

CLUSTERING PENCAPAIAN TARGET PENJUALAN RUMAH PARA KARYAWAN MARKETING MENGGUNAKAN RAPID MINER DAN ALGORITMA K-MEANS

Muhammad Fahmi Julianto¹, Sofian Wira Hadi², Setiaji³, Windu Gata⁴, Rangga Pebrianto⁵
STMIK Nusa Mandiri^{1,2,3,4,5}

14002389@nusamandiri.ac.id¹, 14002361@nusamandiri.ac.id², setiaji.sej@nusamandiri.ac.id³,
windu@nusamandiri.ac.id⁴, 14002396@nusamandiri.ac.id⁵

Abstrak - Dalam persaingan dunia bisnis saat ini, kita dituntut untuk senantiasa mengembangkan bisnis agar selalu bertahan dalam persaingan. Fachry PropertyLand merupakan salah satu bidang usaha yang fokus di bidang property. Di mana toko ini memenuhi kebutuhan konsumen yang sedang trend saat ini. Pada Fachry Property Land terdapat beberapa permasalahan yang selalu muncul terkait penjualan. Dimana banyak karyawan yang tidak memenuhi target penjualannya. Berdasarkan hal tersebut maka diharapkan bisa mempermudah pihak Fachry Property Land dalam menilai kelayakan atas karyawannya dalam menentukan karyawan yang memenuhi target, belum memenuhi target dan tidak memenuhi target, dalam proses pengelompokan maka akan digunakan sebuah metode pengelompokan menggunakan Algoritma K-Means Clustering dengan implementasinya menggunakan sebuah software Data Mining yaitu, RapidMiner Studio versi 9.2. Dengan adanya aplikasi Rapid Miner Studio ini diharapkan pemilik Fachry Propertyland dapat melihat hasil pengelompokan karyawan yang memenuhi target, belum memenuhi target dan tidak memenuhi target. Sehingga pemilik Fachry Propertyland dapat melakukan tindakan kepada karyawannya tersebut.

Kata Kunci : Clustering, Rapid Miner, K-Means

Abstract - In the competition of the business world today, we are required to always develop business in order to always be successful in competition. Fachry PropertyLand is one of the business fields engaged in the sale of homes. Everywhere this shop must meet the needs of customers who are currently trending. On Land Fachry Property Around the issue that always appears regarding sales. Where many employees do not meet their sales targets. Based on this, it is expected to facilitate the Land Fachry Property in assessing the appropriateness of its employees in determining employees who have met the target, has not met the target and does not meet the target, in the grouping process, the grouping method will be used using the K-Me Clustering Algorithm as a method of manual replacement and in its implementation the Data Mining software uses RapidMiner Studio version 9.2. With the application of Rapid Miner Studio, it is expected that the owner of Fachry Propertyland can see the results of the grouping that meets the target, does not meet the target and does not meet the target. It is expected that the owner of Fachry Propertyland can take action on these employees.

Keywords: Clustering, Rapid Miner, K-Means

I. PENDAHULUAN

Kemajuan teknologi informasi pada saat ini sudah semakin berkembang pesat disegala aspek bidang kehidupan. Banyak sekali data yang diperoleh dari teknologi informasi yang canggih, mulai dari bidang industri, ekonomi, ilmu dan teknologi maupun berbagai aspek kehidupan lainnya. Dalam persaingan dunia bisnis saat ini, kita dituntut untuk senantiasa mengembangkan bisnis agar selalu bertahan dalam persaingan (M.Hasyim Siregar, S.Kom., 2018). Fachry PropertyLand merupakan salah satu bidang usaha yang fokus di bidang Property. Dimana toko ini harus memenuhi kebutuhan konsumen yang sedang trend saat ini. Untuk dapat melakukan hal tersebut, maka dibutuhkan sumber informasi untuk dapat dianalisa lebih lanjut. Pada Fachry Property Land terdapat beberapa permasalahan yang selalu muncul mengenai penjualan. Berdasarkan hal tersebut maka penulis berharap bisa mempermudah pihak Fachry Property Land

