



**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

Escuela Académico Profesional de Ingeniería Económica

TESIS

**DESARROLLO DE UN MODELO CREDIT SCORING PARA
PREDECIR EL COMPORTAMIENTO DE PAGOS DE LOS
CLIENTES DEL SEGMENTO MICROEMPRESA**

**PARA OPTAR TÍTULO PROFESIONAL DE INGENIERO
ECONOMISTA**

Autor:

Bach. Barreto Gamarra Dante Ulises

Asesora:

Dra. Campos Díaz Yonira Olinda

Línea de Investigación:

Ingeniería Financiera Empresarial

Pimentel - Perú

2020

**DESARROLLO DE UN MODELO CREDIT SCORING PARA
PREDECIR EL COMPORTAMIENTO DE PAGOS DE LOS
CLIENTES DEL SEGMENTO MICROEMPRESA**

APROBACIÓN DEL JURADO:

Mg. Juan Manuel Raunelli Sander
PRESIDENTE DEL JURADO

Mg. Carlos José Carmona Brenis
SECRETARIO DEL JURADO

Mg. Campos Díaz Yonira Olinda
VOCAL DEL JURADO

ENERO DEL 2020

DEDICATORIA

En primer lugar, a Dios Padre, Jesucristo hijo y Espíritu Santo, por permitirme concluir mi carrera profesional, guiarme por las sendas del amor, esperanza e inteligencia, y sobre todo por la vida y salud que me da día a día para cumplir mis metas.

A mis amados hijos; Dante, Daniel, Daniela Dayara, y mi esposa Junnelly, por ser los motivos de mi crecimiento profesional y laboral.

A mis padres por su amor sin condiciones, sus consejos sabios y oportunos en mi desarrollo personal y profesional.

A mis hermanos Anthony, Franklin, Percy, Macklyn, por su estima y aliento.

Bach. Dante Ulises Barreto Gamarra

AGRADECIMIENTO

A Dios padre celestial, por acompañarme, guiarme, y estar siempre presente en mi caminar cotidiano, por permitirme culminar este trabajo de investigación, muy importante para el desarrollo de mi profesión.

A la Escuela Académica Profesional de Ingeniería Económica de la Universidad Señor de Sipán, a sus docentes académicos, por haber impartido sus conocimientos, experiencias, valores éticos profesionales, para mi continuo crecimiento profesional.

Profundo agradecimiento a mi asesora Dra. Yonira Olinda, Campos Díaz, por su acertado conocimiento para el desarrollo y culminación del presente trabajo de investigación.

El Autor

ÍNDICE

DEDICATORIA.....	III
AGRADECIMIENTO.....	IV
ÍNDICE DE TABLAS.....	VII
RESUMEN.....	VIII
ABSTRACT.....	IX
I. INTRODUCCIÓN.....	1
1.1. Realidad problemática.....	1
1.2. Antecedentes del problema.....	3
1.2.1. A nivel internacional.....	3
1.2.2. A nivel nacional.....	5
1.3. Teorías relacionadas al tema.....	8
1.3.1. Modelo Credit Scoring.....	8
1.3.1.1. Conceptos básicos.....	8
1.3.1.2. La técnica del Credit Scoring: Modelo Logit.....	8
1.3.1.2.1. Especificación del modelo Logit.....	8
1.3.1.2.2. Estimación del efecto marginal de las variables independientes.....	13
1.3.1.2.3. Prueba de significancia estadística individual.....	13
1.3.1.2.4. Interpretación de los signos del modelo.....	14
1.3.1.2.5. Medidas de bondad de ajuste.....	16
1.3.1.2.6. Porcentaje de aciertos estimados en el modelo.....	17
1.3.2. Comportamiento de pagos de los clientes.....	19
1.3.2.1. Conceptos básicos.....	19
1.3.2.2. Riesgo de crédito.....	19
1.3.3. Modelo financiero.....	20
1.4. Formulación del problema.....	21
1.5. Justificación e importancia del estudio.....	21
1.6. Hipótesis.....	22
1.7. Objetivos.....	22
1.7.1. Objetivo general.....	22
1.7.2. Objetivos específicos.....	22
II. MATERIAL Y MÉTODOS.....	23
2.1. Tipo y diseño de investigación.....	23
2.1.1. Tipo de investigación.....	23
2.1.2. Diseño de investigación.....	23
2.2. Población y muestra.....	23
2.2.1. Población.....	23
2.2.2. Muestra.....	23
2.3. Variables y operacionalización.....	23
2.3.1. Variables.....	23
2.3.2. Operacionalización.....	24
2.4. Técnicas e instrumentos de recolección de datos.....	24
2.5. Procedimiento para la recolección de datos.....	24
2.6. Análisis estadístico e interpretación de datos.....	25
2.7. Principios éticos.....	25
2.8. Criterios de rigor científico.....	25

III. RESULTADOS	26
3.1. Resultados en tablas y figuras.....	26
3.1.1. Identificación de las variables que permitan tomar decisiones de otorgamiento de créditos Microempresa.	26
3.1.2. Evaluación de la capacidad del modelo para calificar la admisión de clientes..	30
3.2. Discusión de resultados	33
IV. CONCLUSIONES Y RECOMENDACIONES	36
4.1. Conclusiones	36
4.2. Recomendaciones	37
REFERENCIAS	38
ANEXOS.....	42

ÍNDICE DE TABLAS

Tabla 1 Signos esperados de los estimadores del modelo	15
Tabla 2 Matriz de predicción de aciertos.....	18
Tabla 3 Operacionalización de las variables	24
Tabla 4 Edad de los clientes de la entidad financiera.....	26
Tabla 5 Sexo de los clientes	26
Tabla 6 Número de hijos de los clientes.....	27
Tabla 7 Número de integrantes por familia.....	27
Tabla 8 Experiencia en el negocio de los clientes en meses	28
Tabla 9 Ingreso de los clientes	29
Tabla 10 Número de préstamos de los clientes	29
Tabla 11 Estimación del modelo Logit para predecir el comportamiento de pagos de los clientes del segmento microempresa	31

RESUMEN

El objetivo de esta investigación fue desarrollar un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento Microempresa.

Con respecto a su metodología, se utilizó una investigación explicativa, cuyo diseño utilizado fue no experimental de corte transversal. Por otro lado, la población de estudio compendió las 2000 solicitudes de crédito aprobadas de la microempresa de la agencia San Juan de Lurigancho de una financiera de setiembre de 2017 hasta agosto 2018. De mismo modo, se utilizó un modelo econométrico Logit para verificar la hipótesis. Finalmente, se empleó el análisis documental como principal técnica de recopilación de datos.

Según los resultados del modelo econométrico Logit, se identificó el R^2 de McFadden, o pseudo R^2 (McFadden R-squared) es 0.3454, lo cual implica que las variaciones de la variable comportamiento de pagos de los clientes, es explicado por un 34.54% por las variables independientes. Así mismo, las variables que se incluyeron en el modelo Logit resultaron ser estadísticamente significativas, teniendo el signo esperado. De esta manera, los determinantes del comportamiento de pagos de los clientes son: La edad, género, número de hijos, número de integrantes por familia, experiencia en el negocio, ingreso y número de préstamos.

Por último, se recomienda a la Agencia Financiera de San Juan de Lurigancho incorporar mayor número de variables en la toma de información de los clientes, para optimizar la predicción del modelo credit scoring.

Palabras clave: Modelo scoring, modelo Logit, comportamiento de pagos de los clientes.

ABSTRACT

The objective of this research was to develop a credit scoring model to predict the payment behavior of customers in the Microenterprise segment.

Regarding its methodology, an explanatory investigation was used, the design of which was non-experimental cross-sectional. On the other hand, the study population comprised the 2000 approved credit applications of the microenterprise of the San Juan de Lurigancho agency from a financial company from September 2017 to August 2018. Similarly, a Logit econometric model was used to verify the hypothesis. Finally, documentary analysis was used as the main data collection technique.

According to the results of the Logit econometric model, the R^2 McFadden or Pseudo R^2 (McFadden R-squared) was identified as 0.3454, which implies that the variations in the variable payment behavior of customers, is explained in 34.54% by independent variables. Likewise, the variables included in the Logit model were statistically significant and with the expected sign. In this way, the determinants of customer payment behavior are: Age, gender, number of children, number of members per family, business experience, income and number of loans.

Finally, it is recommended that the Financial Agency of San Juan de Lurigancho incorporate a greater number of variables in the collection of customer information, to optimize the prediction of the credit scoring model.

Keywords: Scoring model, Logit model, customer payment behavior.

I. INTRODUCCIÓN

1.1. Realidad problemática

Según Velandia (2013), la crisis financiera del 2008, se origina por el desarrollo de la famosa burbuja financiera de los Estados Unidos, gran parte de esta burbuja es ocasionado por los créditos inmobiliarios de alto riesgo, este hecho generó pérdidas considerables al sector financiero, conllevando a un problema global en la desaceleración económica, esto dejó en evidencia un problema en la gestión y administración del riesgo. Como resultado surge la necesidad de incorporar en el acuerdo de Basilea III, capital de calidad, esta medida logró generar colchones de liquidez, dando estabilidad y seguridad al sistema financiero (p.21).

A sí mismo manifiesta que la superintendencia financiera de Colombia ha incorporado nuevas normas y políticas con la finalidad, de modificar e incluir nuevos modelos de gestión del riesgo de créditos, teniendo como objetivo contribuir a la estabilización del sistema financiero con una adecuada administración del riesgo. En tal sentido ha implementado el sistema de administración del riesgo de crédito, para medir el riesgo de crédito cada entidad adecua sus modelos internos, esto permite prevenir la probabilidad de incumplimiento, basado en la información histórica, sectorial, garantías, políticas de créditos, cálculo de provisiones, logrando así mitigar la probabilidad de ocurrencia de siniestros (p.22).

Según el reporte de estabilidad financiera, BCRP (2018), el sistema financiero peruano, en los periodos setiembre 2016 - setiembre 2018, registro un incremento en el índice de morosidad, las medianas empresas han evidenciado un crecimiento de la tasa de morosidad de 9.2 por ciento a 11.1 por ciento , al mismo tiempo se observa un incremento de las Mypes de 9.7 por ciento a 9.8 por ciento , si analizamos el comportamiento de las empresas financieras que no pertenecen al grupo de los bancos, se agudiza para el caso de medianas empresas de 10.4 por ciento a 13.8 por ciento , en el caso de las Mypes 8.8 por ciento a 9.2 por ciento .

BCRP (2018), sostiene, que el incremento de la morosidad del segmento Mypes puede ser explicado por que algunas entidades no bancarias, cuentan con un deficiente modelo de concesión de créditos. (p.12) (Ver Anexo N°1).

La compra de Mibanco por parte del holding Credicorp a través de su subsidiaria Financiera Edyficar. Según Custodio (2014), en su columna de Semana económica señala

que entre el 2011 y el 2013, el ROE de Mibanco cayó de 25.45 por ciento a 5.44 por ciento a causa de tres factores fundamentalmente:

La competencia, la oferta de créditos Microempresa había crecido, Mibanco competía con; Credicorp, Scotiabank y el BBVA, que, a través de sus financieras Edyficar, CrediScotia y Financiera Confianza, respectivamente, venían incrementado su oferta de productos microempresa, en el caso de Edyficar venía captando un mayor número de clientes.

El incremento de la morosidad producto del sobreendeudamiento de los clientes pymes. Para el 2013 el ratio del índice de morosidad de Mi Banco fue 5.24 por ciento, sumando a ello la cartera de alto riesgo (créditos vencidos y judiciales) se encontraba en 10.51 por ciento.

Como último factor, los bajos resultados, decrecieron en 49 por ciento entre el 2012 y el 2013, y sumaron S/.35.209 millones, estos resultados son afectados directamente por un lado las provisiones producto de la cartera de alto riesgo en S/. 312,882 millones, por otro lado, los gastos de administración en S/. 481,438 millones, lo que refleja una menor eficiencia operativa. (p.1).

Según reporte de Mibanco (2014), los estados financieros al 31 de diciembre de 2014 y del 2013, junto al dictamen de los auditores independientes por la empresa Paredes, Zaldívar, Burga & Asociados, Sociedad Civil de Responsabilidad Limitada 17 de febrero del 2015, Podemos observar que en el 2014 las pérdidas fueron de -S/.67.570 millones (p.6) (Ver Anexo N°2).

En tal sentido podemos afirmar que el impacto de la cartera atrasada afecta directamente a los resultados de la empresa, por ello es importante desarrollar metodologías de evaluación econométrica para poder mitigar la probabilidad de no pago de los clientes.

