

PENERAPAN FUZZY C-MEANS DALAM KLASTERISASI RISIKO DAN KINERJA SAHAM PERBANKAN INDONESIA

Muhammad Aditya Pratama, Fajar Suryani, Aprilisa Arum Sari

Teknik Informatika, Universitas Duta Bangsa

Jl. Bhayangkara No. 55, Tipes, Serengan, Kota Surakarta, Jawa Tengah 57154

aditya32r@gmail.com

ABSTRAK

Perkembangan pasar modal Indonesia, khususnya pada sektor perbankan, mendorong investor untuk melakukan analisis saham yang lebih akurat dalam menilai tingkat risiko dan profitabilitas emiten. Namun, karakteristik data keuangan dan risiko saham perbankan yang kompleks serta cenderung tumpang tindih (*overlap*) menyulitkan proses pengelompokan menggunakan metode klasterisasi konvensional yang bersifat tegas. Penelitian ini bertujuan mengelompokkan saham perbankan berdasarkan indikator kinerja dan risiko guna mendukung pengambilan keputusan investasi yang lebih objektif. Metode yang digunakan adalah Fuzzy C-Means (FCM) dengan memanfaatkan data rasio keuangan dan risiko pasar yang meliputi *Return on Assets* (ROA), *Return on Equity* (ROE), *Earnings per Share* (EPS), *Debt to Equity Ratio* (DER), *Non-Performing Loan* (NPL), volatilitas, dan *Value at Risk* (VaR). Tahapan penelitian meliputi eksplorasi data, prapemrosesan, pembentukan model klasterisasi, serta validasi dan evaluasi model. Hasil penelitian menunjukkan bahwa FCM berhasil membentuk tiga klaster saham perbankan dengan karakteristik risiko dan profitabilitas yang berbeda. Nilai Fuzzy *Partition Coefficient* (FPC) sebesar 0,5170 menunjukkan kualitas klasterisasi yang cukup baik meskipun terdapat karakteristik data yang saling tumpang tindih. Dengan demikian, Fuzzy C-Means efektif digunakan untuk mendukung analisis investasi berbasis data pada saham sektor perbankan.

Kata Kunci: Fuzzy C-Means, Klasterisasi Saham, Saham Perbankan, Risiko Investasi, Kinerja Keuangan.

1. PENDAHULUAN

Sektor perbankan merupakan pilar utama ekonomi nasional yang mendominasi kapitalisasi pasar di Bursa Efek Indonesia (BEI). Peran perbankan sebagai lembaga intermediasi keuangan menjadikan performa emiten perbankan sebagai indikator utama stabilitas ekonomi makro. Bagi investor, kemampuan untuk menilai kesehatan dan stabilitas sebuah bank merupakan langkah penting dalam menentukan keputusan investasi yang tepat agar dapat meminimalkan potensi kerugian [1] [2].

Dalam beberapa tahun terakhir, pasar modal Indonesia mencatat peningkatan partisipasi investor ritel yang sangat signifikan, terutama pada saham-saham perbankan kategori *blue chip* maupun bank *digital*. Namun, pertumbuhan antusiasme investasi ini tidak selalu diiringi dengan tingkat literasi keuangan yang memadai bagi para pelaku pasar. Berdasarkan Survei Nasional Literasi dan Inklusi Keuangan (SLINK) oleh Otoritas Jasa Keuangan (OJK), masih terdapat kesenjangan yang signifikan antara akses layanan keuangan dengan pemahaman cara memanfaatkan layanan tersebut secara tepat [3] [4].

Kesenjangan literasi ini menyebabkan banyak investor menghadapi hambatan dalam memetakan profil risiko secara objektif, sehingga keputusan investasi sering kali didasari oleh pertimbangan yang tidak rasional. Situasi ini semakin rumit karena data laporan keuangan perbankan bersifat sangat dinamis dan kerap berada pada zona ketidakpastian. Sebagai contoh, sebuah bank dapat menunjukkan kinerja profitabilitas yang tinggi namun pada waktu

bersamaan memiliki profil risiko volatilitas yang juga kategori tinggi [5]

Permasalahan yang muncul adalah metode pengelompokan yang selama ini digunakan, yaitu hard clustering konvensional seperti K-Means, belum mampu mengakomodasi sifat data keuangan perbankan yang kompleks dan tumpang tindih, K-Means metode ini memang terbukti populer dan efektif untuk segmentasi data yang linier, seperti penggunaannya dalam mengelompokkan kinerja penjualan produk UMKM secara *online* [6]. Namun ketika diterapkan pada data laporan keuangan perbankan, metode tersebut memiliki kelemahan mendasar berupa sifatnya yang sangat tegas dan kaku [3].

Di mana setiap data dipaksa menjadi anggota hanya satu klaster secara mutlak. Selain itu, K-Means sangat rentan terhadap data pencilan seperti nilai rasio *Non-Performing Loan* (NPL) yang ekstrem, yang dapat menggeser pusat klaster (*centroid*) dan menghasilkan pemetaan yang tidak konsisten [7]

Berdasarkan permasalahan tersebut, penelitian ini bertujuan mengelompokkan saham perbankan berdasarkan indikator kinerja dan risiko guna mendukung pengambilan keputusan investasi yang lebih objektif. Rencana penyelesaian masalah yang diusulkan adalah mengimplementasikan metode Fuzzy C-Means (FCM) sebagai solusi untuk mengatasi keterbatasan-keterbatasan tersebut. FCM memberikan kemungkinan setiap saham untuk memiliki tingkat keterlibatan pada beberapa klaster secara bersamaan dengan skala nilai antara 0 hingga 1. Strategi ini terbukti lebih responsif terhadap dinamika volatilitas

di pasar, dan mampu menghasilkan pengelompokan yang lebih efektif mencerminkan situasi fundamental perbankan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Beberapa studi terdahulu telah mengeksplorasi pengembangan sistem bantuan keputusan dengan menerapkan berbagai teknik komputasi cerdas. Keandalan sistem komputasi untuk memandu keputusan pengguna ini telah terbukti aplikatif di berbagai bidang, mulai dari sistem rekomendasi personalisasi destinasi pariwisata [8], hingga pengelompokan data yang rumit di pasar modal [2], [5].

