

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS ADMINISTRATIVAS

UNIDAD DE TITULACIÓN

**DESARROLLO DE UN MODELO DE SCORE DE ORIGINACIÓN
DE CRÉDITO PARA PERSONAS NATURALES DEL SEGMENTO
DE MICROCRÉDITO PARA EL BANCO BANECUADOR B.P**

**TESIS PREVIA A LA OBTENCIÓN DEL GRADO DE MAGÍSTER EN
ADMINISTRACIÓN DE EMPRESAS CON MENCIÓN EN OPERACIONES DE
SECTORES ESTRATÉGICOS**

DANILO SANTIAGO CRIOLLO CHÁVEZ

danilo.criollo@epn.edu.ec

DIRECTOR: PhD (c). JOSÉ LUIS ROMÁN VÁSQUEZ

luis.roman@epn.edu.ec

2022

APROBACIÓN DEL DIRECTOR

Como director del trabajo de titulación **DESARROLLO DE UN MODELO DE SCORE DE ORIGINACIÓN DE CRÉDITO PARA PERSONAS NATURALES DEL SEGMENTO DE MICROCRÉDITO PARA EL BANCO BANECUADOR B.P** desarrollado por Danilo Santiago Criollo Chavez, estudiante de la Maestría en Administración de Empresas, mención Operaciones de Sectores Estratégicos, habiendo supervisado la realización de este trabajo y realizado las correcciones correspondientes, doy por aprobada la redacción final del documento escrito para que prosiga con los trámites correspondientes a la sustentación de la defensa oral

PhD (c). José Luis Román Vásquez

DIRECTOR

DECLARACIÓN DE AUTORÍA

Yo, Danilo Santiago Criollo Chávez, declaro que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional, puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Danilo Santiago Criollo Chávez

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Danilo Santiago Criollo Chávez, bajo mi supervisión.

PhD (c). José Luis Román Vásquez

DIRECTOR

AGRADECIMIENTO

Agradezco a todas y cada una de las personas que participaron en la elaboración de este proyecto, durante y después de mis estudios de postgrado, a mis profesores de los cuales me llevo un grato recuerdo.

A las autoridades y a los funcionarios de la Gerencia de Riesgos de BanEcuador B.P, por su respaldo y apertura.

En especial a las personas que me guiaron durante el desarrollo de este trabajo, a mi lector el Ing. Patricio Carrasco Medina y al Director José Luis Román.

DEDICATORIA

A mi Padre, que en vida no pudo verme graduado, pero de seguro sentirá un poco de orgullo desde algún lugar.

Danilo Santiago

ÍNDICE DE CONTENIDO

Capítulo I. Introducción.....	1
1.1. Antecedentes de la investigación	1
1.2. Planteamiento del problema	3
1.3. Objetivos	4
1.3.1. Objetivo general.....	4
1.3.2. Objetivos específicos.....	5
1.4. Pregunta de investigación.....	5
1.5. Justificación	5
2. Capitulo II. Marco Teórico.....	7
2.1. Riesgo Financiero.....	7
2.1.1. Tipos de Riesgos Financieros	8
2.2. BanEcuador como institución financiera.....	12
2.3. Rol en la economía ecuatoriana	13
2.3.1. Objetivos.....	14
2.4. Crédito Bancario	15
2.4.1. Condiciones de otorgamiento de un Crédito Bancario.....	15
2.4.2. Procedimientos en un Crédito Bancario	16
2.4.3. Clasificación de los Créditos Bancarios.....	19
2.5. Modelos de Riesgo de Crédito	22

2.6.	Modelo Scoring	25
2.6.1.	Funcionamiento del Modelo Scoring.....	25
2.6.2.	Elementos del Modelo Scoring	26
2.6.3.	Clasificación del Modelo Scoring.....	27
2.7.	Cartera Bruta.....	28
2.8.	Condición Actual en el Ecuador	28
2.8.1.	Mora crediticia.....	28
Capítulo III. Marco Metodológico.....		30
3.1.	Selección Y Consistencia De La Muestra	31
3.2.	Fuentes De La Data.....	34
3.2.1.	Calidad de la Data	35
3.3.	Determinación del periodo de modelización.....	38
3.4.	Definición de la variable objetivo (bueno y malo)	41
3.4.1.	Matrices de Transición	43
3.5.	Análisis de las variables independientes.....	48
3.5.1.	Análisis Exploratorio de la Data.....	49
3.5.2.	Análisis Uni-variado.....	55
3.5.3.	Análisis Bi - variado	59
3.6.	Poder predictivo de los regresores (WOE)	62
3.6.1.	Test de independencia CHI2 para variables categóricas.....	65
3.7.	Construcción del Modelo Estadístico.....	67

3.7.1. Construcción del modelo de regresión logística.....	68
3.7.2. Resultados – data total	73
3.8. Validación del modelo	74
3.8.1. Matriz de Confusión.....	74
3.8.2. Pruebas Estadísticas.....	77
3.8.3. Curva ROC	79
3.9. Implementación del modelo.....	83
3.9.1. Tabla Scorecard	84
Estrategias.....	86
Capitulo IV. Conclusiones y Recomendaciones.....	89
Bibliografía.....	91

Capítulo I. Introducción

En el desarrollo del presente apartado se presentan los argumentos que contextualizan la problemática de estudio, en tal sentido, se describen los antecedentes investigativos dentro del criterio de vigencia, se realiza la descripción del problema, se formulan los objetivos tanto general como específicos que conllevan a realizar una pregunta de investigación y finalmente la justificación.

1.1. Antecedentes de la investigación

Entre las investigaciones que sirven de referentes y generan aportes a las variables de investigación dentro del criterio de vigencia se alude a la realizada por (Tulcanaza y Noboa, 2021) en la cual realizan una propuesta de originación para el otorgamiento de créditos de consumo, con base a la información que registra una Cooperativa de Ahorro y Crédito del segmento 3, mediante el modelo *Scoring*, de tal manera que le permita realizar una estimación sobre la probabilidad de incumplimiento por parte de un socio.

La metodología utilizada para su ejecución se basó en el método cualitativo y cuantitativo, tomando como población la base de datos del COAC en el cual refleja los créditos vencidos, vigentes y cancelados. Se trata de una investigación de tipo descriptiva, a la cual se le aplicó un aprueba de significancia basada en el modelo estadístico Wald el test de razón de similitud para la comparación de modelos.

Las principales conclusiones indican que las variables que presentan mayor significación en la asignación de un *scoring* se refieren al género, edad, cargas familiares, nivel de educación, situación laboral, ingresos promedio mensual. Seguidamente se realizó la eliminación de variables que no representan consistencia con el signo correspondiente en el modelo *scoring*, lo que podría generar inestabilidad, haciendo necesario que se eliminaran los registros menos pertinentes y cuya información no estaba completa. Cómo resultado se obtiene un modelo que no se corresponde con las necesidades de la entidad financiera considerando para ello el rendimiento y las variables significativas con el signo correcto.

El aporte que se genera con el análisis de la referida investigación, radica en que si bien es cierto el modelo *scoring* de origen nación de crédito permite realizar una predicción sobre la solvencia irresponsabilidad del asociado, su correcta aplicación depende de las necesidades específicas de cada institución, por lo que es preciso analizar todas las aristas que permitan hacer un análisis integral sobre su pertinencia.

Del mismo modo, la investigación desarrollada por (Astudillo, 2022) elaborada para presentar el diseño de un modelo de créditos *scoring* para el otorgamiento de créditos en una entidad dedicada al financiamiento de vehículos, que permita identificar la probabilidad de incumplimiento de los acreedores en el cual se incluyan las variables de mayor relevancia acorde al método utilizado.

La metodología utilizada para el diseño del modelo *Scoring* fue de estimación a través de la regresión logística binaria en el software *Stata* 16, utilizando para ello una población constituida por una base de datos de 12.153 vehículos a otorgados a personas jurídicas y naturales en el periodo comprendido entre enero y abril del año 2020, en el cual se identificaron 51 variables cualitativas y cuantitativas sometidas a un análisis estadístico descriptivo.

Los resultados obtenidos evidencia que la predicción de probabilidad de incumplimiento es de un 83.7%, tomando en cuenta variables explicativas que resultan de relevante consideración para el otorgamiento del crédito, por lo que el investigador recomienda hacer una evaluación sobre el uso y costos que representa la aplicación de esta herramienta financiera a fin de determinar realmente en el flujo de operación, cuál es el valor que agrega esta al otorgamiento, por lo que puede ser aprovechable para estudios de líneas de crédito menos detallados.

El aporte que genera dicho análisis a la presente investigación, radica en que la implementación de este modelo para la oxigenación de créditos en lo que respecta al otorgamiento para la adquisición de vehículos amerita que se identifiquen las brechas entre el modelo y la cartera vigente de personas naturales, así como la importancia de la consideración de variables macroeconómicas.

Igualmente, se alude a la investigación realizada por (Montalván, 2019) en la cual realiza un contraste sobre la hipótesis que involucra el uso de las redes neuronales para actualizar y modernizar el *Credit Scoring* para una cartera de microcrédito, y de esta manera obtener un mejor resultado en la utilización de una metodología de regresión logística.

La metodología empleada se fundamenta en 2 modelos para ser desarrollados con base al programa estadístico R y a partir de este se realiza un análisis comparativo de la capacidad de predicción para diferenciar entre los clientes buenos y malos, adaptando los datos a partir de diferentes destinos estadísticos como KS, coeficiente Gini, matriz de confusión y criterio de información *Akaike*.

Los principales resultados obtenidos a través del desarrollo de esta investigación, revela que la metodología de redes neuronales permite fortalecer el modelo *scoring*, al emplear una regresión logística y permite corroborar que la hipótesis planteada es cierta, siendo eficiente para mejorar y modernizar el modelo empleado.

El aporte relevante que se deriva del análisis de la anterior investigación se centra en que el modelo *scoring* puede no ser adecuado para predecir el incumplimiento por parte del deudor, en el proceso de originación de crédito para personas naturales, sin embargo, es posible adaptar el modelo a través del fortalecimiento que genera la incorporación de una regresión logística.

1.2. Planteamiento del problema

Las entidades financieras ponen a disposición de los clientes diferentes tipos de crédito, quienes podrán obtener ventajas significativas a través de la adquisición de bienes y servicios. En tal sentido, las entidades bancarias a través de la Superintendencia de Bancos, establece términos y condiciones para asegurar la solvencia crediticia de sus clientes, reservándose el derecho de realizar un análisis crediticio que puede ser manual, o apoyado en herramientas basadas en sistemas de información, con el fin de establecer un criterio certero para la asignación de créditos de consumo.

En el caso de las personas naturales, la asignación de microcréditos por parte de las entidades financieras y bancarias genera un incentivo para la reactivación de la economía nacional, con la perspectiva de reducir los índices de pobreza y generar mejores ingresos familiares para la satisfacción de las necesidades, repercutiendo positivamente en la calidad de vida a través de la materialización de emprendimientos (Carvajal y Espinoza, 2020)

En tal sentido, las entidades bancarias podrían emplear el modelo experto, el cual se basa en los referentes empíricos del experto crediticio, que puede basarse en el criterio de “carácter, capital, capacidad, colateral y condiciones, denominado las 5 “C” (Tulcanaza y Noboa, 2019). Este método representa una probabilidad de riesgo elevado en el otorgamiento de créditos, debido a que el criterio podría estar viciado por subjetividades naturales del proceso.

Ante estas posibles eventualidades de insolvencia por parte el deudor, las entidades bancarias han establecido políticas adecuadas para reducir los riesgos, pero a la vez, deberán ser consecuentes con los clientes para atraer la inversión, por lo que se ha hecho necesario incorporar procesos basados en sistemas de información que generen el fortalecimiento de las políticas crediticias y permitan realizar una predicción efectiva en la originación de créditos.

En tal sentido, la presente investigación se avoca a realizar una propuesta de desarrollo de un modelo para la originación de créditos destinado a las personas naturales para el segmento de microcrédito para el Banco BanEcuador B.P. basándose para ello en el Modelo *Scoring*, tomando en cuenta datos cualitativos y cuantitativos y basados en diferentes variables pertinentes para obtener los mejores resultados.

1.3. Objetivos

1.3.1. Objetivo general

Desarrollar un modelo de Scoring de Originación de Crédito para personas naturales del segmento de Microcrédito para BanEcuador B.P.

1.3.2. Objetivos específicos

- Determinar la situación de la cartera bruta de microcrédito de BanEcuador.
- Identificar la variable “objetivo” en función de la mora, así como de las cosechas de la cartera de microcrédito.
- Determinar las variables que conformar el *scoring* de microcrédito por medio de análisis multivariado.
- Desarrollar el *scoring* para el producto de microcrédito para el banco BanEcuador. B.P
- Realizar las simulaciones del funcionamiento e incorporar ajustes al *scoring* para el segmento de microcrédito y las pruebas de estrés para la validación del modelo.

1.4. Pregunta de investigación

¿El modelo *Scoring* es efectivo para la originación de crédito para personas naturales del segmento microcrédito para el bando BanEcuador B.P.?

1.5. Justificación

El presente estudio se justifica desde el punto de vista teórico, ya que a través de la compilación de los referentes que sirven de contexto a las variables de estudio pueden obtenerse inferencias que permitan desarrollar el modelo de *Scoring* para el otorgamiento de créditos a personas naturales, abordada desde una perspectiva científica, ya que se apoya en fuentes confiables dentro del criterio de vigencia.

Desde la perspectiva práctica, el desarrollo de la investigación genera un aporte institucional que ofrece soluciones reales para el otorgamiento de microcréditos personas naturales, tomando como referente el modelo *Scoring*, de tal manera que sea aplicado a los procesos crediticios en el Banco BanEcuador B.P., minimizando los posibles riesgos propios de la actividad.

Finalmente, desde el punto de vista metodológica, la presente investigación se fundamenta en el método científico para la obtención, organización y procesamiento de los datos necesarios para el desarrollo del modelo *Scoring* para el otorgamiento de microcréditos a personas naturales del

Banco BanEcuador B.P., permitiendo la sistematización de datos cualitativos y cuantitativos y abordar los indicadores adecuados para la obtención de resultados acordes a las necesidades de la institución bancaria.

Capítulo II. Marco Teórico

2.1. Riesgo Financiero

Se puede definir al riesgo financiero como la viabilidad de producirse un acontecimiento que genere unas series de pérdidas capaces de afectar en una determinada institución su valía económica, en otras palabras, la posibilidad de padecer pérdidas de índole económico. Como complemento, se puede indicar que el riesgo financiero, es asociado a las dificultades que se generan, cuando se otorga un préstamo monetario y el prestatario pierde la capacidad de realizar los pagos correspondientes a la deuda contraída (Nevárez et al., 2021).

Pero en un sentido más amplio, el riesgo financiero se entiendo como la probabilidad de obtener un rendimiento que diverge del previsto, o explicado de otra forma, una desviación en lo contemplado para un resultado, que debe sus causas a cambios en las variables de índole financiero, se puede sintetizar lo anterior, definiendo el riesgo financiero como la posibilidad de alcanzar réditos diferentes a los concebidos por efecto de los cambios en las variables financieras (Martínez, 2019).

Otra definición importante de riesgo financiero a ser citada es la siguiente: “Cualquier actividad empresarial que implica incertidumbre y que, por tanto, pueda ocasionar alguna consecuencia financiera negativa para la organización. Por tanto, esta situación puede provocar la pérdida de capital” (Banco Santander, 2022, p. 1).

En relación a lo anterior (Orellana et al, 2019) comentan:

Se puede relacionar al riesgo con cualquier tipo de actividad que se realice, y, por ende, va a estar presente en las empresas y su entorno. En la actualidad, la intervención de más variables ha provocado que exista un mayor nivel de volatilidad, y por ende un mayor riesgo en el mercado. (p. 10)

Si se entiende, que toda organización con un fin productivo o económico debe evaluar constantemente, sus capacidades para realizar sus operaciones, por medio solo de, sus propios recursos o si necesita recurrir a una institución financiera para recibir un apoyo económico que aporte liquidez a la organización.

En el último caso, cuando se opta por recibir recursos externos, uno de los caminos más utilizados es la obtención de un crédito bancario, por lo que se hace necesario evaluar los riesgos financieros en los que se incurre al aplicar a dicho crédito.

2.1.1. Tipos de Riesgos Financieros

Para (Ceballos, 2022), el riesgo financiero contempla:

El riesgo de interés, de cambio y de reinversión dentro de ese riesgo sistemático que no se puede controlar y además contempla, el riesgo de endeudamiento, de crédito y de liquidez que hace parte de ese riesgo específico que si podemos controlar. (p. 1)

Para las instituciones bancarias, los riesgos pueden dividirse en riesgos de negocios y de procesos, los riesgos en los negocios, son inseparables de cualquier ejercicio bancario resultando tanto inevitables como previsibles, los niveles de exposición a estos riesgos los definen usualmente los directorios de las entidades bancarias, se enumeran entre estos riesgos el de liquidez, el de crédito, el de mercado, el riesgo país y el de tasa de interés.

Los riesgos de procesos, son causados por inapropiadas administraciones de los riesgos de negocios, por su constitución resultan evitables, se pueden subrayar entre este tipo de riesgos, al riesgo legal, al riesgo operativo, al riesgo por ausencia de control y al riesgo por la fragilidad del sistema de información (Martínez, 2019).

La tabla 1, realiza un resumen descriptivo de los riesgos indicados:

Tabla 1.

Tipos de Riesgos Financieros

Riesgo	Descripción
De Negocios	(Gaytán, 2018) lo define como: La pérdida potencial por cambios en las variables que inciden sobre la valuación de las posiciones por operaciones activas, pasivas o causantes de pasivo contingente, pudiendo mencionar entre otros las tasas
De Mercado	

Riesgo	Descripción
	<p>de interés, tipos de cambio e índices de precios entre otros (p. 123).</p> <p>La amplitud del alcance en este tipo de riesgos es alta, debido a que se origina tanto por las dinámicas entre la demanda y la oferta, como por el grado de incertidumbre en variables económicas. Debido a esta incertidumbre manifiesta en los riesgos de mercado de siempre tomarse en cuenta en los cálculos de los riesgos ya que, de una correcta valoración de los riesgos y las tasas de retorno, las operaciones financieras podrán ser realizadas por las instituciones de forma correcta. Este entonces se subdivide en:</p> <p>Riesgos de interés. Originados por las fluctuaciones en los tipos de interés en un determinado tiempo.</p> <p>Riesgos de los instrumentos financieros. Como su nombre lo indica, viene asociado a las variaciones en las valoraciones de acciones o bonos.</p> <p>Riesgos Cambiarios. Surge cuando se ejecutan operaciones donde se requiere un cambio en las divisas, factores como la posición y la volatilidad monetaria deben ser tomados en cuenta.</p>
De Liquidez	<p>Rey (2019) indica sobre el riesgo de liquidez, que: “se encuentra asociado a que, aun disponiendo de los activos y la voluntad de comerciar con ellos, no se pueda efectuar la compraventa de los mismos, o no se pueda realizar lo suficientemente rápido y al precio adecuado” (p. 1).</p> <p>Se entiende entonces, que dado un impedimento para la renovación de pasivos o de convenir distintos en situaciones regulares, o debido a la venta adelantada y forzada de activos con precios por debajo del valor esperado de venta para el cumplimiento de</p>

Riesgo	Descripción
	<p>obligaciones financieras, se coloca a la institución en una posición donde o vende activos con pérdida u obtiene recursos con un costo alto.</p> <p>Se espera que, la disponibilidad de la liquides sea adecuada para solucionar las obligaciones de la organización, sin recurrir de forma constante a fondeos de instancia ultima que resultan más onerosos.</p> <p>Se puede indicar que el riesgo de liquidez a la vez se divide en:</p> <p>Riesgo de liquidez por fondos. Se da cuando una organización es incapaz de cumplir con sus obligaciones de pago, ya sea por crisis financiera o por gestión inadecuada, lo que implica un descenso en los fondos.</p> <p>Riesgo de liquidez por mercados. Se vincula a situaciones donde la venta o compra de activos no puede realizarse en el precio esperado, lo que conlleva a una transacción a precios no adecuados o a aguardar un tiempo para realizar la transacción.</p>
De Crédito	<p>Se entiende como la probabilidad de un detrimento económico causado la contravención del prestatario o del prestamista, de una obligación contractual financiera, ya sea al no pagar, no hacerlo en el plazo establecido y pagar de forma inadecuada.</p> <p>Por lo general, los acreedores asumen este riesgo de incumplimiento de los deudores, manejando dos tipos de perdidas, las que cubre la provisión contable del prestamista se llaman esperadas, y las sobrepasan la provisión establecida son las inesperadas.</p> <p>Este riesgo puede ser subdividido en:</p> <p>Riesgo de crédito del emisor. Implica la probabilidad de perdida por incumplimiento en los pagos del</p>

Riesgo	Descripción
	<p>comprometido directo, para un emisor, el grado de riesgo dependerá de aspectos como el tipo de garantía y el plazo.</p> <p>Riesgo de crédito de la contraparte. Se refiere a la probabilidad de pérdida por incumplimiento en los pagos de un intermediario, al adquirir por ejemplo un título de deuda de una casa de bolsa, se dan las situaciones donde esta contraparte no esté en la capacidad de cubrir el título a su vencimiento, ya sea por una falla operativa o la quiebra de la entidad.</p> <p>Riesgo País. Es el riesgo referido a las pérdidas potenciales por inversiones en terceros países, ya sea que en este se encuentre el emisor o su contraparte, la calificación de este tipo de riesgo implica el estudio situacional de dicha nación, desde su forma de gobierno, las variables macroeconómicas que maneja, el control de divisas y las primas de riesgo.</p>
<p>De</p> <p>Procesos</p> <p>Legal</p>	<p>Para (Gaytán, 2018):</p> <p>Es la pérdida potencial por el incumplimiento de las disposiciones legales y administrativas aplicables, la emisión de resoluciones administrativas y judiciales desfavorables y la aplicación de sanciones, en relación con las operaciones que las instituciones llevan a cabo (p. 126).</p> <p>Se refiere entonces tanto, al grado de incertidumbre que generar una variación de importancia en las leyes y normativas de una nación o región, como las interpretaciones no adecuadas o deficientes de dichas leyes y normativas, lo cual deja a la institución en una situación desfavorable, y aunque en la mayoría de los</p>

Riesgo	Descripción
Operacional	<p>casos, la institución no puede tener una intervención directa, deben contemplarse para mitigar sus efectos.</p> <p>De acuerdo a (Rey, 2019) el riesgo operacional es: “la posibilidad de ocurrencia de pérdidas financieras originadas por fallos o insuficiencias de procesos, personas, sistemas internos, tecnología, y en la presencia de eventos externos imprevistos” (p. 1).</p> <p>Entonces la probabilidad de contraer pérdidas económicas deriva de proceder de forma ineficaz o deficiente, ante determinadas circunstancias, pudiéndose enumerar: al accionar personal no calificado o capacitado, actividades delictivas, la carencia de control interno, errores humanos, desastres naturales, accidentes y fallas tecnológicas.</p> <p>Estos riesgos en su mayoría se clasifican como no cuantificables dado lo complejo que resulta evaluar las probabilidades que ocurran y las consecuencias que generarían.</p> <p>Para la apropiada gestión de los riesgos operativos, se requiere definir de forma pertinente: procedimientos, puestos de trabajo, responsabilidades, medios y modos de control interno y los planes de acción ante eventualidades.</p>

Nota: Elaborado a partir de (Rey, 2019) y (Gaytán, 2018).

2.2. BanEcuador como institución financiera

El proceso de creación de la institución financiera conocida como BanEcuador comienza en el año 2015, cuando el expresidente Rafael Correa lo anuncia como un banco público, orientado a la inclusión, con horarios y créditos a las condiciones, necesidades y actividades del sector productivo popular, conformado por campesinos, pequeños comerciantes y productores, para el trece de mayo del mismo año, en el decreto Ejecutivo 677, el banco es

incorporado a la banca ecuatoriana, para el año 2016, el antiguo Banco Nacional de Fomento transfirió en cumplimiento a un nuevo decreto ejecutivo, la cartera con mejor calificación, los depósitos, los activos y los pasivos a BanEcuador, de manera que el 9 de mayo del 2016 comienza sus operaciones abiertas al público (BanEcuador B.P., 2023).

El fin de la institución será prestar una serie financieros para el ahorro, el crédito y la inversión, de acuerdo a la pauta de intermediar financieramente recursos de carácter privado y público, brindando atención a organizaciones asociativas en áreas productivas y a las medianas empresas, enfocándose en las áreas comerciales, de servicios y agropecuaria, y enmarcándose dentro del Plan Nacional de Desarrollo, específicamente en el ámbito de lo local y áreas urbanas marginadas, por medio de instrumentos bancarios de primero y segundo nivel (Gob.Ec, 2022).

2.3. Rol en la economía ecuatoriana

De acuerdo a (Lara, 2021) sobre BanEcuador:

La institución pública diseña sus estrategias considerando aspectos decisivos para los clientes, como impulsar los procesos productivos de tal manera que se alcance una efectiva sostenibilidad de la economía a través de procedimientos óptimos, capaces de competir en los mercados con estándares de calidad sobre los productos. (p. 5)

Los aportes de financiamiento de la institución bancaria BanEcuador, han resultado significativos en muchas regiones, convirtiéndose en una contribución al crecimiento económico, sobre todo de los microempresarios, siendo las principales actividades económicas con financiamiento: el sector comercial, el sector agrícola y el sector pecuario (Alava y Bravo, 2020)

Siendo BanEcuador un ente operante de las políticas públicas del poder ejecutivo dentro del gobierno nacional ecuatoriano, ha transferido una cantidad de recursos a sectores prioritarios, por medio de servicios financieros de crédito con bajas tasas de interés y líneas ajustadas a las particularidades de ciertos sectores productivos, pero como indica (Ponce, 2021) “se verifica que continúa aplicando metodologías e instrumentos crediticios de la banca convencional, al

solicitar al cliente requisitos y garantías de los que no disponen, ocasionando exclusión de sectores” (p. 127) entre ellos los de la llamada Economía Popular Solidaria.

De acuerdo al informe de gestión del año 2021 de BanEcuador (2022) se señala que la institución:

Concedió créditos por un monto de USD 540,662.3 a través de 110,831 operaciones de los cuales USD 520,162.3 miles en 110,789 operaciones corresponde a Crédito de Primer Piso (incluye original, novaciones, refinanciamientos y reestructuras) y USD 20,500.0 miles en 42 operaciones a Crédito de Segundo Piso. (p. 154)

Además, en el mismo informe se menciona sobre la participación de en los microcréditos, en el segmento 1 y el 2, que el banco alcanza la segunda posición con respecto a las Cooperativas de Ahorro y Crédito, además de la banca privada.

Para (Vera & Bravo, 2022) aun cuando BanEcuador representa un importante papel en el financiamiento agrícola este resulta insuficiente requiriéndose una participación más directa y decidida, que emplee una dirección más técnica que política a la gestión de la institución bancaria pública no sin obviar la rentabilidad requerida.

También sobre BanEcuador, de acuerdo al diario El Universo (2022):

Esta entidad bancaria pública también lleva adelante las medidas de alivio financiero acordadas en la mesa de diálogo sobre banca luego del paro nacional que vivió el país en junio pasado, para reestructurar o refinanciar deudas. (p. 1)

2.3.1. Objetivos

Unos de sus principales propósitos es desarrollar “productos y servicios financieros especiales y específicos para cubrir las necesidades de financiamiento de los sectores de la micro y pequeña empresa, en condiciones financieras preferentes” (p. 1).

Señala (BanEcuador, 2023) que su visión consiste en: “Ser un banco de desarrollo innovador y referente regional con enfoque agropecuario comprometido con el crecimiento sostenible e inclusivo de nuestro país” (p. 1).

Además, también señala que la misión es:

Somos un banco público de desarrollo que impulsa la producción, la inclusión, la asociatividad y la mejora de la calidad de vida de los micro, pequeños y medianos empresarios principalmente de agro negocios, comercio y servicios de los sectores rural y urbano popular, con productos financieros innovadores, eficientes y sostenible. (BanEcuador, 2023)

2.4. Crédito Bancario

La Superintendencia de Bancos del Ecuador (2019) define a este tipo de crédito bancario como “un contrato por el cual una entidad financiera pone a disposición del cliente cierta cantidad de dinero, el cual deberá de devolver con intereses y comisiones según los plazos pactados” (p. 1).

