JURNAL STATISTIKA DAN SAINS DATA
Volume 2 Nomor 2, April, 2025, 121-135
https://journal.student.uny.ac.id/index.php/jssd
http://dx.doi.org/10.21831/jssd.v2i2.19964

CLUSTERING SAHAM SEKTOR KONSUMEN PRIMER DENGAN METODE *FUZZY C-MEANS* DAN *SELF ORGANIZING MAP*

Ratna Triutami*, Universitas Negeri Yogyakarta
Agus Maman Abadi, Universitas Negeri Yogyakarta
*e-mail: ratnatriutami.2019@student.uny.ac.id

Abstrak. Saham merupakan bukti penyertaan modal dari seseorang/pihak (badan usaha) pada suatu perusahaan atau perseroan terbatas. Dengan demikian pihak tersebut berhak mendapatkan klaim atas aset dan pendapatan perusahaan. Tujuan dari penelitian ini adalah untuk mengelompokkan saham sektor konsumen primer sebagai pertimbangan bagi calon investor yang akan berinvestasi. Digunakan data saham sektor konsumen primer yang terdaftar di BEI pada Juni 2023, data diambil dari yahoo finance dan website resmi BEI. Data terdiri dari 110 saham dengan 3 variabel, meliputi *return on equity*, *return on assets*, dan *earning per share*. Langkah-langkah penelitian meliputi pengumpulan data, *preprocessing* data, analisis klaster metode *Fuzzy C-Means* dan *Self Organizing Map*, validasi klaster optimal, dan perbandingan hasil klaster. Dalam penelitian ini digunakan Indeks Dunn sebagai metode validasi klaster. Pengelompokkan dengan metode FCM menghasilkan jumlah klaster optimal sebanyak 2 klaster, sedangkan dengan metode SOM diperoleh jumlah klaster optimal sebanyak 3 klaster. Metode FCM dengan pengelompokan 2 klaster memiliki nilai indeks Dunn sebesar 0,7383. Sedangkan metode SOM dengan pengelompokan 3 klaster memiliki nilai indeks Dunn sebesar 0,6064.

Kata kunci: saham, sektor konsumen primer, Clustering, Fuzzy C-Means, Self Organizing Map.

Abstract. *Stock is a proof of capital participation from a person/party (business entity) in a company or limited liability company. Thus the party is entitled to a claim on the assets and income of the company. The purpose of this research is to classify the consumer non-cyclicals sector stocks as a consideration for potential investors who will invest. This research used data from consumer non-cyclicals sector stocks that listed on the IDX in June 2023, data was taken from yahoo finance and the official IDX website. The Data consists of 110 stocks with 3 variables, including return on equity, return on assets, and earnings per share. The research steps include data collection, data preprocessing, cluster analysis using Fuzzy C-Means and Self Organizing Map methods, optimal cluster validation, and comparison of cluster results. In this research, Dunn Index is used as a cluster validation method. Clustering with FCM method produces 2 clusters as the optimal clusters, while the SOM method obtained 3 clusters as the optimal clusters. The FCM method with 2 cluster grouping has a Dunn Index value of 0.7383. While the SOM method with grouping 3 clusters has Dunn index value of 0.6064.*

Keywords: stock, consumer non-cyclicals, Cluster, Fuzzy C-Means, Self Organizing Map.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

PENDAHULUAN

Saham merupakan bukti penyertaan modal seseorang/pihak (badan usaha) pada sebuah perusahaan atau perseroan terbatas. Seseorang atau pihak yang ikut serta dalam permodalan memiliki hak atas harta dan pendapatan perusahaan, serta hak untuk menghadiri Rapat Umum Pemegang Saham (RUPS) (Abi, 2016). Salah satu cara berinvestasi saham adalah melalui pasar modal. Bursa Efek Indonesia (BEI) bertugas menjalankan dan mengawasi pasar modal negara. Selama dua tahun terakhir, terjadi peningkatan pola investasi, khususnya di kalangan investor milenial yang menunjukkan minat lebih besar pada investasi pasar modal. Telah diamati bahwa sekitar 60% investor milenial berkontribusi pada pasar modal. Minat investor dalam berinvestasi semakin meningkat karena pasar modal menawarkan peluang keuangan yang fleksibel dan transparan untuk pembagian risiko dan kemungkinan pengembalian yang positif.

Ada berbagai macam indeks dan sektor saham yang terdaftar di BEI, salah satunya yaitu sektor konsumen primer atau *consumer non-cyclicals*. Sektor konsumen primer meliputi perusahaan yang memproduksi atau memasok barang-barang yang sering dijual kepada konsumen, tetapi hanya untuk barang-barang primer atau anti-siklis, dengan demikian permintaan akan barang dan jasa tersebut tidak terpengaruh oleh pertumbuhan ekonomi (Bursa Efek Indonesia). Sektor konsumen primer dinilai sebagai saham defensif atau saham yang kinerjanya tidak terpengaruh oleh kondisi ekonomi. Dengan demikian saham di sektor ini dapat terus tumbuh di era ketidakpastian ekonomi. Terjadinya inflasi memang berpengaruh terhadap daya beli masyarakat, tetapi efeknya dinilai tidak terlalu signifikan terhadap saham konsumen primer karena masyarakat akan tetap fokus untuk memenuhi kebutuhan primer yang disediakan oleh saham-saham di sektor ini. Selain itu, dilansir dari kompas.com, sejak awal tahun hingga Rabu, 8 Maret 2023, sebagian besar indeks sektoral telah terkoreksi, menurut data dari Mirae Asset Sekuritas. Dari sebelas indeks sektoral, hanya tiga sektor yang menunjukkan pertumbuhan, salah satunya sektor konsumen primer dengan kenaikan sebesar 1,59%.