1.2. Antecedentes del problema

1.2.1. A nivel internacional.

Evidencia para el caso Mexicano se encuentra en Albarran, Montes y Meza (2015), en su artículo de investigación denominada: *“El riesgo de no pago en una institución microfinanciera del municipio de Tenancingo estado de México 2011-2014”*. El objetivo general de esta investigación fue diseñar un modelo de Credit Scoring para calcular el riesgo de no pago en una institución microfinanciera del municipio de Tenancingo, Estado de México durante 2011-2014, utilizando el modelo de regresión logística. En relación con la metodología, se empleó un estudio de tipo explicativo con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituida por 148 clientes. Así mismo, para el diseño del modelo de Credit Scoring se empleó la técnica de regresión logística con la finalidad de calcular probabilísticamente la posibilidad del no pago de los clientes y como instrumentos de recopilación de datos se utilizaron el cuestionario y la ficha bibliográfica. En cuanto a los resultados, aplicando la regresión logística, se ha diseñado un modelo de calificación estadística capaz de predecir correctamente 98.5 por ciento de los clientes de la cartera de esta institución, por lo que se concluye que el modelo es válido y podría aplicarse en cualquier IMF rural de México.

Para el caso de Ecuador, Oña (2015), en su tesis denominada: *“Desarrollo de un modelo de gestión de riesgo de crédito para instituciones de microfinanzas del Ecuador”*. El objetivo principal es estructurar el modelo de gestión de riesgo de crédito para IMF, mismo que se adapte al contexto del mercado financiero ecuatoriano. En relación a la metodología, se empleó un estudio de tipo explicativo con un diseño no experimental de corte transversal. Así mismo, se utilizó para validar la hipótesis el modelo discriminante y se usó como instrumentos de recopilación de datos la ficha bibliográfica. El análisis entre los resultados obtenidos en la aplicación del modelo de Scoring de crédito versus los datos reales almacenados en la data permite concluir que dicho modelo de Scoring de crédito tiene un considerable poder discriminatorio, con el cual se logró identificar a 576 operaciones que incurrirían en incumplimiento. Adicionalmente, reduce los créditos aprobados a 9,545 de los cuales 196 presentarían

incumplimiento, es decir el 2,05%. Por lo tanto, el modelo de Scoring de crédito de comportamiento minimiza la probabilidad de la selección adversa y se convierte en una herramienta en la mitigación del riesgo de crédito.

Para Colombia, Valencia (2017), en su tesis denominada: *“Modelo Scoring para el otorgamiento de crédito de las pymes”*. El objetivo general de esta investigación fue desarrollar un modelo Scoring para el segmento pyme, permitiendo de forma ágil y adecuada, el análisis de riesgo crediticio y así otorgar créditos con parámetros definidos. En relación a la metodología, se empleó un estudio de tipo cuantitativo con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituida por 662 observaciones. Por otro lado, utiliza el modelo Logit para verificar la hipótesis y se usó como instrumentos de recopilación de datos el cuestionario y la ficha bibliográfica. En cuanto a los resultados; después de haber aplicado el modelo Logit, se identificaron que el plazo al cuadrado, el cupo, ventas netas, activos totales, acierta plus y prueba ácida, son los determinantes para el acceso de las pymes a crédito en Colombia.

Continuando con Ecuador se encuentra en Paucar (2018), en su tesis denominada: *“Desarrollo de un Scoring de microfinanzas para las cooperativas de ahorro y crédito para los segmentos 2A5 Caso COAC textil 14 de marzo”*. El objetivo general de esta investigación fue desarrollar un scoring de crédito para microfinanzas para las Cooperativas de Ahorro y Crédito de los segmentos 2 A 5. Caso COAC textil 14 de marzo. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituido por 2771 observaciones. Así mismo, se usó como técnica de recopilación de información el análisis documental. En cuanto a los resultados; la calificación central de riesgos, edad, tipo de vivienda, nivel de educación y antigüedad de residencia son los determinantes para el acceso de un crédito.

Nuevamente para Colombia, Millán y Caicedo (2018), en su tesis denominada: *“Modelos para otorgamiento y seguimiento en la gestión de riesgo de crédito”*. Esta investigación tuvo como propósito principal emplear las variables que se utilizan con mayor frecuencia en los sistemas de scoring de crédito (calificación y clasificación). En relación a la metodología, se empleó un estudio de tipo explicativo, con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituida por

673 registros de clientes de una entidad financiera en Colombia en un periodo de doce meses entre los años 2014 y 2015. Así mismo, se utilizó para validar la hipótesis los modelos análisis discriminante, el de regresión logística y el de redes neuronales, y se usó como instrumentos de recopilación de datos la ficha bibliográfica. De los resultados obtenidos el modelo con mejor desempeño con respecto a la tasa de aciertos es el de redes neuronales con 86.9%, en comparación con el modelo de regresión logística 81.0% y análisis discriminante 71.6%. Para el modelo logístico las variables determinantes para determinar el incumplimiento de pagos están comprendidas en, genero, edad, actividad económica, estado civil, tipo de vivienda, número de personas a cargo, préstamo, plazo, ingreso, línea de crédito, tasa de interés, garantía.

1.2.2. A nivel nacional.

Medina y Ulfe (2015), en su tesis denominada: *“Modelo de credit scoring para predecir el otorgamiento de crédito personal en una cooperativa de ahorro y crédito”*. Este estudio tuvo como objetivo general construir un modelo de credit scoring que permita predecir el otorgamiento de crédito personal en la cooperativa de ahorro y crédito. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituido por 3000 clientes que forman parte de la cartera de crédito personal de la Cooperativa de Ahorro y Crédito a julio del 2013. Así mismo, se usó como técnica de recopilación de información el análisis documental. En cuanto a los resultados, concluye que el credit scoring propuesto es una herramienta útil en la evaluación del sujeto a manera de sugerencia de aceptación o rechazo de una futura solicitud de crédito, de esta manera se identifica con una mayor eficiencia, aquellos clientes que se les puede otorgar crédito, logrando de esta manera la optimización y automatización del proceso crediticio en la institución, previniendo el sobreendeudamiento e incumplimiento de los clientes.

Navarro (2015), en su tesis titulada: *“Diseño de un modelo de credit scoring aplicado a créditos recurrentes para una caja municipal”*. Este trabajo tuvo como propósito diseñar un modelo matemático de credit scoring aplicable en la dación de créditos para clientes recurrentes de una caja municipal. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Por otro lado, la muestra estuvo constituido por 1500 clientes

sintéticos empleando la metodología de simulación de sistemas, utilizando la teoría de números aleatorios siendo generados según sus distribuciones estadísticas, generando de esta manera datos aleatorios. Así mismo, se usó como técnica de recopilación de información el análisis documental. En cuanto a los resultados, se desarrolló y evaluó el modelo de credit scoring con el modelo matemático logit, permitiendo obtener resultados relativamente aceptables el R^2 de McFadden 0.003529, el porcentaje de predicciones correctas es 56%, estos resultados son justificados por la data utilizada para generar el modelo, ya que en algunos casos la data que representaba los resultados de la variable a evaluar no eran los esperados.

Pantoja (2016), en su tesis titulada: *“Propuesta de un modelo Logit para evaluar el riesgo crediticio en las cajas municipales de ahorro y crédito: caso de la caja municipal de Huancayo, periodo 2011-2015”*. Este estudio tuvo como objetivo principal contribuir a la innovación y por ende a la reducción de los niveles de riesgo de crédito de la CMAC Huancayo, una IMF representativa en el sistema de las cajas municipales, dada su actual metodología crediticia. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Así mismo, se usó como técnica de recopilación de información el análisis documental. El resultado de la investigación muestra la estrategia y metodología, así como etapas necesarias para diseñar un modelo logit y la implementación de esta metodología crediticia avanzada en la CMAC de Huancayo, contribuyéndose a reducir los niveles de riesgo crediticio y el cálculo del ahorro en provisiones exigidas por las altas tasas de morosidad para ejemplo de la administración del riesgo de crédito para otras CMAC que busquen al igual que la CMAC Huancayo, mayor participación y rentabilidad ofreciendo asistencia técnica de calidad.

Sarco (2017), en su tesis titulada: *“Factores que determinan el otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri, 2015”*. Este trabajo tuvo como objetivo principal identificar los factores que determinan significativamente en el momento de aprobar el otorgamiento de crédito de la Financiera Credinka en la ciudad de Ayaviri. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Así mismo, la muestra estuvo conformado por 61 clientes, y se usó como técnica de recopilación de información el análisis documental y la encuesta. En cuanto a los resultados, los

factores influyentes para el otorgamiento de crédito son: el ingreso, gasto, número de hijos, edad, seguro, vivienda y material de la vivienda, el cual explica en un 78.7% su fiabilidad.

Matos (2017), en su tesis titulada: *“El riesgo crediticio hipotecario en las cajas municipales de ahorro y crédito aplicando credit scoring y VAR”*. Este estudio tuvo como propósito demostrar que una mejor calificación crediticia reduce el valor de la pérdida esperada en los préstamos hipotecarios financiados por las Cajas Municipales de Ahorro y Crédito. Con respecto a la metodología, se empleó una investigación explicativa con un diseño no experimental de corte transversal. Así mismo, la muestra estuvo conformado por 8119 trabajadores dependientes con un crédito hipotecario vigente para el mes de agosto del 2016 financiado a través de las Cajas Municipales de Ahorro y Crédito, y se usó como técnica de recopilación de información el análisis documental. En cuanto a los resultados, con el modelo de regresión logística se comprobó que existe relación causalidad entre la calidad crediticia del prestatario de préstamo hipotecario con su perfil y comportamiento evaluado para el año 2016.

Farfán (2019), en su tesis denominada: *“Caracterización de las variables de riesgo en la evaluación de créditos para la determinación de un modelo scoring en la cooperativa CACSA, de la ciudad de Juliaca en el año 2017”*, tiene como objetivo general , caracterizar las variables de riesgo en la evaluación de créditos para la determinación de un modelo scoring para la Cooperativa CACSA de la Ciudad de Juliaca, en el año 2017, los objetivos específicos fueron, identificar las variables de riesgo en la evaluación de créditos en la Cooperativa CACSA de la Ciudad de Juliaca, en el año 2017, además de construir un modelo scoring en la evaluación de créditos en la Cooperativa de CACSA, de la Ciudad de Juliaca, en el año 2017, esta investigación es de tipo descriptiva, no experimental y no longitudinal, se considera una población de estudio de 468 créditos, con 11 variables explicativas inicialmente, finalmente se seleccionaron 7 variables explicativas, monto del préstamo, número del crédito, edad, sector económico, estado civil, tiempo de residencia y tipo de vivienda, el investigador utiliza la técnica de regresión logística binaria, presenta un R^2 de Nagelkerke 53.3 por ciento, el porcentaje correcto de clasificación es 88.20 por ciento, el área debajo de la curva ROC, 61.2 por ciento, con una sensibilidad de 96.2 por ciento y una especificidad de 53.3 por ciento.

1.3. Teorías relacionadas al tema

1.3.1. Modelo Credit Scoring.

1.3.1.1. Conceptos básicos.

Según Hand y Henley (1997), “son métodos estadísticos utilizados para clasificar a los solicitantes de crédito, o incluso a quienes ya son clientes de la entidad evaluadora, entre las clases de riesgo ‘bueno’ y ‘malo’”. (como se citó en Lara, 2010, p. 134).

Según Flores (2012), “trata de identificar, mediante el empleo de técnicas estadísticas, los factores clave asociados a la probabilidad de incumplimiento de los prestatarios, a partir de un conjunto de variables relevantes (generalmente, ratios empresariales)”. (p. 330).

Rayo, Lara, y Camino, (2010), “El credit scoring estima el momento en el que se está llevando a cabo la solicitud, cuál será el comportamiento del crédito hasta su vencimiento”. (p. 91).

Según Lara (2010), “Consiste en clasificar a individuos que solicitan un crédito o préstamo en clientes potencialmente buenos o malos para la entidad prestamista en relación a ciertos datos cuantitativos medibles”. (p. 135).

1.3.1.2. La técnica del Credit Scoring: Modelo Logit.

1.3.1.2.1. Especificación del modelo Logit.

El tipo de modelo escogido para poder estimar la probabilidad de rechazar o aprobar, una solicitud de crédito microempresa en la Financiera es el modelo de elección discreta binaria. Para ello, se asume una distribución logística, en donde la variable dependiente es una variable no observable que posee una función de distribución Bernoulli que toma los valores 0 y 1.

En primer lugar, existen dos enfoques principales para interpretar la utilidad de los modelos de elección discreta.

El primero se refiere a la posibilidad de modelización de una variable latente a través de una función índice. El segundo hace referencia al aspecto de la maximización de la utilidad esperada de la decisión de una alternativa sobre otra.

