medir el riesgo de impago de las personas o entidades que solicitan el crédito. Se tiene que realizar un análisis de riesgo y construir o aplicar modelos que se adecuen a sus escenarios, para que estos les indiquen si pueden aceptar o no este tipo de riesgos. Una de las dificultades a las que se enfrentan es desarrollar un modelo en el cuál se calcule la probabilidad de los casos de incumplimiento de pago con los datos que se tengan del cliente ya que algunas veces resultan insuficientes. Además, evidenciar los riesgos y cuantificarlos mejor, al ajustar y readecuar los sistemas de evaluación actuales no es una tarea sencilla. Anteriormente, los sistemas de calificación utilizados eran

únicamente cualitativos; donde se sabía que un crédito “A” era menos riesgoso que un crédito “B”, pero no se tenía más información.

De esta manera, el proyecto que se desarrolla tiene la finalidad de analizar la modelación y cuantificación del riesgo, que contienen sólidos fundamentos matemáticos. Adicionalmente se pretende encontrar alguna aplicación que nos permita valorar el riesgo de crédito de un caso específico de una microfinanciera.

La importancia de abordar este problema surge del notable aumento actual de créditos otorgados, estas instituciones deben ir a la vanguardia, y así lograr que su cartera de clientes se incremente de manera razonable con el menor riesgo posible, para esto se deben buscar el modelo que mejor se ajuste y mida el nivel de riesgo al que se expone.

I.2 Objetivo General

Valorar el riesgo de crédito en una institución microfinancieras e implementar técnicas actuariales y financieras basadas en conocimientos matemáticos y estadísticos, realizando una correcta aplicación de un modelo de score para la cuantificación del riesgo de crédito con los datos recabados y enfocados en un grupo de mujeres jóvenes.

I.3 Objetivos Específicos

- Recopilar información sobre el riesgo de crédito, historia y causas de la existencia de los riesgos de crédito a los que nos encontramos expuestos en la actualidad.
- Investigar y recopilar información referente a modelos de cuantificación del riesgo, así como, ver la ejecución de modelos matemáticos actuales para la valoración del riesgo de crédito que cuenten con fundamentos matemáticos firmes.
- Aplicar el modelo de score para obtener la cuantificación del riesgo de crédito de la institución.
- Brindar conclusiones y observaciones que permitan aplicar los resultados de este análisis para así dar opciones de mejoras.

1.4. Hipótesis

Aplicar un modelo Logit sobre una específica muestra histórica de datos, como metodología crediticia permite predecir el riesgo de crédito o estimar las probabilidades de impago de los clientes de una institución microfinanciera.

1.5. Marco Temporal y Espacial.

El desarrollo del presente proyecto fue llevado a cabo durante 2019 y 2020 con datos reales de 2016 y 2017 de créditos otorgados en el país, recopilados de una microfinanciera inglesa con operaciones en cinco países incluyendo México.

1.6 Marco metodológico

Se genera un modelo de scoring para la cartera de microcredito de una microfinanciera inglesa que tiene operación en diversos estados de México. El modelo está enfocado en el proceso de otorgamiento de créditos, por lo que se categoriza como modelo de originación.

La aplicación del modelo y análisis se realiza con la información proporcionada por la microfinanciera, estos datos se toman como base para la construcción del modelo. La institución clasificó previamente a sus clientes en buenos y malos según sus criterios establecidos internamente. El modelo Logit y las tareas que se aplican son de acuerdo a los conocimientos obtenidos durante la carrera, reforzados con investigaciones propias.

CAPÍTULO I.

MARCO TEÓRICO

En este capítulo se describirán los antecedentes más relevantes referentes al riesgo de crédito, así como las definiciones relevantes y necesarias para la comprensión del tema.

I.1 Antecedentes del Riesgo de crédito

A lo largo de la historia el riesgo de crédito ha tomado gran relevancia debido a dos razones principales, primero, por ser considerado como el riesgo más antiguo cuyo origen data desde 1800 a.C., y en segundo lugar se debe al número de casos elevados de un mal manejo y administración del riesgo provocando enormes pérdidas.

Desde inicios de la historia el hombre se ha enfrentado a la incertidumbre, debido a que en muchas ocasiones no puede prevenir, cuantificar o eliminar ciertas situaciones, es decir, que el hombre siempre ha estado expuesto al riesgo. La curiosidad humana hacia el juego y las apuestas se ha canalizado al crecimiento económico, el mejoramiento de la calidad de vida, y el progreso tecnológico. Sin embargo, sin el

conocimiento matemático, estadístico, el dominio de la probabilidad y de otras herramientas de la administración de riesgos, los ingenieros nunca hubieran diseñado los grandes puentes que cruzan los ríos más anchos, no existirían aeroplanos, las enfermedades y epidemias seguirían afectando a gran parte de la población e incrementándose, etc. Las complejidades del mundo actual nos hacen estar prevenidos de descomposturas, accidentes y errores, en términos muy simples, siempre estamos expuestos al riesgo en nuestra vida cotidiana.

Los estudios sobre riesgo de crédito están relacionados ampliamente con la probabilidad y la estadística, ya que estos son las bases para los modelos aplicados en este ámbito. Se puede decir que los primeros estudios relacionados con probabilidad fueron desarrollados en el siglo XVI. Sin embargo, el avance en estos temas fue lento en un principio, más adelante, en la época del Renacimiento se iniciaron los estudios más serios sobre el riesgo.

Hasta 1654, donde Pascal fué retado a resolver un acertijo de un juego de apuesta, se originó la teoría de probabilidades. Años más tarde, en 1703 Jacob Bernoulli inventó la Ley de los grandes números y métodos de muestreo estadístico.

Relacionado con la probabilidad en el año de 1725 los matemáticos construyen las primeras tablas de esperanza de vida y emerge el negocio de los seguros marítimos en Londres, el cual se considera el primer tipo de seguro que se utilizó en el mundo.

En el año de 1730, surge la idea de la distribución normal por el matemático francés Abraham de Moivre y define el concepto de desviación estándar. Thomas Bayes y Joseph Lagrange continuaron con los estudios de Moivre y realizaron grandes contribuciones en el campo de la probabilidad, pero sería hasta 1812 que Pierre-Simon Laplace define por primera vez el concepto de probabilidad en su obra conocida como "*Théorie analytique des probabilités*". Años más tarde, en 1952 se demuestra de forma matemática las ventajas de la diversificación del riesgo por Markowitz. En 1994, se propuso una metodología denominada Riskmetrics por la empresa JP Morgan, y el concepto de "valor en riesgo" como modelo para medir cuantitativamente los riesgos de mercado en instrumentos financieros o portafolios con varios tipos de instrumentos.

I.2 Definiciones

- **Microfinanzas**

Este término se refiere a la acción de otorgar a las personas o familias de bajos recursos, servicios financieros como apoyo económico en actividades productivas.

- **Microfinancieras**

Organizaciones o instituciones, que brindan servicios microfinancieros como ahorro, crédito, seguros, remesas y pago de servicios; estos pueden ser individuales o grupales. También pueden ofrecer servicios no financieros (capacitación, asesoría, de negocio).

- **Riesgo**

El significado de “Riesgo” está completamente relacionado con términos como pérdida, siniestro o peligro. Sin embargo, para los seres humanos el riesgo es parte inevitable en los procesos de toma de decisiones y por esto es necesario reconocer las probabilidades de caer en riesgo. La

medición de éste se establece de acuerdo a la probabilidad asociada a una pérdida en contextos de incertidumbre.

- **Riesgo de crédito**

Se define como la posible pérdida debida a la probabilidad de incumplimiento de pago ocasionado por cambios en la capacidad o intención de la contraparte de cumplir sus obligaciones.

Es importante relacionar al riesgo con el ámbito financiero ya que se puede convivir con él a través de un incentivo, es decir, aceptaremos más riesgo en la medida que haya más recompensa, es por ello que existe una relación muy estrecha entre riesgo y rentabilidad. Para una institución financiera la inversión crediticia representa el mayor porcentaje de la inversión total de su actividad, de ahí la importancia de estudiar y minimizar la incertidumbre al conceder los préstamos o créditos.

Se determina que existen dos tipos de riesgo de crédito: el riesgo de incumplimiento, que se refiere a la pérdida potencial derivada de que la contraparte no pueda cumplir con sus obligaciones financieras en las condiciones definidas contractualmente; y el riesgo de mercado que se define como la pérdida potencial que podría sufrir un tenedor de un portafolio de

préstamos, instrumentos financieros o derivados, como consecuencia de que el valor de mercado de estos disminuya. La segunda definición, plantea exposición al riesgo de crédito aún en el caso de que la contraparte no sufra quebranto alguno. (Comité de Supervisión Bancaria de Basilea, 1999 pp52.)

Existe un riesgo inherente en el otorgamiento de un crédito, el cual se deriva de dos tipos de pérdidas: Pérdida Esperada y Pérdida No Esperada.

- **Riesgo de Incumplimiento**

El riesgo de incumplimiento se considera como la probabilidad de que haya alguna informalidad por parte del acreditado en el momento que le corresponde realizar su pago.

- **Riesgo de Exposición**

Se genera por la incertidumbre respecto a los montos probables a riesgo, especialmente cuando los créditos tienen la libertad de liquidarse de manera anticipada y no existe penalización, ya que no se conoce con exactitud el plazo de pago.

- **Modelos de Cuantificación del Riesgo**

Son modelos utilizados para la medición del riesgo de crédito que buscan cuantificar el riesgo de enfrentar el incumplimiento de pago de los créditos otorgados, es decir, su objetivo es medir la probabilidad de incumplimiento. El cálculo de esta probabilidad tiene implícitas bases matemáticas para generar aplicaciones en la vida real de la teoría financiera y económica.

- **VaR de Crédito**

El Valor en Riesgo (VaR) es un modelo estadístico utilizado para evaluar el riesgo de una determinada posición o cartera, basado en la teoría de probabilidad. Es la máxima pérdida esperada de incumplimiento del crédito otorgado dado un horizonte de tiempo específico a un determinado nivel de confianza.

- **Pérdida esperada**

Resulta del crédito que se estima perder debido al deterioro actual de la cartera y se determina con la calidad de cada uno de los acreditados por

medio de su calificación. Estas deben ser cubiertas con las reservas preventivas de la institución financiera.

- **Pérdida no esperada**

Estima cuanto capital debe de reservar la institución para asegurar su solvencia. Está compuesta por riesgos crediticios en la cartera que no se consideran como probables o estimables en la fecha de evaluación, pero existe el riesgo de ocurrencia en el futuro.

- **Default**

Cuando se aplica un modelo de scoring, cuyo propósito es identificar a los clientes malos y separarlos de los buenos, primero se define que es bueno o malo, respecto a cada institución. Se consideran dos factores para esta definición: cuando supera un límite en los días de atraso en un pago; y el desempeño que se observa en la peor situación de dicho cliente en un determinado lapso de tiempo. La marca de mora¹ determina la situación de default durante el periodo de desempeño.

1 Mora – Situación en la que el acreditado se ha retrasado en el pago de la deuda o de los intereses acumulados. Se considera una situación de alto riesgo para el acreedor debido a que está a un paso de caer en default.

- **Administración de riesgos**

Es un proceso aplicado sistemáticamente por las Instituciones financieras para identificar, analizar, medir, vigilar, limitar, controlar, revelar y dar tratamiento a los distintos riesgos a los que se encuentran expuestas tanto ellas como sus subsidiarias financieras.

I.3 Basilea

Es una organización internacional la cual comenzó sus actividades en 1930, fomenta la cooperación financiera y monetaria a nivel internacional, además funciona como banco para los bancos centrales.

En el año de 1988 se publicó el primer acuerdo de Basilea I debido a la baja calificación crediticia de las entidades, donde se eligió el modelo de regulación prudencial, éste establecía el cálculo de requerimientos de capital con la intención de que fuera aplicado a nivel mundial. Posteriormente en 2004, el Comité de Supervisión Bancaria publica el acuerdo definitivo internacional sobre regulación y supervisión bancaria denominado “Nuevo Acuerdo de Capital”, por el Comité de Basilea, y conocido como Basilea II. Esto resultó ser un gran avance en la gestión de riesgos porque obliga a los sistemas bancarios a capitalizarse y a determinar su exposición al riesgo.

Basilea II estableció los tres pilares básicos para la administración de riesgos.

- Pilar 1, denominado "requerimientos mínimos de capital", en el que se determina el coeficiente de solvencia o capital mínimo regulador.
- Pilar 2, conocido como "revisión supervisora", en el que se comprueban que los criterios seguidos en el pilar 1 se llevan a cabo correctamente.
- Y, por último, Pilar 3 o "disciplina del mercado", basado en la transparencia informativa entre todos los componentes del sistema financiero.

Sin embargo, la metodología de Basilea para la regulación de capital resulta un tanto deficiente ya que no toma en cuenta la diversificación y la calidad crediticia de la contraparte. Se deben tomar diversos parámetros de medida dependiendo de las condiciones económicas de cada país, para solucionar este tema, durante los últimos años se han desarrollado diferentes metodologías cuantitativas para la medición del riesgo permitiendo cuantificar la exposición al riesgo de acuerdo al tipo de producto, concentración geográfica, entre otros factores. Algunos de estas metodologías son CreditMetrics y CreditRisk.

I.4 Modelo Credit Risk

Credit Risk es el nombre del modelo de cuantificación del riesgo creado en 1997 por la compañía Credit Suisse Financial Products (CSFP), éste modelo resuelve muchas de las críticas que han recibido los esquemas regulatorios actuales para la medición de las pérdidas por riesgo. Tiene la función de medir el riesgo de impago sin realizar supuesto alguno sobre las causas que originaron el incumplimiento, considerando a éste como una variable aleatoria continua e incorporando la volatilidad de la tasa con la finalidad de tomar en cuenta la incertidumbre.

En conclusión, el modelo calcula la distribución completa de las pérdidas esperadas de un portafolio que se encuentra expuesto al riesgo de crédito previendo situaciones como el posible volumen de exposición, la calidad del crédito y el riesgo de incumplimiento por parte del deudor. Después de aplicar el modelo el resultado obtenido nos indica la reserva que se debe considerar, es decir, determina el monto de capital que se requiere para cubrir el riesgo de la pérdida por incumplimiento de los deudores y no caer en ruina.

I.5 Aspectos a considerar en la evaluación de un crédito

En el proceso de evaluación para otorgar un crédito se debe realizar una serie de pasos para obtener una evaluación profunda y adecuada tanto de los aspectos cualitativos como cuantitativos del cliente. Es necesario considerar el antecedente histórico del cliente en de la misma institución y en las demás instituciones donde haya solicitado un crédito.

Se evalúan y analizan los créditos mediante las probabilidades de pérdida que están dadas por la incertidumbre acerca de los factores y variables que pueden afectar en el futuro a los clientes y que puedan aumentar el peligro de cumplimiento del pago.

De acuerdo con Basilea II, las entidades financieras y de microfinanzas supervisadas deben adoptar procesos que sean capaces de medir internamente el riesgo de crédito, además deben actuar bajo los parámetros de Basilea II. Con estos parámetros se ven obligados a disponer de herramientas adecuadas que les permitan elegir, modelos de medición con el propósito de discriminar a los clientes según su perfil de riesgo, modelos de evaluación, de exposición y definir la severidad en el riesgo de crédito.