dalam menilai kelayakan atas karyawannya dalam pencapaian target penjualan. Tidak hanya itu penulis juga mengharapkan bisa memberikan rangsangan kepada pengelola data supaya lebih efisien dalam menggali informasi yang dibutuhkan oleh Fachry Property Land, sehingga bisa menjawab kebutuhan dari perusahaan tersebut. Penulis berharap dapat membantu memberikan informasi mengenai hasil pengklasteran data karyawan yang memenuhi target dan tidak memenuhi target penjualan. Dengan demikian, dibutuhkan analisa yang menggunakan software data mining guna menunjang arus data dan informasi sesuai dengan kebutuhan dari proses-proses tersebut. *Knowledge discovery in Database* (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses KDD menghasilkan sebuah proses data mining (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasil

tersebut secara akurat menjadi suatu informasi yang mudah dipahami (Asroni Asroni, 2015). Proses KDD secara garis besar terdiri dari Data Selection, Pre-processing/Cleaning, Transformation, Data mining, dan Interpretation / Evaluation. Algoritma K-Means merupakan algoritma non hirarki yang berasal dari metode *data clustering*. metode K-Means ini membagi data kedalam kelompok sehingga data mempunyai karakteristik sama dimasukkan kedalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain (Prasetyo, 2012). Menerapkan algoritma pengelompokan k-means tradisional dan jarak pengukuran Euclidean dari kesamaan dipilih untuk digunakan dalam analisis nilai (Garg, 2016). Algoritma K-means dapat bekerja dengan baik untuk beberapa dokumen tertentu (Jayant Tikmani, Sudhanshu Tiwari, 2015). Adapun tujuan dari diadakannya pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok pemilihan metode K-Means dikarenakan metode ini harus menggunakan data fisik yang bersifat jelas, hal ini sesuai dengan data yang akan digunakan pada permasalahan di dalam pengelompokan penjualan property. Selain itu, metode ini bersifat fleksibel karena pengguna dapat menentukan jumlah kluster yang akan dibuat (Mardalius, 2018). Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means clustering*. Dengan menggunakan metode ini, data yang sudah didapatkan dapat dibagi kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama (Ong, 2013).

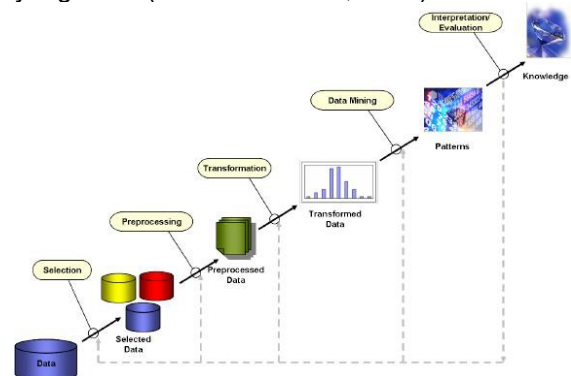
II. METODOLOGI PENELITIAN

Dalam metode penelitian ada beberapa tahapan yang harus dilakukan, tahapan kerja ini merupakan tahapan yang harus dilewati dan dikerjakan sesuai dengan pokok masalah agar tidak keluar jalur dari batasan masalah yang ditentukan dan penelitian ini berjalan dengan baik (M. Hasyim Siregar, S. Kom., 2018). Pada kerangka penelitian yang dipakai yaitu mempelajari literatur, mengumpulkan data, menganalisa data, menganalisa metode K-Means, menguji hasil, dan menganalisa hasil.

2.1. KDD (*Knowledge Discovery In Database*)

Data mining dan *knowledge discovery in database* (KDD) sering digunakan secara

bergiliran untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar (Nasari & Darma, 2015).



