Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Eligibility Analysis of Non-Cash Food Assistance Recipients in Rajadesa Village Using the K-Means Clustering Method

Ati Sumiati[1*], Rini Astuti[2], Willy Prihartono[3]

[1,3]STMIK IKMI Cirebon
[2]STMIK LIKMI Bandung
atisumiat2633@gmail.com[1], riniastuti@likmi.ac.id[2], Willy@likmi.ac.id[3]*

Abstract

The distribution of Non-Cash Food Assistance (BPNT) in Rajadesa Village often encounters challenges in accurately determining recipient eligibility. The selection process lacks objectivity due to limited data and suboptimal verification, leading to uneven distribution. To address this issue, the study employs the K-Means Clustering algorithm to enhance the efficiency of BPNT recipient selection. The assessment is based on attributes such as ID number (NIK), name, address, occupation, and the amount of assistance received.The methodology adopts the Knowledge Discovery in Databases (KDD) approach, involving stages such as data selection, preprocessing, transformation, data mining, and evaluation. Data processing is carried out using RapidMiner version 10.5. The clustering results are evaluated using the Davies Bouldin Index (DBI), yielding the best model with three clusters and a DBI value of 0.346.
This approach successfully identifies recipient groups more accurately, enabling a more targeted distribution of food assistance. Consequently, the study provides significant contributions to ensuring the eligibility of BPNT recipients in Rajadesa Village through the application of K-Means Clustering-based analytical technology.

Keywords: *Non-Cash Food Assistance, K-Means Clustering, Recipient Eligibility, Knowledge Discovery in Database, Davies Bouldin Index.*

1. Introduction

Non-Cash Food Assistance (BPNT) is a government aid program in the form of food designed to reduce the financial burden and provide better balanced nutrition to the recipient families (KPM) in a timely and targeted manner, distributed every month [1] This assistance is distributed to Beneficiary Families (KPM) through invitations provided to the recipients, which can be used to exchange food items at local villages or Post Offices partnered with the government [2].

The distribution of Non-Cash Food Assistance (BPNT) often faces difficulties in determining eligible recipients in a fair manner. In Rajadesa Village, Rajadesa District, Ciamis Regency, many people receive BPNT, but it is still challenging to identify those who are truly in need. Challenges in the BPNT distribution process often arise due to a lack of data and inadequate eligibility verification. Incomplete or inaccurate data can lead to BPNT not being distributed to the right recipients or cause the assistance to not meet expectations, increasing the possibility of aid being given to those who do not qualify or not reaching those who genuinely need it. This issue is particularly relevant given the importance of distributing food assistance equitably to meet the needs of the less fortunate [3].

Therefore, the implementation of the K-Means Clustering method becomes a potential solution to address the eligibility of BPNT recipients based on certain factors such as location, occupation, or amount of assistance. By utilizing K-Means Clustering for BPNT eligibility, it is hoped that issues related to eligibility or mis-targeting can be minimized [4]

Previous research has used the K-Means clustering method to assist in the eligibility process for Non-Cash Food Assistance (BPNT) in various regions. For example, a study by [3]. showed that data mining and the K-means algorithm can be used for clustering or grouping social assistance data and the Family Hope Program in 2022. The data processing results were also carried out using the RapidMiner application, which grouped the data into 2 clusters: cluster 0, which includes the high category with 19 provinces, and cluster 1, which includes the low category with 15 provinces.

1.1 Data Mining

To ensure that social assistance is targeted correctly, data mining is an appropriate method for collecting and analyzing data to identify hidden patterns within the dataset [5]. To carry out the data mining process, social assistance recipient data can be grouped using the K-Means Clustering approach based on certain criteria to ensure that social assistance is truly deserving and reaches the right recipients.

1.2 Clustering

Clustering is a data mining technique used to identify groups of objects with common characteristics in large datasets. The primary goal of the clustering method is to group data or objects into clusters in such a way that each cluster contains data that is as similar as possible. Clustering organizes data based on the similarities between the objects [6].

