Analysis of the C4.5 Algorithm in Determining the Eligibility for Annual Employee Bonus Reception at the BMKG Regional Office of North Sumatra

Albertus Tua Simanullang, Khairul

Abstract

Determining the eligibility for annual employee bonuses is an important form of performance evaluation in human resource management. This study aims to analyse the application of the C4.5 algorithm in determining the eligibility for annual employee bonuses within the BMKG Regional Office of North Sumatra. The method used is quantitative research with a data mining approach, utilising employee data in the form of performance indicators, attendance, and project contributions as input attributes. The C4.5 algorithm is applied to build a decision tree that can predict employees eligible for the annual bonus. The research results show that the C4.5 algorithm is capable of producing a prediction model with a high accuracy level and providing clear decision rules for management in bonus decision-making. This research is expected to serve as a reference for organisations in enhancing the transparency and objectivity of employee bonus determination.

Keywords: *C4.5, Decision Tree, Annual Bonus, Data Mining, BMKG*

Albertus Tua Simanullang[1]
[1]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: albe7ua@gmail.com[1]

Khairul[2]
[2]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: khairul@dosen.pancabudi.ac.id[2]
2nd International Conference on Islamic Community Studies (ICICS)
Theme: History of Malay Civilisation and Islamic Human Capacity and Halal Hub in the Globalization Era
https://proceeding.pancabudi.ac.id/index.php/ICIE/index

Introduction

Employee performance evaluation is an important aspect of human resource management (HRM) aimed at improving productivity, motivation, and job satisfaction of employees. One common form of reward given by organisations is an annual bonus. Determining eligibility for receiving bonuses often becomes a challenge, especially when it involves many employees with various complex performance criteria. The manual approach to assessing bonus eligibility carries risks of subjectivity, inconsistency, and time consumption, which can lead to employee dissatisfaction and reduce organisational transparency, especially at the BMKG Regional Office in North Sumatra.

Managing human resources (HR) is one of the responsibilities and purposes of industrial management. The quality of management performance and the company depends on the success of management in managing human resources. Therefore, the challenge that management must face is how to build the best possible human resource management strategy. Competent human resource management is intended to enhance the contributions that workers can make towards achieving organisational goals [1]. According to Prawirosentono in [2], employee performance is viewed as the work produced by individuals or groups within an organization, following existing authority and responsibilities, without violating laws and in accordance with existing ethics and morals in order to achieve the organization's goals. It can also be explained that performance is the level of success an individual achieves in carrying out their work. There are also ways to improve employee performance in the review that the author highlighted, such as annual bonuses and skills.

Previous researchIt is possible to determine whether marketing candidates passed or failed by looking at the study of new marketing placements. Using an algorithm that is C4.5 is one of the issues that arise while testing the data mining classification approach to choose the algorithm to predict marketing feasibility. Following testing, the C4.5 algorithm yields an accuracy score of 91.10% and an AUC value of 0.921 with an Excellent Classification diagnosis level. In order to determine whether marketing placement is feasible, the C4.5 algorithm is an excellent tool to use. Connect a newly created node to a node that has a weight of 0 [3].

The accuracy of high school pupils' prediction skills is determined using the C4.5 algorithm. The characteristics that influence high school pupils' proficiency in mathematical studies are the feature selection parameters. The Decision Tree C4.5 algorithm can predict high school students' final grades with 60% accuracy, according to testing and analytical results (Rizky Haqmanullah 2018). This research is an application built using C4.5 that can support neonatal mortality risk analysis techniques. The decision tree that was built is still limited to four levels with a maximum of three branches for each node. Therefore, for the further development of the application, it is expected to be able to analyse more than four variables (levels). Classification modeling with the C4.5 Algorithm on Rapidminer obtained an accuracy of 95.19%. [4] [5].

In the digital and big data era, the application of data mining becomes a solution to enhance objectivity and efficiency in decision-making. The C4.5 algorithm is one of the data mining methods widely used to build decision trees that can predict classifications based on certain attributes. This algorithm has the ability to generate clear decision rules, handle numerical and categorical data, and manage attributes with missing values.

