Jurnal Ilmu Komputer dan Informatika (JIKI)
Vol. 1, No. 2, Desember 2021, Hal. 141-156
https://jiki.jurnal-id.com

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Pengelompokan Data Penjualan Mie Berdasarkan Bulan Dengan Menggunakan Algoritma *K-Medoids*

Riahta Ulina Br. Barus[*1], Indra Gunawan[2], Bahrudi Effendi Damanik[3], Iin Parlina[4], Widodo Saputra[5]

[1,2,3,4,5]Teknik Informatika, STIKOM Tunas Bangsa Pematangsiantar, Indonesia
Email: [1]riahtabarus@gmail.com, [2]indra@amiktunasbangsa.ac.id, [3]bahrudiefendi@gmail.com, [4]iin@amiktunasbangsa.ac.id, [5]widodo@amiktunasbangsa.ac.id

Abstrak

Mie merupakan salah satu keanekragaman kuliner yang memiliki potensi sangat besar untuk dikembangkan sebagai peningkatan ekonomi di pematangsiantar. Penjualana Mie di Kota pematangsiantar juga sangat pesat, namun para pedagang masih mneggunakan cara manual dalam pengelompokan data hasil penjualan dari berbagai macam jjenis mie yang di jual sehingga tingkat akurasi keuntungannya menjadi tidak Nampak dengan jelas. Berdasarkan itu makan penulis membuat penelitian ini dengan menggunakan Algoritma *K-Medoids* untuk mengcluster data penjualan mie dari tahun 2018-2019 berdasarkan tahun dan bulan. Hasil yang diperoleh dari produksi mie di tahun 2018 terbanyak adalah mie pangsit dengan *centroid* 304670 pada bulan Januari dan mie pangsit dengan *centroid* 290476 pada bulan Desember, sedangkan di tahun 2019 adalah mie pangsit *centroid* 311532 bulan Januari dan 311532 pada bulan Desember. Tingkat akurasi dari pengolahan *RapidMiner* dengan pengolahan yang dilakukan penulis menunjukkan 12 data yang digunakan. Untuk *Cluster* C1 atau kelompok Tinggi diperoleh akurasi sebesar 17% dengan jumlah data yang sama 2 data, *Cluster* C2 atau kelompok Sedang diperoleh akurasi sebesar 8% dengan jumlah data yang sama 1 data dan *Cluster* C3 atau kelompok Rendah diperoleh akurasi sebesar 75% dengan jumlah data yang sama 9 data.

Kata kunci: data mining, K-Medoids, kuliner, mie, pengelompokan, Rapidminer

Abstract

Noodles are one of the culinary diversity that has enormous potential to be developed as an economic improvement in Pemangsiantar. Sales of noodles in Pematangsiantar City are also very fast, but traders still use manual methods in grouping sales data from various types of noodles that are sold so that the accuracy of profits is not clearly visible. Based on that, the author made this research using the K-Medoids Algorithm to cluster noodle sales data from 2018-2019 based on year and month. The results obtained from the most noodle production in 2018 were wonton noodles with a centroid of 304670 in January and wonton noodles with a centroid of 290476 in December, while in 2019 were wonton noodles centroid 311532 in January and 311532 in December. The accuracy level of RapidMiner processing with the processing carried out by the author shows 12 data used. For Cluster C1 or the High group, an accuracy of 17% was obtained with the same amount of data 2 data, Cluster C2 or the Medium group obtained an accuracy of 8% with the same amount of data 1 data and Cluster C3 or the Low group obtained an accuracy of 75% with the amount of data the same 9 data.

Keywords: *culinary, data mining, grouping, K-Medoids, noodles, Rapidminer.*

1. PENDAHULUAN

Data Mining merupakan proses yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan. *Data Mining* memilki 5 fase, yaitu Estimasi, Klustering, Prediksi, Klasifikasi dan Asosiasi. *Data Mining* Klustering memiliki salah satu metode yaitu *K-Medoids*. *Data Mining* Algoritma *K-Medois* merupakan salah satu metode yang banyak digunakan peneliti dengan membagi rangkaian data-data menjadi beberapa kelompok yang menggunakan nilai tengah yang disebut *medoid* dan perhitungan jarak dihitung dari jarak antar masing-masing data. *K-Medoids* dapat digunakan untuk mengelompokkan beberapa data, salah satunya mengelompokkan data penjualan mie di pematangsiantar.

Mie merupakan salah satu keanekaragaman kuliner di Pematangsiantar membuat kuliner mi di Pematangsiantar menyimpan potensi yang sangat besar untuk dikembangkan sebagai peningkatan ekonomi di Pematangsiantar. Dalam era otonomi daerah, sektor wisata kuliner memegang peranan penting dalam meningkatkan perekonomian suatu daerah karena memiliki keterkaitan sumber pertumbuhan ekonomi daerah. Pengembangan wisata kuliner memberikan efek ganda terhadap sektor ekonomi lainnya melalui peningkatan nilai tambah dan kenaikan pendapatan masyarakat. Peningkatan intensitas pemakaian tenaga kerja dalam pengembangan kuliner tidak hanya diharapkan dapat meningkatkanpendapatan masyarakat, tetapi juga mampu menciptakan kesempatan kerja dan mengurangi tingkat kemiskinan. Pematangsiantar merupakan salah satu wilayah daerah yang memiliki banyak ciri khas wisata kuliner berbagai jenis mie, yaitu mie kuning, mie hun, mie lidi, mie ktiaw, ifu mie dan mie pangsit.. Keberagaman jenis mie yang dijual merupakan kuliner mie menjadi dominan masyarakat di Kota Pematangsiantar. Banyak berbagai macam mie yang dijual di wilayah Pematangsiantar dengan berbagai variasi, harga dan ukuran.

Dari penjelasan diatas banyak cabang kecerdasan buatan dalam ilmu komputer yang dapat menyelesaikan permasalahan tersebut secara kompleks diantaranya sistem pendukung keputusan, sistem pakar, *data mining* dan lain sebagainya. Beberapa penelitian tentang *data mining*, salah satunya yang dilakukan oleh Pramesti, dkk [1] dalam penelitian ini metode yang digunakan dalam penelitian ini adalah *K-Medoids Clustering* yang dapat diimplementasikan pada pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas *(hotspot)* menghasilkan jumlah *cluster* dan jumlah data mempengaruhi terhadap hasil kualitas dari cluster berdasarkan proses pengujian yang dengan jumlah 2 cluster dan jumlah data 7352. Penelitian selanjutnya dilakukan oleh Mustafa [2] Dalam penelitian ini dilakukan pengujian model dengan menggunakan *K-Medoids Chebyshev* dalam pengelompokan data EDGI *E-government Surve* 2014 kedalam 4 status EDGI yeng menghasilkan model mendapatkan nilai *Bouldin Index* dari setiap algoritma yang lebih optomal dalam penentuan pengelompokan EDGI *E-govermnet Surve* 2014 ke dalam 4 status EDGI di bandingkan menggunakan *mahattan* atau *ecludian.*

Penelitian ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkarya teori yang digunakan dalam mengkaji penelitian mengelompokkan data penjualan mie di kota Pematangsiantar. Tingkat penjualan antar perusahaan pembuat kuliner mie cenderung berubah karena banyaknya usaha kuliner mie. Seiring berubahnya tingkat penjualan berdampak pada usaha kuliner mie sehingga membuat pemborosan modal yang tidak terpakai dalam penjualan mie. Pemborosan modal tersebut karena banyaknya pengusaha kuliner mie di Kota Pematangsiantar, sehingga mengakibatkan dampak ketidaksesuaian kebutuhan konsumen dengan tingkat produksi pembuatan mie.

