

SISTEM REKOMENDASI PORTOFOLIO BERDASARKAN PROFIL RISIKO DENGAN K-MEANS DAN MEAN VARIANCE OPTIMIZATION

Gusti Bagus Genta Purusa Arimbawa¹, Putu Indah Ciptayani², Ni Nyoman Harini Puspita³

Program Sarjana Terapan Teknologi Rekayasa Perangkat Lunak, Jurusan Teknologi Informasi

Politeknik Negeri Bali^{1,2,3}

Jalan Kampus Bukit Jimbaran, Kuta Selatan, Kabupaten Badung

gentaarimbawa@gmail.com¹, putuindah@pnb.ac.id², Mangari@pnb.ac.id³

Abstrak

Meningkatnya partisipasi investor ritel di Indonesia belum diimbangi oleh kualitas pengambilan keputusan investasi yang memadai. Kondisi ini tercermin dari perilaku fear of missing out (FOMO) dan ketidaksesuaian profil risiko, terutama pada investor pemula. Penelitian ini mengembangkan sistem rekomendasi portofolio saham terpersonalisasi dengan mengintegrasikan K-Means Clustering dan Mean-Variance Optimization (MVO) berbasis profil risiko investor. Analisis dilakukan terhadap 39 saham LQ45 yang likuid menggunakan data historis selama empat tahun. Hasil klasterisasi menunjukkan nilai optimal K=4 dengan silhouette score sebesar 0,274 yang menghasilkan empat kelompok saham, yaitu Low Performers, Value & High Dividend, High ROE, dan Aggressive Growth. MVO menghasilkan tiga portofolio dengan Sharpe ratio antara 1,19 hingga 1,21. Portofolio agresif memberikan expected return sebesar 24,79% dengan volatilitas 20,53%. Sistem diimplementasikan dalam aplikasi web yang mengintegrasikan kuesioner profil risiko berbasis standar BCA dan konversi bobot portofolio ke satuan lot Bursa Efek Indonesia. Berbeda dengan pendekatan berbasis Fuzzy C-Means yang menggunakan satu saham representatif dari setiap kluster sebagai input MVO, penelitian ini mengoptimalkan seluruh 39 saham dan mengintegrasikan kalibrasi profil risiko investor sehingga menghasilkan rekomendasi portofolio yang dapat langsung dieksekusi.

Kata Kunci: Sistem Rekomendasi, Portofolio Saham, Profil Risiko, K-Means Clustering, Mean-Variance Optimization

Abstract

The increasing participation of retail investors in Indonesia has not been accompanied by adequate investment decision-making quality. This issue is reflected in fear of missing out behavior and mismatches between investment choices and risk profiles, particularly among beginner investors. This study develops a personalized stock portfolio recommendation system by integrating K-Means Clustering and Mean-Variance Optimization based on investor risk profiles. The analysis covers 39 liquid LQ45 stocks using four years of historical data. The clustering process identified an optimal K value of 4 with a silhouette score of 0.274, resulting in four stock groups: Low Performers, Value & High Dividend, High ROE, and Aggressive Growth. MVO generated three portfolios with Sharpe ratios between 1.19 and 1.21. The aggressive portfolio achieved an expected return of 24.79% with a volatility of 20.53%. The system was implemented as a web application integrating a BCA-standard risk profiling questionnaire and portfolio weight conversion into exchange-compliant lot units. Unlike a previous Fuzzy C-Means-based approach that used one representative stock from each cluster as MVO input, this study optimizes all 39 stocks and incorporates investor risk profile calibration, producing directly executable portfolio recommendations.

Keywords: Recommendation System, Stock Portfolio, Risk Profile, K-Means Clustering, Mean-Variance Optimization

PENDAHULUAN

Pasar modal Indonesia mencatat pertumbuhan signifikan partisipasi investor ritel dalam dua tahun terakhir. Berdasarkan data KSEI, total *Single Investor Identification (SID)* mencapai 19 juta pada akhir Oktober 2025 (*PT Kustodian Sentral Efek Indonesia - Data Statistik KSEI, 2025*). Fenomena ini membawa tantangan struktural. Kompleksitas pemilihan saham di Bursa Efek Indonesia (BEI) meningkat seiring 950+ emiten yang tersedia (*Daftar Saham, 2025*). Hal ini melampaui kapasitas analitis investor pemula. Akibatnya, keputusan investasi sering didorong tekanan sosial. Informasi pasar yang tidak terverifikasi juga mempengaruhi pilihan. Analisis fundamental menjadi terabaikan. Penelitian perilaku keuangan mengidentifikasi dua pola disfungsional dominan. Pertama, Fear of

Missing Out (FOMO) mendorong pembelian impulsif saat harga tinggi. Kedua, ketidaksesuaian antara toleransi risiko individu dengan volatilitas saham yang dipilih (Binu, 2024; Purwaningrum et al., 2025).

Sejumlah penelitian telah mencoba menjawab tantangan ini melalui berbagai pendekatan prediktif. Nabipour et al. membuktikan keunggulan deep learning seperti LSTM. Model ini lebih akurat dalam prediksi tren harga dibanding algoritma konvensional. (Nabipour et al., 2020). Singh dan Jenq membangun sistem rekomendasi berbasis ARIMA dan indikator teknikal untuk menghasilkan sinyal beli atau jual otomatis (Singh & Jenq, 2025). Aprihartha dan Prihandono mengintegrasikan K-Means Clustering dengan MVEPOC untuk pembentukan portofolio. Kombinasi ini terbukti menghasilkan Sharpe Ratio lebih tinggi dibanding alokasi bobot rata. (Aprihartha & Prihandono, 2025). Penelitian tersebut tidak mengakomodasi profil risiko investor sebagai variabel input, output berupa persentase bobot teoretis yang tidak dapat langsung dieksekusi dalam satuan lot riil, dan tidak tersedia antarmuka sistem bagi pengguna. Dari keterbatasan ini menjadi celah yang diisi oleh penelitian ini.

