



# PROSIDING SEMINAR NASIONAL SISFOTEK (Sistem Informasi dan Teknologi)

Padang, 4–5 September 2018

ISSN Media Elektronik 2597-3584

## Aplikasi Pengelompokan Penjualan Dengan Clustering Data Mining Pada Toko Retail Kota Padang

Eka Praja Wiyata Mandala<sup>a</sup>, Musli Yanto<sup>b</sup>, Dewi Eka Putri<sup>c</sup>

<sup>a</sup>Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, ekaprajawm@upiypk.ac.id

<sup>b</sup>Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, musli\_yanto@upiypk.ac.id

<sup>c</sup>Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, dewieka@upiypk.ac.id

### Abstract

*Retail is one or more activities that add value to the product to the consumer either for family needs or for personal use. Retail can sell products depending on current market needs. The goods we enjoy today are not apart from retail services, retail helps producers / distributors and consumers so that every need will be fulfilled. In this problem the author tries to do retail store research in the city of Padang. Retail store in question is minimarket. This study aims to assist retail stores in grouping the sale of goods. Data mining can be a solution in overcoming the problem. By using Clustering Data Mining, retailers will be able to classify the sales of goods in their retail stores. The results of the grouping gives an idea to the retail manager to be able to determine what items need to be held for the future. Applications are born this will not only be able to do the grouping process that occurs in this retail store alone, but it can also be penetrated into all aspects of retail both retail goods and retail services. This application can help to determine the procurement of goods in the sales process that will minimize the losses that occur in every sales activity.*

*Keywords: Clustering, Data Mining, Sales, Retail Store*

### Abstrak

Retail adalah satu atau lebih aktivitas yang menambah nilai produk kepada konsumen baik untuk kebutuhan keluarga atau untuk keperluan pribadi. Retail bisa menjual produk tergantung kebutuhan pasar saat ini. Barang yang kita nikmati saat ini tidak terlepas dari jasa retail, retail membantu produsen/distributor dan konsumen agar setiap kebutuhan akan keduanya dapat terpenuhi. Dalam permasalahan ini penulis mencoba melakukan penelitian toko retail yang ada di Kota Padang. Toko retail yang dimaksud adalah minimarket. Penelitian ini bertujuan untuk membantu toko retail dalam melakukan pengelompokan penjualan barang. Data mining dapat dijadikan solusi dalam mengatasi permasalahan yang tersebut. Dengan menggunakan Clustering Data Mining, pihak retail akan bisa mengelompokkan penjualan barang-barang di toko retail mereka. Hasil pengelompokkan memberikan gambaran kepada pengelola retail untuk bisa menentukan barang apa saja yang perlu diadakan untuk kedepannya. Aplikasi yang dilahirkan ini nantinya tidak hanya bisa melakukan proses pengelompokkan yang terjadi pada toko retail ini saja, akan tetapi bisa juga akan merambah ke semua aspek retail baik retail barang maupun retail jasa. Aplikasi ini dapat membantu untuk menentukan pengadaan barang dalam proses penjualan yang nantinya akan meminimalisir kerugian yang terjadi dalam setiap aktifitas penjualan.

Kata kunci : Clustering, Data Mining, Penjualan, Toko Retail

### 1. Pendahuluan

Perkembangan industri retail di Indonesia sangat pesat, namun tidak diiringi oleh informasi yang menunjang, sangat susah mencari informasi tentang perkembangan retail di internet. Retail (usaha perdagangan kecil) adalah bisnis yang akan tetap menjadi bisnis yang menarik sepanjang masa dengan tingkat pertumbuhan tak terbatas. Retail adalah kegiatan jual beli baik barang maupun jasa secara langsung kepada konsumen.

Konsumen yang membeli barang tersebut akan langsung menggunakannya (bukan untuk kepentingan bisnis).

Perusahaan retail di Indonesia ini sudah banyak dan menjamur dimana-mana. Wujud dari retail ini adalah minimarket yang setiap jengkal bisa kita temui. Minimarket merupakan salah satu bentuk retail yang menyediakan barang kebutuhan sehari-hari. Dengan banyaknya minimarket yang bermunculan, konsumen

dimudahkan dalam mencari barang kebutuhan sehari-hari. Namun, tidak begitu dengan retail itu sendiri, barang yang mereka sediakan banyak yang menumpuk karena tidak terjual. Ini lah yang menjadi masalah utama bagi para pengusaha retail. Akibatnya, barang-barang yang tersedia akan habis masa kadaluarsanya sebelum sempat terjual.

