



Penerapan Data Mining Dalam Pengelompokan Data Penjualan Paket Internet di Telkomsel Authorized Partner (TAP) Deli Tua Dengan Algoritma K-Means

Tesa Aurelia Siregar^{1*}, Mesran², Dito Putro Utomo³

Universitas Budi Darma, Medan, Sumatera Utara, Indonesia

Email: ^{1*}tesaaurelia951@gmail.com, ²mesran.skompom@gmail.com, ³ditoputro12@gmail.com

Abstract

The increase in internet package sales made Telkomsel Authorized Partner (TAP) Deli Tua have to be more careful in maintaining internet package products. The process that can be done to maintain internet package products at TAP Deli Tua is to group the products into two, namely Best Selling Products and Non Selling Products. The process of determining best-selling products and non-selling products based on sales results obtained by resellers. However, the grouping process must be carried out carefully and precisely, so that the results obtained are as expected by TAP Deli Tua. If a grouping process is not carried out carefully and precisely, TAP Deli TAP Deli Tua will suffer losses. So with that, the process of grouping internet packages can be done with data mining. In data mining there is a technique that can help in grouping internet packages, namely the clustering technique. In the clustering technique there is also an algorithm that obtains results in clusters, namely the K-Means algorithm. The K-Means algorithm is an algorithm that can obtain different groups of data in a dataset that does not have a label. The k-means algorithm can also divide a different group into several groups based on their characteristics where if the characteristics have something in common then they can be put together or united. So that this can help TAP Deli Tua in grouping internet package data at TAP Deli Tua. So by doing this research with the k-means algorithm it can be shown that data mining can obtain the results of a more appropriate decision in grouping internet package data so as to obtain results from clusters of best-selling products and non-selling products which have 15 internet package sales data obtaining 8 products selling in clusters 0 and 7 non-selling products in cluster 1. And this research has also been tested using a tool, namely RapidMiner.

Keywords: Data mining, K-Means, Clustering, Internet Package Sales, TAP Deli Tua

Abstrak

Meningkatnya penjualan paket internet membuat Telkomsel Authorized Partner (TAP) Deli Tua harus lebih teliti dalam mempertahankan produk paket internet. Proses yang dapat dilakukan untuk mempertahankan produk paket internet pada TAP Deli Tua ialah dengan mengelompokkan produk menjadi dua yaitu Produk Laris dan Produk Non Laris. Proses dalam menentukan produk laris dan produk non laris berdasarkan hasil penjualan yang di peroleh reseller. Akan tetapi dalam proses pengelompokan harus dilakukan secara teliti dan tepat, agar hasil yang di peroleh sesuai dengan yang diharapkan oleh TAP Deli Tua. Jika suatu proses pengelompokan tidak dilakukan secara teliti dan tepat, akan mengalami kerugian pada TAP Deli TAP Deli Tua. Maka dengan itu, proses pengelompokan paket internet dapat dilakukan dengan data mining. Pada data mining terdapat suatu teknik yang dapat membantu dalam pengelompokan paket internet yaitu teknik clustering. Didalam teknik clustering juga terdapat suatu algoritma yang memperoleh hasil dalam cluster yaitu algoritma K-Means. Algoritma K-Means salah satu algoritma yang dapat memperoleh kelompok yang berbeda dari data dalam sebuah dataset yang tidak memiliki sebuah label. Algoritma k-means juga dapat membagi suatu kelompok yang berbeda menjadi beberapa kelompok berdasarkan karakteristiknya yang dimana jika karakteristik memiliki kesamaan maka dapat dijadikan satu atau di persatukan. Sehingga hal tersebut dapat membantu TAP Deli Tua dalam pengelompokan data paket internet pada TAP Deli Tua. Maka dengan dilakukan penelitian ini dengan algoritma k-means dapat memperlihatkan bahwa data mining dapat memperoleh hasil suatu keputusan yang lebih tepat

dalam pengelompokan data paket internet sehingga memperoleh hasil dari cluster dari produk laris dan non produk laris yang terdapat 15 data penjualan paket internet memperoleh 8 produk laris yang berada di cluster 0 dan 7 non-produk laris yang berada di cluster 1. Dan penelitian ini juga sudah di uji dengan menggunakan sebuah tool yaitu RapidMiner.

Kata kunci: Data mining, K-Means, Clustering, Data Penjualan Paket Internet, TAP Deli Tua

1. PENDAHULUAN

Data penjualan merupakan suatu kumpulan informasi yang sudah dianalisis dalam bentuk laporan dari hasil penjualan. Bentuk laporan hasil penjualan dapat berupa harian, bulanan, dan bahkan tahunan. Dalam melakukan transaksi jual beli, laporan data penjualan sangat penting agar kita dapat mengetahui berapa hasil dari penjualan produk yang sudah habis terjual[1]-[2]. Salah satu produk yang dipasarkan ialah paket internet. Paket internet merupakan salah satu produk yang banyak dipasarkan dimana saja. Maka dengan itu produk tersebut sangat diperlukan dalam mendata hasil dari penjualan setiap harinya agar tidak terjadinya kehabisan stok. Sama seperti halnya di lakukan oleh Telkomsel Authorized Partner (TAP) Deli Tua.

