

# APLIKASI PENGELOMPOKAN PELANGGAN PADA UMS STORE MENGUNAKAN ALGORITMA K-MEANS

Muhammad Abdul Ghofar<sup>1)</sup>, Yogie Indra Kurniawan<sup>2)</sup>

<sup>1), 2)</sup> Program Studi Informatika, Fakultas Komunikasi dan Informatika  
Universitas Muhammadiyah Surakarta  
mvghofar1@gmail.com<sup>1)</sup>, yogiek@ums.ac.id<sup>2)</sup>

## Abstract

*UMS Store is an official trading business unit owned by Muhammadiyah University of Surakarta that provides various categories of books, journals, office stationery and official merchandise. UMS Store is also a voucher exchange center for students. Of the many voucher redemption transactions and cash purchases, UMS Store has abundant data and will continue to grow over time. The abundant data if left unchecked would be a pile of only stored data. Actually, if the data is excavated will produce valuable information. UMS Store need a customer grouping application that will be used to provide continuous treatment such as giving discounts or vouchers to their best customers. This research is done to create the application, where application can make customer grouping with UMS Store data and can give recommendation through potential group that formed. This application was developed by utilizing K-means algorithm, which is one of clustering method in data mining technique. Groupings made in the application are limited to 3 large groups of data with restrictions using only student data using UMS Store vouchers. Variables used consist of NIM, year force, discount, sub total, total paid, total item and date. The results of this study is an application used to classify customers using K-means method. The results of this study indicate that if the application is used to create three groups, it will form three clusters, ie clusters of potential customers, normal customers and unlikely customer clusters.*

**Keywords:** *Clustering, Data Mining, K-Means, UMS Store*

## 1 PENDAHULUAN

UMS Store merupakan sebuah unit usaha dagang resmi yang dimiliki oleh Universitas Muhammadiyah Surakarta. Selain menyediakan berbagai macam kategori buku, tersedia pula jurnal, alat tulis kantor serta *official merchandise* berupa barang yang mencirikan UMS. UMS Store juga menjadi pusat layanan program penukaran *voucher* bagi mahasiswa aktif yang diselenggarakan Universitas Muhammadiyah Surakarta. *Voucher* tersebut diberikan kepada mahasiswa setiap semesternya dari semester satu hingga tujuh.

Teknologi yang berkembang pesat membuat kita hidup didunia yang penuh dengan data. Setiap hari, sistem dapat mengumpulkan data dalam jumlah besar dari transaksi pelanggan (Khandre & Alvi, 2016). Sejak awal berdirinya UMS Store telah melakukan banyak proses transaksi, baik penukaran *voucher* maupun pembelian tunai. Dengan kata lain UMS Store memiliki data yang amat berlimpah dan pasti akan terus bertambah seiring berjalannya waktu. Banyaknya data yang tersimpan tersebut membuat pengklasifikasian data menjadi lebih

sulit dan rumit, khususnya dalam hal kepentingan organisasi. Selain itu proses pengelompokan data tersebut juga tidak mungkin dilakukan, mengingat keterbatasan kemampuan yang dimiliki manusia pada bila menggunakan cara manual (Nugroho & Haryati, 2015).

Banyak sekali perusahaan yang tidak menyadari bahwa tumpukan data yang selama ini hanya disimpan sebenarnya sangat berharga. Bahkan banyak diantara data tersebut hanya dipandang sebagai arsip semata dan kemudian menjadi data tak terpakai yang akhirnya dibuang. Dengan penanganan yang tepat, apabila data tersebut diproses maka akan diperoleh informasi strategis yang berguna bagi masa depan perusahaan. Salah satu solusi yang dapat digunakan untuk mengatasi masalah ini ialah dengan menggunakan data *mining* dengan teknik klustering atau pengelompokan. Data *mining* dapat digunakan sebagai alternatif dalam pengolahan data menjadi sumber informasi strategis yang dapat pula sebagai pendukung pengambilan keputusan suatu perusahaan dimasa mendatang.