*Data mining* dapat memberikan solusi untuk masalah pengelompokan barang yang ada di toko retail tersebut. *Data mining* dengan menggunakan metode *Clustering* dapat mengelompokkan barang-barang menjadi beberapa kelompok sesuai yang diinginkan oleh pengusaha retail tersebut. Sehingga pengusaha retail dapat menentukan barang mana yang pengadaannya harus ditambah atau dikurangi.

Berdasarkan latar belakang diatas, maka penulis dapat merumuskan permasalahan yaitu bagaimana aplikasi pengelompokan penjualan ini dapat mengatasi masalah penumpukkan barang yang terjadi di toko-toko retail dengan cara mengelompokkan barang-barang yang terjual ?

Penelitian ini memiliki beberapa tujuan, diantaranya :

1. Membantu pengusaha retail dalam melakukan pengelompokkan barang-barang yang ada di toko retail mereka.
2. Membantu pengusaha retail dalam mengurangi penumpukkan barang yang tidak laku terjual, sehingga pengusaha retail tidak mengalami kerugian.

## 2. Tinjauan Pustaka

### 2.1 Knowledge Discovery and Database

*Knowledge Discovery in Database (KDD)* merupakan proses pencarian pengetahuan yang bermanfaat dari kumpulan data. Proses KDD bersifat interaktif dan iteratif, meliputi sejumlah langkah dengan melibatkan pengguna dalam membuat keputusan dan dapat dilakukan pengulangan diantara dua langkah [1].

Proses KDD terdiri dari beberapa langkah sebagai berikut [2] :

- a. *Data cleaning*, menghilangkan noise dan data yang tidak konsisten.
- b. *Data integration*, mengintegrasikan beberapa sumber data yang dapat digabungkan.
- c. *Data selection*, menyeleksi data yang relevan dengan tugas analisis akan diambil dari database.
- d. *Data transformation*, proses dimana data ditransformasikan atau dikonsolidasi dalam format yang sesuai untuk diproses dalam data mining.
- e. *Data mining*, sebuah proses esensial dimana metode diaplikasikan untuk mengekstrak pola data.

f. *Pattern evaluation*, proses untuk mengidentifikasi pola-pola yang menarik untuk direpresentasikan ke dalam knowledge based.

g. *Knowledge Presentation*, proses visualisasi dan teknik representasi pengetahuan yang digunakan untuk menyajikan pengetahuan yang berguna kepada pengguna.

### 2.2 Data Mining

Data Mining merupakan bagaimana menjelaskan masa lalu dan memprediksi masa depan dengan cara analisis data. Data mining adalah bidang disiplin ilmu yang menggabungkan statistik, pembelajaran mesin, kecerdasan buatan dan teknologi database.

Data mining adalah melakukan ekstraksi untuk mendapatkan informasi penting yang sifatnya implisit dan sebelumnya tidak diketahui, dari suatu data [3].

Data mining adalah analisis dari pengamatan set data (yang biasanya besar) untuk menemukan hubungan tak terduga dan untuk meringkas data dengan cara baru yang dapat dimengerti dan berguna bagi pemilik data [2].

Pengertian diatas diperkuat lagi dengan Jurnal Internasional yang mengatakan Data mining adalah metodologi analisis data yang digunakan untuk mengidentifikasi pola-pola yang tersembunyi dalam satu set data yang besar [4].

### 2.3 Clustering

Clustering menganalisis objek data tanpa mencari keterangan pada label kelas yang diketahui. Pada umumnya, label kelas tidak ditampilkan di dalam latihan data simply, karena mereka tidak tahu bagaimana memulainya. Clustering dapat digunakan untuk menghasilkan label-label.

Salah satu teknik yang di kenal dalam data mining yaitu clustering. Pengertian clustering adalah pengelompokkan sejumlah data atau objek kedalam cluster (group) sehingga setiap cluster akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster yang lainnya [5].

Clustering adalah membagi data ke dalam grup-grup yang mempunyai obyek yang karakteristiknya sama [10].

Clustering mengacu pada pengelompokkan data, observasi, atau kasus kedalam kelas objek yang sama. Cluster adalah kumpulan data yang mirip satu sama lain dan berbeda dengan data dari kelompok lain. Pengelompokkan dalam cluster tidak untuk mengklasifikasi, memperkirakan, atau memprediksi

nilai variable target. Sebaliknya, algoritma klusterisasi untuk semua data di tetapkan menjadi subkelompok yang relatif homogen, di man kesamaan data dalam cluster dimaksimalkan, dan kesamaan dengan data dari luar cluster diminimalkan [6].

