

IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL

Muhammad Rafi Nahjan, Nono Heryana, Apriade Voutama

Program Studi Sistem Informasi S1, Fakultas Ilmu Komputer

Universitas Singaperbangsa Karawang, Jalan HS. Ronggo Waluyo, Karawang, Indonesia

11910631250024@student.unsika.ac.id

ABSTRAK

Oj Cell merupakan sebuah toko yang bergerak dalam bidang penjualan produk digital seperti pulsa, topup e-wallet dan paket data, Oj Cell juga memiliki toko offline di daerah Tegal, Jawa Tengah. Toko offline Oj Cell menjual produk tersebut dan juga menjual beberapa produk dalam bentuk voucher dan juga ada beberapa aksesoris dan sparepart gadget. Namun demikian dari berbagai jenis produk digital tersebut ada perbedaan keuntungan dan perbedaan soal tingkat terjualnya produk tersebut. Dibuatnya penelitian ini diharapkan dapat menghasilkan sebuah data yang dapat dipakai untuk kepentingan pemasaran yang nantinya diharapkan dapat membuat toko berkembang. Dengan metode Clustering K-Means dipakai untuk mengelompokan jenis transaksi yang sangat laris, laris dan kurang laris. Yang membuat kita tahu apa saja provider atau paket data yang digemari pada masyarakat sekitar Oj Cell, penelitian ini juga memakai aplikasi RapidMiner agar mempermudah dan membuat hasil lebih akurat.

Kata kunci: *Data Mining, Rapid Miner, Clustering K-Means.*

1. PENDAHULUAN

Usaha di bidang konter saat ini sangat digemari banyak orang. Terutama yang mengedepankan penjualan kuota, pulsa dan kebutuhan digital lainnya. Di era digital sekarang ini memang hal yang wajar bagi setiap orang untuk membeli kuota karena kuota sekarang ini sudah menjadi kebutuhan pokok bagi setiap orang. Salah satu konter yang mengedepankan penjualan kuota adalah OJ Cell, Oj Cell berlokasi di daerah Tegal, Jawa Tengah yang dimana keberadaan toko hp atau yang kita tahu dengan sebutan “Konter HP” itu masih sedikit, jadi karena tergolong bisnis yang jarang ditemukan pada daerahnya membuat Oj Cell memiliki omset atau keuntungan yang tidak rendah.

Tetapi jarang bukan berarti tidak ada pesaing, di suatu daerah pasti terdapat konter hp lainnya yang merupakan pesaing bisnis yang membuat pemilik Oj Cell khawatir akan bisnisnya akan tersaingi dan imbasnya adalah omset menurun, oleh karena itu pemilik Oj Cell ingin mencari cara agar Oj Cell dapat memberikan promo yang membuat pelanggan tertarik ingin bertransaksi atau membeli kebutuhan digital tersebut di Oj Cell. Oleh karena itu maka, dilakukannya penelitian ini agar data yang ada dapat digunakan dan diharapkan nantinya dapat dimanfaatkan untuk pengembangan strategi pemasaran. Seharusnya data yang telah tersedia dapat dijadikan sebagai sistem pengambilan keputusan untuk solusi bisnis serta dukungan di bidang teknologi yang merupakan penyebab munculnya suatu teknologi yaitu data mining.

Data mining adalah proses mengumpulkan dan mengolah data yang dimaksud untuk mendapatkan atau mengekstrak informasi penting dalam data tersebut [1]. Dan rapidminer adalah sebuah tools atau

software untuk mempermudah pengolahan data, rapid miner tersedia dalam bentuk software free to use atau gratis jika pemakaiannya untuk kepentingan edukasi atau pembelajaran [2].

Adapun Maksud dari penelitian ini adalah mengimplementasikan konsep data mining dengan algoritma Clustering K-Means membagi data transaksi Oj Cell kedalam tiga cluster, yang dimana tiga cluster tersebut untuk menentukan produk yang paling laris, laris dan kurang laris sebagai bahan tolak ukur untuk meningkatkan penjualan [3].