Berdasarkan beberapa uraian tersebut menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkarya teori yang digunakan dalam mengkaji penelitian mengelompokkan data penjualan mie di kota Pematangsiantar. Untuk mengetahui pengelompokkan tingkat penjualan mie di Kota Pematangsiantar pada berdasarkan bulan, penulis menggunakan teknik Data Mining dalam proses pengolahan data dengan Algoritma *K-Medoids*.

Metode *K-Medoids clustering* sangat cocok dalam mengelompokan data- data hasil penjualan mie dimana *K-Medoids* ini menggunakan kelompok metode *partitional Clustering* yang dapat meminimalisir jarak antar titik berlabel yang ada dalam *cluster* yang juga menggunakan objek sebagai perwakilan tidak dengan angka rata-rata sehingga algoritma ini dapat diterapkan pada data penjualan mie di Kota Pematangsiantar, dengan adanya pengelompokan ini dapat mempermudah penjual mie dalam mengelompokan data-data hasil penjualan masing-masing mienya yang berbeda dengan mudah dan akurat. Diharapkan penelitian ini dapat memberikan informasi kepada pengusaha Mie serta pemerintah untuk mengatasi ketidaksesuaian kebutuhan konsumen dan tingkat produksi mie di kota Pematangsiantar.

2. TINJAUAN PUSTAKA

2.1. Data Mining

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Data mining adalah proses analitik yang dirancang untuk memeriksa sejumlah data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga dan konsisten.[3] Istilah lain data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan secara otomatis. *Data mining* digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases* (KDD) yaitu tahapan yang dilakukan dalam menggali pengetahuan dari sekumpulan data. Proses penggalian informasi tersembunyi dalam suatu basis data yang besar sering menggunakan istilah *data mining* dan *knowledge discovery in databases* (KDD). [4] Merupakan tahapan untuk menentukan pola atau informasi delam sekumpulan data dengan menggunakan teknik dan algoritma tertentu. [5]

2.3. Clustering

Clustering adalah salah satu teknik *data mining* yang bertujuan untuk mengidentifikasi sekelompok objek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok objek lainnya, sehingga objek yang berada dalam kelompok yang sama relatif lebih homogen dari pada objek yang berada pada kelompok yang berbeda.[6]

Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma yaitu dengan Euclidean Distance. Analisis kluster ialah metode yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya.[7]

Tujuan dari pengelompokkan sekumpulan data objek kedalam beberapa kelompok yang mempunyai *karakteristik* tertentu dan dapat dibedakan satu sama lainnya adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan. Model yang diambil diasumsikan bahwa data yang dapat digunakan adalah data yang berupa data interval, frekuensi dan biner.

2.4. Metode *K-Medoids Clustering*

Metode *k-medoids* adalah metode pengelompokkan yang berkaitan dengan metode *k-means* dan metode *medoidshift.*[8] Metode *k-medoids* di usulkan pada tahun 1987. Metode *k-medoids* dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw. Metode *k-means* dan *k-medoids* yang partitional (melanggar dataset ke dalam kelompok) dan kedua upaya untuk meminimalkan jarak antara titik berlabel berada dalam *cluster* dan titik yang ditunjuk sebagai pusat *cluster*. Berbeda dengan *k-means*, metode *k-medoids* memilih datapoint sebagai pusat (*medoids* atau *eksemplar*) dan bekerja dengan matriks sewenang-wenang dari jarak antara datapoints.

Penggunaan Algoritma K-Medoids bekerja dengan baik karena setiap objek pada setiap *cluster* memiliki mutu yang baik, dimana setiap objek telah dikelompokan sesuai dengan tingkat kemiripan yang tinggi serta K-Medoids lebih baik dalam melakukan pengelompokan data dibandingkan dengan algoritma K-Means berdasarkan nilai validitasnya.[9]

Algoritma *K-Medoids* merupakan algoritma yang mirip dengan *K-Means* karena kedua algoritma partitional yang memecah dataset menjadi kelompok – kelompok. Perbedaannya terletak pada penentuan pusat cluster, di mana algoritma K-*Means* menggunakan nilai rata –rata *(means)* dari setiap *cluster* sebagai pusat *cluster* dan algoritma *K-Medoids* menggunakan objek data sebagai perwakilan *(medoid)* sebagai pusat *cluster.*[10]

Perbedaan dari kedua metode ini yaitu metode *k-medoids* menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap cluster, sedangkan *k-means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster.* Metode *k-medoids* memiliki kelebihan untuk mengatasi kelemahan pada pada metode *k-means* yang *sensitive* terhadap *noise* dan *outlier*, dimana objek dengan nilai yang besar yang *memungkinkan* menyimpang pada dari distribusi data. Kelebihan lainnya yaitu hasil proses *clustering* tidak bergantung pada urutan masuk dataset.

