CONSTRUCCIÓN DE RELACIONES DE SIMILARIDAD BORROSA BASADA EN LA MEDIDA CALIDAD DE LA SIMILARIDAD

Lenniet Coello*, Mabel Frias*, Yumilka Fernandez*, Yaima Filiberto*, Rafael Bello**, Yailé Caballero*[1]
*Departamento de Computación, Universidad de Camagüey, Cuba
** Departamento de Computación, Universidad Central de Las Villas, Cuba

ABSTRACT:
In this research it proposed used fuzzy relations to define a new metric called Measure Quality of Similarity. New metric used a binary fuzzy relation, which quantify the strength of the similarity relationship in a range of [0.1]. The fuzzy relation is characterized by a membership function, which in this case is defined by a similarity function based on Extended Rough Set theory. This research it also proposed a method called PSO+RST+FUZZY which combines the Measure Quality of Similarity and metaheuristic Particule Swarm Optimization (PSO) to calculate features weights; and its impact for improve the performance of the k-Nearest Neighbors. Experimental results in international database show an effective behavior for accuracy.

KEYWORDS: binary fuzzy relations, Measure Quality of Similarity, features weights

MSC: 90C70

RESUMEN:
En esta investigación se propone el uso de relaciones borrosas para definir una nueva métrica denominada Medida de la Calidad de la Similaridad Borrosa. La nueva medida utiliza una relación borrosa binaria, que cuantifican la fortaleza de la relación de semejanza entre dos objetos en un rango de [0,1]. La relación borrosa está caracterizada por una función de pertenencia, que en este caso se define por una función de semejanza basada en la Teoría de los Conjuntos Aproximados Extendida. Se propone además un método llamado PSO+RST+FUZZY que combina la medida Calidad de la Similaridad Borrosa y la metaheurística Particle Swarm Optimization (PSO) para el cálculo de los pesos de los rasgos; y se estudia su impacto para mejorar el desempeño de los algoritmos k-Vecinos más Cercanos. Los resultados experimentales en bases de datos internacionales muestran un comportamiento efectivo en cuanto a precisión.

1. INTRODUCCIÓN:

Los Conjuntos Borrosos (CB o Difusos como también se les llama) y la Lógica Difusa (LD) son potentes herramientas matemáticas para modelar sistemas con incertidumbre en la industria, la naturaleza y fenómenos cotidianos. Son facilitadores para el razonamiento de sentido común en la toma de decisiones en ausencia de información completa y precisa. Sus funciones son significativas cuando se aplica a fenómenos complejos que no se describen fácilmente por métodos matemáticos tradicionales, especialmente cuando el objetivo es encontrar una buena solución aproximada. Un enfoque racional hacia la toma de decisiones debe tener en cuenta la subjetividad humana, en lugar de solo emplear medidas de probabilidad objetiva [1].
Los Conjuntos Borrosos y la Lógica Difusa se han investigado ampliamente derivando en exitosas aplicaciones, así los abordan diversos autores: [1-22]. La Teoría de los Conjuntos Borrosos, permite representar de forma más adecuada el conocimiento humano, ya que permite que los fenómenos y observaciones tengan más de dos estados lógicos, como lo son en la lógica clásica que sólo tiene verdadero o falso, o bien 0 ó 1. Una de las aplicaciones de la Lógica Borrosa es cuando se quiere construir Conjuntos Borrosos para ser usados en sistemas inteligentes [23].

[1]{lenniet.coello, mabel.frias, yumilka.fernandez, yaima.filiberto, yaile.caballero} @reduc.edu.cu, rbellop@uclv.edu.cu

En un conjunto clásico se asigna el valor 0 ó 1 a cada elemento para indicar la pertenencia o no a dicho conjunto. Esta función puede generalizarse de forma que los valores asignados a los elementos del conjunto caigan en un rango particular, y con ello indicar el grado de pertenencia de los elementos al conjunto en cuestión [23]. A dicha función se le denomina “función de pertenencia” y el conjunto por ella definida “conjunto borroso”. Esta función de pertenencia es definida por μA para un conjunto borroso A, determinado por el rango de los números entre [0,1], definido de la siguiente manera: $\mu A= X\rightarrow [0.1]$.
A diferencia de un conjunto clásico, los elementos pertenecen o no a él totalmente, por ejemplo, un número puede pertenecer o no al conjunto de números pares, pero no pertenecerá con un determinado grado al conjunto, como se hace en los Conjuntos Borrosos en donde hay grados de pertenencia en referencia a un universo local.
La minería de datos a través de grandes volúmenes de datos revela información útil en forma de nuevas relaciones, patrones o clusters, para la toma de decisiones por parte de un usuario. Los Conjuntos Borrosos apoyan una búsqueda centrada en los datos, especificada en términos lingüísticos. Esto ayuda a evitar la búsqueda de patrones sin sentido o triviales en una base de datos, y a descubrir dependencias o relaciones entre los datos en formato cualitativo o semicualitativo [9].

