

Penerapan Algoritma K-Means untuk Mengelompokkan Penjualan Mebel

Mutiara Rita Avivah
Universitas Buana Perjuangan
Karawang, Indonesia
if15.mutiaraavivah@mhs.ubpkarawang.ac.id

Ahmad Fauzi Universitas Buana
Perjuangan
Karawang, Indonesia
ahmadfauzi@ubpkarawang.ac.id

Yana Suryana
Universitas Buana Perjuangan
Karawang, Indonesia
yanasuryana@ubpkarawang.ac.id

Abstract—

Toko mebel menjual barang furniture untuk keperluan rumah tangga, kantor atau pun instansi. Dalam transaksi jual beli selama ini toko mebel masih menggunakan laporan transaksi secara manual pada buku laporan. Cara tersebut memiliki kelemahan diantaranya buku laporan tersimpan secara acak dan mudah hilang. Sehingga, pemilik toko mebel sulit untuk mengetahui transaksi penjualan. Maka dalam mengatasi permasalahan tersebut dilakukan penelitian untuk pencatatan tertata rapih dan mengelompokkan barang dengan menggunakan Algoritma K-Means. Penelitian dilakukan dengan cara perhitungan manual menggunakan Microsoft Excel kemudian dievaluasi menggunakan tools Rapidminer 9.2. Penelitian ini menghasilkan tiga clustering. Clustering 1 yaitu cluster sedang yang mendapatkan hasil 14 kode barang, clustering 2 yaitu cluster tinggi yang mendapatkan hasil empat kode barang, dan cluster 3 yaitu clustering rendah yang mendapatkan hasil 98 kode barang.

Kata kunci — *Clustering, K-Means, Transaksi, Toko mebel,*

I. PENDAHULUAN

Penjualan merupakan kegiatan yang dilakukan oleh penjual dalam memperdagangkan barang atau jasa dengan harapan akan memperoleh laba dari adanya transaksi-transaksi tersebut [1]. Selama ini, masih banyak toko mebel yang mencatat transaksi perdagangan dalam bentuk catatan pada buku laporan secara acak sehingga informasi kurang jelas dalam mengelompokkan data penjualan. Cara mengatasi permasalahan tersebut dilakukan menggunakan teknik *data mining* dengan metode K-Means karena telah banyak penelitian yang membuktikan bahwa metode K-Means mampu menyelesaikan masalah [2, 3, 4,5].

Penelitian yang terkait dengan metode K-Means tentang pengelompokan penyakit Demam Berdarah *Dengue* (DBD) yang telah dilakukan oleh Fatmawati dan Windarto [2] mampu menghasilkan tiga *cluster*. *Cluster* pada penelitian tersebut berdasarkan provinsi yang memiliki pasien penyakit DBD di seluruh Indonesia. Terdapat empat Provinsi *cluster* tingkat tinggi yang memiliki penyakit DBD, 13 provinsi *cluster* tingkat sedang dan 17 provinsi *cluster* tingkat rendah. Berikutnya pada penelitian Sulastri dan Gufroni [3] tentang Penderita *Thalassaemia* yang telah berhasil dilakukan menjadi tiga kelompok. Kelompok tersebut terdiri dari kelompok 1 berjumlah 214 data, kelompok 2 berjumlah 138 data, dan kelompok 3 berjumlah 23 data. Di samping itu Yunita [4] juga melakukan pengelompokan penerimaan mahasiswa baru dengan menggunakan algoritma k-means. Penelitian tersebut menggolongkan data penerimaan mahasiswa menjadi tiga *cluster*. *Cluster* pertama terdapat 195 data dengan rata-rata nilai UN 43,29. *cluster* dua 271 data dengan rata-rata nilai UN 43,33, dan cluster tiga 50 data dengan rata-rata nilai UN 43,77. Pada penelitian dalam pengelompokan Kunjungan Wisata Ke Objek Wisata Unggulan yang dilakukan Maulida [5]. Menghasilkan lima objek wisata unggulan yang berada *cluster* paling rendah menjadi catatan bagi pemerintah Prov. DKI. Jakarta. Dikarnakan dengan adanya perbaikan sarana dan prasarana objek wisatawan unggulan dapat meningkatkan jumlah kunjungan wisatawan yang berdampak pada pengenalan objek wisata dan peningkatan devisa Negara

