

# 38

Fecha de presentación: febrero, 2022

Fecha de aceptación: mayo, 2022

Fecha de publicación: junio, 2022

## CREDIT SCORING

COMO ALTERNATIVA PARA MINIMIZAR RIESGO DE CRÉDITO EN INSTITUCIONES MICROFINANCIERAS EN COVID-19

### **CREDIT SCORING AS AN ALTERNATIVE TO MINIMIZE CREDIT RISK IN MICROFINANCE INSTITUTIONS DURING COVID-19**

Milagro Baldemar Quiroz Calderón<sup>1</sup>

E-mail: [mquirozc@uladech.edu.pe](mailto:mquirozc@uladech.edu.pe)

ORCID: <https://orcid.org/0000-0002-2286-4606>

Roxana Milagros Torres Guzmán<sup>1</sup>

E-mail: [rtorresg@uladech.edu.pe](mailto:rtorresg@uladech.edu.pe)

ORCID: <https://orcid.org/0000-0003-1471-9879>

Julio Javier Montano Barbuda<sup>1</sup>

E-mail: [jmontanob@uladech.edu.pe](mailto:jmontanob@uladech.edu.pe)

ORCID: <https://orcid.org/0000-0002-1620-5946>

Luis Alberto Torres García<sup>1</sup>

E-mail: [ltorresg@uladech.edu.pe](mailto:ltorresg@uladech.edu.pe)

ORCID: <https://orcid.org/0000-0002-5958-494X>

Wilberto Fernando Rubio Cabrera<sup>1</sup>

E-mail: [wrc@uladech.edu.pe](mailto:wrc@uladech.edu.pe)

ORCID: <https://orcid.org/0000-0002-3426-2768>

<sup>1</sup> Universidad Católica Los Ángeles de Chimbote. Perú.

#### Cita sugerida (APA, séptima edición)

Quiroz Calderón, M. B., Torres Guzmán, R. M., Montano Barbuda, J. J., Torres García, L. A., & Rubio Cabrera, W. F., (2022). Credit scoring como alternativa para minimizar riesgo de crédito en instituciones microfinancieras en COVID-19. *Revista Universidad y Sociedad*, 14(S3), 376-385.

#### RESUMEN

En el Perú, el éxito y el desarrollo sostenible de las microfinanzas ha tenido como base aspectos tan importantes como el marco regulatorio institucional de apoyo, especialmente normas de información financiera y transparencia, centrales de riesgo, difusión de tasas de interés, tecnologías crediticias apropiadas, la promoción de la transparencia de precios y competencia en el mercado. Sin embargo, este nicho del mercado se hace cada vez más competitivo por efectos naturales de la globalización, expresado en el ingreso de las empresas oligopólicas imponiéndose con nuevos capitales y fundamentalmente con tecnologías actualizadas en la gestión de los riesgos. El objetivo de este trabajo es proponer un modelo de Credit Scoring para minimizar el riesgo de crédito de la cartera de microcréditos del sector de microfinanzas en Perú.

**Palabras clave:** Riesgo, riesgo crediticio, microfinanzas, Credit Scoring, tecnologías, mercado.

#### ABSTRACT

In Peru, the success and sustainable development of microfinance has been based on such important aspects as the institutional regulatory framework of support, especially the financial information and transparency standards, credit bureaus, interest rate dissemination, appropriate credit technologies, and the promotion of price transparency and market competition. However, this market niche is becoming more and more competitive due to the natural effects of globalization, expressed in the entry of oligopolistic companies imposing themselves with new capital and, fundamentally, with updated technologies in risk management. The objective of this paper is to propose a Credit Scoring model to minimize the credit risk of the microcredit portfolio of the microfinance sector in Peru.

**Keywords:** Risk, credit risk, microfinance, credit scoring, technologies, market.

## INTRODUCCIÓN

En diciembre de 2019 el mundo conoció la aparición de un nuevo coronavirus, causante de un tipo de gripe potencialmente mortal, el cual se expandió rápidamente en China, y en los meses sucesivos se extendió al resto de los países del planeta, alcanzando la categoría de pandemia esta nueva enfermedad denominada Covid-19. Los efectos de su expansión se hicieron sentir progresivamente en Asia, luego Europa, América y el resto del mundo, dejando a su paso la muerte de miles de personas y causando una crisis sanitaria sin precedentes en la última centuria, así como importantes afectaciones en los ámbitos económico y social (García, 2021; Iturralde & Duque, 2021).

En correspondencia, la crisis se extendió a todos los sectores de la economía por la paralización casi total de las producciones y servicios no esenciales, medidas gubernamentales de aislamiento social, entre otras de emergencias que conllevaron al cierre de negocios, y con ello el deterioro de la situación financiera y la capacidad de honrar las obligaciones crediticias contraídas con anterioridad. De acuerdo con el informe especial Covid-19, emitido por (Comisión Económica para América Latina (CEPAL), 2020) en la mayoría de los países, los servicios no esenciales han sido paralizados para asegurar el distanciamiento social (Becerra et al., 2021; La Hora, 2021; Quintero, 2020).