A. Similaridad entre Conjuntos Borrosos
La similaridad entre dos conjuntos depende del punto de vista de los diferentes autores, lo cual hace difícil decir cuál definición es la mejor entre todas las medidas de similitud para dos Conjuntos Borrosos. En [24] se proponen dos nuevas definiciones de medida de similitud entre dos Conjuntos Borrosos. Estas definiciones también se extienden a definir las medidas de similitud entre dos elementos en los Conjuntos Borrosos.
Dos medidas de similaridad entre dos Conjuntos Borrosos A y B son definidas en [24] por la expresión (1) y (2).

$$S(A,B)=\left\{\sum_{i=1}^{n}\left[\frac{m(\mu_A(X_i),\mu_B(X_i))}{M(\mu_A(X_i),\mu_B(X_i))}\right]\right\}/n \qquad (1)$$

Donde $m(a, b)$ y $M(a, b)$ denotan los valores mínimos y máximos de a y b, respectivamente. Para evitar que el denominador pueda ser cero, el autor asume 0/0 = 1 en el empleo de la expresión (2).

$$S(A,B)=\frac{\sum_{i=1}^{n}[1-|\mu_A(X_i)-\mu_B(X_i)|]}{n} \qquad (2)$$

Si la similitud $S(A,B)=1$, entonces se cumple que $\mu_A(X_i)=\mu_A(X_i)$, para toda $X_i \in X$, y se dice que $A=B$. Para cualquier dos Conjuntos Borrosos cualesquiera A y B, $A=B$ no significa que el conjunto A es exactamente el “mismo” que el conjunto B. Definir del término “mismo” o “igual” entre Conjuntos Borrosos es muy difícil, pues la información tiene cierto grado de incertidumbre [24]. La definición de medidas de similitud que satisfacen la propiedad $S(A,B)=1$, se consideran medidas más fiables.

B. Relaciones Borrosas
En [20] una relación borrosa binaria R se define como una colección borrosa de pares ordenados. Por lo tanto, si $X=\{x\}$ e $Y=\{y\}$ son colecciones de objetos denotados genéricamente por x e y, una relación borrosa de X a Y o, equivalentemente, una relación borrosa en $X\cup Y$, es un subconjunto borroso de $X\times Y$, caracterizado por una función de membresía (característica) μR. Esta función asocia a cada par (x, y) su “grado de pertenencia”, en R. Para simplificar, en este trabajo supondremos que el rango de μR es el intervalo [0, l] y se referirá al número $\mu R(x, y)$ como la fuerza de la relación entre x e y.

C. Teoría de los Conjuntos Aproximados Extendida
La Teoría de los Conjuntos Aproximados (RST) fue propuesta por Pawlak en 1982. La filosofía de los Conjuntos Aproximados se basa en la suposición que cada objeto x del universo de discurso U tiene asociada cierta información (datos y conocimiento), expresados por medio de atributos que describen el objeto [25]. Entre las ventajas de la RST para el análisis de datos están: se basa en los datos originales y no necesita información externa, por lo que no es necesario realizar ninguna suposición acerca de los datos; y permite el análisis tanto de rasgos cualitativos como cuantitativos [26].
Los Conjuntos Aproximados se basan en una relación R que define como inseparables a pares de objetos que tienen igual valor para un subconjunto de rasgos B, es decir, los objetos (x,y) son inseparables según B si $\alpha_i(x)=\alpha_i(y)$, para todo i en B, donde $\alpha_i(x)$ denota el valor del rasgo i en el objeto x. Definida de esta forma, R es una relación de equivalencia. En ocasiones cuando el dominio de los rasgos en B no es discreto, una relación de inseparabilidad definida de esta forma no es aplicable. En estos casos es necesario emplear otros tipos de relaciones de

inseparabilidad entre los objetos del universo U. Para ello se define una relación de semejanza en lugar de una relación de equivalencia, remplazando la relación de equivalencia por una relación binaria más débil. Una relación definida como, $xRy \leftrightarrow F(x,y) \geq E, x,y \in U$ donde $F(x,y)$ es una función de semejanza, y E es un umbral. De esta forma se obtiene una extensión del enfoque clásico de la RST [26]. Las relaciones de semejanza no inducen una partición del universo U, sino generan clases de semejanza para cualquier objeto $x \in U$

D. Conjuntos Aproximados y Conjuntos Borrosos

La Teoría de los Conjuntos Aproximados y la Teoría de los Conjuntos Borrosos son complementarias. Es natural combinar los dos modelos de incertidumbre (la vaguedad de los Conjuntos Borrosos y la aproximación de los Conjuntos Aproximados) para conseguir una precisión más exacta trabajando con información imperfecta.