Data *mining* merupakan sekumpulan proses yang dilakukan dalam menemukan dan menggali nilai tambah dari sekumpulan data yang selama ini sulit diketahui secara manual. Kata *mining* dalam pengertiannya berarti usaha untuk mendapatkan sedikit pengetahuan berharga dari sejumlah besar material dasar. Dengan demikian data *mining* yang diterapkan memiliki tujuan yaitu untuk mengungkap dan menggali pola tersembunyi dari sejumlah data. Proses data *mining* sangat diperlukan, terutama untuk mengolah data dalam jumlah yang besar demi memberikan informasi strategis yang akurat (Bahar, Pramono, & Sagala, 2016). Data *mining* dilakukan untuk mengekstrak informasi berharga dari sebuah *dataset* dan kemudian memaparkannya dalam format yang mudah dimengerti oleh manusia dengan tujuan untuk mengambil sebuah keputusan (Ezenkwu, Ozuomba, & Kalu, 2015)

Dalam penelitian sebelumnya Sumadikarta, dkk (2016) mengembangkan sebuah aplikasi dengan konsep data *mining* menggunakan algoritma *k-means* yang berfungsi untuk mengelompokkan data pelanggan serta data produk pada PT Mega Arvia Utama. Hasil dari sistem ini adalah membentuk klaster yang digunakan sebagai bahan pertimbangan dalam menentukan strategi penjualan perusahaan yaitu, dengan menghilangkan produk yang menempati klaster terbawah dan memberikan reward kepada pelanggan potensial yang menempati klaster teratas. Bahar, dkk (2016) juga mengembangkan sistem penentuan strategi yang dibuat untuk menerapkan metode data *mining* dengan menggunakan algoritma *K-Means Clustering*, yaitu dengan cara melakukan pembagian data yang ada menjadi satu atau lebih *cluster*. Tujuan dari penelitian tersebut adalah untuk membagi data ke dalam beberapa kelompok, berfokus pada produk yang paling diminati.

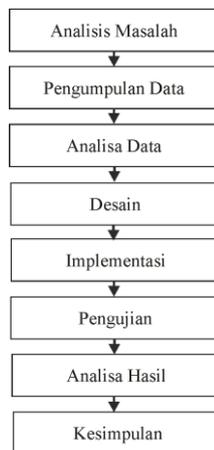
Tikmani, dkk (2015) melakukan penelitian untuk mengelompokkan pelanggan dan bagaimana algoritma *k-means* dapat digunakan untuk membagi pelanggan ke dalam kelompok yang berbeda. Penelitian ini menggambarkan mengenai segmentasi pelanggan dan menjelaskan algoritma *clustering* secara rinci. Secara

husus juga menggambarkan penggunaan teknik *k-means clustering* dan menjelaskan penggunaannya dalam segmentasi pelanggan perusahaan televisi berdasarkan berbagai variabel. Hasil simulasi menunjukkan pembagian delapan puluh pelanggan televisi ke dalam empat kelompok. Pada penelitian serupa, Lanjewar, dkk (2013) menguraikan teknik pengelompokan data untuk melakukan segmentasi pelanggan pada perusahaan produk susu. Proses klustering dapat membantu untuk mengidentifikasi kebutuhan dan perilaku pelanggan dan meningkatkan layanan demi kepuasan pelanggan. Dalam industri apapun langkah pertama untuk menemukan dan menghasilkan pelanggan yang menguntungkan ialah dengan pengelompokan. Hasil dari sistem tersebut memberikan profil dan pemahaman yang lebih baik terhadap pelanggan. Dalam penelitian serupa, Jain, dkk (2014) mengelompokkan konsumen berdasarkan perilaku belanja konsumen. Data yang dikumpulkan dari 1014 responden dilakukan pengelompokan dengan metode *k-means* dan menghasilkan empat buah kelompok pelanggan. Mengelompokkan pelanggan ke dalam segmen tertentu dapat membantu pemasaran dalam penargetan konsumen dengan lebih terfokus.

Berdasarkan latar belakang tersebut, maka penelitian ini akan mengembangkan sebuah aplikasi klustering dengan teknik data mining menggunakan algoritma *K-means*. Aplikasi ini dibuat untuk melakukan pengelompokan pelanggan pada *UMS Store* untuk menemukan kelompok pelanggan potensial, yang nantinya akan digunakan untuk memberikan rekomendasi berupa *treatment* dan penanganan lebih lanjut.

