

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

Tesis doctoral realizada por

Patricia Rosalía Jimbo Santana

Universidad Central del Ecuador

prjimbo@uce.edu.ec

dirigida en cotutela por

Dra. Laura C. Lanzarini

Instituto de Investigación en

Informática LIDI, Facultad de

Informática,

Universidad Nacional de la Plata,

La Plata, Buenos Aires, Argentina

laural@lidi.info.unlp.edu.ar

Dr. Aurelio Fernández Bariviera

Departament of Business, Universitat

Rovira i Virgili,

Avenida de la Universitat,1 Reus, Spain

aurelio.fernandez@urv.cat

Universidades donde se desarrolló la tesis doctoral:

- **Universidad Nacional de la Plata, La Plata, Buenos Aires, Argentina**
- **Universidad Rovira i Virgili, Reus, España**

Fecha de exposición

14 de Julio 2020

RESUMEN

Esta tesis tiene por objetivo principal contribuir al área conformada por la Minería de datos y el Riesgo Financiero especialmente en el área de crédito, disponiendo de estrategias capaces de generar automáticamente reglas de clasificación difusas, resultando de suma utilidad en cualquier proceso de toma de decisiones. El método desarrollado denominado FRvarPSO, es capaz de operar sobre atributos nominales y numéricos, para obtener reglas de clasificación difusas que combina una red neuronal competitiva con una técnica de optimización basada en un cúmulo de partículas de población variable. La función de aptitud que controla el movimiento de las partículas utiliza un criterio de votación que pondera, de manera difusa, la participación de los atributos numéricos. La eficiencia y eficacia de este método se encuentran fuertemente condicionadas por la manera en que se determinen las funciones de pertenencia de cada uno de los conjuntos difusos, se consideran varias alternativas entre las que mencionamos particionando el rango de cada atributo numérico en intervalos de igual longitud, centrando, en cada uno de ellos, una función triangular con un solapamiento adecuado, utilizando el criterio del experto, y utilizando Fuzzy C-Means para la obtención de los conjuntos difusos. Uno de los principales aportes es en el área del riesgo financiero, ya que asocia el peso de la regla difusa al riesgo que puede asumir el usuario en la interpretación de la regla. Los resultados obtenidos fueron comparados con la versión anterior sobre 11 bases de datos del repositorio UCI y 3 bases de datos reales del sistema financiero ecuatoriano, una de las cuales es una cooperativa de ahorro y crédito y las dos restantes pertenecen a bancos encargados de dar crédito de consumo productivo, no productivo y microcrédito. Los resultados obtenidos fueron satisfactorios.

Palabras Claves: Riesgo, Reglas de Clasificación Difusas (Fuzzy Classification Rules), Optimización mediante Cúmulo de partículas tamaño variable (Variable Particle Swarm Optimization), Minería de datos.

1. MOTIVACION

En los últimos años gracias al avance de la tecnología, las organizaciones han almacenado gran cantidad de información. Esto las ha llevado a la necesidad de incorporar técnicas que permitan procesar y obtener información útil de los datos. El proceso de Extracción del Conocimiento, conocido como proceso de KDD (Knowledge Discovery in Databases), está formado por varias fases que van desde la recolección y transformación de los datos hasta la identificación e interpretación de patrones o relaciones subyacentes sumamente útiles en la toma de decisiones.

Dentro del proceso de KDD, la Minería de Datos (Data Mining) es considerada la fase más importante, ya que agrupa a las técnicas capaces de modelizar la información disponible. A partir del uso o comprensión del modelo generado es posible extraer conocimiento. Este conocimiento que se genera resulta de gran interés para las organizaciones, debido a que constituye una herramienta sumamente importante para la toma de decisiones tácticas y estratégicas, lo cual se convierte en una ventaja competitiva.

Según el tipo de tarea que se quiera resolver existe en la literatura una gran cantidad de técnicas disponibles entre las que podemos mencionar: las algebraicas y estadísticas, estocásticas y difusas, técnicas bayesianas, árboles de decisión, redes neuronales artificiales, máquinas de vectores de soporte y técnicas de optimización entre otras. Cada una de ellas tiene características diferentes.

Para poder seleccionar el tipo de técnica es necesario conocer el problema que se va a resolver teniendo en cuenta que no todas ellas pueden operar con datos faltantes, outliers o datos en cualquier formato. Por otra parte, los problemas que se abordan pueden ser de diferente tipología como, por ejemplo: clasificación, categorización, regresión o agrupamiento.

