

## Fuzzy c-Means Clustering untuk Pengenalan Pola Studi kasus Data Saham

### *Fuzzy c-Means Clustering for Pattern Recognition on Case Studies in Holdings Data*

Sylvia Jane A. Sumarauw<sup>1)</sup> Marvel Grace Maukar<sup>2)</sup>  
[sylviasumarauw@unima.ac.id](mailto:sylviasumarauw@unima.ac.id)

Universitas Negeri Manado

#### Abstrak

*Fuzzy Clustering* merupakan salah satu peran dari lima peran yang digunakan para *expert* data mining untuk mentransformasikan sejumlah besar data kedalam informasi yang berguna, dan salah satu metode yang sering dan banyak digunakan adalah *Fuzzy c-Means (FCM) Clustering*. FCM merupakan teknik pengklasteran data dimana keberadaan masing-masing titik data dalam klaster berdasarkan derajat keanggotaan. Penelitian ini bertujuan untuk melihat pola sampel data atau kategori data menggunakan FCM *clustering*. Data yang dianalisis adalah data saham pada Bursa Efek Jakarta (BEJ) bidang Property dan Real Estate (kelompok Emiten). Proses-proses data mining mengikuti *Cross Industry Standard Process Model for Data mining Proses (Crisp-DM)*, dengan beberapa tahapan, dimulai dengan tahapan mengenal proses bisnis (*Business Understanding*) kemudian mempelajari data (*Data Understanding*), lanjut dengan tahapan *Data Preparation*, tahap Pemodelan, tahap Evaluasi dan terakhir tahap *Deployment*. Pada tahap Pemodelan digunakan model FCM. Data mining model FCM *clustering* dapat menganalisis data pada basis data yang besar dengan banyak variabel dan rumit, terutama untuk mendapatkan pola dari data. Kemudian dibangun *Fuzzy Inference System (FIS)* berdasarkan pola yang sudah diketahui untuk simulasi data masukkan (input) menjadi data keluaran (output) berdasarkan logika fuzzy

**Kata kunci:** *fuzzy c-means klastering, pengenalan pola*

#### Abstract

*Fuzzy Clustering is one of the five roles used by data mining experts to transform large amounts of data into useful information, and one method that is often and widely used is Fuzzy c-Means (FCM) Clustering. FCM is a data clustering technique where the existence of each data point in the cluster is based on the degree of membership. This study aims to see the pattern of data samples or data categories using FCM clustering. The analyzed data is stock data on Jakarta Stock Exchange (BEJ) in the Property and Real Estate sector (issuer group). The data mining processes comply Cross Industry Standard Process Model for Data mining Process (Crisp-DM), with several stages, starting with the stage of getting to know the business process (Business Understanding) then studying the data (Data Understanding), continuing with the Data Preparation stage, Modeling stage, Evaluation stage and finally the Deployment stage. In the modeling stage, the FCM model is used. FCM clustering model data mining can analyze data in large databases with many variables and complicated, especially to get patterns from the data. Then a Fuzzy Inference System (FIS) was built based on a known pattern for simulating input data into output data based on fuzzy logic.*

**Keywords:** *fuzzy c-means clustering, pattern recognition*

## PENDAHULUAN

Data mining adalah proses berulang yang kemajuannya ditentukan oleh penemuan, baik secara otomatis atau manual, dengan melakukan pencarian baru terhadap informasi berharga dari data, dalam volume yang besar dan rumit. (Kantardzic, 2020). Data mining juga diartikan sebagai suatu proses eksplorasi dan penemuan pola dan informasi dari *big data* (Zadissa & Apweiler, 2022). Data mining memiliki peranan yang mendasar dalam proses penemuan informasi yang berguna, yang diperoleh dari proses analisis suatu himpunan data yang besar dengan informasi yang berarti dari data tersebut (Wiemer dkk., 2019). Hannachi (2021) mengemukakan bahwa data mining, pada dasarnya adalah suatu proses yang berkaitan dengan penggunaan berbagai alat analisis data dan *software*, untuk mengidentifikasi, mengetahui dan menemukan sesuatu yang penting yang tersembunyi (terpendam) yang berhubungan dengan data yang beragam, yang bisa meningkatkan pemahaman terhadap suatu sistem. Hal ini berdasarkan pada kekuatan komputerasi dan dapat menjelaskan satu bagian dimensi dataset yang besar melalui reduksi dimensi (Hannachi, 2021).

