

PENERAPAN ALGORITMA K-MEANS UNTUK MENGANALISIS DATA PENJUALAN PADA TOKO AYU COLLECTION BERBASIS WEB

APPLICATION OF K-MEANS ALGORITHM TO ANALYZE SALES DATA AT WEB-BASED AYU COLLECTION STORE

Wahyu Tities Pambudi¹⁾ Arita Witanti²⁾

Email: ¹⁾wahyutities100@gmail.com, ²⁾arita@mercubuana-yogya.ac.id

^{1,2} Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Mercu Buana

Abstrak

Toko Ayu Collection menjual berbagai jenis pakaian dan aksesoris. Salah satu cara untuk menjaga kepuasan pelanggan adalah dengan menjaga stok barang agar tidak ada barang yang kosong. Penjual harus menganalisis data barang mana yang laris dan data barang mana yang tidak laku, berdasarkan data laporan penjualan. Masalah ini dapat diselesaikan dengan menggunakan salah satu teknik dalam data mining yaitu algoritma K-Means Clustering. Penelitian ini dimaksudkan untuk membantu Ayu Collection, sebuah toko di Kota Blora yang menjual pakaian dan aksesoris, mengklasifikasikan data penjualannya untuk memaksimalkan pengelolaan stoknya. Variabel yang digunakan adalah nama barang, data barang masuk, data barang keluar, dan stok barang. Pemilik toko bisa melihat hasil pengelompokan baju dan aksesoris yang laris manis dan kurang laris. Oleh karena itu, jika ada produk rendah penjualannya, pemilik toko dapat mencari alternatif lain agar pakaian dan aksesoris yang kurang laku dapat dijual. Metode yang digunakan dalam pengumpulan data adalah observasi dan wawancara dengan pemilik toko Ayu Collection dengan tingkat Akurasi sistem mencapai 83,33%

Kata Kunci : Analisis Penjualan, K-Means, Website, Data Mining

Abstract

The Ayu Collection store sells various types of clothing and accessories. One way to maintain the customers' satisfaction is by keeping the stock of goods so that no items are empty. The seller must analyze which item data are selling well and which item data is not selling well, based on the sales report data. This problem can be solved by using one of the techniques in data mining, namely the K-Means Clustering algorithm. This research was intended to help the Ayu Collection, a shop in Blora City that sells clothing and accessories, classify its sales data to maximize its stock management. The variables used were the name of the goods, the data of incoming goods, the data of outgoing goods, and the stock of goods. The shop owner can see the results of grouping clothes and accessories that are best-selling and not selling well. Therefore, if there are products that are not selling well, the shop owner can look for other alternatives so that clothes and accessories that are not selling well can be sold. The methods used in the data collection were observation and interviews with the shop owner of the Ayu Collection store. The accuracy of this system reached 83.33%

Keywords : Selling Analysis, K-Means, Website, Data Mining

1. PENDAHULUAN

Teknik Informatika merupakan ilmu yang mempelajari Teknologi Informasi dimana pengimplementasiannya mengarah kepada perkembangan teknologi pada saat ini. Pesatnya kemajuan teknologi memberikan kemudahan salah satunya adalah dalam mendapatkan informasi yang dibutuhkan untuk membantu dalam proses pengambilan keputusan[1]. Informasi diperoleh akibat dari adanya komunikasi, jadi dapat disimpulkan

bahwa adanya informasi didahului dengan adanya komunikasi antara dua individu atau kelompok[2] Komunikasi dan informasi dibutuhkan untuk kelangsungan suatu produksi dari perusahaan[3]. Data dan informasi dibutuhkan untuk membantu penjualan dan promosi pada suatu perusahaan atau toko untuk mencapai suatu tujuan tertentu[4]. Penjualan merupakan kegiatan utama perusahaan dalam menghasilkan pendapatan dan keuntungan. Salah satu cara penjualan dan pemasaran adalah dengan menggunakan sistem berbasis web untuk meningkatkan jumlah keuntungan dimana kondisi tersebut membuat produk menjadi lebih kompleks sehingga memberikan fungsi yang besar untuk membuat pasar yang lebih baik[5].

