Obtención de reglas de clasificación difusas utilizando técnicas de optimización - Caso de estudio Riesgo Crediticio

PATRICIA ROSALIA JIMBO SANTANA
Tesis doctoral en cotutela defendida el **14 de julio de 2020**
realizada bajo la dirección de:

Dra. **Laura Lanzarini**
Universidad Nacional de la Plata
La Plata (Argentina)

Dr. **Aurelio F. Bariviiera**
Universitat Rovira i Virgili
Reus (España)

INTRODUCCIÓN Y MOTIVACÍON

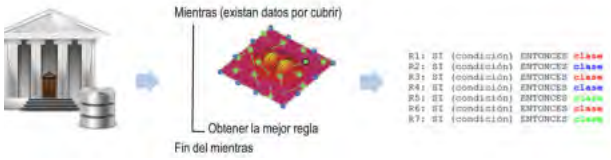
Una característica deseable de los modelos construidos con las técnicas de la Minería de Datos es que el conocimiento extraído pueda ser expresado en términos comprensibles. Esto es un aspecto fundamental ya que, quienes deben tomar decisiones, exigen conocer el fundamento de la respuesta arrojada automáticamente por el modelo. En este sentido, las reglas de clasificación difusas resultan ser sumamente útiles a la hora de justificar la respuesta obtenida ya que emplean expresiones de fácil comprensión, basadas en variables lingüísticas.
El objetivo central de esta tesis es contribuir a la Minería de Datos con un nuevo método para obtener reglas de clasificación difusas, y al área de riesgo financiero a través de la incorporación de métricas que faciliten la evaluación y selección de las reglas más adecuadas para la toma de decisiones.

APORTES DE LA TESIS

El aporte central de esta tesis es la definición de un nuevo método denominado **FRvarPSO** (Fuzzy Rules variable Particle Swarm Oprmization) capaz de generar un conjunto de reglas de clasificación difusas de fácil interpretación, baja cardinalidad y una buena precisión.

FrvarPSO combina una red neuronal competitiva con una técnica de optimización basada en cúmulo de partículas de población variable para la obtención de reglas de clasificación difusas. La conformación de las condiciones difusas requiere conocer el grado de pertenencia a los conjuntos difusos que definen a cada variable lingüística. Esta tesis propone tres alternativas distintas para resolver este punto: el criterio de un experto, intervalos de igual longitud y Fuzzy C Means. Uno de los aportes de esta tesis radica en la definición de la función fitness de cada partícula basada en un "Criterio de Votación" que pondera de manera difusa la participación de las condiciones en la conformación del antecedente.

Una transferencia tecnológica concreta fue realizar la medición sobre tres casos reales en el área de crédito del Sistema Financiero del Ecuador, asociadas al área de riesgo crediticio considerando un conjunto de variables micro y macro económicas, se realizó una consideración especial en la morosidad del cliente teniendo en cuenta los días de vencimiento de la cartera otorgada; esto fue posible debido a que se tenía información del cliente en un horizonte de tiempo, una vez que el crédito se había concedido. El resultado de aplicar el método FRvarPSO genera un conjunto de reglas difusas de baja cardinalidad logrando así reducir el tiempo de selección de la regla a aplicar por parte del oficial de riesgo. Además, dado que el antecedente de cada regla está conformado por la conjunción de condiciones difusas, el valor del menor grado de pertenencia del caso en estudio resulta una métrica proporcional al riesgo de la aplicación de dicha regla. Es decir que el oficial de crédito podrá decidir el otorgamiento del crédito al cliente que representa al menor riesgo para la institución financiera.



LINEAS DE I+D FUTURAS

Se debe considerar la optimización de la función de pertenencia, con la finalidad de identificar automáticamente los parámetros de la función.

Otra variante que se debe tener en cuenta es incorporar técnicas para el manejo del desbalance de las clases.

Resulta de interés utilizar FRvarPSO en otras áreas como por ejemplo la medicina, ya que podría ser de utilidad para determinar el riesgo que un paciente tiene en contraer una enfermedad como es el caso de Covid-19.