



UN MODELO DE CALIFICACIÓN ESTADÍSTICA PARA UNA INSTITUCIÓN MEXICANA ESPECIALIZADA EN MICROCRÉDITOS

Rafael Morales Ibarra. rmoralesi@uaemex.mx ^{*1}

Ma. Eugenia Valdez Pérez*

Gandhi González Guerrero*

Universidad Autónomas del Estado de México

México

Resumen.

El presente trabajo aborda desde un contexto regional el funcionamiento de las instituciones microfinancieras, catalogadas como entidades de apoyo al combate de la pobreza y bajo el supuesto de que los microcréditos son empleados para emprender y desarrollar alguna actividad productiva que permita mejorar las condiciones económicas de los usuarios de dicho servicios.

Los microcréditos fueron concebidos como un mecanismo de ayuda a la población más pobre del planeta, no obstante su carácter social, las microfinancieras no están exentas de enfrentar altos índices de morosidad, por ello el presente trabajo se fija como objetivo construir un modelo de calificación estadística o “*Credit Scoring*” una herramienta que analiza los factores y características de los usuarios, así como de las cantidades prestadas lo que vendría a apoyar la toma de decisiones en instituciones microfinancieras. El modelo se asume bajo el principio de parsimonia y una buena potencia predictiva para préstamos en plazos menores a un año.

PALABRAS CLAVE: MICROCRÉDITO, RIESGO, *CREDIT SCORING*, MICROFINANZAS.

¹ Profesores investigadores del Centro Universitario UAEM Tenancingo, de la Universidad Autónoma del Estado de México



Abstract.

This work is from a regional context the operation of microfinance institutions, considered combat support companies of poverty and under the assumption that microcredit will be used for productive activities to improve the economic conditions of the users of such services.

The credits in smaller amounts were conceived as a mechanism to assist the world's poorest people, despite their social, microfinance institutions are not exempt from face high default rates, so the present work sets the goal to build a scoring model or "Credit Scoring" to analyze the factors and characteristics of users and the amounts loaned, which would support decision making in microfinance institutions. The model assumes the principle of parsimony and good predictive power for loans in terms of less than one year.

Keywords: microcredit, risk, credit scoring, microfinance.

Introducción

El empleo de los modelos de *credit scoring* aplicados a instituciones microfinanzas (MF) en México no son muy comunes debido a poca información que manejan dichas organizaciones. Por lo que el objetivo es construir un modelo de *credit scoring* para una microfinanciera ubicada en el sur del Estado de México, empleando variables y características de los clientes donde se muestra que dichas técnicas no desplazan de manera absoluta el juicio humano, debido a que cuando un cliente obtiene puntaje adverso el préstamo no se otorga.

Las técnicas comúnmente empleadas en los modelos de *Credit Scoring* han sido el análisis multivariante tales como el análisis discriminante, componentes principales, clases latentes. Técnicas no paramétricas; redes neuronales, modelos logit, etc. En este ejercicio, se emplea la regresión logística, una herramienta estadística que ayuda a delinear un modelo de *credit scoring* aplicada a una institución microfinanciera, el modelo busca estimar la probabilidad de



impago de futuros créditos potenciales. Se toma una muestra de 100 microcréditos, y 40 variables iniciales, aplicando una combinación adecuada de los conceptos de sensibilidad.

Destacan dos variables macroeconómicas estadísticamente significativas, el tipo de interés (Ti) y el producto interno bruto (PIB). Se observa un ajuste aceptable al emplear el test de Hosmer y Lemeshow, así como R^2 de Cox y Snell y R^2 de Nagelkerk. La variabilidad mínima de los β 's indica que los "outliers" no son significativos. Por otra parte, la variables predictoras del modelo se agrupan en tres bloques: perfil del cliente, tipología de los créditos y variables macroeconómicas. Al incorporar las variables macroeconómicas mejora sustancialmente los resultados del modelo.

El Problema.

El empleo del *credit scoring* en la evaluación del riesgo de portafolios financieros, tiene sus orígenes en los 70's sin embargo, estas técnicas se hacen más populares en los 90's. Generalmente asociados a modelos de minería de datos y en general a técnicas estadístico-matemática. Aplicaciones de *credit scoring* en microfinanzas no son comunes, principalmente porque se tiene una limitada información, lo que impide construir modelos con cierta bondad de ajuste, hecho que se refleja en la escasa literatura y/o trabajos empíricos sobre modelos de *credit scoring* en microfinancieras y todavía más escaso la aplicación de técnicas a nivel regional.