Dalam pemasaran berbasis web, menciptakan interaktif jenis penjualan yang berbeda dan signifikan dari cara pemasaran tradisional. Penjualan dengan berbasis web membentuk persepsi pembeli tentang produk yang akan dibeli, dan bagaimana cara menganalisis barang dan produk mana saja yang paling laris dipasarkan sehingga toko bisa menyiapkan stok lebih banyak barang yang dibutuhkan

Permasalahan operasional yang sering dihadapi oleh sebuah perusahaan adalah ketersediaan barang, sama halnya dengan Ayu Collection yang merupakan salah satu toko yang menjual berbagai produk busana dan aksesoris untuk semua kalangan dimana toko ini setiap harinya harus memenuhi kebutuhan konsumen dengan stok yang berbeda-beda. Untuk dapat melakukan hal tersebut, maka dibutuhkan sebuah sistem untuk menganalisis kebutuhan konsumen yang paling banyak dicari. Banyak cara yang dapat digunakan untuk menganalisis hal tersebut yaitu salah satunya dengan Metode K-Means. Metode ini dipilih karena memiliki kelebihan diantaranya mudah diimplementasikan selain itu metode ini bersifat fleksibel sehingga pengguna dapat menentukan jumlah cluster yang akan dibuat dengan relatif cepat[6]

2. TINJAUAN PUSTAKA

E-Commerce merupakan suatu website untuk menjual berbagai kebutuhan secara online tanpa harus datang ketempat. Pertumbuhan pesat sektor E-commerce di Indonesia didukung dengan adanya smartphone dan internet yang terus meningkat. Apalagi pada masa pandemi seperti sekarang yang mengharuskan untuk dirumah saja membuat situs e-commerce dibanjiri oleh para customer yang tidak bisa berbelanja ke luar rumah. Selain itu jumlah penduduk Indonesia yang besar dan daya beli yang terus meningkat serta populasi masyarakat yang selalu melek akan teknologi sehingga mereka cepat menyesuaikan diri dengan teknologi baru[7]. E-commerce mencakupi segala proses termasuk pemasaran, penjualan, dan pengiriman serta memperluas jaringan mitra bisnis.

E-Commerce disebut juga sebagai wajah baru dalam dunia bisnis yang menghasilkan manajemen konsumen dan analisis yang baik serta jangkauan komoditi dan operasional yang lebih efisien[8].

Penelitian mengenai penerapan algoritma K-Means sudah banyak dilakukan oleh beberapa penulis yaitu dalam penelitian [9] menggunakan K-Means untuk mengetahui tingkat persebaran kasus COVID-19 yang dikategorikan tinggi, sedang, dan rendah pada masing-masing provinsi yang tersebar di seluruh Indonesia, selain itu penelitian juga [10] membahas mengenai K-Means yang diimplementasikan untuk digunakan sebagai bahan dasar untuk menentukan strategi pemasaran yang paling tepat. Model K-Means juga dapat digunakan untuk melakukan pengelompokan mahasiswa berdasarkan minat belajarnya seperti yang dilakukan oleh [11]. K-Means digunakan pada penelitian ini [12] adalah untuk menentukan interpretasi dan karakteristik dari setiap cluster yang terbentuk pada data obligasi. Penelitian selanjutnya dilakukan oleh [13] yang meneliti tentang status gizi balita dengan hasil penelitian dapat disimpulkan bahwa algoritma K-Means memiliki nilai akurasi sebesar 34% benar.

Berdasarkan fenomena yang terjadi maka penulis melakukan penelitian dengan membedakan study kasus dimana K-Means akan diterapkan pada Toko Ayu Collection dimana toko ini menjual berbagai aksesoris dan perlengkapan baju untuk semua

kalangan dan hasil dari penelitian ini diharapkan pemilik toko dapat memanajemen stok tokonya berdasarkan tinggi rendahnya suatu penjualan pada suatu produk tertentu.