Sedangkan tujuan dari penelitian ini adalah membantu Oj Cell dalam menentukan produk penjualan yang termasuk kedalam kategori yang paling laris, laris dan kurang laris. Dimana data tersebut dapat menjadi sebuah acuan untuk Oj Cell dalam menentukan strategi pemasaran ataupun promosi yang bermanfaat untuk perkembangan Oj Cell.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah proses mengekstrak informasi yang berguna dari data yang besar dan kompleks. Ini mencakup teknik statistik, algoritma, dan proses yang digunakan untuk menemukan pola dan hubungan dalam data [7]. Tujuan dari data mining adalah untuk menemukan informasi yang berguna dan mengubahnya menjadi bentuk yang dapat diterima oleh pengguna. Data mining digunakan dalam berbagai bidang, seperti bisnis, ilmu pengetahuan, teknologi informasi, dan sebagainya untuk menemukan pola dalam data yang dapat digunakan untuk membuat keputusan yang lebih baik.

2.2. Clustering

Clustering adalah metode pembelajaran mesin yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau kluster yang sesuai. Tujuan dari clustering adalah untuk menemukan struktur di data yang tidak diketahui sebelumnya [7]. Ada banyak algoritma clustering yang tersedia, seperti K-Means, Hierarchical Clustering, dan DBSCAN. K-Means adalah algoritma yang paling populer dan sederhana, yang mengelompokkan data ke dalam jumlah kluster yang ditentukan sebelumnya [8]. Hierarchical Clustering mengelompokkan data ke dalam hierarki kluster, sedangkan DBSCAN mengelompokkan data berdasarkan jarak spasial.

2.3. K-Means

K-Means adalah salah satu algoritma clustering yang paling populer dan sederhana. Algoritma ini bekerja dengan cara mengelompokkan data ke dalam jumlah kluster yang ditentukan sebelumnya. Proses K-Means diawali dengan pemilihan jumlah kluster yang diinginkan (k) dan pemilihan titik awal yang digunakan sebagai centroid dari masing-masing kluster. Kemudian, setiap data akan diklasifikasikan ke dalam kluster yang centroidnya paling dekat dengan data tersebut. Setelah semua data diklasifikasikan, centroid dari setiap kluster akan diperbarui berdasarkan rata-rata dari semua data yang terklasifikasi ke dalam kluster tersebut. Proses ini akan diulangi beberapa kali hingga tidak ada perubahan lagi pada klasifikasi data [3].

K-Means memiliki beberapa kelemahan, seperti sensitif terhadap pemilihan titik awal centroid dan tidak dapat mengelompokkan data yang memiliki bentuk kluster yang tidak berbentuk bulat. Namun, algoritma ini masih sering digunakan karena mudah diimplementasikan dan cepat dalam mengelompokkan data besar.

2.4. Rapidminer

RapidMiner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan beragam alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. RapidMiner dirancang untuk mudah digunakan dan memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan tanpa pengalaman pemrograman [1].

RapidMiner menawarkan antarmuka drag-and-drop yang memungkinkan pengguna untuk membangun alur kerja untuk memproses dan menganalisis data. Ini mendukung beragam sumber data, termasuk file datar, basis data, dan platform big data seperti Hadoop dan Spark. Perangkat lunak ini juga mencakup beragam operator yang sudah dibangun, yang merupakan blok bangunan dari alur kerja, yang mencakup semua tahap proses data mining, seperti pembersihan data, pemilihan fitur, dan pemodelan.

3. METODE PENELITIAN

Metode penelitian disusun secara sistematis, penelitian ini menggunakan data transaksi Oj Cell dalam Bulan April 2022. Dari data yang dikumpulkan tersebut didapatkan 139 record transaksi yang valid dan telah diproses dan telah memenuhi kriteria penelitian ini. Dalam penelitian ini juga diperlukan software dan hardware untuk memproses penelitian agar mendapatkan hasil penelitian yang akurat dan maksimal yaitu:

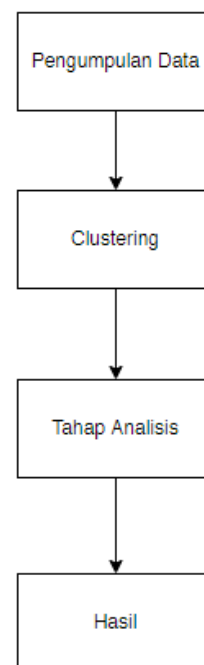
Tabel 1. Spesifikasi *Software*.

Software
Sistem Operasi : Windows 10 Pro Versi 2JH2
<i>Software Data Mining</i> : <i>RapidMiner Studio</i> 9.10.1.0

Tabel 2. Spesifikasi *hardware*.