Por otra parte, esta organización obliga a las entidades financieras a adaptar sus sistemas de cálculo del consumo de capital, además deben

modificar sus sistemas de análisis y reporting. Debido a la importancia de los informes financieros y el análisis de los mismos, éstos son la clave de Basilea II para gestionar grandes bases de datos, capaces de ofrecer la información exacta para cuantificar los riesgos de cada operación, lo que supone un verdadero reto para los bancos y, especialmente, para las entidades de microfinanzas.

I.6 Factores y causas que miden el Riesgo de Crédito.

Dentro del riesgo de crédito influyen diversos factores que es necesario conocer para lograr estimar la probabilidad de pérdida de manera más precisa. Estos factores son heterogéneos, y aunque existen algunos de carácter general, en su gran mayoría se individualizan por tipo de cliente. Por otra parte, la calidad de una cartera de préstamos depende del nivel de riesgo crediticio, que ésta a su vez depende esencialmente de dos tipos de factores:

- Factores internos, que dependen directamente de la administración propia y/o capacidad de los ejecutivos de cada institución. Entre éstos están: el volumen de crédito y las políticas de créditos.

- Factores externos, que no dependen de la administración, tales como depreciación, inflación, desastres o fenómenos naturales, etc.

Si el análisis de crédito no se realiza correctamente, se estará expuesto a una mayor probabilidad de pérdida, al otorgar créditos a personas equivocadas, lo que da como resultado que la probabilidad de incumplimiento será muy alta. Sin embargo, aún después de otorgar el crédito, es decir, en el proceso de formalización, seguimiento y control, sigue estando presente la probabilidad de pérdida.

Una vez que se han analizado los factores, éstos deben ser considerados para evaluar el nivel de riesgo de crédito, debido a que forman parte del resultado de la experiencia del historial de otorgamiento de préstamos a estos clientes.

Algunas de las principales causas del riesgo de crédito son: insolvencia por parte del cliente, inconsistencias jurídicas en la formalización y errores operativos o humanos en el proceso de tramitación.

I.7 Modelos de Credit Scoring

El Scoring estadístico se basa en los antecedentes históricos de pago del cliente y analiza comportamientos pasados para pronosticar

comportamientos futuros de los créditos. Un modelo de Scoring utiliza la misma lógica que el analista de crédito, pues se basa en experiencias y seguimientos de créditos otorgados en el pasado, mediante un análisis de las características de los nuevos solicitantes, con el fin de calificar o descalificar los perfiles. Un analista puede identificar si otorga un crédito a un desempleado basándose en su experiencia, pero estos casos resulta tener un alto riesgo de caer en incumplimiento, debido a que la decisión del analista de crédito depende en gran medida de lo que éste considere “bueno” o “malo”.

Es un método para evaluar el riesgo de crédito basándose en datos históricos de solicitudes de préstamos o créditos ya otorgados o rechazados anteriormente. El objetivo es aislar el efecto de una serie de características personales o propias del producto y relacionarlo con la probabilidad de impago del cliente, utilizando datos históricos y técnicas estadísticas. El modelo da como resultado una puntuación o "score" que la entidad puede utilizar para calificar al cliente y con esta tomar una decisión respecto a la concesión o no.

Los modelos más conocidos para el cálculo de riesgo de crédito son: Regresión lineal, Análisis de discriminante, Regresión logística, Modelos Logit, Modelos Probit.

CAPÍTULO II.

SISTEMA CREDITICIO EN MÉXICO

Debido a los problemas económicos a los que nos enfrentamos día con día, es necesario buscar medios alternos que nos permita alcanzar nuestras perspectivas económicas, tanto individuales como colectivas. Es cuando se recurre a solicitar un crédito, ya sea para cumplir con las necesidades básicas de subsistencia, para mejorar el nivel social, o para llevar a cabo un negocio entre otras.

Durante la prehistoria de México, el trueque fue la primera forma de pago que se utilizó para el comercio entre artesanos y mercaderes, además del pago en trabajo agrícola. Sin embargo, el primer antecedente certero del surgimiento de las operaciones crediticias o de préstamos se dieron hasta la época colonial, en esta época había pocas instituciones de crédito ya que no se contaba con tanta riqueza y no existían bancos, solamente la Casa de Moneda donde se controlaba el flujo monetario.

Cuatro de las ramas del crédito fueron impulsadas por los principales prestamistas coloniales:

- Agrícola, impulsado por la Iglesia.
- Hipotecario, donde principalmente influyo la Iglesia.
- Industrial, impulsado tanto por particulares como por algunos proyectos de la Corona.
- Prendario, a cargo de pequeños particulares, pero mayormente involucrado el Monte de Piedad de Animas.

Para el año de en 1775, la fundación del Sacro Monte de Piedad de Animas, resultó relevante para la historia del crédito en México, ya que al poco tiempo de haberse fundado, la institución realizaba cuatro clases principales de operaciones; préstamos con garantía de prenda, custodia de depósitos confidenciales, admisión de secuestros o depósitos judiciales y mandados por otras autoridades, por último, venta pública en Almoneda de las prendas no desempeñadas ni refrendadas.

Esta institución comenzó sin cobrar intereses con el objetivo de beneficiar al prestatario. Sin embargo, en 1782 comenzó a cobrar un impuesto sobre los depósitos a modo de interés equivalente al 6.25% anual. Para 1815 se incrementó al doble esta cantidad debido a las afectaciones que sufrió a causa de la guerra de Independencia. En el año de 1821, se habían otorgado un promedio de 35,000 préstamos cuyos valores fluctuaban entre 400,000 y 600,000 pesos, obteniendo una utilidad aproximada de 15,000 a 20,000 pesos por año.

Por otra parte, para quitarle poder a la iglesia y su relación con algunos comerciantes, se crearon las reformas borbónicas. España fundó instituciones de crédito, así, en 1782 se creó el Banco Nacional de San Carlos que estaba destinado a fomentar el comercio en general. Mientras que en 1784 se creó el Banco de Avío de Minas, se dedicaba principalmente al crédito de avío, el cual resulta ser el antecedente del Banco de México Independiente.

Después de la Independencia, en el siglo XIX, la situación económica en México fue en declive debido a la gran deuda al nuevo gobierno y las fuentes crediticias se agotaron, produciendo una crisis económica lo que llevó al comienzo de una gran deuda externa e interna. Durante la segunda mitad de este siglo, se inicia el surgimiento de nuevas instituciones de crédito y los primeros bancos. En el gobierno de Maximiliano se creó el Banco de Londres, el primer banco comercial del país, con la finalidad de solucionar la crisis económica por la que se estaba atravesando. Por otra parte el Monte de Piedad tomó también la función de banco de emisión de moneda y depósito.

El nacimiento de los bancos como instituciones de crédito surgió durante el régimen de Porfirio Díaz, debido a la necesidad de la creciente industria. Sin embargo, se generó un descontrol en la creación de bancos, fue entonces que se decidió promulgar el Código de Comercio

con el cual se retiró la facultad para emitir billetes. Como consecuencia surgió el Banco Nacional de México, éste sería el único con la facultad de emitir billetes.

Con el objetivo de reorganizar el sistema bancario, en 1897 fue creada la Ley General de Instituciones de Crédito. Donde se establecía que solo los bancos de emisión serían instituciones de crédito, es decir, únicamente los bancos hipotecarios ofrecían créditos a largo plazo. En los ámbitos relacionados con agricultura, industria y manufacturas, para el otorgamiento de créditos, se recurriría a los bancos refaccionarios ¹.

En 1994, ocurrió la devaluación del peso mexicano, hubo un incremento de las tasas de interés, la inflación y adicionalmente se presentó una caída en la actividad económica de los salarios reales en el país. Debido a este suceso, el sistema financiero mexicano se vio en una gran crisis económica donde parte de esta problemática se debía a la quiebra virtual del sistema bancario mexicano provocado por la alta exposición al riesgo de crédito al que las instituciones financieras estaban expuestas.

¹ *Bancos Refaccionarios - Aquellos establecimientos destinados a facilitar las operaciones mineras, agrícolas e industriales, por medio de préstamos privilegiados, pero sin hipoteca, otorgando su garantía para operaciones determinadas, y emitiendo títulos de crédito, con intereses y son pagaderos en día fijo.*

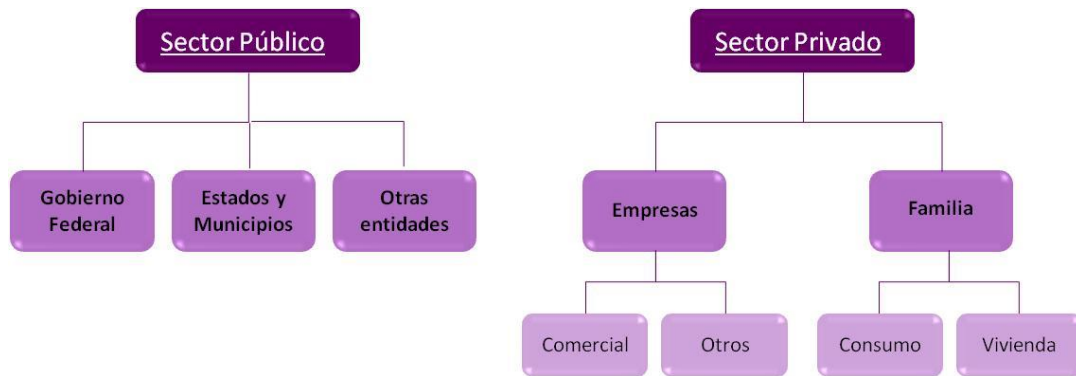
Seguido de la crisis en el país, el crédito de la banca comercial al sector privado cayó en más del 40% en tan solo un año, y se necesitaron de cinco años para recuperarse y que nuevamente creciera.

El crecimiento real del PIB fortaleció el crecimiento del crédito al sector privado, así, para el año 2006 se presentó un crecimiento real anual del 30% reflejado principalmente en el aumento del crédito al consumo, sin dejar de lado la aportación del crédito a las empresas y a la vivienda ya que facilitan el flujo de recursos hacia las empresas, mediante las ventas de productos y servicios.

En 2014 se estableció un sistema financiero regulado por la Comisión Nacional Bancaria y de Valores, buscando que se ofrezcan prestamos mayores a menores costos. El crédito a las empresas ha tomado una tendencia ascendente en los últimos años, impulsada en gran parte por el crecimiento del crédito al sector servicio, donde su participación pasó de alrededor del 50% en 2003 a más de 56% a inicios de 2007.

Actualmente el mercado de crédito en México se puede clasificar en dos principales sectores, el sector público y el sector privado, donde el primero se divide en crédito al Gobierno Federal, a Estados y Municipios y a otras entidades de Gobierno. Por otra parte, tenemos el crédito privado dividido en crédito a empresas (crédito comercial y crédito a otras

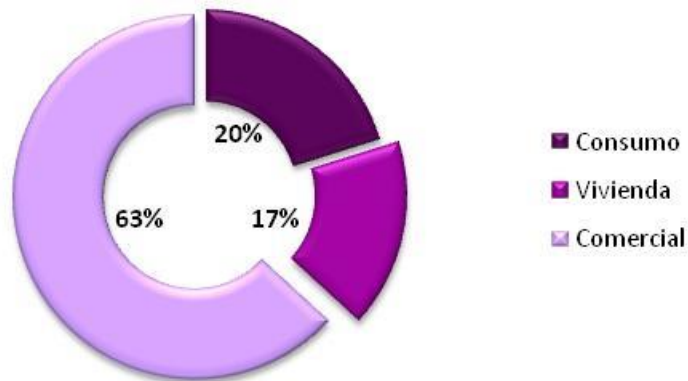
entidades financieras como intermediarios financieros no bancarios) y crédito a las familias (crédito al consumo y a la vivienda).



Elaboración propia.

Gráfico 01.

Integración del Crédito al Sector Privado 2019



Elaboración propia con datos de CNBV.

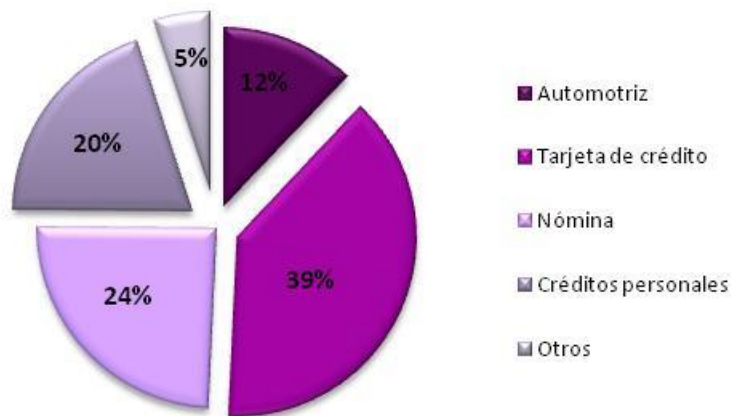
Gráfico 02.

En esta gráfica se ilustra la composición del crédito al sector privado separado por tipo de crédito, donde se observa como predomina el crédito comercial.

El canal más importante para el financiamiento al sector privado es la banca múltiple. Existen distintos tipos de crédito al consumo de la Banca, que se clasifican en: tarjeta de crédito, nómina, personales, automotriz y otros. Para el otorgamiento de este tipo de crédito es necesario realizar una solicitud de crédito con ciertos documentos como la identificación del solicitante y comprobante de ingresos. Seguido de la entrega de dichos

documentos, se realiza una investigación verificando la reputación del solicitante respecto a sus compromisos de crédito con otros intermediarios y se procede a un análisis de los datos para determinar si se otorga o no el crédito.

Estructura del Crédito al Consumo de la Banca Múltiple por Producto 2019



Elaboración propia con datos de CNBV.

Gráfico 03.

En la gráfica anterior se muestra la integración de la cartera de crédito al consumo de la banca por producto, donde se puede observar que el producto predominante es la tarjeta de crédito.

México se rige por las reglas de capitalización de Basilea en la cual se establecen los requisitos mínimos de capital, 8% sobre la suma de activos ponderados por su riesgo de crédito, que deben cumplir las instituciones financieras para hacer frente al riesgo de crédito asumido. Fue necesario adecuar este acuerdo debido a que en el país han existido varias crisis financieras cuyo origen en gran parte ha sido por el mal desempeño y administración de la cartera de crédito de los bancos, para calcular la exposición sujeta a riesgo de los activos, para esto, basándonos en los criterios de Basilea se utilizan ponderaciones desde 0% para las menos riesgosas, hasta el 100% para las de mayor riesgo.

Actualmente los préstamos o créditos bancarios son de gran utilidad como apoyo para adquirir productos con altos costos que no podrías pagar de contado fácilmente, como puede ser la compra de una casa, un automóvil o el capital necesario para iniciar un negocio. Sin embargo, para obtener un crédito es necesario contar con características que te acrediten como una persona responsable con capacidad e intención de pago.