## 2 METODOLOGI

Bab ini menjelaskan berbagai tahapan yang dilakukan dalam melakukan penelitian, sehingga dapat sesuai dengan tujuan penelitian. Adapun alur tahapan dalam metodologi penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Tahapan metodologi penelitian

### 2.1 Analisis Masalah

Banyaknya transaksi pelanggan pada UMS Store menjadikan data yang dimiliki semakin berlimpah setiap harinya. Data yang berlimpah ini apabila dibiarkan akan menjadi tumpukan data yang tersimpan saja. Sebenarnya data tersebut apabila digali akan memberikan informasi yang berharga. Sehingga perlu dikembangkan sebuah Aplikasi yang dapat digunakan untuk melakukan pengelompokan data di UMS Store. Penelitian ini dilakukan dengan tujuan melakukan pengelompokan pelanggan potensial yang diambil dari data transaksi UMS Store. Obyek penelitian ini dipilih karena transaksi pelanggan didominasi oleh mahasiswa UMS. Hasil dari penelitian ini digunakan untuk memberikan rekomendasi pada UMS Store terhadap pelanggan potensial yang ada. Berdasarkan rekomendasi tersebut, UMS Store dapat memberikan penanganan lebih lanjut pada kelompok potensial yang terbentuk berupa perlakuan istimewa atau *treatment* belanja seperti memberikan diskon untuk pembelian selanjutnya. Hal ini nantinya akan menambah minat mahasiswa untuk berbelanja di UMS Store. Pengelompokan pada penelitian ini dibatasi pada data transaksi mahasiswa dengan pembelian menggunakan *voucher* UMS Store saja, sehingga proses pengklasteran dapat lebih spesifik.

### 2.2 Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi serta data yang dibutuhkan agar dapat mencapai tujuan akhir penelitian. Pengumpulan data memiliki

tujuan mendapatkan data akurat yang akan digunakan dalam penelitian guna menentukan hasil akhir yang baik (Kurniawan & Windisani, 2017). Data yang digunakan dalam penelitian ini diperoleh dari pihak UMS Store sesuai kebutuhan data penelitian ini.

### 2.3 Analisa Data

Tahap ini dilakukan untuk menentukan kebutuhan data dalam pengembangan aplikasi. Total keseluruhan data yang diperoleh dari UMS Store sebanyak 9031 data. Data yang akan digunakan dibatasi dengan hanya mengambil data pelanggan yang berstatus mahasiswa dan menggunakan voucher saja dibulan Mei 2017, sehingga diperoleh sampel data akhir sejumlah 357 untuk dilakukan pengelompokan. Pembersihan data juga dilakukan agar data sesuai dengan kebutuhan aplikasi, sehingga data nantinya tidak memiliki *noise* atau ketidakkonsistenan saat pengujian dilakukan.

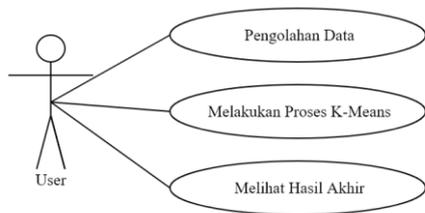
Variabel yang digunakan dalam proses pengelompokan ini dipilih berdasarkan tujuan penelitian seperti yang tertera pada tabel 1.

Tabel 1. Daftar variabel yang digunakan

Nama Variabel	Tipe Data	Deskripsi
NIM	Nominal	Kode mahasiswa
Tahun Angkatan	Numeric	Tahun angkatan mahasiswa
Discount	Numeric	Total diskon yang diperoleh
Subtotal	Numeric	Total harga keseluruhan
Total Paid	Numeric	Total harga yang harus dibayarkan setelah dipotong diskon
Total Item	Numeric	Jumlah barang yang dibeli
Tanggal	Date	Tanggal transaksi pembelian

### 2.4 Desain

Pada tahap desain, dilakukan perancangan dari aplikasi yang akan dibangun. Perancangan menggunakan metode *Unified Modeling Language* (UML) dengan diagram *usecase*. Diagram *usecase* menggambarkan interaksi yang dapat dilakukan oleh pengguna di dalam aplikasi (Lihat pada Gambar 2).



Gambar 2. Usecase diagram aplikasi

Pada Gambar 2, terdapat sebuah aktor yang berperan didalam aplikasi yaitu *user* atau pengguna aplikasi. Interaksi yang dapat dilakukan pengguna dalam aplikasi yang pertama adalah pengolahan data, sebagai contoh pengguna dapat menambah, mengubah, menghapus data, dan melakukan *upload* data dalam bentuk sebuah file csv. Selanjutnya pengguna dapat melakukan proses pengelompokan *K-means* dengan menjalankan fungsi proses. Selain itu pengguna juga dapat melihat hasil akhir dari proses pengelompokan. Aplikasi pengelompokan ini dibuat menggunakan database Mysql dan bahasa pemrograman PHP dan Mysql untuk melakukan perancangan antarmuka aplikasi.