Una característica deseable de los modelos construidos con las técnicas de la Minería de Datos es que el conocimiento que se extrae se exprese en términos comprensibles. En este sentido, las reglas de asociación son consideradas, por quienes deben tomar decisiones, como una de las formas más comprensibles que puede ser utilizada para representar el conocimiento, ya que tienen la capacidad de explicarse por sí mismas. Una regla de asociación es una expresión de la forma "SI condic1 ENTONCES condic2" donde ambas condiciones son conjunciones de proposiciones de la forma (atributo=valor) cuyos atributos/variables pueden ser cuantitativas o cualitativas. Por su estructura condicional, puede afirmarse que el objetivo principal de una regla es identificar relaciones entre los atributos. Cuando el conjunto de reglas de asociación presenta en el consecuente el mismo atributo, se dice que se trata de un conjunto de "reglas de clasificación". Si a esto se le suma que las reglas de clasificación empleen lógica difusa a través de conjuntos difusos para describir los valores de sus atributos, se obtienen "Reglas de Clasificación Difusas", facilitando aún más su comprensión, permitiendo además el manejo de la incertidumbre, aproximándonos cada vez más al razonamiento humano. Tal como se menciona en (Terano, Asai y Sugeno, 1992), *"cuanto más humano debe ser un sistema, más lógica difusa contendrá"*.

La utilización de la lógica difusa para la generación de las reglas ha permitido que los investigadores obtengan formas de mejorar el desempeño de las metaheurísticas, básicamente para acelerar la convergencia y obtener una mejor calidad en las soluciones planteadas.

Las reglas de clasificación difusas son reglas que tienen sentencias lingüísticas que describen la manera en que los sistemas de inferencia difusa toman las decisiones sobre un conjunto de variables de entrada. Estas reglas de clasificación poseen un antecedente formado por variables nominales y/o condiciones difusas, y un consecuente conocido con anticipación denominado clase.

El aporte central de esta tesis es la definición de un nuevo método capaz de generar un conjunto de reglas de clasificación difusas de fácil interpretación, baja cardinalidad y una buena precisión. Estas características ayudan a identificar y comprender las relaciones presentes en los datos facilitando de esta forma la toma de decisiones.

2. FRvarPSO

El conocimiento que se genera utilizando técnicas de minería de datos resulta de gran interés para las organizaciones, ya que les va a permitir la toma de decisiones tácticas y estratégicas, lo que se convierte en una ventaja competitiva. Existe un conjunto de técnicas presentes en la minería de datos, cada una de ellas tiene características diferentes. Para poder seleccionar el tipo de técnica es fundamental tener conocimiento del problema que se va a resolver. En el estado del arte se pueden encontrar referencias que indican que se obtienen mejores resultados cuando se utilizan modelos híbridos, aquí se enmarca el modelo desarrollado.

El nuevo método propuesto en esta investigación se denomina **FRvarPSO (Fuzzy Rules variable Particle Swarm Optimization)**, 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, capaces de operar sobre atributos nominales y numéricos. Los antecedentes de las reglas están formados por atributos nominales y/o condiciones difusas. La conformación de estas últimas 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.

Con respecto a la forma de obtención de las reglas, el método propuesto utiliza un proceso iterativo por medio del cual se van cubriendo los ejemplos de una clase a la vez hasta lograr la cobertura deseada. Por lo tanto, el consecuente de la regla queda determinado por la clase seleccionada y es el antecedente el que se extrae a través de la técnica de optimización.

Cada partícula de la población utiliza una representación de longitud fija, donde sólo se codifica el antecedente de la regla y dado el enfoque adoptado, se efectuará un proceso iterativo asociando todos los individuos de la población con una clase predeterminada. El hecho de que todos los individuos pertenezcan a la misma clase hace innecesaria la codificación del consecuente dentro de la partícula.

El proceso de extracción de cada regla inicia con la generación del cúmulo. Las posiciones iniciales de las partículas tienen una gran incidencia en la capacidad exploratoria de cada una de ellas. Si se encuentran demasiado próximas, su conocimiento social puede provocar la convergencia prematura evitando que se alcancen buenas soluciones. Por otro lado, si inician muy lejos del óptimo se requerirán un mayor número de iteraciones o pasos de búsqueda para alcanzarlo. Por todo lo anterior, se decidió utilizar una red neuronal competitiva para identificar las zonas más prometedoras. Esta información fue utilizada para inicializar el cúmulo.