Karena meningkatnya kebutuhan masyarakat, bisnis di sektor produk konsumen primer telah berkembang dengan sangat cepat. Direktorat Jenderal Kependudukan dan Pencatatan Sipil (Dukcapil) melaporkan peningkatan jumlah penduduk setiap tahunnya, yang menunjukkan bahwa akan ada permintaan yang berkelanjutan dan terus meningkat untuk produk konsumen primer. Salah satu faktor yang berkontribusi terhadap peningkatan profitabilitas adalah tingkat permintaan pelanggan untuk barang-barang manufaktur. Karena itu, investor menganggap sub-sektor konsumen sebagai salah satu kategori investasi yang menarik, terutama bagi mereka yang tujuan utamanya adalah saham jangka panjang. Karena salah satu informasi terpenting yang dipertimbangkan investor saat memilih saham untuk dibeli adalah profitabilitas perusahaan (Susanty & Pangestuti, 2022). Namun karena banyaknya saham yang terdaftar dalam sektor ini tentu akan sulit untuk membandingkan kinerjanya satu-persatu, oleh karena itu dilakukan analisis klaster untuk mengelompokkan perusahaan dengan karakteristik kinerja yang sama.

Mendapatkan *return*/laba atas investasinya adalah motivasi utama bagi seorang investor. Seorang investor mengalami kerugian jika tingkat pengembalian lebih rendah dari risiko yang diasumsikan. Agar tidak mengalami kerugian saat berinvestasi maka seorang investor harus mengamati dan mempertimbangkan saham mana yang benar-benar berkualitas melalui rasio keuangannya. Rasio keuangan dihitung berdasarkan komponen-komponen dalam laporan keuangan. Rasio keuangan digunakan dalam evaluasi kinerja perusahaan, perbandingan dari periode-ke-periode, dan perbandingan kinerja antar perusahaan. Salah satu rasio yang sering diperhatikan investor yaitu rasio profitabilitas. Rasio profitabiltas merupakan rasio yang menggambarkan efektifitas perusahaan untuk mendapatkan keuntungan berdasarkan sumber

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

daya yang dimiliki. Tingkat profitabiltas yang semakin tinggi menunjukkan bahwa perusahaan tersebut mampu menggunakan sumber dayanya dengan baik guna menghasilkan keuntungan (Christine & Winarti, 2022).

Menurut penelitian Widayanti & Colline (2017), harga saham sangat dipengaruhi oleh rasio *debt to equity* dan *earning per share*, tetapi tidak oleh *current ratio* dan total perputaran aset. Oleh karena itu, sebelum melakukan investasi, investor harus meninjau laporan keuangan perusahaan dan rasio keuangan, terutama *debt to equity* dan *earning per share*. Astuti & Di Asih (2021), melakukan penelitian terkait analisis pengelompokan saham berdasarkan rasio keuangan. Dalam penelitian ini digunakan metode pengelompokan *Hierarchical clustering* dengan *average linkage*. Variabel yang dipakai meliputi data aktivitas dari laporan keuangan perusahaan dan indikator rasio likuiditas, solvabilitas, dan profitabilitas. Astuti & Di Asih (2021) menyatakan bahwa rasio keuangan merupakan instrumen evaluasi yang krusial untuk menilai kinerja perusahaan.

Analisis pengelompokan/*clustering* dalam kasus ini digunakan untuk mengelompokkan saham-saham dengan karakteristik yang sama. Rahmatika, dkk (2015), menerapkan clustering untuk pembentukan portofolio, adapun metode yang digunakan adalah *Gustafson Kessel Clustering* (GK) dan *Fuzzy C-Means*. Penelitian tersebut mengkaji penerapan komparatif metode pengelompokan FCM dan GK dalam studi kasus pengelompokan saham dalam LQ45 berdasarkan rasio saham EPS dan *price earning ratio* (PER). Pada beberapa kasus, hasil penelitian menunjukkan bahwa GK lebih baik daripada FCM, khususnya dalam menghasilkan fungsi objektif dan rasio standar deviasi kelompok minimum. Dalam penelitian ini, penulis ingin melakukan analisis mengenai pengelompokan atau *clustering* saham di sektor konsumen primer dengan analisis klaster *Fuzzy C-Means* (FCM) dan *Self Organizing Map* (SOM). Prinsip dasar FCM adalah minimalisasi fungsi objektif. Pada kondisi awal pusat klaster belum akurat sehingga dilakukan perbaikan pusat klaster dan derajat keanggotaan secara berulang berdasarkan minimalisasi fungsi objektifnya. Fungsi objektif yang menggambarkan jarak data ke pusat klaster yang terbobot oleh derajat keanggotaannya. Sedangkan SOM bekerja dengan memperbarui bobot secara berulang dan menghitung jarak bobot baru dengan bobot lama. Vektor bobot setiap unit klaster berfungsi sebagai contoh dari input pola yang terkait dengan klaster tersebut. Selama proses pengelompokan klaster yang bobotnya sesuai dengan pola vektor input yang paling dekat dipilih sebagai pemenang. Adapun perbedaan penelitian ini dengan penelitian-penelitian sebelumnya adalah data dan variabel yang digunakan, karena belum banyak peneliti yang melakukan *clustering* pada sektor konsumen primer. Hasil dari penelitian ini diharapkan dapat menjadi pertimbangan bagi calon investor dalam memilih saham yang akan diivestasikan nantinya.

METODE

Penelitian ini menggunakan data rasio keuangan saham sektor konsumen primer. Terdapat 120 saham yang terdaftar dalam sektor konsumen primer di BEI, tetapi data dari 10 saham diantaranya tidak lengkap, sehingga digunakan 110 saham yang memiliki data lengkap. Data rasio keuangan tersebut diperoleh dari laporan keuangan yang diterbitkan oleh perusahaan melalui website BEI, khususnya dari laporan laba rugi dan neraca. Dalam penelitian ini digunakan laporan keuangan Triwulan II tahun 2023, yaitu pada bulan Juni 2023. Periode tersebut dipilih karena merupakan data terbaru, sehingga diharapkan dapat menggambarkan kinerja perusahaan yang lebih aktual. Adapun variabel yang digunakan, yaitu *return on assets* (ROA), *return on equity* (ROE), dan *earning per share* (EPS), ketiganya dalam satuan persen.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

Ketiga variabel tersebut dipilih karena ketiga rasio ini sangat penting untuk mengetahui kesehatan dan kinerja keuangan perusahaan. Selain itu, karena tujuan investor berinvestasi adalah untuk memperoleh keuntungan maka dipilih rasio-rasio yang juga menggambarkan kemampuan perusahaan dalam menghasilkan keuntungan/laba. Meskipun ROE dan ROA menawarkan sudut pandang yang berbeda, namun menggabungkan data keduanya akan menghasilkan pemahaman yang komprehensif tentang seberapa baik kinerja manajemen perusahaan. Perusahaan dikatakan dapat memperoleh keuntungan dengan modal pemegang saham jika memiliki ROE dan ROA yang tinggi, serta dapat mengelola hutang. Namun, ROE yang tinggi juga tidak menjamin kinerja keuangan baik jika perusahaan memiliki ROA yang rendah dan tingkat utang yang tinggi.

