



Uji Performansi Algoritma K.Means Dalam Mencari Cluster Terbaik Pada Data Sales, Stock Produk Food Beverage (Studi Kasus: PT.Airmas International)

Haris Fadhillah¹, Achmad Hindasyah², Joni Prasetyo³

^{1,2,3}Tehnik Informatika, Universitas Pamulang, Tangerang Selatan, Banten

Email: ¹harisfadhillah111@gmail.com, ²ahindasyah@gmail.com, ³dosen01728@unpam.ac.id

ABSTRACT

This research was conducted with the aim of knowing the products that are most in demand by customers, also knowing the products that sell less and knowing the accuracy of testing the performance of sales, stock data. The qualitative method used focuses on interpretation and understanding of interviews or observations to collect sales data, stock of 23 food beverage product items. The results of manual calculation of the k-means algorithm with sales data, stock is to find out the number of products that are in demand by customers with categories: very salable = 3 product items, salable = 5 product items, less salable = 15 product items, then sales data, stock is tested with RapidMiner and Python applications for clustering / grouping what products are in demand by customers the results are the same as the categories: very salable = 3 product items, salable = 5 product items, less salable = 15 product items and looking for the best number of clusters is 3.

Keywords: Algorithm, Cluster, K-Means, Performance, Test

ABSTRAK

Penelitian ini dilakukan bertujuan ingin mengetahui produk yang paling diminati oleh *customer*, juga mengetahui produk yang kurang laku penjualannya dan mengetahui akurasi pengujian performansi data sales, stock. Metode kualitatif yang digunakan *focus* pada interpretasi dan pemahaman wawancara atau *observasi* untuk mengumpulkan data sales, stock sebanyak 23 item produk *food beverage*. Hasil perhitungan manual algoritma k-means dengan data sales, stock adalah mengetahui jumlah produk apa saja yang diminati oleh *customer* dengan kategori : laku sekali = 3 item produk, laku = 5 item produk, kurang laku = 15 item produk, selanjutnya data sales, stock dilakukan pengujian dengan aplikasi RapidMiner dan Python untuk mengclustering/pengelompokan produk apa saja yang diminati oleh *customer* hasilnya sama dengan kategori : laku sekali = 3 item produk, laku = 5 item produk, kurang laku = 15 item produk dan mencari jumlah *cluster* yang terbaik adalah 3.

Kata Kunci : Algoritma, Cluster, K Means, Performansi, Uji

1. PENDAHULUAN

Penelitian ini bertujuan untuk mengelompokkan data penjualan produk *import food beverage* yang dijual pada supermarket *modern trade* menjadi beberapa cluster untuk mengetahui produk mana yang diminati *customer*, sehingga data sales, stock tersebut akan dilakukan pengujian performansi. Algoritma yang digunakan pada penelitian ini adalah Algoritma K-Means dengan menggunakan program RapidMiner & Google Colab Python untuk mengelompokkan data sekaligus melakukan pengujian data sales, stock produk *import food beverage* dalam mencari *cluster* terbaik dan sebagai pembanding dalam akurasi data yang akurat.

Hasil akhir dari penelitian adalah berupa suatu program aplikasi RapidMiner & Google Colab Python yang dapat mengolah data *sales*, *stock* produk *import food beverage* yang diminati *customer* dengan kategori produk : laku sekali, laku, kurang laku. Dengan latar belakang di atas maka peneliti tertarik untuk melakukan pengkajian data *sales*, *stock* produk *import food beverage* lebih lanjut terhadap permasalahan yang dihadapi PT.Airmas International. Penjelasan *novelty* dari penelitian sebelumnya yang dirujuk dalam tesis ini kelebihan dan kebaharuan dari penelitian ini adalah menggunakan browser Google Colab dengan bahasa pemrogram Python dalam menguji performansi dataset *sales*, *stock* untuk di clustering/pengelompokan. Dengan ini penulis mengambil dengan judul “UJI PERFORMANSI ALGORITMA K.MEANS DALAM MENCARI CLUSTER TERBAIK PADA DATA SALES, STOCK PRODUK FOOD BEVERAGE (Studi kasus : PT.Airmas International)”.

