

IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING PADA APLIKASI PENCARI PELANGGAN POTENSIAL PADA RESTORAN XYZ

Sabar Rudiarto

*Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana
Jl. Raya Meruya Selatan, Kembangan, Jakarta 11650
E-mail : sabar.rudiarto@mercubuana.ac.id*

ABSTRACT

Persaingan dalam usaha, membuat Restoran XYZ mengalami kesulitan dalam mempertahankan pelanggan setianya. Untuk itu diperlukan analisa untuk mengetahui mana pelanggan yang potensial, maka dirancanglah sebuah aplikasi menggunakan algoritma K-Means (Clustering) untuk mengelompokkan data pelanggan. Metode K-Means Clustering mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama. Data pelanggan dikelompokkan menjadi dua cluster yaitu tidak potensial, dan potensial. Kemudian setiap cluster diklasifikasikan berdasarkan kriteria mana yang lebih diprioritaskan. Cluster dengan nilai terbesar pada centroid akhir merupakan cluster yang direkomendasikan menerima reward, sedangkan cluster dengan nilai terkecil pada centroid akhir merupakan cluster yang tidak berhak menerima reward. Hasil dari proses ini membentuk cluster-cluster yang digunakan untuk memberi saran pertimbangan dalam menentukan strategi penjualan yakni memberikan reward untuk pelanggan dengan posisi cluster teratas.

Kata Kunci: K-Means, Clustering, Layanan Pelanggan, Sistem Pengambil Keputusan

PENDAHULUAN

Restoran XYZ merupakan salah satu rumah makan yang menyediakan makanan daerah dan berbagai macam menu makanan dan menyediakan pemesanan delivery order melalui pemesanan by phone yang berada di kawasan Jakarta. Pelanggan merupakan salah satu posisi penting dalam pengembangan strategi bisnis di restoran XYZ, pelanggan juga merupakan salah satu sumber keuntungan dalam perusahaan. Untuk itu diperlukan suatu pemahaman yang baik tentang pelanggan. Pemahaman yang baik terhadap pelanggan dapat digunakan perusahaan untuk mempertahankan pelanggan yang memang harus dipertahankan. Masalah yang sering dihadapi adalah kesulitan dalam menentukan pelanggan yang

tepat secara cepat, hal tersebut dapat mengakibatkan perusahaan kehilangan nasabah yang seharusnya dipertahankan.

Segmentasi pelanggan adalah metode yang populer yang digunakan untuk memilih pelanggan atau nasabah yang tepat untuk memulai promosi. Dengan segmentasi pelanggan berdasarkan prilakunya, kita dapat menargetkan tindakan mereka dengan lebih baik. Seperti memberi loyalitas yang bisa memenuhi harapan pelanggan sehingga pelanggan akan tetap bertahan karena puas terhadap layanan yang diberikan. Proses penentuan pelanggan secara manual dengan menginputkan satu persatu data pelanggan kemudian melakukan sorting data pelanggan seringkali menimbulkan beberapa permasalahan, antara lain membutuhkan waktu yang

lama dan ketelitian yang tinggi. Selain itu proses komputasi penentuan pelanggan juga menjadi salah satu permasalahan. Sehingga dibutuhkan suatu sistem yang dapat membantu dalam proses pengambilan keputusan siapa saja pelanggan yang direkomendasikan menerima penghargaan berdasarkan kriteria-kriteria yang telah ditentukan.

Maka perlu diterapkan sebuah algoritma, penerapan algoritma clustering untuk penentuan pelanggan potensial berdasarkan data - data yang telah ada. Ada dua jenis data clustering yang sering digunakan dalam proses pengelompokan data yaitu hierarchial (hirarki) dan non-hierarchial (non hirarki) data clustering. Dalam penelitian sistem pendukung keputusan penentuan pelanggan potensial ini penulis lebih fokus pada pengelompokan data non hirarki clustering Algoritma Clustering pada non hirarki dengan pendekatan partisi tersebut ada beberapa macam, diantaranya Fuzzy C-Means ,K-Means dan lain sebagainya. Penulis menggunakan algoritma clustering yaitu algoritma K-Means Clustering. Algoritma K-Means adalah salah satu metode data clustering non hirarki yang mempartisi data ke dalam bentuk satu atau lebih cluster/kelompok, sehingga data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan data yang memiliki karakteristik berbeda dikelompokkan ke dalam kelompok lain (Agusta, 2007). Algoritma ini akan membantu pengguna dalam penentuan pelanggan potensial menjadi lebih cepat. Algoritma K-Means Clustering ini selain dapat memangkas waktu penentuan pelanggan potensial juga dapat pula melakukan penentuan secara tepat dan akurat.