### 2.4 Algoritma K-Means

K-Means merupakan metode penglompokkan data nonhierarki yang berusaha mempartisi data kedalam dua bentuk atau lebih kelompok. Metode ini mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok lain [7].

Tujuan dari pengelompokkan ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokkan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [8].

Proses clustering dimulai dengan mengidentifikasi data yang akan di cluster,  $x_{ij}$  ( $i=1, \dots, n; j=1, \dots, m$ ) dengan  $n$  adalah jumlah data yang akan di cluster dan  $m$  adalah jumlah variabel. Pada awal iterasi, pusat setiap cluster di tetapkan secara bebas (sembarang),  $c_{kj}$  ( $k=1, \dots, k; j=1, \dots, m$ ). Kemudian dihitung jarak antara setiap data dengan setiap pusat cluster.

Untuk melakukan perhitungan jarak data ke- $i$  ( $x_i$ ) pada pusat cluster ke- $k$  ( $c_k$ ) diberi nama  $d_{ik}$ , dapat digunakan formula Euclidean, yaitu :

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2} \quad (1)$$

## 3. Metodologi Penelitian

Kerangka penelitian yang digunakan dalam penelitian ini adalah sebagai berikut :



Gambar 1. Kerangka Penelitian

Adapun tahapan dari penelitian yang dilakukan adalah sebagai berikut :

### 3.1 Studi Pendahuluan

Pada tahapan ini peneliti melakukan analisa permasalahan yang akan dibahas pada pembangunan aplikasi pengelompokkan data penjualan.

### 3.2 Mempelajari Literatur

Untuk mencapai tujuan yang dicapai, diperlukan literatur-literatur yang berkaitan dengan permasalahan untuk membantu penyelesaian masalah yang diteliti. Sumber dapat berupa buku ataupun jurnal yang berhubungan dengan sistem yang akan dirancang.

### 3.3 Pengumpulan Data

Dalam pengumpulan data ini, digunakan beberapa metode yang mendukung antara lain :

1. Interview yaitu melakukan penelusuran untuk mendapatkan data dan informasi melalui tanya jawab dan wawancara dengan pihak orang yang berkompeten terhadap permasalahan yang diteliti.
2. Studi Pustaka yaitu pengumpulan data dan penelusuran informasi dengan cara membaca dan mempelajari buku-buku dan jurnal-jurnal penelitian yang berkaitan dan menunjang, baik dalam menganalisa data dan informasi maupun pemecahan masalah secara keseluruhan.

### 3.4 Analisa

Analisa yang dilakukan terdiri dari analisa data, analisa proses dan analisa sistem. Pada tahap analisa data dilakukan pengumpulan fakta-fakta yang mendukung perancangan sistem dengan mengadakan wawancara dengan pengusaha retail dan membandingkan hasil penelitian dengan buku penuntun yang ada. Pada tahap analisa proses, dilakukan dengan menggunakan metode clustering K-Means sebagai cara untuk mengelompokkan data penjualan. Dan pada tahap analisa sistem, akan dijelaskan seperti apa bentuk aplikasi pengelompokkan akan dibuat.

### 3.5 Perancangan

Pada tahap ini akan dilakukan dua tahap perancangan yaitu perancangan model dan perancangan interface. Pada tahap perancangan model, akan digunakan alat bantu perancangan yaitu UML (*Unified Modelling Language*) yang akan terdiri dari beberapa diagram diantaranya use case diagram, class diagram, sequence diagram, collaboration diagram, state chart diagram, activity diagram dan deployment diagram. Kemudian

akan dilakukan tahap selanjutnya yaitu perancangan tampilan user interface program.

### 3.6 Implementasi

Implementasi merupakan proses merubah perancangan yang telah dibuat menjadi program yang dikembangkan dengan bahasa pemrograman yang akan digunakan. Bahasa pemrograman yang digunakan adalah PHP dan didukung dengan database MySQL.

### 3.7 Pengujian

Pada tahap pengujian ini akan dilakukan beberapa pengujian diantaranya pengujian secara multi-tier, pengujian aplikasi dan pengujian interface. Pengujian multi-tier merupakan pengujian secara online agar bisa diakses oleh seluruh pengusaha retail di Kota Padang. Pengujian aplikasi merupakan program yang telah dibangun akan diuji untuk mengetahui apakah program tersebut sudah berjalan dengan benar dan sesuai dengan perancangan yang dilakukan. Dan pengujian interface merupakan pengujian untuk melihat interface aplikasi yang dihasilkan sesuai atau tidak dengan interface yang dirancang.

