



Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

Jefri Junifer Pangaribuan^{1*}, Alexander Putra²

Universitas Pelita Harapan, Indonesia^{1,2}

Corresponding Email: jefri.pangaribuan@uph.edu*

Received: 23-05-2024 Reviewed: 10-06-2024 Accepted: 23-06-2024

Abstract

Donating blood is an altruistic act driven by concern for others and personal commitment to health. It is crucial for patients needing transfusions due to excessive bleeding. However, there has been a decline in blood donations globally. To address this, the medical community needs a method to predict whether a donor will donate again, enabling proactive measures to ensure an adequate blood supply. This study utilizes data from the University of California, Irvine (UCI) Machine Learning Repository, specifically the Blood Transfusion Service Data Set, employing the Decision Tree method with the C4.5 algorithm. C4.5, an improvement over Iterative Dichotomiser 3 (ID3), can handle missing values, pruning, and continuous data. The aim is to classify blood donor eligibility accurately. The aim of this study is to explore how the utilization of the C4.5 algorithm in decision tree classification can predict whether an individual will donate blood again or not. The analysis identifies five key attributes—Recency, Frequency, Monetary, Time (Months), and Decision—as determinants of repeat donation likelihood. Using a confusion matrix to assess accuracy, the C4.5 algorithm achieved a 77.68% accuracy, with an error rate of 22.32%, a sensitivity of 30.19%, and a specificity of 92.40%.

Keywords: Blood Donation, C4.5 Algorithm, Classification, Data Mining, Decision Tree

Introduction

In these days and age, developing countries must consider health challenges, one of which is the provision of enough safe blood supplies for transfusions, which must be a priority when someone experiences excessive bleeding, during surgery, or when cancer patients are undergoing chemotherapy (*Why Blood Donation Is Important – and Who Benefits*, 2021). Hospitals collaborating with many parties or volunteer blood donor organizations typically have challenges, as contacting multiple individuals randomly and not in a pattern to solicit

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

blood donations is less efficient. This is because the information collected from potential donors sometimes meets blood donation guidelines.

According to research by (Kaptoge et al., 2019), the blood donation interval could be shortened to alleviate the global blood supply crisis. Because of this, a classification system is required to identify the selection of eligible prospective blood donors based on preset criteria to determine whether a donor will donate again or not. In general, donors must meet six requirements (Ayodonor - Palang Merah Indonesia, 2016): (1) have healthy physical and mental, (2) with an age range between 17 to 65 years, (3) have a weight body weight of 45 kg, (4) blood pressure in the range of 100-170 systolic and/or 70-100 diastolic, (5) the percentage of hemoglobin levels between 12.5g% and 17.0g%, and (6) history of donor activity blood was counted at least 12 weeks (equivalent to 3 months) ago. This study used data obtained from the UCI Machine Learning Repository database with a data set called the Blood Transfusion Service Center Data Set to find data that complies with predetermined requirements (Yeh, 2008).

In health information systems, the availability of a branch of computer science expertise known as data mining or Knowledge Discovery in Databases is anticipated to solve the current challenges (Pangaribuan et al., 2017; Pangaribuan & Suharjito, 2014). This relates to the fundamental nature of data mining, which can be used as a reference for analysis to discover unrealized yet significant and meaningful knowledge, patterns, and information (Liao et al., 2022; Maimon & Rokach, 2005; Priyasadie & Isa, 2021; Ramadani et al., 2023). Due to the numerous classifications of the imported data, the decision tree approach is required to break down previously massive data sets into smaller record sets by applying a series of decision rules that can be used to predict or clarify an occurrence (Resti et al., 2023; Rusyana et al., 2023; Syukmana et al., 2020; Yang et al., 2020). In this research, implementing the C4.5 data mining method is anticipated to become an alternate decision support system for creating the necessary data. Algorithm C4.5 is required to generate rules in the form of a decision tree by analyzing (1) five attributes of possible donors, such as (a) Recency; (b) Frequency; (c) Monetary; (d) Time; and (e) Decisions (Barus, Nathasya, et al., 2023); (2) Attribute value; (3) Entropy; and (4) Gain (Sumiati et al., 2023). The list is then utilized to determine if an individual would donate again.