Penelitian ini membangun sistem rekomendasi portofolio saham terpersonalisasi berbasis profil risiko investor. Sistem ini mengintegrasikan K-Means Clustering dan Mean-Variance Optimization (MVO) dalam antarmuka web. Penelitian ini memiliki empat kontribusi utama yang belum ditemukan secara terintegrasi. Pertama, integrasi K-Means dan MVO yang dikalibrasi terhadap profil risiko investor. Kedua, interpretabilitas kluster berbasis karakteristik fundamental saham. Ketiga, konversi bobot portofolio ke satuan lot riil. Keempat, sistem web yang dapat dioperasikan investor pemula secara mandiri.

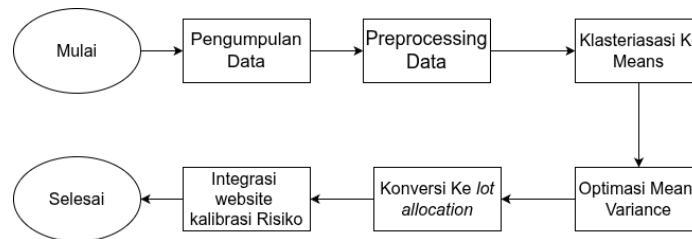
PENELITIAN RELEVAN

Beberapa penelitian sebelumnya telah mengintegrasikan klusterisasi dengan optimasi portofolio. Aprihartha dan Prihandono menggabungkan K-Means untuk seleksi saham berdasarkan dua kriteria: kapitalisasi pasar dan jumlah saham beredar. Selanjutnya, mereka menentukan bobot menggunakan metode Mean-Variance Efficient Portfolio One Constrain (MVEPOC). Hasilnya menunjukkan portofolio optimal menghasilkan ekspektasi imbal hasil sebesar 0,3589% dengan tingkat risiko sebesar 1,549% (contoh saham: BYAN, GOTO, GGRM). (Aprihartha & Prihandono, 2025) Nasution dan Al-khowarizmi membuktikan keunggulan integrasi K-Means dan model Markowitz pada saham LQ45. Portofolio yang dihasilkan lebih resilien, dengan penurunan nilai hanya 8,02% dibanding indeks LQ45 yang turun 15,60% (Akbar Nasution, 2025). Hal ini didukung oleh penelitian Ghifari et al. yang menunjukkan keunggulan pengelompokan saham dengan K-Means sebelum optimasi Markowitz. Pendekatan ini menghasilkan return 0,00162, lebih optimal dibanding pemilihan saham acak atau bobot setara. (Ghifari et al., 2024). Gubu et al. membandingkan Fuzzy C-Means dan K-Means dengan MVO pada 45 saham LQ45. Hasil menunjukkan Fuzzy C-Means konsisten menghasilkan Sharpe ratio portofolio lebih tinggi untuk semua nilai risk aversion. Nilai maksimum Fuzzy C-Means mencapai 0,260, sementara K-Means 0,226. Namun, penelitian tersebut menggunakan pendekatan representasi kluster yang membatasi analisis pada satu saham terbaik per kluster sebagai input MVO. Penelitian itu memiliki tiga keterbatasan: pertama, tidak mempertimbangkan profil risiko investor sebagai variabel kalibrasi; kedua, tidak menghasilkan rekomendasi yang dapat langsung dieksekusi dalam satuan lot riil; ketiga, tidak tersedia antarmuka sistem bagi pengguna (Gubu et al., 2024). Pengembangan dalam penelitian ini berfokus pada perbaikan keterbatasan studi sebelumnya dengan mengintegrasikan variabel profil risiko investor secara komprehensif ke dalam sistem rekomendasi untuk menghasilkan alokasi portofolio yang lebih personal.

METODE PENELITIAN

Tahapan Penelitian

Penelitian ini dirancang dalam empat tahap utama yang saling berurutan, mulai dari pengumpulan data hingga pembentukan portofolio optimal berbasis profil risiko investor. Keseluruhan alur kerja metodologi diilustrasikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Penelitian ini dilaksanakan melalui enam tahapan utama yang saling terintegrasi sebagaimana disajikan pada Gambar 1. Setiap tahap dirancang untuk memastikan data berkualitas dan pengelompokan saham objektif. Optimasi portofolio harus menghasilkan solusi optimal dan dapat diakses mudah oleh investor. Berikut adalah tahapan penelitian yang dilakukan:

1. Pengumpulan Data

Data sekunder diperoleh dari Yahoo Finance, mencakup 45 emiten indeks LQ45.. Pemilihan indeks LQ45 didasarkan pada likuiditas tinggi dan representativitas pasar modal Indonesia. Data historis mencakup periode 4 tahun (2022-2026). Analisis portofolio dilakukan untuk periode proyeksi November 2025 - Januari 2026. Variabel utama yang diambil adalah harga penutupan harian yang diekstraksi menjadi *annualized return* dan *volatility*. Fitur fundamental berupa *Price to Book Ratio (PB Ratio)*, *Return on Equity (ROE)*, dan *Dividend Yield*. BI Rate tahun 2025 sebesar 4,75% ditetapkan sebagai *risk-free rate* dalam evaluasi portofolio (Gubu et al., 2024).