En lo que se refiere al tamaño de la población, se optó por un cúmulo de tamaño variable con capacidad para incorporar individuos en las zonas menos pobladas. Como se dijo anteriormente, en PSO las partículas tienden a reunirse alrededor de las buenas posiciones encontradas por sus vecinos esto hace que el cúmulo pierda diversidad y varias zonas del espacio de búsqueda queden despobladas. La técnica para modificar el tamaño de la población trabaja sobre dos conceptos: por un lado, detecta las zonas densas y elimina individuos muy similares y por otro identifica las zonas muy despobladas y agrega partículas adecuadamente. Esto permite iniciar con pocos individuos e ir modificando el tamaño de la población durante el proceso iterativo de búsqueda.

Volviendo a la representación de la partícula, a la información referida al movimiento del individuo se le agregan dos cuestiones referidas a la construcción del antecedente de la regla. En primer lugar, debe indicarse cuáles serán los atributos que formarán parte del antecedente y esto debe hacerse utilizando información binaria. Por otro lado, es preciso representar el grado de pertenencia de los atributos originalmente numéricos en cada conjunto difuso para conocer cómo operar con las correspondientes variables lingüísticas, a través de una representación difusa.

Uno de los aportes de esta tesis radica en la definición de la función de aptitud o fitness de cada partícula basada en un "Criterio de Votación" que pondera de manera difusa la participación de las condiciones difusas en la conformación del antecedente. Su valor se obtiene a partir de los grados de pertenencia de los ejemplos que cumplen con la regla y se utiliza para reforzar el movimiento de la partícula en la

dirección donde se encuentra el valor más alto. Con la utilización de PSO las partículas compiten entre ellas para encontrar a la mejor regla de la clase seleccionada.

La eficiencia y eficacia de FRvarPSO se encuentran fuertemente condicionadas por la manera en que se determinen las funciones de pertenencia de los conjuntos difusos. En el marco de las investigaciones de esta tesis se han utilizado diferentes opciones. Uno de estas opciones fue particionar el rango de cada atributo numérico en intervalos de igual longitud, y centrando en cada uno de ellos una función triangular con un solapamiento adecuado. Otra de las formas para obtener los conjuntos difusos ha sido utilizando el método Fuzzy C-Means. Adicionalmente, se utilizó también como técnica el conocimiento de un experto para la definición de los conjuntos difusos, y su correspondiente valor de pertenencia.

El desempeño del método propuesto FRvarPSO fue comparado con versiones previas del mismo como son SOM + PSO, SOM + varPSO (PSO con población variable), LVQ + PSO, así como otros métodos de extracción de reglas de clasificación como PART y C4.5. Adicionalmente se comparó FRvarPSO con las versiones SOM + PSO Difuso, SOM + varPSO Difuso, LVQ + PSO Difuso, que son las versiones anteriores a las que se les aplicó lógica difusa al momento de construir el antecedente junto con el criterio de votación utilizado en la función fitness.

La medición se realizó sobre doce bases de datos del repositorio UCI (Machine Learning Repository).

3. RIESGO

Además del aporte en el área informática, se realizó una contribución significativa en el área de riesgo financiero ya que se probó el método propuesto FRvarPSO en tres casos reales en el área de crédito del Sistema Financiero del Ecuador asociadas al riesgo crediticio considerando un conjunto de variables micro y macroeconómicas. Dentro de las bases de datos del Sistema Financiero del Ecuador se contó con la información de una cooperativa de ahorro y crédito, y las otras dos bases de datos correspondieron a bancos que otorgan crédito de consumo productivo, no productivo y microcrédito, que se encuentran en el mercado por más de 25 años.

FRvarPSO se empleó en bases de datos reales en el proceso de análisis para la concesión de crédito. Es importante remarcar la ardua tarea realizada en las tareas de limpieza y transformación de la información, antes de la aplicación de los métodos. Se analizó la función de distribución de cada atributo a fin de detectar valores anómalos, faltantes y/o fuera de rango. Se transformaron variables con la finalidad de tener indicadores aplicados al cliente como, por ejemplo: la capacidad de pago, la capacidad de endeudamiento y la calificación crediticia. Esta última tuvo un análisis particular, debido a que se disponía de información de las características generales del cliente, así como del comportamiento de los créditos otorgados en un horizonte de tiempo, con lo cual se obtuvo el nivel de morosidad.

Otro de los aportes de esta tesis fue haber realizado 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. De esta forma, se consideraron tres casos: la cartera vencida, la que no devenga intereses y la cartera castigada.