Salah satu peran dari data mining adalah *clustering*. Clustering dapat didefinisikan sebagai klasifikasi model tanpa pengawasan ke dalam *cluster* (Singh, 2015). Klastering dilakukan dengan mempartisi titik-titik data ke dalam kelompok-kelompok natural yang disebut *cluster*, sehingga titik-titik dalam suatu kelompok sangat mirip, sedangkan titik-titik antar kelompok yang berbeda dibuat seberbeda mungkin (M. Abdalla dkk., 2011). Tujuan dari *clustering* adalah membuat objek suatu klaster mirip satu dengan lainnya dan objek di antara klaster berbeda satu dengan yang lain dan *clustering* direpresentasikan oleh *center* dari setiap kelas (Zhu dkk., 2019).

Analisis klastering merupakan satu dari beberapa teknik dalam analisis statistika yang dikembangkan untuk menyelesaikan keberagaman variabel (Pellegrini dkk., 2018). Analisis klaster menjadi sarana dasar untuk beberapa aplikasi yang bergantung pada suatu pola rekognisi (Bezdek, 2022). Teknik ini mampu menyelesaikan masalah analisis data yang besar dengan variabel yang banyak (analisis multivariat). Analisis klastering memberi peringkat objek (contoh produk, responden, atau entitas lainnya) pada serangkaian karakteristik yang terpilih (pengelompokan variabel) (Hair dkk., 2018). Berdasarkan pendapat Hair dkk (2018) klaster yang dihasilkan harus homogen di dalam klaster dan heterogen antar klaster. Jika klasifikasi berhasil, objek dalam klaster akan menjadi dekat satu sama lain ketika secara geometri diplot, sedangkan klaster yang berbeda akan jauh satu sama lain (Hair dkk., 2018).

*Fuzzy* klastering adalah satu dari beberapa teknik analisis *clustering* untuk menentukan klaster yang optimal pada sebuah ruang vektor, untuk

jarak antar vektor yang berdasar pada bentuk normal *Euclidean* (Rahakbauw dkk., 2017). Fuzzy klastering juga disebut sebagai *soft* klastering (Ni dkk., 2017). FCM merupakan salah satu algoritma *clustering* data yang pertama kali dikembangkan oleh J.C. Dunn untuk kasus khusus yang selanjutnya didefinisikan secara umum oleh Bezdek pada tahun 1973 (Ruspini dkk., 2019). FCM adalah satu teknik pengklasifikasian data dimana, setiap titik data milik dari satu klaster untuk beberapa derajat keanggotaan berdasarkan nilai keanggotaan (Bora & Gupta, 2014).

Algoritma FCM adalah tipikal dasar dari *soft* partisi (Yiping dkk., 2021). Algoritma FCM pada dasarnya melakukan pengelompokan *image* dengan mencari jarak minimum antara piksel dan pusat *cluster* (Zhi & Liu, 2020). Keuntungan yang menonjol dari algoritma FCM adalah bahwa proses segmentasi tidak diawasi dan algoritma dapat diterapkan ke semua *noise image* (Miao dkk., 2020). Beberapa contoh aplikasi dari penggunaan algoritma FCM adalah Segmentasi dari Gambar MR Otak (Dubey & Mushrif, 2016), clustering objek geological (Efendiyev dkk., 2016), data *stream clustering* (Gao dkk., 2017), *Locating Visible Video Watermarks* (Al. Embaby dkk., 2020), tomografi nodul kaca tanah (Zhang dkk., 2022) dan penempatan pengontrol dalam jaringan yang ditentukan perangkat lunak (Thalapala & Guravaiah, 2022). Tujuan penelitian ini adalah untuk melihat pola sampel data atau kategori data menggunakan FCM klastering.