Debido a que ya se han establecido reglas generales para el otorgamiento de créditos, las cuales son impuestas por la Secretaría de Hacienda y Crédito Público y la Comisión Nacional Bancaria y de Valores, se debe cumplir con lo siguiente:

- **Edad.**- Tener entre 25 y 65 años, debido al alto riesgo de incumplimiento que existe a los 85 años por la alta probabilidad de fallecimiento.
- **Comprobante de domicilio.**- Debe presentarse a nombre del solicitante.
- **Antigüedad en el domicilio mínima de dos años.**- Esto muestra estabilidad en el solicitante.
- **Solicitud de otorgamiento de crédito.**- Llenar la solicitud con la información correspondiente.
- **Registro Federal de Contribuyentes.**- Demostrar que se está afiliado por un mínimo de dos años.
- **Contabilidad positiva.**- En una empresa su contabilidad debe ser positiva al menos los últimos dos o tres años, es decir, no debe presentar pérdidas significativas en su patrimonio.
- **Acta Consecutiva y Poderes Legales.**- Dan facultad al solicitante para realizar los actos correspondientes.
- **Obligado solidario.**- En caso de ser una empresa el solicitante, los obligados del principal accionista serán el resto de los socios y sus

cónyuges sí se encuentran casados bajo el régimen de sociedad conyugal. En caso de las personas físicas con actividad empresarial, el obligado solidario puede ser cualquier persona.

- **Comprobante de ingresos.**- Garantiza que el cliente cuenta con la solvencia económica necesaria para realizar el pago.

El crédito es fundamental para el crecimiento de la banca en México, pero éste se debe hacer con un adecuado procedimiento al otorgar los créditos para así asegurar su crecimiento de manera sostenida durante un largo periodo. Presentar un alto número de créditos es un indicio del crecimiento en la actividad económica del país, pero el crecimiento de créditos comerciales también puede interpretarse como deudas significativas en algunas compañías.

En el país existen tres Sociedades de Información Crediticia (SIC's) autorizadas por la SHCP, por otra parte Banco de México se encarga de emitir las reglas para estas sociedades y a su vez son reguladas por la Comisión Nacional Bancaria y de Valores (CNBV).

- **Trans Union.**- para personas físicas.
- **Dun & Bradstreet.**- para personas morales.

- **Círculo de Crédito.**- se encarga de administrar la información de personas físicas y morales.

Estas instituciones se encargan de recopilar información del comportamiento crediticio de los individuos que en algún momento han solicitado crédito. A si mismo ayudan a disminuir el riesgo al que se exponen las entidades financieras cuando toman la decisión de otorgar un crédito, también se otorgan incentivos a los acreditados para que sean puntuales en sus pagos.

Para mitigar el riesgo, la CNBV ha establecido un plazo fijo para determinar cuándo se debe considerar una cartera como vencida provocado por el incumplimiento de los créditos otorgados, este plazo se estableció un máximo de tres meses para la cartera comercial y 180 días para la cartera hipotecaria. Estos plazos pueden ser variables dependiendo de cada institución

II.1 Tarjeta de crédito

Para una institución de crédito una de las principales y más importantes necesidades que tiene es contar con herramientas y criterios confiables para determinar el otorgamiento del crédito. El objetivo siempre es reducir

lo mayormente posible el riesgo al que se está expuesto, sobre todo cuando se pretende adquirir nuevos clientes.

Durante 1949 en Nueva York surgió la primera tarjeta de crédito llamada Diner's Club, al principio la tarjeta solo era aceptada en 14 restaurantes de Nueva York, pero años más tarde, más de 20 mil personas ya la utilizaban y el número de negocios en los cuales eran aceptada crecía cada vez más. En México las tarjetas de crédito comenzaron a aparecer en 1956, sin reglamentos específicos aplicables. Actualmente se intenta legislar reglamentos para proteger a los consumidores, debido a la gran demanda de las tarjetas de crédito y al aumento en la cantidad de clientes que dejan de pagar lo cual ocasiona un gran riesgo para las entidades financieras.

Basado en la información que proporciona Banco de México hace seis años 95% de la población pagaba principalmente en efectivo sus gastos personales como transporte y alimentos; mientras que el 8% realizaba sus pagos con tarjeta de débito, 4% con tarjeta de crédito y un porcentaje menor utilizaba vales de despensa. Para 2017 no se presentó una variación significativa en la distribución de los métodos de pagos antes mencionados, únicamente incremento 2% el pago con tarjeta de débito y 1% con tarjeta de crédito.

La revista mexicana 'Nexos' informó en Abril de 2019 lo siguiente:

De acuerdo con la Encuesta Nacional de Inclusión Financiera (ENIF 2018) del Instituto Nacional de Estadística y Geografía (INEGI), el porcentaje de la población mayor de 18 años que cuenta con al menos un producto financiero —entendido como cuentas, créditos, seguros o ahorro para el retiro— pasó del 56% (39.4 millones) en 2012 al 68% (54 millones) para 2018, lo que representa un incremento del 37%. Aunado a ello, el 45% de los encuestados (35.4 millones) reportó contar con más de un producto de este tipo.

En cuanto a los productos empleados como medios de pago, se observa que el 34% de los mexicanos (26.9 millones) cuenta con una tarjeta de crédito, mientras que alrededor del 47% (37.3 millones) posee una tarjeta de débito asociada a una cuenta bancaria. En este punto, resulta imprescindible tener en cuenta que 6.5 millones de adultos reciben por este medio algún tipo de apoyo gubernamental. Mejía, Omar., Villegas, Rodrigo. (11 de abril de 2019). *La inclusión financiera en México más allá de la bancarización*. <https://economia.nexos.com.mx/?p=2214>.

En nuestro país uno de los principales instrumentos de crédito son las tarjetas de crédito. Para el primer trimestre de 2019, de acuerdo a los reportes de uno de los bancos más grandes de México (Situación Banca México 1S2019 - BBVA) se registraron más de 27 millones de tarjetas de

crédito. Sin embargo, durante el tercer trimestre del mismo año la cifra disminuyó a 24 millones debido a cambios en las metodologías implementadas por algunos bancos del país lo cual genera una variación en la información recopilada.

Aunque existe un alto número de tarjetas emitidas, alrededor del 60% de los mexicanos no poseen una tarjeta de crédito, esto provoca que exista un mercado potencial para este producto.

Del total de la población que cuenta con tarjetas, de acuerdo con el reporte anteriormente mencionado, solo el 8% cumplen puntualmente con el pago de sus préstamos, 9% corresponde a cartera vencida, 50% hace los pagos mínimos y cerca de 30% trabajan con planes de 6 meses sin intereses. Para evitar que una persona presente un desequilibrio en sus finanzas, debe dedicar menos del 35 % del ingreso al pago de sus deudas, de lo contrario lo más probable es que deje de pagar y sus deudas incrementen de forma exponencial.

A pesar de los avances tecnológicos y de su accesibilidad, los medios de pago digital, tarjetas virtuales, transferencias interbancarias, así como pagos con teléfonos inteligentes continúan sin ser explotados para sustituir al efectivo, debido a que en 2019 aproximadamente el 32% de la población realizó algún tipo de transacción por un medio digital. Estamos

muy por debajo del promedio observado en América Latina y el Caribe (46%).

Las tarjetas de crédito se pueden clasificar por su alcance, forma de pago, accesibilidad y objetivo:

- **Alcance:**

- Locales.- Aplican solo en el país de origen.
- Internacionales.- Se pueden ocupar en cualquier país del mundo.

- **Emisor:**

- Departamentales.- Son proporcionadas por establecimientos comerciales para otorgar productos o servicios que ellos ofrecen.
- Bancarias.- Son proporcionadas por bancos para la compra de bienes, servicios a cargo de terceros o proveer dinero en efectivo de manera directa.

- **Forma de pago:**

- Clásica o estándar.- El monto total se cobra después de un mes vencido. El límite establecido no se sobrepasa y se paga a tiempo y en forma, no se cobra interés, de lo contrario se cobrará de acuerdo a la tasa de interés establecida. La línea de crédito es baja al igual que el costo anual de la tarjeta.

Por el riesgo mayor al que se expone con este tipo, el interés es alto.

- Oro.- Similar a la clásica. Se otorga a clientes con ingresos altos, generalmente después de dos años de historia crediticia en la institución.
- Platinum.- Destinada a clientes VIP, los cuales suelen hacer uso muy frecuente de la tarjeta. Tienen línea de crédito y anualidad alta y bajos intereses.

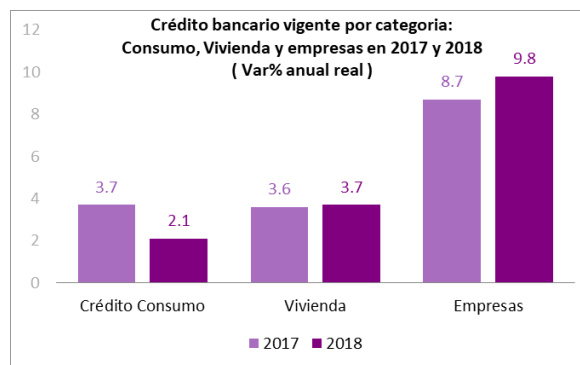
Se manejan tres tipos de interés en las tarjetas de crédito:

- **Revolvente**.- Es la tasa de interés anual que se cobra por hacer uso de la tarjeta. Está en función de la TIIE(Tasa de Interés Interbancaria).
- **Fijo**.- Depende del plazo del cliente, pero no cambia con el tiempo.
- **Cero**.- Tiene cero porcentaje de interés aplicado sobre la deuda durante un periodo determinado.

II.2 Situación actual de la cartera crediticia en México

En relación con los datos registrados en (CNBV, 2012), se observa que el número de créditos otorgados aumentó un 12.8% a pesar de la situación económica de ese periodo. Para 2013 el incremento del crédito total de ese año apenas rebasaba 6.3%. Sin embargo, en los últimos años la actividad económica del país no ha aumentado, y aunque el crédito debería mostrar un crecimiento prácticamente nulo o, incluso, negativo, el crédito ha continuado creciendo, lo cual ha generado problemas de cartera vencida con mayor frecuencia.

En 2018 el mayor crecimiento en la cartera de crédito fue el otorgado a empresas con un promedio real anual de 9.8%, siendo 0,9% arriba del 2017, Por parte del crédito a la vivienda se registró 3.7% para 2018, casi igual comparado con el año anterior (3.6%). En cambio, disminuyó 1.6% el crédito al consumo para 2018 con 2.1%.



Fuente: BBVA Research con datos de Banco e Inegi.

Gráfico 04.

Por otra parte, de acuerdo con datos proporcionados por (CNBV, 2018), para el primer trimestre de 2018, el sector de banca múltiple se integró por 50 instituciones en operación, donde el índice de morosidad de la cartera total se situó en 2.19%, es decir, creció 0.01 puntos porcentuales al nivel observado en 2017.

Durante 2018 la cartera vencida rebasó los 60 mil millones de pesos, el monto más elevado registrado en los últimos años, la cifra superó en 16% a la cantidad registrada en 2017, lo que implicó un incremento de aproximadamente siete mil millones de pesos. Es decir, el monto de los créditos no pagados se ha elevado a un ritmo mucho mayor que el total de los créditos otorgados, lo cual desemboca en un aumento del índice de morosidad¹.

La cartera de las sociedades financieras de objeto múltiple componen a la cartera de crédito total, éstas registraron un incremento anual de 4.6% para 2019 con un saldo de 5,414 mmdp².

¹ *Índice de morosidad - (IMOR) es la cartera de crédito vencida como proporción de la cartera total.*

² *mmdp – miles de millones de pesos.*

En Agosto de 2019, con información recabada en (CNBV, 2019), se aprecia que la cartera de crédito vigente tuvo una participación de 52.1%, del total de activos, con un saldo de 5,296 mmdp, mismo que se incrementó un 4.6% respecto al año anterior. La cartera de crédito vencida tuvo un aumento de 4.2% registrando un saldo de 118 mmdp.

La cartera de crédito comercial creció 4.2% respecto al año anterior, representando el 63.9% de la cartera total con un saldo final de 3,461 mmdp. Por otra parte el crédito al consumo registró un incremento de 3.1% con un saldo de 1,075 mmdp, equivalente al 19.9% del total de la cartera.

Dentro de la categoría de crédito al consumo se registra el crédito que se otorgó a través de tarjetas de crédito, el cual creció 2.5% anual con un saldo de 414 mmdp y una participación de 38.5% del total.

Otro dato importante es el índice de Morosidad el cuál te indica la relación de la cartera vencida entre la cartera total. Para la cartera comercial se obtuvo un nivel de 1.37%, la cartera al consumo registró un nivel de 4.36%, obteniendo así para la cartera de crédito total un IMOR de 2.17%

Cartera de Crédito Total Banca Múltiple Saldos nominales en mmdp	Agosto 2018	Julio 2019	Agosto 2019	Variación real %	
				Anual	Mensual
Cartera de crédito total	5,018	5,336	5,414	4.6	1.5
Créditos comerciales	3,218	3,399	3,461	4.2	1.8
Empresas	2,461	2,675	2,728	7.5	2
Entidades financieras	211	204	205	-5.8	0.4
Entidades gubernamentales	546	519	527	-6.4	1.6
Créditos al gobierno federal o con su garantía	29	30	30	-1.3	-0.6
Créditos a estados y municipios o con su garantía	307	289	292	-7.9	1.3
Créditos a organismos descentralizados o desconcentrados	44	42	47	4.3	13.3
Créditos a empresas productivas del estado	166	160	159	-7.3	-0.5
Consumo	1,010	1,066	1,075	3.1	0.8
Tarjeta de crédito	392	411	414	2.5	0.8
Personales	214	214	215	-2.8	0.5
Nómina	238	259	261	6.3	1
ABCD	145	160	161	7.7	0.8
Automotriz	134	147	148	7	0.7
Adquisición de bienes muebles	12	14	14	16.3	1.9
Operaciones de arrendamiento capitalizable	-	-	-	4.5	-0.1
Otros créditos de consumo	21	23	23	7.2	1.2
Vivienda	789	871	878	7.8	0.8
Media y residencial	688	767	775	9.3	1.2
De interés social	10	9	8	-16.1	-1.3
Créditos adquiridos al INFONAVIT o el FOVISSSTE	88	91	90	-0.7	-1.4
Garantizados por la Banca de Desarrollo o Fideicomisos Públicos	4	5	4	-0.5	-8.2

Fuente: www.gob.mx/cnbv

Gráfico 05.

Cabe mencionar que el incremento que se presenta en la cartera vencida de los créditos otorgados a los clientes es debido al aumento de tres principales factores: la inflación, el desempleo y la tasas de interés, ya que esto incrementa la falta de capacidad de pago por parte de los clientes.

II.3 Microfinancieras en México y su impacto

En México existe una enorme desigualdad económica, por un lado puedes encontrar desde uno de los hombres más ricos del mundo y por otro se tiene una pobreza extrema en zonas marginadas. La pobreza ha sido uno de los principales rasgos de la economía mexicana.

El objetivo de las microfinanzas es proporcionar servicios financieros a una población de bajos recursos, están diseñadas para ayudar en la educación financiera del ahorro y microcréditos de éstas personas. A las instituciones que prestan estos servicios se les conocen como microfinancieras, algunas veces son comerciales y existen otras sin fines de lucro. La base de estas instituciones es su creencia de que la mayoría de las personas pobres no logran un cambio económico en sus vidas y continúan siendo pobres, esto ocasionado por la falta de acceso a los servicios financieros ya que en instituciones bancarias les es difícil conseguir el otorgamiento de un crédito. A pesar de que los microcréditos que ofrecen las microfinancieras son cantidades pequeñas de dinero, estos pueden tener un gran impacto las personas que lo solicitan, podrían permitir a una persona desempleada tener la inversión necesaria para comenzar un negocio pequeño o mejorar un negocio ya establecido para aumentar las ganancias y así poder destinar estos ingresos para la

alimentación de una familia o para mejorar la salud y educación de sus hijos.