En actividades financieras, históricamente subsiste el riesgo de recuperar los montos financiados, las microfinanzas no son la excepción, el problema de la morosidad tiene el mismo nivel de importancia que en cualquier otra institución bancaria, por ello tanto las grandes instituciones financieras como las microfinanciera han requerido de eficientes mecanismos que reduzcan los índices



del no pago, el *credit scoring* viene a coadyuvar al criterio humano en la difícil tarea de filtrar solicitudes de crédito.

El presente ejercicio construye un modelo de *credit scoring* para una microfinanciera (MF)², con el empleo de variables y características de los usuarios de dichos servicios financieros, así como de montos prestados. Se identifica que las técnicas aplicadas no sustituyen de forma determinante el juicio y criterio en la decisión de autorizar o no un microcrédito, si bien el modelo “*recomienda*” que los clientes calificados con puntajes inferiores no son merecedores del crédito, Se tiene que los modelos de *credit scoring* en microfinancieras ayudan a eliminar los tiempos perdidos en papeleos por parte del cliente, lo que lleva a eficientar las funciones de cobranza de las instituciones.

Marco Teórico

Los microcréditos son pequeños préstamos realizados a personas de menores ingresos o pobres, para *Muhammad Yunus* indica que estos pequeños préstamos pueden producir cambios sustanciales en las condiciones de vida de la población más desfavorecida que es igualmente el estrato poblacional que los bancos tradicionales no han atendido. La esencia de los microcréditos radica en que las personas sin recursos puedan acceder a financiamiento de proyectos productivos que conlleven a mejorar sus condiciones de vida.

El microcrédito es la parte fundamental de la microfinanciación, dentro del que se encuentran otros servicios tales como los microseguros, ahorros u otros. Generalmente se otorgan mediante el modelo de créditos grupales, dicho modelo permite asegurar la recuperación de los préstamos, mediante el sistema de “*grupos solidarios*”; los integrantes del grupo actúan para garantizar la devolución

²Sucursal de una Institución microfinanciera ubicada en la región sur del Estado de México (CompartamosBanco)



del préstamo y se apoyan unos a otros en el esfuerzo de mejorar sus condiciones de vida.

A mediados de los 70's, aparecen Organismos No Gubernamentales (ONG) como instituciones dedicadas a otorgar créditos con fondos que provenientes de donaciones realizadas por particulares y fundamentalmente dedicadas a administrar dichos fondos, para los 80's se identifica a los micropréstamos como una actividad rentable, lo que origina que bancos locales fondearan a dichos organismos, lo que lleva en los 90's a involucrar directamente a los bancos privados, dando origen a las microfinancieras privadas.

Para los grandes bancos comerciales como para las microfinancieras, el tema de cartera vencida o tasas de morosidad es muy importante, por ello contar con un sistema de evaluación de solicitudes de crédito, préstamos hipotecarios, al consumo, otorgamiento de tarjetas de crédito etc. Permitirá cuantificar en términos probabilísticos la posibilidad de que un cliente no cumpla o se atrase con sus obligaciones.

Para Hand and Henley (1997) el *credit scoring* es un procedimiento estadístico empleado para clasificar a los solicitantes de un crédito como usuario de riesgo 'bueno' y 'malo'. Mientras que para Schreiner (2011) el término *Scoring* se refiere al uso de conocimiento sobre desempeño y características de un préstamos en el pasado para pronosticar el desempeño de préstamos en el futuro y más específicamente, refiere que el *scoring estadístico* es uso del conocimiento cuantitativo del desempeño y características de un préstamos pasados registrados en una base de datos electrónica.

Así, cuando el analista de crédito compara una solicitud de crédito en el presente contra otra solicitud parecida, está aplicando *scoring*, aunque sea un *scoring* estadístico o subjetivo. De igual manera, cuando una microfinanciera aplica la política de no renovar préstamos a clientes que han tenido atrasos en su préstamo, está aplicando *scoring*, aunque sea un *scoring* simple y unidimensional.



El *credit scoring*, suele ser llamado también *score-cards* o *classifiers*, en los que todos ellos son algoritmos que ayudan en la evaluación del riesgo de otorgar un crédito, uno de los aspectos que lo diferencia de herramientas de medición del riesgo de créditos tales como los *modelos de cartera* y los *VaR marginales*, análisis discriminante, que entre otras cosas son correlacionan la calidad crediticia de los deudores contra una cartera de préstamos, este último una técnica estadística con buena potencia.