Khusus pada sektor finansial, sistem ini dirancang agar investor dapat memetakan kondisi dasar perusahaan secara rasional, sehingga risiko potensial bisa dikenali lebih dini. Selain kapabilitas dalam mengolah informasi laporan keuangan, pembangunan model pengelompokan saham yang andal juga harus mempertimbangkan sifat data pasar yang tumpang tindih (*overlapping*) dan berubah-ubah [5].

Penggunaan beberapa parameter sekaligus dalam kajian saham perbankan menjadi salah satu cara penting untuk meningkatkan keandalan hasil analisis yang dihasilkan. Informasi rasio keuntungan seperti ROA dan ROE dapat memberikan gambaran tentang seberapa efisien bank dalam mengelola operasinya [9], sementara ukuran-ukuran risiko seperti NPL dan *Value at Risk* (VaR) berfungsi sebagai indikator utama guna menilai sejauh mana stabilitas perusahaan terhadap perubahan pasar. Mengingat data-data tersebut kerap bersifat tidak jelas atau berada pada zona ketidakpastian, maka diperlukan pendekatan yang mampu menangkap metode *Fuzzy C-Means* (FCM) untuk mengolah gabungan data rasio keuangan dan profil risiko tersebut, sehingga status pencapaian kinerja dan tingkat risiko saham perbankan dapat ditentukan dengan cara yang lebih luwes dan sistematis dibandingkan metode pengelompokan rigid yang sudah ada selama ini [7].

Penelitian mengenai penerapan algoritma pengelompokan dalam investasi perbankan telah berkembang pesat mengikuti kompleksitas data pasar modal [10].

Beberapa studi menunjukkan keandalan metode *Fuzzy C-Means* (FCM) dalam memetakan saham perbankan berdasarkan data harga historis dan mengidentifikasi kategori kinerja yang berbeda-beda. Pengembangan selanjutnya memanfaatkan rasio keuangan fundamental seperti EPS, ROA, ROE, dan NPM sebagai variabel evaluasi kinerja perbankan, menghasilkan klasifikasi kinerja yang lebih komprehensif [9].

Tantangan dalam menangani ambiguitas data pasar menjadi fokus penelitian berikutnya. Penelitian sebelumnya menerapkan algoritma *K-Means* untuk memetakan struktur bank berdasarkan aset dan laba,

namun metode *hard clustering* menunjukkan keterbatasan dalam menangani data dengan karakteristik tumpang tindih atau samar-samar [1].

Studi komparatif kemudian membuktikan bahwa fleksibilitas nilai keanggotaan pada FCM lebih akurat dalam menangkap profil risiko yang kompleks dibandingkan metode pengelompokan kaku [7]. Penelitian lanjutan mengkonfirmasi validitas FCM sebagai dasar penentuan strategi investasi dan optimasi portofolio di bursa saham [11].

2.2. Investasi Saham Perbankan

Saham merupakan surat berharga yang mencerminkan hak kepemilikan seorang investor atas sebuah perusahaan sesuai dengan modal yang telah mereka investasikan. Sektor perbankan di Indonesia memiliki posisi yang sangat penting sebagai lembaga perantara keuangan yang menjadi penopang pertumbuhan ekonomi nasional [2], [12].

Mengingat beragamnya jenis pelaku di sektor ini, mulai dari bank BUMN, bank swasta, hingga bank digital, kemampuan mengevaluasi kondisi kesehatan sebuah bank menjadi hal yang tidak dapat ditawar bagi investor untuk membuat keputusan investasi yang benar dan menguntungkan. Investasi di sektor perbankan diketahui membawa profil risiko yang cukup besar, sehingga diperlukan penelitian mendalam terhadap kondisi fundamental untuk mengurangi kemungkinan terjadinya kerugian dan meraih keuntungan dari apresiasi harga saham maupun pembagian dividen [12] [2].

2.3. Data Mining dan Klasterisasi

Pengelompokan data atau *clustering* merupakan salah satu teknik dalam data *mining* yang termasuk dalam kategori pembelajaran tanpa pengawasan (*unsupervised learning*), yang bertugas mengorganisir objek-objek data secara natural berdasarkan kesamaan karakteristik mereka tanpa membutuhkan penentuan label kelas terlebih dahulu. Maksud dari teknik ini adalah memastikan bahwa objek-objek yang berada dalam satu kelompok memiliki tingkat keserupaan yang tinggi di antara sesama anggota (*homogenitas intra-klaster*) sementara objek-objek dari kelompok berbeda menunjukkan perbedaan yang jelas dan nyata (*heterogenitas antar-klaster*) [13].

Metode pengelompokan data dapat dibedakan menjadi dua kategori utama. Metode pengelompokan tegas atau *hard clustering* seperti *K-Means* yang menempatkan data ke dalam satu kelompok secara mutlak dan kaku, dan pengelompokan lunak atau *soft clustering* seperti FCM yang dirancang untuk menangani ketidakjelasan data dengan atribut-atribut yang saling tumpang tindih atau bersifat samar-samar [5].

2.4. Rasio Profitabilitas (Kinerja)

Rasio profitabilitas berfungsi untuk mengukur seberapa besar kemampuan perusahaan dalam menciptakan keuntungan dari berbagai strategi dan

keputusan operasional yang telah diterapkan. Indikator utama yang menjadi acuan adalah *Return on Assets* (ROA) yang mengukur sejauh mana efisiensi bank dalam menghasilkan keuntungan dari keseluruhan aset yang dimilikinya, serta *Return on Equity* (ROE) yang mengukur berapa besar tingkat pengembalian atas modal yang telah diinvestasikan oleh para pemegang saham. Nilai ROA dan ROE yang tinggi mencerminkan performa bank yang sehat dan beroperasi secara efisien [14].