2. Preprocessing Data

Preprocessing bertujuan membersihkan data, mengekstrak fitur relevan, dan mentransformasi data untuk klasterisasi. Pertama, data cleaning dilakukan dengan mengisi missing values menggunakan metode forward fill. Emiten dengan hari transaksi kurang dari 95% di eliminasi. Fitur *annualized return* dan *volatility* diekstraksi dari harga penutupan harian yang dengan faktor 252. Fitur fundamental (*PB Ratio*, *ROE*, dan *Dividend Yield*) diperoleh langsung dari Yahoo Finance. Outliers ditangani dengan winsorization pada persentil ke-5 dan ke-95. (Ardhani et al., 2025) Distribusi data kemudian dinormalkan menggunakan transformasi Yeo-Johnson. Yeo-Johnson dipilih karena mampu menangani distribusi data bernilai positif maupun negatif tanpa batasan data harus bernilai positif (Riani et al., 2023). Seluruh fitur distandardisasi dengan Z-score sehingga memiliki rata-rata 0 dan deviasi standar 1. Standardisasi mencegah dominasi variabel berskala besar dalam perhitungan jarak KMeans.

3. Klasterisasi K-Means

Algoritma *K-Means Clustering* diterapkan untuk mengelompokkan ke-39 saham berdasarkan kemiripan karakteristik risiko, *return* dan fundamental. Jumlah kluster optimal ditentukan dengan Silhouette Coefficient, menguji k dari 2 hingga 10. Dipilih k yang menghasilkan skor tertinggi dan struktur kluster yang kuat (Yuliasih et al., 2024). Setiap saham dialokasikan ke kluster terdekat berdasarkan jarak Euclidean terkecil.

4. Optimasi Mean-Variance (MVO)

Portofolio optimal dibentuk pada setiap kluster menggunakan pendekatan *Mean-Variance Optimization* yang dikembangkan oleh Markowitz. Proses ini meminimalkan variansi portofolio untuk tingkat *return* yang diharapkan tertentu. Kendala total bobot bernilai satu dan tidak diperbolehkan adanya *short selling*. Ekspektasi *return* dihitung sebagai rata-rata aritmetika *return* historis, sedangkan risiko portofolio dihitung dari matriks kovarians. Kualitas portofolio dievaluasi dengan *Sharpe Ratio* menggunakan *risk-free rate* 4,75%. Nilai *Sharpe Ratio* yang lebih tinggi menunjukkan efisiensi portofolio yang lebih baik (Devi et al., 2023).

5. Konversi ke Lot Allocation

Bobot alokasi dalam persentase dikonversi ke jumlah lot riil. Sesuai ketentuan pasar modal Indonesia, di mana 1 lot setara dengan 100 lembar saham. Konversi ini didasarkan pada modal yang diinputkan oleh investor sehingga rekomendasi portofolio dapat langsung dieksekusi (Pratiwi et al., 2025).

6. Integrasi Website Kalibrasi Risiko

Seluruh tahapan preprocessing, klasterisasi, dan optimasi portofolio diintegrasikan ke dalam aplikasi web interaktif berbasis Streamlit. Pustaka Python seperti Pandas, NumPy, yFinance, Scikit-learn, dan PyPortfolioOpt digunakan dalam pengembangan. Kuesioner tujuh pertanyaan mengklasifikasikan investor ke tiga profil: Konservatif, Moderat, atau Agresif. Profil Konservatif memprioritaskan keamanan modal dan meminimalkan fluktuasi. Profil Moderat menyeimbangkan potensi imbal hasil dengan perlindungan modal. Profil Agresif berfokus pada imbal hasil maksimal dengan memberi ruang lebih besar pada aset bertumbuh tinggi meskipun volatilitasnya meningkat (Rois et al., 2022).

Hasil akhir berupa rekomendasi portofolio terkalibrasi dan konversi lot ditampilkan secara visual melalui Plotly dan Matplotlib, siap digunakan oleh investor.

HASIL DAN PEMBAHASAN

Data Preprocessing

Proses *data cleaning* diawali dengan penanganan *missing values* menggunakan metode *forward fill* pada 45 emiten LQ45. Selanjutnya, emiten dengan jumlah hari transaksi kurang dari 95% total hari perdagangan maksimum dieliminasi untuk memastikan likuiditas data. Hasil penyaringan mengeliminasi enam saham, yaitu AADI, AMMN, GOTO, MBMA, NCKL, dan PGEO, sehingga tersisa 39 saham yang digunakan dalam analisis selanjutnya. Total data yang dihasilkan sebanyak 38.844 baris dengan rentang periode 3 Januari 2022 hingga 27 Februari 2026, di mana setiap saham memiliki rentang hari transaksi beragam sebanyak 996 hari.

Setelah pembersihan, dilakukan *winsorization* pada persentil ke-5 dan ke-95 untuk menangani *outliers* dapat dilihat pada Persamaan (1).

$$x_{winsorized} = \begin{cases} P_5 & x < P_5 \\ P_{95} & \text{jika } x > P_{95} \\ \text{lainnya} & \text{jika} \end{cases} \quad (1)$$

Nilai *winsorized* ditetapkan menjadi P_5 jika data lebih kecil dari persentil ke-5, menjadi P_{95} jika data lebih besar dari persentil ke-95, dan dibiarkan tidak berubah jika berada di antara keduanya (Ardhani et al., 2025).

