Integration of the Hybrid Decision Support System and Machine Learning Algorithm to Determine Government Assistance Recipients: A Case Study in the Indonesian Funding Program

Indra Rusyadi Adiwijaya[1,✉]**, Sapto Wahyu Indratno**[1]**, Manahan Siallagan**[2]**, Agus Widodo**[3]**, Eka Gandara**[3]

[1]Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Indonesia
[2]School of Business and Management, Institut Teknologi Bandung, Indonesia
[3]National Research and Innovation Agency, Indonesia

indra.rusyadi@gmail.com[✉], 20921001@mahasiswa.itb.ac.id[✉], sapto@math.itb.ac.id, manahan@sbm-itb.ac.id

Abstract

The Indonesian government provides incentives to facilitate community development through various funding programs to improve the economy and restore the national economy. However, there were many obstacles in determining the proper target beneficiaries. This study aims to assist decision-makers in determining targeted and accountable beneficiary candidates. In this study, a hybrid Analytical Hierarchy Process (AHP) method with Simple Additive Weighting (SAW) was used and integrated with machine learning modeling using Logistic Regression (LR). The AHP approach is used to determine the weight of each criterion, and the SAW method is used to sort out each available alternative with the help of an expert team's assessment. Instead, the LR method is used for the predictive analysis and classification of the resulting data.

1 Introduction

Currently, the Indonesian government, both central and regional, is committed to implementing a government based on the norms of good governance, including system governance, methods and work procedures that adhere to good governance principles and are transparent, efficient, effective, and quantifiable. This is mentioned in the Grand Design of Bureaucratic Reform Presidential Regulation of the Republic of Indonesia No. 81 of 2010. The development of excellent governance with a competent and high integrity government bureaucracy is the goal for 2025 [11]. The effective use of resources and better informed decision-making are encouraged by good governance in the public sector [14]. Public services are a benchmark for the success of carrying out tasks and measuring government performance through the bureaucracy. Public services as the primary mover are also considered necessary by all actors from the elements of good governance. The government must thus execute effective governance, which is backed by an accountability framework, appropriate and accurate information, and efficiency in the management of resources and the provision of public services [30].

On the other hand, the National Research and Innovation Agency (BRIN), one of the government institutions, is committed to implementing the principles of good governance in its institutional control. This is implied in the direction and target of BRIN Head in 2022 [12]. In public services, BRIN has pioneered the innovation ecosystem through the synergy of various parties through its program in the form of an Innovating Village, which is a collaboration between BRIN and the community and local government. This program is an incentive to facilitate development for rural communities through legal entities or those determined by communal/community-based authorized institutions that can be used to increase the added value of innovation-based superior products or services to contribute to economic improvement and community economic recovery [13].

However, an essential part of the village program process to innovate is determining the recipients of incentives who must be well-targeted and accountable. A multi-criteria decision support system, also known as Multi-Criteria Decision Making (MCDM), is required to complete the process of selecting and evaluating incentive recipients since it contains several components or criteria that are being evaluated [28]. Many popular methods are used for selection and evaluation problems with the MCDM approach, such as Analytical Hierarchy Process (AHP) Zhao et al. [33], Analytical Network Process (ANP) Rodrigues et al. [23], Elimination et Choix Traduisant la Realité (ELECTRE) Azzeddine et al. [4], Preference Ranking Organization Method for Enrichment of Evaluations (PROMETHEE) Lopes and Rodriguez-Lopez [18], Technique for Order of Prefer-

ence by Similarity to Ideal Solution (TOPSIS) Dursun and Oguncl [8] and VIseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR) Kim and Ahn [17]. In a study conducted by Azhar et al. [3] reviewing and surveying the use of these popular methods in various cases, the results of each process have its advantages and disadvantages depending on the complexity and structure of the problem.