Telkomsel Authorized Partner (TAP) Deli Tua merupakan salah satu perusahaan kecil yang bergerak dibidang layanan internet. TAP Deli Tua salah satu proses bisnis yang dimana TAP Deli Tua memasarkan salah satu produk internet yang cukup luas karena jangkauannya bisa melebihi batas maksimum dalam pemasaran. Proses yang dapat dilakukan agar tidak terjadinya kehabisan stok pada TAP Deli Tua ialah dengan cara pengelompokan terhadap data penjualan menjadi 2 kelompok yaitu produk laris dan produk non laris berdasarkan data penjualan yang dihasilkan oleh TAP Deli Tua. Produk laris yaitu suatu kelompok yang memberikan keuntungan bagi TAP Deli Tua, karena produk habis terjual[4]. Maka dengan itu TAP Deli Tua harus lebih memperbanyak produk yang diinginkan oleh konsumen, dan juga harus mengurangi produk yang kurang diminati oleh konsumen. Karena semakin banyaknya produk yang dipasarkan TAP Deli Tua akan memberikan keuntungan yang lebih besar. Maka dengan itu sangat dibutuhkan suatu pengolahan dari hasil penjualan agar TAP Deli Tua mengetahui hasil produk yang habis terjual.

Dengan meningkatnya hasil penjualan paket internet pada TAP Deli Tua, maka harus lebih teliti dalam pengolahan data hasil penjualan tersebut agar tidak terjadinya kehabisan stok yang mengakibatkan kerugian pada TAP Deli Tua. Dalam penentuan produk terlaris penjualan paket internet, jika pengelompokan tersebut dilakukan secara tidak tepat maka terjadinya kerugian pada TAP Deli Tua. Sehingga tingkat penjualan pada TAP Deli Tua akan semakin rendah dan mengakibatkan kerugian yang cukup besar pada TAP Deli Tua. TAP Deli Tua juga harus lebih teliti dalam pengolahan data agar tidak terjadinya hal-hal yang merugikan bagi TAP Deli Tua. Dalam melakukan proses pengelompokan, harus mengumpulkan data penjualan paket internet agar dapat dianalisis dan menghasilkan informasi yang akan membantu dalam pengambilan keputusan[3].

Proses penyelesaian masalah dalam pengelompokan dapat dilakukan dengan melihat dari data penjualan pada TAP Deli Tua. Dimana pada data penjualan

tersimpan seluruh data penjualan yang terjadi pada TAP Deli Tua. Dengan itu TAP Deli Tua harus mengumpulkan data penjualan agar dapat di olah dengan baik, sehingga menghasilkan informasi-informasi tersebut. Dimana informasi tersebut akan membantu pihak TAP Deli Tua dalam pengambilan keputusan serta penyelesaian pengelompokan pada data penjualan paket internet. Maka dalam melakukan pengelompokan data penjualan paket internet dibutuhkan suatu peran data mining dalam membantu penyelesaian pengelompokan[4]-[5].

Data mining merupakan salah satu proses dalam menemukan informasi yang dihasilkan oleh data yang telah di olah menjadi data yang besar agar data tersebut dapat dipergunakan untuk membantu dalam pengambilan keputusan dimasa depan. Data mining juga salah satu kegiatan pengumpulan data yang dapat dilakukan dengan salah satu proses perhitungan atau aritmatika yang terdapat pada teknik dalam menemukan hasil pengelompokan data tersebut[6]-[7]. Di dalam data mining terdapat teknik yang dimana teknik tersebut ialah teknik clustering. Clustering merupakan pengumpulan suatu observasi secara otomatis yang dimasukan ke dalam kelompok, yang dimana jumlah kelompok belum diketahui. Pada teknik clustering ini, terdapat beberapa algoritma salah satunya ialah algoritma k-means[8]-[9]-[10]-[11]. Dalam penelitian ini penulis menggunakan algoritma yang terdapat pada data mining yaitu algoritma K-Means, yang merupakan pengelompokan untuk mempartisi kumpulan dataset menjadi beberapa *cluster* dan terpisah sehingga kriteria pengelompokan tersebut dioptimalkan[12]. Kriteria yang paling banyak digunakan ialah kriteria dengan kesalahan pengelompokan untuk setiap titik dalam melakukan perhitungan jarak kuadratnya dari pusat *cluster* yang sesuai. Dan juga melakukan penjumlahan untuk jarak pada semua titik dalam pengumpulan sebuah dataset[13]. Pada penelitian ini penulis memilih menggunakan algoritma K-Means karena algoritma tersebut sangat cocok untuk menyelesaikan permasalahan dan juga langkah dalam pengerjaanya sangat mudah dipahami dari pada algoritma lainnya.

Pada penelitian ini, terdapat beberapa penelitian terkait yang menjelaskan mengenai metode K-Means sehingga dapat dijadikan sebuah artikel pendukung atau referensi agar penulis dapat menyelesaikan sebuah penelitian yang sedang dilakukan dengan menerapkan metode K-Means. Penelitian yang telah di telaah oleh Reza Gustrianda, dan Dadang Iskandar Mulyana dilakukan pada tahun 2022 yang membahas tentang produk unggulan dengan menerapkan sebuah algoritma yang akan mendapatkan hasil *clustering* yang sesuai dengan data yang sudah dibuat oleh penulis. Algoritma K-Means digunakan sebagai pengelompokan sebuah data yang akan memperoleh hasil *cluster* yang sesuai.[14]. Penelitian yang telah di kupas oleh Muhammad Syahril, dkk pada tahun 2023 membahas mengenai rekomendasi jurusan dengan menerapkan sebuah metode yang digunakan untuk pengelompokan data, metode tersebut yaitu K-Means. Dalam penelitian ini, terdapat 60 data siswa yang akan digunakan untuk pengelompokan dengan menerapkan algoritma K-Means[15]. Penelitian yang dilakukan Andri Yani, dkk pada tahun 2023 mengenai data penjualan dengan menggunakan algoritma K-Means. Tujuan dari menggunakan algoritma K-Means, yaitu agar mempermudah dalam memperoleh hasil pengelompokan data penjualan tersebut. Maka dengan