STUDI LITERATUR

PELANGGAN

Definisi dari pelanggan adalah adalah semua orang yang menuntut kita (atau perusahaan) kita untuk memenuhi suatu standar kualitas tertentu dan oleh karena itu akan memberikan pengaruh pada kinerja (performance) kita (atau perusahaan kita) .Namun pengertian dari pelanggan potensial adalah pelanggan yang loyal terhadap perusahaan (Zulkarnain, 2012). Pelanggan yang potensial umumnya akan melanjutkan pembelian produk atau jasa tersebut walaupun dihadapkan pada banyak

alternatif produk atau jasa yang lebih unggul dipandang dari berbagai sudut atributnya. Menurut Kotler dalam bukunya Zulkarnain, konsumen yang loyal tidak diukur dari berapa banyak dia membeli, tapi dari berapa sering dia melakukan pembelian ulang, termasuk merekomendasikan orang untuk membeli (Zulkarnain, 2012).

LAYANAN PELANGGAN

Menurut penelitian layanan pelanggan dapat dianggap sebagai investasi usaha atau bisnis. Pelanggan bagi perusahaan adalah aset, karena itu peningkatan kualitas pelayanan diupayakan terus-menerus untuk memenuhi kebutuhan dan harapan pelanggan/kolega, “bila suatu produk atau jasa tertentu diciptakan tanpa memperhatikan perencanaan pelayanan bagi pembeli. Perencanaan pelayanan yang baik bertujuan memberikan pelayanan kepada pelanggan/kolega sehingga tidak akan terjadi masalah/konflik/komplain. Berdasarkan hal tersebut maka Pengertian Pelayanan adalah suatu pelayanan terbaik dalam memenuhi harapan dan kebutuhan pelanggan. Artinya pelayanan yang memenuhi standar kualitas. Pelayanan yang memenuhi standar kualitas adalah suatu pelayanan yang sesuai dengan harapan dan kepuasan pelanggan/masyarakat.

CLUSTERING

Clustering adalah mengelompokkan item data ke dalam sejumlah kecil grup sedemikian sehingga masing-masing grup mempunyai sesuatu persamaan yang esensial (Andayani, 2007). Ada beberapa pendekatan yang digunakan dalam mengembangkan metode clustering. Dua pendekatan utama adalah clustering dengan pendekatan partisi dan clustering dengan pendekatan hirarki. Clustering dengan pendekatan partisi atau sering disebut dengan partition-based clustering mengelompokkan data dengan memilah-milah data yang dianalisa ke dalam cluster-cluster yang ada. Clustering dengan pendekatan hirarki atau sering disebut dengan hierarchical clustering mengelompokkan data dengan membuat suatu hirarki berupa kurva yang menggambarkan pengelompokan cluster dimana data yang mirip akan ditempatkan pada hirarki yang berdekatan dan yang tidak pada hirarki yang berjauhan. Algoritma clustering dibagi ke dalam beberapa kelompok besar, antara lain:

1. Partitioning algorithms: algoritma dalam kelompok ini membentuk bermacam partisi dan kemudian mengevaluasinya dengan berdasarkan

beberapa kriteria. Diantaranya : K-Means Clustering, Mixture-Based Density, Graph Theory-Based Clustering, Fuzzy Clustering.