In recent years researchers have used the hybrid MCDM method to get maximum results, such as the study conducted by Xu, Li and Ren [31] combining the AHP and ANP methods to evaluate decision making using sustainable government data. Similarly, Jain et al [15] and Wang et al. [27] combine the Fuzzy AHP and TOPSIS methodologies to aid in supplier decision-making. Furthermore, research conducted by Büyüközkan, Göçer and Karabulut [5] combines AHP and VIKOR to support decision making in hazardous waste treatment. Furthermore, several MCDM methods have been integrated with other approaches, one of which is widely used to incorporate the MCDM method with machine learning. Research conducted by Hu, Chen and Zhu [10] combines the MCDM AHP technique with regression modeling, Support Vector Machine (SVM) classification modeling, and data mining modeling to assess and analyze the company's credit risk level. Mohammed et al. [21] combined SVM classification modelling with TOPSIS using Entropy criteria weight calculations to assess and classify the COVID-19 diagnosis model. The study conducted by Jena et al. [16] applied an integrated AHP model with Artificial Neural Network (ANN) to measure and assess earthquake risk areas. In the study, Yu et al. [32] to determine the degree of susceptibility and map the landslide-prone area combined multiple machine learning models (logistic regression, decision tree, support vector machines, and random forest) with AHP.

Several integration procedures for the hybrid MCDM method have been widely applied in various fields based on previous research. One hybrid method that is popular and interesting to research is the integration of AHP and SAW. AlNawaiseh et al. [1]; Cahyapratama and Sarno [6]; Saputra, Sitompul and Sihombing [26]; Angelina et al. [2]; Wijayanto et al. [29]; Melvin, Sutrisno and Herwindiati [20]; Noviyanto, Tarmuji and Hardianto [22]; Chen [7]; Macieira et al. [19] widely used in assessment and selection issues is the merging of AHP and SAW. Therefore, in this study, we combine the AHP method with SAW in a hybrid way and integrate it into machine learning modelling using a logistic regression algorithm for the problem of selecting recipients of the BRIN Innovation Village program funding incentives.

In this study, the utilization of the SAW technique as a ranking approach in multi-attribute decision making was examined. The AHP method was used to establish the weight values for various characteristics, and it was supported by eigenvalues. The Consistency Ratio is another metric that the AHP technique calculates, and it

may be used to identify the best option among several others. By integrating the expert team's assessment input for each weighted characteristic, the alternative evaluation of the SAW technique is created. The scoring that is produced from the data is utilized to forecast who will receive financial incentives using machine learning modeling that employs logistic regression.

2 Research Method

2.1 Area of Study

The research was conducted using data from the Innovated Village Program of the National Research and Innovation Agency in 2021, with a total of 2242 applicants who submitted funding proposals. In particular, there are several stages to determining program funding recipients in the selection process. In the first stage of 2242 applicants, 771 applicants officially submitted proposals, followed by the initial selection stage for the recommendations, which resulted in 138 proposals that passed the selection, then continued to the administrative selection stage by producing 96 proposals. And in the last step of this amount, 80 recipients of funding assistance were made. In this study, the data that will be observed is data on the results of the substance selection assessment of 96 prospective funding recipients. The results consist of 80 alternatives that pass or funded proposals and the remaining 16 recommendations that do not give. The methodology used for this study is shown in Figure 1, below.

2.2 Method of the Analytical Hierarchy Process (AHP)

Thomas L. Saaty, a mathematician from Pittsburg, Pennsylvania, created the Analytical Hierarchy Process (AHP) in the 1970s [24]. In a hierarchy, this decision support model will depict a challenging multi-factor or multi-criteria situation. The input utilized to solve this problem is human intuition because the structure of the problem is unclear and proper statistical data and information are not readily available. The AHP technique deconstructs complicated, unstructured situations into their component pieces, organizes the components or variables in a hierarchical pattern, gives numbers to judgements about the relative importance of each variable, synthesizes diverse factors, and increases dependability. AHP as a tool for making decisions.

The following are the processes and techniques for applying the AHP approach to solve problems:

