

Implementasi Algoritma *K-Mean Clustering* dan *Mean-Variance Efficient Portfolio One Constrain* dalam Optimasi Bobot Portofolio Saham

Moch. Anjas Aprihartha¹, Adi Prihandono²

¹Program Studi S1 PJJ Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

²Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
e-mail: anjas.aprihartha@dsn.dinus.ac.id

Abstrak

Klasterisasi (*Clustering*) merupakan teknik yang digunakan dalam analisis data untuk mengelompokkan sekumpulan objek berdasarkan karakteristik tertentu, sedemikian hingga objek dalam satu kelompok lebih mirip satu sama lain dibandingkan dengan objek dalam kelompok yang lain. Metode *k-means* merupakan metode klasterisasi berbasis jarak yang populer dalam data mining. Salah satu contoh penerapan algoritma *k-means* adalah memilih saham pada klaster tertentu dalam proses membangun portofolio. Berinvestasi saham merupakan kegiatan yang memiliki risiko tinggi, sehingga penentuan portofolio optimal perlu dilakukan agar tujuan investasi dapat tercapai. Membangun bobot portofolio optimal dapat dilakukan dengan metode *mean-variance efficient portfolio one constrain*. Metode ini berguna dalam meminimalisir risiko pada tingkat profit yang diharapkan. Penelitian ini bertujuan untuk menciptakan portofolio optimal dengan menggabungkan metode *k-means clustering* dan metode *mean-variance efficient portfolio one constrain*. Hasil penelitian diperoleh dari 50 saham dibagi menjadi tiga klaster berbeda. Masing-masing klaster dengan ekspektasi *return* tertinggi dimasukkan ke dalam portofolio diantaranya adalah BYAN, GOTO, dan GGRM. Hasil analisis diperoleh bobot portofolio saham BYAN, GOTO, dan GGRM masing-masing sebesar 0,556693; 0,219429; dan 0,223879. Nilai ekspektasi *return* portofolio diperoleh sebesar 0,003589 yang berarti potensi rata-rata keuntungan yang diharapkan investor dari portofolio ini sebesar 0,3589%. Risiko yang diterima investor dalam portofolio tersebut relatif rendah dan stabil, yaitu sebesar 1,549%.

Abstract

Clustering is a technique used in data analysis to group objects based on certain characteristics, so that objects in one group are more similar to each other than objects in another group. The k-means method is a distance-based clustering method that is popular in data mining. One example of the application of the k-means algorithm is selecting stocks in a specific cluster in the process of building a portfolio. Investing in stocks is a high-risk activity, so determining the optimal portfolio needs to be carried out so that investment goals can be achieved. Building optimal portfolio weights can be carried out with the mean-variance efficient portfolio one constrain method. This method is useful in minimizing risk at the expected profit level. This study aims to create an optimal portfolio by combining the k-means clustering method and the mean-variance efficient portfolio one constrain method. The results of the study were obtained from 50 stocks divided into three different clusters. Each cluster with the highest expected return is included in the portfolio, including BYAN, GOTO, and GGRM. The results of the analysis obtained the weight of BYAN, GOTO, and GGRM stock portfolios are 0.556693; 0.219429; and 0.223879. The expected value of the portfolio return is obtained at 0.003589 which means that the average potential profit expected by investors from this portfolio is 0.3589%. The risk received by investors in the portfolio is relatively low and stable, which is 1.549%.

Keywords: clustering, k-means, mean-variance, portfolio, stocks

1. INTRODUCTION

Klasterisasi (*Clustering*) merupakan teknik yang digunakan dalam analisis data untuk mengelompokkan sekumpulan objek berdasarkan karakteristik tertentu, sedemikian hingga objek dalam satu kelompok lebih mirip satu sama lain dibandingkan dengan objek dalam kelompok yang lain [1]. Metode *unsupervised learning* ini sangat penting dalam mengidentifikasi pengelompokan intrinsik dalam sekumpulan data yang tidak berlabel, mengungkap pola, dan struktur yang tersembunyi dalam data [2]. Analisis klasterisasi memainkan peran penting sebagai alat eksplorasi yang telah banyak digunakan di berbagai bidang, termasuk analisis data, pemrosesan gambar, bioinformatika, pembentukan portofolio, dll. Tujuannya untuk mengatur kumpulan item data, yang disebut sebagai titik data, ke dalam klaster item yang serupa. Metode ini menilai kesamaan di antara titik data berdasarkan metrik jarak tertentu [3]. Terdapat tiga metodologi dalam klasterisasi yaitu klasterisasi berbasis jarak, klasterisasi berbasis kepadatan, dan hirarkis.

