

Penerapan *Data Mining* Menggunakan *K-Means* Untuk Penentuan Penerima Bantuan Langsung Tunai (BLT) Dana Desa Pada Desa Cempaka Timur

Huswatun Hasanah¹, Pakarti Riswanto², Sidik Rahmatullah³

^{1,2}Program Studi Teknologi Komputer, Institut Teknologi Bisnis dan Bahasa Dian Cipta Cendikia

³Program Studi Sistem Informasi, Institut Teknologi Bisnis dan Bahasa Dian Cipta Cendikia

^{1,2,3,4}Jl. Lintas Sumatera No.03, Candi MAS, Lampung Utara, Lampung

E-mail: huswatunhna@gmail.com¹, tutiriswanto@gmail.com², sidik@dcc.ac.id³

Abstrak

Bantuan Tunai Langsung adalah pemberian hadiah keuangan kepada mereka yang kurang beruntung. Dengan bantuan ini, masyarakat miskin akan mampu memerangi virus corona (Covid19). Dalam pendistribusian Bantuan Langsung Tunai ini sering terjadi masalah sebagaimana pendistribusiannya tidak sepenuhnya tepat sasaran. Penerima BLT tidak tepat sasaran karena proses penentuan masih menggunakan proses pencatatan pembukuan manual dan proses komputerisasinya menggunakan MS. Excel sehingga tingkat keakurasian pemrosesan data masih diragukan. Oleh karena itu untuk memprediksi penentuan penerima BLT menggunakan metode K-Means serta implementasinya menggunakan MS. Excel dan Python. Mengingat hal tersebut merupakan salah satu strategi yang digunakan untuk mengelompokkan data sebagai panduan dalam mengambil keputusan untuk mengelompokkan data dalam jumlah yang cukup besar, maka pendekatan K-Means Clustering sangat cocok untuk mengimplementasikan prediksi penerima Bantuan BLT-DD. Dari hasil yang didapatkan dalam pengolahan data Bantuan Langsung Tunai menggunakan atribut yang sangat mempengaruhi yaitu pekerjaan, status tempat tinggal, penghasilan dan jumlah tanggungan dengan memperoleh dua Cluster yaitu layak atau tidak layak. Pada Tahun 2020 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 38 orang dan kategori tidak layak 72 orang, Tahun 2021 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 30 orang dan kategori tidak layak 70 orang, Tahun 2022 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 29 orang dan kategori tidak layak 61 orang.

Kata Kunci: Data Mining, K-Means, Bantuan Langsung Tunai Dana Desa.

Abstract

Direct Cash Assistance is the provision of a sum of money to underprivileged people. This assistance is to help underprivileged communities in dealing with the corona virus (Covid19). In the distribution of Direct Cash Assistance, problems often occur as the distribution is not fully on target. BLT recipients were not on target because the determination process still used manual bookkeeping and the computerized process used MS. Excel so that the level of accuracy of data processing is still in doubt. Therefore to predict the determination of BLT recipients using the K-Means method and its implementation using MS. Excel and Python. Given that this method is one of those utilized to group data as a reference in making decisions on clustering sizable amounts of data, it is quite ideal for applying the predictions of BLT-DD Assistance recipients. From the results obtained in the processing of Direct Cash Assistance data using attributes that greatly influence, namely work, status of residence, income and number of dependents by obtaining two clusters, namely feasible or not feasible. In 2020 there were 38 eligible recipients of direct cash assistance and 72 inappropriate categories, in 2021 there were 30

utama dari pemrosesan data. Penambangan data, juga dikenal sebagai penemuan pengetahuan dalam basis data (KDD), adalah metode yang menggunakan teknik seperti klasifikasi asosiasi dan pengelompokan untuk secara otomatis mencari data dalam jumlah yang sangat besar untuk mendeteksi pola. Untuk mengidentifikasi desa penerima bantuan langsung tunai (BLT) Desa Cempaka Timur, dapat diterapkan data mining dengan menggunakan K-Means [5]. Tahapan metode KDD ada 6 yaitu :

1. Selection data
2. Processing data
3. Transformasi data
4. Data mining
5. Evaluation data
6. knowledge

2.5 Bantuan Langsung Tunai (BLT)

Menurut Selviana, Irwan, dan Risfaisal (2016), Bantuan Langsung Tunai (BLT) adalah pemberian uang kepada mereka yang kurang mampu. Program bantuan pemerintah yang pendanaannya berasal dari APBN kemudian dilaksanakan melalui keuangan desa di seluruh Indonesia (Suhu, Suaib, Abdul, & Wance, 2021) [2].

