The Utility of Decision Tree and Analytics Hierarchy Process in Prioritizing of Social Aid Distribution due to Covid-19 Pandemic in Indonesia

Saucha Diwandari[1], Enny Itje Sela[1*], Briyan Efflin Syahputra[1], Nathaniela Aptanta Parama[1] & Anindita Septiarini[2]

[1]Faculty of Science and Engineering, Universitas Teknologi Yogyakarta, Jalan Siliwangi, Ring Road Road, Jombor, Sleman, Yogyakarta, 55528, Indonesia
[2]Department of Computer Science, Universitas Mulawarman, Jalan Kuaro, Gn. Kelua, Kec. Samarinda Ulu, Kota Samarinda, Kalimantan Timur, 75123, Indonesia
*E-mail: ennysela@uty.ac.id

Abstract. The Indonesian government provided various social assistance programs to local governments during Covid-19. One of the difficulties for the local governments in determining candidates for social aid is ensuring that the number of candidates is in balance with the available quota. Therefore, the local governments must select the most eligible candidates. This study proposes a priority model that can provide recommendations for candidates who meet the criteria for social assistance. The six parameters used in this study were: number of dependents, occupation, income, age, Covid status, and citizen status. The model operates in two stages, namely classification followed by ranking. The classification stage is conducted using a decision tree, while the ranking stage is performed conducted using the Analytical Hierarchy Process (AHP) algorithm. The decision tree separates two classes, namely, eligible and non-eligible. In addition, the classification process is also used to determine the dominant attributes and played a role in the modeling. The proposed model generates a list of the most eligible candidates based on our research. These are sorted by weight from greatest to most eligible using five dominant parameters: number of dependents, income, age, Covid status, and citizen status.

Keywords: *analytical hierarchy process; classification; decision tree; ranking; social aid.*

1 Introduction

The Indonesian government provided various social assistance programs during the Covid-19 pandemic through the Jaring Pengaman Sosial. These various programs are aimed at helping people who are affected by the pandemic [1,2]. Local governments can apply for social aid funds. The number of candidates and the type of aid provided by a local government is in accordance with the needs proposed by the local government in question. Furthermore, the local government will distribute it according to what has been proposed beforehand. One of the

Received July 7th, 2022, Revised October 27th, 2022, Accepted for publication January 11th, 2023.
 ISSN: 2337-5787, DOI: 10.5614/itbj.ict.res.appl.2023.17.1.6

difficulties of local governments is prioritizing eligible candidates who genuinely need the aid, considering the limited quota provided, as the number of candidates will keep increasing while the quota remains same. This growing imbalance necessitates the local governments to select the most eligible candidates [3]. To overcome this problem, our proposed model processes the candidates' data in two stages, namely classification followed by ranking. The classification stage is used to divide the candidates into two classes, namely eligible (class A) and not eligible (class B). In addition, this stage also identifies the dominant parameters in the model of social aid distribution. Only the data and dominant parameters of class A are classified and processed in the ranking stage to obtain the recommended candidates. Thus, candidates who most need social aid will receive priority. Based on field experience, this study aimed to explore the utility of the C4.5 decision tree and Analytical Hierarchy Process (AHP) algorithms. Decision tree is used to classify the candidates, while AHP is used to produce the most eligible candidates. In previous works, decision tree has resulted in accuracy above 80% in the process of selecting dominant features in the health sector, [4,5], classification in the social field [6-8], diagnosis of diabetes [9], and some other fields [10,11].

Previous research on candidates for social aid, humanitarian aid, and scholarships have been widely carried out using various methods. In education, several studies have also produced good models, especially related to scholarships. Reference [12] discusses prospective university scholarship recipients using the Technique For Others Reference by Similarity to Ideal Solution (TOPSIS) and product weight. The recommendation system for Bidik Misi scholarship candidates was developed using the Mamdani Fuzzy Inference System with the Elbow method, K-means clustering, and Pearson's correlation [13]. Reference [14] was conducted to determine the priority of livelihood activities towards poverty reduction in developing countries. Reference [15] also conducted research to provide candidates for Movement for Foster Parents scholarships with three methods, namely AHP, Support Vector Machine (SVM), and TOPSIS, which also gives a ranked recommendation. However, these studies only used ranking and did not carry out classification. Reference [13] was improved by [16,17] to increase its accuracy using a combination of Backpropagation, Mamdani FIS with the Elbow method, K-means clustering, and Pearson's correlation. The parameters used in the research above were grouped into economic status and academic status. The three studies implemented classification and ranking, achieving very good accuracy but they used all existing parameters, which could result in biased or invalid results [18].

The purpose of the present study was to develop a candidate prioritization model for social aid distribution due to the COVID-19 pandemic. The parameters to be used to generate the model were selected, meaning that these parameters are dominant or play an important role in the model.