According to the research findings, the C4.5 algorithm with optimization utilizing the stratified sample + forward selection approach has a higher accuracy level (81.75%) than the C4.5 algorithm without optimization (80.23%). The study's conclusion is that the C4.5 algorithm with optimization, which makes use of forward selection and stratified sampling techniques, is more efficient and can overcome the drawbacks of the C4.5 algorithm without optimization. This study aims to predict the achievement of online learning with the C4.5 algorithm (Bentar Candra 2024) [6]. The basic C4.5 algorithm and the improved C4.5 algorithm with the forward selection method. The test results of the C4.5 without enhancement achieved an average accuracy of 83.33%, a class recall of 88.24% for true smooth, and 76.02% for true risk of congestion. The test results of the enhanced C4.5 achieved an average accuracy of 87.59%, categorised as good classification [7].

The environment of the BMKG Regional Office in North Sumatra has a considerable number

of employees with various performance indicators, ranging from attendance, achievement of work targets, to project contributions. The use of the C4.5 algorithm is expected to assist management in determining which employees are eligible for the annual bonus in a transparent, objective, and data-driven manner. This research aims to analyse the effectiveness of the C4.5 algorithm in the process of determining employee bonus eligibility, thereby providing a systematic decision-making guideline and reducing subjective bias.

The application of the C4.5 algorithm is one of the case-solving solutions frequently used in problem-solving within classification techniques, characterised by the process of determining the entropy value and gain value from the possibilities of each criterion that serves as the basis for decisions, followed by the ranking process of the decision results. The output of the C4.5 algorithm is in the form of a decision tree. A decision tree is a structure that can be used to transform data into a decision tree that will generate decision rules. In solving existing problems, it will be simpler and easier to work on, because the decision-making area that was previously complex and very global can be transformed into something simpler and more specific. By applying the C4.5 algorithm to manual calculations, the calculations will be easier, as it will eliminate unnecessary calculations and samples will be tested based only on certain criteria or classes.

Literature Review

2.1 Data Mining

According to [8], data mining is the process of discovering knowledge in a database or the analytical step of discovering knowledge in a database, referred to as KDD. Knowledge can take the form of data schemas or relationships (previously unknown) between valid data.

Data mining is a combination of many computer science disciplines defined as the process of discovering new patterns from a very large dataset, including methods such as slices of AI (artificial intelligence), machine learning, statistics, and database systems. Data mining aims to extract (take the essence of) knowledge from a dataset to obtain a structure that can be understood by humans, including database and data management, data processing, model consideration and reasoning, interest measurement, complexity consideration, discovery structure postprocessing, visualisation, and online updating.

According to [9], "Generally, there are 5 (five) roles in data mining, namely estimation, prediction, classification, clustering, and association." When examined more closely, there are differences among the five roles mentioned, namely:

a. Estimation It is the process of guessing an unknown value based on the characteristics and attributes of a group of data.
b. Prediction It is the process of analysing historical data and forecasting what may happen in the future based on patterns and variables in the data.
c. Classification It is the process of grouping data into several classes or categories based on certain characteristics or attributes.
d. Clustering It is the process of grouping data into several similar groups based on certain characteristics or attributes.
e. Association It is the process of finding relationships between objects in the data, such as products that are often bought together in a store.

There are several stages involved in the data mining process, including [10]:

a. Data selection This is the process of selecting the data that will be used in the data mining process.
b. Data preprocessing Performing data preprocessing such as cleaning data from noise and outliers, filling in missing values, and transforming data.
c. Modelling Building a model to find patterns or information from the data.
d. Model evaluation Evaluating the built model to see how well it finds patterns or information from the data.
e. Result interpretation Interpreting the results of the data mining process and extracting new knowledge from the data.

2.2 C4.5 Algorithm

The C4.5 algorithm is a well-known and widely used algorithm for classifying data with numerical and categorical attributes. The results of the classification process in the form of rules can be used to predict the values of discrete-type attributes from new records. [11] In general, the C4.5 algorithm for building decisions is as follows:

1. Select an attribute as the root node
2. Create branches for each value 3. Split cases within the branches
3. Repeat the process for each branch until all cases in the branch have the same class

To calculate the gain value, the formula is used as shown in the following equation:

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

Explanation: S = Set of Cases A = Attribute n = Number of partitions of attribute A | = Number of cases in the i-th partition |S| = Number of cases in S

To calculate the value of Entropy, the formula used is as follows:

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

Explanation: S = Set of cases n = Number of partitions of S pi = Proportion of Si to S