Entonces un crédito bancario es concedido por una institución crediticia, usualmente un banco, que puede ser de índole público o privado, a través de creación de un contrato por el que surge un endeudamiento, la diferencia con un préstamo, es que en este último el deudor recibe una cantidad monetaria y de cancelar intereses por la totalidad, no importando cuando de eso gaste realmente, mientras que en el crédito, el individuo cuanto con un capital a su disposición y solo cancela los intereses del monto usado.

2.4.1. Condiciones de otorgamiento de un Crédito Bancario

Señala (Sigvas, 2019):

Cada institución financiera u organismo crediticio tienen diferentes políticas para el otorgamiento de créditos; lo que buscan es minimizar el riesgo de impagos de los compromisos contraídos mediante el estudio de la solvencia del cliente. El estudio de la solvencia implica una sucesión de procesos de carácter administrativo. (pp. 32-33)

Como en la nación ecuatoriana tiene presencia la Superintendencia de Bancos del Ecuador, ente encargado de la supervisión y control de los servicios

financieros prestados tanto por la banca privada como la pública, pero en ambas bancas, las normas de riesgo crediticio son aplicadas de igual forma.

Las normas regulatorias de los bancos e las naciones de Latinoamérica se adecuan a los acuerdos de Basilea, lo que da forma también a los procedimientos para otorgar crédito, donde se exigen a los potenciales prestatarios a presentar una serie de requisitos y garantías, de acuerdo al tipo de crédito, que permitan la adecuada gestión del riesgo financiero, estas metodologías son propias de cada institución financiera, siempre que cumplan con las regulaciones de la Superintendencia de Bancos, del Banco Central del Ecuador y demás entes regulatorios (Ponce, 2021).

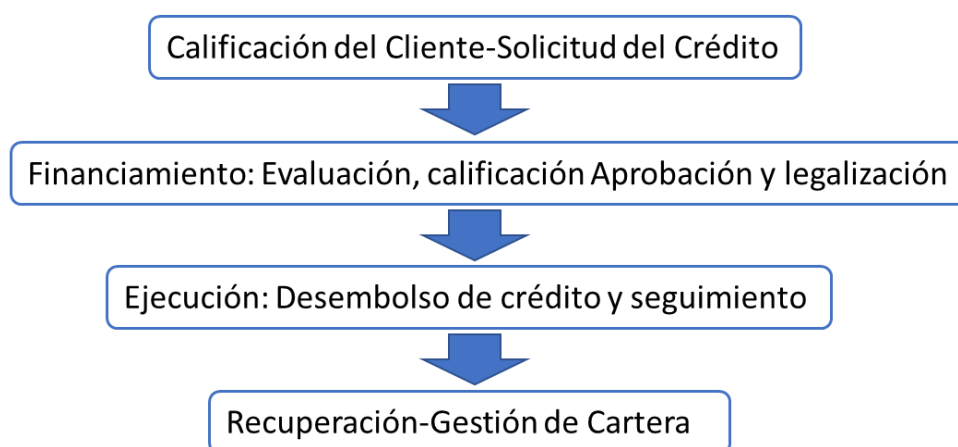
Se entiende, de acuerdo a (Guevara, 2020), que el crédito bancario “se basa en la comprobación de la solvencia del cliente verificando la misma a través de procesos rigurosos que den constancia de la misma” (p. 22) .

2.4.2. Procedimientos en un Crédito Bancario

La figura 1 muestra las fases del Ciclo de Crédito, de acuerdo al Banco del Desarrollo del Ecuador B. P.

Figura 1.

Fases del Ciclo de Crédito



Fuente: (Banco de Desarrollo del Ecuador B.P, 2021)

En la tabla 2 se desarrollan los procedimientos señalados, en las fases del ciclo de crédito:

Tabla 2.*Procedimientos del Ciclo de Crédito*

Procedimiento	Descripción
Calificación de los clientes	<p>La entidad bancaria cuenta con una metodología para calificar los riesgos crediticios y además de terminar el cupo de endeudamiento, usualmente, estas calificaciones, para un determinado cliente, tiene una vigencia de un año, la clasificación es evalúa el desempeño la transparencia del capital el nivel de endeudamiento, e indicadores macroeconómicos.</p>
Cupo de endeudamiento	<p>En este tipo de procedimientos, los métodos que son utilizados para definir el cupo de endeudamiento, tienen como fin definir el monto disponible para el otorgamiento, así como el monto efectivo de trabajo, existiendo un balance entre los ingresos y gastos que no puede ser destinados a la cobertura del crédito. Esto puede aplicar tanto a entes grandes, a empresas, a particulares y otros.</p>
Solicitud de Crédito	<p>Proceso mediante el cual el cliente da razón del cumplimiento de los requisitos de elegibilidad en función a la calificación del riesgo presentando la información requerida para el respectivo análisis y evaluación.</p> <p>El proceso comienza con la recepción de la solicitud de crédito y todos los documentos correspondientes en una sucursal del Banco, para lo cual el mismo debe entregar al solicitante toda la información pertinente para cumplir con los pasos solicitados.</p> <p>Si la organización que solicita el crédito no cuenta con un estudio adecuado que respalde su petición de crédito el Banco podría asesorar a dicha</p>

Procedimiento	Descripción
	organización para que solicite de forma adecuada el crédito bancario.
Evaluación del Crédito	Dentro de los aspectos técnicos, financiero y legal, la entidad bancaria evaluará la solicitud de crédito desarrollando un informe de evaluación de crédito que debe contener un perfil financiero y la decisión de aprobación. Todo incumplimiento del marco normativo vigente correspondiente al crédito bancario.
Aprobación	La aprobación del crédito la realiza el operador financiero correspondiente, lo cual se delimita en función al monto de financiamiento y a la calificación recibida en función al riesgo crediticio.
Legalización y contrato	El encargado o Junta jurídica del ente bancario, genera el contrato del crédito bancario, que incluye términos como: tasa de interés, plazo, forma de pago y garantías, además de otras condiciones de crédito, Y se proceda a la legalización y firma en un plazo máximo estipulado por el ente bancario luego de la notificación de aprobación del crédito bancario.
Desembolsos	<p>En todo contrato de un crédito bancario se establecen tanto los periodos de disponibilidad del crédito como los plazos para solicitar desde el primer hasta el último desembolso.</p> <p>Cumpliendo los requisitos establecidos en el contrato y las normas internas de la entidad bancaria el cliente presentará la filosofía la solicitud suscrita por su representante legal para tu desembolso.</p>
Forma de Pago, Recuperación y Garantías	Por medio de retenciones, pignoración de rentas, hoy inclusión de la deuda en un capítulo de deuda se podrá recuperar el crédito, siempre y cuando, no se cumplan con la forma de pago convenidas, lo que

Procedimiento	Descripción
	conlleve a requerir el uso de las garantías establecidas.
Terminación del contrato	La culminación usual de un contrato de crédito se da cuando dentro del plazo descrito se da el pago total del crédito además que el prestatario cubriera la totalidad de la deuda antes de la fecha de vencimiento o que éste renuncie voluntariamente a la totalidad o parte del crédito.

Fuente (Banco de Desarrollo del Ecuador B.P, 2021)

2.4.3. Clasificación de los Créditos Bancarios

Una clasificación de los créditos bancarios deriva en 3 tipos de crédito: el crédito comercial el crédito empresarial y los microcréditos, dentro del primero se brindan líneas de crédito para el consumo a hogares o individuos, en el segundo, se solventan los requerimientos de los empresarios, a través de líneas de crédito especializadas, en el tercero, los créditos son orientados a los microempresarios y personas con ingresos menores que requieren financiamiento para las producciones que ejecutan a menor nivel (Rivera, 2019).

Otra clasificación a ser descrita se presenta en la tabla 3, en base a Resolución 043-2015-F de la Junta de Política y Regulación Monetaria y Financiera (2015), sobre la segmentación de las carteras de crédito en la banca ecuatoriana:

Tabla 3.

Clasificación de los Créditos Bancarios

Tipo de Crédito	Descripción
Crédito Productivo	Este se otorga a personas naturales, las cuáles, están obligadas a llevar contabilidad o también personas jurídicas, con plazos superiores a un año para el financiamiento de proyectos productivos. Este crédito se subdivide en crédito productivo corporativo crédito productivo empresarial y crédito

Tipo de Crédito	Descripción
	productivo pymes, la diferencia entre dichos créditos es el registro de ventas anuales que para el corporativo debe ser superior a 5 millones de dólares para la empresarial entre un millón y 5 millones y para las pymes entre 100.000 y un millón de dólares
Crédito Comercial	El crédito comercial también se otorga a también se otorga a personas naturales y jurídicas cuyo registro de venta anual sea mayor a 100.000 dólares, al igual que el crédito productivo, este se subdivide, en crédito de comercial ordinario y comercial prioritario, el primero se destina para la compra y venta de vehículos livianos, el segundo para los servicios y bienes no categorizados en el anterior.
Crédito de Consumo	Este se entrega a personas naturales y se utiliza, de acuerdo a su suscripción en ordinario y prioritario, ya sea para la compra y venta de vehículos livianos de combustible fósil, o la compra de servicios y bienes que no tiene relación con una actividad productiva comercial incluyendo los créditos prendarios de joyas.
Crédito de Vivienda de interés público	Se entrega con una garantía hipotecaria a individuos naturales para la adquisición o construcción de una vivienda que sea única y de primer uso.
Crédito Educativo	Implica los servicios de crédito entregados a personas naturales para su formación técnica o profesional o para el financiamiento para la capacitación de un determinado talento humano.
Crédito de Inversión Pública	Se utiliza para el financiamiento de programas obras proyectos o servicios cuyo fin es el suministro de servicios públicos, cuya prestación es responsabilidad del Estado, de forma directa o por o ,, como intermediario de empresas ejecutoras.

Tipo de Crédito	Descripción
Crédito Inmobiliario	<p>A diferencia del crédito de vivienda este se para construir reparar remodelar y mejorar un inmueble propio, o también, para adquirir terrenos que se destinan a la construcción de una vivienda propia, y finalmente para la compra de una vivienda terminada que no se categorice dentro del crédito de vivienda de interés público.</p>
Microcréditos	<p>El microcrédito se entrega AA todas las personas jurídicas o naturales que posean un nivel de venta anual menor o equivalente a 100.000 dólares, guapa o un grupo vasco sección de prestatarios con una garantía solidaria, cuyo fin sea el financiamiento de tareas de producción o comercialización, en escala pequeña. Este crédito se subdivide en:</p> <p>Microcrédito minorista. Implica un saldo de adeudamiento menor igual a 1000 dólares incluyéndose el monto de la operación solicitada.</p> <p>Microcrédito de acumulación simple. En este el saldo de adeudamiento oscila entre 1.000 dólares y 10.000 dólares.</p> <p>Microcrédito de acumulación ampliada. Los saldos adeudados en este microcrédito son superiores a los 10.000 dólares.</p> <p>Sobre los microcréditos Vallejo & Ochoa (2019) refieren:</p> <p>Tienen mayor participación en las entidades financieras de la economía popular y solidaria que la banca pública y privada, porque dichas entidades se mantuvieron más restrictivas en el monto, plazo y tasa de interés al momento de otorgar el microcrédito a diferencia de las cooperativas de ahorro y crédito que estuvieron menos restrictivas tanto por la</p>

Tipo de Crédito	Descripción
	competencia de otras entidades como por los cambios de normativa legal (p. 146).

Fuente: (Espinoza, 2020) y (Asencio et al., 2021)

2.5. Modelos de Riesgo de Crédito

Para realizar un recuento de evolución de los modelos de riesgo de crédito se presenta la siguiente tabla:

Tabla 4.

Modelos de Riesgo de Crédito

Modelo de Riesgo	Descripción
Modern Portfolio Theory (MPT)	<p>Es un método de gestión de carteras para reducir el riesgo que tiene su origen en un artículo de 1952 del Premio Nobel Harry Markowitz. La teoría afirma que, dado un nivel de riesgo deseado, un inversor puede optimizar los rendimientos esperados de una cartera mediante la diversificación. Esto se consigue invirtiendo en activos menos correlacionados y agrupando los activos correlacionados con los que se mueven en direcciones opuestas entre sí, a fin de reducir el riesgo para una rentabilidad dada. En un gráfico, el conjunto de carteras que maximizan la rentabilidad esperada para una desviación típica dada se representa mediante la "frontera eficiente".</p> <p>La teoría moderna de carteras exige que se especifique una rentabilidad esperada para cada activo, pero esto puede resultar difícil. Aunque la rentabilidad esperada puede estimarse a partir de datos históricos, el pasado no es necesariamente indicativo del futuro. Según medidas estadísticas como la varianza y la correlación, el rendimiento de una sola inversión es menos importante que el impacto que tiene en toda la cartera.</p>

Capital Asset Pricing Model (CAPM)	<p>El modelo de fijación de precios de activos de capital fue desarrollado por los economistas financieros William Sharpe, John Lintner, Jack Treynor y Jan Mossin de forma independiente en la década de 1960.</p> <p>Este modelo describe la relación entre el riesgo sistemático, o los peligros generales de invertir, y el rendimiento esperado de los activos, en particular las acciones.</p> <p>Es un modelo financiero que establece una relación lineal entre el rendimiento requerido de una inversión y el riesgo.</p> <p>El modelo se basa en la relación entre la beta de un activo, la tasa libre de riesgo y la prima de riesgo de las acciones, o el rendimiento esperado en el mercado menos la tasa libre de riesgo.</p> <p>CAPM evolucionó como una forma de medir este riesgo sistemático. Se usa ampliamente en las finanzas para valorar valores de riesgo y generar rendimientos esperados para los activos, dado el riesgo de esos activos y el costo del capital.</p>
Value-at-Risk, VaR	<p>El valor en riesgo (VaR) es una estadística que cuantifica el alcance de las posibles pérdidas financieras dentro de una empresa, cartera o posición durante un período de tiempo específico.</p> <p>Esta métrica es la más utilizada por los bancos comerciales y de inversión para determinar el alcance y las probabilidades de pérdidas potenciales en sus carteras institucionales.</p> <p>Los gestores de riesgos utilizan el VaR para medir y controlar el nivel de exposición al riesgo</p> <p>El VAR se puede calcular utilizando diferentes técnicas. Bajo el método paramétrico, también conocido como método de varianza-covarianza, el VAR se calcula en función de la</p>

	<p>media y la varianza de la serie de retornos, asumiendo una distribución normal.</p> <p>Con el método histórico, el VAR se determina tomando los retornos pertenecientes al quintil más bajo de la serie (identificado por el nivel de confianza) y observando el más alto de esos retornos. El método Monte Carlo simula un gran número de escenarios para la cartera y determina el VAR observando la distribución de las trayectorias resultantes</p>
Stress Testing	<p>La prueba de estrés es una herramienta de gestión de riesgos que implica analizar los impactos de los escenarios extremos que son poco probables pero factibles. Resulta una técnica de simulación por computadora utilizada para probar la resiliencia de las instituciones y las carteras de inversión frente a posibles situaciones financieras futuras.</p> <p>La pregunta principal para las instituciones financieras es si cuentan con el capital y los activos líquidos adecuados para sobrevivir en tiempos de tensión. Las pruebas de estrés se realizan con fines regulatorios o para la gestión interna de riesgos por parte de las instituciones financieras</p>
Expected Shortfall (ES)	<p>Es una Medida de Riesgo utilizada en el contexto de la Gestión Cuantitativa del Riesgo de Mercado o del Riesgo de Crédito en un contexto de Gestión de Carteras. El Expected Shortfall a nivel es la Pérdida Esperada experimentada en la cartera en el peor de los escenarios.</p> <p>El Expected Shortfall es una alternativa al Valor en Riesgo que aborda algunos atributos ampliamente criticados del VaR como medida del riesgo. Las medidas de ES son susceptibles a la inclusión o exclusión de eventos extremos de muy baja probabilidad, por lo que las medidas de ES suelen ser menos estables que las medidas de VaR correspondientes</p>

Fuente: (Coronado y Carabrias, 2018)

La puntuación crediticia y la puntuación conductual son las técnicas que ayudan a las organizaciones a decidir si conceden o no crédito a los consumidores que les solicitan. Hay dos tipos de decisiones que deben tomar las instituciones financieras que confieren créditos. En primer lugar, deberían otorgar crédito a un nuevo cliente, las herramientas que ayudan en esta decisión, se denominan métodos de calificación crediticia. El segundo tipo de decisión es cómo tratar con los clientes existentes (Oquendo, 2022).

Si un cliente existente quiere aumentar su límite de crédito, ¿debería la entidad estar de acuerdo con eso? Si el cliente comienza a atrasarse en sus pagos, ¿qué acciones debe tomar la entidad? Las técnicas que ayudan con estas decisiones se denominan puntuación conductual.

Actualmente, la calificación crediticia se basa en métodos estadísticos o de investigación operativa. Las herramientas estadísticas incluyen el análisis discriminante, que es esencialmente regresión lineal, una variante de esta llamada regresión logística y árboles de clasificación, a veces llamados algoritmos de partición recursiva (Skura, 2020).

2.6. Modelo Scoring

Para (Moncada & Muñoz, 2020), el Modelo Scoring representa:

Una herramienta de discriminación utilizada para determinar un puntaje individual establecido a cada cliente, evaluando su alta o baja calidad de cumplimiento por el préstamo concedido por la institución financiera, con el fin de optimizar sus carteras de crédito para mejorar los indicadores de rentabilidad de la entidad y prevenir el origen de riesgo crediticio. (p. 14)

Se entiende entonces cómo un análisis de la data o información que provista sobre el ejercicio y características de un préstamo en el pasado para la pronosticación del cometido de préstamos a futuro. Un modelo de Scoring adecuado es capaz de identificar con un porcentaje mínimo de error a los buenos y malos clientes en los pagos de crédito-

2.6.1. Funcionamiento del Modelo Scoring

Los modelos Scoring para (Garzón, 2020) son:

(...) algoritmos que de manera automática evalúan el riesgo de crédito de un solicitante de financiamiento o de alguien que ya es cliente de la entidad. Tienen una dimensión individual, ya que se enfocan en el riesgo de incumplimiento del individuo o empresa, independientemente de lo que ocurra con el resto de la cartera de préstamos. (pp. 35-36)

Los bancos suelen conceder préstamos sobre la base de un modelo de calificación crediticia que combina el análisis cualitativo y cuantitativo. La calificación crediticia se basa en métodos estadísticos, gracias a los cuales es posible predecir la probabilidad de que ocurra un determinado evento en el futuro, en este caso, el incumplimiento de un préstamo.

El proceso de puntuación utiliza información sobre el cliente, recopilada en la etapa de solicitud, principalmente datos que caracterizan al cliente, pero también información sobre su comportamiento anterior. Cada institución de crédito considera un conjunto diferente de características y les asigna diferentes valores en puntos.

Por ejemplo, una persona con un alto nivel educativo generalmente obtendrá una puntuación más alta que una que abandonó la universidad, pero el valor exacto del punto y su impacto en la puntuación final pueden variar de un banco a otro. La suma de puntos de características particulares suele ser la puntuación final. El rango de puntuación posible lo determina cada banco o institución (Perez, 2019).

2.6.2. Elementos del Modelo Scoring

La evaluación del riesgo de las personas físicas se basa en sus características personales. Los más utilizados son: edad, estado civil, educación, número de dependientes, antigüedad, forma de empleo, ocupación, etc. La evaluación del riesgo de las personas físicas se basa en sus características personales.

Las características financieras también juegan un papel muy importante, como los ingresos mensuales, posibles ingresos adicionales o información sobre gastos (reembolso de otros préstamos, costos de vida, facturas pagadas, etc.). Si es posible, también se tienen en cuenta datos sobre el historial financiero del cliente, como el número de compromisos anteriores adquiridos y el historial de

reembolso de los mismos o información sobre posibles descubiertos. Dependiendo del tipo de crédito solicitado, la lista de características a tener en cuenta en la evaluación del riesgo puede variar (Garrido, 2019).

En el caso de las pequeñas y medianas empresas, el riesgo de quiebra e insolvencia se examina asignando puntos. Se tienen en cuenta su capital, su endeudamiento y su estrategia de desarrollo. En el caso de las empresas, el cálculo del crédito se realiza en función de: las características del sector, las características de la empresa, sus resultados financieros anteriores. El análisis del riesgo de crédito también se realiza teniendo en cuenta la estructura de la empresa, la fuente de financiación, la competencia, e incluso se examina la cualificación de los empleados, principalmente los de alto nivel.

El modelo de evaluación del crédito en el caso de una empresa también tiene en cuenta la situación financiera de la empresa, sus proyectos previstos, la liquidez, el pasivo financiero y la evaluación del riesgo del sector (Tulcanaza, 2021).

Sin embargo, en el caso de las empresas más pequeñas, muy a menudo, además de tener en cuenta los parámetros que caracterizan a la empresa, también se tiene en cuenta al propietario de la empresa examinada. Resulta que el perfil del propietario y su historial crediticio personal son más importantes que las cifras que describen su negocio. Esto es especialmente frecuente en el caso de las entidades más pequeñas que operan en el mercado desde hace poco tiempo.

2.6.3. Clasificación del Modelo Scoring

Los modelos de Scoring se pueden clasificar según diferentes criterios. Así, podemos hablar de un Scoring de personas físicas o jurídicas, división en función de la entidad evaluada, o de una tarjeta de crédito, efectivo o hipotecario, según el tipo de producto solicitado por el cliente.

Teniendo en cuenta quién creó y gestionó el modelo de Scoring, podemos hablar de Scoring interno, preparado por los bancos para sus propias necesidades, o Scoring externo, creado y puesto a disposición por instituciones especializadas, por ejemplo, oficinas de información crediticia (Vimos, 2021).

Hay una línea divisoria muy clara entre la aplicación y la calificación conductual. La primera está diseñada para evaluar nuevos clientes en base a los datos proporcionados por el cliente en la solicitud de crédito. El Scoring conductual, por otro lado, se determina sobre la base del historial de comportamiento del cliente en relación con el servicio de productos financieros.

Por lo tanto, se calcula para clientes habituales, principalmente para revender nuevos productos o cambiar los términos y condiciones de los productos existentes, por ejemplo, aumentar el límite de la tarjeta de crédito (Barreto, 2020).

2.7. Cartera Bruta

Para la Asociación de Bancos del Ecuador (2019) la cartera bruta es:

Es el total de créditos concedidos por un banco, esta se compone de la cartera por vencer más la cartera improductiva. A su vez, la cartera improductiva se compone de la cartera vencida y la cartera que no devenga intereses. En la cartera vencida se registran las cuotas de los créditos que dejan de ganar intereses o ingresos por más de 30 días. Una vez que la cuota se traslada a cartera vencida, las demás cuotas de ese crédito son transferidas a la cartera que no devenga intereses. (p. 2)

2.8. Condición Actual en el Ecuador

La Cartera Bruta de la Banca Privada Ecuatoriana de acuerdo a la Asociación de Bancos del Ecuador (2022):

Al cierre de octubre, la cartera bruta registró un saldo de USD 37.947 millones equivalente a un crecimiento mensual de 0,8% (USD 308 millones). De manera anual, la cartera bruta creció en 16,7%, lo que significó USD 5.441 millones más en la cartera de créditos. (p. 12)

2.8.1. Mora crediticia

Se producirá una mora crediticia, si un individuo u organización no cancela una deuda esperada, como un cargo por servicios públicos, una factura de teléfono móvil, un pago de tarjeta de crédito, un préstamo o un crédito bancario.

Pero los pagos perdidos o atrasados no siempre se convierten en incumplimientos. Se debe tener un plazo máximo de atraso en un pago de mínimo, y el proveedor de crédito debe haber tomado medidas para recuperar el dinero, si las deudas siguen sin pagarse, es probable que el proveedor se comunique con una agencia de calificación crediticia para informar el incumplimiento, que luego aparecerá en su archivo de crédito.

Un incumplimiento de puntaje de crédito permanecerá en su informe durante unos años y puede tener un efecto significativo en la capacidad para acceder a hipotecas, tarjetas de crédito, préstamos y otras formas de crédito.

Incluso si los prestamistas miran más allá del puntaje de crédito, es posible que no se obtengan las mejores ofertas y es probable que se deba enfrentar a tasas de interés más altas.

Lo que es peor, los incumplimientos permanecen en su informe durante años, incluso si se paga la deuda original. Sin embargo, eso no significa que deba dejar de pagar; cualquier deuda que se elimine, también se registrará en un informe de crédito, para que los prestamistas sepan que se está haciendo un esfuerzo.

Capítulo III. Marco Metodológico

Dada la naturaleza financiera del problema planteado, se utiliza un enfoque cuantitativo. En la primera etapa se selecciona el periodo en el cual se extraerá la data para la construcción del modelo, el periodo de análisis y los clientes en los cuales se validará el modelo.

La complejidad del estudio se encuentra en la minería de datos, donde se realizará el análisis de selección de las variables que se utilizarán en el modelo econométrico que se desea construir, para esto las herramientas de investigación son meramente cuantitativas (estadísticas) y no experimentales, ya que el listado de variables que participan del análisis no es extraído de manera aleatoria, son las que están vigentes y se consideren de mayor utilidad en el periodo de estudio.

Debido a que el análisis involucra medidas de tendencia central y de dispersión, es evidente que el diseño de la investigación es correlacional. Esta correlación estadística de variables permite la identificación de información significativa y que finalmente será usada para la validación de resultados, lo que abre el terreno a futuras investigaciones.

A fin de cumplir con los objetivos y la comprobación de la hipótesis se utilizará modelos econométricos de corte probabilístico, el que llevará a realizar contrastes de hipótesis, test de validación del modelos y demás evaluaciones que permitirán identificar los determinantes del acceso al crédito bancario. Este modelo de regresión probabilístico permitirá cuantificar la probabilidad que tienen las personas naturales para acceder o no al crédito.

La construcción del scoring de crédito será mediante un análisis de regresión logística con información proveniente de las solicitudes de crédito donde se registren la información socio demográfica de los solicitantes de microcrédito, así como su información proveniente del buró de crédito que captura el comportamiento del cliente en el sistema financiero, lo que permite elegir un perfil de cliente sujeto de crédito y un perfil de cliente no sujeto de crédito, para dar paso al procesamiento de una operación de crédito, esta herramienta de carácter predictiva ayuda a escoger clientes con probabilidades

de pago mayores y discrimina a potenciales clientes con menores probabilidades de no pago.

La base de datos que se utilizará será proporcionada por BanEcuador. B.P, la fuente principal será la solicitud de crédito, además de la información proveniente del Core bancario que captura el comportamiento e información adicional proveniente de buro de crédito.

Las etapas de construcción de un modelo de score estadístico varían de una institución a otra, dependen mucho de la madurez de los procesos del negocio o tecnológicos, dependen de las tecnologías disponibles, así como sus capacidades analíticas y experiencia que se tengan.

De manera muy sucinta se describen las siguientes etapas para la construcción:

- Selección y consistencia de la muestra
- Determinación del periodo de modelización.
- Definición de cliente bueno y malo
- Análisis de las variables independientes.
- Análisis del Poder Predictivo de las Variables Explicativas
- Construcción del Modelo Estadístico.
- Validación del modelo
- Implementación del modelo.

3.1. Selección Y Consistencia De La Muestra

En las instituciones financieras la data puede provenir de diferentes fuentes y puede ser recolectada de diferentes maneras, por eso es importante conocer de donde proviene cada grupo de variables, cuando es capturada y como es almacenada esa información. Para el desarrollo de un modelo de score se debe contar con una muestra de clientes previos, y su respectiva información crediticia; la cual permitirá explicar su comportamiento.

La data utilizada debe ser representativa y capturar a aquellos futuros o potenciales clientes que aplicarán a un crédito a futuro y ser suficientemente variada que capture la heterogeneidad de la población.

La data es recolectada en un formato “wide” o matricial, es decir; que cada fila representa una operación o un cliente y que cada columna represente una variable o característica. Si la data es a nivel de préstamo u operación, la información de los clientes asociada a esa operación debe reflejar información de esa época y no del último crédito. Si un cliente posee varios créditos en la institución la variable de “ciclo” debe ser creada y se debe seleccionar aquellos clientes cuyo ciclo sea igual a uno, debido a que captura la información de su primera aplicación en el caso de un score de originación, aquellos clientes con más de un ciclo de seguro tendrán más información y un mejor perfil de riesgo.