**2.5 Implementasi**

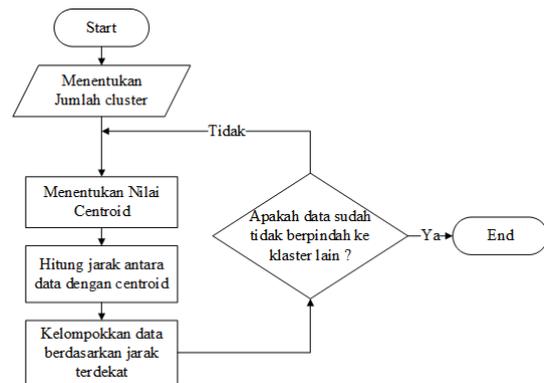
Salah satu algoritma yang dapat menyelesaikan permasalahan dalam melakukan pengklasteran adalah dengan menggunakan algoritma *K-means*. Huruf k pada *k-means* mewakili jumlah kelompok yang dipilih. Dalam algoritma *k-means* setiap kumpulan data diperlakukan sebagai objek yang memiliki beberapa lokasi. *K-means* menemukan partisi sehingga objek di dalam kelompok *cluster* sangat dekat satu sama lain, dan sangat jauh dan berebeda dengan kelompok lain. (Tikmani, Tiwari, & Khedkar, 2015). *K-means* menggunakan pendekatan iteratif untuk meminimalkan jumlah jarak antara setiap objek dengan *centroid*. *K-means* digunakan secara luas untuk melakukan analisis data dengan volume besar yang dihasilkan oleh berbagai sistem modern (Khandre & Alvi, 2016)

Berikut langkah pada algoritma *k-means*:

1. Menentukan nomor *cluster* (jumlah Kelompok)
2. Menentukan nilai *centroid* awal pada setiap klaster (dipilih secara acak)

3. Ulangi langkah berikut sampai anggota tidak berubah, dimana tidak ada data pada setiap *cluster* yang dapat berpindah ke *cluster* lain:
  - a. Tentukan nilai untuk *centroid* baru
  - b. Hitung jarak di antara setiap data dengan *centroid*
  - c. Kelompokkan data berdasarkan jarak terdekat dengan *centroid*

*Flowchart* dari algoritma *k-means* tersebut dapat dilihat melalui Gambar 3.



Gambar 3. *Flowchart* algoritma *k-means*

*K-means* dapat membagi data menjadi beberapa kelompok, sehingga data yang mempunyai kesamaan karakteristik akan dikelompokkan dalam satu *cluster*, sedangkan data yang berbeda karakteristiknya dikelompokkan pada *cluster* lain. Dalam membagi kelompok tersebut digunakan persamaan *Euclidean distance space* (persamaan 1) yang digunakan untuk menghitung jarak terpendek antara dua titik yang diperhitungkan yaitu data dengan *centroid* (Nugroho & Haryati, 2015).

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

(1)

Keterangan:

- $d_{ij}$  : *Euclidean distance*, Jarak antara objek i dan j
- $p$  : Dimensi data yang digunakan
- $x_{ik}$  : Koordinat data ke-i
- $x_{jk}$  : Koordinat *centroid* ke-j

Sedangkan untuk memperoleh titik *centroid* baru dilakukan dengan cara menghitung nilai rata-rata dari data yang ada pada *cluster* yang sama menggunakan persamaan 2.

$$C_k = \frac{1}{n_k} \sum d_i \quad (2)$$

Dimana:

$n_k$  : Jumlah data dalam *cluster*

$d_i$  : Jumlah nilai jarak yang masuk dalam setiap *cluster*

## 2.6 Pengujian

Setelah aplikasi pengelompokan pelanggan selesai dibuat, tahapan selanjutnya adalah pengujian. Pengujian ini memiliki tujuan untuk mengetahui apakah aplikasi sudah berjalan dengan baik sesuai dengan penelitian. Pengujian dengan metode *blackbox* dilakukan terhadap aplikasi pengelompokan untuk mengetahui apakah telah berfungsi dengan baik ataukah belum. Pengujian tersebut memiliki tujuan dalam menemukan fungsi yang gagal atau hilang, kesalahan struktur data, *interface* dan akses database (Prabiantissa, Ririd, & Asmara, 2017). Pengujian terhadap hasil pengelompokan juga dilakukan untuk membuktikan kebenaran hasil pengelompokan yang telah dilakukan sebelumnya (Handoko, 2016). Metode pengujian terhadap hasil pengelompokan dilakukan dengan cara membandingkan hasil yang pengelompokan aplikasi dengan hasil pada *software* RapidMiner.