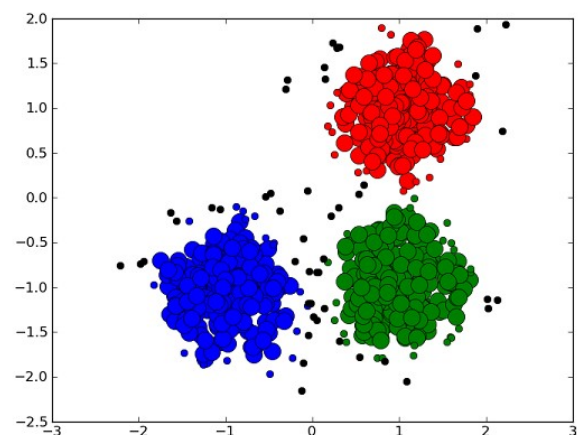
Gambar 1. Aliran Informasi Data Mining

a. Data Mining

Data Mining merupakan analisa terhadap sebuah data untuk menemukan hubungan yang nyata serta menyimpulkannya dengan yang belum diketahui sebelumnya dengan cara terkini sehingga mudah dipahami dan dapat berguna bagi pemilik data tersebut (Widodo, 2004).

b. Clustering

Clustering merupakan metode yang digunakan untuk membagi suatu rangkaian data menjadi beberapa kelompok berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya (Dhuhita, 2015).



Gambar 2. Contoh Pada Clustering

c. Algoritma K-Means

Algoritma K-Means merupakan suatu *algoritma* yang membagi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam bagian kelompok yang lain (Prasetyo, 2012).

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut (Alfian, Sandi, Raharjo, & Putra, 2018):

1. menentukan berapa jumlah cluster
2. menempatkan data ke dalam cluster secara acak
3. Hitung *centroid* rata-rata dari data yang ada di masing-masing cluster
4. menempatkan masing-masing data ke *centroid*/rata-rata terdekat
5. Kembali ke langkah 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai *centroid*, ada yang di atas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada object function yang digunakan di atas nilai *threshold* yang ditentukan.

d. *Cleaning*

Sebelum masuk ke dalam proses *data mining*, perlu dilakukan proses pembersihan pada data yang menjadi focus KDD. Proses pembersihan bermaksud untuk membuang *duplikasi data*, memeriksa data yang *inkosisten*, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*)(Nasari & Darma, 2015).

e. *Interpretation / Evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang dapat dimengerti oleh pihak yang berkepentingan. Langkah ini merupakan bagian dari proses KDD yang disebut *interpretation*. langkah ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya(Nasari & Darma, 2015).

III. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data yang di uji bersumber dari data penjualan pada farchy property land, dalam data yang diambil terdapat atribut-atribut seperti nama marketing, target penjualan, satuan, unit terjual, harga, persentase penjualan, dan lokasi perumahan yang terjual. Periode yang digunakan dari data tersebut bulan januari sampai desember 2019.

2. Praproses Data

Pada tahapan ini data yang telah dikumpulkan akan masuk pada tahap pemilihan atribut yang akan dipilih untuk di olah nanti nya.

Tabel 1. Tabel Penjualan

NO	Nama Attribut	Keterangan
1	Nama Marketing	Nama marketing penjualan
2	Jumlah Target	Jumlah target yang harus tercapai.
3	Satuan	Unit/rumah
4	Unit terjual	Unit yang berhasil dijual oleh marketing

5	Harga	Harga rumah
6	Persentasi Pencapaian	Persentasi yang diakumulasikan dengan unit terjual dan harga
7	Lokasi Perumahan	Letak perumahan yang tersedia

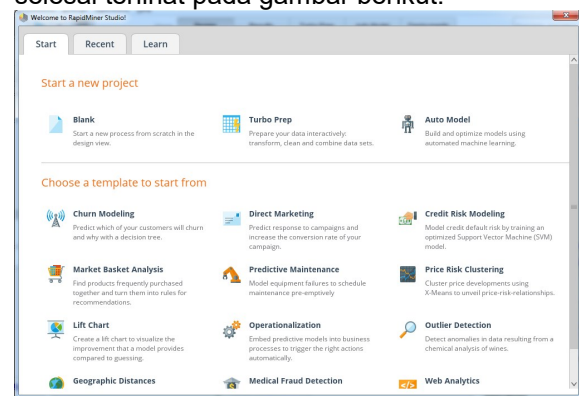
3. Proses Klastering menggunakan algoritma K-Means.

Tahapan dalam pemakaian algoritma K-means adalah menentukan banyaknya cluster, cluster yang akan dibuat sebanyak 3 cluster. Banyaknya cluster harus lebih kecil dari pada banyaknya data.