Aunque los trabajadores formales que trabajan en sectores esenciales aún reciben sus ingresos, muchos trabajadores formales ya no reciben ingresos o los reciben de manera parcial. Esta situación incluye a asalariados y a dueños de micro, pequeñas y medianas empresas. Las medidas de protección social para los trabajadores formales son de dos tipos. El primer tipo de medidas está destinado a reducir la exposición al virus de los trabajadores y garantizar la continuidad de la actividad económica, y el segundo, a asegurar los ingresos o los puestos de trabajo (Bermúdez et al., 2020; España, 2020).

El impacto negativo de la crisis en la región latinoamericana es complejo debido al alto índice de informalidad en el aparato productivo de sus países, hecho que agrava la situación de vulnerabilidad de un amplio sector de la población que carece de recursos económicos y no cuenta con el acceso a la infraestructura institucional necesaria para superar la crisis (Bapuji et al., 2020). En el Perú ha ocasionado una aguda crisis económica y social a causa de la paralización de las actividades económicas aunque el sistema financiero peruano se encontraba sólido cuando llegó la crisis, lo que permitió enfrentar ese escenario

de la mejor manera, según indicó Luis Baba Nakao, consultor de microfinanzas (Pérez, 2020).

Este experto dio a conocer el desempeño de cada sector crediticio entre marzo y junio, periodo en el que se desarrolló la cuarentena. A diferencia de otras crisis, en esta oportunidad, el sistema financiero afronta la crisis del COVID-19 bastante sólido, bien capitalizado, con bastante liquidez y buenas previsiones, poco dependiente de los préstamos del banco, y, sobre todo, apoyado en forma oportuna por las autoridades. Señaló que las colocaciones entre marzo y junio del 2020 se incrementaron en 25.1 millones (7.27%), pasando de \$ 346 393.538 en marzo a \$ 371 584,305 al cierre de junio. Dicho incremento, dijo el experto, se debió principalmente a la inyección de liquidez del gobierno mediante diversos programas se trató de reestablecer la economía peruana (Pérez, 2020).

En el sistema financiero peruano, el uso de modelos se enfoca principalmente en la estimación de la calidad crediticia de operaciones y deudores (scoring y rating) y en la inferencia de ingresos. Si bien el uso de este último tipo de modelo (inferencia de ingresos) obedece en parte al alto nivel de informalidad existente en el país -el mismo que limita la posibilidad de obtener sustentos adecuados de todos los ingresos de un deudor en varios segmentos-, en sectores con menos informalidad, representa una herramienta eficiente que permite reducir los tiempos de evaluación. Hasta febrero de 2020, el sistema financiero contaba con 328 modelos, de los cuales el 68% correspondía a modelos de scoring, el 11% a modelos de rating y el 21% a modelos de inferencia de ingresos, que son utilizados en diversas fases de la gestión, según se muestra en la Tabla 1.

Tabla 1. Usos de los modelos financieros.

Categoría	Fase
Gestión de la cartera de créditos	Admisión
	Seguimiento
	Cobranzas y recuperaciones
	Segmentación
Otros	Cálculo de provisiones IFRS9 y capital económico

Fuente: Elaboración propia a partir de información proporcionada por las empresas del sistema financiero a (Superintendencia de Banca y Seguro, 2021).

Ante la emergencia por Covid-19 estos modelos de gestión de riesgo de crédito requirieron ser ajustados para tomar en cuenta el deterioro de la actividad económica por la emergencia sanitaria y las medidas de aislamiento

social obligatorio, así como el efecto que los programas del Gobierno y las medidas regulatorias tuvieron sobre sus principales variables. En consecuencia, las empresas del sistema financiero tomaron diversas medidas de corto plazo consistentes en ajustar a las nuevas circunstancias la gestión de sus modelos y la revisión de sus políticas.

En lo que se refiere a la gestión de modelos, según (Superintendencia de Banca y Seguro, 2021) las estrategias más utilizadas incluyeron un monitoreo más frecuente sobre su desempeño, el congelamiento de los resultados obtenidos antes del inicio de la pandemia, el estrés de los resultados del modelo original, así como la creación de nuevos modelos o su reemplazo. Entre los principales cambios a las políticas, se encuentran la incorporación de nuevas segmentaciones de cartera, la restricción de desembolsos en algunos productos o actividades económicas, así como la revisión de la estrategia en campañas de créditos. También, se revisaron los puntajes mínimos a partir de los cuales se otorga un crédito.

