NEWTON: Networking and Information Technology
Vol. 4 No. 3 February 2025, Page.132 - 140
E-ISSN: 2797-0728

Implementation of data mining to predict BLT receipts in Kedungbetik village using the c4.5 algorithm

Gilang Dwi Saputra[1], Achmad Agus Athok Miftachuddin [2]
[1]Sistem Informasi, Universitas KH. A. Wahab Hasbullah
[2]Dosen Sistem Informasi, Universitas KH. A.Wahab Hasbullah
*Email: gilangdong@gmail.com

ABSTRACT
Receiving Direct Cash Assistance (BLT) is a government program to help poor people meet basic needs. This research aims to implement data mining techniques to predict BLT receipts in Kedungbetik Village using the C4.5 algorithm. The C4.5 algorithm was chosen because of its ability to build efficient and accurate decision trees. The data used includes attributes such as ID, name, address, type of work, BLT criteria and class. The data is analyzed to find patterns and relationships that are relevant to the BLT acceptance criteria. The research results show that the C4.5 algorithm can build accurate prediction models with a high success rate. This model is expected to help village governments identify residents who are entitled to receive BLT in a more targeted manner. This research contributes to the development of a more transparent and accountable BLT recipient selection method, and can be applied in other villages with similar characteristics. In this way, it is hoped that the distribution of BLT will be more even and effective, helping to reduce the level of poverty in society.

***Keywords**: Data Mining, C4.5 Algorithm, BLT Receipt, Kedungbetik Village, Decision Tree*

INTRODUCTION

Social assistance is assistance provided by local governments to the community in the form of funds or living necessities. The nature of this assistance is for a certain period of time, not continuously. This assistance is selected selectively in the form of currency, and distribution is adjusted to regional financial capabilities. The aim is to support local government programs and activities to make people's lives better and more prosperous by prioritizing the principles of justice, togetherness and the interests of the community.

However, of course the terms and conditions apply, and all the selection criteria for determining the community in obtaining social assistance are carried out based on existing data, then analyzed, but sometimes the results obtained are not appropriate. Thus, in this research the author uses the Data Mining Classification Algorithm, namely the C4.5 Algorithm, to predict how people will receive educational financial assistance. The data sample was taken from data on social assistance recipients in Kedungbetik Village, Kesamben subdistrict, Jombang district.

Prediction is a function that can find certain patterns from data. These patterns can be identified from various variables in the data. When you have found a pattern, the pattern obtained can be used to predict other variables whose value or type is not yet known.

The C4.5 algorithm is an algorithm used to form decision trees. Decision Tree is a very powerful and well-known method for classification and prediction. The Decision Tree method converts very large facts into a Decision Tree that represents rules. Decision Trees are also useful for exploring data, finding hidden relationships between several candidate input variables and target variables.

METHOD

A. Research methods

1. Planning

   Experimental methods can be used to study the effects of aid programs on aid recipients. In an experiment, a group of aid recipients is compared with a control group who did not receive aid to see differences in desired outcomes. Experiments can help understand the effectiveness of aid programs in increasing aid acceptance.

2. Analysis

The analysis stage in systems development involves several important activities to ensure the system can handle the identified problems:

a. Literature Study: Reviewing relevant literature to find similar cases that can provide insight into how to solve the problem.

b. Case Selection: Determining the most appropriate case to model based on relevance, level of difficulty, and potential solutions.

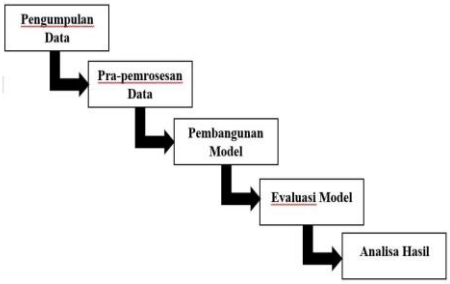
c. Problem, Opportunity, and Solution Classification: Groups problems, opportunities, and possible solutions to determine priorities and next steps.

d. Requirements Analysis: Performs an in-depth analysis of system requirements, including identification of key features, technical specifications, and limitations.

e. Defining System Requirements: Define system requirements clearly and in detail, including descriptions of functionality, performance, and standards that must be met.