$$ROA = \frac{\text{Laba Bersih}}{\text{Total Aset}} \times 100 \quad (1)$$

$$ROE = \frac{\text{Laba Bersih}}{\text{Total Ekuitas}} \times 100 \quad (2)$$

2.5. Rasio Risiko Operasional dan Kredit

Rasio risiko dimanfaatkan untuk mengevaluasi tingkat ancaman yang dihadapi oleh bank, khususnya berkaitan dengan kapasitas untuk memenuhi kewajiban finansial dan kualitas dari aset-aset yang dimilikinya. *Debt to Equity Ratio* (DER) membandingkan jumlah keseluruhan utang terhadap modal sendiri guna melihat seberapa besar risiko ketidakmampuan bayar dari sebuah emiten. Di sisi lain, *Non-Performing Loan* (NPL) merupakan indikator utama risiko dalam pemberian kredit yang menampilkan proporsi pinjaman yang mengalami kesulitan pelunasan, di mana para regulator umumnya menentukan batas keamanan di angka kurang dari 5% [4].

$$DER = \frac{\text{Total Liabilitas}}{\text{Total Ekuitas}} \quad (3)$$

$$NPL = \frac{\text{Total Kredit Bermasalah}}{\text{Total Kredit}} \times 100 \quad (4)$$

2.6. Analisis Risiko Pasar

Profil risiko diukur untuk memahami seberapa luas sebaran fluktuasi harga saham dengan menggunakan metrik volatilitas dan *Value at Risk* (VaR). Volatilitas dihitung berdasarkan simpangan baku dari return saham guna mengidentifikasi instrumen investasi yang mampu memberikan hasil pengembalian yang cukup untuk mengkompensasi risiko yang ditanggung. Adapun VaR digunakan untuk memperkirakan besarnya kerugian maksimal yang kemungkinan dapat terjadi pada suatu portofolio investasi dalam periode waktu dan tingkat keyakinan tertentu [3].

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (R_i - \bar{R})^2}{n - 1}} \quad (5)$$

$$VaR = W_0 \times Z_\alpha \times \sigma \times \sqrt{t} \quad (6)$$

2.7. Metode Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) didasarkan pada prinsip-prinsip logika *Fuzzy* yang menggunakan tingkat kebenaran (*degree of truth*) sebagai pengganti nilai-nilai yang mutlak untuk memproses data yang bersifat tidak jelas atau ambigu. Mekanisme inti dari FCM adalah menentukan tingkat afiliasi (u_{ij}) untuk setiap data terhadap setiap kelompok dalam rentang nilai

mulai dari 0 hingga 1. Algoritma ini bekerja dengan cara meminimalkan fungsi objektif J_m melalui proses iteratif dengan melakukan pembaruan secara berulang pada matriks afiliasi dan pusat-pusat dari setiap kelompok. Keunggulan pendekatan ini memungkinkan sebuah saham bank untuk memiliki tingkat keterlibatan sebagian pada lebih dari satu kategori profil risiko sekaligus, sehingga menghasilkan analisis yang lebih adaptif dengan perubahan kondisi pasar [5]

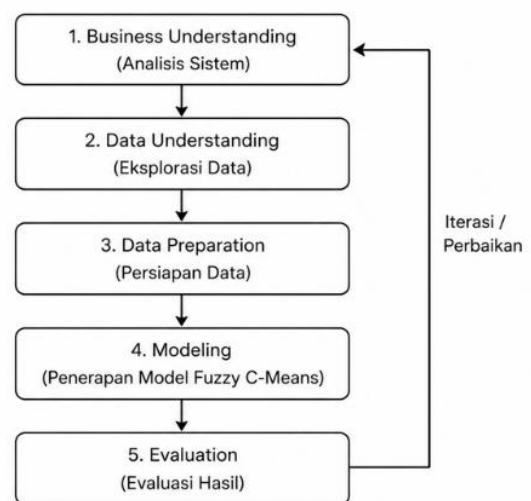
$$J_m = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|x_i - v_j\|^2 \quad (7)$$

3. METODE PENELITIAN

Studi ini menerapkan metodologi kuantitatif dengan memanfaatkan teknik *Unsupervised Learning* guna melakukan pengelompokan terhadap profil risiko dan kinerja dari emiten-emiten perbankan. Pemilihan kerangka CRISP-DM didasarkan pada kemampuannya untuk mendukung proses iteratif dalam pengolahan data, memungkinkan pencapaian hasil klusterisasi yang optimal.

3.1. Tahapan Penelitian

Tahapan penelitian dirancang secara sistematis dan berurutan agar proses pengembangan model berjalan terarah, meliputi persiapan data, eksplorasi karakteristik dataset, prapemrosesan data, pembentukan model, serta validasi dan evaluasi hasil. Alur keseluruhan tahapan penelitian disajikan pada gambar 1.



Gambar 1. Tahapan pengembangan sistem berdasarkan CRISP-DM

Pada gambar 1 ditunjukkan tahapan alur pengembangan sistem, Penjelasan setiap tahapan penelitian adalah sebagai berikut :

a. Analisis Sistem

Tahap awal dilakukan untuk mengidentifikasi sumber data, mengumpulkan data rasio keuangan dan indikator pasar dari emiten perbankan, serta

melakukan persiapan awal terhadap kualitas data yang akan digunakan dalam analisis.

b. Eksplorasi Data

Tahap ini berfokus pada pemahaman mendalam terhadap karakteristik dataset, distribusi nilai variabel, serta identifikasi potensi data anomali atau nilai ekstrim yang memerlukan penanganan khusus.

c. Prapemrosesan Data

Tahap prapemrosesan mencakup pembersihan data untuk menangani nilai yang hilang atau anomali, normalisasi untuk menyamakan skala antar variabel, serta pemisahan dataset untuk keperluan model development dan validasi.

d. Pembentukan Model Klasterisasi

Tahap ini merupakan inti dari implementasi algoritma. Model *Fuzzy C-Means* diterapkan terhadap data yang telah dipreparasi dengan melakukan iterasi dan penyesuaian parameter hingga diperoleh hasil klasterisasi yang stabil dan optimal.

e. Validasi dan Evaluasi Model

Tahap validasi dilakukan untuk menilai keandalan dan kualitas hasil klasterisasi yang dihasilkan melalui perhitungan metrik-metrik validitas serta analisis interpretasi klaster.