La Teoría de los Conjuntos Aproximados y la Teoría de los Conjuntos Borrosos crean muchas posibilidades para la hibridación. El número de resultados reportados en esta dirección están aumentando continuamente; algunos ejemplos los podemos encontrar en [8],[9],[12],[19],[21],[22],[25],[27],[28],[29],[30],[31],[32].

Resulta interesante establecer una comparación entre las definiciones de conjuntos clásicos, Conjuntos Borrosos y los Conjuntos Aproximados. Los conjuntos clásicos es una noción primitiva y se define de manera intuitiva. Los Conjuntos Borrosos se definen mediante el empleo de la función de pertenencia difusa, lo que implica estructuras matemáticas avanzadas, número y funciones. Los Conjuntos Aproximados se definen por aproximaciones que requieren también conceptos matemáticos avanzados.

Los métodos de los Conjuntos Aproximados proporcionan descripciones aproximadas de conceptos y pueden ser empleados para construir una descripción aproximada de los conceptos difusos también. Esto es muy importante para la representación más comprimida de conceptos, reglas y patrones en los procesos de Descubrimiento de Conocimiento en Base de Datos (KDD por su siglas en Ingles Knowledge Discovery in Databases) porque el uso de conceptos difusos puede describir estos elementos de una manera más compacta. Estas descripciones, por otra parte, pueden ser más adecuadas para la comunicación con el ser humano [25].

Para tratar con los datos numéricos y discretos, o una mezcla de ambos, en conjuntos de datos, el "modelo aproximado borroso" se propuso por primera vez por Dubois y Prade [33] combinando los Conjuntos Aproximados y Borrosos. Las aproximaciones inferiores y superiores en lo "aproximado borroso" intentan asignar una función de pertenencia a cada objeto del conjunto de datos.

Como se mencionó anteriormente la Teoría de los Conjuntos Aproximados y la Teoría de los Conjuntos Borrosos brindan posibilidades para la hibridación. En este trabajo se presenta la idea de utilizar una función de semejanza basada en la Teoría de los Conjuntos Aproximados en su enfoque extendido como función de pertenencia para caracterizar una relación binaria difusa.

2. RELACIONES BORROSAS EN LA MEDIDA DE CALIDAD DE LA SIMILARIDAD

En esta investigación se plantea la nueva medida denominada Medida de la Calidad de la Similaridad Borrosa, una prolongación de la Medida de Calidad de la Similaridad propuesta en [26], que se respalda en el uso de las relaciones binarias borrosas para obtener una medida de mayor robustez y tolerancia a la imprecisión. Las relaciones binarias borrosas son caracterizadas por una función de pertenencia que indica en un intervalo del [0,1] la fortaleza de la relación entre un par de objetos. En la medida de Calidad de la Similaridad Borrosa, la función de pertenencia se define por la función de semejanza basada en la Teoría de los Conjuntos Aproximados Extendida entre dos objetos.

Sea un sistema de decisión DS= (U, A ∪ {d}), donde *A* denota el conjunto de rasgos de condición y *d* el rasgo de decisión, y el dominio de los rasgos en *A, {d}* pueden ser valores discretos o continuos. Sean las relaciones de similaridad borrosa R_1 y R_2 para todo par de objetos (*x, y*) en *U x U*, definidas por las expresiones (3) y (4). En la Medida de Calidad de la Similaridad Borrosa siempre va a existir una relación entre los objetos *x* e *y*, esta relación asume valores en el intervalo [0; 1] que representa la fuerza de la relación entre *x* e *y*.

Para todo par de objetos *(x, y)* en *U x U*, se cumplen las expresiones (3) y (4).

$$xR_1y = F_1(x,y) \tag{3}$$
$$xR_2y = F_2(x,y) \tag{4}$$

Donde R_1 y R_2 son relaciones borrosas binarias para cuantificar el nivel de semejanza entre los objetos *x* e *y* respecto a los rasgos de condición y el rasgo de decisión respectivamente. La parte interesante de Medida de Calidad de la Similaridad Borrosa radica en que el concepto de relaciones binarias borrosas R_1 y R_2 -basado en la Teoría de los Conjuntos Borrosos- quedan definidas por las funciones de semejanza F_1 y F_2 basada en la Teoría de los Conjuntos

Aproximados. De esta forma se está aprovechando las posibilidades para la hibridación de ambas teorías. F_1 considera los rasgos en *A* y F_2 calcula la similaridad entre dos objetos de acuerdo al valor del rasgo *d*. F_1 y F_2 están definidas por las expresiones 5 y 6.

$$F_1(x,y) = \sum_{i=1}^{n} w_i * \partial_i (x_i, y_i) \qquad (5)$$

$$F_2(x,y) = \partial (d(x), d(y)) \qquad (6)$$

Donde

$$\partial_i(x_i, y_i) = \begin{cases} 1 - \dfrac{|x_i - y_i|}{Max(\alpha_i) - Min(\alpha_i)} & si\ i\ es\ continuo \\ 1 & si\ i\ es\ discreto\ y\ x_i = y_i \\ 0 & si\ i\ es\ discreto\ y\ x_i \neq y_i \end{cases} \qquad (7)$$