## 2.7 Analisis Hasil

Pada penelitian ini, aplikasi pengelompokan pelanggan dengan menggunakan algoritma *K-means* setelah diterapkan pada data UMS *Store* akan menghasilkan tiga buah *cluster*. Berdasarkan hasil dari pengelompokan ini akan diketahui anggota kelompok manakah yang menjadi rekomendasi pelanggan potensial yang nantinya akan mendapatkan treatment lebih lanjut oleh UMS *Store*. Dilihat dari hubungan dengan pelanggan pengelompokan ini dapat menentukan pelanggan potensial yang dimiliki oleh UMS *Store*, sehingga apabila ada suatu *event* atau diskon dapat melibatkan pelanggan tersebut. Dilihat dari

segi sistem, aplikasi ini memudahkan UMS *Store* dalam menentukan pelanggan potensial tanpa memerlukan tenaga tambahan dalam melakukan analisa.

## 2.8 Kesimpulan

Pada tahap ini, penulis merumuskan suatu kesimpulan sesuai dengan penelitian dan hasil yang telah dilakukan dalam pembuatan tugas akhir ini.

## 3 HASIL DAN PEMBAHASAN

### 3.1 Implementasi Hasil

Aplikasi pengelompokan pelanggan pada UMS *Store* telah selesai diimplementasikan. Aplikasi ini terdiri atas beberapa halaman. Halaman utama merupakan tampilan awal dimana pengguna telah memasuki aplikasi, halaman ini berisi sedikit penjelasan mengenai aplikasi pengelompokan. Terdapat beberapa menu yang dapat diakses pengguna antara lain menu *Home*, *Project* dan *About*.

Gambar 4 menunjukkan Menu *Project*, menu ini merupakan menu yang berperan dalam melakukan proses utama pada pengelompokan. Pada menu ini pengguna dapat melihat tampilan data yang akan diproses serta melakukan pengolahan data seperti *edit* dan hapus data. Terdapat beberapa *button* untuk melakukan proses lain diantaranya *input* data, *upload* data dan proses *K-means*.

ID	NAME	Tahun Angkutan	Discount	Sub Total	Total Paid	Total Status	Tanggal	Aksi
1.	4020130128	2013	-100000	100000	0	2	2	[Edit] [Delete]
2.	4020130073	2013	-100000	100000	500	9	2	[Edit] [Delete]
3.	5030130273	2013	-100000	100000	0	5	2	[Edit] [Delete]
4.	5030130273	2013	-100000	100000	500	4	2	[Edit] [Delete]
5.	6000180200	2013	-100000	113000	13000	2	2	[Edit] [Delete]
6.	6000180200	2013	-100000	104000	4000	2	2	[Edit] [Delete]
7.	2014401008	2014	-100000	100000	0	3	2	[Edit] [Delete]
8.	4020140128	2014	-100000	100000	5000	3	2	[Edit] [Delete]
9.	2000140128	2014	-100000	100000	0	3	2	[Edit] [Delete]
10.	2000140200	2014	-100000	112000	12000	3	2	[Edit] [Delete]

Gambar 4. Tampilan menu *Project*

*Button Input* data yang berada pada bagian kiri memiliki fungsi untuk menambah data. Dan apabila dijalankan maka akan ditampilkan sebuah *form* yang digunakan untuk pengisian data yang akan ditambahkan, seperti pada Gambar 5. Data tersebut akan disimpan kedalam database.

The form titled 'Tambah Data' contains several input fields: 'NIM' (Masukkan NIM), 'Discount' (Masukkan Jumlah Diskon), 'Subtotal' (Masukkan Jumlah Subtotal), and 'Total Item' (Masukkan Tota Item). There are 'Reset' and 'Tambah' buttons at the bottom.