Tabel 1 menunjukkan perbandingan nilai maksimum dan minimum sebelum dan sesudah *winsorization*.

Tabel 1. Perbandingan Nilai Ekstrem Sebelum dan Sesudah Winsorization

Fitur	Max Sebelum	Max Sesudah	Min Sebelum	Min Sesudah
Annualized Return	0.9597	0.9445	-0.2249	-0.2010
Annualized Volatility	0.6622	0.6612	0.2329	0.2346
PB Ratio	32.9065	6.8011	0.4997	0.7536
ROE	1.5092	0.3891	-0.0257	0.0030
Dividend Yield	0.4750	0.2803	0.0011	0.0046

Fitur PB_Ratio mengalami penurunan nilai maksimum paling signifikan, dari 32,91 menjadi 6,80, yang mengindikasikan adanya *outliers* ekstrem yang berhasil dibatasi tanpa menghapus data.

Transformasi Yeo-Johnson diterapkan untuk memperbaiki distribusi data agar mendekati normal, diikuti standarisasi Z-score untuk menyamakan skala seluruh fitur. Keberhasilan proses ini dikonfirmasi oleh nilai rata-rata setiap fitur yang mendekati nol dan standar deviasi yang mendekati satu, sebagaimana ditunjukkan pada Tabel 2. Setelah standarisasi, rentang nilai seluruh fitur berada dalam interval -2,22 hingga 2,34, yang menunjukkan bahwa data telah bebas dari *outliers* ekstrem dan siap digunakan dalam klasterisasi.

Tabel 2. Statistik Deskriptif

Statistik	Annualized Return	Annualized Volatility	PB Ratio	ROE	Dividend Yield
Count	39	39	39	39	39
Mean	0.124976	0.397167	2.037941	0.158248	0.058723
Std	0.242567	0.110996	1.401928	0.086675	0.069586
Min	-0.200965	0.234625	0.753555	0.003049	0.004572
25%	-0.000469	0,311980	1.309260	0.102733	0.019614
50%	0.097654	0.386376	1.676569	0.148626	0.040526
75%	0.199042	0,447993	2.042284	0.197092	0.060295
Max	0.944455	0.661166	6.801094	0.389082	0.280253

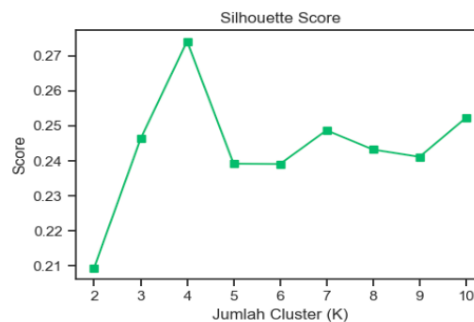
Klasterisasi K-Means

Penentuan jumlah kluster optimal dilakukan dengan menguji nilai k dari 2 hingga 10 menggunakan metrik evaluasi, yaitu *Silhouette Coefficient*. *Silhouette Coefficient* mengukur kohesi (cohesion, $a(i)$) dan separasi (separation, $b(i)$) antar kluster dengan rentang -1 hingga 1. Untuk perhitungan *Silhouette Coefficient* dapat di lihat Persamaan (2) :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

di mana nilai *silhouette score* lebih tinggi menandakan struktur kluster yang lebih baik (Yuliasih et al., 2024).

Pada Gambar 2 menunjukkan bagaimana nilai *silhouette score* dari kluster 2 hingga 10.



Gambar 2. Grafik Silhouette Score

Gambar 2 menampilkan silhouette score untuk $k=2$ hingga 10. Nilai $k=4$ menghasilkan silhouette score tertinggi, yaitu 0,274. Skor ini termasuk kategori struktur lemah (0,26–0,50). Nilai tersebut tertinggi di antara semua k yang diuji. Jadi, $k=4$ adalah konfigurasi paling optimal. Skor 0,274 menunjukkan separasi kluster relatif lemah. Hal ini wajar terjadi pada data finansial. Overlap perilaku pasar dan korelasi saham menjadi penyebabnya. Dinamika volatilitas saham LQ45 pun saling berdekatan. Oleh karena itu, $k=4$ tetap dipilih karena menghasilkan nilai tertinggi. Segmentasi sahamnya mudah diinterpretasikan. Ini cocok untuk rekomendasi portofolio. Empat kluster cukup mewakili karakteristik risiko-return. Pendekatan ini tidak menghasilkan segmen terlalu banyak (Ghifari et al., 2024).

Sebagai perbandingan, penelitian Gubu et al. menggunakan Fuzzy C-Means pada data LQ45 dan menghasilkan 6 kluster tanpa melaporkan silhouette score. K-Means dalam studi yang sama juga menghasilkan 6 kluster. Perbedaan mendasarnya terletak pada mekanisme keanggotaan. Fuzzy C-Means memungkinkan satu saham memiliki derajat keanggotaan parsial di beberapa kluster. Sementara K-Means mengalokasikan setiap saham secara eksklusif ke satu kluster. (Gubu et al., 2024). Untuk konteks sistem rekomendasi berbasis profil risiko investor pemula, alokasi eksklusif K-Means dipilih karena menghasilkan segmentasi yang lebih interpretatif. K-Means dapat dipetakan langsung ke kategori profil risiko tanpa ambiguitas keanggotaan. Setelah jumlah kluster optimal ditetapkan sebanyak empat. Algoritma K-Means mengalokasikan 39 saham ke dalam

klaster-klaster tersebut berdasarkan kemiripan lima fitur yang telah distandardisasi. Proses dimulai dengan menentukan k titik pusat klaster (*centroid*) secara acak, kemudian setiap titik data dialokasikan ke klaster dengan centroid terdekat berdasarkan jarak *Euclidean* melalui persamaan (3) (Novika & Rahayu, 2024) :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

$d(x, y)$ adalah jarak Euclidean antara dua titik data, x_i dan y_i adalah nilai fitur ke- i dari masing-masing titik, dan n adalah jumlah fitur yang digunakan semakin kecil nilai d , semakin mirip kedua titik tersebut sehingga akan dialokasikan ke klaster yang sama (Novika & Rahayu, 2024).

