

Application of Data Mining with C5.0 Algorithm for Recommendations for Prosperous Family Card (KKS) Recipients

Nurul Fadilla^{1)*}, Ilka Zufria²⁾, M. Fakhri³⁾

^{1,2,3)} State Islamic University of North Sumatra Medan

¹⁾nurul.fadilla@uinsu.ac.id, ²⁾ilkazufria@uinsu.ac.id, ³⁾fakhri@uinsu.ac.id

ABSTRACT

Poverty is a social problem that still often occurs in various regions in Indonesia, including in Silau Laut District which consists of several villages such as Bangun Sari, Silo Bonto, Silo Lama, Lubuk Palas, and Silo Baru. Although the area is quite large, there are still many families who are classified as poor and unable to meet their basic needs. To overcome this, the government launched the Prosperous Family Card (KKS) program as a form of social assistance. However, the process of determining prospective KKS recipients still faces various obstacles, such as a random selection method based on data sent by each village to the central government. This raises concerns about the inaccuracy of the target in the distribution of aid, so that the aid is not received by the families who really need it. In addition, a lot of data has not been optimally utilized in the selection process. Therefore, this study aims to design a website-based information system that can assist Silau Laut Regency in providing recommendations for prospective recipients of KKS assistance by utilizing data mining techniques. The algorithm used is C5.0, because it is able to produce a decision tree with high accuracy, while the system development method used is Rapid Application Development (RAD) to speed up the system development process. The result of this research is an information system that can process community data and provide recommendations for prospective recipients of KKS assistance in a more objective and targeted manner in the next period.

Keywords: Prosperous Family Card (KKS), Poverty, Data Mining, C5.0, RAD.

INTRODUCTION

Indonesia is one of the poorest countries in the world. Based on the Central Statistics Agency (BPS), there are 25.14 million Indonesians who are classified as underprivileged. This is a problem that must be considered by the government, because the poverty rate in Indonesia is still relatively large and has not been completely resolved. KKS is one of the government's initiatives to accelerate poverty alleviation, which is regulated in Presidential Regulation (Perpres) Number 166 of 2014 concerning Poverty Alleviation and Presidential Instruction of the Republic of Indonesia Number 7 of 2014 concerning the implementation of the Prosperous Family Savings Program, Smart Indonesia Program, and Healthy Indonesia Program, with the aim of creating more productive families. (www.kominfo.go.id, accessed at 11.10 WIB on March 1, 2020).

Reporting from Online News, *Detik.com* on 01/03/2020 stated that the budget used in providing KKS assistance amounted to IDR 6.2 trillion and each family received IDR 200,000/month. The data collection in this program is based on the results of data collection from the previous program, namely PPP which was carried out in July 2013. This figure refers to the Integrated Data Base (BDT) managed by TNP2K (National Team for the Acceleration of Poverty Alleviation). In determining the recipients of KKS assistance in Silau Laut Regency, the selection system is carried out by each village sending data to the central government and the central government selects the recipients of assistance at random. This concern arises because it can cause confusion and inaccuracy of the target in determining aid recipients, so that KKS assistance is not distributed to people who really need help and the time needed to choose prospective KKS assistance recipients is 2-3 months. With the KKS program, it is hoped that it can maximize the socio-economic conditions of the community, and KKS is expected to not only be able to reduce the poverty rate in Indonesia but can also be one of the ways to break the poverty chain in Indonesia.

confusion and inaccuracy of the target in determining aid recipients, so that KKS assistance is not distributed to people who really need assistance and the time needed to select prospective KKS assistance recipients is 2-3 months. With the KKS program, it is hoped that it can maximize the socio-economic conditions of the community, and KKS is expected to not only be able to reduce the poverty rate in Indonesia but can also be one of the ways to break the poverty chain in Indonesia.

* Suitable author



Previous research was conducted by Muhammad Amirsyah Tanjung, Poningsih, and Hendry Qurniawan in their scientific paper entitled Analysis of the eligibility of Family Hope Program (PKH) recipients using the C4.5 algorithm. The study discussed PKH assistance recipients using the C4.5 algorithm where the algorithm produced 6 rules to determine the eligibility of PKH assistance recipients. In this study, the researcher used the C5.0 algorithm to develop from the C4.5 algorithm and also ID3, the advantages of the C5.0 algorithm can form rule simpler or more concise, the level of accuracy is maximized (Dewi et al., 2019). The purpose of this research is to recommend residents who are eligible for assistance by utilizing data techniques Mining and C5.0 algorithms and implementations by utilizing Web, is expected to be able to assist the agency in determining recommendations for prospective KKS recipients in the next period.

LITERATURE REVIEW

Data Mining

Data is a set of facts that are recorded without meaning. On the other hand, data can be understood as a collection of facts presented in various forms, such as letters, numbers, symbols, or characters, which are then processed to create information (Nofriansyah, 2019). According to (Heriyanto, 2018) Data is information that has not yet given meaning to the recipient and needs further processing. Data can be conditions, sounds, images, numbers, letters, mathematical symbols, language, or other symbols that can be used to describe a certain environment, event, object or concept.

Data Mining is defined as the process of extracting valuable information from large databases. One of the key aspects of data engineering Mining i.e. finding a pattern of large frequencies among the set of items, known as the association rule (Rahim et al., 2018). According to Feelders, Daniels, and Holsheimer in Gunadi (2008), Mining namely the stage of retrieving information from the dataset by utilizing algorithms and techniques that involve disciplines such as database management systems, statistics, machine learning, as well as mathematics. This process allows in generating crucial information as well as knowledge that is not seen in large databases. Data availability Mining, we can find a "treasure" in the form of knowledge contained in existing data (Zaman, 2016).

According to Pramudiono (2008) in Kusrini's book, Emha Taufiq Luthfi, entitled "Data Algorithm Mining" Defining data mining that is, the automatic analysis of complex or large data that aims to identify important patterns and trends that are often unaware of their existence (Kusrini, 2018).

C5.0 algorithm

The C5.0 algorithm is a classification algorithm in data mining which is specifically used in decision tree techniques. C5.0 is a development of the previous algorithm created by Ross Quinlan in 1987, namely ID3 and C4.5. In this algorithm, the selection of attributes is done using the gain ratio. For the selection of attributes that want to separate objects into classes, the attribute that produces the highest gain ratio will be designated as the parent for the next node. C5.0 results in a tree with a diverse total of branches per node. Physically, the difference is that the trees created by the C5.0 algorithm will be less when compared to the C4.5 algorithm. Therefore, the tree development process in this algorithm is not long compared to the C4.5 algorithm (Pratiwi et al., 2020).

