

# PENERAPAN METODE CLUSTERING PADA PENJUALAN BARANG

Muhammad Rifkah

Fakultas Rekayasa Sistem, Universitas Teknologi Sumbawa

Email: muhammadrifkah46@gmail.com

**Abstrak:** Pada minimarket, produk merupakan bahan pokok yang dijual belikan, produk di minimarket terdapat antara produk yang laku dan yang tidak laku. Dengan adanya masalah ini, perlu untuk menciptakan sistem yang dapat mengelompokkan produk laku dan produk tidak laku, yang dilakukan di minimarket UD. RATNA Sumbawa dan dilaksanakan pada bulan Juni sampai Juli tahun 2022. Algoritma K-Means tidak terpengaruh terhadap urutan objek yang digunakan, hal ini dibuktikan ketika penulis mencoba menentukan secara acak titik awal pusat cluster dari salah satu objek pada permulaan perhitungan, tujuan membuat aplikasi pada swalaya UD. RATNA kota Sumbawa menggunakan metode K-Means. Algoritma K-Means Cluster Analysis pada dasarnya dapat diterapkan pada permasalahan dalam memahami perilaku konsumen, mengidentifikasi peluang produk baru dipasaran dan Algoritma K-Means ini juga dapat digunakan untuk meringkas objek dari jumlah besar sehingga lebih memudahkan untuk mendeskripsikan sifat-sifat dan karakteristik dari masing-masing kelompok

**Kata kunci:** *clustering, K-Means, Algoritma Data mining, Produk*

## 1 PENDAHULUAN

Data mining merupakan suatu langkah dalam KKD yang memiliki teknik menganalisa data untuk digali informasi tersembunyi dalam jumlah besar dan kompleks, sehingga menghasilkan output karakteristik atau pola dari data tersebut. Clustering merupakan metode analisis data yang tujuannya mengelompokkan data dengan karakteristik yang sama ke suatu wilayah yang sama. Salah satu pendekatan yang digunakan dalam mengembangkan metode clustering yaitu metode K-Means, dimana metode ini merupakan salah satu metode pengelompokan data yang berusaha mempartisi data ke dalam dua bentuk dua atau lebih kelompok (cluster) yang berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama.

Data mining sudah banyak diterapkan dalam berbagai bidang, salah satu metode clustering khususnya K-Means, antara lain oleh Andika (2008) menjelaskan tentang algoritma *K-Means clustering* dapat diimplementasikan dengan model perangkat lunak untuk verifikasi citra sidik jari poin *minutiae*. Suprihatin (2011) menyatakan bahwa teknik *clustering K-Means* digunakan untuk menentukan nilai huruf ujian akhir pada Universitas Teknologi Sumbawa

Berdasarkan hal tersebut maka diharapkan bisa mempermudah pihak UD. RATNA dalam menemukan informasi dari tumpukan-tumpukan data yang berguna bagi UD. RATNA. Tidak hanya itu penulis juga mengharapkan bisa memberikan rangsangan kepada pengelola data supaya lebih aktif dalam menggali informasi yang

dibutuhkan oleh UD RATNA, sehingga bisa menjawab kebutuhan dari Penulis berharap dapat membantu memberikan informasi mengenai hasil *Clustering* data penjualan.

Dengan demikian, dibutuhkan analisis yang terkomputerisasi menggunakan *software* data mining yang menunjang arus data dan informasi sesuai dengan kebutuhan dari proses-proses tersebut

## 2 TINJAUAN PUSTAKA

### A) Algoritma klasifikasi K-Means

K-Means merupakan algoritma clustering yang berulang-ulang. Algoritma K-Means dimulai dengan pemilihan secara acak K, K disini merupakan banyaknya Cluster yang ingin dibentuk. Kemudian tetapkan nilai-nilai K secara random, untuk sementara nilai tersebut menjadi pusat dari cluster atau biasa disebut dengan centroid, mean atau "means". Hitung jarak setiap data yang ada terhadap masing-masing centroid menggunakan rumus Euclidian hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Klasifikasikan setiap data

### B) Data Mining

menurut Sulianta (2010:17) Data mining adalah metoda yang digunakan untuk mengekstraksi informasi prediktif tersembunyi pada database, ini adalah teknologi yang sangat potensial bagi perusahaan yang sangat potensial bagi perusahaan dalam memberdayakan data *warehouse*.