Uno de los modelos empleados es el credit scoring, que es un sistema de calificación de créditos que intenta automatizar la toma de decisiones en cuanto a conceder o no una determinada operación de riesgo, normalmente un crédito. Un credit rating/credit score (calificación crediticia) es una puntuación que otorgan las agencias de rating a los créditos o deudas de diferentes empresas, gobiernos o personas, según su calidad crediticia (que mide la probabilidad de que esos créditos sean impagados). La calificación crediticia desempeña un papel fundamental en muchas áreas, como negocios y las finanzas (Fang & Chen, 2019).

Si bien las organizaciones de microfinanzas desempeñan un papel importante en las economías en desarrollo, los modelos de apoyo a la toma de decisiones para la calificación crediticia de las microfinanzas no han sido suficientemente cubiertos en la literatura, particularmente para las empresas de microcréditos (Gici & Subasi, 2019). En revisión a la literatura especializada en el tema se ha encontrado descrito el empleo de un modelo de calificación crediticia utilizando regresión logística y análisis discriminante multivariante aplicado a 1500 préstamos individuales. Los resultados mostraron la importancia de tener una muestra más grande, un historial lo suficientemente profundo sobre el comportamiento del cliente y también la necesidad de más información sobre las variables relacionadas con la actividad del cliente y su rendimiento para predecir mejor el incumplimiento (Bennouna & Tkiouat, 2019).

De igual modo, se encontró descrito el modelo de calificación crediticia aplicado por una institución de

microfinanzas en Bosnia y Herzegovina, mostrando cómo se identificaron los atributos más relevantes para su implementación (Nali & Švraka, 2018). Otro modelo de calificación crediticia revisado proporciona evidencia de una gran institución de microfinanzas (IMF) en India, donde se ha aplicado tanto el método de calificación crediticia como el método de red neuronal (NN) y con posterior comparación de los resultados (Viswanathan & Shanthi, 2017). Esta revisión bibliográfica también encontró que los datos no tradicionales pueden usarse para construir algoritmos que puedan identificar a los buenos prestatarios como en la banca tradicional (Ruiz & Gama, 2017).

En este entorno se identifica como problema que sustenta esta investigación el deterioro de la situación financiera en Perú a partir de la crisis económica-financiera generada por la Covid-19 y con ella una disminución de la capacidad de honrar las obligaciones crediticias tanto de personas naturales como jurídicas. En este contexto se reconoce la existencia de varios modelos utilizados por la Banca comercial que sin embargo son limitados en su aplicación a nivel de las Microfinanzas, con una notoria necesidad de su empleo a partir de la mencionada crisis. Por tanto, la presente investigación tiene como objetivo proponer un modelo estadístico de Credit Scoring para la cartera de microcréditos del sector de Microfinanzas en Perú postpandemia. Si el modelo demuestra su utilidad, entonces permitirá mejorar la toma de decisiones para el analista de crédito y reducir el indicador de morosidad de las Microfinancieras. Si se demuestra esta hipótesis se espera desarrollar una estrategia que asegure la sostenibilidad postpandemia del proyecto a largo plazo como un programa.

## MATERIALES Y MÉTODOS

Como metodología se aplica la Regresión Logística Binaria. Los modelos de regresión logística son modelos estadísticos en los que se desea conocer la relación entre una variable dependiente cualitativa dicotómica (regresión logística binaria o binomial) y una o más variables explicativas independientes, o covariables, ya sean cualitativas o cuantitativas. También es posible analizar una variable dependiente cualitativa con más de dos valores (regresión logística multinomial), aunque se analiza la regresión logística binaria. La ecuación inicial del modelo es de tipo exponencial, si bien su transformación logarítmica permite su uso como una función lineal. El objetivo primordial que resuelve esta técnica es el de modelar cómo influye en la probabilidad de aparición de un suceso, habitualmente dicotómico, la presencia o no de diversos factores y su valor o nivel (Berlanga, 2014).

En esta investigación la variable respuesta o dependiente es una variable discreta dicotómica, con un valor de cero (0) para el escenario de cliente que paga (no default) y uno (1) cuando no paga (default) y además, se definen como variables independientes o predictores. La población se conforma por 997 clientes que integran parte de la cartera de microcréditos de una Caja Municipal de la ciudad de Piura. Las variables de interés para el estudio son: Número de Cuotas Vencidas, Número de Días de Morosidad, Número de Cuotas del Crédito (log), Tasa de Crédito, Edad, Cuotas Pagadas, Antigüedad del Cliente, Número de Créditos en la Institución, entre otros.