1. **Identify the issue and decide on the ideal remedy.**
   Prioritization requires that the problem be able to be broken down into the objectives (goals) of an activity, the identification of alternatives (alternatives), and the creation of criteria (criteria) for picking priorities.
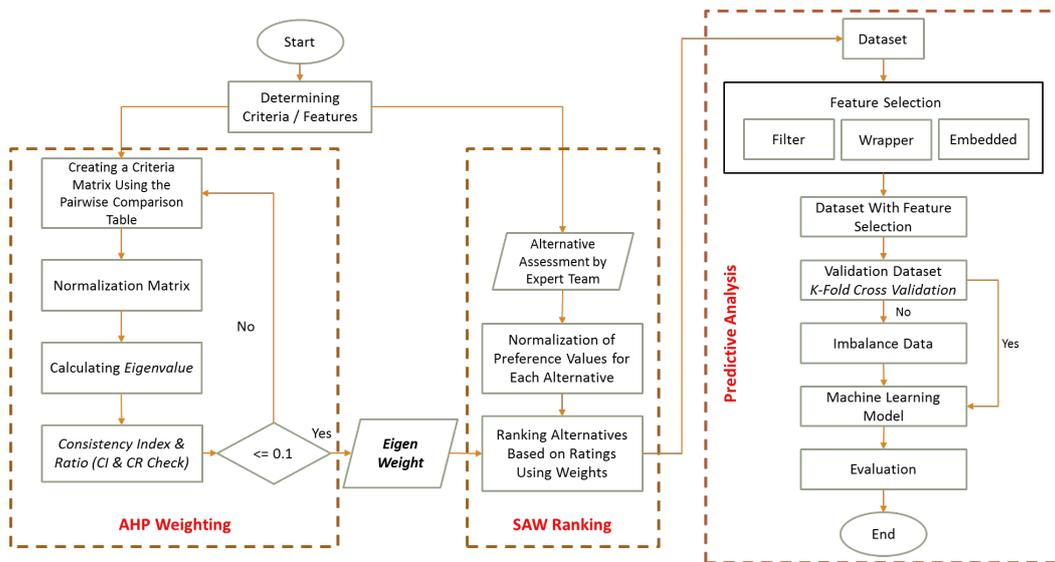
Adiwijaya, et al.: Integration of the Hybrid Decision Support System and Machine Learning Algorithm ...

MENDEL
Soft Computing Journal

Figure 1: Stages of Research.

Figure 2: Flow Chart AHP.

Figure 3: AHP Hierarchy Structure.
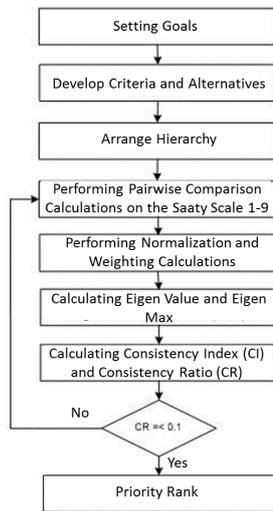
2. **Create a hierarchy, starting with the main objective.**
An abstraction of a system's structure called hierarchy is used to study how components interact and have an influence on one another. The system components or choice alternatives are described using a hierarchy or decision structure layout. The objective is the first level of a hierarchy, which, according to Saaty [25], is followed by levels of elements, criteria, sub-criteria, and so on all the way down to the last level of alternatives. The formulation of a priority-setting activity's goals is the first phase. As indicated in Figure 3, a hierarchy level will be created underneath the main aim, with sufficient criteria for examining or evaluating the provided options and deciding these alternatives, followed by sub-criteria.
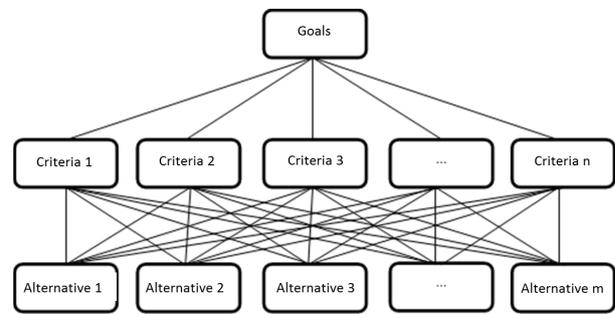
3. **Establish a pairwise comparison rating scale to create a criteria matrix.**
The following step is evaluating or comparing aspects, specifically comparisons between criteria using a paired comparison rating scale, after determining the hierarchy of objectives to alternatives. The weight of each metric is supposed to be determined by comparison between the criteria. Saaty [24] asserts that a scale of 1 to 9 is ideal for expressing thoughts on a range of issues. On Saaty's scale of 1 to 9, each pairwise comparison is rated as follows:

4. **Matrix Normalization**

$$w_i = \sum_{i=1}^{n} a_{ij}/n \qquad (1)$$

Definition. w_1 is weighted value. a_{ij}/n is row matrix normalization

5. **Test consistency by using the Consistency Index (CI)**

$$CI = \frac{(\lambda_{max} - n)}{(n-1)} \qquad (2)$$

Table 1: Pairwise Comparison Matrix.

