

PENERAPAN ALGORITMA K-MEANS UNTUK MENGELOMPOKKAN DATA PENGIRIMAN PAKET DI KANTOR POS CIREBON

Aby Febrian¹, Nana Suarna², Gifthera Dwilestari³

^{1,2,3}Program Studi Teknik Informatika, STMIK IKMI Cirebon

Email: ¹aby11febrian@gmail.com, ²st_nana@yahoo.com, ³ggdwilestari@gmail.com

Masuk: 15 Maret 2022, Revisi masuk: 15 Agustus 2022, Diterima: 15 Agustus 2022

ABSTRACT

In the current industrial era, the development of shipping services has increased relatively rapidly, making trade routes for goods and services increased to meet consumer needs. One of the shipping service expeditions is PT Pos Indonesia which is a State-Owned Enterprise (BUMN). The Cirebon Post Office currently has many competitors including JNE, TIKI, Sicepat, and J&T. One of the main factors in the delivery process is that accuracy is needed in classifying shipping data, so an appropriate calculation process is needed, in order to achieve accurate results. In this study, the clustering process with the K-Means algorithm was carried out with the aim of getting the best group on packet delivery data based on the results of the DBI evaluation. The results of the cluster data packet delivery using the K-means algorithm and the Davies Bouldin Index calculation, the value closest to 0 with cluster 2 to cluster 10 experiments produces the best k value in cluster 3, namely 0.104 with the number of Cluster members 0: 414 items, Cluster 1: 280 items, Cluster 2: 6 items.

Keywords: Data Mining, Clustering, K-Means, Post Office.

INTISARI

Pada era industri sekarang ini perkembangan ekspedisi jasa kirim mengalami peningkatan yang relatif pesat, membuat jalur perdagangan barang maupun bidang jasa menjadi meningkat untuk memenuhi kebutuhan konsumen. Salah satu ekspedisi jasa kirim yaitu PT Pos Indonesia yang merupakan Badan Usaha Milik Negara (BUMN). Kantor Pos Cirebon saat ini mempunyai banyak pesaing diantaranya yaitu JNE, TIKI, Sicepat, dan J&T. Salah satu faktor utama dalam proses pengiriman yaitu dibutuhkan keakuratan dalam mengelompokkan data pengiriman, maka diperlukan proses perhitungan yang tepat, agar dapat mencapai hasil yang akurat. Pada penelitian ini dilakukan proses clustering dengan algoritma K-Means dengan tujuan untuk mendapatkan kelompok terbaik pada data pengiriman paket berdasarkan hasil evaluasi DBI. Hasil cluster data pengiriman paket menggunakan algoritma K-means serta dengan perhitungan Davies Bouldin Index nilai yang paling mendekati angka 0 dengan percobaan cluster 2 sampai cluster 10 menghasilkan nilai k terbaik pada cluster 3 yaitu 0.104 dengan jumlah anggota Cluster 0: 414 items, Cluster 1: 280 items, Cluster 2: 6 items.

Kata-kata kunci: Data Mining, Clustering, K-Means, Kantor Pos.

PENDAHULUAN

Pada era industri sekarang ini perkembangan ekspedisi jasa kirim mengalami peningkatan yang relatif pesat, membuat jalur perdagangan barang maupun bidang jasa menjadi meningkat untuk memenuhi kebutuhan konsumen. Efek yang ditimbulkan dari persaingan tersebut menyebabkan persaingan yang semakin ketat antar perusahaan yang bergerak di bidang ekspedisi jasa kirim. Salah satu ekspedisi jasa kirim yaitu PT Pos Indonesia yang merupakan Badan Usaha Milik Negara (BUMN). Bidang usaha yang dikelola PT Pos Indonesia yaitu bidang usaha pengiriman paket dan surat, bidang usaha

jasa keuangan dan logistik. PT Pos Indonesia mempunyai berbagai jenis produk bidang usaha yang beroperasi, namun PT. Pos Indonesia memiliki bidang usaha utama dari sejak dulu sampai saat ini yaitu bidang usaha pengiriman paket dan surat, oleh karena itu bidang usaha pengiriman paket dan surat merupakan bisnis utama dan harus menjadi profit center PT Pos Indonesia. PT Pos Indonesia yang sebelumnya berdiri sendiri tanpa persaingan, kini mulai menghadapi persaingan dari pihak swasta yang bergerak di bidang yang sama.