Secara garis besar, data mining dapat dikelompokkan menjadi 2 kategori utama, yaitu:

*Descriptive mining*, yaitu proses untuk menemukan karakteristik penting dari data dalam satu basis data. Teknik data mining yang termasuk *descriptive mining* adalah *clustering, asosiasi, dan sequential mining*.

*Predictive*, yaitu proses untuk menemukan pola dari data dengan menggunakan beberapa variabel lain di masa depan. Salah satu teknik yang terdapat dalam *predictive mining* adalah klasifikasi. Secara sederhana data mining biasa dikatakan sebagai proses penyaringan atau "menambang" pengetahuan dari sejumlah data yang besar. Istilah lain untuk data mining adalah *Knowledge Discovery in Database (KDD)*.

### C) Clustering

Menurut Widodo (2013:9) *Clustering* atau klasifikasi adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya.

Dengan menggunakan clustering ini, kita dapat mengklasifikasikan daerah yang padat, menemukan pola-pola

distribusi secara keseluruhan, dan menemukan keterkaitan yang menarik antara atribut data. Dalam data mining, usaha difokuskan pada metode-metode penemuan untuk cluster pada basis data berukuran besar secara efektif dan efisien.

#### D) Metode K-means

*K-Means* merupakan salah satu metode data clustering non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbedadikelompokkan ke dalam kelompok yang lainnya.

*K-Means* adalah metode clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numeric. Algoritma *K-Means* termasuk partitioning clustering yang memisahkan data ke  $k$  daerah bagian yang terpisah. Algoritma *K-Means* sangat terkenal karena kemudahan dan kemampuannya untuk mengcluster data yang besar dan data outlier dengan sangat cepat.

Algoritma *K-Means* merupakan metode non-hierarki yang pada awalnya mengambil sebagian banyaknya komponen populasi untuk dijadikan pusatcluster awal. Pada tahap ini pusat cluster dipilih secara acak dari sekumpulan populasi data. Berikutnya *K-Means* menguji masing-masing komponen di dalam populasi data dan menandai komponen tersebut ke salah satu pusat cluster yang telah didefinisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap cluster. Posisi pusat cluster akan dihitung kembali sampai semua komponen data digolongkan kedalam tiap-tiap pusat cluster dan terakhir akan terbentuk posisi pusat cluster yang baru.

#### E) Produk

Produk menurut Kotler dan Amstrong (1996:274) adalah segala sesuatu yang ditawarkan ke pasar untuk mendapatkan perhatian, dibeli, dipergunakan dan yang dapat memuaskan keinginan atau kebutuhan konsumen.

suatu produk adalah kumpulan dari atribut-atribut yang nyata maupun tidak nyata, termasuk di dalamnya kemasan, warna, harga, kualitas dan merk ditambah dengan jasa dan reputasi penjualannya.

### 3 METODOLOGI PENELITIAN

Metodelogi yang digunakan dalam kasus ini adalah metode clustring yang didasarkan dari metode yang ada pada data maining adapun langka-langka yang dilakukan dalam metode ini

#### A) Model Pengembangan Sistem

Dalam penelitian ini model pengembangan sistem yang digunakan yaitu model *incremental* yang merupakan model pengembangan sistem pada *software engineering* berdasarkan *requirement software* yang dipecah menjadi beberapa fungsi atau bagian sehingga model pengembangannya secara bertahap.

#### B) Proses metode clustering k-means

Pertama akan kita bahas dulu metode *cluster* secara statistik untuk non hirachical method yaitu: *K-Means Clustering*.

Algoritma:

- 1) Partisi item menjadi K initial cluster
- 2) Lakukan proses perhitungan dari daftar item, tandai item untuk kelompok yang mana berdasarkan pusat (mean) yang terdekat (dengan menggunakan *distance* dapat digunakan *Euclidean distance*). Hitung kembali pusat centroid untuk item baru yang diterima pada cluster tersebut dari cluster yang kehilangan item.
- 3) Ulangi step 2 hingga tidak ada lagi tempat yang akan ditandai sebagai cluster baru.

