

Clustering K-Means Dalam Pengelompokan Penjualan Produk Pada RTO Group

K-MEANS CLUSTERING IN PRODUCT SALES GROUPING IN THE RTO GROUP

Imam Rosyadi^{*1}, Hadwitya Handayani Kusumawardhani²,
Fenilinas Adi Artanto³, Andaru Alwan A. Hardani⁴, Farah Nafilaturrosyidah⁵

^{1,2,3,4,5} Universitas Muhammadiyah Pekajangan Pekalongan:

Jl. Raya Ambokembang No. 08, Kab. Pekalongan, 51173, Indonesia

^{1,2,4,5} Manajemen Informatika, UMPP, ³ Informatika, UMPP

e-mail: ^{*1}mamrosya@yahoo.com, ²hadwitya.hk@gmail.com,

³ fenilinasadi@gmail.com, ⁴ andaru.aalwan@umpp.ac.id, ⁵ farah.nafila@umpp.ac.id

Abstrak

RTO Grup merupakan sebuah perusahaan yang bergerak dalam bidang penjualan batik tulis. Dengan memiliki banyak produk maka RTO Grup membutuhkan pengelompokan untuk penjualan produknya. Pengelompokan penjualan produk RTO grup dapat digunakan untuk mengetahui produk-produk mana yang laku dan dapat memaksimalkan strategi penjualan dalam produknya. Selain itu pengelompokan produk dapat digunakan sebagai dasar dalam metode penyimpanan produk agar penataan stok bisa menjadi lebih mudah dan tertata. Dalam mengelompokan produk digunakan metode datamining. Dengan menggunakan metode datamining maka akan didapatkan informasi dari sekumpulan data penjualan yang sudah didapatkan oleh RTO Group. Dalam menerapkan konsep datamining digunakan metode K-Means. Metode K-Means digunakan karena mudah untuk diimplementasikan dan dijalankan, waktu yang diperlukan dalam pembelajarannya juga relatif cepat dan mudah diadaptasikan. Dalam mengestimasi data berdasarkan algoritma K-Means digunakan aplikasi *datamining*. Dari hasil estimasi didapatkan pembagian klaster sebanyak 4 *clustering* dengan jumlah *items* terbanyak berada di cluster 2. Dari *boxplot* juga didapatkan bahwa cluster ke 4 menjadi kelompok untuk barang produksi yang memiliki jumlah terjual paling banyak, sehingga dapat ditingkatkan dalam proses produksinya, sedangkan klaster 3 menjadi kelompok dengan jumlah barang laku yang paling sedikit, yang pastinya menjadi poin utama dalam peningkatan penjualan.

Kata kunci— Data Mining, Clustering, K-Means, Orange Datamining, Penjualan

Abstract

RTO Group is a company that operates in the field of selling written batik. By having many products, the RTO Group requires grouping for product sales. RTO group product sales grouping can be used to find out which products are selling well and can maximize sales strategies for their products. Apart from that, product grouping can be used as a basis for product storage methods so that stock management can be easier and more organized. In grouping products, datamining methods are used. By using the data mining method, information will be obtained from a set of sales data that has been obtained by the RTO Group. In applying the datamining concept, the K-Means method is used. The K-Means method is used because it is easy to implement and run, the time required to learn is also relatively fast and easy to adapt. In estimating data based on the K-Means algorithm, a datamining application is used. From the estimation results, it was found that the clustering was divided into 4 clusters with the highest number of items being in cluster 2. From the boxplot it was also found that the 4th cluster was the group for production goods that had the highest number of sales, so that it could be improved in the production process, while cluster 3 was the group with the smallest number of selling items, which is definitely the main point in increasing sales.

Keyword— Data Mining, Clustering, K-Means, Orange Datamining, Sale

1. PENDAHULUAN

Sebuah perusahaan dapat meningkatkan kualitas perusahaan dapat dilihat dari banyaknya produk yang dijual. Dengan mengetahui penjualan sebuah produk maka perusahaan

dapat menerapkan metode atau strategi pemasaran untuk produknya. RTO Group merupakan sebuah perusahaan yang bergerak di industri batik tulis dan juga pengadaan material [1]. Saat ini RTO Group sedang memfokuskan produknya di bidang batik tulis dengan mengeluarkan beberapa desain batik. Dengan sudah terjualnya beberapa produk desain batik maka RTO perlu adanya pengelompokan penjualan batik tulisnya. Agar RTO dapat menerapkan strategi Pemasaran produknya dengan sesuai. Untuk itu digunakan konsep *data mining*.