The C4.5 algorithm is one of the classification algorithms in data mining, which uses a decision tree model. The C4.5 algorithm is the successor to ID3 (Iterative Dichotomiser) which adopts a greedy/non-backtracking approach where the decision tree is built top-down, recursively, and using a divide and conquer method [12]. The C4.5 algorithm has the same basic working principle as the ID3 algorithm, but has several differences, namely: Can handle discrete and numeric/continuous attributes, can handle training data with missing values, the decision tree results are pruned after formation, attribute selection is done using gain ratio. The C4.5 algorithm uses the pessimistic pruning method, which prunes subtrees based on the error rate of the training dataset and does not require a prune set. In selecting the default class, C4.5 chooses the class with the fewest training tuples.

2.3 Understanding Employees

According to Hasibuan (2007), an employee is anyone who works by selling their labour (physical and mental) to the company and receives compensation in accordance with the agreement. Meanwhile, according to Widjaja, A (2006), employees are the labour force of both physical and mental (spiritual and intellectual) human beings, which are always needed and thus become one of the essential capitals in cooperative efforts to achieve specific goals (organisations). Next, employees are individuals who work within a specific organization, whether in government institutions or business entities. Meanwhile, according to the Indonesian dictionary, an employee is a person who works in an institution (office, company) and receives a salary (wage).

According to Musanef (1984), employees as workers are those who are directly driven by a superior to act as executors who will carry out the work, thereby producing the expected outputs in the effort to achieve the established organisational goals. According to Suharno (2008), an employee is someone assigned as a worker in a company to carry out the company's operations. They work for a salary and serve as the main driving force of every organization. Without them, the organization and other resources would never mean anything. These factors greatly influence work productivity, work quality, work discipline, and employee loyalty to the company.

2.4 Bonus Theory

A bonus is an additional remuneration given to employees as recognition of their exemplary performance, which also benefits the organization. Bonuses can be given to employees as recognition of their commitment to the organization or to achieve specific company goals. In general, monetary compensation is used as a means to enhance employee loyalty and motivation

(Hanafi, 2019). Despite its diverse nature, bonuses are usually set within annual variable compensation plans that depend on the fulfilment of previously established goals or performance targets (Haussmann, 2016).

The provision of bonuses serves as a means to encourage personnel to contribute to achieving higher organisational profits (Rachmat, Halim, 2019). Employees are encouraged to improve the quality of products and services through the implementation of a bonus system (Bugdol & Jedynak, 2020). Efforts to motivate employees can be disrupted if incentives are reduced (Kroos et al., 2018). The purpose of this bonus is to encourage individuals or groups to exceed the predetermined goals and to acknowledge their extraordinary efforts and results. (Amelia GR, 2022

Court, Gwinner, and Meuter state that loyalty bonuses are given to employees who have shown sustained dedication to the organization. As a form of appreciation for the employees' tenure with the organization, these bonuses are often calculated after several years of service. These bonuses can be given on various occasions, including the end of the fiscal year, the organization's anniversary, or the employee's own birthday (Meuter, Gwinner, & Court, 2018, p. 27).

Research Methodology

This research uses a quantitative approach with experimental methods and data mining analysis. The objective of this approach is to analyse the capability of the C4.5 algorithm in determining the eligibility for annual employee bonuses based on performance and administrative attribute data within the BMKG Regional Office of North Sumatra. The quantitative approach was chosen because this research focuses on processing numerical and categorical data to obtain objective decision patterns.

The data source for this research is secondary, obtained directly from the BMKG Regional Office of North Sumatra. The data used includes employee performance data, such as achievement scores, attendance rates, contribution to organisational targets, as well as administrative data, including rank, length of service, and position. Additionally, historical data on annual employee bonuses is also collected as the target variable to be classified. The available data consists of numerical and categorical attributes, making it suitable for analysis using the C4.5 algorithm, which is capable of handling both types of data.

Data analysis is conducted using the C4.5 algorithm, which operates on the principles of entropy and information gain to form a decision tree. The cleaned dataset is divided into a training set and a testing set with a 70:30 ratio. The algorithm then builds a decision tree model based on the available attributes, and the model is tested to predict bonus eligibility on the testing data. Model evaluation is conducted using standard classification metrics, such as accuracy, precision, recall, and F1-score, as well as the confusion matrix to understand the classification performance in more detail.