Generalmente la data es:

- Documental: Esta es tomada o verificada por los documentos oficiales como el documento nacional de identificación o pasaporte, esta constituye data oficial y legal pero usualmente desactualizada.
- Transaccional: Proviene del core bancario de las instituciones financieras, como: frecuencia de pagos y montos entregados o adeudados, la utilidad de esta data depende en qué parte del ciclo de crédito se encuentre el cliente.
- Auto – Reportada: Data auto declarada por el cliente y capturada por los formularios de aplicación al crédito, usualmente es información sesgada debido a que el cliente desea potenciar su aplicación al crédito.

La calidad de la data es uno de los problemas más importante que se debe solventar, ya que determinan la calidad del modelo de score, se la mide en función de la completitud (datos incompletos “missings” por cada variable y por cada individuo), unicidad de los registros como por ejemplo operaciones o clientes repetidos, y la validez, que las variables estén en los rangos creíbles o en los formatos correspondientes.

Como se mencionó inicialmente una de las dificultades que puede afrontar un modelo scoring es la escasez de una buena base de datos o incluso la ausencia de esta. Es importante, contar con un sistema de información adecuadamente construido con una lógica de almacenamiento de datos que permita contar con una codificación o representación numérica de las

características cualitativas y cuantitativas que servirán para la aplicación de técnicas estadísticas.

La definición y selección de los datos a incluir en el modelo requiere identificar las escalas de medida que pueden presentarse. Existen dos grupos de variables a estudiar: las cualitativas y las cuantitativas.

Las primeras son aquellas que no aparecen en forma numérica, sino como categorías o atributos (género, actividad económica, tipo de vivienda) y sólo pueden ser nominales u ordinales.

Las nominales permiten establecer frecuencias en cada atributo y la igualdad o desigualdad entre los diferentes casos, no es posible jerarquizar sus modalidades, estas variables se denotan asignando a cada atributo dentro de la variable un número.

La medida estadística que permite ver el grupo con mayor frecuencia es la moda. Por ejemplo, la variable sexo se define como Hombre: 1; Mujer: 0, donde los números son indicadores de pertenencia a una clase y no reflejan ninguna relación u orden.

Las variables ordinales recogen la idea de orden, pero carece de lógica realizar operaciones aritméticas con ellas o entre ellas, ya que no existe concepto de distancia entre una categoría y otra. En este tipo de variables se puede establecer igualdad y desigualdad y relaciones de mayor y menor, establecen orden, pero no hay métricas (distancias) dentro de ese orden.

Las variables cuantitativas, en cambio son aquellas cuyas categorías pueden expresarse numéricamente. Su naturaleza numérica permite un tratamiento estadístico más elaborado debido a las operaciones matemáticas que permiten.

Estas variables pueden ser discretas o continuas. Las discretas son aquellas cuyas categorías sólo pueden tomar valores enteros. Por ejemplo, la variable número de cargas familiares (no existen 2,3 cargas familiares). Las continuas son aquellas cuyas categorías pueden fraccionarse según cualquier entero, por ejemplo, la variable salario.

Una vez definido los tipos de datos con los que se puede trabajar se procede a seleccionar la información en función de las características disponibles.

3.2. Fuentes De La Data

Al ser un modelo de originación, es necesario que la data utilizada para la construcción del modelo de credit score, en su mayoría provenga de la solicitud de crédito y capture las condiciones iniciales con las que un cliente aplica a un crédito, junto con la información del buró de crédito constituyen las fuentes de información primordial para la construcción de este tipo de modelos, y se debe recalcar que en el sector financiero, bancario y comercial se considera un error usar información transaccional para modelos de originación, debido a que la información transaccional captura el comportamiento de pagos del cliente y no su estado inicial al momento de la solicitud del crédito.

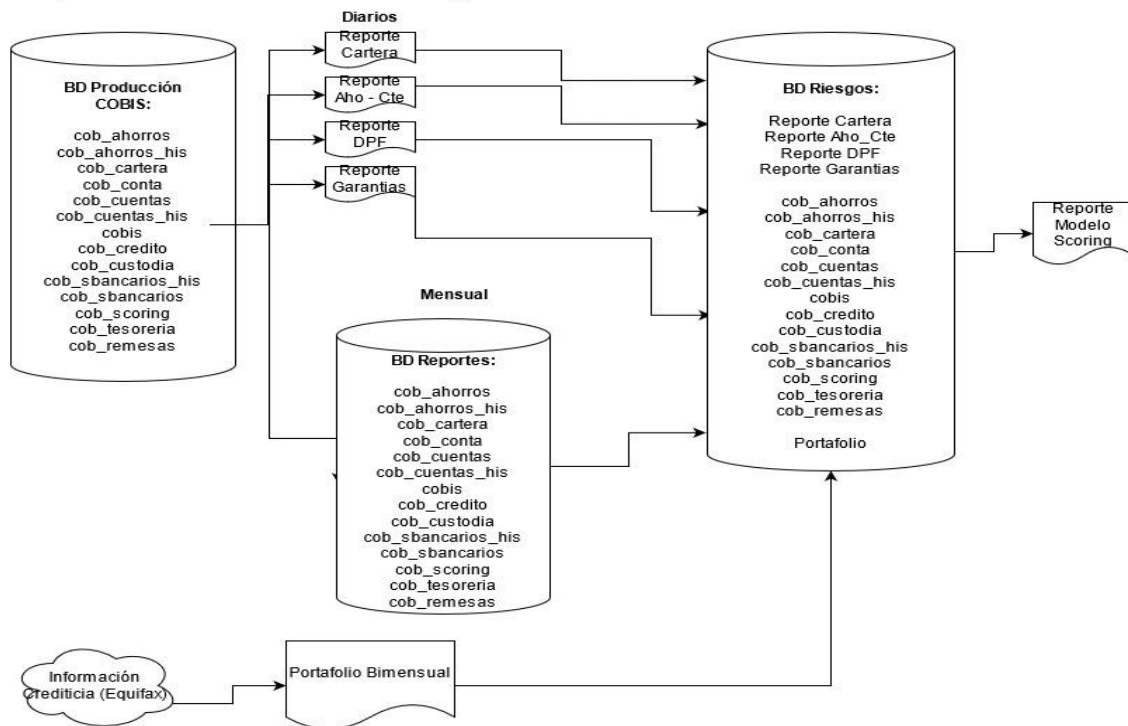
El archivo con la data que se usa para la elaboración del modelo de score, proviene de la base de datos BD RIESGOS de BanEcuador, que a su vez se alimenta de dos bases, una “BD Producción COBIS” que es información del core bancario, y que básicamente provee información de la cartera de créditos del Banco con un corte diario, debido a que es un ambiente de producción, la información se la obtiene de una réplica llamada “BD Reportes” que provee la información de cliente y de cartera (estado de las operaciones de crédito), que junto con la información crediticia obtenida del buró de créditos genera el archivo: “base_score.txt”, con 206 variables y 669000 mil registros con fecha de corte 2020-12-30 que se usará para el desarrollo del modelo estadístico.

En el Caso de BanEcuador, con fecha al 31 de agosto del 2021, el segmento micro representa alrededor del 97.27% de la cartera de crédito del Banco medido por número de operaciones, seguido por la cartera comercial con un 1.41% y con un 1.32% por la cartera de consumo. Si se realiza la misma comparación en términos de saldo, la cartera de microcrédito representa el 86.41%, la cartera comercial el 13.09% y la cartera de consumo tan solo el 0.50%.

Por lo que, el análisis de este segmento es de vital importancia para la institución, por esta razón la implementación de un modelo de score de originación es necesario.

Este archivo se lo somete a un proceso de depuración, en el cual se realiza eliminación de variables similares o duplicadas, se elimina variables del tipo string, o variables con codificaciones innecesarias, se elimina variables de comportamiento, se eliminan variables con presencia de registros nulos superior al 60%, se modifica el tipo de variables acorde a su naturaleza numérica, y se ajustan formatos como fechas etc.

Esquema Ambiente de Riesgos



Fuente: Riesgo de Crédito y Control Metodológico

Elaboración: Autor

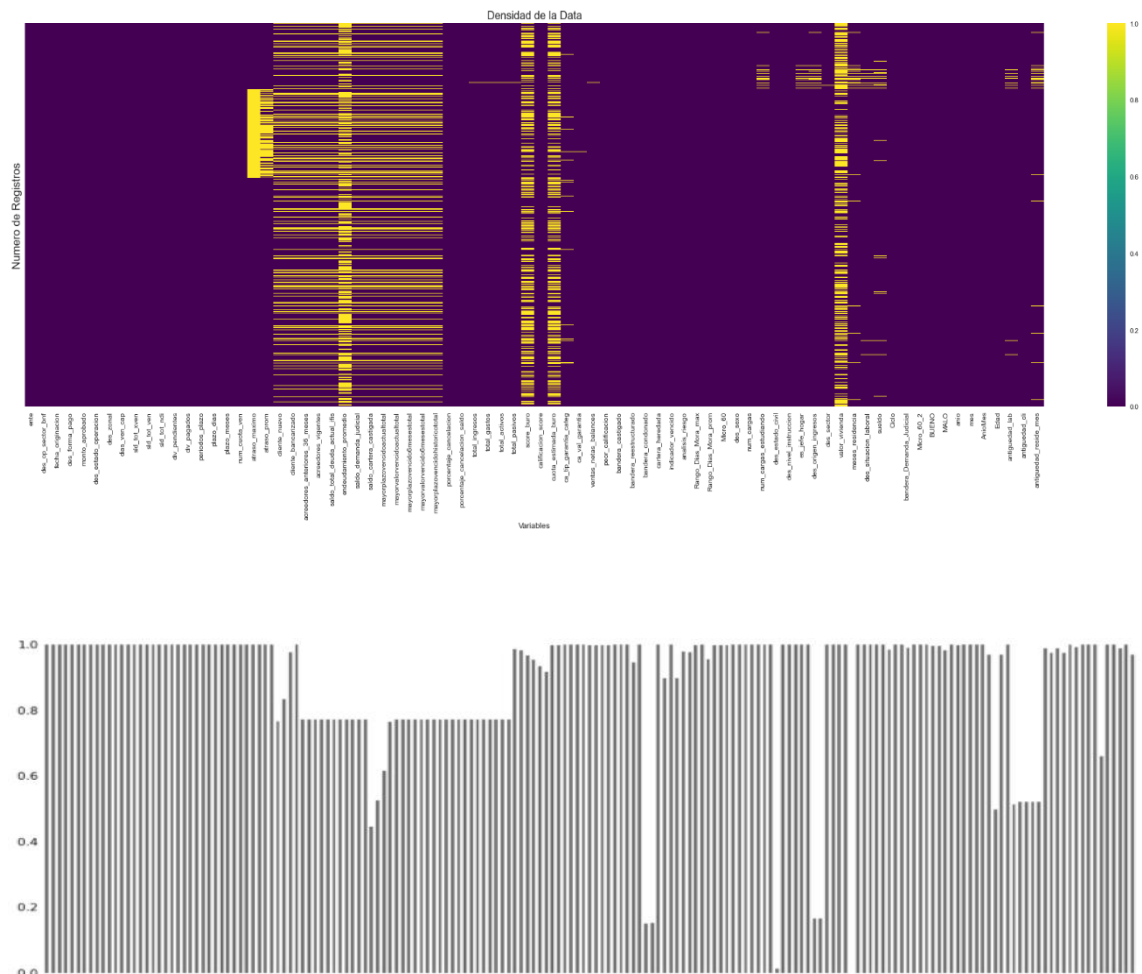
3.2.1. Calidad de la Data

Registros Duplicados. - Una de las primeras tareas es revisar la duplicidad de registros, y se verifica que el archivo no posee registros duplicados:

```
df_cliente.duplicated().sum() = 0
```

Además, un cliente puede tener más de un crédito, por lo que se selecciona los créditos que mayor mora presenten de cada cliente (el peor crédito), por lo que el número de registros se reduce a 426510.

Porcentaje De Nulos Por Variable. - Con respecto a la población de cada una de las variables del archivo con el que se va a trabajar, de manera descriptiva se tiene lo siguiente:



Se realizó un control de la presencia de valores nulos en cada una de las variables del archivo, y se encontró variables prácticamente vacías, por lo que se procede con su eliminación.

Var_Cliente	Num_Nulos	Nulos_Percen
p_pasaporte	426510	100.00%
codigo_analista_riesgo	421273	98.80%
fabrica_riesgo	421273	98.80%
cod_oficina_analista_riesgo	421273	98.80%
gerente_zonal	358955	84.20%
gerente_sucursal	357901	83.90%

Como se observa la base de datos presenta variables que cuentan con más del 60% de datos nulos por lo que se ordenan y se eliminan estas variables y variables que aportan información similar.

En BanEcuador el identificador único para cada cliente se llama 'ENTE', y el identificador único de cada operación es 'Nro_operacion', por lo que se eliminan columnas como: "cedula_ruc", "p_pasaporte", "nro_tramite", "nombre del cliente", etc.

Se eliminan variables que presentan códigos como: "cod provincia", "cod cantón", "cod_ profesión" etc, ya que la información relevante proviene de la vectorización de cada una de las variables categóricas como Provincia, cantón, profesión etc.

Con variables temporales como fecha de nacimiento, fecha de creación de clientes, fecha de ingreso laboral, se generaron variables como edad, antigüedad laboral, antigüedad de cliente, antigüedad de residencia etc., solo se mantiene la variable "Fecha contabilización" que será usada para segmentar la data en data de desarrollo y data de testeo.

Con respecto a las variables obtenidas del buró (endeudamiento promedio, mayor plazo vencido, acreedores vigentes, acreedores anteriores) se eliminan variables referentes a segmentos comercial o consumo, se mantienen variables de carácter global que incluyen deudas totales y no por segmento de crédito, por lo que deudas de los clientes en el sistema financiero (incluido segmento micro), estarían implícitas en la deuda del sistema financiero total, por lo que también se eliminan variables del segmento micro pese a que se desea construir un modelo de credit score para éste segmento.

Adicionalmente, las variables que BanEcuador obtiene del buró tienen un corte histórico, otro corte con información 36 meses atrás y el último es un corte vigente, por lo que de preferencia se toma el corte o variables históricas con el fin de reducir colinealidad o variables que estadísticamente aporten la misma información al modelo.

Se eliminan variables relacionadas a tasas de interés, ya que el ejercicio y la data es de corte transversal y no temporal, además que las tasas han sido fijas.

3.3. Determinación del periodo de modelización.

Con la finalidad de desarrollar un sistema scoring es necesario contar con una muestra de clientes que contenga información histórica, el comportamiento futuro será un reflejo del comportamiento pasado, es el supuesto más importante de este tipo de metodologías.

Este requerimiento genera una dicotomía entre dos elementos de decisión: a) la muestra debe ser representativa de aquellas personas (clientes potenciales) que deseen aplicar a una línea de crédito en el futuro y b) la muestra debe incorporar información suficiente acerca de las diferentes conductas de pago de los clientes (buena y mala conducta de pago) a fin de identificar las mejores características que lograrán recoger estos comportamientos en los clientes potenciales.

Por lo que se debe contar con un periodo de tiempo al cual se denomina periodo de construcción o desarrollo, en el que se observa la colocación y maduración de la cartera originada, para esto se debe usar un portafolio estable en términos de tasa de morosidad. Usualmente se selecciona como indicador a la tasa de morosidad definida como:

$$\text{Tasa de morosidad}_t = \frac{\# \text{ de malos clientes}_t}{\# \text{ total de clientes}_t}$$

En este periodo, el número de créditos considerados en default no debe crecer o la presencia de créditos malos no sea excesiva o nula, en ella, no se puede considerar grupos de créditos recientes pues no se contaría con el tiempo suficiente para observar el comportamiento de pago del cliente. Aquí se captura la información que será validada en el periodo de desempeño.

Acorde al grafico de tasa de morosidad histórica, se puede comprobar que el periodo de modelización parte desde el 2016-01 hasta el 2018-12, periodo en

el que se aprecia que la tasa de morosidad definida anteriormente es estable y se ajusta a la definición de madurez de la muestra.

Para evaluar el periodo de testeo del modelo, se usan operaciones comprendidas en el periodo 2019-01 hasta el 2019-12 (previa pandemia), en este periodo se aprecia que el número de operaciones es mínimo y por lo tanto puede sesgar los resultados, además, en microfinanzas existe el concepto que los modelos de calificación de originación deben considerar a aquellos clientes que obtienen un crédito por primera vez (CICLO 1).



Fuente: Riesgo de Crédito y Control Metodológico
Elaboración: Autor

Evaluando estas definiciones, para el periodo 2019, la tasa de morosidad decrece de manera rápida, lo que muestra poca madurez de la cartera y poca presencia de créditos, por lo que este criterio se descarta para ser usado con la cartera de Banecuator.



Fuente: BDD - Riesgos
Elaboración: Autor

Como se aprecia en el grafico anterior, la elección de una cartera madura no necesariamente implica una tasa de morosidad decreciente.

La elección de periodos con tasas de morosidad decrecientes no implica necesariamente una mejora en el comportamiento de la cartera, puede ser que estas cosechas por ser cercanas a la fecha actual estén reflejando carteras poco maduras y por ende no comparables con otros periodos de análisis.

Por tanto, una buena cosecha será aquella que mostrando una adecuada maduración tiene un bajo porcentaje de malos clientes y una cosecha será mejor que otra (en el mismo producto o productos similares) siempre que su porcentaje de malos sea menor.

Para validar, el periodo de desempeño del modelo estadístico, se usa aquellos créditos del segmento micro, que hayan pagado al menos el 50% de sus cuotas, lo que se considera una cartera madura



Fuente: BDD - Riesgos
Elaboración: Autor

El periodo de observación es el tiempo t en el que el investigador decide situarse y observar el desempeño del cliente. Es este periodo de desempeño o performance el que va a ser empleado para predecir el comportamiento futuro de los potenciales clientes. En el punto de resultado se asigna una calificación (bueno o malo) al cliente con base en un resumen del comportamiento en el periodo de desempeño.

De ahí la importancia de la madurez de la cartera para no calificar como bueno a un cliente que es malo, pero que no logra denotar un comportamiento porque inicia a pagar su obligación. En este periodo se evalúan los créditos originados en el periodo de construcción, usualmente fijado en un lapso de 12 meses para un modelo de originación, tiempo en el cual se puede definir el comportamiento de pago, y evaluar al crédito como bueno o malo.

Por lo que para concluir se resalta que, para el caso de BanEcuador, el periodo de desarrollo del modelo será el comprendido entre 2016-07-01 hasta el 2018-10-31, esto debido a que la tasa de morosidad de la cartera madura se mantiene estable en ese periodo de tiempo, además de que el plazo promedio de una operación de crédito medido en meses es de 39, por lo que ésta ventana de tiempo estaría acorde al plazo. Y como periodo de prueba para el modelo sería el comprendido entre: 2018-11-01 hasta 2019-12-31, esto debido a que a diciembre del 2019 la tasa de morosidad desciende drásticamente.

3.4. Definición de la variable objetivo (bueno y malo)

Cuando se habla de clientes malos, se busca describir al conjunto de clientes que BanEcuador, no quiere seleccionar para su actividad crediticia en función de la experiencia obtenida en su pasado reciente. Para el caso de los modelos de riesgo, usualmente esta definición hace alusión la cartera que de conocer su comportamiento no se hubiese aceptado.

De aquí se sigue que la definición de bueno y malo se basará prácticamente en el comportamiento de pago de los clientes: mora máxima histórica, mora promedio, contadores de mora (número de veces que ha caído en mora o reincidencia). Es evidente que estas definiciones pasan por un grado de subjetividad siendo tan conservadoras como lo quiera la entidad o viceversa.

En el caso de contar con varios productos se puede optar por una definición de bueno y malo para cada uno, los productos financieros no son homogéneos y sus características pueden influir en el comportamiento de pago de los clientes.

Por ejemplo, la definición de bueno para el caso de un crédito hipotecario puede ser más estricta que la requerida para una tarjeta de crédito, es habitual

encontrar moras de 30 días en tarjeta habientes, por lo que la definición de bueno en este caso no requiere ser tan ácida (muy bajo atraso medio y atraso máximo). Aunque finalmente todo dependerá de los objetivos en reducción del riesgo que se quiera plantear la institución.

Una aproximación para hallar el corte en días de mora, en el cual el cliente muestra un deterioro significativo y se lo puede clasificar como mal pagador, es el uso de las matrices de transición, estas generalmente usan dos estados un inicial y uno final, de preferencia este estado final debe ser a un año de otorgado el crédito o de clientes que al menos hayan pagado el 50% de sus cuotas, con el fin de evidenciar un comportamiento estable de pagos.

En el caso de BanEcuador, las matrices de transición se construyeron en base a las operaciones mensuales de microcrédito y se extrapola para créditos con frecuencia de pago diferentes a la mensual. Primero se muestra la transición a un año, luego a dos años y finalmente a tres años, esto debido a que el plazo de los créditos vigentes en BanEcuador es mayor al año y medio. La ejecución se efectuó como comparativo del stock de cartera a enero 2017 respecto a diciembre 2019, en pares mensuales. A continuación, se muestra el promedio de las matrices mensuales por año:

3.4.1. Matrices de Transición

DÍAS PROMEDIO	DÍAS PROMEDIO 2017 (1 AÑO)						
Día Morosidad	0	[1-30]	[31-60]	[61-90]	[91-120]	[121-360]	[+360]
0	88.48%	8.18%	2.16%	1.15%	0.00%	0.00%	0.03%
[1-30]	41.16%	32.67%	17.41%	8.19%	0.00%	0.00%	0.57%
[31-60]	18.24%	40.67%	22.29%	17.19%	0.00%	0.00%	1.62%
[61-90]	13.84%	56.15%	15.21%	12.97%	0.00%	0.00%	1.82%
[91-120]	10.21%	66.83%	15.06%	6.41%	0.00%	0.00%	1.49%
[121-360]	9.45%	36.10%	50.42%	3.62%	0.00%	0.00%	0.41%
[+360]	9.10%	44.07%	19.37%	0.18%	0.00%	0.00%	0.01%

DÍAS PROMEDIO	DÍAS PROMEDIO 2018 (2 AÑOS)						
Día Morosidad	0	[1-30]	[31-60]	[61-90]	[91-120]	[121-360]	[+360]
0	78.55%	13.93%	5.35%	2.10%	0.00%	0.00%	0.06%
[1-30]	40.72%	30.84%	19.98%	7.93%	0.00%	0.00%	0.53%
[31-60]	20.44%	32.77%	35.31%	10.94%	0.00%	0.00%	0.55%
[61-90]	19.41%	26.96%	34.98%	17.52%	0.00%	0.00%	1.12%
[91-120]	25.05%	27.64%	31.37%	15.04%	0.00%	0.00%	0.91%
[121-360]	21.35%	25.01%	35.15%	18.49%	0.00%	0.00%	0.00%
[+360]	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

DÍAS PROMEDIO	DÍAS PROMEDIO 2019 (3 AÑOS)						
Día Morosidad	0	[1-30]	[31-60]	[61-90]	[91-120]	[121-360]	[+360]
0	66.85%	17.30%	9.00%	6.40%	0.00%	0.00%	0.44%
[1-30]	34.60%	27.93%	18.73%	16.16%	0.00%	0.00%	2.59%
[31-60]	18.89%	28.93%	16.15%	30.12%	0.00%	0.00%	5.92%
[61-90]	8.48%	34.74%	10.51%	29.73%	0.00%	0.00%	16.54%
[91-120]	28.01%	37.14%	6.93%	16.17%	0.00%	0.00%	11.75%
[121-360]	15.24%	42.46%	4.28%	18.60%	0.00%	0.00%	19.43%
[+360]	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Fuente: BDD - Riesgos

Elaboración: Autor

Se observa que, para el caso de BanEcuador, la mayor concentración se ubica 31-60 días; es decir que la mayor concentración o deterioro, se ubica de 31-60 días, por lo que se escoge ese corte como rango en el que un cliente de BanEcuador es considerado como pagador. Adicionalmente, se considera los ejercicios de score realizados por el ente regulador, en que la definición de incumplimiento se la realiza mediante la metodología de roll rate, con la que se determinó que a partir de los 90 días de morosidad la mayor parte de los créditos del sistema financiero ecuatoriano tienden a empeorar su situación.

Acorde a la normativa vigente y detallada por el ente regulador, en el Libro I de normas generales para las instituciones del sistema financiero, en el en el título IX de los activos y límites de crédito, en su capítulo II en lo referente a calificación de activos de riesgo y constitución de provisiones por parte de las instituciones controladas, en su sección II , artículo 5, literal I de cartera de créditos y contingentes, se menciona que, en lo referente a créditos del segmento micro, "...la cobertura de la calificación cubrirá la totalidad de las

operaciones de microcrédito concedidas por la institución ... y en base a los siguientes parámetros:"

Tabla 5.

Cobertura de Calificación de Microcréditos

CATEGORIAS	DIAS DE MOROSIDAD
A-1	0
A-2	1 - 8
A-3	9 - 15
B-1	16 -30
B-2	31 - 45
C-1	46 - 70
C-2	71 - 90
D	91- 120
E	+ 120

Fuente: Superintendencia de Bancos del Ecuador – Libro 1 Normas Generales

Elaboración: Autor

Con el fin de tener alineado el puntaje del score de originación, la calificación del cliente y los días mora que presenta el cliente en su crédito, se construye la matriz de atraso promedio y atraso máximo, la cual es una técnica sencilla para definir los clientes buenos y clientes malos en función del atraso promedio y atraso máximo que consiste en listar en filas los rangos de atraso máximo y en columnas los rangos de atraso medio, en las celdas de la matriz se van situando los clientes distribuidos de acuerdo con el par ordenado de valores (atraso medio; atraso máximo) de una cartera madura.

Tabla 6.

Matriz atraso promedio atraso máximo

Rango_Dias_Mora_max	0	0-1	1-8	8-15	15-30	30-45	45-60	60-90	90-120	+121	All
Rango_Dias_Mora_prom											
0	23.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	23.0%
0-1	0.0%	3.0%	11.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	15.0%
1-8	0.0%	0.0%	7.0%	6.0%	7.0%	2.0%	1.0%	0.0%	0.0%	0.0%	23.0%
8-15	0.0%	0.0%	0.0%	2.0%	2.0%	3.0%	1.0%	1.0%	0.0%	0.0%	10.0%
15-30	0.0%	0.0%	0.0%	0.0%	1.0%	2.0%	1.0%	2.0%	1.0%	1.0%	8.0%
30-45	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	0.0%	1.0%	1.0%	1.0%	4.0%
45-60	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	2.0%
60-90	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	0.0%	2.0%	3.0%
90-120	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.0%	2.0%
+120	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%	10.0%
All	23.0%	3.0%	18.0%	9.0%	11.0%	8.0%	4.0%	5.0%	3.0%	17.0%	100.0%

Fuente: BDD - Riesgos

Elaboración: Autor

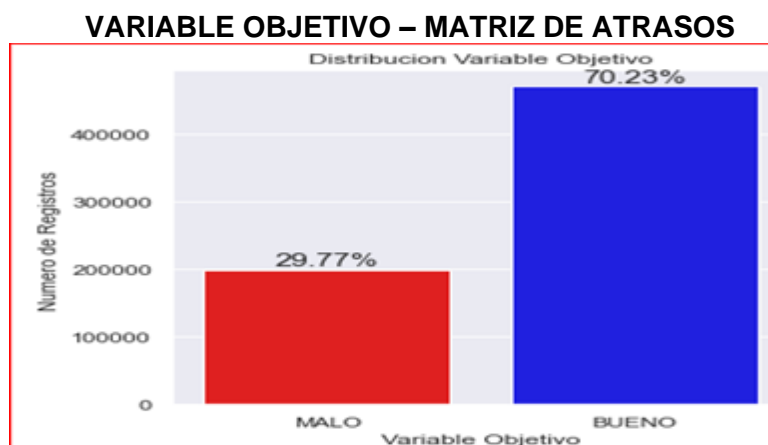
Esto, se lo realiza con el fin de no perjudicar a clientes que periódicamente realizan pagos dentro del plazo correspondiente y que alguna vez se hayan atrasado por motivos de enfermedad o fuerza mayor y en el caso contrario de no beneficiar a clientes que continuamente realizan pagos atrasados fuera del periodo de pago.

