ISSN: 2007 8102

Una mejora del algoritmo Fuzzy c-Means utilizando la función de distribución gaussiana

A. Mexicano Santoyo[1], P.N. Montes Dorantes[2], E.G. Pérez Estrada[1], J.C. Carmona Frausto[1], S. Cervantes Alvarez[3]*
[1]División de Estudios de Posgrado e Investigación, Tecnológico Nacional de México / Instituto Tecnológico de Cd. Victoria, Boulevard Emilio Portes Gil #1301 Pte. A.P. 175 C.P. 87010 Cd. Victoria, Tamaulipas.
[2]Tecnológico Nacional de México / Instituto Tecnológico de Saltillo, Blvd. Venustiano Carranza, Priv. Tecnológico 2400, 25280 Saltillo, Coah.
[3]Departamento de Ciencias Computacionales e Ingenierías, Universidad de Guadalajara, UdeG, Ameca, Jalisco.
[*]jesus.cf@cdvictoria.tecnm.mx

***Área de participación**: Sistemas Computacionales*

Resumen
Fuzzy c-Means (FCM) es un algoritmo de agrupamiento suave no supervisado, altamente utilizado en diversas áreas de estudio como el reconocimiento de patrones, segmentación, análisis de datos, etc. No obstante, una de sus limitantes es el número de cálculos que se realizan para obtener el grado de membresía de cada objeto hacia cada centroide, lo que implica un alto costo computacional. En este trabajo se propone reducir el número de cálculos utilizando la función de distribución gaussiana para calcular el grado de membresía de cada objeto hacia cada centroide y para el cálculo de los centroides en cada iteración se calculó la media, considerando los principios del algoritmo *K-means*. Para evaluar la propuesta se utilizaron dos instancias sintéticas y una real. Los resultados mostraron una reducción en el tiempo de ejecución de hasta el 78.39% y una reducción de la calidad del agrupamiento del 7.78% en comparación con el algoritmo estándar.

***Palabras clave**: Fuzzy c-Means, función de membresía, distribución gaussiana*

Abstract
Fuzzy c-Means (FCM) is an unsupervised smooth clustering algorithm, widely used in various study areas such as pattern recognition, segmentation, data analysis, etc. However, one of its limitations is the number of calculations that are performed to obtain the degree of membership of each object towards each centroid, which implies a high computational cost. In this work it is proposed to reduce the number of calculations using the Gaussian distribution function to calculate the degree of membership of each object towards each centroid, also for the calculation of the centroids in each iteration the average was calculated, considering the principles of the K-means algorithm. To evaluate the proposal, two synthetic instances and one real were used. The results showed a reduction in execution time of up to 78.39% and a reduction in clustering quality of 7.78% compared against the standard algorithm.
***Key words**: Fuzzy c-Means, membership function, Gaussian distribution*

Introducción
El agrupamiento de datos es una técnica no supervisada, la cual es muy utilizada en diversas áreas tales como reconocimiento de patrones, procesamiento de imágenes, negocios, minería de datos, aprendizaje automático, entre otros [1]. Debido al crecimiento exponencial de datos en los últimos años el agrupamiento ha cobrado mucha relevancia, convirtiéndose en un objeto de estudio relevante para el aprendizaje automático. El objetivo del agrupamiento consiste en dividir un conjunto de objetos en grupos, donde los objetos que pertenecen a un grupo son muy similares entre sí, y también son diferentes en cuanto a los objetos que pertenecen a otro grupo. Con el tiempo, se han desarrollado diferentes algoritmos de agrupamiento y se han categorizado como agrupamiento basado en centroide [2], agrupamiento jerárquico [3], agrupamiento basado en distribución [4] y agrupamiento basado en densidad [5]. FCM forma parte del agrupamiento basado en centroide, este algoritmo es muy popular y forma parte de los algoritmos de lógica difusa. Un inconveniente de FCM es su alto costo computacional, lo cual

es una razón para generar nuevas mejoras. En este trabajo se propone una mejora al algoritmo FCM que consiste en reducir el costo computacional del algoritmo, sin una pérdida significativa en la calidad del agrupamiento. Para ello se sustituyó la función del cálculo de membresía de cada objeto hacia cada centroide por la función de distribución gaussiana y adicionalmente, para el cálculo de los centroides en cada iteración, se utilizó el cálculo de la media.