Langkah-langkah metode *K-Medoids* adalah sebagai berikut :

a. Inisialisasi pusat *cluster* sebanyak *k* (jumlah *cluster*)

Alokasikan setiap data (objek) ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan:

$$d_{ij} = \sqrt{\sum_{a=1}^{p}(x_{ia} - x_{ja})^2} = \sqrt{(\mathrm{x}_i - \mathrm{x}_j)'(\mathrm{x}_i - \mathrm{x}_j)} \quad (1)$$

dimana i=1,…..,n; j=1,…..,n dan p adalah banyak variable, serta V adalah matrik varian kovarian.

b. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru.
c. Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat *medoid* baru.
d. Hitung total simpangan (*S*) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan *k* objek baru sebagai *medoid*.
e. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoid*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

2.5. Rapid Miner

RapidMiner adalah salah satu alat penambangan data yang digunakan untuk menganalisis informasi yang diakses web. Ini digunakan untuk penelitian, pendidikan, pembuatan prototipe cepat, aplikasi pengembangan, dan aplikasi industri.[11]

Dengan menggunakan Rapid Miner, tidak dibutuhkan kemampuan koding khusus, karena semua fasilits sudah disediakan. Rapid Miner dikhususkan untuk penggunaan data mining. Model yang disediakan juga cukup banyak dan lengkap, seperti *Model Bayesian, Modelling, Tree Induction, Neural Network* dan lain-lain[12]

Pengaruh dalam penilaian akurasi ini dinilai baik dengan data dan juga dengan algoritma yang dipakai. Akurasi dari sebuah algoritma tentunya dapat dihitung baik menggunakan manual seperti *Confussion Matrix* atau dapat dilihat dengan bantuan *software* seperti *Rapidminer*. Terdapat beberapa jenis aplikasi atau *Software* data mining terutama pada klasifikasi seperti *Rapidminer*, Weka*, Orange*, KNIME, SPSS *Climatte* dan sebagainya. Aplikasi yang dianggap mirip, mudah digunakan dan tidak perlu bingung dengan penggunaan bahasa pemogramanyang rumit adalah *Rapidminer* dan *Weka*. Langkah-langkah keduanya hampir sama yaitu hanya memasukkan data *(import data)* ke dalam aplikasi kemudian memilih algoritma lalu mendapatkan hasilnya, kedua aplikasi bersifat terbuka sehingga dapat digunakan tanpa pembayaran[13]

2.6. *Unifield Modelling Language (UML)*

UML (*Unified Modelling Language*) adalah bahasa spesifikasi standar untuk mendokumentasi, merancang dan membuat software orientasi objek. UML ini merupakan bahasa visual untuk pemodelan *bahasa* berorientasi objek, sehingga semua elemen dan diagram berbasiskan pada paradigma *object oriented*. UML menawarkan sebuah standard untuk merancang model sebuah sistem. UML mendefenisikan notasi dan *syntax*/semantik.[14] Pemilihan UML karena UML merupakan pemrograman berorientasi objek yang memiliki kemampuan dalam menganalisa dan menjabarkan sistem secara rinci. [15]

3. METODE PENELITIAN

Pada metode penelitian ini penulis menjelaskan bagaimana menguraikan cara ilmiah untuk menyelesaikan masalah-masalah penelitian.

3.1. Rancangan Penelitian

Flowchart penelitian pada skripsi ini dapat dideskripsikan sesuai gambar 1.

Pada gambar 1 diatas dijelaskan bahwasanya rancangan penelitian yang dilakukan untuk menentukan pengelompokkan data penjualan mie berdasarkan bulan di Kota Pematangsiantar terdiri dari :

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Gambar 1. Rancangan Penelitian

a. Identifikasi Masalah
   Mengidentifikasi masalah yang terkait dalam pengelompokkan data penjualan mie berdasarkan bulan di Kota Pematangsiantar dari tahun 2018-2019.
b. Mempelajari Literatur
   Penelitian ini dilandasi rujukan atau referensi yang terkait guna mendapatkan informasi pendukung dalam penelitian.Pengumpulan Data
c. Analisa Hasil
   Proses yang dilakukan untuk mengelompokkan data penjualan mie berdasarkan bulan di Kota Pematangsiantar dengan menggunakan metode *K-Medoids*.
d. Pengujian
   Penulis melakukan pengujian dengan menggunakan aplikasi *RapidMiner* untuk mengelompokkan data penjualan mie.
e. Keputusan
   Hasil yang diberikan oleh sistem dan analisa penulis dapat memberikan masukan untuk pihak pengusaha mie untuk melihat data kebutuhan masyarakat konsumsi mie di Kota Pematangsiantar supaya tidak terjadi pemborosan modal.

3.2. Prosedur pengumpulan data

Penulis mengumpukan data dengan menggunakan metode Studi Kepustakaan (*Libarry Research*) yaitu memanfaatkan perpustakaan, buku, prosiding atau jurnal sebagai referensi dalam menentukan

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

parameter yang digunakan Penelitian Lapangan (*Field Work Research*) Yaitu penelitian yang dilakukan secara langsung dengan menggunakan beberapa teknik wawancara dengan penjual mie yang ada di Kota Pematangsiantar

3.3. Analisis Data

Data yang yang di dapatkan dikumpulkan dari melakukan isian kuesioner terhadap penjual mie yang ada di Kota Pematangsiantar yang menggunakan beberapa parameter, yang terdiri dari bulan dari tahun 2018 dan 2019.

Berikut data yang dikumpulkan sebanyak 12 data dengan data perbulan dari ± 80 usaha mie di Kota Pematangsiantar.

Tabel 1.Data Penelitian penjualan Mie Tahun 2018

BULAN	MIE KUNING	MIE HUN	MIE LIDI	MIE KTIAW	IFU MIE	MIE PANGSIT
Januari	204750	184750	134750	198750	201608	304670
Februari	176653	176653	176653	176653	176653	176653
Maret	182082	182082	150085	182082	151698	182082
April	150085	187511	136716	150085	126743	150085
Mei	136716	192940	170741	136716	101788	136716
Juni	170741	198369	344265	170741	170741	100463
Juli	168879	198332	168879	168879	168879	164579
Agustus	167017	172652	153482	193635	153432	153983
September	165155	165155	165155	165155	165155	165155
Oktober	173293	173293	176828	173293	168879	173293
November	181431	165155	188501	136716	167017	150085
Desember	240088	240088	240088	240088	240088	290476
Jumlah	2116890	2236980	2206143	2092793	1992681	2148240

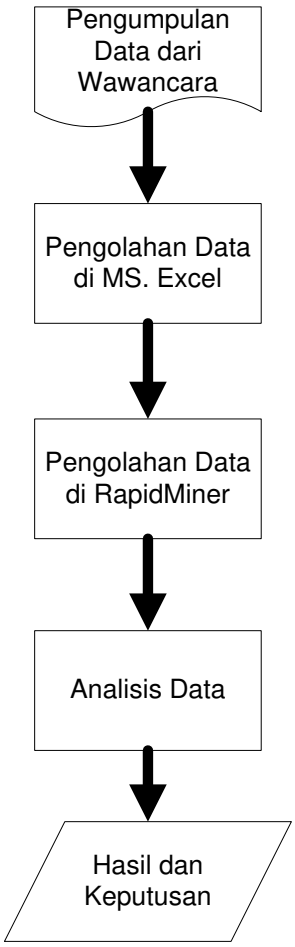
Tabel 2.Data Penelitian penjualan Mie Tahun 2019