3.2. Kebutuhan Sistem

Kebutuhan pada penelitian ini mencakup data dan lingkungan pemrograman yang diperlukan untuk implementasi model. Data yang digunakan adalah data sekunder berupa laporan keuangan dan data pasar dari emiten-emiten perbankan, sedangkan lingkungan pemrograman meliputi bahasa pemrograman, library, dan platform yang mendukung proses analisis data.

Tabel 1. Kebutuhan Lingkungan Pemrograman

No	Komponen	Spesifikasi
1	Bahasa Pemrograman	Python 3.8
2	Text Editor/IDE	Visual Studio Code
3	Library Utama	Pandas, NumPy, Scikit-learn
4	Library Visualisasi	Streamlit
5	Database/Penyimpanan	CSV

Tabel 2. Parameter Data yang Digunakan

No	Kategori	Indikator	Sumber
1	Kinerja Keuangan	ROA, ROE, NPM, EPS	Laporan Keuangan Tahunan
2	Risiko Operasional	DER, NPL	Laporan Keuangan Tahunan
3	Risiko Pasar	Volatilitas, VaR	Data Harga Saham Historis

3.3. Arsitektur Model

Arsitektur model klasterisasi yang dikembangkan dalam penelitian ini dirancang untuk mengintegrasikan data dari berbagai dimensi finansial dan mempresentasikannya dalam bentuk klaster yang bermakna. Model menggunakan *Fuzzy C-Means* sebagai algoritma utama dengan input berupa vektor

multidimensi yang terdiri dari indikator-indikator kinerja dan risiko.

Proses kerja model dapat dijelaskan sebagai berikut :

- Data input yang telah dinormalisasikan dimasukkan ke dalam algoritma FCM.
- Algoritma melakukan iterasi untuk menghitung tingkat keanggotaan fuzzy setiap observasi terhadap masing-masing klaster.
- Pusat-pusat klaster diperbarui secara iteratif hingga mencapai kondisi konvergensi.
- Setiap emiten diklasifikasikan ke dalam klaster berdasarkan nilai keanggotaan Fuzzy tertinggi, namun tetap mempertahankan informasi tingkat keterlibatan terhadap klaster lain.

3.4. Spesifikasi Algoritma Fuzzy C-Means

Fuzzy C-Means (FCM) merupakan algoritma pengelompokan yang bekerja dengan meminimalkan fungsi objektif melalui proses iterative. Fungsi objektif FCM didefinisikan sebagai :

$$J_m = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|x_i - v_j\|^2 \quad (8)$$

Di mana :

- u_{ij} = Derajat keanggotaan data ke-I pada klaster ke-j (rentang 0-1)
- m = Parameter *fuzziness* (umumnya $m = 2$)
- x_i = Data ke-i
- v_j = Pusat Klaster ke-j
- $\|\cdot\|$ = Norm Euclidean
- n = Jumlah Data
- c = Jumlah klaster

3.5. Proses Prapemrosesan Data

Prapemrosesan data dilakukan melalui serangkaian tahapan sebagai berikut :

a. Pembersihan Data

Tahap pembersihan mencakup identifikasi dan penanganan nilai yang hilang (*missing values*) serta deteksi data anomali atau *outlier*. Untuk nilai yang hilang, dilakukan interpolasi atau penghapusan baris data sesuai proporsi ketiadaan. Data *outlier* diidentifikasi menggunakan metode *Interquartile Range* (IQR) dan ditangani sesuai konteks bisnis.

b. Normalisasi Data

Mengingat variabel-variabel input memiliki satuan dan magnitude yang berbeda-beda (misalnya ROA dalam persen, NPL dalam nilai rasio), dilakukan normalisasi menggunakan metode *Min-Max scaling* untuk menyamakan rentang nilai ke dalam interval $[0,1]$.

c. Pemisahan Dataset

Dataset dibagi menjadi dua bagian: data training untuk pengembangan model (70%) dan data validasi untuk evaluasi model (30%).

3.6. Metrik Validasi dan Evaluasi

Evaluasi model dilakukan menggunakan metrik-metrik validitas klaster yang meliputi :

a. Silhouette Coefficient

Metrik ini mengukur seberapa baik setiap observasi berada dalam klusterisasi dibandingkan dengan kluster lain. Nilai berkisar -1 hingga 1, dengan nilai positif menunjukkan klusterisasi yang baik.

b. Davies-Bouldin Index

Metrik ini mengukur rasio rata-rata similarity antara masing-masing kluster dan kluster terdekatnya. Nilai yang lebih rendah menunjukkan klusterisasi yang lebih baik.

c. Calinski-Harabasz Index

Metrik ini menghitung rasio antara dispersi antar-kluster terhadap dispersi dalam-kluster. Nilai yang lebih tinggi menunjukkan kluster yang lebih terpisah dan kohesif.

4. HASIL DAN PEMBAHASAN

4.1. Analisis sistem

Sistem yang dikembangkan pada penelitian ini dirancang untuk melakukan analisis klusterisasi saham perbankan berdasarkan indikator risiko dan kinerja keuangan menggunakan metode *Fuzzy C-Means*. Sistem menerima data input berupa rasio keuangan dan risiko pasar, kemudian melakukan proses prapemrosesan, normalisasi, dan perhitungan derajat keanggotaan *fuzzy* sebelum menghasilkan kategori kluster saham.

Metode Fuzzy C-Means (FCM)

Input Data Preprocessing Modeling Hasil

Tp	ROA	ROE	NPM	DER	NPL	Volatilitas	VaR	EPS
BBCA	0.48	22.28	46.80	479.04	1.88	0.2299	0.0223	378.4
BBRI	0.49	16.78	22.95	502.89	0.05	0.2047	0.0260	349.0
MEDIA	0.47	16.48	81.97	5.37	1.45	0.3390	0.0273	34.1
BBNI	0.22	13.28	28.84	645.44	2.55	0.2071	0.0275	541.4
BRIS	0.16	16.41	18.72	694.23	2.23	0.4463	0.0391	123.2
BKIA	0.32	12.39	31.80	578.81	2.36	0.2204	0.0394	253.2
BKRN	1.46	7.34	16.47	3.44	2.22	0.2418	0.0221	324.0
ARTO	0.50	1.76	7.88	171.61	0.80	0.6233	0.0549	6.4
BBTN	0.92	13.25	9.56	14.26	3.28	0.3382	0.0303	244.0
BUMI	1.27	13.82	14.74	1082.07	1.75	0.1844	0.0370	146.0

Gambar 2 . Tampilan Input data awal

Pada gambar 2, ialah tampilan input data awal di sistem, implementasi sistem dilakukan menggunakan framework Streamlit berbasis Python sehingga proses analisis dilakukan secara interaktif dan real-time. Pengguna menginputkan nilai awal rata-rata dalam rentang waktu tertentu dari bank yang ingin diklusterisasikan, mulai dari nilai ROA, ROE, NPM dan variabel yang lainnya, sebelum ke proses selanjutnya.