Metode *k-means* merupakan metode klasterisasi berbasis jarak yang dijalankan di bawah prinsip *unsupervised learning*. Metode ini beroperasi menggunakan jarak *euclidean* melalui proses iteratif untuk mengelompokkan titik data ke dalam klaster berdasarkan ukuran kesamaan yang ditentukan. Salah satu contoh penerapan algoritma *k-means* adalah memilih saham pada klaster tertentu dalam proses membangun portofolio. Beberapa penelitian terkait penggunaan algoritma *k-means* dalam pembentukan portofolio. Penelitian oleh Amanah *et al.* [4], menerapkan algoritma *k-means* dalam proses menciptakan portofolio model Black-Litterman. Amanah *et al.* [4] mengemukakan, pemilihan metode *k-means* dilakukan karena dianggap efisien untuk menangani data besar. Penelitian oleh Baiti & Prasetya [5] tentang penggunaan algoritma klasterisasi non-hirarkis dalam penentuan bobot portofolio. Dalam penelitiannya membandingkan metode *k-means* dan *k-medoids* untuk menentukan metode terbaik dalam pengelompokan. Hasil penelitian diperoleh algoritma *k-means* lebih efektif dibandingkan dengan *k-medoids* berdasarkan kriteria *within sum of square*.

Berinvestasi saham merupakan kegiatan yang memiliki risiko tinggi, sehingga penentuan portofolio optimal perlu dilakukan agar tujuan investasi dapat tercapai [6]. Untuk mengurangi risiko portofolio, investor wajib mengaplikasikan mekanisme diversifikasi. Diversifikasi diartikan sebagai cara investor dalam mendistribusikan dana untuk menciptakan berbagai kombinasi portofolio sehingga risiko dapat diturunkan tanpa mengurangi *return* yang akan diterima. Beberapa cara membentuk portofolio optimal, salah satunya dengan metode *mean-variance efficient portfolio one constrain* (MVEPOC). MVEPOC bertujuan untuk meningkatkan profit yang diinginkan pada tingkat risiko yang relatif rendah atau sebaliknya meminimalisir risiko pada tingkat profit yang diharapkan.

Penelitian ini bertujuan untuk menciptakan portofolio optimal dengan menggabungkan metode *k-means clustering* dan metode *mean-variance efficient portfolio one constrain* (MVEPOC). Proses diawali dengan membuat klaster saham menggunakan algoritma *k-means*. Kemudian dipilih saham-saham dengan ekspektasi *return* tertinggi dari masing-masing klaster. Selanjutnya saham-saham yang terpilih tersebut akan ditentukan masing-masing bobot dengan algoritma *mean-variance efficient*

portfolio one constrain. Hasil penelitian ini diharapkan dapat menjadi gambaran bagi investor dalam mewujudkan portofolio optimal saham dengan resiko yang kecil.

2. METODE PENELITIAN

2.1 Sumber Data dan Variabel Penelitian

Data yang diterapkan dalam penelitian ini berasal dari website resmi Bursa Efek Indonesia (BEI) dan *Yahoo Finance*. Data tersebut merupakan laporan tentang 50 perusahaan dengan kapitalisasi pasar terbesar pada bulan Maret tahun 2023. Variabel yang digunakan dalam analisis klaterisasi adalah kode saham, nilai total saham yang tersebar (*market capital*), dan jumlah saham yang beredar (*shares*). Selanjutnya informasi riwayat pergerakan harga saham harian selama periode Maret tahun 2023 diperoleh dari platform resmi *Yahoo Finance*. Variabel yang digunakan dalam metode *mean-variance efficient portfolio one constrain* adalah kode saham dan harga penutupan saham.

2.2 Metode K-Means Clustering

Algoritma *k-means* dijalankan berdasarkan prinsip partisi iteratif. Dimulai dengan memilih secara acak k titik pusat klaster (*centroid*), dengan k mewakili jumlah klaster yang telah ditentukan sebelumnya. Selanjutnya, titik data ditetapkan ke klaster yang titik pusat klasternya paling dekat, biasanya diukur dengan jarak Euclidean dengan persamaan berikut:

$$||f - \mu_n|| = \sqrt{\sum_{f \in C_n} (f - \mu_n)^2} \quad (1)$$

di mana μ_n adalah *centroid* dari klaster C_n , dan f mewakili titik data dalam klaster tersebut. *Centroid* kemudian dihitung ulang sebagai rata-rata titik dalam setiap klaster, dan proses berulang hingga mencapai konvergensi [7].