Attribute	cluster_0	cluster_1	cluster_2
Unit Terjual	0.249	2.281	4.459
Harga (Rp)	40595135.908	400179856.115	784672131.148
Persentasi Pencapaian	10.250	100.259	196.361

Gambar 3. Centroid Table

Gambar 3 adalah gambar centroid table yang dihasilkan menggunakan aplikasi *Rapid Miner Studio v.9.2*. untuk bagaimana mendapatkan hasil tersebut akan dijelaskan dibawah ini. Pada gambar 4 merupakan halaman utama dari aplikasi *RapidMiner Studio v.9.2* saat pertama kali membuka aplikasi setelah proses loading selesai terlihat pada gambar berikut:

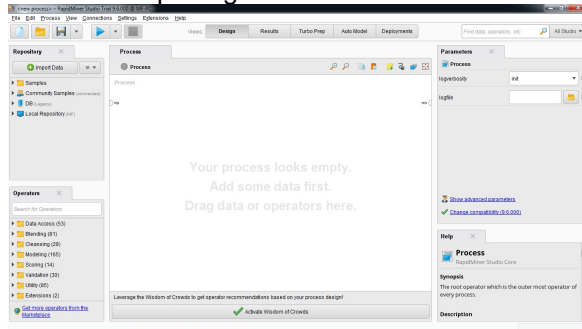


Gambar 4. Halaman Utama *Rapid Miner Studio V 9.2*.

Pada tampilan halaman utama ada tiga menu yang akan digunakan yaitu :

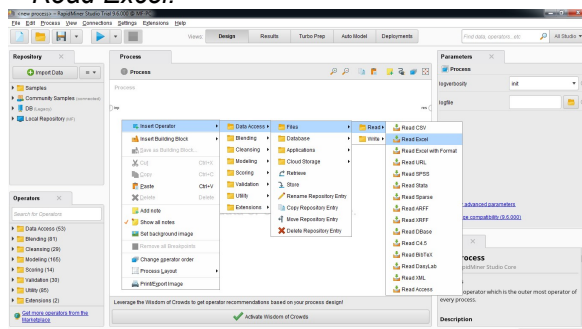
1. *Icon New Process* berfungsi untuk membuat halaman proses kerja *Data Mining* yang baru.
2. *Open Process* berfungsi untuk membuka proses yang sudah ada direpository sebelumnya.
3. *Learn* berisikan petunjuk-petunjuk menggunakan *RapidMiner Studio v.9.2*.

Berikut ini adalah tampilan *New Process* yang akan digunakan untuk memulai proses kerja baru terlihat pada gambar 5



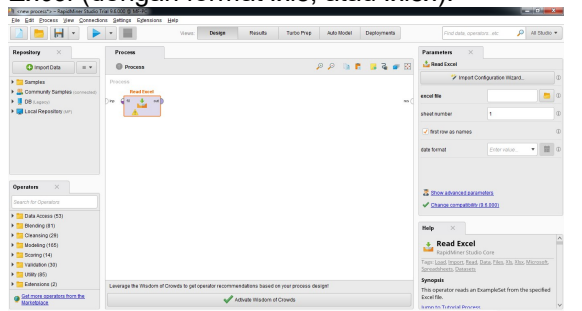
Gambar 5. *New Process RapidMiner Studio v.9.2*

Tampilan *new process* adalah untuk membuat halaman kerja pada *RapidMiner Studio v.9.2*. *Import data* dilakukan untuk memasukkan data yang akan diuji dalam bentuk format *.xls* atau *.xlsx*. Berikut adalah cara untuk melakukan import file Microsoft Excel. Untuk membuat mengimport data yang akan diproses, maka dilakukan *New process*, dengan cara klik kanan > *Insert Operator* > *Data Access* > *Files* > *Read* > *Read Excel*.



Gambar 6. Tampilan *Import Data* Dengan Membuat Operator Baru

Setelah selesai maka akan tampil *read excel* pada *RapidMiner Studio v.9.2*, Seperti terlihat pada Gambar 8. Setelah itu lanjutkan dengan mengklik Tab *Import Configuration Wizard* untuk melakukan menginputkan data dari *Microsoft Excel* (dengan format *.xls*, atau *.xlsx*).