Kantor Pos Cirebon saat ini mempunyai banyak pesaing diantaranya yaitu JNE, TIKI, Sicepat, dan J&T. Keunggulan pesaing berkaitan dengan layanan paket dan surat, metode packing, segi biaya, kecepatan dan keakuratan waktu pengiriman (Yunaida, 2017). Salah satu faktor utama dalam proses pengiriman yaitu dibutuhkan keakuratan dalam mengelompokkan data pengiriman, maka diperlukan proses perhitungan yang tepat, agar dapat mencapai hasil yang akurat (Nas, 2020). Suatu tahapan perlu diterapkan guna menganalisis data pengiriman paket, salah satu yang bisa dilakukan adalah dengan mengelompokkan data pengiriman paket sesuai dengan karakteristik masing-masing dari setiap wilayah pengiriman. Tujuan utama dari penelitian ini adalah untuk mengetahui kelompok terbaik dari data pengiriman paket di Kantor Pos Cirebon menggunakan metode clustering dengan algoritma K-Means berdasarkan hasil evaluasi DBI.

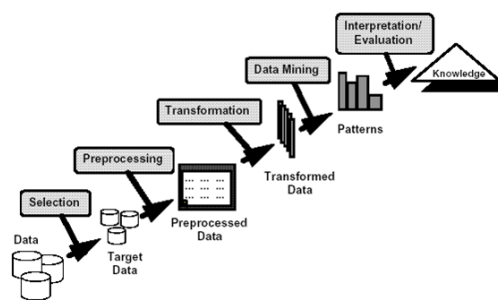
Dataset

Pada proses clustering diperlukan dataset yang akan digunakan pada tahapan pengujian. Dataset yang digunakan adalah data pengiriman paket yang ada di Kantor Pos Cirebon pada rentang bulan Juli 2020. Jumlah data yang digunakan pada tahapan pengujian ini sebanyak 700 data dengan 9 atribut diantaranya diantaranya No, KD Item, KD Produk, KD NOPEND Asal, Waktu Asal, KD NOPEND Tujuan, Waktu Tujuan, SLA, SWP.

Metode

Pada penelitian ini metode yang digunakan yaitu deskriptif dengan pendekatan kuantitatif dan analisis berupa *clustering* menggunakan algoritma K-Means. *Clustering* adalah salah satu teknik yang termasuk dalam salah satu fitur data mining, algoritma *clustering* adalah algoritma yang mengelompokkan beberapa data ke dalam kelompok data tertentu. *Clustering* yaitu teknik dalam mengelompokkan data berdasarkan kesamaan dari karakteristik data (Aditya et al., 2020). Dengan menggunakan algoritma K-Means yaitu metode pengelompokan data non-hirarki yang mencoba untuk mempartisi data menjadi dua atau lebih kelompok. Metode ini membagi data ke dalam kelompok-kelompok sehingga data dengan karakteristik yang

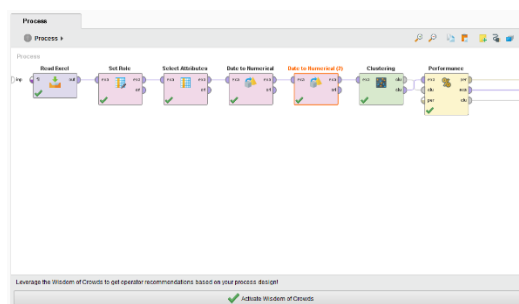
sama digolongkan ke dalam satu kelompok yang sama sedangkan data dengan karakteristik yang tidak sama dikelompokkan ke dalam kelompok lain (Annur, 2019). Dalam menentukan kelompok terbaik menggunakan hasil evaluasi DBI. Davis Bouldien Index merupakan metrik untuk meninjau/mengevaluasi hasil dari algoritma pengelompokan. Penggunaan DBI pada sebuah cluster akan dianggap memiliki skema pengelompokan yang optimal yaitu dengan DBI yang memiliki hasil minimal (Kamila et al., 2019). Metode data mining yang digunakan dalam menganalisis data adalah dengan proses tahapan KDD (*Knowledge Discovery in Database*). KDD adalah proses komputasi menggunakan algoritma matematika untuk mengekstrak data dan menghitung probabilitas kemungkinan suatu tindakan di masa mendatang. Hasil KDD berupa suatu pengetahuan yang potensial dan bermanfaat (Muttaqin & Defriani, 2020). Tahapan proses KDD tertampil pada gambar 1 berikut :



Gambar 1. Tahapan Proses KDD

HASIL & PEMBAHASAN

Tahapan pengujian ini menggunakan algoritma K-Means dengan *tools* RapidMiner Versi 9.10, didapatkan model proses data mining memakai algoritma K-Means seperti pada gambar 2 berikut :



Gambar 2. Model Proses K-Means

Pada percobaan clustering yang telah dilakukan, diperoleh hasil 3 cluster terbaik yang nilai DBI nya mendekati 0 atau terkecil dengan menggunakan algoritma K-Means.