#### 3.2.8 Kesimpulan

Membuat laporan atau hasil dari analisa dan perancangan kedalam format penulisan penelitian yang disertai dengan kesimpulan akhir.

## 4. Hasil dan Pembahasan

### 4.1. Analisa

Pada tahap analisa dilakukan analisa terhadap data yang diperoleh dari salah satu toko retail. Untuk proses analisa data ini, digunakan hanya beberapa data sampel untuk memastikan aplikasi data mining yang dibuat sesuai dengan proses perhitungan dengan metode Clustering Algoritma K-Means.

Berikut sampel data yang digunakan terdiri dari 7 item barang yaitu makanan ringan dan minuman.

Tabel 1. Tabel data sampel

Item (m)	Stok (x)	Terjual (y)
FRESTEA Lemon	45	25
MIZONE Passion Fruit	39	21
Biskuit GOOD TIME Cheese	65	10
Minute Maid PULPY ORANGE	62	15
AQUA 600 ml	32	26
GERY Saluut Chocolate	37	23
Wafer NISSIN Chocolate	56	10

Dari tabel diatas akan dilakukan pengelompokkan barang yang terjual menjadi 2 buah cluster (kelompok) yaitu barang yang **Laris** dan barang yang **Kurang Laris**.

Untuk cluster yang Laris (cluster 1), akan dibangkitkan centroid dengan stok 30 dan terjual 26. Sedangkan, untuk cluster yang Kurang Laris (cluster 2), akan dibangkitkan centroid dengan stok 55 dan terjual 13. Jadi, nilai  $C1 = \{30, 26\}$  dan  $C2 = \{55, 13\}$ .

Kemudian hitung jarak antara masing-masing item ke masing-masing centroid dengan menggunakan rumus korelasi antar dua objek yaitu Euclidean Distance sehingga diperoleh jarak antara masing-masing item (M) ke C1 dan ke C2 adalah sebagai berikut :

Tabel 2. Tabel Jarak Item (M) ke C1 dan C2

	M1	M2	M3	M4	M5	M6	M7
C1	15,03	10,30	38,48	33,84	2,00	7,62	30,53
C2	15,62	17,89	10,44	7,28	26,42	20,59	3,16

Langkah selanjutnya adalah bandingkan jarak dari masing-masing item ke C1 dan C2 dan kemudian kelompokkan, misalnya M1 lebih dekat ke C1 daripada ke C2, sehingga M1 menjadi anggota C1.

Tabel 3. Tabel Kelompok Data Masing-Masing Cluster

Anggota	
C1	M1, M2, M5, M6
C2	M3, M4, M7

Kemudian tentukan nilai centroid baru untuk masing-masing cluster dengan menghitung jumlah masing-masing stok barang dan barang terjual per masing-masing cluster.

$$C1_{baru} = \frac{M1x + M2x + M5x + M6x}{4}, \frac{M1y + M2y + M5y + M6y}{4}$$

$$= \frac{45 + 39 + 32 + 37}{4}, \frac{25 + 21 + 26 + 23}{4}$$

$$C1_{baru} = \mathbf{38,25 ; 23,75}$$

$$C2_{baru} = \frac{M3x + M4x + M7x}{3}, \frac{M3y + M4y + M7y}{3}$$

$$= \frac{65 + 62 + 56}{3}, \frac{10 + 15 + 10}{3}$$

$$C2_{baru} = \mathbf{61,00 ; 11,67}$$

Dikarenakan nilai C1baru dan C2baru tidak sama dengan nilai C1 dan C2, maka dilakukan kembali perhitungan jarak dari masing-masing item M ke C1baru dan C2baru sehingga diperoleh jaraknya seperti tabel berikut :

Tabel 4. Tabel Jarak Item (M) ke C1baru dan C2baru

	M1	M2	M3	M4	M5	M6	M7
C1	6,86	2,85	30,08	25,31	6,64	1,46	22,45
C2	20,83	23,90	4,33	3,48	32,35	26,54	5,27

Langkah selanjutnya adalah bandingkan jarak dari masing-masing item ke C1baru dan C2baru dan kemudian kelompokkan, misalnya M1 lebih dekat ke C1baru daripada ke C2baru, sehingga M1 menjadi anggota C1baru.