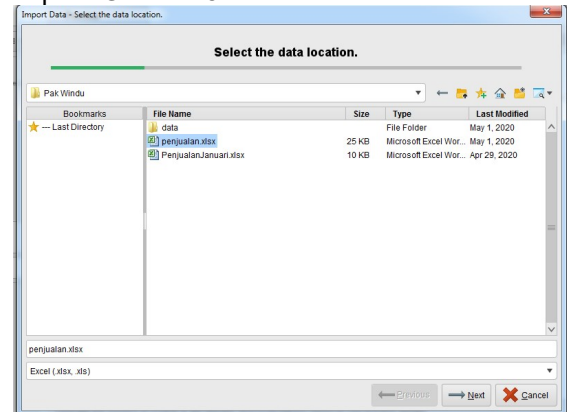


Gambar 7 Tampilan Operator Baru *Read Excel*.

Dalam melakukan *import data* pada aplikasi *RapidMiner Studio v.9.2* terdiri dari 3 tahap, adapun langkah-langkahnya sebagai berikut:

1. *Data Import Wizard – Step 1 of 3*

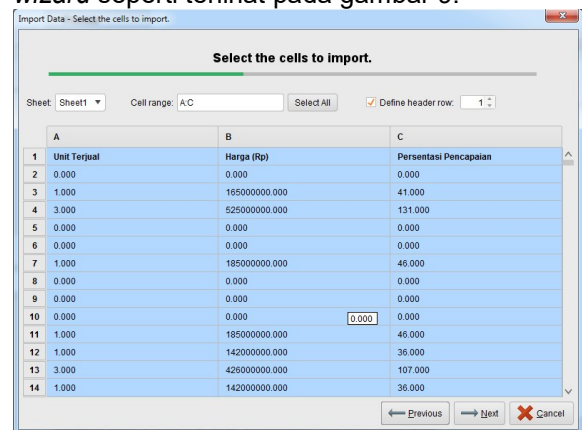
Tahap pertama adalah mencari lokasi file yang sudah ada sebelumnya dengan format *.xlsx* atau *.xls*, pilih dan kemudian simpan. Pada pengujian ini data yang akan diuji disimpan dengan nama *PENJUALAN.xlsx* lalu pilih diklik, seperti Gambar 8



Gambar 8. Tampilan *Data Import Wizard - Step 1 of 4*

2. *Data Import Wizard – Step 2 of 3*

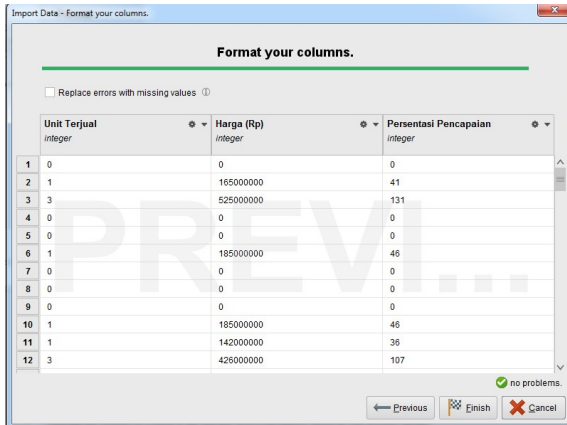
Kemudian pilih *next* maka lanjut ke tahap berikutnya yaitu *import wizard step 2 of 4*. Pada tahap ini akan dipilih *Sheet* yang berisikan sumber data pada *Microsoft Excel* yang digunakan. Maka akan terlihat data *import wizard* seperti terlihat pada gambar 9.



Gambar 9 Tampilan *Data Import Wizard - Step 2 of 4*

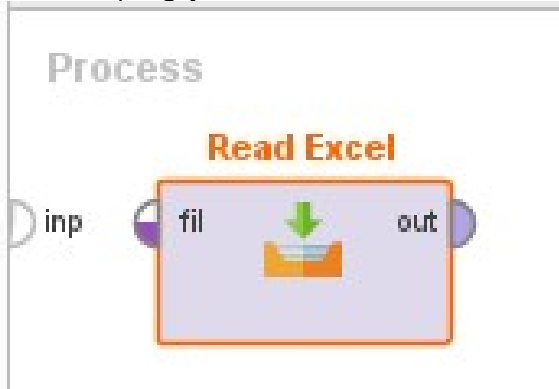
3. *Data Import Wizard – Step 3 of 3*

Pada *Step 3*, tahap ini merupakan tahap penentuan tipe data dan atribut. Sebenarnya *RapidMiner Studio* akan memberikan tipe data yang tepat secara otomatis. Namun, jika kita merasa tipe data yang diberikan *RapidMiner Studio* tidak cocok, kita bisa mengubahnya. Kemudian klik tombol *Next*.