Se verificó que con este análisis las reglas difusas obtenidas a través de FRvarPSO permiten que el oficial de crédito de respuesta al cliente en menor tiempo, y principalmente disminuya el riesgo que representa el otorgamiento de crédito para las instituciones financieras. Lo anterior fue posible, debido a que al aplicar una regla difusa se toma el menor grado de pertenencia promedio de las condiciones difusas que forman el antecedente de la regla, con lo que se tiene una métrica proporcional al riesgo de su aplicación. Con esta observación el oficial de crédito puede tomar la decisión de conceder el crédito incrementando la tasa de interés, las garantías y/o colaterales (activo), con la finalidad de disminuir el riesgo asociado.

Los resultados obtenidos fueron comparados mediante tests de diferencia de medias, verificándose que los modelos difusos presentan en la mayor parte de los casos una precisión superior a la del método PART,

pero ligeramente mayor a la alcanzada por el método C4.5. Sin embargo, si se observa la cantidad de reglas generadas para alcanzar dicha precisión, los métodos difusos utilizan una cantidad promedio de reglas mucho menor, que las reglas de C4.5 y PART. Esto último ratifica el énfasis puesto en la sencillez del modelo y en su facilidad de interpretación, gracias a la utilización de la lógica difusa.

4. RESULTADOS

En la figura 1 se ilustra un análisis comparativo de los resultados obtenidos al aplicar los métodos a las 12 bases de datos del repositorio UCI así como a las 3 bases de datos reales del Sistema Financiero Ecuatoriano. Aquí se puede observar la relación que existe entre la precisión de cada uno de los métodos con la mejor solución encontrada. Dicho valor corresponde al cociente entre la precisión promedio obtenida por el método y la mejor precisión obtenida, es decir, la mayor precisión de la base de datos analizada. Aquí se puede observar como el método propuesto tiene una mejor precisión, que los otros métodos analizados.

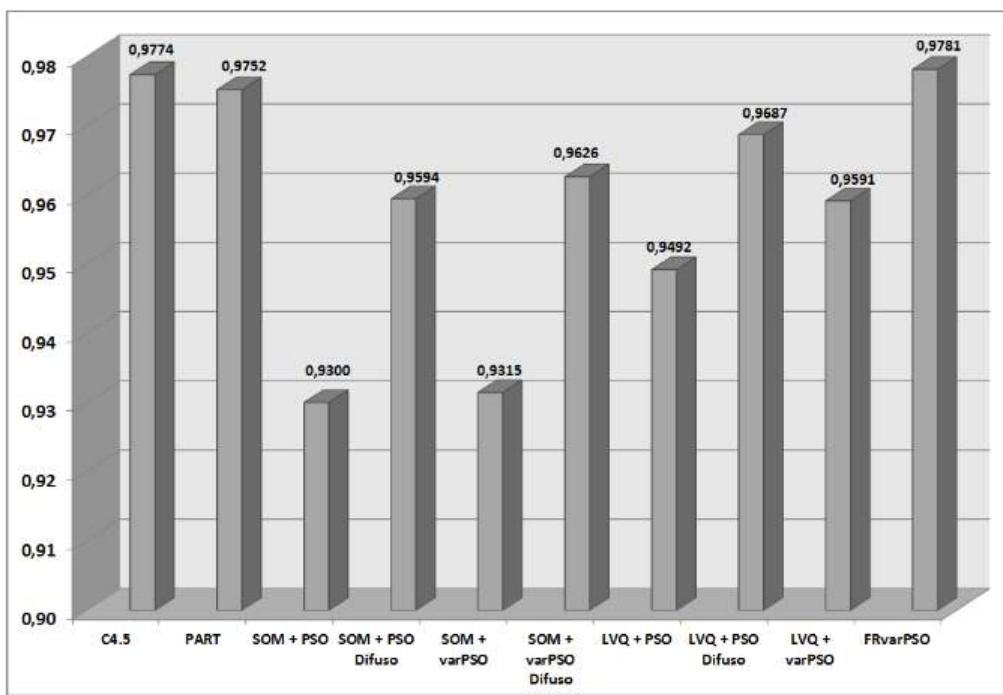


FIGURA 1: Comparación de la Precisión de los métodos

En la figura 2 se puede observar la relación que existe entre la cardinalidad (número de reglas) de cada uno de los métodos con la cardinalidad de la mejor solución encontrada. Dicha relación corresponde al cociente entre el promedio del valor del número de reglas difusas obtenidas por el método y la solución que tiene el menor número de reglas difusas de la base de datos analizada. Los resultados obtenidos permiten afirmar que el método propuesto es el que obtiene el menor número de reglas.