2. Hierarchy algorithms: pembentukan dekomposisi hirarki dari sekumpulan data menggunakan beberapa kriteria.

K-MEANS

Metode K-Means pertama kali diperkenalkan oleh MacQueen JB pada tahun 1976. Metode ini adalah salah satu metode non hierarki yang umum digunakan. Metode ini termasuk dalam teknik penyekatan (partition) yang membagi atau memisahkan objek ke k daerah bagian yang terpisah. Pada K-Means, setiap objek harus masuk dalam kelompok tertentu, tetapi dalam satu tahapan proses tertentu, objek yang sudah masuk dalam satu kelompok, pada satu tahapan berikutnya objek akan berpindah ke kelompok lain. Hasil cluster dengan metode K-Means sangat bergantung pada nilai pusat kelompok awal yang diberikan. Pemberian nilai awal yang berbeda bisa menghasilkan kelompok yang berbeda. Ada beberapa cara memberi nilai awal misalnya dengan mengambil sampel awal dari objek, lalu mencari nilai pusatnya, memberi nilai awal secara random, menentukan nilai awalnya atau menggunakan hasil dari kelompok hierarki dengan jumlah kelompok yang sesuai. K-Means adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Dengan kata lain, metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu cluster dan memaksimalkan variasi dengan data yang ada di cluster lainnya.

Penelitian mengenai metode K-Means telah dilakukan oleh banyak peneliti dengan berbagai macam kasus, diantaranya adalah:

- (1) Pada penelitian (Abdul Haris Lubis, 2016), algoritma Kernel K-Means Clustering ini digunakan untuk mendapatkan suatu model aturan menentukan pelanggan terbaik dan potensial Kantor Pos Medan. Model aturan yang diperoleh menunjukkan bahwa kategori pelanggan terbaik dapat diperoleh jika

transaksinya banyak dan besar uangnya sedang dan tinggi.

- (2) Penelitian berikutnya (Gita Febrina Wulandari, 2014), metode yang digunakan adalah CRISP-DM dengan melalui tahapan business understanding, data understanding, data preparation, modeling, evaluasi dan development. Algoritma yang digunakan untuk pembentukan kluster adalah algoritma K-Means, dengan tools Rapidminer 5.3 yang merepresentasikan jumlah pelanggan di setiap kluster dengan menggunakan atribut RFM (Recency, Frequency dan Monetary). Kelompok pelanggan yang terbentuk adalah empat kelas dengan kluster pertama 4 pelanggan, kluster kedua 190 pelanggan, kluster ketiga 23 pelanggan dan kluster keempat dengan jumlah pelanggan 115.

- (3) Penelitian yang dilakukan oleh (Teguh Iman Hermanto), metode yang digunakan pada penelitian ini adalah pengolahan data transaksi pembelian menggunakan algoritma Association Rule dan pengolahan data pelanggan menggunakan K-Means clustering. Pengolahan data transaksi dilakukan dengan menentukan nilai support dan confidence sebagai nilai kepastian dari pola data pembelian yang dihasilkan. Pengolahan data pelanggan dilakukan dengan menentukan nilai Recency, Frequency, dan Monetary sebagai variabel untuk menentukan segmentasi pelanggan. Hasil dari penelitian ini adalah pengolahan data menggunakan algoritma Association Rule dengan nilai optimal confidence dan support sebesar 8% dan pengolahan data pelanggan menggunakan algoritma K-Means dengan nilai below zeros sebanyak 2 pelanggan, nilai migrator sebanyak 6 pelanggan, nilai most growable sebanyak 97 pelanggan, dan nilai most valuable sebanyak 31 pelanggan.

ANALISA SISTEM

ANALISA SISTEM

Pada perancangan suatu sistem diperlakukan analisa yang tepat, sehingga proses pembuatan sistem dapat berjalan dengan lancar dan sesuai seperti yang diinginkan. Setelah dilakukan analisis perlu adanya desain sistem (design interface) yang bertujuan untuk memberikan gambaran tentang jalannya sistem tersebut.

ANALISA MASALAH

Penentuan pelanggan potensial untuk mengetahui mana pelanggan yang potensial, mana yang tidak berdasarkan transaksi yang ada saat ini masih secara manual. Analisa sistem berjalan ini dilakukan bertujuan untuk mengetahui permasalahan yang ditimbulkan oleh sistem yang berjalan pada saat ini. Permasalahan tersebut dapat digunakan untuk menganalisis kebutuhan pengguna terhadap sistem yang baru. Kemudian, setelah permasalahan dan kebutuhan sistem dianalisis dengan baik, maka hasil analisis tersebut dapat digunakan sebagai dasar untuk perancangan sistem yang baru, sehingga dapat dihasilkan sistem baru yang dapat memenuhi kebutuhan pengguna.