Gambar 10 Tampilan *Data Import Wizard - Step 4 of 4*

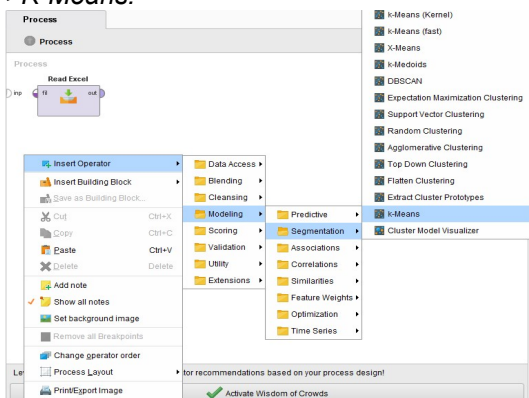
Tahap *import* data selesai dengan memilih *Finish*, maka pada *Main Process* terlihat sebuah operator baru *Read Excel* yang sudah berisi file *PENJUALAN.xlsx* yang di *import* langsung dari *file excel*. Data tersebut siap dilakukan pengujian.



Gambar 11 Tampilan Operator Data Siap Import

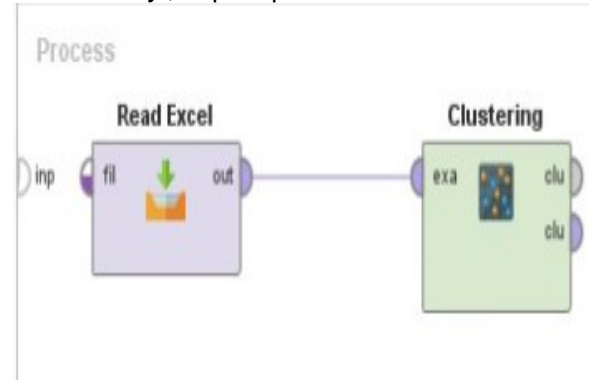
Pada gambar 12 menampilkan data yang sudah di import, sehingga tombol pada operator *read excel* tidak ada lagi tanda seru warna kuning (menandakan operator telah berisi data dan siap diolah).

Langkah selanjutnya adalah menambahkan operator *K-Means* dengan cara klik kanan -> *Insert Operator -> Modeling -> Segmentation -> K-Means*.

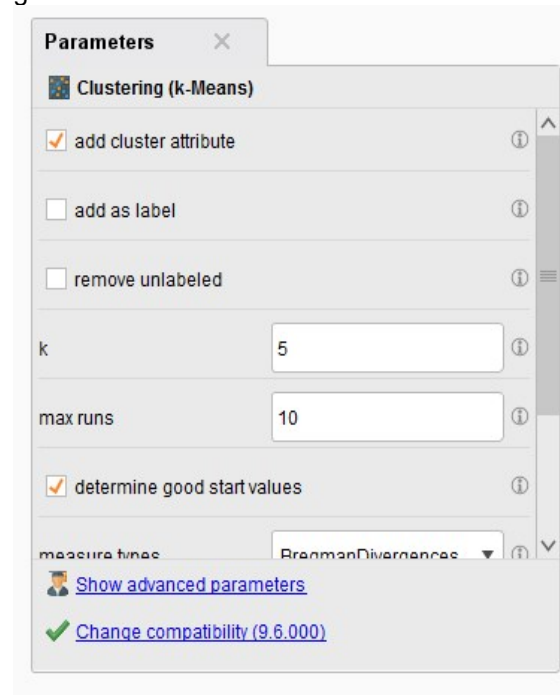


Gambar 12. Menambahkan Operator *K-Means*

Hubungkan data dengan *clustering k-means* untuk mengetahui output ke arah *result*. Apabila sudah terhubung maka tentukan jumlah *cluster* dan *maksimal runs* yang diberikan. Terakhir klik *Tombol Play*, seperti pada Gambar 13