Las instituciones dedicadas a las microfinanzas tuvieron origen en la India en el año de 1973 con la creación del Greameen Bank por parte de un economista llamado Muhamad Yunus. En Latinoamérica llegaron inicialmente en Bolivia y Perú.

Después de la crisis económica en México, gran parte de la población quedó desprotegida financieramente, provocando que Organismos No Gubernamentales se enfocaran en actividades financieras dirigidas a este sector mediante préstamos, para principios de los 90's éstas organizaciones se formalizaron como instituciones financieras.

Con los años, las microfinancieras han evolucionado sus funciones y estructuras, bajo el esquema legal se consolidaron y son conocidas como SOFOM (Sociedad Financiera de Objeto Múltiple) o SOFIPO (Sociedades Financieras Populares), operan mediante la autorización de la CNBV. Estas instituciones siguen desarrollándose y creciendo debido a que existe una fuerte tendencia de mercado dirigido al sector de bajos recursos donde inclusive la banca privada tradicional busca crear sus propias empresas para ingresar a este segmento convirtiéndose en competencia en el mundo de las microfinanzas.

En la actualidad se calculan más de tres mil sucursales de microfinancieras en el país, siendo Jalisco, Guanajuato, Ciudad de México, Michoacán, Querétaro y Veracruz los principales estados donde se encuentran la mayoría de estas instituciones.

Sucursales en operación 2019			
Estado	Sucursales SOFOM	Sucursales SOFIPO	Total Sucursales
Aguascalientes	33	9	42
Baja California	1	21	22
Baja California Sur	1	12	13
Campeche	9	19	28
Coahuila	28	16	44
Colima	54	7	61
Chiapas	25	121	146
Chihuahua	26	34	60
Ciudad de México	10	78	88
Durango	43	10	53
Guanajuato	312	96	408
Gerrero	40	27	67
Hidalgo	37	37	74
Jalisco	435	50	485
México	50	173	223
Michoacán	166	31	197
Morelos	39	30	69
Nayarit	77	7	84
Nuevo León	114	38	152
Oaxaca	108	17	125
Puebla	68	66	134
Querétaro	105	58	163
Quintana Roo	18	28	46
San Luis Potosí	72	30	102
Sinaloa	22	13	35
Sonora	2	33	35
Tabasco	7	27	34
Tamaulipas	26	25	51
Tlaxcala	5	15	20
Veracruz	88	75	163
Yucatán	93	28	121
Zacatecas	37	5	42
Nacional	2,151	1,236	3,387

Elaboración propia con datos de CNBV

Tabla 01.

Instituciones en operación al cierre de diciembre de 2019	
Sector	Instituciones en operación
SOFOM	156
SOFIPO	39

Elaboración propia con datos de CNBV
Tabla 02.

Se dividen en dos segmentos de acuerdo a la metodología con la que trabajan las instituciones para el otorgamiento de crédito.

- **Créditos individuales:**

Están dirigidos a un mercado formal, las exigencias son similares a las de banca tradicional pero son más flexibles. Las transacciones se realizan en las sucursales de manera personal. Pueden contar con áreas como finanzas, crédito, auditoría y mercadotecnia.

- **Créditos solidarios:**

Principalmente se enfocan en ciertos segmentos de población como mujeres, comerciantes informales o microempresarios. Se caracteriza por el otorgamiento de créditos un grupo de personas que se reúnen para solicitar el crédito y apoyarse para pagar.

Estas instituciones generan un alto impacto social debido a que están dirigidas a la base de la pirámide poblacional en la cual se encuentra la mayor parte de la población del país, que a su vez no es aceptada por la banca tradicional. Un ejemplo del impacto que tienen es Bansefi (Banca del Ahorro Nacional y Servicios Financieros), generando acuerdos globales para adquirir recursos, como software a menor costo, o para brindar capacitaciones, asesorías, asistencia técnica, entre otros servicios. Además estas entidades ayudan a impulsar la agricultura, estimula el autoempleo, igualdad de género, acceso a los beneficios de un seguro de vida a bajo costo.

Muchas personas temen solicitar un crédito pero resulta ser una opción muy viable siempre y cuando se tomen en cuenta las condiciones financieras y el alcance de pago, estos tipos de préstamos suelen ocurrir cuando no es posible solventar los gastos que se tienen provocado por la relación negativa de ingresos entre egresos, y están dirigidos principalmente a personas o microempresas que desean emprender un negocio que le ayude a obtener un ingreso el cual debe ser suficiente para pagar la deuda adquirida.

Siempre es conveniente revisar e informarse de todas las condiciones y características que implican solicitar un préstamo, para así poder tomar una decisión racional e informada. Se pueden solicitar créditos en varias instituciones ya sea en banca tradicional o instituciones de microfinanzas, las condiciones y requisitos dependerán de la institución que se escoja. Por ejemplo, existen microfinancieras que realizan todo el proceso online,

las cuales otorgan préstamos desde mil pesos hasta un máximo de cuarenta mil pesos, los plazos para realizar los pagos y las tasas de interés son establecidos por cada entidad con un máximo de diez años dependiendo directamente del monto solicitado.

II.4 Requisitos para solicitar un microcrédito

Los requisitos dependen directamente de la entidad financiera que se escoja, así como a la cantidad de dinero que se solicite. Sin embargo, la mayoría de las entidades coinciden en los siguientes puntos:

- Contar con Registro Federal de Causantes (RFC) de la persona (o personas en el caso del microcrédito grupal) o microempresa que hace la solicitud.
- Ser mayor de 21 años.
- Documento de identificación del o los solicitantes. En el caso de ser una empresa se entrega el documento de identificación del representante legal de ésta.
- Un comprobante del domicilio de la empresa, el o los solicitantes, que tenga una antigüedad de tres años como máximo.
- Las últimas dos declaraciones de impuestos de quién o quiénes hacen la solicitud, así como la empresa, en caso de que aplique.
- Los documentos que certifican la garantía prendaria. En este caso puede ser el título de propiedad de un automóvil, un departamento,

una casa u otro inmueble. Cuando no se tiene esta garantía se consigna un aval¹.

- Por lo general las entidades bancarias piden los estados de cuenta y otros datos financieros del o los solicitantes, para asegurar que éstos tienen capacidad de pago y verificar su posición crediticia.

Para los microcréditos, las tasa de interés son muy altas, esto se debe al costo elevado del proceso previo al otorgamiento del crédito. Estas tasas pueden llegar a superar el 100% anual, dejando un margen de ganancia alto para la entidad, sin embargo, esto no siempre resulta ser así cuando los clientes caen en incumplimiento de pago.

¹ *Aval - es un activo que una persona u organización presenta como fianza o promesa de pago de la deuda de otra persona en caso de incumplimiento.*

CAPÍTULO III.

METODOLOGIAS PARA CALCULAR CREDIT SCORING

Los métodos que se describirán a continuación fueron creados con base en técnicas estadísticas. Inicialmente se utilizaba el análisis del discriminante, posteriormente, estos modelos fueron enriquecidos a través de técnicas matemáticas, econométricas y de inteligencia artificial. Durante la construcción de un modelo de Credit Scoring es necesario información interna y externa del cliente, estos datos se deben encontrar en la solicitud del crédito.

Los modelos de Credit Scoring se pueden definir como algoritmos avanzados que de forma automática evalúan el probable riesgo de crédito del solicitante de financiamiento, ya sea un cliente de la entidad con historial en ella o uno nuevo. Estos modelos tienen una dimensión individual debido a que se enfocan en el riesgo de incumplimiento del individuo o empresa, independiente de lo que pueda ocurrir con el resto de la cartera de crédito.

El objetivo de los modelos de Credit Scoring es contar con una herramienta que sea capaz de ayudar eficientemente en la toma de

decisiones en la concesión de crédito a personas físicas o morales, así como para la construcción de scordcards. Se asigna una medida que permita comparar y ordenar a los solicitantes en función de su riesgo, y clasificarlos en buenos o malos de acuerdo al análisis realizado, permitiendo así, obtener alguna estimación de la probabilidad de incumplimiento del deudor asociada a su score.

La teoría detrás de cualquier modelo de Credit Scoring se resume en lo siguiente:

- Se considera un problema dinámico.
- La información inicial que se utiliza es el comportamiento crediticio de otros individuos, con características similares, a los que se les ha concedido un préstamo en el pasado.
- Consiste en estimar el valor presente descontado de los posibles beneficios o pérdidas derivadas de la concesión de un crédito, en todos los períodos del horizonte temporal considerado.
- Se concede el crédito solicitado, si la esperanza de este valor presente es positiva, de lo contrario no se concede el crédito. La cantidad máxima de crédito a conceder puede determinarse igualando dicho valor presente a cero.

El objetivo al aplicar este modelo es asignar puntajes a cada perfil de cliente de la institución financiera, para determinar, dadas las características cuantitativas y cualitativas, si el cliente tendrá una alta o baja probabilidad de caer en incumplimiento de pago, y de esta manera agilizar el trámite de un crédito y hacerlo menos subjetivo. Este tipo de procedimientos se denomina minería de datos, pues permite extraer información relevante y encontrar patrones de comportamiento de los datos; para posteriormente utilizar esta información relevante, en un análisis descriptivo y predictivo de los datos.

Se conoce como un buen modelo de scoring aquel que identifica como malos a aquellos que incumplieron con su obligación, e identifica como buenos a los clientes que no presentaron incumplimiento alguno en el pago del crédito. Se debe tomar en cuenta que el modelo puede llegar a tener incidencias y calificar como buenos a clientes que han entrado en default. Por ello es importante realizar análisis adecuados que permitan minimizar dichos errores.

Por otra parte, este modelo nos ayuda a ahorrar tiempo requerido para tomar una decisión sobre la admisión o no de un préstamo. El tiempo ahorrado dependerá de cuánto confíe la institución financiera en el modelo, y qué grado de atención quiera emplear en aquellos clientes que se acercan a un comportamiento de probable incumplimiento. Es

importante aclarar que no se establecen criterios de aceptación en base a características personales de índole racial, familiar, de género o cualquier aspecto con una sensibilidad especial.

Además, se debe considerar que la muestra no debe contener únicamente préstamos previamente concedidos, ya que de esta manera se puede provocar que el modelo este sesgado hacia una población en particular. Tomando en cuenta lo anterior, para obtener una mejor estimación se deben incluir los préstamos que fueron denegados. Igualmente, el modelo puede presentar inconsistencias si la población con la que fue calibrada no es suficientemente diversa, esto provoca que el poder predictivo del modelo sea muy pobre para determinadas sub-poblaciones y no se puede determinar adecuadamente la probabilidad de caer en default de un cliente cuyas características difieran de la muestra. No todos los factores de la muestra pueden ser incluidos en un modelo de scoring ya que no todos resultan significativos para discriminar a los clientes.

En instituciones dedicadas a las microfinanzas podemos encontrar el riesgo de crédito manifestado de la misma manera que en la banca tradicional. Sin embargo, la documentación para este tipo de instituciones es muy escasa lo que provoca dificultades a la hora de construir los modelos debido a que se requiere un sistema de gestión que minimice los costos que implican estos procedimientos. Los modelos deben ser

adecuados a las necesidades de análisis de cada entidad financiera, siempre se busca el modelo más eficiente, pero esto es subjetivo ya que la eficiencia de los modelos depende de la calidad de los datos y las variables elegidas para la construcción.

III.1 Técnicas para evaluar el riesgo crediticio

Una entidad financiera tiene como objetivo principal maximizar los beneficios que se pueden obtener derivados de la intermediación crediticia. Cada entidad debe tomar decisiones críticas acerca de conceder o no cada uno de los créditos individuales basándose en el rendimiento esperado. Mismo que se determina como:

$$(1 - P) r + P (-w) \geq 0$$

Donde:

P: probabilidad de que el individuo no cumpla con su obligación o presente algún tipo de incumplimiento.

r: rendimiento o tipo de interés en el mercado.

w: fracción del importe que se pierde por incumplimiento.

Los modelos de Credit Scoring tienen como finalidad mejorar o actualizar el criterio de selección, los modelos más utilizados son los siguientes:

- Regresión lineal
- Análisis de discriminante
- Regresión logística
- Modelos Logit
- Modelos Probit

Los procedimientos de regresión lineal y análisis de discriminante tienen como objetivo clasificar a los individuos en dos grupos, aquellos de los que se espera que cumplan con el pago del crédito (buenos) y aquellos de los que no (malos). Sin embargo, los modelos de regresión lineal han caído en desuso por sus desventajas técnicas y baja confiabilidad. Los modelos Probit, Logit y la regresión logística son superiores al análisis discriminante ya que proveen para cada deudor una probabilidad de incumplimiento, en tanto que éste sólo clasifica a los deudores en grupos de riesgo.

En los modelos de regresión logística se realiza una clasificación de los clientes entre buenos y malos, esta clasificación se efectúa analizando el comportamiento un conjunto de variables independientes. Resulta tener un gran beneficio ya que para la implementación de este modelo no es necesario plantear una hipótesis inicial, como suponer la normalidad de la distribución de las variables, de esta manera se mejora el tratamiento de las variables cualitativas o categóricas. Adicional, el modelo te permite

medir la probabilidad de incumplimiento al mantener la variable dentro de un rango de variación entre cero y uno.

La diferencia entre los modelos Probit y Logit principalmente resulta ser la función de distribución que se utiliza para mapear el resultado de la regresión lineal al intervalo $[0,1]$. Donde el modelo Probit utiliza una distribución normal y el modelo Logit una logística.

A estos modelos se les puede implementar distintas técnicas para mejorar la valuación del riesgo de crédito, pero estos son complejos de aplicar, algunas de estas técnicas son:

- Métodos no paramétricos de suavizado
- Métodos de programación matemática
- Modelos basados en cadenas de Markov
- Algoritmos genéticos
- Redes neuronales

Los modelos basados en cadenas de Markov, permiten diseñar scorecards mejor adaptadas a las necesidades de la entidad crediticia y trabajar con una gran cantidad de variables, se basan en optimizar un

criterio objetivo como por ejemplo el porcentaje de solicitantes bien clasificados.

Las redes neuronales y los algoritmos genéticos, son poco intuitivos y de difícil implementación, se utiliza cuando se desconoce la probable forma de relación de los datos y no se comportan de manera lineal.

En general, para todos los modelos el uso de la regresión lineal es la base de análisis, debido a su eficiencia para la discriminación de información. Cada modelo se diferencia por sus propios supuestos de la función de distribución de la información.

- **Modelo de Regresión lineal**

Es un modelo que permite establecer la relación entre dos variables, x variable dependiente y la variable independiente y , cuando se conoce un conjunto de datos x que indican el comportamiento de y se busca establecer una relación entre las variables de la forma $f(x)$ y suponemos linealidad en la función. Por tanto, tenemos un modelo donde la relación entre los datos se comporta de manera:

$$\hat{y} = b_0 + b_1x + e$$

buscamos:

$$E(Y_i | X_i) = P_i = \beta^T X_i$$

La ventaja de utilizar este modelo es que se puede estimar e interpretar de forma simple, pero la probabilidad no es lo más preciso ya que la varianza condicional varía según las observaciones, por lo que las estimaciones de β no son eficientes.