Posisi centroid diperbarui secara iteratif menggunakan nilai rata-rata seluruh data pada masing-masing cluster hingga proses mencapai kondisi konvergen, di mana tidak ada lagi perubahan signifikan dalam alokasi anggota cluster atau posisi centroid antar iterasi. Tabel 3 menyajikan hasil klusterisasi menggunakan k optimal ($k = 4$).

Tabel 3. Hasil Klusterisasi

Klaster	N	Annual return	Annual Volatility	PB Ratio	ROE	Devidend Yield	Daftar Saham	Label
0	10	-0.0366	0.4404	1.3132	0.0663	0.0300	ACES, BBTN, CTRA, EMTK, EXCL, INCO, INKP, MDKA, SCMA, SMGR	Low Performers
1	12	0.1644	0.3228	1.3558	0.1925	0.1443	ADRO, AKRA, ASII, BBNI, BBRI, BMRI, INDF, ITMG, JPFA, PGAS, PTBA, UNTR	Value & High Dividend
2	9	-0.0234	0.3252	6.8899	0.3132	0.0257	AMRT, BBKA, CPIN, HEAL, ICBP, KLBF, TLKM, TOWR, UNVR	High ROE
3	8	0.4336	0.5356	1.7401	0.1839	0.0275	ADMR, ANTM, BRPT, BUMI, DSSA, ISAT, MAPI, MEDC	Aggressive Growth

Keempat klaster memiliki pergerakan yang berbeda dapat di lihat pada Gambar 3.

Perbandingan Karakteristik Antar Klaster (Radar Chart)



Gambar 3. Radar Chart

Radar chart menunjukkan annualized return dan volatility sebagai dimensi pembeda utama antar cluster. Dividend yield dan ROE memperkuat interpretasi kelompok saham defensif dan growth-oriented. Analisis mengidentifikasi empat klaster dengan diferensiasi signifikan: Klaster 3 (Aggressive Growth, 8 saham) menunjukkan return tertinggi (0,4336) dan volatilitas tertinggi (0,5356); Klaster 2 (High ROE, 9 saham) memiliki ROE tertinggi (0,3132) namun return negatif (-0,0234), mencerminkan overvaluation; Klaster 1 (Value & High Dividend, 12 saham) menawarkan dividend yield tertinggi (0,1443), return positif (0,1644), dan volatilitas terendah (0,3228); Klaster 0 (Low Performers, 10 saham) menunjukkan return negatif (-0,0366), ROE terendah (0,0663), dan volatilitas tinggi (0,4404). Segmentasi klaster ini memfasilitasi konstruksi portofolio terdiferensiasi sesuai profil risiko investor, memungkinkan pemilihan kombinasi klaster yang selaras dengan preferensi risiko individual(Akbar Nasution, 2025).

Optimasi Mean-Variance

Optimasi portofolio dilakukan pada seluruh 39 saham menggunakan kerangka *Mean-Variance Optimization (MVO)*. MVO bekerja melalui beberapa tahap perhitungan. Pertama, ekspektasi return setiap saham dihitung sebagai rata-rata aritmetika return historis sesuai persamaan (4), di mana $E(R_i)$ adalah ekspektasi return saham i , $r_{i,t}$ adalah return saham i pada periode t , dan T adalah total periode pengamatan (González-Bueno et al., 2025).

$$E(R_i) = \frac{1}{T} \sum_{t=1}^T r_{i,t} \quad (4)$$

Variansi masing-masing saham dihitung melalui persamaan (5), di mana σ_i^2 adalah variansi return saham i dan $R_{i,t}$ adalah return saham i pada periode :

$$\sigma_i^2 = \frac{1}{T-1} \sum (R_{i,t} - E(R_i))^2 \quad (5)$$

Kovarians antar pasangan saham dihitung untuk mengukur dependensi linier, di mana σ_{ij} adalah kovarians antara saham i dan j . Nilai kovarians ini kemudian digunakan untuk menghitung risiko portofolio melalui Persamaan (6), di mana σ_p^2 adalah variansi portofolio, w_i dan w_j adalah bobot saham i dan j :

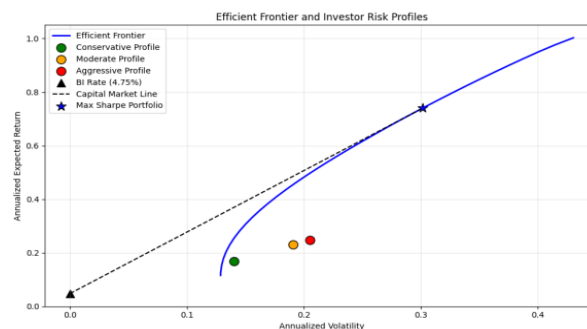
$$\sigma_p^2 = \sum \sum w_i w_j \sigma_{ij} \quad (6)$$

Selanjutnya, bobot optimal ditentukan dengan meminimalkan Persamaan (6) untuk memperoleh portofolio efisien, di mana w adalah vektor bobot dan Σ adalah matriks kovarians, dengan kendala seluruh bobot berjumlah satu dan tidak ada posisi short ($w_i \geq 0$) sesuai Persamaan (7) (Devi et al., 2023):

$$\min_w w^T \Sigma w \text{ subject to } \begin{cases} \sum w_i = 1 \\ w_i \geq 0 \ (i = 1, 2, \dots, n) \end{cases} \quad (7)$$

Batasan target *return* tertentu sesuai karakteristik profil (konservatif: target return rendah; agresif: target return tinggi)(Rois et al., 2022).