Tabel 5. Tabel Kelompok Data Masing-Masing Cluster Anggota

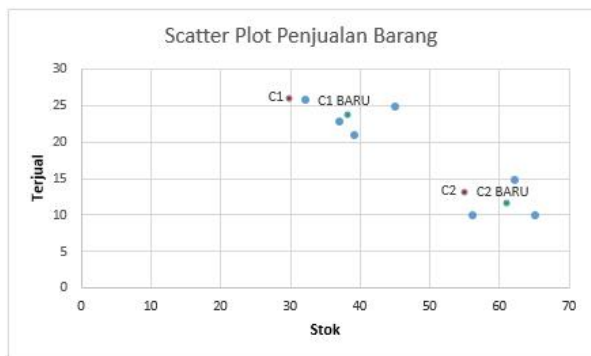
Anggota	
C1baru	M1, M2, M5, M6
C2baru	M3, M4, M7

Dari tabel diatas, dapat dilihat bahwa tidak terjadi perubahan anggota untuk masing-masing cluster. Jika tidak terjadi perubahan anggota, maka proses dihentikan. Sehingga diperoleh pengelompokan penjualan seperti berikut :

Tabel 6. Tabel Kelompok Data Penjualan Anggota

Anggota	
C1baru	M1 FRESTEA Lemon, M2 MIZONE Passion Fruit, M5 AQUA 600 ml, M6 GERY Saluut Chocolate
C2baru	M3 Biskuit GOOD TIME Cheese, M4 Minute Maid PULPY ORANGE, M7 Wafer NISSIN Chocolate

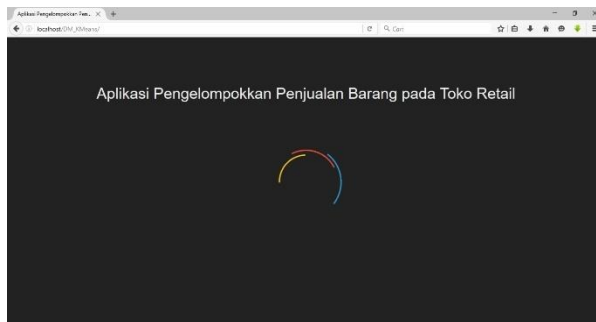
Dari tabel diatas, dapat dibuatkan grafik scatter plot untuk menggambarkan kelompok data penjualan tersebut diatas.



Gambar 2. Scatter Plot Cluster Penjualan Barang

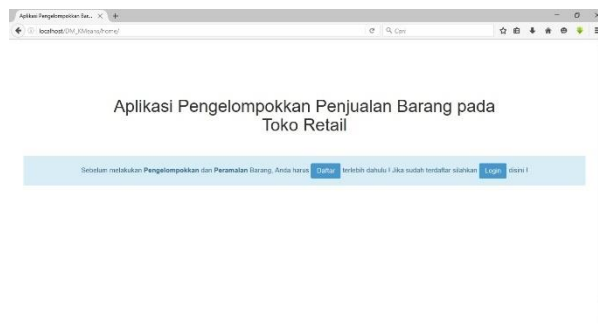
#### 4.2 Implementasi

Pada tahap implementasi, analisa yang sudah dilakukan sebelumnya, akan diimplementasikan ke dalam bentuk aplikasi data mining untuk melakukan pengelompokan penjualan barang untuk masing-masing toko retail yang ada di kota Padang.



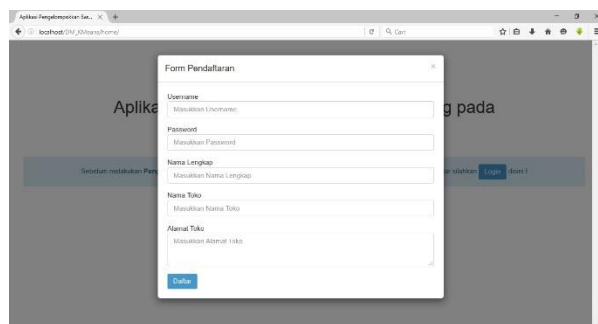
Gambar 3. Tampilan Awal Aplikasi Data Mining

Sebelum melakukan proses pengelompokan barang, maka pengguna akan diarahkan ke halaman ini



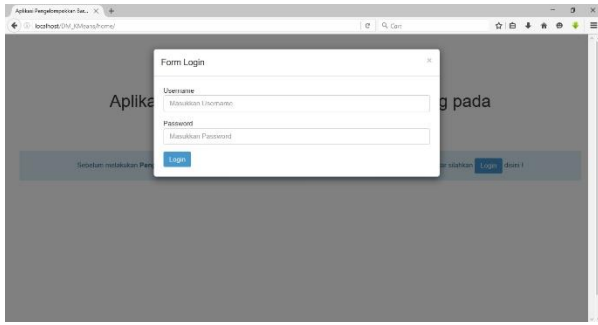
Gambar 4. Halaman Utama Aplikasi Data Mining

Pihak retail diharuskan melakukan pendaftaran terlebih dahulu.