- **Análisis discriminante**

Es una técnica estadística multivariante que permite estudiar la relación de comportamiento de múltiples variables independientes y poder clasificarlas en grupos frente a una variable dependiente, separando a los buenos o malos clientes para los modelos de Credit Scoring. Este método tiene el objetivo de obtener las mejores combinaciones lineales de variables independientes maximizando la diferencia entre los grupos.

Para realizar este análisis, se suele dividir en dos la muestra a estudiar, denominadas como muestra de entrenamiento la cual es utilizada para obtener las funciones discriminantes, y muestra de validación donde se determina la capacidad predictiva de modelo implementado. Se busca una serie de modelos tales que:

$$X = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_i x_i$$

Donde

λ_i Son los parámetros de la relación de combinaciones de variables.

X_i Es la cantidad de n modelos posibles, de acuerdo a las combinaciones entre las variables.

Una vez encontrados los parámetros se toma la desviación estándar quedando:

$$D = \lambda_1 d_1 + \lambda_2 d_2 + \dots + \lambda_i d_i$$

Luego, se busca la función de varianzas y covarianzas que determine el modelo con menos valores residuales entre ellos.

$$S = \sum_{p=1}^n \sum_{q=1}^n \lambda_p \lambda_q S_{pq}$$

Existe una variante de este análisis conocida como Análisis Discriminante Lineal o Modelo Z - Score, enfocado a la probabilidad de quiebra de una entidad, también resultó útil para la estimación de impagos en entidades financieras:

$$Z = \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$$

Donde

Z es el índice general.

B_k son los parámetros de cambio que afectan la probabilidad de impago.

X_{ik} son las variables cualitativas o cuantitativas.

- **Modelo Probit**

Los modelos Probit se clasifican como modelos de elección discreta binaria. El modelo postula como variable a observar si se otorga o no el crédito a la persona. Con esta información subsecuente el modelo reproduce una variable latente, que define la probabilidad de que la persona caiga en incumplimiento.

Se observa el siguiente modelo:

$$P(y=1 | x) = G(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) = G(\beta_0 + \beta_x)$$

Donde:

G: es una función que adopta valores entre cero y uno para todos los números reales.

En el modelo Probit, G representa la función de distribución acumulativa normal. La estimación de los parámetros se hace a través del método de máxima verosimilitud ya que es un modelo de variable dependiente limitada. Este método induce a elegir como estimados los valores de los parámetros que maximicen el logaritmo de la función de verosimilitud

El logaritmo de la función de verosimilitud para una muestra de tamaño n se define entonces como:

$$\ln L = \sum_{i=1}^n \ln f_i(\beta)$$

- **Modelo Tobit**

Se considera a este modelo como una extensión del modelo Probit, este toma tanto la información de la probabilidad del punto de interés, como los valores de la variable latente. El uso de este modelo es menos comercial debido a su complejidad.

En casos de modelos de regresión lineal siempre se conocen todos los valores de las variables a utilizar, sin embargo, se puede presentar el caso donde la muestra se encuentre truncada o censurada. El modelo Tobit proporciona un mejor ajuste de la realidad ya que considera estos casos, condicionando la variabilidad de los errores.

El modelo depende totalmente de los supuestos de normalidad y homocedasticidad, es decir, que la varianza de los errores es constante, del término de error. La estimación e inferencia es inválida si alguno de los dos supuestos es incorrecto

$$y_i^* = \alpha + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

Donde

y_i^ es la variable inobservable, cuyos resultados pueden ser:*

$$y_i = y_i^*, \text{ si } y_i > 0$$

$$y_i = 0, \text{ si } y_i \leq 0$$

En cualquier caso, se debe obtener la función que indica el resultado óptimo

$$\ln L = \sum_{y_i > 0} [-\ln(\sigma) + \ln(\varphi) (y_i - \alpha + \beta_2 x_{i2} + \dots + \beta_k x_{ik})] + \sum_{y_i > 0} \ln [1 - \varphi (\frac{(\alpha + \beta_2 x_{i2}) + \dots + \beta_k x_{ik}}{\sigma})]$$

Donde

y_i es la respuesta bivariada de riesgo de impago.

α , β_k son parámetros de cambio que afectan la probabilidad de impago.

x_{ik} son variables cualitativas y cuantitativas.

- **Redes Neuronales**

Es una metodología no paramétrica enfocada para modelos con relaciones no lineales, utiliza un sistema artificial que trata de imitar al sistema nervioso humano. Permite identificar comportamientos de varias variables al mismo tiempo para realizar predicciones eficientes. Se conforma por una serie de procesadores conectados entre sí, conocidos como nodos, los cuales tienen la finalidad de responder a cada señal de entrada y dar una salida, estudian el comportamiento de los datos a modo de entrenamiento examinando registros individuales y una vez entrenada la red es capaz de definir el posible resultado de nuevos datos.

Los nodos toman el papel de una neurona artificial, estas son modelos computacionales que surgen del comportamiento de una neurona real. Sin embargo, el proceso que realiza una neurona es muy complejo para modelarlo, estas reciben una señal a través de la sinapsis¹ lo cual se considera la entrada en un modelo artificial. Cuando la señal recibida es lo suficientemente fuerte, la neurona se activa y emite una señal a través del axón², que a su vez podría ser enviada a otra sinapsis y activar más neuronas. Matemáticamente, se debe calcular una función para cada entrada que determine la activación de la neurona y posteriormente calcular otra función para dar la salida.

Dentro del Credit Scoring esta técnica resulta ser complicada ya que requiere conocimientos especializados, Bajo este tema se considera nodo de entrada a las variables de la operación de crédito y a la probabilidad de incumplimiento se le denomina nodo de salida.

- **Árboles de Decisión**

Es un método gráfico de representar los posibles eventos que pueden ocurrir a partir de una decisión asumida y permite tomar una mejor decisión basada en la probabilidad de los eventos, los árboles basados en funciones bayesianas pueden ser utilizados como modelos de discriminación en el Credit Score. Este método tiene de beneficio que no asume ninguna distribución ni tiene supuestos estadísticos. Se hace uso de la regresión lineal para su ejecución, se conforma de nodos de decisión y nodos de probabilidades los cuales señalan los eventos sujetos a riesgo asociados a las diferentes variables, además estos nodos se unen por ramas, representando las posibles alternativas y el riesgo que implica cada una.

¹ *Sinapsis – Zona de transmisión de impulsos nerviosos eléctricos entre dos neuronas.*

² *Axón – Canal de comunicación entre la parte central de la neurona con otra parte del sistema nervioso.*

Presenta de manera visual las relaciones entre las variables, los grupos de la variable respuesta y el riesgo; por ello, es muy usado en el Credit Scoring.

$$(\tau) = \sum (Y_i - \bar{Y}(\tau))^2$$

$$Y_i = F(x_1, x_2, \dots, x_p) = \alpha + \sum_{j=1}^p \beta_k x_j = \alpha + \sum_{j=1}^p \beta_k g_j(x) = f(x)$$

Donde

(τ) es la probabilidad de éxito del nodo τ .

Y_i es la probabilidad de riesgo de impago.

α, β_k son parámetros de cambio que afectan la probabilidad de impago.

x_j variables cualitativas.

- **Modelo Logit**

Este modelo permite identificar los factores de riesgo que determinan las probabilidades de ocurrencia de un evento, así como la correlación que éstos tienen. Este tipo de modelo arroja como resultado un índice que facilita realizar ordenaciones o clasificaciones en las que se le asocia a cada elemento una calificación. Existen diversos criterios para llevar a cabo la asociación índice - calificación, el caso más simple es cuando solo

se presenta una variable explicativa y se trata de encontrar la relación que existe entre ésta y la endógena, el cual es conocido como método de regresión lineal donde se asume que la relación entre las variables tiene un comportamiento lineal, sin embargo esto en la mayoría de los casos no se cumple. El proceso aplicado para el modelo Logit es similar a la regresión lineal a diferencia que utiliza como función de estimación la función logística en vez de la lineal.

Para este modelo tenemos las siguientes expresiones:

$$Z_t = \alpha + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

$$Y_i = \frac{1}{1 + \lambda^{(-\alpha - \beta_k - X_{ik})}} + \varepsilon_i = \frac{\lambda^{(-\alpha - \beta_k - X_{ik})}}{1 + \lambda^{(-\alpha - \beta_k - X_{ik})}} + \varepsilon_i$$

Donde

Z_t nivel de discriminación.

Y_i es una variable dicotómica, que puede tomar solamente dos valores (1 y 0).

α, β son parámetros de cambio que afectan la probabilidad de impago.

X_{ik} son variables cualitativas o cuantitativas.

Existen distintos tipos de modelos Logit en función de las características que definen a la variable endógena, ésta permite medir el número de

grupos que puedan existir en el análisis, los modelos Logit se pueden clasificar así:

- Logit dicotómico: Se utiliza cuando el número de alternativas es igual a dos y son excluyentes entre sí.
- Logit de respuesta múltiple: Se utiliza cuando las alternativas a modelar es superior a dos.
- Logit con datos no ordenados: Se utiliza cuando la variable endógena no indican ningún orden.
- Logit multinomial: Se utiliza cuando los indicadores del modelo hacen referencia a las observaciones muestrales, por lo que varían entre observaciones pero no entre alternativas.
- Logit condicional: Se utiliza cuando los indicadores del modelo hacen referencia a las alternativas, por lo que sus valores varían entre alternativas pudiendo hacerlo o no entre observaciones.
- Logit con datos ordenados: Se utiliza cuando la variable endógena establece un orden entre sus características.

Para determinar cuáles de las variables resultan significativas para el modelo se debe calcular la estadística de Wald. Se puede presentar que la variable sea categórica o no categórica, para cada uno de los cuales:

- Si la variable no es categórica, se distribuye chi cuadrado con un grado de libertad.
- Si la variable es categórica, se distribuye chi cuadrado, con n grados de libertad igual al número de parámetros estimados.

El estadístico de Wald tiene una gran relevancia porque a través de él se puede determinar la significancia del parámetro en el modelo y la relación que tiene.

Se consideran dos hipótesis para este modelo: nula (H_0) y alternativa (H_1).

- $H_0: \beta_i = 0$ donde la variable X_i no es importante para establecer variaciones en el modelo.
- $H_1: \beta_i \neq 0$ donde la variable X_i si es importante para establecer variaciones en el modelo.

Una vez efectuado el contraste, sí se llega a rechazar la hipótesis nula, se observa a qué nivel de confianza se efectúa el rechazo, y dependiendo del

valor se le da el peso a la variable en el modelo, según el siguiente criterio:

- 30% - 95% Poco significativa.
- 97.5% Significante.
- 99% Muy significativa.
- 99.9% Altamente significativa.

III.2 Base de datos

La elaboración de una base de datos tiene el objetivo de obtener información fiable y detallada sobre el comportamiento de clientes a los que se les ha otorgado un crédito. Es la fuente de información para la construcción del modelo. La base debe ser funcional para cualquier software estadístico, para esto los datos necesitan estar en forma de matriz dentro de un archivo plano o de texto, donde cada columna corresponde a una variable.

Al crear la base de datos se debe considerar el tamaño de muestra disponible, el cambio histórico en la situación del cliente, la calidad de los datos disponibles, las fuentes externas o internas y las variables a utilizar.

La mayoría de las bases de datos presentan inconsistencias en sus datos, por lo cual es importante realizar una limpieza de la base antes de cualquier análisis, ya sea que existan datos faltantes o los valores no tienen sentido para una característica en particular. Las situaciones más comunes son: no fueron capturados, no disponibles, no fueron llenados por los clientes o casos extremos.

Para tratar con datos faltantes se puede:

- Excluir características o registros que tienen datos faltantes significativos.
- Incluir características con una pequeña proporción de datos faltantes.
- Estimar los datos faltantes.

Los datos faltantes pueden ser parte de una tendencia, la cual puede ser valuada en el mal comportamiento predictivo y afectar los resultados del modelo. Si la característica o registros con datos faltantes son excluidos, datos importantes pueden ser perdidos. Cuando se tienen registros totalmente vacíos o solamente se llenó el campo del nombre, se elimina. Sin embargo, cuando solo alguno de los campos está vacío, mediante un análisis respecto a la combinación del resto de los campos, se puede estimar que valor le corresponde.

Para la construcción del modelo se deben escoger qué variables son significativas para el análisis requerido. El tipo de variables empleadas depende del modelo que se aplique.

En el caso de Modelos para individuos y PyMEs se utilizan como variables explicativas a un conjunto de características relativas a individuos que han recibido un crédito con anterioridad. Estas variables son de tipo socioeconómico y por lo general se clasifican en:

- **Personales.**- edad, estado civil, dependientes económicos, antigüedad en el empleo, tiempo de permanencia en el domicilio actual, nivel educativo, etc.
- **Económicas.**- tipo de vivienda, gastos mensuales promedio, ingresos mensuales promedio, tipo de ocupación, si tiene tarjeta de crédito, cuenta corriente o caja de ahorro.
- **Referentes al préstamo solicitado.**- número de consultas en los buros de crédito y cómo está calificado en ellos.

Las variables o categorías se refieren a las preguntas y las posibles respuestas se denominan atributos.

- **Variable:** Tipo de vivienda

- **Atributo:** Propia/Rentada/Familiares

Los atributos utilizados deben ser generales y mutuamente excluyentes, es decir, toda pregunta debe ser contestada y solo un atributo debe responder a la pregunta.

La cantidad excesiva de variables a utilizar en el modelo puede ser un factor negativo, ya que puede existir alta correlación entre algunas variables y esto generar que el modelo no sea del todo confiable. Para evitar variables excesivas, se puede realizar un análisis de correlación entre variables o detectar variables neutras y así poderlas descartar del modelo.

Algunas entidades financieras complementan sus bases con datos externos como en el caso del país, aunque existen varias instituciones con esta información, la mayoría utiliza los datos proporcionados por el Buró Nacional de Crédito, en el cual puedes obtener el historial crediticio de toda persona que haya adquirido un crédito.

CAPÍTULO IV.

SELECCIÓN DE VARIABLES

En este capítulo se estudiará el caso específico de los créditos otorgados a mujeres jóvenes de 20 a 35 años dentro de una microfinanciera (no se permite divulgar el nombre por temas de confidencialidad), dicha institución utiliza SAS como software para construir sus scorecards. En promedio tardan 6 meses en la construcción, con alrededor de 4 horas al día.

La base de datos que se utilizó para realizar este proyecto fue proporcionada por la empresa, la cual contiene datos reales con un total de 14,000 clientes en la cartera. Los pagos de créditos que realizan los deudores varía, la mayoría se realizan de manera semanal, aunque existen otros clientes que lo hacen de forma quincenal y mensual. De acuerdo a la base de datos, se tiene un histórico de seguimiento de créditos con 35 variables registradas de cada cliente de la entidad. Para este análisis se filtraron los datos donde la variable sexo es igual a "F" y la edad este en el rango de 20 a 35 años, quedando un total de 4,093 datos para analizar.

De acuerdo a las características observadas de los clientes mediante el modelo aplicado, se obtuvieron diversas formas para mitigar los riesgos crediticios de la microfinanciera, y así tener la capacidad de predecir anticipadamente el comportamiento de sus clientes para que la institución pueda tomar decisiones oportunas.

IV.1 Descripción de las variables

La variable dependiente será default, que es el número de días en el cual la institución financiera considerará que se entra en incumplimiento de las obligaciones. Ésta es utilizada en las instituciones financieras para evaluar la salud de la cartera y mediante este análisis administrar mejor el riesgo de probabilidad de impago. Se define como un default bueno aquel que esté por debajo de este número de días y uno malo aquel que se encuentre por encima. El nivel de default se determina de acuerdo con las políticas de cada entidad.