Se utilizó el Software SPSS como programa estadístico (métodos automáticos “por pasos”) versión 24, el modelo de calificación estadística capaz de predecir correctamente la probabilidad de default, teniendo en cuenta que SPSS Statistics es un conjunto de programas. Este software proporciona a los investigadores herramientas que permiten consultar datos y formular hipótesis de forma rápida, ejecutar procedimientos para aclarar las relaciones entre variables, identificar tendencias y realizar predicciones. Esta información servirá para el diseño del Software de aplicación. Se utilizará la  $R^2$  de Cox y Snell y Nagelkerke, la CURVA ROC (Receiver Operating Characteristic).

El modelo de regresión logística puede formularse como (Ecuación 1):

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

Donde:

$p$ : La probabilidad de ocurrencia del evento de interés, en este caso impago o default (Ecuación 2).

$$p = P(\text{default}) = P(\text{impago}) = P(\text{mora}) = P(\text{Incumplimiento}) \quad (2)$$

$1-p$ : Es el complemento de la probabilidad de ocurrencia del evento de interés, en este caso es la probabilidad que el cliente pague o no default (Ecuación 3).

$$\frac{p}{1-p} \quad (3)$$

Se le llama Odds a la proporción del número esperado de veces que ocurra un evento y el número esperado de veces que no ocurra. El odds ratio (OR) expresa si la probabilidad de ocurrencia de un evento o enfermedad: caso/no caso difiere o no en distintos grupos, por lo general catalogados de alto o bajo riesgo o también con relación a su calificación en una encuesta: resultado positivo/

resultado negativo, pero debido a que no posee límites claros es difícil interpretarlo. Su información es fundamentalmente descriptiva, aunque si su intervalo de confianza (IC) no incluye al 1 se concluye que la asociación es estadísticamente significativa. Es decir, que la cantidad de casos que posee el grupo de alto riesgo es significativamente más grande que la cantidad de casos que ubicados en el grupo de bajo riesgo (Domínguez, 2019).

Dado el valor de las variables independientes, la probabilidad de default, puede ser calculada directamente de la siguiente forma (Ecuación 4):

$$p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \quad (4)$$

## RESULTADOS Y DISCUSIÓN

El nivel de significación es un umbral que permite determinar si el resultado de un estudio se puede considerar estadísticamente significativo después de realizar las pruebas estadísticas planificadas. Para ello se determinó la significancia conjunta e individual, como se muestra a continuación.

**Significación de chi-cuadrado del modelo en la prueba ómnibus:** Si la significación es menor de 0,05 indica que el modelo ayuda a explicar el evento, es decir, las variables independientes explican la variable dependiente.

Evaluación de la significancia conjunta se muestra en las Tablas 2 y 3.

Tabla 2. Pruebas ómnibus de coeficientes de modelo.

		Chi-cuadrado	gl	Sig.
Paso 1	Paso	705,855	12	0
	Bloque	705,855	12	0
	Modelo	705,855	12	0

Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del Software SPSS versión 24.

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots = \beta_k$$

$$H_1: \beta_i \neq \beta_j \text{ para todo } i \neq j$$

Dado que el valor de significancia del modelo es  $0.000 < 0.05$ , se rechaza la hipótesis; por lo tanto, existirá por lo menos un . Es decir, existe por lo menos una variable independiente que explique el comportamiento de la variable dependiente (default).

Evaluación de la significancia individual

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$

$$\text{Test de Wald} = \left(\frac{\hat{\beta}}{S_{\hat{\beta}}}\right)^2 \sim \chi_1^2$$

La Tabla 3 muestra las variables en la ecuación.

Tabla 3. Variables en la ecuación.

Variables en la ecuación	B	Error estándar	Wald	gl	Sig.
Número de Cuotas Vencidas	-7,181	4,086	3,088	1	0,079
Número de Días de Morosidad	13,045	1,625	64,419	1	0
Número de Cuotas del Crédito	-5,803	3,366	2,972	1	0,085
Tasa de Crédito	-15,253	4,189	13,257	1	0
Edad	3,014	3,116	0,936	1	0,333
Cuotas Pagadas	-3,584	2,082	2,963	1	0,085
Antigüedad del Cliente	0,452	1,295	0,122	1	0,727
Número de Créditos en la Institución	0,999	1,573	0,403	1	0,525
Variación Anual del IPC.	-21,485	12,309	3,047	1	0,081
Sector Económico(1)	0,052	0,397	0,017	1	0,895
Estado Civil (1)	0,524	0,415	1,595	1	0,207
Sexo(1)	0,883	0,406	4,725	1	0,03
Constante	16,373	6,307	6,738	1	0,009

Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del Software SPSS versión 24.

Variables especificadas en el paso 1: Número de Cuotas Vencidas, Número de Días de Morosidad, Número de Cuotas del Crédito (log), Tasa de Crédito, Edad, Cuotas Pagadas, Antigüedad del Cliente, Número de Créditos en la Institución, Variación Anual del IPC., Sector Económico Dico 0,1, Estado Civil Dico 0,1, Sexo.