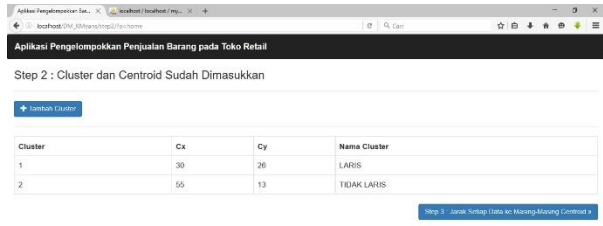


Gambar 5. Halaman Pendaftaran Aplikasi Data Mining

Setelah melakukan pendaftaran, maka pihak retail bisa melakukan login.



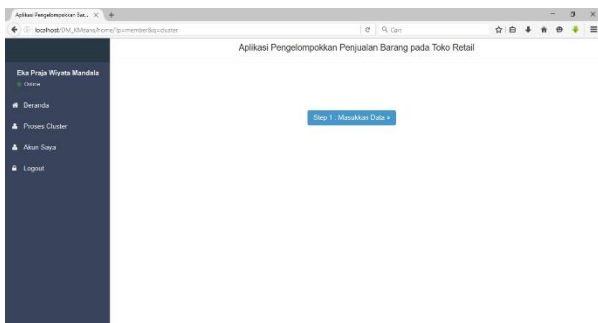
Gambar 6. Halaman Login Aplikasi Data Mining



Gambar 9. Halaman Step 2 : Tentukan Cluster dan Centroidnya

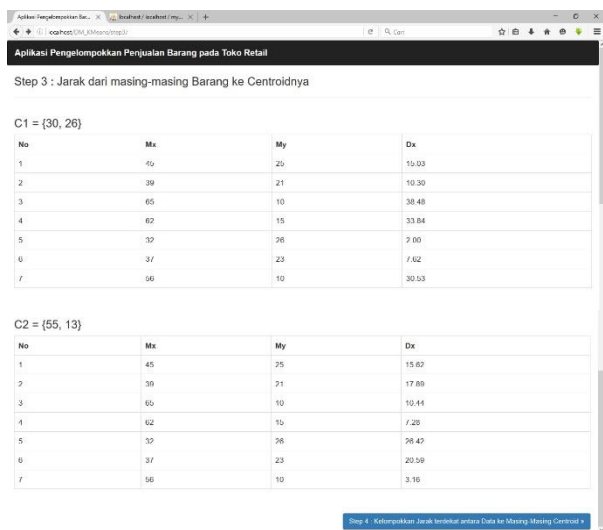
Setelah berhasil login, maka pihak retail sudah bisa melakukan pengelompokan barang melalui halaman ini.

Klik tombol Step 3, maka akan dihitung jarak masing-masing item barang ke masing-masing centroid yang sudah ditentukan.

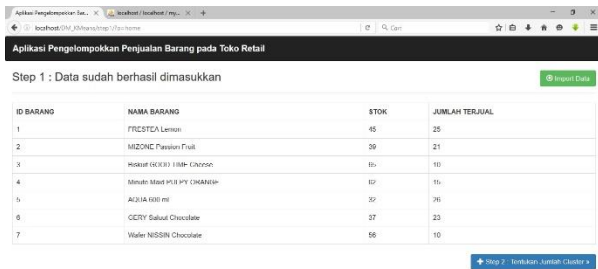


Gambar 7. Halaman Proses Cluster Aplikasi Data Mining

Klik tombol Step 1, maka langkah pertama yang dilakukan adalah proses import data dari file excel sesuai dengan format yang sudah ditentukan. Dalam hal ini, data yang akan digunakan adalah data yang sama dengan data sampel pada bab sebelumnya.