Las variables independientes explican y determinan la variable dependiente del modelo, de las treinta y cinco variables con las que se cuenta y los clientes previamente clasificados en buenos y malos, una vez aplicados los filtros mencionados anteriormente, se tomaron el resto de

las variables que guardan una mejor relación con la variable dependiente, estas se definieron en categorías o rangos, cada uno debe contener al menos el 5% de la población y el número de clientes buenos o malos dentro de cada rango debe ser mayor a cero.

Para evaluar las variables de la base, se calcula el WoE (Weight of Evidence), con ayuda de este valor se determina el poder predictivo de una variable independiente en relación a una variable dependiente, es decir, nos indica que tan bien se separan los buenos de los malos en los rangos creados.

$$\text{WoE} = \ln \left(\frac{\text{Buenos}}{\text{Malos}} \right) \quad (100)$$

Si al realizar los cálculos, se obtienen valores negativos del WoE quiere decir que se tienen proporciones altas de malos sobre los buenos en la base de datos por lo tanto los solicitantes en ese grupo presentan un mayor riesgo de crédito.

A continuación se muestra un ejemplo de los resultados al aplicar la fórmula mencionada anteriormente en los respectivos rangos de cada variable.

Cuota / Utilidad al periodo					
Rangos	Malos	Buenos	% Malos	% Buenos	WoE
≤15%	347	649	28.49%	31.57%	0.1026
15% - 30%	91	284	7.47%	13.81%	0.6146
30% - 45%	193	367	15.85%	17.85%	0.1191
45% - 60%	185	307	15.19%	14.93%	-0.0171
60% - 75%	130	246	10.67%	11.96%	0.1142
> 75%	272	203	22.33%	9.87%	-0.8161
	1218	2056	100.00%	100.00%	

Vivienda					
Rangos	Malos	Buenos	% Malos	% Buenos	WoE
DUEÑO/PROPIA	689	1193	56.57%	58.03%	0.03
RENTADA/ES INQUILINO	102	158	8.37%	7.68%	-0.09
VIVE CON PARIENTES/AMIGOS/OTROS	427	705	35.06%	34.29%	-0.02
	1218	2056	100.00%	100.00%	

Estado Civil					
Rangos	Malos	Buenos	% Malos	% Buenos	WoE
CASADO(A)	764	1273	62.73%	61.92%	-0.01
SOLTERO(A)	428	717	35.14%	34.87%	-0.01
VIUDO / DIVORCIADO	26	66	2.13%	3.21%	0.41
	1218	2056	100.00%	100.00%	

Resultados WOE por variable.

Tabla 03.

Posteriormente se debe calcular el Information Value (IV) la cual nos describe que tan significativa es la variable para la separación de buenos y malos.

$$IV = \sum_{i=1}^n (\text{Buenos} - \text{Malos}) \left(\frac{\text{WoE}}{100} \right)$$

Seguido de realizar estos pasos, los resultados mostraron sólo nueve variables significativas para aplicar el modelo. Para las variables que obtuvieron un IV similar pero menor a 0.02, estas variables se consideran no predictivas y se deben excluir del modelo, sin embargo, si a la variable le falta poco para ser predictiva se puede incluir y usarla en el modelo a desarrollar ya que se considera una variable importante para la toma de decisión de otorgamiento de un crédito. Las variables elegidas son:

- **Cuota / Utilidad al periodo:** Relación del pago que se realiza cada periodo entre la utilidad equivalente al mismo periodo.
- **Vivienda:** Tipo de vivienda en la que vive el prestatario.
- **Estado Civil:** Situación civil en la que se encuentra actualmente el cliente.
- **Tipo de Trabajo:** Régimen de empleo del prestatario.
- **Frecuencia:** Este atributo nos indica la modalidad con la que el cliente paga.
 - o Semanal
 - o Quincenal
 - o Mensual
 - o Otro
-

- **Pago total a realizar:** Atributo que indica el total de lo que el cliente debe a la financiera en ese momento.
- **Préstamo / Utilidad:** Relación de la cantidad prestada al cliente entre su utilidad mensual.
- **Préstamo solicitado:** Indica cuanto se le prestó al cliente.
- **Deuda:** Atributo que indica el total de lo que me va a pagar el cliente en el momento 0.

Por tanto, las variables anteriormente descritas fueron seleccionadas bajo el criterio establecido del IV y utilizadas para determinar, dadas las características cuantitativas y cualitativas, si el cliente tiene una alta o baja probabilidad de incumplimiento.

Information Value	Poder Predictivo
< 0.02	No es predictivo
0.02 - 0.1	Débil
0.1 - 0.3	Medio
0.3 - 0.5	Fuerte
> 0.5	Sospechoso

Reglas del criterio IV.

Tabla 04.

CAPÍTULO V.

APLICACIÓN DE MODELO A CASO REAL

Como se mencionó anteriormente, existen diversos modelos propuestos para evaluar el riesgo de crédito, entre ellos destacan los modelos de Credit Scoring los cuales proponen automatizar el proceso de gestión de créditos en cuanto a conceder o no un crédito sujeto a un conjunto de variables relacionadas al solicitante de éste. El éxito de este modelo se basa en la calidad del algoritmo utilizado y de la existencia de un sistema eficiente de análisis de datos. El objetivo de aplicar un modelo de scoring consiste en predecir el riesgo de impago, es decir, estimar la probabilidad de insolvencia para una cartera minorista de créditos, de esta manera se busca conocer las medidas o modificaciones que se deben implementar a los modelos conocidos para reducir los niveles de riesgo de crédito.

En este proyecto se utilizó el modelo Logit para calcular la probabilidad de que una persona sea un cliente bueno o malo para decidir si se le otorga el crédito solicitado, es decir, nos interesa saber la probabilidad de que caiga en incumplimiento de pago. De acuerdo al comportamiento de las variables independientes de cada persona, se realiza la clasificación de esta. Se trata de un modelo paramétrico similar al análisis

discriminante, pero la ventaja que los diferencia es que no necesita plantear hipótesis iniciales como en la distribución de las variables, lo cual permite mejorar el tratamiento de las variables categóricas y cualitativas ya que usualmente los problemas a los que se enfrentan las entidades financieras presentan alguna variable cualitativa que no permite el correcto funcionamiento de los modelos que se basan en hipótesis de normalidad.

Para llevar a cabo la ejecución del modelo de Credit Scoring, primero se construyó una base de datos con la información necesaria que utiliza la microfinanciera para decidir a quién le otorgará crédito. Una vez que se obtiene la base de datos se debe llevar a cabo una limpieza de ésta, ya que puede contener errores de captura o elementos no válidos para las variables; como se menciona en capítulos anteriores, estos deben ser corregidos o eliminados. Para este proyecto la limpieza de la base fue realizada en Excel, analizando cada variable de acuerdo a su tipo e importancia.

Generalmente se considera que un cliente tiene un estatus “malo” cuando se registran cuatro o más pagos vencidos, cuando un cliente cae en incumplimiento y comienza a incrementar su deuda, lo más probable es que continúe en ese estado. Se necesita clasificar a los clientes en buenos y malos para llevar a cabo los siguientes pasos del modelo.

Después de realizar la clasificación de buenos y malos, se eligen las variables categóricas para calcular su WoE y con estos nuevos valores poder realizar la regresión logística más adelante. Para realizar el cálculo de este indicador se lleva a cabo el siguiente procedimiento; para cada variable se crean rangos que contengan al menos el 5% del total de los datos, se registra en una columna el número de clientes que fueron clasificados como buenos con las características de cada rango y el número de clientes que se definieron como malos; posteriormente se aplica la fórmula descrita en capítulos anteriores.

Una vez calculado el WOE, se deben seleccionar las variables que son significativas para los cálculos de la predicción, para esto se calcula Information Value (IV). Para obtener este indicador, es necesario multiplicar el valor del WOE obtenido en cada rango de las variables, por la diferencia del porcentaje de bueno menos el porcentaje de malos.

La suma de los IV's de cada una de las variables significativas es el valor que nos interesa y en el que centraremos nuestra atención ya que nos ayuda a decidir si la variable es predictiva o no para el modelo. Para esto se toma el siguiente criterio, cuando el $IV < 0.02$ se dice que la variable no es predictiva y por tanto se descarta.

En nuestro caso de estudio se hicieron algunas excepciones debido a que se consideraron variables predictivas adicionales dada su importancia para nuestros objetivos a pesar de que se obtuvo un IV ligeramente menor a 0.02.

Cuota / Utilidad al periodo						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
≤15%	347	649	28.49%	31.57%	0.1026	0.0032
15% - 30%	91	284	7.47%	13.81%	0.6146	0.0390
30% - 45%	193	367	15.85%	17.85%	0.1191	0.0024
45% - 60%	185	307	15.19%	14.93%	-0.0171	0.0000
60% - 75%	130	246	10.67%	11.96%	0.1142	0.0015
> 75%	272	203	22.33%	9.87%	-0.8161	0.1017
	1218	2056	100.00%	100.00%		0.1477

Vivienda						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
DUEÑO/PROPIA	689	1193	56.57%	58.03%	0.03	0.0004
RENTADA/ES INQUILINO	102	158	8.37%	7.68%	-0.09	0.0006
VIVE CON PARIENTES/AMIGOS/OTROS	427	705	35.06%	34.29%	-0.02	0.0002
	1218	2056	100.00%	100.00%		0.0011

Estado Civil						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
CASADO(A)	764	1273	62.73%	61.92%	-0.01	0.0001
SOLTERO(A)	428	717	35.14%	34.87%	-0.01	0.0000
VIUDO / DIVORCIADO	26	66	2.13%	3.21%	0.41	0.0044
	1218	2056	100.00%	100.00%		0.0045

Tipo de Trabajo						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
Tiempo Completo	341	545	28.00%	26.51%	-0.055	0.0008137
Autoempleado	560	913	45.98%	44.41%	-0.035	0.0005458
Ama de Casa	134	232	11.00%	11.28%	0.025	0.0000716
Otros	183	366	15.02%	17.80%	0.170	0.0047095
	1218	2056	100.00%	100.00%		0.00614

Frecuencia						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
Semanal	1037	1804	85.14%	87.74%	0.030	0.0008
Quin - Men	77	117	6.32%	5.69%	-0.105	0.0007
Otro	104	135	8.54%	6.57%	-0.263	0.0052
	1218	2056	100.00%	100.00%		0.0066

Pago total a realizar						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
0 - 2500	171	345	11.29%	13.38%	0.170	0.0036
2501 - 5500	102	277	6.73%	10.74%	0.467	0.0188
5500 - 7000	201	397	13.27%	15.40%	0.149	0.0032
7001 - 9000	225	335	14.85%	12.99%	-0.134	0.0025
9000 - 11500	257	365	16.96%	14.16%	-0.181	0.0051
11501 - 13500	180	247	11.88%	9.58%	-0.215	0.0049
>13500	82	90	5.41%	3.49%	-0.439	0.0084
	1218	2056	80.40%	79.75%		0.0241

Prestamo / Utilidad						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
0 - 0.5	226	482	18.56%	23.44%	0.2339	0.0114
0.5 - 1	380	646	31.20%	31.42%	0.0071	0.0000
1 - 1.5	275	495	22.58%	24.08%	0.0642	0.0010
1.5 - 2	172	230	14.12%	11.19%	-0.2330	0.0068
> 2	165	203	13.55%	9.87%	-0.3163	0.0116
	1218	2056	100.00%	100.00%		0.0309

Prestamo solicitado						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
0 - 2000	212	490	13.99%	19.01%	0.3062	0.0154
2001 - 4000	364	702	24.03%	27.23%	0.1252	0.0040
4001 - 6000	404	564	26.67%	21.88%	-0.1980	0.0095
> 6000	238	300	15.71%	11.64%	-0.3001	0.0122
	1218	2056	80.40%	79.75%		0.0122

Deuda						
Rangos	Malos	Buenos	% Malos	% Buenos	WoE	IV
1000 - 4000	98	370	6.47%	14.35%	0.7969	0.0628
4001 - 7000	283	529	18.68%	20.52%	0.0939	0.0017
7001 - 10000	435	638	28.71%	24.75%	-0.1486	0.0059
> 10000	402	519	26.53%	20.13%	-0.2761	0.0177
	1218	2056	80.40%	79.75%		0.0177

Resultados IV's por variable.

Tabla 05.

Una vez realizados los pasos anteriores se genera una base únicamente con las variables predictivas y con la variable dependiente, donde a la variable dependiente se le asigna un valor igual a 1 si el cliente se clasifica como bueno, o de lo contrario tiene valor igual a 0 si el cliente se clasifica como malo. En el caso de las variables explicativas su valor será igual a sus WOE's de las características no excluidas.

Cuo_Util	Edo_Civil	Tipo_Trab	Frecuencia	Vivienda	Pago_total	Prest / Util	Prest_sol	Deuda	Bueno/Malo
0.11424491	0.408006026	-0.034753081	0.030122314	0.025432973	-0.1807772	-0.23296735	-0.1979591	-0.27615	0
0.10255576	-0.01298837	-0.034753081	-0.262668299	-0.08592996	0.467446227	0.233856936	0.30622065	0.796937	1
0.11911948	-0.01298837	-0.054648861	0.030122314	0.025432973	-0.13356834	0.007076073	-0.1979591	-0.14861	1
-0.01706026	-0.00759953	-0.034753081	0.030122314	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	1
0.10255576	-0.01298837	-0.054648861	0.030122314	-0.08592996	0.149032907	0.007076073	0.12518107	0.093943	1
0.11911948	-0.01298837	-0.034753081	-0.262668299	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	0
0.10255576	-0.01298837	-0.034753081	0.030122314	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	1
0.10255576	-0.00759953	-0.054648861	0.030122314	-0.02213839	-0.1807772	0.064234487	-0.1979591	-0.14861	1
-0.81614827	-0.01298837	-0.034753081	-0.262668299	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	0
0.10255576	-0.01298837	-0.034753081	-0.262668299	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	0
0.10255576	-0.00759953	-0.034753081	0.030122314	-0.02213839	0.170282395	0.233856936	0.30622065	-0.14861	1
-0.81614827	-0.01298837	0.025345393	0.030122314	0.025432973	-0.1807772	-0.31629167	-0.1979591	-0.14861	0
0.10255576	0.408006026	-0.034753081	0.030122314	0.025432973	0.149032907	0.007076073	0.12518107	0.093943	1
0.11911948	-0.01298837	-0.054648861	0.030122314	0.025432973	0.149032907	0.233856936	0.12518107	0.093943	1
0.11424491	-0.01298837	0.169595002	-0.262668299	0.025432973	-0.1807772	-0.23296735	-0.1979591	-0.27615	1
0.11911948	-0.01298837	-0.034753081	0.030122314	-0.08592996	0.149032907	0.007076073	0.12518107	0.093943	1
-0.01706026	-0.01298837	0.169595002	0.030122314	0.025432973	-0.21516698	0.064234487	-0.3000867	-0.27615	0
0.10255576	-0.00759953	0.169595002	0.030122314	-0.02213839	-0.13356834	-0.23296735	-0.1979591	-0.14861	0
0.10255576	-0.01298837	-0.034753081	0.030122314	0.025432973	-0.21516698	0.007076073	-0.3000867	-0.27615	0
0.11424491	-0.01298837	-0.054648861	0.030122314	0.025432973	0.149032907	0.007076073	0.12518107	0.093943	1
0.11911948	-0.01298837	-0.034753081	0.030122314	0.025432973	0.170282395	0.233856936	0.30622065	0.796937	1
0.10255576	-0.01298837	-0.034753081	0.030122314	0.025432973	-0.21516698	-0.23296735	-0.3000867	-0.27615	0
0.11911948	-0.01298837	0.169595002	0.030122314	0.025432973	0.170282395	0.233856936	0.30622065	0.093943	1
-0.81614827	-0.01298837	-0.034753081	-0.262668299	0.025432973	-0.1807772	-0.23296735	-0.1979591	-0.14861	0
0.10255576	-0.01298837	0.169595002	0.030122314	-0.02213839	-0.21516698	-0.23296735	-0.3000867	-0.27615	0
0.10255576	-0.00759953	-0.034753081	0.030122314	0.025432973	0.170282395	0.233856936	0.30622065	-0.14861	0
0.11911948	-0.01298837	-0.034753081	0.030122314	0.025432973	-0.13356834	0.007076073	0.12518107	-0.14861	0
0.61456255	-0.00759953	0.169595002	-0.262668299	0.025432973	-0.43850804	0.007076073	-0.3000867	-0.27615	1
0.11911948	-0.01298837	-0.054648861	0.030122314	0.025432973	-0.13356834	0.007076073	0.12518107	-0.14861	1
-0.01706026	-0.00759953	-0.034753081	-0.262668299	-0.08592996	0.170282395	0.233856936	0.30622065	0.796937	1

Ejemplo de base para cálculo de regresión.

