|Jurnal Ilmu Komputer dan Sistem Informasi |JIKSI|Vol. 4| No. 3|Oct|Year 2023|

Implementation of Data Mining for Social Assistance Funds Using the C4.5 Algorithm in the Empat Lawang Regency Government

Pahmi Ladianto[1*]

Abstract

The Government of Empat Lawang Regency seeks to improve public welfare through social assistance programs, including the Home Renovation Assistance Program for residents living in uninhabitable houses. Many residents still live below the housing eligibility line, creating the need for a fair, data-driven selection mechanism. Traditionally, the selection process begins at the village level, where village heads identify residents with inadequate housing and low daily income levels. These data are then submitted to the Regency Government, where the Verification Team determines the final recipients. However, manual data processing makes it difficult to ensure transparency and efficiency. This study applies the C4.5 Decision Tree Algorithm to classify eligibility for social assistance using existing data from the Empat Lawang Regency Government. A web-based system was developed to support automated classification and decision-making in line with the criteria set by the Minister of Public Works and Housing Regulation No. 13/PRT/M/2016. The system enhances accuracy, transparency, and efficiency in determining eligible beneficiaries for social housing renovation programs.

Keywords

Data Mining, Social Assistance Fund, C4.5 Algorithm

Article History

Received 02 August 2023
Accepted 30 October 2023

How to Cite

Ladianto, P. (2023). Implementation of Data Mining for Social Assistance Funds Using the C4.5 Algorithm in the Empat Lawang Regency Government. Jurnal Ilmu Komputer dan Sistem Informasi (JIKSI), 4(3), [77-85].

[1*] Universitas Bina Darma, Indonesia, Corresponding email: ladiantopahmi@gmail.com

Introduction

In the era of globalization and digital transformation, technological advancement has become a central driver of change in nearly every aspect of governance and public administration. The rapid development of information and communication technology (ICT) has revolutionized how governments manage, store, and utilize data for decision-making and service delivery. One of the most significant transformations is seen in the management of public welfare programs, where the integration of digital systems enables more efficient and accountable distribution of social assistance. In such programs, accurate data management is essential to ensure that aid reaches only those who are truly eligible, thereby promoting equity and minimizing misuse of public funds.

However, in many local government institutions, data collection and selection processes are still performed manually, which often leads to inefficiencies and inaccuracies. Manual approaches are vulnerable to human error, data duplication, and subjective judgment, particularly in the determination of eligibility for housing renovation assistance—a form of social aid intended to improve living conditions for low-income households. These procedural weaknesses can cause delays, discrepancies, and even potential conflicts within communities. To overcome such limitations, the integration of information technology and data analytics offers a promising solution to automate and objectify the decision-making process.

Information technology, particularly in the form of data mining, enables governments to analyze large volumes of social data and extract actionable insights for policy implementation. Data mining refers to the process of discovering meaningful patterns and relationships from datasets through computational algorithms. For the Department of Housing and Settlement (Dinas Perumahan dan Permukiman) of Empat Lawang Regency, which manages social housing renovation programs, the adoption of data mining can enhance decision quality by identifying eligibility patterns based on quantifiable criteria. This transformation supports evidence-based governance, where data-driven insights replace intuition or manual assessment in determining social assistance recipients.

This study specifically applies the C4.5 Decision Tree Algorithm as the analytical method for classifying citizens' eligibility to receive housing renovation assistance. The C4.5 algorithm is one of the most widely used decision tree methods due to its high interpretability and accuracy in handling categorical and continuous data. It operates by constructing a decision tree based on information gain, allowing classification of individuals into eligible and ineligible categories according to various socioeconomic indicators such as income, house condition, number of dependents, and ownership status. By using this approach, the system can process verified citizen data and produce objective, transparent, and data-based classifications, thereby minimizing bias and improving fairness in aid distribution.