Por tanto, la probabilidad de rechazar una solicitud microempresa se puede expresar de la siguiente manera:

$$P(Y_i) = p^{Y_i} (1 - p)^{(1-Y_i)}$$

En donde Y_i es una función del conjunto de variables X de los solicitantes y se representa de la siguiente manera:

$$Y_i = \beta_1 + \beta_2 X_1 + u_i \quad (1)$$

En el supuesto de que $E(u_i) = 0$, para obtener estimadores insesgados, obtenemos

$$E(Y_i|X_i) = \beta_1 + \beta_2 X_i \quad (2)$$

Asimismo, u_i es la perturbación aleatoria que tiene una función de distribución conocida y Y_i a pesar de no poder observarse, se entiende que implica una preferencia de una opción, que brinda mayor utilidad, sobre otra.

Según Gujarati y Porter. (2010).

Y_i	Probabilidad
0	$1 - P_i$
1	P_i

Es así, que Y_i sigue la distribución de probabilidades de Bernoulli.

Por la definición de esperanza matemática, obtenemos

$$E(Y_i) = 0(1 - P_i) + 1P_i = P_i \quad (3)$$

Igualemos las ecuaciones (2) y (3)

$$E(Y_i|X_i) = \beta_1 + \beta_2 X_i = P_i \quad (4)$$

La esperanza de una variable aleatoria Bernoulli está dada por la probabilidad de que esa variable sea igual a 1, tengamos en cuenta que, si existen n intentos independientes, cada uno con una probabilidad p de éxito y una probabilidad (1 - p) de fracaso, y X de tales intentos representa el número de éxitos, se dice que X sigue una distribución binomial. La media de la distribución binomial es np, y su varianza, np(1-p). Como la probabilidad P_i debe encontrarse entre 0 y 1, tenemos la restricción. (p. 543).

$$0 \leq E(Y_i|X_i) \leq 1 \quad (5)$$

Según Gujarati y Porter. (2010), Tomando la ecuación (4) donde la matriz X son las

variables de los clientes y $P_i = E(Y = 1_i | X_i)$ significa la probabilidad en caer en default y se consigna como clientes malos. En tal sentido se considera la siguiente representación de clientes malos.

$$P_i = \frac{e^{(\beta_1 + \beta_2 x_1)}}{1 + e^{-(\beta_1 + \beta_2 x_1)}} \quad (6)$$

Por facilidad de simplificar la ecuación (6) escribimos como **función de distribución logística**

Donde $Z_1 = \beta_1 + \beta_2 x_1$

$$P_i = \frac{1}{1 + e^{-z_1}} = \frac{e^z}{1 + e^z} \quad (7)$$

Si P_i , es la probabilidad de caer en default, está dada por (7), entonces $(1 - P_i)$, es la probabilidad que no caiga en default y se consigna como clientes buenos.

$$1 - P_i = \frac{1}{1 + e^{z_1}} \quad (8)$$

Podemos asumir.

$$\frac{P_i}{1 - P_i} = \frac{e^{z_i}}{1 + e^{z_i}} = e^{z_i} \quad (9)$$

Por tanto $P_i / 1 - P_i$ es **la razón de las probabilidades** en favor de que un cliente caiga en default: la razón de la probabilidad de que un cliente caiga en default respecto de la probabilidad de que no caiga en default.

Ahora, si tomamos el logaritmo natural de la ecuación (9), obtendremos un resultado muy interesante, a saber, es decir, L , el logaritmo de la razón de las probabilidades, no es sólo lineal en las variables independientes X , sino también (desde el punto de vista de estimación) lineal en los parámetros.

L se llama logit, y de aquí el nombre modelo logit para el siguiente modelo

$$L_1 = Z_1 = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_1 + \beta_2 X_1 \quad (10)$$

Por tanto;

Tomando de referencia la Ecuación (7), se debe desarrollar un método para estimar

β_1 y β_2 a partir de una muestra de n observaciones, dado que en este caso la variable respuesta es binaria, se usa el método de Máxima Verosimilitud para la estimación de los parámetros. (p. 543).

Estimación de máxima verosimilitud del modelo logit para datos individuales (no agrupados)

Según Gujarati y Porter. (2010), Al estimar la probabilidad de que un cliente sea considerado como cliente que cae en default, cliente malo, dado la matriz de variables X de dicho cliente. Esta probabilidad se expresa mediante la función logística, que a continuación se detalla.

$$P_i = \frac{e^{(\beta_1 + \beta_2 x_1)}}{1 + e^{-(\beta_1 + \beta_2 x_1)}} \quad (11)$$

En realidad, no observamos P_i , sino sólo el resultado ($Y_i = 1$), si un cliente es considerado como moroso o malo, y ($Y_i = 0$) si es considerado como cliente bueno.

Como cada Y_i es una variable aleatoria Bernoulli, se expresa

$$Pr(Y_i = 1) = P_i \quad (12)$$

$$Pr(Y_i = 0) = (1 - P_i) \quad (13)$$

Suponga que tenemos una muestra aleatoria de n observaciones. Sea la función $f_n(Y_i)$ tal que denote la probabilidad de que $Y_i = 1$ o $Y_i = 0$; la probabilidad conjunta de observar los n valores Y_i , es decir, $f(Y_1, Y_2, \dots, Y_n)$, se expresa como:

$$f(Y_1, Y_2, \dots, Y_n) = \prod_{i=1}^n f(Y_i) = \prod_{i=1}^n P_i^{Y_i} (1 - P_i)^{1 - Y_i} \quad (14)$$

Aplicando el operador producto; se observa la función de densidad de probabilidades conjuntas como producto de las funciones de densidad individuales, pues cada Y_i se obtiene de manera independiente y cada Y_i tiene la misma función de densidad (logística). La probabilidad conjunta dada en la ecuación (14) se conoce como **función de verosimilitud (FV)**.

De la ecuación (14). Tomaremos su logaritmo natural, obteniendo de esta manera la **función log de verosimilitud (FLV)**:

$$\begin{aligned}
\ln f(Y_1, Y_2, \dots, Y_n) &= \sum_{i=1}^n [Y_i \ln P_i + (1 - Y_i) \ln(1 - P_i)] \\
&= \sum_{i=1}^n [Y_i \ln P_i - Y_i \ln(1 - P_i) + \ln(1 - P_i)] \quad (15) \\
&= \sum_{i=1}^n \left[Y_i \ln \left(\frac{P_i}{1 - P_i} \right) \right] + \sum_{i=1}^n \ln(1 - P_i)
\end{aligned}$$

De (11) resulta fácil verificar que

$$(1 - P_i) = \frac{1}{1 + e^{(\beta_1 + \beta_2 x_1)}} \quad (16)$$

Así como

$$\ln \left(\frac{P_i}{1 - P_i} \right) = \beta_1 + \beta_2 x_1 \quad (17)$$

Mediante (16) y (17) expresamos la FLV (15) como:

$$\ln f(Y_1, Y_2, \dots, Y_n) = \sum_{i=1}^n Y_i (\beta_1 + \beta_2 x_1) - \sum_{i=1}^n \ln(1 + e^{(\beta_1 + \beta_2 x_1)}) \quad (18)$$

Finalmente, para obtener los estimadores (β) de máxima verosimilitud derivamos $L(\beta)$ con respecto de cada uno de los parámetros β_j con $j= 1, 2, \dots, p$, e igualamos a cero. En términos de matrices

$$\frac{\partial L(\beta)}{\partial \beta} = \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \left(\frac{e^{(\beta_1 + \beta_2 x_1)}}{1 + e^{(\beta_1 + \beta_2 x_1)}} \right) \quad (19)$$

Igualando 19 al vector cero.

$$\sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \left(\frac{e^{(\beta_1 + \beta_2 x_1)}}{1 + e^{(\beta_1 + \beta_2 x_1)}} \right) = \sum_{i=1}^n x_i p_i \quad (20)$$

Si $\hat{\beta}$ es el vector de parámetros que cumple el sistema en términos de matrices, calculamos p_i en términos de esos estimadores y de aquí se obtiene una estimación para y_i , tal que $\hat{y}_i = \hat{p}_i$, así. (p. 590).

$$\sum_{i=1}^n y_i x_{ij} = \sum_{i=1}^n x_{ij} \hat{y}_i$$

de aquí

$$\sum_{i=1}^n x_{ij} e_i = \sum_{i=1}^n x_{ij} (y_i - \hat{y}_i) = 0$$

1.3.1.2.2. Estimación del efecto marginal de las variables independientes

Para el cálculo del efecto marginal de las variables sobre la variable dependiente se pueden considerar tres escenarios. El primero se refiere al efecto del vector de las variables independientes para un individuo representativo de la muestra, para lo cual se toman los valores promedios de cada variable y se reemplaza en la función F y esta función estimada se reemplaza en la función p, la cual tiene la siguiente forma:

$$\hat{p}_i = \frac{e^{\bar{x}_i \hat{\beta}}}{1 + e^{\bar{x}_i \hat{\beta}}}$$

El segundo escenario se presenta cuando la variable explicativa es categórica, por lo que el efecto marginal se obtiene como la derivada de la probabilidad de ocurrencia del evento ($Y_i = 1$) dado un cambio unitario en dicha variable. Expresado formalmente, se tiene:

$$\frac{\partial P[Y_i = 1]}{\partial X_j} = F(X_i \hat{\beta}) \beta_j = \Lambda(X_i \hat{\beta}) \beta_j [1 - \Lambda(X_i \hat{\beta})] \beta_j = \frac{e^{\bar{x}_i \hat{\beta}}}{1 + e^{\bar{x}_i \hat{\beta}}} = \hat{p}_i (1 - \hat{p}_i) \beta_j$$

Finalmente, el tercer escenario corresponde cuando la variable independiente es una dummy que toma valores 0 ó 1; en este caso, el efecto marginal o efecto impacto se calcula mediante la diferencia entre la probabilidad de ocurrencia del evento ($Y_i = 1$) cuando la variable es 1 y la probabilidad de ocurrencia del evento ($Y_i = 1$) cuando la misma variable es 0. La expresión toma la siguiente forma:

$$\frac{\partial P[Y_i = 1]}{\partial X_j} = P(Y_i = 1 | X_j = 1) - P(Y_i = 1 | X_j = 0) = \left. \frac{e^{\bar{x}_i \hat{\beta}}}{1 + e^{\bar{x}_i \hat{\beta}}} \right|_{X_j=1} - \left. \frac{e^{\bar{x}_i \hat{\beta}}}{1 + e^{\bar{x}_i \hat{\beta}}} \right|_{X_j=0}$$

Para las variables categóricas, el cálculo es similar al de las variables dummies, pero se escoge una de las alternativas como referencia para realizar las diferencias y cuantificar el efecto impacto.

1.3.1.2.3. Prueba de significancia estadística individual

Estadístico de Wald Una vez especificado el modelo, es necesario validar que los estimadores de máxima verosimilitud de cada variable independiente son significativos en la explicación de la variable dependiente; por lo que, las hipótesis que se plantean son:

$$\begin{cases} H_0 : \beta_i = 0 \\ H_1 : \beta_i \neq 0 \end{cases}$$

Estos estimadores se distribuyen asintóticamente según una distribución normal; por tanto, para tamaños de muestras relativamente grandes se tiene:

$$\hat{\beta}_i \sim N(\beta_i, var(\hat{\beta}_i))$$

Por ello, puede contrastarse la hipótesis nula sobre la significancia de una variable mediante:

$$Z = \frac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}}$$

Es así que, si el estadístico z calculado es mayor al z tabular, se rechaza la hipótesis nula y la variable justifica su presencia en el modelo.

1.3.1.2.4. Interpretación de los signos del modelo

En los modelos de elección discreta, los estimadores no se interpretan como la magnitud del impacto de dicha variable sobre el evento, ya que la probabilidad p es una función F de estos valores. Es por este motivo que sólo es posible hacer una interpretación de los signos, los cuales indican la dirección de la influencia de las variables a los que están asociados.

En este caso, un signo positivo indica que, si la variable independiente se incrementa, la probabilidad de rechazar a un solicitante se incrementa; en tanto que, un signo negativo refleja que cambios en dicha variable favorece la probabilidad de aceptar al solicitante.

A continuación se presenta en la Tabla 1, los signos esperados de las variables independientes en la estimación econométrica del modelo Logit.

Tabla 1
Signos esperados de los estimadores del modelo

Grupo	Variable	Definición	Escala	Concepto	Signo esperado
Social	EDAD	Edad	Años	Desde 22 a más	-
	GENERO	Genero	Unidades	Masculino=0, Femenino=1	+/-
	NUM_HIJ	Numero de hijo	Unidades	No tiene hijos =0, Menor a 18 años =1, mayor a 18 años = 2	-
	NUM_FAMIL	Numero de familiares	Unidades	Desde 0 a mas	+
Económico	EXP_NEG	Experiencia en el negocio	Meses	Desde 6 a mas	-
	ING	Ingreso	Soles	Desde 950 a mas	+
Interno Financiera	NUM_PREST	Numero Prestamos	Unidades	Desde 0 a mas	+

Nota. Elaboración propia.