In conducting classification, the Decision Tree method is the most frequently employed approach (Irawan, 2021; Sathiyarayanan et al., 2019; Zulfikar et al., 2018). This method is also utilized because in this research, a dataset sourced from UCI Machine Learning Repository is used. According to previous studies, this method can process datasets from the machine learning repository with the highest accuracy rate of 99.93% (Charbuty & Abdulazeez, 2021).

Based on the results of the preceding description, this research aims to determine the method and obtain the results of rules from using a decision tree with the C4.5 algorithm in classifying whether someone will donate blood again or not. Five attributes are utilized as study criteria, and testing is performed using a confusion matrix to determine the validity or accuracy of the data and projected results. The benefits of this research are expected to enhance knowledge and insights regarding the application of the C4.5 Algorithm in the medical field, particularly in blood donation.

Literature Review

Blood donation

In the human body, there is usually about 5 liters of blood circulating, which also depends on body size. If there is a severe blood loss, also known as hemorrhage, of 30-40% of the total blood volume, the individual must be given a blood transfusion immediately. For a blood transfusion to be given, of course, blood donors are needed. Blood donors are people who voluntarily donate their blood. For people to donate blood, of course, organizers are needed.

Decision tree

Decision tree, which first appeared in the 1960s, are one of the most widely used methods in data mining when it comes to classification problems. As the name of the method suggests, decision trees perform classification into branch-shaped segments that look like an inverted tree with 3 nodes or nodes, namely:

1) Root Node

This node is located at the top of the decision tree and only has output.

2) Internal Node

This node is located after the root node, where one output from the root node or the previous internal node becomes input in the internal node and produces a minimum of 2 outputs.

3) Leaf Node

The last node of a decision tree which usually means that the data has been classified. This node also does not have any more output.

C4.5 algorithm

One of the most popular algorithms for creating decision trees is the C4.5 algorithm. The C4.5 algorithm is an algorithm used to solve classification problems from datasets. This algorithm is a development and refinement of the ID3 algorithm, so now the algorithm can handle missing values, pruning, and continuous data. The steps taken by the C4.5 algorithm to build a decision tree are as follows:

- 1) Choose an attribute as the root node.
- 2) Create a branch for each value.
- 3) Divide each branch into cases.
- 4) Repeat these steps for each branch until all have the same class.

Confusion matrix

Confusion Matrix is one of the methods used to measure the performance of classification results by comparing the results of predicted values with factual values. The confusion matrix table will have a size of $n \times n$, where the value of n depends on how many output classes (Visa et al., 2011). If the value of $n = 2$, then the confusion matrix table will be as follows:

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

Table 1. Confusion matrix table

Actual	Predicted	Total
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)
Total	TP + FP	FN + TN

Here is a breakdown of the table:

- True Positive (TP): The number of data points that were correctly classified as positive.
- False Negative (FN): The number of data points that were incorrectly classified as negative.
- False Positive (FP): The number of data points that were incorrectly classified as positive.
- True Negative (TN): The number of data points that were correctly classified as negative.

The confusion matrix can be used to calculate several metrics, such as:

- Accuracy: The proportion of data points that were correctly classified.
- Precision: The proportion of positive predictions that were correct.
- Recall: The proportion of actual positives that were correctly identified.
- F1-score: The harmonic mean of precision and recall.

Research Method

The process begins with data collection and division, followed by the determination of criteria for each attribute. The data was obtained from the UCI Machine Learning Repository and is freely available for use. This dataset was collected from one of the blood donation service centers in Hsin-Chu City, Taiwan. The service center then sent their blood donation service bus to one of the universities in Hsin-Chu City to conduct regular blood donation drives every three months. This data was provided to the UCI Machine Learning Repository on October 3, 2008 (Yeh, 2008). The data is divided into 30% for training data and 70% for testing data.