1.3 K-Means

The K-Means algorithm is a non-hierarchical data clustering technique that seeks to divide data into two or more groups based on shared characteristics. In the K-Means method, data is divided so that objects with similar traits are grouped together, while objects with differing characteristics are placed in separate groups. The primary goal of this clustering process is to minimize a given objective function, which typically aims to reduce variation within each group and maximize variation between groups [7].

2. Research Method

2.1 Research Method

This study uses the K-Means Clustering algorithm to assess the eligibility of Non-Cash Food Assistance (BPNT) recipients. The data analysis process follows the Knowledge Discovery in Databases (KDD) approach, which aims to uncover hidden patterns in data related to the needs of assistance recipients [8]. Figure 1:

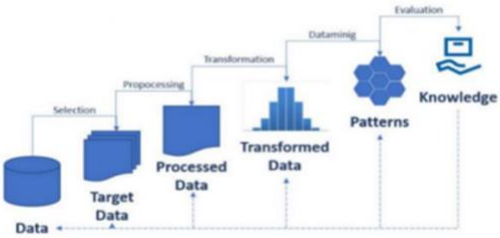

Figure 1: Knowledge Discovery in Databases

2.2 Research Supporting Data

Table 1: Data BNPT Rajadesa

No	Nama Kolom	Keterangan
1	Nama	Berisi nama calon penerima BPNT dengan jumlah sebanyak 386 data
2	Nik	Berisi angka Nik calon penerima masing-masing pada calon penerima
3	Alamat	Menyatakan tempat tinggal pada masing-masing calon penerima seperti Dusun Sirnasari RT 04 RW 08 Dusun Cibingbin RT 06 RW 03 Dusun Jagamulya RT 04 RW 05 Dusun Jagamulya RT 03 RW 05 Dusun Jagamulya RT 07 RW 04 Dusun Sirnasari RT 06 RW 09 Dusun Cibingbin RT 01 RW 01 Dusun Sirnasari RT 01 RW 07 Dusun Cibingbin RT 05 RW 02 Dusun Sirnasari RT 07 RW 08 Dusun Cibingbin RT 04 RW 02 Dusun Sirnasari RT 03 RW 08 Dusun Sirnasari RT 02 RW 07 Dusun Jagamulya RT 02 RW 04 Dusun Sirnasari RT 05 RW 09 Dusun Cibingbin RT 07 RW 03 Dusun Jagamulya RT 06 RW 06 Dusun Jagamulya RT 01 RW 04 Dusun Cibingbin RT 03 RW 01 Dusun Jagamulya RT 05 RW 05 Dusun Cibingbin RT 02 RW 01 Dusun Jagamulya RT 05 rw 06
4	Pekerjaan	Menyatakan berberapa jenis pekerjaan calon penerima seperti Wiraswasta Ibu Rumah Tangga Petani/pekebun Buruh Pedagang Pegawai
5	Jumlah (KG)	Menyatakan jumlah berat pangan yang didapatkan.

3. Result and Discussion

3.1. Data Selection

In the data selection phase of the modeling process, the sequence of operators involves using Read Excel to load data from an Excel file, Set Role to define the roles of attributes such as labels and identifiers, and Select Attributes to filter the necessary data, ensuring readiness for subsequent analysis or modeling tasks. Figure 2:

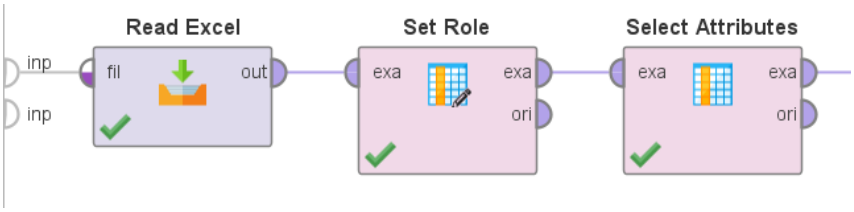

Figure 2: Data Selection Stage Modeling Process

3.2. Preprocessing Data

In the preprocessing stage, addressing missing values or verifying the absence of data in Non-Cash Food Assistance (BPNT) recipient records is essential to ensure a clean dataset for analysis. This step begins with identifying any missing or null values within particular attributes of the data.