#### C) DFD (Data Flow Diagram)

proses *clustering k-means* dan proses *clustering k-means* membalas dengan mengirimkan data barang penjualan dan data barang penjualan mengirimkan ke aplikasi, dari aplikasi membuat laporan penjualan laporan barang kemudian ke pimpinan.

#### D) Perancangan File Tabel

#### E) Rancangan Menu

Pada tahap ini akan dilakukan proses utama yaitu segmentasi atau pengelompokkan data Penjualan barang yang diakses dari database, yaitu sebuah metode clustering algoritma K-Means. Berikut ini merupakan diagram flowchart dari algoritma K-Means dengan asumsi bahwa parameter input adalah jumlah dataset sebanyak  $n$  data dan jumlah inialisasi centroid  $K=3$  sesuai dengan penelitian. Dari banyak data penjualan yang diperoleh, diambil 14 jenis barang untuk dijadikan sampel untuk penerapan algoritma k-means dalam penjurusan mahasiswa. Percobaan dilakukan dengan menggunakan parameter-parameter berikut.

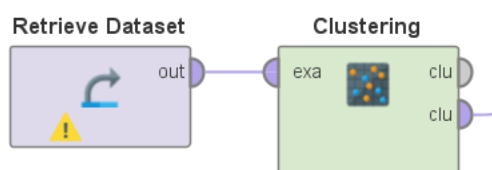
#### F) Perancangan Pengujian

Dalam tahapan ini user melakukan pengujian dengan metode pengujian yang dipakai adalah *black box testing*. *black box testing* atau test fungsional adalah pengujian program yang dilakukan oleh pengembang (Programmer) dengan memberikan input tertentu dan melihat hasil yang didapatkan dari input tersebut. dengan kata lain, *black box testing* terfokus pada fungsionalitas sistem.

### 4 PEMBAHASAN

#### A) Analisa Sistem Aktual

Dari hasil penelitian yang dilakukan di UD. RATNA Sumbawa, didapatkan sistem yang sudah berjalan dan digunakan saat ini masih manual. Disamping itu, pemilik UD. RATNA Sumbawa mempunyai kesulitan dalam mengklasifikasi produk yang laku dengan yang tidak laku. Hal tersebut tentunya akan sulit karena banyaknya produk yang masuk dan keluar setiap harinya.



#### B) Analisa Sistem Baru

Dari kekurangan sistem yang sedang berjalan tersebut maka dibuatlah suatu aplikasi pengelompokkan penjualan produk dengan menerapkan metode k-means dengan

menggunakan bahas pemograman mengklasifikasikan penjualan produk yang laku dan kurang laku. Sehingga pemesanan barang yang kurang laku dapat dikurangi.

### C) Perhitungan jarak pusat cluster

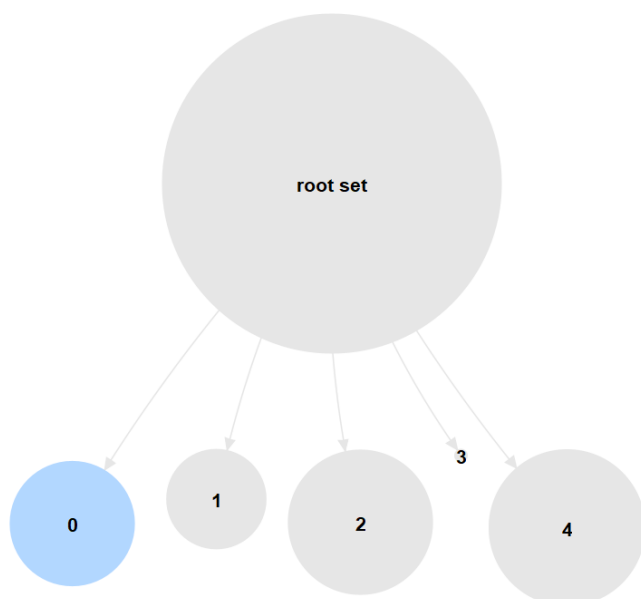
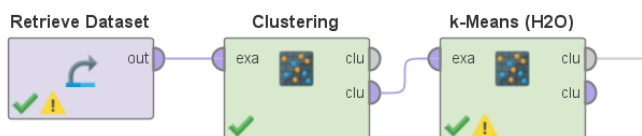
Pengelompokan objek dilakukan dengan menggunakan metode K-MeansCluster yaitu mengelompokkan n buah objek ke dalam k kelas berdasarkan jaraknya dengan pusat kelas. Hasil dari penelitian ini terhadap 20 objek sampel diperoleh 3 kelompok mahasiswa berdasarkan nilai BMI ( Body Mass Index) dan ukuran rangka, yaitu BMI normal dengan kerangka besar, BMI obesitas sedang dengan kerangka sedang, dan BMI obesitas berat dengan kerangka kecil.