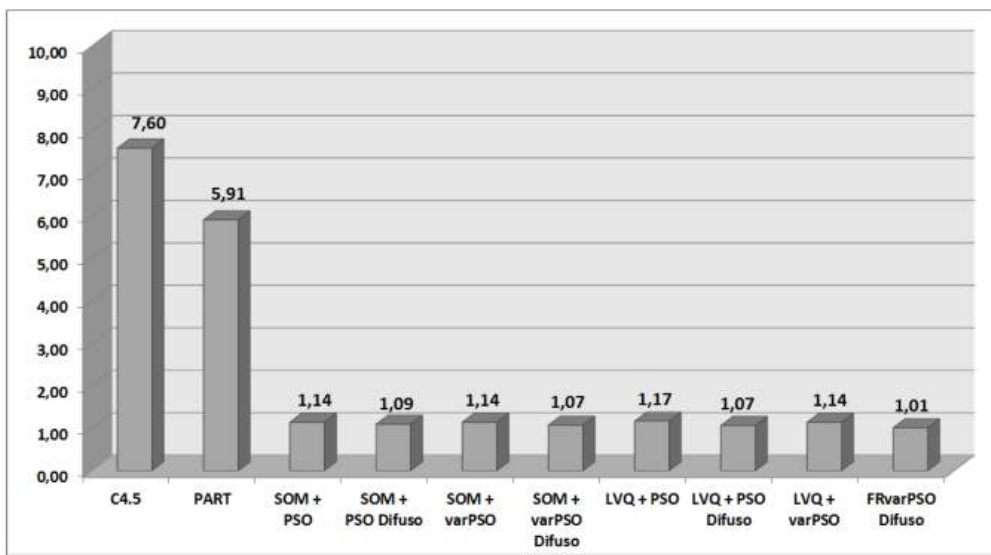


FIGURA 2: Comparación de la Cardinalidad de los métodos

En la figura 3 se puede observar la relación entre la precisión y la cardinalidad de los métodos analizados.