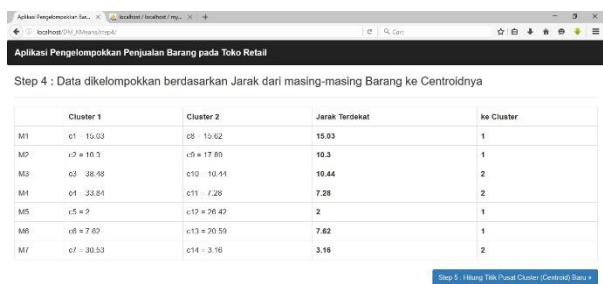
Gambar 10. Halaman Step 3 : Jarak dari Item ke Centroid



Gambar 8. Halaman Step 1 : Import Data

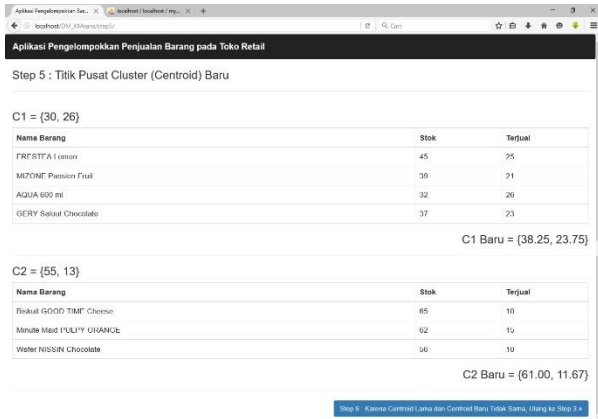
Klik tombol Step 2, maka langkah yang harus dilakukan adalah menentukan berapa jumlah cluster yang akan dibuat dan tentukan nilai centroid untuk masing-masing clusternya.

Klik tombol Step 4, maka akan dikelompokkan barang-barang tersebut per masing-masing cluster.



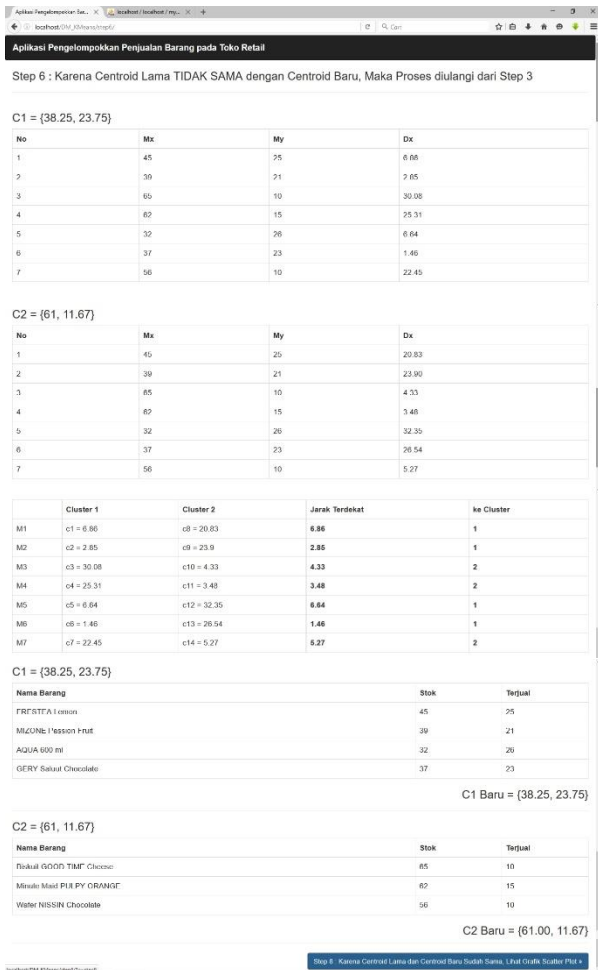
Gambar 11. Halaman Step 4 : Kelompok Barang per Cluster

Klik tombol Step 5, akan muncul nilai centroid baru untuk masing-masing cluster.



Gambar 12. Halaman Step 5 : Nilai Centroid Baru

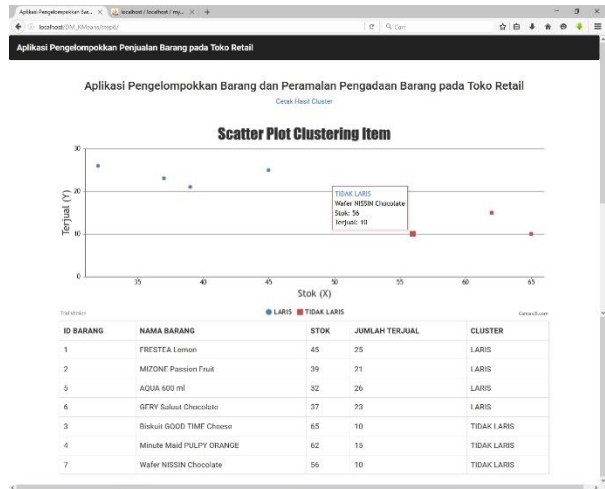
Klik tombol Step 6, proses perhitungan akan diulang kembali karena nilai centroid baru tidak sama dengan nilai centroid lama, sehingga harus dilakukan kembali proses perhitungan jarak item ke masing-masing cluster.