2. METODE

Metode merupakan pendukung penelitian baik dari segi pengumpulan data, tempat penelitian, jenis dan sumber penelitian, tahapan penelitian, metode pengembangan, dan lainnya yang diperlukan dalam metode. Metode yang digunakan dalam penelitian ini adalah *mixed methods* adalah pendekatan penelitian yang menggabungkan elemen-elemen dari metode kualitatif dan kuantitatif dalam satu studi. Tujuan dari pendekatan ini adalah untuk memanfaatkan kelebihan masing-masing metode dan mendapatkan pemahaman yang lebih komprehensif tentang fenomena yang diteliti. Berikut adalah beberapa aspek penting dari *mixed methods*:

2.1. Analisis Kebutuhan Data

Analisis kebutuhan data adalah proses sistematis untuk mengidentifikasi dan mengevaluasi jenis data yang diperlukan untuk mendukung pengambilan keputusan, perencanaan, dan pelaksanaan berbagai kegiatan dalam suatu organisasi atau proyek. Proses ini melibatkan beberapa langkah utama:

a. Observasi

Teknik *observasi* ini digunakan untuk melakukan pengamatan langsung dikantor (gudang cabang Jakarta) dalam merekap data *sales*, data *stock*. Teknik pengumpulan data dengan *observasi* digunakan apabila penelitian berkenaan dengan perilaku manusia, proses kerja, gejala-gejala alam dan bila responden yang diamati tidak terlalu besar.

b. Wawancara

Wawancara adalah percakapan dengan maksud tertentu. Percakapan itu dilakukan oleh dua pihak, yaitu pewawancara (*interviewer*) yang mengajukan pertanyaan dan terwawancara (*interviewee*) yang memberikan jawaban atas pertanyaan itu. Maksud dari digunakannya metode ini adalah untuk memperoleh keterangan atau data yang diperlukan untuk tujuan penelitian. Data yang akan dicari melalui wawancara ini berkaitan dengan pendapat atau pernyataan dari sumber data.

c. Studi Kepustakaan

Dalam studi kepustakaan ini penulis mengumpulkan dan mempelajari berbagai teori dan konsep dasar yang berhubungan dengan masalah yang diteliti. Teori dan konsep dasar tersebut penulis peroleh dengan cara pembacaan literatur atau sumber-sumber tertulis seperti buku-buku, penelitian terdahulu, makalah, jurnal, artikel, hasil laporan dan majalah yang berkaitan dengan penelitian.

2.2. Analisis Kebutuhan Perangkat

Analisis kebutuhan perangkat adalah proses mengidentifikasi dan mendokumentasikan semua kebutuhan dan persyaratan perangkat keras dan perangkat lunak yang dibutuhkan untuk suatu proyek atau sistem. Proses ini melibatkan beberapa langkah, termasuk:

a. Perangkat Keras (*Hardware*)

Perangkat keras adalah semua bagian fisik komputer, dan ini dibedakan dengan data yang berada di dalamnya.

Tabel 2.1 Perangkat Keras

| Spesifikasi Hardware | Keterangan |
|----------------------|-------------------------|
| Processor | Intel core i3 |
| RAM | 4 Gb |
| System Type | 64-bit Operating System |
| Harddisk | 160Gb |

b. Perangkat Lunak (*Software*)

Perangkat lunak adalah istilah khusus untuk data yang diformat dan disimpan secara digital, termasuk program komputer, dokumentasinya, dan berbagai informasi yang bisa dibaca, dan ditulis oleh komputer. Peralatan dalam penelitian ini meliputi

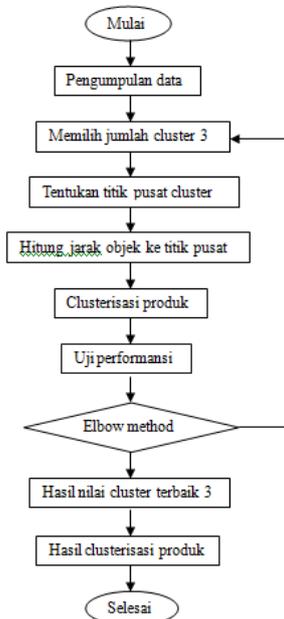
kebutuhan perangkat lunak dibawah ini merupakan peralatan atau *tools* yang dibutuhkan, diantaranya:

Tabel 2.2 Perangkat Lunak

| Versi | Versi | Fungsi |
|----------------------------------|----------|--|
| Sistem Operasi Microsoft Windows | 10 | Sebagai sistem operasi penelitian ini |
| Microsoft Office Word | 2007 | Digunakan untuk mengolah laporan hasil penelitian |
| Microsoft Office Exel | 2007 | Perhitungan manual Clusterisasi k means dataset |
| RapidMiner | 10.1.3.0 | Digunakan untuk uji performance dataset dalam mencari cluster terbaik & mengelompokan dataset produk yang diminati <i>customer</i> |
| Google Colab Python | Browser | Digunakan untuk uji performance dataset dalam mencari cluster terbaik |

c. Perancangan Penelitian

Perancangan penelitian adalah langkah-langkah penelitian yang terstruktur, ekonomis dan sesuai dengan tujuan penelitian sehingga data-data yang didapatkan adalah data yang akurat. Adapun tahapan yang dilakukannya diantaranya :



Gambar 2.3 Flowchart Perancangan

Dari gambar diatas dapat dijelaskan data produk *sales*, *stock* produk *food beverage* dikumpulkan semua, kemudian data tersebut dipilih menjadi 3 cluster dalam menentukan titik *cluster centroid* (C1, C2, C3), setelah itu dihitung jarak objek ke titik

pusat, data produk kemudian di clusterisasi/kelompokan, langkah selanjutnya data diuji performansi adalah proses evaluasi yang digunakan untuk mengukur seberapa baik model data bekerja dalam memprediksi atau mengklasifikasikan data baru dengan menggunakan program RapidMiner & Google Colab dengan bahasa pemrograman Python, hasil dari pengujian data adalah mendapatkan jumlah nilai cluster terbaik adalah 3 dan hasil berikutnya adalah jumlah produk yang diminati oleh *customer* dengan kategori : laku sekali = 3 item produk, laku = 5 item produk, dan kurang laku = 15 item produk.

3. HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai cara pengujian dataset untuk mengukur performansi Algoritma K-Means dalam mencari *cluster* terbaik pada data *sales*, *stock* produk *food beverage*. Berikut langkah pembahasannya dibawah ini:

3.1. Mengumpulkan Dataset Sales, Stock Produk Food Beverage

Dataset yang digunakan sebagai dataset adalah sebanyak 23 item produk. Data set yang digunakan dalam penelitian sebagai berikut:

| NO | BARCODE | Description | STOCK AWAL | SALES | STOCK AKHIR | |
|----|---------------|--|------------|-------|-------------|----|
| 1 | 8850632205159 | HartBeat Salt Candy Lime | 36 | 1 | 35 | C3 |
| 2 | 8850632205722 | HartBeat Salt Candy Gac & Passion Fruit | 32 | 2 | 30 | |
| 3 | 8850632232063 | HartBeat Salt Candy Pineapple | 40 | 1 | 39 | |
| 4 | 8710666036900 | Dr. Tom's Eucalyptus Menthol | 59 | 33 | 26 | |
| 5 | 8717399040860 | Dr. Tom's Mint Menthol | 70 | 34 | 36 | |
| 6 | 8717399041911 | Dr. Tom's Honey Lemon | 72 | 32 | 40 | |
| 7 | 8717399042406 | Dr. Tom's Ginger Orange | 66 | 22 | 44 | |
| 8 | 4009077072527 | Dr. C. Soldan Em-eukal Lemon Pouch | 130 | 126 | 4 | |
| 9 | 4009077072541 | Dr. C. Soldan Em-eukal Ginger Orange Pouch | 114 | 108 | 6 | |
| 10 | 4009077072510 | Dr. C. Soldan Em-eukal Salvia Pouch | 64 | 48 | 16 | |
| 11 | 4791038200000 | Te Reval Earl Grey | 59 | 49 | 10 | |
| 12 | 4791038016809 | Te Reval English Breakfast | 56 | 43 | 13 | |
| 13 | 4791038009757 | Te Reval Jasmine Green Tea | 53 | 41 | 12 | |
| 14 | 4791038004752 | Te Reval Chamomile Fresh | 92 | 89 | 3 | |
| 15 | 8423371002459 | Salysol Fried Almond | 42 | 0 | 42 | |
| 16 | 8423371000479 | Salysol Pistachio | 53 | 5 | 48 | |
| 17 | 8850189071153 | Flower Food Sesame Bars With Honey | 346 | 303 | 43 | |
| 18 | 8850189071139 | Flower Food Pumpkin Bars With Honey | 403 | 374 | 29 | C1 |
| 19 | 8850189071146 | Flower Food Sunflower Bars With Honey | 356 | 347 | 9 | |
| 20 | 8851004418030 | Tasto Potato Chips Devil Emperor Chili | 36 | 36 | 0 | C2 |
| 21 | 8851004408055 | Tasto Potato Chips Pla Sam Rod | 36 | 36 | 0 | |
| 22 | 8855917008616 | Otori Baked Coin Snack Cheese 38gr | 94 | 94 | 0 | |
| 23 | 8855917007961 | Otori Baked Coin Snack Chocolate 45gr | 95 | 95 | 0 | |