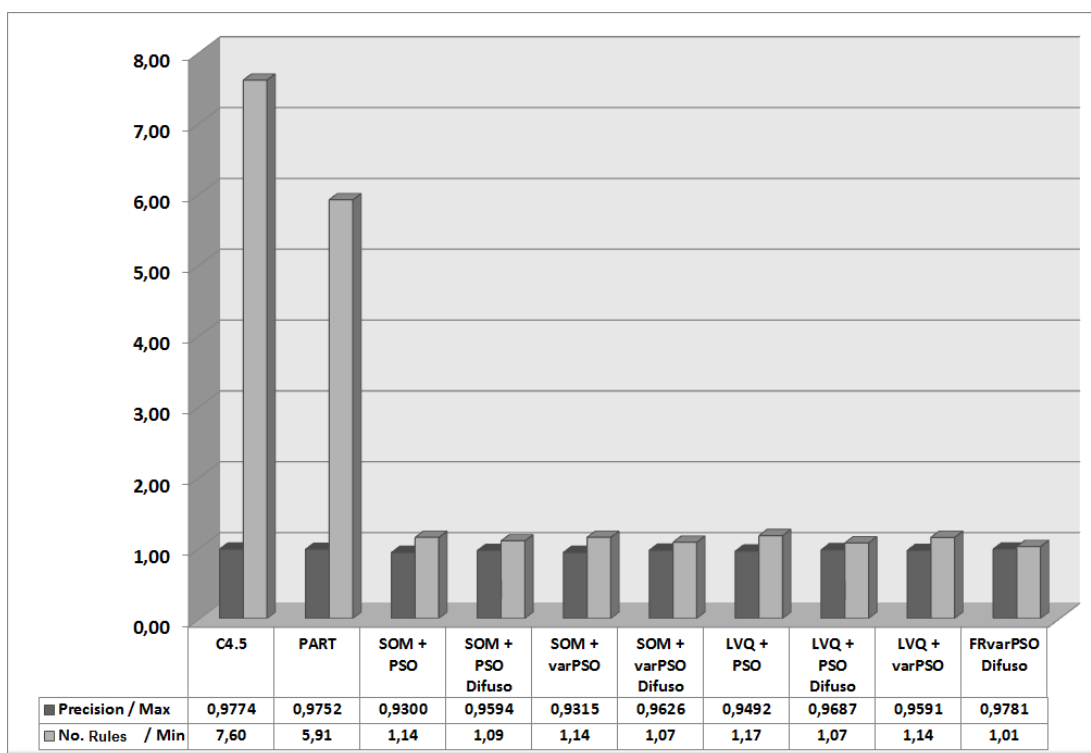


Figura 3: Comparación de la Precisión y Cardinalidad de los métodos

En las figuras 4, 5, 6 se puede observar claramente la relación que existe entre la precisión y el número de reglas de las tres bases de datos reales del Sistema Financiero Ecuatoriano utilizando FRvarPSO y generando los conjuntos difusos por medio de particiones equitativas, Fuzzy C-Means y con conocimiento del experto.

Finalmente, como resultado, se puede indicar que al utilizar FRvarPSO y no contar con el criterio del experto para la determinación de las funciones de pertenencia y el conjunto difuso definido por ella, el utilizar FCM, da como resultado un conjunto de reglas difusas con mayor precisión, menor cardinalidad y fácil interpretación, que cuando se utilizan conjuntos equitativos. A base de los resultados obtenidos se puede decir que el objetivo de la tesis se ha cumplido.

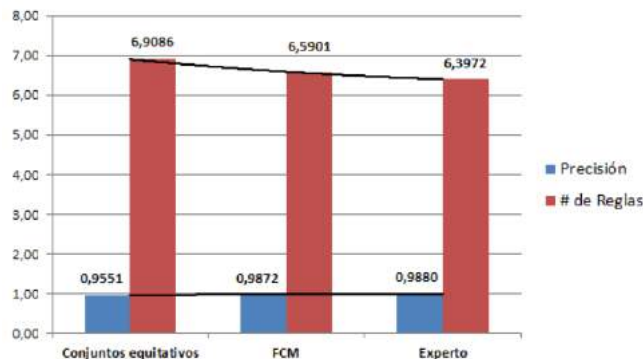


FIGURA 4: Banco encargado de microcrédito

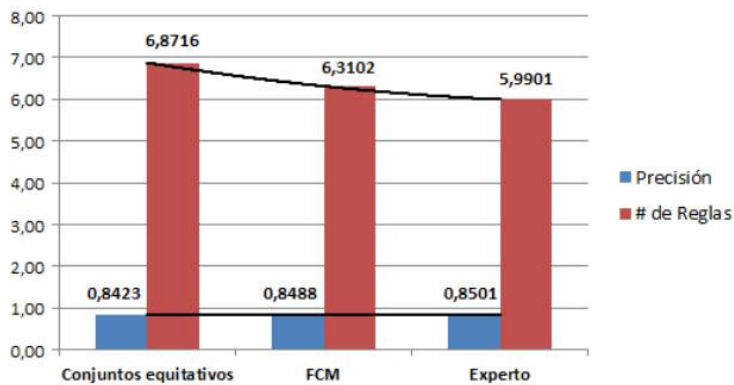


FIGURA 5: Banco encargado de crédito de consumo y crédito productivo empresarial

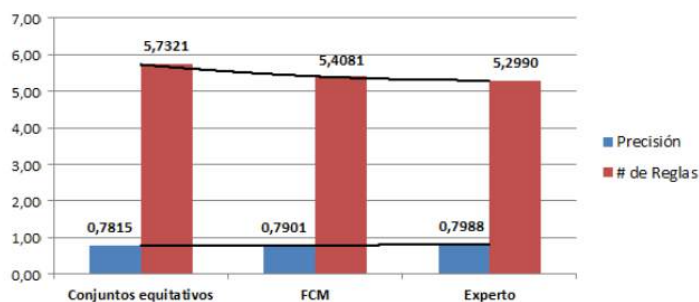


FIGURA 6: Cooperativa de Ahorro y Crédito del Ecuador que encuentra en el segmento 2 dentro de la Superintendencia de Economía Popular y Solidaria, activos mayor a 20'000.000,00 hasta 80'000.000,00

5. CONCLUSIONES Y TRABAJOS FUTUROS

Uno de los objetivos de esta tesis es la revisión de los diferentes métodos que permiten la obtención de reglas de clasificación, presentando un nuevo método que encuentra una mejor solución comparado con otras técnicas. El método propuesto denominado FRvarPSO (Fuzzy Rules Variable Particle Swarm Optimization), permite la representación del conocimiento en términos difusos a través de reglas de clasificación difusas extraídas de un conjunto de datos de entrada.