Dalam metode *k-means clustering*, menentukan jumlah klaster setara dengan menentukan titik pusat klaster (*centroid*). Salah satu cara yang banyak diaplikasikan dalam memilih jumlah klaster optimal dengan metode *Elbow* [8]. Metode *Elbow* bekerja dengan menganalisis *within-cluster sum of squares* (WCSS) berdasarkan jumlah klaster (N_c). WCSS dinyatakan sebagai ukuran total jarak kuadrat antara setiap titik data dengan titik pusat klaster (*centroid*). WCSS dirancang untuk meminimalkan, artinya dalam mengelompokkan data memperhatikan jarak minimum antara titik data dengan masing-masing *centroid* yang. WCSS dihitung menggunakan persamaan berikut [9]:

$$WCSS = \sum_{n=1}^{N_c} \sum_{f \in C_n} ||f - \mu_n||^2 \quad (2)$$

Ketika jumlah klaster meningkat maka nilai WCSS menurun karena membagi data menjadi lebih banyak klaster dapat memperkecil jarak antara titik data dengan masing-masing *centroid*. Disisi lain, saat penurunan WCSS menjadi lebih kecil, terdapat titik yang disebut sebagai titik *elbow*, yang menunjukkan keseimbangan optimal antara jumlah klaster dengan kesederhanaan model [9]. Titik ini sebagai acuan dalam memutuskan jumlah klaster optimal suatu model.

2.3 Ekspektasi dan Variansi Return Portofolio

Return saham dinyatakan sebagai proporsi perubahan harga suatu saham selama periode tertentu yang dapat menjadi salah satu faktor dalam memotivasi investor untuk berinvestasi [10]. *Return* dapat dihitung untuk berbagai periode seperti harian, mingguan, bulanan, atau tahunan [11]. Rumus dasar untuk menghitung *return* saham ke- i periode ke- t sebagai berikut:

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}}; i = 1, 2, \dots, m \text{ dan } t = 1, 2, \dots, n \quad (3)$$

dengan $R_{i,t}$ merupakan, $P_{i,t}$ merupakan harga saham ke- i pada akhir periode ke t , dan $P_{i,t-1}$ merupakan harga saham ke- i pada awal periode ke t .

Return portofolio dari m saham pada periode ke- t dapat ditulis sebagai berikut:

$$R_{P,t} = w_1 R_{1,t} + w_2 R_{2,t} + \dots + w_m R_{m,t} \quad (4)$$

dengan $w_1, w_2, \dots, w_m > 0$ merupakan bobot portofolio dan $w_1 + w_2 + \dots + w_m \geq 1$. Dalam membentuk portofolio maka perlu menganalisis tingkat keuntungan dan resiko dengan menghitung ekspektasi *return* dan variansi *return*. Ekspektasi *return* dinyatakan sebagai hasil rata-rata *return* pada saham ke- i .

$$E(R_i) = \bar{R}_i = \frac{\sum_{t=1}^n R_{i,t}}{n} \quad (5)$$

Ekspektasi *return* portofolio dari m saham dapat dinyatakan sebagai berikut:

$$E(R_P) = w_1 E(R_1) + w_2 E(R_2) + \dots + w_m E(R_m) \quad (6)$$

Perhitungan variansi *return* dan kovariansi *return* disajikan dalam formula sebagai berikut:

$$Var(R_i) = \frac{\sum_{t=1}^n (R_{i,t} - \bar{R}_i)^2}{n-1} \quad (7)$$

$$Cov(R_i, R_j) = \frac{\sum_{t=1}^n (R_{i,t} - \bar{R}_i)(R_{j,t} - \bar{R}_j)}{n-1} \text{ dengan } i \neq j; i = 1, 2, \dots, m \quad (8)$$

Sehingga variansi *return* portofolio dari m saham dapat ditulis sebagai berikut:

$$Var(R_P) = \sum_{i=1}^m w_i^2 Var(R_i) + \sum_{i=1}^m \sum_{j=1}^m w_i w_j Cov(R_i, R_j) \quad (9)$$

2.4 Metode Mean-Variance Efficient Portfolio One Constraint

Portofolio optimal merupakan salah satu portofolio terbaik yang dipilih investor diantara himpunan portofolio efisien yang tersedia [5]. Salah satu cara dalam menciptakan portofolio optimal dengan menggunakan metode *mean-variance efficient portfolio* (MVEP). Metode *mean-variance efficient portfolio* dinyatakan sebagai metode pembentukan bobot portofolio dengan mempertimbangkan variansi/ resiko minimum diantara semua kemungkinan susunan bobot portofolio yang dapat terjadi, pada tingkat ekspektasi *return* yang sama [13]. Dalam teknik MVEP dengan satu kendala, proses pembobotan portofolio optimal dilakukan dengan membuat variansi/ resiko menjadi minimal melalui batasan jumlah saham dalam portofolio dan faktor kendala (*constrain*) yang disimbolkan dengan λ .