Dengan adanya pengelompokan produk akan memudahkan RTO untuk menganalisa produk mana yang terjual lebih banyak dan produk mana yang penjualan masih sedikit, sehingga RTO Group dapat menerapkan strategi yang tepat dalam penjualan. Selain itu dengan pengelompokan produk akan mempermudah RTO dalam penyusunan stock dari produknya, dimana stock produk yang laku akan lebih mudah untuk dicari atau ditata dilokasi yang lebih mudah dijangkau.

Data mining merupakan sebuah konsep dimana kita mencari sebuah informasi dari banyaknya data yang telah didapatkan [2]. Data mining sering juga mendapatkan istilah *Knowledge Discovery in Database* (KDD) yang memiliki artian bahwa datamining adalah kegiatan dalam mencari sebuah pola dalam data yang berjumlah besar [3]. Hasil dari data mining dapat digunakan dalam proses untuk memperbaiki ataupun konsep dari pengambilan keputusan di masa depan. Dalam data mining terbagi menjadi enam tahapan yaitu, pembersihan data (*data cleaning*), integrasi data (*data integration*), seleksi data (*data selection*), transformasi data (*data transformation*), proses *mining*, dan presentasi pengetahuan (*knowledge presentation*) [4][5].

Dalam konsep data mining memiliki metode *Clustering* salah satunya adalah K-Means. *Clustering* adalah teknik dalam data mining yang dipakai untuk menganalisa data dalam menyelesaikan permasalahan pengelompokan data [6]. Algoritma K-Means merupakan algoritma clustering yang termasuk dalam kelompok *unsupervised learning* yang berfungsi untuk mengelompokan data dengan metode partisi yang berbasis pada titik pusat (*centroid*) [7]. K-Means merupakan metode pengelompokan data non-hirarki yang mempartisi data ke dalam bentuk dua ataupun lebih cluster [8]. Metode K-Means memiliki keunggulan seperti digunakan karena mudah untuk diimplementasikan dan dijalankan, waktu yang diperlukan dalam pembelajarannya juga relatif cepat dan mudah diadaptasikan [9]. Pada K-Means data atau objek yang diteliti tidak mempunyai label kelas yang akan dikelompokan tanpa mengetahui terlebih dahulu target kelansnya [10].

Pada penelitian Darmi (2017) melakukan pengelompokan produk pada swalayan dengan menggunakan algoritma K-Means. Pengelompokan yang dilakukan dimaksudkan untuk mengetahui produk mana yang laku dan tidak laku. Dengan menggunakan K-Means maka data penjualan batik pada RTO dapat dikelompokan. Dengan adanya pengelompokan maka akan mudah digunakan dalam pengambilan keputusan dalam pengembangan perusahaan RTO group.

2. METODE PENELITIAN

Metode Penelitian dalam Penelitian dapat dilihat pada gambar dibawah ini:



Gambar 1. Bagan Alir Penelitian

Dari masalah yang telah terindifikasi yaitu kebutuhan akan pengelompokan produk pada RTO group maka dalam mengatasi masalah tersebut diperlukan metode pengelompokan data yaitu dengan menggunakan algoritma K-Menas. Data yang digunakan dalam Penelitian ini adalah data dari penjualan produk pada RTO group yang berisi tentang nama produk, stock, harga, dan penjualan produk. Data yang telah didapatkan akan diestimasi menggunakan algoritma K-Means, langkah-langkah dari algoritma K-Means adalah sebagai berikut [10][12]:

1. Tentukan jumlah kluster (k).
2. Bangkitkan nilai sebagai pusat cluster awal sebanyak (k)
3. Hitung jarak antara data dan titik pusat cluster menggunakan rumus:

$$D_{i,j} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Dimana:

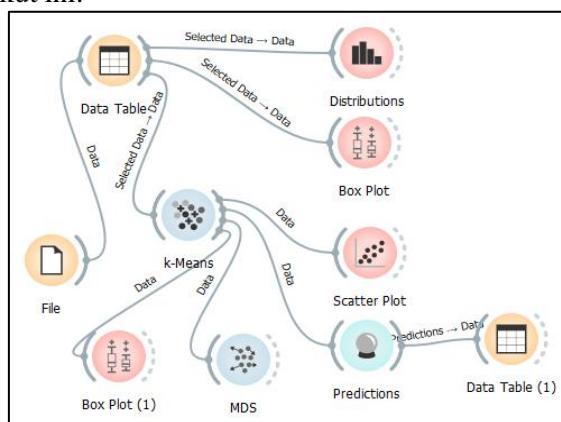
$D_{i,j}$ = jarak data ke i ke pusat cluster j