4.2. Eksplorasi Data

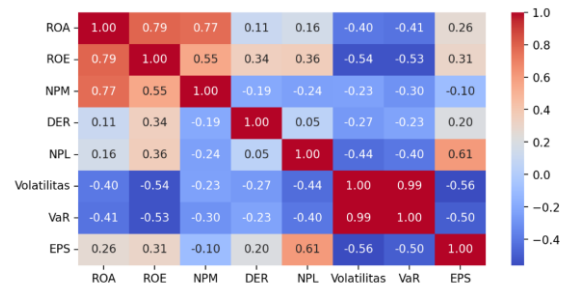
Tahap ini dilakukan untuk memastikan kualitas data sebelum masuk ke proses kalkulasi *fuzzy*. Subjek penelitian terdiri dari 10 emiten perbankan yang tercatat di Bursa Efek Indonesia.

a. Analisis Deskriptif

Berdasarkan eksplorasi data, variabel diukur menggunakan 8 parameter yang mencakup kinerja keuangan (ROA, ROE, NPM, EPS), risiko operasional (DER, NPL), dan risiko (Volatilitas, VaR).

b. Uji Korelasi

Analisis menggunakan *heatmap* menunjukkan adanya hubungan linier yang signifikan antar variabel profitabilitas. Hal ini mengonfirmasi bahwa bank dengan manajemen aset yang efisien cenderung menghasilkan pengembalian modal yang stabil.



Gambar 3 . Heatmap korelasi variabel keuangan

Gambar 3 menampilkan sebuah matriks korelasi dalam format *heatmap* yang berguna untuk melihat keterkaitan linear di antara berbagai variabel finansial dan risiko di sektor perbankan. Perhitungan yang digunakan adalah koefisien korelasi Pearson, dengan nilai berkisar dari -1 sampai 1. Pola warna dalam visualisasi ini cukup intuitif, semakin merah gelap warnanya, semakin kuat hubungan positif antar variabel, sementara warna biru gelap mengindikasikan adanya hubungan negatif yang kuat di antara keduanya.

4.3. Prapemrosesan Data

Pemodelan tahap prapemrosesan data dilakukan untuk memastikan kualitas data yang digunakan dalam proses klusterisasi berada pada kondisi optimal. Tahapan ini bertujuan mengurangi noise, menangani ketidakonsistenan data, serta menyamakan rentang nilai antar variabel agar proses perhitungan pada algoritma *Fuzzy C-Means* dapat berjalan secara lebih stabil dan akurat.

a. Pembersihan Data

Algoritma Proses pembersihan data dilakukan dengan mengidentifikasi nilai yang hilang (*missing value*) serta mendeteksi data ekstrem (*outlier*) pada setiap variabel penelitian. Deteksi outlier dilakukan menggunakan pendekatan *Interquartile Range* (IQR) untuk mengetahui data yang berada di luar batas distribusi normal. Berdasarkan hasil pemeriksaan, data yang digunakan tidak menunjukkan anomaly signifikan sehingga seluruh data tetap digunakan dalam proses analisis.

b. Normalisasi Data

Mengingat adanya perbedaan satuan (persentase pada ROA/NPL vs nilai nominal pada EPS), dilakukan normalisasi menggunakan metode *Min-Max Scaling* ke dalam interval [0,1]. Proses ini krusial untuk menjaga stabilitas jarak *Euclidean* pada algoritma FCM.

Data Normalisasi

Ticker	ROA	ROE	NPV	DER	NPV	Volatilitas	Vol	EPS
0 BBKA	0.9907	1	0.5589	0.4703	0.4055	0.3306	0.224	0.6852
1 BBRI	1	0.732	0.2139	0.3603	0.3673	0.2385	0.2042	0.6338
2 MEGA	0.9903	0.7271	1	0.3603	0.2621	0.3522	0.2719	0.9305
3 BBNI	0.5703	0.3614	0.3636	0.6429	0.6855	0.2568	0.2762	1
4 BRIS	0.5502	0.7139	0.1706	0.8309	0.5766	0.5946	0.5946	0.2125
5 BNGA	0.6807	0.518	0.3209	0.5702	0.429	0.1027	0.362	0.4533
6 BDMN	0.3213	0.2719	0.127	0	0.5726	0.1307	0.1382	0.5844
7 ARTO	0	0	0	0.3884	0	1	1	0
8 BBTN	0.1405	0.5099	0.0343	0.0338	1	0.3305	0.3306	0.4378
9 BJBR	0.2575	0.5977	0.1038	1	0.3833	0	0	0.2594

Gambar 4. Hasil Normalisasi Data Saham

Pada gambar 4, menampilkan hasil dari normalisasi data emiten perbankan yang telah ditransformasikan ke dalam rentang skala 0 hingga 1. Tahapan ini berfungsi untuk menyetarakan bobot seluruh variabel penelitian sehingga data siap diolah menggunakan algoritma *Fuzzy C-Means* tanpa adanya bias akibat perbedaan skala nilai nominal.