BULAN	MIE KUNING	MIE HUN	MIE LIDI	MIE KTIAW	IFU MIE	MIE PANGSIT
Januari	217456	165155	134750	240088	240641	311532
Februari	176653	173293	176653	181431	176653	176653
Maret	182082	181431	150085	172353	151698	182082
April	184750	172652	136716	150085	126743	150085
Mei	176653	165155	170741	136716	101788	136716
Juni	182082	173293	344043	170741	170741	193643
Juli	187511	198332	168879	168879	243253	153532
Agustus	167017	172652	153482	193635	153432	153983
September	165155	165155	165155	165155	165155	165554
Oktober	173293	173293	136716	173293	168879	173293
November	181431	154383	170741	136716	167017	150085
Desember	240088	193464	201036	233872	269485	306953
Jumlah	2234171	2088258	2108997	2122964	2135485	2254111

3.4. Instrumen Penelitian

Penelitian ini menggunakan penelitian dengan metode kuantitatif. Metode kuantitatif merupakan penelitian yang menggunakan data berupa angka sebagai alat menganalisis keterangan dengan menggunakan teknik wawancara. Alur instrumen penelitian dapat dilihat pada gambar 3 sebagai berikut :

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591



Gambar 2. Alur instrument penelitian

Pada gambar 2 diatas menunjukkan alur penelitian yang terdiri dari pengguna dan sistem. menjelaskan penulis sebagai pengguna melakukan analisa masalah dan tujuan masalah, kemudian mengumpulkan data yang diperoleh dari hasil melakukan wawancara, data yang telah dikumpulkan kemudian diolah menggunakan sistem yaitu *software RapidMiner*, hasil yang telah diperoleh dari sistem *software RapidMiner* menghasilkan pengelompokkan data penjualan mie di Kota Pematangsiantar. Penulis sebagai pengguna membua keputusan dari hasil penelitian pengelompokkan data penjualan mie dan membuat keputusan terhadap penelitian.

3.5. Pemodelan Algoritma

Permodelan yang digunakan menggunakan *RapidMiner*. Pada lembar *main process* operator *K-Medoids* menggunakan parameter yang digunakan dimana *k* merupakan jumlah *cluster* yang digunakan, max *runs* merupakan penentukan jumlah maksimum berjalan *K-Medoids* dengan inisialisasi acak yang dilakukan, *max optimization steps* merupakan jumlah iterasi maksimum yang dilakukan untuk satu kali *K-Medoids*. *Measure types* merupakan jenis pengukuran yang akan digunakan untuk mengukur jarak antara titik dan *mixed measure* merupakan parameter yang tersedia ketika parameter tipe pengukuran diatur ke "*mixed measure*"

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian sesuai data yang digunakan pada saat pengumpulan data. Data yang telah dikumpulkan untuk diolah atau ditransformasikan ke format data *Ms.Excel 2010*. Data yang ditransformasikan ersebut digunakan sebagai syarat dalam pengolahan Algoritma *K-Medoids*.

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

4.1. Algoritma *K-Medoids*

Penulis menentukan analisa pengelompokkan tingkat penjualan mie di Kota Pematangsiantar menggunakan Algoritma *K-Medoids*. Berikut langkah-langkah dalam pengolahan data menggunakan Algoritma *K-Medoids* menggunakan data tahun 2018 dan tahun 2019.

a. Menentukan Data yang diolah
   Data yang digunakan adalah data pembuatan mie tahun 2018 dan tahun 2019 di Kota Pematangsiantar.data dapat dilihat pada tabel 1 dan tabel 2.
b. Menentukan Jumlah Cluster, atau k yang digunakan sebanyak 3 yaitu Cluster Tinggi, Cluster Sedang dan Cluste Rendah
c. Menentukan Pusat Medoid
   Menentukan Pusat Medoid dilakukan secara acak pada masing-masing Cluster. Pusat Medoid yang digunakan adalah :
   Pusat Medoid awal Rendah tahun 2018 = (136716, 192940, 170741, 136716, 101788, 136716)
   Pusat Medoid awal Sedang tahun 2018 = (167017, 172652, 153482, 193635, 153432, 153983)
   Pusat Medoid awal Tinggi tahun 2018 = (204750, 184750, 134750, 198750, 201608, 304670)
   Pusat Medoid awal Rendah tahun 2019 = (150335, 187761, 136966, 150335, 126993, 150335)
   Pusat Medoid awal Sedang tahun 2019 = (176903, 169346, 172366, 170001, 180362, 173256)
   Pusat Medoid awal Tinggi tahun 2019 = (205000, 185000, 135000, 199000, 201858, 304920)
d. Menghitung Nilai *Euclidian*
   Untuk menghitung jarak antara titik Medoid dengan titik tiap objek menggunakan *Euclidian Distance*. Rumus untuk menghitung Nilai *Euclidian* adalah :

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

   Sehingga didapat tabel jarak dari *Medoid* dan mencari nilai minimal dari ketiga *Medoid*. Tabel Jarak dari *Medoid* adalah sebagai berikut :

Tabel 3. Hasil Jarak *Medoid* Iterasi 1 tahun 2018

BULAN	C1	C2	C3	Jarak Dekat
Januari	219116	164240	0	0
Februari	103392	44579	141811	44579
Maret	95968	35404	136209	35404
April	48437	58418	186733	48437
Mei	0	88502	219116	0
Juni	196276	201895	297786	196276
Juli	85901	43171	155644	43171
Agustus	88502	0	164240	0
September	85117	35627	157474	35627
Oktober	94540	40126	147887	40126
November	86720	70213	180992	70213
Desember	232468	171296	48314	48314

   Dari tabel 3 dapat dilihat hasil jarak *Medoid* iterasi 1 untuk tahun 2018 yang akan digunakan untuk menentukan nilai *Cost*. Nilai Jarak Dekat ini yang akan digunakan sebagai letak *cluster* atau kelompok berdasarkan nilai terkecil perbaris.
e. Cara Menghitung Nilai *Cost*
   Nilai *Cost* diperoleh dari total penjumlahan nilai jarak dekat *Medoid* yang diperoleh yaitu :

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Cost = (0 + 4459 + 35404 + 48437 + 0 + 196276 + 43171 + 0 + 35627 + 40126 + 70213 + 48314) = 562147

f. Lakukan Ulang Langkah d dan e dengan nilai Pusat *Medoid* baru
Mengulangi proses meghitungan nilai jarak dekat *Medoid* dengan nilai pusat *Medoid* baru secara acak. Untuk data tahun 2018 yaitu :
Pusat *Medoid* baru Rendah tahun 2018 = (136716, 192940, 170741, 136716, 101788, 136716)
Pusat *Medoid* baru Sedang tahun 2018 = (170741, 198369, 344265, 170741, 170741, 100463)
Pusat *Medoid* baru Tinggi tahun 2018 = (2040088, 178573, 149846, 196338, 226349, 290476).
Dari pengolahan menggunakan pusat *Medoid* baru diperoleh hasil sebagai Tabel berikut :