B. Research design

This research uses a quantitative approach with data mining methods to predict receipt of aid. The method used is the C4.5 algorithm, which is one of the machine learning algorithms for building decision trees. This research design consists of several stages, namely:



Gambar 1 Alur Diagram

1. Data Collection: Collect relevant data from specified sources.
2. Data Pre-processing: Perform data transformation so that it is ready for analysis.
3. Model Building: Applying the C4.5 algorithm to build a prediction model.
4. Model Evaluation: Measure model performance and assess prediction accuracy.
5. Analysis of Results: Analyze the prediction results and summarize the findings from the research.

C. Data collection

The data source in this research was taken from [mention the data source, for example: "data provided by the Kedung Betik District Social Service", which includes information related to receipt of assistance. The data used includes the following attributes:

1. id: Unique identification for each entity.
2. name: Name of aid recipient.
3. address: The recipient's residence address.
4. job_type: Job type of the aid recipient.
5. blt_criteria: Criteria used to determine eligibility for assistance.
6. class: Category of aid receipt (e.g., accepted or not received).

Id	Nama penerima	Alamat	Jenis pekerjaan	Kriteria blt	kelas
1	SISWATI	Ngemprak	Petani pemilik lahan	1	Tidak
2	RAKIM	Ngemprak	Buruh Tani	3	Layak
3	MA'RUF SUKARNI	Guyangan	Petani pemilik lahan	1	Tidak

4	AKUP	Dero	Lainnya	1	Tidak
5	M THOYIBIN	Dero	Buruh tani	1	layak
6	KUSWATI	Kedungbetik	Lainnya	2	Tidak
7	AMANAH	Kedungbetik	Buruh Tani	3	Layak
8	SITI IKAYAROH	Kalanganyar	Lainnya	2	Tidak
9	KUSMIYATI	Kalanganyar	Buruh Tani	1	Layak
10	DARSE	Kedungmacan	Lainnya	2	Tidak
11	TUMINAH	Kedungmacan	Buruh Tani	3	Layak
12	LAMINAH	Sidowengku	Lainnya	2	Tidak
13	SUMINI	Sidowengku	Buruh Tani	3	Layak
14	MUSAMAH	Kandangsapi	Lainnya	1	Tidak
15	NUR MALIKI	Kandangsapi	Buruh Tani	3	Layak

Tabel 1 Data Penerima Bantuan

D. Pre-processing

data involves several important steps to prepare the data before analysis. These steps include:

a. Data Transformation: Transforming data into a format suitable for analysis, including encoding categorical variables and normalizing numeric data.

b. Data Splitting: Splitting data into a training set and a test set to build and test a model.

id	**Nama penerima**	**Alamat**	**Jenis pekerjaan**	**Kriteria blt**	**kelas**
1	SISWATI	Ngemprak	1	1	Tidak
2	RAKIM	Ngemprak	3	3	Layak
3	MA'RUF SUKARNI	Guyangan	1	1	Tidak
4	AKUP	Dero	99	1	Tidak
5	M THOYIBIN	Dero	3	1	layak
6	KUSWATI	Kedungbetik	99	2	Tidak
7	AMANAH	Kedungbetik	3	3	Layak
8	SITI IKAYAROH	Kalanganyar	99	2	Tidak
9	KUSMIYATI	Kalanganyar	3	1	Layak
10	DARSE	Kedungmacan	99	2	Tidak
11	TUMINAH	Kedungmacan	3	3	Layak
12	LAMINAH	Sidowengku	99	2	Tidak
13	SUMINI	Sidowengku	3	3	Layak
14	MUSAMAH	Kandangsapi	99	1	Tidak
15	NUR MALIKI	Kandangsapi	3	3	Layak

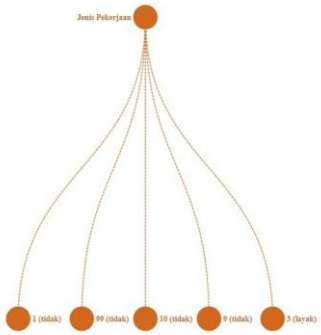
Tabel 2 Data Setelah Transformasi

E. Model building

The prediction model was built using the C4.5 algorithm, which is a decision tree algorithm that functions to classify data based on available attributes. Model building stages include:

a.Decision Tree Building: Applying the C4.5 algorithm to build a decision tree based on training data.

b. Determining Splitting Criteria: Using criteria such as entropy and information gain to determine the attributes to be used on each branch of the decision tree.