The tree-building process in the C5.0 algorithm is in harmony with C4.5, including in terms of calculation of the frequency of events, entropy and information gain. However, the main difference is that C4.5 only stops at the calculation information gain, while C5.0 continues with counting gain ratio, information gain and entropy that have been calculated beforehand. C5.0 is a commercial development of C4.5 that is widely used in various software data mining like Clementine and RuleQuest. The test results showed that the C5.0 was more efficient in memory usage, even about 90% faster than the C4.5 (Wijaya et al., 2018).

Prosperous Family Card (KKS)

The Prosperous Family Card (KKS) is a government initiative to accelerate poverty alleviation, which is contained in Presidential Regulation (Perpres) Number 166 of 2014 concerning Poverty Alleviation and Presidential Instruction of the Republic of Indonesia Number 7 of 2014 related to the implementation of the Prosperous Family Savings Program, Smart Indonesia Program, and Healthy Indonesia Program, which aims to build a more active family. The KKS acts as a sign that cardholders are entitled to receive financial assistance from the government. With the KKS program, it is hoped that it can maximize the socio-economic conditions of the community, and KKS is expected to not only be able to reduce poverty rates in Indonesia but also be one of the ways to break the poverty chain in Indonesia (Muhazir et al., 2017).

* Suitable author



Unified Modeling Language (UML)

According to Windu and Grace "Unified Modeling Language (UML) is the basic language used in documenting, specifying, and developing software. UML functions as a method for the development of object-oriented systems as well as a tool in helping the process of system development" (Suendri, 2018). UML also offers a standard guideline for designing a system's framework, which includes business process concepts, class writing in a specific programming language, database structure, and elements required in software creation (Zufria, 2013).

Rapid Application Development (RAD)

Rapid Application Development (RAD) is a method to develop information systems in a shorter time. Usually, the development of a standard information system takes a minimum of 180 days, but by using the RAD method, a system can be completed in as little as 30-90 days. According to (Irnawati, 2018), *Fast App Development* (RAD) is a software development model that uses an incremental approach, primarily designed for projects with short completion durations. The (Trimahardhika, 2017) states that RAD is an object-focused approach to system development, encompassing development methods as well as the software tools used in the process.

Table 1. References for Similar Studies

Ye s	Name	Heading	Method	Debilitation	Excess
1	Muhammad Amirsyah Tanjung, Poningsih, Hendry Qurniawan	Eligibility of Family Hope Program (PKH) Recipients Using C4.5 Algorithm	Classification Method with C4.5 Algorithm	The results of the implementation of this research have not used a website-based information system so that the presentation of data is still classified as less dynamic.	This study used the classification method with al-goritma C4.5 and testing with the Rapidminer application to get maximum results.
2	Apriani Candra Wijaya, Nelly Astuti Hasibuan, Princess Ramadhani	Implementation of the C5.0 Algorithm in Classification Community Income (Case Study: Urban Village Medan Kota District Mosque)	Classification Method with C5.0 Algorithm	The results of the implementation of this research have not used a website-based information system	This study uses the C5.0 algorithm to form a decision tree in classifying people's income using 6 attributes and testing with the Weka application, getting maximum results.
3	Dyah Ayu Wahyuning Dewi, Priest Cholissodin, Sutrisno	Classification of Child Growth and Development Deviations Using the C5.0 Algorithm	Classification Method with C5.0 Algorithm	The results of the implementation of this research have not used a website-based information system	This study applied a C5.0 algorithm, which at the test stage produced the highest accuracy of 100%. The average accuracy of the test set showed a value of 95.99% in the testing of the total training data, 97.33% in the testing of the number of trials, and the comparison between the C4.5 and C5.0 algorithms showed an accuracy of 93.33% for C5.0, while the C4.5 reached 87.61%. These results show

					that C5.0 provides better accuracy compared to C4.5. One of the factors that affect the accuracy of the ad is the boosting of the use of C5.0, which contributes to the increase in accuracy
4	Khairul Zaman	Application of Data Mining Using C4.5 Algorithm to Determine the Eligibility of Recipients of Social Rehabilitation Assistance for Uninhabitable Houses (Case Study in South Solok Regency Government)	Classification Method with C4.5 Algorithm	The results of the implementation of this research have not used a website-based information system	This research utilizes the C4 Algorithm. 5 to identify the eligibility of social rehabilitation assistance for unsuitable housing, as well as to use Weka and RapidMiner soft tools that can produce accurate decisions in the process of sending messages

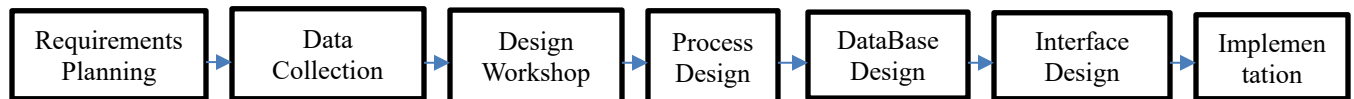
METHOD

Data Source

The data used in this study is in the form of observation, interview, and literature study data. This data was collected based on the number of residents who received KKS assistance in Silau Laut Regency from August 5 to August 20, 2020. The results of the observations obtained were in the form of data on residents who received KKS assistance in Bangun Sari Village, Lubuk Palas Village, and Silo Bonto Village, Silau Laut District for the last 2 years.

Data Analysis Techniques

The data analysis techniques used are in the form of research methods *Mixed Methods Research* which is a combination of qualitative and quantitative research, by applying the *Rapid Application Development* (RAD). The following are the steps in the development of the system that were carried out:



RESULT

Running System Analysis

The analysis of the current system aims to gain a deeper understanding of how the system works and the problems it faces, which can later be the basis for better system design. The following is an overview of the system currently implemented in Silau Laut District:

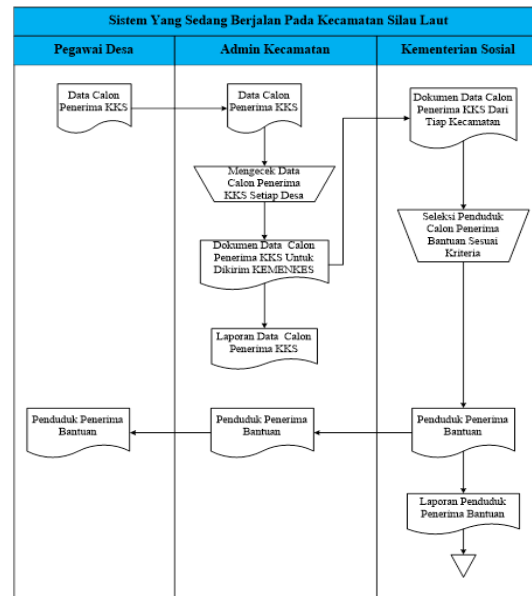


Figure 1. Running System in Silau Laut District

Proposal System Analysis

Based on the results of the analysis of the ongoing system, the author proposes to build a data *mining* system with the C5.0 algorithm, which is expected to be able to help predict prospective recipients of Prosperous Family Card (KKS) assistance in the next period. The following is a picture of the proposed system.