ALGORITMA K-MEANS CLUSTERING

Pada analisis tentang clustering untuk penentuan pelanggan potensial sebelumnya, ketika proses pengelompokan data membutuhkan algoritma clustering untuk pengelompokan data sesuai karakter-karakter yang diinput dari user dengan database informasi data transaksi. Untuk itu berikut merupakan analisa terhadap pencarian Clustering dengan menggunakan Algoritma K-Means Clustering dalam pengelompokan data. Pada artikel yang diterbitkan Elsevier "Data clustering: 50 years beyond K-means" (Jain, 2009) diungkapkan bahwa pengorganisasian data ke dalam suatu kluster merupakan suatu model yang paling mendasar untuk pemahaman dan pembelajaran. Analisis kluster adalah studi formal untuk mengelompokkan, atau clustering benda-benda sesuai dengan karakteristik yang diukur berdasarkan kemiripan satu sama lain. Clustering adalah pengelompokan menggunakan teknik unsupervised learning dimana tidak diperlukan pelatihan pada metode tersebut atau dengan kata lain, tidak ada fase learning serta tidak menggunakan pelabelan pada setiap kelompok. Pada jurnal yang sama diungkapkan pula perihal penggunaan K-means sebagai algoritma yang sejak dulu bahkan sampai saat ini cukup handal digunakan untuk proses clustering. Algoritma K-Means Clustering merupakan metode clustering berbasis jarak yang membagi data-data ke dalam sejumlah

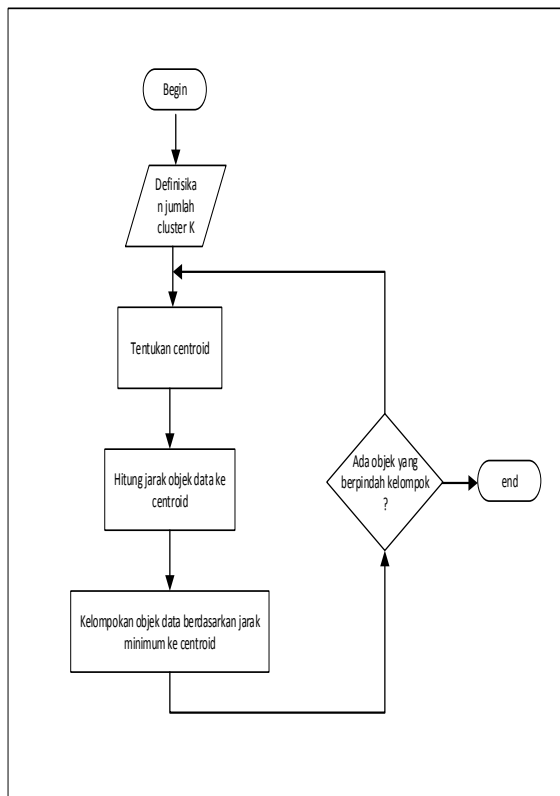
cluster dimana proses clustering tersebut dilakukan dengan memperhatikan kumpulan dari data-data yang akan dikelompokkan. Pada algoritma ini, pusat cluster atau centroid dipilih pada tahap awal secara acak dari sekumpulan koleksi (populasi) data. Kemudian K-Means menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu centroid yang telah didefinisikan sebelumnya berdasarkan jarak minimum antara komponen (data) dengan masing-masing centroid. Posisi centroid akan dihitung kembali sampai semua komponen data dikelompokkan ke setiap centroid dan terakhir akan terbentuk posisi centroid baru. Iterasi ini akan terus dilakukan sampai tercipta kondisi konvergen.

CARA KERJA ALGORITMA K-MEANS

Adapun cara kerja algoritma dalam melakukan pencarian pelanggan potensial adalah sebagai berikut:

1. Definisikan jumlah K cluster.
2. Inisialisasi K pusat cluster (centroid) sebagai seed points (prototipe cluster awal). Centroid ini dapat diperoleh secara acak atau dipilih dari K objek data pertama.
3. Untuk setiap komponen data, hitung dan tandai jarak (distance) ke centroid awal kemudian masukkan data tersebut ke centroid yang paling dekat jaraknya 2
4. Hitung dan ubah kembali centroid tiap cluster sebagai rata-rata dari seluruh anggota kelompok (group) cluster tersebut.
5. Cek semua data kembali dan taruh setiap data yang terdekat dengan centroid baru. Jika anggota tiap cluster tidak berubah (konvergen), maka langkah berhenti dan jika masih berubah, kembali ke langkah 2.