Membangun vektor bobot ($\mathbf{w}$) dengan variansi minimum maka dapat diselesaikan dengan fungsi Lagrange. Asumsikan semua dana yang dialokasikan dihabiskan untuk membeli saham, sehingga jumlah vektor bobot ($\mathbf{w}$) dapat dinyatakan sebagai berikut:

$$\mathbf{w}'\mathbf{1}_p = 1 \quad (10)$$

Secara matematis, permasalahan optimisasi pada (10) dapat diselesaikan dengan fungsi Lagrange.

$$L(\mathbf{w}, \lambda) = \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} - \lambda(\mathbf{w}'\mathbf{1}_p - 1) \quad (11)$$

Untuk mendapatkan penyelesaian nilai optimal dari $\mathbf{w}$, persamaan (11) diturunkan parsial terhadap $\mathbf{w}$, kemudian hasilnya disamakan dengan nol.

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} &= 0 \\ \frac{\partial}{\partial \mathbf{w}} \left(\frac{1}{2}\mathbf{w}'\Sigma\mathbf{w} - \lambda(\mathbf{w}'\mathbf{1}_p - 1) \right) &= 0 \\ \Sigma\mathbf{w} - \lambda\mathbf{1}_p &= 0 \\ \Sigma\mathbf{w} &= \lambda\mathbf{1}_p \end{aligned}$$

maka diperoleh

$$\mathbf{w} = \Sigma^{-1}\lambda\mathbf{1}_p \text{ atau } \mathbf{w}' = \lambda\mathbf{1}_p'\Sigma^{-1} \quad (12)$$

kemudian mensubstitusikan persamaan (12) ke fungsi Lagrange (11) sehingga menjadi

$$\begin{aligned} L(\mathbf{w}, \lambda) &= \frac{1}{2}\lambda\mathbf{1}_p'\Sigma^{-1}\Sigma\Sigma^{-1}\lambda\mathbf{1}_p - \lambda(\lambda\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p - 1) \\ &= \lambda - \frac{1}{2}\lambda^2\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p \end{aligned} \quad (13)$$

Untuk mendapatkan nilai optimal dari λ , persamaan di atas diturunkan parsial terhadap λ , kemudian hasilnya disamakan dengan nol.

$$\begin{aligned} \frac{\partial}{\partial \lambda} \left(\lambda - \frac{1}{2}\lambda^2\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p \right) &= 0 \\ \lambda\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p &= 1 \\ \lambda &= \frac{1}{\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p} \end{aligned} \quad (14)$$

dengan mensubstitusikan λ pada persamaan (14) ke persamaan (12), diperoleh vektor bobot optimal ($\mathbf{w}$) sebagai berikut:

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}_p}{\mathbf{1}_p'\Sigma^{-1}\mathbf{1}_p} \quad (15)$$

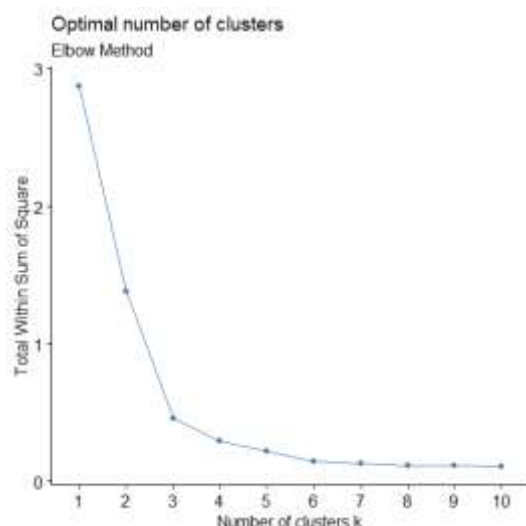
dengan Σ^{-1} merupakan invers matriks varians dan covarians, $\mathbf{1}_p$ merupakan matriks yang semua elemennya bernilai 1, dan $\mathbf{1}_p'$ merupakan transpose dari matriks $\mathbf{1}_p$