2 Material and Method

The research data were primary data provided by the local government of Sukoharjo, Ngaglik, Sleman, Daerah Istimewa Yogyakarta, Indonesia. The number of data obtained in this study was 550 data, divided into 215 data in class A and 335 data in class B. The parameters used in this study were in accordance with the criteria from the government for distribution of social aid due to the pandemic. These parameters were: number of dependents (C1), occupation (C2), income (C3), age (C4), Covid status (C5), resident status (C6), and class label (C7). The model to produce a candidate's ranking for social aid is presented in Figure 1. Generally, the process is divided into two stages, namely classification and ranking.

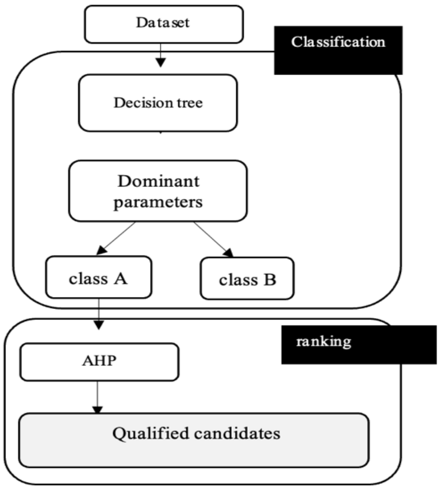

Figure 1 Proposed model.

2.1 Decision Tree Classification

The input for the classification process is a dataset consisting of 7 columns and 550 rows. An explanation of each parameter is presented in Table 1. Furthermore, the dataset is processed using the decision tree method with gain ratio selection criteria. This method is generally referred to as the C4.5 algorithm and the results of the algorithm are visualized in the form of a decision tree. This method was chosen because it can process discrete and continuous data and has been used in various previous studies with high accuracy [18-25]. This method can also handle imbalanced data [22,26,27].

Table 1 Data description.

Code	Parameter	Description
C1	Number of dependents (person)	[0…6]
C2	Occupation	[A = government employees, B = private employees, C = farmer, D = entrepreneur, E = unemployed]
C3	Income (millions of Rupiah)	[0 …. 5]
C4	Age (year)	[20 …. 80]
C5	Covid status	[SH = not affected, SK = affected]
C6	Resident status	[PD = resident, PT = non-resident]
C7	Class label	[A = eligible, B = non-eligible]

The C4.5 algorithm was proposed by Ross Quinlan. This algorithm is an enhanced version of ID3. The enhancement that distinguishes ID3 from C4.5 is that C4.5 can handle numeric parameters or features, pruning trees, and derive a set of rules.

The C4.5 algorithm uses gain ratio criteria in selecting features that are node splits in the tree. For building a decision tree, the first thing to do is to select attributes as the roots. Then a branch is generated for each value of the root. The next step is to divide the cases in branches and then repeat the process for each branch until all the cases in the branch have the same class. Gain ratio (GR) takes the information gain (IG) and normalizes it with entropy [2,3]. The formula for entropy (H) is given by Eq. (1):

$$H = -\sum_{i=1}^{m}(p_i\ log_2 p_i) \tag{1}$$

where p_i is the proportion of classes in the dataset.

IG equals the entropy subtracted by the weighted sum of the sub-entropies. The weights equals the proportion of samples being moved to the sub-dataset. The formula for IG is shown in Eq. (2):

$$IG = H - \left(\sum_{j=1}^{v}\frac{|D_j|}{|D|} * H_j\right) \tag{2}$$

where:

1. D is the dataset.
2. D_j is the j-th sub-dataset after being split.
3. $|D|$ and $|D_j|$ are the numbers of samples to the original dataset and the sub-dataset, respectively.
4. H_j is the entropy of the j-th sub-dataset.

The formula for GR is shown in Eq. (3):

$$GR(A) = \frac{Gain(A)}{SplitInfo(A)} \tag{3}$$

To calculate the split entropy (SplitInfo), the following equation is used:

$$SplitInfo(A) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} * log_2 \frac{|D_j|}{|D|} \tag{4}$$

The resulting decision tree can separate the data into class A and class B. Furthermore, the dominant parameters and data in class A become the input for the ranking process. In the ranking stage, the data entered in class A will be processed using AHP to obtain a list for the order of the distribution of the social aid.

2.2 AHP Ranking

Analytical Hierarchy Process (AHP) is a DSS method. DSS has been widely used to solve problems in different fields, such as assessment and installation projects [29,30], real-time systems [31,32], supply-chain management [33,34], scheduling [35-37], hazard mitigation [38,39], and energy transition [40].

At this stage, the decision support system architecture (Figure 2) aims to describe the design of the data management, model management, knowledge-based system, and user interface. The data management subsystem uses an internal dataset from the local government.