Una vez establecida la definición de buenos, se entiende como malos al complemento, es decir al resto de la población distribuido en las restantes edades de mora.

Considerando que BanEcuador es un Banco de desarrollo y que la evidencia de las cosechas muestra deterioros de cartera significativos superiores a los 36 meses, inicialmente se define a un cliente como “bueno” a aquel cliente que muestre una mora máxima y una mora promedio menor a los 60 días.

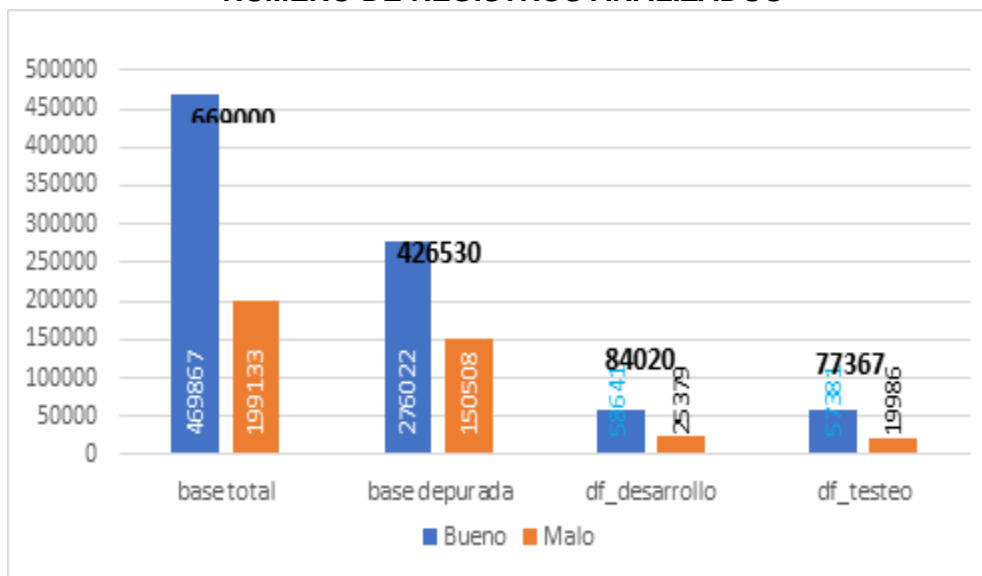
Y acorde a la tabla de cobertura del ente regulador, se estaría aceptando como clientes buenos hasta una porción de clientes con calificación C1, debido a que BanEcuador es un banco de desarrollo y debe ir a segmentos y poblaciones más necesitadas y por lo tanto el banco asume una exposición y un riesgo superior en comparación a otras instituciones financieras en este mismo segmento.

En esta etapa del desarrollo del modelo, es imprescindible conocer la variable target u objetivo que es la de buenos y malos, para solventar los potenciales problemas de balance de la muestra, es decir que el número de clientes buenos y de malos no sea desproporcionado.



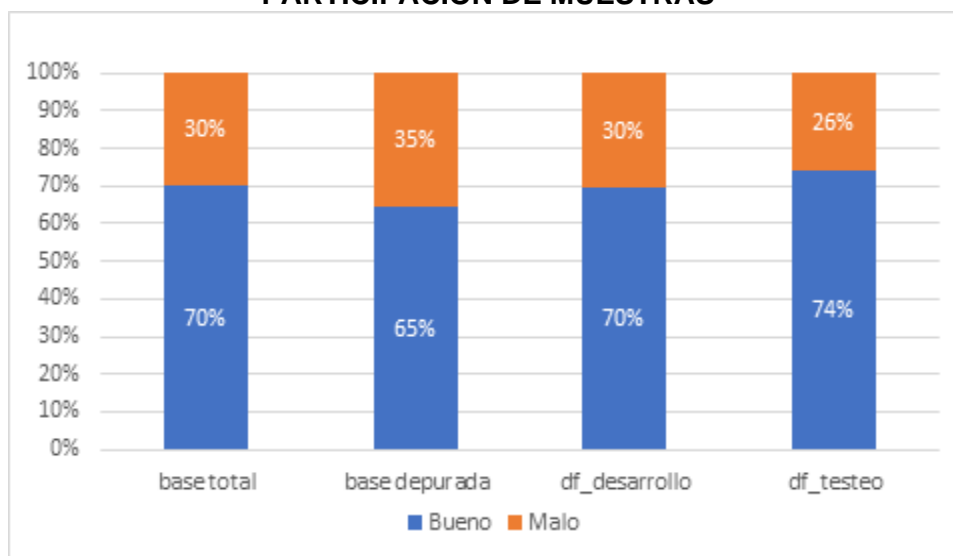
Fuente: BDD - Riesgos
Elaboración: Autor

NÚMERO DE REGISTROS ANALIZADOS



Fuente: BDD - Riesgos
Elaboración: Autor

PARTICIPACIÓN DE MUESTRAS



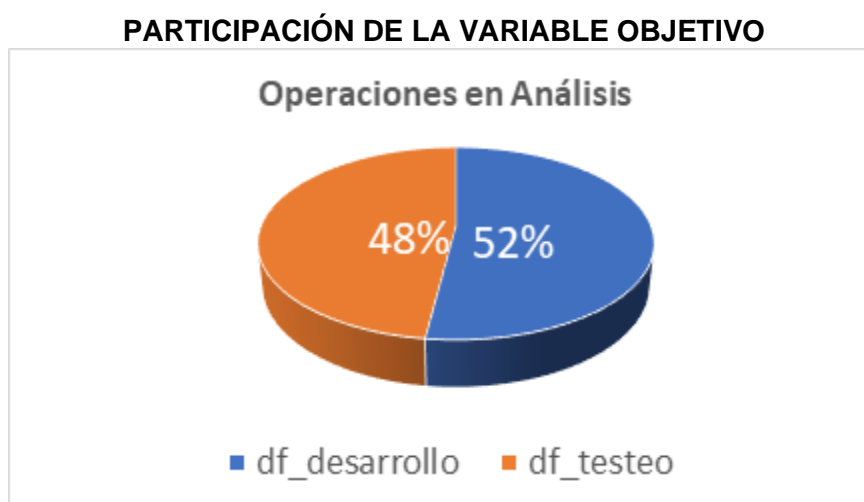
Fuente: BDD - Riesgos
Elaboración: Autor

Del total de operaciones otorgadas entre junio del 2016 hasta diciembre 2019, se cuenta con un total de registros 161387 de los cuales el 52% pertenece a la muestra de desarrollo donde el promedio del índice de cartera vencida es del 29% y el 48% de operaciones corresponden a la muestra de testeo con un índice de cartera vencida del 26%.

Este índice de cartera vencida es menor en la muestra de testeo, precisamente por tratarse de operaciones más recientes. Estas son las

operaciones a las que se les debe calcular el score construido con la data 2016 – 2018 y ver su desempeño 12 meses después del otorgamiento.

En el caso del segmento de micro finanzas y siendo BanEcuador un Banco de Desarrollo, las operaciones cuentan con plazos más amplios que microcréditos con similares montos, por lo que el corte o estado de la operación debe mirarse mucho más tarde que los 12 meses planteados en las metodologías de credit score.



Fuente: BDD - Riesgos
Elaboración: Autor

En la definición de malos clientes es necesario hacer algunas modificaciones y se debe incluir como clientes malos a todos aquellos clientes que tengan operaciones reestructuradas, con demanda judicial, castigados o con intereses condonados. En el caso de créditos castigados y con demanda judicial, internamente BanEcuador registra este estado y estos clientes, por lo que se debe añadir los clientes que presenten estas características cuya información proviene del buró, de esta manera, la proporción de malos clientes que se encuentra en el archivo para la elaboración del modelo pasa de un 29.7% a un 42.7%, con lo que se tiene una muestra más verídica y balanceada para la construcción del modelo estadístico.

3.5. Análisis de las variables independientes

En esta sección se realiza una verificación de la integridad de la información contenida en la base de datos, realizar un control de calidad de la base de datos, mediante un análisis univariante de todas las variables con las que se cuenta e identificar posibles errores en la base y de ser factible corregirlos, además de conocer el estado de las variables en la base de datos, es decir, identificar características de los datos.

El análisis univariante debe diferenciar las variables nominales de las continuas, para las primeras bastará con un estudio de frecuencias y para las segundas una identificación de máximos, mínimos y promedios; esto permitirá tener una primera impresión de la calidad de la base de datos.

- **Distribución de frecuencias:** La distribución de frecuencias indica el número de casos que hay en cada categoría de la variable. A partir de dichos valores, en una tabla de frecuencias, se calcula el porcentaje (respecto del total de observaciones), porcentaje válido (excluido los valores perdidos) y el porcentaje acumulado (porcentaje de la primera categoría, luego de este más el de la segunda categoría y así sucesivamente). Se aplica para variables nominales, ordinales
Además de la tabla de frecuencias también es posible hacer representaciones gráficas tales como: gráfico de barras, gráfico circular, o un polígono de frecuencias.
- **Medidas de tendencia central:** Las medidas de tenencia central dan cuenta del tipo de distribución que tiene la variable respecto de un valor típico, o puntuación central, a partir del cual se agrupan.
La media o promedio corresponde a la suma de todas las puntuaciones de la variable dividida por el número total de casos. La mediana es el valor que divide por la mitad a las puntuaciones de la variable: los que están por debajo de éste y los que están por encima. Es decir, es el valor que divide en dos mitades a las observaciones. La moda es el valor que más se repite del conjunto de observaciones, pudiendo haber más de una moda (bimodal o multimodal).

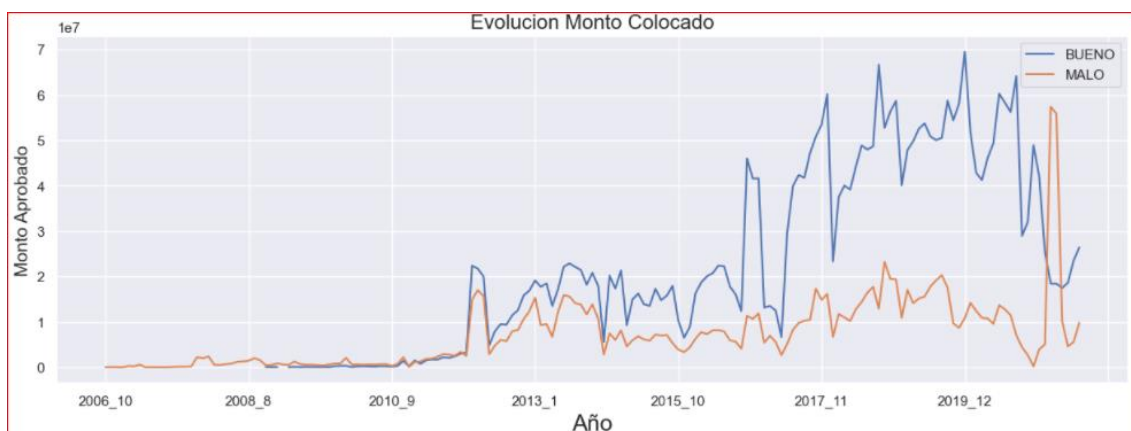
- **Medidas de dispersión:** Las medidas de dispersión indican el grado de variabilidad de los datos respecto de la media (promedio). La varianza es el valor promedio del cuadrado de las puntuaciones respecto de la media. La desviación estándar es el promedio de desviación de los valores de las observaciones respecto a la media, expresada en los valores originales de la medición de la variable.

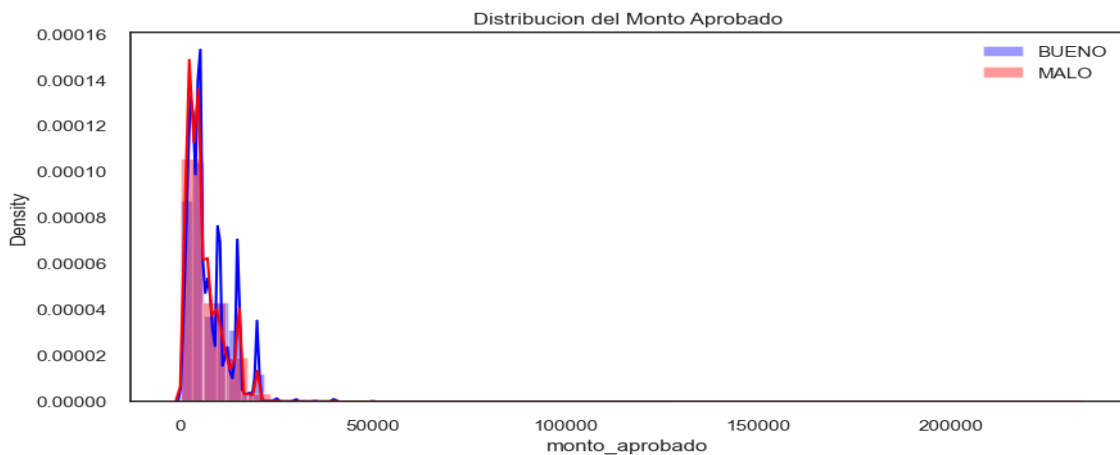
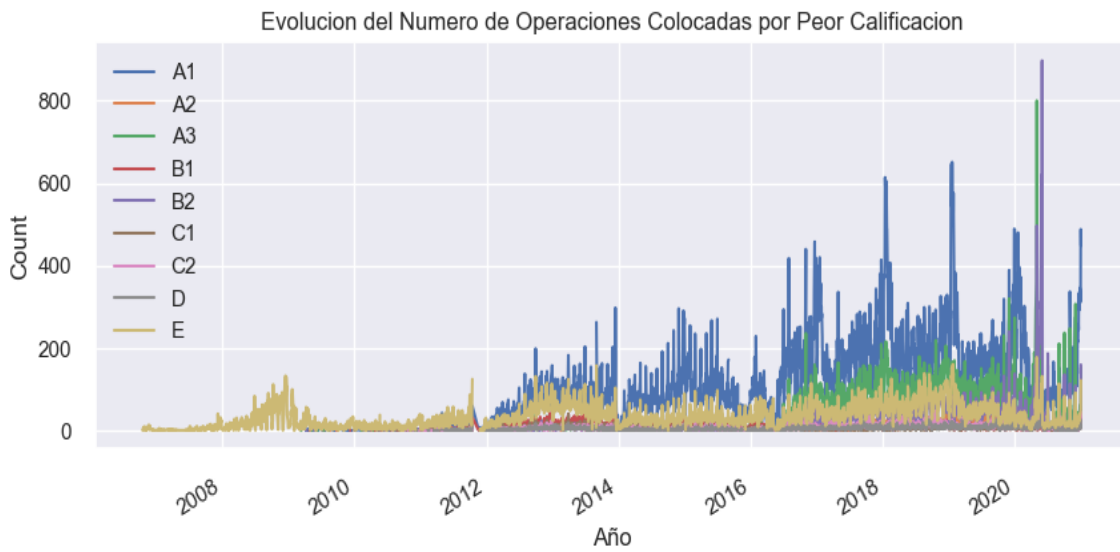
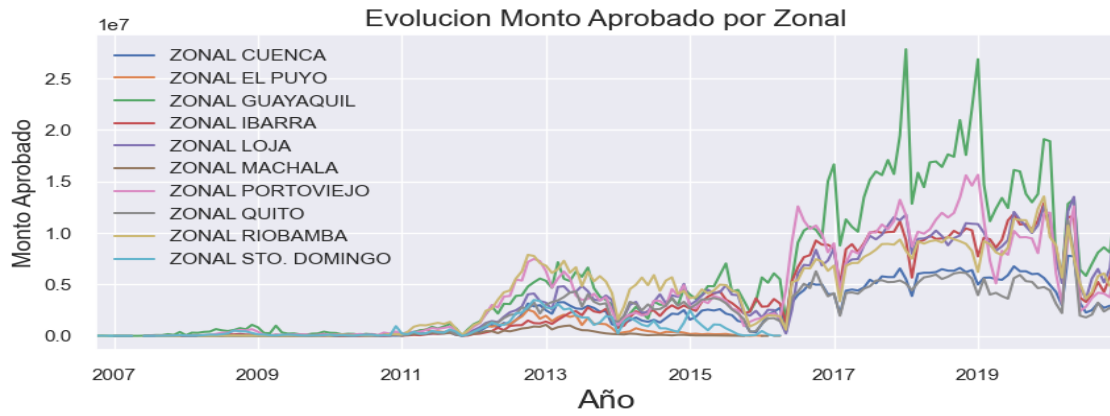
En esta etapa de la construcción del modelo score se concentra el mayor esfuerzo por cuanto se debe realizar una verificación de la integridad de la información contenida en las bases de datos, análisis preliminar, que incluye transformación de variables y tratamiento de datos nulos. Se debe decidir sobre esquemas de segmentación y una selección preliminar de variables.

3.5.1. Análisis Exploratorio de la Data

Consiste en la realización de estadísticas univariadas y multivariadas para cada variable de interés, con el fin de evaluar el número de aplicaciones, número de defaults, montos otorgados, calificaciones actuales, sus tendencias, volatilidad y demás información que pueda ayudar a entender el estado actual de la originación de crédito y así validar hipótesis acerca del comportamiento de los clientes al momento de la originación del crédito, ésta evaluación preliminar tiene como objetivo ver la relación entre montos otorgados y riesgo actual existente sin modelo de score y luego evaluar como cambiaría esta situación ex post con modelo.

EVOLUCIÓN DE LOS MONTOS OTORGADOS POR BUEN Y MAL CLIENTE





Fuente: BDD - Riesgos
Elaboración: Autor

De este grupo de gráficos se puede deducir que la colocación de BanEcuador se acelera o crece a partir del 2016 debido a que antes de ese año

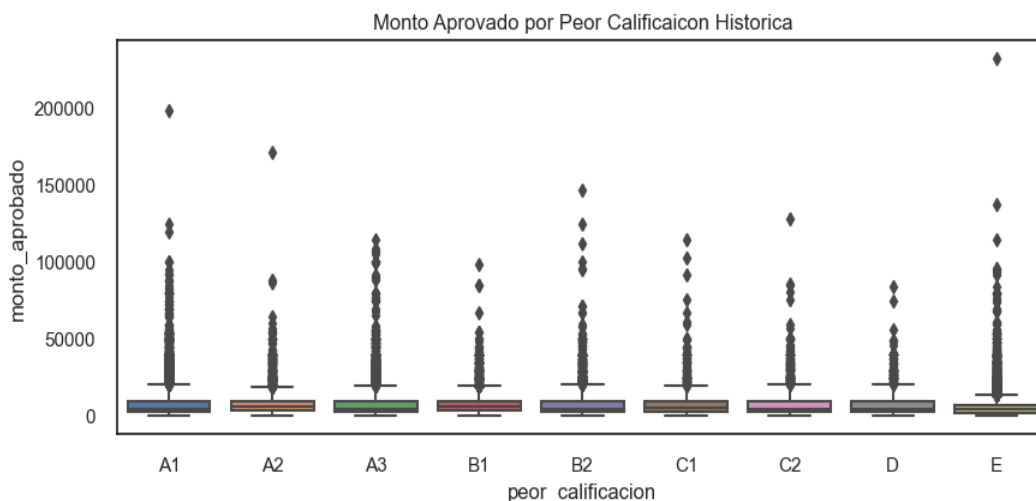
era Banco Nacional de Fomento. A partir de este año, se cuenta con más capital, una mejor tecnología crediticia, y una mayor presencia nacional.

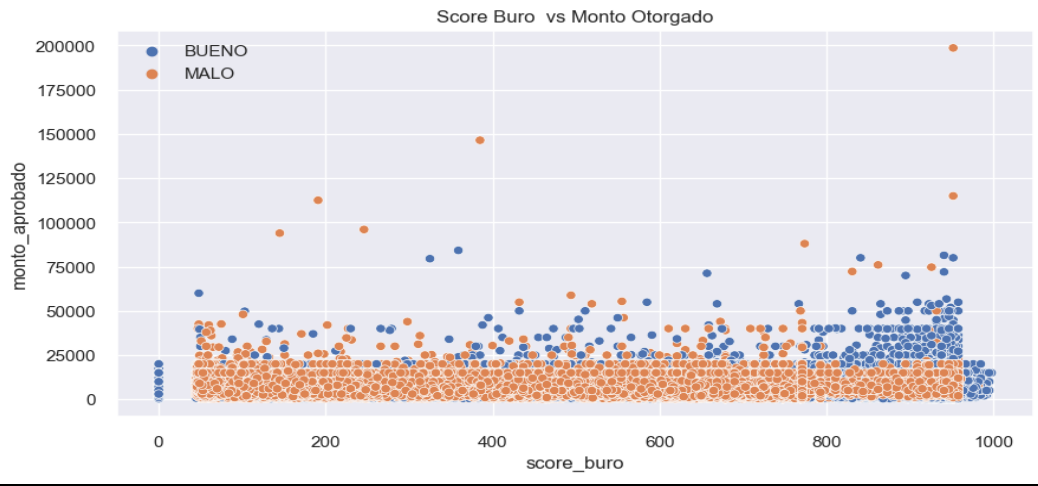
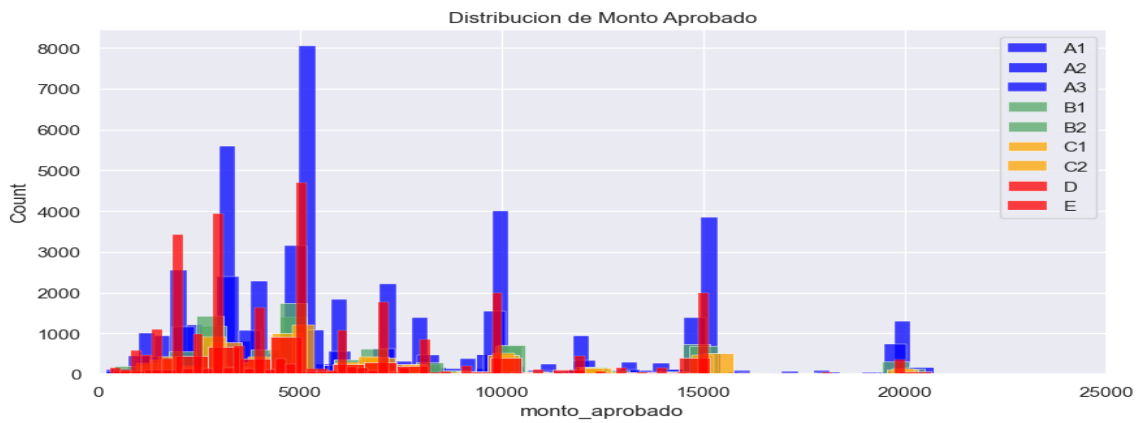
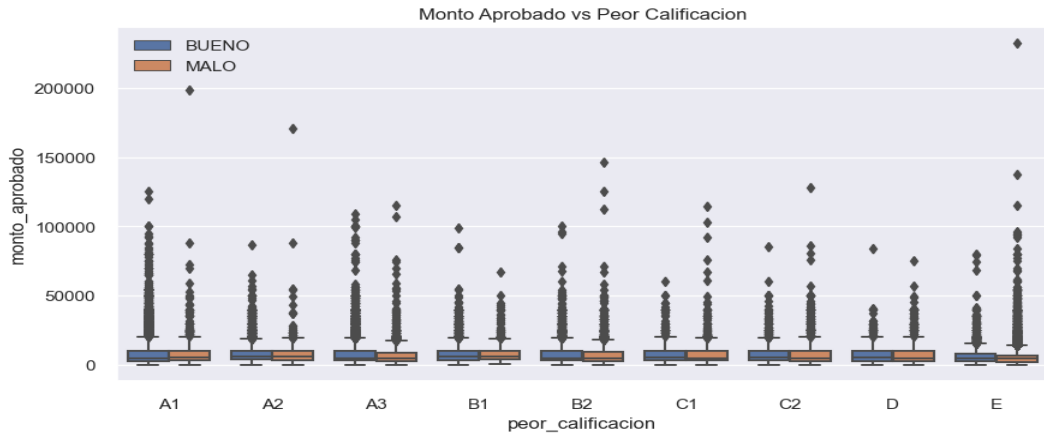
Se aprecia también que la originación del crédito se deterioró a partir del 2020, mayor colocación en monto a clientes malos que a clientes buenos, las zonas de Guayaquil y Portoviejo son las que demandan montos más altos y las que deberían presentar un mayor riesgo.

El Monto otorgado no difiere del monto solicitado y que en promedio bordea los 25 mil dólares para este segmento. Es importante aclarar la gran dispersión o desviación estándar de la media que presentan los créditos, debido a como se aprecia en el grafico existen montos otorgados cercanos a los 200 mil dólares, lo cual deja una duda de si se trata completamente de una cartera de microcrédito y de la calidad de la data.

También se evidencia que casi la totalidad de los créditos se originan con una calificación A1, lo cual muestra un ligero desajuste en la originación debido a que la calidad de cartera micro en la actualidad no es así de buena, no muestra esa mora que se ajuste a la calificación de A1.

MONTO VS PEOR CALIFICACIÓN





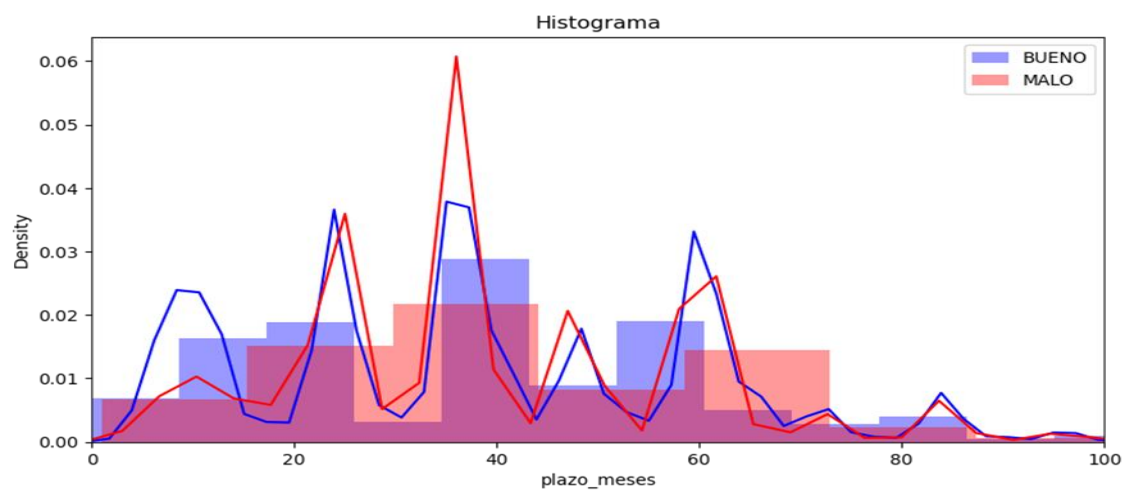
Fuente: BDD - Riesgos
 Elaboración: Autor

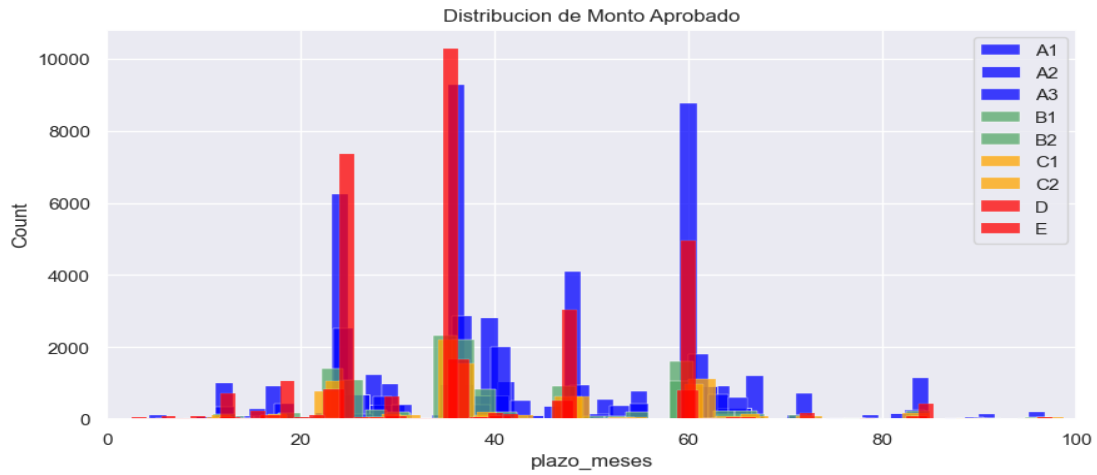
Como se aprecia en los gráficos de este bloque, la presencia de clientes malos no es exclusiva de las categorías D o E, ni específica de montos grandes, en los gráficos de cajas los clientes malos se distribuyen de similar manera acorde a la calificación, mientras que en el gráfico de distribución los clientes malos se acumulan en montos menores a los 10 mil dólares, pero esto se debe a que hay mayor número de créditos.