3. HASIL DAN PEMBAHASAN

3.1 Klasterisasi dengan K-Mean

Data yang dikumpulkan untuk membentuk klaster terdiri dari 50 perusahaan

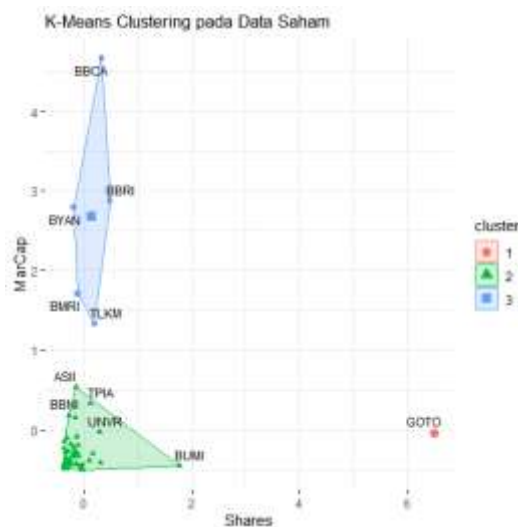
dengan kapitalisasi pasar terbesar yang dihimpun oleh Bursa Efek Indonesia (BEI) tahun 2023. Data yang tersedia adalah nilai total saham perusahaan yang tersebar di pasar saham (*market capital*) dan jumlah lembar saham yang beredar (*shares*). Selanjutnya dilakukan proses klusterisasi dengan algoritma *k-means*. Nilai *k* yang diuji berada pada rentang 1 sampai 10. Tujuan pengujian ini untuk menghasilkan jumlah kelompok yang terbaik dengan memperhatikan grafik *elbow*. Hasil uji ditampilkan pada Gambar 1.



Gambar 1. Grafik *Elbow*

Pada Gambar 1, sumbu *x* menyatakan jumlah klaster dari $k = 1$ sampai $k = 10$ sedangkan sumbu *y* menyatakan *total within sum of square*. Penurunan tajam terjadi dari $k = 1$ ke $k = 2$. Hal serupa terjadi pada $k = 2$ ke $k = 3$. Setelah memasuki $k = 3$ penurunan menjadi landai. Kemudian pada k seterusnya *total within sum of square* menjadi lebih landai dan stabil. Oleh karena itu, *elbow point* atau titik *k* optimum dipilih pada $k = 3$ dimana penurunan *total within sum of square* mulai melambat. Hasil pengelompokan dengan algoritma *k-means* untuk $k = 3$ ditampilkan pada Gambar 2.

Berdasarkan Gambar 2, dari 50 saham perusahaan terbagi menjadi tiga kelompok. Pada tiap-tiap kelompok memiliki karakteristik yang berbeda. Kelompok pertama hanya terdapat satu perusahaan dengan kode GOTO. GOTO memiliki jumlah lembar saham tertinggi dibandingkan dengan perusahaan lainnya. Kelompok kedua terdiri dari 44 perusahaan. Perusahaan-perusahaan tersebut memiliki nilai total saham dan jumlah lembar saham pada level menengah. Sedangkan kelompok terakhir terdiri dari lima perusahaan dengan kode BBKA, BBRI, BYAN, BMRI, dan TLKM yang memiliki nilai total saham perusahaan tertinggi dibandingkan dengan perusahaan lainnya.

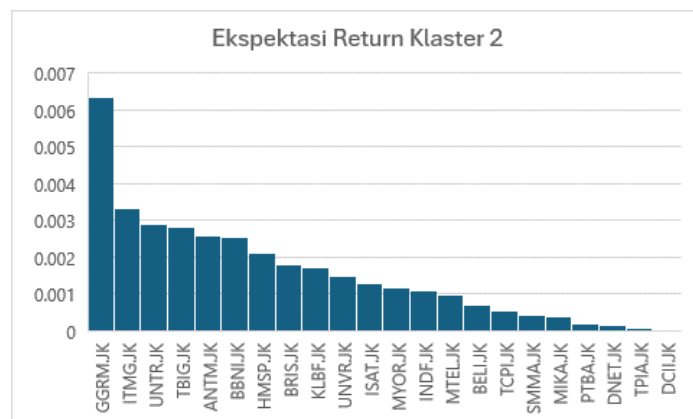


Gambar 2. Pengelompokan dengan Metode K-Means

3.2. Optimasi Bobot Portofolio dengan Mean-Variance Efficient Portfolio One Constrain

Berdasarkan data 50 perusahaan yang telah dihimpun ke dalam 3 klaster. Langkah selanjutnya dipilih satu perusahaan pada tiap-tiap klaster yang memiliki nilai ekspektasi *return* terbesar. Pada proses perhitungan *return*, data yang digunakan adalah harga penutupan saham harian periode Maret tahun 2023.