Gambar 5. Tampilan form pengisian data

Untuk mengupload data yang berasal dari file csv, pengguna dapat menjalankan *button upload* data yang apabila dijalankan nantinya akan menampilkan sebuah *form* untuk memilih file dari komputer pengguna seperti pada Gambar 6.

The form titled 'Import File CSV:' has a 'Choose Files' button and a file selection area showing 'No file chosen'. Below it is a green 'UPLOAD' button.

Gambar 6. Tampilan upload data CSV

Pada bagian kanan menu *Project* terdapat *button* Proses Data yang digunakan untuk melakukan inialisasi proses. Apabila *button* dijalankan maka akan tampil sebuah *form*, dimana pengguna diminta untuk memasukkan jumlah *cluster* dan jumlah maksimal iterasi yang akan diproses pada *form* tersebut. Gambar 7 menunjukkan tampilan hasil *centroid* yang diambil dari data secara acak. Kemudian *centroid* berikutnya diperoleh dari perhitungan nilai rata-rata dengan menggunakan persamaan yang telah dibahas pada bab 2.

The screenshot shows 'Centroid pada Proses' with 'Jumlah Iterasi: 20'. It displays 'Centroid Awal' and 'Centroid 2' with columns for 'Tahun\_angkatan', 'Discount', 'Sub Total', 'Total Paid', 'Total Item', and 'Tanggal'.

Gambar 7. Tampilan Centroid pada proses

Kemudian untuk dapat melihat hasil akhir pengelompokan tersebut, terdapat *button* Lihat Hasil Cluster. Pengguna akan diarahkan menuju halaman Hasil Cluster seperti yang ditunjukkan pada Gambar 8.

Halaman ini menampilkan tabel berisi keterangan anggota kelompok beserta hasil *cluster*.

The screenshot shows 'Hasil Cluster' with a table listing members across three clusters. Buttons for 'Anggota Cluster 1', 'Anggota Cluster 2', and 'Anggota Cluster 3' are visible.

Gambar 8. Tampilan hasil cluster

Berdasarkan Gambar 8, untuk dapat melihat anggota dari setiap *cluster* tersedia *button* yang akan menampilkan tabel sesuai dengan *cluster* yang dipilih. Pembahasan

### 3.2 Pembahasan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diperoleh proses perhitungan menggunakan algoritma *k-means*. Perhitungan ini sesuai dengan langkah-langkah pada algoritma *k-means* yang telah dijelaskan sebelumnya. Pertama ialah menentukan nilai *centroid* awal yang diambil dari data secara *random* atau acak. Daftar *centroid* awal yang telah diproses dapat dilihat pada tabel 2.

Tabel 2. Daftar centroid awal

K	Centroid				
	Tahun	Discount	Sub Total	Total Paid	Total Item
A	2015	-100000	113000	13000	2
B	2015	-100000	116000	16000	2
C	2015	-100000	107000	7000	2

Langkah selanjutnya adalah menghitung jarak terdekat antara setiap data dan *centroid*. Agar lebih mudah diambil contoh data pertama pada data percobaan yang dapat dilihat pada tabel 3.

Tabel 3. Data pertama pada data percobaan

Tahun Angkatan	Discount	Sub Total	Total Paid	Total Item	Tanggal
2013	-100000	100000	0	2	2

Dengan menggunakan persamaan *euclidean distance space*, diperoleh perhitungan jarak sebagai berikut:

$$d(c_1) = \sqrt{(2013 - 2014)^2 + (-100000 - -100000)^2 + (100000 - 100000)^2 + (0 - 0)^2 + (2 - 2)^2 + (2 - 2)^2} = 18384,77642$$

$$d(c_2) = \sqrt{(2013 - 2015)^2 + (-100000 - -100000)^2 + (100000 - 100000)^2 + (0 - 16000)^2 + (2 - 2)^2 + (2 - 2)^2} = 22627,41709$$

$$d(c_3) = \sqrt{(2013 - 2015)^2 + (-100000 - -100000)^2 + (100000 - 100000)^2 + (0 - 7000)^2 + (2 - 2)^2 + (2 - 2)^2} = 9899,495139$$

Jarak terdekat berdasarkan perhitungan diatas adalah jarak pada *centroid* ketiga, sehingga data tersebut menjadi anggota *cluster* 3. Demikian seterusnya hingga setiap data selesai dihitung. Selanjutnya tentukan *centroid*

baru dan lakukan langkah diatas sampai diperoleh hasil dengan jarak yang konstan dan anggota *cluster* tidak mengalami perubahan.