Berikut adalah *flowchart* algoritma *K-Means Clustering* pada Penentuan pelanggan potensial.



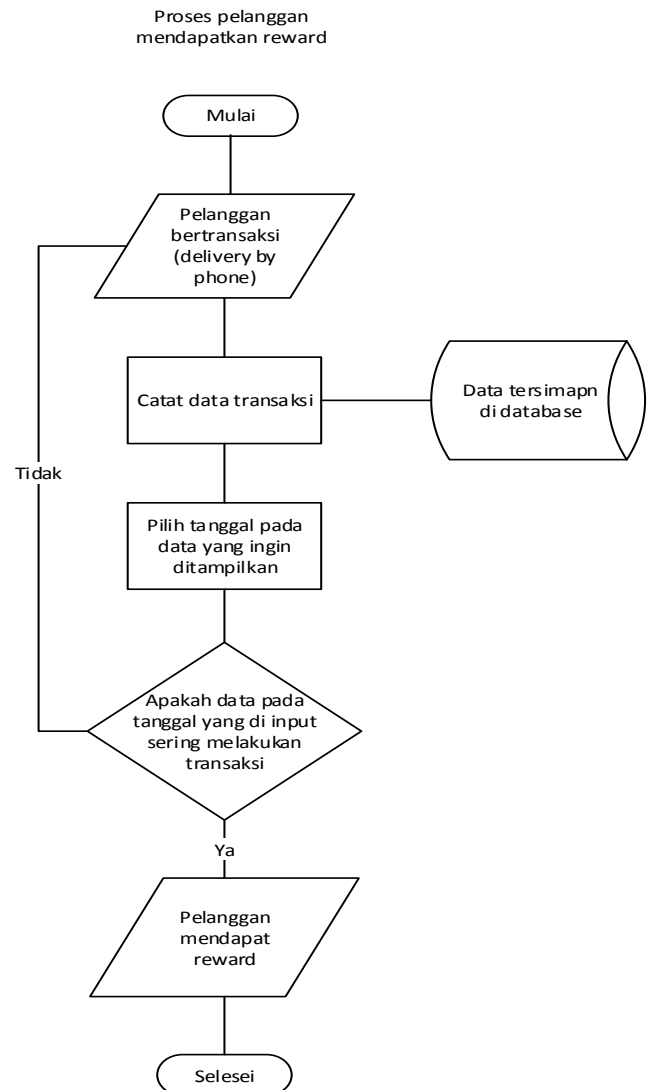
Gambar 1 Flowchart cara kerja Algoritma K-Means Clustering pada pencarian pelanggan

HASIL DAN DISKUSI

PERANCANGAN SISTEM PENCARIAN PELANGGAN POTENSIAL

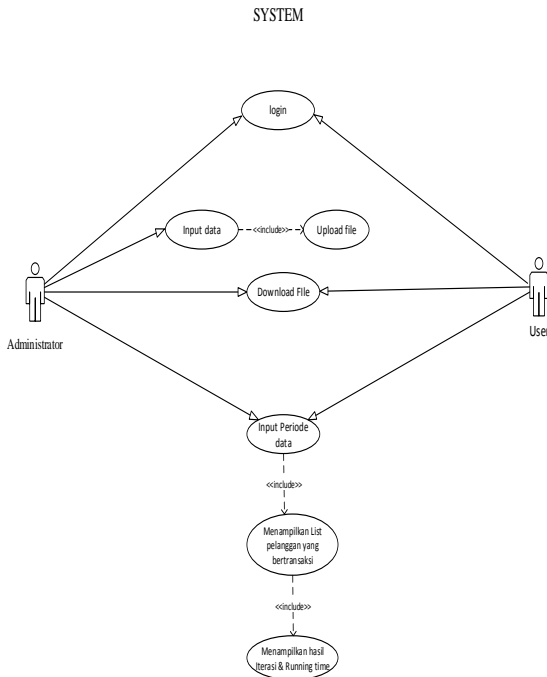
Pada Restoran XYZ akan dibuat sebuah Sistem Aplikasi pencarian pelanggan potensial berbasis web. Perancangan aplikasi ini dibangun bertujuan untuk memudahkan owner secara cepat dan tepat dalam mendapatkan informasi pelanggan potensial, cukup dengan cara membuka website, informasi mengenai siapa yang mendapat pelanggan potensial dapat dilihat dengan baik dan jelas.