1.3.1.2.5. Medidas de bondad de ajuste

El ajuste de un modelo mide la capacidad del modelo para explicar la variable dependiente. Esto es medido por el R^2 en los modelos tradicionales lineales; sin embargo, dado que la variable a estimar de este modelo no es observable, la medida a emplear será el pseudo R^2 de McFadden, que se basa en la función de verosimilitud. El supuesto detrás de este indicador es que un modelo estimado sólo por la constante equivale a un modelo con el peor ajuste posible.

El estadístico se calcula mediante la siguiente expresión de acuerdo con Greene (2003) (como se citó en Valdivia et.al, 2009, p. 232).

$$pseudo R^2 = 1 - \frac{LnL}{LnL_0}$$

Medina (2003), En donde, LnL es la función de verosimilitud para un modelo sin restringir, que en este caso es el modelo propuesto y LnL_0 la misma función con la restricción de que todos los parámetros sean nulos.

El ratio calculado tendrá valores entre los rangos comprendidos de 0 y 1.

Los valores próximos a 0 se obtendrán cuando L_0 sea muy parecido a L , en este punto las variables incluidas en el modelo son poco significativas, es decir. La estimación de los parámetros β no mejora el error que se comete si dichos parámetros se igualan a 0. Por lo que en este caso la capacidad explicativa del modelo será muy reducida.

Cuanto mayor sea la capacidad explicativa del modelo, mayor será el valor de L sobre el valor de L_0 , debido a que los logaritmos de la función de log-verosimilitud son siempre negativos -, y más se aproxima al ratio de verosimilitud calculado al valor 1. (p. 590).

De acuerdo con lo recomendado por McFadden (1973), el valor del *pseudo* R^2 debe encontrarse dentro del intervalo entre 0.2 y 0.4 para considerar un ajuste relativamente alto del modelo, sin embargo, otros autores como Bateman et.al. (2002), indican que debido a que no existe un acuerdo sobre el valor de este estadístico que exprese un buen ajuste del

modelo, un resultado por encima de 0.1 reflejara un ajuste satisfactorio (como se citó en Valdivia et.al, 2009, p. 233).

1.3.1.2.6. Porcentaje de aciertos estimados en el modelo

Medina (2003), una de las vías utilizadas para determinar la bondad de modelo logit es predecir con el modelo los valores de la variable endógena Y_i de tal manera que:

$$\hat{y}_i \begin{cases} 1, si F(X_i \beta) > c \\ 0, si F(X_i \beta) \leq c \end{cases}$$

Generalmente, el valor que se asigna a c para determinar si el valor de la predicción es igual a 1 o 0 es de 0.5, puesto que parece lógico que la predicción sea 1 cuando el modelo dice que es más probable obtener un 1 que un 0. Sin embargo, la elección de un umbral igual a 0.5 no siempre es la mejor alternativa. En el caso que la muestra presente desequilibrios entre el número de unos y el de ceros la elección de un umbral igual a 0.5 podría conducir a no predecir ningún 1 o ningún 0.

Además, con cualquier tipo de regla predictiva similar se cometerán dos errores: habrá ceros que se clasifiquen incorrectamente como unos y viceversa. Si se reduce el umbral por debajo de 0.5 aumentará el número de veces que se clasifican correctamente observaciones para las que $Y_i = 1$, pero también aumentará el número de veces en que se clasifiquen observaciones como unos para las que $Y_i = 0$. Por lo que, cambios en el valor del umbral reducirá siempre la probabilidad de un error de un tipo y se aumentará la probabilidad del otro tipo de error.

Es así como, el valor que debe tomar el umbral depende de la distribución de datos en la muestra. Una vez seleccionado el nivel del umbral, y dado que los valores reales de Y_i ; son conocidos, basta con contabilizar el porcentaje de aciertos para decir si la bondad del ajuste es elevada o no. A partir de este recuento se puede construir el Cuadro 1 de predicción:

Tabla 2
Matriz de predicción de aciertos

		Pronóstico del modelo		Total
		$\hat{Y}_{=0}$	$\hat{Y}_{=1}$	
Valores Observados	$Y_i = 0$	$N_{0\hat{0}}$	$N_{0\hat{1}}$	N_0
	$Y_i = 1$	$N_{1\hat{0}}$	$N_{1\hat{1}}$	N_1
Total		$N_{\hat{0}}$	$N_{\hat{1}}$	N

Nota. Elaboracion propia.

En la matriz $N_{0\hat{0}}$ y $N_{1\hat{1}}$ corresponden a las predicciones correctas, mientras que $N_{0\hat{1}}$ y $N_{1\hat{0}}$ corresponden a las predicciones erróneas (valores 1 mal predichos en el primer caso y valores 0 mal predichos en el segundo caso). (p. 590).

Con los valores de la matriz de clasificación se pueden construir los siguientes índices:

Tasa de aciertos: Cociente entre las predicciones correctas y el total de predicciones.

$$TA = \frac{N_{0\hat{0}} + N_{1\hat{1}}}{N}$$

Tasa de errores: Cociente entre las predicciones incorrectas y el total de predicciones.

$$TE = \frac{N_{1\hat{0}} + N_{0\hat{1}}}{N}$$

Especificidad: Proporción entre la frecuencia de valores 0 correctos y el total de valores 0 observados.

$$E = \frac{N_{0\hat{0}}}{N_0}$$

Sensibilidad: Razón entre los valores 1 correctos y el total de valores 1 observados.

$$S = \frac{N_{1\hat{1}}}{N_1}$$

Tasa de falsos ceros: Proporción entre la frecuencia de valores 0 incorrectos y el total de valores 0 observados.

$$TF = \frac{N_{0\hat{1}}}{N_0}$$

Tasa de falsos unos: Proporción entre la frecuencia de valores 1 incorrectos y el total de valores 1 observados.

$$TF_1 = \frac{N_{1\hat{0}}}{N_1}$$

1.3.2. Comportamiento de pagos de los clientes

1.3.2.1. Conceptos básicos

Según Rayo, Lara, y Camino (2010), manifiestan que el comportamiento de pago o reembolso se puede medir mediante una puntuación que mide el riesgo de un prestatario y/o de la operación (p.91). Adema detallan que al analizar el comportamiento de pagos de un cliente de microcrédito se requiere de una base de datos amplia que recoja el historial de los préstamos que resultaron impagos (p.96).

Según Laitinen (1999), afirma que el comportamiento previo al pago se centra en variables como retrasos en los pagos o alteraciones en los pagos, especialmente con respecto a la probabilidad de dificultades financieras en el futuro (como se citó en Back, 2005, p.844).

1.3.2.2. Riesgo de crédito

Viene hacer la posibilidad de que no se materialice la esperanza de obtener la devolución de una determinada cantidad de dinero que se prestó en el pasado, en una fecha concreta. (Bessis, 2002, p. 121).

Se define también como la pérdida asociada al evento de incumplimiento del prestatario o bien al evento del deterioro de su calidad crediticia. (Peña, 2002, p. 121).

El riesgo de crédito es la pérdida potencial que se registra con motivo del incumplimiento de una contraparte en una transacción financiera, o en alguno de los términos y condiciones de la transacción. (De Lara, 2003, p. 163).

Se entiende por Riesgo de crédito el riesgo derivado de cambios en la calificación crediticia del emisor derivado de la probabilidad de incurrir en pérdidas derivadas del impago en tiempo o forma de las obligaciones crediticias de uno o varios clientes. (Lara, 2010, p. 132).

Según Flores (2012), “El riesgo de crédito constituye el componente más importante de los riesgos a los que se enfrentan las entidades financieras, tanto por lo que respecta a su

volumen como a sus consecuencias para el futuro de la organización. No obstante, las características de este tipo de riesgo se han modificado sustancialmente en las últimas décadas, resultando cada vez más complejo y difícil de cuantificar, particularmente respecto a las posiciones frente a empresas” (p. 1096).

Podríamos definir como la perdida esperada potencial frente al incumplimiento del deudor, de acuerdo a al deterioro de su calificación crediticia.

1.3.3. Modelo financiero

Partiendo del previo análisis teórico - estructural, el modelo financiero a probar con sentido financiero y para fines de nuestra investigación es el siguiente:

$$CC_i = \beta_1 + \beta_2(\text{edad}_{2i}) + \beta_3(\text{genero}_{3i}) + \beta_4(\text{num_hij}_{4i}) + \beta_5(\text{num_fam}_{5i}) \\ + \beta_6(\text{exp_neg}_{6i}) + \beta_7(\text{ing}_{7i}) + \beta_8(\text{num_pres}_{8i})$$

Donde:

- CC_i : Comportamiento de pagos de los clientes, donde toma valores 0 = clientes buenos y 1= clientes malos.
- $Edad_{2i}$: Edad de los clientes, desde 22 a más
- $genero_{3i}$: Genero del cliente, donde toma valores de Masculino=0, Femenino=1
- num_hij_{4i} : Número de hijo de los clientes, donde va a tomar valores de no tiene hijos =0, Menor a 18 años =1, mayor a 18 años = 2
- num_fam_{5i} : Número de integrantes por familiar, donde va a tomar valores desde 0 a más.
- exp_neg_{6i} : Experiencia en el negocio, donde va a tomar valores desde 6 meses a más.
- ing_{7i} : Ingreso de los clientes, donde va a tomar valores desde S/ 950 a más.
- num_pres_{8i} : Representa el número préstamos de los clientes, donde va a tomar valores desde 0 a más.

Según Aglietta (2000), para el desarrollo de un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento microempresa son la edad, genero,

número de hijos, números de integrantes por familia, experiencia en el negocio, ingreso y número de préstamos.

Por otro lado, Aglietta (2000), afirma que los signos del modelo financiero son:

La variable edad es negativo, lo cual indica que un año adicional de los clientes, disminuye la probabilidad de incumplimiento de ser un cliente etiquetado como malo. Por otra parte, la variable genero su coeficiente es negativo, lo cual indica que ante un cambio de género masculino a género femenino, la probabilidad de incumplimiento de ser un cliente etiquetado como malo, disminuye. Así mismo, la variable número de hijos tiene un coeficiente positivo al incrementarse el número de hijos, la probabilidad de incumplimiento aumenta. Por otro lado, la variable número de integrantes por familia tiene signo positivo indica que el riesgo aumenta conforme aumenta en una persona en el hogar. En cuanto, a la variable experiencia en el negocio tiene signo negativo, esto indica que el riesgo disminuye conforme aumenta una unidad de meses de experiencia en el negocio. Con respecto la variable ingreso, tiene signo negativo y esto indica que el riesgo disminuye conforme aumenta los ingresos en una unidad de soles. Finalmente, la variable número de préstamos tiene signo positivo y esto indica que el riesgo aumenta conforme aumenta los números de préstamos.

1.4. Formulación del problema

¿Es posible desarrollar un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento Microempresa?

1.5. Justificación e importancia del estudio

Esta investigación es importante debido a los siguientes aspectos:

Justificación Teórica: El presente trabajo de investigación se basará en el valor teórico del modelo credit scoring, en la cual se podrá identificar con mayor facilidad a los clientes con buen comportamiento de pagos, y por consiguiente una mejor calidad de cartera, con ello se tiene una ventaja competitiva.

Justificación Práctica: La implicancia práctica del estudio, infiere que los resultados de la investigación pueden ser aplicados en el uso diario por los funcionarios, analistas, asesores de crédito ligados al segmento de créditos microempresa, siendo una herramienta útil para

aumentar el mayor número de operaciones de créditos, con riesgo bajo de probabilidad de impago.

Justificación metodológica: Para lograr los objetivos de la investigación se utilizan técnicas de investigación científica fundamentados en antecedentes estadísticos, modelo econométrico Logit, con una revisión bibliografía teórica que nos permitan refrendar los resultados.

1.6. Hipótesis

La hipótesis de la presente investigación es:

Hipótesis nula: El modelo credit scoring no se puede predecir el comportamiento de pagos de los clientes del segmento Microempresa.

Hipótesis alternativa: El modelo credit scoring puede predecir el comportamiento de pagos de los clientes del segmento Microempresa.

1.7. Objetivos

1.7.1. Objetivo general

Desarrollar un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento Microempresa.

1.7.2. Objetivos específicos

1. Identificar un conjunto de variables que permitan tomar decisiones de otorgamiento de créditos Microempresa.
2. Evaluar la capacidad del modelo para calificar la admisión de clientes.