#### Evaluación del ajuste del modelo.

**R-cuadrado de Cox y Snell, y R-cuadrado de Nagelkerke:** Indica la parte de la varianza de la variable dependiente explicada por el modelo. Hay dos R-cuadrados en la regresión logística, y ambas son válidas. Se acostumbra a decir que la parte de la variable dependiente explicada por el modelo oscila entre la R-cuadrado de Cox y Snell y la R-cuadrado de Nagelkerke. Cuanto más alto es la R-cuadrado más explicativo es el modelo, es decir, las variables independientes explican la variable dependiente. La determinación de R<sup>2</sup> de Cox y Snell y Nagelkerke es un coeficiente de determinación generalizado que se utiliza para estimar la proporción de varianza de la variable dependiente explicada por las variables predictoras. Su valor fluctúa entre 0 y 1, pero en la práctica no llega a 1. En la Tabla 4 se muestran los valores obtenidos de R<sup>2</sup> de Cox y Snell y Nagelkerke.

Tabla 4. Resumen de los resultados de R<sup>2</sup> de Cox y Snell y Nagelkerke.

Paso	Logaritmo de la verosimilitud -2	R <sup>2</sup> de Cox y Snell	R <sup>2</sup> de Nagelkerke
1	238,758 <sup>a</sup>	0,507	0,829

a) La estimación ha terminado en el número de iteración 8 porque las estimaciones de parámetro han cambiado en menos de 0,001.

Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del **Software** SPSS versión 24.

Para evaluar la bondad del ajuste del modelo de regresión logística, deben observarse los indicadores:

- $R^2$  de Cox y Snell: este indicador muestra que el modelo explica el 50.7% de la variabilidad de la variable dependiente (default).
- $R^2$  de Nagelkerke: este indicador muestra que el modelo explica el 82.9% de la variabilidad de la variable dependiente (default).

**Porcentaje global correctamente clasificado:** Este porcentaje indica el número de casos que el modelo es capaz de predecir correctamente. En base a la ecuación de regresión y los datos observados, se realiza una predicción del valor de la variable dependiente (valor pronosticado). Esta predicción se compara con el valor observado. Si acierta, el caso es correctamente clasificado. Si no acierta, el caso no es correctamente clasificado. Cuantos más casos clasifica correctamente (es decir coincide el valor pronosticado con el valor observado) mejor es el modelo, más explicativo, por tanto, las variables independientes son buenas predictoras del evento o variable dependiente. Si es modelo clasifica correctamente más del 50% de los casos, el modelo se acepta; en caso contrario se seleccionan nuevas variables independientes. En la Tabla 5 se observan los resultados de la determinación del porcentaje global.

Tabla 5. Tabla de clasificación.

Observado	Pronosticado			
	Calificación del Cliente		Porcentaje correcto	
	No Default	Default		
Calificación del Cliente	No Default	807	9	98,9
	Default	24	157	86,7
Porcentaje global				96,7

El valor de corte es 0.500

Fuente: Módulo de regresión logística binaria del **Software** SPSS versión 24.

$$\text{Sensibilidad} = \frac{157}{181} = 86.7\%$$

Sensibilidad: Porcentaje de **default** que son clasificados como **default**.

$$\text{Especificidad} = \frac{807}{816} = 98.9\%$$

Especificidad: Porcentaje de **no default** que son clasificados como **no default**.

- Según los resultados de la tabla de clasificación, de los 181 clientes que no cumplen con sus pagos (**default**), el modelo estimó correctamente 157 de ellos; es decir, el 86.7%. Por tanto, el modelo tiene una sensibilidad del 86.7%.
- De los 816 clientes que cumplen con sus pagos (**no default**), el modelo estimó correctamente a 807 de ellos; es decir al 98.9%. Por tanto, el modelo tiene una especificidad del 98.9%.
- Finalmente, si se suman 807 (no default) más 157 (default) es igual a 964 clientes que fueron seleccionados correctamente, del total de 997. Por tanto, la probabilidad de acertar con este modelo de regresión logística es del 96.7%.

Seguidamente se elaboró la curva ROC (Receiver Operating Characteristic). Una curva ROC es una representación gráfica que ilustra la relación entre la sensibilidad y la especificidad de un sistema clasificador para diferentes puntos de corte. Fue desarrollada por ingenieros eléctricos en la II Guerra Mundial, para medir la eficacia de la detección de objetos enemigos en el campo de batalla mediante señales de radar. Su uso está muy extendido en medicina, para validar técnicas diagnósticas. Más recientemente con el auge de las técnicas de aprendizaje automatizado, se han empleado las curvas ROC para evaluar diferentes algoritmos de clasificación.