Previous studies have demonstrated the effectiveness of data mining techniques in similar decision-support contexts. Yunita et al. (2019) successfully developed a decision support system for housing assistance distribution using the Simple Additive Weighting (SAW) method, which facilitated the ranking of candidates based on weighted criteria. Meanwhile, Swastina (2018) applied the C4.5 algorithm to predict academic major suitability with an accuracy rate of 93.31%, confirming the algorithm's reliability in classification and pattern recognition tasks. These studies collectively illustrate the viability of decision tree-based algorithms for solving complex classification problems involving social or educational data.

The success of such approaches supports their adaptation in government welfare programs, where fairness and accuracy are critical determinants of success.

Building upon these findings, this research develops a web-based Decision Support System (DSS) utilizing the C4.5 Decision Tree Algorithm to assist the Empat Lawang Government in objectively determining recipients of social housing renovation aid. The system is designed to automate data analysis, provide transparent selection reports, and support decision-makers in verifying eligibility based on established criteria. Through this implementation, the study aims to improve the efficiency, consistency, and accountability of the housing assistance selection process, aligning with the government's vision of transparent and technology-driven public service. Ultimately, the integration of data mining into social welfare management contributes not only to better governance practices but also to the broader goals of digital transformation and sustainable community development in Empat Lawang Regency.

Methodology

This section describes the analysis, design, and implementation of data mining for social assistance eligibility. The research employs an experimental method comprising three main stages: data collection, preprocessing, and model development.

Data Collection

Two types of data were used:

1. Primary data, obtained from experimental results of applying the C4.5 algorithm.
2. Secondary data, collected from official government documents, books, and journals related to Decision Tree C4.5 and data mining, as well as population data from Empat Lawang Regency regarding residents below housing eligibility standards.

Data Preprocessing

The dataset was obtained from the Department of Housing and Settlement of Empat Lawang Regency, containing information such as name, National Identification Number (NIK), Family Card Number, address, monthly income, and other relevant attributes. These data were standardized according to predefined eligibility assessment forms to ensure consistency before analysis.

Proposed Model

The Decision Tree C4.5 algorithm was used to determine social assistance eligibility. The steps are as follows:

1. Prepare training data
2. Determine the root node using entropy:

$$Entropy(S) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

3. Calculate the gain value:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} Entropy(S_i)$$

4. Repeat steps 2-3 until all tuples are partitioned
5. The decision tree construction stops when: All records in node *N* belong the same class, No further attributes remain for partitioning.

System Design

The system was designed using several modeling tools:

1. Flowchart: Describes the flow of user activities from login, data upload (Excel format), decision tree generation, and classification results.

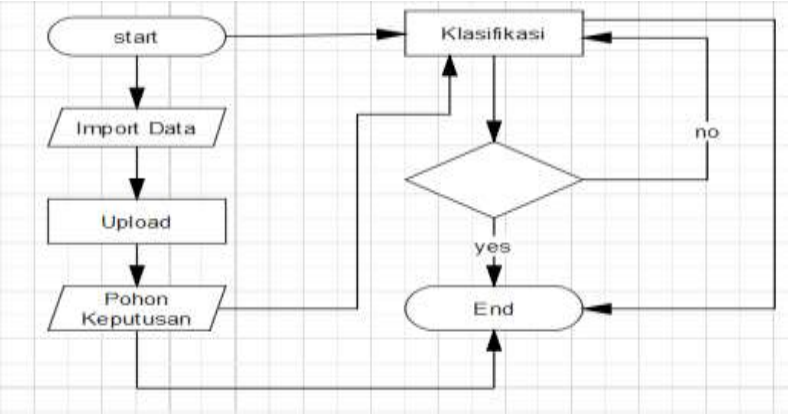

Figure 1: Flowchart

2. Use Case Diagram: Illustrates system functionalities including login, data import, decision tree generation, classification, and logout.