itu, penelitian ini menghasilkan *Clustering* 1 terdapat 29 item, *Clustering* 2 terdapat 29 item, dan *Clustering* 3 terdapat 2 item[16]. Penelitian yang ditelaah Elsa Ramadanti, dan Muhamad Muslih dilakukan pada tahun 2022 mengenai *Clustering* pada populasi ayam petelur di Indonesia. Pada penelitian ini menerapkan algoritma algoritma K-Means. Metode K-Means salah satu metode yang mudah dalam menemukan hasil dari clustering tersebut[17]. Selanjutnya penelitian yang kupas Cornelia Selvi Dinta, dkk pada tahun 2022 yang berjudul "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi" dapat diberikan kesimpulan dalam menerapkan algoritma K-Means yang akan memperoleh hasil yang diinginkan penulis[18]. Dari penjelasan diatas, terdapat masalah yang terjadi pada penelitian ini sehingga penulis harus memecahkan permasalahan yang terjadi didalam penelitian ini. Permasalahannya yaitu pengelompokan data penjualan terbagi menjadi dua diantaranya produk laris dan produk *non* laris.

2. METODOLOGI PENELITIAN

2.1. Data Mining

Data mining merupakan salah satu cara dalam menemukan suatu informasi yang didapatkan dari sebuah data yang sudah di olah menjadi data yang lebih besar dan data tersebut bisa dijadikan dalam pengambilan suatu keputusan dimasa depan. [19]- [20].

2.2. Data Penjualan

Data penjualan merupakan suatu kumpulan informasi yang sudah dianalisis dalam bentuk laporan dari hasil penjualan. Bentuk laporan hasil penjualan dapat berupa harian, bulanan, dan bahkan tahunan. Dalam melakukan transaksi jual beli, laporan data penjualan sangat penting agar kita dapat mengetahui berapa hasil dari penjualan produk yang sudah habis terjual[1].

2.3. Data Mining

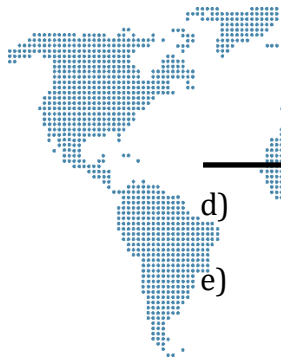
Algoritma K-Means yaitu salah satu algoritma yang termasuk kedalam *Non Hierarchical Clustering* yang mempartisi kumpulan dataset menjadi beberapa *cluster* dan terpisah sehingga kriteria pengelompokan tersebut dioptimalkan[21]-[22]-[23]-[10]-. Dalam algoritma K-Means terdapat beberapa langkah-langkah penyelesaian sebagai berikut:

- Menentukan jumlah *clustering* yang akan di bentuk.
- Menentukan nilai *cetroid* awal secara acak atau random pada data yang sudah ada. Rumus yang akan di gunakan sebagai berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (1)$$

- Menghitung jarak data dengan *cetroid* awal. Dapat menggunakan rumus dibawah ini:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (2)$$



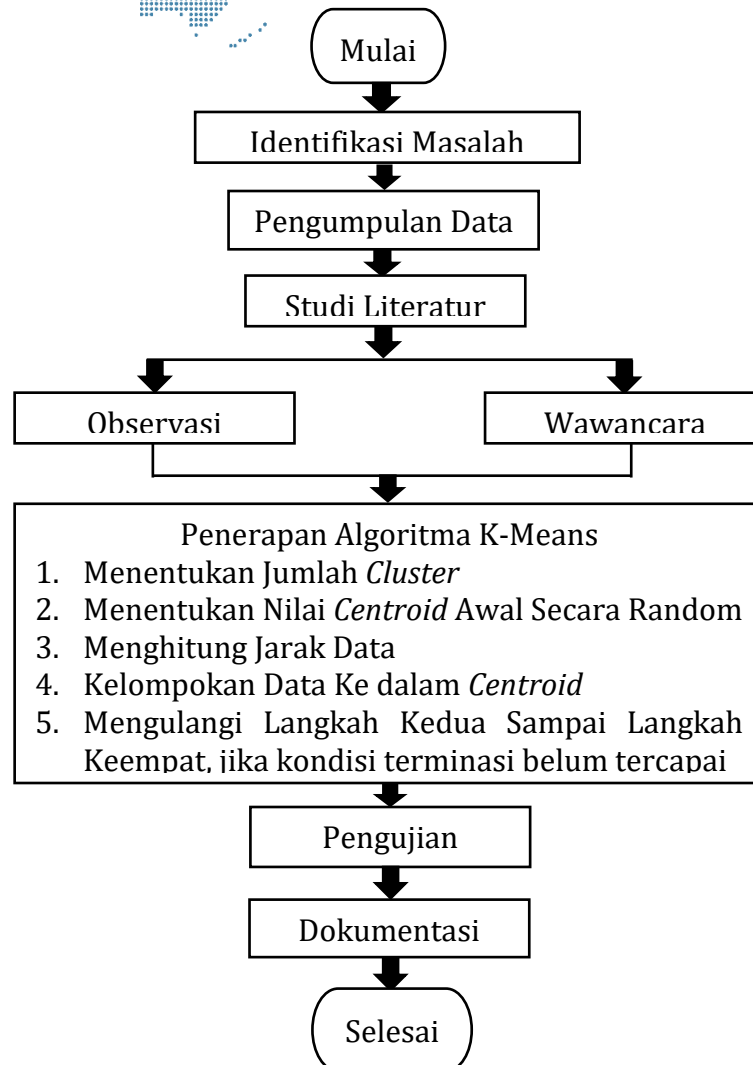
- d) Kelompokkan data kedalam *centroid* yang bersangkutan berdasarkan jarak terdekat untuk membentuk sebuah *cluster*.
- e) Jika kondisi terminasi belum tercapai maka dapat di ulangi pengerjaan langkah ke-2 sampai langkah ke-4 sampai kondisi nya sudah tercapai baru berhenti dalam pengerjaan tersebut.