Con respecto al monto otorgado versus el score se evidencia que son totalmente independientes, es decir; a mayor score o calificación, no implica un mayor monto otorgado, los clientes malos se distribuyen de manera uniforme desde puntajes pequeños hasta puntajes grandes y montos de hasta 25 mil USD.

Los modelos estadísticos son de carácter predictivo, capturan comportamientos pasados y predicen el futuro en función de las relaciones estadísticas o tendencias de data histórica, por lo que usar los modelos de score para otorgar más monto en función de un puntaje no es recomendable.

PLAZO DEL CRÉDITO VS. CALIFICACIÓN



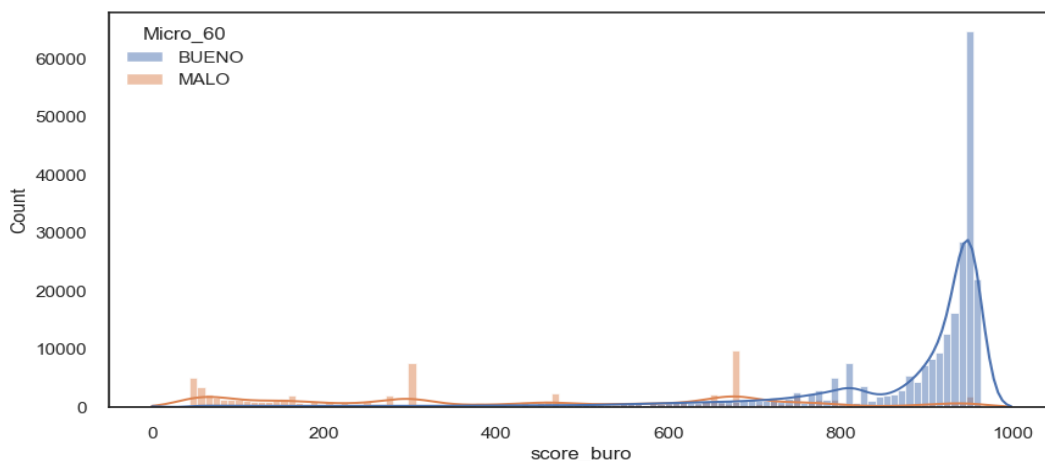


Fuente: BDD - Riesgos
Elaboración: Autor

Con respecto al monto, plazo y calificación, se ve que la distribución de malos clientes y la distribución de clientes buenos es la misma indistintamente del plazo, por lo tanto, el plazo no es un determinante del crédito.

Finalmente, se muestra que el score actual vigente otorgado a los clientes de BanEcuador califica con puntaje más alto a los clientes buenos (mora < 60 días) y los clientes considerados malos se acumulan en scores bajos con una ligera presencia en puntajes altos. Por lo que se espera igual comportamiento para el score de originación socio demográfico.

SCORE EQUIFAX



Fuente: BDD - Riesgos
Elaboración: Autor

3.5.2. Análisis Uni-variado

La estrategia de análisis consiste en separar las variables continuas de las variables categóricas. Para las variables continuas se ejecutan pruebas de correlación (Pearson, Spearman y Kendall) con el objetivo de evaluar si los valores de una variable se mueven o se comportan de manera lineal con respecto a los valores de las otras variables y por lo tanto una correlación alta implica un aporte de información redundante al modelo estadístico (multi colinealidad).

El primer problema que aparece cuando existe multi colinealidad en las variables usadas para la construcción del modelo es que los coeficientes se vuelven muy sensibles a pequeños cambios. Si se modifica un parámetro, los de las variables relacionadas también se verán afectados, es decir; modelos ligeramente diferentes ofrecen conclusiones completamente diferentes.

Otro problema es una reducción de la precisión de los coeficientes, lo que debilita el modelo. Provocando que los p-valores no sean estadísticamente significativos. Esto se puede ver a la hora de ajustar un modelo. En estos casos lo más probable es que los valores de los coeficientes cambian de forma significativa, llegando incluso a cambiar de signo, a incluir o algunas de las variables independientes en el modelo. Además de esto, se dificulta confiar en los p-valores para seleccionar aquellas variables que se deben incluir en el modelo. Haciendo complicado tanto la definición como la posterior justificación de este.

El primer criterio que se usa para eliminar variables, es un criterio experto, la mayoría de las variables continuas presentes en el archivo son variables de comportamiento (variables financieras relacionadas a saldos, por lo que se eliminan) y no sociodemográficas como se requiere, ya que el objetivo es construir un score de originación; por ejemplo: Saldo por Vencer, Saldo No Devenga Intereses, Saldo Vencido o Número de cuotas pagadas, número de cuotas por pagar, adicionalmente se generan Ratios Financieros como: Activos/ Pasivos e Ingresos / Gastos, por lo que las variables originales de Activos, Pasivos, Ingresos y Gastos se eliminan.

Las variables de atraso promedio y atraso máximo se eliminan, debido a que con estas se construye la variable independiente o target de Buenos y Malos clientes, así como las variables de demanda judicial y saldo de cartera castigada, ya que estas se consideran como clientes malos y están por definición incluidas en la variable target.

Las variables plazo días, plazo meses y periodos plazo evidentemente están correlacionadas y se elimina plazo días y plazo meses, se mantiene periodos plazo, al igual que meses residencia y antigüedad de residencia y de igual manera para número de cargas y número de cargas estudiando, manteniendo las variables más pobladas.

Variabes como Mayor Plazo Vencido y Mayor Valor Vencido, para los periodos: actual, hace 6 meses e histórico presentan una alta correlación entre sí, por lo que tanto para las variables plazo como valor se escoge las variables históricas, además como se mencionó anteriormente, las variables Valor y Saldo, Montos son consideradas de comportamiento por lo que se eliminan.

Correlaciones y colinealidad para variables continuas

Al estimar un modelo de regresión, se busca la relación entre cada una de las variables independientes con la variable dependiente o target de lo que se considera buen y mal cliente para BanEcuador. Cada uno de los coeficientes del modelo de regresión indica el cambio que se espera observar en la variable dependiente al modificar una unidad en la variable independiente, asumiendo que se mantienen constantes todas las demás variables independientes. Pero, si dos o más variables independiente están relacionadas esto no se observa en los datos.

Se Calcula la correlación entre las variables regresoras, para lo cual se debe tomar en cuenta la naturaleza de la variable (continua o discreta), primero se lo hace para el caso continuo, con lo que se elimina variables con altas correlaciones entre si (de acuerdo con test de hipótesis), ya que pueden generar multi colinealidad y/o endogenidad.

Se considera variables colineales para valores reportados superiores a 0.5, se realizan las tres pruebas mencionadas y se comparan los valores

reportados para evaluar sensibilidades en cada prueba, ya que una prueba es más estricta que otra, por lo que se observa cambios con cada una de ellas y de esta manera se seleccionan variables que se encuentran con valores de correlación en la frontera de 0.5.

A medida que aumenta la relación entre las variables, el grado de multicolinealidad es mayor, se hace más complicado observar cambios en una variable sin que se modifique la otra. Lo que hace difícil para el modelo estimar la relación que existe entre cada una de las variables independiente y la dependiente ya que algunas de las independientes cambian a la vez.

Selección de las variables continuas

La metodología para elegir las variables que se usarán en el modelo estadístico, es la de RANDOM FOREST, esta difiere de otras similares o frecuentemente usadas como son los árboles de decisión, ya que evitan el sobre ajuste (sobre ajuste); es decir, llegan a un punto donde el algoritmo se estanca o empieza a reproducir los mismos patrones de clusterización, en términos técnicos el algoritmo empieza a “memorizar”. Para evitar esto el algoritmo realiza lo siguiente:

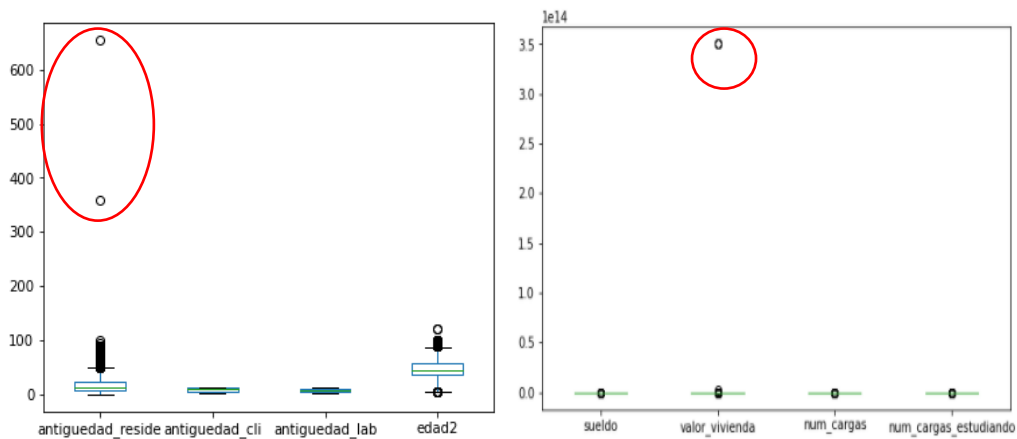
- Seleccionar k columnas de las m totales (siendo k menor a m) y crea un árbol de decisión con esas k características.
- Crea n árboles, variando siempre la cantidad de k columnas y también varía la cantidad de muestras que pasan a esos árboles (“bootstrap sample”)
- Se toma cada uno de los n árboles y se hace una misma clasificación, se guarda el resultado de cada árbol y se obtiene n salidas.
- Se calculan los “votos” obtenidos para cada “clase” seleccionada y se considera a la más votada como la clasificación final del “bosque”.
- Se garantiza la aleatoriedad en la elección de K y la elección de n .

Corrección missings y outliers

Para la corrección de datos atípicos existen varias maneras, usualmente relacionadas con el diagrama de caja, que consiste en hallar datos fuera de los

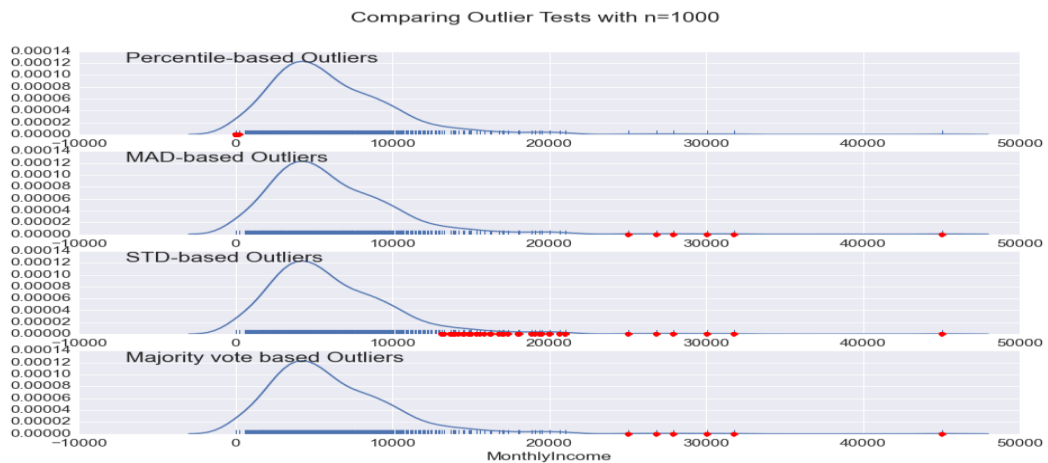
quintiles de cada variable, estos se identifican y se eliminan o se imputan estos valores con el valor máximo reportado.

Para remover datos atípicos se usa el rango inter-cuartílico que es una medida de dispersión de un conjunto de datos que expresa la distancia entre el primer y el tercer cuartil. En otras palabras, el rango inter-cuartílico es la diferencia entre el penúltimo y el primer cuartil de una distribución utilizado en el diagrama de caja.



Fuente: BDD - Riesgos
Elaboración: Autor

El rango Inter cuartil es considerado un estadístico robusto por su baja exposición a valores extremos. Esto es debido a que solo se consideran las observaciones entre el tercer cuartil y el primer cuartil. Todas las observaciones fuera de ese rango quedan excluidas del cálculo y, por tanto, solo se tienen en cuenta las observaciones más cercanas a la mediana.



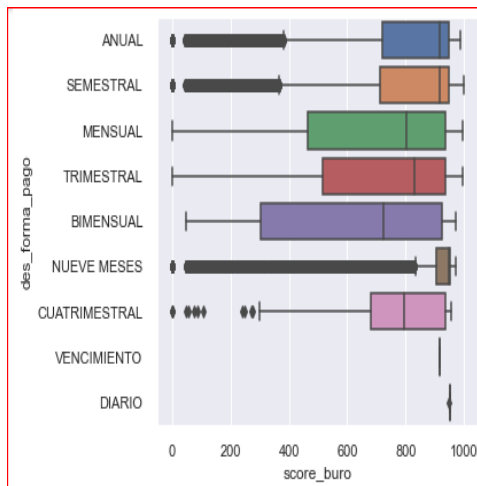
3.5.3. Análisis Bi - variado

Variable objetivo vs. Variables categóricas

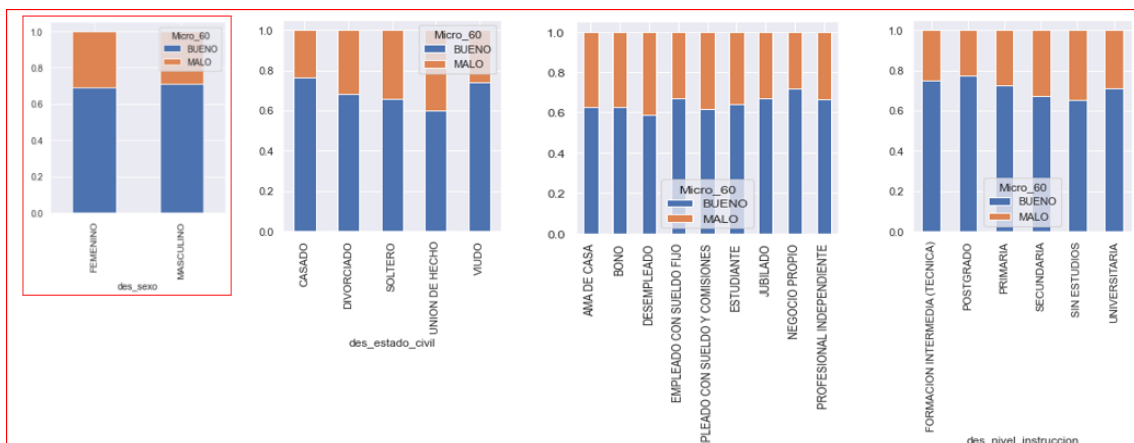
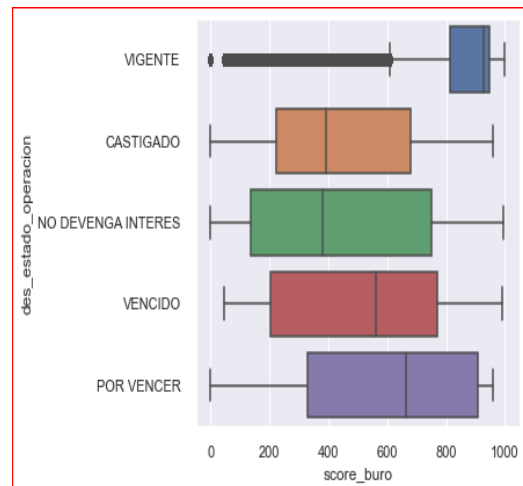
Se pueden usar variables continuas o discretas, aunque se recomienda categorizar las variables continuas debido a que es más fácil tratar los datos atípicos (outliers), lo que implica contar con más conocimiento de la cartera, se obtiene un mayor control en el desarrollo del modelo, ya que formando grupos permite al analista entender el comportamiento del riesgo y desarrollar mejores estrategias.

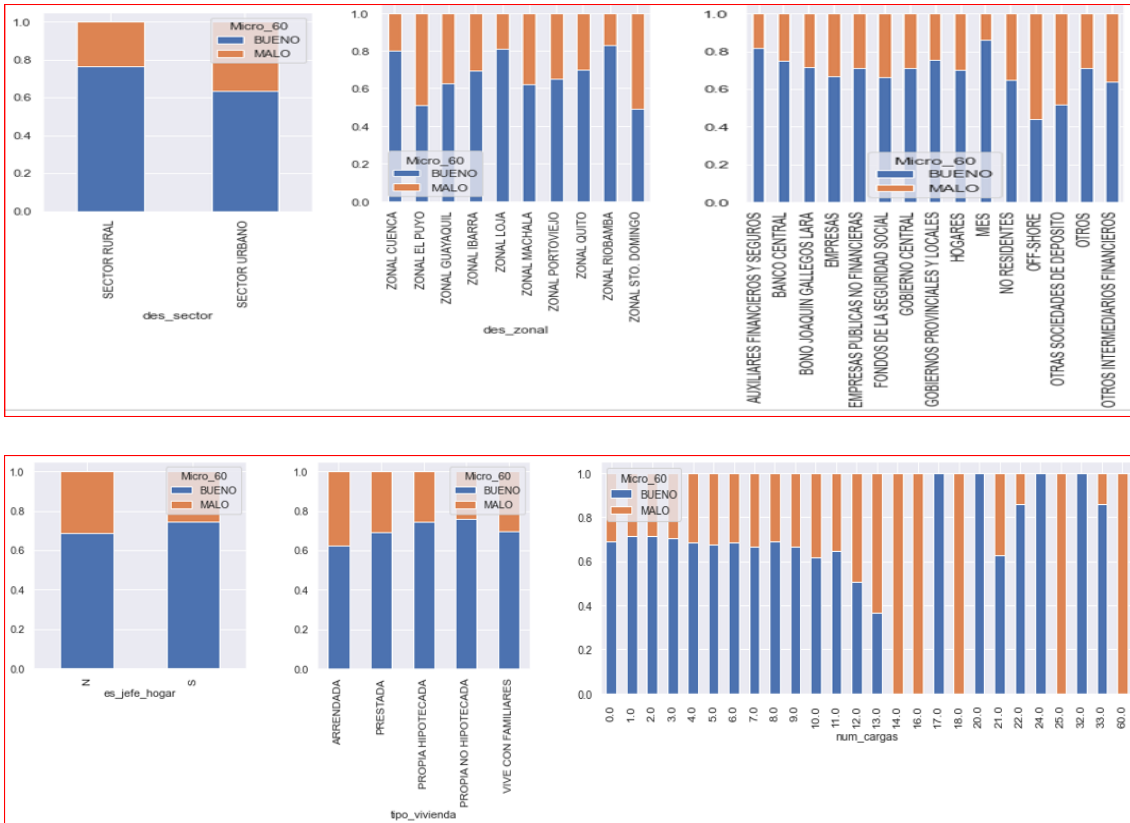
El estado actual de la originación no solo es evolutivo sino también de corte transversal y también incluye análisis categórico que se resume a continuación:

FORMA DE PAGO DEL CRÉDITO VS SCORE BURÓ



ESTADO DEL CRÉDITO VS SCORE BURÓ

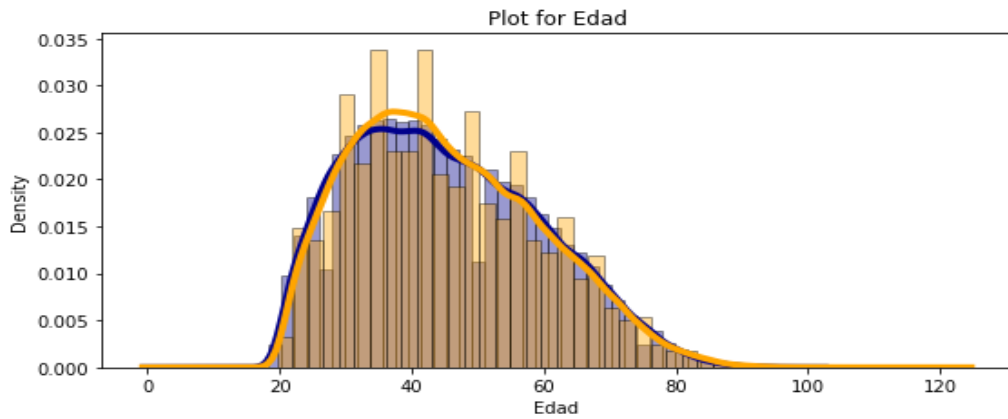




En esta etapa del desarrollo es importante identificar categorías que se pueden agrupar, como aquellas que tengan pocos registros o poca presencia de clientes malos, que no son relevantes para el modelo final, esto se lo hace de tal manera que se encuentre un orden ascendente o descendente de clientes malos con el fin de encontrar algún patrón que permita predecir o diferenciar la variable objetivo, debido a que se considera que cada categoría es un potencial coeficiente de la ecuación final y por temas de implementación no es recomendable tener demasiados coeficientes, como por ejemplo la variable estado civil.

Variable objetivo vs. Variables continuas

Se procede a discretizar las Variables continuas, para lo cual se hace uso de los percentiles de cada variable, adicionalmente se usa la metodología de Arboles de Decisión para comparar el comportamiento de los clientes Buenos y Malos en cada corte de la variable continua, se busca patrones ascendentes o descendentes acorde al concepto económico de cada variable continúa analizada, el objetivo es que los clientes buenos crezcan en número y los clientes malos se reduzcan en número acorde a cada clúster realizado



```

count      426508.000000
mean       45.194416
std        14.208642
min         3.000000
25%        34.000000
50%        44.000000
75%        55.000000
max        121.000000
Name: Edad, dtype: float64

```

```

1 def edad_B(row):
2     if row['Edad'] <= 35:
3         return '_35'
4     elif (row['Edad'] >35) & (row['Edad'] <= 45):
5         return '_35_45'
6     elif (row['Edad'] >45) & (row['Edad'] <= 55):
7         return '_45_55'
8     elif (row['Edad'] >55) & (row['Edad'] <= 65):
9         return '_55_65'
10    elif (row['Edad'] > 65) :
11        return '_65'
12    else:
13        return 0
14
15 df_cliente['edad_B'] = df_cliente.apply(edad_B, axis=1)
16 pd.crosstab(df_cliente['edad_B'],df_cliente['Micro_60'])

```

El objetivo primordial es encontrar segmentos o rangos de la variable en los que el número de clientes malos crezca o decrezca en dirección contraria al número de clientes buenos, con el objetivo de que cada variable discrimine o separe clientes (buenos de malos), por lo que se puede inferir que esta variable no es buena para predecir o discriminar clientes, por lo que no será usada como regresor en el modelo estadístico.

	Micro_60	BUENO	MALO
edad_B			
_35	72348	52079	
_35_45	60573	48155	
_45_55	50393	37591	
_55_65	36740	26655	
_65	24363	1761	

Se debe determinar el poder predictivo de cada característica individualmente respecto al default (Bueno/Malo), para esto es necesario excluir de la data aquellas variables con poder predictivo débil y con poco sentido lógico y agrupar o ranquear las variables de acuerdo con su poder predictivo.

3.6. Poder predictivo de los regresores (WOE)

El poder predictivo de una característica se mide de acuerdo con 4 criterios:

1. El poder predictivo de cada categoría de una variable, para esto se usa el peso de la evidencia (WOE).

$$WOE = \ln\left(\frac{\text{Event\%}}{\text{Non Event\%}}\right)$$

2. El poder predictivo de la variable, para esto se usa la métrica “Information Value” (IV), permite escoger las variables más importantes en un modelo predictivo.

$$IV = \sum (\text{Event\%} - \text{Non Event\%}) * \ln\left(\frac{\text{Event\%}}{\text{Non Event\%}}\right)$$

3. El rango y tendencia del WOE dentro de cada característica y Consideraciones operacionales y de negocio.

Information Value	Poder Predictivo
<0.02	No es Útil para predecir
0.02 – 0.1	Débil Poder Predictivo
0.1 – 0.3	Poder predictivo Medio
0.3 – 0.5	Fuerte Poder predictivo
>0.5	Posible Colinealidad

Esto permite realizar estrictas transformaciones lineales de las variables continuas que penalizan la presencia de malos clientes en forma decreciente. A continuación, se lista el WoE calculado de Cada Variable que participa en la Ecuación estimada Final.

variable	IV Score:
mpv_histot_B__3	0.61
des_op_sector_bnf_AGRO_GANADERO	0.24
des_forma_pago_MENSUAL	0.24
cartera_heredada_SI	0.13
des_sector_SECTOR_URBANO	0.11
des_op_sector_bnf_COMERCIO	0.1
numperiodos_B__5	0.1
numperiodos_B__35	0.09
des_op_sector_bnf_OTROS	0.08
mpv_acttot_B__1	0.07
mvv_acttot_B__1	0.07
des_forma_pago_ANUAL	0.07
cuotestburo_B__122	0.07
des_forma_pago_SEMESTRAL	0.06
des_zonal_ZONAL_PORTOVIEJO	0.05
des_zonal_ZONAL_RIOBAMBA	0.05
act_pas_quintil__472	0.05
endeudprom_B__0	0.05
des_estado_civil_SOL_VIU_DIVO	0.04
des_zonal_ZONAL_GUAYAQUIL	0.04
des_zonal_ZONAL_LOJA	0.04
des_forma_pago_NUEVE_MESES	0.04
tamano_ventas__6500	0.04
cuotestburo_B__712	0.04
numperiodos_B__12_35	0.04
plazomeses_B__24_36	0.04
ing_gas_quintil__1449	0.03
des_op_sector_bnf_SERVICIOS	0.03
sueldo_B__400	0.03
numperiodos_B__6_12	0.03
plazomeses_B__60	0.03
Ciclo_B__2	0.02

cliente_nuevo_SI	0.02
des_zonal_ZONAL_CUENCA	0.02
monto_B__3000	0.02
cuotestburo_B__334_712	0.02
endeudprom_B__0_1264	0.02
endeudprom_B__1264	0.02
es_jefe_hogar_S	0.01
acreedores_vigentes_B__3	0.01
des_op_sector_bnf_ACUICOLA	0.01
des_op_sector_bnf_INDUSTRIAL	0.01
tipo_vivienda_ARRENDADA	0.01
monto_B__10000	0.01
sueldo_B__400_800	0.01
sueldo_B__800_1250	0.01
tamano_ventas__1161	0.01
tamano_ventas__1161_2500	0.01
act_pas_quintil__11813	0.01
act_pas_quintil__1444_11813	0.01
des_nivel_instruccion_BASICA	0.01
des_nivel_instruccion_MEDIA	0.01
dessexo_MASCULINO	0
numcargas_est_B__3	0
des_op_sector_bnf_ARTESANAL	0
des_op_sector_bnf_TURISMO	0
des_zonal_ZONAL_IBARRA	0
des_zonal_ZONAL_QUITO	0
des_forma_pago_OTROS	0
des_forma_pago_TRIMESTRAL	0
tipo_vivienda_PRESTADA	0
tipo_vivienda_PROPIA_HIPOTECADA	0
tipo_vivienda_PROPIA_NO_HIPOTECADA	0
tipo_vivienda_VIVE_CON_FAMILIARES	0
edad_B__35	0

edad_B__35_45	0
edad_B__45_55	0
edad_B__55_65	0
edad_B__65	0
monto_B__3000_5000	0
monto_B__5000_10000	0
sueldo_B__1250	0
tamano_ventas__2500_6500	0
act_pas_quintil__472_1444	0
cuotestburo_B__122_334	0
plazomeses_B__24	0
plazomeses_B__36_60	0
des_nivel_instruccion_SUPERIOR	0

Fuente: BDD - Riesgos
Elaboración: Autor

3.6.1. Test de independencia CHI2 para variables categóricas

Chi2, es una métrica de las diferencias existentes entre dos variables, en caso de haber diferencias significativas, estas se deben al azar en el contraste de hipótesis. Es usada para determinar cuándo hay una diferencia estadísticamente significativa entre las frecuencias esperadas y observadas de una o más categorías en una tabla de contingencia.