To ensure the stability and generalisation capability of the model, this study uses 10-fold cross-validation, which allows the entire data to be tested and validated alternately. The results obtained were analysed to identify the most influential attributes in determining the eligibility for receiving bonuses. Finally, all findings are used to provide evidence-based recommendations and conclusions regarding the effectiveness of the C4.5 algorithm in the context of annual employee bonus decision-making.

Results

In order to improve the quality and loyalty of employees at the BMKG Regional Office of North Sumatra, the company leadership awards bonuses to outstanding employees each year based on established criteria. Due to the numerous criteria and employee data that need to be analysed, a company leader must work hard to determine which employees are eligible to receive

this annual bonus. Therefore, further data processing is needed to find new information/knowledge that can be useful as an aid in decision-making.

Therefore, the formation of a decision pattern is very necessary as it can avoid such problems. The application of data mining is considered suitable to solve this problem; by digging into the existing data, new knowledge can be obtained which can be used as a decision pattern that will later be used to analyse the eligibility of employees for the annual bonus. The application of the C4.5 algorithm to perform calculations with existing data will form a decision tree. From this decision tree, new knowledge in the form of rules or regulations will be derived, which will be used to create a decision pattern. This algorithm is considered quite easy to implement, so if there are changes in the data in the future, it will also be easy to update the new decision patterns.

Table 1. Employee Performance Data

No	Work Period	Presence	Loyalty	Work Results
1	10 year	80%	65%	90
2	5 year	80%	65%	90
3	7 year	75%	75%	90
4	3 moon	85%	80%	60
5	3 year	85%	80%	70
6	10 moon	65%	70%	70
7	2 year	77%	80%	65
8	8 year	65%	80%	65
9	5 moon	65%	70%	80
10	9 moon	70%	85%	70
11	2 year	70%	60%	75
12	3 year	70%	60%	75
13	9 year	80%	70%	90
14	6 year	80%	65%	80
15	4 year	65%	60%	65
16	2 year	65%	65%	65
17	11 moon	80%	85%	70
18	5 moon	70%	75%	70
19	3 year	70%	75%	75
20	2 year	80%	80%	85
21	10 moon	70%	65%	60
22	5 year	70%	65%	65
23	8 year	80%	75%	80
24	7 year	70%	85%	85
25	8 moon	70%	75%	70
26	5 year	65%	65%	75
27	10 year	65%	85%	80
28	5 year	75%	65%	65
29	8 moon	80%	75%	80
30	8 year	85%	70%	85

31	7 moon	70%	65%	70
32	7 moon	80%	80%	70
33	3 year	65%	80%	85
34	5 year	75%	80%	85
35	3 year	80%	70%	75

From the table above, it can be explained that the data obtained is the annual processed data, which is the performance data of all employees at the BMKG Regional Office of North Sumatra due to data limitations, capabilities, and time, as well as the personal nature of the data. At this stage, data cleaning is not performed because the data is considered ready for use. Next, the data transformation process will be carried out, where the data is converted into a format suitable for processing in data mining so that the data is ready to be calculated using the C4.5 algorithm. Below is the table of the results from the employee performance data transformation. Total Entropy is calculating the value of Eligible and Ineligible Results.

Given:

Total Cases = 35

Suitable Results = 9

Unsuitable Results = 26

Entropy (Total) $= \left(-\frac{9}{35} * \log_2\left(\frac{9}{35}\right)\right) + \left(-\frac{26}{35} * \log_2\left(\frac{26}{35}\right)\right)$

Entropy (Total) $= 0.822404$

Entropy of the Work Tenure Attribute consists of 2 values, namely: MS and TS, where the entropy values are as follows:

a. Work Tenure = MS
   Given: Number of MS cases = 24
   Eligible Results = 9
   Ineligible Results = 15

Entropy (MS) $= \left(-\frac{9}{24} * \log_2\left(\frac{9}{24}\right)\right) + \left(-\frac{15}{24} * \log_2\left(\frac{15}{24}\right)\right)$

Entropy (MS) $= \mathbf{0.954434}$

b. Working Period = TS
   Given: Number of TS cases = 11
   Eligible Results = 0
   Ineligible Results = 11

Entropy (TS) $= \left(-\frac{0}{11} * \log_2\left(\frac{0}{11}\right)\right) + \left(-\frac{11}{11} * \log_2\left(\frac{11}{11}\right)\right)$