Tabel 3.1 Dataset Sales & Stock Produk

3.2. Menentukan Nilai Pusat Cluster

Dalam menentukan “n” buah pusat *cluster* awal dilakukan pembangkitan bilangan acak atau *random* yang mempresentasikan urutan data input. Pusat awal *cluster* didapatkan dari data itu sendiri, bukan dengan menentukan titik baru, yaitu dengan me-random pusat awal dari data. Berikut data nilai pusat clusternya.

Tabel 3.2 Nilai Pusat Cluster

| | | | | |
|----|-----|-----|----|-------------|
| C1 | 403 | 374 | 29 | Laku sekali |
| C2 | 36 | 36 | 0 | Laku |
| C3 | 36 | 1 | 35 | Kurang laku |

3.3. Menghitung Manual Jarak Data Dengan Pusat Cluster

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidian distance*. Algoritma perhitungan jarak data dengan pusat *cluster*. Dalam perhitungan ini semua perhitungan telah dilakukan berdasarkan rumus *Euclidean distance* bisa dilihat dibawah ini:

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

Atau disederhanakan :

$$\text{data 1,1} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_1)^2 + (x_3 - y_1)^2}$$

$$\text{data 1,2} = \sqrt{(x_1 - y_2)^2 + (x_2 - y_2)^2 + (x_3 - y_2)^2}$$

$$\text{data 1,3} = \sqrt{(x_1 - y_3)^2 + (x_2 - y_3)^2 + (x_3 - y_3)^2}$$

Keterangan rumus *Euclidean distance* :

Dimana : De adalah *Euclidean Distance* ; adalah banyaknya objek

(x, y) merupakan koordinat *object*

(x_i, y_i) merupakan koordinat *centroid*

Berikut cara perhitungannya dibawah ini :