Rasio keuangan adalah rasio yang dihitung dengan menggunakan laporan keuangan sebagai patokan untuk mengevaluasi kinerja dan kesehatan keuangan perusahaan. Rasio keuangan merupakan nilai yang diperoleh dari membandingkan satu item laporan keuangan dengan item lain yang terkait (Hery, 2016). Investor dapat menggunakan rasio keuangan ini sebagai alat untuk menilai profitabilitas perusahaan tergantung pada jumlah saham yang dimiliki. Permintaan investor akan menaikkan harga saham perusahaan jika kinerja keuangannya menunjukkan prospek yang positif. Hal ini menunjukkan bahwa harga saham perusahaan dan keberhasilan finansial berkorelasi positif (Raharjo & Muid, 2013).

Hanafi dan Halim (2003) menyatakan bahwa ROA adalah rasio keuangan yang menilai kapasitas perusahaan untuk menghasilkan keuntungan pada jumlah modal, aset, dan pendapatan tertentu. Secara umum, aset adalah sesuatu yang dimiliki oleh seseorang, perusahaan, atau entitas lain dan memiliki nilai ekonomis. Di sisi lain, aset perusahaan adalah semua uang atau sumber daya yang dimilikinya dan berpotensi untuk diproduksi di masa depan. Aset perusahaan memiliki nilai ekonomi yang dapat diukur secara objektif dan dapat berwujud atau tidak berwujud. Dalam ilmu akuntansi aset adalah hasil penjumlahan modal dan hutang. Misalkan suatu perusahaan memiliki nilai ROA sebesar 6% berarti laba bersih yang mampu dihasilkan sebesar 6% dari total aset yang dimiliki. Nilai ROA yang baik adalah 5% atau lebih. Semakin baik perusahaan menggunakan asetnya untuk menghasilkan keuntungan, semakin tinggi pulai nilai ROA-nya.

Tandelilin (2001) menyatakan bahwa ROE menunjukkan seberapa menguntungkan suatu perusahaan dalam menjalankan operasinya. ROE mengevaluasi seberapa baik perusahaan mengelola dan menggunakan modalnya untuk menghasilkan keuntungan. Otoritas Jasa Keuangan (OJK) mendefinisikan modal sebagai sejumlah uang yang biasanya diperoleh melalui penerbitan saham, serta digunakan untuk menjalankan operasi perusahaan. Modal dalam konteks bisnis adalah item neraca yang menunjukkan jumlah kekayaan bersih perusahaan yang bukan berasal dari operasinya. Modal dan aset berbeda karena meskipun modal merupakan komponen dari aset, tidak semua aset merupakan modal sebab adanya aset tidak berwujud yang tidak bisa dijual. Apabila sebuah perusahaan memiliki nilai ROE sebesar 20% menandakan jika setiap investasi Rp1.000 akan memberikan keuntungan bersih setidaknya 20%. Perusahaan dikatakan baik apabila memiliki ROE 15% atau lebih, semakin nilai ROE mendekati 100% maka semakin baik. Perbedaan utama antara ROA dan ROE adalah cara memperhitungkan hutang. Bunga tidak perlu ditambahkan ke pembilang saat menghitung ROE karena hutang tidak termasuk dalam perhitungan.

Kasmir (2008) berpendapat bahwa EPS menunjukkan laba yang dihasilkan oleh setiap saham yang beredar. Perusahaan dengan nilai EPS tinggi biasanya lebih disukai oleh investor karena menunjukkan bahwa investor akan menerima dividen yang lebih besar.

Setelah data diperoleh kemudian dilakukan *prepocessing data* yaitu dengan melakukan standarisasi pada data. Pada setiap atribut dalam kumpulan data seringkali ditemukan rentang

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

nilai yang berbeda, standarisasi berupaya menyamakan rentang tersebut. Peran karakteristik dalam kumpulan data menjadi tidak berfungsi ketika perbedaan rentang nilai cukup menyimpang (Whendasmoro & Joseph, 2022). Metode z-score adalah pendekatan yang dapat digunakan ketika nilai aktual terendah dan tertinggi dari data tidak diketahui (Nasution, dkk, 2019).

Selanjutnya data yang telah distandarisasi akan dianalisis dengan metode *Fuzzy C-Means* (FCM) dan *Self Organizing Map* (SOM). FCM dipilih agar dapat mengatasi masalah outlier yang sering terjadi pada data saham. Kinerja setiap perusahaan yang berbeda-beda terkadang memunculkan outlier yang tidak dapat diatasi oleh analisis klaster metode K-means. Algoritma klastering *Fuzzy C-Means* (FCM) diinisiasi oleh Dunn (1973), kemudian Bezdek (1981) mengembangkannya menjadi lebih baik. FCM merupakan suatu metode klastering dimana derajat keanggotaan menjadi penentu letak titik dalam suatu klaster. Menurut Kusumadewi (2010) berikut merupakan langkah-langkah algoritma FCM:

1. Menentukan input dalam bentuk matriks $n \times p$, dengan n adalah banyaknya data dan p adalah variabel yang digunakan.

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

2. Menentukan jumlah cluster (c), *fuzzifier* / pangkat pembobot $(m > 1)$, maksimum iterasi (*MaxIter*), error terkecil yang diharapkan (ε), dan fungsi objektif awal $P_0 = 0$.
3. Membangkitkan bilangan random u_{ik} yang merupakan elemen matriks partisi U, dimana $i = 1,2, \dots, n$ dan $k = 1,2, \dots, c$.
4. Menghitung pusat klaster atau *centroid*

$$v_{kj} = \frac{\sum_{i=1}^{n} (u_{ik})^m x_{ij}}{\sum_{i=1}^{n} (u_{ik})^2} \quad (2)$$

dengan $k = 1,2, \dots, c$ dan $j = 1,2, \dots, p$.