After the data is collected and divided, the decision variable is determined based on the data that is divided into attributes. There are 5 criteria:

- 1) Recency - when was the last blood donation in terms of months.
- 2) Frequency - how many times blood was donated.
- 3) Monetary - the amount of blood donated in terms of cc.
- 4) Time - how many months since the first blood donation.
- 5) Decision - whether the person will donate blood again or not.

Table 2 is the criteria table for the collected data.

Table 2. Criteria table

Criteria	Condition	Classification
R (Recency)	Recency \leq 3 months	Recently donated blood
	Recency $>$ 3 months	Not recently donated blood
F (Frequency)	Frequency \leq 4 months	Infrequently
	Frequency $>$ 4 months	Frequently

	Monetary < 500cc	Little
M (Monetary)	Monetary 500cc – 1,750cc	Medium
	Monetary > 1,750cc	Plenty
T (Time – Months)	Months <= 28 months	New donor
	Months > 28 months	Existing donor
Decision	0	Non-donor
	1	Donor

Next, the data is subjected to calculations, during which the highest Gain value is sought from the Entropy results. The final step is to generate the decision tree rules and evaluate the prediction results using a confusion matrix (Angraini et al., 2020; Barus, Romindo, et al., 2023). By following these steps, it is possible to classify data using the decision tree method and C4.5 algorithm accurately. Calculating the decision tree method using the C4.5 algorithm involves several stages, as outlined in Figure 1.

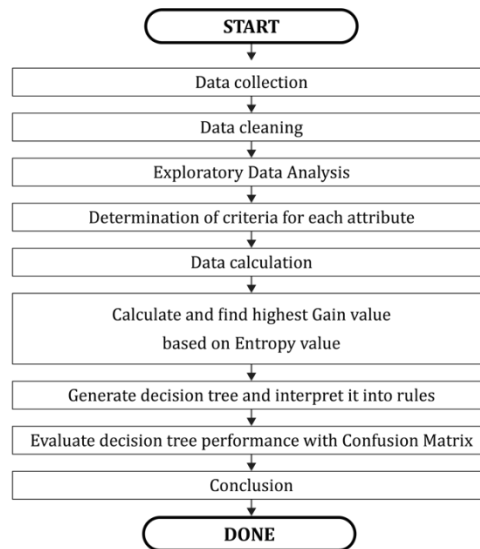


Figure 1. Completion stages with the decision tree method

Result/Findings

To get early insights, data that has been visually represented is subjected to exploratory data analysis. Following is the histogram for each attribute that has been. Figure 2 indicates that the majority of people in the dataset are not recent blood donors.

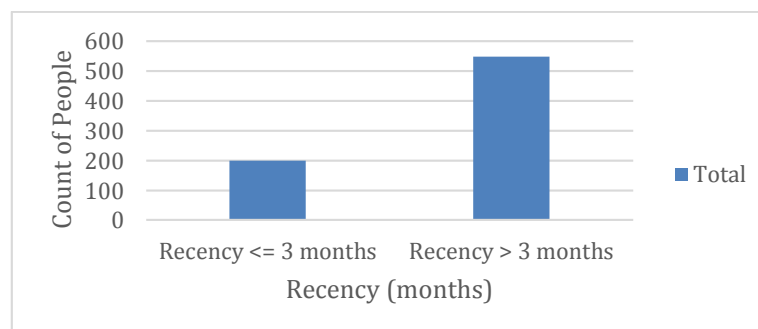


Figure 2. Histogram attribute for recency

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

The histogram in Figure 3 indicates that there are more people donating blood at lower frequencies than those donating at higher frequencies. There are 419 instances where individuals donated blood less than or equal to every 4 months. Meanwhile, there are 329 instances where individuals donated blood more than every 4 months.