Value	Information
1	Both Criteria/Alternative A and B are equally significant
3	B is less significant than A
5	A is somewhat more significant than B
7	Obviously, A is more significant than B
9	A is unquestionably more significant than B
2,4,6,8	Between two nearby values, if in doubt
Opposite	If alternative 1 is compared to alternative 2 the value is 3, then alternative 2 is compared to alternative 1 the value is 1/3

Table 2: Random Consistency Index.

n	1	2	3	4	5	6	7	8	9	10
R	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Definition. λ_{max} is eigenvalue maximum. n is number of matrices

6. **Calculating Consistency Ratio (CR)**

$$CR = \frac{CI}{RI} \qquad (3)$$

Definition. CR is Consistency Ratio. RI is Random Consistency Index (Table 2)

2.3 Method of the Simple Additive Weighting (SAW)

The weighted addition technique is another name for the SAW method. Finding the weighted total of the performance ratings for each alternative across all criteria is the fundamental tenet of the SAW technique. The choice matrix (X) must be normalized for the SAW technique so that it may be compared to all other available alternative ratings on the same scale [9]. The formula for determining the normalized matrix is given below.

$$r_{ij} = \begin{cases} \dfrac{x_{ij}}{max\ x_{ij}} & ; \text{ If } j \text{ is a benefit attribute,} \\[2ex] \dfrac{min\ x_{ij}}{x_{ij}} & ; \text{ If } j \text{ is a cost attribute} \end{cases} \qquad (4)$$

Definition. r_{ij} is value of a normalized performance rating. x_{ij} is the attribute value that each criteria has. $max\ x_{ij}$ is each criterion's highest value. $min\ x_{ij}$ is each criterion's lowest value. $Benefit$ if highest value is best then $Cost$ if the smallest value is the best where r_{ij} is the normalized performance rating of alternative on attribute $C_j; i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

Steps to solve using the SAW procedure:

1. Choose the standards that will serve as a guide while making decisions.

2. Establish the weighted average of each criterion that has previously been determined.

3. Ascertain each alternative's appropriateness score in relation to each criterion.

4. Create a decision matrix based on the criteria, normalize the matrix using an equation modified to account for the kind of attribute (profit attribute and cost attribute), and create a normalized matrix R.

5. Give each option (V_i) a preference score using the following formula:

$$V_i = \sum_{j=1}^{n} w_j r_{ij} \qquad (5)$$

Definition. V_i is rank for each alternative. w_j is weight value of each criterion. r_{ij} is normalized performance rating value.

2.4 Data Preparation

2.4.1 Feature Selection

Before a dataset is used to train a machine learning model, a series of steps need to be performed on the data. This series of processes is commonly referred to as data preparation or data preparation. Data preparation is an essential part of improving data quality and minimizing noise because the data generated from this series of processes will determine the efficiency of training and the performance of the resulting model.

Feature selection is one way of preparing data to improve accuracy in a machine learning model. The feature selection process reduces the number of features or input variables by selecting the features that are considered most relevant to the model. There are two types of feature selection, namely supervised and unsupervised. Supervised methods consist of the wrapper, filter, and intrinsic/embedded methods.
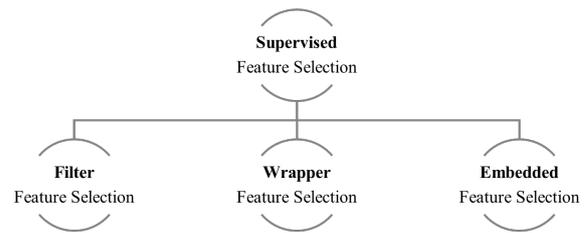

Figure 4: Feature Selection Method.

2.4.2 Imbalanced Data Handling

Unbalanced classes are a common problem in machine learning classification. Where imbalance class is a disproportionate distribution between classes in a dataset, one class has a considerable amount of data (majority class) compared to other classes (minority class). The significant difference in the amount of data between classes can result in the classification model being often

unable to predict the minority class accurately. A lot of test data that should be in the minority class is mispredicted by the classification model. Some solutions to overcome this include using the right evaluation metrics, resampling Over-Sampling and Under-Sampling.