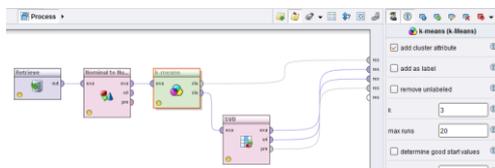
### 3.3 Pengujian

Pengujian menggunakan metode *blackbox* dilakukan untuk menemukan kesalahan dan ketidaksesuaian fungsi dalam aplikasi, dimana pengujian dilakukan terhadap setiap fungsi yang terdapat pada program. Fokus utama pengujian ini adalah pada spesifikasi fungsional dalam aplikasi. Tabel 4 memperlihatkan hasil pengujian *blackbox* yang menunjukkan bahwa aplikasi pengelompokan pelanggan telah berjalan dengan baik.

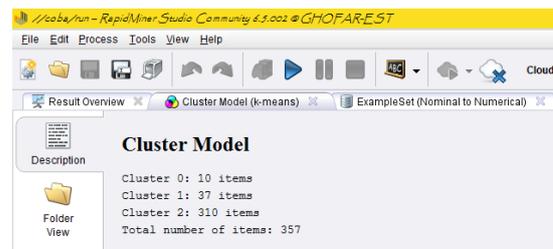
Tabel 4. Hasil pengujian *blackbox* aplikasi pengelompokan pelanggan

Fungsi	Skenario	Input	Output	Keterangan
Input Data	Menambahkan data baru kedalam database	NIM, Discount, Subtotal, Total Item	Data berhasil ditambahkan ke database	Valid
Upload File	Mengunggah dokumen dengan format CSV kedalam database	Dokumen CSV	Dokumen berhasil diunggah	Valid
Hapus Data	Menghapus yang berada dalam database	Klik Tombol Hapus	Data berhasil dihapus	Valid
Edit Data	Mengedit Data yang berada dalam database	Klik Tombol Edit	Data berhasil di-edit	Valid
Input Proses <i>K-means</i>	Memasukkan jumlah <i>cluster</i> dan jumlah iterasi untuk proses <i>clustering</i>	Jumlah <i>cluster</i> dan jumlah iterasi	Nilai tersimpan dalam variabel K dan <i>Maxruns</i>	Valid
Proses <i>K-means clustering</i>	Mengelompokkan data kedalam <i>cluster</i>	Klik tombol proses.	Data diproses, menampilkan <i>centroid</i> proses.	Valid
Lihat Hasil <i>Cluster</i>	Menampilkan hasil pengelompokan keseluruhan.	Menekan Tombol Lihat Hasil <i>Cluster</i>	Dta berhasil ditampilkan dalam tabel	Valid
Lihat Anggota <i>Cluster</i>	Menampilkan hasil untuk setiap <i>cluster</i> .	Menekan Tombol Anggota <i>Cluster</i>	Data berhasil ditampilkan dalam tabel	Valid

Pengujian kedua adalah untuk membuktikan keakuratan hasil pengelompokan pelanggan dalam aplikasi. Pengujian dilakukan dengan cara membandingkan anggota *cluster* pengelompokan dari aplikasi dengan hasil yang diperoleh dari software *RapidMiner*.



Gambar 9. Tampilan *RapidMiner* dengan proses *K-means*



Gambar 10. Tampilan *Cluster Model* pada *RapidMiner*

Gambar 9 dan Gambar 10 menunjukkan proses dan hasil dari pengelompokan menggunakan Software *Rapidminer*. Berdasarkan hasil pengelompokan pelanggan baik menggunakan aplikasi pengelompokan ini maupun

*software Rapidminer* pada akhirnya mendapatkan hasil yang sama.

### 3.4 Interpretasi Hasil

Pada penelitian ini, aplikasi pengelompokan pelanggan menggunakan algoritma *k-means* telah diimplementasikan pada data percobaan dan sesuai dengan yang diharapkan. Berdasarkan sampel sebanyak 357 data, setelah dilakukan pengelompokan dengan jumlah kelompok sebanyak tiga *cluster* diperoleh hasil sebagai berikut: anggota *cluster 1* dengan titik *centroid* akhir (2013.967; -100000; 101687.741; 1687.741; 3.351; 7.429) berjumlah 310 data, *cluster 2* dengan titik *centroid* (2014.270; -100000; 125908.108; 25908.108; 1.891; 8.432) berjumlah 37 data dan *cluster 3* dengan titik *centroid* akhir (2014.500; -100000; 162390; 62390; 1.200; 6.600) berjumlah 10 data. Sehingga dengan membandingkan nilai atribut dalam setiap kelompok dapat diambil kesimpulan bahwa kelompok pelanggan potensial berada dalam *cluster 3*. Kelompok inilah yang akan menjadi rekomendasi pada UMS Store.