Un gráfico de curva ROC ilustra la sensibilidad y especificidad de cada uno de los posibles puntos de corte de un test diagnóstico cuya escala de medición es continua. El análisis ROC, permite evaluar la precisión de algunos modelos estadísticos, tales como: Regresión Logística, Análisis Discriminante, etc. Es un gráfico en el que en el eje de las X se representa ( $1 - \text{Especificidad}$ ) y en el eje de las Y, la **Sensibilidad**.

Información contenida en la curva:

- Si la prueba fuera perfecta, es decir, sin solapamiento, hay una región en la que cualquier punto de corte tiene sensibilidad y especificidad iguales a 1: la curva sólo tiene el punto (0,1).
- Si la prueba fuera inútil: ambas *fdp*'s coinciden y la sensibilidad (verdaderos positivos) es igual a la proporción de falsos positivos, la curva sería la diagonal de (0,0) a (1,1).
- Las pruebas habituales tienen curvas intermedias.

La curva ROC cuantifica la capacidad de un indicador, en este caso el default, para discriminar entre buenos y malos pagadores. Es decir, mientras mayor sea el área bajo la curva ROC, el modelo de Regresión Logística tiene un mayor poder de discriminación. El resultado arrojado para este modelo fue de 96.6%. En la Tabla 6 se muestran los resultados para ambos estados.

Tabla 6. Resumen de procesamiento de casos.

Calificación del Cliente	N válido (por lista)
Positivo <sup>a</sup>	181
Negativo	816

El estado real positivo es Default.

Nota: Los valores más grandes de las variables de resultado de prueba indican una prueba mayor para un estado real positivo. Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del **Software** SPSS versión 24.

A continuación, se muestra en la Figura 1 el gráfico de curva ROC obtenido.

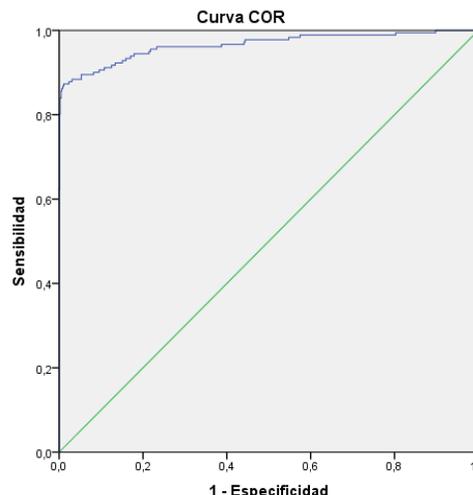


Figura 1. Gráfico de curva ROC. Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del **Software** SPSS versión 24.

La Tabla 7 muestra el resultado de la probabilidad pronosticada.

Tabla 7. Probabilidad pronosticada.

Variables de resultado de prueba: Probabilidad pronosticada
Área
0,966

Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del **Software** SPSS versión 24.

Interpretación del modelo:

Para la interpretación del modelo se tiene que analizar simultáneamente los  $\beta_i$  y sus correspondientes  $\text{Exp}(\beta_i)$ .

Donde, además:

Entonces:

- Si  $\beta_i > 0$ , implica que el OR  $> 0$ . (la variable  $X_i$  aumenta la probabilidad de impago).
- Si  $\beta_i < 0$ , implica que el OR  $< 0$ . (la variable  $X_i$  disminuye la probabilidad de impago).
- Si  $\beta_i = 0$ , implica que el OR = 1 (la variable  $X_i$  no ejerce ningún efecto sobre la probabilidad de impago).

$X_1$  = Número de Cuotas Vencidas. Por cada unidad de incremento del número de cuotas vencidas, es 0.001 veces menos probable que el cliente caiga en default, frente a no default.

$X_2$ = Número de Días de Morosidad. Por cada día de aumento en la Morosidad del cliente, es 462,998.4 veces más probable que el cliente caiga en default, frente a no default.

$X_3$ = Número de Cuotas del Crédito. Por cada unidad de incremento del número de cuotas del Crédito, es 0,003 veces menos probable que el cliente caiga en default, frente a no default.

$X_4$ = Tasa de Crédito. Por cada unidad de incremento en la tasa del crédito, es 0,000 veces menos probable que el cliente caiga en default, frente a no default.

$X_5$ = Edad. Por cada año de aumento en la edad del cliente, es 20.4 veces más probable que el cliente caiga en default, frente a no default.

$X_6$ = Cuotas Pagadas. Por cada unidad de incremento en el número de cuotas pagadas del cliente, es 0,028 veces menos probable que el cliente caiga en default, frente a no default.

$X_7$ = Antigüedad del Cliente en años. Por cada año de aumento en la antigüedad del cliente, es 1.571 veces más probable que el cliente caiga en default, frente a no default.

$X_8$ = Número de Créditos en la Institución. Por cada unidad de aumento del número de créditos del cliente, es 2.716 veces más probable que el cliente caiga en default, frente a no default.