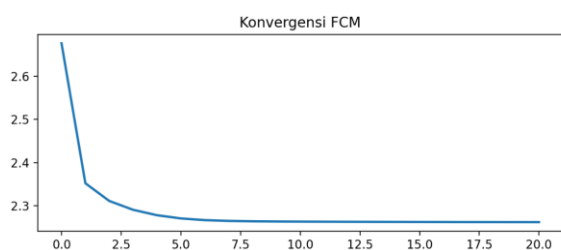
c. Kesiapan Data untuk Klasterisasi

Hasil prapemrosesan menunjukkan bahwa dataset telah berada dalam kondisi yang siap digunakan pada tahap pembentukan model klasterisasi. Seluruh variabel telah memiliki distribusi nilai yang lebih seimbang dan konsisten sehingga dapat meningkatkan stabilitas proses iterasi pada algoritma *Fuzzy C-Means*.

4.4. Pembentukan Model Klasterisasi

Pada tahap pembentukan model. Algoritma *Fuzzy C-Means* dijalankan dengan membagi data ke dalam tiga kelompok klaster. Penetapan jumlah tiga klaster ini didasarkan pada kebutuhan untuk mengklasifikasikan saham menurut profil risiko dan kinerjanya, yang mencakup kategori *low risk-high return*, *medium risk-medium return*, dan *high risk-low return*. Parameter *Fuzziness Exponent (m)* ditetapkan sebesar 2 dengan batas *error tolerance* sebesar 0,005 untuk mencapai konvergensi yang stabil.

Proses iterasi berjalan secara berulang dengan memperbarui pusat klaster dan derajat keanggotaan data hingga perubahan nilai fungsi objektif berada di bawah ambang toleransi yang ditetapkan, sehingga dihasilkan pusat klaster yang *konvergen* dan representatif terhadap karakteristik data saham perbankan.



Gambar 5. Grafik Konvergensi Fungsi Objektif FCM

Gambar 5 memperlihatkan tren penurunan nilai fungsi objektif selama proses iterasi *Fuzzy C-Means* berlangsung, di mana nilai tersebut terus menurun secara bertahap hingga mencapai kestabilan pada iterasi tertentu. Kondisi ini mengindikasikan bahwa

model telah berhasil *konvergen*, sehingga hasil klasterisasi layak digunakan sebagai dasar analisis pada tahap berikutnya.

Inti dari metode FCM terletak pada nilai derajat keanggotaan (u_{ij}), yang memungkinkan setiap emiten memiliki tingkat afiliasi pada masing-masing klaster alih-alih hanya tergabung secara mutlak dalam satu kelompok. Berdasarkan nilai keanggotaan tertinggi yang diperoleh pada matriks *membership degree*, klasifikasi terbagi sebagai berikut :

a. Klaster 3 (*High Risk – Low Return*)

Kelompok ini dihuni oleh BBKA, BBRI, MEGA, dan BBNI, yang secara keseluruhan dapat dikategorikan sebagai pemimpin industri. Kinerjanya tercermin dari rata-rata ROA sebesar 3,17% dan ROE 17,25%, dengan tingkat volatilitas terendah di angka 0,2876 yang mencerminkan stabilitas pasar paling solid di antara kelompok lainnya.

b. Klaster 2 (*Medium Risk – Medium Return*)

Terdiri atas emiten lapis kedua seperti BDMN dan BBTN, serta beberapa emiten yang berada di zona transisi dengan nilai keanggotaan yang cukup bersaing. Kelompok ini mencerminkan karakteristik perbankan dengan tingkat profitabilitas yang berada di kisaran menengah, yakni ROA 1,63% dan ROE 12,64%. Kelompok ini menawarkan potensi pertumbuhan yang cukup menjanjikan seiring dengan eksposur risiko yang moderat

c. Klaster 1 (*Low Risk – High Return*)

Posisi ini terdiri atas BRIS, ARTO, dan BJBR yang tergabung dalam Klaster 1 (*Low Risk – High Return*). Kelompok ini memiliki rata-rata *Return on Assets* (ROA) sebesar 1,03%, *Return on Equity* (ROE) sebesar 11,13%, dan volatilitas sebesar 0,500. Hasil perhitungan *Fuzzy C-Means* menunjukkan bahwa ketiga emiten memiliki karakteristik yang lebih dekat dengan kategori *Low Risk – High Return* berdasarkan kombinasi variabel profitabilitas dan risiko yang dianalisis.

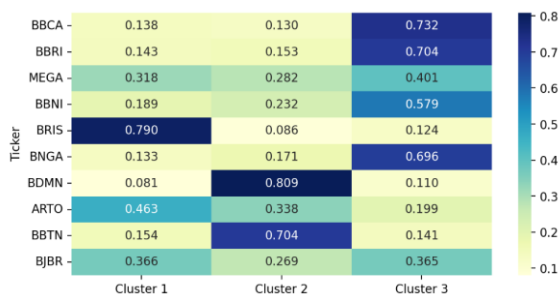
Membership Degree

Ticker	Cluster 1	Cluster 2	Cluster 3	Cluster
0 BBKA	0.2376	0.13	0.7504	3
1 BBRI	0.3426	0.1529	0.7545	3
2 MEGA	0.3177	0.2815	0.4008	3
3 BBNI	0.3889	0.2318	0.5792	3
4 BRIS	0.7962	0.0861	0.1187	1
5 BNGA	0.1333	0.1711	0.6956	3
6 BDMN	0.0896	0.8993	0.1101	2
7 ARTO	0.4031	0.1381	0.1988	1
8 BBTN	0.1242	0.7045	0.1613	2
9 BJBR	0.3058	0.2089	0.3853	1

Gambar 6. Matriks Derajat Keanggotaan Saham

Pada gambar 6, menyajikan derajat keanggotaan (*membership degree*) untuk masing-masing emiten perbankan pada tiga klaster yang berbeda. Setiap baris merepresentasikan seberapa besar kemungkinan suatu saham termasuk ke dalam klaster tertentu, yang ditentukan berdasarkan karakteristik finansial dan risiko yang telah diproses melalui algoritma *Fuzzy C-Means*. Kolom terakhir menampilkan penugasan klaster final untuk setiap emiten, ditetapkan

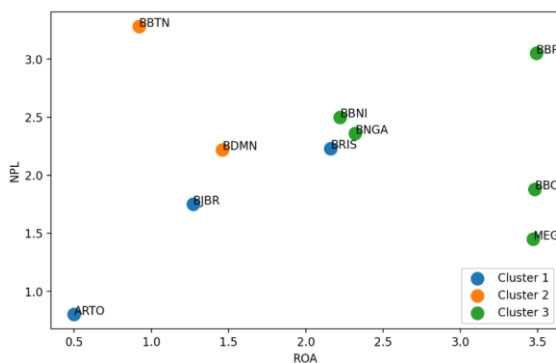
berdasarkan nilai keanggotaan yang paling tinggi, dan inilah yang menjadi dasar utama untuk mengklasifikasikan performa bank tersebut.