3. METODOLOGI PENELITIAN

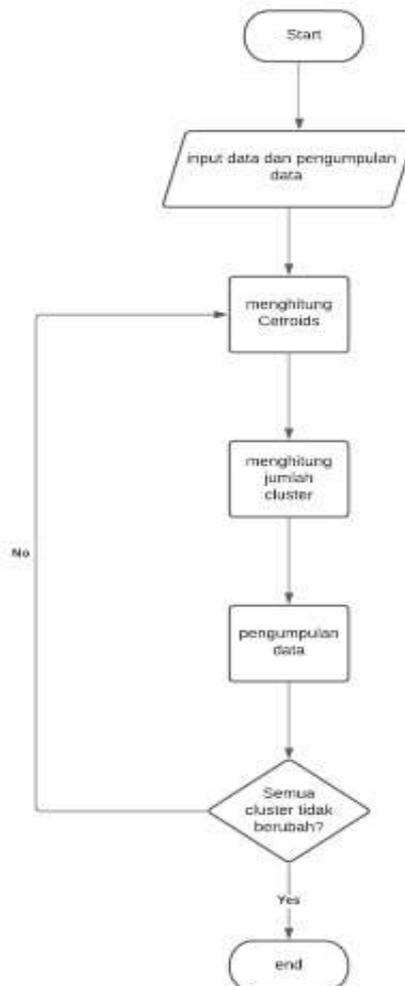
3.1 Pengumpulan Data

Pada penelitian ini dilakukan beberapa tahap pengumpulan data yaitu dengan wawancara atau dengan cara tanya jawab secara langsung dengan pemilik toko[14]. Observasi dilakukan dengan mengumpulkan data secara langsung ke objek yang diteliti[15]. Selanjutnya untuk mendapatkan data teoritis dilakukan dengan studi pustaka dimana penulis mengumpulkan data dari buku, journal dan referensi yang terakait dengan penelitian[16].

3.2 Pre-Processing Data

Setelah semua data terkumpul, lalu di lakukan preprocessing agar data bisa di olah untuk dilanjutkan ke proses selanjutnya pada sistem yang akan di buat. Pada kasus ini data yang terkumpul pada toko Ayu Collaction yang memiliki masing-masing data yang berbeda pada setiap item barang. Integrasi Data yang diperoleh adalah data dari tahun 2017 sampai dengan 2018 yang digabungkan sehingga menjadi satu data yang sudah terintegrasi salanjutnya dilakukan Cleaning data untuk pembersihan data-data yang kosong dan memastikan data tersebut relevan satu dengan yang lainnya untuk menghasilkan informasi yang diinginkan.

3.3 Membuat rancangan sistem



Gambar 1. Flowchart sistem

Dalam membuat rancangan sistem bisa dilihat pada gambar 1 dimana terdapat tahapan-tahapan yang dimulai dari pengumpulan dan inputan data barang yang tinggi penjualannya dan yang tidak, setelah itu akan dihitung rata-ratanya dan masing-masing akan dimasukkan ke dalam label, kemudian data akan dialokasikan ke dalam centroid dan dicari dengan nilai rata-rata terdekat.

3.4 Perhitungan K-Means

K-Means merupakan merupakan salah satu teknik pengelompokan yang bekerja berdasarkan partitioned clustering[17]. Prinsip kerja dari pengelompokan hierarchical clustering dilakukan secara bertahap. Dan disetiap iterasi dari pengelompokan hierarchical clustering hanya ada satu pemilihan penggabungan suatu item terhadap item lainnya. K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengelompokan data besar dan outlier dengan sangat cepat[18].

Dalam metode K-Means setiap data harus termasuk ke cluster tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke cluster yang lain. Algoritma K-Means adalah metode non-hierarki yaitu mengambil sebagian banyaknya komponen populasi pada tahapan awal untuk dijadikan pusat cluster pertama. Pada tahap ini pusat cluster dipilih secara acak dari sekumpulan populasi data. K-Means menguji masing-masing komponen di dalam populasi data menandai komponen tersebut ke dalam salah satu pusat cluster. Posisi pusat cluster akan di hitung kembali sampai semua komponen data digolongkan kedalam tiap-tiap pusat cluster dan terakhir akan terbentuk posisi pusat cluster yang baru.

