

ANALISIS CLUSTERING DENGAN K-MEANS UNTUK PENGELOMPOKAN PENJUALAN PRODUK PADA HOTEL NEWTON

Event¹Intan Utnasari²¹ Mahasiswa Program Studi Sistem Informasi, Universitas Putera Batam² Dosen Program Studi Sistem Informasi, Universitas Putera BatamEmail: pb171510008@upbatam.ac.id

ABSTRACT

Customer satisfaction is the main goal of many companies wants to achieve. One of the business fields that focus on customer satisfaction is hotel business. Apart from serving accommodation for customers, hotels also provide variety of products for sale. Inventory management is very important is hotel business because one of many ways to maintain customer satisfaction is by keeping stock so no items are empty when needed. In addition, a good inventory management will not let the company to experiences losses due to outdated inventory. Using data mining K-means clustering algorithm, we can group goods based on the salable goods and the less salable goods. This research aims to assist hotel Newton to improve their inventory management. Data that was used is sales that have been made which divided into 3 trimesters. Data also will be evaluated using RapidMiner application. The results obtained from the research are 3 clusters where the clusters consist of very salable goods, medium salable goods and not salable goods.

Keywords: data mining; clustering; k-means algorithm; inventory management

PENDAHULUAN

Di era globalisasi, perkembangan teknologi sangatlah cepat. Perkembangan teknologi yang pesat ini merupakan salah satu aspek yang sepatutnya kita manfaatkan untuk mempermudah pekerjaan kita dalam sehari-hari, baik dalam bidang apapun. Penjualan yang telah terjadi dalam sebuah perusahaan dapat disimpan dalam bentuk digital dalam sebuah komputer. Penjualan yang dilakukan oleh sebuah perusahaan dapat disebut juga sebagai data. Data dapat ditemukan baik dalam aspek pekerjaan ataupun saat berada di rumah.. Data

yang paling sederhana untuk diperhatikan adalah data penjualan atau data pembelian suatu kantor atau perusahaan. Segala bentuk pembelian dan penjualan yang terjadi dapat dijadikan data dan jika data diolah dengan baik, dapat membantu meringankan pekerjaan manusia. Data yang tidak diolah lagi akan mengakibatkan penumpukkan data yang tidak berguna. Salah satu bidang usaha yang melakukan pembelian dan penjualan barang adalah bidang perhotelan. Dengan jumlah pelanggan yang sangat banyak dan pelanggan

yang keluar masuk setiap harinya, tidak menutup kemungkinan bahwa penjualan yang terjadi dalam suatu hotel sangatlah banyak. Penjualan yang sangat banyak tentunya menghasilkan data dengan jumlah yang sangat besar. Pada Hotel Newton, pendataan sudah memanfaatkan komputer, namun hanya berupa pendataan saja, tanpa adanya pengolahan data tambahan. Dengan memanfaatkan data mining dalam pengolahan data yang ada, diharapkan data penjualan yang banyak tersebut akan menghasilkan sebuah manfaat yang dapat mempermudah pekerjaan manusia. Salah satu manfaat yang akan terlihat jika data penjualan hotel Newton dilakukan pengolahan data dengan data mining, adalah pengelompokan produk yang laris dan tidak laris. Hotel Newton sering mengalami kendala dimana persediaan barang yang dimiliki tidak sesuai dengan tingkat penjualan produk tersebut. Sering terjadinya kasus dimana persediaan barang kurang daripada kebutuhan sehingga mengakibatkan kehabisan persediaan barang. Kasus yang serupa terjadi sebaliknya dimana persediaan barang yang terlalu banyak dibandingkan dengan penjualan barang sehingga persediaan barang menumpuk dan akhirnya melewati masa kadaluarsa. Dengan melakukan pengelompokan tersebut, maka masalah seperti kekurangan atau kelebihan stok barang yang dijual dapat dihindari. Jika persediaan barang yang berlebihan dapat menyebabkan penurunan kualitas barang seiring dengan waktu yang lama akibat tidak digunakan dan untuk persediaan barang seperti makanan dan minuman dapat kadaluarsa. Cara untuk menangani masalah manajemen