Hasil optimasi menggunakan *Mean-Variance Optimization (MVO)* berdasarkan Persamaan (4) samapai (7) menghasilkan kurva *efficient frontier* seperti disajikan pada Gambar 4.



Gambar 4. Kurva Efficient Frontier

Penerapan *framework Mean-Variance Optimization (MVO)* menghasilkan kurva *efficient frontier* melalui minimisasi fungsi objektif dengan kendala non-negatif bobot dan total bobot sama dengan satu. Estimasi ekspektasi return dan variansi individual saham diperoleh dari data historis (Persamaan 4-5), sedangkan matriks kovarians dibangun untuk mengukur dependensi linier antar aset (Persamaan 6). Substitusi bobot optimal ke dalam fungsi objektif menghasilkan tiga profil investasi yang merepresentasikan segmen berbeda pada *efficient frontier*. Dari kurva tersebut dipilih tiga titik yang merepresentasikan profil risiko berbeda seperti konservatif, moderat. Dan agresif. Kinerja setiap profil disajikan pada Tabel 4.

Tabel 4. Hasil Optimasi Sesuai Profil Risiko

Profil	Expected Return (%)	Volatility (%)	Sharpe Ratio	N
Konservatif	16.70	14,01	1.19	28
Moderat	23.09	19.07	1.21	21
Agresif	24.79	20.53	1.21	18

Tabel 4 menampilkan hasil optimasi dengan tiga profil. Profil Konservatif mencakup 28 saham dengan expected return 16,70% dan volatilitas 14,01% (Sharpe ratio 1,19). Moderat mencakup 21 saham dengan return 23,09% dan volatilitas 19,07% (Sharpe ratio 1,21). Agresif mencakup 18 saham dengan return 24,79% dan volatilitas 20,53% (Sharpe ratio 1,21). Nilai Sharpe ratio yang sama pada profil Moderat dan Agresif mengindikasikan bahwa peningkatan return dari profil Moderat ke Agresif tidak diikuti peningkatan risk-adjusted return secara proporsional. Ini mencerminkan trade-off risiko-return yang karakteristik pada segmen upper efficient frontier.

Tabel 5. Komposisi Pembobotan Klaster Saham Setiap Profil

Profil	Klaster 0 (%)	Klaster 1 (%)	Klaster 2 (%)	Klaster 3 (%)	N
Konservatif	1,40	48,06	26,86	23,69	28
Moderat	0,00	47,47	4,94	47,59	21
Agresif	0,00	43,55	47,59	53,31	18

Tabel 5 menunjukkan bahwa Klaster 1 (Value & High Dividend) menjadi fondasi di semua profil (43–48%). Seiring peningkatan selera risiko, alokasi bergeser dari Klaster 2 (High ROE) yang dominan di konservatif (26,86%) menuju Klaster 3 (Aggressive Growth) yang menguasai profil moderat dan agresif (47–53%). Klaster 0 (Low Performers) hanya muncul dengan bobot kecil (1,40%) di profil konservatif sebagai penyeimbang diversifikasi, dan hilang di profil lainnya. Pola ini menegaskan bahwa sistem rekomendasi mengalokasikan aset secara adaptif dari saham defensif ke agresif sesuai preferensi risiko.

Integrasi Website Kalibrasi Risiko

Profil risiko investor dikelompokkan menjadi tiga kategori Konservatif, Moderat, dan Agresif berdasarkan hasil kuesioner pada Gambar 5.

Parameter 2: Jelaskan pendapat Anda tentang berinvestasi.

- ☒ Untuk menyimpan modal, saya lebih menyukai bunga yang rendah namun modal saya aman.
- ☐ Saya berharap investasi saya dapat menghasilkan profit sesuai dengan keinginan dan modal saya relative stabil.
- ☐ Saya mengerti bahwa nilai investasi saya dapat berfluktuasi. Saya siap menghadapi fluktuasi atas modal saya, demi mendapatkan profit yang lebih besar dalam jangka panjang
- ☐ Saya siap menerima risiko lebih besar demi mendapatkan profit yang lebih besar dalam jangka panjang.

Parameter 3: Untuk mencapai tujuan jangka panjang anda, apa pendapat anda jika dalam beberapa tahun pertama terjadi penurunan nilai modal yang Anda investasikan?