Trabajos relacionados

En la literatura se pueden encontrar diversos trabajos que se han desarrollado para mejorar el algoritmo *Fuzzy c-Means* (FCM), algunos de ellos van enfocadas en cuanto a la calidad del agrupamiento, para ello en [1], [6], [7], [8] y [9], se emplean algoritmos de inteligencia de enjambres buscando solucionar el principal inconveniente que presenta FCM que es sensible a la inicialización de los centroides, lo cual puede resultar en una convergencia lenta o no óptima, algunos de los algoritmos utilizados son PSO (*Particle Swarm Optimization*), GWO (*Grey Wolf Optimizer*), WOA (W*hale Optimization Algorithm*) y WSA (*Weighted Superposition Attraction Algorithm*). Otro enfoque al cual se han desarrollado es en busca de la reducción de tiempo de ejecución del algoritmo como el que se muestra en [10] el cual considera el parámetro m como un hiperparámetro clave que afecta el rendimiento de la agrupación, se propone un algoritmo FCM mejorado que permita variar el parámetro de difusividad en la función objetivo mostrando que mejora el rendimiento de la agrupación en grandes dimensiones, además el tiempo que tarda en ejecutarse se reduce considerablemente. En la literatura también se pueden encontrar distintos trabajos que implementan sistemas de lógica difusa utilizando distintas funciones de membresía, como se muestra en [11], donde el uso de una función de membresía de distribución gaussiana favorece ante el problema de la incertidumbre por lo que brindan una mejor interpretabilidad de los datos que se estén manejando, en [12] es aplicada una función triangular para la clasificación de una fruta utilizando reglas difusas para llevar a cabo la clasificación, donde el método propuesto reduce el tiempo y el esfuerzo de clasificación.

Fuzzy c-Means

El algoritmo FCM fue presentado por Bezdeck en [13] basándose en la teoría de conjuntos difusos que fue propuesta por Zadeh [14] [15]. Donde se expresa que un elemento puede tener una membresía parcial a múltiples grupos, expresado mediante un valor difuso entre 0 y 1. FCM es un algoritmo de optimización no lineal alternativo [1], como medida de la calidad del conglomerado busca minimizar el valor de la función objetivo [16] que se muestra en la Ecuación 1, la cual considera la distancia euclidiana entre los elementos del conjunto de datos y los centroides, tomando en cuenta los grados de membresía.

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij} \parallel x_i - c_j \parallel^2 \tag{1}$$

Donde N es el número de elementos, C es el número de grupos requeridos, μ_{ij} representa el grado de membresía para el iésimo elemento x_i en el grupo j , $\parallel x_i - c_j \parallel^2$ denota la distancia euclidiana entre el elemento x_i perteneciente al grupo c_j.

Para minimizar la función objetivo el algoritmo realiza iteraciones, en donde se van actualizando la matriz de membresía y así mismo el valor de los centroides.

La función de membresía de FCM [17] se describe mediante la Ecuación 2:

$$\mu_{ij} = \left[\sum_{t=1}^{c} \left(\frac{\|x_j - v_i\|_A}{\|x_j - v_t\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \tag{2}$$

μ_{ij} es el valor de membresía, el número de grupos está representado por c , x_j representa el j-ésimo objeto, v_i es el i-ésimo centroide sobre el que se calculará el grado de pertenencia. $\|.\|_A$ es la función de la norma y $\sum_{t=1}^{c} \mu_{tj} = 1$, lo que significa que la suma de los grados de membresía de cualquier grupo es uno.

La actualización de los centroides para cada iteración se lleva a cabo por medio de la Ecuación 3:

 ISSN: 2007 8102

$$c_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \quad (3)$$

Cabe destacar que el parámetro m, influye en la suavidad de la asignación del grado de membresía y en la nitidez de los grupos. Con un valor mayor de m se obtendrán resultados más suaves, mientras que con un valor bajo se obtendrán resultados más nítidos, este valor tiene que ser mayor a 1, aunque diversos trabajos en la literatura emplean un valor de $m = 2$.

Como condición de paro se pueden utilizar distintos criterios, uno muy conocido es el de un umbral de diferencia en los grados de membresía entre una iteración y otra, esto se denota mediante la Ecuación 4.

$$||\mu_{\text{ij}}^{k+1} - \mu_{ij}^{k}|| < \varepsilon \quad (4)$$

El Algoritmo 1, muestra las fases del algoritmo *Fuzzy c-Means*

Algoritmo 1: Fuzzy c-Means

Entrada: Datos, número de centroides c, constante difusa m, *valor del umbral* ε.
Salida: Centroides c_j, matriz de membresía $U = \mu_{ij}$, resultado del agrupamiento $\boldsymbol{G} = \{G_1, G_1, G_1\}$.

