

# Pengaruh Cluster Emiten terhadap Return Saham JSX Berbasis Parameter Rasio Analisa Fundamental (*JSX Stock Issuer Cluster Analysis Based on Fundamental Analysis Ratio Parameters*)

Berlian Karlina<sup>1\*</sup>, Ario Menak Sanoyo<sup>2</sup>

Fakultas Ekonomi dan Bisnis, Universitas Budi Luhur Jakarta, Jakarta Selatan<sup>1</sup>

Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Noverber, Surabaya<sup>2</sup>

[berlian.karlina@budiluhur.ac.id](mailto:berlian.karlina@budiluhur.ac.id)<sup>\*</sup>, [ariomenak@yahoo.com](mailto:ariomenak@yahoo.com)<sup>2</sup>



Riwayat Artikel

Diterima pada 3 Mei 2021

Direvisi pada 21 Mei 2021

Disetujui pada 27 Mei 2021

## Abstract

**Purpose:** This research aimed to find the effect of cluster techniques in determining stock selection to maximize return and minimize risk in the stock market.

**Research Methodology:** The methodology consists of two of several algorithmic approaches of the clustering method to find hidden patterns in a group of datasets, i.e., Partitioning clusters (k-means) defined by the dataset object and its central area, and hierarchical clusters that group data through varying scales to be implemented into cluster trees or dendrogram. Dataset summary analysis of the fundamental ratio of stocks in the study was obtained from IDX stock data.

**Results:** This study's classification has been obtained that consists of three zones: green, blue and red zone. The significance obtained provides an alternative form of stock categorization, creating an investment decision support system based on Cluster Analysis, the search for correlations and patterns between ratios of the Financial Statements as complementary tools of Investment Risk Management.

**Limitations:** This research uses only two clustering algorithm methods to analyze the effect of clustering in maximizing return and minimizing risk and only used variables of financial reports for the company listed on the Indonesia Stock Exchange.

**Contribution:** The risk management portfolio is a crucial part of being analyzed for investors and management to improve financial performance. As a complement to decision support, the risk management systems have to be analyzed based on cluster analysis and subsequent data mining to know the potential stock valuation in the market.

**Keywords:** *Stock, Partitioning Clusters, Hierarchical Clusters, Financial Statements, Fundamental Analysis Ratio*

**How to cite:** Karlina, B., & Sanoyo, A. M. (2021). Pengaruh Cluster Emiten terhadap Return Saham JSX berbasis Parameter Rasio Analisa Fundamental. *Jurnal Akuntansi, Keuangan, dan Manajemen*, 2(4), 279-291.

## 1. Pendahuluan

Pertumbuhan bisnis saat ini sedang mengalami peningkatan yang cukup tinggi. Tidak hanya terjadi pada sektor tertentu, pertumbuhan tersebut dirasakan oleh berbagai sektor industri saat ini. Pertumbuhan ini akibat dari dampak globalisasi yang saat ini sedang terjadi, dimana kemudahan melakukan bisnis dapat dirasakan yang mengakibatkan tingkat kompetisi pasar semakin ketat. Sehingga untuk dapat bertahan di pasar di tengah kompetisi yang sangat ketat diperlukan strategi khusus untuk terus mengembangkan inovasi agar bisa terhindar dari kebangkrutan. Seperti yang kita

tahu, setiap perusahaan masing-masing memiliki tujuan. Salah satu tujuan utama perusahaan ialah memperoleh laba seoptimal mungkin. Dengan keberhasilan dalam mencapai tujuan utama perusahaan, maka akan membawa pengaruh positif terhadap pemilik perusahaan dan para pemegang saham.

Dengan semakin *competitive* nya dunia pasar modal yang ada di Indonesia, maka memberikan tantangan bagi investor atau bahkan calon investor untuk menganalisa jenis saham yang layak untuk mereka investasikan. Diperlukan analisa yang mendalam untuk dapat menentukan saham yang baik dan kurang baik, sehingga dapat menghasilkan potensi keuntungan yang tinggi untuk investor. Keuntungan dalam saham sendiri berasal dari dua sumber, yaitu keuntungan dari capital gain dan pembagian dividen. Dalam keuntungan yang diperoleh dari capital gain bersumber dari perbedaan antar harga jual dan harga beli dimana dalam hal ini harga jual saham harus lebih mahal dari harga beli. Sementara pembagian dividen dibagikan oleh perusahaan kepada berdasarkan dari keuntungan perusahaan. Keuntungan perusahaan sendiri dapat investor proyeksikan berdasarkan dari kinerja keuangan perusahaan.

Harga saham yang merupakan cerminan dari pertumbuhan keuntungan bagi investor merupakan variabel yang sangat penting untuk dianalisa. Harga saham yang merupakan bagian dari nilai perusahaan memberikan *perspective* terhadap investor yaitu dengan melihat tingkat keuntungan yang tercermin dari harga saham perusahaan tersebut (M. Sholihun, 2019). Nilai perusahaan memberikan gambaran tentang kinerja keuangan perusahaan yang dapat memberikan pengaruh terhadap persepsi investor dalam menentukan investasinya. Tingginya nilai perusahaan akan mengakibatkan keyakinan investor akan semakin tinggi terhadap perusahaan tersebut. *Price to book value* (PBV) merupakan salah satu rasio yang dapat menentukan nilai perusahaan. PBV memberikan sinyal bagi manajemen perusahaan tentang bagaimana investor memiliki gambaran terhadap suatu saham dilihat dari tingkat risiko dan potensi pertumbuhan perusahaan di masa depan.

Maka, untuk dapat memproyeksikan data-data tersebut dibutuhkan analisa yang cukup mendalam. Dalam penelitian ini menggunakan analisa *cluster*. Dimana analisa *cluster* merupakan salah satu teknik statistik nonparametrik yang sangat berperan dalam pengklasifikasian obyek dan data. Pengaplikasian ini sudah banyak di terapkan di dalam berbagai *real-case* obyek seperti *signaling*, medis, internet big data, social science, ekonometrik, biologi, berita dan lain-lainnya. Dalam penelitian ini, akan di bahas penggunaan analisa klaster sebagai sistem pendukung keputusan yang diterapkan dalam dunia perdagangan dan investasi saham. Ada beberapa jenis parameter yang dapat di jadikan input untuk teknik analisa klaster, baik itu menggunakan data berbasis fundamental keuangan maupun teknikal yang berbasis *time series*.