X_{ki} = data ke i pada atribut ke k

X_{kj} = titik pusat ke j pada atribut ke k

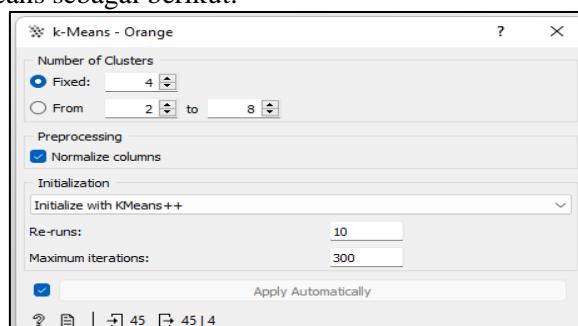
4. Kelompokan setiap data berdasarkan kedekatan dengan centroid (jarak terkecil)
5. Lakukan pengulangan sampai tidak ada anggota cluster yang berubah
6. Jika pada langkah ke 5 terpenuhi maka pengelompokan data digunakan nilai pusat cluster pada iterasi terkecil.

Nantinya data yang telah didapatkan akan di estimasikan menggunakan aplikasi *Orange Datamining* dengan algoritma K-Means Clustering, dan digunakan desain gambar estimasi dari program *orange datamining* berikut ini:



Gambar 2. Desain Orange Data Mining

Dengan settingan K-Means sebagai berikut:



Gambar 3. Setingan K-Means

Dipilih jumlah cluster sebanyak 4 cluster dengan pengulangan sebanyak 10 kali dan maksimal iterasi sebanyak 300.

3. HASIL DAN PEMBAHASAN

Pada Penelitian terdahulu digunakan aplikasi *rapidminer* dalam mengelompokan hasil dari penjualan, sedangkan pada penelitian ini digunakan *orange datamining*. Digunakanya orange data mining dikarenakan lebih mudah dan simpel dalam penggunaan dan juga tidak memerlukan spesifikasi komputer yang tinggi.

Dari data yang sudah diestimasi dengan menggunakan aplikasi orange datamining didapatkan hasil pembagian cluster sebanyak 4 kelompok kluster dengan jumlah items sebagai berikut:

Tabel 1. Hasil Pembagian Clustering K-Means

| Clustering | Items |
|------------|-------|
| 1 | 9 |
| 2 | 26 |
| 3 | 5 |
| 4 | 5 |

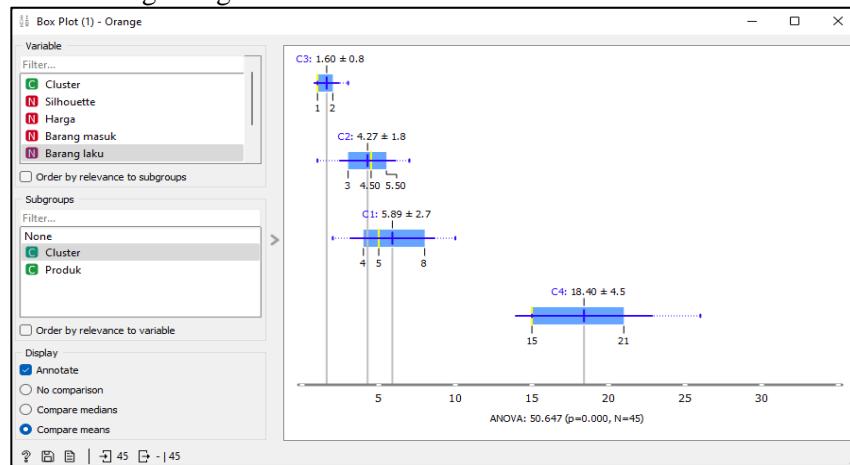
Dari ke empat cluster tersebut mendapatkan pembagian cluster sebagai berikut:

Tabel 2. Hasil Pembagian Clustering

| Cluster | Silhouette | Harga | Barang masuk | Barang laku | Stock | Produk |
|---------|------------|--------|--------------|-------------|-------|--------|
| C1 | 0.631493 | 50000 | 24 | 6 | 18 | CHL5 |
| C1 | 0.635125 | 50000 | 24 | 3 | 21 | CHL4 |
| C1 | 0.636685 | 50000 | 24 | 4 | 20 | CHL3 |
| C1 | 0.631414 | 50000 | 24 | 2 | 22 | CHL2 |
| C1 | 0.6356 | 50000 | 24 | 5 | 19 | CHL1 |
| C1 | 0.613197 | 75000 | 24 | 8 | 16 | CHP5 |
| C1 | 0.583581 | 75000 | 24 | 10 | 14 | CHP4 |
| C1 | 0.635304 | 75000 | 24 | 5 | 19 | CHP3 |
| C1 | 0.583581 | 75000 | 24 | 10 | 14 | CHP1 |
| C2 | 0.597259 | 120000 | 12 | 1 | 11 | JK4 |
| C2 | 0.630812 | 120000 | 12 | 4 | 8 | JK3 |
| C2 | 0.597259 | 120000 | 12 | 1 | 11 | JK2 |
| C2 | 0.611546 | 120000 | 12 | 2 | 10 | JK1 |
| C2 | 0.640362 | 250000 | 12 | 6 | 6 | BT27 |
| C2 | 0.627035 | 250000 | 12 | 3 | 9 | BT23 |
| C2 | 0.639197 | 250000 | 12 | 5 | 7 | BT22 |
| C2 | 0.64131 | 300000 | 12 | 6 | 6 | BT21 |
| C2 | 0.615806 | 250000 | 12 | 2 | 10 | BT20 |
| C2 | 0.63743 | 500000 | 12 | 4 | 8 | BT19 |
| C2 | 0.640816 | 400000 | 12 | 7 | 5 | BT18 |
| C2 | 0.627035 | 250000 | 12 | 3 | 9 | BT16 |

| Cluster | Silhouette | Harga | Barang masuk | Barang laku | Stock | Produk |
|---------|------------|---------|--------------|-------------|-------|--------|
| C2 | 0.640203 | 300000 | 12 | 5 | 7 | BT15 |
| C2 | 0.633929 | 600000 | 12 | 7 | 5 | BT14 |
| C2 | 0.642644 | 450000 | 12 | 6 | 6 | BT13 |
| C2 | 0.63743 | 500000 | 12 | 4 | 8 | BT12 |
| C2 | 0.640203 | 300000 | 12 | 5 | 7 | BT11 |
| C2 | 0.637721 | 200000 | 12 | 7 | 5 | BT10 |
| C2 | 0.531488 | 1000000 | 6 | 1 | 5 | BT8 |
| C2 | 0.63429 | 700000 | 12 | 4 | 8 | BT7 |
| C2 | 0.639517 | 600000 | 12 | 6 | 6 | BT6 |
| C2 | 0.641564 | 500000 | 12 | 5 | 7 | BT5 |
| C2 | 0.629583 | 400000 | 12 | 3 | 9 | BT4 |
| C2 | 0.639197 | 250000 | 12 | 5 | 7 | BT3 |
| C2 | 0.635909 | 300000 | 12 | 4 | 8 | BT2 |
| C2 | 0.640203 | 300000 | 12 | 5 | 7 | BT1 |
| C3 | 0.646222 | 3000000 | 3 | 1 | 2 | BT26 |
| C3 | 0.626675 | 2000000 | 6 | 1 | 5 | BT25 |
| C3 | 0.56591 | 1500000 | 6 | 2 | 4 | BT24 |
| C3 | 0.651411 | 2500000 | 3 | 1 | 2 | BT17 |
| C3 | 0.565974 | 1500000 | 6 | 3 | 3 | BT9 |
| C4 | 0.541743 | 100000 | 36 | 15 | 21 | JN4 |
| C4 | 0.618105 | 100000 | 36 | 26 | 10 | JN3 |
| C4 | 0.620585 | 100000 | 36 | 21 | 15 | JN2 |
| C4 | 0.541743 | 100000 | 36 | 15 | 21 | JN1 |
| C4 | 0.499298 | 75000 | 24 | 15 | 9 | CHP2 |

Hasil plot dari clustering sebagai berikut:



Gambar 4. Hasil Boxplot Clustering

Dari plot grafik diatas dapat dilihat bahwa cluster 4 yang memiliki banyak barang laku, sehingga nantinya pihak RTO dapat meningkatkan produksi ataupun penataan barang pada produk di cluster 4, sedangkan pada cluster 3 banyak produk yang belum laku, sehingga pihak

RTO dapat menerapkan beberapa strategi pemasaran pada produk cluster 3 dan mengurangi produksinya terlebih dahulu. Terlihat cluster 4 berisi produk dengan inisial JN yang artinya produk jeans lebih diminati oleh konsumen RTO grup. Selain itu produk pada kluster 3 berisi produk dengan inisial BT dengan desain 24, 25, 26 yang masih belum laku, maka dapat disimpulkan bahwa desain batik 24, 25 dan 26 kurang diminati oleh para konsumen RTO group.

