

VIGÉSIMO PROGRAMA DE BECAS
PARA DESARROLLO DE TESIS - 2009

“Desarrollo de un Prototipo de Credit Scoring, Utilizando Técnicas de Aprendizaje Automático”.

El alumno señala que un importante reto para las instituciones que entregan créditos a empresas o clientes individuales, es disminuir, de la mejor forma posible, el riesgo de entregar un préstamo. El problema surge al momento de evaluar al cliente: ¿Cómo evaluarlo?, ¿Qué características del mismo considerar y cuáles no?. Esto, con la finalidad de predecir el comportamiento del cliente y evitar una posible pérdida de dinero. Así, el Credit Scoring surge como un conjunto de modelos de decisiones y técnicas basadas en análisis estadístico, que instituciones financieras utilizan para evaluar la solvencia de una persona. El Credit Scoring es usado como apoyo en el momento de decidir qué cliente recibe un crédito, y el monto que le podría ser entregado.

Los modelos de Credit Scoring han sido ampliamente discutidos y empleados para resolver este problema, llegando a ser un tema importante para muchos investigadores. En la literatura se han propuesto varias técnicas para construir un modelo de Credit Scoring, las que incluyen Análisis de Discriminación Lineal (Estadística Paramétrica), Árboles de Decisión (Estadístico No-Paramétrico) y Técnicas de Inteligencia Artificial.

El tesista sostiene que en nuestro país el proceso de análisis de créditos se realiza en forma semi-manual, utilizando software o planillas de excel y bajo la mirada de un evaluador experto. Este considera una serie de variables discretas como carga financiera, nivel de endeudamiento, score SINACOFI y calificaciones del sistema financiero relacionadas con comportamientos anteriores, sobregiros no pactados, atraso en el pago de tarjetas de crédito etc. Con esta información, las instituciones pueden calcular una tasa de riesgo para el cliente en estudio, la cual determinará cuál es la probabilidad de incumplimiento del mismo.

El trabajo apunta a desarrollar un prototipo de Credit Scoring, utilizando técnicas de Inteligencia Artificial. El alumno plantea que en investigaciones recientes se han comparado Modelos Estadísticos y Redes Neuronales, demostrándose que al usar estas últimas se comete un menor porcentaje de errores –definir a un cliente como bueno cuando es malo, o viceversa- en la clasificación de solicitantes de créditos, el que se ha reducido desde el 24% obtenido con un Modelo Estadístico, a sólo 3% con la Red Neuronal. De acuerdo a un estudio citado por el tesista, las técnicas de “Simulated Annealing” y “K-nearest neighbor” arrojan mejores resultados, entre un 2% y un 30%, que los modelos actualmente usados en Credit Scoring (Análisis de Discriminación Lineal y Análisis de Discriminación Cuadrática).

Considerando lo anterior, se propone el desarrollo de un modelo de Credit Scoring, empleando un ensamble de técnicas de Aprendizaje Automático, puesto que han demostrado ser superiores a los Modelos Estadísticos tradicionales.

El alumno define los siguientes objetivos de su tesis.

Objetivo General:

- Desarrollar un prototipo de Credit Scoring, utilizando un ensamble de técnicas de aprendizaje automático.

Objetivos Específicos:

- Investigar sobre los diferentes métodos y técnicas de Inteligencia Artificial para resolver el problema de Credit Scoring.
- Construir un prototipo de sistema de Credit Scoring, utilizando información financiera real.
- Evaluar el prototipo construido.

Universidad : Universidad de Concepción
Facultad : Facultad de Ingeniería
Carrera : Ingeniero Civil Informático
Postulante : Rodrigo Ignacio Parada Vásquez
Profesor Guía: Dr. John Atkinson Abutridy

“Inferencia de Rechazos en Modelos de Riesgo Crediticio”.