Algoritma juga mudah untuk diimplementasikan dan memiliki kompleksitas waktu dan ruang yang relatif kecil. Algoritma ini juga merupakan algoritma yang cukup efisien dalam komputasinya dan memberika hasil yang cukup baik dan memuaskan jika clusternya compact, hyperspherical dalam shape dan mampu memisahkan fitur-fitur ruangnya dengan baik. Berdasarkan[19] Langkah-langkah dasar untuk algoritma K-Means adalah menentukan nilai k cluster sesuai yang diinginkan. Pilih titik-titik atau sampel yang menjadi anggota cluster secara acak. Tentukan nilai centroid atau titik tengah dari cluster tersebut dengan rumus:

$$M_k = \left(\frac{1}{n_k} \right) \sum_{i=1}^{n_k} x_{ik} \dots \dots \dots (1)$$

Di mana x_{ik} = sampel yang ada yang ada dalam cluster C_k , kemudian Hitung square error untuk tiap cluster c_k yang merupakan jumlah kuadrat dari jarak Euclidean antara tiap sampel dalam c_k dan titik tengahnya(centroid). Error ini dikenal juga dengan nama within chister variation (WCV), yaitu :

$$e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - m_k)^2 \dots \dots \dots (2)$$

Selanjutnya jumlah dari keseluruhan error dari k-cluster juga dihitung dengan rumus

$$E_K^2 = \sum_{k=1}^k e_k \dots \dots \dots (3)$$

Kelompokkan kembali semua sampel berdasarkan jarak minimum dari masing-masing pusat M_1, M_2, \dots, M_k Sehingga diperoleh distribusi baru dari sampel sesuai clusternya. Untuk memperoleh distribusi sampel baru tersebut dapat dilakukan dengan menghitung jarak di (M_k, x_k) . Perhitungan jarak dari masing-masing titik tersebut dapat menggunakan beberapa metode. Kemudian hasil dari anggota cluster baru sesuai dengan hasil yang diperoleh. Ulangi langkah ke-3 sampai beberapa iterasi sehingga nantinya ditemukan nilai total square error turun secara signifikan.

4 HASIL DAN PEMBAHASAN

Pengujian dilakukan menggunakan beberapa cluster yang berbeda untuk mendapatkan jumlah cluster yang paling optimal dalam penjualan di toko Ayu Collections yang akan di implementasikan pada proses selanjutnya dengan jumlah data sebanyak 2350 yang mana data tersebut merupakan data penjualan produk pada satu periode penjualan yaitu bulan Agustus 2017 sampai dengan Mei 2018. Dimana sistem akan mengelompokkan produk yang tinggi penjualan dengan yang tidak. Pada gambar 2. terdapat tampilan dashboard yang terdiri dari beberapa menu yaitu supplier, customer, produk, biaya, sales, karyawan, komentar, dan laporan



Gambar 2. Halaman Dashboard

Pada gambar 3. merupakan hasil data barang yang tinggi penjualannya untuk satuan minggu

Product data

id	kode_barang	tanggal	harga	stock	jumlah	
0	34838	GANG	2017-09	5000	45	1
1	34839	GIS	2017-09	175000	47	1
2	34842	MKGR	2017-09	25000	48	1
3	34843	SWG	2017-09	10000	40	1
4	34844	ELUK	2017-09	6000	35	1
5	34852	USTIK	2017-09	20000	50	1
6	34855	BSS	2017-09	3500	60	1
7	34857	KCIRAN	2017-09	4000	52	1
8	34862	LEKAN	2017-09	2500	47	1
9	34863	RS	2017-09	133000	62	1
10	34864	RDC	2017-09	6750	55	1
11	34865	LEKAN	2017-09	14000	50	1
12	34867	ESHAGAW	2017-09	35000	37	1
13	34870	GIS	2017-09	175000	48	1
14	34879	KUNJ	2017-09	22000	70	1
15	34880	GANG	2017-09	5000	68	1
16	34882	PFUN	2017-09	10000	45	1
17	34883	BREL	2017-09	5000	45	1
18	34884	ELUN	2017-09	12000	72	1
19	34885	AWNG	2017-09	8500	50	1
20	34887	ST	2017-09	50000	95	1

Gambar 3. Tampilan Data Barang dengan penjualan yang tinggi



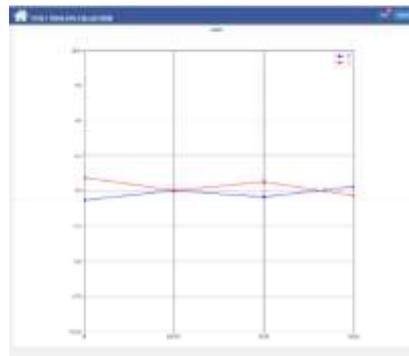
Gambar 4. Tampilan Perhitungan Cluster K-Means

Pada gambar 5. Merupakan tampilan proses data untuk menentukan proses cluster k-means.



