

“SISTEMA DE APROBACIÓN DE MICROCRÉDITO BASADO EN PATRONES DE POBREZA”

Fabrício Echeverría ¹, María Eugenia Andrade Ramírez ², Enrique Guido Anchundia Aguirre ³, Néstor Alejandro Uyaguari Espinoza ⁴.

¹ Director de Tópico, Magíster en Sistemas de Información Gerencial, 2006, Profesor de ESPOL; e-mail: pechever@espol.edu.ec

² Ingeniera en Computación Especialización Sistemas de Información, 2006; e-mail: mandrade@espol.edu.ec

³ Ingeniero en Computación Especialización Sistemas Tecnológicos, 2006; e-mail: eanchund@espol.edu.ec

⁴ Ingeniero en Computación Especialización Sistemas Tecnológicos, 2006; e-mail: nuyaguar@fiec.espol.edu.ec

Resumen

Una de las discusiones más relevantes en las microfinanzas actuales se refiere a la necesidad de un enfoque de los servicios microfinancieros a los más pobres, dando lugar, a la superación económica y social de dichos individuos en las sociedades del mundo, especialmente en el Ecuador. Por tal motivo, seguros de aportar con una herramienta útil para este fin, el presente trabajo esta basado en el desarrollo de un sistema de aprobación de microcrédito basado en patrones de pobreza (SAMBP).

El “SAMBP” pretende facilitar y enfocar la aprobación de microcrédito a las clases más desprotegidas de la sociedad ecuatoriana, mediante la clasificación de los individuos como “pobres” o “muy pobres”, utilizando una de la metodologías propuestas por la Cumbre de Microcrédito desde sus inicios en 1997, denominado Índice de la Vivienda de Cashpor, el cual analiza las características físicas de la vivienda; y adicionalmente, usando técnicas de minería de datos descriptivas y predictivas para la extracción de patrones de pobreza tales como: el algoritmo Enhanced K-Means y el algoritmo ABN (Adaptive Bayes Network) respectivamente.

Palabras Claves: *microfinanzas, enfoque en los más pobres, microcrédito, Enhanced K-Means, patrones de pobreza, Cashpor, ABN.*

Abstract

One of the most important discussions in current micro-finance theory is the necessity to focus micro-financial services to the poorest, helping to the economic and social improvement of the people in the societies of the world, in Ecuador specially. Therefore, ensure of contribute with a helpful tool; the present work is based on the development of a “Microcredit Approval System poverty model-based”.

The “SAMBP” pretend to facilitate and focus the microcredit approval to the Ecuadorians poorest, through people classification as “poor” or “poorest”, using one of the methodologies proposed by The Microcredit Summit since 1997, named Cashpor Housing Index, which analyze the physical features of the housing; and moreover, using descriptive and predictive data mining techniques for the extraction of poor model such as: Enhanced K-Means algorithm and ABN (Adaptive Bayes Network) algorithm respectively.

Keywords: *micro-finance, focusing in the poorest, microcredit, Enhanced K-Means, poor models, Cashpor, ABN*

1. Introducción

La pobreza es uno de los grandes males que aquejan a las sociedades del mundo, impidiendo el desarrollo económico y social de sus habitantes. A pesar de que instituciones financieras poseen programas de microcrédito, la realidad muestra que no benefician a los más desfavorecidos. Dando lugar, a la necesidad de establecer un enfoque hacia los grupos considerados vulnerables solucionando el error de excluirlos.

Por tal motivo, a partir de 1997, la cumbre de microcrédito establece la importancia que el microcrédito sea una salida para los sectores de menores ingresos, que por lo general, no tienen acceso a ningún tipo de financiamiento bancario, debido que no cumplen con los requisitos que normalmente una institución financiera de primer plano solicita. Por medio del uso de metodologías que ayuden a indicar la medida de pobreza de una persona, tales como: el índice de la vivienda de Cashpor y la calificación participativa del patrimonio de Small Enterprise Foundation. [1]

Dado este conocimiento “El sistema de aprobación de microcrédito basado en patrones de pobreza” (SAMBP), utiliza como herramienta una de las metodologías para medir los niveles de pobreza denominado Índice de Cashpor, la cual se basa en el análisis de la estructura de la vivienda como medio para distinguir los niveles económicos de los hogares y para identificar a los más pobres. Y a su vez, complementada con la utilización de técnicas predictivas y descriptivas de minería de datos tales como el algoritmo Enhanced K-means y ABN, dan lugar a la generación de patrones de pobreza que proporcionan el conocimiento necesario para establecer si una persona puede ser considerada para la aprobación de microcrédito. Por lo que “SAMBP” logra ser una herramienta útil de análisis de pobreza, y medio referencial de decisión para aprobar microcrédito a los más pobres.