2.4.3 Logistics Regression

Logistics regression is a data analysis technique in statistics that determines the relationship between several variables. The response variable is categorical, either nominal or ordinal, with the explanatory variable being categorical or continuous. Binary logistic regression is a mathematical model approach used to analyze the relationship between several factors and a binary variable. In logistic regression, if the response variable consists of two categories, $Y = 1$ indicates the results obtained are "successful," and $Y = 0$ shows the results obtained are "failed." The logistic regression uses binary logistic regression.

Similar methods and processes are used in the logistic regression method and the linear regression method. The model used to determine the logistic equation is:

$$\pi(x) = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}}{1 + \sum_{j=1}^{p} \beta_j x_j} \tag{6}$$

From $equation(8)$ obtained $1-\pi(x)$ as follows:

$$1 - \pi(x) = 1 - \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}}{1 + \sum_{j=1}^{p} \beta_j x_j} \tag{7}$$

$$1 - \pi(x) = \frac{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j} - e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}}{1 + \sum_{j=1}^{p} \beta_j x_j} \tag{8}$$

$$1 - \pi(x) = \frac{1}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}} \tag{9}$$

So that $\frac{\pi(x)}{1-\pi(x)}$ as follows

$$\frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j} \tag{10}$$

So the logistic equation is :

$$g\hat{(}x) = ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) \tag{11}$$

$$g\hat{(}x) = ln\left(e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}\right) \tag{12}$$

$$g\hat{(}x) = \beta_0 + \sum_{j=1}^{p} \beta_j x_j \tag{13}$$

3 Results and Discussion

To help in the best decision-making process when choosing the recipients of funding assistance for the Desa Berinovasi program at BRIN, the results and application of the integration of the decision support system algorithm using the hybrid AHP and SAW methods with predictive analysis using machine learning will

Table 3: Assessment Criteria.

Code	Assessment Aspect
K1	Locus Profile
K2	Regional Featured Product Profile
K3	Technology and Innovation Profile
K4	Stakeholder/Stakeholders Support
K5	Human Resources
K6	Management Agency
K7	Activity Implementation Method
K8	Work Plan and Implementation Strategy
K9	Detailed Budget

Table 4: Assessment Criteria Result.

	K1	K2	K3	K4	K5	K6	K7	K8	K9
K1	1.00	0.14	0.20	0.16	0.33	0.33	1.00	0.33	1.00
K2	7.00	1.00	5.00	3.00	3.00	3.00	7.00	3.00	7.00
K3	5.00	0.20	1.00	0.20	1.00	1.00	3.00	1.00	3.00
K4	6.00	0.33	5.00	1.00	5.00	5.00	7.00	5.00	7.00
K5	3.00	0.33	1.00	0.20	1.00	1.00	5.00	1.00	5.00
K6	3.00	0.33	1.00	0.20	1.00	1.00	5.00	1.00	5.00
K7	1.00	0.14	0.33	0.14	0.20	0.20	1.00	0.20	1.00
K8	3.00	0.33	1.00	0.20	1.00	1.00	5.00	1.00	5.00
K9	1.00	0.14	0.33	0.14	0.20	0.20	1.00	0.20	1.00
Total	30.00	2.96	14.86	5.25	12.73	12.73	35.00	12.73	35.00

be explained in the results and discussion. Each criterion's weight is determined using the AHP technique, after which the SAW method, using the AHP weight, determines the rating value of each option based on all of the requirements and the assessment team's findings. In contrast, the resultant data is subjected to predictive analysis and categorization using machine learning techniques. processing data with Python programming and Microsoft Excel.

3.1 Weighting Criteria Using AHP Method

The AHP approach was used to calculate each criterion's weight. The chief executive of the program serves as the respondent during the brainstorming process for the weighting of the criteria, and Table 3 of the pairwise comparison matrix used in the AHP computations contains the weight values.

A pairwise comparison matrix is first created before the criteria are weighted, as shown in Table 4. The matrix will then be normalized, with the results of that normalization shown in Table 5. Table 6 shows the results of the computation of the eigenvectors, which are the weighted values.