$X_9$ = Variación Anual del IPC. Por cada unidad de incremento en el IPC anual, es 0,000 veces menos probable que el cliente caiga en default, frente a no default.

$X_{10}$  = Sector Económico (1). Que un cliente no pague su deuda (default) es 1.054 veces más probable en todos que los clientes que se dedican al comercio.

$X_{11}$  = Estado Civil (1). Que un cliente no pague su deuda (default) es 1,689 veces más probable en los soltero que los clientes casados, viudos, divorciados, etc.

$X_{12}$ = Sexo (1). Que un cliente no pague su deuda (default) es 2.419 veces más probable en los hombres que las mujeres. La tabla 8 muestra los resultados de la aplicación del modelo para cada una de las variables.

Tabla 8. Variables en la ecuación.

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
-7,181	4,086	3,088	1	0,079	0,001	0	2,29	
Número de Días de Morosidad	13,045	1,625	64,419	1	0	462998,4	19145	11196913
Número de Cuotas del Crédito	-5,803	3,366	2,972	1	0,085	0,003	0	2,214
Tasa de Crédito	-15,25	4,189	13,257	1	0	0	0	0,001
Edad	3,014	3,116	0,936	1	0,333	20,4	0	9155
Cuotas Pagadas	-3,584	2,082	2,963	1	0,085	0,028	0	1,644
Antigüedad del Cliente	0,452	1,295	0,122	1	0,727	1,571	0,124	19,892
Número de Créditos en la Institución	0,999	1,573	0,403	1	0,525	2,716	0,124	59,289
Variación Anual del IPC.	-21,49	12,309	3,047	1	0,081	0	0	14,01
Sector Económico (1)	0,052	0,397	0,017	1	0,895	1,054	0,484	2,294
Estado Civil (1)	0,524	0,415	1,595	1	0,207	1,689	0,749	3,807
Sexo(1)	0,883	0,406	4,725	1	0,03	2,419	1,091	5,365
Constante	16,373	6,307	6,738	1	0,009	12903240	16,373	6,307

Fuente: Elaboración propia a partir del Módulo de regresión logística binaria del **Software** SPSS versión 24.

Finalmente, dado el valor de las variables independientes, la probabilidad de **default**, fue calculada directamente de la siguiente forma (Ecuación 5 y 6):

$$p = P(\text{default}) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \quad (5)$$

Donde:

$$z = 16,373 - 7,181x_1 + 13,045x_2 - 5,803x_3 - 15,253x_4 + 3,014x_5 - 3,584x_6 + 0,452x_7 + 0,999x_8 - 21,485x_9 + 0,052x_{10} + 0,524x_{11} + 0,883x_{12} \quad (6)$$

Además:

$X_1$ = Número de Cuotas Vencidas

$X_2$ = Número de Días de Morosidad

$X_3$ = Número de Cuotas del Crédito

$X_4$ = Tasa de Crédito

$X_5$ = Edad

$X_6$ = Cuotas Pagadas

$X_7$ = Antigüedad del Cliente en años

$X_8$ = Número de Créditos en la Institución

$X_9$ = Variación Anual del IPC.

$X_{10}$  = Sector Económico (1)

$X_{11}$  = Estado Civil (1)

$X_{12}$ = Sexo (1)

El tratamiento de la base de datos de la cartera de microcréditos de la Caja Municipal de la ciudad de Piura, mediante el módulo de regresión logística binaria del **Software** SPSS versión 24, se obtuvo como resultado la probabilidad de **default**, conjuntamente con la ecuación de puntuación. En cuanto a la valoración del modelo y a la vista de los coeficientes del mismo mostrados en la tabla 05, se observa que 05 variables influyen positivamente y 07 influyen negativamente en la probabilidad de que un cliente sea moroso. Cabe destacar, que la variable "Variación Anual del IPC", correspondiente al grupo de variables macroeconómicas, forma parte de este modelo.

Respecto a la bondad de ajuste, según la tabla 03, se encontró que los  $R^2$  de Cox y Snell y  $R^2$  de Nagelkerke, fueron de 50.7% y 82.9%; respectivamente. Estos indicadores muestran que el modelo explica el 82.9% de la variabilidad de la variable dependiente (**default**). A su vez, es interesante valorar la capacidad o eficacia predictiva del modelo, lo que se realiza a partir de la tabla 04. La sensibilidad del modelo fue del 86.7%. Así mismo, el porcentaje correcto de clasificación obtenido para este modelo fue del 96.7%, valor que es concordante con área del 0.966 obtenida en la figura 01 de la curva ROC.