- ☒ Saya tidak menginginkan adanya penurunan terhadap nilai investasi saya.
- ☐ Jika penghasilan yang saya terima tidak berubah, saya tidak terlalu peduli dengan menurunnya modal jangka pendek.
- ☐ Pada dasarnya saya berinvestasi untuk jangka panjang, namun saya cukup peduli dengan penurunan ini.
- ☐ Saya berinvestasi untuk jangka panjang dan dapat memahami adanya fluktuasi tersebut sebagai bagian dari keinginan saya untuk memperoleh keuntungan lebih besar dengan

Gambar 5. Kuesioner Kalibrasi Risiko Pada Website

Kuesioner profil risiko yang diadaptasi dari BCA Life (*Satu-Satunya Asuransi Jiwa Anak Usaha Grup BCA*, n.d.) Gambar 5 menampilkan tujuh pertanyaan untuk menentukan profil risiko investor. Pertanyaan mencakup tujuan investasi, toleransi risiko, respons terhadap penurunan nilai, target imbal hasil, kepedulian terhadap inflasi, pemahaman pasar modal, dan preferensi investasi. Setiap jawaban diberi skor. Total skor digunakan untuk mengklasifikasikan investor sebagai konservatif, moderat, atau agresif. Hasil klasifikasi menjadi dasar rekomendasi portofolio yang sesuai dengan profil risiko investor.

Gambar 6 menunjukkan hasil kalibrasi profil risiko kategori agresif.

☒ Sinkronisasi profil berhasil! AI telah memetakan DNA investasi Anda.

Investor ID
Genta

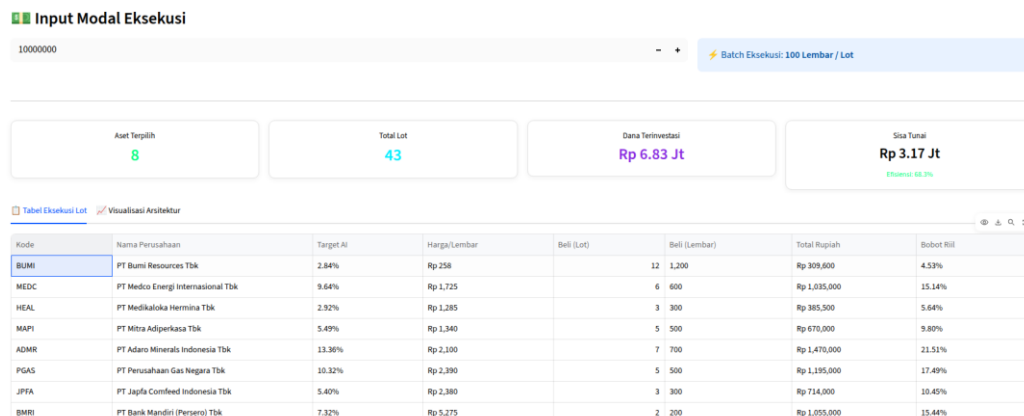
Skor Toleransi
31

Klasifikasi AI
AGRESIF

Gambar 6. Hasil Kalibrasi Kuesioner Profil Risiko

Konversi Lot Allocation

Gambar 7 menampilkan konversi bobot portofolio hasil MVO ke dalam jumlah lot saham (1 lot = 100 lembar) sesuai ketentuan BEI. Sistem membulatkan alokasi ke lot terdekat dengan batasan modal (Pratiwi et al., 2025).



Gambar 7. Konversi Lot Allocation

Hasil optimasi portofolio dikonversi menjadi jumlah lot saham. Konversi mengikuti ketentuan Bursa Efek Indonesia, yaitu 1 lot setara 100 lembar saham. Proses ini memastikan hasil optimasi dapat diterapkan secara praktis. Pada profil agresif dengan modal Rp10 juta, sistem mengalokasikan dana ke 8 saham. Total alokasi mencapai 43 lot saham. Dana yang terinvestasi mencapai 68,3% dari total modal. Sisa modal sebesar 31,7% tidak teralokasi sepenuhnya. Kondisi ini disebabkan oleh batas minimum lot dan harga saham yang diskrit.

SIMPULAN

Penelitian ini berhasil membangun sistem rekomendasi portofolio saham berbasis profil risiko investor. Sistem mengintegrasikan K-Means Clustering dan Mean-Variance Optimization (MVO). Klusterisasi terhadap 39 saham LQ45 menghasilkan empat kelompok saham. Kelompok tersebut meliputi Low Performers, Value & High Dividend, High ROE, dan Aggressive Growth. Nilai optimal klaster diperoleh pada $K=4$ dengan silhouette score 0,274. Optimasi MVO menghasilkan tiga portofolio sesuai profil risiko investor. Portofolio konservatif menghasilkan return 16,70% dengan volatilitas 14,01%. Portofolio moderat menghasilkan return 23,09% dengan volatilitas 19,07%. Portofolio agresif menghasilkan return 24,79% dengan volatilitas 20,53%. Seluruh portofolio memiliki Sharpe ratio di atas 1,19. Sistem juga mengonversi bobot portofolio menjadi satuan lot saham. Konversi mengikuti ketentuan Bursa Efek Indonesia. Pada profil agresif, modal Rp10 juta dialokasikan ke 8 saham. Total alokasi mencapai 43 lot dengan efisiensi 68,3%. Hasil penelitian menunjukkan bahwa sistem mampu menghasilkan rekomendasi yang sesuai dengan profil risiko investor. Sistem ini dapat membantu investor pemula mengambil keputusan investasi secara lebih terstruktur.