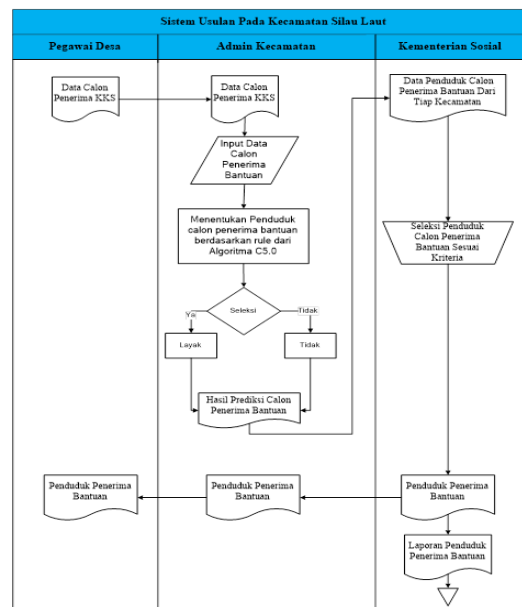


Figure 2. Proposal System in Silau Laut District

Results of the Calculation of the Determination of KKS Recipients with the C5.0 Algorithm

Below is a table of training data for Silau Laut Regency residents. This training data table will later be used to find the *value of entropy*, *gain*, and *gain* and form a pattern on the decision tree rules to recommend prospective KKS recipients in the next period.

Table 2. Sample Training Data

Not	Name of the Head of Family	Age	School Child's Ownership	Toddler Ownership	Work	Types of House Walls	Income	Information
1	A.Dahlan Yunan Tampubolon	36	1	0	Self employed	Wall	1800000	No
2	Abd Azis	64	2	0	Self employed	Kayu	1300000	No
3	WE. Jamil	62	2	0	Farmer	Tembok	1000000	No
4	WE. Yahmin	33	2	0	Self employed	Kayu	1000000	Proper
5	Snoop Dogg	36	3	0	Self employed	Tembok	2000000	No
6	Abdul Azis	46	3	0	Self employed	Tembok	2000000	Proper
7	Abdul Hakim	37	1	0	Self employed	Kayu	3000000	No
8	Abdul Hakim Pulungan	36	1	0	Self employed	Tembok	1500000	No
9	Abdul Karim	36	2	0	Private Tutor	Tembok	1500000	Proper
10	Adenan Saragi	54	1	0	Self employed	Tembok	2400000	No
11	Adi Sahputra	40	2	0	Merchant	Kayu	1500000	No
12	Ahmad	45	1	0	Self employed	Kayu	2000000	Proper
13	Ahmad Arifin Sitorus	46	1	0	Self employed	Tembok	2500000	No
.....								
.....								
99	Zulkkifli	55	2	0	Farmer	Tembok	2000000	No
100	Zulpadli	41	2	0	Self employed	Tembok	1500000	Proper

Calculation of Entropy and Gain in Data Mining

In the process of developing the prediction system for KKS assistance recipients, entropy and gain calculations are carried out to determine the best attributes as the separating nodes in the decision tree. The training data used consisted of 100 cases, with 40 cases being feasible and 60 not feasible.

Calculation of total entropy:

$$\text{Entropy (Total)} = \left(-\frac{40}{100} \times \log_2 \left(\frac{40}{100} \right) \right) + \left(-\frac{60}{100} \times \log_2 \left(\frac{60}{100} \right) \right)$$

$$\text{Entrophies (Total)} = 0.971$$

Calculation of Entropy attributes:

$$\text{Entropy (24-35)} = \left(-\frac{9}{21} \times \log_2 \left(\frac{9}{21} \right) \right) + \left(-\frac{12}{21} \times \log_2 \left(\frac{12}{21} \right) \right)$$

Gain value calculation:

Here's the calculation of the gain value of each attribute:

$$\text{Gain (Usia)} = 0,9710,985 + 0,994 + 0,811 = -\left(\frac{21}{100} \times \left(\frac{55}{100} \times \left(\frac{24}{100} \times 0.023 \right) \right) \right)$$

Calculation of gain and gain ratio in Data Mining

This document describes in detail the process of calculating the Gain and Gain Ratio for attributes used in the prediction system for KKS assistance recipients. This calculation aims to determine the best attribute as a separator node in the decision tree. The training data used consisted of 100 cases with 40 eligible cases and 60 unsuitable cases.

* Suitable author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

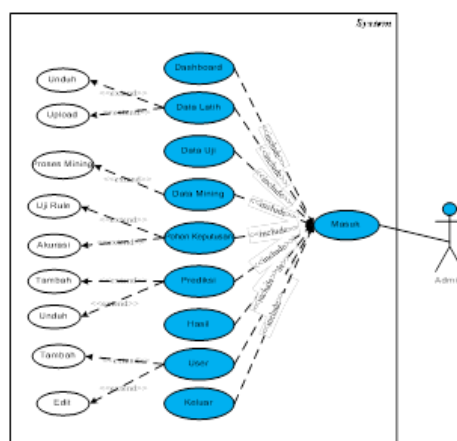
The total entropy is calculated as 0.971, and from the results of the Profit calculation, attributes such as age, occupation, and income show varying Gain values, with the age and occupation attributes having the highest Gain. This shows that these attributes are most relevant in the classification process.

This Gain and Gain Ratio calculation is important because it helps in selecting the most informative attributes, so that the decision tree built becomes more efficient and accurate. The attribute with the highest gain will be selected as the main node (root) in the decision tree, which is then used to determine potential beneficiaries automatically.

Workshop Design

Use Case Diagram

The use case diagram is made to identify the functions contained in the information system and the parties who have the right to access these functions. This diagram is used to illustrate the relationship between actors and the system to be developed (Fridayantie & Mahdiati, 2016).



Gambar 3. Use Case Diagram

Activity Diagram

After the stage *use case diagram*, the next step is to go to the *activity diagram* to describe the flow of activity of the system. *Activity diagram* aims to describe the workflow or activities that exist in the system (Fridayantie & Mahdiati, 2016). The following is an illustration of the *activity diagram* for the implementation of data mining on the prediction of KKS assistance recipients.