Gambar 5. data barang yang terjual berdasarkan bulan

Dari hasil data yang telah diolah maka bisa diketahui produk yang memiliki penjualan paling tinggi dan paling rendah berdasarkan hasil presentase jumlah produk yang dapat dilihat pada gambar 6.



Gambar 6. Hasil pengolahan data

Berdasarkan perbandingan Cluster data didapat 20 barang yang memiliki hasil Cluster yang berbeda

4.1 Hasil Pengujian

Dalam proses cluster sistem ini akan menghitung rata-rata pada cluster seperti berikut:

Tabel 1. Rata-Rata Cluster

No	Id	Jumlah	Stok	harga
0	-0.001	0.0	-0.10	0.002
1	0.14	0.0	8.46	-0.15

Hasil rata-rata pada tabel diatas akan dialokasikan masing-masing data ke rata-rata terdekat:

Tabel 2. Rata-Rata Terdekat

No	Id	Jumlah	Stok	Harga	Prediksi
1	-0.001	0.0	-0.10	0.002	0
2	0.14	0.0	8.46	-0.15	1

Maka setelah itu akan didapat data train dan data testing yang sudah dikelompokan sesuai dengan clusternya seperti berikut:

Tabel 3. Data Latih dan Data Uji

Kode barang	Data latih	Data uji
AING	0	0
ATON	1	1
BANG	0	1
BBEL	1	0
BDUL	0	1
BIK	1	1
BSS	1	0
CCIN	0	0
CET	0	1
CLEK	1	1
CLIK	1	0
DKRIP	0	0
DPET	1	1
DRON	1	0
EPS	1	0
ESHADOW	1	0
GANG	1	1

Berdasarkan perbandingan Cluster data didapat 22 barang yang memiliki hasil Cluster yang berbeda, sehingga dengan demikian dapat dihitung unjuk kerja sistem dengan rumus sebagai berikut:

$$U_k = ((TD - JDM)/TD) \times 100\%$$

Keterangan: U_k : Unjuk kerja sistem

TD : Total data

JDM : Jumlah data meleset

Dengan ini hasil unjuk kerja sistem dalam melakukan pengelompokan barang yang tinggi dan barang yang rendah dalam penjualannya adalah sebagai berikut :

$$U_k = ((156 - 22)/156) \times 100\% = 81.21\%$$

Maka dari hasil perhitungan menggunakan rumus diatas diperoleh nilai unjuk kerja sistem ini adalah 81.21%

5. KESIMPULAN

Hasil dari penelitian ini, penulis mengambil beberapa kesimpulan diantaranya:

- a. Implementasi K-Means Clustering berbasis web ini dapat mengelompokkan data barang yang tinggi penjualannya dengan yang tidak dengan nilai unjuk kerja 81.21 %
- b. Algoritma K-Means dapat menghasilkan cluster rasio dengan melihat kemiripan nilai dalam setiap clusternya
- c. Dengan menggunakan metode ini bisa dilihat produk mana yang paling tinggi dan rendah penjualannya sehingga pemilik toko bisa menyiapkan stok barang untuk barang yang tinggi penjualannya dan memasang strategi pemasaran baru untuk barang yang rendah penjualannya