En estas expresiones *n* es la cantidad de rasgos; W_i es el peso del rasgo *i*; x_i, y_i son los valores del rasgo i en los objetos *x*, *y* respectivamente. En (7) aparece reflejada ∂_i como función de comparación para el rasgo *i*, sea *i* un rasgo predictor o un rasgo de decisión.
A partir de las relaciones borrosas R_1 y R_2 se definen los Conjuntos Borrosos $N_1(x)$ y $N_2(x)$ para cualquier objeto *x* en el universo U según las expresiones 8 y 9 respectivamente, donde $N_1(x)$ y $N_2(x)$ son los conjuntos borrosos de objetos similares a *x* de acuerdo a las relaciones R_1 y R_2 respectivamente

$$N_1(x) = \{(y, \mu R_1(x,y))\ para\ todo\ y \in U\} \qquad (8)$$

$$N_2(x) = \{(y, \mu R_2(x,y))\ para\ todo\ y \in U\} \qquad (9)$$

El problema de encontrar una relación de semejanza adecuada para la RST extendida consiste en encontrar las relaciones R_1 y R_2 que logren mayor similaridad posible entre los conjuntos borrosos $N_1(x)$ y $N_2(x)$. El grado de similaridad entre ambos conjuntos para un objeto *x* se calcula como la semejanza entre los Conjuntos Borrosos $N_1(x)$ y $N_2(x)$ mediante la expresión 10.

$$\varphi(x) = \frac{\sum_{i=1}^{n}[1 - |\mu R_1(x_i) - \mu R_2(x_i)|]}{n} \qquad (10)$$

Usando la expresión 10 la medida calidad de la similaridad borrosa de un DS con un universo de *n* objetos se define por la expresión 11.

$$\theta(DS) = \left\{\frac{\sum_{i=1}^{n} \varphi(x)}{n}\right\} \qquad (11)$$

La medida $\theta(DS)$ representa el grado en el cual la similaridad entre los objetos usando los rasgos en *A* es equivalente a la similaridad que se obtiene de acuerdo al rasgo de decisión *d*. Esta medida representa el grado de similaridad entre los objetos de un sistema de decisión heterogéneo y constituye una herramienta importante para el análisis inteligente de los datos.

3. MÉTODO PARA CÁLCULO DE PESOS PARA LOS RASGOS BASADO EN LA MEDIDA DE MEDIDA DE CALIDAD DE LA SIMILARIDAD BORROSA

Hibridando la metaheurística Particle Swarm Optimization (PSO) como método de búsqueda para encontrar el conjunto de pesos que maximicen la Medida de Calidad de la Similaridad Borrosa se desarrolló un nuevo método denominado PSO+RST+FUZZY, el cual es una extensión con carácter borroso del método PSO+RST propuesto por [26] con el cual se obtuvieron excelentes resultados.
El método PSO+RST+FUZZY permite construir relaciones de similaridad borrosas a partir de funciones de semejanza definidas como una suma pesada de la comparación a nivel de rasgo (expresión 5) constituyendo un nuevo procedimiento de aprendizaje de pesos cuya aplicación no está limitada a problemas de clasificación. En este caso, el objetivo del algoritmo de aprendizaje es encontrar el conjunto de pesos W= $\{w_1, w_2, \ldots, w_n\}$, que maximiza

la medida calidad de la Medida de Calidad de la Similaridad Borrosa, es decir la metaheurística PSO buscará el conjunto de pesos que maximiza la expresión (11).

La optimización basada en partículas (Particle Swarm Optimization PSO) es un algoritmo de inteligencia de enjambre basada en la metaheurística propuesta por Kennedy y Eberhart [34]. PSO tiene tres componentes principales, los componentes sociales y cognoscitivos de las partículas, y la velocidad de las partículas. Cada partícula tiene una medida de calidad, así como una posición y velocidad en el espacio de la búsqueda, donde la posición determina el contenido de la posible solución. El aprendizaje de las partículas proviene de dos fuentes, una fuente es la propia experiencia de la partícula, llamada aprendizaje cognitivo y la otra fuente es el aprendizaje combinado de todo el enjambre, llamado aprendizaje social. El aprendizaje cognitivo está representado por la mejor experiencia del individuo (*pBest*) y el aprendizaje social está representado por la mejor experiencia del enjambre, el valor global (*gBest*) [35]. Juntos el aprendizaje cognitivo y social se utilizan para calcular la velocidad de las partículas a su próxima posición. En procesos de selección de rasgos, PSO junto a los Conjuntos Aproximados ya se ha estudiado por otros autores reportando muy buenos resultados [26]. Además ha demostrado tener un buen desempeño en los problemas de optimización continua.