2.4. Tahapan Penelitian

Adapun penjelasan dari setiap tahapan-tahapan yang dilakukan oleh penulis dapat dilihat di bawah ini sebagai berikut:

- a) Mengidentifikasi Masalah
Dalam melakukan setiap penelitian, hal utama yang sangat diperlukan ialah mengidentifikasi suatu permasalahan yang terjadi pada TAP Deli Tua agar dapat menemukan ide pokok dari objek yang diangkat oleh penulis.
- b) Literatur Review
Pada tahap literatur review, penulis harus melakukan sebuah pencarian referensi mengenai objek penelitian agar pengetahuan penulis mengenai objek tersebut semakin luas. Literatur review dapat dilakukan dengan pencarian suatu referensi yang didapat pada setiap perpustakaan atau pun google scholar.
- c) Pengumpulan Data
Dalam melakukan sebuah penelitian, penulis harus melakukan pencarian atau pengumpulan data yang berkaitan dengan objek yang sedang di teliti. Pengumpul data dapat dilakukan dengan dua cara yaitu observasi dan wawancara. Pengumpulan data ini dilakukan untuk acuan atau referensi dalam mendukung penyelesaian permasalahan ini.
 - 1) Observasi
Observasi merupakan proses yang dilakukan dalam pencarian data mengenai objek yang diangkat oleh penulis, yang dimana penulis tersebut langsung mengamatin penelitian tersebut langsung ke lapangan.
 - 2) Wawancara
Wawancara merupakan proses yang dilakukan oleh penulis dalam informasi mengenai objek penelitian yang diangkat, akan tetapi hal tersebut langsung di tanyak atau interview kepada manager TAP Delia Tua tersebut.
- d) Penerapan Algoritma K-Means
Pada penelitian ini, sangat di perlukan dalam penerapan metode agar mendapatkan atau memperoleh hasil yang sesuai penulis inginkan. Dimana penelitian ini menggunakan sebuah algoritma yang dapat menemukan hasil *cluster* yang sesuai yaitu K-Means.
- e) Pengujian
Dalam penelitian ini, penulis melakukan pengujian algoritma K-Means menggunakan sebuah *tools* yang dimana *tools* tersebut sangat membantu penulis dalam menemukan atau memperoleh hasil dari *cluster*. *Tools* yang digunakan oleh penulis ialah RapidMiner.
- f) Laporan Penelitian
Tahap terakhir ialah pembuatan laporan penelitian yang dimana setelah penulis menyelesaikan penelitian tersebut, wajib membuat laporan penelitian.

Berikut ini dalam penyelesaian penelitian ini, terdapat beberapa tahapan dalam penelitian dapat dilihat pada gambar dibawah ini sebagai berikut:



Gambar 1. Dijelaskan Bahwa Diatas Terdapat Kerangka Dalam Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

3.1. *Preprocessing* Data

Sebelum melakukan *preprocessing* data di perlukan beberapa sampel data yang ingin di proses. Maka dapat dilihat dibawah sebagai berikut:

Tabel 1. Dijelaskan Bahwa Saat Melakukan *Preprocessing* Diperlukan Sampel Data Order Reseller TAP Deli Tua Tahun 2022

No	Nama_Produk	Januari	Februari	Maret	April	Mei	...	Desember
1	1.5gb3d	20000	25000	26600	28000	30000	...	29700
2	1.5gb5d	25000	27000	21000	26000	18000	...	13000
3	1.5gb7d	17000	15500	16500	20000	14300	...	10000

No	Nama_Produk	Januari	Februari	Maret	April	Mei	...	Desember
4	2gb3d	15000	9000	9600	10000	13000	...	7700
...	...	...	...	...	...	...	...	...
15	vo_kosong	600	640	445	550	750	...	670

Setelah data yang akan di proses sudah tersedia, maka selanjutnya dilakukan tahap preprocessing data. Nilai range tersebut dapat dilihat pada Tabel 2 yang telah dibuat oleh penulis sebagai berikut:

Tabel 1. Dijelaskan Terdapat Range Nilai Bobot Untuk Melakukan *Preprocessing Data*

Range Produk	Nilai
≥ 30000	14
≥ 20000	12
≥ 10000	10
≥ 1000	8
≥ 100	6
> 0	4
0	0

Dimana dalam pemberian nilai bobot kriteria tidak ada aturan baku, akan tetapi jika nilai bobot pada setiap kriteria semakin tinggi, maka attribute tersebut semakin penting. Maka dengan itu jika nilai attributenya 30000 maka nilai bobotnya 14 dan begitu juga seterusnya. Maka dengan itu hasil dari data yang telah di *preprocessing* dapat dilihat pada Tabel 3 sebagai berikut:

Tabel 2. Dijelaskan Bahwa Dari Terdapat Hasil *Preprocessing Data* Penjualan Paket Internet

Nama_Produk	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sept	Okt	Nov	Des
1.5gb3d	12	12	12	12	14	14	12	12	14	14	14	12
1.5gb5d	12	12	12	12	10	10	12	10	10	12	10	10
1.5gb7d	10	10	10	12	10	10	10	12	10	10	10	10
2gb3d	10	8	8	10	10	10	8	8	8	8	8	8
2gb5d	10	10	10	12	12	10	10	10	10	10	10	10
2.5gb5d	10	10	10	10	12	10	10	10	10	10	10	10
3gb3d	8	8	8	8	8	8	8	8	8	8	8	8
3gb7d	10	10	10	12	10	10	10	10	10	10	10	10
3.5gb3d	8	8	8	8	8	8	8	8	8	8	8	8
...	...	...	...	...	...	...	...	...	...	...	...	...
vo_kosong	6	6	6	6	6	6	8	8	8	1	4	6

3.2. Penerapan Algoritma K-Means

Algoritma K-Means merupakan salah satu algoritma yang termasuk kedalam *Non Hierarchical Clustering* yang mempartisi kumpulan dataset menjadi beberapa *cluster* dan terpisah sehingga kriteria pengelompokan tersebut dioptimalkan.

a) Menentukan jumlah *clustering* yang akan di bentuk

Dalam proses pengelompokan pada algoritma K-Means, terlebih dahulu harus menentukan jumlah kelompok.