Berdasarkan masalah dan penelitian-penelitian yang telah membuktikan menggunakan *data mining*, maka penelitian ini menerapkan K-Means dalam mengambil keputusan menentukan pengelompokan mebel yang sering diminati. Sehingga mampu menambah pengetahuan tentang transaksi penjualan, serta membantu pihak toko dalam mengelompokkan penjualan mebel.

Penelitian ini dijelaskan lebih terperinci pada bagian kedua menguraikan teori-teori yang menjadi referensi. Bagian ketiga mengimplementasikan Algoritma K-Means pada penjualan mebel selanjutnya Bagian empat menunjukkan hasil perhitungan, akhirnya kesimpulan dijelaskan pada bagian lima.

II. ALGORITMA K-MEANS

Algoritma K-Means dikenalkan oleh James MacQueen pada tahun 1976 [6]. K-Means merupakan algoritma untuk mengclusterkan perhitungan yang secara berulang-ulang. Algoritma K-Means memilih *cluster* dengan secara acak.. Kemudian menetapkan nilai *cluster* secara acak dan sementara nilai tersebut

akan menjadi *centroid*. Kelebihan dari Algoritma K-Means yaitu mudah untuk dipelajari dengan waktu komputasi yang cepat dan umum untuk digunakan secara umum[7].

Langkah-langkah melakukan *clustering* dengan metode K-Means yang pertama menentukan k untuk jumlah pengelompokan, lalu menentukan tiga nilai awal yang dipilih secara acak. Selanjutnya melakukan alokasi semua data ke *cluster* terdekat, setelah melakukan kedekatan pada suatu data ke *cluster* untuk menentukan jarak antara data dengan pusat *cluster*. Cara menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang ditunjukkan pada persamaan 1.

$$D(i,j)=\sqrt{(x_{1i}-x_{1j})^2+(x_{2i}-x_{2j})^2+\dots+(x_{ri}-x_{rj})^2} \quad (1)$$

Di mana:

$D(ij)$ = jarak data ke i ke pusat *cluster* j

X_{ki} = data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

Sehingga menghitung kembali pusat *cluster* dengan anggota *cluster* yang baru. Pusat *cluster* yaitu rata-rata dari semua data dalam *cluster* tertentu. Cara menentukan posisi *cluster* suatu dokumen. Adapun rumus iterasi lainnya didefinisikan pada persamaan 2 :

$$C(i) = \frac{x_1+x_2+x_3+\dots+x_n}{\sum x} \quad (2)$$

Keterangan :

x_1 = Nilai data *record* ke 1

x_2 = Nilai data *record* ke 2

$\sum x$ = jumlah data *record*

Setelah setiap objek memakai pusat *cluster* yang baru, maka pusat *cluster* tidak berubah lagi maka proses *clustering* selesai

III. IMPLEMENTASI ALGORITMA K-MEANS TERHADAP PENGELOMPOKAN PENJUALAN MEBEL

A. Data Penelitian

Data yang didapatkan dalam penelitian ini adalah Toko Mebel Kosambi dengan melakukan wawancara dengan pemilik usaha dan pegawai Toko Mebel Kosambi. Data penjualan mebel yang diperoleh pada tahun 2016. Data yang didapat terdiri dari 1364 transaksi, 116 nama barang, dan 16 atribut. 16 atribut yaitu Tanggal Penjualan, Kode Barang Nama Barang, Harga, Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, Desember. Berdasarkan data tersebut dilakukan penyeleksian data untuk memfokuskan penelitian dengan mengambil data penjualan yang berjumlah 14 atribut yaitu Kode Barang, Nama Barang, Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, Desember.