Tabla 06.

La base que se menciona anteriormente es necesaria su construcción adecuada para el cálculo de los coeficientes de regresión logística, ya que estas son fundamentales para realizar el modelo de predicción para los nuevos clientes, entendiéndose que el coeficiente β_i ($i = 0,1,2,\dots, 9$) representa el cambio en el Logit resultante al aumentar una unidad en

cada una de las variables. Para ello utilizamos al algoritmo de regresión logística en la aplicación RStudio. Para estimar los coeficientes β_i necesarios se utilizó la función glm del paquete 'Stats' de la librería que nos proporciona R.

El modelo Logit, el cual forma parte del grupo de los modelos lineales generalizados, es un caso particular de los modelos lineales generalizados en los que la distribución es binomial. La función glm() de R permite ajustar modelos lineales de muchos tipos. La sintaxis básica para obtener un modelo lineal con esta función es la siguiente:

*glm (dependiente~independiente1+independiente2,
family=binomial(),data=datos")*

Donde, family = binomial() especifica la función de probabilidad a utilizar.

Posteriormente, con las variables seleccionadas como predictivas, se debe calcular su Score correspondiente de cada rango, para esto se utiliza la β_i calculada anteriormente para cada una de estas variables. Los resultados obtenidos en R se muestran a continuación.