Entropy (TS) $= \mathbf{0.000000}$

The Entropy of the Attendance Attribute consists of 2 values: GOOD and NOT GOOD, where the entropy values are as follows:

a. Attendance = GOOD
   Given: Number of GOOD cases = 17
   Eligible Results = 9 Ineligible Results = 8

Entropy (GOOD) $= \left(-\frac{9}{17} * \log_2\left(\frac{9}{17}\right)\right) + \left(-\frac{8}{17} * \log_2\left(\frac{8}{17}\right)\right)$

Entropy (GOOD) $= 0.997503$

b. Attendance = NOT GOOD
   Given: Number of NOT GOOD cases = 18
   Eligible Results = 0
   Ineligible Results = 18

Entropy (NOT GOOD) $= \left(-\frac{0}{18} * \log_2\left(\frac{0}{18}\right)\right) + \left(-\frac{18}{18} * \log_2\left(\frac{18}{18}\right)\right)$

Entropy (NOT GOOD) $= 0.000000$

The Entropy of the Loyalty Attribute consists of 3 values: GOOD, FAIR, and POOR, where

the entropy values are as follows:

a. Loyalty = GOOD
   Given: Number of GOOD cases = 12
   Acceptable Outcome = 2
   Unacceptable Outcome = 10

Entropy (GOOD) $= \left(-\frac{2}{12} * \log_2\left(\frac{2}{12}\right)\right) + \left(-\frac{10}{12} * \log_2\left(\frac{10}{12}\right)\right)$

Entropy (GOOD) $= 0.650022$

b. Loyalty = SUFFICIENT
   Given: Total cases SUFFICIENT = 11
   Eligible results = 4
   Ineligible results = 7

Entropy (SUFFICIENT) $= \left(-\frac{4}{11} * \log_2\left(\frac{4}{11}\right)\right) + \left(-\frac{7}{11} * \log_2\left(\frac{7}{11}\right)\right)$

Entropy (SUFFICIENT) $= 0.945660$

c. Loyalty = LOW
   Given: Number of LOW cases = 12
   Eligible results = 3
   neligible results = 9

Entropy (LOW) $= \left(-\frac{3}{12} * \log_2\left(\frac{3}{12}\right)\right) + \left(-\frac{9}{12} * \log_2\left(\frac{9}{12}\right)\right)$

Entropy (LOW) $= 0.811278$

The Entropy of the Work Result Attribute consists of 3 values: GOOD, FAIR, and POOR, where the entropy values for each are as follows:

a. Work Result = GOOD
   Given: Number of GOOD cases = 14
   Eligible Results = 9
   Ineligible Results = 5

Entropy (GOOD) $= \left(-\frac{9}{14} * \log_2\left(\frac{9}{14}\right)\right) + \left(-\frac{5}{14} * \log_2\left(\frac{5}{14}\right)\right)$

Entropy (GOOD) $= 0.940286$

b. Work Result = SUFFICIENT
   Given: Number of SUFFICIENT cases = 13
   Eligible Result = 0
   Ineligible Result = 13

Entropy (SUFFICIENT) $= \left(-\frac{0}{13} * \log_2\left(\frac{0}{13}\right)\right) + \left(-\frac{13}{13} * \log_2\left(\frac{13}{13}\right)\right)$

Entropy (SUFFICIENT) $= 0.000000$

c. Work Result = INSUFFICIENT
   Given: Number of INSUFFICIENT cases = 8
   Eligible Result = 0
   Ineligible Result = 8

Entropy (INSUFFICIENT) $= \left(-\frac{0}{8} * \log_2\left(\frac{0}{8}\right)\right) + \left(-\frac{8}{8} * \log_2\left(\frac{8}{8}\right)\right)$

Entropy (INSUFFICIENT) $= 0.000000$

Here is the calculation of the gain value:

1. *Gain* (Total, Work Period)

   Gain (Total, Work Period) $= 0.822404 - \left(\frac{24}{35} * 0.954434\right) + \left(-\frac{11}{35} * 0.000000\right)$

   Gain (Total, Work Period) $= 0.167935$

2. *Gain* (Total, Work Period)

Gain (Total, Presence) $= 0.822404 - \left(\frac{17}{35} * 0.997503\right) + \left(-\frac{18}{35} * 0.000000\right)$