II. MATERIAL Y MÉTODOS

2.1. Tipo y diseño de investigación

2.1.1. Tipo de investigación.

La presente investigación es de tipo explicativa porque se diseñó mediante un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento microempresa.

2.1.2. Diseño de investigación.

El presente estudio es una investigación que tiene un diseño no experimental de corte transversal, porque la variable dependiente (Comportamiento de pagos de los clientes) y la variable independiente (Modelo Credit Scoring) son analizadas en su contexto real, sin ser manipulados.

2.2. Población y muestra

2.2.1. Población.

La población de estudio compendió los datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018, contando con 2000 solicitudes de crédito aprobadas (Ver Anexo N°10).

2.2.2. Muestra.

Por ser un estudio de econometría de corte transversal, la muestra es la misma de la población, o sea comprende las 2000 solicitudes de crédito aprobadas de la microempresa de la agencia San Juan de Lurigancho de una financiera de setiembre de 2017 hasta agosto 2018 (Ver Anexo N°10).

2.3. Variables y operacionalización

2.3.1. Variables.

Las variables del presente estudio son.

Variable dependiente: Comportamiento de pagos de los clientes

Variables independientes: Modelo Credit Scoring

2.3.2. Operacionalización.

A continuación, se presenta en la Tabla 3, la operacionalización de las variables.

Tabla 3
Operacionalización de las variables

Variabes	Dimensión	Indicadores	Técnicas e instrumentos de recolección de datos
Comportamiento de pagos de los clientes (Variable dependiente)	Calidad Crediticia	0 = clientes buenos	Análisis documental/ Reporte estadístico
		1= clientes malos	
Modelo Credit Scoring (Variable independiente)	Social	Edad	Análisis documental/ Reporte estadístico
		Genero	
	Económico	Nº hijos	Análisis documental/ Reporte estadístico
		Nº Familiares	
Interno Financiera	Experiencia Negocio	Análisis documental/ Reporte estadístico	
	Ingreso		
		Numero Prestamos	Análisis documental/ Reporte estadístico

Nota. Elaboración Propia.

2.4. Técnicas e instrumentos de recolección de datos

La técnica que se ha utilizado en la presente investigación es el análisis documental. Se usó para adquirir, el marco teórico y la información estadística con el fin de sustentar el objeto de estudio, además del diseño del informe de tesis, el análisis e interpretación de los resultados obtenidos.

Por otro lado, el instrumento de recolección de datos que se empleó son los datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018, contando con 2000 solicitudes de crédito aprobadas.

2.5. Procedimiento para la recolección de datos

A continuación, se detalla el procedimiento de recopilación de datos.

- a) Recopilación de los datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.
- b) Se tabuló los datos recolectados mediante la utilización del programa Excel versión 13 y se interpretó los resultados.
- c) Se estimó el modelo econométrico Logit y su interpretación utilizando el programa econométrico Stata versión 13.

2.6. Análisis estadístico e interpretación de datos

Se ha utilizado la estadística descriptiva, para analizar e interpretar los datos, utilizando las tablas y figuras. Se utilizó la hoja de cálculo excel versión 13 como también el procesador de textos word versión 13, para procesar la información, para efectuar la estimación del modelo econométrico Logit se utilizó el programa econométrico Stata versión 13.

2.7. Principios éticos

Los principios éticos en la presente investigación son:

- a) **Confidencialidad.** Se asegurará la protección de la identidad de financiera que participa como informante de la investigación.
- b) **Originalidad.** Se citarán las fuentes bibliográficas de la información mostrada, a fin de demostrar la inexistencia de plagio intelectual.

2.8. Criterios de rigor científico

Los criterios de científicidad sobre los cuales se sustenta en el estudio son los siguientes:

- a) **Juicio crítico:** Toda información y material a presentar antes de ser ingresado a nuestro trabajo de investigación, fue validada por tres especialistas (Ver Anexo N°07, Anexo N°08 y Anexo N°09).
- b) **Validez:** La información mostrada es verdadera, ya que toda la información de datos estadísticos es proporcionada por la agencia San Juan de Lurigancho de una financiera.

III. RESULTADOS

3.1. Resultados en tablas y figuras

3.1.1. Identificación de las variables que permitan tomar decisiones de otorgamiento de créditos Microempresa.

En las variables demográficas; Edad se observa que los clientes de la entidad financiera en donde el 27.0% tienen edades de 22 a 31 años; el 36.0% tienen edades de 32 a 41 años; el 24% tienen edades de 42 a 51 años y el 9% tienen edades de 52 a 61 años. Significa que los clientes de entidad financiera tienen mayor edades entre 32 a 41 años (Ver Tabla 4).

Tabla 4
Edad de los clientes de la entidad financiera

Rangos	Frecuencia	Porcentaje
22 -31	532	27
32 -41	721	36
42 - 51	488	24
52 - 61	171	9
62 - 71	63	3
72 - 83	25	1
Total	2000	100

Nota. Elaboración Propia

En el variable demográfico sexo, se observa que los clientes de la entidad financiera en donde el 58.2% son de sexo masculino y el 41.8% son de sexo femenino. Significa que los clientes de la entidad financiera están conformados más por el sexo masculino (Ver Tabla 5).

Tabla 5
Sexo de los clientes

Rangos	Frecuencia	Porcentaje
Masculino	1164	58.2
Femenino	836	41.8
Total	2000	100.0

Nota. Elaboración Propia.

Por otro lado, como se puede observar en la Tabla 6, el 52.9% de los clientes tienen hijos menores de 18 años, el 24.5% tienen hijos mayores a 18 años y sólo el 22.7% de los clientes no tienen hijos (Ver Tabla 6).

Tabla 6
Número de hijos de los clientes

Nº de hijos	Frecuencia	Porcentaje
No tiene hijos	454	22.7
Menor a 18 años	1057	52.9
Mayor a 18 años	489	24.5
Total	2000	100.0

Nota. Elaboración Propia.

Así mismo, en la Tabla 7, se puede observar que el 23.1% de los clientes tienen 2 miembros en sus hogares, el 18.3% de los clientes tienen 3 miembros y sólo un 8.0% de los clientes tienen un solo miembro en sus hogares.

Tabla 7
Número de integrantes por familia

Nº Integrantes	Frecuencia	Porcentaje
1	159	8.0
2	462	23.1
3	365	18.3
4	314	15.7
5	304	15.2
6	309	15.5
7	87	4.4
Total	2000	100.0

Nota. Elaboración Propia.

En lo que respecta a la experiencia del negocio de los clientes, como se puede apreciar en la Tabla 8, el 30.4% de los clientes tienen experiencia hasta 36 meses en el negocio de créditos y sólo un 1.8% de los clientes tienen entre 6 y 7 meses de experiencia.

Tabla 8
Experiencia en el negocio de los clientes en meses

Experiencia en el negocio	Frecuencia	Porcentaje
6	35	1.8
7	35	1.8
8	23	1.2
9	47	2.4
10	30	1.5
11	154	7.7
12	195	9.8
13	32	1.6
14	9	0.5
15	4	0.2
16	4	0.2
18	9	0.5
19	13	0.7
20	21	1.1
21	20	1.0
22	31	1.6
23	27	1.4
24	22	1.1
25	9	0.5
26	65	3.3
27	6	0.3
28	45	2.3
29	15	0.8
30	57	2.9
31	16	0.8
32	62	3.1
33	7	0.4
34	30	1.5
35	73	3.7
36	607	30.4
37	47	2.4
38	17	0.9
39	33	1.7
40	25	1.3
41	13	0.7
42	15	0.8
43	33	1.7
44	11	0.6
46	43	2.2
47	44	2.2
48	15	0.8
112	1	0.1
Total	2000	100.0

Nota. Elaboración Propia.

Como se puede observar en la Tabla 9, el 36% de los ingreso de los clientes de la entidad financiera fluctúan entre 950 a 3026 soles, el 20% de los ingresos están entre 3029 a 3568 soles.

Tabla 9
Ingreso de los clientes

Rangos	Frecuencia	Porcentaje
950 - 3026	725	36
3029 - 3568	392	20
3572 - 4525	312	16
4527 - 5119	287	14
5126 - 6740	284	14
Total	2000	100

Nota. Elaboración Propia.

Finalmente, como se puede observar en la Tabla 10, el 40.5% de los clientes tienen cero préstamos, el 30.7% de los clientes tienen 1 préstamo, el 13.8% de los clientes tienen 2 préstamos.

Tabla 10
Número de préstamos de los clientes

N° Prestamos	Frecuencia	Porcentaje
cero préstamo	810	40.5
1 préstamo	613	30.7
2 préstamo	275	13.8
3 prestamos	165	8.3
4 préstamo	77	3.9
5 préstamo	29	1.5
6 préstamo	10	0.5
7 préstamo	21	1.1
Total	2000	100.0

Nota. Elaboración Propia.

3.1.2. Evaluación de la capacidad del modelo para calificar la admisión de clientes

De acuerdo a los resultados que se pueden observar en la Tabla 11, los coeficientes del modelo logit estimado, representan la relación directa o indirecta de la variable dependiente, comportamiento de pagos de los clientes con respecto a las variables independientes, modelo credit scoring, estos coeficientes no expresan las probabilidades del desarrollo del evento o suceso. Los coeficientes del modelo logit no son directamente interpretables, debido a que el modelo logit no es lineal. En tal sentido el análisis de los resultados se desarrolló en los efectos marginales de las variables independientes, modelo credit scoring, con respecto a la probabilidad de ocurrencia de la variable dependiente, comportamiento de pago de los clientes (buenos o malos), en tal sentido la razón de ventajas o probabilidad (odds ratio), se convertirán como semielasticidades, para ello se calculan las derivadas parciales cuyo cambio en las variables independientes de la matriz X_i , ocurre a partir de un valor medio, siendo estas interpretadas en términos del cambio proporcional, es decir ya sea en incremento o disminución.

El modelo Logit tiene una distribución logística. Así mismo, este modelo garantiza que el rango de probabilidad se encuentre entre 0 y 1, además que las variables explicativas modelo credit scoring, no se encuentren relacionadas linealmente. La Tabla 1 se observan los resultados del modelo Logit. De mismo modo, el R^2 de McFadden, o pseudo R^2 (McFadden R-squared), del modelo logit fue 0.3454, lo cual implica que las variaciones de la variable comportamiento de pagos de los clientes, es explicado en un 34.54% por las variables explicativas, modelo credit scoring.

De mismo modo, el valor obtenido de la razón de verosimilitud, la prueba de LR Chi^2 (07) fue 654.36 del modelo (Ver Anexo N°03), en tal sentido los coeficientes del modelo logit en forma global son significativos en forma conjunta. Por otro lado, La probabilidad de aceptar la hipótesis nula es 0.000, por lo tanto, aceptamos la hipótesis alterna, esto significa que los datos de las variables del modelo se ajustan para interpretar al modelo de regresión logística. El porcentaje de aciertos correctos del modelo consiste en observar el porcentaje de veces en el que el valor de la variable dependiente coincide con su predicción, es decir: el modelo Logit predice adecuadamente el 83.65% de las

observaciones, lo cual nos muestra también que existe un buen ajuste del modelo (Ver Anexo N°05).

Tabla 11
Estimación del modelo Logit para predecir el comportamiento de pagos de los clientes del segmento microempresa

Variables independientes	Modelo Logit			
	Estimación		Efectos marginales	
	Estimador	P-valué	%	P-valué
constante	4.7988	0.000		
edad	-0.1378	0.000	-0.98%	0.000
genero	-1.2444	0.000	-8.41%	0.000
num_hij	1.3954	0.000	9.90%	0.000
num_fam	0.1287	0.004	0.91%	0.004
exp_neg	-0.0106	0.068	-0.08%	0.071
Ing	-0.0009	0.000	-0.01%	0.000
num_prest	0.1559	0.002	1.11%	0.002

Evaluación del modelo econométrico Logit	
Teoría Económica	Estadística
Los signos son correctos de acuerdo a la teoría económica del modelo credit scoring.	Prueba de significancia estadística individual: Todos los parámetros estimados son estadísticamente significativos al 1% y con un nivel de confianza de 99%, a excepción del parámetro experiencia del negocio (exp_neg) que es estadísticamente significativo al 10% y un nivel de confianza de 90%.
	Prueba de significancia estadística global: A nivel global todos los parámetros son estadísticamente significativos al 1% y un nivel de confianza de 99%, ya que [Prob > chi2 = 0.0000](Ver Anexo N°01)
	Medida de bondad de ajuste: El valor de la bondad de ajuste es 0.3454 [Pseudo R2 = 0.3454], que significa que el 34.54% de las variables independientes explican el comportamiento de pagos de los clientes (Ver Anexo N°03).
	Porcentaje de aciertos estimados en el modelo: El criterio del porcentaje de predicciones correctas del modelo, consiste en observar el porcentaje de veces en el que el valor de la variable dependiente coincide con su predicción, es decir: el modelo Logit predice adecuadamente el 83.65% de las observaciones, lo cual nos muestra también que existe un buen ajuste del modelo (Ver Anexo N°05).
	Valor de la Curva ROC: El ajuste del modelo a los datos es adecuado, se observa en la curva ROC cubriendo un área de 89.09% (Ver Anexo N°06).

Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.
 Elaboración propia con Stata versión 13.

De acuerdo a los efectos marginales estimados por el modelo Logit, se concluye lo siguiente: La variable edad es negativo con un nivel de 99% de confianza, lo cual indica que un año adicional de los clientes, disminuye la probabilidad de incumplimiento de ser un cliente etiquetado como malo en un 0.98%. Por otra parte, la variable genero su coeficiente es negativo y estadísticamente significativo al 1% y un nivel de confianza de

99%, lo cual indica que ante un cambio de género masculino a género femenino, la probabilidad de incumplimiento de ser un cliente etiquetado como malo, disminuye en un 8.41%.

Así mismo, la variable número de hijos tiene un coeficiente positivo y es estadísticamente significativo al 1% y con un nivel de confianza de 99%. Esta variable fue clasificada como categórica donde se codificó para aquellos clientes que no tienen hijo (0), hijos menor a 18 años (1), tiene hijos mayores a 18 años (2), ante un cambio de clientes que no tienen hijos a clientes con hijos menores a 18 años, el logaritmo de probabilidad de incumplimiento de ser un cliente etiquetado como malo, se incrementa en 9.90%. Por otro lado, la variable número de integrantes por familia es estadísticamente significativo al 1% con un 99% de confianza. El signo positivo indica que el riesgo aumenta conforme aumenta en una persona en el hogar, por lo tanto, el logaritmo de probabilidad de incumplimiento de ser un cliente etiquetado como malo, aumenta en 0.91% conforme se incrementa una persona en el hogar.

En cuanto, a la variable experiencia en el negocio (*exp_neg*) tiene signo negativo, y es estadísticamente significativo al 10% con un 90% de confianza. El signo negativo indica que el riesgo disminuye conforme aumenta una unidad de meses de experiencia en el negocio, por lo tanto, el logaritmo de probabilidad de incumplimiento de ser un cliente etiquetado como malo disminuye en 0.08%, conforme se incrementa en una unidad de meses la experiencia en el negocio.

Con respecto la variable ingreso, estadísticamente es significativo al 1% con un 99% de confianza. Este signo negativo indica que el riesgo disminuye conforme aumenta los ingresos en una unidad de soles, por lo tanto, el logaritmo de probabilidad de incumplimiento de ser un cliente etiquetado como malo, disminuye en 0.01%, conforme se incrementa los ingresos en una unidad de soles. Finalmente, la variable número de préstamos (*num_prest*), estadísticamente es significativo al 1% con un 99% de confianza. Este signo positivo indica que el riesgo aumenta conforme aumenta los números de préstamos, por lo tanto, el logaritmo de probabilidad de incumplimiento de ser un cliente etiquetado como malo, aumenta en 1.11%, conforme se incrementa los préstamos.

3.2. Discusión de resultados

El objetivo general de esta investigación fue desarrollar un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento Microempresa, donde la variable a explicar es el comportamiento de pagos de los clientes microempresa y las variables explicativas o también llamadas independientes son la edad, género, número de hijos, número de integrantes por familia, experiencia en el negocio, ingreso y número de préstamos.

Por otro lado, los resultados obtenidos a través de la estimación de modelo probabilístico Logit, se identificó el R cuadrado McFadden o Pseudo R2 (McFadden R-squared) es 0.3454, lo cual implica que las variaciones de la variable comportamiento de pagos de los clientes, es explicado en un 34.54% por las variables independientes.

Las variables incluidas en el modelo Logit resultaron ser estadísticamente significativas y con el signo esperado. De esta manera, los determinantes del comportamiento de pagos de los clientes son: La edad, género, número de hijos, número de integrantes por familia, experiencia en el negocio, ingreso real y número de préstamos. De esta manera, se rechazó la hipótesis nula y se aceptó la hipótesis alternativa, donde el modelo credit scoring puede predecir el comportamiento de pagos de los clientes del segmento Microempresa.

Con respecto al objetivo específico 1: Identificar un conjunto de variables que permitan tomar decisiones de otorgamiento de créditos Microempresa. En el presente estudio se obtuvo que los clientes de la entidad financiera el 36.0% tienen edades de 32 a 41 años. Con respecto al sexo, el 58.2% son de sexo masculino y el 41.8% son de sexo femenino. Así mismo, el 52.9% de los clientes tienen hijos menores de 18 años, el 24.5% tienen hijos mayores a 18 años y sólo el 22.7% de los clientes no tienen hijos. Por otro lado, el 23.1% de los clientes tienen 2 miembros en sus hogares, el 18.3% de los clientes tienen 3 miembros y sólo un 8.0% de los clientes tienen un solo miembro en sus hogares. De mismo modo, el 30.4% de los clientes tienen experiencia hasta 36 meses en el negocio de créditos y por último, el 40.5% de los clientes tienen cero préstamos. Resultados similares encontramos para Ecuador, Paucar (2018), en su tesis denominada: “Desarrollo de un Scoring de microfinanzas para las cooperativas de ahorro y crédito para los segmentos 2A5 Caso COAC textil 14 de marzo”; en donde se obtuvo que el 12.6% de los clientes tienen edades de 30 años. Así mismo, el 20.3% son de sexo femenino y el 79.7% son de sexo masculino. Por otro lado, el 39.3% de los clientes tienen 2 personas a cargo

y el 54.7% de los clientes cuentan con estudios primarios.

Con respecto al objetivo específico 2: Evaluar la capacidad del modelo para calificar la admisión de clientes. En el presente estudio según los resultados de la estimación econométrica del modelo Logit, se encontró que la variable edad es negativo, lo cual significa que disminuye la probabilidad de incumplimiento de ser un cliente etiquetado como malo en un 0.98%. Por otra parte, la variable genero su coeficiente es negativo, lo cual indica que ante un cambio de género masculino a género femenino, la probabilidad de incumplimiento de ser un cliente etiquetado como malo, disminuye en un 8.41%. Así mismo, la variable número de hijos tiene un coeficiente positivo, que significa que el incumplimiento de ser un cliente etiquetado como malo, se incrementa en 9.90%. Por otro lado, la variable número de integrantes por familia tiene signo positivo, lo cual significa la probabilidad de incumplimiento de ser un cliente etiquetado como malo, aumenta en 0.91%. En cuanto, a la variable experiencia en el negocio tiene signo negativo, que significa la probabilidad de incumplimiento de ser un cliente etiquetado como malo disminuye en 0.08%. Con respecto la variable ingreso tiene signo negativo, que indica que el riesgo disminuye en 0.01%. Por último, la variable número de préstamos tiene signo positivo, que significa la probabilidad de incumplimiento de ser un cliente etiquetado como malo, aumenta en 1.11%. Resultados similares encontramos en el estudio realizado al comparar el estudio de investigación con los antecedentes recopilados y evidenciamos lo siguiente:

México:

Albarran, Montes y Meza (2015), en su artículo de investigación denominada: *“El riesgo de no pago en una institución microfinanciera del municipio de Tenancingo estado de México 2011-2014”*. En cuanto a los resultados, aplicando la regresión logística, se ha diseñado un modelo de calificación estadística capaz de predecir correctamente 98.5 por ciento de los clientes de la cartera de esta institución, por lo que se concluye que el modelo es válido y podría aplicarse en cualquier IMF rural de México.

Ecuador:

Paucar (2018), en su tesis denominada: *“Desarrollo de un Scoring de microfinanzas para las cooperativas de ahorro y crédito para los segmentos 2A5 Caso COAC textil 14 de marzo”*. En cuanto a los resultados; la calificación central de riesgos, edad, tipo de vivienda, nivel de educación y antigüedad de residencia son los determinantes para el acceso de un crédito.

Colombia:

Valencia (2017), en su tesis denominada: “*Modelo Scoring para el otorgamiento de crédito de las pymes*”. En cuanto a los resultados; después de haber aplicado el modelo Logit, se identificaron que el plazo al cuadrado, el cupo, ventas netas, activos totales, acierta plus y prueba ácida, son los determinantes para el acceso de las pymes a crédito en Colombia.

Millán y Caicedo (2018), en su tesis denominada: “*Modelos para otorgamiento y seguimiento en la gestión de riesgo de crédito*”. En cuanto a los resultados obtenidos luego de emplear tres modelos, el modelo de regresión logística tiene un performance del 81.0% con respecto a la tasa de aciertos en la clasificación, para los tres modelos se emplearon catorce variables, los cuales informan, las características socioeconómicas del prestatario, las características propias de la operación crediticia. Con respecto al ámbito de la gestión financiera, estos resultados son importante dado que puede complementarse con el cálculo de la probabilidad de incumplimiento, con los montos expuestos en cada operación de crédito y con la tasa de recuperación de la entidad para a si de esta manera establecer el valor de las perdidas esperadas a nivel individual y a nivel del portafolio de créditos de la entidad.

IV.CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

Después de haber concluido el presente trabajo de investigación se llega a las siguientes conclusiones:

1. En las variables demográficas; Edad se observa que los clientes de la entidad financiera el 36.0% tienen edades de 32 a 41 años. Con respecto al sexo, el 58.2% son de sexo masculino y el 41.8% son de sexo femenino. Así mismo, el 52.9% de los clientes tienen hijos menores de 18 años, el 24.5% tienen hijos mayores a 18 años y sólo el 22.7% de los clientes no tienen hijos. Por otro lado, el 23.1% de los clientes tienen 2 miembros en sus hogares, el 18.3% de los clientes tienen 3 miembros y sólo un 8.0% de los clientes tienen un solo miembro en sus hogares. De mismo modo, el 30.4% de los clientes tienen experiencia hasta 36 meses en el negocio de créditos y por último, el 40.5% de los clientes tienen cero préstamos.
2. Las variables incluidas en el modelo Logit resultaron ser estadísticamente significativas y con el signo esperado. De esta manera, los determinantes del comportamiento de pagos de los clientes son: La edad, género, número de hijos, número de integrantes por familia, experiencia en el negocio, ingreso y número de préstamos.

4.2. Recomendaciones

Frente a la realidad identificada y posteriormente analizado en cada una de los objetivos específicos, el investigador considera las siguientes recomendaciones:

1. A la comunidad científica

Se recomienda a los futuros investigadores utilizar este estudio referido al desarrollo de un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento microempresa, utilizando un modelo econométrico Logit.

2. A la entidad financiera de la agencia San Juan de Lurigancho

Para gestionar de una manera óptima y eficiente, el riesgo de crédito, se requieren desarrollar modelos predictivos internos, con respecto al comportamiento de pago de los clientes del segmento microempresa, estos tienen que sustentarse en la información histórica de los clientes de la financiera, para clasificar de esta manera a los clientes con buen comportamiento de pagos, de aquellos clientes con mal comportamiento de pagos (buenos y malos), es importante mencionar que el modelo credit scoring logit se convierte en una herramienta fundamental al tomar la decisión de aprobar o rechazar un crédito, en el segmento microempresa se encuentran clientes sin antecedentes crediticios históricos, como también sin información económica financiera, por tanto es importante levantar esta información in situ en el negocio y la vivienda de los clientes, de esta manera se garantiza una información de calidad, para mejorar la capacidad predictiva del modelo.

El modelo econométrico logit, clasifico erróneamente 153 clientes malos como clientes buenos, representando el 42.15% del total de clientes malos, en tal sentido, se recomienda incorporar mayor número de variables en la toma de información de los clientes, para optimizar la predicción del modelo credit scoring de crédito.

3. A la Universidad Señor de Sipán

Publicar este trabajo de investigación en el repositorio institucional para darse a conocer sobre la importancia del desarrollo de un modelo credit scoring para predecir el comportamiento de pagos de los clientes del segmento microempresa, utilizando un modelo econométrico Logit.