Call:

```
glm(formula = tabla$Buena.Malo ~ tabla$Cuo_Util + tabla$Vivienda +
     tabla$Edo_Civil + tabla$Tipo_Trab + tabla$Frecuencia +
     tabla$Pago.total.a.realizar +
     tabla$Prestamo...Utilidad + tabla$Prestamo.solicitado +
     tabla$Deuda.actual,
     family = binomial)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.1688	-1.2973	0.7587	0.9782	1.5949

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	0.54495	0.03774	14.438	< 2e-16	***
tabla\$Cuo_Util	1.05454	0.10136	10.404	< 2e-16	***
tabla\$Vivienda	1.60829	1.14069	1.410	0.1586	
tabla\$Edo_Civil	1.20917	0.57772	2.093	0.0363	*
tabla\$Tipo_Trab	0.94489	0.48401	1.952	0.0509	.
tabla\$Frecuencia	0.93054	0.45556	2.043	0.0411	*
tabla\$Pago.total.a.realizar	0.05109	0.31274	0.163	0.8702	
tabla\$Prestamo...Utilidad	-0.21404	0.25944	-0.825	0.4094	
tabla\$Prestamo.solicitado	0.03606	0.31507	0.114	0.9089	
tabla\$Deuda.actual	1.11485	0.18483	6.032	1.62e-09	***

La constante para la variable dependiente (Buena.Malo=Intercept) es 0.54495. Es decir, el logaritmo nos indica que la probabilidad de que $y=1$ es 0.54495. Sin embargo, interpretar un logaritmo de razón de

probabilidad no resulta lo más intuitivo. Al menos por el signo positivo sabemos que la probabilidad de ($y=1$) es mayor a la de ($y=0$), en otras palabras, hay mayor número de clientes clasificados en el modelo como buenos que clientes clasificados como malos. Para hacer más intuitiva la interpretación procedemos a convertir el coeficiente en una razón de probabilidad, esto es, exponenciar el logaritmo. En R lo hacemos con la función ‘ $\exp(0.54495)$ ’, la cual nos regresa el exponencial de un logaritmo. Entonces tenemos que el exponencial de 0.54495 es igual a 1.724522, es decir, por cada cliente clasificado malo, hay 1.724522 clientes buenos.

Estos puntajes se usan para calcular el score de un cliente, esto a su vez es utilizado para decidir si se le otorga o no el crédito solicitado. Cada uno de estos describe la beta a utilizar para el cálculo del Score:

β_i	
β_0	0.54495
β_1 (Cuo_Util)	1.05454
β_2 (Vivienda)	1.60829
β_3 (Edo_Civil)	1.20917
β_4 (Tipo_Trab)	0.94489
β_5 (Frecuencia)	0.93054
β_6 (Pago.Total.a.realizar)	0.05109
β_7 (Prestamo...Utilidad)	-0.21404
β_8 (Prestamo.solicitado)	0.03606
β_9 (Deuda)	1.11485

Resultados de β por variable.

Tabla 07.

En la práctica es común dividir la muestra en dos submuestras, la primera es conocida como muestra de entrenamiento, se considera el 80% de los datos totales y se utiliza para estimar el modelo; la segunda se conoce como muestra de validación donde se considera solo el 20% restante. La división de la muestra se debe realizar de manera aleatoria, en este caso, utilizando las herramientas de Excel. Se toma un total de 3,274 registros para la muestra de entrenamiento.

La scorecard se construye con los estimadores de los parámetros de la regresión logística. Para cada atributo, con ayuda del valor de su WoE y el coeficiente de regresión de su característica, al realizar el producto de estos valores se puede obtener el score del atributo. La score total de un solicitante sería proporcional al logaritmo de las probabilidades malas entre las buenas predichas de ese solicitante.

Los puntajes del score son resultado de una traslación y reescalamiento del modelo logístico, dado por la ecuación:

$$\text{Score} = \frac{\text{Offset}}{n} + \left(\left(\frac{\text{Factor}}{\text{Ln}(2)} \right) \left(\left(\frac{\beta_i}{n} \right) - \left(\beta_0 \right) \left(\text{WOE}_j \right) \right) \right)$$

Donde offset es un término de compensación y Factor es un término de reescalamiento.

El Offset y Factor deben satisfacer condiciones que son impuestas por microfinanciera. Este procedimiento permite la estandarización del score para que diferentes scorcards puedan ser comparadas. Se considera escalar el score de tal modo que el score total de 500 puntos corresponda al valor de las probabilidades buenas entre las malas sea de 50 a 1 y un aumento del score de 20 puntos corresponda a una duplicación de las probabilidades buenas entre malas.

El scorecard obtenido para cada variable se muestra en los siguientes cuadros.

Cuota / Utilidad al periodo	
Rangos	Score
≤15%	48.4519
15% - 30%	32.8727
30% - 45%	47.9479
45% - 60%	52.0915
60% - 75%	48.0962
> 75%	76.4058

Estado Civil	
Rangos	Score
CASADO(A)	52.0255
SOLTERO(A)	51.8375
VIUDO / DIVORCIADO	37.3374

Vivienda	
Rangos	Score
DUEÑO/PROPIA	50.3922
RENTADA/ES INQUILINO	55.5600
VIVE CON PARIENTES/AMIGOS/OTROS	52.5997

Tipo de Trabajo	
Rangos	Score
Tiempo Completo	53.0623
Autoempleado	52.5199
Ama de Casa	50.8814
Otros	46.9486

Frecuencia	
Rangos	Score
Semanal	50.7636
Quin - Men	54.3965
Otro	58.6250

Pago total a realizar	
Rangos	Score
0 - 2500	51.3214
2501 - 5500	50.8833
5500 - 7000	51.3527
7001 - 9000	51.7693
9000 - 11500	51.8389
11501 - 13500	51.8896
>13500	52.2188

Resultados Score por variable.

Tabla 08.

Prestamo / Utilidad	
Rangos	Score
0 - 0.5	53.0167
0.5 - 1	51.6161
1 - 1.5	51.9691
1.5 - 2	50.1336
> 2	49.6190

Prestamo solicitado	
Rangos	Score
0 - 2000	51.2538
2001 - 4000	51.4421
4001 - 6000	51.7784
> 6000	51.8846

Deuda	
Rangos	Score
1000 - 4000	25.9367
4001 - 7000	48.5505
7001 - 10000	56.3527
> 10000	60.4554

Finalmente, para determinar si el modelo aplicado predice correctamente los valores, se analiza el p-value. En la tabla de análisis de varianza que arroja el algoritmo utilizado en R, el p-valor es inferior a 0.01 por lo que se puede concluir que existe una relación funcional de los clientes buenos y malos explicada por el modelo ajustado. Por otra parte, el p-value correspondiente a cada variable es de igual manera menor a 0.01 por lo que todas las variables son aptas para el modelo que se desea aplicar.

Para hacer la validación del modelo de clasificación de buenos y malos se calcula la distribución de los clientes de acuerdo al score asignado. Para ello consideramos intervalos de longitud igual a 15 puntos de score para obtener un número considerable de clases, con esto se realizó la validación de la scorecard al calcular el índice de Gini y el estadístico K-S. El estadístico de Komogorov Smirnov (K-S) consiste en medir que tan diferentes son las funciones de distribución de buenos y malos para cada valor de score. Por otro lado, el índice de Gini es un estadístico utilizado para medir que tan bien se distinguen los buenos de los malos dentro del modelo scoring.

$$\mathbf{K - S} = \text{Max}_i (P_m(i) - P_b (i))$$

$$\mathbf{GINI} = 1 - \sum_{i=L}^H (P_m (i + 1) - P_m (i)) (P_b (i + 1) - P_b (i))$$

Donde:

i : es el valor del score, en el rango L - H, con $L \leq i \leq H$.

Pb (i), Pm (i) : Proporción de buenos y malos con score menor o igual a i.

Los resultados se muestran en la siguiente tabla.

Score	Buenos	Malos	FB	FM	AB	AM	KS	GINI
< 430	77	10	3.75%	0.82%	3.75%	0.82%	2.92%	
430 - 445	311	66	15.13%	5.42%	18.87%	6.24%	9.71%	0.9918
445 - 460	478	211	23.25%	17.32%	42.12%	23.56%	5.93%	0.9597
460 - 475	965	633	46.94%	51.97%	89.06%	75.53%	5.03%	0.7561
475 - 490	118	137	5.74%	11.25%	94.80%	86.78%	5.51%	0.9935
> 490	107	161	5.20%	13.22%	100.00%	100.00%	8.01%	0.9931
TOTAL	2056	1218	100%	100%			37.11%	94%

Resultados Gini por grupo de score.

Tabla 09.

Con la imagen anterior se puede observar que se obtuvo un Gini = 94% y un K-S = 37.11% para la muestra de entrenamiento.

Este mismo proceso se hizo para la muestra de validación con 819 registros correspondiente al 20% de los datos a analizar.

Score	Buenos	Malos	FB	FM	AB	AM	KS	GINI
< 430	74	12	14.18%	4.04%	14.18%	4.04%	10.14%	
430 - 445	116	33	22.22%	11.11%	36.40%	15.15%	11.11%	0.9753
445 - 460	191	95	36.59%	31.99%	72.99%	47.14%	4.60%	0.8830
460 - 475	92	74	17.62%	24.92%	90.61%	72.05%	7.29%	0.9561
475 - 490	27	42	5.17%	14.14%	95.79%	86.20%	8.97%	0.9927
> 490	22	41	4.21%	13.80%	100.00%	100.00%	9.59%	0.9942
TOTAL	522	297	100%	100%			51.70%	96%

Resultados Gini por grupo de score.

Tabla 10.

Obteniendo un Gini= 96% y un K-S=51.70% para la muestra de validación.

CAPÍTULO VI.

ANÁLISIS DE RESULTADOS

Un buen modelo de scoring debe satisfacer al menos dos condiciones, la primera es que tenga una fuerte capacidad predictora y la segunda es que la estimación de los parámetros tenga una alta precisión. Adicional a estas condiciones, existe otra de gran importancia ya que siempre se busca que el modelo que se planea aplicar sea lo más sencillo posible, es decir, que contenga el mínimo de variables explicativas y que satisfaga las dos condiciones anteriores. Teniendo en cuenta dichas condiciones surge la pregunta, ¿todas las variables explicativas son necesarias para construir el modelo?, o si alguna de ellas puede ser excluida. Para determinar que variables podrán ser excluidas se realiza una prueba de hipótesis. Con ayuda del modelo de regresión ejecutado en el software 'R', se obtuvo como resultados que las variables con mayor influencia en la calificación del cliente que determinan la probabilidad de incumplimiento para la microfinanciera fueron Cuota/Utilidad, Vivienda, Estado Civil, Tipo de Trabajo, Frecuencia del pago, Pago total a realizar, Préstamo/Utilidad y Deuda. Una vez identificadas estas variables proseguimos con el desarrollo del modelo de scoring, para analizar y dar

una interpretación de los resultados que se obtuvieron, se utilizó el índice de Gini en conjunto con la curva de Lorentz.

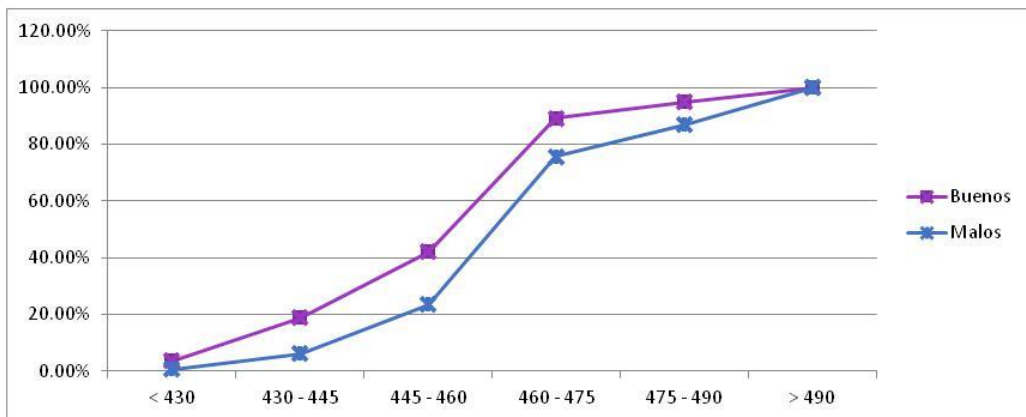
El Índice o coeficiente de Gini es un método utilizado en la estadística desarrollado por Corrado Gini, para obtener la desigualdad que representa la distancia entre la curva de Lorentz y la bisectriz, que corresponde a una distribución perfecta. El coeficiente de Gini es un número entre 0 y 1, donde 0 representa que toda la población analizada cuenta con la misma riqueza o ingresos y, el valor 1 señala el punto máximo de desigualdad. Este indicador no utiliza como parámetro de referencia la calificación media de la distribución, como es el caso del cálculo para la desviación media, la varianza y el coeficiente de variación. El índice de Gini es el coeficiente de Gini expresado en porcentaje.

La curva de Lorentz muestra gráficamente la relación que existe entre las proporciones acumuladas de población ($\%P_i$) y las proporciones acumuladas de ingreso ($\%Y_i$). De tal manera que si se presenta el caso en que a cada porcentaje de la población le corresponda el mismo porcentaje de ingresos ($P_i=Y_i; \forall i$), se forma una línea de 45° . En términos simples, la curva representa el porcentaje acumulado de ingresos de grupos de población ordenados de forma ascendente según a la cuantía de su ingreso.

De acuerdo con lo anterior, se ordenó a la población en forma ascendente respecto al valor de su score ($y_1 \leq y_2, \dots, \leq y_n$), una vez ordenados se forman grupos de igual tamaño, llamados percentiles, y con estos datos se realiza la gráfica con la curva de Lorentz.

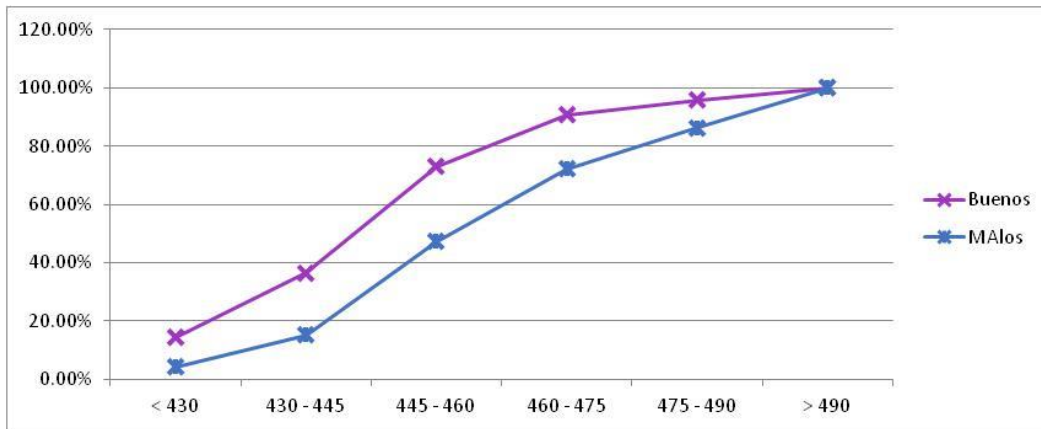
El estadístico 'K-S' mide la diferencia entre la media de las distribuciones de buenos y malos, se utiliza este indicador para saber que tanto difieren las poblaciones de buenos y malos. El modelo requiere de monitoreo conforme pasa el tiempo. Debe ser evaluado generalmente cada año para comprobar que se encuentra dentro de los márgenes de clasificación con respecto de la nueva población de lo contrario se debe ajustar.

A continuación se muestran los resultados obtenidos de forma gráfica:



Distribución de Buenos y Malos para la muestra de entrenamiento.

Gráfica 06.

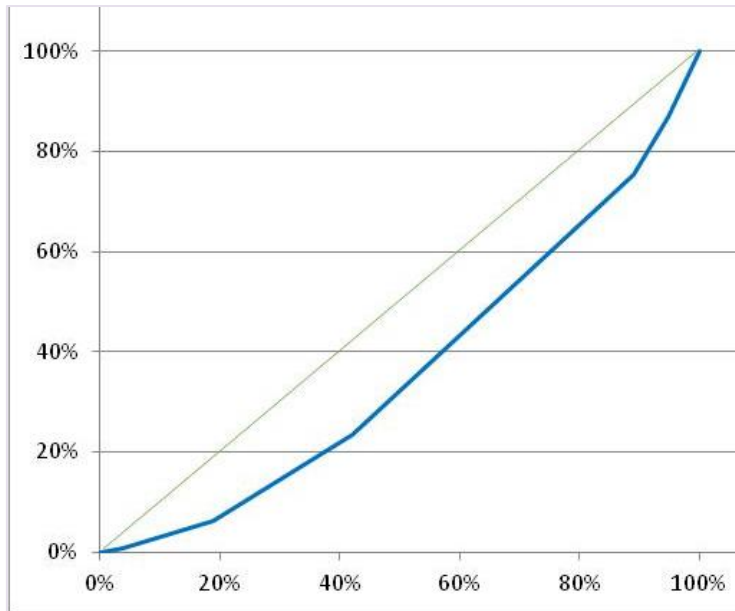


Distribución de Buenos y Malos para la muestra de validación.

Gráfica 07.

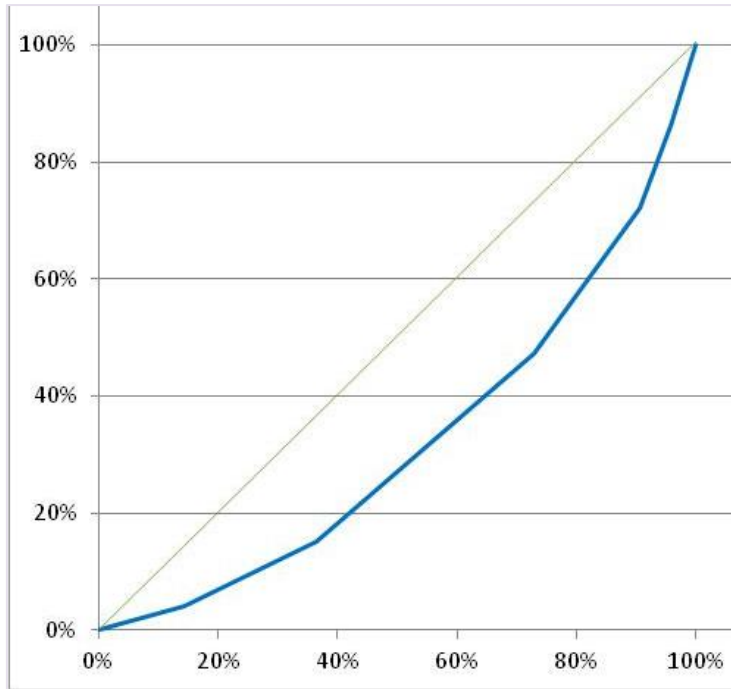
Se considera que un modelo con un KS menor a 20% es sospechoso, de 20% a 40% es satisfactorio, de 41% a 50% es bueno, de 51% a 60% se considera muy bueno y mayor a 60% se dice que es extraordinario. En las gráficas anteriores se puede observar la distribución de clientes buenos y malos, en donde para la primera gráfica se tiene un KS igual a 37.11% el cual corresponde a la muestra de entrenamiento, por otra parte en la segunda imagen correspondiente a la muestra de validación se tiene un KS igual a 51.70% por lo que se puede decir que es un muy buen modelo de scoring.

Otro de los métodos de analizar los resultados es mediante el índice de Gini y la curva de Lorenz, se ilustran a continuación las curvas obtenidas para ambas muestras:



Curva de Lorenz para la muestra de entrenamiento.

Gráfica 08.



Curva de Lorenz para la muestra de validación.

Gráfica 09.

Se considera que para un modelo de scoring, si el coeficiente de Gini es igual a 1, entonces el modelo separa perfectamente a buenos y malos. Un coeficiente menor a 35% se considera sospechoso, mientras que uno mayor o igual a 50% es más que satisfactorio.

En la imagen anterior se observa un Gini igual a 94% para la muestra de entrenamiento, mientras que para la muestra de validación se obtuvo un

Gini igual a 96%, en ambos casos se considera que se tiene un excelente modelo. Con los datos anteriores se realiza el cálculo para obtener el Cut off o punto de corte, el cual nos indica el score que la microfinanciera debe considerar para conceder un crédito. De acuerdo a los resultados mostrados, el solicitante debe obtener un score mayor a 499.82.

Es importante señalar que posiblemente la muestra utilizada y los resultados obtenidos están sesgados a una población en específico, debido a que la microfinanciera no incluye los datos de personas a las que no se les ha concedido crédito anteriormente, solo existen datos de clientes a los que se les ha otorgado uno o más créditos.

CAPÍTULO VII.

CONCLUSIONES

La idea de crear un modelo de scoring que se ajuste a las pyme's surgió de la necesidad de entender cómo se comportan los pagos realizados en este tipo de empresas, basándose en estudios y herramientas estadísticas, todo esto debido a que el crecimiento de estas empresas va en aumento en la economía mexicana.

Después de realizar este análisis se concluye que los sistemas financieros tanto en banca tradicional como microfinanzas son muy importantes para el crecimiento y desarrollo económico del país, la relevancia del sector microfinanciero recae en brindar servicios financieros a la población de bajos recursos que normalmente son excluidos del financiamiento tradicional y a su vez ayudan a combatir la pobreza en México, impulsando el desarrollo económico. Además estas instituciones apoyan los préstamos a mujeres lo que ha provocado un aumento de poder dentro de la sociedad, se manifiesta principalmente en las mujeres que han creado sus propias microempresas.

Debido a que se consideran a los modelos de scoring como herramientas sólidas para la toma de decisiones finales en el otorgamiento de solicitudes de crédito o rechazo de estas, se necesita un desarrollo preciso a la hora de aplicar un modelo. A pesar de que estos modelos no logran determinar exactamente el comportamiento de un cliente en específico, si logran dar un estimado del comportamiento promedio de la población que se evalúa y definir el punto de corte para rechazar a los clientes que no cumplan con el score requerido. La capacidad y eficiencia del modelo depende en gran medida de los datos de la muestra que se utiliza para la construcción, además del método utilizado en la determinación de los clientes “buenos” y “malos”. Adicional a estos puntos, se debe tomar en cuenta los factores económicos de cada región que podrían afectar de diversas formas a los modelos.

De los capítulos que conforman la tesis, el cinco se considera uno de los más interesantes, puesto que en él se plantea una situación práctica, con datos reales, correspondientes a una pyme, y en su desarrollo se aplicaron los conceptos mencionados en los capítulos anteriores. Se estudió el caso específico de créditos otorgados a mujeres jóvenes de un rango de edad de 20 a 35 años, los resultados arrojados nos muestran que esta población en particular suele tener una tendencia a cumplir

mayormente sus pagos, lo cual las convierte en buenos clientes para este tipo de empresas.

El primer paso en el proceso del Credit Scoring es la preparación de la base de datos, esto consiste en la limpieza de la base de datos el cual resulta ser un proceso bastante largo y tedioso que se intenta reducir con el uso de un software adecuado para esta tarea, en este proyecto nos auxiliamos de Microsoft EXCEL.

La determinación de los clientes como buenos o malos fue dada por la propia empresa que nos proporciono su base de datos. Analizando los datos se detecta que los malos clientes son los que tienen un desempeño menor o igual al 50% de sus pagos los cuales resultan ser más del 90% clientes nuevos que adquirió la microfinanciera, con esta información se aconseja modificar su método actual para dicha determinación. Probablemente este resultado está relacionado con el hecho de que la población en su base de datos esta sesgada al no considerar a las personas a las que en el pasado se rechazó su solicitud.

En general, se considera que el modelo y los resultados obtenidos son adecuados para ser utilizados en la práctica por la empresa. Aunque se obtuvo el resultado óptimo, éste se puede mejorar tomando en cuenta las observaciones que se acaban de mencionar.

BIBLIOGRAFÍA

- ❖ Pedro Carrasco, “La economía prehispánica”, Ensayos sobre el desarrollo económico de México y América Latina.
- ❖ Blanco H Gonzalo, “El Sistema Financiero en México”. (1996).
- ❖ Espin-García, O., & Rodríguez- Caballero, C., “Metodología para un Scoring de clientes sin referencias crediticias”. Cuadernos De Economía, (59), 139 -166. (2013).
- ❖ Finlay, S. “Credit scoring, response modelling and insurance rating”. Basingstoke: Palgrave Macmillan,. (2012).
- ❖ Mays, E., & Lynas, N. “Credit Scoring for Risk Managers”. [S.I.]: The Handbook for Lenders, (2010).
- ❖ Lara Rubio, “La Gestión del Riesgo de Crédito en las Instituciones de Microfinanzas”. (2010).
- ❖ Gómez Fernández P. & Partal Uraña A., “Gestión y Control del Riesgo de Crédito en la Banca”. Finanzas para la nueva economía.
- ❖ Santos Téllez Carla, “Riesgo de Crédito en México”. (1999).
- ❖ Comité de Supervisión Bancaria de Basilea, “Principios Básicos para una supervisión bancaria eficaz”. 1999, pp52.
- ❖ Galicia Romero, Martha, “Nuevos enfoques de Riesgo de crédito” México (2003).

- ❖ Elizondo Alan, I. Altman Edward, “Medición integral del riesgo de crédito”. (2004).
- ❖ Comunicado de prensa Núm. 600/18, 23 de noviembre 2018, págs., 1,2. Recuperado 29 de Septiembre 2019, < <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2018/OtrTemEcon/ENIF2018.pdf>
- ❖ Reporte sobre el Sistema Financiero 2018. (2018). Recuperado 19 de Mayo 2019, < <http://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-el-sistema-financiero/%7B7C5ED3AB-B558-0157-C1AF-168EAFB04EC7%7D.pdf> >
- ❖ Breve Historia del Sistema de Créditos en México, 15 de julio 2015. Recuperado 14 de Mayo 2019, <<https://prestadero.com/blog/la-historia-del-sistema-de-creditos/>>
- ❖ Situación Banca México (2018). Recuperado 19 de Mayo 2019, < https://www.bbva.com/wpcontent/uploads/2018/02/1802_SituacionBancaMexico_1S18.pdf >
- ❖ La inclusión financiera en México más allá de la bancarización. 11 de abril de 2019. Recuperado 19 de Mayo 2019 < <https://economia.nexos.com.mx/?p=2214>>

Gráficas

- 01 - Mercado de Crédito en México.
- 02 - Integración del Crédito al Sector Privado 2019.
- 03 - Estructura del Crédito al Consumo de la Banca Múltiple por producto 2019.
- 04 - Crédito bancario vigente por categoría: Consumo, Vivienda y empresas en 2017 y 2018 (Var% anual real).
- 05 - Cartera de Crédito Total Banca Múltiple.
- 06 - Distribución de Buenos y Malos para la muestra de entrenamiento.
- 07 - Distribución de Buenos y Malos para la muestra de validación.
- 08 - Curva de Lorenz para la muestra de entrenamiento.
- 09 - Curva de Lorenz para la muestra de validación.

Tablas

- 01 - Sucursales en operación 2019.
- 02 - Instituciones en operación al cierre de diciembre de 2019.
- 03 - Resultados WOE por variable.
- 04 - Reglas del criterio IV.
- 05 - Resultados IV's por variable.
- 06 - Ejemplo de base para cálculo de regresión.

- 07 - Resultados de β por variable.
- 08 - Resultados Score por variable.
- 09 - Resultados Gini por grupo de Score para la muestra de entrenamiento.
- 10 - Resultados Gini por grupo de Score para la muestra de validación.