## METODE

Penelitian ini menggunakan proses data mining yang telah dibuat oleh CRISP-DM. Metode CRISP-DM merupakan standar *de facto* untuk melakukan analisis data pada aplikasi industri (Wiemer dkk., 2019). Perspektif CRISP-DM dan metodologi terkait adalah bahwa data mining, adalah proses yang dimulai dari tujuan bisnis yang relatif jelas dan data yang telah dikumpulkan dan tersedia untuk pemrosesan komputasi lebih lanjut (Martinez-Plumed dkk., 2021). Tahapan-tahapan CRISP-DM dimulai dengan tahapan mengenal proses bisnis (*Business Understanding*) kemudian mempelajari data (*Data Understanding*), lanjut dengan tahapan *Data Preparation*, tahap Pemodelan, tahap Evaluasi dan terakhir tahap *Deployment*. Studi kasus data saham dengan variabel; *Date, Day, Code, Company, Volume, Amount* dan *Individual Index* dari *stock summary* data saham kelompok *real estate and property* pada Bursa Efek Jakarta. FCM menggunakan metode *Max-Min Mamdani* dan *Defuzzifikasi metode Centroid*. Metode Mamdani diperkenalkan oleh Ebrahim Mamdani untuk membangun sistem kendali dari mesin uap (Arif dkk., 2016). Mamdani FIS digunakan untuk mengembangkan model *fuzzy* yang komprehensif untuk gerakan awal dari suatu partikel butiran sedimen

tanpa kohesi yang mengalami aliran searah yang stabil (Bizimana & Altunkaynak, 2020).

Algoritma *clustering* dapat dikategorikan pada *clustering* yang berbasis *centroid* (Wu dkk., 2015). Metode *centroid* merupakan metode *defuzzifikasi* yang paling sering digunakan (Arun & Mohan, 2017). Metode ini juga biasa disebut dengan *center of area method* (COA) (K. Wang, 2001). Menurut (K. Wang, 2001) metode ini menentukan pusat area himpunan *fuzzy* dan mengembalikan nilai crisp yang sesuai. Himpunan *fuzzy* yang merupakan Input adalah *Individual Index*, *Volume* dan *Amount*. Dengan demikian ada 3 pengelompokan data input berdasarkan kriteria yang sudah ditentukan (sumber data *understanding*).

**Tabel 1. Kriteria Variabel Input**

Variabel	Kriteria		
	Rendah/Kecil	Menengah/Sedang	Tinggi/Besar
Individual Index	< 69	69-100	>100
Volume	< 112.000.000	112.000.000-138.000.000	>138.000.000.
Amount	>24.000.000.000	24.000.000.000-30.000.000.000	>30.000.000.000

Selain pengelompokan kriteria data input, output juga dikelompokkan menurut kriteria; beresiko tinggi, beresiko sedang dan beresiko rendah. Berdasarkan pengelompokan kriteria data input dan output diatas, dibentuk membership function (fungsi keanggotaan) untuk setiap variabel. Kemudian dibentuk aturan *fuzzy* untuk mendapatkan *fuzzy decision* yang pada akhirnya akan menghasilkan output keputusan apakah perusahaan termasuk beresiko tinggi, beresiko sedang atau beresiko rendah. Sesuai dengan kriteria data input dan kriteria data output maka ditentukan aturan fuzzy sebanyak 27 (tabel 2)