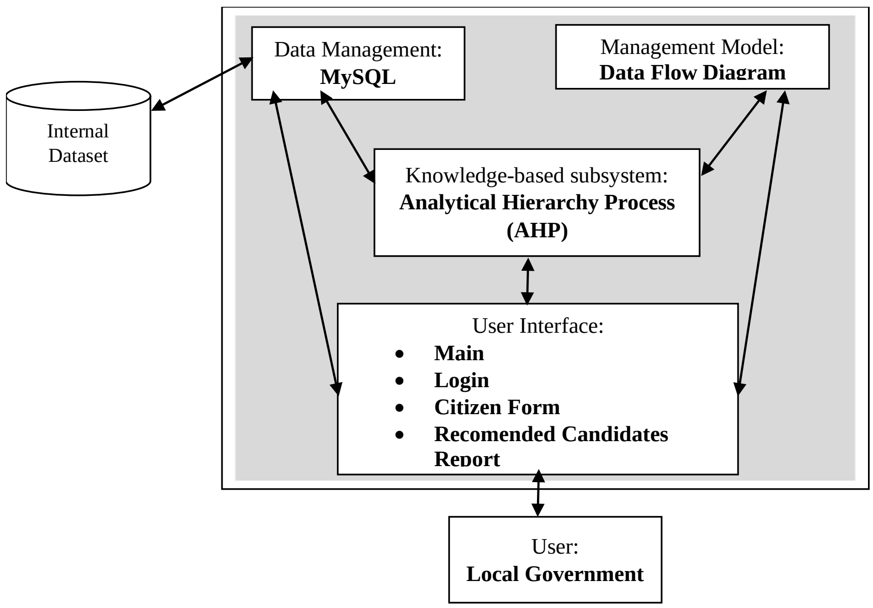

Figure 2 Architecture of the proposed system.

The dataset was saved in a MySQL database, including data on criteria, sub-criteria, alternatives, criteria and sub-criteria weights, comparison criteria, and comparisons of sub-criteria. The model management sub system uses a data flow diagram (DFD. The knowledge-based subsystem uses the Analytical Hierarchy Process (AHP) method to rank the candidates.

The AHP process is carried out using the following procedure [31,41-43]:

1. The first stage of the AHP is defining the problem. In AHP, the dominant parameters generated in the classification stage are referred to as criteria, while the alternatives are the eligible candidates (class A).
2. Preparing a pairwise comparison matrix constitutes the creation of an n*n dimensional pairwise comparison matrix of the conditioning factors.
3. Determining a consistency ratio (CR) index is used to examine the consistency pairwise comparison matrix (Table 2).

Table 2 Random inconsistency indices.

n	3	4	5	6	7	8	9	10
R1	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

CR is the consistency index (CI) divided by the random index (RI). The formula for CR is as follows:

$$CR = \frac{CI}{RI} \tag{5}$$

CI is the consistency index and RI is the random inconsistency index. The formula for CI is as follows:

$$CI = \frac{(\lambda - n)}{(n-1)} \tag{6}$$

The flowchart of the AHP ranking system can be seen in Figure 3. The inputs of the application are the period and quota for the limitation of the number of qualified candidates, the dataset of the classification results, and the priority scale of the criteria. Furthermore, the AHP procedure is applied accordingly, starting with the creation of a pairwise comparison matrix, synthesis, computing λ_{max}, CI and CR, and checking the value of CR. If CR > 0.1, then the AHP process will start from the beginning. However, if the CR value <= 0.1, then there is consistency and the system will provide a list prioritizing the candidates based on the largest to the smallest weight. The system is also able to provide the number of potential recipients according to the available quota.

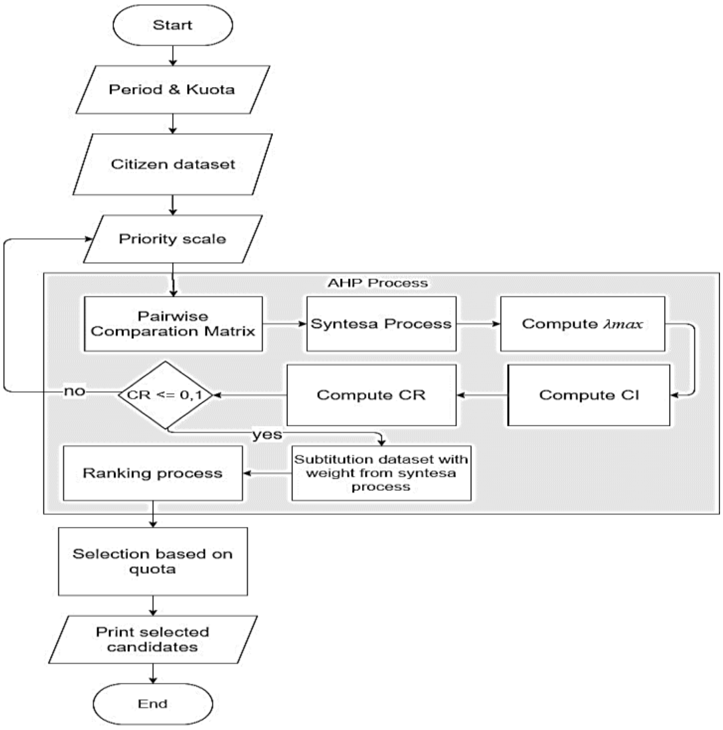

Figure 3 Flowchart of the AHP system.