5. Menghitung fungsi objektif pada iterasi ke-t (P_t)

$$P_t = \sum_{i=1}^{n} \sum_{k=1}^{c} (u_{ik})^m d_{ik} \quad (3)$$

dengan d_{ik} merupakan jarak euclidean yang dihitung dengan persamaan

$$d_{ik} = \sqrt{\sum_{j=1}^{p} (x_{ij} - v_{kj})^2} \quad (4)$$

6. Menghitung perubahan matriks keanggotaan $u_{i_r k}$

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

$$u_{ik} = \left[\frac{\left[\sum_{j=1}^{p}(x_{ij}-v_{kj})^2\right]^{\frac{1}{m-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{p}(x_{ij}-v_{kj})^2\right]^{\frac{1}{m-1}}} \right]^{-1} \quad (5)$$

7. Memeriksa apakah $|P_t - P_{t-1}| < \varepsilon$ atau $t > MaxIter$. Iterasi dapat dihentikan jika kondisi tersebut terpenuhi. Namun apabila tidak terpenuhi maka diulang dari langkah keempat.

Tahun 1982 Teuvo Kohonen memperkenalkan algoritma *Self Organizing Map* (SOM) sebagai metode *unsupervised learning* dalam jaringan saraf tiruan. Fungsi utama jaringan SOM adalah memetakan data input dari ruang berdimensi lebih rendah ke ruang berdimensi lebih tinggi (seringkali plot satu atau dua dimensi) sambil mempertahankan keterkaitan topologi asli. Lokasi fisik titik-titik pada peta menunjukkan kesamaan relatif antara titik-titik di ruang multi-dimensi (Kiang, 2001). Lapisan input dan output pada arsitektur SOM masing-masing terdiri dari satu lapisan. Dengan bobot konektivitas, setiap unit di lapisan input terhubung ke semua unit di lapisan output (Ghozali & Sugiharto, 2017). Algoritma dari metode *Self Organizing Map* adalah sebagai berikut (Halim & Widodo, 2017):

1. Menginisialisasi bobot (w_{kj}) dengan acak bagi setiap node, kemudian masukkan jaringan input (x_i).
2. Melakukan perhitungan jarak vektor d_{ki} jika input sudah diterima, dengan cara menambahkan selisih vektor boot (w_{kj}) dengan vektor input (x_{ij}).

$$d_{ki} = \sqrt{\sum_{j=1}^{p}(w_{kj} - x_{ij})^2} \quad (6)$$

3. Menentukan nilai terkecil jrak vektor d_{ki}, kemudian dilakukan perubahan bobot.

$$w_{kj}(baru) = w_{kj}(lama) + \alpha[x_{ij} - w_{kj}(lama)] \quad (7)$$

4. Bobot baru diperoleh berdasarkan nilai *learning rate* (α), yaitu $0 \leq \alpha \leq 1$. Pada masing-masing iterasi, nilai *learning rate* akan menurun menjadi $\alpha_{(t+1)} = 0{,}5\alpha$.
5. Penghentian dilakukan dengan menghitung selisih bobot baru dan bobot lama. Pengujian dikatakan konvergen apabila selisih tidak signifikan, dan algoritma bisa dihentikan.

Setelah melakukan analisis klaster kemudian dilakukan validasi klaster. Memilih jumlah klaster optimal, membandingkan berbagai teknik pengelompokan, dan mengevaluasi stabilitas klaster merupakan fungsi validasi klaster. Indeks Dunn adalah kriteria validasi klaster yang digunakan dalam penelitian ini. Pada tahun 1973, J. C. Dunn memperkenalkan indeks Dunn sebagai ukuran validitas klaster. Diperhitungkan rasio diameter klaster tertinggi dengan jarak minimum antar klaster. Jika hanya ada sedikit perbedaan di antara anggota klaster, sebuah klaster dikatakan kompak; jika klaster berjauhan, mereka dikatakan terpisah dengan baik. Adapun jarak yang digunakan adalah jarak euclidean.
Terdapat tiga langkah menghitung indeks indeks Dunn menurut Ncir, dkk (2021):

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

1. Kekompakan dihitung dengan cara mengevaluasi jarak antar dua objek dalam suatu klaster, kemudian dihitung nilai maksimumnya.

$$\Delta_k = \max_{x_{i1}, x_{i2} \in C_k} d(x_{i1}, x_{i2}) \quad (8)$$

dimana Δ_k adalah kekompakan, C_k adalah klaster, $d(x_{i1}, x_{i2})$ adalah jarak Euclidean data x_{i1} dengan data x_{i2}.

2. Derajat pemisahan antar pasangan klaster dihitung dengan menentukan jarak terdekat antar dua objek pada klaster yang berbeda. Namun karena perhitungannya yang terlalu kompleks dan rentan terhadap *outlier*, maka pemisahan ini biasanya dihitung berdasarkan *centroid* klaster.

$$\delta_{k1,k2} = d(v_{k1}, v_{k2}) \quad (9)$$

dimana $\delta_{k1,k2}$ adalah pemisahan pasangan klaster (C_{k1}, C_{k2}), v_i dan v_j adalah centroid klaster C_i dan C_j.

3. Mencari nilai indeks Dunn (*ID*) dengan menghitung rasio nilai minimum pemisahan dibagi dengan nilai maksimum kekompakan.

$$ID = \frac{\min_{1 \le k1 \ne k2 \le c} \delta_{k1,k2}}{\max_{1 \le k \le c} \Delta_k} \quad (10)$$

Dalam kasus perbandingan beberapa metode klaster, menurut Gupta & Panda (2019) nilai indeks Dunn untuk yang lebih besar menunjukkan bahwa metode tersebut adalah lebih baik dari yang lain. Dalam penelitiannya Gupta membandingkan metode K-means dan CLARA, indeks Silhouette dan indeks Dunn dapat memvalidasi pengelompokan, dan disimpulkan bahwa CLARA merupakan pengelompokan yang lebih baik daripada Pengelompokan K-Means karena memiliki nilai indeks yang lebih tinggi. Misuraca, dkk (2018) juga membandingkan nilai indeks Dunn dan indeks BMS untuk versi K-means yang berbeda. Nilai yang sama dengan 1 berarti indeks telah mengidentifikasi suatu partisi sebagai solusi terbaik di antara semua partisi. BMS, serta indeks Dunn menunjukkan bahwa Bisecting K-Means merupakan solusi terburuk