**Tabel 2. Aturan Fuzzy**

Individual Index	Volume	Amount	Output
Rendah	Kecil	Kecil	Beresiko Tinggi
Rendah	Kecil	Sedang	Beresiko Tinggi
Rendah	Kecil	Besar	Beresiko Tinggi
Rendah	Sedang	Kecil	Beresiko Tinggi
Rendah	Sedang	Sedang	Beresiko Tinggi
Rendah	Sedang	Besar	Beresiko Tinggi
Rendah	Besar	Kecil	Beresiko Tinggi
Rendah	Besar	Sedang	Beresiko sedang
Rendah	Besar	Besar	Beresiko sedang
Menengah	Kecil	Kecil	Beresiko Tinggi
Menengah	Kecil	Sedang	Beresiko sedang
Menengah	Kecil	Besar	Beresiko sedang
Menengah	Sedang	Kecil	Beresiko sedang
Menengah	Sedang	Sedang	Beresiko sedang
Menengah	Sedang	Besar	Beresiko sedang
Menengah	Besar	Kecil	Beresiko sedang
Menengah	Besar	Sedang	Beresiko sedang
Menengah	Besar	Besar	Beresiko Rendah
Tinggi	Kecil	Kecil	Beresiko sedang
Tinggi	Kecil	Sedang	Beresiko Rendah

Tinggi	Kecil	Besar	Beresiko Rendah
Tinggi	Sedang	Kecil	Beresiko Rendah
Tinggi	Sedang	Sedang	Beresiko Rendah
Tinggi	Sedang	Besar	Beresiko Rendah
Tinggi	Besar	Kecil	Beresiko Rendah
Tinggi	Besar	Sedang	Beresiko Rendah
Tinggi	Besar	Besar	Beresiko Rendah

### Algoritma Fuzzy c-Means Clustering

Konsep dasar FCM dimulai dari menentukan pusat kluster yang menandai lokasi rata-rata setiap kluster. Pusat kluster pada kondisi awal belum akurat, kemudian dengan memperbaiki pusat kluster demikian juga dengan memperbaiki derajat keanggotaan secara berulang dari setiap titik data, maka pusat kluster akan bergerak menuju tempat yang tepat, perulangan berdasarkan pada minimisasi fungsi objektif yang memanifestasikan selisih antara titik data yang diberikan ke dengan pusat kluster yang terbobot oleh derajat keanggotaan dari titik data tersebut.

Tahapan algoritma FCM klustering yakni: penentuan banyaknya cluster, penentuan matriks partisi awal, penentuan nilai W, epsilon dan MaxIter awal, penentuan fungsi objektif awal dan cek konvergensi (Kusumadewi, 2002).

Apabila terdapat suatu himpunan data seperti berikut;

$U = (u_1, u_2, u_3, \dots, u_N)$ , derajat keanggotaan suatu titik data ke- $k$  di *cluster* ke- $i$  adalah:

$$\mu_{ik}(u_k) \in [0,1], \text{ dengan } (1 \leq i \leq c; 1 \leq k \leq N)$$

Metode FCM matriks partisi didefinisikan sebagai berikut:

$$\mu_f(c) = \begin{bmatrix} \mu_{11}[u_1] & \mu_{21}[u_1] & \cdots & \mu_{c1}[u_1] \\ \mu_{12}[u_2] & \mu_{22}[u_2] & \cdots & \mu_{c2}[u_2] \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{1N}[u_N] & \mu_{2N}[u_N] & \cdots & \mu_{cN}[u_N] \end{bmatrix}$$

dengan  $\sum_{i=1}^c \mu_{ik} = 1, \dots \dots \dots (1)$

yang berarti banyaknya nilai keanggotaan satu data pada semua kluster harus sama dengan 1

Fungsi objektif iterasi ke- $t$   $P(c)$  pada matriks partisi adalah:

$$P_t(c) = \sum_{k=1}^N \sum_{i=1}^c (\mu_{ik})^w |u_k - v_{fi}|^2 \dots \dots \dots (2) \text{ dengan}$$

$v_{fi}$  adalah pusat vektor pada kluster *fuzzy* ke- $i$ ,

$$v_{fi} = \frac{\sum_{k=1}^N (\mu_{ik})^w u_k}{\sum_{k=1}^N (\mu_{ik})^w}, \dots \dots \dots (3)$$



dan  $w$  adalah bobot-bobot pada nilai keanggotaan,

$|u_k - v_{fi}|$  adalah bentuk normal *Euclidian*, digunakan sebagai jarak diantara  $u_k$  dan  $v_{fi}$  (Algoritma mengikuti, Kusumadewi, 2002).