persediaan barang adalah dengan melakukan pengolahan historis transaksi yang pernah terjadi dan dengan memanfaatkan data mining untuk melakukan pengolahan historis transaksi tersebut, maka akan semakin cepat dan efisien. Dengan memanfaatkan data mining untuk memproses pengolahan data penjualan yang ada, diharapkan dapat membantu hotel Newton selaku lokasi penelitian untuk menghindari kasus manajemen persediaan barang seperti kekurangan ataupun kelebihan barang persediaan.

KAJIAN TEORI

1. Knowledge Discovery in Database (KDD)

Knowledge discovery in database (KDD) merupakan salah satu tahapan atau mempunyai kegiatan yang meliputi kegiatan pengumpulan data dan diikuti oleh pemakaian data historis supaya menemukan sebuah pola yang ada dalam hubungan antar data dalam sebuah database yang berukuran besar (K. Handoko, 2016). Dalam melakukan KDD, terdapat beberapa ciri penting dalam sebuah pencarian pengetahuan, yaitu (Aranski & Handoko, 2019):

1. Data yang dikerjakan dalam jumlah yang besar.
2. Volume data yang dikerjakan harus diefisiensikan.
3. Akurasi menjadi prioritas utama.
4. Menggunakan bahasa tingkat tinggi.
5. Memanfaatkan beberapa bentuk pembelajaran otomatis.
6. Hasil yang dihasilkan berupa pengetahuan yang bermanfaat.

Secara keseluruhan, proses KDD dapat dibagi menjadi beberapa tahapan yaitu

sebagai berikut (K. Handoko & Lesmana, 2018):

1. Data Selection / Pemilihan Data
2. Data Cleaning / Pembersihan Data
3. Data Transformasi Data
4. Data Mining
5. Evaluasi

2. Data Mining

Data mining adalah suatu rangkaian proses untuk dimana data mining ini mencari nilai tambah dari sekumpulan data yang serupa pengetahuan yang selama ini tidak diketahui secara manual (K. Handoko, 2016). Tujuan utama data mining adalah memanfaatkan data suatu basis data dengan mengolahnya sehingga mendapatkan informasi baru yang berguna (lin Parlina, Agus Perdana Windarto, Anjar Wanto, 2018). Data mining sendiri sering disebut dengan istilah knowledge discovery karena tujuan utama dari dilakukannya data mining adalah untuk mendapatkan sebuah pengetahuan yang baru (Suryadi, 2019). Data mining bisa juga bisa memberikan dampak negatif maupun positif bergantung pada penggunaannya data miningnya. Data mining memiliki dua sudut pandang dalam pemanfaatannya, yaitu secara komersial dan keilmuan. Dari sudut pandang komersial, pemanfaatan data mining dapat digunakan untuk menangani meledaknya volume data dan dengan memanfaatkan data mining, diharapkan dapat menyelesaikan permasalahan ini, sedangkan pada sudut pandang keilmuan data mining dapat digunakan untuk mengcapture, menganalisis dan menyimpan data yang bersifat real time dan sangat besar. Data mining memiliki tugas untuk

mempesifikasi pola atau pattern yang ada dalam sebuah data. Dalam melakukan tugasnya, data mining memiliki kategori utama yang terdiri dari 2 yaitu (Vulandari, 2017):

1. Prediktif
2. Deskriptif

Data mining memiliki beberapa metode berdasarkan jenis tugas-tugas yang akan dilakukan oleh data mining tersebut. Beberapa metode data mining yaitu (Hermawati, 2013):

1. Estimation / Estimasi
2. Prediction / Prediksi
3. Classification / Klasifikasi
4. Clustering / Pengelompokkan
5. Association / Asosiasi
6. Description / Deskripsi