3 Result and Discussion

3.1 Result

3.1.1 Classification

The dataset consisting of 6 parameters, 1 class label, and 550 rows was trained and tested using the RapidMiner software with decision tree operators and gain ratio selection criteria. Dataset validation was carried out using cross validation (k = 25). The decision tree obtained at this stage can be seen in Figure 4. From the figure, it can be seen that there were 5 dominant parameters in the classification process, namely C1, C3, C4, C5 and C6. These parameters appear as nodes in the decision tree. Table 3 shows the final performance of the classification using the C4.5 algorithm in the form of a confusion matrix. Classification performance was measured in terms of the percentage of accuracy. The value of accuracy is better if it is closer to 100%.

From the confusion matrix in Table 3, it can be seen that the number of correct data that was successfully predicted was 449 out of 550 data. This means that

81.64% of the data was successfully predicted by the model. In class A, 178 of 215 data were predicted correctly, while in class B the number of data that was correctly predicted was 271 out of 335 data. This shows that the recall value (sensitivity), which is the ratio of the correct predictions for class A compared to the overall positive data, was 82.79%. The specificity value obtained was 80.89% and the average of F1-score was 77.89%. The data and parameters of the results of this stage were processed in the next stage.

```
C3 > 2250000
|   C1 > 2.500
|   |   C3 > 2625000
|   |   |   C1 > 3.500
|   |   |   |   C5 = SH: 0 {B}
|   |   |   |   C5 = SK: 1 {A}
|   |   |   C1 ≤ 3.500: 0 {B}
|   |   C3 ≤ 2625000: 0 {B}
|   C1 ≤ 2.500: {B}
C3 ≤ 2250000
|   C3 > 1375000
|   |   C1 > 3.500: 1 {A}
|   |   C1 ≤ 3.500
|   |   |   C3 > 1875000
|   |   |   |   C4 > 46.500: 0 {B}
|   |   |   |   C4 ≤ 46.500: 1 {A}
|   |   |   C3 ≤ 1875000: 0 {B}
|   C3 ≤ 1375000
|   |   C6 = PD
|   |   |   C4 > 35.500: 1 {A}
|   |   |   C4 ≤ 35.500: 0 {B}
|   |   C6 = PT
|   |   |   C1 > 4.500: 1 {A}
|   |   |   C1 ≤ 4.500: 0 {B}
```

Figure 4 Decision tree.

Table 3 Confusion matrix.

observed / **predicted**	Class A	Class B	Total
Class A	178	64	242
Class B	37	271	308
Total	215	335	550

3.1.2 Ranking

The ranking stage is performed using the AHP method to produce a priority list of candidates. The implementation of AHP in this study was carried out by software developed by the researchers. The input data was a classification result dataset consisting of 5 criteria and 178 alternatives. Figure 5 expresses a map hierarchy showing the proposed goals, criteria, and alternatives. The criteria used in this stage were the dominant parameters generated by the decision tree, namely

C1, C3, C4, C5 and C6. The alternatives are the potential beneficiaries (class A), symbolized by A1, A2,…, Ai (i = 178), where Ai indicates the name of the ith candidate.

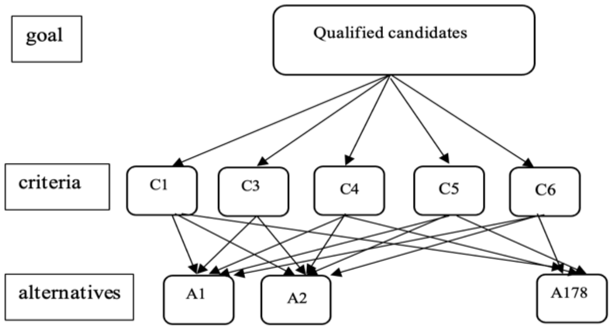

Figure 5 Hierarchical map of the problem.

The value of the criteria data used to process AHP can be seen in Table 4. The values in Table 4 are slightly different from the data values in Table 1 because processing using AHP requires categorical data type.

Table 4 AHP's criteria.

Code	Criteria	Description
C1	Number of dependents (persons)	0-1; 2-3; >4
C3	Income (millions of Rupiah)	0 – 1; 1 – 2.5; 2.5-5; >5
C4	Age (year)	15 – 24; 25 – 34; 35 – 44; 45 – 54; 55 – 64; >64
C5	Covid status	affected; not affected
C6	Resident status	resident; non-resident

3.2 Discussion

Table 5 presents the 5 x 5 pair comparison result of the criteria in view of the overall goal of the assessment. This study resulted in a consistency ratio (CR) value of 0.082. Table 6 shows an example of the ranking results along with the final score obtained for each alternative. Rank 1 has the highest priority, meaning that it takes precedence over the sequences after it. The order is selected based on the total value from the largest to the smallest.