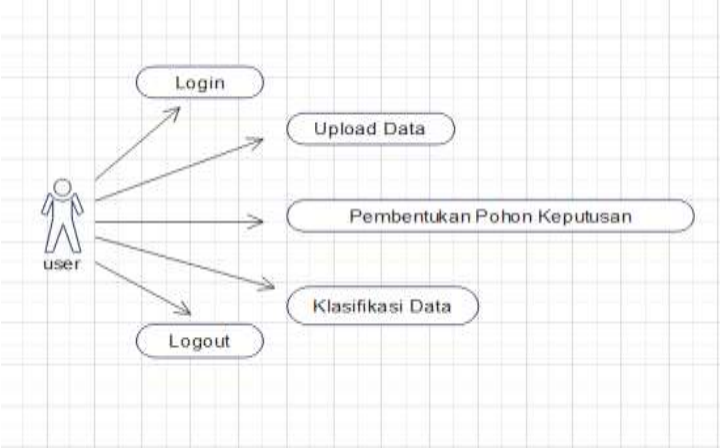

Figure 2. Use Case Diagram

3. Activity Diagram: Depicts user interactions and decision-making steps, from uploading data to viewing classification results.

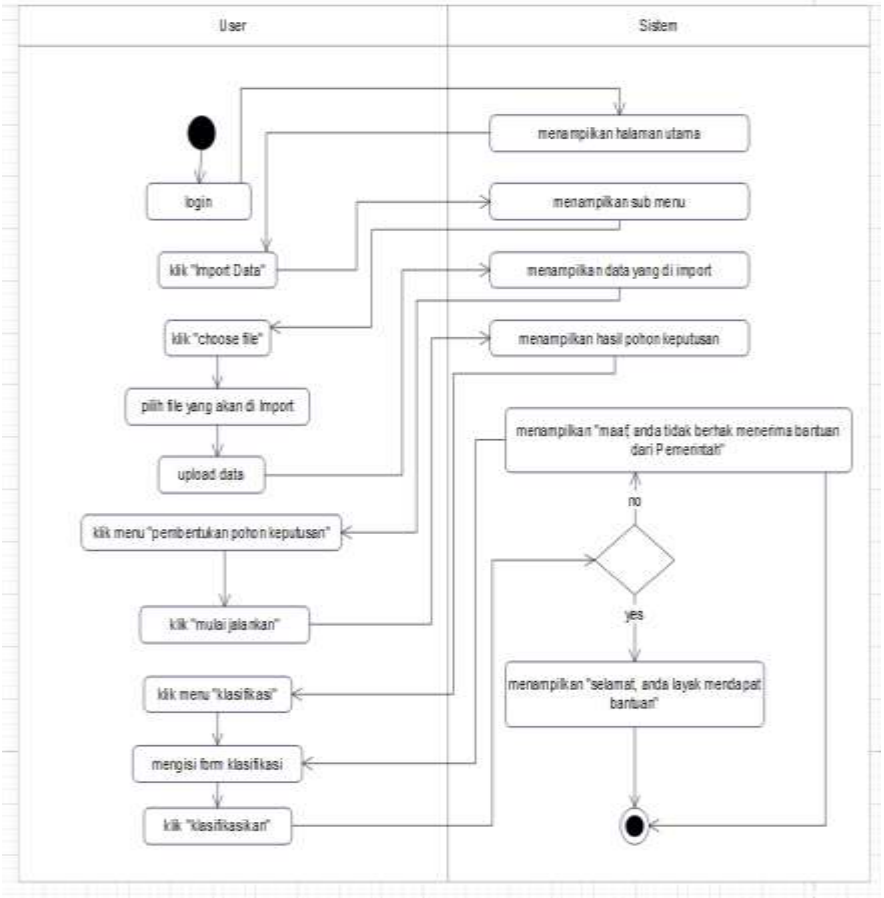

Figure 3. Activity Diagram

4. Class Diagram: Defines object relationships between users, data imports, decision tree models, and classification results.