Tabel 3. Dijelaskan Bahwa Terdapat 2 Kelompok Data Penjualan

No	Cluster	Keterangan
1	C0	Produk Laris
2	C1	Produk <i>Non</i> Laris

- b) Mengalokasikan data secara random/acak
Dalam memperoleh nilai *centroid* awal pada algoritma K-Means dilakukan dengan cara menentukan nilai secara random/acak.

Tabel 4. Dijelaskan Bahwa Sebelum Melakukan Perhitungan K-Means Harus Menentukan Nilai *Centroid* Awal

C0	1.5gb3d	12	12	12	12	14	14	12	12	14	14	14	12
C1	vo_telepon	6	6	6	6	6	6	6	6	6	6	6	6

- c) Menghitung jarak data dengan *centroid* awal
Pada penelitian ini, penulis akan melakukan perhitungan iterasi ke-1 pada *cluster* C0 dan C1.

Cluster 0

$$d(x_{1i}, c_0) = \sqrt{(12 - 12)^2 + (12 - 12)^2 + (12 - 12)^2 + (12 - 12)^2 + (14 - 14)^2 + (14 - 14)^2 + (12 - 12)^2 + (12 - 12)^2 + (14 - 14)^2 + (14 - 14)^2 + (14 - 14)^2 + (12 - 12)^2} = 0$$
$$d(x_{2i}, c_0) = \sqrt{(12 - 12)^2 + (12 - 12)^2 + (12 - 12)^2 + (12 - 12)^2 + (10 - 14)^2 + (10 - 14)^2 + (12 - 12)^2 + (10 - 12)^2 + (10 - 14)^2 + (12 - 14)^2 + (10 - 14)^2 + (10 - 12)^2} = 8,72$$

Berdasarkan pencarian iterasi ke-2 yang dimulai dengan *cluster* 0 dan dimana langkah dalam pengerjaanya sama, maka dengan itu penulis hanya membuat 2 cara saja dalam pengerjaanya. Setelah melakukan pencarian *cluster* 0 selanjutnya melakukan pencarian terhadap *cluster* 1 dapat dilihat sebagai berikut:

Cluster 1

$$d(x_{1i}, c_1) = \sqrt{(12 - 6)^2 + (12 - 6)^2 + (12 - 6)^2 + (12 - 6)^2 + (14 - 6)^2 + (14 - 6)^2 + (12 - 6)^2 + (12 - 6)^2 + (14 - 6)^2 + (14 - 6)^2 + (14 - 6)^2 + (12 - 6)^2} = 23,92$$
$$d(x_{2i}, c_1) = \sqrt{(12 - 6)^2 + (12 - 6)^2 + (12 - 6)^2 + (12 - 6)^2 + (10 - 6)^2 + (10 - 6)^2 + (12 - 6)^2 + (10 - 6)^2 + (10 - 6)^2 + (12 - 6)^2 + (10 - 6)^2 + (10 - 6)^2} = 17,66$$

Berdasarkan hasil dari perhitungan dengan algoritma K-Means dalam mencari hasil jarak terhadap nilai *centroid* awal di iterasi-1 pada *cluster* 0 dan 1, maka dengan itu dapat di lihat hasil dari pencarian nilai *cluster* 0 dan 1 pada iterasi ke-1 pada Tabel 6 sebagai berikut:

Tabel 6. Dijelaskan Bawah Terdapat Rekapitulasi Hasil Perhitungan Iterasi-1

nama_produk	dc0	dc1	Cluster
1.5gb3d	0	23,92	0
1.5gb5d	8,72	17,66	0
1.5gb7d	10	15,23	0
2gb3d	15,10	9,80	1
2gb5d	9,59	15,23	0
2.5gb5d	9,80	14,56	0
3gb3d	17,09	6,93	1
3gb7d	10,20	14,56	0
3.5gb3d	17,09	6,93	1
3.5gb7d	17,09	6,93	1
5gb3d	17,09	6,93	1
5gb7d	11,31	13,42	0
10gb30d	8,94	15,87	0
vo_telepon	23,92	0	1
vo_kosong	22,45	3,46	1

Setelah melakukan perhitungan iterasi ke-1, selanjutnya akan melakukan perhitungan iterasi ke-2. Langkah yang akan dilakukan sama seperti iterasi ke-1 harus membuat nilai centroid terlebih dahulu sebagai berikut:

Tabel 7. Dijelaskan Bahwa Sebelum Melakukan Perhitungan K-Means Harus Menentukan Nilai *Centroid* Awal

C0	10,50	10,50	10,50	11,25	11,00	10,75	10,75	10,50	10,25	10,75	10,75	10,25
C1	7,71	7,43	7,43	7,71	7,71	7,71	7,71	7,71	7,71	7,43	7,43	7,43