Perancangan aplikasi website Restoran XYZ yang dibangun ini bersifat object oriented (berorientasi objek) dengan menggunakan Unified Modeling Language (UML) sebagai bahasa pemodelan.



Gambar 2 Flowchart Proses Pelanggan Mendapatkan reward

USE CASE DIAGRAM



Gambar 3 Use Case Diagram Sistem penentuan pencarian pelanggan potensial

a. Aktor

Adapun aktor-aktor yang terlibat dalam Aplikasi Pencarian Pelanggan Potensial sebagai berikut

Tabel 1 Definisi Actor Proses Pencarian Pelanggan Potensial

No	Aktor	Deskripsi
1	Admin	1. Mengelola / memanipulasi (Insert, Update, dan Delete) seluruh content dan informasi yang akan ditampilkan pada halaman website Pencarian pelanggan potensial
2	User	2. Melihat informasi data pada halaman web

b. Definisi Use Case

Tabel 2 Tabel Definisi Use Case Proses Pencarian Pelanggan Potensial

No	Use Case	Deskripsi
1	Login Admin	Validasi untuk Admin sistem sebelum masuk ke dalam sistem
2	Admin mengelola sistem	Melakukan penambahan, penghapusan, dan edit data di menu yang ada
3	Login User	Melakukan login, melihat data dan melakukan input tanggal untuk penentuan pelanggan potensial

c. Skenario Use Case

Skenario (flow of event) use case dari Aplikasi Pencarian Pelanggan Potensial adalah sebagai berikut

Tabel 3 Tabel Skenario Use Case Proses Pencarian Pelanggan Potensial

No	Use Case	Deskripsi
1	Login Admin	Admin harus melakukan Login menggunakan Username dan Password.
2	Input Data	Merupakan proses pengelolaan Input data transaksi
3	Input Periode Data	Merupakan proses pengelolaan untuk menampilkan hasil data yang di olah oleh admin dan user
4	Login User	User harus melakukan login untu melihat informasi data yang ada
5	Download File	Admin dan User mencetak file data berdasarkan periode yang di inputkan

HASIL UJI PROGRAM

Tabel 4 Table Hasil Uji Program

Jenis Data	Data Transaksi
Tahun Data	Januari 2016 s/d Juni 2016
Total Transaksi	600 Record
Jmlh Customer thdp Transaksi	± 650 Record
Jumlah Cluster	2
Proses Clustering	± 10 Menit
Jumlah Iterasi	10 Iterasi

KESIMPULAN DAN STUDI LANJUTAN

Beberapa kesimpulan yang dapat diambil dari tugas akhir ini adalah sebagai berikut:

1. Algoritma K-Means Clustering dapat digunakan untuk mengelompokkan data pelanggan sebagai pendukung keputusan penentuan pelanggan potensial.
2. Data pelanggan dikelompokkan menjadi dua cluster (potensial dan tidak potensial). Kemudian setiap cluster diklasifikasikan berdasarkan kriteria mana yang lebih diprioritaskan yaitu salah satu dari kriteria jumlah berapa kali transaksi dan jumlah nominal transaksi. Cluster dengan nilai terbesar pada centroid akhir merupakan cluster yang direkomendasikan menerima reward, sedangkan cluster dengan nilai terkecil merupakan cluster yang tidak berhak menerima reward.
3. Diterapkan aplikasi dengan konsep Pengambil keputusan menggunakan Algoritma K-Means guna mengelompokkan data pelanggan untuk mengetahui data yang memiliki potensi dengan melakukan proses perhitungan jumlah transaksi untuk mengelompokkan data pelanggan, Hasil ini dapat digunakan untuk memberi saran pertimbangan dalam menentukan strategi

penjualan yaitu memberikan reward untuk pelanggan dengan posisi cluster teratas.