Row No.	Hari	Tanggal	Kegiatan	Curah Huja...	Penjualan
1	2	1	0	1.400	0
2	3	2	1	1.200	27
3	4	3	1	15.200	458
4	5	4	1	13.400	20
5	6	5	1	7.700	7
6	0	6	1	0.600	0
7	1	7	1	1	4
8	2	8	1	0	350
9	3	9	1	0	1
10	4	10	1	3.200	240
11	5	11	1	0	13
12	6	12	1	6.200	169
13	0	13	0	0	0

ExampleSet (121 examples, 0 special attributes, 5 regular attributes)

### D) Hasil Perancangan Sistem

Metode analisis yang digunakan adalah metode clustering yang dimana metode ini mengumpulkan objek-objek data yang semisal satu sama yang lainnya dalam cluster



### E) Pengujian Sistem

Tahap pengujian atau testing merupakan salah satu tahap yang harus ada dalam sebuah siklus pengembangan

```

Cluster 0: 28 items
Cluster 1: 22 items
Cluster 2: 33 items
Cluster 3: 2 items
Cluster 4: 36 items
Total number of items: 121
  
```

## 5 KESIMPULAN

### A) Kesimpulan

Berdasarkan hasil dan pembahasan, maka kesimpulan dari penelitian ini antara lain :

- 1) Pengklasteran dengan menggunakan metode K-Means didapatkan untuk produk yang laku dengan yang tidak laku.
- 2) Semakin banyak data penjualan barang yang diinput, maka semakin banyak clustering centroid yang terbentuk di proses metode k-means.
- 3) Produk yang laku terdiri dari makanan, minuman dan produk yang tidak laku terdiri dari kebanyakan dari kosmetik

### B) Saran

Adapun saran untuk penelitian selanjutnya yaitu :

- 1) Ada beberapa metode klastering yang lainnya yang masih perlu dianalisis untuk dijadikan sebagai referensi untuk mengembangkan metode k-means
- 2) Menggunakan metode atau algoritma klas-tering yang lainnya untuk mengelompokkan produk

## 6 DAFTAR PUSTAKA

Agusta, Y. 2007. K-means-Penerapan, Permasalahan dan Metode Terkait. Jurnal Sistem dan Informatika Vol.3 (Februari 2007): 47-60.

Ems, TIM. 2012. Web Progreming for Beginners.

Jakarta: PT Elex Media Komputindo.

Feri Sulianta, Dominikus juju, (2010), Data Mining Meramalkan Bisnis. Perusahaan, Jakarta : Elex Media Komputindo

Hidayat, T., dan Istiadah, N. 2011 . Panduan Lengkap Menguasai SPSS 19 untuk Mengolah Data Statistik Penelitian. Jakarta: Media Kita.

Jurnal by Ahmad Afif. Clustering Data Penjualan dan Persediaan Barang pada PT. Sayap Mas Utama dengan Metode K-Means. STMIK GI MDP

Jurnal by Johan Oscar Ong. Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University. Program Studi Teknik Industri, President University, Jln. Jln. Ki Hajar Dewantara, Kota Jababeka, Cikarang Baru, Bekasi

Kurniawan, Rulianto.2010. PHP dan MySQL. Palembang. Maxikom

Pahevi, Said Mirza. 2013. Pembangunan Basis Data.  
Jakarta: PT. Elex Media Komputindo

Santosa, B. 2007. Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu

Widodo.2004. Psikologi Belajar. Jakarta: RinekaCipta.