Dalam dunia investasi, kita mengenal adanya pengelompokan jenis-jenis saham berdasarkan parameter yang dapat di kategorikan seperti antara lain:

1. *Blue chip stocks*

Merupakan jenis saham yang memiliki value terbaik dengan likuiditas tinggi, konsistensi pendapatan, dividen dan *annual growth* yang secara historis terjamin serta lebih resistant terhadap krisis.

2. *Speculative stocks*

Saham jenis ini tidak secara terus menerus memperoleh keuntungan setiap tahunnya. Namun, memiliki tingkat probabilitas keuntungan yang tinggi pada masa mendatang walaupun kemungkinan tersebut tidak pasti.

3. *Cyclical stocks*

Saham jenis ini dapat dengan mudah dipengaruhi oleh pertumbuhan ekonomi secara umum. Nilai dari saham yang termasuk dalam jenis ini cenderung untuk turun saat terjadi perubahan makroekonomi atau sistematis dalam perekonomian secara keseluruhan.

4. *Growth Stocks*

Saham jenis ini dimiliki oleh perusahaan yang mengalami pertumbuhan yang lebih tinggi daripada industrinya sendiri. Namun, saham ini tidak mempunyai catatan (*proven track record*), sehingga pertumbuhan pada saham jenis ini memiliki risiko yang lebih tinggi jika dibandingkan dengan

saham lain, tetapi memiliki kemungkinan kenaikan harga yang potensial di masa yang akan datang.

#### 5. *Emerging Growth Stock*

Saham jenis ini perusahaan yang relatif lebih kecil namun mempunyai kemampuan yang kuat walaupun di tengah trend ekonomi yang kurang baik. Meskipun demikian, harga saham pada jenis ini biasanya sangat fluktuatif sehingga risiko pada saham jenis ini cukup tinggi.

#### 6. *Defensive Stocks*

Saham jenis ini adalah kebalikan dari jenis Saham Siklus (*cyclical stocks*). *Defensive stock* dapat bertahan sekalipun pada masa resesi. Perusahaan yang tergolong pada *defensive stock* adalah perusahaan yang memproduksi obat-obatan, makanan, minuman, *insurance*, dan jenis kebutuhan dasar dimana kebutuhan ini tidak akan terpengaruh di masa resesi atau bahkan dalam krisis.

Penelitian ini mencoba memformulasikan proses pengelompokan data emiten saham berbasis fundamental seperti laporan keuangan dengan menggunakan algoritma *data mining*, pengenalan pola berbasis statistik, sehingga diharapkan dapat menemukan bentuk korelasi dan menghasilkan pola *cluster* yang berbeda dari bentuk pengkategorian tersebut. *Data mining* diklasifikasikan menjadi prediktif dan deskriptif. Metode prediksi membuat prediksi tentang nilai-nilai data dan metode deskriptif mengidentifikasi hubungan antar data. *Data Mining* terutama ditujukan untuk ekstraksi informasi yang sebelumnya tidak diketahui dan sangat membantu dari kumpulan data yang tersedia. Kategori utama dari data mining (Bini dan Mathew 2016). Teknik meliputi pengelompokan, klasifikasi dan regresi. Dalam penelitian ini dilakukan untuk mendirikan sistem berbasis IT dengan menggunakan penerapan metode *k-means clustering*.

Berdasarkan latar belakang yang telah dibahas diatas, permasalahan yang dapat dirumuskan pada penelitian ini:

1. Bagaimana cara mengimplementasikan metode *k-means clustering*, dalam mencari portofolio saham yang memberikan risiko minimal dengan *return* tertentu?
2. Bagaimana pengaruh dari penambahan teknik *clustering* terhadap kinerja portofolio yang terbentuk?

Dalam konteks penelitian ini, parameter-parameter yang akan di angkat adalah rasio rasio saham berbasis analisa fundamental, seperti *Price Earning Ratio* (PER), *Price to Book Value* (PBV), *Earning per Share* (EPS), dan *Return on Equity* (ROE). Rasio tersebut merupakan *summary* dari laporan keuangan yang mencerminkan kondisi kesehatan perusahaan pada periode tertentu. Para *trader*, *investor* dan *analyst* berpedoman pada rasio-rasio keuangan tersebut sebagai dasar informasi pendukung keputusan berinvestasi dan digunakan juga sebagai materi rujukan dalam mengeluarkan rekomendasi pasar pada waktu tertentu. Rasio keuangan tersebut merupakan elemen yang dinamis, berubah seiring waktu sesuai dengan pergerakan harga saham dan kondisi pasar saat itu.

## 2. Tinjauan pustaka dan Hipotesis

### 2.1. Teori Clustering

*Clustering* adalah metode untuk mengekstrak pengetahuan tentang struktur tersembunyi dalam data atau mempartisinya menjadi kelompok (*cluster*) tanpa pengetahuan atau pengalaman sebelumnya tentangnya (Nesic, Radojevic 2015). Inti dari setiap metode pengelompokan adalah cara kita membandingkan objek. Objek dibandingkan sehubungan dengan posisi, karakteristik atau perilakunya (Aleksandar 2020). Menurut Everitt et al 2011, tidak ada ukuran universal yang dapat mempertimbangkan aspek kedekatan yang berbeda. Sebaliknya, seseorang harus memilih salah satu yang sesuai dari serangkaian luas pengukuran jarak, kesamaan / ketidaksamaan, atau korelasi yang berbeda.

Di dalam *business intelligence*, *clustering* bisa digunakan pengguna ke dalam beberapa kategori yang dibutuhkan. Contohnya mengklasifikasikan pelanggan ke dalam beberapa jenis kelompok berdasarkan tingkat karakteristik kesamaan yang tinggi. *Clustering* dapat diartikan sebagai segmentasi data dimana **clustering** mengelompokkan beberapa jenis data set ke dalam sejumlah kelompok berdasarkan kesamaan yang kuat.

Dalam *finance intelligence, clustering* dapat digunakan sebagai penentuan jenis investasi yang sesuai dengan portofolio investor. Seperti dalam penelitian ini, menggunakan data rasio keuangan untuk mengklasifikasikan jenis saham berdasarkan kinerja keuangan perusahaan. *Clustering* di sini dapat dijadikan rekomendasi bagi investor dalam berinvestasi.