$$\text{data 1,1} = \sqrt{(36 - 403)^2 + (1 - 374)^2 + (35 - 29)^2}$$

$$\text{data 1,2} = \sqrt{(32 - 36)^2 + (2 - 36)^2 + (30 - 0)^2}$$

$$\text{data 1,3} = \sqrt{(40 - 36)^2 + (1 - 1)^2 + (39 - 35)^2}$$

$$\text{data 1,1} = \sqrt{(523,311+49,4975 + 0}$$

$$\text{data 1,2} = \sqrt{(525,382+45,5192 + 6,480741}$$

$$\text{data 1,3} = \sqrt{(520,575+52,5547 + 5,656854}$$

$$\text{data 1,1} = \sqrt{572,8085}$$

$$\text{data 1,2} = \sqrt{577,381941}$$

$$\text{data 1,3} = \sqrt{578,786554}$$

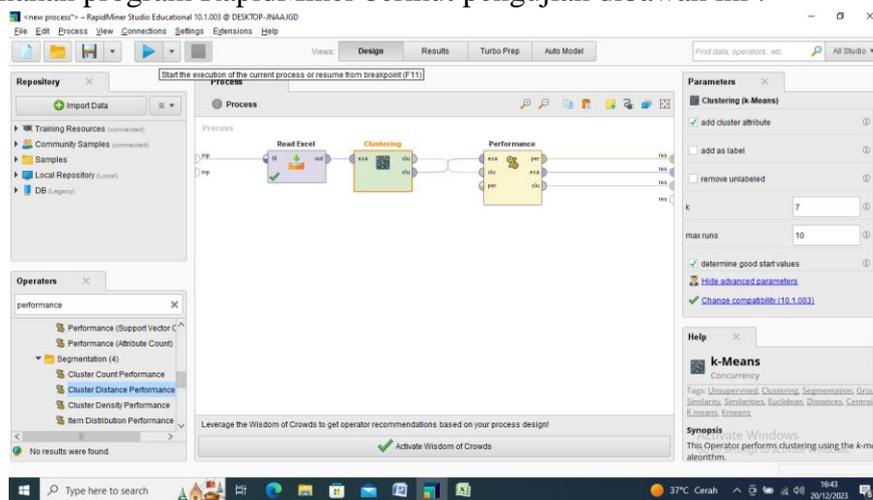
$$\text{data 1,1} = 23,933418$$

$$\text{data 1,2} = 24,028773$$

data 1,3 = 24,057983

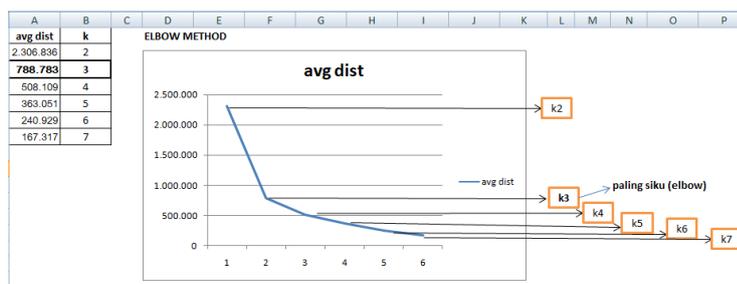
3.4. Pengujian Dalam Mencari Cluster Terbaik dan Clustering K.Means dengan Program RapidMiner

Berikut ini pengujian dalam mencari cluster terbaik dengan menggunakan dataset yang akan diuji dapat dilihat pada dataset sales, stock produk *food beverage* dengan menggunakan program RapidMiner berikut pengujian dibawah ini :



Gambar 3.3 Pengujian Memilih Cluster (K-7)

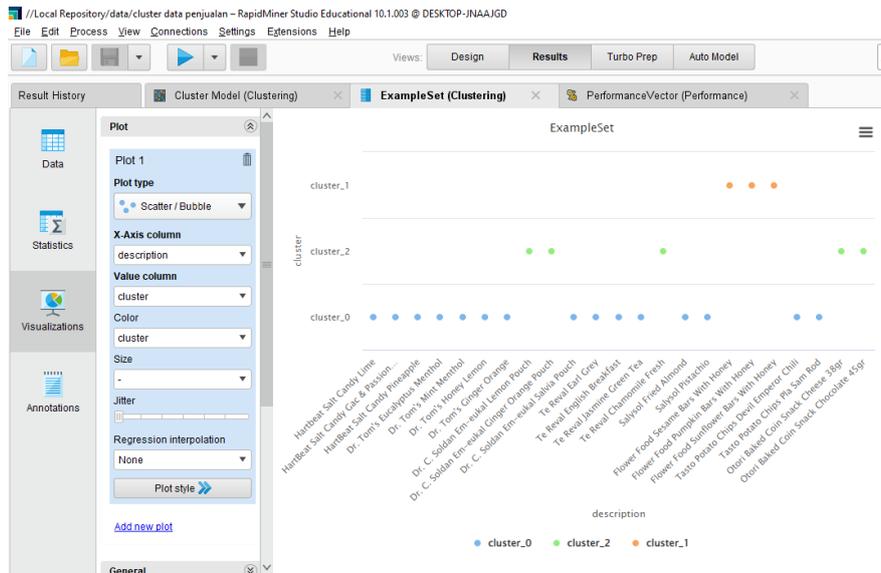
Penjelasan gambar diatas adalah pengujian mencari *cluster* terbaik dari *cluster* 1 sampai 7 dengan menggunakan program aplikasi RapidMiner, berikut gambar dibawah ini hasil pengujian seberapa baik cluster yang harus digunakan.