En el PSO cada individuo está constituido por *n* componentes en dependencia de la cantidad de rasgos del problema a tratarse. El PSO se emplea como procedimiento de generación de subconjuntos; y la Medida de Calidad de la Similaridad Borrosa es empleada como función de evaluación de los subconjuntos. En este caso las partículas, que representan el vector W, tienen n componentes (una por cada rasgo en A). La calidad de las partículas se calcula usando la expresión (11). Al final del proceso de búsqueda desarrollado por el método PSO, la mejor partícula constituye el vector de pesos W a usar en la función de semejanza F_1 (expresión 5), y con esta queda construida la relación de semejanza R_1 (expresión 3).

A continuación se describe el funcionamiento del algoritmo PSO+RST+FUZZY:

Paso 1: Inicializar una población de partículas con posiciones y velocidades aleatorias en un espacio *D*-dimensional.

Paso 2: Para cada partícula, evaluar la medida calidad de la similaridad empleando la expresión 12, en *D* Variables.

$$max \rightarrow \left\{ \frac{\sum_{\forall x \in U} \varphi(x)}{|U|} \right\} \qquad (12)$$

Paso 3: Comparar la medida calidad de la similaridad actual de cada partícula con la medida calidad de la similaridad de su mejor posición anterior *pBest*. Si el valor actual es mejor que el de *pBest*, entonces asignarle a *pBest* el valor actual, y *Pi =Xi*, es decir, la localización actual pasa a ser la mejor hasta el momento.

Paso 4: Identificar la partícula en la vecindad con mayor valor para la medida calidad de la similaridad y asignar su índice a la variable *g* y asignar el mejor valor de la medida calidad de la similaridad a *m*.

Paso 5: Ajustar la velocidad y posición de la partícula según las ecuaciones 13 y 14 (para cada dimensión)

$$v_i(t+1) = \alpha * v_i(t) + U(0,\varphi 1)\big(pbest(t) - x_i(t)\big) + U(0,\varphi 2)\big(gbest(t) - x_i(t)\big) \qquad (13)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \qquad (14)$$

Paso 6: Verificar si se cumple el criterio de parada (máximo número de iteraciones o si lleva cinco iteraciones sin mejorar la medida calidad de la similaridad *global (m)*), si no, ir al Paso2.

A los efectos de demostrar la efectividad de los pesos que se calculan por PSO+RST+FUZZY para construir la relación de similaridad se muestran el caso de aplicación del método de aprendizaje automático: k-Vecinos más Cercanos, los resultados aparecen reflejados en el próximo apartado.

4. RESULTADOS:

A. k-Vecinos más Cercanos

Una forma práctica y de fácil aplicación para predecir o clasificar un nuevo dato, basado en observaciones conocidas o pasadas, es la técnica del Vecino más Cercano [36]. Este método utiliza funciones de distancia o similitud para generar predicciones a partir de ejemplos almacenados. Este algoritmo pertenece a la clase de algoritmo perezoso (*Lazy Learning Algoritms*), pues almacena el conjunto de entrenamiento y deja todo el procesamiento para la fase de clasificación.

En este estudio se denota al método como k-Vecinos más Cercanos o k-NN, aunque se está empleando el criterio de los k-Vecinos más Similares. En la propuesta de esta investigación basada en la regla de los k-Vecinos más Similares, se utiliza como función de semejanza la ecuación 15, que vendría siendo para calcular la similaridad entre los objetos, donde los valores de *Wi* son los pesos calculados por el método PSO+RST+FUZZY, con el objetivo de mejorar la recuperación de objetos similares.

$$\delta(x,y) = \sum_{i=1}^{n} w_i * \partial_i (X_i, Y_i) \quad (15)$$

A partir de las salidas de los vecinos más similares se calculan las salidas de un objeto usando las ecuaciones 16 y 17 según el problema sea de aproximación de funciones o de clasificación respectivamente.

$$x(d) = \frac{\sum_{y \in N(x)} y(d) * sim(x,y)^2}{\sum_{y \in N(x)} sim(x,y)^2} \quad (16)$$

$$\gamma(y,c,N(x)) = \frac{\sum_{y \in N(x)} \lambda(y(d),c) * sim(x,y)^2}{\sum_{y \in N(x)} sim(x,y)^2} \quad (17)$$

En la ecuación 16 se calcula el valor de decisión de *x (d)* a partir de los valores de decisión de los objetos en *N*(*x*), siendo *y(d)* el valor de decisión del objeto *y*. En la ecuación 17 para el caso de los problemas de clasificación, para cada clase calcular la medida *γ(y,c,N(x))*-1 si *y(d)* igual a *c*, 0 en otro caso- y se asigna como valor de decisión a *x(d)*, la clase que alcanza el valor de γ máximo. En este trabajo se propone emplear conceptos tomados de la Teoría de los Conjuntos Borrosos a través del método PSO+RST+FUZZY con el objetivo de lograr una mayor eficacia del k-NN para el problema de la aproximación de funciones.