Paso 1. Se inicializan los centroides de manera aleatoria, se establecen los valores de c, m y ε.
Paso 2. Cálculo del grado de membresía μ_{ij} de cada objeto i hacia cada centroide j
Paso 3. Actualización de los centroides usando la Ecuación 3.
Paso 4. Validación de la condición de paro, tomando en cuenta la diferencia entre el grado de membresía y el grado de membresía anterior $||\mu_{\text{ij}}^{k+1} - \mu_{ij}^{k}|| < \varepsilon$, en caso de que la diferencia sea mayor que ε, el algoritmo se detiene, en caso contrario, $k = k + 1$ y se repite a partir del Paso 2.
Paso 5. Se asignan los objetos al grupo con mayor grado de membresía y se obtiene $\boldsymbol{G} = \{G_1, G_1, G_1\}$

Observaciones realizadas al algoritmo Fuzzy c-Means

Por medio de observaciones realizadas al comportamiento del algoritmo FCM, se observó que para obtener el grado de membresía de un objeto a un grupo se requiere calcular la distancia euclidiana del objeto hacia los diferentes centroides de los grupos existentes (ver Ecuación 2), lo cual resulta muy costoso, de tal forma se optó por sustituir dicha función por una función de distribución gaussiana (Ecuación 6) dado a que esta solamente requiere el cálculo de la media de los datos, la desviación estándar y la distancia del objeto al centroide que se desea obtener el grado de membresía, lo cual reduce la cantidad de cómputo requerido. Por otra parte, se observó que la etapa de la actualización de los centroides de FCM requiere elevar el grado de membresía de cada objeto a cada grupo a la potencia del coeficiente de difusividad (${\mu_{ij}}_{ij}^{m}$) y posteriormente obtener el producto por cada una de las características del objeto x_i lo cual también representa alto número de cálculos (ver Ecuación 3), por lo tanto, se decidió cambiarlo por el cálculo de la media de los objetos asignados a un grupo, inspirándose en el algoritmo *k-Means*.

Propuesta de mejora del algoritmo Fuzzy c-Means

La etapa de clasificación es muy importante, debido a que es donde se asignan los objetos a un grupo de acuerdo con sus características. El algoritmo FCM emplea una función de membresía propia en donde se utiliza la distancia euclidiana para poder realizar dicho cálculo. En busca de reducir los cálculos requeridos para cumplir con la tarea de agrupamiento, se propone la utilización de una función de membresía de distribución gaussiana en busca de reducir los cálculos requeridos para el agrupamiento. Esto contribuirá dado a que para dicha función no será necesario el cálculo de la distancia euclidiana entre cada uno de los objetos al centroide de cada grupo.

 ISSN: 2007 8102

Para la medición de la calidad del agrupamiento se emplea una función objetivo, la cual evalúa la cohesión que existe en los grupos que el algoritmo generó (Ecuación 5).

$$J = \sum_{l=1}^{k} \sum_{x_j \in C_l} \left\| x_j - C_l \right\|^2 \quad (5)$$

Donde k representa el número de grupos, j el j-ésimo elemento y C el l-ésimo centroide.

Para el cálculo del grado de membresía del conjunto de datos se utilizó la Ecuación 6, tomando como universo del discurso las variables del conjunto de datos a evaluar.

$$\mu_{x_i} = \text{Gauss}(x; c, \sigma) = e^{-\frac{1}{2}\left(\frac{x_i - c}{\sigma}\right)} \quad (6)$$

donde x_i representa las variables de entrada, c representa la media del conjunto de datos y σ representa la desviación estándar del conjunto difuso.

Tomando en cuenta la membresía de cada elemento a los diferentes centroides, se asigna al que tenga mayor grado de membresía. Una vez que se han asignado todos los elementos a un grupo, se recalcula el centroide de cada grupo. Dicha actualización se lleva a cabo obteniendo la media de todos los elementos asignados a determinado grupo, lo cual contribuye a la reducción de cálculos en la fase de recalculo de centroides, en comparación con el algoritmo FCM que requiere el uso de la Ecuación 3.

El algoritmo se repite iterativamente hasta cumplir con la condición de paro, la cual puede ser un número máximo de iteraciones o un umbral en los grados de membresía entre una iteración y otra, como se muestra en la Ecuación 4. El Algoritmo 2 representa la propuesta de mejora.

Algoritmo 2 Propuesta de mejora del algoritmo Fuzzy c-Means

Entrada: Datos, número de centroides c, constante difusa m, *valor del umbral ε.*
Salida: Centroides c_j, resultados del agrupamiento $\boldsymbol{G} = \{G_1, G_1, G_1\}$.

Paso 1. Inicialización de los centroides de manera aleatoria, se establecen los valores de c, m y ε.
Paso 2. Cálculo del grado de membresía con una función de membresía de distribución gaussiana mediante la Ecuación 6, se asignan los elementos al centroide que obtenga un mayor número de grado de membresía se obtiene $\boldsymbol{G} = \{G_1, G_1, G_1\}$
Paso 3. Actualizar los centroides, inspirándose en el funcionamiento del algoritmo K-Means, donde se obtiene la media de todos los elementos asignados a determinado grupo.
Paso 4. Se valida la condición de paro, tomando en cuenta la diferencia entre el grado de membresía y el grado de membresía anterior $||\mu_{ij}^{k+1} - \mu_{ij}^{k}|| < \varepsilon$, en caso de que la diferencia sea mayor que ε el algoritmo se detiene, en caso contrario $k = k + 1$ y se repite a partir del Paso 2.