Gambar 3.4 Grafik Elbow Method

Penjelasan gambar diatas Grafik garis atau line adalah grafik yang menggunakan titik sebagai poin data dan kemudian titik tersebut disambung dengan titik berikutnya. Grafik ini merupakan salah satu grafik yang paling sering digunakan setelah grafik batang. Cara membuat grafik line di Excel tidak jauh berbeda dengan “Cara Membuat Grafik Batang”. Secara umum grafik line digunakan untuk membandingkan suatu data berdasarkan deret waktu tertentu. Dari gambar diatas Metode *Elbow* merupakan

suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* yang akan membentuk siku pada suatu titik. *Sum of Square Error* (SSE) sering digunakan dalam acuan penelitian yang berhubungan dengan *optimalisasi* kluster. Dari hasil grafik *Elbow method* diatas ini garis yang paling siku menunjukkan pada grafik **k3** yang menunjukkan nilai *cluster/k = 3* adalah yang terbaik.



Gambar 3.5 Hasil Clustering Produk Food Beverage

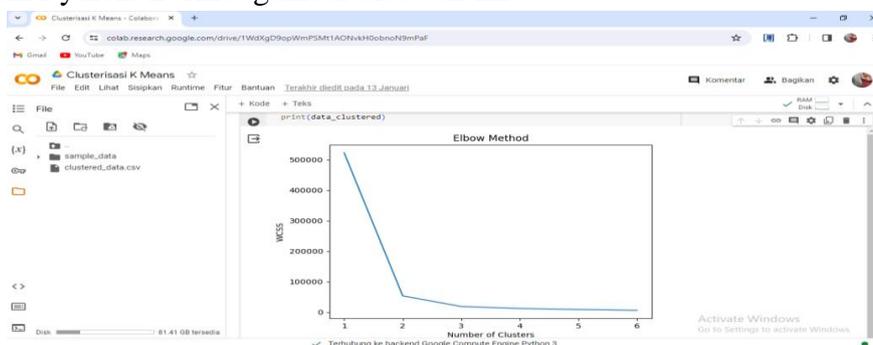
Penjelasan gambar diatas mengenai hasil *visualizations* di RapidMiner digunakan untuk memahami dan menganalisis data secara *grafis*. Kita juga dapat membuat berbagai jenis visualisasi seperti *scatter plots*, *bar charts*, dan *line charts*, untuk mendapatkan wawasan yang lebih baik tentang pola dan hubungan dalam data anda. Pada bagian *visualizations* muncul grafik hasil *clustering* $K=3$

- a. Cluster 1 berwarna kuning menunjukkan ada 3 item produk yang paling diminati oleh customer dengan kategori : Laku sekali. Berikut nama produk yang Laku sekali:
 1. Flower Food Sesame Bars With Honey
 2. Flower Food Pumpkin Bars With Honey
 3. Flower Food Sunflower Bars With Honey
- b. Cluster 2 berwarna hijau menunjukkan ada 5 item produk yang sedang/lumayan diminati oleh customer dengan kategori : Laku. Berikut nama produk yang Laku :
 1. Dr. C. Soldan Em-eukal Lemon Pouch
 2. Dr. C. Soldan Em-eukal Ginger Orange Pouch
 3. Te Reval Chamomile Fresh

4. Otori Baked Coin Snack Cheese 38gr
 5. Otori Baked Coin Snack Chocolate 45gr
- c. Cluster 0/3 berwarna biru menunjukkan ada 15 item produk yang tidak disukai oleh customer dengan kategori : Kurang laku. Berikut nama produk yang Kurang laku.
1. Hartbeat Salt Candy Lime
 2. HartBeat Salt Candy Gac & Passion Fruit
 3. HartBeat Salt Candy Pineapple
 4. Dr. Tom's Eucalyptus Menthol
 5. Dr. Tom's Mint Menthol
 6. Dr. Tom's Honey Lemon
 7. Dr. Tom's Ginger Orange
 8. Dr. C. Soldan Em-eukal Salvia Pouch
 9. Te Reval Earl Grey
 10. Te Reval English Breakfast
 11. Te Reval Jasmine Green Tea
 12. Salysol Fried Almond
 13. Salysol Pistachio
 14. Tasto Potato Chips Devil Emperor Chili
 15. Tasto Potato Chips Pla Sam Rod

3.5. Pengujian Dalam Mencari Cluster Terbaik dan Clustering K.Means dengan Google Colab Python

Berikut ini pengujian dengan menggunakan dataset untuk mencari *cluster* terbaik dan clustering produk *food*, *beverage* dengan menggunakan Google Colab dan bahasa pemrogram Python. Berikut gambar dibawah ini.