5. DAFTAR RUJUKAN

- [1] M. Nawang, L. Kurniawati, D. Duta, K. Akuntansi, S. Informasi, and K. Akuntansi, "Rancang Bangun Sistem Informasi Pengolahan Data Persediaan Barang Berbasis Dekstop Dengan Model," vol. 13, no. 2, pp. 233–238, 2017.
- [2] R. P. Hastanti, "Analisis Dan Perancangan Sistem Penjualan Berbasis Web (E-Commerce)," pp. 1–8, 2011.
- [3] M. Taufani, R. Riyadi, and R. Dewantara, "Analisis Dan Desain Sistem Informasi Pemasaran (Studi pada Sistem Informasi Pemasaran untuk Promosi CV. Intan Catering)," *J. Adm. Bisnis S1 Univ. Brawijaya*, vol. 38, no. 2, pp. 1–10, 2016.
- [4] B. S. Sutejo, "Internet Marketing : Konsep Dan," *J. Manaj.*, vol. 6, no. 1, pp. 41–57, 2006.
- [5] D. T. Hernandhi, E. S. Astuti, and S. Priambada, "Desain Sistem Informasi Pemasaran Berbasis Website Untuk Promosi (Studi Kasus pada Kedai Ayam Geprak & Sambal Bawang Malang)," *J. Adm. Bisnis*, vol. 55, no. 1, pp. 1–10, 2018.
- [6] M. H. Siregar, "Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan)," *J. Teknol. Dan Open Source*, vol. 1, no. 2, pp. 83–91, 2018.
- [7] E. Muningsih and S. Kiswati, "Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang," *J. Bianglala Inform.*, vol. 3, no. 1, pp. 10–17, 2015.
- [8] A. E. Permana, A. M. Reyhan, H. Rafli, and N. Aini, "Analisa Transaksi Belanja Online Pada Masa Pandemi," *J. TEKNOINFO*, vol. 15, no. 1, pp. 32–37, 2021.
- [9] D. N. P. Sari and Y. L. Sukestiyarno, "Analisis Cluster dengan Metode K-Means pada Persebaran Kasus Covid-19 Berdasarkan Provinsi di Indonesia," *Prism. Pros. Semin. Nas. Mat.*, vol. 4, pp. 602–610, 2021.
- [10] A. A. Rismayadi, N. N. Fatonah, and E. Junianto, "Algoritma K-Means Clustering Untuk Menentukan Strategi Pemasaran Di Cv. Integreet Konstruksi," *J. Responsif Ris. Sains dan Inform.*, vol. 3, no. 1, pp. 30–36, 2021.
- [11] S. Suliman, "Implementasi Data Mining Terhadap Prestasi Belajar Mahasiswa Berdasarkan Pergaulan dan Sosial Ekonomi Dengan Algoritma K-Means Clustering," *Simkom*, vol. 6, no. 1, pp. 1–11, 2021.
- [12] D. R. Ningrat, D. A. I. Maruddani, and T. Wuryandari, "Analisis cluster dengan algoritma K-Means dan Fuzzy C-Means clustering untuk pengelompokan data obligasi korporasi," *None*, vol. 5, no. 4, pp. 641–650, 2016.
- [13] W. Dhuhita, "Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita," *J. Inform. Darmajaya*, vol. 15, no. 2, pp. 160–174, 2015.
- [14] I. N. Rachmawati, "Pengumpulan Data Dalam Penelitian Kualitatif: WaRachmawati, I. N. (2007). Pengumpulan Data Dalam Penelitian Kualitatif: Wawancara. Jurnal Keperawatan Indonesia, 11(1), 35–40. <https://doi.org/10.7454/jki.v11i1.184wawancara>," *J. Keperawatan Indones.*, vol. 11, no. 1, pp. 35–40, 2007.
- [15] H. Hasanah, "Teknik-Teknik Observasi (Sebuah Alternatif Metode Pengumpulan Data Kualitatif Ilmu-ilmu Sosial)," *At-Taqaddum*, vol. 8, no. 1, p. 21, 2017.

- [16] A. T. Mirzaqon and B. Purwoko, “Studi Kepustakaan Mengenai Landasan Teori Dan Praktik Konseling Expressive Writing Library,” *J. BK UNESA*, pp. 1–8, 2017.
- [17] C. Oktarina, K. A. Notodiputro, and I. Indahwati, “Comparison of K-Means Clustering Method and K-Medoids on Twitter Data,” *Indones. J. Stat. Its Appl.*, vol. 4, no. 1, pp. 189–202, 2020.
- [18] N. Harianto Kristanto, A. LA Christopher, and H. S. Budi, “Implementasi K-Means Clustering untuk Pengelompokan Analisis Rasio Profitabilitas dalam Working Capital,” *Juisi*, vol. 02, no. 01, 2016.
- [19] N. Wakhidah, “Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering),” *Fak. Teknol. Inf.*, vol. 21, no. 1, pp. 70–80, 2014.