## CONCLUSIONES

Una revisión bibliográfica de la literatura sobre modelos para instituciones de Microfinanzas ha puesto de manifiesto la insuficiencia de trabajos en este campo de estudio, lo que permite afirmar la necesidad de investigaciones futuras que aborden este tema y establezcan propuestas de modelos eficientes. En el presente trabajo se ha planteado un modelo de **credit scoring** para la cartera de microcréditos de una Caja Municipal de la ciudad de Piura; en la que, aplicando la regresión logística binaria, se ha logrado obtener un modelo de calificación estadística capaz de predecir correctamente el 96.7% de los créditos de la cartera de la Caja Municipal.

## REFERENCIAS BIBLIOGRÁFICAS

Bapuji, H., de Bakker, F., Brown, J., Higgins, C., Rehbein, K., & Spicer, A. (2020). Business and society research in times of the corona crisis. *Business and Society*, 59(6), 10+17-1078. <https://doi.org/10.1177/0007650320921172>.

- Becerra, M. F., Valencia, E. G., & Revelo, R. E. (2021). Análisis del desempleo durante la pandemia COVID-19 y el impacto en diferentes sectores económicos del Ecuador. *593 Digital Publisher CEIT*, 6(3), 442-451. <https://doi.org/10.33386/593dp.2021.3.454>
- Bennouna, G., & Tkiouat, M. (2019). Scoring in microfinance: credit risk management tool—Case of Morocco. *Procedia computer science*, 148, 522-531.
- Berlanga, V. (2014). Cómo obtener un Modelo de Regresión Logística Binaria con SPSS. *REIRE : revista d'innovació i recerca en educació*, 7(2), 105-118. <http://hdl.handle.net/11162/111356>
- Bermúdez, J. M. M., Palacios, A. J. P., Farias, G. K. A., & León, J. A. T. (2020). La contratación pública en el Ecuador, situación actual ante la pandemia de Covid-19. *Universidad y Sociedad*, 12(S (1)), 217-224.
- Comisión Económica para América Latina (CEPAL). (2020). El desafío social en tiempos del COVID-19. *Informe Especial Covid 19*, 1-22.
- Domínguez, S. A. (2019). El odds ratio y su interpretación como magnitud del efecto en investigación. *Revista Educación Médica*, 19(1 Enero - Febrero 2018), 65-66. [www.elsevier.es/edumed](http://www.elsevier.es/edumed)
- España, S. (2020). La pandemia deja al 83% de los trabajadores en Ecuador en el desempleo o con condiciones precarias. *El País*.
- Fang, F., & Chen, Y. (2019). A new approach for credit scoring by directly maximizing the Kolmogorov–Smirnov statistic. *Computational Statistics & Data Analysis*(133), 180-194.
- García, L. (2021). COVID-19 y educación a distancia digital: preconfinamiento, confinamiento y posconfinamiento *RIED. Revista Iberoamericana de Educación a Distancia*, 24(1), 09-32. <https://doi.org/http://dx.doi.org/10.5944/ried.24.1.28080>
- Gici , A., & Subasi, A. (2019). Calificación crediticia para un conjunto de datos de microcrédito utilizando la técnica de sobremuestreo de minorías sintéticas y clasificadores de conjuntos. *Expert Systems* , 36 (2), e12363.
- Iturralde, C., & Duque, L. (2021). Precarización del teletrabajo en Ecuador en contexto de covid-19: variables de análisis desde el enfoque marxista. *Revista Chakiñan de Ciencias Sociales y Humanidades*(14), 146-162.
- La Hora. (2021). *Ambato: sector comercial no superó ni el 60% de ventas prepandemia*. <https://www.lahora.com.ec/categoria/tungurahua/destacado-tungurahua/>
- Nalić, J., & Švraka, A. (2018). Using data mining approaches to build credit scoring model: Case study—Implementation of credit scoring model in microfinance institution. *In 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*, (pp. 1-5). IEEE.
- Pérez, S. (2020). ¿Cuál fue el desempeño del sistema financiero peruano durante la cuarentena? *Diario Gestión*. <https://www.gestion.pe>
- Quintero, J. J. (2020). El Efecto del COVID-19 en la Economía y la Educación: Estrategias para la Educación Virtual de Colombia. *Revista Scientific*, 5(17), 280-291. <https://doi.org/10.29394/Scientific.issn.2542-2987.2020.5.17.15.280-291>
- Ruiz, S., Gomes, P., Rodrigues, L., & Gama, J. (2017). Credit Scoring in Microfinance Using Non-traditional Data. *In EPIA Conference on Artificial Intelligence*, (pp. 447-458). Springer, Cham.
- Superintendencia de Banca y Seguro. (2021). Sistema financiero: la pandemia y su impacto en los modelos de gestión de riesgo de crédito *Boletín Semanal SBS Informa*(29 Julio 2021). <https://sbs.gob.pe>
- Viswanathan, P. K., & Shanthi, S. K. (2017). Modelling Credit Default in Microfinance—An Indian Case Study. *Journal of Emerging Market Finance*, 16(3), 246-258.