Gambar 7. Heatmap Derajat Keanggotaan

Gambar 7, memvisualisasikan *heatmap* distribusi derajat keanggotaan dengan menggunakan gradasi warna, sehingga memudahkan pembacaan hasil klastering. Melalui intensitas warna yang berbeda-beda, dapat dilihat dengan jelas emiten mana yang jelas termasuk dalam suatu klaster tertentu dan mana yang nilai keanggotaannya lebih terdistribusi di antara beberapa klaster atau berada di zona transisi antar kelompok. Pendekatan visual ini terbukti sangat berguna untuk melakukan *risk profiling* dengan cepat dan efisien, terutama dalam membedakan profil risiko dan kinerja di antara berbagai saham perbankan yang menjadi fokus penelitian.

4.5. Validasi dan Evaluasi Model



Gambar 8. Visualisasi Scatter Plot ROA vs NPL

Hasil klasterisasi ini memberikan sejumlah wawasan yang relevan bagi pengelolaan portofolio secara strategis. Masuknya BBKA dan BBRI ke dalam kelompok *Low Risk* menjadi bukti bahwa pendekatan logika *fuzzy* cukup andal dalam mendeteksi kestabilan fundamental suatu emiten di tengah dinamika pasar yang terus berfluktuasi. Di sisi lain, keberadaan ARTO sebagai satu-satunya penghuni Klaster 1 mengisyaratkan adanya anomali kinerja yang semestinya mendapat perhatian lebih serius dari perspektif manajemen risiko. Gambaran ini semakin diperkuat melalui visualisasi *scatter plot* antara ROA dan NPL, di mana terlihat pemisahan yang cukup tegas antarkelompok bank-bank dengan fundamental yang kokoh cenderung berkumpul pada zona profitabilitas

tinggi sekaligus mencatatkan eksposur risiko kredit yang relatif rendah.

Gambar 8 menyajikan visualisasi sebaran (*scatter plot*) posisi emiten perbankan berdasarkan dua indikator utama, yaitu profitabilitas (ROA) dan risiko kredit (NPL). Melalui visualisasi ini, terlihat jelas pemisahan karakteristik antarkelompok saham. Kelompok saham pada Klaster 3 (warna hijau) berkumpul secara konsisten pada zona kanan bawah, yang mengindikasikan tingkat profitabilitas tinggi dengan eksposur risiko kredit yang relatif rendah. Sebaliknya, ARTO pada Klaster 1 (warna biru) berada terisolasi di area kiri bawah, menunjukkan profitabilitas yang rendah dengan volatilitas tinggi, sementara Klaster 2 (warna oranye) menjembatani kedua kondisi tersebut di area menengah.

Hasil Klasterisasi

Ticker	ROA	ROE	NPL	DPR	NPM	Profitabilitas	Valu	EPS	Cluster	Profit
0 BBCA	3.48	22.28	48.68	479.04	1.88	0.2289	0.0213	378.4	3	High Risk - Low Return
1 BBRI	3.49	16.78	22.95	503.89	3.05	0.2847	0.0266	349.6	3	High Risk - Low Return
2 MEGA	3.47	16.68	81.57	5.37	1.45	0.339	0.0273	14.07	3	High Risk - Low Return
3 BBNI	2.22	13.28	29.64	645.44	2.5	0.2971	0.0275	548.4	3	High Risk - Low Return
4 BRIS	2.16	16.41	19.72	834.23	2.23	0.4453	0.0391	123.16	1	Low Risk - High Return
5 BNGA	2.32	12.39	31.6	578.81	2.36	0.2284	0.0194	253.17	3	High Risk - Low Return
6 BDMN	1.46	7.34	16.47	3.44	2.22	0.2458	0.0235	324	2	Medium Risk - Medium Return
7 ARTO	0.5	1.76	7	171.61	0.8	0.0333	0.0549	6.41	1	Low Risk - High Return
8 BBTN	0.92	13.25	9.56	14.26	3.28	0.3362	0.0303	244.8	2	Medium Risk - Medium Return
9 BJBR	1.27	13.62	14.74	1062.97	1.75	0.3844	0.017	106.86	1	Low Risk - High Return

Gambar 9. Hasil Klasterisasi

Pada gambar 9, menyajikan hasil akhir pengelompokan emiten perbankan berdasarkan profil risiko dan kinerjanya, yang diperoleh melalui pengolahan data dengan algoritma *Fuzzy C-Means*. Emiten terbagi menjadi tiga klaster kategori, *High Risk - Low Return* (Klaster 3) untuk saham dengan risiko tinggi namun return rendah, *Medium Risk - Medium Return* (Klaster 2) untuk saham dengan profil menengah, dan *Low Risk - High Return* (Klaster 1) untuk saham dengan risiko rendah dan return tinggi. Tabel ini merangkum kesimpulan dari seluruh analisis dan menjadi acuan utama dalam memetakan karakteristik kinerja saham-saham perbankan yang terdaftar di Bursa Efek Indonesia (BEI) dalam penelitian ini.