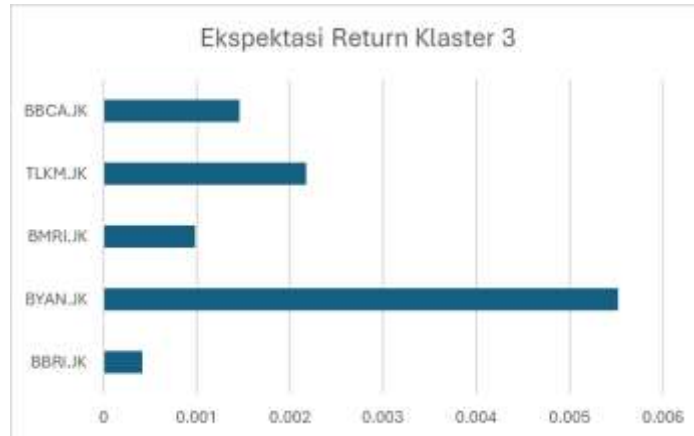
Pada klaster pertama, satunya-satunya perusahaan dengan kode GOTO memiliki nilai ekspektasi *return* sebesar -0.0041. Pada klaster kedua, nilai ekspektasi *return* tiap-tiap perusahaan ditampilkan pada Gambar 3.



Gambar 3. Ekspektasi Return Klaster 2

Diperlihatkan nilai ekspektasi *return* pada klaster kedua paling tinggi adalah perusahaan dengan kode GGRM yang senilai 0,006333 sedangkan paling rendah adalah perusahaan dengan kode EMTK senilai -0,01002. Sementara itu, pada Gambar 3 menyajikan ekspektasi *return* klaster ketiga, perusahaan dengan kode BYAN memberikan ekspektasi *return* paling tinggi sebesar 0,0055 sedangkan yang paling

rendah perusahaan dengan kode BBRI sebesar 0,0004.



Gambar 3. Ekspektasi *Return* Kluster 3

Ketiga kode perusahaan dengan ekspektasi *return* tertinggi pada tiap-tiap kluster diperlihatkan pada Tabel 1.

Tabel 1. Ekspektasi *Return* Tertinggi Setiap Kluster

Kode Saham	Ekspektasi <i>Return</i>
BYAN	0,005515
GOTO	-0,0041
GGRM	0,006333

Pada Tabel 1, saham BYAN memiliki nilai ekspektasi *return* sebesar 0,005515 (0,5515%) yang berarti saham BYAN diperkirakan akan mengalami kenaikan sebesar 0,5515% dalam periode tersebut. Saham GOTO memiliki nilai ekspektasi *return* sebesar -0,0041 (-0,41%) yang menunjukkan saham tersebut diperkirakan akan mengalami penurunan sebesar 0,41%. Saham GGRM dengan ekspektasi *return* terbesar yaitu 0,006333 (0,6333%) yang berarti saham tersebut diprediksi akan mengalami kenaikan sebesar 0,6333%. Ini berarti investasi pada saham BYAN dan GGRM memiliki potensi keuntungan berdasarkan estimasi *return* historis.

Kemudian dipilih saham dengan ekspektasi *return* tertinggi di masing-masing kluster, selanjutnya menghitung varians dan kovarians. Hasil perhitungan disajikan dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} 0,000358 & 0,000109 & 7,57813 \times 10^{-5} \\ 4,00917 \times 10^{-5} & 0,000815 & 0,000175 \\ 7,57813 \times 10^{-5} & 0,000175 & 0,000713 \end{bmatrix}$$

Setelah disusun dalam bentuk matriks varians dan kovarians, kemudian menghitung invers matriks. Hasil invers matriks varians dan kovarians disajikan sebagai berikut:

$$\begin{bmatrix} 2874,043148 & -335,01 & -223,016327 \\ -80,02966 & 1304,434 & -311,511336 \\ -285,569187 & -284,434 & 1501,409853 \end{bmatrix}$$

Tahap selanjutnya menghitung bobot portofolio saham BYAN, GOTO, dan GGRM. Hasil perhitungannya bobot portofolio variansi minimum tersaji dalam Tabel 2.

$$\begin{bmatrix} 2874,043148 & -335,01 & -223,016327 \\ -80,02966 & 1304,434 & -311,511336 \\ -285,569187 & -284,434 & 1501,409853 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2316,017316 \\ 912,8929792 \\ 931,406332 \end{bmatrix}$$

$$[1 \quad 1 \quad 1] \times \begin{bmatrix} 2874,043148 & -335,01 & -223,016327 \\ -80,02966 & 1304,434 & -311,511336 \\ -285,569187 & -284,434 & 1501,409853 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 4160,316628$$

$$w = \frac{1}{4160,316628} \times \begin{bmatrix} 2316,017316 \\ 912,8929792 \\ 931,406332 \end{bmatrix} = \begin{bmatrix} 0,556693 \\ 0,219429 \\ 0,223879 \end{bmatrix}$$