## 2.2. Rasio Keuangan

Analisis rasio keuangan merupakan bagian terpenting dalam melakukan perkiraan laporan keuangan yang dapat dijadikan dasar dalam mengevaluasi kondisi kesehatan keuangan perusahaan. Perhitungan rasio keuangan akan lebih bermanfaat dengan membandingkan rasio keuangan perusahaan dari tahun ke tahun untuk mengetahui trend yang sedang berlaku pada perusahaan yaitu kenaikan atau penurunan (Hery, 2015).

Rasio keuangan sendiri adalah rasio yang dapat ditentukan dengan memperhitungkan rasio berdasarkan laporan keuangan perusahaan dan dapat berfungsi sebagai parameter dalam menilai kondisi keuangan perusahaan pada suatu waktu (Hery, 2017). Dalam bahasa sederhana rasio merupakan membandingkan total nilai dimana dengan total nilai dapat terlihat perbandingan antar periode sehingga perbandingan tersebut dapat dijadikan bahan sebagai sumber analisa sebelum memberikan keputusan (Fahmi, 2018).

Manfaat dari rasio keuangan dapat dirasakan oleh perusahaan, investor, atau kreditor. Berikut beberapa manfaat yang bisa dari rasio keuangan (Fahmi, 2018), yaitu:

1. Analisis rasio keuangan dapat digunakan sebagai parameter dalam penilaian kinerja dan pertumbuhan perusahaan pada suatu periode.
2. Analisis rasio keuangan dapat dijadikan rujukan atau sumber data dalam membuat suatu perencanaan bagi manajemen atau investor.
3. Analisis rasio keuangan adalah parameter yang digunakan dalam melakukan evaluasi kondisi keuangan perusahaan pada satu periode.
4. Analisis rasio keuangan dapat bermanfaat untuk berbagai pihak, yaitu tidak hanya bermanfaat bagi manajemen dan investor tetapi juga memberikan manfaat bagi kreditor. Kreditor dapat menentukan potensi risiko perusahaan berdasarkan dari analisis rasio keuangan pada suatu periode.
5. Analisis rasio keuangan merupakan sumber penilaian bagi stakeholder perusahaan dalam membuat keputusan.

Rasio keuangan sendiri dibagi menjadi enam (Kasmir, 2016), yaitu:

1. Rasio Likuiditas dapat memberikan gambaran terkait kesanggupan perusahaan dalam memenuhi kewajiban jangka pendeknya. Ada beberapa jenis rasi likuiditas, yaitu:
  - a. *Current Ratio* adalah rasio untuk mengetahui kemampuan perusahaan dalam membayar kewajiban jangka pendeknya.
  - b. *Quick Ratio* adalah rasio untuk mengetahui kemampuan perusahaan dalam memenuhi pembayaran kewajiban jangka pendeknya dengan aktiva lancar dengan mengabaikan nilai *inventory* perusahaan tersebut.
  - c. *Cash Ratio* adalah rasio yang digunakan untuk mengetahui ukuran uang kas yang tersedia dalam membayar kewajiban perusahaan.
  - d. *Cash Turnover* adalah rasio yang digunakan untuk mengukur tingkat ketersediaan kas dalam melakukan pembayaran kewajiban dan biaya lain yang terkait dengan penjualan perusahaan.
  - e. *Inventory to Net Working Capital* adalah rasio yang digunakan untuk mengetahui tingkat ketersediaan dengan modal kerja perusahaan.
2. Rasio Solvabilitas adalah rasio yang digunakan dalam mengetahui besaran aktiva perusahaan yang dibiayai dengan utang. Beberapa jenis rasio solvabilitas yang sering digunakan dalam sumber analisis adalah:
  - a. *Debt to Asset Ratio* adalah rasio yang digunakan untuk mengetahui besaran aktiva perusahaan yang dibiayai oleh utang atau dapat dikatakan seberapa besar utang perusahaan memberikan pengaruh terhadap pengelolaan aktiva dalam suatu periode tertentu.

- b. *Debt to Equity Ratio* adalah rasio ini sering digunakan untuk melihat proporsi utang terhadap ekuitas. Dalam menentukan DER yaitu dengan membagi seluruh utang dengan seluruh ekuitas.
  - c. *Long Term Debt to Equity Ratio* adalah perbandingan antara rasio kewajiban jangka panjang dengan equity atau modal sendiri dalam suatu periode.
  - d. *Times Interest Earned* adalah rasio untuk mengetahui perolehan bunga perusahaan.
  - e. *Fixed Charge Coverage* tetap adalah rasio yang mirip dengan *Times Interest Earned* dimana dapat mengetahui perolehan bunga, namun dalam rasio ini dapat ditentukan jika perusahaan memiliki hutang jangka panjang dan terdapat kontrak sewa yang dilakukan oleh perusahaan.
3. Rasio Aktivitas adalah rasio untuk mengetahui tingkat efektivitas perusahaan dalam menggunakan aktiva yang dimilikinya, atau rasio ini merupakan alat ukur efisiensi dalam penggunaan sumber daya perusahaan. Ada beberapa jenis rasio aktivitas yaitu:
- a. *Receivable Turnover* adalah rasio digunakan untuk dapat mengetahui lamanya penagihan utang perusahaan dalam satu periode. Lama proses penagihan akan mempengaruhi kinerja perusahaan
  - b. *Days of Receivable* adalah rata-rata penagihan piutang yang dibayarkan dalam satu periode dilihat dari seberapa cepat proses penagihan piutang dapat dilakukan.
  - c. *Inventory Turnover* adalah rasio yang digunakan untuk mengetahui tingkat frekuensi dana yang ditanam pada inventory dalam satu periode tertentu.
  - d. *Working Capital Turnover* adalah rasio untuk menentukan tingkat efektifitas modal kerja yang dimiliki oleh perusahaan dalam satu periode.
  - e. *Fixed Assets Turnover* adalah rasio untuk mengetahui seberapa sering dana yang ditanam dalam aktiva tetap dalam satu periode.
  - f. *Assets Turnover* merupakan rasio untuk mengetahui seberapa sering perputaran semua aktiva yang dimiliki perusahaan.
4. Rasio Profitabilitas adalah rasio untuk mengetahui tingkat kemampuan perusahaan dalam menghasilkan profit. Rasio profitabilitas juga dapat menentukan tingkat ukuran efektifitas perusahaan. Hal ini dapat terlihat dari keuntungan yang dimiliki perusahaan dari penjualan dan investasi. Ada beberapa jenis rasio profitabilitas yang terdiri dari:
- a. *Profit Margin* adalah rasio yang dapat mengetahui ukuran laba atas penjualan perusahaan.
  - b. *Return on Investment (ROI)* atau dikenal juga *Return on Total Assets* adalah rasio yang digunakan untuk memperlihatkan hasil investasi atas jumlah aktiva yang dimiliki perusahaan.
  - c. *Return on Equity* adalah rasio yang digunakan untuk mengetahui tingkat laba bersih sesuai pajak dengan perbandingan modal sendiri yang dimiliki perusahaan.
  - d. *Earning Per Share* atau dikenal dengan rasio nilai buku adalah rasio yang digunakan untuk mengetahui keberhasilan manajemen dalam memberikan keuntungan kepada investor atau pemegang saham.
  - e. *Growth Ratio* adalah rasio yang digunakan untuk mengetahui kemampuan perusahaan dalam mempertahankan kinerja bisnisnya ditengah pertumbuhan perekonomian industri usahanya.
5. Rasio Penilaian (*Valuation Ratio*) adalah rasio yang digunakan untuk memberikan gambaran tentang tingkat kesanggupan perusahaan dalam menentukan nilai pasar usahanya di atas biaya investasi.