HASIL DAN PEMBAHASAN

Hasil

1. Metode *Fuzzy C-Means*

Data input yang digunakan dalam penelitian ini adalah data hasil standarisasi dengan tiga variabel, yaitu EPS, ROA, dan ROE. Beberapa parameter yang digunakan, antara lain jumlah klaster $(c = 2)$, pangkat pembobot $(m = 2)$, iterasi maksimum $(MaxIter = 1000)$, dan nilai error terkecil yang diharapkan $(\varepsilon = 10^{-9})$. Pada metode FCM dengan $m = 2$. Setelah diperoleh matriks *U* akan dihitung centroid ke-1 menggunakan persamaan (1).

$$v_{1j} = \frac{\sum_{i=1}^{110} (u_{i1})^2 x_{ij}}{\sum_{i=1}^{110} (u_{i1})^2}$$

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

$$= \frac{1}{46{,}0763}[-14{,}32505 \quad -2{,}9877 \quad -23{,}777]$$
$$= [-0{,}31038 \quad -0{,}06473 \quad -0{,}51567]$$

Dengan langkah yang sama dihitung centroid ke-1, sehingga diperoleh nilai centroid akhir dari klastering metode FCM dengan $m = 2$ yang dinyatakan dalam matriks berikut

$$v = \begin{bmatrix} -0{,}31038 & -0{,}06473 & -0{,}51567 \\ 0{,}27009 & 0{,}15083 & 0{,}63920 \end{bmatrix}$$

Pusat klaster tersebut digunakan untuk memperbarui derajat keanggotaan matriks U. Namun, sebelum memperbarui derajat keanggotaan akan dihitung jarak antara data dengan pusat klasternya (jarak euclidean) terlebih dahulu dengan persamaan (3).

$$d_{1,1}^2 = \left(-0{,}3245 - (-0{,}31038)\right)^2 + \left(0{,}18084 - (-0{,}06473)\right)^2 + \left(-0{,}0777 - (-0{,}51567)\right)^2$$
$$d_{1,1}^2 = 0{,}2523$$
$$d_{1,2}^2 = \left(-0{,}3245 - (0{,}27009)\right)^2 + \left(0{,}18084 - (0{,}15083)\right)^2 + \left(-0{,}0777 - (0{,}63920)\right)^2$$
$$d_{1,2}^2 = 0{,}8684$$

Selanjutnya dihitung derajat keanggotaan yang baru dengan persamaan (4).

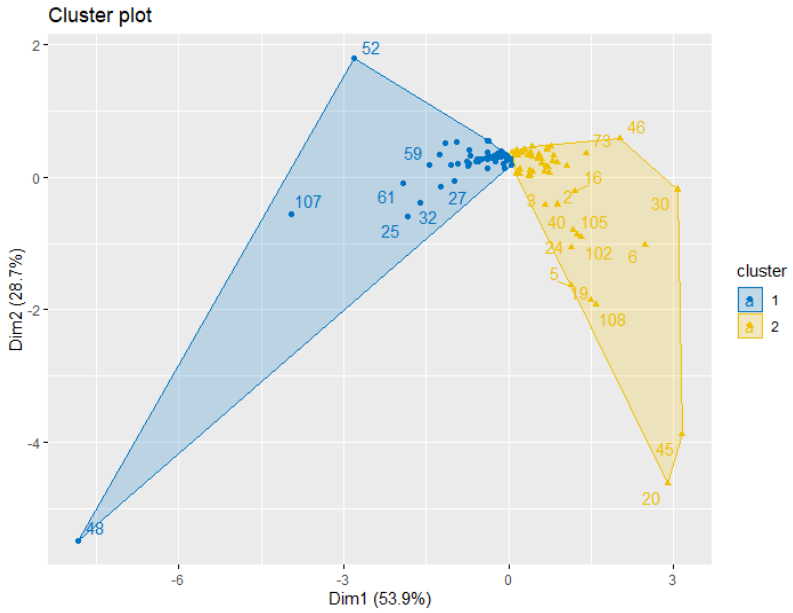
$$u_{1,1} = \frac{0{,}2523^{-1}}{0{,}2523^{-1} + 0{,}8684^{-1}} = 0{,}77486$$
$$u_{1,2} = \frac{0{,}8684^{-1}}{0{,}2523^{-1} + 0{,}8684^{-1}} = 0{,}25514$$

Berdasarkan perhitungan dengan program RStudio, diperoleh nilai fungsi objektif iterasi ke-78 sebesar $152{,}5828$. Adapun perubahan nilai fungsi objektifnya yaitu $|P_{77} - P_{78}| = |152{,}5828 - 152{,}5828| = 0$, karena kondisi optimal $|P_{77} - P_{78}| < \varepsilon$ sudah terpenuhi maka proses iterasi dihentikan. Dengan demikian pengelompokan metode FCM dengan $m = 2$ dan $c = 2$ mencapai kondisi optimal pada iterasi ke-78. Pada metode FCM, matriks partisi U akhir digunakan untuk menentukan suatu perusahaan masuk ke dalam klaster 1 atau 2. Hasil analisis klaster dengan metode FCM disajikan dalam Tabel 1.

Tabel 1. Hasil klaster metode FCM

Klaster	Jumlah Anggota
1	57
2	53

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649



Gambar 1. Visualisasi hasil klaster metode FCM

Gambar 1 merupakan visualisasi hasil klastering 110 saham sektor konsumen primer dengan 3 varibel menggunakan metode FCM dengan banyak klaster 2. Warna biru merupakan klister 1 yang beranggotakan 57 perusahaan, diantaranya BWPT, JPFA, AISA, ALTO, dan lannnya. Sedangkan warna kuning merupakan klaster 2 yang beranggotakan 53 perusahaan, diantaranya CEKA, EPMT, AALI, ADES, dan lainnya.

2. Metode *Self Organizing Map*

Langkah awal pembentukan klaster metode *Self Organizing Map* (SOM) setelah data distandarisasi adalah menentukan jumlah neuron. Jumlah neuron digunakan untuk mengklasifikasi data input ke dalam beberapa klaster yang jumlahnya sama dengan neuron input. Dalam penelitian ini data yang digunakan sebanyak 110 dengan 3 variabel, jumlah klaster sebanyak 3, dan maksimum iterasi 1000.