Ejemplo motivacional

Con la finalidad de mostrar el comportamiento de los algoritmos FCM y la mejora (denotada en adelante por "Mejora"), se hicieron pruebas con una pequeña instancia sintética de 25 objetos distribuidos de manera uniforme a la cual se le generaron 3 centroides pseudoaleatorios. Para este ejemplo se midió la calidad del agrupamiento mediante la Ecuación 5 para ambos algoritmos, también se consideró medir el tiempo que demoró en ejecutarse cada uno de los algoritmos, el algoritmo fue detenido al cumplir con un umbral entre los grados de membresía de ε =0.1.

Figura 1 se observa el comportamiento de los algoritmos FCM y el comportamiento de la Mejora, en ambos casos se observa la asignación de grupos por iteración, mostrando en cada iteración el valor de la calidad del

 ISSN: 2007 8102

agrupamiento denotado por J que corresponde al error medio cuadrado, donde se observa que la propuesta de mejora muestra un agrupamiento similar al obtenido por medio del algoritmo FCM, no obstante, la Mejora requiere una menor cantidad de cálculos para su funcionamiento.

El desempeño del algoritmo se puede observar en la *Tabla 1*, la columna $\tau\%$ representa el porcentaje de la reducción del tiempo de la Mejora con respecto a FCM donde se observa una diferencia del 8.57%. La columna $E\%$ representa la diferencia en la calidad del agrupamiento y se observó una disminución del 1.51% en comparación con la versión estándar. Los valores de $\tau\%$ y $E\%$ fueron obtenidos mediante las Ecuaciones 7 y 8, respectivamente.

$$\tau\% = \frac{(T_c - T_g)x100}{T_c} \tag{7}$$

Donde T_c denota el tiempo promedio en ejecutarse el algoritmo FCM y T_g el tiempo promedio en ejecutarse el algoritmo Fuzzy G-Mean.

$$E\% = \frac{(E_c - E_g)x100}{E_g} \tag{8}$$

Donde E_c denota la media del error al cuadrado del algoritmo FCM y E_g el error medio cuadrado de la Mejora.

Tabla 1 Desempeño de los algoritmos FCM y la Mejora, al agrupar una instancia de 25 objetos distribuidos uniformemente en tres grupos.

Tiempo (s)			**Error medio al cuadrado**		
FCM	**Mejora**	**τ %**	**FCM**	**Mejora**	***E* %**
0.0068	0.0063	8.5708	33.4735	33.9793	1.5112

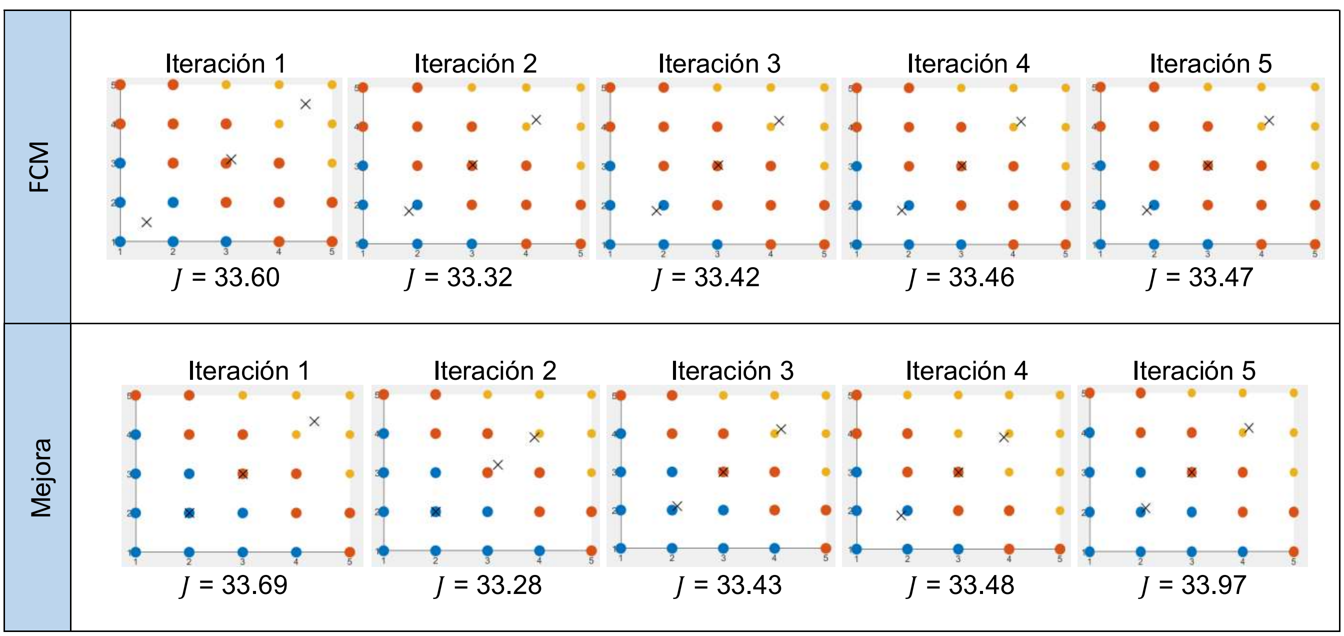

Figura 1 Ejemplo Motivacional, tomando de referencia el comportamiento del algoritmo FCM y la Mejora, donde se muestra un agrupamiento similar en cada iteración entre ambos algoritmos y el valor del error obtenido por cada iteración.