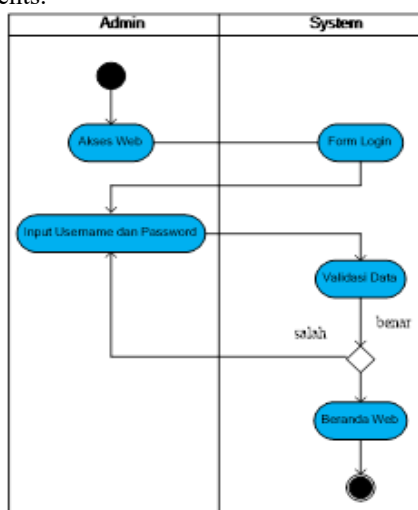


Figure 4. Activity Diagram Login

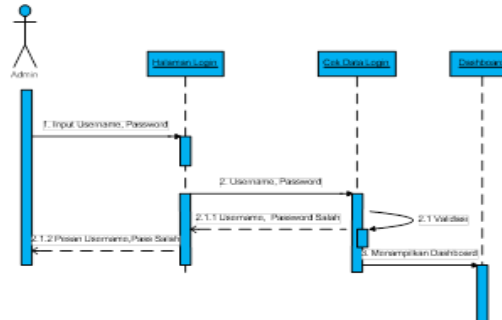
Sequence Diagram

* Suitable author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

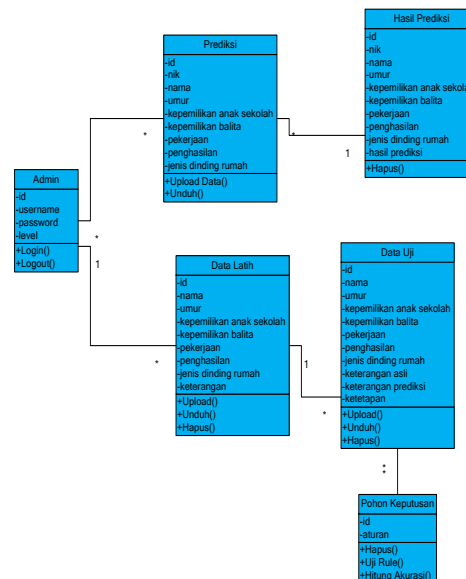
At this stage, it will be described how the interaction between objects in sending a series of messages will be described. In the journal (Suendri, 2018b), states that according to (Haviluddin, 2011) *sequence Diagram* describes the steps in sequence, including a logical chronology of changes that need to be made to produce the output that corresponds to the *Use case Diagram*.



Gambar 5. Sequence Diagram Login

Class Diagram

Class diagram describe the relationships between classes as well as an in-depth explanation of each class in a system design model. This diagram also shows the rules and responsibilities of entities that affect the behavior of the system (Irawan & Simargolang, 2018)



Gambar 6. Class Diagram

DISCUSSION

System Implementation and Testing

After the analysis and design stage of the system is completed, the next step is the implementation of the design into a data information system *Mining Based Copyright* © which uses the C5.0 algorithm, which can display the pattern of recommendations for prospective KKS recipients for the next period. The system testing applied in this study is the *blackbox testing*. This test is intended to focus on what functional needs are in the system.

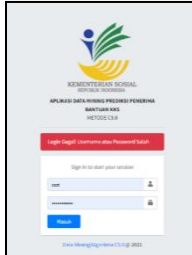
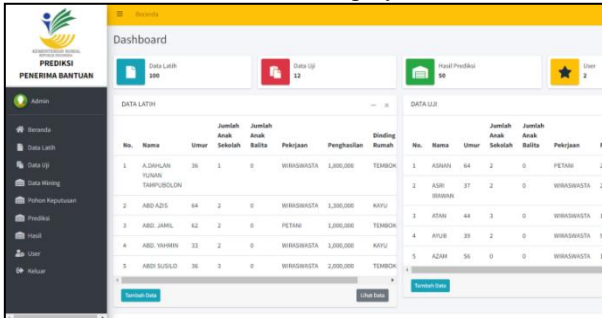
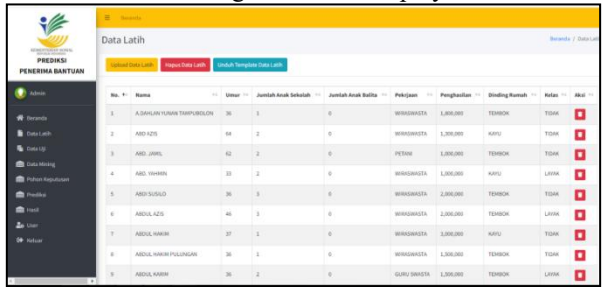
Table 2. Black Box Testing

Yes	Input/Output Design	Expected results	Result Current
-----	---------------------	------------------	----------------

* Suitable author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

1.	Opening the website,	Log in to formlogin	OKAY
	Login view :  Figure 2.1 Displaying the Login Page		
2.	Click Login (incorrect username and password)	View message (failed to sign-in)	OKAY
	Login failure display :  Figure 2.2 Displaying Failed Login Pages		
3.	Click Login (username and password are correct)	Log in to the dashboard	OKAY
	Dashboard Display :  Figure 2.3 Displaying the Dashboard Page		
4.	Click Training Data	View training data pages	OKAY
	Training data menu display:  Figure 2.4 Displaying the Train Data Page		
5.	Click Upload Training Data	Display the training data upload page	OKAY

Coach data upload view :

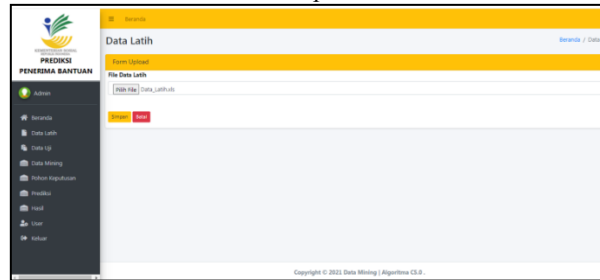


Figure 2.5 Displaying Page Upload Training Data

6. Click Clear Training Data

Display message successfully delete training data

OKAY

Message view successfully delete training data:

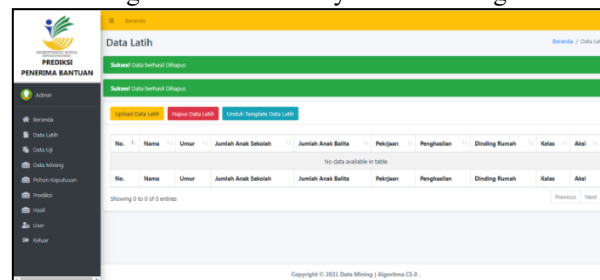


Figure 2.6 Showing Deleted Messages Successfully Training Data

7. Click Download Training Data

View training data download results

OKAY

Train data download view:

No.	Nama	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Kates	Aksi
1	ADAMULYAN YUSMAN TANJUNGPINANG	35	1	0	WIRASWASTA	1,800,000	TEMBOK	TIDAK	
2	ABDI AZIZ	64	2	0	WIRASWASTA	1,300,000	KATU	TIDAK	
3	ABDI JAVEL	62	2	0	PEKAWAN	1,000,000	TEMBOK	TIDAK	
4	ABDI NAWABEN	33	2	0	WIRASWASTA	1,000,000	KATU	LAINNYA	
5	ABDI RUSLED	38	3	0	WIRASWASTA	2,300,000	TEMBOK	TIDAK	
6	ABDUL AZIZ	46	3	0	WIRASWASTA	2,300,000	TEMBOK	LAINNYA	
7	ABDUL NAWABEN	37	1	0	WIRASWASTA	3,000,000	KATU	TIDAK	

Figure 2.7 Showing the Train Data Download Results

8. Click Test Data

View test data page

OKAY

Test data menu display:

No.	Nama	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Kates	Aksi
1	ADAMULYAN YUSMAN TANJUNGPINANG	35	1	0	WIRASWASTA	1,800,000	TEMBOK	TIDAK	
2	ABDI AZIZ	64	2	0	WIRASWASTA	1,300,000	KATU	TIDAK	
3	ABDI JAVEL	62	2	0	PEKAWAN	1,000,000	TEMBOK	TIDAK	
4	ABDI NAWABEN	33	2	0	WIRASWASTA	1,000,000	KATU	LAINNYA	
5	ABDI RUSLED	38	3	0	WIRASWASTA	2,300,000	TEMBOK	TIDAK	
6	ABDUL AZIZ	46	3	0	WIRASWASTA	2,300,000	TEMBOK	LAINNYA	
7	ABDUL NAWABEN	37	1	0	WIRASWASTA	3,000,000	KATU	TIDAK	

Figure 2.8 Showing the Test Data Page

9. Click Upload Test Data

View the test data upload page

OKAY

Test the display of data uploads:

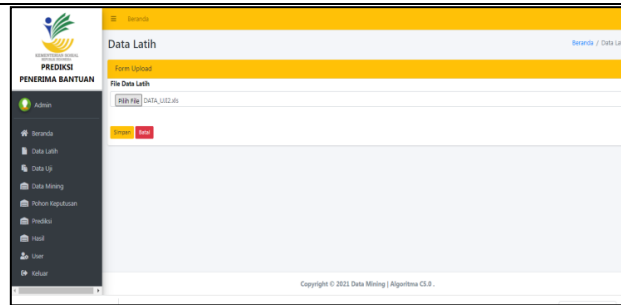


Figure 2.9 Showing the Test Data Upload Page

10.

Click Clear Test Data

Display a message successfully
delete test data

OKAY

The message view successfully clears the test data:

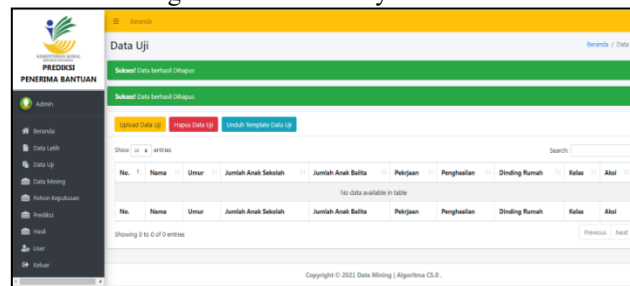


Figure 2.10 Displaying the message Successfully Clear Test Data

11.

Click Download Test Data

View test data download
results

OKAY

Test data download view:

No.	Nama	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Kotor	Aki
1	ADHAN	64	2	0	PETANI	2.000.000	TEMBOKE	LAKUK	
2	ADRI BRABIAN	57	2	0	WIRASWASTA	2.900.000	KAYU	TEKAK	
3	ADRI	48	3	0	WIRASWASTA	1.700.000	KAYU	LAKUK	
4	ARJUN	39	2	0	WIRASWASTA	9.000.000	KAYU	TEKAK	
5	ADZAR	56	0	0	WIRASWASTA	1.500.000	KAYU	TEKAK	
6	BAHREN	35	0	1	WIRASWASTA	3.000.000	TEMBOKE	TEKAK	
7	BAMBANG BRABIAN	52	1	0	SOPIR	2.000.000	TEMBOKE	TEKAK	
8	BEATI SHAMUAN	45	3	0	IBU RUMAH-TANGGA	2.500.000	TEMBOKE	LAKUK	

Figure 2.11 Showing Test Data Download Page

12.

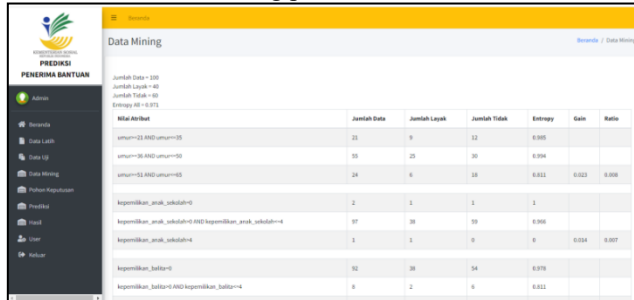
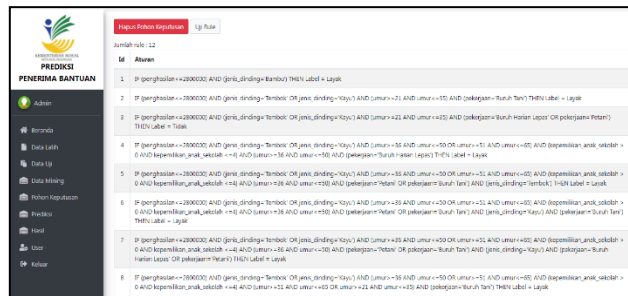
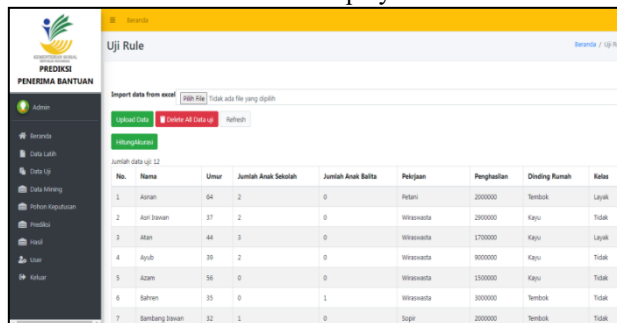
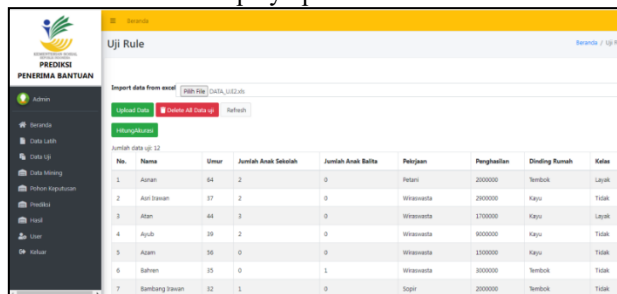
Click the Data Mining menu