Berdasarkan dari penjelasan diatas, maka rasio keuangan menjadi ringkasan item yang sangat krusial sebagai bahan perencanaan dan alat monitoring parameter kesehatan kinerja sebuah perusahaan. Dalam perspektif manajemen, trend, efisiensi, hutang, profitabilitas dan tingkat likuiditas perusahaan dapat dilihat dari rasio-rasio tersebut. Para investor dan analis berpedoman pada rasio fundamental saham tersebut. Rasio keuangan yang kurang baik akan berdampak pada *financing cost* yang terlalu tinggi sedangkan Rasio keuangan yang baik akan memberikan dampak positif terhadap minat investor untuk menanamkan modalnya (Tambunan, 2008). Beberapa rasio yang digunakan pada penelitian ini antara lain:

#### 1. *Price-Earning Ratio (PER)*

Rasio ini menunjukkan tingkat nilai kewajaran harga saham apabila dilihat dari perbandingan harga dan *earning* yang dihasilkan perusahaan. Dapat dikatakan PER menjadi gambaran tentang

besarnya harga pasar saat ini berdasarkan dari pendapatan laba di masa lalu dan perkiraan laba yang akan datang.

2. *Price to Book Value (PBV)*

Rasio ini adalah membandingkan harga saham saat ini dengan nilai buku perusahaan. *Price to Book Value* dapat ini membantu investor dalam menilai harga saham per lembar saham saat ini dengan nilai suatu perusahaan.

3. *Debt-to-Equity Ratio (DER)*

Rasio adalah untuk mengetahui perbandingan antara asset dan hutang yang sedang berlaku dalam perusahaan. Nilai DER yang tinggi akan memberikan stimulant yang buruk bagi investor. Rasio ini juga sering dikatakan sebagai rasio leverage atau rasio pengungkit yang artinya rasio yang digunakan dalam menentukan tingkat ukuran suatu nilai investasi yang ada pada perusahaan.

4. *Return on Equity (ROE)*

*Return On Equity* adalah rasio yang digunakan untuk menentukan tingkat kelayakan perusahaan yang ditanami modal. Dengan kata lain, investor dapat memutuskan yang terbaik dari kebijakan perusahaan tersebut.

### 2.3. *Portfolio Saham*

Portofolio saham adalah merupakan kumpulan aset investasi baik berupa saham atau jenis investasi lainnya yang dimiliki oleh perorangan atau perusahaan. Sebuah portofolio saham memberikan data tentang profil risiko dari investor tersebut. Profil risiko investor yang tinggi atau rendah dapat dilihat dari portofolio saham yang dimiliki oleh investor. Karakter investor yang berbeda-beda mengakibatkan pemilihan saham juga berbeda berdasarkan tingkat profil risiko masing-masing investor. Maka, mengetahui portofolio saham menjadi penting untuk menentukan saham yang sesuai berdasarkan karakter masing-masing pemegang saham.

Dengan mengetahui portofolio saham, investor dapat menentukan jenis saham yang dibutuhkan. Maka, dapat dikatakan jenis saham didapat berdasarkan portofolio saham investor. Investor yang memiliki portofolio dengan risiko tinggi masih dapat mempertimbangkan saham-saham yang termasuk dalam speculative stock, dimana memiliki keuntungan yang cukup tinggi pada masa mendatang walaupun kemungkinan keuntungan tersebut tidak pasti. Sementara investor dengan portofolio saham yang rendah akan memilih saham-saham yang termasuk dalam kategori *Blue chip stocks* dimana saham jenis ini memiliki value terbaik dengan likuiditas tinggi, konsistensi pendapatan, dividen dan *annual growth* yang secara historis terjamin serta lebih resistant terhadap krisis.

### 2.4. *Instrumen Penelitian*

Alat yang digunakan dalam melakukan kegiatan penelitian seperti: pengukuran dalam pengumpulan data, atau lembar observasi disebut dengan instrument penelitian . Instrumen penelitian berdasarkan definisinya adalah suatu alat yang digunakan mengukur fenomena alam maupun sosial ([Sugiyono, 2018](#)). Dalam penelitian ini dapat di gunakan software yang memiliki kapabilitas perhitungan statistic seperti Excel, SPSS, atau yang memiliki fitur lebih *customizable* dan *advance* seperti *Matlab*, *R Language*. Dapat juga kita bagi fungsi dan perannya seperti excel untuk data preparation dan Matlab scripting language untuk melakukan komputasi numeris, visualisasi dan simulasi data.

## 3. *Metode Penelitian*

### 3.1. *Metode*

Metode Analisa klaster (*Cluster Analysis*) merupakan metode yang cukup populer dalam aspek kalkulasi keuangan modern di karenakan kemampuannya untuk mengolah data dalam jumlah besar untuk menghasilkan sebuah informasi yang berharga. Metode ini menggunakan beberapa pendekatan algoritma untuk menemukan pola tersembunyi dalam suatu group dataset, sehingga dapat ditemukan aspek *similarity* yang tinggi dalam satu lingkup *cluster* dan *gap dissimilarity* antar *cluster*. Tujuan utama dari mekanisme clustering ini adalah pengelompokan sejumlah data, sehingga dalam tiap clusternya terdapat data-data yang dapat dikatakan memiliki karakteristik persamaan yang tinggi.