Tabel 2. Bobot Portofolio Optimum

Kode Saham	Bobot Portofolio
BYAN	0,556693
GOTO	0,219429
GGRM	0,223879

Berdasarkan Tabel 2, Saham BYAN memiliki bobot terbesar dalam portofolio sebesar 0,556693 (55,6693%) yang menunjukkan bahwa saham tersebut memiliki pengaruh besar dalam terhadap portofolio. Saham GOTO dan GGRM memiliki bobot yang cukup rendah masing-masing sebesar 21,9429% dan 22,3879% namun tetap memberikan peran signifikan dalam pengurangan resiko total. Sehingga, total dana investasi dialokasikan ke saham BYAN sebesar 55,6693%, saham GOTO sebesar 21,9429%, dan saham GGRM sebesar 22,3879%.

Tabel 3. Ekspektasi Return Portofolio

Saham	Bobot Portofolio (BP)	Ekspektasi Return (ER)	BP × ER
BYAN	0,556693	0,005515	0,00307
GOTO	0,219429	-0,0041	-0,0009
GGRM	0,223879	0,006333	0,001418
Total	1	0,007748	0,003589

Bobot tertinggi dalam portofolio diberikan pada saham BYAN, namun ekspektasi return tertinggi adalah saham GGRM. Ini dikarenakan strategi pembentukan bobot portofolio tidak cukup hanya mempertimbangkan ekspektasi return tetapi mempertimbangkan pengaruh lainnya seperti resiko dan korelasi antar saham. Ekspektasi

return portofolio merupakan penjumlahan bobot portofolio dikalikan ekspektasi *return* saham pada portofolio. Hasil ekspektasi *return* portofolio yang diperoleh pada Tabel 3 sebesar 0,003589 (0,3589%) yang berarti bahwa potensi rata-rata keuntungan yang diharapkan investor dari portofolio ini sebesar 0,3589% terhadap total dana yang dialokasikan dalam periode bulanan, dengan asumsi tidak ada perubahan signifikan dalam kondisi pasar. Selanjutnya mengukur resiko minimum portofolio dengan melakukan perhitungan variansi. Hasil perhitungan sebagai berikut:

$$[0,5566931 \quad 0,219429 \quad 0,223879] \times \begin{bmatrix} 0,000358496 & 0,000109 & 7,57813 \times 10^{-5} \\ 4,00917 \times 10^{-5} & 0,000815 & 0,000175069 \\ 7,57813 \times 10^{-5} & 0,000175 & 0,00071362 \end{bmatrix} \times \begin{bmatrix} 0,556693 \\ 0,219429 \\ 0,223879 \end{bmatrix} = 0,00024$$

Nilai variansi minimum sebesar 0,00024 menunjukkan bahwa portofolio memiliki penyebaran *return* terhadap ekspektasi *return* yang relatif lebih rendah. Sementara itu, standar deviasi adalah $\sqrt{0,00024} = 0,01549$ yang berarti resiko yang relatif rendah dan stabil diberikan pada portofolio tersebut sebesar 1,549%.

4. KESIMPULAN

Data yang dikumpulkan terdiri dari 50 perusahaan dengan kapitalisasi terbesar yang dihimpun oleh Bursa Efek Indonesia (BEI) tahun 2023. Berdasarkan data tersebut dibentuk klaster berdasarkan *market capital* dan jumlah lembaran saham yang beredar. Proses pembentukan klaster menggunakan metode *k-means*. Hasil penelitian dengan metode tersebut memberikan gambaran tiga klaster yang memiliki karakteristik berbeda. Klaster pertama dapat diartikan sebagai klaster dengan jumlah lembar saham tertinggi, anggota klaster pertama adalah saham GOTO. Klaster kedua adalah klaster yang memiliki nilai total saham dan jumlah lembar saham pada level menengah, terdiri dari 44 kode saham. Terakhir, klaster ketiga merupakan klaster yang memiliki nilai total saham perusahaan tertinggi, terdiri dari saham BBKA, BBRI, BYAN, BMRI, dan TLKM. Selanjutnya terpilih 1 saham dari tiap-tiap klaster dengan *return* tertinggi, yaitu saham BYAN, GOTO, dan GGRM untuk dimasukkan kedalam portofolio. Kemudian, dilakukan optimasi bobot portofolio dengan metode varians minimum. Hasil analisis diperoleh bobot portofolio saham BYAN, GOTO, dan GGRM masing-masing sebesar 0,556693; 0,219429; dan 0,223879. Sebesar 55,6693% dana untuk investasi dialokasikan untuk saham BYAN dan sisanya 21,9429% untuk saham GOTO dan 22,3879% untuk saham GGRM. Nilai ekspektasi *return* portofolio untuk saham GOTO, BYAN, dan GGRM sebesar 0,003589 yang berarti potensi rata-rata keuntungan yang diharapkan investor dari portofolio ini sebesar 0,3589%. Resiko yang diterima investor dalam portofolio tersebut relatif rendah dan stabil, yaitu sebesar 1,549%.