With the increasing number of social aid programs provided to residents through local governments, an adequate model is needed to ensure that the social aid is on target. Thus, the aid provided can have a positive impact on people's lives and encourage national economic growth. This research has succeeded in producing a list of the most eligible candidates recommend for social aid from the government. This list can be used as a reference for local governments to propose or distribute social aid. This list ensures that the prioritized candidates are genuinely in need of one particular social aid, also candidates who only receive one type of aid (if the criteria for social aid have these conditions).

Table 5 Paired matrix and its normalized values.

criteria	Paired Matrix					Normalized					Total	Relative Weight
	C1	C3	C4	C5	C6	C1	C3	C4	C5	C6		
C1	1	0.25	3	5	4	0.18	0.13	0.40	0.37	0.22	1.28	0.258
C3	4	1	3	4	7	0.70	0.51	0.40	0.29	0.39	2.28	0.456
C4	0.33	0.33	1	3	4	0.06	0.17	0.33	0.22	0.22	0.80	0.161
C5	0.2	0.25	0.33	1	2	0.04	0.13	0.04	0.07	0.11	0.39	0.078
C6	0.25	0.14	0.25	0.5	1	0.04	0.07	0.03	0.04	0.06	0.24	0.048
											CR =	0.082

Table 6 Candidate's ranking (quota = 15).

R	QC	C1	C3	C4	C5	C6	Total
1	A50	0.17	0.28	0.06	0.02	0.04	0.57
2	A79	0.17	0.28	0.06	0.02	0.04	0.57
3	A133	0.17	0.28	0.06	0.02	0.04	0.57
4	A137	0.17	0.28	0.06	0.02	0.04	0.57
5	A144	0.17	0.28	0.06	0.02	0.04	0.57
6	A152	0.17	0.28	0.06	0.02	0.04	0.57
7	A1	0.17	0.28	0.02	0.06	0.04	0.57
8	A8	0.17	0.28	0.02	0.06	0.04	0.57
9	A99	0.17	0.28	0.02	0.06	0.04	0.57
10	A37	0.17	0.28	0.02	0.06	0.04	0.56
11	A128	0.17	0.28	0.02	0.06	0.04	0.56
12	A71	0.17	0.28	0.06	0.02	0.01	0.55
13	A170	0.17	0.28	0.04	0.02	0.04	0.55
14	A178	0.17	0.28	0.04	0.02	0.04	0.55
15	A45	0.17	0.28	0.02	0.02	0.04	0.53

R = Ranking; QC = Qualified Candidate

The two stages of this research were classification and ranking. In the classification stage, this study also made a comparison using three other methods, namely Naïve Bayes, Neural Network and Logistic Regression. Table 7 presents the results of a comparison of the accuracy values of the respective classification processes.

The highest accuracy results were obtained using the C4.5 method, i.e., 81.64%. As many as 178 candidates were declared eligible to receive social aid from the local government and were then further processed to determine recommendations for prospective beneficiaries using AHP. In the ranking process, (Table 5), it can be seen that C3 and C1 are the two major factors that influenced the prioritization in the social aid distribution. These are represented by relative weights of 0.456

and 0.258, respectively. This means that the type of income and the number of dependents are factors that influence the distribution model of social aid.

Table 7 Comparation of classification's accuracy.

Method	Accuracy (%)
C4.5	81.64
Naïve Bayes	78.18
Neural Network	78.91
Logistic Regression	70.55

The CR value of 0.082 (less than 0.1) in this research means that the criteria and data processed using this AHP are reliable [30].

Table 8 shows an accuracy comparison between this work and similar works.

Table 8 Comparison between this work and similar works.

Ref.	Classification Process		Selection Process		Contribution
	Method	Accuracy (%)	Method	Accuracy (%)	
[4]	-	-	Mamdani FIS with the Elbow method, K-means clustering	71.4	Perform recommended candidates without classification
[5]	-	-	FMADM with TOPSIS and WP	-	Perform recommended candidates without classification
[6]	SVM	89.9	FMADM with AHP and TOPSIS	-	Perform classification using SVM and recommended candidates using some methods
[7]	BPNN	91.3	Mamdani FIS with the Elbow method, K-means clustering, Pearson's correlation, and matching process	85.6	Perform classification using backpropagation neural network (BPNN) and recommended candidates using several methods
This work	C4.5	81.64	AHP	-	Perform classification using C4.5 and recommendation of candidates based on dominant parameters using AHP

From Table 8 it can be seen that this study used only two methods, namely C4.5 and AHP for classification and ranking, respectively, compared to the other studies, who used a combination of more than three methods [15,16]. However, the classification accuracy of this study was lower than that of these other studies. References [12] and [13] provide recommendations without classification, so that the application of their methods is not able to solve the problem of this research.

4 Conclusion

The most eligible candidates for social aid distribution were successfully obtained based on dominant parameters using a combination of decision tree and AHP algorithms. The combination of the two methods produced an accuracy value of 81.64% in the classification stage. This model can be used as an alternative for making a list of eligible candidates for social aid distribution. Future research will be done to develop an automated recommendation system using the proposed model.