Cluster 0

$$d(x_{1i}, c_0) = \sqrt{(12 - 10,50)^2 + (12 - 10,50)^2 + (12 - 10,50)^2 + (12 - 11,25)^2 + (14 - 11,00)^2 + (14 - 10,75)^2 + (12 - 10,75)^2 + (12 - 10,50)^2 + (14 - 10,25)^2 + (14 - 10,75)^2 + (14 - 10,75)^2 + (12 - 10,25)^2} = 3,80$$

$$d(x_{2i}, c_0) = \sqrt{(12 - 10,50)^2 + (12 - 10,50)^2 + (12 - 10,50)^2 + (12 - 11,25)^2 + (10 - 11,00)^2 + (10 - 10,75)^2 + (12 - 10,75)^2 + (10 - 10,50)^2 + (10 - 10,25)^2 + (12 - 10,75)^2 + (10 - 10,75)^2 + (10 - 10,25)^2} = 3,60$$

Cluster 1

$$d(x_{1i}, c_1) = \sqrt{(12 - 7,71)^2 + (12 - 7,43)^2 + (12 - 7,43)^2 + (12 - 7,71)^2 + (14 - 7,71)^2 + (14 - 7,71)^2 + (12 - 7,71)^2 + (12 - 7,71)^2 + (14 - 7,71)^2 + (14 - 7,43)^2 + (14 - 7,43)^2 + (12 - 7,43)^2} = 18,47$$

$$d(x_2, c_1) = \sqrt{\begin{matrix} (12 - 7,71)^2 + (12 - 7,43)^2 + (12 - 7,43)^2 + \\ (12 - 7,71)^2 + (10 - 7,71)^2 + (10 - 7,71)^2 + \\ (12 - 7,71)^2 + (10 - 7,71)^2 + (10 - 7,71)^2 + \\ (12 - 7,43)^2 + (10 - 7,43)^2 + (10 - 7,43)^2 \end{matrix}} = 12,33$$

Maka dari hasil perhitungan iterasi 1 dan 2 dengan menggunakan algoritma K-Means tidak terjadinya perubahan cluster, sehingga perhitungan pada pengelompokan data penjualan paket internet di TAP Deli Tua dengan menggunakan algoritma K-Means telah selesai tidak ada pengulangan ke langkah awal lagi karna hasil nya sama seperti iterasi ke-1. Maka dengan itu hasil yang di dapatkan penulis dengan menggunakan algoritma K-Means dimulai dari hasil iterasi ke -1 hingga iterasi ke-2 dapat dilihat pada Tabel 8 sebagai berikut:

Tabel 8. Dijelaskan Bahwa Terdapat Hasil *Cluster* iterasi 1 dan iterasi 2

nama_produk	Iterasi	
	1	2
	<i>Cluster 0</i>	<i>Cluster 1</i>
1.5gb3d	0	0
1.5gb5d	0	0
1.5gb7d	0	0
2gb3d	1	1
2gb5d	0	0
2.5gb5d	0	0
3gb3d	1	1
3gb7d	0	0
5gb3d	1	1
5gb7d	0	0
10gb30d	0	0
Votelepon	1	1
Vokosong	1	1

Maka dengan itu akan menghasilkan 2 kelompok yaitu kelompok produk laris dan kelompok produk non laris dapat dilihat pada Tabel 9 sebagai berikut:

Tabel 9. Dijelaskan Bahwa Terdapat Hasil Akhir Perhitungan

<i>Cluster 0</i>	<i>Cluster 1</i>
1.5gb3d	2gb3d
1.5gb5d	3gb3d
1.5gb7d	3.5gb3d
2gb5d	3.5gb7d
2.5gb5d	5gb3d
3gb7d	votelepon
5gb7d	vokosong
10gb30d	

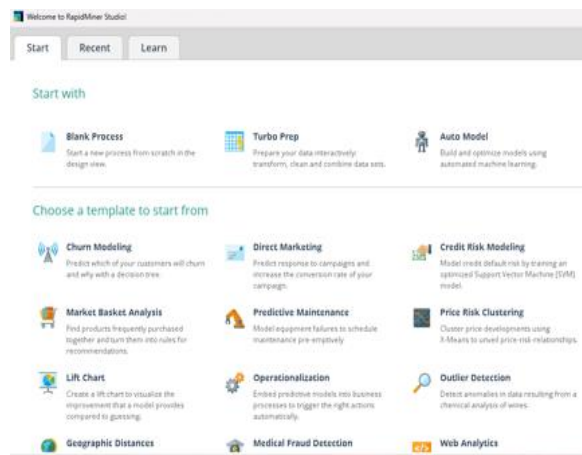
Dari Tabel 9 dapat dijelaskan bahwa *cluster 0* ialah kelompok produk laris sedangkan *cluster 1* ialah kelompok produk *non* laris. Maka dari 15 data penjualan paket internet terdapat 8 produk laris yang berada di *cluster 0* dan 7 produk *non* laris yang berada di *cluster 1*.

3.3. Pengujian

Pada penelitian yang telah dilakukan oleh penulis, maka dengan itu penulis akan menampilkan hasil pengujian dengan algoritma K-Means mengenai data penjualan paket internet pada TAP Deli Tua menggunakan aplikasi RapidMiner. Dapat dilihat dibawah ini sebagai berikut:

a) Membuka Aplikasi RapidMiner

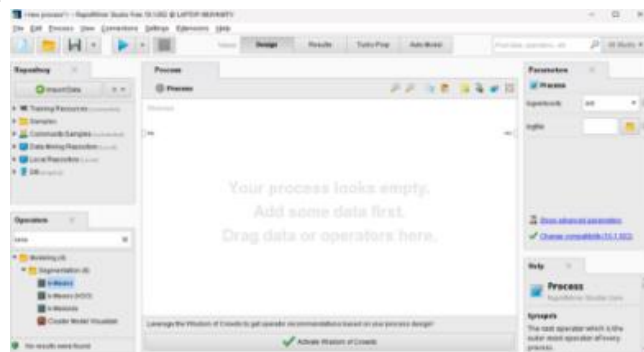
Pertama penulis membuka aplikasi RapidMiner untuk memasukan data yang telah dibuat dan memperoleh hasil setelah di proses.