Tabel 4. Hasil Jarak Medoid Iterasi 2 tahun 2018

BULAN	C1	C2	C3	Jarak Dekat
Januari	219116	297786	48314	48314
Februari	103392	185675	143385	103392
Maret	95968	212727	144576	95968
April	48437	220105	200308	48437
Mei	0	196276	232468	0
Juni	196276	0	287840	0
Juli	85901	186766	160407	85901
Agustus	88502	201895	171296	88502
September	85117	193552	162640	85117
Oktober	94540	184349	150941	94540
November	86720	170628	178605	86720
Desember	232468	287840	0	0

Dari table 4 dapat dilihat hasil jarak *Medoid* iterasi 2 untuk tahun 2018 yang akan digunakan untuk menentukan nilai *Cost*. Nilai Jarak Dekat ini yang akan digunakan sebagai letak *cluster* atau kelompok berdasarkan nilai terkecil perbaris.
Dari tabel 4 dapat diperoleh nilai *Cost* yaitu :
Cost = (48314 + 103392 + 95968 + 48437 + 0 + 0 + 85901 + 88502 + 85117 + 94540 + 86720) = 73689

g. Mencari Nilai *S*
Nilai *S* diperoleh dengan cara mengurangi nilai *Cost* pada iterasi yang baru kepada iterasi awal. Jika nilai S < 0 maka pengolahan diteruskan dengan menggunakan nilai pusat medoid baru. Jika Nilai S > 0 atau nilai *Cost* iterasi baru lebih besar daripada nilai *Cost* iterasi lama maka proses dihentikan. Sehingga nilai S diperoleh :
S = Nilai Cost Baru – Nilai Cost Lama = 736891 – 562147 = 174745
Karena Nilai Cost Baru > Nilai Cost Lama, maka iterasi dihentikan.

h. Mencari *Cluster* atau Pengelompokan
Untuk menentukan *Cluster* diperoleh dengan menentukan nilai Cluster 1, 2 dan 3 terdekat dengan Nilai Jarak *Distance* sehingga diperoleh Cluster untuk tahun 2018 seperti pada tabel 5.
Dari tabel 5 dapat dilihat hasil *cluster* yang diperoleh dari iterasi terakhir untuk tahun 2018. Diperoleh Cluster Rendah pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November. *Cluster* Sedang pada bulan Juni, dan *Cluster* Tinggi pada bulan Januari dan Desember.
Untuk data tahun 2019 diperoleh hasil yang sama seperti data tahun 2018 dengan nilai S = 54677 iterasi terakhir pada tabel 6.

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Tabel 5. Cluster Tahun 2018

BULAN	C1uster
Januari	3
Februari	1
Maret	1
April	1
Mei	1
Juni	2
Juli	1
Agustus	1
September	1
Oktober	1
November	1
Desember	3

Tabel 6. Hasil Jarak Medoid Iterasi 2 tahun 2019

BULAN	C1	C2	C3	Jarak Dekat
Januari	180992	297406	48314	48314
Februari	45771	189547	145786	45771
Maret	71269	212458	144576	71269
April	77232	220532	200308	77232
Mei	86720	197030	232468	86720
Juni	171731	0	287545	0
Juli	54526	186842	160121	54526
Agustus	70213	201317	171296	70213
September	39999	186778	164239	39999
Oktober	48521	185755	147989	48521
November	0	171731	178605	0
Desember	178605	287545	0	0

Dari tabel 6 diperoleh pengelompokkan atau cluster seperti pada tabel 7:

Tabel 7. Cluster Tahun 2019

BULAN	C1uster
Januari	3
Februari	1
Maret	1
April	1
Mei	1
Juni	2
Juli	1
Agustus	1
September	1
Oktober	1
November	1
Desember	3

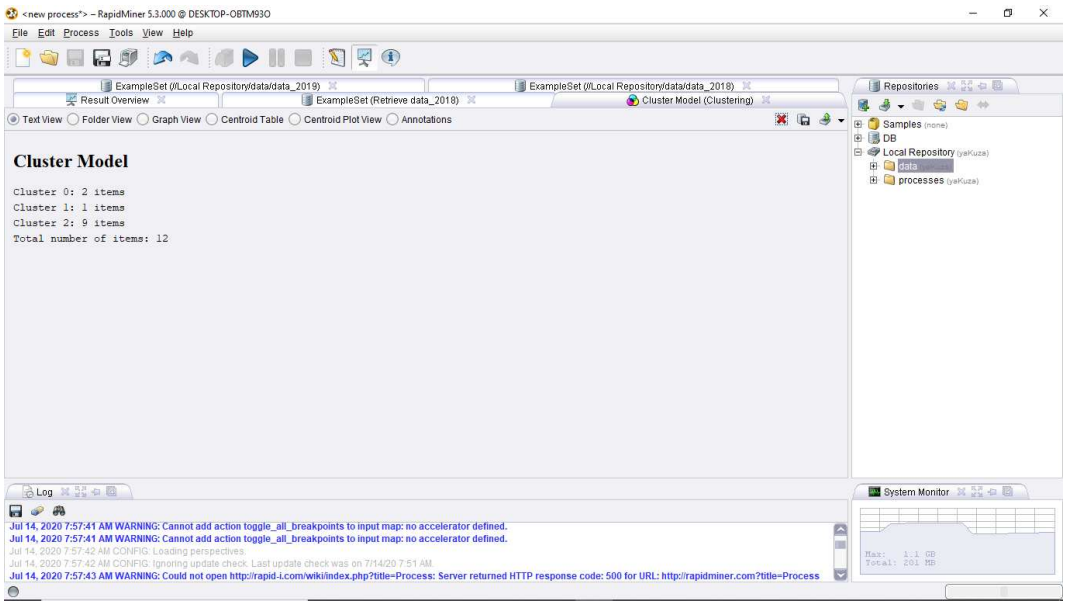
DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Dari tabel 7 dapat dilihat hasil *cluster* yang diperoleh dari iterasi terakhir untuk tahun 2019. Diperoleh *Cluster* Rendah pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November. *Cluster* Sedang pada bulan Juni, dan *Cluster* Tinggi pada bulan Januari dan Desember.