Por lo tanto, la calificación de créditos se aplica para obtener una caracterización de un cliente en términos de los siguientes factores:

- a) El historial financiero en los productos solicitados e índice de morosidad
- b) Relación riesgo - rentabilidad.
- c) Coste de la operación.

Metodología

Como ya se ha dicho, el *scoring* es una técnica cuantitativa multivariante que analiza el desempeño y características del historial de préstamos con el fin de pronosticar el desempeño de préstamos futuros. La evaluación del riesgo de incumplimiento de pago de quienes trabajan por cuenta propia es uno de los desafíos más importantes de las microfinanzas. Algunas de las novedades han sido el trabajo con grupos solidarios y su evaluación detallada del solicitante.

Tradicionalmente los modelos de *Credit Scoring* en la banca comercial son muy comunes, las técnicas comúnmente empleados son; Análisis Discriminante (AD) técnica multivariante que estudia simultáneamente el comportamiento de un grupo de variables independientes con la intención de clasificar una serie de casos en grupos previamente definidos y excluyentes entre sí (Fisher, 1936).



Cartagena de Indias, Colombia, 8, 9, y 10 de agosto de 2013

De igual manera los modelos de probabilidad lineal (regresión mínimo cuadrática) con variable dependiente dicotómica en la cual $Y=1$ si un cliente es fallido y $Y=0$ si el cliente cumple con su obligación de pago, según Orgler, (1971) al emplear estos modelos un cliente se puede clasificar en un grupo de liquidez, rentabilidad, apalancamiento y actividad.

Un tercer método empleado en los *Credit Scoring* han sido los modelos de regresión logística, los cuales tienen la bondad de calcular la probabilidad que un cliente pague o no un crédito.

Hand (1981), Showers y Chakrin (1981) y Kolesar y Showers (1985) son algunos autores que mejor uso han dado a la metodología no paramétrica conocida como la Programación Lineal (PL), modelos que presentan una buena validez cuando se desconoce la forma que pueda mantener la relación funcional entre las variables, al tiempo de permitir una asignación de *rating* sin olvidarse del criterio de optimización de clientes.

Algunas aplicaciones igualmente catalogadas como no paramétricas son las redes neuronales artificiales que han sido empleadas por Davis, Edelman y Gammerman (1992), Ripley (1994) y Rosenberg y Gleit (1994). El proceso de *credit scoring* mediante el uso de esta técnica resulta complicado el hecho de buscar modelar el sistema nervioso, requiere de conocimientos especializados, en la actualidad el avance tecnológico ha permitido el diseño de sistemas avanzados para la ordenación clientes potenciales a un crédito.

La técnica de *credit scoring* bajo la lógica de los Árboles de Decisión ofrece de manera similar un buen perfil de riesgo, así lo concluye el trabajo de Breiman, Friedman, Olshen y Stone (1984), así mismo como la clasificación u ordenamiento de clientes en términos de *credit scoring*, Makowski (1985), Coffman (1986), Carter y Catlett (1987), una comparación entre técnicas paramétricas (análisis discriminantes) contra no paramétricas (árboles de decisión) se encuentra en Boyle, Crook, Hamilton y Thomas (1992).



El Modelo

Se usa la regresión logística para el modelo de *credit scoring* para una sucursal de una microfinanciera asentada en la región sur del Estado de México, el modelo econométrico tiene fines explicativos como predictiva, es decir, busca estimar la probabilidad de impago de futuros créditos potenciales. Se toma una muestra de 100 microcréditos, y 40 variables iniciales, aplicando una combinación adecuada de los conceptos de *sensibilidad y especificidad*; Se entiende por sensibilidad la probabilidad de clasificar correctamente a un cliente que atiende el pago de sus deudas. Por el contrario, la especificidad es la probabilidad de clasificar correctamente a un cliente fallido.

El planteamiento de un modelo cuya variable respuesta o dependiente es binaria –el cliente paga (0) contra el cliente no paga (1) – insta optar por un modelo de Regresión Logística Binaria, debido a:

1. Cuenta con propiedades estadísticas más adecuadas comparada con las que ofrecen los modelos lineales.
2. Regularmente proporcionan estimadores ineficientes.
3. La información cualitativa complementa la escasez de variables cuantitativas.
4. La Regresión Logística admite variables categóricas
5. El método permite estimar la probabilidad del no pago empleando los *Odds Ratio* (OR)³.