Gambar 2 Plotting pohon keputusan

F. Model Evaluation

Evaluasi Model evaluation aims to measure the performance of the model that has been built. The evaluation methods used include:

a. Evaluation Metrics: Use metrics such as accuracy to assess model performance.

b. Model Quality Test: Uses the test set to evaluate the accuracy of predictions and compare them with actual results.

G. Aanalysis of results

Analysis of the results was carried out to evaluate the effectiveness of the model in predicting aid receipt. Analysis steps include:

a. Decision Tree Interpretation: Examining the structure of a decision tree to understand the decisions made by the model.

b. Discussion of Findings: Compare predicted results with actual data and discuss the implications of the findings for policy or further action.

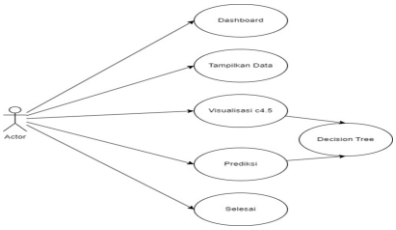
1. System design

a.System Design

In software system development, careful planning is essential to ensure that the software developed has quality documentation.

b. Designing Use Case Diagrams

Use Case Diagrams, or also known as behavior diagrams, are used to briefly describe the relationship between BLT acceptance cases, admin, and the system. This diagram also helps in understanding how the system works.

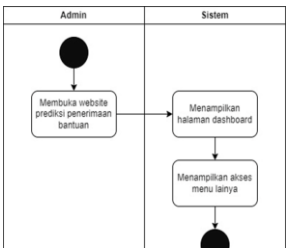


Gambar 3 Perancangan Use Diagram

c. Design Activity Diagram

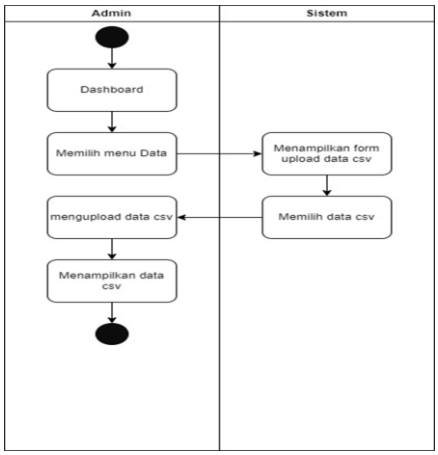
The activity diagram describes the flow of activities that occur in the BLT receipt system. Each activity can represent operations in progress and model the actions to be taken when an operation is executed, as well as the expected results.

1.Dashboard Display



Gambar 4 Tampilan dashboard

1. Show Data

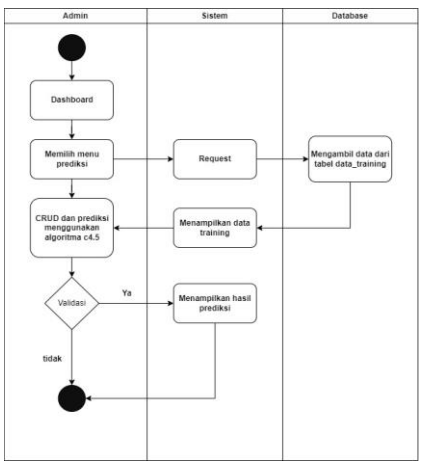


Gambar 5 Tampilkan Data

2. Visualisasi C4.5

Gambar 6 Visualisasi c4.5

3. Prediction



Gambar 7 Prediksi

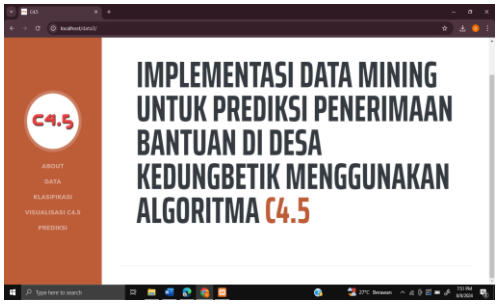
RESULT AND DISCUSSION

Through the process of gathering needs and system design, a Direct Cash Assistance (BLT) Receipt Classification System has been developed which utilizes the C4.5 Algorithm.