Interface Design

The user interface includes:

1. Login Page – for user authentication.
2. Home Menu – displaying the main dashboard.
3. Import Data Menu – to upload Excel data files.
4. Decision Tree Menu – displaying generated tree structures.
5. Classification Menu – allowing users to input classification data and view eligibility results.

Results

System Implementation

The system was developed using XAMPP, EdrawMax, and a web-based PHP-MySQL framework. Users can upload datasets, generate decision trees, and classify new applicants. The interface includes pages for login, main dashboard, data import, decision tree visualization, and classification results.

For example, after uploading data, the system automatically generates a decision tree showing patterns that determine eligibility. The classification form produces one of two outcomes: "Eligible for Assistance" , "Not Eligible for Assistance."

System Testing

Testing was conducted using 500 household records validated by the Department of Housing and Settlement. Classification rules were developed for attributes such as income, roof type, flooring, wall material, citizenship status, family role, home ownership, and prior aid status.

Table 1. Monthly Income Classification

No.	Monthly Income (IDR)	Rank
1	< 800,000	4
2	800,000–1,600,000	3
3	1,600,000–2,400,000	2
4	> 2,400,000	1

Similar classification tables were applied for roof, floor, wall, family status, and home ownership attributes, each ranked from 1 (highest quality) to 4 (lowest quality). Testing confirmed that the algorithm accurately classified eligibility based on attribute combinations. For instance:

1. Applicants with low income, wooden walls, and rented homes were classified as eligible.
2. Applicants with higher income and permanent housing materials were classified as ineligible.

Discussion

The implementation of the C4.5 Decision Tree algorithm in this study proved highly effective in classifying and predicting eligibility for social housing assistance. The algorithm's structured logic allows for a systematic evaluation of multiple socioeconomic attributes, translating complex datasets into interpretable decision rules. By constructing a hierarchical decision tree, the model visualizes the flow of decisions from the root node to the terminal leaves, showing precisely how each attribute contributes to the final eligibility classification. This visual representation is crucial not only from a computational standpoint but also from an administrative and ethical perspective, as it promotes transparency, traceability, and accountability in government decision-making.

The analysis results indicate that several attributes play a dominant role in determining eligibility, including monthly income, citizenship status, family role, and previous aid status.

Among these, monthly income emerged as the most influential factor, which aligns with the government's welfare policy emphasizing support for low-income households. Citizenship status ensures that assistance is directed to legal residents in accordance with regulatory requirements, while family role and prior aid status function as fairness indicators to prevent duplicate or overlapping benefits. This outcome reinforces the idea that algorithmic decision models, when appropriately designed, can reflect policy priorities and enhance the integrity of welfare distribution mechanisms.

The resulting decision tree model demonstrates high interpretability, making it an effective decision-support tool for both technical staff and policymakers. Each branch of the tree represents a rule derived from empirical data—such as "If monthly income ≤ Rp 1,000,000 and no prior aid, then eligible"—which can be easily understood without specialized data science expertise. This interpretability distinguishes the C4.5 algorithm from more opaque methods like neural networks, thereby fostering trust and acceptance among government stakeholders and the public. Transparency is particularly important in social welfare contexts, where public scrutiny is high and perceptions of fairness directly influence institutional credibility.

Furthermore, the integration of the C4.5-based Decision Support System (DSS) significantly improves the operational efficiency of the Department of Housing and Settlement of Empat Lawang Regency. The automation of data processing and classification reduces manual workloads that previously required extensive human labor, especially during data validation and verification stages. By minimizing manual intervention, the system helps reduce human bias, ensuring that eligibility decisions are made based on objective, quantifiable criteria rather than subjective interpretation. In addition, the DSS accelerates the overall verification process, allowing aid distribution to be completed in a timelier and more consistent manner. This automation aligns with modern e-government principles emphasizing efficiency, accountability, and data-driven governance.