Ada beberapa pendekatan yang terlibat dalam melakukan metode clustering. Dalam *clustering* ada dua pendekatan utama yang sering digunakan: pendekatan partisi dan pendekatan hirarki. *Partition-based clustering* atau dikenal dengan pendekatan partisi atau adalah mengelompokkan data dengan cara memilah-milah data yang dianalisa ke dalam *cluster* yang ada. Sementara pendekatan yang lain adalah clustering dengan pendekatan hirarki adalah teknik pengelompokan data dengan membuat suatu hirarki berdasarkan dendrogram. Dalam teknik ini data yang memiliki kesamaan yang tinggi akan ditempatkan pada hirarki yang berdekatan, sementara data yang memiliki kesamaan rendah akan ditempatkan pada tempat yang berjauhan.

Pendekatan metode clustering memiliki beberapa algoritma yang populer antara lain adalah:

1. Algoritma *Cluster Partitioning (k-means)*, dimana tiap partisi *cluster* akan ditentukan oleh *dataset object* dan area tengahnya (*central*). kmean menghitung centroid cluster secara berbeda untuk tiap *distance-metric* nya
2. Algoritma *cluster* hierarki, pengelompokan data melalui skala yang bervariasi untuk di implementasikan ke dalam pohon *cluster* atau *dendrogram*.
3. *Gaussian mixture models*: permodelan cluster sebagai campuran dari komponen normal density multivariate
4. *Self-organizing maps*: menggunakan *neural networks* untuk mempelajari topologi dan distribusi data Namun, dua metode terbesar dan yang paling banyak diaplikasikan adalah *partitioning cluster (k-means)*, dan *hierarchical cluster* ([Santosa, 2007](#)).

### 3.2. Hierarchical Clustering

Berdasarkan penjelasan diatas, ada dua pendekatan *clustering* yang dapat digunakan: *partition-based clustering* dan *Hierarchical Clustering*. Dimana pada *Hierarchical Clustering* dilakukan pembangunan hirarki pengelompokan data yang terlibat. Strategi pengelompokan yang dilakukan memiliki 2 jenis yaitu *Agglomerative (Bottom-Up)* dan *Devisive (Top-Down)*. Berikut beberapa langkah yang digunakan dalam *Algoritma Agglomerative Hierarchical Clustering*:

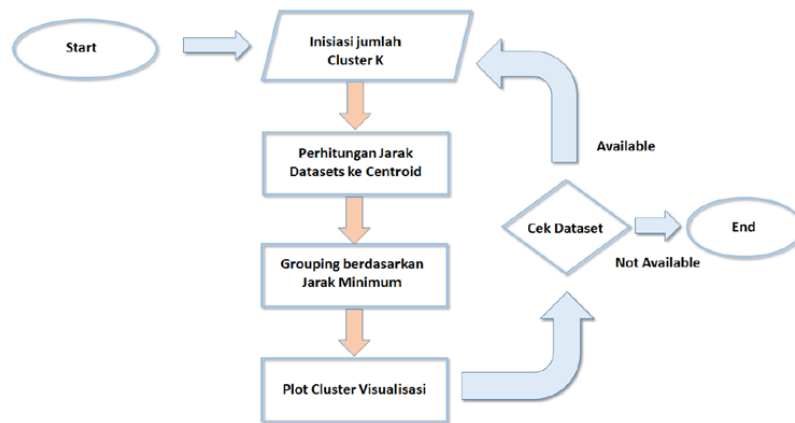
1. Menghitung total jarak tiap data.
2. Menggabungkan dua kelompok data yang memiliki jarak terdekat yang ditentukan berdasarkan jarak kedekatan yang telah ditentukan.
3. Memperbarui matrik jarak antar data dengan tujuan untuk menunjukkan jarak anantara kelompok yang baru dengan kelompok yang masih ada.
4. Mengulangi proses sebelumnya hingga hanya menyisakan satu kelompok saja.

Dalam *Hierarchical clustering*, data tidak di partisi atau di bagi dalam satu tahap saja, tetapi melalui beberapa tahapan, yang terpecah dari *single cluster*, terpartisi menjadi beberapa *sub cluster*. Metode pembagian *hierarchical clustering* di presentasikan dalam 2 dimensi diagram yang di sebut dengan dendrogram.

### 3.3. Partitioning Cluster

Analisis *Clustering* adalah sebuah aktifitas yang dilakukan untuk mengelompokkan data ke dalam beberapa kategor berdasarkan dari tingkat kesamaannya. Data yang ada dalam satu kelompok memiliki kesamaan yang tinggi dilihat dari beberapa kriteria tertentu. Sementara *Partitional clustering* adalah data yang dikelompokkan ke dalam beberapa *cluster* dengan mengabaikan *hierarcial structural* antar satu data dengan data lainnya. Metode *partitional clustering* yang dilakukan memiliki beberapa *cluster*, dimana *cluster* tersebut memiliki titik pusat *cluster* (centroid). Dalam metode ini, *partitional clustering* memiliki tujuan untuk mendekati jarak (*dissimilarity*) dari seluruh data ke pusat *cluster* masing-masing. Contoh metode *partitional clustering*: *K-Means*, *Fuzzy K-means* dan *Mixture Modelling*. *Partitioning cluster* menggunakan algoritma *k-means* yang cukup sederhana dan relatif mudah di implementasikan dalam menyelesaikan berbagai macam permasalahan *clustering*. Metode ini merupakan partisi sebuah *cluster* yang menggunakan rata-rata jarak terdekat.

Flow diagram dibawah dapat memberikan gambaran terkait proses perhitungan *k means cluster*.



Gambar 1. Flow Diagram Algoritma

Berdasarkan *flow* diagram diatas, penentuan jumlah *cluster* dilakukan pada tahap awal. Kemudian dilakukan perhitungan antara jarak data sets ke *centroid*, dimana dalam hal ini adalah ROE, DER, PER, kepada PBV. PBV menjadi pusat *clustering*, sehingga penentuan jarak dilakukan antara variabel ROE, DER dan PER terhadap titik pusat clustering yaitu PBV. Setelah jarak sudah ditentukan, maka data akan diklasifikasikan berdasarkan dari jarak minimum. Jarak minimum mencerminkan bahwa data memiliki karakteristik yang sama. Proses selanjutnya adalah *plot cluster* visualisasi, dimana akan tampil hasil dari kalifikasi yang sudah ditentukan pada proses sebelumnya.