Dari hasil pengujian didapat nilai k terbaik ada pada *cluster 3*, hasil cluster model dapat dilihat pada gambar 3 berikut :

Untuk tabel hasil dbi dan cluster dapat dilihat pada tabel 1 berikut :

Tabel 1. Hasil Clustering dan Nilai DBI

Cluster Set	Jumlah Anggota Cluster	Hasil Nilai DBI
2	Cluster 0: 414 items	0,115
3	Cluster 1: 286 items Cluster 0: 414 items Cluster 1: 280 items Cluster 2: 6 items	0,104
4	Cluster 0: 414 items Cluster 1: 50 items Cluster 2: 6 items Cluster 3: 230 items Cluster 0: 414 items Cluster 1: 209 items	0,406
5	Cluster 2: 22 items Cluster 3: 6 items Cluster 4: 49 items	0,425
6	Cluster 0: 209 items Cluster 1: 288 items Cluster 2: 6 items Cluster 3: 49 items Cluster 4: 22 items Cluster 5: 126 items	0,49

Cluster Model

```
Cluster 0: 414 items
Cluster 1: 280 items
Cluster 2: 6 items
Total number of items: 700
```

Gambar 3. Cluster Model

Hasil *cluster* terbaik diperoleh 3 *cluster* terbaik diantaranya *cluster 0* yaitu 414 *items*, *cluster 1* yaitu 280 *items* dan *cluster 2* yaitu 6 *items*. Masing-masing anggota cluster dan jumlah *items* data pengiriman paket dapat dilihat pada tabel 2,3 dan 4 berikut :

Tabel 2. Anggota Cluster 0

Cluster 0		
NO	Kode Pos Tujuan	Jumlah
1	45600 = Sumber	412 items
2	45652 = Dukupuntang	1 items
3	45658 = Karangkendal	1 items

Tabel 3. Anggota Cluster 1

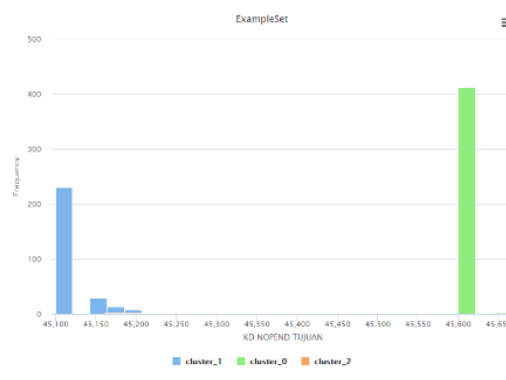
Cluster 1		
NO	Kode Pos Tujuan	Jumlah
1	45100 = Kode Pos Kota Cirebon	230 Items
2	45153 = Kode Pos Kedawung	8 Items
3	45154 = Kode Pos Plered	5 Items

Cluster 1		
4	45155 = Kode Pos Plumbon	6 Items
5	45156 = Kode Pos Klangeran	1 Items
6	45161 = Kode Pos Palimanan	1 Items
7	45162 = Kode Pos Arjawinangun	6 Items
8	45164 = Kode Pos Gegesik	2 Items
9	45166 = Kode Pos Susukan	1 Items
10	45167 = Kode Pos Ciwaringin	1 Items
11	45172 = Kode Pos Beber	2 Items
12	45173 = Kode Pos Mundu	2 Items
13	45181 = Kode Pos Astanajapura	3 Items
14	45184 = Kode Pos Karangwareng	4 Items
15	45186 = Kode Pos Karangsembung	1 Items
16	45191 = Kode Pos Babakan	3 Items
17	45192 = Kode Pos Losari	2 Items
18	45194 = Kode Pos Pabuaran	2 Items