Gambar 3.6 Hasil Visual Elbow Method

Penjelasan gambar diatas *elbow method* adalah sebuah metode yang digunakan untuk menentukan jumlah optimal dari *cluster* dalam analisis klusterisasi data. Pada hasil visual di *google colab*, *elbow method* umumnya diterapkan pada grafik yang menunjukkan inersia (*inertia*) kluster terhadap jumlah kluster yang berbeda. Jumlah kuadrat dalam cluster (WCSS) adalah metrik yang digunakan dalam algoritma *clustering*, seperti *k-means*, untuk mengukur kekompakan *cluster*. Ini mewakili jumlah jarak kuadrat antara setiap titik data dalam sebuah *cluster* dan pusat massa *cluster* tersebut. Hasil grafik *Elbow method* menunjuka garis yang paling siku adalah **k3** atau **cluster 3** yang terbaik.

| | description | stock_awal | sales | stock_akhir | Cluster | Cluster_Name |
|----|--|------------|-------|-------------|---------|--------------|
| 0 | HartBeat Salt Candy Lime | 36 | 1 | 35 | 0 | kurang laku |
| 1 | HartBeat Salt Candy Gac & Passion Fruit | 32 | 2 | 30 | 0 | kurang laku |
| 2 | HartBeat Salt Candy Pineapple | 40 | 1 | 39 | 0 | kurang laku |
| 3 | Dr. Tom's Eucalyptus Menthol | 59 | 33 | 26 | 0 | kurang laku |
| 4 | Dr. Tom's Mint Menthol | 70 | 34 | 36 | 0 | kurang laku |
| 5 | Dr. Tom's Honey Lemon | 72 | 32 | 40 | 0 | kurang laku |
| 6 | Dr. Tom's Ginger Orange | 66 | 22 | 44 | 0 | kurang laku |
| 7 | Dr. C. Soldan Em-eukal Lemon Pouch | 130 | 126 | 4 | 2 | laku |
| 8 | Dr. C. Soldan Em-eukal Ginger Orange Pouch | 114 | 108 | 6 | 2 | laku |
| 9 | Dr. C. Soldan Em-eukal Salvia Pouch | 64 | 48 | 16 | 0 | kurang laku |
| 10 | Te Reval Earl Grey | 59 | 49 | 10 | 0 | kurang laku |
| 11 | Te Reval English Breakfast | 56 | 43 | 13 | 0 | kurang laku |
| 12 | Te Reval Jasmine Green Tea | 53 | 41 | 12 | 0 | kurang laku |
| 13 | Te Reval Chamomile Fresh | 92 | 89 | 3 | 2 | laku |
| 14 | Salysol Fried Almond | 42 | 0 | 42 | 0 | kurang laku |
| 15 | Salysol Pistachio | 53 | 5 | 48 | 0 | kurang laku |
| 16 | Flower Food Sesame Bars With Honey | 346 | 303 | 43 | 1 | laku sekali |
| 17 | Flower Food Pumpkin Bars With Honey | 403 | 374 | 29 | 1 | laku sekali |
| 18 | Flower Food Sunflower Bars With Honey | 356 | 347 | 9 | 1 | laku sekali |
| 19 | Tasto Potato Chips Devil Emperor Chili | 36 | 36 | 0 | 0 | kurang laku |
| 20 | Tasto Potato Chips Pla Sam Rod | 36 | 36 | 0 | 0 | kurang laku |
| 21 | Otori Baked Coin Snack Cheese 38gr | 94 | 94 | 0 | 2 | laku |
| 22 | Otori Baked Coin Snack Chocolate 45gr | 95 | 95 | 0 | 2 | laku |

Gambar 3.7 Hasil Visual Clusterisasi Produk *Food Beverage*

Penjelasan gambar diatas dari hasil visual clusterisasi/pengelompokan jumlah produk yang diminati oleh *customer* dengan kategori : laku sekali = 3 item produk, laku = 5 item produk, kurang laku = 15 item produk dengan pemrograman Google Colab Python.