#### 4. Hasil dan Pembahasan

##### 4.1. Algoritma Kluster Hirarki

Pengukuran *similarity* tiap objectnya untuk dikelompokkan ke dalam *cluster* tertentu dapat menggunakan rumusan persamaan sebagai berikut:

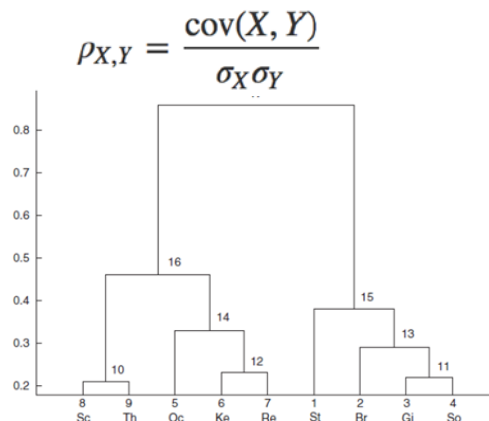
1. Cosinus antara dua titik

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$

2. Covarian

$$\text{Cov}(X, Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{N}$$

3. Koefisien Korelasi



Gambar 2. Dendrogram Algoritma Hirarki



Pembagian *hierarchical clustering* di presentasikan dalam dendrogram. Sumbu y menunjukkan nomor objek sedangkan sumbu X menunjukkan jarak antar objek / *cluster*. Untuk mengukur tingkat akurasi hasil *clustering* dapat digunakan rumusan yang disebut koefisien korelasi *cophentic* dan menghasilkan persamaan sebagai berikut karena parameter ini mengukur korelasi jarak yang dihitung selama proses pembentukan dendrogram:

$$r_{\text{coph}} = \frac{\sum_{i=1}^{n-1} \sum_{j>i}^n (c_{ij} - \bar{c})(d_{ij} - \bar{d})}{\left( \sum_{i=1}^{n-1} \sum_{j>i}^n (c_{ij} - \bar{c})^2 \right)^{1/2} \left( \sum_{i=1}^{n-1} \sum_{j>i}^n (d_{ij} - \bar{d})^2 \right)^{1/2}}$$

#### 4.2. Algoritma Cluster Partitioning (k-means)

*Partitioning cluster* menggunakan algoritma *k-means* yang cukup sederhana dan relatif mudah di implementasikan dalam menyelesaikan berbagai macam permasalahan *clustering*. Pengelompokan data dengan karakteristik yang sama antar data dilakukan pada tahap ini. Pengelompokan dilakukan berdasarkan dari perhitungan jarak terdekat ke titik pusat *cluster*. Maka dapat dikatakan, metode ini merupakan partisi sebuah cluster yang menggunakan rata-rata jarak terdekat. Berikut rumusan Euclidian distance yang digunakan:

$$\arg \min_s \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2$$

*Means* dianggap sebagai *centroid* dari *cluster* dari data set yang di observasi dalam dataset. Hasil perhitungan di tempatkan dalam *cluster* yang memiliki jarak mean terkecil. Secara garis besar tahapan perhitungannya adalah sebagai berikut:

1. Inisiasi *k means* untuk *cluster* k
2. Perhitungan dari *dissimilarity* antar *object* dataset dan *mean* dari *cluster*
3. Penempatan dan alokasi sebuah *object* ke dalam *cluster* dari pehitungan terdekat poin nomor 2
4. Re-kalkulasi mean *cluster* dari *object* dataset yang masih ada

#### 4.3. Cara Kerja Algoritma Klaster

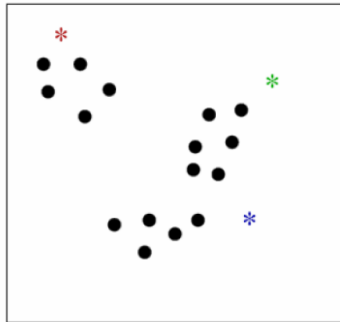
Berdasarkan cara kerja algoritma *kmeans cluster* dalam proses yang telah di jelaskan sebelumnya, akan di visualisasikan pada gambar berikut ([Santoso, 2007](#)):

1. Penentuan jumlah *cluster* k

Penentuan jumlah cluster ditentukan dari berapa banyak karakteristik yang akan dibedakan dalam data yang ada. Dalam hal ini jumlah cluster terdiri dari 3 yaitu meliputi: *green zone*, *blue zone*, dan *red zone*. *Green zone* adalah untuk jenis saham yang memiliki kinerja keuangan yang baik, terlihat dari data PER, PBV, DER, dan ROE yang proporsional pada perusahaan. Sementara *red zone* adalah kategori saham yang memiliki kinerja keuangan kurang baik, berdasarkan dari data PER, PBV, DER, dan ROE. *Blue zone* adalah berada diantara *green zone* dan *red zone*, yang merupakan klasifikasi yang memiliki kinerja keuangan yang biasa saja berdasarkan dari variabel yang digunakan pada penelitian ini.

2. Inisialisasi centorid k-means

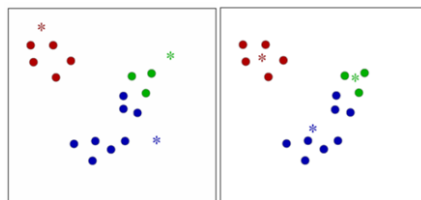
Dalam hal ini inisialisasi *centorid k-means* dapat dilakukan dengan berbagai metode, salah satu metode yang sangat umum dilakukan adalah dengan teknik insiasi random. Pada gambar 2 dibawah dapat terlihat bagaimana inisiasi *central cluster* di tentukan untuk kemudian dilakukan pengukuran jarak sebagai penentuan lokasi cluster dari tiap *object*. Pengukuran jarak ini dilakukan untuk menentukan jarak ke pusat *cluster*. Hasil dari pengukuran ini akan mempengaruhi klasifikasi *cluster* yang akan ditentukan pada tahap selanjutnya.