The findings of this research are consistent with the results of prior studies. Yunita et al. (2019) demonstrated that the Simple Additive Weighting (SAW) method improved fairness in determining housing assistance eligibility through quantitative weighting of decision criteria. Similarly, Swastina (2018) showed that the C4.5 algorithm achieved a high level of accuracy—up to 93.31%—in classifying educational data, confirming the algorithm's reliability for structured classification tasks. In both cases, data mining methodologies enhanced the accuracy, transparency, and consistency of decisions in domains characterized by complex, multidimensional datasets. The alignment of these studies with the present findings supports the validity of using C4.5 Decision Trees in social policy analytics.

Beyond its immediate application, the successful use of the C4.5 algorithm in this study has broader implications for the digital transformation of local governance. It illustrates how algorithmic tools can support public institutions in making data-driven, auditable, and equitable decisions, especially in resource-sensitive programs such as social housing. As local governments increasingly adopt digital technologies, integrating data mining into decision-making processes can enhance policy precision, reduce administrative inefficiencies, and strengthen public trust. In the long term, such systems could be expanded to include predictive analytics for proactive policy planning, thereby enabling more responsive and evidence-based public service delivery in Empat Lawang and other regions.

Conclusion and Recommendations

.

Based on analysis and implementation results, the following conclusions were drawn:

1. The C4.5 Decision Tree Algorithm was successfully implemented in a web-based system to assess eligibility for social housing assistance in Empat Lawang Regency.
2. The model achieved reliable accuracy, making it suitable as a decision-support tool for selecting social aid recipients.
3. The most influential attributes in classification include monthly income, citizenship status, family role, and previous aid history.
4. The system processes verified Excel-based datasets, providing efficient and transparent decision-making.

Disclosure Statement

The authors declare no conflicts of interest related to this research.

Acknowledgments

The authors would like to thank Allah SWT for His blessings, their parents for continuous support, Mr. Firamon Syakti as research advisor, and the staff of Universitas Bina Darma, Palembang, for their contributions and encouragement throughout this study.

References

Hartanto Kamagi, D., & Hansun, S. (2014). Implementation of Data Mining with the C4.5 Algorithm to Predict Student Graduation Rates. Journal of Informatics Engineering, Tangerang, Indonesia.

Surya, Y. (2017). The Effect of Government Assistance, Experience, and Business Environment on Fishermen Productivity in Sinar Mutiara Samudra Cooperative, Swarangan Village. Journal of Economic Studies, 17(3), 466–486.

Yunita, P., et al. (2019). Decision Support System for Selecting Home Renovation Assistance in Purnama Village Using the SAW Method.

Pembudi, S., et al. (2018). Application of the C4.5 Algorithm to Predict Student Graduation Scores. Journal of Information Technology and Computer Science Development, 2639–2640.

Sekardiana, D., et al. (2017). Implementation of Data Mining for Creditworthiness Determination Using the C4.5 Algorithm. Journal of Satya Negara Indonesia University, 46.

Putra, Y. B. A. (2018). Design of Scholarship Registration and Selection Application Using the SAW Method. Institute of Business and Informatics Surabaya.

Azizah, N. (2019). Decision-Making for Non-Habitable Housing Aid Selection Using the Naïve Bayes Algorithm. Pelita Bangsa University, Bekasi.

Hijriana, N., & Rasyidan, M. (2017). Application of the C4.5 Decision Tree Algorithm for University Scholarship Selection. Al-Ulum Science and Technology Journal, 3(1).

Alfiansyah, I. (2014). Implementation of Data Mining with the C4.5 Algorithm to Predict Elective Course Enrollment. STMIK AMIKOM Yogyakarta.

Haqmanullah, R., et al. (2018). Application of the C4.5 Algorithm to Predict High School Student Graduation Based on External Factors. Journal of Information Technology and Computer Science Development, 2(7), 2637–2643.

Biographical Notes

PAHMI LADIANTO Undergraduate Student, Department of Informatics Engineering, Faculty of Computer Science, Universitas Bina Darma, Palembang. Research interests focus on data classification, social information systems, and application development using machine learning algorithms.