Acknowledgement

The authors would like to thank the *Lembaga Pengelola Dana Pendidikan* (LPDP), Ministry of Finance of the Republic of Indonesia for funding this research in the *Program Riset Keilmuan scheme of Riset Mandiri Dosen* in 2022.

References

[1] Susilawati, S., Falefi, R. & Purwoko, A., *Impact of COVID-19's Pandemic on the Economy of Indonesia*, Bp. Int. Res. Crit. Inst. BIRCI-J. Humanit. Soc. Sci., **3**(2), pp. 1147-1156, 2020. DOI: 10.33258/birci. v3i2.954.

[2] Firasari, E., Khasanah, N., Khultsum, U., Kholifah, D.N., Komarudin, R. & Widyastuty, W., *Comparation of K-Nearest Neighbor (K-NN) and Naive Bayes Algorithm for the Classification of the Poor in Recipients of Social Assistance*, J. Phys. Conf. Ser., **1641**(1), 012077, 2020. DOI: 10.1088/1742-6596/1641/1/012077.

[3] Mohd, S., Syazli Fathi, M. & Nahar Harun, A., *Information Management for Humanitarian Aid Distribution System in Malaysia*, IOP Conf. Ser. Mater. Sci. Eng., **513**(1), 012012, 2019. DOI: 10.1088/1757-899X/513/1/012012.

[4] Sela, E.I. & Pulungan, R., *Osteoporosis Identification Based on the Validated Trabecular Area on Digital Dental Radiographic Images*, Procedia Comput. Sci., **157**, pp. 282-289, 2019. DOI: 10.1016/j.procs.2019.08.168.

[5] Sela, E.I. & Sutarman, *Extracting the Potential Features of Digital Panoramic Radiograph Images by Combining Radio Morphometry Index, Texture Analysis and Morphological Features*, J. Comput. Sci., **14**(2), pp. 144-152, 2017. DOI: 10.3844/jcssp.2018.144.152.

[6] Sela, E.I., *Determination of Indicators that Play a Role in Poverty Identification using Data Mining*, J. Ris. Drh. Ed. Khusus, **2015**, pp. 16-32, 2015. (Text in Indonesian)

[7] Wang, X., Zhou, C. & Xu, X., *Application of C4.5 Decision Tree for Scholarship Evaluations*, Procedia Comput. Sci., **151**(2018), pp. 179-184, 2019. DOI: 10.1016/j.procs.2019.04.027.
[8] Majid, M.A. & Sela, E.I., *Performance Evaluation of Combined Consistency-Based Subset Evaluation and Artificial Neural Network for Recognition of Dynamic Malaysian Sign Language*, J. Theor. Appl. Inf. Technol., **95**(11), pp. 248-2496, 2017.
[9] Liu, J., Ning, B. & Shi, D., *Application of Improved Decision Tree C4.5 Algorithms in the Judgment of Diabetes Diagnostic Effectiveness*, J. Phys. Conf. Ser., **1237**(2), 022116, 2019. DOI: 10.1088/1742-6596/1237/2/022116.
[10] Diwandari, S. & Hidayat, A.T, *Comparison of Classification Performance Based on Dynamic Mining of User Interest Navigation Pattern in e-Commerce Websites*, J. Phys. Conf. Ser., **1844**(1), 012025, Mar. 2021. DOI: 10.1088/1742-6596/1844/1/012025.
[11] Fachrie, M. & Ardiani, F., *Predictive Model for Regional Elections Results based on Candidate Profiles*, in 2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), pp. 247-252, Oct. 2021. DOI: 10.23919/EECSI53397.2021.9624256.
[12] 'Uyun, S. & Riadi, I., *A Fuzzy Topsis Multiple-Attribute Decision Making for Scholarship Selection*, Telkomnika, **9**(1), pp. 37-46, 2011. DOI: 10.12928/telkomnika. v9i1.643.
[13] Latumakulita, L.A., Purnama, F., Usagawa, T., Paturusi, S. & Prima, D.A., *Indonesia Scholarship Selection Framework Using Fuzzy Inferences System Approach: Case Study: 'Bidik Misi' Scholarship Selection*, in 2016 International Conference on Information & Communication Technology and Systems (ICTS), pp. 107-113, 2016. DOI: 10.1109/ICTS.2016.7910282.
[14] Baffoe, G., *Exploring The Utility of Analytic Hierarchy Process (AHP) In Ranking Livelihood Activities for Effective and Sustainable Rural Development Interventions in Developing Countries*, Eval. Program Plann., **72**, pp. 197-204, 2019. DOI: 10.1016/j.evalprogplan.2018.10.017.
[15] Putra, M.G.L., Ariyanti, W. & Cholissodin, I., *Selection and Recommendation Scholarships Using AHP-SVM-TOPSIS*, J. Inf. Technol. Comput. Sci., **1**(1), pp. 1-13, 2016. DOI: 10.25126/jitecs.2016111.
[16] Latumakulita, L.A. & Usagawa, T., *Indonesia Scholarship Selection Model Using a Combination of Back-Propagation Neural Network and Fuzzy Inference System Approaches*, Int. J. Intell. Eng. Syst., **11**(3), pp. 79-90, 2018. DOI: 10.22266/IJIES2018.0630.09.
[17] Latumakulita, L.A. & Usagawa, T., *A Combination of Backpropagation Neural Network on Fuzzy Inference System Approach in Indonesia Scholarship Selection Process: Case Study: 'Bidik Misi' Scholarship Selection*, in 2017 13th International Conference on Natural Computation,

Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 1309-1314, 2017. DOI: 10.1109/FSKD.2017.8392955.
[18] Shah, D., Wang, J. & He, Q.P., *Feature Engineering in Big Data Analytics for Iot-Enabled Smart Manufacturing – Comparison between Deep Learning and Statistical Learning*, Comput. Chem. Eng., **141**, 106970, 2020. DOI: 10 1016/j.compchemeng.2020.106970.
[19] Song, Y.Q., Yao, X., Liu, Z., Shen, X. & Mao, J., *An Improved C4.5 Algorithm in Bagging Integration Model*, IEEE Access, **8**, pp. 206866-206875, 2020. DOI: 10.1109/ACCESS.2020.3032291.
[20] Liu, W., Fan, H. & Xia, M., *Step-Wise Multi-Grained Augmented Gradient Boosting Decision Trees for Credit Scoring*, Eng. Appl. Artif. Intell., **97**, 104036, 2021, DOI: 10.1016/j.engappai.2020.104036.
[21] Muslim, M.A., Nurzahputra, A. & Prasetiyo, B., *Improving Accuracy of C4.5 Algorithm Using Split Feature Reduction Model and Bagging Ensemble for Credit Card Risk Prediction*, in 2018 International Conference on Information and Communications Technology (ICOIACT), pp. 141-145, 2018. DOI: 10.1109/ICOIACT.2018.8350753.
[22] Sela, E.I., Hartati, S., Harjoko, A., Wardoyo, R. & Mudjosemedi, M., *Segmentation on the Dental Periapical X-Ray Images for Osteoporosis Screenin*g, Int. J. Adv. Comput. Sci. Appl., **4**(7), pp. 147-151, 2013. DOI: 10.14569/ijacsa.2013.040720.
[23] Fitri, V.A., Andreswari, R. & Hasibuan, M.A., *Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia Using Naïve Bayes, Decision Tree, and Random Forest Algorithm,* Procedia Comput. Sci., **161**, pp. 765-772, 2019. DOI: 10.1016/j.procs.2019.11.181.
[24] Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H. & Hong, H., *Application of Alternating Decision Tree with Adaboost and Bagging Ensembles for Landslide Susceptibility Mapping*, Catena, **187**, 104396, 2020. DOI: 10.1016/j.catena.2019.104396.
[25] Mienye, I.D., Sun, Y. & Wang, Z., *Prediction Performance of Improved Decision Tree-Based Algorithms: A Review*, Procedia Manuf., **35**, pp. 698-703, 2019. DOI: 10.1016/j.promfg.2019.06.011.
[26] Lee, J.S., *AUC4.5: AUC-Based C4.5 Decision Tree Algorithm for Imbalanced Data Classification*, IEEE Access, **7**, pp. 106034-106042, 2019. DOI: 10.1109/ACCESS.2019.2931865.
[27] Trabelsi, A., Elouedi, Z. & Lefevre, E., *Decision Tree Classifiers for Evidential Attribute Values and Class Labels*, Fuzzy Sets Syst., **366**, pp. 46-62, 2019. DOI: 10.1016/j.fss.2018.11.006.
[28] Wang, H.B. & Gao, Y.J., *Research on C4.5 Algorithm Improvement Strategy based on MapReduce*, Procedia Comput. Sci., **183**, pp. 160-165, 2021. DOI: 10.1016/j.procs.2021.02.045.

[29] Maceika, A., Bugajev, A., Šostak, O.R. & Vilutienė, T., *Decision Tree and Ahp Methods Application for Projects Assessment: A Case Study*, Sustain. Switz., **13**(10), pp. 1-33, 2021. DOI: 10.3390/su13105502.

[30] Maceika, A., Bugajev, A. & Šostak, O.R., *The Modelling of Roof Installation Projects Using Decision Trees and the AHP Method*, Sustain. Switz., **12**(1), pp. 1-21, 2020. DOI: 10.3390/SU12010059.

[31] Polo-Castañeda, M., Gómez-Rojas, J. & Linero-Cueto, J., *Application of AHP and GIS for Determination of Suitable Wireless Sensor Network Zones for Oceanographic Monitoring in the South Caribbean Sea Upwelling Zone*, Int. J. Adv. Sci. Eng. Inf. Technol., **11**(5), pp. 1696-1703, 2021. DOI: 10.18517/ijaseit.11.5.14293.