Las reglas de clasificación difusas obtenidas a través de FRvarPSO se caracterizan por tener cardinalidad baja y precisión adecuada, y especialmente por ser fáciles de interpretar gracias a la incorporación de la lógica difusa. Para conseguir esto se consideraron dos aspectos importantes; el primero tiene que ver con la capacidad del método de operar con variables lingüísticas y el segundo se relaciona con la inserción de información basada en grados de pertenencia tanto en la evaluación de la función fitness, así como en la manera en que se realiza la búsqueda por medio de la técnica de optimización.

El método inicia con la determinación de los conjuntos difusos, y las funciones de pertenencia de cada variable difusa asociada a cada uno de ellos, es aquí donde radica la eficiencia y eficacia de este método. Para esto se utilizaron dos variantes. La primera variante se encarga de particionar el rango de cada atributo numérico en intervalos de igual longitud centrando, en cada uno de ellos, una función triangular con un solapamiento adecuado. En este caso se consideraron particiones difusas uniformes con el mismo número de etiquetas para cada una de las variables y se definieron los conjuntos difusos de tal manera que el valor de una variable difusa sólo puede pertenecer a dos conjuntos de forma simultánea. En la segunda variante se mejora el proceso de obtención de los conjuntos difusos utilizando el algoritmo FCM con tres clusters, donde cada uno de ellos representaba un conjunto difuso.

Posteriormente se utiliza una técnica de optimización de población variable varPSO la misma que emplea una red neuronal competitiva LVQ, con la finalidad de obtener las zonas más prometedoras para realizar de búsqueda. Esta información suministrada por dicha red es utilizada para inicializar el cúmulo. El método utiliza el grado de pertenencia de las variables difusas, incidiendo directamente a través de la técnica de optimización varPSO, en el vector velocidad que controla el movimiento de las partículas. Ninguna de las técnicas por separado logra obtener los resultados que se indican en esta investigación. Los resultados obtenidos permiten afirmar que el modelo híbrido propuesto FRvarPSO tiene un mejor desempeño que sus versiones anteriores a raíz de la incorporación de la lógica difusa.

Una de las ventajas del método propuesto es que además de probarlo con bases de datos del repositorio UCI, se realizaron pruebas con base de datos reales pertenecientes a tres instituciones financieras. Aquí, para la determinación de los conjuntos difusos se utilizó el criterio de un experto en el área de crédito. Adicionalmente, se incorporó en el modelo no solo las variables del sujeto de crédito, si no también aquellas variables que hacen referencia a la economía del país en su conjunto, que son las variables macroeconómicas. Esto da una situación más real para la toma de decisiones, por parte del oficial de crédito en las instituciones financieras.

En este caso FRvarPSO no solo va a influir en la determinación del scoring de crédito, permitiendo analizar el riesgo que representa para la institución financiera la concesión del crédito, si no que en el caso de que el análisis indique que el crédito debe ser negado, este dependiendo del peso que tiene la regla de clasificación difusa puede ser concedido incrementado garantías o aumentando la tasa de interés. Todas estas características hacen que FRvarPSO sea una buena opción para análisis de riesgo crediticio. Es importante indicar que actualmente FRvarPSO está siendo utilizado en instituciones financieras del Ecuador.

FRvarPSO no solo obtiene un modelo más simple ya que utiliza menor cantidad de reglas que otros métodos, si no que presenta una buena precisión, y especialmente gracias a la incorporación de la lógica difusa mejora la interpretabilidad de la regla. Luego, este método aporta al área informática y se ha demostrado que también realiza aportes en el área de la economía, a través del análisis del riesgo crediticio, incorporando para ello variables macro económicas.

Es importante que en investigaciones futuras se consideren:

Una optimización de la función de pertenencia. Dentro de esta una alternativa puede ser la utilización de algoritmos genéticos, o de la misma optimización por cúmulo de partículas, con la finalidad de identificar automáticamente los parámetros de dicha función, sin perder de lado el objetivo que es tener un conjunto de reglas reducidas, pero conservando la interpretabilidad con las variables lingüísticas.

Un segundo aspecto que se debe considerar es la obtención de una nueva forma de representación de los atributos nominales, con el objetivo de reducir el tiempo computacional, y disminuir la longitud de la representación del antecedente dentro de cada partícula.

Otra variante que se debe tener en cuenta es incorporar técnicas para el manejo del desbalance de las clases, ya que en el área de riesgo crediticio la clase que corresponde a los créditos otorgados es la que tiene la mayor cantidad de ejemplos.