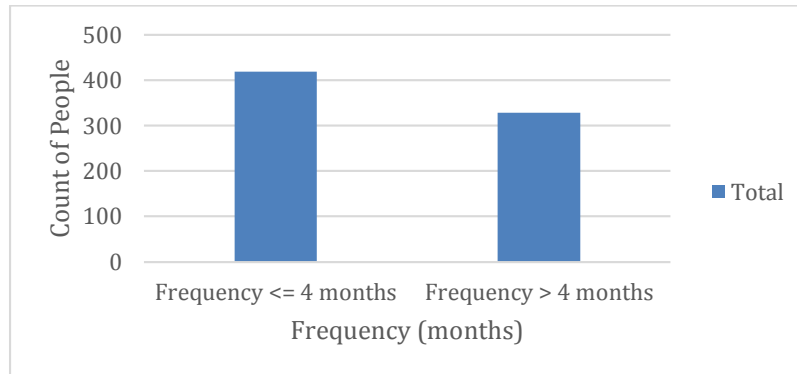


Figure 3. Histogram attribute for frequency

From Figure 4, the histogram indicates that there are individuals who have donated significantly less blood than others. The histogram below shows that 418 data points donated a total of 500cc to 1750cc of blood. Meanwhile, there are 172 data points that donated more than 1750cc of blood, and 158 data points that donated less than 500cc.

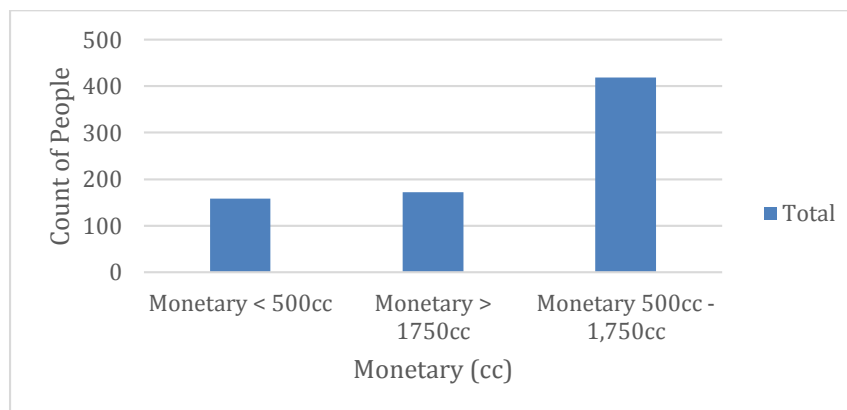


Figure 4. Histogram attribute for monetary

Figure 5 indicates that there are more new donors than old donors. This histogram indicates that the new donors who donated blood less than or equal to the last 28 months are 392 data. Meanwhile, old donors who donated blood more than 28 months ago are 356 data.

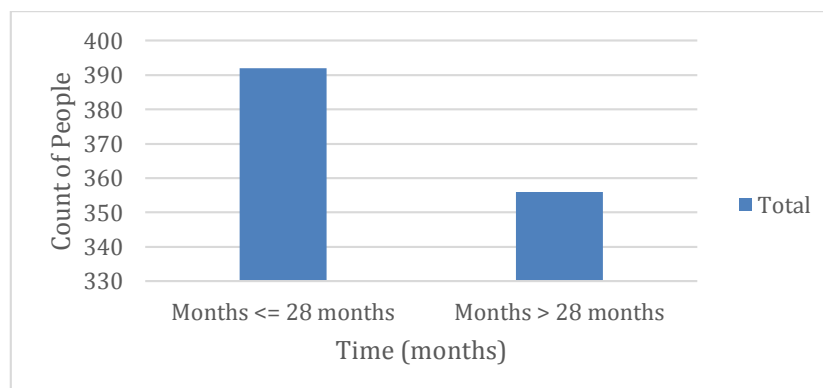


Figure 5. Histogram attribute for time (months)

Using the scikit-learn function library, checking for null data, skewness and kurtosis may be calculated from the data visualization. Figure 6 displays the code along with the outcomes with the following captions:

1. Figure 6 (a) shows the code for checking null data and its result.
2. Figure 6 (b) displays the code for calculating skewness and its result.
3. Figure 6 (c) exhibits the code for calculating kurtosis and its result.