Result

a. Main page display

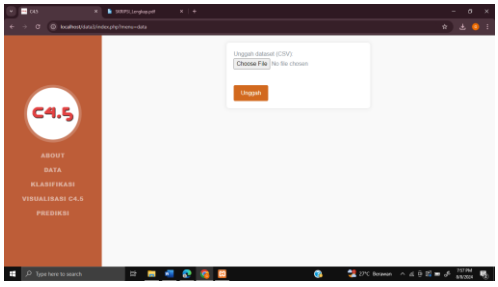
The main page displays the home or home page.



Gambar 8 Main page display

b. Display form page for inputting CSV data

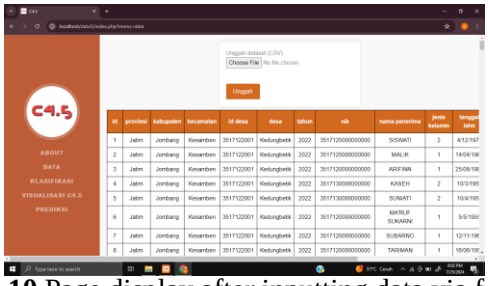
The form page display for inputting CSV data contains a form display for inputting the CSV data that you want to process.



Gambar 9 Display form page for inputting CSV data

c. Page display after inputting data via form

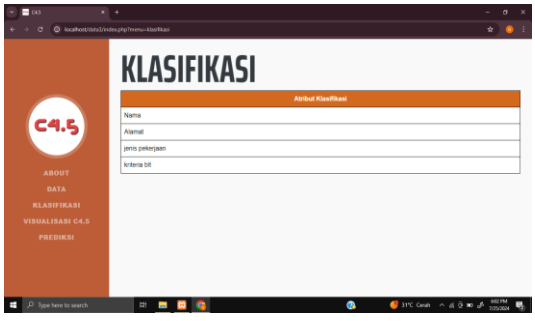
The page display after inputting data via this form contains information about the data that has been input via the form



Gambar 10 Page display after inputting data via form

d. Classification page display

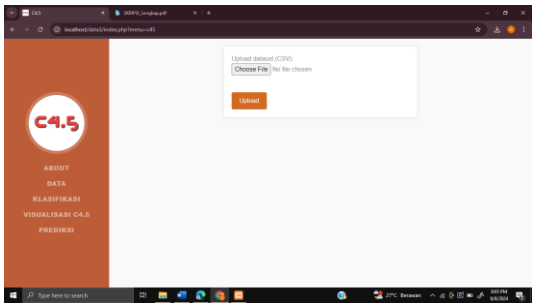
This classification page display contains information about the attributes of the data being classified, which can be seen in Figure 4.5



Gambar 11 Classification page display

e. The form page displays for inputting data that wants to go through the c4.5 algorithm calculation process

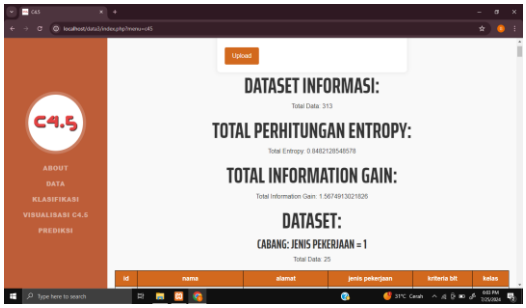
The form page display for inputting data that wants to go through the c4.5 algorithm calculation process contains a form display for inputting spreadsheet data that wants to be calculated using the c4.5 algorithm



Gambar 12 The form page displays for inputting data that wants to go through the c4.5 algorithm calculation process

f. algorithm calculation page display

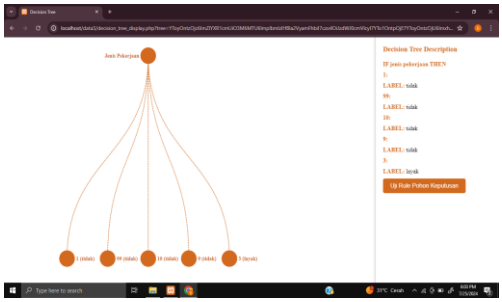
The c4.5 algorithm calculation page display contains the calculation results of the c4.5 algorithm