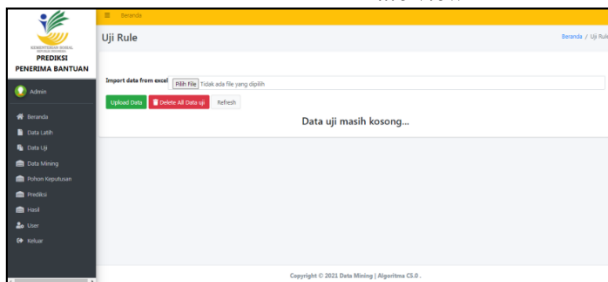
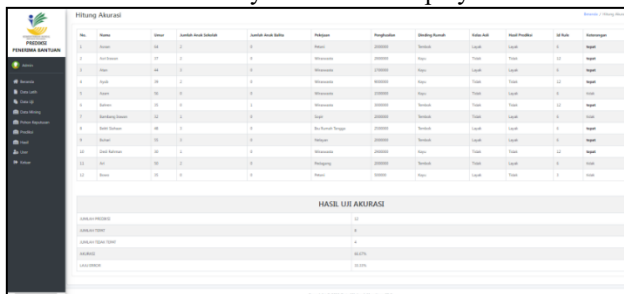
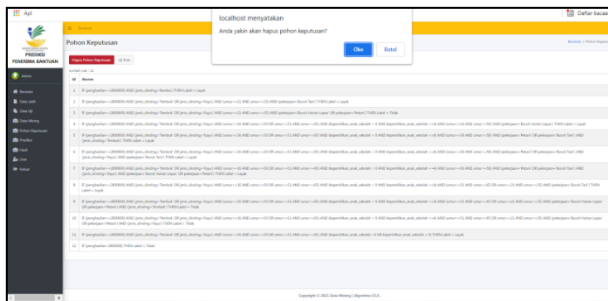
Show data mining page

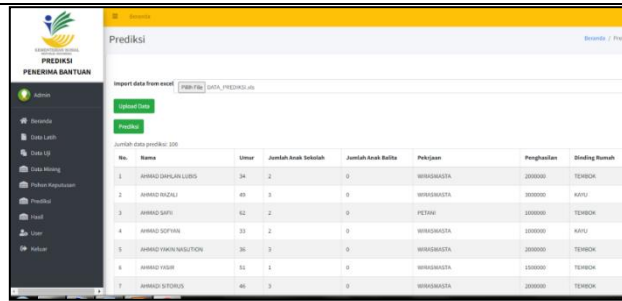
OKAY

Mining data display:

Figure 2.12 Showing the Data Mining Page

13.	Click the Mining Process button	Display the mining process page	OKAY
	Mining process view :  Figure 2.13 Showing the Mining Process Page		
14.	Click the Decision Tree menu	View the decision tree page	OKAY
	Decision tree view:  Figure 2.14 Displaying the Page Decision Tree		
15.	Click the Test Rules button	Display test rules page	OKAY
	Test rule display :  Figure 2.15 Showing Page Test Rules		
16.	Click the Upload test data rule button	Show test rule upload page	OKAY
	DisplayUpload Rules:  Figure 2.16 Showing Page Upload Test Rules		

17.	Click the Clear test rule data button	Show message successfully delete data test rules	OKAY
	Delete the data test rule view: 		
18.	Click the Calculate Accuracy button	Display accuracy calculation page	OKAY
	Accuracy calculation display: 		
19.	Click the Delete Decision Tree	View messages that successfully delete the decision tree	OKAY
	Remove the view of the decision tree: 		
20.	Click the Upload Population Prediction	View the population prediction upload page	OKAY
	See population prediction:		

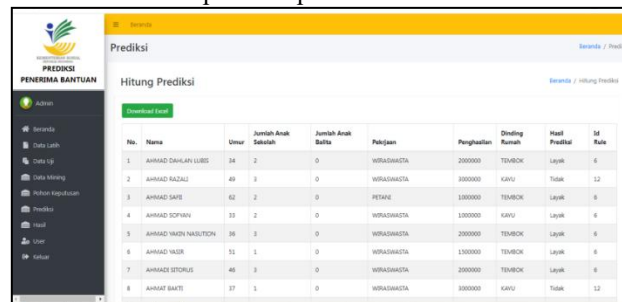


No.	Name	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah
1	AMIRAH DAHLIA LUBIS	34	2	0	WIRASWASTA	200000	TEMBOK
2	AMIRAH RAZALI	49	3	0	WIRASWASTA	300000	KAYU
3	AMIRAH SARI	62	2	0	PETANI	100000	TEMBOK
4	AMIRAH SOPHAN	33	2	0	WIRASWASTA	100000	KAYU
5	AMIRAH YANZI NASEUTON	36	3	0	WIRASWASTA	200000	TEMBOK
6	AMIRAH YASIR	51	1	0	WIRASWASTA	150000	TEMBOK
7	AMIRAH STORUS	46	3	0	WIRASWASTA	200000	TEMBOK

Figure 2.20 Showing Page Upload Population Prediction

21. Click the Predict button Showing the prediction page OKAY

Population prediction view:

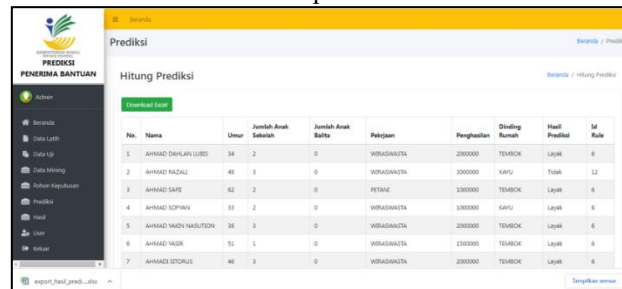


No.	Name	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Hasil Prediksi	Sk. Rata
1	AMIRAH DAHLIA LUBIS	34	2	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
2	AMIRAH RAZALI	49	3	0	WIRASWASTA	300000	KAYU	TIDAK	12
3	AMIRAH SARI	62	2	0	PETANI	100000	TEMBOK	LAYAK	6
4	AMIRAH SOPHAN	33	2	0	WIRASWASTA	100000	KAYU	LAYAK	6
5	AMIRAH YANZI NASEUTON	36	3	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
6	AMIRAH YASIR	51	1	0	WIRASWASTA	150000	TEMBOK	LAYAK	6
7	AMIRAH STORUS	46	3	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
8	AMIRAH SARI	37	1	0	WIRASWASTA	300000	KAYU	TIDAK	12

Figure 2.21 Showing Page Population Prediction

22. Click Download Excel Display the prediction results download page OKAY

Showdownload prediction results:

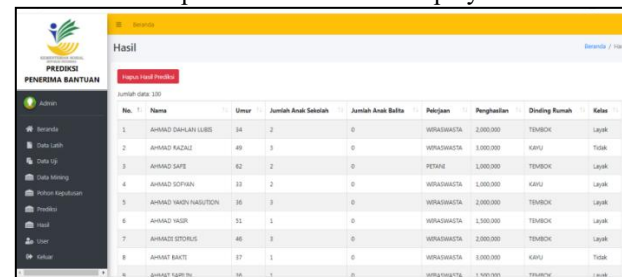


No.	Name	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Hasil Prediksi	Sk. Rata
1	AMIRAH DAHLIA LUBIS	34	2	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
2	AMIRAH RAZALI	49	3	0	WIRASWASTA	300000	KAYU	TIDAK	12
3	AMIRAH SARI	62	2	0	PETANI	100000	TEMBOK	LAYAK	6
4	AMIRAH SOPHAN	33	2	0	WIRASWASTA	100000	KAYU	LAYAK	6
5	AMIRAH YANZI NASEUTON	36	3	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
6	AMIRAH YASIR	51	1	0	WIRASWASTA	150000	TEMBOK	LAYAK	6
7	AMIRAH STORUS	46	3	0	WIRASWASTA	200000	TEMBOK	LAYAK	6
8	AMIRAH SARI	37	1	0	WIRASWASTA	300000	KAYU	TIDAK	12

Figure 2.22 Showing the Population Prediction Download Page

23. Click Results Menu Show results page OKAY

The predicted results are displayed:



No.	Name	Umur	Jumlah Anak Sekolah	Jumlah Anak Balita	Pekerjaan	Penghasilan	Dinding Rumah	Rata
1	AMIRAH DAHLIA LUBIS	34	2	0	WIRASWASTA	2.000.000	TEMBOK	LAYAK
2	AMIRAH RAZALI	49	3	0	WIRASWASTA	3.000.000	KAYU	TIDAK
3	AMIRAH SARI	62	2	0	PETANI	1.000.000	TEMBOK	LAYAK
4	AMIRAH SOPHAN	33	2	0	WIRASWASTA	1.000.000	KAYU	LAYAK
5	AMIRAH YANZI NASEUTON	36	3	0	WIRASWASTA	2.000.000	TEMBOK	LAYAK
6	AMIRAH YASIR	51	1	0	WIRASWASTA	1.500.000	TEMBOK	LAYAK
7	AMIRAH STORUS	46	3	0	WIRASWASTA	2.000.000	TEMBOK	LAYAK
8	AMIRAH SARI	37	1	0	WIRASWASTA	3.000.000	KAYU	TIDAK
9	AMIRAH SARI	34	1	0	WIRASWASTA	1.000.000	TEMBOK	LAYAK

Figure 2.23 Showing the Prediction Results Page

24. Click the Delete Prediction Result button Display the delete prediction results page OKAY

Delete the prediction results view:

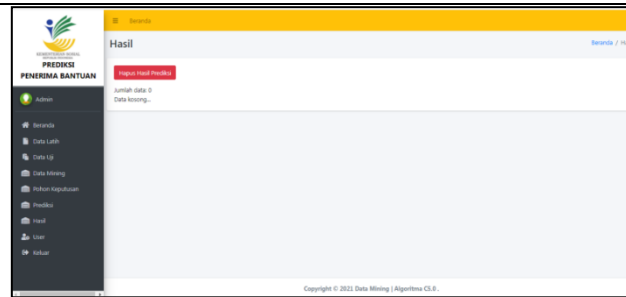


Figure 2.24 Showing the Delete Prediction Results Page

25. Click the User Menu Display user pages OKAY

User page view :

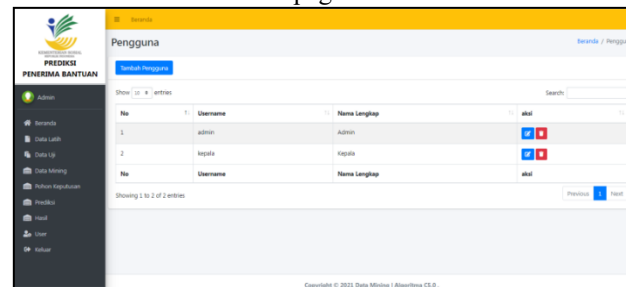


Figure 2.25 Displaying the User Page

26. Click the Add User button Display the add user form page OKAY

User addition page view:



Figure 2.26 Showing the Add User Form Page

27. Click the Save button Show save user additions OKAY

User save page view:

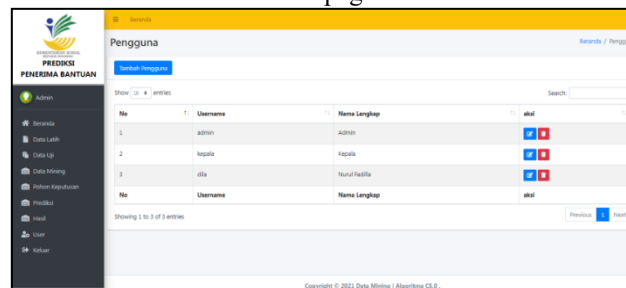


Figure 2.27 Showing the Save Add User Page

28. Click the Edit User button View user edit forms OKAY

User edit form page view:

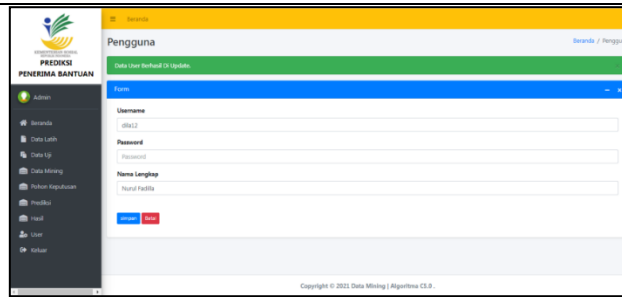


Figure 2.28 Showing Page Form Edit Users

29. Click the **Save button**

Display users edit saved messages successfully

OKAY

Saved view messages successfully edit the user:

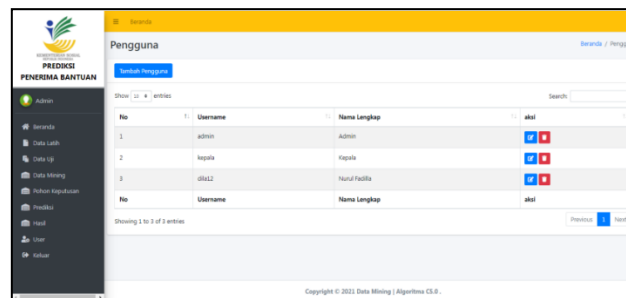


Figure 2.29 Showing the User's Save Edit Results

30. Click the **Remove User button**

View message successfully delete user data

OKAY

See Successfully Delete User Data:

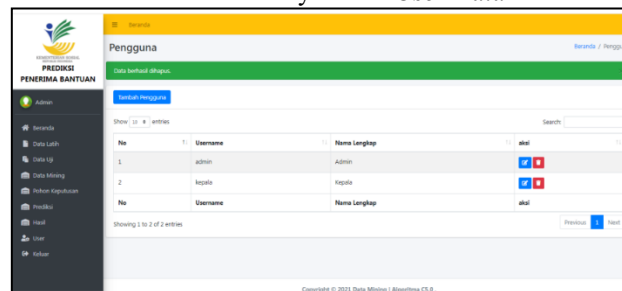


Figure 2.30 Show message successfully delete data user

31. Click **Exit Menu**

Displays messages that exit the system.

OKAY

"Sure you want to log out of the system" message:

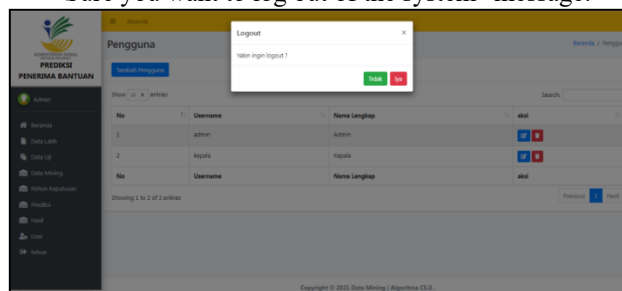


Figure 2.31 Show out-of-system messages

CONCLUSION

Based on the research that has been conducted, the application of the C5.0 Algorithm method can help the Silau Laut Regency Office in recommending prospective recipients of KKS assistance for the next period. The data used in the data *Mining* using the C5.0 algorithm using previous KKS assistance recipient data and a combination of population data so that attributes or parameters can determine the recommended pattern of KKS assistance recipients in the next period. With the use of the C5.0 algorithm, a decision pattern is produced in the form of *rule* which can be used as a guideline for determining recommendations for prospective recipients of KKS assistance in the next period. The weakness of this system lies in its reliance on previous data, which if not up-to-date can affect the accuracy of predictions. In the future, ensemble methods such as boosting or hybrid C5.0 – Naïve Bayes can be tested.

REFERENCES

- Dewi, D. A. W., Imam, C., & Sutrisno. (2019). *Classification of Child Growth and Development Deviations Using*. 3(10), 10258–10265.
- Heriyanto, Y. (2018). The design of a web-based car rental information system at PT. APM Rent Car. *Journal of Intra-Tech*, 2(2), 64–77.
- Irawan, M. D., & Simargolang, S. A. (2018). Implementation of e-Archives in the Informatics Engineering Study Program. *Journal of Information Technology*, 2(1).
- Kusrini, E. T. L. & (2018). Data Mining Algorithm. In *Algorithms of Data Mining* (pp. 4–12). CV Andi Offset.
- Muhazir, A., Fakhriza, M., & Sutejo, E. (2017). Implementation of sequential methods in searching for the distribution of goods in the cargo integration system. *Journal & Research of Informatics Engineering Volume 2 Number 2, October 2017*, 2, 24–30.
- Nofriansyah, D. (2019). *Data Mining and Testing Algorithms*. Deepublish Publisher.
- Pratiwi, R., Hayati, M. N., & Prangga, S. (2020). *Comparison of C5 Algorithm . 0 Classification With Classification And Regression Tree (Case Study: Social Data of Family Heads of Teluk Baru Village, Muara Ancalong District in 2019) Comparison of C5 . 0 Algorithm Classification with Classificat.* 14(2), 267–278.
- Rahim, R., Zufria, I., Kurniasih, N., Simargolang, M. Y., Hasibuan, A., Sutiksno, D. U., Nanuru, R. F., Anamofa, J. N., Ahmar, A. S., & Achmad Daengs, G. S. (2018). C4.5 classification data mining for inventory control. *International Journal of Engineering and Technology(UAE)*, 7, 68–72. <https://doi.org/10.14419/ijet.v7i2.3.12618>
- Suendri. (2018a). *Diktat of Web-Based Programming Lectures*.
- Suendri. (2018b). Implementation of UML (Unified Modelling Language) Diagram in the Design of Lecturer Remuneration Information System with Oracle Database (Case Study: UIN North Sumatra Medan). *Journal of Computer Science and Informatics*, 3(1), 1–9.
- Trimahardhika, R. and E. S. (2017). User of the Rapid Application Development Method in the Design of Library Information Systems. *Journal of Informatics*, 4 no. 2(2), 249. <http://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/2226>
- Wijaya, A. C., Hasibuan, N. A., & Ramadhani, P. (2018). Implementation of the C5 Algorithm. 0 In the Classification of Community Income (Case Study: Mosque Village, Medan City District). *Scientific Information and Technology (INTI)*, 13, 192–198.
- Zaman, K. (2016). The application of data mining uses the C4.5 algorithm to determine the eligibility of recipients of social rehabilitation assistance for uninhabitable houses (Case Study in the South Solok Regency Government). *UPI YPTK Journal of KomTekInfo*, 3(2), 12–24. <http://lppm.upiypk.ac.id/komtekinfo/index.php/KOMTEKINFO/article/download/58/51>
- Zufria, I. (2013). UML (Unified Modeling Language) Based Modeling with User Centered Design (UCD) Orientation Engineering Strategy in the Education Administration System. *State Islamic University of North Sumatra Medan, January 2013*.