Setelah menentukan jumlah neuron kemudian dilakukan proses *training* pada jaringan, dengan tujuan meng-*update* bobot awal secara acak. *Training* pada jaringan dilakukan dengan maksimum 1000 iterasi. Hasil *training* menunjukkan iterasi konvergen terjadi pada iterasi ke-500. Setelah training mencapai maksimum selanjutnya ditentukan bobot akhir yang memenuhi *training*. Bobot akhir digunakan untuk menentukan jarak data input dengan neuron. Jarak data input dengan neuron dihitung menggunakan jarak euclidean, yang nantinya juga digunakan untuk menentukan klaster. Tabel 2 menunjukkan bobot akhir metode SOM dengan 3 klaster yang diperoleh dengan bantuan program RStudio.

Tabel 2. Bobot akhir SOM dengan 3 klaster

	EPS	ROE	ROA
W1	2,4841705	0,2565553	0,8034479
W2	-0,5000435	-4,8458090	-4,0621547
W3	-0,2759038	0,1133321	0,0568053

Selanjutnya dihitung jarak inter-cluster untuk masing-masing klaster menggunaan persamaan (5)

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

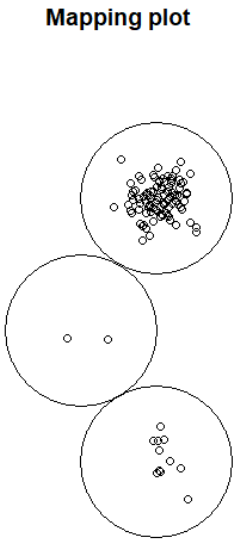
$$d_{1,1} = \sqrt{\left(2{,}4842 - (-0{,}3245)\right)^2 + (0{,}2565 - 0{,}1808)^2 + \left(0{,}8034 - (-0{,}0777)\right)^2}$$
$$d_{1,1} = 2{,}94$$
$$d_{1,2} = \sqrt{\left(-0{,}5004 - (-0{,}3245)\right)^2 + (-4{,}8458 - 0{,}1808)^2 + \left(-4{,}0621 - (-0{,}0777)\right)^2}$$
$$d_{1,2} = 6{,}42$$
$$d_{1,3} = \sqrt{\left(-0{,}2759 - (-0{,}3245)\right)^2 + (0{,}1133 - 0{,}1808)^2 + \left(0{,}0568 - (-0{,}0777)\right)^2}$$
$$d_{1,3} = 0{,}16$$

Berdasarkan perhitungan tersebut diperoleh jarak antara data input dengan neuron 1 sebesar 2,94, jarak dengan neuron 2 sebesar 6,42, dan jarak dengan neuron 3 sebesar 0,16. Dengan langkah yang sama dilakukan perhitungan jarak hingga data ke-110. Dari jarak tersebut kemudian dilakukan pembentukan klaster. Data input dikelompokkan berdasarkan jaraknya, semakin kecil jarak berarti semakin dekat dengan neuron. Hasil analisis klaster dengan metode SOM disajikan dalam Tabel 3.

Tabel 3. Hasil klaster metode SOM

Klaster	Jumlah Anggota
1	11
2	2
3	97

Hasil klastering data saham sektor konsumen primer dengan banyak data 110 dan 3 variabel menggunakan metode SOM dengan 3 klaster, diperoleh sebanyak 11 saham masuk ke dalam klaster 1, 2 saham masuk ke dalam klaster 2, dan 97 saham lainnya masuk ke dalam klaster 3. Klaster plot metode SOM dengan 3 klaster disajikan pada Gambar 2.



Gambar 2. Plot hasil klaster metode SOM

Selanjutnya akan dilakukan uji validitas klaster. Indeks validitas digunakan untuk menentukan optimal atau tidaknya hasil klaster yang terbentuk. Indeks validitas yang

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

digunakan untuk menentukan jumlah klaster optimal metode FCM dan SOM adalah Indeks Dunn. Nilai koefisien indeks Dunn yang semakin tinggi menunjukkan bahwa semakin optimal pula klister yang dihasilkan. Dalam penelitian ini hanya digunakan satu indeks validitas dengan maksud untuk menghindari adanya perbedaan hasil kluster optimal dengan metode lain.

Nilai validitas berdasarkan indeks Dunn hasil klastering data saham sektor konsumen primer menggunakan metode FCM dengan banyak klaster 2, 3, 4, 5, dan 6 tertera pada Tabel 4.

Tabel 4. Indeks Dunn metode FCM

	FCM $k=2$	FCM $k=3$	FCM $k=4$	FCM $k=5$	FCM $k=6$
Indeks Dunn	0,7383	0,7166	0,6632	0,6247	0,6051

Berdasarkan Tabel 4 dapat dilihat bahwa nilai indeks Dunn tertinggi yang dihasilkan yaitu sebesar 0,7383 pada banyak klaster 2 ($k = 2$). Dengan demikian jumlah klaster optimal peengelompokan saham sektor konsumen primer dengan metode FCM adalah sebanyak 2 klaster.

Nilai validitas berdasarkan indeks Dunn hasil klastering data saham sektor konsumen primer menggunakan metode SOM dengan banyak klaster 2, 3, 4, 5, dan 6 tertera pada Tabel 5.

Tabel 5. Indeks Dunn metode SOM

	SOM $k=2$	SOM $k=3$	SOM $k=4$	SOM $k=5$	SOM $k=6$
Indeks Dunn	0,4154	0,6064	0,431	0,3604	0,4005

Tabel 5 menunjukkan bawa nilai indeks Dunn tertinggi yang dihasilkan yaitu sebesar 0,6064 pada banyak klaster 3 ($k = 3$). Dengan demikian jumlah klaster optimal pengelompokan saham sektor konsumen primer dengan metode SOM adalah sebanyak 3 klaster.

Pembahasan

Klaster optimal yang hasilkan dari pengelompokan dengan metode FCM terhadap data saham sektor konsumen primer yaitu 2 klaster, klaster 1 beranggotakan 57 saham dan klaster 2 beranggotakan 53 saham. Sedangkan klaster optimal yang dihasilkan dari pengelompokan dengan metode SOM terhadap data saham sektor konsumen primer yaitu 3 klaster, klister 1 sebanyak 11 saham, klister 2 sebanyak 2 saham, dan klaster 3 sebanyak 97 saham. Setelah memperoleh klaster optimal berdasarkan Indeks Dunn akan dilihat statistik deskriptif setiap klaster untuk mengkategorikan jenis klaster.