Gambar 13 algorithm calculation page display

g. visualization page display

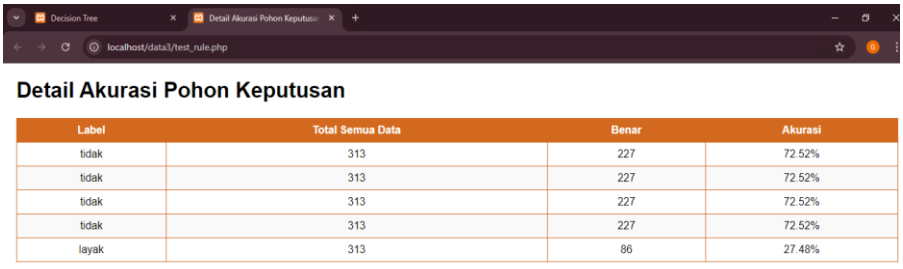
This display contains visualization information or decision trees from data that has gone through the c4.5 algorithm calculation process



Gambar 14 visualization page display

h. Decision tree rule test page display

This display contains the results of rule testing on the Decision tree that was built previously and for determining accuracy in Calculation and Construction of Decision Trees using the c4.5 algorithm

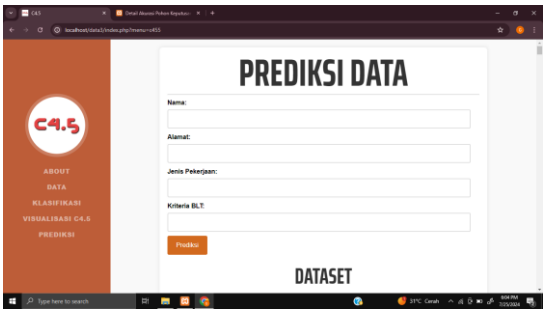


Label	Total Semua Data	Benar	Akurasi
tidak	313	227	72.52%
tidak	313	227	72.52%
tidak	313	227	72.52%
tidak	313	227	72.52%
layak	313	86	27.48%

Gambar 15 Decision tree rule test page display

i. Page views for c4.5 algorithm predictions

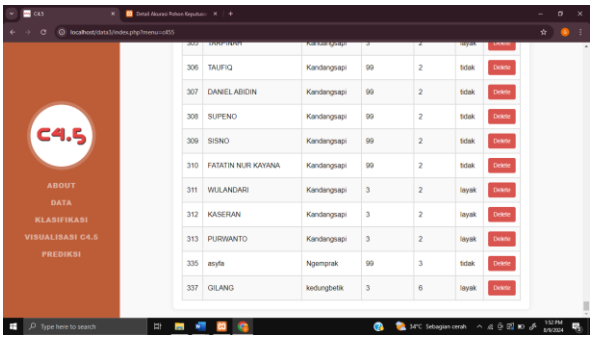
This page displays a form for filling in data where the data will be predicted and calculated using the c4.5 algorithm.



Gambar 16 Page views for c4.5 algorithm predictions

j. Page display of c4.5 algorithm prediction results

This page shows the predicted results of the c4.5 algorithm calculation results



Gambar 17 Page display of c4.5 algorithm prediction results

CONCLUSIONS

His research aims to implement the C4.5 algorithm in predicting the receipt of aid in Kedung Betik Village. Based on the results and analysis that have been carried out, several important points can be concluded as follows:

1. The C4.5 algorithm has proven to be effective in assisting the prediction process of receiving social assistance. This algorithm is able to process historical data and produce a classification model that is accurate enough to determine aid recipients based on predetermined criteria.

2. Important patterns generated by the C4.5 algorithm, such as type of work, are key variables in determining eligibility for receiving assistance. This pattern provides an objective basis for the decision making process.

3. The prediction system built through this research speeds up and simplifies the selection process for aid recipients, thereby reducing the workload of village officials or related agencies.

4. The use of a data mining-based system can increase transparency and fairness in the process of providing assistance, because decisions are made based on valid data and an automatic process without subjective intervention.

5. Implementation of the C4.5 algorithm in the case of receiving aid in Kedungbetik Village shows that data mining technology has great potential to be applied in various social policies, especially those that require accurate data classification and prediction.