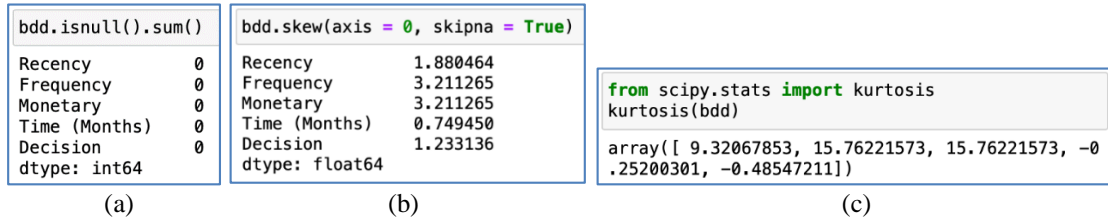


Figure 6. (a): Code to check for null data and result; (b): Code to calculate skewness value and result; (c): Code for calculating kurtosis value and result

Derived from Figure 6, the data used has no null value. As demonstrated by the data presented in Figure 7, most blood donors in the sample had recently made donations, while a minority had not. Figure 7 represents the statistical data for all attributes.

bdd.describe()					
	Recency	Frequency	Monetary	Time (Months)	Decision
count	748.000000	748.000000	748.000000	748.000000	748.000000
mean	9.506684	5.514706	1378.676471	34.282086	0.237968
std	8.095396	5.839307	1459.826781	24.376714	0.426124
min	0.000000	1.000000	250.000000	2.000000	0.000000
25%	2.750000	2.000000	500.000000	16.000000	0.000000
50%	7.000000	4.000000	1000.000000	28.000000	0.000000
75%	14.000000	7.000000	1750.000000	50.000000	0.000000
max	74.000000	50.000000	12500.000000	98.000000	1.000000

Figure 7. Statistics data for all attributes

Additionally, most blood donors had a low frequency of donations, resulting in a smaller cumulative volume of blood donated per individual. In contrast, a smaller group of donors exhibited a higher frequency of donations, leading to a greater cumulative volume of blood donated per individual. The average volume of blood donated per person was 1,000 cc, but some individuals had donated up to 12,500 ccs. The average time since a person first donated their bloods is 28 months, although some are new donors, who just started donating 2 months ago, while some have donated for 98 months.

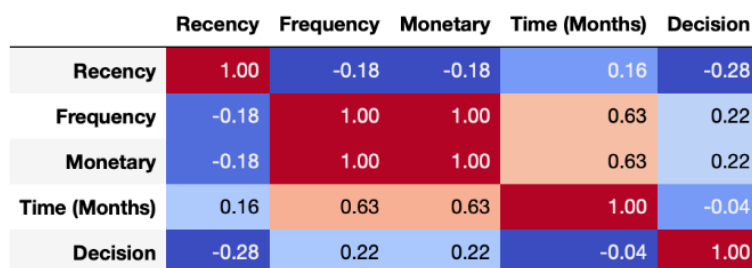


Figure 8. Heatmap of correlations between attributes

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

The results of the correlation heatmap shown in Figure 8, indicate that there is a strong positive correlation between the Monetary and Frequency attributes. This suggests that individuals who donate more frequently also tend to donate larger cumulative volumes of blood. Additionally, there is a moderate positive correlation between the Time (Months) attribute and the Monetary and Frequency attributes. This suggests that individuals who have been donating for a more extended period and who donate more frequently also tend to donate larger cumulative volumes of blood. Based on the correlations between each attribute, all are used in making the decision tree.