B. Tahap Implementasi Algoritma K-Means

Tahap implementasi algoritma K-Means dilakukan dengan perhitungan data dengan menggunakan tiga *cluster*. Berikutnya menentukan tiga nilai awal yang dipilih secara acak. Setelah menentukan nilai awal, maka melakukan perhitungan pada setiap data barang ke *centroid*, agar menghasilkan nilai jarak. Nilai jarak dihitung dengan menggunakan persamaan 1. Contoh diambil dari “Nama Barang Meja Belajar Olympic”, pada penjualan mebel. Cara perhitungan manual “Nama Barang Meja Belajar Olympic” untuk menghitung *centroid* 1 menggunakan persamaan 1.

$$D(1,1) = \sqrt{\frac{((0-1)^2 + (1-2)^2 + (0-2)^2 + (1-2)^2 + (2-2)^2 + (2-4)^2 + (2-2)^2 + (2-3)^2 + (0-1)^2 + (1-0)^2 + (0-2)^2 + (1-0)^2)}{12}} = 4,3589$$

Cara perhitungan manual “Nama Barang Meja Belajar Olympic” untuk menghitung *centroid* 2 menggunakan persamaan 1.

$$D(2,1) = \sqrt{\frac{((0-1)^2 + (1-3)^2 + (0-7)^2 + (1-4)^2 + (2-5)^2 + (2-4)^2 + (2-3)^2 + (2-3)^2 + (0-3)^2 + (1-0)^2 + (0-0)^2 + (1-0)^2)}{12}} = 9.43398$$

Cara perhitungan manual “Nama Barang Meja Belajar Olympic” untuk menghitung *centroid* 1 menggunakan persamaan 1

Kode Barang	Nama Barang	Centroid 1	Centroid 2	Centroid 3
-				
-				
KT02	Kursi Tamu Sudut Minimalis	5.92807	7.02673	6.73289

Setelah mengetahui hasil *clustering* pada Tabel 1, disamping itu mengurutkan nilai *centroid* terkecil untuk menentukan *cluster* tinggi, *cluster* sedang dan *cluster* rendah. Sehingga mendapatkan hasil yang ditunjukkan pada Tabel 2.

Tabel 2 Hasil perhitungan manual dan evaluasi secara komputasi

Cluster	Kode Barang	Jumlah
Tinggi	JM01, SB06, SB10, SB12	4 Kode Barang
Sedang	MB02, LPB30, MR01, LP03, SB11, SB15, MM05, RP01, RP02, RP03, RP05, RP07, MKT 01, KT02	14 Kode Barang
Rendah	MB01, MB03, MB04, MB05, MB06, MB07, MB08, MB09, MB10, MB11, MB12, MB13, MB14, LPB01, LPB02, LPB03, LPB04, LPB05, LPB06, LPB07, LPB08, LPB09, LPB10, LPB11, LPB12, LPB13, LPB14, LPB15, LPB16, LPB17, LPB18, LPB19, LPB20, LPB21, LPB22, LPB23, LPB24, LPB25, LPB26, LPB27, LPB28, LPB29, LPB31, MR02, MR03, MR04, MR05, MR06, MR07, MR08, RTV01, RTV02, RTV03, RTV04, RTV05, RTV06, RTV07, RTV08, RTV09, RTV10, RTV11, RTV12, RTV13, LP01, LP02, LP04, LP05, LP06, LP07, SB01, SB02, SB03, SB04, SB05, SB07, SB08, SB09, SB13, SB14, KB01, KB02, JM02, MM01, MM02, MM03, MM04, RP04, RP06, MKT02, S01, S02, S03, S04, S05, S06, S07, S08, KT01	98 Kode Barang