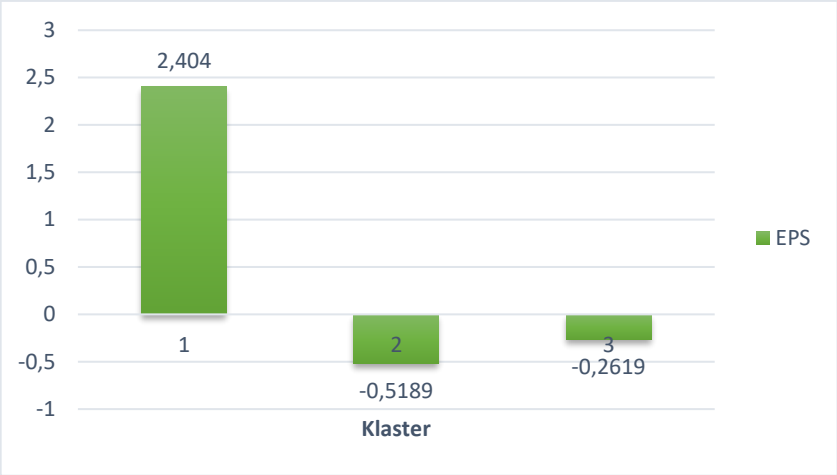
Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

Gambar 3. Perbandingan rata-rata EPS klaster FCM

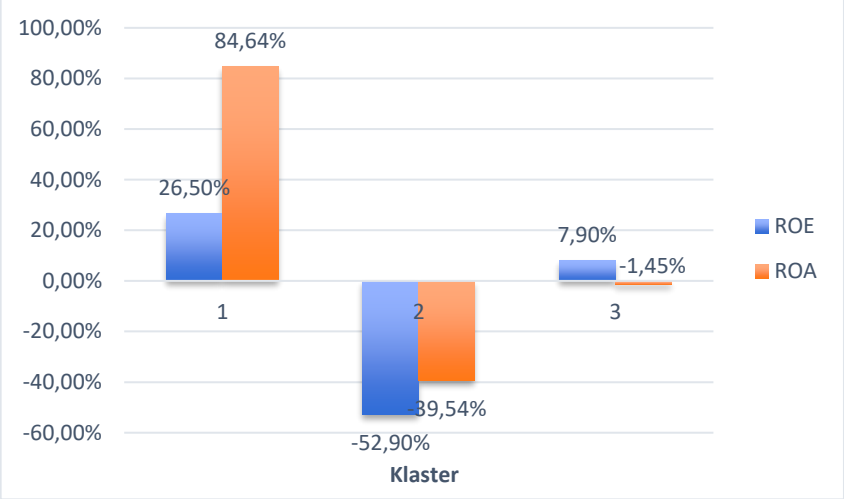
Gambar 4. Perbandingan rata-rata ROE dan ROA klaster FCM

Dari Gambar 3 dapat disimpulkan bahwa rata-rata EPS klaster 1 lebih rendah daripada pada klaster 2, maka dapat dikatakan bahwa perusahan-perusahan di klaster 2 lebih baik dari pada klaster 1. Selain itu, berdasarkan Gambar 4 terlihat bahwa klaster 2 juga memiliki rata-rata ROE 0,19019 atau 19,019% dan rata-rata ROA 0,08579 atau 8,579%. Dengan demikian klaster 2 merupakan kelompok perusahan dengan kinerja keuangan yang baik karena nilai ROE >15% dan nilai ROA >5%. Sedangkan klaster 1 memiliki rata-rata ROE sebesar -35,5% dan rata-rata ROA sebesar 0,31%, sehingga perusahaan-perusahaan di klaster 1 memiliki kinerja yang kurang baik.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649



Gambar 5. Perbandingan rata-rata EPS klaster SOM



Gambar 6. Perbandingan rata-rata ROE dan ROA klaster SOM

Dari Gambar 5 dapat disimpulkan bahwa klaster 1 memiliki rata-rata EPS paling tinggi di antara ketiga klaster, sedangkan klaster 2 memiliki rata-rata EPS terendah. Berdasarkan Gambar 6, nilai ROE rata-rata tertinggi juga dimiliki oleh klaster 1 sebesar 26,5%, kemudian klaster 3 sebesar 7,9%, dan terendah klaster 2 sebesar -52,9%. Sedangkan rata-rata nilai ROA tertingi dimiliki oleh klaster 1 sebesar 84,64%, klaster 3 sebesar -1,44%, dan terendah klaster 2 sebesar -39,5%. Dengan demikian dapat dikatakan bahwa klaster 1 merupakan kelompok perusahaan dengan kinerja keuangan baik, klaster 3 merupakan kelompok perusahaan dengan kinerja keuangan kurang baik, sedangkan, klaster 2 merupakan perusahaan dengan kinerja keuangan buruk.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

Tabel 6. Daftar saham klaster 2 FCM dan klaster 1 SOM

Klaster 2 FCM	CEKA, EPMT, AALI, ADES, AMRT, BISI, CPIN, CPRO, DLTA, DSNG, FISH, GGRM, HMSP, ICBP, LSIP, MLBI, MRAT, MYOR, SGRO, SKLT, SMAR, UNVR, WIIM, CLEO, CAMP, PANI, GOOD, ITIC, KEJU, PSGO, CSRA, PGUN, KMDS, VICI, FAPA, TAPG, CMRY, STAA, TLDN, TRGU, EURO, BUAH, CBUT, WINE, NAYZ, ROTI, SSMS, STTP, TBLA, TGKA, ULTJ, INDF, dan DEWI
Klaster 1 SOM	AALI, ADES, FISH, GGRM, ICBP, MLBI, SGRO, SMAR, STTP, TGKA, dan INDF

Bagi investor yang ingin berinvestasi di saham sektor konsumen primer disarankan untuk memilih saham-saham yang termasuk dalam klaster 1 SOM atau klaster 2 FCM. Berdasarkan Tabel 16 seluruh saham dari Klaster 1 SOM juga termasuk dalam Klaster 2 FCM, sehingga tidak ada kontradiksi dari hasil kedua metode klastering. Kedua klaster tersebut merupakan kelompok perusahan dengan kinerja keuangan yang baik, dikatakan baik karena memiliki rata-rata *Return on Assets* (ROA) di atas 5%, rata-rata *Return on Equity* (ROE) di atas 15%, dan rata-rata *Earning per Share* (EPS) cukup tinggi dibanding klaster lainnya.