4. KESIMPULAN

4.1. Kesimpulan

Berdasarkan data yang sudah di uji menggunakan aplikasi RapidMiner dan Google Colab Python menunjukkan hasil yang sama *visualizations* grafik *clustering* $K=3$ terhadap produk yang diminati oleh customer dengan kategori : laku sekali = 3 item produk, laku = 5 item produk, dan kurang laku = 15 item produk. Dan menguji performansi kelayakan dalam mencari cluster terbaik dari 2 sampai 7 hasil dari grafik *Elbow method* menunjukkan dalam menentukan K atau *Cluster* terbaik adalah $K/Cluster = 3$.

4.2. Saran

Saran untuk penelitian selanjutnya agar dilakukan evaluasi pengembangan terhadap data penjualan untuk mengidentifikasi tren, pola pembelian dan memprediksi permintaan produk dimasa depan, memungkinkan perencanaan stok yang lebih baik dan pengurangan risiko kelebihan atau kekurangan persediaan.

5. DAFTAR PUSTAKA

- [1] A. T. Rahman, W. Wiranto, dan R. Anggrainingsih, "Coal Trade Data Clustering Using K-Means (Case Study Pt. Global Bangkit Utama)," *ITSMART J. Teknol. dan Inf.*, vol. 6, no. 1, hal. 24–31, 2017.
- [2] A. Prasetya, R. Salkiawati, dan A. D. Alexander, "Analisis Cluster K-Means dengan Metode Elbow untuk Menentukan Pola Penjualan Produk Traffic Room Summarecon Mal Bekasi," *J. Students' Res. Comput. Sci.*, vol. 4, no. 1, hal. 105–118, 2023, doi: 10.31599/jsrscs.v4i1.2480. [6]
- [3] A. Nugraha, O. Nurdiawan, dan G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, hal. 849–855, 2022, doi: 10.36040/jati.v6i2.5755.
- [4] D. O. Dacwanda dan Y. Nataliani, "Implementasi k-Means Clustering untuk Analisis Nilai Akademik Siswa Berdasarkan Nilai Pengetahuan dan Keterampilan," *Aiti*, vol. 18, no. 2, hal. 125–138, 2021, doi: 10.24246/aiti.v18i2.125-138.
- [5] E. Rahmawati, Y. H. Chrisnanto, dan A. Maspupah, "Identifikasi Kemampuan Akademik Mahasiswa Menggunakan K-Means Clustering," *Pros. SEMNAS Inov. Teknol.*, hal. 87–92, 2019, [Daring]. Tersedia pada:

- <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/518>
- [6] M. R. Nugroho, I. E. Hendrawan, dan P. P. Purwantoro, “Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI,” *Nuansa Inform.*, vol. 16, no. 1, hal. 125–133, 2022, doi: 10.25134/nuansa.v16i1.5294.
- [7] R. K. Dinata, S. Safwandi, N. Hasdyna, dan N. Azizah, “Analisis K-Means Clustering pada Data Sepeda Motor,” *INFORMAL Informatics J.*, vol. 5, no. 1, hal. 10, 2020, doi: 10.19184/isj.v5i1.17071.
- [8] S. Dewi dan S. Fatimah, “Implementasi Metode Clustering Algoritma K-Means untuk Menentukan Kelompok Tahfidz dan Tahsin di Pesantren Siswa Al-Ma’soem,” *J. Account. Inf. Syst.*, vol. 6, no. 1, hal. 10–18, 2023, doi: 10.32627/aims.v6i1.701.
- [9] S. Syahidatul Helma *et al.*, “Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means,” *Puzzle Res. Data Technol. Fak. Sains dan Teknol.*, vol. 1, no. November, hal. 4, 2019.
- [10] S. I. Murpratiwi, I. G. Agung Indrawan, dan A. Aranta, “Analisis Pemilihan Cluster Optimal Dalam Segmentasi Pelanggan Toko Retail,” *J. Pendidik. Teknol. dan Kejuru.*, vol. 18, no. 2, hal. 152, 2021, doi: 10.23887/jptk-undiksha.v18i2.37426.