4.2. Hasil Percobaan

a. Implementasi Algoritma *K-Medoids* dengan *RapidMiner*
Pada bagian ini berisikan tampilan sistem dengan menggunakan Algoritma K-Medoids menggunakan RapidMiner. Berikut tampilan proses dari implementasi Algoritma K-Medoids :

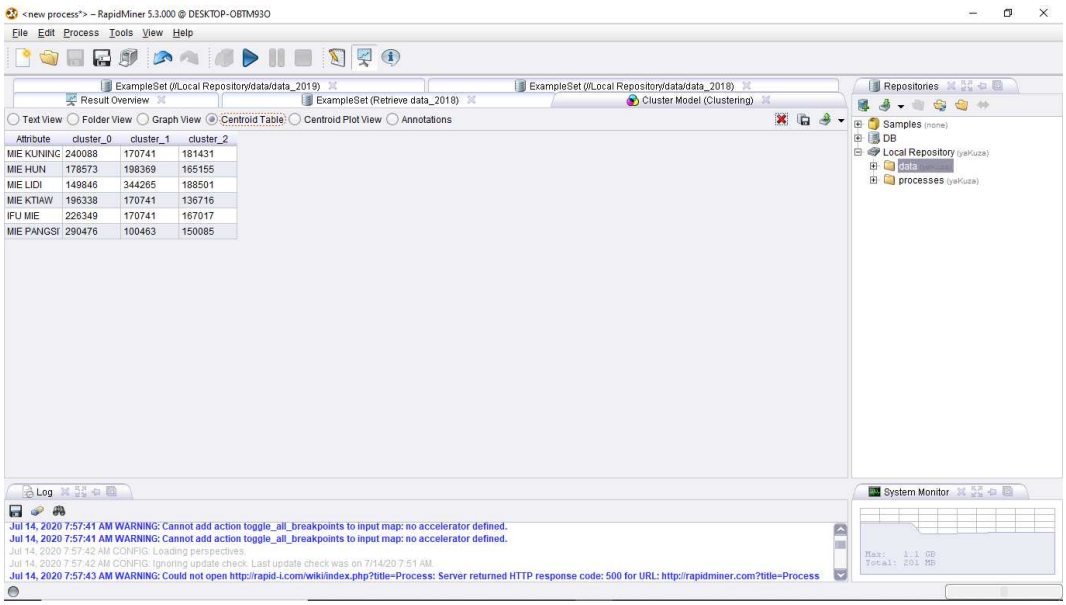
1. Tahapan *Importing Data*
   Tahapan awal yaitu pemasukkan data yang akan digunakan dalam penelitian. Penelitian ini menggunakan dua data yaitu data penjualan mie tahun 2018 dan data penjualan mie tahun 2019.
2. Tahapan Pengolahan Operator *K-Medoids*
   Tahapan berikutnya memasukkan operator yang berkaitan dengan Algoritma *K-Medoids* kedalam *Main Process* pada *RapidMiner* untuk mencari hasil yang diperoleh. Proses yang dilakukan sebanyak dua kali yaitu data tahun 2018 dan tahun 2019
3. Hasil Pengolahan *RapidMiner* data tahun 2018
   Hasil yang diperoleh dari pengolahan Algoritma *K-Medoids* pada *RapidMiner* untuk data tahun 2018 adalah sebagai berikut :



Gambar 3. Hasil Pengolahan *RapidMiner Cluster* Model

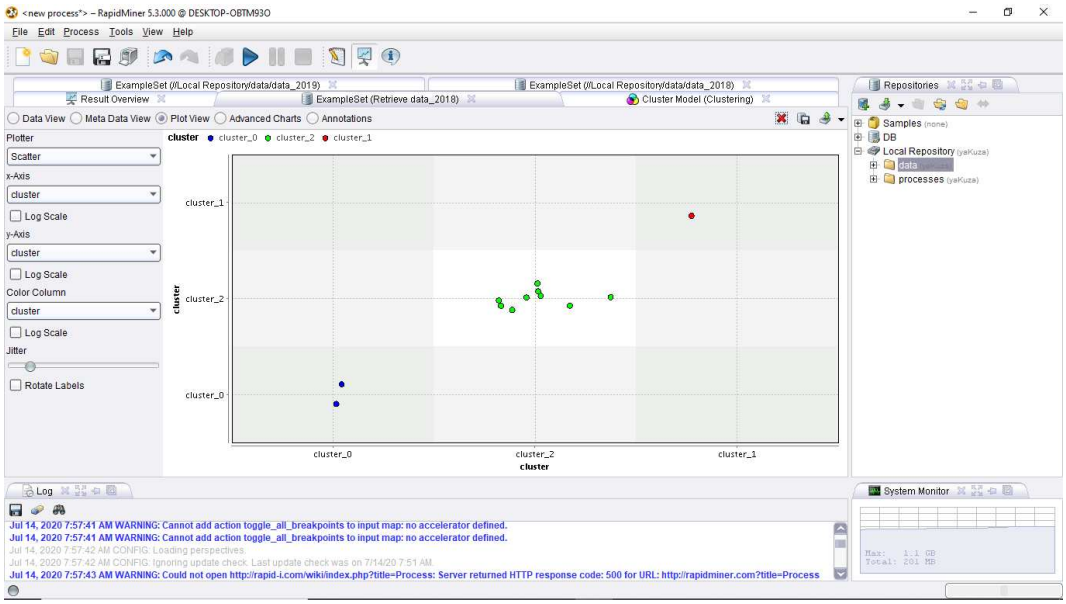
Gambar 3 menjelaskan hasil pengolahan *RapidMiner* dengan menampilkan *Cluster* Model dengan *Cluster* 0 sebagai *Cluster* Tinggi dengan jumlah 2 *items*, *Cluster* 1 sebagai *Cluster* Sedang dengan jumlah 1 *items*, dan *Cluster* 2 sebagai *Cluster* Rendah dengan jumlah 9 *items*.

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591



Gambar 4 Hasil Pengolahan *RapidMiner Centroid Table*

Gambar 4 menjelaskan hasil pengolahan *RapidMiner* dengan menampilkan hasil *Centroid Table.*