Gambar 13. Halaman Step 6 : Perhitungan Jarak dengan Centroid Baru

Tombol Step 7 tidak muncul karena tidak ada perubahan anggota kelompok, sehingga nilai centroid sudah bisa dipastikan sama dengan sebelumnya.

Klik tombol Step 8, maka akan muncul scatter plot dari pengelompokan barang yang dilakukan sebelumnya.



Gambar 14. Halaman Step 8 : Scatter Plot dan Kelompok Barang

Dari gambar diatas, dapat dilihat hasil pengelompokan dari barang-barang yang terjual di toko retail tersebut.

## 5. Kesimpulan

### 5.1 Simpulan

Dari penelitian yang sudah dilakukan diatas, maka diperoleh beberapa kesimpulan, diantaranya :

1. Aplikasi data mining yang dihasilkan dari penelitian ini, sudah dapat membantu toko retail dalam mengelompokkan penjualan barang di masing-masing toko retail mereka.
2. Dengan adanya aplikasi data mining ini, sudah bisa mengurangi penumpukan barang di masing-masing toko retail karena data penjualan barang sudah dikelompokkan dengan baik.

### 5.2 Saran

Dari kesimpulan diatas, dapat diusulkan beberapa saran, diantaranya :

1. Diharapkan adanya penelitian lanjutan untuk mengembangkan aplikasi data mining ini, dengan menambahkan proses peramalan penjualan barang setelah dilakukan pengelompokan.
2. Diharapkan aplikasi ini bisa mencakup seluruh toko retail di Indonesia, sehingga dapat membantu semua toko retail dalam pengadaan barang.

## Ucapan Terima Kasih

Penelitian ini merupakan Penelitian Dosen Pemula tahun Anggaran 2018. Peneliti mengucapkan terima kasih kepada Kementerian Riset Teknologi dan Pendidikan Tinggi yang sudah mendanai penelitian ini. Peneliti juga mengucapkan terima kasih pada Lembaga

Penelitian dan Pengabdian Masyarakat Universitas Putra Indonesia “YPTK” Padang, yang sudah memfasilitasi terlaksananya penelitian ini melalui kontrak penelitian dengan Nomor kontrak : 05/UPI-YPTK/LPPM/KP/PDP/II/2018.

### Daftar Rujukan

- [1] Ginting, S.L.B., 2010., Konstruksi Struktur Bayesian Network dalam Data Mining untuk Basis Data Incomplete Menggunakan Algoritma CB\*, Jurnal Spektrum., pp 111-127
- [2] Han, J and Kamber, M., 2011., Data Mining Concepts and Techniques Third Edition, Elsevier
- [3] Witten, I.H., Eibe, F, and Hall, M.A. 2011., Data mining Practical Machine Learning Tools and Techniques 3rd Edition, Elsevier
- [4] Durairaj, M., and Vijitha, C., 2014., Educational data mining for prediction of student performance using clustering algorithms. International Journal of Computer Science and Information Technologies
- [5] Alfina, T., Santosa, B., and Barakbah, A.R., 2012., Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Cluster Data (Studi Kasus: Problem Kerja Praktek Teknik Industri ITS)., Jurnal Teknik ITS
- [6] Larose, D.T., 2005., Discovering Knowledge in Data an Introduction to Data Mining, John Wiley & Sons
- [7] Putra, R.R, and Wadisman, C., 2018., Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. INTECOMS: Journal of Information Technology and Computer Science pp: 72-77
- [8] Putri, D.E., 2015., Metode Non Hierarchy Algoritma K-Means Dalam Mengelompokkan Tingkat Kelarisan Barang (Studi Kasus: Koperasi Keluarga Besar Semen Padang)., Prosiding Senatkom
- [9] Mas`udia, P.E, Arinie, F, and Mustafa, L.D., 2018., Clustering Data Remunerasi Dosen Untuk Penilaian Kinerja Menggunakan Fuzzy c-Means., Jurnal RESTI
- [10] Hakim, L and Seruni, H., 2018., Indikasi Penyimpangan Laporan Keuangan Akademik Universitas XYZ Menggunakan Algoritma Greedy dan K-Means., Jurnal RESTI
- [11] Mandala, E.P.W., 2015., Web Programming Project : e.p.w.m Forum., Yogyakarta., Andi