Gambar 2. Dijelaskan Bahwa Terdapat Tampilan Depan RapidMiner

b) Tampilan Layar Untuk Proses Data

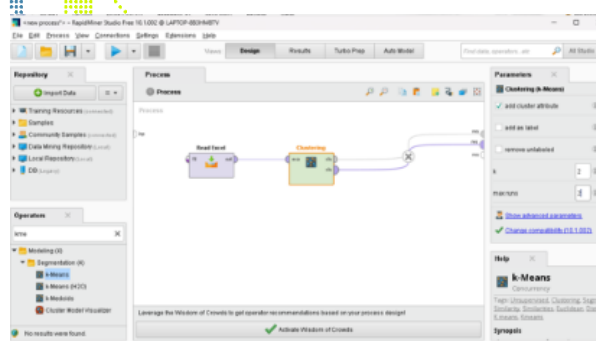
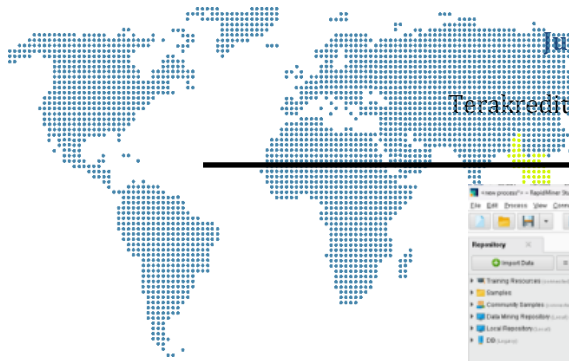
Tampilan layar ini digunakan saat penulis ingin melakukan hasil dari proses data tersebut.



Gambar 3. Dijelaskan Bahwa Terdapat Tampilan Untuk Proses Data

c) Proses Memperoleh Hasil Data

Utuk mengetahui hasil *cluster* data penjualan maka dapat dilakukan proses pada RapidMiner sebagai berikut:



Gambar 4 Dijelaskan Bahwa Terdapat Proses Memperoleh Hasil Data

d) Hasil Data Setelah Di Proses

Dari hasil data yang telah di uji menggunakan RapidMiner akan menghasilkan *luster 0* dan *7 cluster 1*.

Row No.	nama_produk	cluster_0	cluster_1	lat	long	age	sex
1	1. Sg30d	cluster_0	12	12	12	12	14
2	1. Sg30d	cluster_0	12	12	12	12	10
3	1. Sg30d	cluster_0	10	10	10	12	10
4	2. Sg30d	cluster_1	10	8	8	10	10
5	2. Sg30d	cluster_0	10	10	10	12	12
6	2. Sg30d	cluster_0	10	10	10	10	12
7	3. Sg30d	cluster_1	8	8	8	8	8
8	3. Sg30d	cluster_0	10	10	10	12	10
9	3. Sg30d	cluster_1	8	8	8	8	8
10	3. Sg30d	cluster_1	8	8	8	8	8
11	3. Sg30d	cluster_1	8	8	8	8	8
12	3. Sg30d	cluster_0	10	10	10	10	10
13	1. Sg30d	cluster_0	10	10	10	10	10

Gambar 5. Dijelaskan Bahwa Terdapat Hasil Data Penjualan Paket Internet

4. SIMPULAN

Berdasarkan dari penelitian yang telah dilakukan oleh penulis bahwa Proses yang dapat dilakukan agar tidak terjadinya kehabisan stok pada TAP Deli Tua ialah dengan cara pengelompokan terhadap data penjualan menjadi 2 kelompok yaitu produk laris dan produk non laris berdasarkan data penjualan yang dihasilkan oleh TAP Deli Tua. Produk laris yaitu suatu kelompok yang memberikan keuntungan bagi TAP Deli Tua, karena produk habis terjual. Maka dengan itu TAP Deli Tua harus lebih memperbanyak produk yang diinginkan oleh konsumen, dan juga harus mengurangi produk yang kurang diminati oleh konsumen. Karena semakin banyaknya produk yang dipasarkan TAP Deli Tua akan memberikan keuntungan yang lebih besar. Maka dengan itu sangat dibutuhkan suatu pengolahan dari hasil penjualan agar TAP Deli Tua mengetahui hasil produk yang habis terjual. Maka dalam melakukan pengelompokan data penjualan paket internet dibutuhkan suatu peran data mining dalam membantu penyelesaian pengelompokan. Data mining merupakan salah satu proses dalam menemukan informasi yang dihasilkan oleh data yang telah di olah menjadi data yang besar agar data tersebut dapat dipergunakan untuk membantu dalam pengambilan keputusan dimasa depan. Data mining juga salah satu kegiatan pengumpulan data yang dapat dilakukan dengan salah satu proses perhitungan atau aritmatika yang terdapat pada teknik dalam menemukan hasil pengelompokan data tersebut. Pada

data mining terdapat suatu teknik yang dapat membantu dalam pengelompokan paket internet yaitu teknik clustering. Didalam teknik clustering juga terdapat suatu algoritma yang memperoleh hasil dalam cluster yaitu algoritma K-Means. Algoritma K-Means salah satu algoritma yang dapat memperoleh kelompok yang berbeda dari data dalam sebuah dataset yang tidak memiliki sebuah label. Algoritma k-means juga dapat membagi suatu kelompok yang berbeda menjadi beberapa kelompok berdasarkan karakteristiknya yang dimana jika karakteristik memiliki kesamaan maka dapat dijadikan satu atau di persatukan. Sehingga hal tersebut dapat membantu TAP Deli Tua dalam pengelompokan data paket internet pada TAP Deli Tua. Maka dengan dilakukan penelitian ini dengan algoritma k-means dapat memperlihatkan bahwa data mining dapat memperoleh hasil suatu keputusan yang lebih tepat dalam pengelompokan data paket internet sehingga memperoleh hasil dari cluster dari produk laris dan non produk laris yang terdapat 15 data penjualan paket internet memperoleh 8 produk laris yang berada di cluster 0 dan 7 non-produk laris yang berada di cluster 1. Dan penelitian ini juga sudah di uji dengan menggunakan sebuah tool yaitu RapidMiner.