 ISSN: 2007 8102

Experimentación y Resultados

La implementación se realizó en MATLAB R2022b, usando una computadora con un procesador Intel CORE i9 de 3.42 GHz, con 64 GB de RAM y con un sistema operativo Windows 11 Pro. Se realizaron experimentos con los algoritmos FCM y la Mejora. Se utilizaron tres instancias, las primeras dos son instancias sintéticas con datos uniformemente distribuidos en dos dimensiones con 2500 y 40000 objetos y la tercera es un dataset real denominado *The concrete compressive strength,* con 1030 objetos en 8 dimensiones extraído de *UCI Machine Learning* [18]. Para las tres instancias se le generaron 100, 200 y 400 grupos. Para medir el desempeño de los algoritmos se calculó la sumatoria de las distancias al cuadrado (Ecuación 5) y también se midió el tiempo de ejecución de cada prueba. Para las pruebas realizadas se tomó en cuenta como criterio de paro un umbral entre los grados de membresía ε = 0.01.

Las columnas tres y cuatro de la Tabla 2 y Tabla 3 están divididas en tres subcolumnas donde se muestran los resultados del algoritmo FCM y la Mejora; la subcolumna $\tau\%$ corresponde a la reducción del tiempo (expresado en segundos) respecto a FCM (Ecuación 7). La subcolumna $E\%$ corresponde al porcentaje de la reducción en la calidad calculada por el error medio cuadrado (Ecuación 8). Los resultados presentados corresponden al promedio de 30 ejecuciones. Los resultados muestran que la Mejora reduce el tiempo de ejecución en el mayor de los casos hasta un 78.39% y además en el caso de la instancia de 40000 objetos al generarle 200 centroides la calidad tuvo una ligera reducción de 7.78%, para el caso de las instancias sintéticas. En el caso de la instancia real la mayor reducción de tiempo obtenida fue de 72.20 %, mientras que en calidad del agrupamiento se obtuvo una reducción del 4.33% respecto a FCM.

Tabla 2 Comparación de los resultados al evaluar el algoritmo FCM y la Mejora con instancias sintéticas de 2500 y 40000 objetos, donde se resaltan los valores más significativos.

Dataset	Número de Grupos	*Tiempo (s)*			Error medio al cuadrado		
		FCM	Mejora	$\tau\%$	FCM	Mejora	$E\%$
2500	100	44.37	46.68	-4.94	10951.01	10952.21	0.01
	200	433.96	406.96	6.63	5674.59	5785.70	1.96
	400	1349.19	756.32	**78.39**	3661.70	3646.39	-0.42
40000	100	841.24	844.32	-0.36	2808510.46	2848619.36	1.43
	200	4783.46	3865.25	23.76	1414684.95	1524785.64	**7.78**
	400	25941.60	15856.38	63.60	729378.91	739705.87	1.42

Tabla 3 Comparación de los resultados al evaluar el algoritmo FCM y la Mejora con el dataset Concrete con 1030 objetos, donde se resaltan los valores más significativos.

Instancia	**Número de Grupos**	***Tiempo (s)***			**Error medio al cuadrado**		
		FCM	**Mejora**	**τ %**	**FCM**	**Mejora**	$J\%$
Concret	100	10.30	9.98	3.25	83281.61	83363.22	0.10
	200	54.95	25.36	56.53	71978.32	73593.90	2.24
	400	143.70	52.98	**72.20**	57293.81	59775.39	**4.33**