## HASIL DAN PEMBAHASAN

Data Saham yang diteliti adalah entitas *Stock Summary* dengan 7 variabel yang bisa dijadikan tolok ukur untuk penilaian perusahaan. Variabel tersebut adalah; *Date, Day, Code, Company, Individual Index, Volume* dan *Amount*. *Date* adalah tanggal transaksi, *Day* hari transaksi dalam hal ini hanya ada hari Senin, Selasa, Rabu, Kamis dan Jumat. *Code* adalah kode perusahaan, *Company* adalah nama perusahaan, *Individual Index* adalah pergerakan harga saham individual, *Volume* adalah banyaknya transaksi jual beli saham/hari/perusahaan dan *Amount* adalah total nilai transaksi dalam rupiah perhari per perusahaan. Dari sekian perusahaan yang tercatat pada Bursa Efek Jakarta dipilih 35 perusahaan kelompok *Property* dan *Real Estate*. Dari 35 perusahaan hanya 34 perusahaan yang dianalisis karena 1 perusahaan tidak ada nilai transaksi *volume* dan *amount*, perusahaan tersebut langsung dikategorikan menjadi beresiko tinggi. Berikut contoh hasil deskripsi dan eksplorasi data *Individual Index* dari 35 perusahaan.

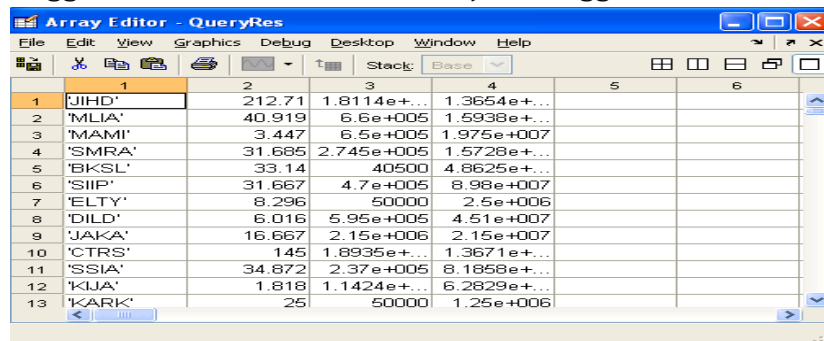
<i>Individual Index</i>	
Mean	47,25970027
Standard Error	0,589236739
Median	30,489
Mode	1,48
Standard Devi:	50,9613139
Sample Variar	2597,055514
Kurtosis	4,720926631
Skewness	2,075157133
Range	314,107
Minimum	0,893
Maximum	315
Sum	353502,558
Count	7480

**Gambar 1. Deskripsi IndIndex**

Pada proses data *preparation* ada beberapa tahapan yang pada dasarnya bertujuan untuk *cleaning* data, Hasil data *cleaning* ditampilkan sesuai dengan kebutuhan melalui perintah *query* (Gambar 2)

Algoritma FCM Klustering menentukan pusat kluster yang optimal. Pusat kluster belum akurat pada kondisi awal, iterasi akan berulang sampai konvergen. Setelah konvergen matriks partisi untuk kluster optimal yaitu derajat keanggotaan masing-masing titik data ke kluster. Banyaknya kolom menunjukkan banyaknya cluster dan banyaknya baris menunjukkan banyaknya data. Jumlah keanggotaan adalah satu, titik data akan menjadi anggota kluster pada keanggotaan terbesar. Sebagai contoh beberapa data

pada 3 (tiga) cluster yang sudah ditentukan berdasarkan aturan yakni klaster 1, klaster 2 dan klaster 3. Misal data pada urutan pertama bisa saja menjadi anggota cluster 2 karena derajat keanggotaan di cluster 2 adalah 0,99188 derajat keanggotaan tertinggi, data tersebut juga bisa menjadi anggota cluster 1 dan cluster 3 tetapi dengan nilai keanggotaan yang kecil yaitu masing-masing 0,000419 dan 0,0077034. Demikian seterusnya semua data menjadi anggota cluster berdasarkan derajat keanggotaan.