2.6 Dana Desa

Bantuan Langsung Tunai (BLT) adalah bantuan yang berasal dari alokasi dana desa pada Anggaran Belanja Desa (APB Desa) yang akan diberikan oleh pemerintah kepada masyarakat yang kehidupan mata pencaharian karena pandemik virus corona atau covid 19 [6].

2.7 Implementasi

Implementasi berasal dari Bahasa Inggris “to implement” yang artinya mengimplementasikan. Implementasi bukan hanya suatu aktivitas, tetapi implementasi juga merupakan suatu kegiatan yang direncanakan serta dilaksanakan dengan serius dan mengacu pada norma-norma tertentu, guna mencapai tujuan kegiatan [7].

2.8 Metode K-Means

Metode K-Means adalah metode clustering yang paling sederhana dan umum. Hal ini dikarenakan K-Means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relative cepat dan efisien [8]. Clustering adalah metode mencari dan mengkluster data yang mempunyai kesamaan karakteristik (similarity) antara beberapa data [9].

Berikut merupakan langkah-langkah pengolahan data menggunakan metode K-Means [10] :

1. Menentukan jumlah cluster k yang dibentuk
2. Tentukan k centroid (titik pusat cluster) secara acak. Penentuan centroid awal dilakukan dilakukan secara acak berdasarkan data-data yang tersedia sebanyak jumlah cluster yang sudah ditentukan pada langkah pertama

3. Menghitung jarak setiap objek ke masing-masing centroid dari setiap cluster. Hitung dengan rumus Euclidian distance jarak yang paling terdekat dengan data terhadap titik pusat. Rumus Euclidian sebagai berikut:

Dist(x,y) =

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

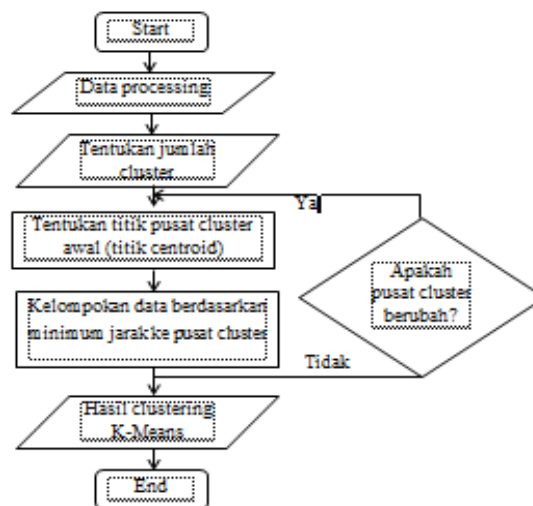
Dimana :

Dist (x,y)= jarak data ke x ke pusat cluster y

x1 = data ke 1 pada atribut data x

y1= titik pusat ke 1 pada atribut ke y

4. Mengelompokkan setiap data kedalam cluster dengan jarak terkecil. Suatu data akan menjadi anggota sebuah cluster apabila jarak data ke pusat cluster bernilai paling kecil
5. Memperbaharui pusat cluster dengan cara mengitung nilai rata-rata dari data yang menjadi anggota pada cluster tersebut
6. Ulangi langkah 1,2,3 sampai tidak ada data yang pindah ke cluster yang lain atau sampai semua anggota cluster tidak lagi berubah [11]. Tahapan perhitungan pada metode K-Means adalah sebagai berikut:



Gambar 1. Flowchart algoritma K-Means [12].

Python adalah bahasa pemrograman terkenal yang mendukung pemrograman berorientasi objek secara efektif dan dapat digunakan dengan banyak sistem operasi yang berbeda, termasuk PSc, Macintosh, dan UNIX [14]. Bahasa pemrograman Python memiliki beberapa manfaat, seperti berikut ini:

- a. Pengembangan program selesai dengan cepat dan dengan kode minimal
- b. Mendukung platform yang berbeda
- c. Python memiliki sistem pemrosesan memori otomatis, mudah dipelajari, dan merupakan bahasa pemrograman berorientasi objek.

2.9 Google Colabs

Program atau editor kode yang disebut Google Colabs atau Google Collaboratory menggunakan

sistem penyimpanan cloud gratis untuk penelitian. Aspek yang terinspirasi Jupyter disertakan dalam desain Google Colab, dan dengan semua pustaka input yang tersedia, hampir semua hal dapat digunakan untuk belajar, termasuk penelitian penambangan data [15].