[32] Fertier, A., Barthe-Delanoë, A. M., Montarnal, A., Truptil, S. & Bénaben, F., *A New Emergency Decision Support System: The Automatic Interpretation and Contextualisation of Events to Model a Crisis Situation in Real-Time*, Decis. Support Syst., **133**, no. December 2020, 2020, doi: 10.1016/j.dss.2020.113260.

[33] Allaoui, H., Guo, Y. & Sarkis, J., *Decision Support for Collaboration Planning in Sustainable Supply Chains*, J. Clean. Prod., **229**, pp. 761-774, 2019. DOI: 10.1016/j.jclepro.2019.04.367.

[34] Butdee, S. & Phuangsalee, P., *Uncertain Risk Assessment Modelling for Bus Body Manufacturing Supply Chain Using AHP and Fuzzy AHP*, Procedia Manuf., **30**, pp. 663-670, 2019. DOI: 10.1016/j.promfg.2019.02.094.

[35] Kandakoglu, A., Sauré, A., Michalowski, W., Aquino, M., Graham, J. & McCormick, B., *A Decision Support System for Home Dialysis Visit Scheduling and Nurse Routing*, Decis. Support Syst., **130**, November 2019, 113224, 2020. DOI: 10.1016/j.dss.2019.113224.

[36] Wolnowska, A.E. & Konicki, W., *Multi-Criterial Analysis of Oversize Cargo Transport through the City, Using The AHP Method*, Transp. Res. Procedia, **39**, pp. 614-623, 2019, DOI: 10.1016/j.trpro.2019.06.063.

[37] Petruni, A., Giagloglou, E., Douglas, E., Geng, J., Leva, M.C. & M. Demichela, *Applying Analytic Hierarchy Process (AHP) to Choose a Human Factors Technique: Choosing the Suitable Human Reliability Analysis Technique for the Automotive Industry*, Saf. Sci., **119**, pp. 229-239, 2019. DOI: 10.1016/j.ssci.2017.05.007.

[38] Amir-Heidari, P. & Raie, M., *Response Planning for Accidental Oil Spills in Persian Gulf: A Decision Support System (DSS) based on Consequence Modeling*, Mar. Pollut. Bull., **140**, pp. 116-128, 2019, DOI: 10.1016/j.marpolbul.2018.12.053.

[39] Xu, H., Windsor, M., Muste, M. & Demir, I., *A Web-Based Decision Support System for Collaborative Mitigation of Multiple Water-Related Hazards Using Serious Gaming*, J. Environ. Manage., **255**(1), 109887, 2020. DOI: 10.1016/j.jenvman.2019.109887.

[40] Shah, S.A.A., Solangi, Y.A. & Ikram, M., *Analysis of Barriers to the Adoption of Cleaner Energy Technologies in Pakistan using Modified Delphi and Fuzzy Analytical Hierarchy Process*, J. Clean. Prod., **235**, pp. 1037-1050, 2019. DOI: 10.1016/j.jclepro.2019.07.020.

[41] Hammami, S., *Application of the GIS based Multi-criteria Decision Analysis and Analytical Hierarchy Process (AHP) in the Flood Susceptibility Mapping (Tunisia)*, Arab. J. Geosci., **12**(21), 653, 2019, DOI: 10.1007/s12517-019-4754-9.

[42] Li, X. & Mo, X., *Application of AHP Based on Mathematical Operational Research in Teaching Evaluation System*, J. Phys. Conf. Ser., **1650**(3), 032015, 2020. DOI: 10.1088/1742-6596/1650/3/032015.

[43] Syahputra, F., Muslim, A.M., Talaat, W.I.A.W. & Irsalinda, N., *Analytical Hierarchy Process (AHP) in Selecting Suitable Marine Protected Area (MPA) site in Pulo Breuh (Breuh Island), Indonesi*a, J. Phys. Conf. Ser., **1373** (1), 012005, 2019. DOI: 10.1088/1742-6596/1373/1/012005.

[44] 'Uyun, S. & Riadi, I., *A Fuzzy Topsis Multiple-Attribute Decision Making for Scholarship Selection*, Telkomnika, **9**(1), pp. 37-46, 2011, DOI: 10.12928/telkomnika. v9i1.643.

[45] Putra, M.G.L., Ariyanti, W. & Cholissodin, I., *Selection and Recommendation Scholarships Using AHP-SVM-TOPSIS*, J. Inf. Technol. Comput. Sci., **1**(1), pp. 1-13, 2016. DOI: 10.25126/jitecs.2016111.

[46] Latumakulita, L.A. & Usagawa, T., *Indonesia Scholarship Selection Model Using a Combination of Back- Propagation Neural Network and Fuzzy Inference System Approaches*, Int. J. Intell. Eng. Syst., **11**(3), pp. 79-90, 2018. DOI: 10.22266/IJIES2018.0630.09.