Hasil Pengelompokan penjualan mebel menggunakan algoritma K-Means dengan perhitungan secara manual membentuk tiga *cluster* dengan *centroid* (C_1, C_2, C_3) dan membentuk tiga *cluster* pada pengujian menggunakan evaluasi komputasi dengan *centroid* (C_1, C_2, C_3), keduanya sama – sama melakukan pengulangan sampai tiga kali pengulangan. Pada pengujian menggunakan Rapidminer menghasilkan nilai jumlah yang sama pada setiap *cluster*. Hasil *cluster* berdasarkan tingkatannya yaitu *cluster* tinggi yaitu C_2 terdapat empat Nama Barang. Pada *cluster* sedang yaitu C_1 terdapat 14 Nama Barang, sedangkan pada *cluster* rendah yaitu C_3 terdapat 98 Nama Barang.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian pada bagian IV, maka dapat disimpulkan bahwa algoritma K-Means mampu mengelompokkan data penjualan mebel. Berdasarkan data yang diolah algoritma tersebut mampu menghasilkan tiga *cluster*. *Cluster* terdiri dari *cluster* tinggi yang menghasilkan empat kode barang, *cluster* sedang menghasilkan 14 kode barang dan *cluster* rendah menghasilkan 98 kode barang. Penelitian ini berhasil untuk menentukan pengelompokan mebel yang sering diminati. Sehingga mampu membantu pihak toko dalam mengelompokkan penjualan mebel. Sehingga saran untuk penelitian selanjutnya adalah menggunakan data transaksi yang bersumber lebih dari satu toko (penjual), agar rekomendasi yang digunakan menjadi lebih baik.

PENGAKUAN

Naskah ilmiah ini sebagian dari penelitian Tugas Akhir milik Mutiara Rita Avivah dengan judul Penerapan *Data mining* Menggunakan Pada Penjualan Furnitur, dibimbing oleh Ahmad Fauzi dan Yana Cahyana

DAFTAR PUSTAKA

- [1] Kusriani dan Luthfi. *Algoritma Data mining*. Yogyakarta : Andi Offset, .2009.
- [2] K. Fatmawati and A.P.Windarto, "Data mining Penerapan RapidMiner dengan K-Means Cluster pada daerah terjangkau Demam Berdarah Dengue berdasarkan provinsi," *Jurnal of Computer Engineering System and Science*, p.175,2018.
- [3] Sulastri, H., & Gufroni, A. I. Penerapan *data mining* dalam pengelompokan penderita thalassaemia. *Nasional Teknologi dan Sistem Informasi*, vol 2, 2017 Siregar, A. M., & Paspahuana, A.. *Pengolahan Data Menjadi Informasi Dengan Rapidminer*. Surakarta: Civi.Kekata Grow. 2017
- [4] Yunita, F. Penerapan *Data mining* Menggunakan Algoritma K-Means Clustering pada Penerimaan Mahasiswa Baru. *Jurnal Sistemasi* Vol. 7 No. 3, pp 238 – 249.2018

- [5] Maulidia, L. Penerapan *Data mining* Dalam Mengelompokkan Kunjungan Wisata Unggulan Di Prov, DKI Jakarta Dengan K-Means. 2018
- [6] MacQueen, J. 1967. *Some methods for clasification and analysis of multivariate observations. In proceedings of the Fufth Berkeley Symposium on Mathematical Statistic and Probability, Volume 1 : Statistic*. Berkeley, Calif.: University of California Press.
- [7] Achmad, Y. Implementasi Automatic Clustering menggunakan Differential Evolution dan CS Measure untuk Analisis Data Kemahasiswaan. *Jurnal Ilmiah NERO* Vol. 1 No. 2
- [8] W.M.W Mobd, A. H. Beg, T. Herawan and K. F. Rabbi, “MaxD K-means: A Clustering Algorithm for Auto-generation of centroids and distance of data points in cluster,” *Communication in Computer and Information Science*, pp. 192-199.2012