To improve the accuracy of the decision tree rules, it is necessary to scale the Monetary attribute so that its range, in the thousands, is more comparable to the ranges of the other attributes, in the hundreds. By using min-max scale method, the Monetary range is scaled from thousands to between 0-1. Figure 9 shows the code and results for scaling the Monetary attribute with min-max scale.

```
from sklearn.preprocessing import minmax_scale
bdd["Monetary"] = minmax_scale(bdd["Monetary"])
with pd.option_context('display.max_rows',10):
    print(bdd)
```

	Recency	Frequency	Monetary	Time (Months)	Decision
0	2	50	1.000000	98	1
1	0	13	0.244898	28	1
2	1	16	0.306122	35	1
3	2	20	0.387755	45	1
4	1	24	0.469388	77	0
..
743	23	2	0.020408	38	0
744	21	2	0.020408	52	0
745	23	3	0.040816	62	0
746	39	1	0.000000	39	0
747	72	1	0.000000	72	0

Figure 9. Code for scaling the monetary value and results

After scaling the monetary range into between 0-1, the data will be multiplied by 100 to have a similar range in the hundreds. Figure 10 shows the codes and the result. Data transformation is performed using the min-max scaling method on the monetary attribute.

```
bdd['Monetary'] = round(bdd['Monetary'].multiply(100))
with pd.option_context('display.max_rows',10):
    print(bdd)
```

	Recency	Frequency	Monetary	Time (Months)	Decision
0	2	50	100.0	98	1
1	0	13	24.0	28	1
2	1	16	31.0	35	1
3	2	20	39.0	45	1
4	1	24	47.0	77	0
..
743	23	2	2.0	38	0
744	21	2	2.0	52	0
745	23	3	4.0	62	0
746	39	1	0.0	39	0
747	72	1	0.0	72	0

[748 rows x 5 columns]

Figure 10. The result of data transformations with min-max scale on monetary attribute

The constructed decision tree can aid in classifying an individual's blood donation based on specific characteristics. Figure 11 shows the constructed and pruned decision tree. After

utilizing the C4.5 method to calculate the decision tree formation, the resulting rules or regulations of the generated decision tree are as follows:

- 1) A donor with Rare Frequency ≤ 4 times will not donate again.
- 2) A donor with Frequent Frequency > 4 times, with Recently Donated < 3 months, followed by the type of New Donor Time ≤ 28 months, will donate again.
- 3) A donor with Frequent Frequency ≥ 4 times, with Recently Donated < 3 months, followed by the type of Old Donor Time ≥ 28 months, will not donate again.
- 4) A donor with Frequent Frequency ≥ 4 times, with Recently Donated < 3 months, will not donate again.

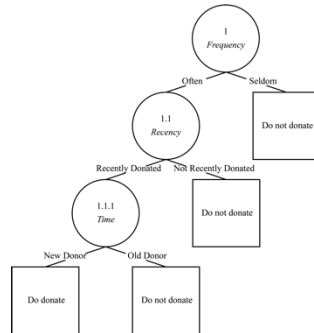


Figure 11. The final result of the decision tree after pruning

The use of a decision tree and confusion matrix can help classify and evaluate the accuracy of predictions made about an individual's likelihood of donating blood again. Next, the confusion matrix method is used to evaluate the accuracy of the decision tree's classification findings using actual data. The confusion matrix for the generated rules is computed with RapidMiner Studio as shown in Figure 12. From Figure 12, we can interpret that from the generated rules. The overall accuracy is 77.68%, the error rate is 22.32%. the sensitivity is 30.19%, and the specificity is 92.40%.

accuracy: 77.68%

	true Donor	true Tidak Donor	class precision
pred. Donor	16	13	55.17%
pred. Tidak Donor	37	158	81.03%
class recall	30.19%	92.40%	