 ISSN: 2007 8102

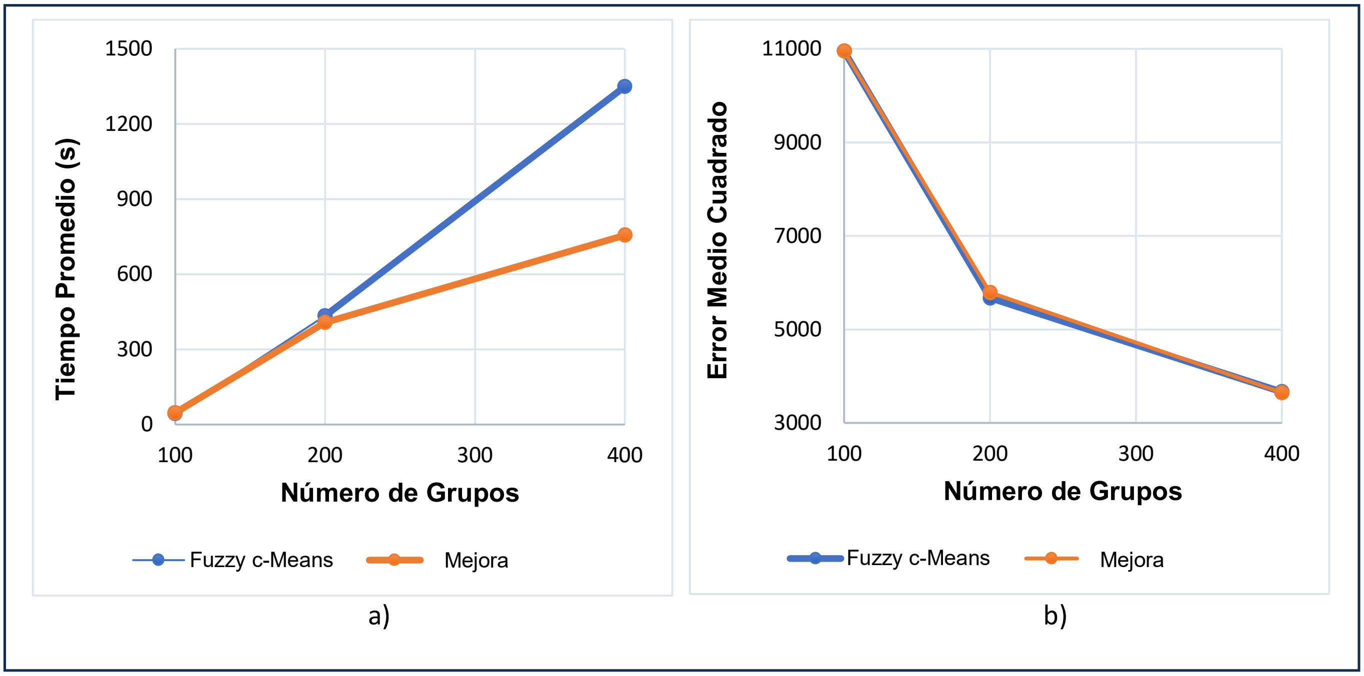

Figura 2 Comparación de los resultados obtenidos de los algoritmos FCM y la mejora para la instancia sintética de 2500 objetos, con respecto a: a) tiempo promedio que demoró en ejecutarse el algoritmo y b) la calidad del agrupamiento.

El tiempo y la medición del error (calidad del agrupamiento) se muestran gráficamente en las Figuras 2, 3 y 4 donde se aprecian los resultados obtenidos en las pruebas realizadas a las instancias de 2500, 40000 y 1030 objetos, respectivamente. En cuestión del tiempo se observa que la Mejora muestra una reducción considerable a diferencia de FCM. En cuanto a la calidad del agrupamiento, para las tres instancias se observa una diferencia pequeña en comparación con el algoritmo FCM.