3. Metode *Data Mining*

Pada penelitian ini, peneliti akan menggunakan metode clustering dalam penelitian ini. Clustering merupakan salah satu kategori dari data mining yang bersifat tanpa arahan / unsupervised. Unsupervised artinya tidak ada target kelompok yang dibicarakan secara spesifik. Clustering adalah proses pemisahan data ke dalam sejumlah kelompok dengan karakteristik yang mirip antara satu data dengan data lainnya pada kelompok yang sama. Target dari clustering adalah untuk melakukan pendistribusian data ke dalam kelompok-kelompok yang sedemikian sehingga derajat tingkat keterhubungan antar anggota kelompok yang sama kuat dan hubungan antar anggota kelompok yang berbeda lemah. Algoritma clustering akan melakukan pembagian terhadap keseluruhan data dan menjadikannya ke dalam kelompok-kelompok yang homogeny (memiliki tingkat kemiripan

maksimal) (K. Handoko, 2016). Salah satu teknik yang dipakai dalam clustering adalah menggunakan algoritma K-means. Algoritma K-means adalah salah satu algoritma yang multivariat dengan dartujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Algoritma K-means menggunakan descriptive model untuk menjelaskan algoritma pengelompokan yang dilakukannya (Putri et al., 2020). Algoritma K-means merupakan algoritma paling banyak digunakan dengan aplikasi penggunaan yang berukuran kecil hingga menengah. Ide dasar algoritma K-means sangatlah sederhana, yaitu dengan meminimalkan Sum of Squared Error antara objek data dengan sejumlah k centroid. Namun kelemahan utama dari algoritma K-means adalah dapat yang terlalu peka terhadap pilihan cluster awal dan perhitungan solusi lokal untuk mencapai kondisi optimal (Purnamasari & Dana, 2017). Jika dijabarkan secara rinci maka langkah-langkah algoritma K-means adalah sebagai berikut (S. Handoko et al., 2020):

1. Tentukan nilai k untuk hasil cluster yang diinginkan.
2. Tentukan nilai centroid awal secara acak.

3. Menghitung jarak setiap data ke masing-masing sebuah centroid yang ada menggunakan rumus korelasi antar dua objek (Euclidean Distance)
4. Mengelompokkan setiap data berdasarkan jarak terdekat yang dihasilkan berupa data tersebut dengan centroid.
5. Tentukan posisi centroid yang baru (k,C) dengan cara menghitung nilai rata-rata dari data yang ada pada centroid yang sama.
6. Jika nilai centroid baru berbeda dengan nilai centroid lama, maka nilai centroid belum stabil. Lakukan lagi langkah ketiga hingga nilai centroid yang menghnilai yang sama dengan nilai centroid yang ada pada proses iterasi sebelumnya.

METODE PENELITIAN

1. Desain Penelitian

Bentuk desain penelitian pada penelitian ini berbentuk sebagai berikut:



(Sumber : Data Penelitian, 2020)

Gambar 1. Desain Penelitian

Berikut penjelasan masing-masing node pada desain penelitian yang ada di atas:

1. Identifikasi Masalah

Pada penelitian ini, permasalahan yang ada adalah data penjualan yang tidak diolah dengan menggunakan data mining

yang mengakibatkan proses stok barang yang tidak efektif dan efisien.

2. Studi Pustaka

Studi pustaka yang dilakukan peneliti pada Referensi yang digunakan antara lain buku, jurnal dan penelitian sebelumnya yang memiliki permasalahan atau memanfaatkan metode yang sama.

3. Analisa Data

Analisa data yang dilakukan peneliti pada penelitian ini adalah melakukan pengumpulan data primer dari lokasi penelitian yang telah ditentukan lalu dilakukan pembersihan pada data dari noise yang sering disebut data cleaning.

4. Data Mining

Data mining pada penelitian ini adalah melakukan proses perhitungan memanfaatkan algoritma K-means.