Gambar 3. Tahap awal K-Means

### 3. Proses Iterasi

Terdapat proses iterasi dalam penentuan pusat *cluster* (gambar 3) menggunakan *mean* atau median rata-rata tiap *object* di *cluster*. Proses iterasi ini dilakukan secara berulang-ulang pada program. Perhitungan pengulangan yang terjadi dalam proses iterasi adalah berdasarkan error. Proses pengulangan ini akan terus dilakukan sampai error kecil, yang dapat diartikan bahwa titik variabel solusi sudah mendekati satu titik (*Konvergen*)



Gambar 4. Iterasi penentuan pusat cluster dan pembentukan cluster

### 4. Kesimpulan Iterasi

Dalam proses iterasi dilakukan sampai nilai pusat cluster tidak berubah. Jika pusat *cluster* masih berubah, maka proses iterasi terus dilakukan. Iterasi dapat dihentikan jika hasil *clustering* memiliki hasil yang sama dengan proses iterasi sebelumnya dan dapat dikonfirmasi klasifikasi hasil cluster menggunakan proses iterasi yang tidak memiliki perubahan.

#### 4.4. Sum of squared-error (SSE)

$$SSE = \sum_{k=1}^K \sum_{\forall x_i \in C_k} \|x_i - \mu_k\|^2$$

SSE adalah perhitungan statistik awal yang dipakai untuk menghitung nilai lain. *Error* pada SSE adalah perbedaan dari observed value (*dependent variable y*) dengan dan *predicted value*. Sehingga pada SSE menunjukkan total kesalahan kuadrat yang terjadi bila data  $x_1, x_2, x_3 \dots$  dst di kelompokkan ke dalam *Cluster K* dengan pusat *cluster*  $\mu_k$ . Semakin kecil nilai SSE maka semakin bagus hasil *clustering* yang sudah di tampilkan.

#### 4.5. Dataset Summary Statistic Results

No	Company Name	Asset	Liabilities	Equity	Sales	EBT	Profit for the Period	Profit Period attr. to Owner's Entit	Annual EPS	Book Value	PER	PBV	DER
1	Bisi International Tbk. [S]	2416	353	2064	1852	454	336	336.0	112.00	688.0	16.96	2.76	0.17
2	Astra Agro Lestari Tbk. [S]	24226	6633	17593	14121	2209	2114	2007.0	1043.00	9141.0	16.09	1.84	0.38
3	Austindo Nusantara Jaya Tbk. [S]	7055	2290	4765	1806	260	124	124.0	37.00	1421.0	53.99	1.40	0.48
4	Eagle High Plantation Tbk.	16254	9995	6259	2542	-336	-391	-390.0	-12.00	199.0	-22.16	1.38	1.60
5	Dharma Satya Nusantara Tbk.	8183	5479	2704	3942	337	252	251.0	24.00	255.0	23.25	2.16	2.03
6	Golden Plantation Tbk. [S]	2376	1402	974	177	15	8	0.8	0.22	266.0	548.56	0.45	1.44
7	Gozco Plantations Tbk.	3547	2403	1145	545	-1594	-1548	-1536.0	-256.00	191.0	-0.29	0.39	2.10
8	Jaya Agra Wattie Tbk.	3291	2241	1050	590	-223	-225	-223.0	-59.00	278.0	-2.15	0.46	2.13
9	PP London Sumatra Indonesia Tbk. [S]	9459	1813	7646	3848	779	593	594.0	87.00	1121.0	19.99	1.55	0.24
10	Multi Agro Gemilang Plantation Tbk. [S]	1151	357	794	1	-10	-10	-10.0	-1.00	88.0	-34.19	0.57	0.45
11	Provident Agro Tbk. [S]	3861	1534	2326	1170	251	219	219.0	31.00	327.0	14.81	1.40	0.66
12	Sampoerna Agro Tbk. [S]	8328	4570	3759	2915	267	459	442.0	234.00	1989.0	8.17	0.96	1.22
13	Salim Ivomas Pratama Tbk. [S]	32538	14919	17618	14531	1395	610	538.0	34.00	1114.0	14.51	0.44	0.85
14	SMART Tbk.	26141	15942	10199	29752	1431	2600	2601.0	906.00	3551.0	4.80	1.22	1.56
15	Sawit Sumbermas Sarana Tbk. [S]	7163	3709	3454	2723	847	592	592.0	62.00	363.0	22.54	3.86	1.07
16	Tunas Baru Lampung Tbk.	12597	9176	3421	6514	803	621	615.0	115.00	640.0	8.59	1.55	2.68
17	Bakrie Sumatera Plantations Tbk.	14700	13503	1198	1565	-321	-485	-483.0	-35.00	87.0	-1.42	0.57	11.27
18	Central Proteina Prima Tbk.	7323	7142	181	8597	-1725	-1997	-1988.0	-49.00	4.0	-1.02	11.19	39.49
19	Dharma Samudera Fishing Industries Tbk. [S]	329	180	149	604	8	6	6.0	3.00	80.0	51.02	1.97	1.21
20	Inti Agri Resources Tbk. [S]	365	85	280	84	-32	-28	-27.0	-8.00	83.0	-307.63	30.09	0.30
21	Bumi Teknokultura Unggul Tbk. [S]	4880	3369	1511	748	2	2	2.0	0.29	261.0	4311.50	4.71	2.23
23	Adaro Energy Tbk. [S]	87633	36766	50867	33916	7343	4577	4495.0	141.00	1590.0	12.06	1.07	0.72
24	Atlas Resources Tbk [S]	4435	3679	756	156	-260	-342	-328.0	-109.00	252.0	-4.75	2.06	4.87
25	Bara Jaya Internasional Tbk. [S]	1586	849	737	10	-292	-288	-288.0	-50.00	128.0	-3.88	1.52	1.15

Gambar 5. Visualisasi hasil dari teknik *clustering*

Terlihat bahwa masing – masing saham data mining diklasifikasikan menjadi prediktif dan deskriptif dan di kategorikan dengan perspektif. Data dari rasio keuangan yang digunakan dalam penelitian ini menjadi ringkasan item yang sangat krusial sebagai bahan perencanaan dan alat monitoring parameter kesehatan kinerja sebuah perusahaan dalam berinvestasi. Rasio keuangan yang digunakan PBV, PER, DER, dan ROE merupakan jenis rasio yang selalu dijadikan pertimbangan oleh investor dalam berinvestasi. Sehingga dengan adanya rekomendasi dari hasil clustering diharapkan memberikan *insight* yang berbeda baik bagi investor atau calon investor sebelum berinvestasi pada pasar modal, dalam hal ini saham yang ada terdaftar di Bursa Efek Indonesia.