3.3. Transformation Data

During the transformation phase, data formats are modified to ensure they are ready for use in modeling or further analytical processes. This includes converting nominal attributes into numerical values for attributes categorized as nominal.

3.4. Data Mining

At this stage, the K-Means Clustering method is utilized to classify individuals as eligible or ineligible for BPNT assistance. The initial step in the data mining process involves setting the primary parameter of the K-Means Clustering algorithm, specifying the number of clusters needed [7]. The data is then organized into clusters based on comparable attributes using the Clustering operator. Figure 3 :

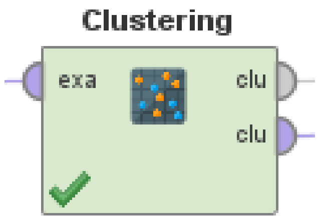

Figure 3: Example of an image with acceptable resolution

The clustering process starts by setting the K value, such as K-2 to K-4, allowing the model to classify the data into 2 to 4 clusters. The parameters for this process are adjusted within the Clustering operator. Table 2:

Table 2: Parameter Clustering

No	Parameter	Keterangan
1.	*K*	2-4
2.	*Max run*	10
3.	*Meansur type*	*Numerical Measure*
4.	*Numerical Measure*	*Euclidean Distance* *Chebychev Distance*
5.	*Max optimization steps*	100

Determining the optimal K value by selecting the smallest possible value and using the K-Means Clustering operator with default configurations. Table 3:

Table 3: Operator Performance Result

Type Numerical Measures	K	Max_Runs	DBI
Euclidean Distance	2	10	0.471
	3	10	0.346
	4	10	0.386
CorelationSimilarity	2	10	0.471
	3	10	0.346
	4	10	0.386

3.5. Evaluation

The analysis results identify cluster 3 as the optimal cluster with a score of 0.346, as presented in Table 4 of the Performance Operator results. According to the DBI, cluster 0 consists of 219 items, cluster 1 comprises 30 items, and cluster 2 contains 137 items.

Table 4: Clustering Test Result

No	Cluster	Keterangan
1	Cluster 0	219 *Items*
2	Cluster 1	30 *Items*
3	Cluster 2	137 *Items*
	Total Number of Items	386 *Items*

The Davies-Bouldin index, calculated from the clustering model, measures the effectiveness of data grouping. A lower value signifies that the clusters are more cohesive and distinctly separated [9]. Figure 4:

Figure 4: Davies Bouldin Best Value

The centroids of the clusters serve to define the distinct characteristics of each group. Table 5:

Attribute	cluster_0	cluster_1	cluster_2
PEKERJAAN	0	3.433	1.175

Figure 5: Centroid Point of Clustering Results

From the processing results, the optimal K value identified using the Davies-Bouldin Index (DBI) divides BPNT recipients into three clusters, representing the lowest score. Cluster 0 indicates eligibility, Cluster 1 represents ineligibility, and Cluster 2 reflects partial eligibility [10].

The K-Means Clustering algorithm results indicate that the smallest score is achieved when the data is grouped into three clusters, analyzed using RapidMiner version 10.5. This research focuses on identifying the optimal K value for a relevant dataset by experimenting with K values ranging from 2 to 4 clusters. Each experiment uses a different cluster count and evaluates performance using the DBI metric. Figure 6:

Figure 6: Results of K-Means Clustering Visualizations

The Scatter/Bubble visualizations comparing DBI scores reveal that three clusters represent the optimal configuration, with the lowest DBI value recorded at 0.346. For the BPNT recipient dataset, cluster 0 comprises 219 items, cluster 1 includes 30 items, and cluster 2 contains 137 items, totaling 386 items. These clustering results are consistent with the findings of prior research conducted by [11]

Testing with the elbow method determined that three clusters are optimal, leading this study to utilize cluster 0, cluster 1, and cluster 2. Similarly, the silhouette coefficient analysis confirms that three clusters are the most effective, with a silhouette coefficient score of 0.489, outperforming other cluster configurations [12].