4. KESIMPULAN

Dari pembagian clustering menggunakan algoritma K-Means didapatkan 4 clustering dengan jumlah items terbanyak berada di cluster 2. Dari boxplot juga didapatkan bahwa cluster ke 4 menjadi kelompok untuk barang produksi yang memiliki jumlah terjual paling banyak, sehingga dapat ditingkatkan dalam proses produksinya, sedangkan kluster 3 menjadi kelompok dengan jumlah barang laku yang paling sedikit, yang pastinya menjadi point utama dalam peningkatan penjualan.

DAFTAR PUSTAKA

- [1] I. Rosyadi, F. A. Artanto, and M. Y. Febrianto, “Pengaruh Kepuasan Kerja terhadap Kinerja Karyawan di RTO dengan Structural Equation Modeling (SEM) dan Partial Least Square (PLS),” *Indones. J. Strateg. Manag.*, vol. 9, no. 1, pp. 8–17, 2021.
- [2] H. H. Kusumawardani, I. Rosyadi, F. A. Artanto, F. I. Arzha, and N. A. Rachmayani, “Analisis Decision Tree dalam Pengaruh Digital Marketing terhadap Penerimaan Siswa Baru,” *Remik*, vol. 6, no. April, pp. 225–231, 2022.
- [3] H. Effendi, A. Syahrial, S. Prayoga, and W. D. Hidayat, “Penerapan Metode K-Means Clustering untuk Pengelompokan Lahan Sawit Produktif pada PT Kasih Agro Mandiri,” *Teknematika*, vol. 11, no. 02, pp. 117–126, 2021.
- [4] I. Rosyadi, F. A. Artanto, S. E. Rahmawati, H. Tri, and B. Joyo, “Decision Tree Dalam Analisis Keputusan Pembelian Program Pada Perkumpulan Penggiat Programmer Indonesia,” *J. Fasilkom*, vol. XII, no. III, pp. 141–144, 2022.
- [5] P. R. Saputra Andri , Subing Mulia, “Perbandingan Metode Naïve Bayes Classifier Dan Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Mengenai Piala Dunia Fifa 2022,” *Teknematika*, vol. 13, no. 01, pp. 22–31, 2023, [Online]. Available: <http://ojs.palcomtech.ac.id/index.php/teknomatika/article/view/616>
- [6] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [7] N. Damanik and M. Sigiro, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru Sebagai Metode Promosi,” *J. Tek. Inform. Komput. Univers.*, vol. 4, no. 2, p. 158, 2021.
- [8] I. Of and K. M. In, “Penerapan Metode K-Means dalam Mengelompokkan Tingkat Kesembuhan Penderita Covid-19,” vol. 11, no. 02, pp. 127–136, 2021.
- [9] F. Putri, A. Hasibuan, S. Sumarno, and I. Parlina, “Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone,” vol. 1, no. 1, pp. 15–20, 2021.
- [10] B. J. M. Putra and D. A. F. Yuniarti, “Analisis Gaya Belajar terhadap Nilai Mahasiswa dengan Menggunakan Metode k-Means,” *Techno.Com*, vol. 21, no. 2, pp. 343–354, 2022, doi: 10.33633/tc.v21i2.5837.
- [11] Y. D. Darmi and A. Setiawan, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk,” *J. Media Infotama*, vol. 12, no. 2, pp. 148–157, 2017, doi: 10.37676/jmi.v12i2.418.
- [12] N L G P Suwirmayanti and I. G. A. D. Saryanti, “Penerapan Teknik Clustering Untuk Pengelompokan Konsetrasi Mahasiswa Dengan Metode K-Means,” *Sintesa*, pp. 519–526, 2019, [Online]. Available: <http://jurnal.undhirabali.ac.id/index.php/sintesa/article/view/884>