La prueba CHI2, se considera de bondad de ajuste, generalmente usada para probar independencia entre variables discretas vs la variable objetivo, estas evalúan la discrepancia entre distribuciones, la observada y otra esperada o teórica.

Para lo cual se construyen las tablas de contingencias en dos versiones, la primera tabla es lo observado en los datos y la otra tabla es lo esperado construida a partir de la primera con la siguiente fórmula.

$$\chi^2 = \sum_i \frac{(\text{observada}_i - \text{teórica}_i)^2}{\text{teórica}_i}$$

Los resultados de este test para ver que variables son independientes son:

	Column			Hypothesis
0	des_op_sector_bnf	Reject	Null	Hypothesis
1	des_zonal	Reject	Null	Hypothesis
2	des_situacion_laboral	Reject	Null	Hypothesis
3	des_origen_ingresos	Reject	Null	Hypothesis
4	des_forma_pago	Reject	Null	Hypothesis
5	des_nivel_instruccion	Reject	Null	Hypothesis
6	tipo_vivienda	Reject	Null	Hypothesis
7	des_estado_civil	Reject	Null	Hypothesis
8	ca_tip_garantia_categ	Reject	Null	Hypothesis
9	des_sector	Reject	Null	Hypothesis
10	es_jefe_hogar	Reject	Null	Hypothesis
11	dessexo	Reject	Null	Hypothesis
12	cliente_nuevo	Reject	Null	Hypothesis
13	cliente_bancarizado	Reject	Null	Hypothesis
14	analisis_riesgo	Reject	Null	Hypothesis
15	cartera_heredada	Reject	Null	Hypothesis
16	indicador_vencido	Reject	Null	Hypothesis
17	bandera_reestructurado	Reject	Null	Hypothesis
18	bandera_Demanda_Judicial	Reject	Null	Hypothesis
19	bandera_castigado	Reject	Null	Hypothesis
20	bandera_condonado	Reject	Null	Hypothesis

Fuente: BDD - Riesgos
Elaboración: Autor

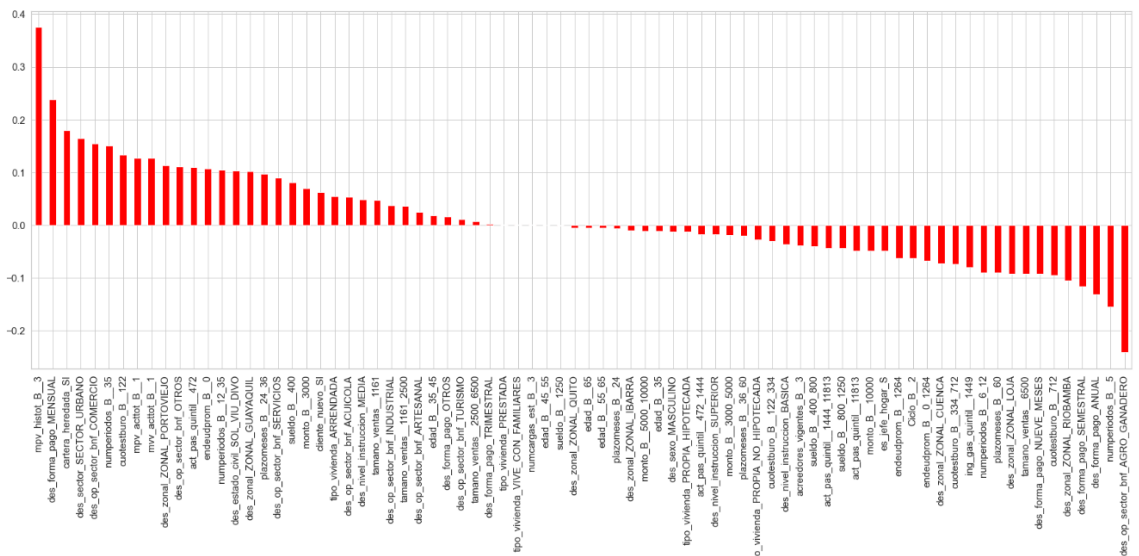
Estos resultados indican que las variables listadas, son independientes de la variable objetivo y pueden ser retiradas del modelo de score, pero se analiza los posibles falsos positivos por categoría. El texto “fail to reject the Null Hypothesis” indica que no hay una diferencia estadísticamente significativa entre la variable categórica y la variable objetivo de BUENOS y MALOS clientes, en base al test de independencia CHI2, por lo que se remueven de los análisis variables como Origen de ingresos, Nivel de instrucción, situación laboral, tipo de vivienda etc.

	Category	Hypothesis
43	des_nivel_instruccion-POSTGRADO	Fail to Reject Null Hypothesis
52	tipo_vivienda-VIVE CON FAMILIARES	Fail to Reject Null Hypothesis
41	des_forma_pago-VENCIMIENTO	Fail to Reject Null Hypothesis
40	des_forma_pago-TRIMESTRAL	Fail to Reject Null Hypothesis
36	des_forma_pago-CUATRIMESTRAL	Fail to Reject Null Hypothesis
33	des_origen_ingresos-RENTISTA	Fail to Reject Null Hypothesis
32	des_origen_ingresos-REMESAS DEL EXTERIOR	Fail to Reject Null Hypothesis
31	des_origen_ingresos-MISION CASA PARA TODOS	Fail to Reject Null Hypothesis
60	ca_tip_garantia_categ-OTRAS GARANTIAS ADECUADAS	Fail to Reject Null Hypothesis
25	des_situacion_laboral-PROFESIONAL INDEPENDIENTE	Fail to Reject Null Hypothesis
49	tipo_vivienda-PRESTADA	Fail to Reject Null Hypothesis
20	des_situacion_laboral-EMPLEADO CON SUELDO FIJO	Fail to Reject Null Hypothesis
23	des_situacion_laboral-JUBILADO	Fail to Reject Null Hypothesis
6	des_op_sector_bnf-MINERIA	Fail to Reject Null Hypothesis
73	analisis_riesgo-NO	Reject Null Hypothesis
80	bandera_reestructurado-SI	Reject Null Hypothesis
58	ca_tip_garantia_categ-GARANTIA PERSONAL	Reject Null Hypothesis
57	des_estado_civil-VIUDO	Reject Null Hypothesis
81	bandera_Demanda_Judicial-NO	Reject Null Hypothesis
56	des_estado_civil-UNION DE HECHO	Reject Null Hypothesis
79	bandera_reestructurado-NO	Reject Null Hypothesis
55	des_estado_civil-SOLTERO	Reject Null Hypothesis

Fuente: BDD - Riesgos
Elaboración: Autor

3.7. Construcción del Modelo Estadístico

IMPORTANCIA DE LAS VARIABLES



3.7.1. Construcción del modelo de regresión logística

Debido a que se tiene dos posibles situaciones, la primera es que el cliente paga (0) y la segunda que el cliente no paga (1), se aplicará una metodología paramétrica, específicamente un modelo de Regresión Logística Binaria y así evitar los inconvenientes que presentan los modelos de Regresión Lineal o de Análisis Discriminante, se elige la regresión logística binaria como técnica empleada debido a las ventajas siguientes:

Las propiedades estadísticas son más adecuadas que las de los modelos lineales en los que, en ocasiones, se obtienen estimadores ineficientes.

Dadas las características del historial crediticio de los clientes de BanEcuador, donde la información cualitativa complementa la escasez de variables cuantitativas, la Regresión Logística Binaria admite variables categóricas con mayor flexibilidad que los modelos lineales. Permite estimar la probabilidad de impago del crédito según los valores de las variables independientes.

Determina la influencia de cada variable independiente sobre la variable dependiente (pago o impago) según el OR (Odd Ratio o ventaja). Éste se define como $\exp(\beta)$, donde \exp es la base de los logaritmos neperianos (una constante cuyo valor es 2,718), y β es el valor del parámetro de regresión de la variable independiente en el modelo.

Una $OR > 1$ indica un aumento en la probabilidad del evento de incumplimiento sobre el hecho de pagar cuando la variable explicativa aumenta en una unidad; inversamente, una $OR < 1$ indica lo contrario. El modelo de regresión logística puede formularse como:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

Donde p es la probabilidad de ocurrencia del evento de interés, en nuestro caso, impago o default. Debido al valor de las variables independientes, la probabilidad señalada puede ser calculada directamente de la siguiente forma:

$$p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$$

Con Z igual a:

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

Para el cálculo del score de otorgamiento, se deben calcular variables binarias (vectores) por cada variable del modelo, el número de vectores creados para cada variable será igual a n categorías existentes, estas categorías se crearon de acuerdo con el comportamiento y tasa de clientes malos presentes en cada una de las variables usadas en el desarrollo del modelo estadístico.

Por ejemplo:

ESTADO CIVIL		VARIABLES DUMMIES				
Nro.	CATEGORIAS	Vector 1	Vector 2	Vector 3	Vector 4	Vector 5
1	SOLTERO	1	0	0	0	0
2	CASADO	0	1	0	0	0
3	DIVORCIADO	0	0	1	0	0
4	VIUDO	0	0	0	1	0
5	UNION DE HECHO	0	0	0	0	1

Para el caso de variables binarias, se tiene:

DES_SEXO		VARIABLES DUMMIES	
Nro.	CATEGORIAS	Vector 1	Vector 2
1	MASCULINO	1	0
2	FEMENINO	0	1

En este caso el vector 2 de esta variable, se considera un vector con información repetida con respecto al vector 1, por lo que se mantiene solo un vector o se mantienen n-1 categorías.

El SCORE es una variable aleatoria calculada, ésta variable captura el comportamiento de pago de un cliente de microcrédito; básicamente consiste en un PUNTAJE otorgado a dicho cliente y que depende de la información propia

del individuo, de la actividad económica que practique y de su historial de crédito, las cuales determinan el perfil crediticio de una persona en particular.

Esta variable Z depende de las k Variables categóricas presentes en el modelo logístico, estas variables son las que poseen información del cliente y que el modelo estadístico aceptó como relevantes.

Del modelo estadístico se obtiene los coeficientes β_i para cada una de las variables categóricas, más un coeficiente para una contante que representa la información no capturada por el modelo. La estimación del modelo se deriva de la aplicación del algoritmo a la base de datos de la cartera de microcréditos a personas naturales, el resultado es una ecuación que arroja la puntuación Z y la probabilidad de impago p para cada uno de los prestatarios de dicha cartera.

El ejercicio, se realiza para la muestra total, para la muestra de desarrollo y se prueba en la muestra de testeo con corte a diciembre 2020 y la misma muestra de testeo con corte a julio 2021.

Los coeficientes de la ecuación del modelo logístico se presentan a continuación:

Variables: Depurada					
Variables en la ecuación	B	Error estándar	Sig.	Exp(B)	Beta_Error
Constante	-2.481	0.091	0.000	0.084	-27.286
mpv_histot_B__3	2.220	0.010	0.000	9.207	225.975
des_sector_SECTOR_URBANO	0.291	0.008	0.000	1.338	34.756
cliente_nuevo_SI	0.160	0.008	0.000	1.174	19.670
des_estado_civil_SOL_VIU_DIVO	0.132	0.008	0.000	1.142	16.863
ing_gas_quintil__1449	-0.327	0.011	0.000	0.721	-31.017
act_pas_quintil__472	0.320	0.010	0.000	1.377	30.498
act_pas_quintil__11813	0.008	0.010	0.437	1.008	0.778
cuotestburo_B__122	0.491	0.013	0.000	1.633	38.124
cuotestburo_B__122_334	0.257	0.016	0.000	1.293	16.057
cuotestburo_B__334_712	0.093	0.016	0.000	1.097	5.867
des_forma_pago_ANUAL	-0.298	0.022	0.000	0.742	-13.276
des_forma_pago_MENSUAL	-0.102	0.020	0.000	0.903	-5.198

des_forma_pago_NUEVE_MESES	-1.322	0.032	0.000	0.267	-40.758
des_forma_pago_OTROS	0.282	0.075	0.000	1.326	3.776
des_forma_pago_SEMESTRAL	-0.252	0.017	0.000	0.777	-14.651
des_op_sector_bnf_ACUICOLA	0.160	0.091	0.078	1.173	1.764
des_op_sector_bnf_AGRO_GANADERO	-0.449	0.086	0.000	0.638	-5.204
des_op_sector_bnf_ARTESANAL	-0.276	0.088	0.002	0.759	-3.135
des_op_sector_bnf_COMERCIO	-0.181	0.086	0.035	0.834	-2.105
des_op_sector_bnf_INDUSTRIAL	-0.240	0.087	0.006	0.787	-2.757
des_op_sector_bnf_OTROS	1.921	0.110	0.000	6.829	17.408
des_op_sector_bnf_SERVICIOS	-0.168	0.087	0.053	0.845	-1.937
des_zonal_ZONAL_CUENCA	-0.140	0.018	0.000	0.869	-7.823
des_zonal_ZONAL_GUAYAQUIL	0.306	0.013	0.000	1.358	23.623
des_zonal_ZONAL_IBARRA	0.021	0.015	0.149	1.022	1.442
des_zonal_ZONAL_LOJA	-0.155	0.014	0.000	0.856	-10.901
des_zonal_ZONAL_PORTOVIEJO	0.620	0.013	0.000	1.859	48.031
des_zonal_ZONAL_QUITO	0.149	0.016	0.000	1.161	9.201
endeudprom_B__0	1.104	0.011	0.000	3.016	103.415
endeudprom_B__0_1264	0.070	0.012	0.000	1.073	5.713
monto_B__3000	-0.159	0.012	0.000	0.853	-13.088
monto_B__3000_5000	-0.234	0.010	0.000	0.792	-22.252
monto_B__10000	0.080	0.011	0.000	1.084	7.009
numperiodos_B__5	-0.256	0.018	0.000	0.774	-14.359
numperiodos_B__12_35	0.292	0.015	0.000	1.339	19.009
numperiodos_B__35	0.337	0.022	0.000	1.401	15.109
plazomeses_B__24	0.612	0.022	0.000	1.844	28.051
plazomeses_B__24_36	0.692	0.017	0.000	1.998	41.214
plazomeses_B__36_60	0.525	0.015	0.000	1.690	35.748
sueldo_B__400	0.337	0.012	0.000	1.401	29.167
sueldo_B__400_800	0.051	0.011	0.000	1.053	4.671
sueldo_B__1250	0.045	0.012	0.000	1.046	3.827
tamano_ventas__1161	0.176	0.012	0.000	1.192	14.773
tamano_ventas__1161_2500	0.036	0.011	0.002	1.036	3.135
tamano_ventas__2500_6500	0.040	0.011	0.000	1.041	3.588

Fuente: BDD - Riesgos
Elaboración: Autor

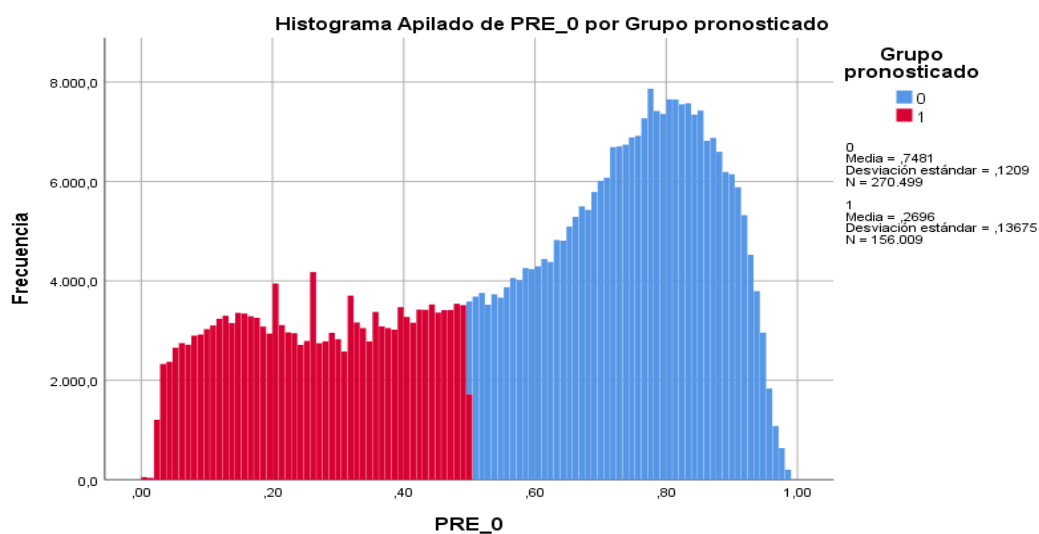
Variables: Desarrollo					
	B	Error estándar	Sig.	Exp(B)	Beta_Error
Constante	-2.019	0.090	0.000	0.133	-22.454
mpv_histot_B__3	2.220	0.010	0.000	9.207	225.975
des_sector_SECTOR_URBANO	0.291	0.008	0.000	1.338	34.756
cliente_nuevo_SI	0.160	0.008	0.000	1.174	19.670
des_estado_civil_SOL_VIU_DIVO	0.132	0.008	0.000	1.142	16.863
ing_gas_quintil__1449	-0.327	0.011	0.000	0.721	-31.017
act_pas_quintil__472	0.320	0.010	0.000	1.377	30.498
act_pas_quintil__11813	0.008	0.010	0.437	1.008	0.778
cuotestburo_B__122	0.491	0.013	0.000	1.633	38.124
cuotestburo_B__122_334	0.257	0.016	0.000	1.293	16.057
cuotestburo_B__334_712	0.093	0.016	0.000	1.097	5.867
des_forma_pago_ANUAL	-0.298	0.022	0.000	0.742	-13.276
des_forma_pago_MENSUAL	-0.102	0.020	0.000	0.903	-5.198
des_forma_pago_NUEVE_MESES	-1.322	0.032	0.000	0.267	-40.758
des_forma_pago_OTROS	0.282	0.075	0.000	1.326	3.776
des_forma_pago_SEMESTRAL	-0.252	0.017	0.000	0.777	-14.651
des_op_sector_bnf_ACUICOLA	0.160	0.091	0.078	1.173	1.764
des_op_sector_bnf_AGRO_GANADERO	-0.449	0.086	0.000	0.638	-5.204
des_op_sector_bnf_ARTESANAL	-0.276	0.088	0.002	0.759	-3.135
des_op_sector_bnf_COMERCIO	-0.181	0.086	0.035	0.834	-2.105
des_op_sector_bnf_INDUSTRIAL	-0.240	0.087	0.006	0.787	-2.757
des_op_sector_bnf_OTROS	1.921	0.110	0.000	6.829	17.408
des_op_sector_bnf_SERVICIOS	-0.168	0.087	0.053	0.845	-1.937
des_zonal_ZONAL_CUENCA	-0.140	0.018	0.000	0.869	-7.823
des_zonal_ZONAL_GUAYAQUIL	0.306	0.013	0.000	1.358	23.623
des_zonal_ZONAL_IBARRA	0.021	0.015	0.149	1.022	1.442
des_zonal_ZONAL_LOJA	-0.155	0.014	0.000	0.856	-10.901
des_zonal_ZONAL_PORTOVIEJO	0.620	0.013	0.000	1.859	48.031
des_zonal_ZONAL_QUITO	0.149	0.016	0.000	1.161	9.201
endeudprom_B__0	1.104	0.011	0.000	3.016	103.415
endeudprom_B__0_1264	0.070	0.012	0.000	1.073	5.713
monto_B__3000	-0.239	0.014	0.000	0.787	-16.549

monto_B__3000_5000	-0.314	0.013	0.000	0.731	-25.065
monto_B__5000_10000	-0.080	0.011	0.000	0.923	-7.009
numperiodos_B__5	-0.594	0.033	0.000	0.552	-17.783
numperiodos_B__6_12	-0.337	0.022	0.000	0.714	-15.109
numperiodos_B__12_35	-0.046	0.016	0.005	0.955	-2.807
plazomeses_B__24	0.612	0.022	0.000	1.844	28.051
plazomeses_B__24_36	0.692	0.017	0.000	1.998	41.214
plazomeses_B__36_60	0.525	0.015	0.000	1.690	35.748
sueldo_B__400	0.293	0.012	0.000	1.340	25.419
sueldo_B__400_800	0.007	0.011	0.531	1.007	0.626
sueldo_B__800_1250	-0.045	0.012	0.000	0.956	-3.827
tamano_ventas__1161	0.176	0.012	0.000	1.192	14.773
tamano_ventas__1161_2500	0.036	0.011	0.002	1.036	3.135
tamano_ventas__2500_6500	0.040	0.011	0.000	1.041	3.588

Fuente: BDD - Riesgos
Elaboración: Autor

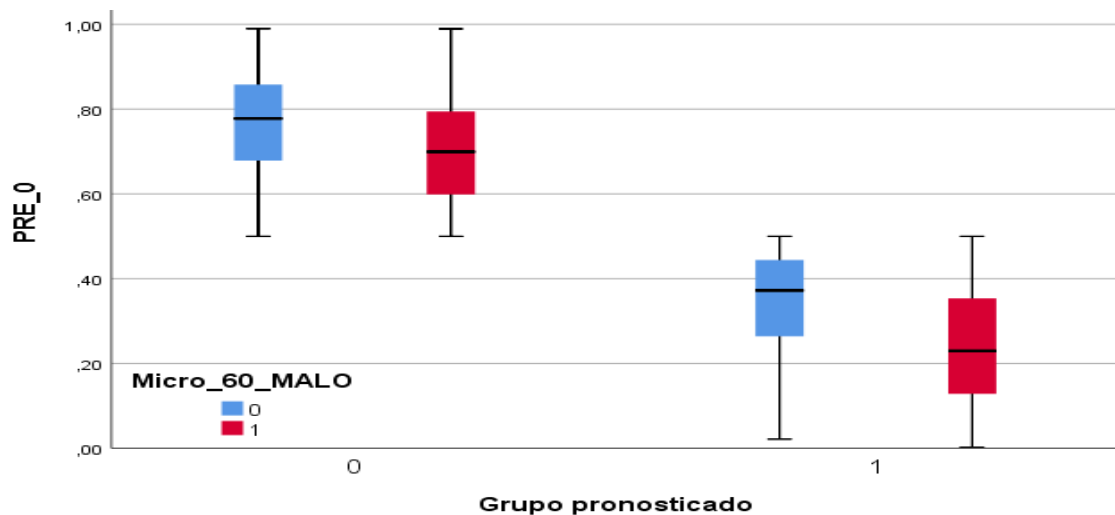
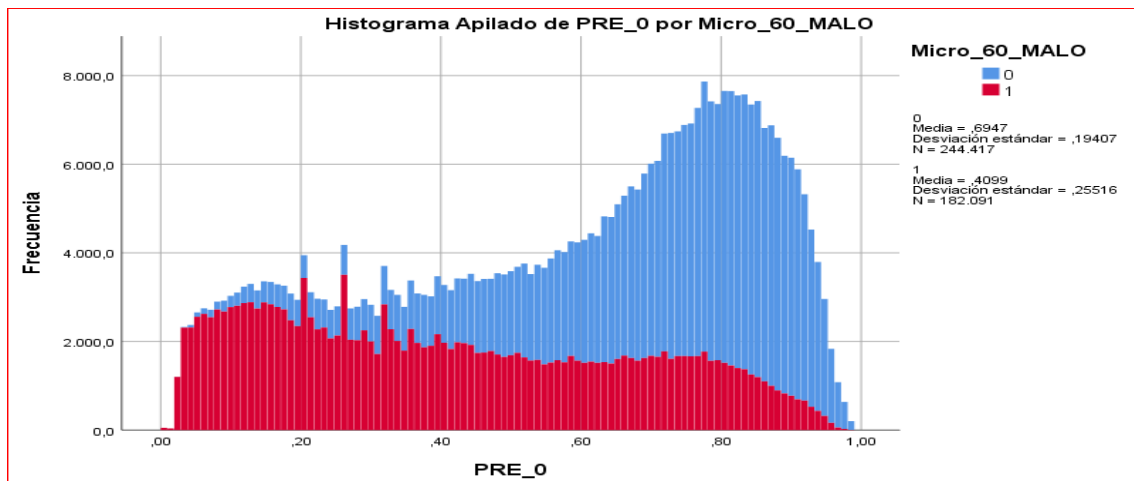
Como se aprecia tanto en la muestra total, como en la muestra de desarrollo, los signos de cada coeficiente son similares, excepto por los resaltados que muestran categorías y signos diferentes, se selecciona el modelo general debido a que los signos de cada una de las categorías no son similares, esto muestra indicios de una mayor segmentación entre categorías de una misma variable.

3.7.2. Resultados – data total



Los resultados muestran que los clientes definidos inicialmente como malos (moras mayores a 60 días) alcanzan puntajes menores a aquellos definidos como buenos.

Si se realiza el mismo ejercicio separado por población de buenos clientes y población de malos clientes, se obtiene que a medida que el puntaje crece, la presencia de clientes malos se reduce, mientras que la presencia de clientes buenos crece con un puntaje más alto.



3.8. Validación del modelo

3.8.1. Matriz de Confusión

Finalmente se tiene la matriz de confusión que es una forma sencilla de evaluar el ajuste del modelo, se usa como indicador de bondad de ajuste. Es una

tabla sencilla de 2x2, en la cual se muestra la distribución de los clientes buenos y malos, es decir; cuando $y=0$ e $y=1$ respectivamente, juntamente con la clasificación a cualquiera de las dos categorías según la probabilidad estimada

La interpretación se hace mediante el porcentaje de individuos bien clasificados, es decir, aquellos que mediante la probabilidad estimada permanecen en su respectiva categoría. También se interpreta el porcentaje de individuos mal clasificados, es decir, aquellos que mediante la probabilidad estimada se asignan a categorías diferentes de la cual fueron observados.

Matriz de Confusión

Micro_60_MALO	0	204102	40315	244417.0
	1	66490	115601	182091.0
Porcentaje global		270592	155916	426508.0

		Pronosticado		
		Micro_60_MALO		
		0	1	Porcentaje correcto
Micro_60_MALO	0	204044	40373	83.5
	1	66455	115636	63.5
			Exactitud	75.0%
			Precisión	74.1%
			Sensibilidad	63.5%
			Especificidad	83.5%

Los valores de la diagonal principal $a=204044$ y $d=40373$ corresponden con los valores estimados de forma correcta por el modelo, tanto los verdaderos positivos (d), como los verdaderos negativos (a). La otra diagonal, por tanto, representa los casos en los que el modelo “se ha equivocado ($c=66455$ falsos negativos, $b=40373$ falsos positivos).

Exactitud (accuracy) representa el porcentaje de predicciones correctas frente al total. Por tanto, es el cociente entre los casos bien clasificados por el modelo (verdaderos positivos y verdaderos negativos, es decir, los valores en la diagonal de la matriz de confusión), y la suma de todos los casos. Cuando los datos no son balanceados, no es una métrica confiable, si el algoritmo clasifica a todos los individuos como buenos, podría ser muy preciso (hasta un 99%).

$$(204044+115636)/(426508)=75\%$$

Precisión: se refiere a lo cerca que está el resultado de una predicción del valor verdadero. Por tanto, es el cociente entre los casos positivos bien clasificados por el modelo y el total de predicciones positivas. Para calcularlo a mano, a partir de la matriz de confusión:

$$(115636)/(115636+40373)=74.1\%$$

Sensibilidad (exhaustividad - recall) representa la tasa de verdaderos positivos. Es la proporción entre los casos positivos bien clasificados por el modelo, respecto al total de positivos.

$$115636/(66455+115636)=63.5\%$$

Es la habilidad del modelo de detectar los casos relevantes. Un 63.5% es claramente un valor aceptable como métrica. Se puede decir que el modelo de clasificación es aceptable, aunque “se le escapan” algunos positivos.

Especificidad: Es la proporción entre los casos negativos bien clasificados por el modelo, respecto al total de negativos (es la tasa de verdaderos negativos).

$$204044/(204044+40373)=83.5\%$$

En este caso, la especificidad tiene un valor muy bueno. Esto significa que su capacidad de discriminar los casos negativos es muy buena. O lo que es lo mismo, es difícil obtener falsos positivos.