Tabel 4. Anggota Cluster 2

Cluster 2		
NO	Kode Pos Tujuan	Jumlah
1	45192 = Kode Pos Losari	1 Items
2	45181 = Kode Pos Astanajapura	1 Items
3	45172 = Kode Pos Beber	1 Items
4	45184 = Kode Pos Karangwareng	1 Items
5	45162 = Kode Pos Arjawinangun	1 Items
6	45184 = Kode Pos Karangwareng	1 Items

Dari hasil pengujian didapatkan hasil visualizations cluster algoritma K-Means bisa dilihat pada gambar 4 berikut :



Gambar 4. Visualisasi Cluster Algoritma K-Means

Algoritma K-Means di cluster 0 ditandai dengan warna histogram hijau, untuk cluster 1 ditandai dengan histogram warna oranye sedangkan untuk cluster 2 ditandai dengan histogram warna biru. Hasil nilai centroid yang didapatkan dari masing-masing atribut bisa dilihat pada tabel 4 berikut :

Tabel 5. Nilai Centroid Tiap Atribut

Atribut	Cluster	Cluster	Cluster
	0	1	2
Nopend	45.100.	45109.	456634
Asal	833	33	.7
Nopend	45.600.	45112.	45179.
Tujuan	266	050	17
SLA	1	1	1
SWP	0.780	0.904	1.167
Waktu	11.889	11.146	13
Asal			
Waktu	12.669	12.050	14.167
Tujuan			

KESIMPULAN

Hasil pengujian dengan tahapan KDD pada penerapan algoritma K-Means clustering data pengiriman paket dapat diterapkan dan didapatkan kelompok terbaik ada pada cluster 3. Hasil cluster data pengiriman paket menggunakan algoritma K-means serta dengan perhitungan Davies Bouldin Index nilai yang paling mendekati angka 0 dengan percobaan cluster 2 sampai cluster 10 menghasilkan nilai k terbaik pada cluster 3 yaitu 0.104 dengan jumlah anggota Cluster 0: 414 items, Cluster 1: 280 items, Cluster 2: 6 items.

DAFTAR PUSTAKA

- Aditya, A., Jovian, I., & Sari, B. N. (2020). Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019. *Jurnal Media Informatika Budidarma*, 4(1), 51. <https://doi.org/10.30865/mib.v4i1.1784>
- Annur, H. (2019). Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering. *Jurnal Informatika Upgris*, 5(1). <https://doi.org/10.26877/jiu.v5i1.3091>
- Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa Dan Manajemen*

- Sistem Informasi*, 5(1), 119. <https://doi.org/10.24014/rmsi.v5i1.7381>
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*, 12(2), 121–129. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- Nas, C. (2020). Data Mining Pengelompokan Bidang Keahlian Mahasiswa Menggunakan Algoritma K-Means (Studi Kasus : Universitas Cic Cirebon). *Syntax : Jurnal Informatika*, 9(1), 1. <https://doi.org/10.35706/syji.v9i1.3472>
- Yunaida, E. (2017). Analisis Strategi Pemasaran untuk Meningkatkan Pendapatan Produk Surat dan Paket Kantor Pos Langsa. *Jurnal Manajemen & Keuangan*, 6(1), 688–699.

BIODATA PENULIS

- Aby Febrian**, lahir di Kuningan pada tanggal 08 Februari 2001, Saat ini tercatat sebagai Mahasiswa Program Studi Teknik Informatika STMIK IKMI Cirebon.
- Nana Suarna, M.Kom**, lahir di Majalengka pada tanggal 17 Desember 1965, menyelesaikan pendidikan S1 bidang ilmu Teknik Informatika dari Universitas STMIK Tulus Cendikia Bandung tahun 2003, S2 bidang ilmu Teknik Informatika, dari Universitas UDINUS Semarang tahun 2010, Saat ini tercatat sebagai Dosen Tetap di STMIK IKMI Cirebon dengan jabatan akademik akademik Lektor pada bidang minat data science.
- Gifthera Dwilestari, S.I.Kom., M.Kom**, lahir di Subang pada tanggal 01 Desember 1994, menyelesaikan pendidikan S1 bidang ilmu Ilmu Komunikasi dari Universitas Telkom tahun 2017, S2 bidang ilmu Sistem Informasi dari STMIK LIKMI tahun 2020. Saat ini tercatat sebagai Dosen Tetap di Program Studi Sistem Informasi STMIK IKMI Cirebon dengan jabatan akademik Tenaga Pengajar pada bidang minat data science.