Gain (Total, Presence) $= 0.337903$

3. *Gain* (Total, Loyalty)

Gain (Total, Loyalty) $= 0.822404 - \left(\frac{12}{35} * 0.650022\right) + \left(-\frac{11}{35} * 0.945660\right) +$
$\left(-\frac{12}{35} * 0.811278\right)$

Gain (Total, Loyalty) $= 0.024179$

4. *Gain* (Total, Work results)

Gain (Total, Work results) $= 0.822404 - \left(\frac{14}{35} * 0.940286\right) + \left(-\frac{13}{35} * 0.000000\right) + \left(-\frac{8}{35} * 0.000000\right)$

Gain (Total, Work results) $= 0.446290$

Conclusion

Based on the research results that have been conducted, several conclusions can be drawn as follows. First, the C4.5 algorithm has proven capable of modelling the eligibility decision for annual employee bonuses with a fairly high level of accuracy. The decision tree generated is capable of identifying the most influential performance and administrative attributes, such as achievement scores, attendance rates, length of service, and positions, thereby facilitating the classification process objectively.

Second, this research shows that employee performance attributes play a dominant role in determining bonus eligibility, while administrative attributes provide additional contributions to strengthen the decision. This emphasises that annual bonus decision-making should consider a proportional combination of performance achievements and administrative factors.

Third, the application of cross-validation and evaluation of classification metrics (accuracy, precision, recall, and F1-score) indicates that the model built using C4.5 is stable and has good generalisation capability to new data. Thus, this algorithm can be used as an effective and transparent decision-making tool in human resource management, particularly in the context of annual bonus distribution. Finally, this research emphasises that the utilisation of data mining methods based on the C4.5 algorithm can enhance efficiency, objectivity, and transparency in the decision-making process, and serve as a foundation for the development of more intelligent human resource management information systems in the future.

References

[1] L. H. Alfajar, "No Title," 2014.

[2] E. Fitriani, R. Aryanti, A. Saepudin, and D. Ardiansyah, "Penerapan Algoritma C4 . 5 Untuk Klasifikasi Penempatan Tenaga Marketing," vol. 22, no. 1, pp. 72–78, 2020.

[3] B. Identifikasi, "p-ISSN : 2579-5201 (Print) PERANCANGAN DAN IMPLEMENTASI ALGORITMA C4 . 5 UNTUK DATA MINING p-ISSN : 2579-5201 (Print)," vol. 3, no. 1, pp. 29–44, 2019.

[4] P. Mata, P. Matematika, D. Z. Azhari, I. S. Damanik, and D. Suhendro, "Penerapan Algoritma C4 . 5 Untuk Klasifikasi Tingkat Pemahaman Siswa," vol. 1, no. 1, pp. 11–20, 2022.

[5] M. Azis, H. Kurnia, P. Kartika, and D. Fanny, "Implementasi Algoritma C4 . 5 Untuk Memprediksi Capaian Pembelajaran Daring (Studi Kasus Siswa MAN 3 Blitar)," vol. 3, no. 1, 2022.

[6] A. Yani, F. Ramadhan, D. Irawan, and A. Wasid, "Implementasi Algoritma C4 . 5 Melalui Pohon Keputusan (Decision Tree) berbasis Metode Forward Selection Untuk Memprediksi Risiko Kredit Macet," vol. 9, no. 4, pp. 1425–1436, 2025.

[7] S. Alhadi, A. Supriyanto, and A. Pendahuluan, "SELF-REGULATED LEARNING CONCEPT :," pp. 333–342, 2017.

[8] A. Halim, "Application of Data Mining with the Least Square Method to Predict Web-Based Drug Inventory," vol. 5, no. 3, pp. 80–85, 2025.

[9] F. M. Sarimole and L. Nurmayanti, "Sistem Data Mining Penentuan Prioritas terhadap Penerima Bantuan Bencana Banjir dengan Metode Naive Bayes dan Klusterisasi K-Means

(Studi Kasus : Wilayah Cengkareng 2025) Jurnal Pengabdian Nasional (JPN) Indonesia," vol. 6, no. 3, pp. 685–697, 2025.

[10] P. Kesehatan, J. Ilmu, and K. Masyarakat, "HIGEIA : JOURNAL OF PUBLIC HEALTH," vol. 1, no. 1, pp. 1–7, 2017.

[11] T. Edition, *No Title*.