REFERENCIAS

- Aglietta, M. (2000). *Macroeconomía Financiera*. Ecuador : ABYA - YALA .
- Albarran, V., Montes, F., y Meza, I.(2015), *El riesgo de no pago en una institución microfinanciera del municipio de Tenancingo estado de México 2011-2014*, Obtenido de http://ri.uaemex.mx/bitstream/handle/20.500.11799/40683/TESIS-IMF_Password_Removed.pdf?sequence=1&isAllowed=y
- Back, P. (2005). *Explaining financial difficulties based on previous payment behavior, management background variables and financial ratios*, *European Accounting Review*, 14:4, 839-868,
- Banco Interamericano de Desarrollo. (2009). *Políticas y buenas prácticas de financiamiento para la pequeña empresa*. Obtenido de https://issuu.com/alvaroaraya/docs/politicas_de_financiamiento_pymes
- Bateman, J., Carson, T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones, M., Loomes, G., Mourato, S., Özdemiroğlu, E., Pearce, D., Sugden, R., & Swanson, J. (2002). *Economic valuation with stated preference techniques: a manual*. Edward Elgar Publishing, Inc. USA.
- Bessis, J. (2002). *Risk Management in Banking*. John Wiley & Sons. Chichester.
- Beltran, M. (2015). *Diseño e implementación de un nuevo clasificador de préstamos bancarios a través de la minería de datos*. (Tesis doctoral, UNED). Obtenido de: http://espacio.uned.es/fez/eserv/tesisuned:CiencEcoEmpMbeltran/BELTRAN_PASCUAL_Ma uricio_Tesis.pdf
- BCRP. (2018). *Reporte de Estabilidad Financiera*. Obtenido de <http://www.bcrp.gob.pe/docs/Publicaciones/Reporte-Estabilidad-Financiera/ref-noviembre-2018.pdf>
- Castillo, O. (2009). *Análisis econométrico para identificar y cuantificar la exclusión social por variables socioeconómicas en el acceso al microcrédito en el Ecuador a partir de la escuela de condiciones de Vida Quinta Ronda 2005-2006*. Facultad de Ciencias de la Escuela Politécnica Nacional. Ecuador 2009. Setiembre del 2009. Obtenido de <http://bibdigital.epn.edu.ec/bitstream/15000/1861/1/CD-2430.pdf>.
- Congreso de la Republica. (2003). *Ley de promoción y formalización de la micro y pequeña empresa N° 28015*. Lima, Perú: PRODLAB. Obtenido de <http://www4.congreso.gob.pe/comisiones/2002/discapacidad/leyes/28015.htm>
- Congreso de la Republica. (2013). *Ley que modifica diversas leyes para facilitar la inversión, impulsar el desarrollo productivo y el crecimiento empresarial N° 30056*. Lima, Perú: Obtenido de <https://busquedas.elperuano.pe/normaslegales/ley-que-modifica-diversas-leyes-para-facilitar-la-inversion-ley-n-30056-956689-1/>
- Custodio, P. (2014). *La venta de Mibanco: ¿Por qué era urgente para el Grupo ACP?* Semana Económica.com. Obtenido de <http://semanaeconomica.com/article/management/negocios/132010-la-venta-de-mibanco-por-que-era-tan-urgente-para-el-grupo-acp/>
- De Lara, A. (2003): *Medición y Control de Riesgos Financieros*. Limusa, México D.F.

- Farfán., J. (2019). *Caracterización de las variables de riesgo en la evaluación de créditos para la determinación de un modelo scoring en la cooperativa CACSA, de la ciudad de Juliaca en el año 2017*. Obtenido de: http://repositorio.uancv.edu.pe/bitstream/handle/UANCV/2991/T036_01335989_M.pdf?sequence=3&isAllowed=y
- Flórez, R. (2012). *Análisis de los determinantes del riesgo de crédito: aplicación de técnicas emergentes en el mercado de los acuerdos de Basilea II y Solvencia II*. Madrid. Editorial Instituto de Contabilidad y Auditoría de Cuentas, Ministerio de Economía y Competitividad.
- Gujarati, D., Porter, D. (2010). *Econometría, Quinta edición*. México: McGraw-Hill.
- Hand, Y., Henley, E. (1997). *Statistical Classification Methods in Costumer Credit Scoring: A review*. Journal of the Royal Statistical Association. Series A. Vol. 160. Part 3. pp. 523-541.
- Hernandez, S. (2014). *Tipos de investigacion*. Obtenido de http://catarina.udlap.mx/u_dl_a/tales/documentos/lad/arenas_m_a/capitulo3.pdf
- Kim, J. (2005). *A Credit Risk Model for Agricultural Loan Portfolios under the New Basel Capital Accord*. Dissertation submitted to the Office of Graduate Studies of Texas A&M University.
- Lara, J. (2010). *La gestión del riesgo de crédito en las instituciones de microfinanzas*. (Tesis doctoral, Universidad de Granada). Obtenido de: <http://hera.ugr.es/tesisugr/18892656.pdf>
- Matos., C. (2017), *El riesgo crediticio hipotecario en las cajas municipales de ahorro y crédito aplicando credit scoring y VAR*. Obtenido de: <http://repositorio.lamolina.edu.pe/bitstream/handle/UNALM/3236/matos-falvy-christian-cesar.pdf?sequence=1&isAllowed=y>
- Medina, H., Ulfe, H. (2015). *Modelo de credit scoring para predecir el otorgamiento de crédito personal en una cooperativa de ahorro y crédito*. Obtenido de <http://repositorio.unprg.edu.pe/bitstream/handle/UNPRG/1339/BC-TES-TMP-172.pdf?sequence=1&isAllowed=y>.
- Medina, E. (2003). *Modelos de elección discreta*. Obtenido de <http://www.cartagena99.com/recursos/alumnos/apuntes/logit.pdf>
- Mibanco. (2014). *Estados financieros al 31 de diciembre de 2014 y de 2013 junto con el dictamen de los auditores independientes*. Obtenido de: <https://www.mibanco.com.pe/repositorioaps/data/1/1/1/jer/eeff-auditados/files/Estados%20Financieros%20Auditados%20al%2031%20Dic%202014.pdf>
- Millan, J., Caicedo, E. (2018). *Modelos para otorgamiento y seguimiento en la gestión de riesgo de crédito*. Obtenido de: <https://dialnet.unirioja.es/servlet/articulo?codigo=6499643>
- Ministerio de Trabajo y promoción del empleo. (2006). *La microempresa: es una propuesta tipológica y ejercicios de aplicación en Lima Sur*. Obtenido de: https://www.trabajo.gob.pe/archivos/file/estadisticas/peel/bel/BEL_34.pdf
- Monteros, E. (2005). *Manual de gestión Microempresarial*. Editorial universitaria. Ecuador.
- McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. In *Frontiers in Econometrics* (Edited by P. Zarembka), 105-42. Academic Press, New York.

- Navarro, H. (2015). *Diseño de un Modelo de credit scoring aplicado a créditos recurrentes para una caja municipal*. Obtenido de <http://repositorio.unp.edu.pe/bitstream/handle/UNP/674/IND-NAV-GUE-15.pdf?sequence=1>.
- Ñaupas, H., Mejia, E., Novoa, E., y Villagomez, (2015). *Metodología de la investigación Cuantitativa - Cualitativa y Redacción de la Tesis*. (4ª ed.). Lima, peru: Ediciones de la U.
- Oña, E. (2015). *Desarrollo de un modelo de gestión de riesgo de crédito para instituciones de microfinanzas del Ecuador*. Obtenido de <http://repositorio.uasb.edu.ec/bitstream/10644/4652/1/T1709-MFGR-O%c3%b1a-Desarrollo.pdf>
- Pantoja, P. (2016). *Propuesta de un Modelo Logit para evaluar el Riesgo Crediticio en las Cajas Municipales de Ahorro y Crédito: Caso de la Caja Municipal de Huancayo, periodo 2011-2015*. Obtenido de http://repositorio.usil.edu.pe/bitstream/USIL/2738/1/2016_Pantoja_Propuesta-de-un-modelo-Logit.pdf
- Paucar, W. (2018). *Desarrollo de un scoring de microfinanzas para las cooperativas de ahorro y crédito para los segmentos 2A5. Caso COAC textil 14 de marzo*. Obtenido de <https://bibdigital.epn.edu.ec/bitstream/15000/20038/1/CD-9462.pdf>
- Peña, J. (2002). *La Gestión de Riesgos Financieros de Mercado y Crédito*. Prentice Hall - Financial Times, Madrid.
- Plonicki, B. (2005). *Modelo de comportamiento y predicción de incumplimiento crediticio: el caso de empresas pyme en Argentina*. En temas de management.Vol.3. Argentina. Obtenido de https://www.researchgate.net/publication/4806984_Modelo_de_Comportamiento_y_Prediccion_de_Incumplimiento_Crediticio_el_caso_de_empresas_Pyme_en_Argentina
- Rayo, S., Lara, J., y Camino, D. (2010). *Un Modelo de Credit Scoring Para Instituciones de Microfinanzas En El Marco de Basilea II*. Journal of Economics, Finance and Administrative, Volume (12), pp 89-124. Obtenido de <http://ezproxy.unal.edu.co/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eds&AN=eds&S2077.18862010000100005&lang=es&site=eds-live>.
- Rodríguez, D., Becerra, J. y Cardona, D. (2017). *Modelos y metodologías de credit score para personas naturales: una revisión literaria*. Revista CEA, 3(5), 13-28. Obtenido de <https://revistas.itm.edu.co/index.php/revista-cea/article/view/645/643>
- Rodríguez., R. (2018). *Modelo credit scoring para obtener la probabilidad de impago de créditos personales futuro en la empresa servicios financieros grupo Buro Huaraz 2018*. Obtenido de http://repositorio.unasam.edu.pe/bitstream/handle/UNASAM/2666/T033_44395919_T.pdf?sequence=5&isAllowed=y
- Sarco, M. (2017). *Factores que determinan el otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri, 2015*. Obtenido de http://repositorio.unap.edu.pe/bitstream/handle/UNAP/4232/Sarco_Yampasi_Maribel_Giovan_a.pdf?sequence=1&isAllowed=y
- Tesen., A. (2017). *Eficacia de los modelos de aprendizaje de máquina para evaluar el riesgo crediticio de personas naturales en una institución financiera de Chiclayo*. Obtenido de <http://repositorio.uns.edu.pe/bitstream/handle/UNS/2980/46297.pdf?sequence=1&isAllowed=y>

- Basulto, S. (2018). Modelo de Credit Scoring para el otorgamiento de microcréditos en Guantánamo. *Revista cubana de Finanzas y precios*, 2(4), 70-81. Obtenido de http://www.mfp.gob.cu/revista_mfp/index.php/RCFP/article/view/06_V2N42018_RHM
- Valencia, A. (2017), *Modelo Scoring para el otorgamiento de crédito de las pymes*. Obtenido de. https://repository.eafit.edu.co/bitstream/handle/10784/12295/Andrea_ValenciaEcheverri_2017.pdf?sequence=2
- Valdivia, R., Cuevas, C., Sandoval, M., Romo, J, (2009), *Estimación conométrica de la disposición a pagar de los consumidores de servicios recreativos turísticos*. *Terra Latinoamericana* 27: 227-235. Obtenido de. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-57792009000300007
- Valle., A. (2017), Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones. Obtenido de: <https://idus.us.es/xmlui/bitstream/handle/11441/63201/Valle%20Benavides%20Ana%20Roc%20c3%ado%20del%20TFG.pdf?sequence=1&isAllowed=y>
- Venero, H, (2002). *¿Operan las instituciones financieras en la frontera de tecnologías crediticias eficientes?* . Obtenido de <http://www.cies.org.pe/sites/default/files/investigaciones/operan-las-instituciones-financieras-en-la-frontera-de-tecnologias-crediticias-eficientes-un-analisis-de-transacciones-de-credito-en-los-sectores-de.pdf>
- Velandia, N. (2013). *Establecimiento de un Modelo Logit para la Medicion del Riesgo de Incumplimiento en Creditos para una Entidad Financiera del Municipio de Arauca*., Obtenido de <http://bdigital.unal.edu.co/10202/1/7709584.2013.pdf>

ANEXOS

ANEXO N°01 RATIOS DE MOROSIDAD DEL SISTEMA FINANCIERO (%)

	Sistema Financiero			Banca			No Banca		
	Set.16	Set.17	Set.18	Set.16	Set.17	Set.18	Set.16	Set.17	Set.18
Total	4.50	4.80	5.00	4.10	4.50	4.50	7.70	7.50	8.50
A Empresas:	4.60	5.00	5.30	4.00	4.50	4.70	8.80	8.60	9.70
Corporativos	0.00	0.20	0.50	0.00	0.20	0.50	0.00	0.00	0.00
Grandes	1.80	2.10	2.90	1.80	2.10	2.90	9.70	6.30	6.10
Medianas	9.20	10.30	11.10	9.10	10.30	10.90	10.40	10.20	13.80
MYPE	9.70	9.50	9.80	10.50	10.40	10.40	8.80	8.60	9.20

Nota. Balances de Comprobación. BCRP.