El modelo de regresión logística puede formularse como:

$$\text{Log} \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_{k1} \quad (2)$$

³ Éste se define como $\exp(\beta)$, donde \exp es la base de los logaritmos neperianos (una constante cuyo valor es 2,718), y β es el valor del parámetro de regresión de la variable independiente en el modelo. Una OR mayor que 1 indica un aumento en la probabilidad del evento de incumplimiento sobre el hecho de pagar cuando la variable explicativa aumenta en una unidad; inversamente, una OR menor que 1 indica lo contrario.

Dónde:

P = probabilidad de ocurrencia del evento de interés, en nuestro caso, probabilidad de que el cliente incurra en no pago.

Dicha probabilidad se obtiene:

$$p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \quad (3)$$

Donde:

$$Z = \beta_0 + \beta_1x_1 + \dots + \beta_kx_{k1} \quad (4)$$

Resultados

Con la base de datos de 100 microcréditos de la MF y con el empleo de la Regresión Logística Binaria, se obtiene como resultado la siguiente ecuación:

$$Z = 3.183 - 0.32Uso_{Cred} - 0.987Lab_{Usf} - 3.062Casa_i - 0.801Pib - 1.13Cta_{Loc} - 2.016Tip_{Garant} - 3.67Ti + 0.147Ad_{cred} \quad (5)$$

Empleando (5) y (6)

$$P(\text{probabilidad de incumplimiento}) = \frac{1}{1+e^{-z}} \quad (6)$$

Se obtiene la probabilidad de no pago P para cada prestatario.

La relación de variables estadísticamente significativas es:

Definición	Variable	Parámetro	p -value ($\alpha=0.05$)
Cuál es el destino o uso del crédito	Uso_cred	-0.324	(0.000)
Actual situación laboral del cliente	Lab_Usf	+0.019	(0.025)
Número de atrasos anteriores	Atr_pas	-6.331	(0.055)
Número de cuotas atrasadas	Num_atrss	-2.795	(0.0670)



Cartagena de Indias, Colombia, 8, 9, y 10 de agosto de 2013

Comportamiento de la economía a nivel nacional	Pib	+1.891	(0.000)
Tasas de interés en el mercado	Ti_m	-0.298	(0.844)
El cliente cuenta con casa propia	Casa_i	+0.423	(0.005)
Estado civil del cliente	E_civil	+0.002	(0.060)
Número de préstamos anteriores	Num_cred	-4.893	(0.055)
Años funcionando del negocio o actividad económica actual	Num_exp	-0.078	(0.667)
Cuenta con local distinto a su casa	Cta_local	+0.689	(0.000)
Ingresos netos	Ynet_s	-0.069	(0.500)
Monto solicitado	Mon_s	+0.0226	(0.650)
Monto otorgado	Mon_o	-0.0240	(0.650)
Tipo de garantía	Tip_garant	-3.561	(0.035)
Analista de crédito	Ad_cred	-2.041	(0.000)
Nivel de estudios del analista	N_est	-0.0027	(0.500)

Nota: las variables exógenas que resultaron estadísticamente significativas presentan signo esperado de influencia en la variable explicativa, por lo que se considera su influencia en el comportamiento de pago del cliente.

El modelo final queda definido de la siguiente manera:

$$Z = 3.183 - 0.32Uso_Cred - 0.987Lab_Usf - 3.062Casa_i - 0.801Pib - 1.13Cta_Loc - 2.016Tip_Garant - 3.67Ti + 0.147Ad_cred \quad (7)$$

Se destaca dos variables macroeconómicas estadísticamente significativas, el tipo de interés (Ti) y el comportamiento de la economía medido por la evolución del producto interno bruto (PIB), lo que significa que el entorno macroeconómico es importante en la evaluación del riesgo.

Se observa una variable con signo positivo y siete con signo negativo por lo que dichas variables influyen en la probabilidad de que un cliente sea clasificado como moroso⁴.

⁴ Cuando el signo es positivo significa que, al aumentar la variable independiente en una unidad, el logaritmo de las probabilidades de ser un cliente moroso aumenta en el valor del coeficiente respectivo, y viceversa para los estimadores con signo negativo.



Bondad de ajuste.

Identificadas las variables su coeficiente se incluye en la ecuación del modelo final (7) por su contribución a la mejora del estadístico χ^2 del modelo.

Para la bondad de ajuste se valoró la prueba de razón de verosimilitud, cuyo valor es de 1469,492, se calculan el R^2 de Cox y Snell y el R^2 de Nagelkerke, cuyos valores respectivos son indicativos de un ajuste aceptable en regresión logística.