DAFTAR PUSTAKA

- [1] N. Afiasari, N. Suarna, and N. Rahaningsih, "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means E-commerce K-Means melakukan analisis penerapan Data Mining dalam mengelompokkan jumlah," vol. 9, pp. 100–110, 2023.
- [2] B. Anggraeni, Hapsari Dita, Saputra, Ragisl, Noranita, "Aplikasi Data Mining Analisis Data Transaksi Penjualan Obat Menggunakan Algoritma Apriori," *Min. Massive Datasets*, vol. 2, no. January 2013, pp. 5–20, 2005, [Online]. Available: https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part.
- [3] Z. Nabila, A. R. Isnain, P. Permata, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 2, pp. 100–108, 2021.
- [4] L. Purba, S. Saifullah, and R. Dewi, "Pengelompokan Kasus Penyakit Aids Berdasarkan Provinsi Dengan Data Mining K-Medoids Clustering," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, 2019.
- [5] A. A. Aldino and A. Rahma, "Comparative Analysis of Apriori Algorithm and Hash-Based Algorithm in Market Basket Analysis," vol. 8, no. 6, pp. 382–385, 2021, doi: 10.30865/jurikom.v8i6.3574.
- [6] F. Panjaitan, A. Surahman, and T. D. Rosmalasari, "Analisis Market Basket Dengan Algoritma Hash-Based Pada Transaksi Penjualan (Studi Kasus: TB. Menara)," *J. Teknol. dan Sist. Inf.*, vol. 1, no. 2, pp. 111–119, 2020, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>.
- [7] J. Ilmiah and W. Pendidikan, "Implementasi Algoritma K-Means Clustering Status Gizi Balita Nurul Rizki Octaviyani 1 , Rini Mayasari 2 , Susilawati 3 Universitas Singaperbangsa Karawang," vol. 8, no. 13, pp. 370–381, 2022.
- [8] I. W. Wijaya, K. Sandi, I. Bagus, and G. Dwidasmara, "Implementasi Algoritma K-Means Clustering dalam Penentuan Klasifikasi Tingkat Pembangunan Perekonomian di Provinsi Bali," vol. 1, no. 2, pp. 761–770, 2023.
- [9] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering



- dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, pp. 25–36, 2021.
- [10] U. D. Rahayu *et al.*, "Analisis Kasus Perceraian Pada Pengadilan Negeri Bekasi Menggunakan Algoritma K-Means Clustering," *J. Univ. Pendidik. Indones.*, vol. 6, no. 1, pp. 165–172, 2022.
- [11] M. Di, T. Erremy, and I. Nawangsih, "Analisa Penjualan Produk Kosmetik Dengan Metode Algoritma K-," vol. 4, no. 1, pp. 140–145, 2023.
- [12] D. Kurniadi, Y. H. Agustin, H. Ilham, N. Akbar, and I. Farida, "Penerapan Algoritma k-Means Clustering untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang," vol. 20, no. 1, pp. 64–77, 2023.
- [13] I. Of, D. Mining, U. The, and K. C. Event, "Penerapan Data Mining Dengan Algoritma K-Means Clustering Penyewaan Alat - Alat Event Pada Studi Kasus Cv . Dipo Rental Creativindo Implementation Of Data Mining Using The K-Means Clustering Event Rental," vol. 2, no. 1, pp. 8–15, 2023, doi: 10.14710/jtk.v2i1.37011.
- [14] R. Gustrianda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [15] M. Syahril, S. Kusnasari, A. Muhazir, and A. Syahputri, "Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Implementasi Data Mining Untuk Rekomendasi Jurusan Menggunakan Algoritma K-Means Clustering Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD," *Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 6, pp. 235–245, 2023.
- [16] A. Yani, Z. Azmi, and D. Suherdi, "Implementasi Data Mining Menganalisa Data Penjualan Menggunakan Algoritma K-Means Clustering," vol. 2, pp. 315–323, 2023.
- [17] E. Ramadanti and M. Muslih, "Penerapan Data Mining Algoritma K-Means Clustering Pada Populasi Ayam Petelur Di Indonesia," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 1, pp. 1–7, 2022, doi: 10.36341/rabit.v7i1.2155.
- [18] C. S. D. B. Sembiring, L. Hanum, and S. P. Tamba, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri)," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 80–85, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2393.
- [19] R. Yanto and R. Khoiriah, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," *Creat. Inf. Technol. J.*, vol. 2, no. 2, p. 102, 2015, doi: 10.24076/citec.2015v2i2.41.
- [20] Q. I. Mawarni and E. S. Budi, "Implementasi Algoritma K-Means Clustering Dalam Penilaian Kedisiplinan Siswa," *J. Sist. Komput. dan Inform.*, vol. 3, no. 4, p. 522, 2022, doi: 10.30865/json.v3i4.4242.
- [21] S. Khairunnisa and M. I. Jambak, "Pengelompokan Cuaca Kota Palembang Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Karakteristik Cuaca," vol. 6, pp. 2352–2360, 2022, doi: 10.30865/mib.v6i4.4810.
- [22] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," vol. 1, no. 1, pp. 20–28, 2023.
- [23] J. Hutagalung, "Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 606–620, 2022, doi: 10.35957/jatisi.v9i1.1516.