	1	2	3	4	5	6
1	'JIHD'	212.71	1.8114e+...	1.3654e+...		
2	'MLIA'	40.919	6.6e+005	1.5938e+...		
3	'MAMI'	3.447	6.5e+005	1.975e+007		
4	'SMRA'	31.685	2.745e+005	1.5728e+...		
5	'BKSL'	33.14	40500	4.8625e+...		
6	'SIIP'	31.667	4.7e+005	8.98e+007		
7	'ELTY'	8.296	50000	2.5e+006		
8	'DILD'	6.016	5.95e+005	4.51e+007		
9	'JAKA'	16.667	2.15e+006	2.15e+007		
10	'CTRS'	145	1.8935e+...	1.3671e+...		
11	'SSIA'	34.872	2.37e+005	8.1858e+...		
12	'KJJA'	1.818	1.1424e+...	6.2829e+...		
13	'KARIK'	25	50000	1.25e+006		

Gambar 2. Sebagian Hasil *Query Code, Individual Index, Volume, Amount*

Output FCM klustering adalah pusat klaster dengan beberapa derajat keanggotaan yang merupakan informasi untuk membangun fuzzy inference system. Pada dasarnya FIS adalah proses yang mensimulasikan input yang diberikan dengan menggunakan logika fuzzy yaitu mengelola sekelompok input menjadi output yang memenuhi aturan-aturan yang telah ditentukan.

Secara umum dalam menganalisis masalah dilakukan melalui beberapa tahapan proses analisis sistem fuzzy. Contoh hasil proses analisis sistem fuzzy untuk data Individual Index (67), Volume (150.000.000) dan Amount (40.000.000.000) adalah *IF individual Index rendah AND volume besar AND Amount besar THEN kriteria beresiko rendah*. Untuk melihat rata-rata transaksi masing-masing perusahaan, data pada output diolah lagi dengan cara menghitung rata-rata output masing-masing perusahaan kemudian hasil terakhir ditentukan apakah perusahaan termasuk beresiko tinggi, beresiko sedang atau beresiko rendah.

Berdasarkan input, dan aturan *fuzzy* yang sudah ditentukan maka diperoleh hasil dari 4360 transaksi selama periode penelitian, hanya ada 22 transaksi yang bisa dikategorikan beresiko rendah, 478 transaksi bisa dikategorikan beresiko sedang dan ada 3860 transaksi yang bisa dikategorikan beresiko tinggi. Berdasarkan output rata-rata transaksi dari masing-masing perusahaan maka diperoleh hasil, dari 35 perusahaan yang diteliti tidak terdapat perusahaan yang bisa dikategorikan beresiko rendah, dan hanya 3 perusahaan beresiko sedang dan ada 32 perusahaan yang bisa dikategorikan beresiko tinggi.

## KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan, dapat disimpulkan bahwa, model FCM klustering dapat menganalisis *big* data dengan banyak variabel dan rumit terutama untuk menemukan pola dari data, dengan pemodelan yang lebih luwes dan mempermudah pemecahan dalam perhitungan berdasarkan perumusan masalah. Data mining sangat menarik dan sangat bermanfaat terutama *big* data, variabel yang banyak dan rumit maka disarankan agar dilakukan penelitian lebih lanjut terkait data *mining* baik model klustering yang berbeda atau dengan model bukan klustering dengan berbagai aplikasi.