Hasil perbandingan antara metode FCM dan SOM untuk klastering saham sektor konsumen primer menunjukkan bahwa metode FCM dengan pengelompokan 2 klaster lebih baik, karena memiliki nilai indeks Dunn sebesar 0,7383. Sedangkan metode SOM dengan pengelompokan 3 klaster memiliki nilai indeks Dunn sebesar 0,6064. Namun perlu diingat bahwa hasil pengelompokan tersebut hanya berdasarkan data rasio profitabilitas, sehingga hanya memberi gambaran kepada calon investor mengenai kemampuan/kinerja perusahaan dalam memperoleh keuntungan. Untuk dapat melihat kinerja perusahaan secara keseluruhan ada baiknya calon investor juga melihat rasio-rasio keuangan lainnya.

SIMPULAN

Dari penelitian yang telah dilakukan mengenai klastering saham sektor konsumen primer menggunakan metode *Fuzzy C-Means* dan *Self oganizing Map*, diperoleh kesimpulan bahwa hasil pengelompokan dengan metode FCM diperoleh banyaknya klaster optimal adalah 2 klaster, klaster 1 beranggotakan 57 saham dan klaster 2 beranggotakan 53 saham. Klaster 1 dikategorikan sebagai kelompok perusahaan dengan kinerja kurang baik, sedangkan klaster 2 dikategorikan sebagai kelompok perusahan dengan kinerja keuangan yang baik Hasil pengelompokan dengan metode SOM diperoleh banyaknya klaster optimal adalah 3 klaster, dengan anggota klaster 1 sebanyak 11 saham, klaster 2 sebanyak 2 saham, dan klaster 3 sebanyak 97 saham. Klaster 1 dikategorikan sebagai kelompok perusahaan dengan kinerja keuangan baik, klaster 2 dikategorikan sebagai kelompok perusahaan dengan kinerja keuangan buruk, sedangkan klaster 3 dikategorikan sebagai kelompok perusahaan dengan kinerja keuangan kurang baik,

Perbandingan antara metode FCM dan SOM untuk klastering saham sektor konsumen primer menunjukkan bahwa metode FCM dengan pengelompokan 2 klaster lebih baik, karena memiliki nilai indeks Dunn sebesar 0,7383. Sedangkan metode SOM dengan pengelompokan 3 klaster memiliki nilai indeks Dunn sebesar 0,6064.

DAFTAR PUSTAKA

Abi, F. P. P. (2016). *Semakin dekat dengan pasar modal Indonesia*. Deepublish.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649

Astuti, T. D., & Di Asih, I. M. (2021). Klasterisasi perusahaan second liner di masa resesi global dengan indikator rasio keuangan. *Jurnal Riset Akuntansi Mercu Buana*, *7*(2).

Christine, D., & Winarti, W. (2022). Pengaruh Return on Assets (ROA), Return on Equity (ROE), dan Earning per Share (EPS) terhadap harga saham:(Studi empiris pada perusahaan subsektor makanan dan minuman yang terdaftar di bursa efek Indonesia Periode 2018-2020). *Owner: Riset dan Jurnal Akuntansi*, *6*(4), 4113-4124.

Gupta, T., & Panda, S. P. (2019). Clustering validation of CLARA and k-means using silhouette & DUNN measures on iris dataset. *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 10-13. doi: 10.1109/COMITCon.2019.8862199.

Halim, N. N., & Widodo, E. (2017). Clustering dampak gempa bumi di Indonesia menggunakan Kohonen Self Organizing Maps (SOM). *Prosiding SI MaNIs (Seminar Nasional Integrasi Matematika dan Nilai-Nilai Islami)*, 1(1), 188-194.

Hanafi, M. M., & Halim, A. (2003). Analisis laporan keuangan edisi revisi. *Yogyakarta: UPP AMP YKPN*, *101*.

Irwansyah, E., & Faisal, M. (2015). *Advanced clustering: Teori dan aplikasi*. Yogyakarta: Deepublish.

Kasmir, 2008. *Analisis laporan keuangan*. Jakarta: PT. Raja Grafindo Persada.

Kiang, M. Y. (2001). Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics & Data Analysis*, *38*(2), 161-180.

Kusumadewi, S. (2010). *Aplikasi logika fuzzy untuk pendukung keputusan*. Graha Ilmu.

Misuraca, M., Spano, M., & Balbi, S. (2019). BMS: An improved Dunn index for document clustering validation. *Communications in statistics-theory and methods*, 48(20), 5036-5049.

Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma K-NN. *CESS (Journal of Computer Engineering, System and Science)*, *4*(1), 78-82.

Ncir, C. E. B., Hamza, A., & Bouaguel, W. (2021). Parallel and scalable Dunn index for the validation of big data clusters. *Parallel Computing*, *102*, 102751.

Raharjo, D., & Muid, D. (2013). Analisis pengaruh faktor-faktor fundamental rasio keuangan terhadap perubahan harga saham. *Diponegoro Journal of Accounting*, 444-454.

Susanty, S. M. D., & Pangestuti, D. C. (2022). Analisis nilai perusahaan sektor barang konsumen primer di bursa efek indonesia. *AKUNTABEL: Jurnal Ekonomi dan Keuangan*, *19*(2), 220-229.

Tandelilin, E. 2001. *Analisis investasi dan manajemen portofolio*. Yogyakarta: BPFE

Whendasmoro, R. G., & Joseph, J. (2022). Analisis penerapan normalisasi data dengan menggunakan Z-score pada kinerja algoritma K-NN. *JURIKOM (Jurnal Riset Komputer)*, *9*(4), 872-876.

Widayanti, R., & Colline, F. (2017). Pengaruh rasio keuangan terhadap harga saham perusahaan LQ 45 periode 2011-2015. *Bina ekonomi*, *21*(1), 35-49.

Copyright © 2025, Jurnal Statistika dan Sains Data, Online ISSN: 3025-9649