4. Conclusion

This study demonstrates that the K-Means Clustering algorithm is effective in improving the accuracy of determining the eligibility of Non-Cash Food Assistance (BPNT) recipients in Rajadesa Village. By grouping data based on recipient attributes, this method successfully divides recipients into three clusters: eligible, ineligible, and partially eligible. Evaluation using the Davies-Bouldin Index (DBI) shows the optimal cluster with a DBI value of 0.346. The use of K-Means Clustering can minimize targeting issues and enhance the efficiency of BPNT distribution, ensuring it is more fair and accurately targeted.

References

[1] M. I. St Nur Rahmah, Muliani.S, Andi Nilwana, "Jurnal Ilmiah Pemerintahan," vol. 12, no. Idm, pp. 27–38, 2024.
[2] T. Pipit Muliyah, Dyah Aminatun, Sukma Septian Nasution, Tommy Hastomo, Setiana Sri Wahyuni Sitepu, "済無No Title No Title No Title," *J. GEEJ*, vol. 7, no. 2, pp. 309–316, 2020.
[3] S. A. Ajun, S. Canon, B. R. Payu, and U. N. Gorontalo, "Analisis ketepatan sasaran penerima bantuan sosial provinsi gorontalo," vol. 7, pp. 8–19, 2024, doi: 10.37600/ekbi.v7i1.1319.
[4] F. Fitriani, R. Kurniawan, and T. Suprapti, "Penerapan Algoritma K-Means Clustering Untuk Identifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan (Pkh) Di Desa Tambaksari Ciamis," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3363–3369, 2024, doi: 10.36040/jati.v7i6.8197.
[5] C. A. Rahayu, "Prediksi Penderita Diabetes Menggunakan Metode Naive Bayes," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3, 2023, doi: 10.23960/jitet.v11i3.3055.
[6] K. Kameshwaran and K. Malarvizhi, "Survey on Clustering Techniques in Data Mining," vol. 5, no. 2, pp. 2272–2276, 2014.
[7] S. G. Pratama, A. Mahudi, and S. Achmadi, "Klasifikasi Penentuan Penerima Bantuan Pangan Non Tunai Menggunakan Metode K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 4, no. 1, pp. 341–348, 2020, doi: 10.36040/jati.v4i1.2360.
[8] F. Juliawati, R. Buaton, R. Saragih, and S. Kaputama, "Pengelompokan Data Mining Penerimaan Bantuan Pangan Non Tunai (BPNT) Menggunakan Metode Clustering (Studi Kasus : Kantor Desa Payabakung Hamparan Perak)," *J. Comput. Sci. Inf. Technol. E-ISSN*, vol. 3, no. 2, p. 69, 2023.
[9] A. Singh, S. Mittal, P. Malhotra, and Y. Srivastava, *Clustering Evaluation by Davies-Bouldin Index(DBI) in Cereal data using K-Means*. 2020. doi: 10.1109/ICCMC48092.2020.ICCMC-00057.
[10] D. A. Tarigan, "Optimization of the K-Means Clustering Algorithm Using Davies Bouldin Index in Iris Data Classification," *Media Online)*, vol. 4, no. 1, pp. 545–552, 2023, doi: 10.30865/klik.v4i1.964.
[11] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," *J. Artif. Intell. Technol. Inf.*, vol. 1, no. 1, pp. 20–28, 2023, doi: 10.58602/jaiti.v1i1.22.
[12] D. A. Setiady and H. Leong, "Implementation of K-Means Algorithm Elbow Method and Silhouette Coefficient for Rainfall Classification," *Proxies J. Inform.*, vol. 4, no. 1, pp. 18–25, 2024, doi: 10.24167/proxies.v4i1.12433.