#### 4.6. Signifikansi

Sesuai dari apa yang telah di tuliskan pada uraian sebelumnya bahwa beberapa manfaat signifikan yang dapat di hasilkan dari penelitian ini, antara lain adalah:

1. Memberikan alternatif bentuk pengkategorian saham dari yang ada sekarang berdasarkan dari data kinerja keuangan yang digunakan yaitu PBV, PER, DER, dan ROE.
2. Pembuatan sistem pendukung keputusan investasi berbasis Analisa *Cluster* kepada inverstor sehingga memberikan potensi keuntungan bagi investor dalam berinvestasi.
3. Pencarian korelasi dan pattern antar rasio-rasio dari laporan keuangan. Dimana kolerasi dan *pattern* belum pernah di lakukan penelitian sebelumnya.
4. Sebagai *tools* pelengkap *Risk Management* Investasi. Dengan mengetahui jenis saham yang sudah terkelompok berdasarkan 3 jenis *clustering*: *green zone*, *blue zone*, dan *red zone* maka memberikan sinyal kepada investor terkait jenis saham yang memiliki potensi keuntungan atau kerugian di masa yang akan datang.
5. Peletakan pondasi untuk pengembangan penelitian berbasis analisa *cluster* dan *data mining* selanjutnya, sehingga penelitian analisa *cluster* dapat diperluas tidak hanya pada data kinerja keuangan namun jenis investasi lain baik pada pasar modal atau pasar uang.

### 5. Kesimpulan

Teknik *cluster* merupakan algoritma *data mining* yang sudah sejak lama digunakan dalam berbagai bidang keilmuan. Penelitian ini mencoba menemukan bentuk korelasi dan *pattern* untuk menemukan bentuk pengkategorian baru dengan laporan keuangan sebagai basis datanya yang diharapkan memiliki manfaat sebagai pelengkap sistem pendukung keputusan dan manajemen resiko dalam bidang investasi dan keuangan. Pada jenis saham yang terdaftar di Bursa Efek Indonesia setelah dilakukan

*clustering* dengan 3 jenis *k* yaitu: *green zone*, *blue zone*, dan *red zone*. Sehingga berdasarkan pengelompokan *clustering* tersebut memberikan rekomendasi kepada investor atau calon investor dalam berinvestasi dalam pasar modal khususnya pasar saham yang ada di Bursa Efek Indonesia.

### **Limitasi dan Studi Lanjutan**

Dalam penelitian hanya menggunakan *software* yang memiliki kapabilitas perhitungan statistic seperti *Excel* dan penggunaan dua dari beberapa pendekatan algoritmik metode pengelompokan untuk menemukan pola tersembunyi dalam sekelompok set data yaitu *cluster* partisi (*k-means*) yang ditentukan oleh objek set data dan area pusatnya, dan klaster hierarkis yang mengelompokkan data melalui berbagai skala untuk diimplementasikan ke dalam pohon *cluster* atau dendrogram.

Dapat menggunakan *software* yang memiliki kapabilitas perhitungan statistic yang memiliki fitur lebih *customizable* dan *advance* seperti *Matlab*, *R Language* agar *scripting language* untuk melakukan komputasi numeris, visualisasi dan simulasi data dapat lebih optimal. Tindak lanjut pengembangan penelitian ini dapat dipadukan dengan algoritma lainnya seperti teknik *forecasting*, *optimization* dan *classification* bahkan algoritma *cluster* yang lain.

### **Ucapan terima kasih**

Dalam proses pembuatan penelitian ini tidak terlepas dari bantuan berbagai pihak yang terlibat secara langsung ataupun tidak langsung. Maka dalam hal ini, penulis akan menyampaikan ucapan terimakasih banyak kepada pihak yang telah memberikan bantuan dalam penyelesaian penelitian baik secara langsung atau tidak langsung dalam menyelesaikan penelitian ini. Yaitu kepada beberapa pihak yang saya sebutkan: partner penulis yaitu Ario Menak Sanoyo yang sudah berkontribusi dalam melahirkan penelitian ini; Keluarga peneliti yang telah membantu dalam penyelesaian penelitian ini yaitu dengan waktu dan dukungan yang diberikan sehingga penelitian ini dapat diselesaikan dengan baik; dan teman sejawat penulis tidak dapat penulis sebutkan satu per satu yang sudah memberikan kontribusi pemikiran melalui diskusi yang dilakukan.

### **Referensi**

- Aleksandar Vasilev. (2020). How Quantitatively Important is Public Investment for Both Business Cycle Fluctuations and Output Growth in Bulgaria (1999–2018)?. *Macroeconomics and Finance in Emerging Market Economies*.
- Bini B. S, Tessa Mathew. (2016). Clustering and Regression Techniques for Stock Prediction. *Procedia Technology* 24, 1248 – 1255.
- Dragan Radojevic, Boris Delibašić, Ksenija Mandić. (2015). An application of the Integrated IBA-TOPSIS Model in supplier selection. *International Journal of Decision Support System Technology* 7(1), 15-30.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5 Ed.). Hoboken, NJ, United States: John Wiley & Sons.
- Hery. (2015). *Analisis Laporan Keuangan*. Edisi 1. Yogyakarta: Center For Academic Publishing Services.
- Hery. (2017). *Auditing dan Asurans*. Jakarta. Grasindo
- Irham Fahmi. (2018). *Pengantar Manajemen Keuangan*. Bandung. Alfabeta.
- Kasmir. (2016). *Analisis Laporan Keuangan*. Jakarta: Raja Grafindo Persada.
- Munawir, S. (2004). *Analisa Laporan Keuangan*. Edisi IV, Liberty, Yogyakarta
- M. Sholihun (2019). Pengaruh Likuiditas terhadap Nilai Perusahaan melalui Struktur Modal pada Perusahaan Manufaktur Indonesia. *Jurnal Mu'allim*, 1(2).  
E-ISSN: 2655-8912 P-ISSN: 2655-8939
- Santosa, B. (2007). *Data Mining Terapan dengan Matlab*. Salemba Empat. hal 8. 2. Indonesia Graha Ilmu.

- Santosa, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Indonesia Graha Ilmu.
- Sugiyono. (2018). *Metode Penelitian Kuantitatif*. Bandung. Alfabeta.
- Tambunan, A. P. (2008). *Menilai Harga Wajar Saham (Stock Valuation)*. Jakarta Elex Media Komputindo.