Gambar 13 Tampilan Untuk Proses *K-Means* Selanjutnya melakukan pengaturan algoritma *K-Means*, yang diatur pada menu *Parameter Clustering K-Means*, seperti terlihat pada gambar 14



Gambar 14 Penentuan Jumlah *Cluster*

Pada gambar 14 di atas dilakukan pengaturan nilai *k*, di mana *k* merupakan nilai yang digunakan untuk menentukan jumlah *cluster* yang akan dibentuk. Di sini jumlah *cluster* yang akan dibentuk adalah sebanyak 3 *cluster* sesuai tingkatan bahaya yaitu rendah, sedang dan tinggi.

Setelah proses *running*, akan tampil *Example Set (Read Excel)*, seperti yang terlihat pada Gambar 15.

Row No.	id	cluster	Unit Terjual	Harga (Rp)	Persentasi ...
1	1	cluster_0	0	0	0
2	2	cluster_0	1	185000000	41
3	3	cluster_1	3	525000000	131
4	4	cluster_0	0	0	0
5	5	cluster_0	0	0	0
6	6	cluster_0	1	185000000	46
7	7	cluster_0	0	0	0
8	8	cluster_0	0	0	0
9	9	cluster_0	0	0	0
10	10	cluster_0	1	185000000	46
11	11	cluster_0	1	142000000	36
12	12	cluster_1	3	426000000	107
13	13	cluster_0	1	142000000	36
14	14	cluster_1	3	457000000	114
15	15	cluster_1	2	390000000	98

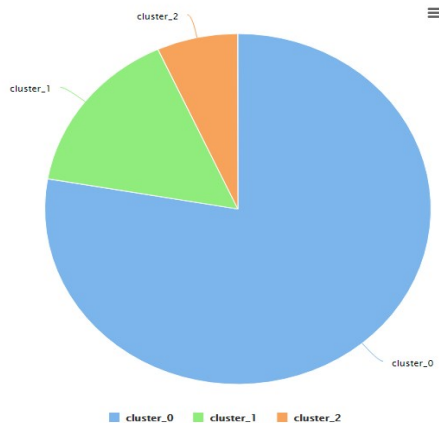
ExampleSet (899 examples, 2 special attributes, 3 regular attributes)

Gambar 15. Tampilan *Running Data*

1. *ExampleSet*

Pada *ExampleSet* dapat dilihat beberapa tampilan hasil *cluster*, yaitu *Data View*. *Data View* merupakan tampilan hasil *cluster* data secara keseluruhan sesuai dengan data yang telah diinputkan

Chart merupakan tampilan grafik hasil pengelompokan atau *cluster* sampel data nilai mata pelajaran ujian nasional dengan 3 *cluster*.



Gambar 16. Tampilan *Pie* pada *Chart*

2. *Cluster Model (Clustering)*

Pada *Cluster Model (Clustering)* dapat dilihat beberapa tampilan hasil *cluster*, yaitu *Text View* yang merupakan tampilan hasil pengelompokan berdasarkan *cluster* dan jumlah anggotanya. Tampilan *Text View* akan terlihat pada gambar 17.

Cluster Model

```
Cluster 0: 699 items
Cluster 1: 139 items
Cluster 2: 61 items
Total number of items: 899
```

Gambar 17. Tampilan *Text View*

Pada gambar 17 dapat kita lihat data berupa : cluster 0 yang memiliki kategori tidak memenuhi target mendapatkan jumlah sebanyak 699 items.

Cluster 1 yang memiliki kategori belum memenuhi target mendapatkan jumlah sebanyak 139 items.

Cluster 2 yang memiliki kategori memenuhi target mendapatkan jumlah sebanyak 61 items.



Gambar 18. Tampilan *Folder View*

Folder View merupakan tampilan data bagian-bagian *cluster* secara keseluruhan, dimana masing – masing anggota cluster menampilkan field.