ANEXO 02
ESTADOS DE RESULTADOS POR LOS AÑOS TERMINADOS EL 31 DE
DICIEMBRE 2014 Y DEL 2013

	Nota	2014 S/0	2013 S/0
Ingresos por intereses	15	911,949	1,034,695
Gastos por intereses	15	-208,663	-215,565
Margen financiero bruto		703,286	819,130
Provisión para créditos de cobranza dudosa, neta de recuperos	5(f)	-310,915	-312,882
Margen financiero neto		392,371	506,248
Ingresos por servicios financieros		42,766	52,218
Gastos por servicios financieros		-14,316	-13,215
Margen financiero neto de ingresos y gastos por servicios financieros		420,821	545,251
Resultados por operaciones financieras	16	10,484	7,039
Gastos de administración	17	-427,316	-481,438
Depreciación	6(a)	-17,789	-18,421
Amortización	7(c)	-30,893	-7,281
Margen operacional neto		-44,693	45,150
Provisiones para contingencias y otros	7(d)	-25,563	-4,374
(Pérdida) utilidad de operación		-70,256	40,776
Otros(gastos) ingresos, neto	18	-4,726	11,685
(Pérdida) utilidad antes del impuesto a la renta		-74,982	52,461
Impuesto a la renta	11	7,412	-17,252
(Pérdida) utilidad neta		-67,570	35,209
(Pérdida) utilidad por acción básica y diluida (en nuevos soles)	19	-0.14	0
Número de acciones promedio ponderado en circulación (en miles de unidades)	19	481,338	481,338

Nota. Mibanco 2014.

ANEXO N°03
ESTIMACIÓN DEL MODELO CREDIT SCORING PARA PREDECIR EL
COMPORTAMIENTO DE PAGOS DE LOS CLIENTES DEL SEGMENTO
MICROEMPRESA

```

. logit calif_cred edad genero num_hij num_famil exp_neg ing_real num_prest

Iteration 0:  log likelihood = -947.32081
Iteration 1:  log likelihood = -676.13719
Iteration 2:  log likelihood = -621.31682
Iteration 3:  log likelihood = -620.1444
Iteration 4:  log likelihood = -620.14146
Iteration 5:  log likelihood = -620.14146

Logistic regression                                Number of obs =      2000
                                                    LR chi2(7)         =    654.36
                                                    Prob > chi2        =    0.0000
Log likelihood = -620.14146                       Pseudo R2         =    0.3454

```

calif_cred	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
edad	-.1377817	.0148569	-9.27	0.000	-.1669008	-.1086627
genero	-1.244435	.1717622	-7.25	0.000	-1.581083	-.9077873
num_hij	1.395378	.1950028	7.16	0.000	1.013179	1.777576
num_famil	.1287055	.0444174	2.90	0.004	.041649	.215762
exp_neg	-.0106437	.0058251	-1.83	0.068	-.0220608	.0007733
ing_real	-.0008988	.00007	-12.84	0.000	-.001036	-.0007616
num_prest	.1559247	.0505094	3.09	0.002	.0569282	.2549212
_cons	4.798816	.4366291	10.99	0.000	3.943038	5.654593

Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.
 Elaboración propia con Stata versión 13.

ANEXO N° 04
MODELO LOGIT: EFECTOS MARGINALES SOBRE LA PROBABILIDAD
COMPORTAMIENTO DE PAGOS DE LOS CLIENTES DEL SEGMENTO
MICROEMPRESA

. mfx							
Marginal effects after logit							
y = Pr(calif_cred) (predict)							
= .07688757							
variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]
edad	-.0097792	.00113	-8.63	0.000	-.012	-.007558	39.2425
genero*	-.0840983	.01177	-7.14	0.000	-.107176	-.06102	.418
num_hij	.0990381	.01448	6.84	0.000	.070657	.127419	1.0175
num_fa^l	.009135	.00317	2.88	0.004	.002916	.015354	3.7085
exp_neg	-.0007554	.00042	-1.80	0.071	-.001576	.000065	27.689
ing_real	-.0000638	.00001	-11.79	0.000	-.000074	-.000053	3493.88
num_pr^t	.0110669	.00361	3.07	0.002	.003994	.018139	1.159

Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.

Elaboración propia con Stata versión 13.

**ANEXO N°05
ESTADÍSTICA DE CLASIFICACIÓN**

```
. estat class
```

Logistic model for calif_cred

Classified	True		Total
	D	~D	
+	153	117	270
-	210	1520	1730
Total	363	1637	2000

Classified + if predicted Pr(D) >= .5
True D defined as calif_cred != 0

Sensitivity	Pr(+ D)	42.15%
Specificity	Pr(- ~D)	92.85%
Positive predictive value	Pr(D +)	56.67%
Negative predictive value	Pr(~D -)	87.86%

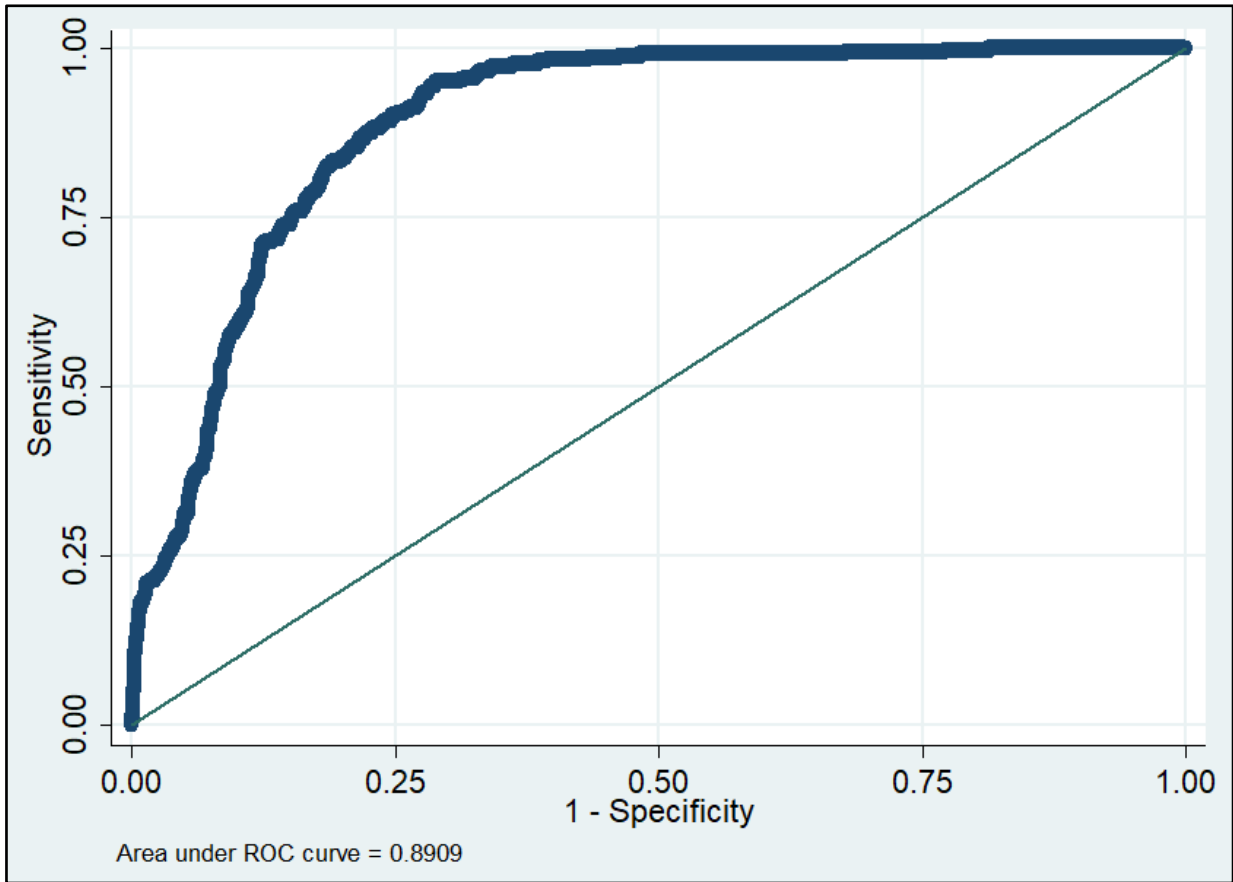
False + rate for true ~D	Pr(+ ~D)	7.15%
False - rate for true D	Pr(- D)	57.85%
False + rate for classified +	Pr(~D +)	43.33%
False - rate for classified -	Pr(D -)	12.14%

Correctly classified	83.65%
----------------------	---------------

Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.

Elaboración propia con Stata versión 13.

**ANEXO N°06
CURVAS ROC**



Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018.
Elaboración propia con Stata versión 13.

ANEXO N°07
CONSTANCIA DE JUICIO DEL EXPERTO

Nombre del experto: Mg. Nelson Alejandro Puyen Farias

Especialidad:

- Magíster en Administración de Negocios y Relaciones Internacionales - Universidad Cesar Vallejo.
- Ingeniero en Organización de Empresas - Universidad Argentina de la Empresa, Buenos Aires, Argentina.

Por medio de la presente hago constar que el trabajo metodológico elaborado por el estudiante Barreto Gamarra Dante Ulises de la Escuela de Ingeniería Económica de la Universidad Señor de Sipán, de la tesis titulada: **“DESARROLLO DE UN MODELO CREDIT SCORING PARA PREDECIR EL COMPORTAMIENTO DE PAGOS DE LOS CLIENTES DEL SEGMENTO MICROEMPRESA”**, que después de haber revisado todo el proceso del trabajo metodológico, doy fe que el proceso de la metodología utilizada en esta tesis cuenta con sustento teórico.

Mg. Nelson Alejandro Puyen Farias
DNI N°06431088

Pimentel, 17 octubre del 2019

ANEXO N°08
CONSTANCIA DE JUICIO DEL EXPERTO

Nombre del experto: Mg. Econ. Maximo Damian Valdera

Especialidad:

- Magíster en Investigación y Docencia - Universidad Nacional Pedro Ruíz Gallo.
- Economista - Universidad Nacional Pedro Ruíz Gallo.
- Especialista en Econometría Aplicada - Universidad Nacional de Ingeniería.

Por medio de la presente hago constar que la estimación del modelo econométrico Logit elaborada por el estudiante Barreto Gamarra Dante Ulises de la Escuela Profesional de Ingeniería Económica de la Universidad Señor de Sipán, de la tesis titulada: “**DESARROLLO DE UN MODELO CREDIT SCORING PARA PREDECIR EL COMPORTAMIENTO DE PAGOS DE LOS CLIENTES DEL SEGMENTO MICROEMPRESA**”, que después de haber revisado las estimaciones como especialista en econometría, doy fe que los resultados del modelo econométrico Logit cuentan con sustento teórico económico, estadístico y econométrico.

Mg. Econ. Maximo Damian Valdera
DNI N°42558156

Pimentel, 10 Noviembre del 2019

ANEXO N° 09
CONSTANCIA DE JUICIO DEL EXPERTO

Nombre del experto: Dra. Yonira Olinda Campos Díaz

Especialidad:

- DOCTORA EN ADMINISTRACION DE LA EDUCACION
UNIVERSIDAD PRIVADA CÉSAR VALLEJO
- ECONOMISTA - UNIVERSIDAD NACIONAL PEDRO RUÍZ GALLO.

Por medio de la presente hago constar que las interpretaciones de los resultados del modelo scoring por el estudiante Barreto Gamarra Dante Ulises de la Escuela Académico Profesional de Ingeniería Económica de la Universidad Señor de Sipán, de la tesis titulada: **“DESARROLLO DE UN MODELO CREDIT SCORING PARA PREDECIR EL COMPORTAMIENTO DE PAGOS DE LOS CLIENTES DEL SEGMENTO MICROEMPRESA”**, que después de haber revisado detalladamente las interpretaciones de las Tablas presentadas, doy fe que los resultados cuentan con sustento teórico.

Dra. Campos Díaz Yonira Olinda
DNI N° 16696421

Pimentel, 25 Octubre del 2019

ANEXO N°10
MATRIZ DE CORRELACION DE LAS VARIABLES DEL MODELO LOGIT

	EDAD	GENERO	NUM_HIJ	NUM_FAMIL	EXP_NEG	ING	NUM_PREST	CALIF_CRED
EDAD	1.000							
GENERO	0.122	1.000						
NUM_HIJ	0.819	0.089	1.000					
NUM_FAMIL	0.235	-0.035	0.266	1.000				
EXP_NEG	0.121	0.084	0.079	0.011	1.000			
ING	0.406	0.160	0.360	0.077	0.176	1.000		
NUM_PREST	-0.019	-0.003	-0.013	0.043	-0.033	-0.024	1.000	
CALIF_CRED	-0.333	-0.246	-0.239	0.011	-0.362	-0.446	0.065	1.000

Nota. Datos estadísticos de la cartera de créditos microempresa de la agencia San Juan de Lurigancho de una financiera, de setiembre de 2017 hasta agosto 2018. Elaboración propia con Stata versión 13.