B. Validación experimental en conjuntos de datos internacionales

Para este estudio se emplearon conjuntos de datos reconocidos internacionalmente, descritos en el Anexo 1, provenientes del UCI Repository de Aprendizaje Automático [37]. En la presente investigación se utiliza el test de Friedman y el Iman-Davenport para obtener los rankings de los algoritmos e identificar si existen diferencias significativas entre los mismos. Se emplea el test de Holm para probar secuencialmente las hipótesis ordenadas según su significancia e identificar las diferencias existentes entre el algoritmo de control y el resto. Este test tiene la ventaja de no realizar ninguna suposición adicional sobre las hipótesis chequeadas y en [38] se considera que el Test de Holm debe aparecer siempre en toda validación experimental.

Se obtuvieron los conjuntos de entrenamiento y control, tomando para el primero el 75% de los casos y para el segundo el 25%, de forma totalmente aleatoria. Siguiendo este principio de selección aleatoria se repitió el proceso diez veces y se obtuvieron para cada conjunto de datos diez conjuntos de entrenamiento y diez conjuntos de control, con el fin de aplicar validación cruzada (cross validation) para una mejor validación de los resultados [39].

El desempeño del método k-NN usando los pesos calculados por el método propuesto en esta investigación, llamado PSO+RST+FUZZY, se comparó con el resultado alcanzado por el método precedente PSO+RST y con los pesos calculados por otros métodos altamente referenciados como: el método del Gradiente Conjugado (KNN_{VSM}) [40], el método Relief [41] y Estándar (1/Cantidad-Rasgos). Estos métodos son los mismos empleados en la validación del PSO+RST en [39].

Para evaluar la precisión de los resultados alcanzados para el caso de aproximación de funciones se utilizó la medida PMD (promedio de las diferencias entre el valor deseado y el producido por el método) [42], definida por la expresión 18, donde: $\boldsymbol{x}_i(d)$ es el valor del rasgo de decisión para el objeto $\boldsymbol{x}_i$; q_i es el valor producido por el método y *m* es el número de objetos.

$$PMD = \frac{\sum_{i=1}^{m} |X_i(d) - q_i|}{m} \quad (18)$$

Para analizar el efecto del empleo de los pesos de los rasgos en el método k-Vecinos más Similares se realizó un experimento, el cual se detallan a continuación.

Experimento: Comparar los resultados del estimador de funciones k-NN (k=3), cuando se utilizan los diferentes métodos de cálculo de pesos: Estándar (1/Cantidad-Rasgos), KNNVSM, PSO+RST y el método de cálculo de pesos propuesto PSO+RST+FUZZY respecto al PMD.

Objetivos: Determinar si el método propuesto PSO+RST+FUZZY es significativamente inferior a los métodos Estándar, KNN_{VSM} y PSO+RST respecto al PMD.

Tabla 1. Resultados del error PMD para el algoritmo k-NN para problemas de aproximación de funciones.

No.	Conjuntos de datos	PMD Mean absolute error			
		Estándar	KNN_{VSM}	PSO+RST	PSO+RST+FUZZY
1	baskball	0.08	0.08	0.07	0.07
2	detroit	36.71	36.71	32.52	30.8
3	elusage	9.02	8.94	8.32	7.782
4	pollution	37.51	42.71	36.38	35.08
5	pwLinear	1.96	2.2	1.7	2.15
6	pyrim	0.06	0.06	0.05	0.05
7	sleep	2.6	2.6	2.41	2.56
8	vineyard	1.95	1.95	1.8	1.66
9	bodyfat	2.56	2.56	1.06	1.26
10	fishcatch	57.02	57.02	40.02	38.52

Tabla 2. Resultados del test de Friedman.

Algoritmos	Ranking
PSO+RST+FUZZY	1.5
PSO+RST	1.6
Estándar	3.35
KNN_{VSM}	3.55

Iman-Davenport (distribución de F con 3 y 27 grados de libertad): 23.967033
Valor de p computado por el test Iman-Davenport: 0.000000089925

Tabla 3. Prueba de Holm para α=0.05 para la medida de error PMD

i	Algoritmos	$z=(R_0-R_i)/SE$	p	Holm	Hipótesis
3	KNN_{VSM}	3.550704	0.000384	0.016667	Se rechaza
2	Estándar	3.204294	0.001354	0.025	Se rechaza
1	PSO+RST	0.173205	0.86249	0.05	No se rechaza

El test de Holm rechaza todas las hipótesis con valor de $p \leq 0.05$.