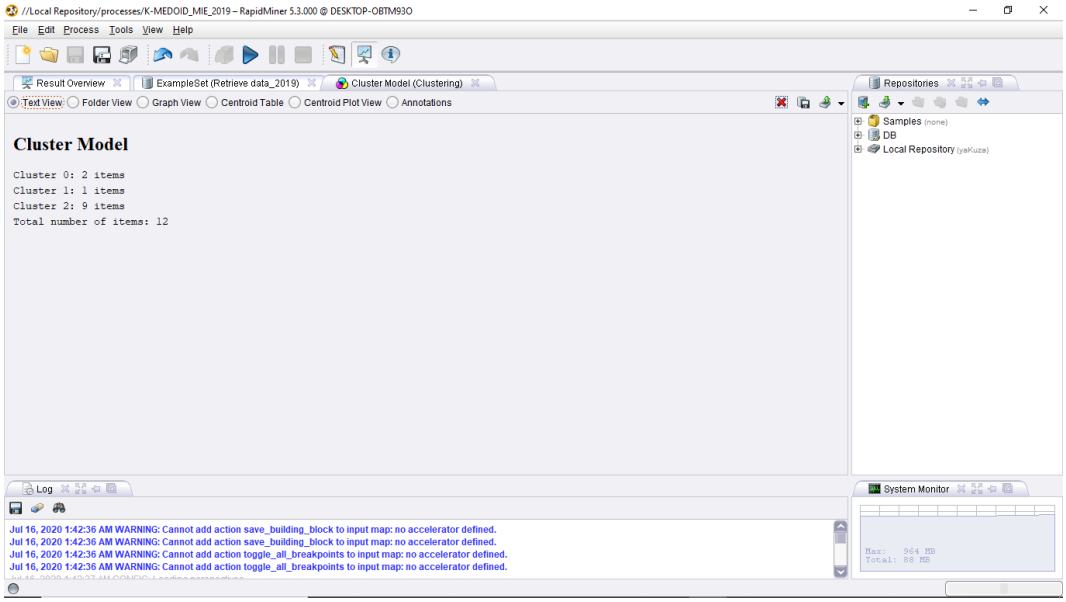


Gambar 5. Hasil Pengolahan *RapidMiner Scatter*

Gambar 5 menjelaskan hasil pengolahan data tahun 2018 pada tampilan Scatter. Gambar 5 dapat dilihat Clutser Tinggi berwarna biru dengan jumlah 2 items, Cluster Sedang berwarna Merah dengan jumlah 1 items dan Cluster Rendah berwarna hijau dengan jumlah 9 items.

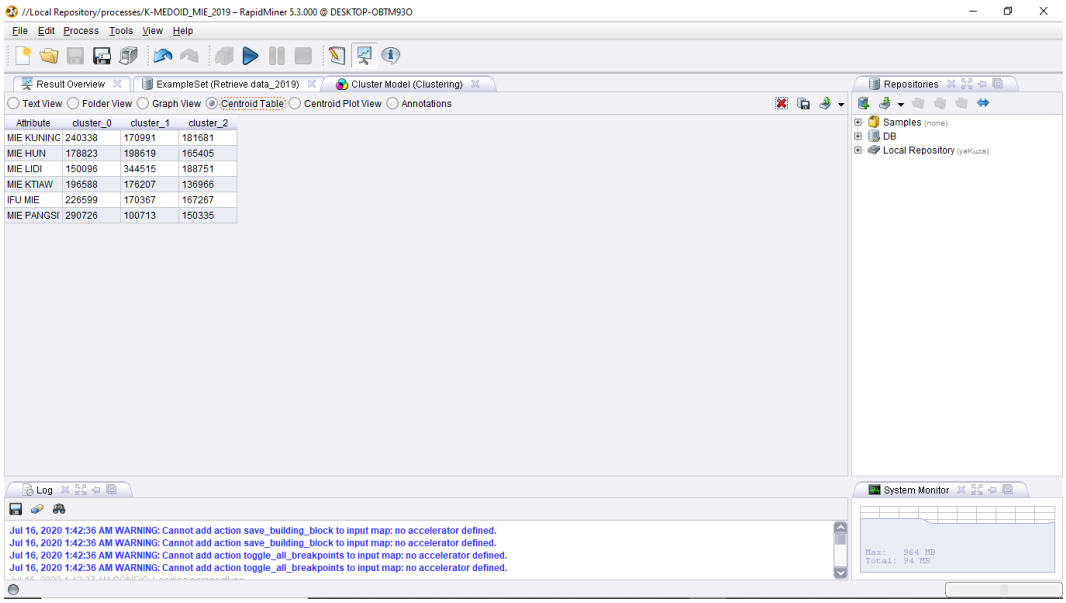
4. Hasil Pengolahan *RapidMiner* data tahun 2019
   Hasil yang diperoleh dari pengolahan Algoritma *K-Medoids* pada *RapidMiner* untuk data tahun 2019 adalah sebagai berikut :

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591



Gambar 6. Hasil Pengolahan *RapidMiner Cluster* Model

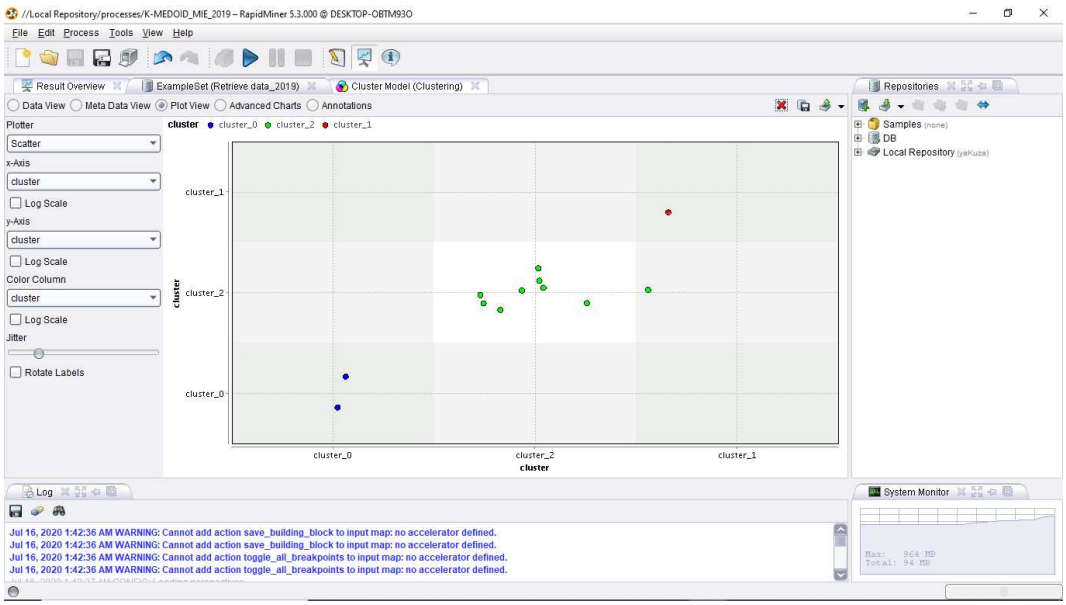
Gambar 6 menjelaskan hasil pengolahan *RapidMiner* dengan menampilkan *Cluster* Model dengan *Cluster* 0 sebagai *Cluster* Tinggi dengan jumlah 2 *items*, *Cluster* 1 sebagai *Cluster* Sedang dengan jumlah 1 *items*, dan *Cluster* 2 sebagai *Cluster* Rendah dengan jumlah 9 *items*



Gambar 7. Hasil Pengolahan *RapidMiner Centroid Table*

Gambar 7 menjelaskan hasil pengolahan *RapidMiner* dengan menampilkan hasil *Centroid Table.*

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591



Gambar 8. Hasil Pengolahan *RapidMiner Scatter*

Gambar 8 menjelaskan hasil pengolahan data tahun 2019 pada tampilan *Scatter*. Gambar 8 dapat dilihat *Clutser* Tinggi berwarna biru dengan jumlah 2 *items*, *Cluster* Sedang berwarna Merah dengan jumlah 1 *items* dan *Cluster* Rendah berwarna hijau dengan jumlah 9 *items*.