DAFTAR PUSTAKA

- [1] G. C. Nutakki and O. Nasraoui, "Clustering Data Streams with Adaptive Forgetting," in *2017 IEEE International Congress on Big Data (BigData Congress)*, IEEE, Jun. 2017, pp. 494–497. doi: 10.1109/BigDataCongress.2017.72.

- [2] C. Abraham, N. Molinari, and R. Servien, "Unsupervised clustering of multivariate circular data," *Statistics in Medicine*, vol. 32, no. 8, pp. 1376–1382, Apr. 2013, doi: 10.1002/sim.5589.
- [3] I. K. Khan, H. Daud, N. Zainuddin, and R. Sokkalingam, "Standardizing reference data in gap statistic for selection optimal number of cluster in K-means algorithm," *Alexandria Engineering Journal*, vol. 118, pp. 246–260, Apr. 2025, doi: 10.1016/j.aej.2025.01.034.
- [4] F. Amanah, F. Roshafara, P. I. Lestari, S. Salsabila, and R. Maharani, "Utilizing K-Means Clustering for Constructing Black-Litterman Portfolio Models," *Jurnal Matematika, Statistika dan Komputasi*, vol. 20, no. 3, pp. 670–679, May 2024, doi: 10.20956/j.v20i3.34165.
- [5] Putri Isnaini Cahyaning Baiti and J. Prasetya, "Analisis Non-Hierarchical Clustering dan Lagrangean Multiplier dalam Penentuan Bobot Portofolio Optimal Saham Perbankan Indonesia," *Jurnal Statistika dan Aplikasinya*, vol. 6, no. 2, pp. 355–365, Dec. 2022, doi: 10.21009/JSA.06219.
- [6] I. Suputra, *Pembentukan Portofolio Optimal Saham-Saham LQ45 Berdasarkan Sentimen Positif Menggunakan Metode DBSCAN dan Markowitz*, Doctoral dissertation, Institut Teknologi Sepuluh Nopember, 2025.
- [7] E.-M. Papia and A. Kondi, "Quantifying subtle color transitions in Mark Rothko's abstract paintings through K-means clustering and Delta E analysis," *Journal of Cultural Heritage*, vol. 72, pp. 194–204, Mar. 2025, doi: 10.1016/j.culher.2025.02.005.
- [8] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster," *IOP Conference Series: Materials Science and Engineering*, vol. 336, p. 012017, Apr. 2018, doi: 10.1088/1757-899X/336/1/012017.
- [9] M. Barzegar, S. M. Cherati, D. J. Pasadas, C. Pernechele, A. L. Ribeiro, and H. G. Ramos, "Baseline-free damage imaging of CFRP lap joints using K-means clustering of guided wave signals," *Mechanical Systems and Signal Processing*, vol. 229, p. 112562, Apr. 2025, doi: 10.1016/j.ymssp.2025.112562.
- [10] B. Setiawaty, H. A. F. Istiadi, Ruhiyat, and W. Erliana, "ESTIMASI TAIL VALUE AT RISK RETURN PORTOFOLIO TIGA OBLIGASI MENGGUNAKAN SIMULASI MONTE-CARLO," *MILANG Journal of Mathematics and Its Applications*, vol. 20, no. 2, pp. 77–87, Dec. 2024, doi: 10.29244/milang.20.2.77-87.
- [11] A. Kautsar, N. N. Thahirrah, A. Safitri, M. F. W. Rahman, D. T. Wijayati and R. Amelia, "Computational Mathematics of Risk and Return Portofolio Banking (Ramadhan Effect 2024)," 2024 12th International Conference on Cyber and IT Service Management (CITSM), Batam, Indonesia, 2024, pp. 1-5, doi: 10.1109/CITSM64103.2024.10775616
- [13] G. P. Gilang, M. Gunawan Supiarmo, and Junaidi, "Analisis Perbandingan Portofolio Saham dengan Metode Mean Varian Satu Kendala dan Dua Kendala," *Mandalika Mathematics and Educations Journal*, vol. 6, no. 2, pp. 739–751, Dec. 2024, doi: 10.29303/jm.v6i2.7258.