Por lo que por medio de este sistema se espera adicionalmente, fomentar la iniciativa que en el Ecuador las instituciones financieras orienten sus servicios hacia la clase más desprotegida de la sociedad. Considerándola como una forma de retribuir la confianza que la sociedad les ha otorgado durante años, mostrando que poseen responsabilidad social corporativa.

1. Definición del Problema y aplicación

El problema consiste en determinar la condición económica de personas específicamente de bajos recursos; en base a patrones de pobreza, extraídos mediante el análisis de las características de la vivienda, tales como: material de las paredes, techo, piso, tenencia de la vivienda etc. Para su posterior aprobación de un posible microcrédito. Es así que el proceso de aprobación de microcrédito comienza, extrayendo información de viviendas analizadas en base a la metodología de Cashpor (estudio de características tales como: tipo vivienda, paredes, techo, piso, tenencia, agua beber, sanitario, basura y combustible). Con dichos datos se procede a la construcción de modelos y/o patrones de pobreza que servirán posteriormente para clasificar nuevas vivienda. Cabe recalcar que los patrones mostrarán su porcentaje de certeza, con lo cual se aumenta la confianza de los resultados extraídos.

El proceso termina cuando el cliente, el cual desea saber si puede ser un posible candidato para la aprobación de un microcrédito, es ingresado por medio de un formulario donde se anotan sus datos personales, seguida de los datos geográficos y características de la vivienda anteriormente señalados. Para luego, mediante los patrones de pobreza previamente extraídos clasificar dicha persona como “pobre” o “muy pobre”, brindado a la persona encargada de la aprobación de microcrédito, información referencial de una posible decisión.

2. Algoritmo Predictivo ABN y su aplicación en la aprobación de microcrédito

Algoritmo Adaptivo de Redes Bayesianas (ABN) es un algoritmo de Oracle Data Mining, provee una rápida y escalable extracción de información predictiva de los datos con respecto a la clase o atributo a predecir. El algoritmo puede describir su modelo en forma de reglas humanamente legibles. Por lo tanto, las reglas producidas por ABN es una de sus principales ventajas debido a que provee de un modelo transparente de reglas que pueden ser entendidas por analistas de mercado o de negocio. ABN predice valores binarios¹ y multiclases², así

mismo usa costos y probabilidades para construir y clasificar. [2]

La aplicación del algoritmo ABN en el sistema “SAMBP” consiste primero en realizar un análisis de cada atributo para establecer que tipo de normalización (convertir los valores de los atributos en rangos comunes) se aplica. En este caso se aplica el método de normalización Top-N, el cual es usado para normalizar valores categóricos (cualitativos) de acuerdo a su frecuencia. Los valores que son excluidos de la normalización son las claves de la tablas (ID _ VIVIENDA) y el atributo a predecir (CONDICION), consecuentemente se construye un modelo en base al atributo predictor, el cual contiene los valores de “pobre” o “muy pobre”, generando reglas/patrones de predicción con su respectivo porcentaje de confianza, cobertura y certeza. (Ver Figura 1).

Detalle de la Reglas:	
Cobertura:	0.4785
Confianza:	0.9978
Antecedente:	
PISO isIn	Concreto/Cemento
COMBUSTIBLE isIn	Gas
TIPO_VIVIENDA isIn	Casa individual
Consecuente:	
CONDICION <i>equal pobre</i>	

Figura 1. Patrones ABN

Finalmente, se aplica el modelo/patrón extraído para predecir la “CONDICION” de nuevas viviendas. Dando como resultado las probabilidades para cada uno de los valores predictores. (Ver Figura 2).

Entendiéndose como resultado favorable para la aprobación de microcrédito, aquel cuya probabilidad del atributo predictor “muy pobre” es mayor.