Berdasarkan kesimpulan dan analisis laporan, saran dari peneliti untuk penelitian lebih lanjut yaitu:

1. Dapat dilakukan penelitian lebih lanjut dengan mengembangkan desain antarmuka sistem, sehingga sistem dapat lebih mudah digunakan.
2. Dapat dilakukan pengembangan lebih lanjut dengan mengcluster pelanggan berdasarkan wilayahnya.

Dengan berkembangnya teknologi mobile saat ini, diharapkan dapat dilakukan pengembangan lebih lanjut untuk akses sistem ini melalui smartphone dengan berbagai platform yang ada.

DAFTAR PUSTAKA

- Algoritma Pemrograman 1. (n.d.). Bagan Yang memperlihatkan urutan prosedur dan proses dari beberapa file di dalam media tertentu. *Definisi Dan Simbol Flowchart*, 1–9.
- Algoritma Pemrograman 1. (n.d.). Bagan Yang memperlihatkan urutan prosedur dan proses dari beberapa file di dalam media tertentu. *Definisi Dan Simbol Flowchart*, 1–9.
- Bell, J. (n.d.). Fungsi Dan Pengertian UML. Retrieved from https://www.academia.edu/4887559/Fungsi_i_Dan_Pengertian_UML
- Christine. (2016). Mengenal Use Case Diagram. Retrieved March 16, 2016, from https://www.academia.edu/5295802/Mengenal_Use_Case_Diagram
- Etunas. (2012). Pengertian MySQL. Retrieved February 20, 2016, from <http://www.etunas.com/web/pengertian-mysql.htm>
- Fadul, F. (2016). Apa Itu Bootstrap? bagaimana Memulai Belajar Bootstrap Untuk Pemula. Retrieved March 17, 2016, from <http://dul.web.id/bootstrap/3/tuts-tips/belajar-bootstrap-untuk-pemula.php>

- Fadli. (2016). K-Means, Clustering . Retrieved June 14, 2014 , from <https://fadlikadn.wordpress.com/2013/06/14/tahap-tahap-k-means-clustering/>
- Febrina Wulandari, G. (2014). Segmentasi Pelanggan Menggunakan Algoritma K-Means Untuk Customer Relationship Management (CRM) Pada Hijab Miulan
- Glandangan, D. (2016). BAB I Model Waterfall. Retrieved March 16, 2016, from https://www.academia.edu/7585995/BAB_I_Model_Waterfall
- Haris Lubis, A. (2016). Model Segmentasi Pelanggan Dengan Kernel K-Means Clustering Berbasis Customer Relationship Management
- Ii, L. F. (n.d.). Laporan Framework II maven spring web 1, 1–6.
- Iman Hermanto, T. Implementasi Algoritma Association Rule Dan K-Means Sebagai Sistem Rekomendasi Produk Pada Website Penjualan Online
- Ismo Broto. (2016). SDLC (System Development Life Cycle). Retrieved March 16, 2016, from https://www.academia.edu/8309937/SDLC_Systems_Development_Life_Cycle_
- Java Application , JBoss , EJB dan Hibernate Software yang dibutuhkan. (n.d.).
- Teorey, T. (2006). The Unified Modeling Language (UML), 3, 33–51. <http://doi.org/10.1016/B978-012685352-0/50003-3>
- Teorey, T. (2006). The Unified Modeling Language (UML), 3, 33–51. <http://doi.org/10.1016/B978-012685352-0/50003-3>
- User, S. (2014). Perpustakaan. Retrieved March 16, 2016, from <http://www.mercubuana.ac.id/id/lembaga/perpustakaan>
- Yuki, S. Pengertian Clustering Retrieved from <https://yukisuhendrasitohang.blogspot.co.id/2010/10/clustering-adalah-proses-mengelompokkan.html>
- Yulisopianti. (2015). Perbedaan Use Case, Activity Diagram, Sequence Diagram. Retrieved March 16, 2016, from <http://yulisopianti.staff.telkomuniversity.ac.id/2015/05/02/perbedaan-use-case-activity-diagram-sequence-diagram/>