III. HASIL DAN PEMBAHASAN

Perhitungan K-Means

Proses perhitungan K-Means clustering dengan menentukan k centroid awal pada penelitian ini diambil datanya secara acak. Pada perhitungan data ini, yang digunakan adalah tahun 2022. Penentuan untuk pusat centroid K01 dan K02 disajikan pada tabel berikut :

Tabel 1. Tabel Hasil Jarak dengan Centroid

| No | Nama | C1 | C2 |
|----|---------------|-----------|-----------|
| 1 | Karno | 500.004 | 500.001 |
| 2 | Hermanto | 1.000.001 | 2 |
| 3 | Sugiarto | 1.000.001 | 2 |
| 4 | Satimin | 1.000.009 | 6 |
| 5 | Juwarno | 500.004 | 500.001 |
| 6 | Sikan | 14 | 1.00.004 |
| 7 | Sulistiyono | 500.009 | 500.004 |
| 8 | Sugeng | 500.004 | 500.001 |
| 9 | Andri Mustopa | 4 | 1.000.001 |
| 10 | Boimin | 1.000.009 | 6 |

Proses pada tabel diatas merupakan proses literasi kesatu, kemudian akan dilakukan perhitungan dengan menentukan centroid baru. Hasil proses literasi kedua akan disajikan pada tabel dibawah ini.

Tabel 2. Tabel Hasil Jarak dengan Centroid literasi-2

| No | Nama | C1 | C2 |
|----|---------------|---------|---------|
| 1 | Karno | 109.756 | 10.204 |
| 2 | Hermanto | 609.757 | 489.798 |
| 3 | Sugiarto | 609.757 | 489.798 |
| 4 | Satimin | 609.757 | 489.796 |
| 5 | Juwarno | 109.756 | 10.204 |
| 6 | Sikan | 390.245 | 510.204 |
| 7 | Sulistiyono | 109.757 | 10.204 |
| 8 | Sugeng | 109.756 | 10.204 |
| 9 | Andri Mustopa | 390.244 | 510.204 |
| 10 | Boimin | 609.757 | 489.796 |

Hasil dari pengelompokkan pada literasi kedua ditemukan ada perubahan data yang berubah, maka proses perhitungan dilanjutkan pada literasi berikutnya yang disajikan pada tabel dibawah ini .

Tabel 3. Tabel Hasil Jarak dengan Centroid literasi-3

| No | Nama | C1 | C2 |
|----|-------|---------|--------|
| 1 | Karno | 109.756 | 10.204 |

| | | | |
|----|---------------|---------|---------|
| 2 | Hermanto | 609.757 | 489.798 |
| 3 | Sugiarto | 609.757 | 489.798 |
| 4 | Satimin | 609.757 | 489.796 |
| 5 | Juwarno | 109.756 | 10.204 |
| 6 | Sikan | 390.245 | 510.204 |
| 7 | Sulistiyono | 109.757 | 10.204 |
| 8 | Sugeng | 109.756 | 10.204 |
| 9 | Andri Mustopa | 390.244 | 510.204 |
| 10 | Boimin | 609.757 | 489.796 |

Hasil dari pengelompokkan pada literasi ketiga tidak ditemukan data yang berubah maka perhitungan dihentikan. Dari hasil literasi 2 sampai literasi 3, maka dapat diambil kesimpulan yaitu :

- Untuk cluster 1 dengan status layak ada 2 orang yang untuk menerima BLT-DD.
- Untuk cluster 8 dengan status layak ada 2 orang yang untuk menerima BLT-DD.

Hasil pengujian yang dilakukan dapat dilihat yang tergolong ke cluster 1 sebagai layak dan cluster 2 sebagai tidak layak seperti yang disajikan pada tabel dibawah ini :

Tabel 4. Tabel hasil pengujian

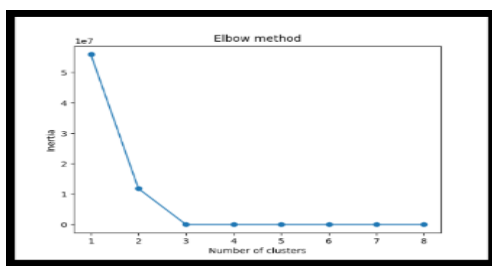
| No | Nama | Cluster | Keterangan |
|----|---------------|---------|-------------|
| 1 | Karno | C2 | Tidak Layak |
| 2 | Hermanto | C2 | Tidak Layak |
| 3 | Sugiarto | C2 | Tidak Layak |
| 4 | Satimin | C2 | Tidak Layak |
| 5 | Juwarno | C2 | Tidak Layak |
| 6 | Sikan | C1 | Layak |
| 7 | Sulistiyono | C2 | Tidak Layak |
| 8 | Sugeng | C2 | Tidak Layak |
| 9 | Andri Mustopa | C1 | Layak |
| 10 | Boimin | C2 | Tidak Layak |

Berdasarkan tabel diatas maka terdapat 2 kepala keluarga yang benar layak untuk mendapatkan BLT-DD dengan tingkat persentase 10%, kepala keluarga yang layak tersebut adalah atas nama Sikan dan Andri Mustopa.