La conveniencia de usar una métrica u otra como medida de bondad de ajuste dependerá de cada caso en particular y, en concreto, del “coste” asociado a cada error de clasificación del algoritmo. En este ejemplo, la sensibilidad = 63.5% y la especificidad = 83.5%. Por tanto, el modelo es mucho más específico que sensible.

Esta situación es de interés cuando el objetivo es evitar a toda costa los falsos positivos. Si el interés es identificar los verdaderos negativos, (evitar falsos positivos) se debe elegir especificidad alta. Por el contrario, si las “falsas alarmas” no son de interés, y lo que se desea evitar son los falsos negativos, una

mayor sensibilidad o recall es lo que se debe buscar, es decir; elegir sensibilidad alta.

Los resultados del modelo son:

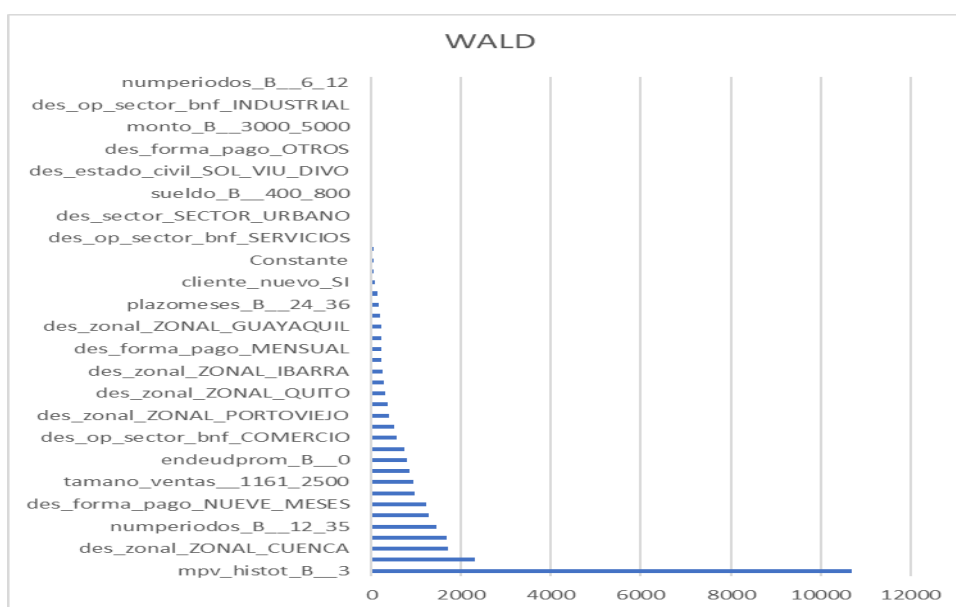
Exactitud	75.0%
Precisión	74.1%
Sensibilidad	63.5%
Especificidad	83.5%

A continuación, se estima la significación estadística de los coeficientes del modelo a través del estadístico de Wald y la bondad del ajuste. Si una variable resulta ser no significativa, se procede inmediatamente a eliminarla del modelo.

3.8.2. Pruebas Estadísticas

Además, existen las pruebas de hipótesis características de aceptación o rechazo de un modelo estadístico entre las que se pueden mencionar de manera breve se tienen:

Prueba de Wald: Esta prueba se utiliza para evaluar la significancia estadística de cada variable explicativa o regresora. Se divide el Coeficiente estimado para el error estándar.



Prueba Omnibus: Esta prueba es útil para analizar la significancia conjunta de las componentes del modelo.

Pruebas ómnibus de coef del modelo

	Chi-cuadrado	gl	Sig.
Paso	128,391,256	56	,000
Bloque	128,391,256	56	,000
Modelo	128,391,256	56	,000

Los coeficientes del modelo, son estadísticamente significativos y distintos de cero.

Prueba de Hosmer-Leshow: Esta es una prueba de bondad de ajuste, aspira a detectar variables no lineales e interacciones no tenidas en cuenta.

Prueba de Hosmer y Lemeshow

Paso	Chi-cuadrado	gl	Sig.
1	668,117	8	,000

Tabla de contingencia para la prueba de Hosmer y Lemeshow

		Micro_60 = BUEN		Micro_60 = MALO		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	37500	38314,461	5151	4336,539	42651
	2	34841	35614,862	7810	7036,138	42651
	3	33530	33619,268	9121	9031,732	42651
	4	32080	31478,551	10571	11172,449	42651
	5	29821	28891,886	12830	13759,114	42651
	6	26717	25582,823	15934	17068,177	42651
	7	21728	21421,878	20923	21229,122	42651
	8	15279	16115,291	27371	26534,709	42650
	9	9625	9653,549	33027	32998,451	42652
	10	3296	3724,432	39353	38924,568	42649

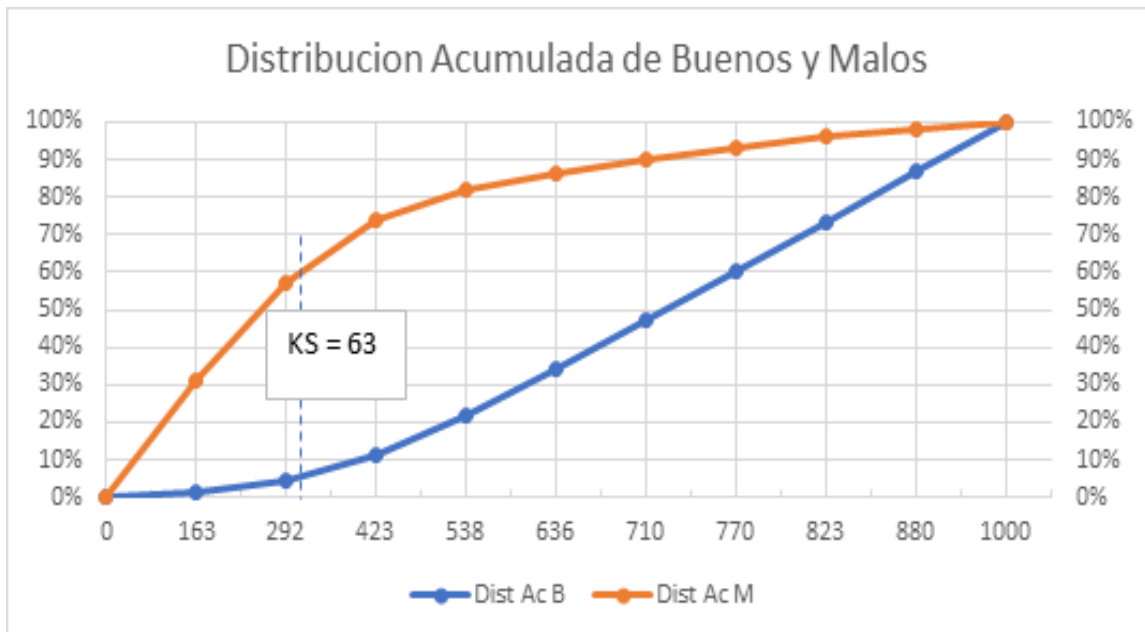
Y finalmente se cuenta con las pruebas gráficas que parten de la construcción de la tabla de clasificación de clientes considerados como buenos y malos y lo que se resume en los siguientes pasos:

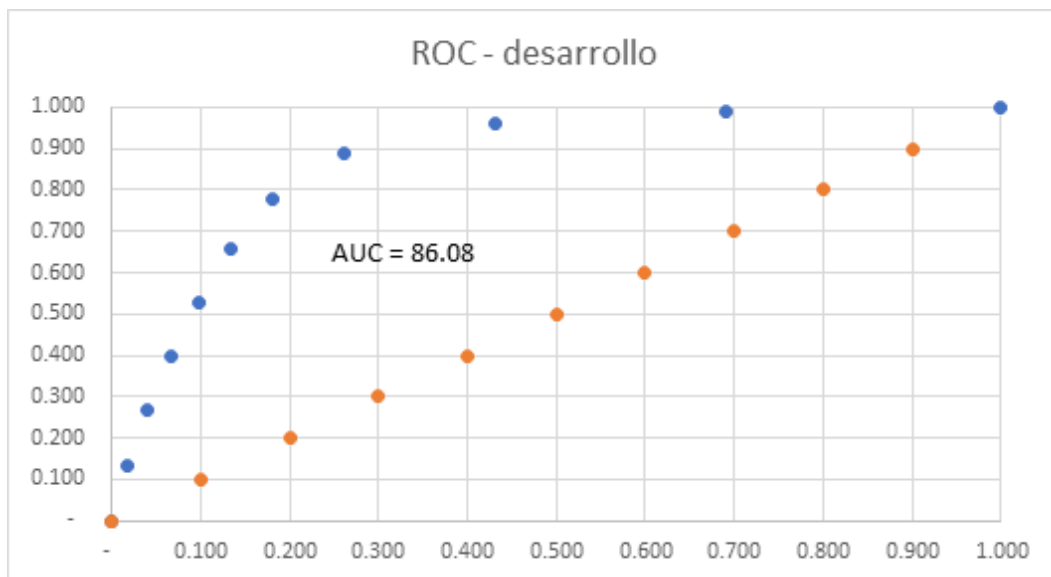
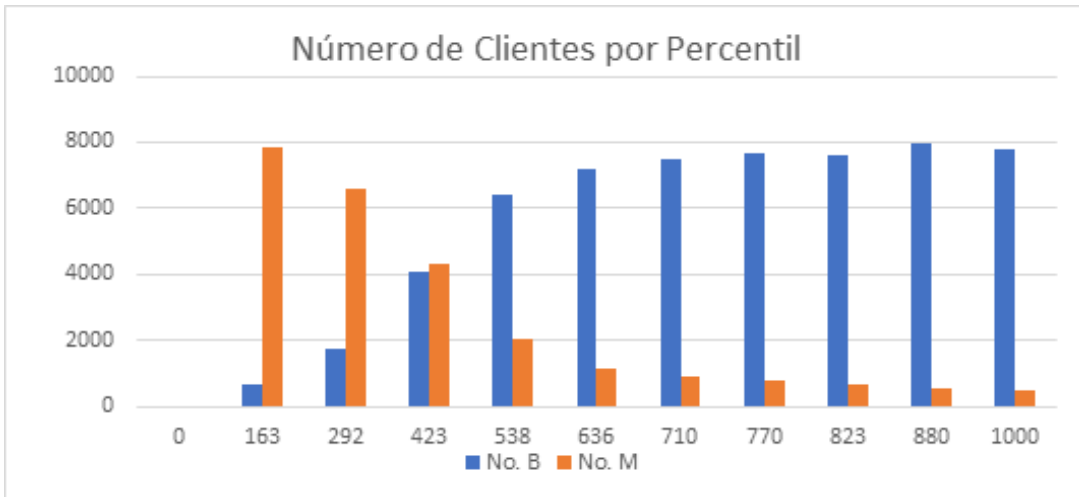
3.8.3. Curva ROC

Mediante esta representación de los pares (1-especificidad; sensibilidad) obtenidos al considerar todos los puntos de corte del score card, la curva ROC proporciona una representación global de la exactitud del modelo. Note que la curva ROC será necesariamente creciente, propiedad que refleja el compromiso entre sensibilidad y especificidad: si se modifica el punto de corte para obtener mayor sensibilidad, sólo puede hacerse a expensas de disminuir al mismo tiempo la especificidad.

RESULTADOS DEL MODELO CON BASE DESARROLLO - CORTE 12-2020

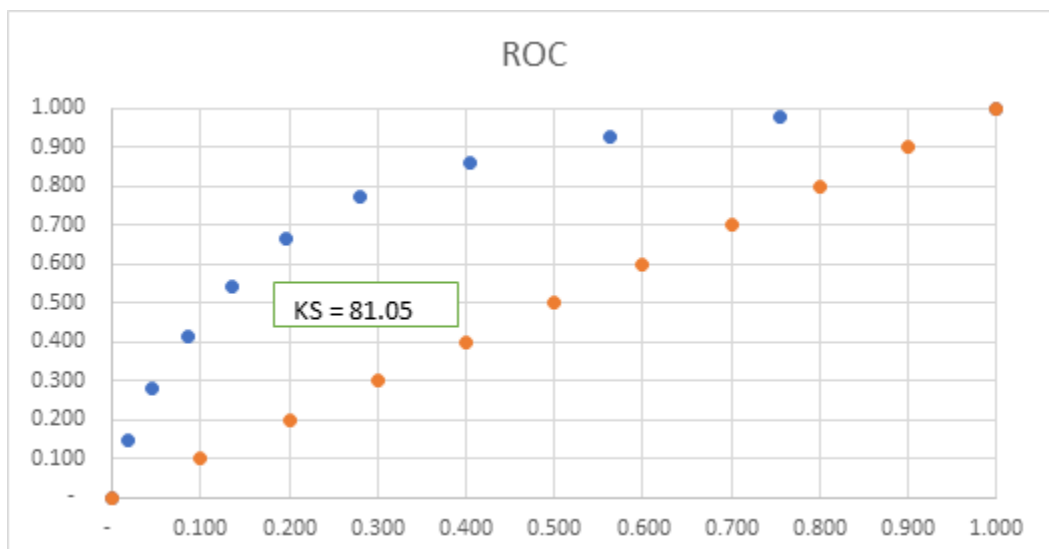
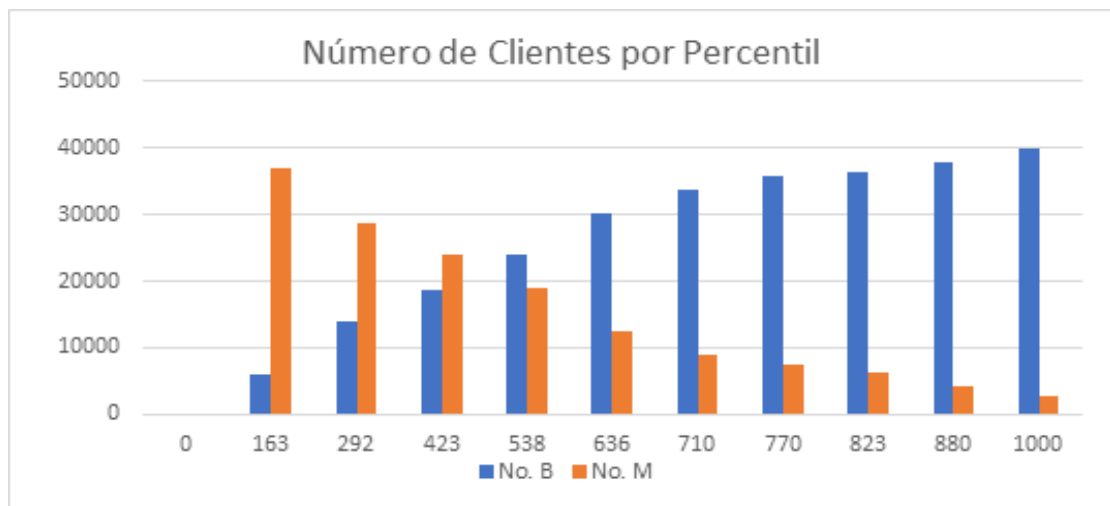
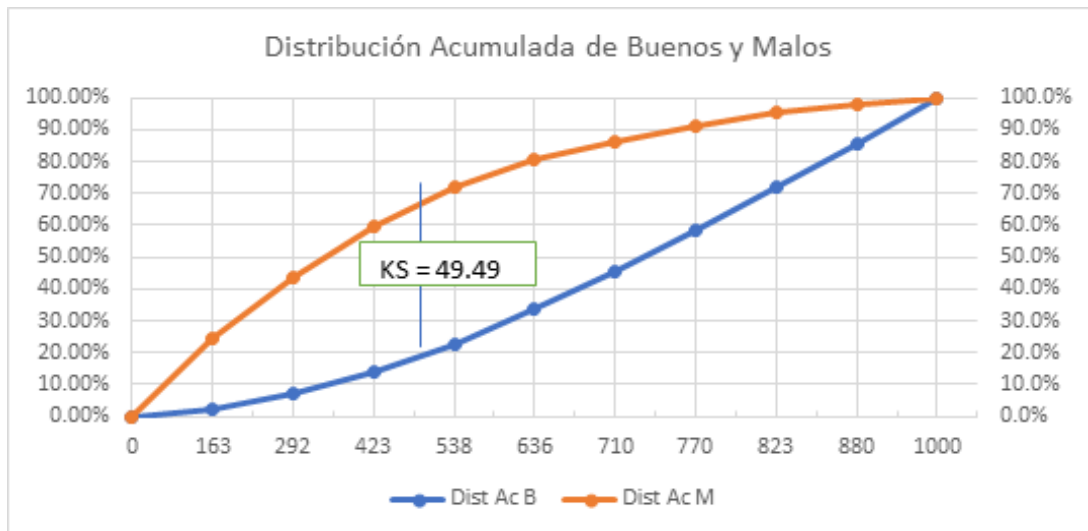
Percentil	Rango score	Total Población	No. B	No. B Ac.	Dist Ac B	Dist Ac B	No. M	No. M Ac.	Dist Ac M	Dist Ac M	ODDS	Tasa de Malos	Gini	K-S	AUC
0	0		0	0	0		0	0	0				0.00%	0.00%	0.93%
1	0 163	8498	642	642	1.1%	1.1%	7,856	7,856	31.0%	31.0%	0:1	92.4%	0.34%	29.86%	1.68%
2	164 292	8331	1,759	2,401	4.1%	5.2%	6,572	14,428	56.9%	25.9%	0:1	78.9%	1.34%	52.76%	2.40%
3	293 423	8394	4,064	6,465	11.0%	15.1%	4,330	18,758	73.9%	17.1%	0:1	51.6%	2.58%	62.89%	2.42%
4	424 538	8465	6,422	12,887	22.0%	33.0%	2,043	20,801	82.0%	8.0%	3:1	24.1%	2.66%	59.99%	1.93%
5	539 636	8354	7,184	20,071	34.2%	56.2%	1,170	21,971	86.6%	4.6%	6:1	14.0%	2.59%	52.34%	1.48%
6	637 710	8409	7,486	27,557	47.0%	81.2%	923	22,894	90.2%	3.6%	8:1	11.0%	2.95%	43.22%	1.08%
7	711 770	8479	7,686	35,243	60.1%	107.1%	793	23,687	93.3%	3.1%	9:1	9.4%	3.35%	33.23%	0.69%
8	771 823	8305	7,623	42,866	73.1%	133.2%	682	24,369	96.0%	2.7%	11:1	8.2%	3.58%	22.92%	0.40%
9	824 880	8514	7,970	50,836	86.7%	159.8%	544	24,913	98.2%	2.1%	14:1	6.4%	3.43%	11.47%	0.12%
10	881 1000	8271	7,805	58,641	100.0%	186.7%	466	25,379	100.0%	1.8%	16:1	5.6%	3.43%	0.00%	0.00%
	Totales	84020	58,641	58,641	100.0%		25,379	25,379	100.0%				73.76%	62.89%	86.88%

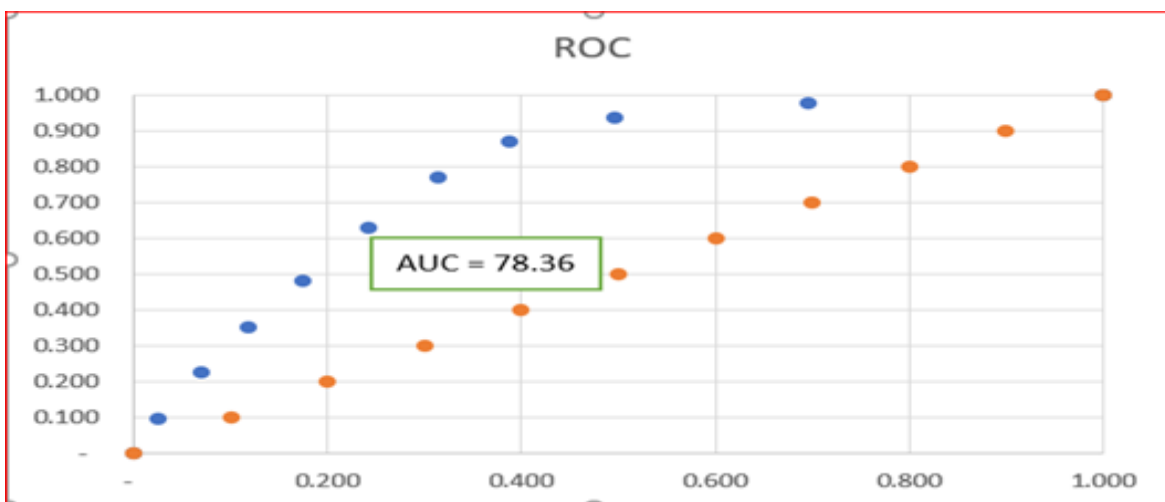
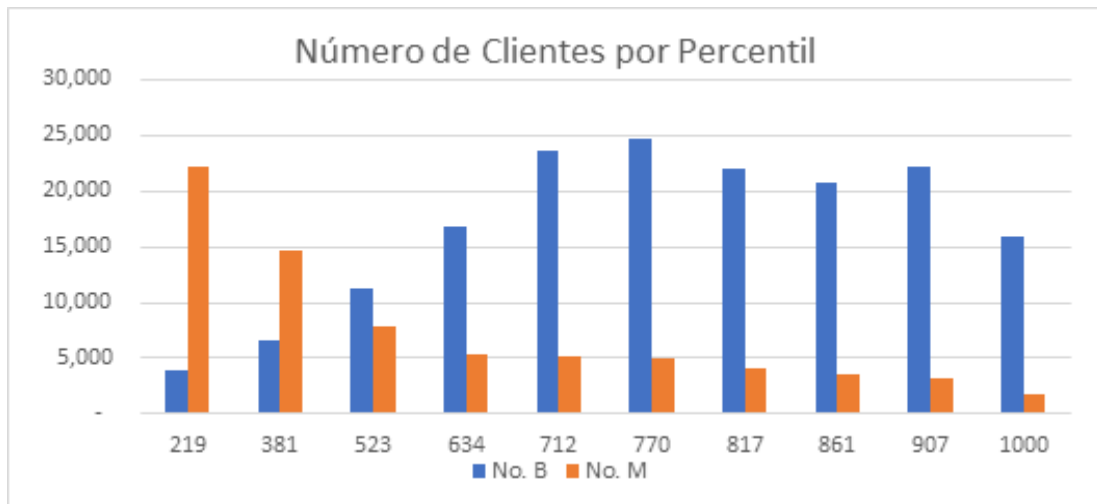




RESULTADOS DEL MODELO CON BASE TOTAL DEPURADA – CORTE 12-2020

Percentil	Rango score	Total Poblacion	No. B	No. B Ac.	Dist Ac B	Dist Ac B	No. M	No. M Ac.	Dist Ac M	Dist Ac M	ODDS	Tasa de Malos	Gini	K-S	AUC
0	0		0	0	0		0	0	0.0%				0.00%	0.00%	1.91%
1	0 163	42877	6022	6,022	2.2%	2.2%	36855	36,855	24.5%	24.5%	0:1	86.0%	0.53%	22.31%	3.32%
2	164 292	42668	13888	19,910	7.2%	9.4%	28780	65,635	43.6%	19.1%	0:1	67.5%	1.80%	36.40%	3.25%
3	293 423	42440	18530	38,440	13.9%	21.1%	23910	89,545	59.5%	15.9%	0:1	56.3%	3.36%	45.57%	2.96%
4	424 538	42793	23873	62,313	22.6%	36.5%	18920	108,465	72.1%	12.6%	1:1	44.2%	4.59%	49.49%	2.59%
5	539 636	42732	30147	92,460	33.5%	56.1%	12585	121,050	80.4%	8.4%	2:1	29.5%	4.69%	46.93%	2.03%
6	637 710	42607	33705	126,165	45.7%	79.2%	8902	129,952	86.3%	5.9%	3:1	20.9%	4.68%	40.63%	1.45%
7	711 770	43225	35734	161,899	58.7%	104.4%	7491	137,443	91.3%	5.0%	4:1	17.3%	5.19%	32.67%	0.87%
8	771 823	42494	36322	198,221	71.8%	130.5%	6172	143,615	95.4%	4.1%	5:1	14.5%	5.35%	23.61%	0.44%
9	824 880	42049	37869	236,090	85.5%	157.3%	4180	147,795	98.2%	2.8%	9:1	9.9%	4.37%	12.66%	0.13%
10	881 1000	42645	39932	276,022	100.0%	185.5%	2713	150,508	100.0%	1.8%	14:1	6.4%	3.34%	0.00%	0.00%
	Totales	426530	276,022	276,022	100.0%		150,508	150,508	100.0%				62.09%	49.49%	81.05%





Estos indicadores de GINI = 56.71%, KS = 48.01 y AUC =78.36% tienen una fecha de corte a julio 2021 y son estadísticamente aceptables y cumplen los parámetros de evaluación para este tipo de modelos como un Gini > 50% y un KS que al menos sea los 2/3 del GINI y desde la fecha de originación que fue el periodo 2019, los mismos no se han deteriorado de manera importante, lo que valida la construcción del modelo y se recomienda su implementación.

3.9. Implementación del modelo

La implementación del modelo de score es importante en el proceso de originación del crédito, pero el comportamiento del cliente depende de todo el proceso existente en BanEcuador, por lo que cuando se proceda con la

implementación del modelo, se revisen todas las etapas del proceso de crédito y se verifique su alineación con la estrategia del negocio.

Hay varias maneras en las cuales un buen modelo estadístico no es garantía de un buen desempeño, por ejemplo: si la data utilizada es desactualizada, o por ejemplo que el corte mínimo de aceptación no es el correcto o sea demasiado exigente en términos de puntaje y termine rechazando a la mayoría de los clientes

3.9.1. *Tabla Scorecard*

Una manera sencilla de implantación se deriva de la estimación de la Ecuación Z previamente definida y la posterior construcción de la SCORECARD

Un scorecard o tabla de puntaje es una tabla que resume la información sobre los pesos y signos asignados a las características relevantes para identificar el comportamiento de pago de un cliente. En su forma más simple es una tabla en la que se listan las variables seleccionadas y sus atributos con el puntaje asignado a cada uno. Un puntaje negativo indica que la variable castiga al cliente que cumpla con este atributo, mientras que una variable positiva estaría premiándolo. Un mayor o menor valor o peso significa cuán importante es la variable en la calificación.