5. Evaluasi

Pada penelitian ini, evaluasi akan menggunakan alat bantu berupa aplikasi RapidMiner.

6. Hasil Penelitian

Pada tahapan ini, hasil penelitian akan dijelaskan agar dapat dimengerti oleh masyarakat.

2. Objek dan Variabel Penelitian

Lokasi penelitian yang digunakan untuk penelitian adalah hotel Newton yang merupakan salah satu hotel di kota Batam, Kepulauan Riau. Untuk sampel penelitian yang digunakan pada penelitian ini adalah data penjualan produk di hotel Newton pada tahun 2018 yang telah dibagi ke dalam bentuk trimester, sedangkan variabel penelitian adalah jenis-jenis produk yang terdapat pada data penjualan produk tersebut.

3. Model Penelitian

Model penelitian bersifat kuantitatif. Model kuantitatif berarti pada penelitian ini akan dilakukan proses perhitungan dengan menggunakan angka-angka dalam menemukan solusi dari permasalahan yang ada. Prosedur yang digunakan dalam penelitian ini adalah *data mining* dengan menggunakan algoritma K-means. Algoritma K-means akan melakukan perhitungan secara berulang atau iterasi hingga mendapatkan pola yang tidak berubah lagi dari data yang diolah. Algoritma K-means akan menggunakan rumus *Euclidian Distance* dalam perhitungannya.

HASIL DAN PEMBAHASAN

Setelah melakukan penelitian, hasil yang didapatkan dari pengolahan data mining clustering algoritma K-means adalah pengelompokan yang dilakukan berhasil dengan menghasilkan 2 buah kelompok. Jumlah kelompok telah ditetapkan oleh peneliti sebelum perhitungan algoritma dilakukan. Kelompok terbagi berdasarkan sifat data yaitu: sangat laku, menengah laku dan tidak terlalu laku. Hasil dari penelitian menghasilkan kelompok dengan data-data yang memiliki sifat yang bersamaan di dalamnya. Cluster 1 adalah kelompok data yang memiliki sifat tidak terlalu laku terjual, cluster 2 adalah kelompok data yang memiliki sifat menengah laku dan cluster 3 berisi data yang sangat laku terjual. Berikut isi data berdasarkan clusternya masing-masing:

Cluster 1: *Pocari Sweat, Marlboro Lights, Sampoerna H, Evolution Red*

Cluster 2: *Green Tea Yeos, Heineken-CAN, GG Surya, Marlboro Menthol, Marlboro Red, Evolution Menthol*

- ▼ cluster_0
 - ▶ Pocari Sweat
 - ▶ Marlboro Lights
 - ▶ Sampoerna H
 - ▶ Evolution Red
- ▼ cluster_1
 - ▶ Green Tea yeos
 - ▶ Heineken-CAN
 - ▶ GG Surya
 - ▶ Marlboro Menthol
 - ▶ Marlboro Red
 - ▶ Evolution Menthol

Sumber : Data Peneliti, 2020

Gambar 2. Hasil K-means Clustering

Berikut adalah tabel data yang digunakan dalam penelitian :

Tabel 1. Tabel data yang digunakan

No	JENIS	Kuartal I	Kuartal II	Kuartal III
1	Pocari Sweat	508	479	469
2	Susu Beruang	143	246	346
3	Heineken-CAN	78	88	72
4	GG Surya	129	192	170
5	Marlboro Lights	403	314	283
6	Marlboro Menthol	143	197	292
7	Marlboro Red	216	229	221
8	Sampoerna H	287	327	444
9	Evolution Red	407	300	279
10	Evolution Menthol	246	122	77

(Sumber: Data Penelitian, 2020)

Lakukan perhitungan *Euclidian Distance* untuk semua jenis produk pada tabel data di atas, maka selanjutnya akan didapatkan posisi *cluster* yang baru.