Hasil Implementasi

Proses perhitungan kesesuaian dalam Penentuan Penerima Bantuan Langsung Tunai (BLT) Dana Desa Pada Desa Cempaka Timur pada tahun 2020-2022 adalah Prediksi penerima Bantuan Langsung Tunai, Total prediksi Bantuan Langsung Tunai, Prediksi berdasarkan tahun 2020 penerima layak sebanyak 38 dan tidak layak 72 orang. Prediksi berdasarkan tahun 2021 penerima layak sebanyak 30 dan tidak layak 70 orang. Prediksi berdasarkan tahun 2022 penerima layak sebanyak 29 dan tidak layak 61 orang.

Selanjutnya penulis akan melakukan perhitungan peramalan dengan menggunakan Python berikut adalah langkah-langkahnya :



Gambar 2. Tampilan proses kesembilan

Pada gambar diatas didapatkan hasil dari coding dan import scatter yang telah kita lakukan berdasarkan variable x dan y.

```
import sys
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

x = [1.500, 2000, 2000, 2000, 1.500, 1000, 1.500, 1.500, 1000, 2000, 1000,
1.500, 1.500, 1.500, 1.500, 2000, 1000, 1.500, 1.500, 1000, 1000, 1.500,
1.500, 1000, 1.500, 1.500, 1.500, 1.500, 1000, 1.500, 1000, 1.500,
1.500, 2000, 1000, 1000, 1.500, 1.500, 1000, 2000, 2000, 2000, 1.500,
1000, 1000, 1000, 1000, 1.500, 1000, 1000, 1000, 1000, 1.500, 1000, 1000,
2000, 2000, 1.500, 1000, 1000, 1000, 1000, 1000, 1.500, 2000, 1000, 2000,
1.500, 1.500, 2000, 1.500, 1.500, 1.500, 1.500, 1.500, 2000, 1.500]
y = [2, 3, 3, 1, 2, 1, 1, 2, 2, 1, 3, 3, 3, 2, 2, 2, 2, 1, 3, 3, 0, 3, 0,
1, 2, 2, 1, 2, 2, 3, 1, 2, 3, 3, 2, 1, 2, 2, 2, 2, 0, 1, 3, 3, 0, 1, 1,
2, 4, 4, 3, 3, 3, 3, 0, 1, 0, 3, 2, 0, 0, 1, 2, 2, 2, 2, 3, 0, 0, 0, 1,
1, 1, 1, 1, 0, 3, 2, 2, 2, 2, 1, 0, 1, 3, 0]

data = list(zip(x, y))

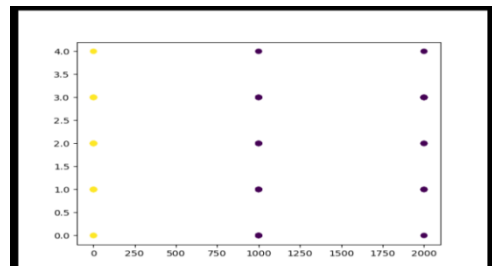
kmeans = KMeans(n_clusters=2)
kmeans.fit(data)

plt.scatter(x, y, c=kmeans.labels_)
plt.show()

#Two lines to make our compiler able to draw:
plt.savefig(sys.stdout.buffer)
sys.stdout.flush()
```

Gambar 3. Tampilan proses kesepuluh

Selanjutnya pada gambar diatas adalah coding untuk menampilkan scatter yang sesuai dengan banyaknya cluster pada data yang di uji.



Gambar 4. Tampilan proses kesebelas

Selanjutnya setelah kita dapat menentukan titik scatter pada tahap ini dapat kita lihat bahwa terdapat dua scatter dengan warna yang berbeda, dengan 2 cluster pada data tersebut yaitu dengan arti layak atau tidak layak.