Figure 12. Confusion matrix from decision tree results

Conclusion

The classification of whether an individual will donate blood again or not is based on 5 attributes, namely Recency, Frequency, Monetary, Time, and Decision which is the decision attribute. From these five attributes, a decision tree is formed using the C4.5 algorithm. From these five attributes, four rules are generated. Out of the four rules obtained from the formed decision tree, one rule yields the decision Donor, and three rules yield the decision Not Donor. The decision tree may be complex and prone to overfitting, so pruning is recommended to improve efficiency and effectiveness. The accuracy of the classification result of the decision tree using the confusion matrix method is 77.68%, indicating that the decision tree is effective in classifying an individual's blood donation status. For future research, other classification

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

methods can be utilized to compare the accuracy of decision trees with alternative classification methods.

Declaration of conflicting interest

The authors declare that there is no conflict of interest in this work.

Funding acknowledgment

The work is fully supported by LPPM (Lembaga Penelitian dan Pengabdian kepada Masyarakat) of Universitas Pelita Harapan.

References

- Angraini, Y., Fauziah, S., & Putra, J. L. (2020). Analisis Kinerja Algoritma C4.5 dan Naive Bayes Dalam Memprediksi Keberhasilan Sekolah Menghadapi UN. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer (JITK)*, 5(2), 285–290. <https://doi.org/10.33480/jitk.v5i2.1233>
- Ayodonor - Palang Merah Indonesia. (2016). Palang Merah Indonesia. <https://ayodonor.pmi.or.id/>
- Barus, O. P., Nathasya, C., & Pangaribuan, J. J. (2023). The Implementation of RFM Analysis to Customer Profiling Using K-Means Clustering. *Mathematical Modelling of Engineering Problems*, 10(1), 298–303. <https://doi.org/10.18280/mmep.100135>
- Barus, O. P., Romindo, & Pangaribuan, J. J. (2023). Classification of Hearing Loss Degrees with Naive Bayes Algorithm. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(4), 751–757. <https://doi.org/10.29207/resti.v7i4.4683>
- Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28. <https://doi.org/10.38094/jastt20165>
- Irawan, Y. (2021). Penerapan Algoritma Decision Tree C4.5 Untuk Memprediksi Kelayakan Calon Pendoron Melakukan Donor Darah Dengan Klasifikasi Data Mining. *JTIM : Jurnal Teknologi Informasi Dan Multimedia*, 2(4), 181–189. <https://doi.org/10.35746/jtim.v2i4.75>
- Kaptoge, S., Di Angelantonio, E., Moore, C., Walker, M., Armitage, J., Ouwehand, W. H., Roberts, D. J., Danesh, J., Thompson, S. G., Kaptoge, S., Di Angelantonio, E., Moore, C., Walker, M., Armitage, J., Ouwehand, W. H., Roberts, D. J., Danesh, J., Thompson, S. G., Donovan, J., ... Roberts, D. J. (2019). Longer-term efficiency and safety of increasing the frequency of whole blood donation (INTERVAL): extension study of a randomised trial of 20 757 blood donors. *The Lancet Haematology*, 6(10), e510–e520. [https://doi.org/10.1016/S2352-3026\(19\)30106-1](https://doi.org/10.1016/S2352-3026(19)30106-1)