Es necesario incorporar al modelo la defuzificación de la variable de salida, siendo esta otra forma de interpretar el riesgo. Este valor crisp que se obtiene de la aplicación de las reglas de clasificación, va a indicar el porcentaje de riesgo que existe. En este caso para las reglas de clasificación aplicables a riesgo crediticio, se va a tener no solamente como salida la concesión o no del crédito, sino también el riesgo que implica el cliente, logrando disminuir considerablemente las consecuencias negativas para la institución financiera. Este mismo criterio puede ser utilizado en otra área de riesgo como el operacional, liquidez entre otros en el área financiera. Incluso el método propuesto puede utilizarse en otras áreas como la medicina, para saber el riesgo que un paciente tiene en contraer una determinada enfermedad.

6. PUBLICACIONES

Publicaciones en Revistas

- [1] **Simpliflyng Credit Scoring Rules using LVQ+PSO.** Lanzarini, Villa Monte, Fernández Bariviera, Jimbo Santana Patricia. *Kybernetes*. Vol.46 No 1. pp.8-16
 - Impact Factor: 0.98. Cuartil 3 Computer Science, Cybernetics
 - SJR 0.29. Cuartil 2 Computer Science (miscellaneous).
 - Número de citas: 11 (Scopus). 7 (Web of Science)

- [2] **Analysis of Methods for Generating Classification Rules Applicable to Credit Risk** Jimbo Santana, Villa Monte, Rucci, Lanzarini, Fernández Bariviera. *Journal of Computer Science & Technology*; vol. 17, No 1. pp.20-28. 2017
 - Indexado en Emerging Sources Citation Index
 - Número de citas: 1 (Web of Science)

- [3] **Data Mining Methods linked to Artificial Intelligence to Credit Risk.** Jimbo Santana, Lanzarini, Villa Monte, Fernández Bariviera. FIGEMPA. Universidad Central del Ecuador: Año V, volumen 2, No 8. 2017
 - Indexada en Latindex

- [4] **Fuzzy Credit Risk Scoring Rules using FRvarPSO,** Jimbo Patricia, Lanzarini Laura, Bariviera Aurelio, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol 26, No 1. pp.39-57. 2018
 - Impact Factor: 1.286. Cuartil 3 Computer Science, Artificial Intelligence
 - SJR 0.39. Cuartil 2 Artificial Intelligence.
 - Número de citas: 2 (Scopus). 1 (Web of Science)

- [5] **Variations of Particle Swarm Optimization for Obtaining Classification Rules Applied to Credit Risk in Financial Institutions of Ecuador.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. *Risks* Vol 8, No.2. 2020
 - Indexado en Emerging Sources Citation Index
 - Indexado en Scopus. CiteScoreTracker 2019:0.89.

- [6] **Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. *Journal of Advances in Information Technology* Vol. 11, No. 4, November 2020
 - Indexado en Scopus.

Capítulos de Libro

- [7] **Obtaining Classification Rules Using LVQ+PSO: an application to Credit Risk.** Lanzarini, Villa Monte, Fernández Bariviera, Jimbo Santana. *Advances in Intelligent Systems and Computing*. Springer-Verlag Berlin Heidelberg. Pp 383-391. ISBN 978-3-319-19704-3. DOI 10.1007/978-3-319-19704-3_31
 - Indexado en Scopus
 - Número de citas 7 (Scopus).

- [8] **Extraction of Knowledge with Population-Based Metaheuristics Fuzzy Rules Applied to Credit Risk.** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez-Bariviera. *Advances in Swarm Intelligence*. Springer. pp 153-163. 2018. ISBN 978-3-319-93817-2. DOI 10.1007/978-3-319-93818-9_15
 - Indexado en Scopus
 - Número de citas 4 (Scopus).

- [9] **FRvarPSO: A Method for Obtaining Fuzzy Classification Rules Using Optimization Techniques,** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernández- Bariviera, *Modelling and Simulation in*

Management Sciences, Springer International Publishing, Cham, pp 112–126, 2020 DOI: 10.1007/978-3-030-15413-4_9

- Indexado en Scopus

Congresos con referato internacional

- [10] **An exploratory analysis of methods for extracting credit risk rules.** Jimbo Santana, Villa Monte, Rucci, Lanzarini, Bariviera. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). pp. 834-841.
- [11] **Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets .** Patricia Jimbo Santana, Laura Lanzarini, Aurelio Fernandez Bariviera. The 12th International Conference on Computer Science and Information Technology (ICCSIT December 2019)