DAFTAR PUSTAKA

- Akbar Nasution, M. (2025). Data-Driven Portfolio Optimization using K-Means and Markowitz Model: Evidence from LQ45 Stocks. In *Economic: Journal Economic and Business* (Vol. 4, Number 2).
- Aprihartha, A., & Prihandono, A. (2025). Implementasi Algoritma K-Mean Clustering dan Mean-Variance Efficient Portfolio One Constrain dalam Optimasi Bobot Portofolio Saham. *JSI: Jurnal Sistem Informasi (E-Journal)*, 17(1), 203–213. <https://doi.org/10.18495/jsi.v17i1.236>
- Ardhani, N. T., Notodiputro, K. A., & Oktarina, S. D. (2025). Winsorization for Outliers in Clustering Non-Cyclical Stocks with K-Means and K-Medoids. *Indonesian Journal of Statistics and Its Applications*, 9(1), 46–60. <https://doi.org/10.29244/ijsa.v9i1p46-60>
- Binu, M. O. U. (2024). The Influence of Heuristic and Herding Behavior on Investment Decisions through Fomo on Retail Investors in Indonesia. *Jurnal Indonesia Sosial Teknologi*, 5(9), 3454–3470. <https://doi.org/10.59141/jist.v5i9.3313>
- Daftar Saham. (2025). <https://www.idx.co.id/>
- Devi, A. A. N., Dharmawan, K., & Tastrawati, N. K. T. (2023). PENGELOMPOKAN SAHAM MENGGUNAKAN K-MEANS DALAM PEMBENTUKAN PORTOFOLIO OPTIMAL. *E-Jurnal Matematika*, 12(4), 302–308. <https://doi.org/10.24843/MTK.2023.v12.i04.p433>
- Ghifari, A. A. G. Al, Toharodin, T., & Hendrawati, T. (2024). FORMATION OF AN OPTIMAL PORTFOLIO OF LQ45 SHARES USING MARKOWITZ METHOD. *JRAK*, 16(1), 85–94. <https://doi.org/10.23969/jrak.v16i1.8445>

- González-Bueno, J., Tamošiūnienė, R., Morales, C. G., & Rueda-Barrios, G. (2025). Applying the mean-variance framework: portfolio optimization and comparative performance analysis in the emerging Colombian capital market. *Business, Management and Economics Engineering*, 23(1), 164–188. <https://doi.org/10.3846/bmee.2025.22695>
- Gubu, L., Cahyono, E., Arman, A., Budiman, H., & Djafar, M. K. (2024). OPTIMASI PORTOFOLIO MEAN-VARIANCE DENGAN ANALISIS KLASSTER FUZZY C-MEANS. *Jurnal Gaussian*, 12(4), 593–604. <https://doi.org/10.14710/j.gauss.12.4.593-604>
- Nabipour, M., Nayyeri, P., Jabani, H., S., S., & Mosavi, A. (2020). Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis. *IEEE Access*, 8, 150199–150212. <https://doi.org/10.1109/ACCESS.2020.3015966>
- Novika, F., & Rahayu, S. (2024). Clustering Sukuk Using the K-Means Algorithm for Allocation of Investors Based on Investment Risk Profile. *E3S Web of Conferences*, 483, 03002. <https://doi.org/10.1051/e3sconf/202448303002>
- Pratiwi, A. S. D., Kurniawan, D. I., Adestiani, F., Aulia, L., & Abdillah. (2025). Analisis Investasi Online Pada Aplikasi Stockbit. *Journal Scientific of Mandalika (JSM) e-ISSN 2745-5955 | p-ISSN 2809-0543*, 6(4), 826–833. <https://doi.org/10.36312/10.36312/vol6iss4pp826-833>
- PT Kustodian Sentral Efek Indonesia - Data Statistik KSEI. (2025). https://web.ksei.co.id/publications/Data_Statistik_KSEI?setLocale=id-ID
- Purwaningrum, E., Ali, Ingriyani, L., Fatimah, F., Aminah, I., & Waersini, S. (2025). Keterkaitan Profil Risiko dengan Preferensi Investasi: Mengungkap Cara Pandang Investor melalui Simulasi Investasi Saham di Bursa Efek Indonesia. *Journal of Applied Community Engagement*, 5, 8–17. <https://doi.org/10.52158/jace.v5i1.980>
- Riani, M., Atkinson, A. C., & Corbellini, A. (2023). Automatic robust Box–Cox and extended Yeo–Johnson transformations in regression. *Statistical Methods & Applications*, 32(1), 75–102. <https://doi.org/10.1007/s10260-022-00640-7>
- Rois, M., Arbainah, S., Nikmatuniayah, N., Korawijayanti, L., & Rikawati, R. (2022). STUDI TENTANG JENIS INVESTASI BERDASARKAN PADA PROFIL RISIKO DI WILAYAH TEMBALANG SEMARANG. *Prosiding Seminar Hasil Penelitian Dan Pengabdian Masyarakat*, 4(1). <https://jurnal.polines.ac.id/index.php/Sentrikom/article/view/3335>
- Satu-Satunya Asuransi Jiwa Anak Usaha Grup BCA. (n.d.). <https://www.bcalife.co.id/Layanan/Formulir/Kuesioner-Hobi-Dan-Pekerjaan/Kuesioner-Profil-Risiko>
- Singh, J., & Jenq, J. (2025). A Stock Recommendation System. *Proceedings of 38th International Conference on Computers and Their Applications*, 155–164. <https://doi.org/10.29007/p27x>
- Yulisasih, B. N., Herman, H., Sunardi, S., & Yuliansyah, H. (2024). Evaluation of K-Means Clustering Using Silhouette Score Method on Customer Segmentation. *ILKOM Jurnal Ilmiah*, 16(3), 330–342. <https://doi.org/10.33096/ilkom.v16i3.2325.330-342>