## 4 PENUTUP

Berdasarkan hasil penelitian yang telah dilakukan dapat diambil beberapa kesimpulan diantaranya:

1. Sistem pengelompokan pelanggan dengan menggunakan algoritma *k-means* berhasil dibuat untuk mengelompokkan pelanggan potensial.
2. Berdasarkan pengujian *Blackbox* aplikasi telah berjalan dengan baik. Sedangkan pengujian hasil dengan menggunakan *software RapidMiner* juga mendapatkan hasil yang sama dengan aplikasi pengelompokan pelanggan.
3. Pengelompokan yang dihasilkan berdasarkan perhitungan jarak terdekat antara setiap data dengan *centroid*. Jumlah *cluster* dalam sistem dapat ditentukan sebanyak 3 *cluster*.

## 5. DAFTAR PUSTAKA

Bahar, A., Pramono, B., & Sagala, L. H. (2016). Penentuan Strategi Penjualan Alat-Alat Tattoo di

Studio Sonyxtattoo Menggunakan Metode K-Means Clustering. *Semantik*, 2(2), 75-86.

- Ezenkwu, C. P., Ozuomba, S., & Kalu, C. (2015). Application of K-Means Algorithm for Efficient Customer Segmentation: A Strategy for Targeted Customer Services. *IJARAI*, 4(10), 40-44. doi:10.14569/IJARAI.2015.041007
- Handoko, K. (2016). Penerapan Data Mining dalam Meningkatkan Mutu Pembelajaran pada Instansi Perguruan Tinggi Menggunakan Metode K-Means Clustering (Studi Kasus di Program Studi TKJ Akademi Komunitas Solok Selatan). *TEKNOSI*, 02(3), 31-40. doi:https://doi.org/10.25077/TEKNOSI.v2i3.2016.31-40
- Jain, N., & Ahuja, V. (2014). Segmenting Online Consumers Using K-means Cluster Analysis. *International Journal of Logistics Economics and Globalisation*, 6(2), 161-178. doi:https://doi.org/10.1504/IJLEG.2014.068274
- Khandre, A., & Alvi, A. S. (2016). Efficient Clustering Algorithm with Improved Clusters Quality. *IOSR-JCE*, 18(6), 15-19. doi:10.9790/0661-1806051519
- Kurniawan, Y. I., & Windisani, P. A. (2017). Sistem Pendukung Keputusan untuk Penentuan Kelulusan Beasiswa Sekolah Menengah Kejuruan (SMK) Menggunakan Metode Fuzzy. *Jurnal Teknik Elektro*, 9(1).
- Lanjewar, R., & Yadav, O. P. (2013). Understanding of Customer Profiling and Segmentation Using K-Means Clustering Method for Raipur Sakhari Dugdh Sangh Milk Products. *IJRCCT*, 2(3), 103-107.
- Nugroho, Y. S., & Haryati, S. N. (2015). Klasifikasi dan Klastering Penjurusan SMA Negeri 3 Boyolali. *Khazanah Informatika*, 1(1), 1-6. doi:https://doi.org/10.23917/khif.v1i1.1175
- Prabiantissa, C. N., Ririd, A. R., & Asmara, R. A. (2017). Sistem Identifikasi Batik Alami dan Batik Sintetis Berdasarkan Karakteristik Warna Citra dengan Metode K-Means Clustering. *Jurnal Informatika Polinema*, 5(2), 26-31.
- Sumadikarta, I., & Abeiza, E. (2016). Penerapan Algoritma K=Means pada Data Mining Untuk Memilih Produk dan Pelanggan Potensial (Studi Kasus : PT Mega Arvia Utama). *Jurnal Satya Informatika*, 1(1), 12-22.
- Tikmani, J., Tiwari, S., & Khedkar, S. (2015). An Approach To Consumer Classification Using K-Means. *IJIRCCE*, 3(11), 10542-10549. doi:10.15680/IJIRCCE.2015.0311029