## DAFTAR PUSTAKA

- Al. Embaby, A., Shalaby, M. A. W., & Elsayed, K. M. (2020). FCM-based approach for locating visible videowatermarks. *Symmetry*, 12(3). <https://doi.org/10.3390/sym12030339>
- Arif, M. F., Anoraga, B., Handoyo, S., & Nasir, H. (2016). Algorithm Apriori Association Rule in Determination of Fuzzy Rule Based on Comparison of Fuzzy Inference System (FIS) Mamdani Method and Sugeno Method. *Business Management and Strategy*, 7(1), 103. <https://doi.org/10.5296/bms.v7i1.9410>
- Arun, N. K., & Mohan, B. M. (2017). Modeling, stability analysis, and computational aspects of some simplest nonlinear fuzzy two-term controllers derived via center of area/gravity defuzzification. *ISA Transactions*, 70, 16–29. <https://doi.org/10.1016/j.isatra.2017.04.023>
- Bezdek, J. C. (2022). *Elementary Cluster Analysis: Four Basic Methods that (Usually) Work* (Vol. 59). River Publishers.
- Bizimana, H., & Altunkaynak, A. (2020). Modeling the initiation of sediment motion under a wide range of flow conditions using a Geno-Mamdani Fuzzy Inference System method. *International Journal of Sediment Research*, 35(5), 467–483. <https://doi.org/10.1016/j.ijsrc.2020.03.009>
- Bora, D. J., & Gupta, D. A. K. (2014). A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm. *International Journal of Computer Trends and Technology*, 10(2), 108–113. <https://doi.org/10.14445/22312803/ijctt-v10p119>
- Dubey, Y. K., & Mushrif, M. M. (2016). FCM Clustering Algorithms for Segmentation of Brain MR Images. *Advances in Fuzzy Systems*, 2016. <https://doi.org/10.1155/2016/3406406>
- Efendiyev, G. M., Mammadov, P. Z., Piriverdiyev, I. A., & Mammadov, V. N. (2016). Clustering of Geological Objects Using FCM-algorithm and Evaluation of the Rate of Lost Circulation. *Procedia Computer Science*, 102(August), 159–162. <https://doi.org/10.1016/j.procs.2016.09.383>
- Gao, T., Li, A., & Meng, F. (2017). Research on Data Stream Clustering Based on FCM Algorithm1. *Procedia Computer Science*, 122, 595–602. <https://doi.org/10.1016/j.procs.2017.11.411>
- Hair, J. F. J., Black, W. C., Babin, B. J., Anderson, R. E., Black, W. C., & Anderson, R. E. (2018). *Multivariate Data Analysis*. <https://doi.org/10.1002/9781119409137.ch4>