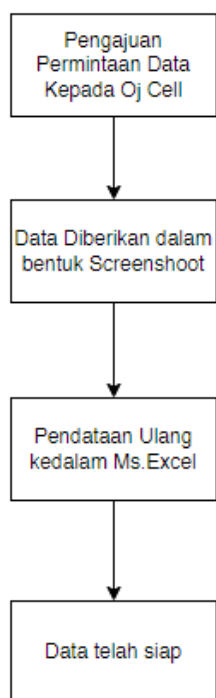
Hardware
CPU : Intel® core™ I5- 7300HQ 2.5 GHz

Tahapan dalam penelitian ini adalah sebagai berikut:



Gambar 1. Alur metode penelitian.

3.1. Tahap Pengumpulan Data



Gambar 2. Alur pengumpulan data.

Dalam penelitian ini sumber data yang digunakan adalah data Oj Cell yang diberikan dalam bentuk Image atau foto, data tersebut lalu diproses kembali atau diinput kedalam Microsoft Excel agar data dapat digunakan untuk proses penelitian, lalu data yang telah diinput ke Microsoft Excel dan dirangkum sedemikian rupa agar data dapat di proses dengan metode clustering, sehingga menjadi data yang bermanfaat.

3.2. Tahap Clustering

Pada tahapan ini akan dilakukan proses *clustering* dengan menentukan bobot- bobot nilai *cluster* dan yang nantinya akan dianalisis lebih lanjut untuk mendapatkan hasil yang sesuai. Tahap *clustering* ini menggunakan perangkat lunak *Rapidminer*.

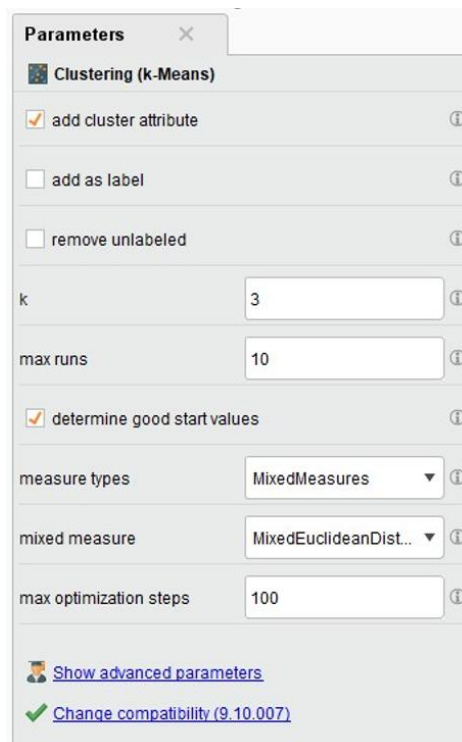
2.3. Tahap Analisis

Data yang telah di proses di perangkat lunak *Rapidminer* di analisis lebih lanjut untuk memahami dan nantinya akan menjadi kesimpulan dari penelitian ini.

4. HASIL DAN PEMBAHASAN

Setelah menginput dan melakukan pengecekan data, peneliti mulai untuk melakukan pemodelan terhadap dataset Oj Cell. Algoritma yang digunakan

adalah *k-means* dan dengan bantuan perangkat lunak *rapidminer*. dimulai dengan input data ke dalam *rapidminer* kemudian data di proses dengan metode *k-means* dengan parameter $k = 3$, *measure type = mixed measure* hal ini dilakukan karena data Oj Cell ada yang berupa kode dan ada yang berupa angka dan menggunakan *Euclidean Distance* untuk standar jarak antar titik klaster.



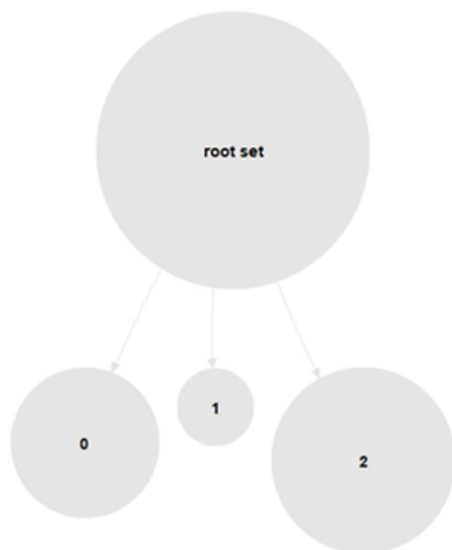
Gambar 3. Parameter pada Rapidminer

Selanjutnya data di proses dengan *rapidminer* dan menghasilkan 3 model klaster yang berisi *cluster 0 = 50 items, cluster 1 = 24 items, cluster 2 = 65 items*. *Cluster* dimulai dari 0 karena pada bahasa pemrograman angka 0 adalah angka pertama dari urutan penomoran.