Las pruebas de Friedman e Iman-Davenport arrojaron que existen diferencias significativas entre los resultados del experimento, donde el mejor ranking es obtenido por el método PSO+RST+FUZZY. La prueba de Holm, tomando como método de control al PSO+RST+FUZZY, demostró que los resultados de las medidas cuando se calculan los pesos por el método PSO+RST+FUZZY son significativamente inferiores a los que se obtienen cuando se calculan los pesos mediante las otras medidas, con excepción del PSO+RST cuya hipótesis no se rechaza. Con respecto a PSO+RST+FUZZY vs PSO+RST la hipótesis nula no se rechaza, esto es equivalente a decir que no hay diferencias significativas entre sus comportamientos y por lo tanto se puede concluir que son igualmente eficaces.

5. CONCLUSIONES

Utilizando los conceptos de la Teoría de los Conjuntos Borrosos y específicamente las relaciones borrosas binarias se planteó una nueva métrica denominada Medida de Calidad de la Similaridad Borrosa. Esta modificación se reutilizó como medida de calidad para evaluar las partículas y redefinir el método para pesado de rasgos PSO+RST+FUZZY. Se demostró la eficacia de los pesos calculados por el PSO+RST+FUZZY para mejorar el desempeño del método de aprendizaje k-NN para aproximación de funciones, obteniendo resultados experimentales satisfactorios en los experimentos realizados.

RECEIVED: JUNE, 2016
REVISED: NOVEMBER, 2016

REFERENCIAS

[1] HELVACIOGLU, S. and OZEN, E. (2014): Fuzzy based failure modes and effect analysis for yacht system design. **Ocean Engineering**, 79, 131–141.
[2] BODENHOFER, U. (2000): A similarity-based generalization of fuzzy orderings preserving the classical axioms. **Uncertainty and Fuzziness Knowledge-Based Systems**, 8, 593–610.
[3] DEGANG, C. and DELI, Z. (2014): Structure of feature spaces related to fuzzy similarity relations as kernels. **Fuzzy Sets and Systems**, 237, 90–95.
[4] DEGANG, C., DELI, Z., JIASHANG, J., CONGXIN, W. and TSANG, E. C. C. (2005): Some notes on equivalent fuzzy sets and fuzzy subgroups. **Fuzzy Sets and Systems,** 152, 403 – 409.
[5] HUSSAIN, M. (2010): Fuzzy Relation. Master of Science in Mathematical Modelling and Simulation. **Institute of Technology School of Engineering.**
[6] IGNJATOVIC, J., CIRICA, M., ŠEŠELJA, B. and TEPAVCEVIC, A. (2015): Fuzzy relational inequalities and equations, fuzzy quasi-orders,closures and openings of fuzzy sets. **Fuzzy Sets and Systems**, 260, 1–24.
[7] KAWAHARA, Y., FURUSAWA, H. and MORI, M. (1999): Categorical representation theorems of fuzzy relations. **Information Sciences**, 119, 235-251.
[8] LI, Z. and CUI, R. (2015): Similarity of fuzzy relations based on fuzzy topologies induced by fuzzy rough approximation operators. **Information Sciences**, 305, 219–233.
[9] MITRA, P., MITRA, S. and PAL, S. (1999): Modular Rough Fuzzy MLP: Evolutionary Design. RSFDGrC, Heidelberg **Springer**, 1711, 128-137.
[10] OVCHINNIKOV, S. (2002): Numerical representation of transitive fuzzy relations. **Fuzzy Sets and Systems**, 126, 225–232.
[11] PEDRYCZ, W. (1997): Fuzzy sets in pattern recognition: Accomplishments and challenges. **Fuzzy Sets and Systems,** 90, 171-176.
[12] QIAN, Y., WANG, Q., CHENG, H., LIANG, J. and DANG, C. (2015): Fuzzy-rough feature selection accelerator. **Fuzzy Sets and Systems,** 258, 61-78.
[13] RAMAKRISHNAN, R. and RAO, C. J. M. (1992): Representations of fuzzy relations with an application to group decision making. **Fuzzy Sets and Systems**, 49, 143-150.
[14] ROYCHOWDHURY, P. and SHUKLA, K. K. (2003): Incorporating fuzzy concepts along with dynamic tunneling for fast and robust training of multilayer perceptrons. **Neurocomputing,** 50, 319 – 340.
[15] VERDEGAY, J. L., YAGER, R. R. and BONISSONE, P. P. (2008): On heuristics as a fundamental constituent of soft computing. **Fuzzy Sets and Systems,** 159, 846-855.
[16] WANG, X., CAO, X., WU, C. and CHEN, J. (2013): Indicators of fuzzy relations. **Fuzzy Sets and Systems**, 216, 91 – 107.
[17] WANG, X. and XUE, Y. (2014): Traces and property indicators of fuzzy relations. **Fuzzy Sets and Systems**, 246 78–90.
[18] YANG, M. S. and SHIH, H. M. (2001): Cluster analysis based on fuzzy relations. **Fuzzy Sets and Systems**, 120, 197–212.
[19] YAO, Y. Y. (1998): A comparative study of fuzzy sets and rough sets. **Journal of Information Sciences**, 109, 227-242.
[20] ZADEH, L. A. (1971): Similarity Relations and Fuzzy Orderings. **Information Sciences**, 3, 177-200.
[21]ZENG, A., LI, T., LIU, D., ZHANG, J. and CHEN, H. (2015): A fuzzy rough set approach for incremental feature selection on hybrid information systems. **Fuzzy Sets and Systems**, 258 39–60.
[22] ZHAO, X. R. and HU, B. Q. (2015): Fuzzy and interval-valued fuzzy decision-theoretic rough set approaches based on fuzzy probability measure. **Information Sciences,** 298, 534–554.
[23] RAMÍREZ, N.V. AND M. LAGUNA. (2012): La lógica borrosa: conjuntos borrosos, razonamiento aproximado y control borroso. **Pistas Educativas**, 100.
[24] WANG (1997): New similarity measures on fuzzy sets and on elements. **Fuzzy Sets and Systems**, 85, 305-309.
[25] KOMOROWSKI, J. and PAWLAK, Z. (1999): Rough Fuzzy Hybridization: A new trend in decision-making. **Rough Sets: A tutorial,** 3-98 . Spinger Press, Berlin.