IV. KESIMPULAN

Berdasarkan uraian yang sudah dikemukakan pada sebelumnya, maka dapat ditarik kesimpulan Metode *Clustering* dengan menggunakan algoritma *K-Means* dapat digunakan untuk mengelompokkan karyawan berdasarkan kinerjanya, yaitu tidak memenuhi target, belum memenuhi target, dan sudah memenuhi target. Sehingga pihak perusahaan dapat mengevaluasi kinerja karyawan yang tidak atau belum memenuhi target penjualan. Apa lagi dapat kita simpulkan bahwa data rekapan karyawan tersebut hanya sebanyak 61 kali dapat mencapai target penjualan. Peneliti menyadari adanya kekurangan dalam penulisan ini, karena keterbatasan penulis baik dalam hal waktu maupun pengetahuan. Dalam rangka memperbaiki kekurangan dan untuk penyempurnaan penelitian ini penulis memberikan beberapa saran Untuk penelitian selanjutnya dapat dilakukan dengan data yang lebih banyak dan parameter yang lebih banyak lagi serta Untuk memaksimalkan waktu proses *clustering* dengan algoritma *K-Means*, penelitian selanjutnya dalam hal pengekrasian data tercapainya target penjualan menerapkan metode *Data Mining* dengan algoritma yang berbeda.

V. REFERENSI

- Alfian, T., Sandi, A., Raharjo, M., & Putra, J. L. (2018). CLUSTERING KESETIAAN PELANGGAN E-RITEL DENGAN MODEL RFM, *14*(2), 239–246. Retrieved from <https://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/74>
- Asroni Asroni, R. A. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang, *18*(1), 76–82. Retrieved from https://www.researchgate.net/publication/329831347_Penerapan_Metode_K-Means_Untuk_Clustering_Mahasiswa_Berdasarkan_Nilai_Akademik_Dengan_Weka_Interface_Studi_Kasus_Pada_Jurusan_Teknik_Informatika_UMM_Magelang_Implementation_Method_for_K-Means_Clusteri
- Duhita, W. M. P. (2015). CLUSTERING MENGGUNAKAN METODE K MEANS UNTUK MENENTUKAN STATUS GIZI BALITA, *15*(2). Retrieved from <https://www.neliti.com/id/publications/103765/clustering-menggunakan-metode-k-mean-untuk-menentukan-status-gizi-balita>
- Garg, S. K. dan K. (2016). Document Clustering using Improved K-means Algorithm. *International Journal of Engineering Research and General Science*, *4*(3). Retrieved from <http://www.indianjournals.com/ijor.aspx?target=ijor:ijrss&volume=6&issue=9&article=017>
- Jayant Tikmani, Sudhanshu Tiwari, S. K. (2015). n Approach to Customer Classification using k-means. *International Journal of Innovative Research in Computer and Communication Engineering*, *3*(11). Retrieved from http://www.ijircce.com/upload/2015/november/29_An.pdf
- M.Hasyim Siregar, S.Kom., M. K. (2018). KLASERISASI PENJUALAN ALAT-ALAT BANGUNAN MENGGUNAKAN METODE K-MEANS, *1*(2), 83–91. Retrieved from <https://ejournal.uniks.ac.id/index.php/JTOS/article/view/24>
- Mardalius. (2018). PENGELOMPOKAN DATA PENJUALAN AKSESORIS MENGGUNAKAN ALGORITMA K-MEANS, *IV*(2), 401–411. Retrieved from https://www.researchgate.net/publication/330609314_PEMANFAATAN_RAPID_MINER_STUDIO_82_UNTUK_PENGELOMPOKAN_DATA_PENJUALAN_AKSESORIS_MENGGUNAKAN_ALGORITMA_K-MEANS
- Nasari, F., & Darma, S. (2015). PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU, 6–8. Retrieved from <https://www.ojs.amikom.ac.id/index.php/seminasteknomedia/article/viewFile/837/801>
- Ong, J. O. (2013). IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING, (April), 10–20. Retrieved from <http://journals.ums.ac.id/index.php/jiti/article/view/651>
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: CV. Andi Offset.
- Widodo. (2004). *Psikologi Belajar*. Jakarta: Rineka Cipta.