REFERENCES

Gultom, J. P., & Rikki, A. (2020). Implementasi Data Mining menggunakan Algoritma C-45 pada Data Masyarakat Kecamatan Garoga untuk Menentukan Pola Penerima Beras Raskin. *Kumpul. Artik. Karya Ilm. Fak. Ilmu Komput*, *2*(01), 11-19.

Anggriani, S. D., Muhammad Syahril, S. E., Kom, M., & Ita Mariami, S. E. (2018). Data Mining Algoritma C4. 5 Untuk Menganalisa Penduduk Penerima Program Keluarga Harapan (PKH)(Studi Kasus: Kelurahan Tualang V). *Jurnal Cyber Tech*, *1*(6).

Ermawati, E. (2019). Algoritma Klasifikasi C4. 5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai. *Jurnal Sistem Informasi*, *8*(3), 513-528.

Widiyati, D. K., Wati, M., & Pakpahan, H. S. (2018). Penerapan Algoritma ID3 Decision Tree Pada Penentuan Penerima Program Bantuan Pemerintah Daerah di Kabupaten Kutai Kartanegara. *J. Rekayasa Teknol. Inf*, *2*(2), 125.

Purnama, A. I., Aziz, A., & Wiguna, A. S. (2020). Penerapan Data Mining untuk Mengklasifikasi Penerima Bantuan PKH Desa Wae Jare Menggunakan Metode Naïve Bayes. *Kurawal-Jurnal Teknologi, Informasi dan Industri*, *3*(2), 173-180.

Asroni, Respati, B. M., & Riyadi, S. (2018, November). Penerapan Algoritma C4.5 Untuk Klasifikasi Jenis Pekerjaan Alumni Di Universitas Muhammadiyah Yogyakarta. SEMESTA TEKNIKA, 21, 158-165.

Purnomo, C. E., & Rikendry, R. (2021). Penerapan Metode C4. 5 Untuk Klasifikasi Warga Miskin Pada Desa Mengandung Sari. *Jurnal Teknologi dan Sistem Informasi*, *2*(3), 14-25.

Susanto, A., & Setiawan, R. (2018). Penerapan Algoritma C4.5 untuk Prediksi Penerimaan Bantuan Sosial Berdasarkan Kriteria Ekonomi dan Demografi. Jurnal Teknologi Informasi, 10(2), 45-55.

Nurhadi, M., & Ramadhan, F. (2019). Penggunaan Data Mining dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Menggunakan Algoritma C4.5. Jurnal Sistem Informasi dan Komputerisasi, 15(1), 25-34.

Hidayat, A., & Pratama, D. (2020). Implementasi Algoritma C4.5 pada Sistem Prediksi Penerimaan Bantuan Pendidikan. Jurnal Informatika dan Komputerisasi, 12(3), 88-98.

Wahyuni, I., & Setiadi, R. (2021). Penggunaan Algoritma C4.5 untuk Klasifikasi Penerima Bantuan Kesejahteraan Sosial. Jurnal Ilmu Komputer dan Informatika, 14(2), 102-110.

Suryadi, E., & Novitasari, L. (2022). Model Prediksi Penerima Bantuan Berdasarkan Algoritma C4.5: Studi Kasus Program Keluarga Harapan. Jurnal Penelitian Teknologi Informasi dan Komputerisasi, 18(4), 120-130.

Yunus, M., Ramadhan, H., Aji, D. R., & Yulianto, A. (2021). Penerapan Metode Data Mining C4. 5 Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP). *Paradigma-Jurnal Komputer dan Informatika*, *23*(2).

Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, *8*(6), 219-225.

Girsang, R., Ginting, E. F., & Hutasuhut, M. (2022). Penerapan Algoritma C4. 5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, *1*(4), 449-459.

Susanto, W., & Mulyani, A. (2022). Analisa Algoritma C4. 5 Terhadap Penentuan Rekomendasi Penerima Beasiswa. *Oktal: Jurnal Ilmu Komputer Dan Sains*, *1*(10), 1607-1619.

Barito, E. E., Beng, J. T., & Arisandi, D. (2022). Penerapan Algoritma C4. 5 untuk Klasifikasi Mahasiswa Penerima Bantuan Sosial Covid-19. *Jurnal Ilmu Komputer dan Sistem Informasi*, *10*(1).