- Hannachi, A. (2021). *Patterns Identification and Data Mining in Weather and Climate* (Cham, Swit). Springer Atmospheric Sciences.  
<http://www.springer.com/series/10176>
- Kantardzic, M. (2020). *Data Mining Concepts, Models, Methods, and Algorithms* (3rd ed.). John Wiley & Sons, Inc.
- Kusumadewi, S. (2002). *Analisis & Desain Sistem Fuzzy Menggunakan Toolbox Matlab*.
- M. Abdalla, A. M., Dress, S., & Zaki, N. (2011). Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network. *International Journal of Computer Science and Information Technology*, 3(3), 176–186. <https://doi.org/10.5121/ijcsit.2011.3312>
- Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., & Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048–3061. <https://doi.org/10.1109/TKDE.2019.2962680>
- Miao, J., Zhou, X., & Huang, T. Z. (2020). Local segmentation of images using an improved fuzzy C-means clustering algorithm based on self-adaptive dictionary learning. *Applied Soft Computing Journal*, 91, 106200. <https://doi.org/10.1016/j.asoc.2020.106200>
- Ni, Q., Pan, Q., Du, H., Cao, C., & Zhai, Y. (2017). A Novel Cluster Head Selection Algorithm Based on Fuzzy Clustering and Particle Swarm Optimization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 14(1), 76–84. <https://doi.org/10.1109/TCBB.2015.2446475>
- Pellegrini, M., Zoghi, M., & Jaberzadeh, S. (2018). Cluster analysis and subgrouping to investigate inter-individual variability to non-invasive brain stimulation: A systematic review. *Reviews in the Neurosciences*, 29(6), 675–697. <https://doi.org/10.1515/revneuro-2017-0083>
- Rahakbauw, D. L., Ilwaru, V. Y. I., & Hahury, M. H. (2017). Implementasi Fuzzy C-Means Clustering Dalam Penentuan Beasiswa. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 11(1), 1–12. <https://doi.org/10.30598/barekengvol11iss1pp1-12>
- Ruspini, E. H., Bezdek, J. C., & Keller, J. M. (2019). Fuzzy clustering: A historical perspective. *IEEE Computational Intelligence Magazine*, 14(1), 45–55. <https://doi.org/10.1109/MCI.2018.2881643>
- Singh, M. (2015). *Study on Clustering of Data*. 11(10), 55–60.
- Thalapala, V. S., & Guravaiah, K. (2022). FCMCP: Fuzzy C-Means for Controller Placement in Software Defined Networking. *Procedia Computer Science*, 201, 109–116. <https://doi.org/10.1016/j.procs.2022.03.017>
- Wang, K. (2001). Computational Intelligence in Agile Manufacturing Engineering. In A. Gunasekaran (Ed.), *AGILE MANUFACTURING: THE 21st CENTURY COMPETITIVE STRATEGY* (pp. 297–315). Elsevier Science Ltd.
- Wang, L.-X. (1997). A Course in Fuzzy Systems and Control. *Design*, 448. <http://portal.acm.org/citation.cfm?id=248374&dl=>
- Wiemer, H., Drowatzky, L., & Ihlenfeldt, S. (2019). Applied Acienes Data Mining Methodology for Engineering Applications ( DMME )— A Holistic

- Extension. *Applied Sciences*.
- Wu, C. H., Ouyang, C. Sen, Chen, L. W., & Lu, L. W. (2015). A new fuzzy clustering validity index with a median factor for centroid-based clustering. *IEEE Transactions on Fuzzy Systems*, 23(3), 701–718. <https://doi.org/10.1109/TFUZZ.2014.2322495>
- Yiping, W., Buqing, S., Jianjun, W., Qing, W., Haowu, L., Zhanxiang, L., Ningning, Z., & Qingchao, C. (2021). An improved multi-view collaborative fuzzy C-means clustering algorithm and its application in overseas oil and gas exploration. *Journal of Petroleum Science and Engineering*, 197(xxxx), 108093. <https://doi.org/10.1016/j.petrol.2020.108093>
- Zadissa, A., & Apweiler, R. (2022). Data Mining, Quality and Management in the Life Sciences. In *Methods in Molecular Biology* (Vol. 2449). [https://doi.org/10.1007/978-1-0716-2095-3\\_1](https://doi.org/10.1007/978-1-0716-2095-3_1)
- Zhang, M., Zhu, L., Sun, Y., Niu, D., & Liu, J. (2022). Computed tomography of ground glass nodule image based on fuzzy C-means clustering algorithm to predict invasion of pulmonary adenocarcinoma. *Journal of Radiation Research and Applied Sciences*, 15(1), 152–158. <https://doi.org/10.1016/j.jrras.2022.01.015>
- Zhi, H., & Liu, S. (2020). Gray image segmentation based on fuzzy c-means and artificial bee colony optimization. *Journal of Intelligent & Fuzzy Systems*, 38(4), 3647–3655. <https://doi.org/10.3233/jifs-179587>
- Zhu, L. F., Wang, J. S., & Wang, H. Y. (2019). A Novel Clustering Validity Function of FCM Clustering Algorithm. *IEEE Access*, 7, 152289–152315. <https://doi.org/10.1109/ACCESS.2019.2946599>