4.3. Pembahasan

Hasil yang dilakukan penulis dalam perhitungan Algoritma *K-Medoids* diperoleh dengan hasil tahun 2018 *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan *Cluster* Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November. Hasil tahun 2019 juga sama diperoleh *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan Cluster Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November.

Hasil yang dilakukan dalam *tools RapidMiner* dalam perhitungan Algoritma *K-Medoids* diperoleh dengan hasil tahun 2018 *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan *Cluster* Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November. Hasil tahun 2019 juga sama diperoleh *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan *Cluster* Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November.

Artinya Hasil dari proses yang dilakukan penulis dan *tools RapidMiner* adalah sesuai dan sama dari proses yang dilakukan. Kesamaan hasil yang dilakukan dapat dijadikan salah satu solusi pemecahan dari permasalahan yang diteliti.

5. KESIMPULAN

Berdasarkan pembahasan di atas dapat disimpulkan bahwa Penerapan data *mining* menggunakan algoritma *K-Medoids* untuk untuk mencari pengelompokkan tingkat penjualan mie di Kota Pematangsiantar diperoleh dengan hasil tahun 2018 *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan *Cluster* Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli,

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

Agustus, September, Oktober, dan November. Serta penerapan data *mining* menggunakan algoritma *K-Medoids* untuk untuk mencari pengelompokkan tingkat penjualan mie di Kota Pematangsiantar tahun 2019 juga sama diperoleh *Cluster* Tinggi dengan jumlah 2 *items* pada bulan Januari dan Desember, *Cluster* Sedang dengan jumlah 1 *items* pada bulan Juni, dan Cluster Rendah dengan jumlah 9 *items* pada bulan Februari, Maret, April, Mei, Juli, Agustus, September, Oktober, dan November.

Pengujian data pada *Rapidminer5.3* dengan menggunakan algoritma *KMedoids* berhasil menampilkan dtinggi *cluster* dari hasil klasifikasi dengan presentase keakuratan sebesar 100%

DAFTAR PUSTAKA

[1] D. F. Pramesti, M. T. Furqon, dan C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan / Lahan Berdasarkan Persebaran Titik Panas (Hotspot)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, hal. 723–732, 2017.

[2] Z. Mustofa dan I. S. Suasana, "ALGORITMA CLUSTERING K-MEDOIDS PADA E-GOVERNMENT BIDANG INFORMATION AND COMMUNICATION TECHNOLOGY DALAM PENENTUAN STATUS EDGI," vol. 9, hal. 1–10, 2018.

[3] D. F. Pramesti, Lahan, M. Tanzil Furqon, dan C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, hal. 723–732, 2017, doi: 10.1109/EUMC.2008.4751704.

[4] S. Defiyanti, M. Jajuli, dan N. Rohmawati, "Optimalisasi K-MEDOID dalam Pengklasteran Mahasiswa Pelamar Beasiswa dengan CUBIC CLUSTERING CRITERION," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 1, hal. 211–218, 2017, doi: 10.25077/teknosi.v3i1.2017.211-218.

[5] M. Ependi, Sopyan; Akbar, "implementasi Data Mining pada Penjualan Produk Elektronik Dengan Algotitma apriori," *implementasi Data Min. pada Penjualan Prod. Elektron. Dengan Algotitma apriori*, vol. 43, no. 5, hal. 10–23, 2013.

[6] H. Zayuka, S. M. Nasution, dan Y. Purwanto, "Perancangan Dan Analisis Clustering Data Menggunakan Metode K-Medoids Untuk Berita Berbahasa Inggris Design and Analysis of Data Clustering Using K-Medoids Method For English News," *e-Proceeding Eng.* , vol. 4, no. 2, hal. 2182–2190, 2017.

[7] H. Ningrum, E. Irawan, dan M. R. Lubis, "Implementasi Metode K-Medoids Clustering Dalam Pengelompokan Data Penyakit Alergi Pada Anak," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.*, vol. 6, no. 1, hal. 130, 2021, doi: 10.30645/jurasik.v6i1.277.

[8] D. Listiyanti, Y. A. Syahbana, dan S. R. Henim, "Perancangan dan Implementasi Aplikasi Android Penentu Salient Area pada Video dengan Algoritma K-Medoids," vol. 2, no. 1, hal. 96–101, 2016, [Daring]. Tersedia pada: http://ars.ilkom.unsri.ac.id.

[9] A. A. D. Sulistyawati dan M. Sadikin, "Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan," *Sistemasi*, vol. 10, no. 3, hal. 516, 2021, doi: 10.32520/stmsi.v10i3.1332.

[10] S. Darma dan G. W. Nurcahyo, "Klasterisasi Teknik Promosi dalam Meningkatkan Mutu Kampus Menggunakan Algoritma K-Medoids," *J. Inform. Ekon. Bisnis*, vol. 3, hal. 89–94, 2021, doi: 10.37034/infeb.v3i3.87.

[11] J. Perbanas *et al.*, "'Towards Economic Recovery by Accelerating Human Capital and Digital Tranformation' Perbanas Institute-SNAP_2021_FULL PAPER_41 ANALISIS DATA RISIKO NASABAH PADA BUSINESS CONTROL (BC) TOOLS MENGGUNAKAN RAPID MINER," hal. 178–189.

[12] B. G. Sudarsono, M. I. Leo, A. Santoso, dan F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, hal. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.

[13] Ainurrohmah, "Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka," *Prisma*, vol. 4, hal. 493–499, 2021, [Daring]. Tersedia pada:

DOI: https://doi.org/10.54082/jiki.15
P-ISSN: 2807-6664
E-ISSN: 2807-6591

https://journal.unnes.ac.id/sju/index.php/prisma/.

[14] G. A. Pranata, H. Tanuwijaya, dan P. Sudarmaningtyas, "Rancang Bangun Sistem Informasi Permintaan Pembelian Barang Berbasis Web Di Stmik Stikom Surabaya," *J. Sist. Inf. dan Komput. Akunt.*, vol. 3, no. 1, hal. 197–203, 2015.

[15] T. Informatika, F. I. Komputer, U. Al, A. Mandar, dan T. Sampah, "PENERAPAN UNIFIED MODELLING LANGUAGE (UML) PADA ANALISIS SISTEM SERTA PERANCANGAN DATABASE," vol. 6, hal. 170–177, 2021.