El uso de los procesos de modelación del sistema de crédito (Credit Scoring), se introdujo en los años 70, generalizándose en los 90 gracias al desarrollo de mejores recursos estadísticos y computacionales. Hoy en día prácticamente todas las entidades financieras emplean estas metodologías para disminuir el riesgo y así maximizar la rentabilidad.

Entre las principales aplicaciones de estos métodos, figura “Applicant scoring”, proceso que permite a las entidades elegir un punto de corte preestablecido, con el fin de optimizar la tasa de aprobación de las solicitudes. Luego, se produce la clasificación en las categorías de riesgo “bajo” y “alto”.

Resulta importante evaluar la capacidad predictiva de cada modelo en particular. Es decir, la precisión de los puntajes que otorga a las solicitudes y las consecuencias de las decisiones de aceptar o rechazar un crédito. Entre la gama de técnicas a utilizar, se encuentra la inferencia de rechazos, la cual tiene como objetivo llevar a cabo la relación Buenos/Malos en una muestra de peticiones desechadas. Con ello, se obtiene una base de datos integrada, que contiene todas las solicitudes acogidas, más la muestra creada por el método “Reject Inference”, para, finalmente, realizar de nuevo el proceso de modelación. El objetivo de este paso es eliminar el sesgo que se produce al predecir valores de buenos y malos a partir de la base de solicitudes aceptadas, sin tener en cuenta los puntajes que daría este modelo en la base de datos de peticiones que fueron denegadas con anterioridad.

Investigaciones anteriores sugieren que la omisión de los solicitantes rechazados tiene un impacto perjudicial sobre el modelo de estimación y predicción. Por ello, es que para corregir el sesgo derivado de la selección muestral y obtener consistencia, se necesita, en principio, utilizar una muestra formada por peticiones aprobadas y denegadas.

Generalmente, el sesgo de selectividad muestral constituye un problema a la hora de desarrollar un modelo de Credit Scoring, ya que las fuentes de información suelen disponer de datos de solicitudes que terminaron transformándose en créditos otorgados luego de pasar con éxito un proceso de evaluación previo. En ese sentido, las muestras disponibles no reflejan totalmente al universo de solicitudes que se presentarán a la oficina de evaluaciones de una entidad bancaria financiera.

“Reject Inference” es un proceso mediante el cual el rendimiento de las solicitudes denegadas anteriormente, se analiza para estimar su comportamiento. Es muy importante revisar -una vez aplicado este método- cuántos “nuevos” malos se han agregado a la muestra.

La tesista señala que este es un tema de interés práctico en nuestro país, donde las pérdidas por incumplimiento que sufren las entidades financieras están aumentando, dada la coyuntura que se está viviendo.

Esta situación lo ha motivado a realizar un estudio de la inferencia de rechazos, para el cual se utilizará una base de datos de una entidad financiera especializada en otorgar créditos de consumo. Dicha base incluye a los clientes nuevos que solicitaron préstamos de consumo entre enero y diciembre de 2006, es decir, considera a toda la población, lo cual constituye una de las características para poder aplicar la inferencia de rechazos.

La alumna se ha planteado los objetivos descritos a continuación.

Objetivo General:

- Estudiar el método de “Reject Inference”, llevando a cabo la inferencia Buenos/Malos en una muestra de solicitudes rechazadas por la entidad financiera bancaria.

Objetivos Específicos:

- Estudiar Distribuciones truncadas.
- Estudiar modelos de Credit Scoring.
- Estudiar el Método de Heckman para corregir el sesgo de muestras.
- A través de software especializado, aplicar la metodología de Reject Inference.
- Comparar resultados de “Reject Inference” con el esquema de análisis tradicional en Credit Scoring.

Universidad : Universidad de Santiago de Chile
Facultad : Facultad de Ciencias
Carrera : Ingeniero Estadístico
Postulante : Nathaly Andrea Valencia Molina
Profesor Guía: Reinaldo González Saavedra