Tabel 3. Hasil dari proses rapidminer

Cluster Model	Jumlah
Cluster 0	50 Items
Cluster 1	24 Items
Cluster 2	65 Items
Total Cluster	139 items

Adapun ada hasil berbentuk *tree* dalam opsi yang disediakan oleh *rapidminer* untuk memberikan gambaran seberapa besar perbandingan antar cluster dengan semua data. dengan opsi ini kita dapat melihat sebuah pemecahan dari data induk atau *root set*, menjadi beberapa data turunan yang dinamai cluster.



Gambar 4. Perbandingan data dengan model tree

Adapun hasil dari proses sebelumnya menjadi menjadi data yang informatif tetapi data yang dihasilkan belum ringkas jadi penulis merangkumnya kedalam sebuah tabel dibawah.

Tabel 4. Kesimpulan hasil dari proses pengolahan data pada rapidminer.

Atribut	Jumlah	Profit
Cluster 0	50	Rp1.607,60
Cluster 1	24	Rp5.549,95
Cluster2	65	Rp2.729,69

5. KESIMPULAN DAN SARAN

Dari penelitian yang sudah dilakukan maka dapat disimpulkan bahwa transaksi yang termasuk klaster 2 sangat diminati dengan jumlah 65 data, sedangkan transaksi klaster 1 adalah yang paling kurang diminati dengan jumlah 24 data. Disamping karena keuntungan yang diambil besar dan dirasa cukup mahal jadi hanya terdapat sedikit konsumen yang membeli produk tersebut. Transaksi pada klaster 0 dirasa cukup dengan jumlah sebanyak 50 data, karena tidak berbeda jauh dengan transaksi klaster 2.

Dengan hasil penelitian ini maka dapat disimpulkan bahwa transaksi yang masuk dalam klaster 2 adalah transaksi yang diminati yaitu yang berisi transaksi isi pulsa dan pembelian paket data dengan nominal diatas Rp 50.000 dan memberi profit rata-rata Rp.2.729 per transaksi.

Dengan dibuatnya penelitian ini maka diharapkan dapat menjadi gambaran untuk Oj Cell dalam menyusun strategi pemasarannya. Adapun hal yang peneliti ingin sarankan adalah sesuaikan lagi harga dan profit yang di ambil dalam transaksi yang termasuk klaster 1. dan juga membuka layanan online 24 jam seperti web atau bot sosial media seperti bot whatsapp yang memiliki fitur transaksi otomatis juga dapat membantu menarik minat dari pelanggan.

DAFTAR PUSTAKA

[1] Rapidminer, "RapidMiner for Academics," [https://rapidminer.com/platform/educational/]. diakses 5 Agustus 2022

[2] R. Setiawan, "Apa itu Data Mining dan Bagaimana Metodenya?" [https://www.dicoding.com/blog/apa-itu-data-mining/], 2022.

[3] H. Haviluddin, S. J. Patandianan, G. M. Putra, N. Puspitasari, and H. S. Pakpahan, "Implementasi Metode K-Means Untuk Pengelompokan Rekomendasi Tugas Akhir," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 16, no. 1, 2021, doi: 10.30872/jim.v16i1.5182.

[4] A. L. R. Putri and N. Dwidayati, "Analisa Perbandingan K-Means Dan Fuzzy C-Means Dalam Pengelompokan Daerah Penyebaran Covid-19 Indonesia," *UNNES J. Math.*, vol. 10, no. 2, 2021.

[5] S. Hussein, "K-means Clustering: Pengertian, Metode Algoritma, Beserta Contoh," [https://geospasialis.com/k-Means-Clustering/], 2021.

[6] N. Normah, S. Nurajizah, and A. Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *J. Tek. Komput.*, vol. 7, no. 2, 2021, doi: 10.31294/jtk.v7i2.10553.

[7] R. W. Sari, A. Wanto, and A. P. Windarto, "IMPLEMENTASI RAPIDMINER DENGAN METODE K-MEANS (STUDY KASUS: IMUNISASI CAMPAK PADA BALITA BERDASARKAN PROVINSI)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, 2018, doi: 10.30865/komik.v2i1.930.

[8] Ardiyansyah, P. A. Rahayuningsih, and R. Maulana, "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner," *J. Khatulistiwa Inform.*, 2018.