Hasil pengujian menunjukkan bahwa algoritma *Fuzzy C-Means* berhasil membentuk tiga klaster saham perbankan dengan profil risiko dan profitabilitas yang terbedakan secara cukup jelas, dengan nilai FPC sebesar 0,5170 yang mengindikasikan kualitas pemisahan klaster yang memadai meskipun data memiliki kecenderungan overlap. Pendekatan ini dinilai lebih sesuai untuk data finansial yang dinamis dibandingkan metode *hard clustering*, sejalan dengan penelitian Purba dan Hartanto [7] yang menegaskan keunggulan fleksibilitas *Fuzzy C-Means* dalam menangani data berambiguitas tinggi.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan metode *Fuzzy C-Means* untuk klasterisasi saham perbankan dengan raihan FPC score sebesar 0,5170 (kualitas sedang). Berdasarkan hasil akhir model,

Klaster 3 (BBCA, BBRI, MEGA, BBNI, BNGA) terkelompokkan ke dalam label *High Risk – Low Return* yang merepresentasikan barisan emiten dengan penguasaan pasar terbesar. Klaster 2 (BBTN, BDMN) menunjukkan profil *Medium Risk – Medium Return* yang mencerminkan kinerja profitabilitas dan eksposur risiko di tingkat menengah. Sementara itu, Klaster 1 (BRIS, ARTO, BJBR) masuk ke dalam kategori *Low Risk – High Return* yang diisi oleh saham-saham dengan karakteristik pergerakan finansial yang lebih dinamis di bursa. Keberadaan visualisasi derajat keanggotaan mendemonstrasikan fleksibilitas *Fuzzy C-Means* dalam memetakan zona transisi dan menangani ketidakpastian data pasar modal secara lebih luwes dibandingkan metode *hard clustering*.

DAFTAR PUSTAKA

- [1] A. N. Hafidzah, H. S. Gulo, S. P. Yolanda, V. P. A. Saban, V. A. Sari, and N. Nanang, "Klasterisasi Bank di Indonesia Berdasarkan Kinerja Finansial Menggunakan Algoritma K-Means," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 5075–5080, Jul. 2025, doi: 10.31004/riggs.v4i2.1387.
- [2] K. A. Kinanti, H. Sukarno, and E. S. Utami, "Banking Clustering Study Based On Fuzzy C-Mean And Fuzzy Gustafson Kessel Volume: 2 Number: 1 Page: 1-6," *International Journal of Environmental, Sustainability, and Social Science*, vol. 2, no. 1, 2021, doi: 10.38142/ijesss.v1i3.58.
- [3] H. Basri and W. Wibowo, "Optimasi Portofolio Dan Analisis Risiko Saham IDX30 Menggunakan K-Means, Markowitz Dan VaR," *Jurnal Ilmiah Akuntansi dan Keuangan (JIAKu)*, vol. 4, no. 2, pp. 42–61, Sep. 2025, doi: 10.24034/jiaku.v4i2.7244.
- [4] N. Jumairi and B. Purnama, "Penerapan Data Mining Dalam Pengelompokan Saham Melalui Rasio Keuangan Pada Saham Papan Utama Bursa Efek Indonesia Menggunakan Metode K-Means Clustering," *Jurnal Mahasiswa Teknik Informatika*, vol. 9, no. 3, p. 4863, 2025, [Online]. Available: www.idx.co.id
- [5] S. J. A. Sumarauw, M. G. Maukar, M. Grace Maukar, U. Negeri, and M. Abstrak, "Fuzzy c-Means Clustering untuk Pengenalan Pola Studi kasus Data Saham Fuzzy c-Means Clustering for Pattern Recognition on Case Studies in Holdings Data," *Jurnal Axioma : Jurnal Matematika dan Pembelajaran*, vol. 7, no. 2, 2022.
- [6] A. Catur Pamungkas, D. Wakhid Nur Sakti, F. Trisna Aqilla, K. Miza, and A. Arum Sari, "Clustering Produk UMKM Berdasarkan Penjualan Online Menggunakan K-Means," *Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, p. 2025, 2025.
- [7] W. N. Purba and R. Hartanto, "Perbandingan Penerapan Algoritma K-Means Dan Fuzzy C-Means Dalam Analisis Clustering Terhadap Pergerakan Harga Historis Saham Bank Rakyat Indonesia," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 7, no. 2, p. 865, Dec. 2024, doi: 10.37600/tekinkom.v7i2.1214.
- [8] F. A. Abdul Aziz, D. Hartanti, and F. Suryani, "Prototype Sistem Rekomendasi Wisata Kabupaten Wonogiri Menggunakan Knowledge-Based Recommendation," *Infotech: Journal of Technology Information*, vol. 10, no. 1, pp. 37–44, Jun. 2024, doi: 10.37365/jti.v10i1.245.
- [9] E. F. Himah and R. Sulaiman, "Tahun 2021 Implementasi Metode Fuzzy C-Means Dan Topsis Dalam Evaluasi Kinerja Keuangan Perusahaan Perbankan Di Indonesia Berdasarkan Rasio Keuangan," *MATHunesa: Jurnal Ilmiah Matematika*, vol. 9, no. 1, 2021.
- [10] W. Aliansa and R. A. Sapuutra, "Perbandingan Harga Saham Perbankan Dengan Metode Fuzzy C-Means Clustering," vol. 1, no. 1, pp. 1–5, 2023.
- [11] R. Gunawan and G. Owen, "Clustering and Portfolio Optimization on LQ45 Stocks with Fuzzy C-Means and Single Index Model Klasterisasi dan Optimalisasi Portofolio pada Saham LQ45 dengan Fuzzy C-Means dan Single Index Model," vol. 21, no. 3, pp. 655–668, 2025.
- [12] D. Tohendry and D. Jollyta, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Saham Berdasarkan Price Earning Ratio Dan Price To Book Value," *Jurnal Mahasiswa Aplikasi Teknologi Komputer dan Informasi*, vol. 5, no. 1, 2023.
- [13] S. Ufia, A. Dwi Nugroho, and T. Wahjoedi Sekolah Tinggi Ilmu Ekonomi Mahardhika Surabaya, "Journal Of Knowledge And Collaboration Meningkatkan Kompetensi Mahasiswa melalui Program Magang Sebagai Upaya Peningkatan Hard Skill dan Soft Skill," vol. 7, no. 1, 2025, [Online]. Available: <https://ojs.arbain.co.id/index.php/jkc/index>
- [14] I. Sekar, D. Hasnida, and R. Kusumawati, "Penerapan Model-Based Clustering ... / Penerapan Model-Based Clustering Pada Pengelompokan Saham Berdasarkan Rasio Keuangan," vol. 9, no. 3, 2025, [Online]. Available: www.yahooofinance.com