ID _ VIVIENDA	PREDICCION	PROBABILIDAD
6363	muy pobre	0.1518
6363	pobre	0.8482

Figura 2. Resultados aplicación modelo ABN

3. Algoritmo Descriptivo Enhanced K-Means y su aplicación en la aprobación de microcrédito

El algoritmo Enhanced K-Means es un algoritmo de Oracle Data Mining, realiza el proceso de clusterización jerárquica de datos, la cual es una técnica útil para la exploración de datos. El análisis de clusterización jerárquica identifica grupos embebidos en los datos formando un árbol de clasificación, donde cada nodo representa a un grupo que contiene descripciones probabilísticas que resume a los elementos clasificados bajo ese nodo.[3]

La aplicación del algoritmo Enhanced K-Means en el sistema “SAMBP” implica el descubrimiento de grupos de elementos,, en este caso viviendas, los cuales son usados para crear reglas que capturan las características principales de los elementos(viviendas) asignados a cada grupo. (Ver Figura 3) Cabe resaltar, que el algoritmo de Enhanced K-Means no proporciona reglas para predecir, sino para describir.

<p>If AGE >= 25 and AGE <= 40 and HEIGHT >= 5.0ft and HEIGHT <= 5.5ft then CLUSTER = 10</p>
--

Figura 3. Reglas descriptivas Enhanced K-Means

4. Resultados Obtenidos

Con la ayuda de este sistema de aprobación de microcrédito basado en patrones de pobreza, se ha logrado predecir la condición económica de una persona en base a características de la vivienda de Cashpor, de manera eficiente y confiable, debido a que el sistema gracias al algoritmo ABN genera patrones que posteriormente servirán para clasificar y/o predecir qué otras personas podrían ser catalogadas como “muy pobres” para la aprobación de microcrédito.

5. Conclusiones

Basados en los resultados y en la experiencia obtenida en la realización de éste sistema podemos concluir lo siguiente:

- 1) Gracias a la minería de datos y a los diferentes modelos que ella ofrece podemos explorar grandes volúmenes de datos, analizar archivos y bitácoras de transacciones, esta trabaja a nivel del conocimiento con el fin de descubrir patrones, relaciones, reglas, asociaciones o incluso excepciones útiles para la toma de decisiones, se podría decir que la minería de datos se encuentra muy ligada a los Data Warehouse
- 2) La minería de datos permite generar conocimiento que ayuda a mejorar la toma de decisiones en los procesos fundamentales de un negocio.
- 3) El campo de la minería de datos es un campo en pleno desarrollo, donde la mayoría de las herramientas utilizadas provienen de otros campos relacionados como el reconocimiento de patrones, la Estadística o la teoría de complejidad.
- 4) El desarrollo de la tecnología de minería de datos a su vez posee algunas trabas. Existen elementos que la hacen operable, pero por otra parte, hay factores que pueden crear un descrédito a esta tecnología, como por ejemplo: los productos a comercializar son, actualmente costosos, y los consumidores pueden hallar una relación costo/beneficio improductiva, se requiere de mucha experiencia para utilizar herramientas de la tecnología, o que sea muy fácil hallar patrones equívocos, triviales o no interesantes.

Referencias

a) Referencias de Internet

[1] Orientación de los servicios microfinancieros hacia los más pobres, www.cmq.edu.mx/documentos/Revista/revista17/est17_6.pdf

[2] Oracle Data Mining 10g Release 2 (10.2), <http://otn.oracle.com/documentation>

[3] Descriptive Data Mining Models, <http://www.stanford.edu/dept/itss/docs/oracle/10g/datamine.101/b10698/4descrip.htm#1006535>

[4] Predictive Data Mining Models, <http://www.stanford.edu/dept/itss/docs/oracle/10g/datamine.101/b10698/3predict.htm>

[5] Oracle Data Mining Administrator's Guide 10g Release 2 (10.2), <http://otn.oracle.com/documentation>

[6] Oracle Data Mining Application Developer's Guide 2 (10.2), <http://otn.oracle.com/documentation>

[7] Oracle Database Installation's Guide 10g Release 2 (10.2) for Microsoft Windows, <http://otn.oracle.com/documentation>

[8] Planet Finance, Microfinance Platform, <http://www.planetfinance.org/ES/ong-socios/ong-socios-red.php>

[9] Banca Ecuatoriana dirige su apoyo al microcrédito, www.elpanamaamerica.com.pa/archive/0213200/finance07

[10] Draft Training Manual on Cost-Effective Targeting, www.microfinancegateway.org/static/561.htm

b) Referencias de libros

- HERNÁNDEZ, J.; RAMÍREZ, M.J. y FERRI, C. Introducción a la Minería de Datos, Prentice Hall, España, primera edición, 2004.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L. y BLACK, W.C. Análisis multivariante, Prentice Hall, Madrid, 1999

¹ Valores binarios indican decisiones de SI/NO (comprar/no comprar, cobrar/no cobrar, etc.)

² Valores multiclase indican una alternativa preferida como: color de suéter, rango de sueldo)