Tabel 2. Titik *centroid* baru iterasi pertama

c0	159,16667	179	196,33333
c1	401,25	355	368,75

(Sumber: Data Penelitian, 2020)

Lakukan proses iterasi berulang-ulang hingga posisi *cluster* pada iterasi tidak lagi mengalami perubahan. Jika posisi tidak berubah lagi, maka iterasi dinyatakan berhenti dan hasil *cluster* akhir telah didapatkan. Pada penelitian ini, iterasi

terjadi hingga 2 kali dan tidak ada perubahan posisi *cluster* yang dialami oleh data dari iterasi ke-1 dan iterasi ke-2 sehingga dapat dinyatakan bahwa iterasi ke-2 sudah menjadi hasil akhir proses *data mining clustering*.

Tabel 3. Hasil iterasi ke-2 / Iterasi akhir

	c0	c1
k1	534,82	191,89
k2	164,77	281,23
k3	174,15	513,65
k4	42,10	374,42
k5	291,87	95,06
k6	98,68	312,33
k7	79,61	268,37
k8	315,57	139,64
k9	287,92	105,42
k10	158,21	404,36

(Sumber: Data Penelitian, 2020)

SIMPULAN

Berdasarkan hasil yang di teliti dapat dilakukan menarik kesimpulan tersebut:

1. Dengan menerapkan *data mining* diharapkan dapat membantu pihak hotel Newton dalam melakukan manajemen persediaan barang dengan baik.
2. Pengelompokan yang dilakukan oleh *data mining* pada penelitian ini menghasil 3 buah *cluster* yang membagi jenis produk ke dalam sifat

sangat laku, menengah laku dan tidak terlalu laku.

DAFTAR PUSTAKA

- Aranski, A. W., & Handoko, K. (2019). *Data Mining Clustering Nilai IQ Siswa Sekolah SMAK Yos Sudarso Batam. September*, 31–36.
- 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). *Jurnal Teknologi Dan Sistem Informasi*, 02(03), 31.
- Handoko, K., & Lesmana, L. S. (2018). Data Mining Pada Jumlah Penumpang Menggunakan Metode Clustering. *Snistek*, 1, 97–102.
- Handoko, S., Fauziah, F., & Handayani, E. T. E. (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1), 76–88. <https://doi.org/10.35760/tr.2020.v25i1.2677>
- Hermawati, F. A. (2013). *Data Mining*. Andi.
- Iin Parlina, Agus Perdana Windarto, Anjar Wanto, M. R. L. (2018). Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center. *Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center Untuk Clustering Program Sdp*, 3(1), 87–93.
- Purnamasari, A. I., & Dana, R. D. (2017). *Analisis Penikmat Kopi Berdasarkan Cita Rasa Kopi Menggunakan Teknik Clustering Algoritma K-Means di Cirebon The Analyzing of Coffee Lover Based on Coffee Flavor Using K-Means Algorithms Clustering Technique in Cirebon*.
- Putri, R. A., Maghfirani, N. I., Setyawan, G. R., Rayhan, A. A., & Rakhmawati, N. A. (2020). Analisis Pengelompokan Peraturan Kementerian dengan Menggunakan K-Means Clustering. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(2), 152. <https://doi.org/10.32736/sisfokom.v9i2.817>
- Suryadi, S. (2019). Penerapan Metode Clustering K-Means Untuk Pengelompokan Kelulusan Mahasiswa Berbasis Kompetensi. *Jurnal Informatika*, 6(1), 52–72. <https://doi.org/10.36987/informatika.v6i1.738>
- Vulandari, R. T. (2017). *Data Mining Teori dan Aplikasi RapidMiner*.

	<p>Biodata Penulis pertama, Event, merupakan mahasiswa Prodi Sistem Informasi, Universitas Putera Batam</p>
	<p>Biodata Penulis kedua, Intan Utnasari, S.Kom., M.Kom., merupakan Dosen Prodi Sistem Informasi, Universitas Putera Batam. Penulis banyak berkecimpung di bidang Sistem Informasi</p>