- Liao, S. H., Widowati, R., & Puttong, P. (2022). Data Mining Analytics Investigate Facebook Live Stream Users' Behaviors and Business Models: The Evidence from Thailand. *Entertainment Computing*, 41, 100478. <https://doi.org/10.1016/j.entcom.2022.100478>
- Maimon, O., & Rokach, L. (2005). Introduction to Knowledge Discovery in Databases. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 1–17). Springer US. https://doi.org/10.1007/0-387-25465-X_1
- Pangaribuan, J. J., & Suharjito. (2014). Diagnosis of Diabetes Mellitus Using Extreme Learning Machine. *2014 International Conference on Information Technology Systems and Innovation, ICITSI 2014 - Proceedings*, 33–38. <https://doi.org/10.1109/ICITSI.2014.7048234>
- Pangaribuan, J. J., Tanjaya, H., & Kenichi. (2017). Mendeteksi Penyakit Jantung Menggunakan Machine Learning dengan Algoritma Logistic Regression. *Journal Information System Development*, 6(2), 40–48.
- Priyasadie, N., & Isa, S. M. (2021). Educational Data Mining in Predicting Student Final Grades on Standardized Indonesia Data Pokok Pendidikan Data Set. *International Journal of Advanced Computer Science and Applications*, 12(12), 212–216. <https://doi.org/10.14569/IJACSA.2021.0121227>
- Ramadani, S., Hidayat, S., & Ramahdanty, R. (2023). Application of Data Mining on Inventory Grouping using Clustering Method. *Jurnal Teknik Informatika C.I.T Medicom*, 15(5), 228–239. <https://doi.org/10.35335/cit.Vol15.2023.608.pp228-239>
- Resti, Y., Aryanto, R., Yahdin, S., & Kresnawati, E. S. (2023). Rain Event Prediction Performance Using Decision Tree Method. *AIP Conference Proceedings*, 2689(1), 120006. <https://doi.org/10.1063/5.0117434>
- Rusyana, N. R., Renaldi, F., & Destiani, D. (2023). Prediction Analysis of Four Disease Risk Using Decision Tree C4.5. *ICCoSITE 2023 - International Conference on Computer Science, Information Technology and Engineering: Digital Transformation Strategy in Facing the VUCA and TUNA Era*, 90–94. <https://doi.org/10.1109/ICCoSITE57641.2023.10127710>
- Sathiyarayanan, P., Pavithra., S., Saranya., M. S., & Makeswari., M. (2019). Identification of Breast Cancer Using The Decision Tree Algorithm. *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, 1–6. <https://doi.org/10.1109/ICSCAN.2019.8878757>
- Sumiati, S., Repi, V. V. R., Hendriyati, P., Anharudin, A., Yusta, A., & Triayudi, A. (2023). Classification of Cardiac Disorders Based on Electrocardiogram Data Using a Decision Tree Classification Approach with the C45 Algorithm. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 12(3), 1128. <https://doi.org/10.11591/ijai.v12.i3.pp1128-1138>
- Syukmana, F., Wahyudi, E., Gata, W., Wahono, H., Febianto, N. I., Kuntoro, A. Y., Nitra, R. O., Effendi, L., Saputra, D. D., & Sulaeman, O. R. (2020). Predicting Relegation Clubs in Italian Serie A with Method based C4.5 Decision Tree Algorithm. *Journal of Physics: Conference Series*, 1471(1), 12016. <https://doi.org/10.1088/1742-6596/1471/1/012016>
- Visa, S., Ramsay, B., Ralescu, A., & Van Der Knaap, E. (2011). Confusion Matrix-based Feature Selection. *Proceedings of The 22nd Midwest Artificial Intelligence and Cognitive Science Conference*.

Blood Donation Classification with Decision Tree Method using C4.5 Algorithm

- Why Blood Donation is Important – and Who Benefits.* (2021, May 19). The American National Red Cross. <https://www.redcrossblood.org/local-homepage/news/article/blood-donation-importance.html>
- Yang, A., Zhang, W., Wang, J., Yang, K., Han, Y., & Zhang, L. (2020). Review on the Application of Machine Learning Algorithms in the Sequence Data Mining of DNA. *Frontiers in Bioengineering and Biotechnology*, 8. <https://doi.org/10.3389/fbioe.2020.01032>
- Yeh, I.-C. (2008). Blood Transfusion Service Center Data Set. In *UCI Machine Learning Repository*. <https://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center>
- Zulfikar, W. B., Gerhana, Y. A., & Rahmania, A. F. (2018). An Approach to Classify Eligibility Blood Donors Using Decision Tree and Naive Bayes Classifier. *6th International Conference on Cyber and IT Service Management, CITSM*, 2018. <https://doi.org/10.1109/CITSM.2018.8674353>