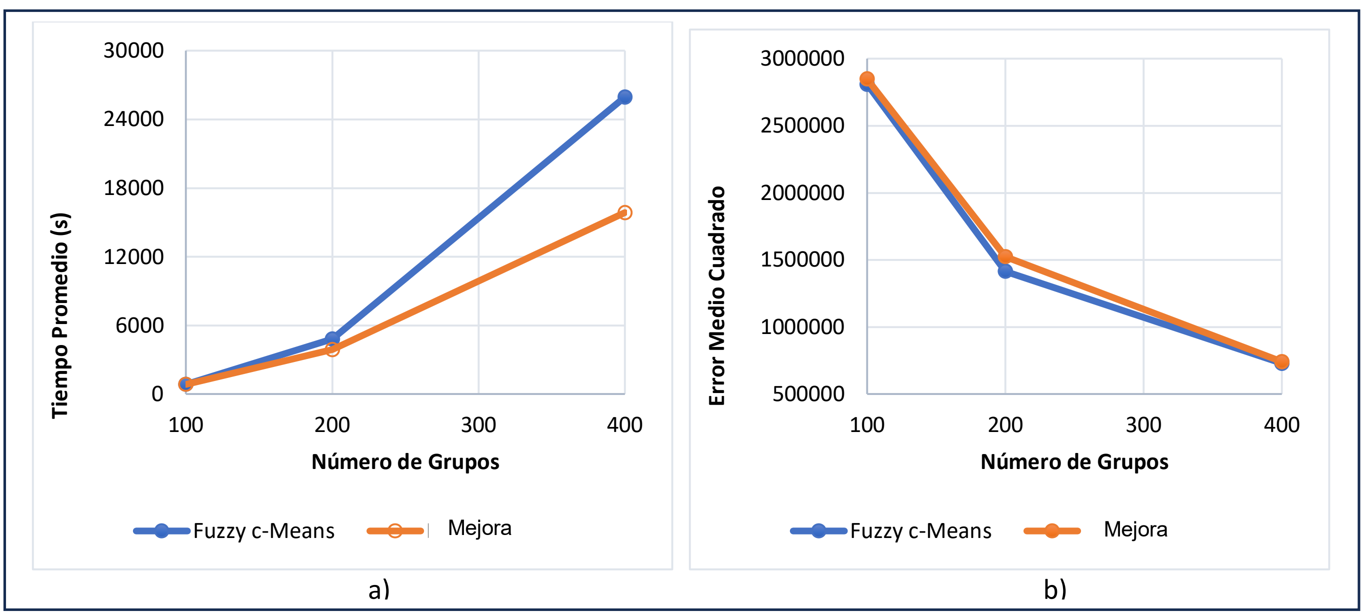

Figura 3 Comparación de los resultados obtenidos de los algoritmos FCM y la mejora para la instancia sintética de 40000 objetos, con respecto a: a) tiempo promedio que demoró en ejecutarse el algoritmo y b) la calidad del agrupamiento.

 ISSN: 2007 8102

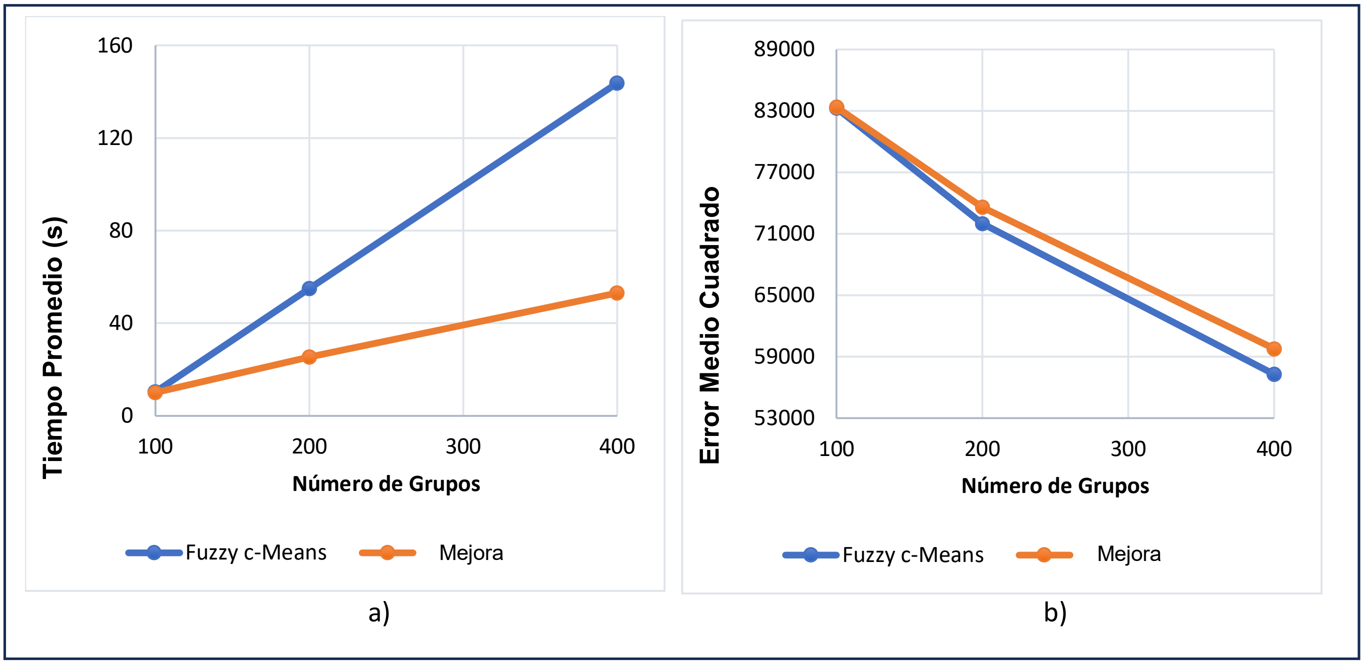

Figura 4 Comparación de los resultados obtenidos de los algoritmos Fuzzy c-Means y la mejora para la instancia real de 1030 objetos, con respecto a: a) tiempo promedio que demoró en ejecutarse el algoritmo y b) la calidad del agrupamiento.

Conclusiones

En este trabajo se muestra una propuesta de mejora para un algoritmo de agrupamiento difuso (FCM), donde lo que se busca es reducir el costo computacional requerido para su ejecución, empelando una función de membresía de distribución gaussiana, buscando con ello la reducción del costo computacional, sin afectar significativamente la calidad el agrupamiento, además la actualización de los centroides se realizó obteniendo la medio de los objetos de cada grupo lo cual también contribuyó a la reducción de tiempo del algoritmo. Al realizar distintas pruebas se pudo observar que el tiempo requerido para su ejecución llego a ser hasta un 78.39% menor comparado a FCM e incluso en la calidad del agrupamiento se obtuvo una ligera reducción en la calidad del agrupamiento de hasta un 7.78%. Como trabajo futuro se pretende la implementación de las funciones de membresía triangular, trapezoidal y de la campana, con el propósito de encontrar la función con la que se obtenga mejores resultados.