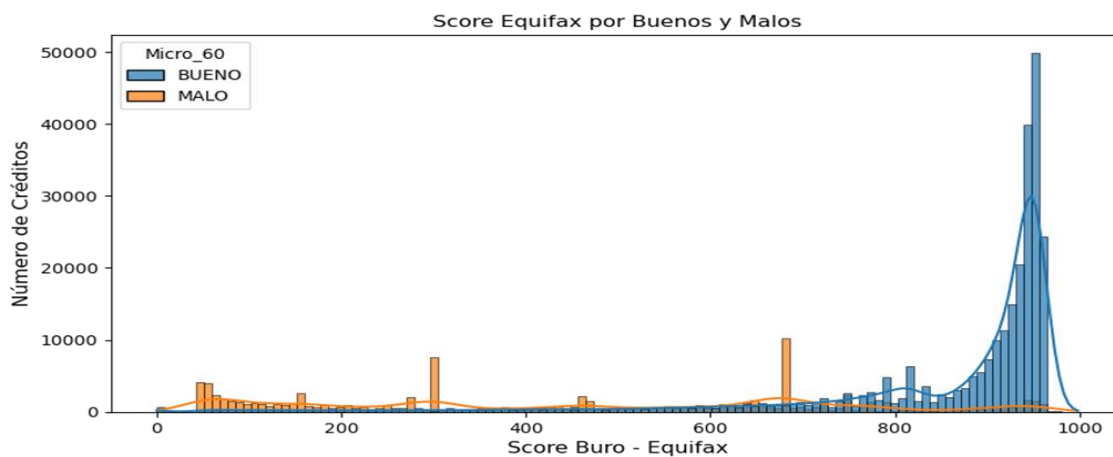
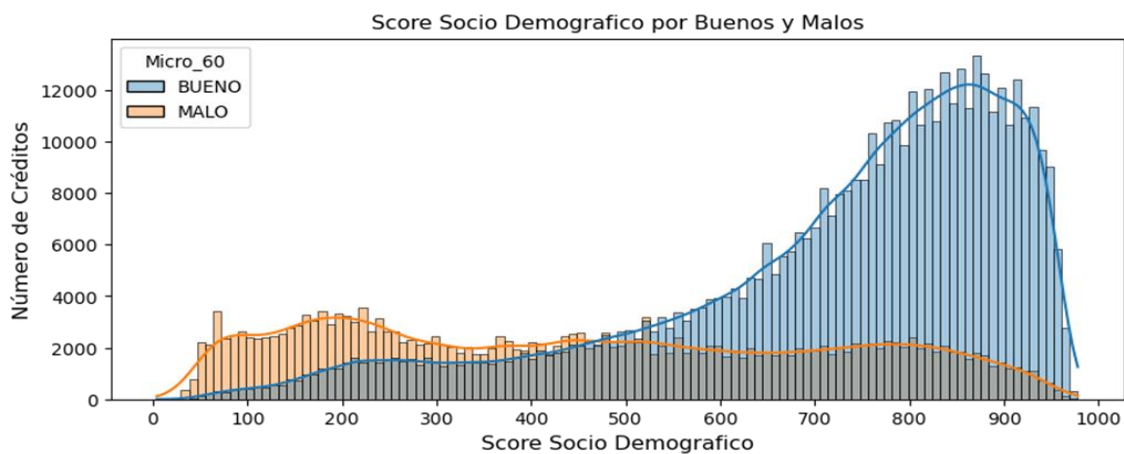
	B	Error estándar	Wald	Sig.	Exp(B)
des_estado_civil_SOL_VIU_DIVO	0.132	0.008	284.357	0.000	1.142
cliente_nuevo_SI	0.160	0.008	386.892	0.000	1.174
des_sector_SECTOR_URBANO	0.291	0.008	1207.958	0.000	1.338
ing_gas_quintil__1449	-0.327	0.011	962.042	0.000	0.721
des_op_sector_bnf_ACUICOLA	0.160	0.091	3.112	0.078	1.173
des_op_sector_bnf_AGRO_GANADERO	-0.449	0.086	27.084	0.000	0.638
des_op_sector_bnf_ARTESANAL	-0.276	0.088	9.829	0.002	0.759
des_op_sector_bnf_COMERCIO	-0.181	0.086	4.433	0.035	0.834
des_op_sector_bnf_INDUSTRIAL	-0.240	0.087	7.602	0.006	0.787
des_op_sector_bnf_OTROS	1.921	0.110	303.028	0.000	6.829
des_op_sector_bnf_SERVICIOS	-0.168	0.087	3.753	0.053	0.845
des_zonal_ZONAL_CUENCA	-0.140	0.018	61.202	0.000	0.869
des_zonal_ZONAL_GUAYAQUIL	0.306	0.013	558.046	0.000	1.358
des_zonal_ZONAL_IBARRA	0.021	0.015	2.079	0.149	1.022

des_zonal_ZONAL_LOJA	-0.155	0.014	118.834	0.000	0.856
des_zonal_ZONAL_PORTOVIEJO	0.620	0.013	2306.977	0.000	1.859
des_zonal_ZONAL_QUITO	0.149	0.016	84.667	0.000	1.161
des_forma_pago_ANUAL	-0.298	0.022	176.253	0.000	0.742
des_forma_pago_MENSUAL	-0.102	0.020	27.018	0.000	0.903
des_forma_pago_NUEVE_MESES	-1.322	0.032	1661.250	0.000	0.267
des_forma_pago_OTROS	0.282	0.075	14.259	0.000	1.326
des_forma_pago_SEMESTRAL	-0.252	0.017	214.656	0.000	0.777
monto_B__10000	0.080	0.011	49.131	0.000	1.084
monto_B__3000	-0.159	0.012	171.296	0.000	0.853
monto_B__3000_5000	-0.234	0.010	495.148	0.000	0.792
sueldo_B__1250	0.045	0.012	14.648	0.000	1.046
sueldo_B__400	0.337	0.012	850.706	0.000	1.401
sueldo_B__400_800	0.051	0.011	21.818	0.000	1.053
tamano_ventas__1161	0.176	0.012	218.254	0.000	1.192
tamano_ventas__1161_2500	0.036	0.011	9.827	0.002	1.036
tamano_ventas__2500_4600	0.040	0.011	12.870	0.000	1.041
act_pas_quintil__11813	0.008	0.010	0.605	0.437	1.008
act_pas_quintil__472	0.320	0.010	930.138	0.000	1.377
cuotestburo_B__122	0.491	0.013	1453.407	0.000	1.633
cuotestburo_B__122_334	0.257	0.016	257.837	0.000	1.293
cuotestburo_B__334_712	0.093	0.016	34.425	0.000	1.097
numperiodos_B__12_35	0.292	0.015	361.353	0.000	1.339
numperiodos_B__35	0.337	0.022	228.292	0.000	1.401
numperiodos_B__5	-0.256	0.018	206.194	0.000	0.774
plazomeses_B__24	0.612	0.022	786.872	0.000	1.844
plazomeses_B__24_36	0.692	0.017	1698.614	0.000	1.998
plazomeses_B__36_60	0.525	0.015	1277.931	0.000	1.690
endeudprom_B__0	1.104	0.011	10694.642	0.000	3.016
endeudprom_B__0_1264	0.070	0.012	32.638	0.000	1.073
mpv_histot_B__3	2.220	0.010	51064.759	0.000	9.207
Constante	-2.481	0.091	744.534	0.000	0.084

Cabe resaltar que el modelo de regresión esta implementado en un ambiente SQL en las bases de riesgos y se está monitoreando de manera trimestral, además ha pasado las pruebas de validación con la gerencia de colocaciones.

Estrategias

Del modelo de score sociodemográfico de originación de crédito para personas naturales del segmento micro, se obtuvo los siguientes indicadores de desempeño:



Este modelo de score se aplica los clientes de la base de prueba, que comprenden a créditos otorgados en el 2019

Percentil	Rango score	No. B	No. B Ac.	Dist Ac B	Dist Ac B	No. M	No. M Ac.	Dist Ac M	Dist Ac M	Gini	K-S	AUC
1	0 219	2,250	2,250	0.03	2.7%	12,204	12,204	38.7%	38.7%	1.06%	35.98%	0.022
2	219 381	3,579	5,829	0.07	9.9%	7,056	19,260	61.1%	22.4%	2.21%	54.00%	0.020
3	381 523	4,981	10,810	0.13	20.3%	3,369	22,629	71.8%	10.7%	2.17%	58.60%	0.021
4	523 634	6,775	17,585	0.21	34.7%	1,752	24,381	77.4%	5.6%	1.93%	55.89%	0.024
5	634 712	9,664	27,249	0.33	54.8%	1,614	25,995	82.5%	5.1%	2.81%	49.20%	0.020
6	712 770	11,070	38,319	0.47	80.1%	1,494	27,489	87.2%	4.7%	3.80%	40.42%	0.014
7	770 817	10,828	49,147	0.60	106.9%	1,186	28,675	91.0%	3.8%	4.02%	30.96%	0.009
8	817 861	10,709	59,856	0.73	133.2%	1,119	29,794	94.5%	3.6%	4.73%	21.42%	0.005
9	861 907	11,844	71,700	0.88	160.7%	1,070	30,864	97.9%	3.4%	5.46%	10.35%	0.001
10	907 1000	10,155	81,855	1.00	187.6%	648	31,512	100.0%	2.1%	3.86%	0.00%	-
Totales		81,855	81,855	100.0%		31,512	31,512	100.0%		67.96%	58.60%	0.8619

Con lo que se obtuvo un coeficiente de Gini del 68% un índice KS del 58.6 % y un Indicador ROC del 86%, lo que se considera estadísticamente aceptable para el uso de estos modelos estadísticos para ser implementados en la gestión de crédito del Banco.

Los segmentos de clientes a considerar se toman a partir de los percentiles del score de originación y se evalúan de manera aleatoria usando algoritmos de segmentación. Para la evaluación Se construyen indicadores adicionales como:

- Relación entre Buenos y Malos Clientes por segmento
- Tasa de Malos, definida como el número de clientes malos sobre el total de la población en cada segmento.
- Pérdida, definida como el producto entre la tasa de malos y el porcentaje de población acumulada en cada percentil
- Pérdida Acumulada por cada percentil

Método 1 – Deciles de población del Score

Percentil	Rango score	Total Poblacion	% Pob.	% Pob Ac	ODDS	M_B	Tasa de M	Pérdida	Pérdida Ac
1	0 219	14454	12.7%	12.7%	0 : 1	5 : 1	84.4%	10.8%	10.8%
2	219 381	10635	9.4%	22.1%	0 : 1	1 : 1	66.3%	14.7%	25.4%
3	381 523	8350	7.4%	29.5%	1 : 1	0 : 1	40.3%	11.9%	26.6%
4	523 634	8527	7.5%	37.0%	3 : 1	0 : 1	20.5%	7.6%	19.5%
5	634 712	11278	9.9%	47.0%	5 : 1	0 : 1	14.3%	6.7%	14.3%
6	712 770	12564	11.1%	58.0%	7 : 1	0 : 1	11.9%	6.9%	13.6%
7	770 817	12014	10.6%	68.6%	9 : 1	0 : 1	9.9%	6.8%	13.7%
8	817 861	11828	10.4%	79.1%	9 : 1	0 : 1	9.5%	7.5%	14.3%
9	861 907	12914	11.4%	90.5%	11 : 1	0 : 1	8.3%	7.5%	15.0%
10	907 1000	10803	9.5%	100.0%	15 : 1	0 : 1	6.0%	6.0%	13.5%
Totales		113367							

De la tabla de validación inicial se deduce que clientes con un score superior a 523 puntos serían sujetos de aprobación automática o que continúen con el proceso de aprobación, debido que es precisamente en este percentil donde se alcanza una diferenciación máxima entre clientes malos y clientes buenos definidos en la metodología de score.

Dado a que la relación ente buenos clientes y malos clientes es 1:1 en el percentil 3, comprendido entre 381 hasta 523 puntos (incluidos), el modelo estadístico no puede diferenciar entre buenos y malos clientes, por lo que es necesario una revisión manual para aquellos clientes que caigan en este rango.

Se recomienda que aquellos clientes con puntajes menores a 381 puntos sean rechazados de manera automática.

Con estos cortes, se estaría restringiendo cerca del 22.1 % de las aplicaciones, esto significa un 25.4% de reducción en las pérdidas que se estarían generando por créditos mal colocados, con una pérdida promedio del 7% en cada decil a partir de los 523 puntos.

La pérdida potencial para los clientes con incertidumbre (< 523) es del 37.3% y cerca del 30% estarían rechazadas. Tomando en cuenta que si tiene un score menor a 381 puntos el rechazo directo es del 22.1%

Como una ampliación de este ejercicio, se realiza una desagregación (un corte más fino) de 20 percentiles.

Percentil	Rango score	Total Poblacion	% Pob.	% Pob Ac	ODDS	M_B	Tasa de M	Perdida	Perdida Ac
1	0 130	5689	5.0%	5.0%	0 : 1	8 : 1	89.0%	4.5%	4.5%
2	130 188	5711	5.0%	10.1%	0 : 1	4 : 1	82.8%	8.3%	12.8%
3	188 248	5605	4.9%	15.0%	0 : 1	3 : 1	78.4%	11.8%	20.1%
4	248 336	5675	5.0%	20.0%	0 : 1	1 : 1	66.1%	13.2%	25.0%
5	336 439	5690	5.0%	25.0%	0 : 1	1 : 1	50.5%	12.6%	25.9%
6	439 533	5670	5.0%	30.0%	1 : 1	0 : 1	34.7%	10.4%	23.1%
7	533 610	5691	5.0%	35.0%	3 : 1	0 : 1	21.2%	7.4%	17.9%
8	610 661	5627	5.0%	40.0%	5 : 1	0 : 1	16.3%	6.5%	14.0%
9	661 700	5738	5.1%	45.1%	6 : 1	0 : 1	14.2%	6.4%	12.9%
10	700 730	5655	5.0%	50.1%	6 : 1	0 : 1	12.6%	6.3%	12.7%
11	730 756	5694	5.0%	55.1%	7 : 1	0 : 1	11.7%	6.4%	12.7%
12	756 780	5756	5.1%	60.2%	8 : 1	0 : 1	10.8%	6.5%	12.9%
13	780 802	5686	5.0%	65.2%	9 : 1	0 : 1	9.8%	6.4%	12.9%
14	802 823	5496	4.8%	70.0%	9 : 1	0 : 1	9.9%	6.9%	13.3%
15	823 845	5888	5.2%	75.2%	9 : 1	0 : 1	9.4%	7.1%	14.0%
16	845 865	5458	4.8%	80.0%	9 : 1	0 : 1	9.4%	7.5%	14.6%
17	865 885	5673	5.0%	85.0%	10 : 1	0 : 1	8.9%	7.6%	15.1%
18	885 906	5877	5.2%	90.2%	12 : 1	0 : 1	7.5%	6.8%	14.4%
19	906 926	5529	4.9%	95.1%	13 : 1	0 : 1	6.9%	6.6%	13.4%
20	926 1000	5559	4.9%	100.0%	18 : 1	0 : 1	5.2%	5.2%	11.8%
	Totales	113367	100.0%						

Lo que se evidencia es que el rechazo automático pasa del 22% al 25% y la pérdida acumulada pase del 25.4 al 26%; es decir; el puntaje mínimo pasa de 523 a 533 manteniendo el 30% potencial de clientes rechazados, por lo cual se considera un escenario más viable con la misma exposición al riesgo.

De este análisis se acordó con la gerencia de Colocaciones y Captaciones que los rangos ideales para rechazar o aceptar a un potencial cliente son los siguientes.

Capítulo IV. Conclusiones y Recomendaciones

4.1. CONCLUSIONES

- El modelo Scoring de Originación de Micro Crédito, dependerá de la calidad de información y la depuración que se le dé a esta, al no existir un modelo de gobernanza de datos, el proceso de depuración es el más demandante en el proceso de construcción del modelo de scoring. Esto es de vital importancia debido a que el score estará al inicio del proceso del crédito.
- En base al modelo logit estimado, se evidencia claramente que los resultados (puntaje estimado) son muy superiores a los obtenidos por el score de buró de crédito.
- El puntaje obtenido en el rango de 300 a 600 puntos logra diferenciar clientes buenos de malos, lo que permite mejorar la calidad de cartera en el mediano plazo.
- Por medio de la identificación de patrones de atrasos en los créditos y las características sociodemográficas subyacentes es posible predecir el comportamiento de pagos de clientes que no poseen información financiera previa, lo que genera una reducción del riesgo de préstamos y por ende una inclusión financiera.
- Con las tablas de validación del score es posible cuantificar criterios de apetito al riesgo, debido a que mediante los deciles se puede identificar límites y variaciones en la participación de clientes malos y buenos y el correspondiente trade off, al modificar límites de un puntaje mínimo.

- Una vez validado un modelo estadístico es posible ponerlo en producción en el core bancario con lo que se acelera el proceso de colocaciones.

4.2. RECOMENDACIONES

Se recomienda la implementación de este modelo en el core bancario y su uso para los clientes nuevos o sin historial crediticio.

Se recomienda la implementación de un modelo de gobernanza de datos con el fin de mejorar la calidad de la data , ajustando las opciones de ingreso de la información de los solicitantes.

Se recomienda el ajuste de la política de crédito vigente, ya que el puntaje mínimo requerido de 300 puntos es perjudicial para el Banco y del análisis desarrollado en esta investigación muestra puntajes superiores, lo que estaría generando perdidas para el Banco.

Se recomienda el desarrollo de modelos estadísticos predictivos para los otros segmentos de crédito.

Se recomienda el uso de plataformas de analítica avanzada y de código libre con el fin de obtener la data almacenada en los servidores del banco. Y permita la aplicación de algoritmos de bigdata

Bibliografía

- Álava, J., & Bravo, V. (2020). La banca pública y su contribución al financiamiento del sector microempresarial caso BanEcuador de la ciudad de Portoviejo. *Dominio de las Ciencias*, 6(3), 533-547. doi:<https://doi.org/10.23857/dc.v6i3.1298>
- Asencio, L., Romero, J., Andrade, C., Caicedo, W., Coca, J., Chamba, L., & Ortiz, D. (2021). *La Banca y el Sistema Financiero Ecuatoriano*. Guayaquil, Ecuador: Editorial Grupo Compás.
- Asociación de Bancos del Ecuador. (2019). *Sistema Bancario Términos Clave*. Retrieved from <https://datalab.asobanca.org.ec/resources/site/terminos/Principales%20Cuentas.pdf>
- Asociación de Bancos del Ecuador. (2022). *Evolución de la banca privada ecuatoriana*. Retrieved from Octubre: <https://asobanca.org.ec/wp-content/uploads/2022/11/Evolucion-de-la-Banca-10-2022-completo-1.pdf>
- Astudillo, V. (2022). *Modelo de scoring para una entidad financiera especializada en el*. Universidad EAFIT. Obtenido de https://repository.eafit.edu.co/bitstream/handle/10784/31543/Valeria_AstudilloGiron_2022.pdf?sequence=2&isAllowed=y
- Banco de Desarrollo del Ecuador B.P. (2021). *Guía de Procesos Crediticios, Formulación y Gestión de Proyectos*. Obtenido de https://bde.fin.ec/wp-content/uploads/2021/06/GUIA_DE_PROCESOS_CREIDTICIOS_-MODULO_1.pdf
- Banco Santander. (2022, junio 29). *Riesgos financieros: qué son, tipos y consejos para enfrentarse a ellos*. Retrieved from <https://www.becas-santander.com/es/blog/riesgos-financieros.html>
- BanEcuador. (2022). *Informe de Gestión 2021*. Banca Pública de Desarrollo, Quito, Ecuador. Retrieved from <https://www.banecuador.fin.ec/mdocs-posts/informe-anual-de-gestion-2021/>

- BanEcuador. (2023). *Misión y Visión*. Retrieved from <https://www.banecuador.fin.ec/historia-banecuador/mision-y-vision/>
- BanEcuador B.P. (2023). *Historia BanEcuador B. P.* Obtenido de <https://www.banecuador.fin.ec/historia-banecuador/>
- Barreto, D. (2020). *Desarrollo de un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento microempresa*. Universidad Señor de Sipán, Facultad de Ingeniería, Arquitectura y Urbanismo. Escuela de Ingeniería Económica, Pimentel, Perú. Obtenido de <https://hdl.handle.net/20.500.12802/6687>
- Carvajal, A., & Espinoza, L. (2020). Microcréditos ecuatorianos: incentivo a la reducción de la pobreza y mejora del ingreso familiar. *Revista Vínculos*, 5(3), 51-65. doi:<https://doi.org/10.24133/vinculosespe.v5i3.1671>
- Ceballos, H. (2022). Financial risk in colombian health sector entities. *TECHNO REVIEW. International Technology, Science and Society Review /Revista Internacional De Tecnología, Ciencia Y Sociedad*, 11(2.3), 1–9. doi:<https://doi.org/10.37467/revtechno.v11.4450>
- Coronado, M., & Carabias, S. (2018). Evolución de la medición del riesgo financiero en los últimos 40 años: una panorámica con especial mención en la banca. *Revista De La Facultad De Derecho*(105), 1-37. doi:<https://doi.org/10.14422/icade.i105.y2018.002>
- Espinoza, G. (2020). El Crédito Bancario y las Pymes en Ecuador. *YACHANA Revista Científica*, 9(2), 77-90. Obtenido de <http://revistas.ulvr.edu.ec/index.php/yachana/article/view/649>
- Garrido, N. (2019). *Construcción de un modelo scoring de aprobación para cartera comercial de una institución financiera pública mediante modelos aditivos generalizados*. Escuela Politécnica Nacional, Facultad de Ciencias. Maestría en Estadística Aplicada, Quito, Ecuador. Obtenido de <http://bibdigital.epn.edu.ec/handle/15000/20756>
- Garzón, H. (2020). *Modelo Scoring de Comportamiento de Cartera de Crédito para una Empresa de Servicios Públicos Domiciliarios*. Universidad de Santander, Facultad de Ciencias Económicas, Administrativa y Contables.

- Maestría en Finanzas, Bucaramanga, Colombia. Retrieved from <https://repositorio.udes.edu.co/handle/001/5090>
- Gaytán, J. (2018). Clasificación de los riesgos financieros. *Mercados y Negocios*, 1(38), 123-136. doi:<https://doi.org/10.32870/myn.v0i38.7274>
- Gob.Ec. (2022, julio 01). *Banca para el Desarrollo Productivo Rural y Urbano*. Retrieved from <https://www.gob.ec/banecuador-bp>
- Guevara, C. (2020). *El crédito bancario como factor de crecimiento económico en el Ecuador, un análisis temporal del período post dolarización*. Universidad Técnica de Ambato, Facultad de Contabilidad y Auditoría. Carrera de Economía, Ambato, Ecuador. Retrieved from <https://repositorio.uta.edu.ec/jspui/handle/123456789/30881>
- Junta de Política y Regulación Monetaria y Financiera. (5 de marzo de 2015). Regulación Resolución 043-2015-F Expídense las normas que regulan la segmentación de la Cartera de Crédito de las entidades del Sistema Financiero Nacional. *Registro Oficial 484*. Obtenido de <https://www.oficial.ec/resolucion-043-2015-f-expidense-normas-que-regulan-segmentacion-cartera-credito-entidades-sistema>
- Lara, G. (2021). *Microcréditos BanEcuador y el desarrollo agrícola en el cantón Guaranda*. Universidad Regional Autónoma de los Andes, Facultad de Sistemas Mercantiles. Maestría en Contabilidad y Auditoría, Ambato, Ecuador. Retrieved from <https://dspace.uniandes.edu.ec/handle/123456789/13882>
- Martínez, G. (2019). Tipología del riesgo financiero y normativa bancaria panameña. *Revista Saberes APUDEP*, 2(2), 21-28. Obtenido de https://revistas.up.ac.pa/index.php/saberes_apudep/article/view/827
- Mhlanga, D. (2020). Industry 4.0 in Finance: The Impact of Artificial Intelligence (AI) on Digital Financial Inclusion. *International Journal of Financial Studies*, 8(3), 1-14. doi:<https://doi.org/10.3390/ijfs8030045>
- Moncada, A., & Muñoz, M. (2020). *Aplicación de la metodología del algoritmo K-Nearest Neighbours del modelo Scoring y su incidencia en la calificación del cliente en el otorgamiento del crédito en la Cooperativa de Ahorro y*

- Crédito San Pío X Limitada, Cajamarca, durante el año 2019.* Universidad Privada del Norte, Facultad de Negocios. Carrera de Contabilidad y Finanzas, Cajamarca, Perú. Retrieved from <https://hdl.handle.net/11537/25138>
- Montalván, C. (2019). *Credit scoring, aplicando técnicas de regresión logística y redes neuronales, para una cartera de microcrédito.* Universidad Andina Simón Bolívar. Obtenido de <https://repositorio.uasb.edu.ec/bitstream/10644/6872/1/T2962-MGFARF-Montalvan-Credit.pdf>
- Nevárez, J., Álvarez, V., Sisalema, A., & Nevárez, V. (2021). Valoraciones teóricas y metodológicas en torno a la gestión de riesgo financiero en el sector financiero popular y solidario. *ULEAM Bahía Magazine (UBM)*, 1(2), 33-58. Retrieved from https://revistas.uleam.edu.ec/index.php/uleam_bahia_magazine/article/view/64
- Oquendo, A. (2022). *Diseño de un modelo de credit scoring que fomente la inclusión social al sistema financiero haciendo énfasis en las características del bien a financiar mediante técnicas de inteligencia artificial explicables.* Escuela Politécnica Nacional, Facultad de Ingeniería en Sistemas Informáticos y de Computación, Quito, Ecuador. Obtenido de <http://bibdigital.epn.edu.ec/handle/15000/22553>
- Orellana, I., Reyes, M., & Cevallos, E. (2019). Evolución de los modelos para la medición del riesgo financiero. *UDA AKADEM*, 1(3), 7–34. doi:<https://doi.org/10.33324/udaakadem.v1i3.201>
- Perez, A. (2019). *Análisis del riesgo de la cartera de microcrédito mediante la construcción de un modelo scoring para una cooperativa de ahorro y crédito del segmento 2 de Cuenca.* Universidad del Azuay, Maestría en Auditoría Integral y Gestión de Riesgos Financieros, Cuenca, Ecuador. Obtenido de <http://dspace.uazuay.edu.ec/handle/datos/9383>
- Ponce, C. (2021). La Banca Pública de Desarrollo y su relación con la Economía Popular y Solidaria en Ecuador. *Otra Economía*, 14(26), 111-132. Retrieved from

<https://www.revistaotraeconomia.org/index.php/otraeconomia/article/view/14917>

Redacción El Universo. (2022, diciembre 29). Gobierno alista crédito productivo al 5 % de hasta \$ 20.000 para sector turístico. *El Universo*. Retrieved from <https://www.eluniverso.com/noticias/economia/gobierno-alista-credito-productivo-al-5-de-hasta-20000-para-sector-turistico-nota/>

Rey, L. (2019, noviembre 30). *Pérdida de rendimientos en la inversión: tipos de riesgos*. Retrieved from <https://www.inesem.es/revistadigital/gestion-empresarial/riesgo-financiero/>

Rivera, Y. (2019). *El crédito bancario como mecanismo de financiamiento para la compra de activos aplicando los sistemas de amortización*. Universidad Técnica de Machala, Facultad de Ciencias Empresariales. Carrera de Administración de Empresas, Machala, Ecuador. Obtenido de <http://repositorio.utmachala.edu.ec/handle/48000/13450>

Siguas, A. (2019). *El proceso de otorgamiento de crédito y su relación con la clasificación crediticia del deudor en la empresa de bienes y servicios para el HOGAR S.A.C., Iquitos, periodo - 2018*. Universidad Privada de la Selva Peruana, Facultad de Ciencias Empresariales. Escuela Profesional de Administración, Iquitos, Perú. Retrieved from <http://repositorio.ups.edu.pe/handle/UPS/81>

Skura, E. (17 de septiembre de 2020). *Behavioral Scorecard with Machine Learning Components*. Obtenido de <https://www.finalyse.com/blog/behavioral-scorecard-with-machine-learning-components>

Superintendencia de Bancos del Ecuador. (2019, febrero 05). *Glosario de Términos*. Retrieved from <https://www.superbancos.gob.ec/bancos/glosario-de-terminos/>

Tulcanaza, M. (2021). *Propuesta de un modelo de score de originación para la cartera de consumo de una cooperativa de ahorro y crédito del segmento 3 en el Ecuador*. Universidad Andina Simón Bolívar, Sede Ecuador, Área de Gestión. Maestría en Finanzas y Gestión de Riesgos, Quito, Ecuador. Obtenido de <http://hdl.handle.net/10644/8115>

- Tulcanaza, M., & Noboa, A. (2021). *Propuesta de un modelo de score de originación para la cartera de consumo de una cooperativa de ahorro y crédito del segmento 3 en el Ecuador*. Universidad Andina Simón Bolívar. Obtenido de <https://repositorio.uasb.edu.ec/bitstream/10644/81115/1/T3536-MFGR-Tulcanaza-Propuesta.pdf>
- Vallejo, J., & Ochoa, J. (2019). Evolución del microcredito del sector popular y solidario frente a la banca pública-privada del Ecuador. *ECA Sinergia*, 10(2), 140-150. doi:https://doi.org/10.33936/eca_sinergia.v10i2.1550
- Vera, M., & Bravo, V. (2022). Impacto de la banca pública en el fomento del sector agrícola de Manabí Caso BanEcuador B.P. *593 Digital Publisher CEIT*, 7(4-1), 241-258. doi:<https://doi.org/10.33386/593dp.2022.4-1.1224>
- Vimos, S. (2021). *Implementación de un modelo de scoring en la calificación de la cartera de microcrédito para mitigar el riesgo crediticio en la Cooperativa de Ahorro y Crédito Lucha Campesina*. Escuela Superior Politécnica de Chimborazo, Instituto de Postgrado y Educación Continua. Maestría en Finanzas, Riobamba, Ecuador. Obtenido de <http://dspace.esPOCH.edu.ec/handle/123456789/14939>