[26] FILIBERTO, Y., BELLO, R., CABALLERO, Y. and LARRUA, R. (2010): Using PSO and RST to Predict the Resistant Capacity of Connections in Composite Structures. **International Workshop on Nature Inspired Cooperative Strategies for Optimization, Egipto**.
[27] NGUYEN, H. S. (2011): Rough sets and fuzzy sets in natural computing. **Theoretical Computer Science**, 412, 5816–5819.
[28] RIZA, L. S., JANUSZ, A., BERGMEIR, C., CORNELIS, C., HERRERA, F., S´LEZAK, D. and BENÍTEZ, J. M. (2014): Implementing algorithms of rough set theory and fuzzy rough set theory in the R package ‘‘RoughSets’’. **Information Sciences**, 287, 68–89.
[29] SUN, B., MAB, W. and CHEN, D. (2014): Rough approximation of a fuzzy concept on a hybrid attribute information system and its uncertainty measure. **Information Sciences**, 284, 60–80
[30] WUA, W., MI, J. and ZHANG, W. (2003): Generalized fuzzy rough sets. **Information Sciences**, 151, 263–282.
[31] KOMOROWSKI, J., PAWLAK, Z., POLKOWSKI, L. and SKOWRON, A. (1996): Rough Sets: A Tutorial. **Fundamenta Informaticae**, 17, 105-115.
[32] PAWLAK, Z. (1984): Rough sets. **International Journal of Computer and Information Sciences**, 11, 341-356.
[33] DUBOIS, D. and PRADE, H. (1987): Two fold fuzzy sets and rough sets- Some issues in knowledge representation. **Fuzzy Sets and Systems**, 23, 3-18
[34] KENNEDY, J. and EBERHART, R. (1995): Particle swarm optimization. IEEE International Conference on Neural Networks. IEEE.
[35] ALAM, S., DOBBIE, G., KOH, Y. S., RIDDLE, P. and REHMAN, S. U. (2014): Research on particle swarm optimization based clustering: A systematic review of literature and techniques. **Swarm and Evolutionary Computation**, 17, 1–13.
[36] MORA, J., MORALES, G. and BARRERA, R. (2008): Evaluación del clasificador basado en los k vecinos más cercanos para la localización de la zona en falla en los sistemas de potencia. **Ingeniería e Investigación**, 28, 25-28.
[37] ASUNCION, A. and NEWMAN, D. (2007): UCI machine learning repository, A study of the behaviour of several methods for balancing machine learning training data. **SIGKDD Explorations**, 6, 2-29.
[38] GARCÍA, S. and HERRERA, F. (2008): An extension on “statistical comparisons of classifiers over multiple data sets” for all pairwise comparisons. **Journal of Machine Learning Research**, 9, 2677-2694.
[39] FILIBERTO, Y. (2012): Métodos de aprendizaje para dominios con datos mezclados basados en la teoría de los conjuntos aproximados extendida. **Universidad Central de Las Villas**.
[40]WETTSCHERECKD, D. (1995): **A description of the mutual information approach and the variable similarity metric**. Sankt Augustin, Germany
[41] KONONENKO, I.(1994): Estimating attributes: Analysis and extensions of RELIEF. **European Conference on Machine Learning, England.**
[42] WEBER, R. AND C. FOIX.(2007): Pronóstico del precio del cobre mediante redes neuronales. **Revista Ingenieria de Sistemas**, 21, 75-76.