IV KESIMPULAN

Dari hasil yang didapatkan dalam pengolahan data Bantuan Langsung Tunai menggunakan Algoritma K-Means, maka diperoleh kesimpulan Metode Clustering dengan Algoritma K-Means dapat digunakan untuk memprediksi Penerimaan Bantuan Langsung Tunai berdasarkan atribut yang sangat memengaruhi yaitu pekerjaan, status tempat tinggal, penghasilan dan jumlah tanggungan dengan menggunakan dua Cluster yaitu layak atau tidak layak. Dengan demikian penentuan Penerimaan Bantuan

Langsung Tunai bisa tepat sasaran tersalurkan kepada keluarga yang berhak menerimanya secara akurat. Tahun 2020 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 38 orang dan kategori tidak layak 72 orang. Tahun 2021 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 30 orang dan kategori tidak layak 70 orang. Tahun 2022 terdapat penerima Bantuan Langsung Tunai dengan kategori layak sebanyak 29 orang dan kategori tidak layak 61 orang.

DAFTAR PUSTAKA

- [1] Y. Filki, "Jurnal Informatika Ekonomi Bisnis Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa," vol. 4, pp. 6–9, 2022, doi: 10.37034/infeb.v4i4.166.
- [2] I. Algoritma, C. Untuk, B. L. T. Di, and D. Tanjung, "178160037 - Sri Juwita - Fulltext," 2022.
- [3] R. A. Farissa, R. Mayasari, and Y. Umaidah, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokkan Data Obat dengan Silhouette Coefficient di Puskesmas Karangsembung," *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 109–116, 2021, doi: 10.30871/jaic.v5i1.3237.
- [4] A. Afandi, D. Nurdianah, P. C. Rejo, N. Bayes, and K. A. Dominan, "Naive Bayes Method and C4 . 5 in Classification of Birth Data," vol. 16, no. 4, pp. 435–446, 2022.
- [5] R. Muliono and Z. Sembiring, "Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 4, no. 2, pp. 2502–714, 2019.
- [6] E. Barito, J. T. Beng, and D. Arisandi, "Penerapan Algoritma C4.5 Untuk Klasifikasi Mahasiswa Penerima Bantuan Sosial Covid-19," *J. Ilmu Komput. dan Sist. Inf.*, vol. 10, no. 1, 2022, doi: 10.24912/jiksi.v10i1.17819.
- [7] E. Dewi *et al.*, "Clustering Wilayah Dan Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Data Satlantas Polres Tasikmalaya Kota," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 8–1, no. 1, pp. 1–11, 2019, doi: 10.36774/jusiti.v8i1.595.
- [8] N. N. F. R, D. S. Anggraeni, and U. Enri, "Pengelompokkan Data Kemiskinan Provinsi Jawa Barat Menggunakan Algoritma K-Means dengan Silhouette Coefficient," *Temat. J. Teknol. Inf. Komun.*, vol. 9, no. 1, pp. 29–35, 2022, [Online]. Available: <http://jurnal.plb.ac.id/index.php/tematik/article/view/901>
- [9] A. Muni, "Analisis Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Penjualan Sepeda Motor Studi Kasus PT. Alfa Scorpii," *Juti Unisi*, vol. 4, no. 1, pp. 1–8, 2020, doi: 10.32520/juti.v4i1.1087.
- [10] I. N. M. Adiputra, "Clustering Penyakit Dbd Pada Rumah Sakit Dharma Kerti Menggunakan

- Algoritma K-Means,” *Inser. Inf. Syst. Emerg. Technol. J.*, vol. 2, no. 2, p. 99, 2022, doi: 10.23887/insert.v2i2.41673.
- [11] O. Kinerja, E. C. U. Study, K. Mobil, and A. Dan, “Implementasi Algoritma K-Means Dan Algoritma Apriori,” vol. 1, no. 2, pp. 81–88, 2021.
- [12] C. Algorithm *et al.*, “Classification and Clustering of Internet Quota Sales Data Using,” vol. 9, no. 2, pp. 268–283, 2023, doi: 10.26555/jiteki.v9i2.25970.
- [13] Raf Knowledge, *MENGENAL MICROSOFT OFFICE 2010*. 2010.
- [14] I. Algoritma and C. Untuk, “Implementasi algoritma c4.5 untuk memprediksi keterlambatan pembayaran sumbangan pembangunan pendidikan sekolah menggunakan python,” vol. 10, pp. 36–44, 2020.
- [15] M. Khandava Mulyadien and U. Enri, “Algoritma K-Means Untuk Pengelompokan Bantuan Langsung Tunai (BLT),” *J. Ilm. Wahana Pendidik.*, vol. 2022, no. 12, pp. 198–210, 2022.