Referencias

[1] H. Verma, D. Verma y P. Kumar, “A population based hybrid FCM-PSO algorithm for clustering analysis and segmentation of brain image,” Expert Systems with Applications, vol. 167, 2021, https://doi.org/10.1016/j.eswa.2020.114121.

[2] A. Mexicano, R. Rodriguez, S. Cervantes, P. Montes, M. Jiménez, N. Almanza y A. Abrego, “The early stop heuristic: A new convergence criterion for K-means,” AIP Conference Proceedings 1738, 310003; 2016, Rodas, Grecia. pp. 22-28, 2015, doi: 10.1063/1.4952103.

[3] Nisha y J. Puneet, “Cluster quality based performance evaluation of hierarchical clustering method,” 2015 1st International Conference on Next Generation Computing Technologies (NGCT), pp. 649-653, 2015, 10.1109/NGCT.2015.7375201.

[4] A. Mann y N. Kaur, “Review Paper on Clustering Techniques,” Global Journal of Computer Science and Technology Software & Data Engineering, vol. 13, pp. 1-7, 2013.

[5] Y. Yang, C. Qian, H. Li, J. Wu, C.-j. Liu y S. Zhao, “An efficient DBSCAN optimized by arithmetic optimization algorithm with opposition-based learning,” The Journal of Supercomputing, vol. 78, pp. 19566–19604, 2022, https://doi.org/10.1007/s11227-022-04634-w.

[6] J. Zhang y L. Shen, “An improved fuzzy c-means clustering algorithm based on shadowed sets and PSO,” Computational Intelligence and Neuroscience, vol. 2014, 2014, https://doi.org/10.1155/2014/368628.

 ISSN: 2007 8102

[7] A. Baykasoglu, I. Golcuk y F. Burcin, "Improving fuzzy c-means clustering via quantum-enhanced weighted superposition attraction algorithm," Hacettepe Journal of Mathematics and Statistics, vol. 48, pp. 859-882, 2019.

[8] S. Emaddedin, F. Gholian y M. Hajiaghaei, "A fuzzy C-means algorithm for optimizing data clustering," Expert Systems with Applications, vol. 227, 2023, https://doi.org/10.1016/j.eswa.2023.120377

[9] M. Mohammdian khoshnoud, A. R. Soltanian, A. Dehghan y M. Farhadian, "Optimization of fuzzy c-means (FCM) clustering in cytology image segmentation using the gray wolf algorithm," BMC Molecular and Cell Biology, vol. 23, no. 9, 2022, https://doi.org/10.1186/s12860-022-00408-7

[10] Y. Chen, S. Zhou, X. Zhang, D. Li y C. Fu, "Improved fuzzy c-Means clustering by varying the fuzzyness parameter", Pattern Recognition Letters, vol. 157, pp. 60-66, 2022, https://doi.org/10.1016/j.patrec.2022.03.017 .

[11] X. Liu, T. Zhao y X. Xie, "Two-dimensional Gaussian hierarchical priority fuzzy modeling for interval-valued data," Information Sciences, vol. 630, pp. 23-29, 2023, https://doi.org/10.1016/j.ins.2023.02.026.

[12] K. A. Ahmad, S. L. Abdullah, M. Othman y M. N. Bakar, "Inducction of membership functions and Fuzzy rules for harumanis classification," Journal of Fundamental and Applied Sciences, vol. 10, no. 1S(2018), pp. 1202-1215, 2018, https://doi.org/10.4314/jfas.v10i1s.89

[13] J. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, New York: Springer, 1981.

[14] L. Zadeh, "Fuzzy Sets," Information and Control, vol. 8, pp. 338-353, 1965.

[15] P. K. TImawari y A. Srivastava, "Zadeh extension principle: A note," Annals of Fuzzy Mathematics and Informatics, vol. 9, no. 1, pp. 37-41, 2014.

[16] J. Nayak, B. Naik y H. Behera, "Fuzzy C-Means (FCM) Clustering Algorithm: A Decade Review from 2000 to 2014," Computational Intelligence in Data Mining, vol. 32, pp. 133-149, 2014.

[17] M. Huang, Z. Xia, H. Wang, Q. Zeng y Q. Wang, "The range of the value for the fuzzifier of the fuzzy c-means algorithm," Pattern Recognition Letters, vol. 33, pp. 2280-2284, 2012.

[18] C. Merz, P. Murphy y D. Aha, UCI Repository of Machine Learning Databases. Departament of Information anda computer Science, University of California, https://archive.ics.uci.edu/datasets, último acceso: julio 2023.