El test de Hosmer y Lemeshow, no es satisfactorio por su nivel de significación, quizás el tamaño de muestra indique la razón. Se asume una muestra aleatoria y el análisis residual lo evidencia En cuanto a la independencia de los residuos del modelo de regresión logística, se asume que la muestra objeto de estudio es aleatoria, por lo que los residuos también mostrarán una pauta similar.

Realizando un ejercicio de la eficacia predictiva del modelo, el porcentaje correcto de clasificación obtenido para el modelo de *credit scoring* de la MF fue de 83.1%⁵.

Finalmente, la mínima variabilidad que presentan los β 's de las variables endógenas indica que los "outliers" hayan sido significativos.

Conclusiones

Idear y aplicar un modelo de *credit scoring* para una sucursal microfinanciera establecida en una comunidad de aproximadamente 25 mil habitantes y una cartera de mil doscientos clientes acreditados, con un índice de representatividad poblacional del 8.3%, aplicando la regresión logística binaria, una herramienta apropiada en el análisis de modelos con variables cuantitativas y cualitativas en las exógenas, se obtienen las siguientes conclusiones:

⁵ Porcentaje que mide la precisión del modelo para discriminar entre clientes que pagan y clientes morosos



**3er Simposio Internacional de Investigación en Ciencias Económicas,
Administrativas y Contables - Sociedad y Desarrollo**



Cartagena de Indias, Colombia, 8, 9, y 10 de agosto de 2013

- Las variables predictoras del modelo son del tipo cuantitativo, de igual manera se agrupan en tres bloques: perfil del cliente, tipología de los créditos y variables macroeconómicas. Al incorporar las variables macroeconómicas mejora sustancialmente los resultados del modelo.
- Se identifica una relación inversa entre el riesgo y el puntaje del score; a mayor puntaje menor riesgo de no pago del cliente.
- Las limitaciones de información se complementan con técnicas mixtas, lo que permitió la flexibilización del análisis principalmente en la inclusión de variables cualitativas en las predictivas.
- La eficacia predictiva del modelo predice correctamente en más del 83% de los créditos de la cartera de MF.
- La técnica de regresión logística tiene la ventaja de calcular la probabilidad de incumplimiento del cliente que solicita un crédito.

En resumen, el modelo de credit scoring solo es un indicativo, la determinación fundamental radica en el recurso humano de las instituciones, de igual manera es necesario que la calificación estadística debe constituirse en herramienta complementaria que permita resumir un conjunto de información cada vez más compleja de analizar y particularmente la presente aplicación requiere validación permanente para calibrar la bondad de los datos y las variables empleadas y por ende la eficiencia y eficacia, así como de la precisión en sus estimaciones de tasa de mora.



Bibliografía

- Altman, E., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy". *Journal of Finance*, September 1968.
- Boyle, M., Crook, J. N., Hamilton, R., & Thomas, L. C. (1992). *Methods for Credit Scoring Applied to Slow Payers*. In Thomas, L. C., Crook, J. N., Edelman, D. B. (Eds.), *Credit Scoring and Credit Control* (pp. 75-90). Oxford, UK: Clarendon.
- Coffman, J. Y. (1986). *The Proper Role of Tree Analysis in Forecasting the Risk Behaviour of Borrowers*. *Management Decision Systems, MDS Reports* 3, 4, 7 & 9.
- Hand, D. J. (1981). *Discrimination and Classification*. Chichester, UK: Wiley.
- Kolesar, P., & Showers, J. L. (1985). *A Robust Credit Screening Model Using Categorical Data*. *Management Science*, 31, 123-133.
- Makowski, P. (1985). *Credit Scoring Branches Out: Decision Tree-Recent Technology*. *Credit World*, 75, 30-37.
- Orgler, Y. E. (1971). *Evaluation of Bank Consumer Loans with Credit Scoring Models*. *Journal of Bank Research*, 2, 31-37.
- Rosenberg, E., & Gleit, A. (1994). *Quantitative Methods in Credit Management: A Survey*. *Operations Research*, 42, 589-613.
- Schreiner, M. (2011). *Ventajas y Desventajas del Scoring Estadístico para las Microfinanzas*. Washington University.
- _____. (2000a) *A Scoring Model of the Risk of Costly Arrears for Loans from Affiliates of Women's World Banking in Colombia*, reporte para el Banco Mundial de la Mujer.