After the matrix has been normalized, the first step with the comparison matrix criteria in the preceding table is used to determine the eigenvalues of each row. An example of the process of calculating the normalization of pairwise comparisons based on $equation(1)$ is as follows:

$$X_{1,1} = \frac{1}{1+7+5+6+3+3+1+3+1} = 0.0333$$

$$X_{2,1} = \frac{7}{1+7+5+6+3+3+1+3+1} = 0.2333$$

The pairwise comparison matrix's first column's normalized value is generated by the calculation shown in the example above. Table 5 displays the final numbers.

Table 5: Matrix of Comparison's Normalization.

	K1	K2	K3	K4	K5	K6	K7	K8	K9	Total
K1	0.0333	0.0482	0.0135	0.0317	0.0262	0.0262	0.0286	0.0262	0.0286	0.2624
K2	0.2333	0.3376	0.3363	0.5712	0.2356	0.2356	0.2000	0.2356	0.2000	2.5853
K3	0.1667	0.0675	0.0673	0.0381	0.0785	0.0785	0.0857	0.0785	0.0857	0.7466
K4	0.2000	0.1125	0.3363	0.1904	0.3927	0.3927	0.2000	0.3927	0.2000	2.4173
K5	0.1000	0.1125	0.0673	0.0381	0.0785	0.0785	0.1429	0.0785	0.1429	0.8392
K6	0.1000	0.1125	0.0673	0.0381	0.0785	0.0785	0.1429	0.0785	0.1429	0.8392
K7	0.0333	0.0482	0.0224	0.0272	0.0157	0.0157	0.0286	0.0157	0.0286	0.2354
K8	0.1000	0.1125	0.0673	0.0381	0.0785	0.0785	0.1429	0.0785	0.1429	0.8392
K9	0.0333	0.0482	0.0224	0.0272	0.0157	0.0157	0.0286	0.0157	0.0286	0.2354

Here is an illustration of how to calculate eigenvalues, which are the weights assigned to each criterion based on (2):

$$\lambda_1 = \frac{0.26}{30} = 0.0292 \ , \ \lambda_2 = \frac{2.58}{2.96} = 0.2873$$

The example calculation above produces eigenvalues which are weight values for criteria 1 and 2. The results of calculating eigenvalues for all requirements are shown in Table 6.

Table 6: Eigenvalue.

	Eigenvalue	Weight Value
K1	0.0292	2.92
K2	0.2873	28.73
K3	0.0830	8.30
K4	0.2686	26.86
K5	0.0932	9.32
K6	0.0932	9.32
K7	0.0262	2.62
K8	0.0932	9.32
K9	0.0262	2.62

Then calculate the index consistency value (CI) and the consistency ratio value (CR) using the following formula:

$$CI = \frac{\lambda_{max} - n}{n-1}, CR = \frac{CI}{RI}$$

Then the index consistency value for all criteria is,

$$CI = \frac{9.7627 - 9}{9 - 1} = 0.0953$$

Based on the Random Consistency Index Table 2 for $n = 9$, the value of $RI = 1.45$, then the consistency ratio's value is,

$$R = \frac{0.0953}{1.45} = 0.0657$$

Due to the fact that $CR = 0.0657 < 0.1$, The evaluation does not need to be revised because the preferred value of the assessment criteria is constant. The SAW approach may be employed in the subsequent computation to use the weight value, or eigenvalue.

3.2 Calculation of SAW Approach

By ranking the significance of each alternative using the weights from the AHP technique, the SAW method is utilized to get the final alternative value. The result is a list of additional possibilities, arranged from the highest to the lowest value. 96 prospective receivers of financial aid are the possibilities mentioned. Finding the weighted total of the performance ratings for each alternative across all criteria is the fundamental tenet of the SAW technique. The decision matrix must be normalized for the SAW approach so that it may be compared to every other alternative rating scale. Table 7 lists the measures that benefit and cost the matrix according to the nine assessment criteria used to choose potential receivers of monetary assistance.

Table 7: Normalization of Cost Benefit.

	Attribute	Value Weight	Rounding
K1	Benefit	2.92	3
K2	Benefit	28.73	29
K3	Benefit	8.30	8
K4	Benefit	26.86	27
K5	Benefit	9.32	9
K6	Benefit	9.32	9
K7	Benefit	2.62	3
K8	Benefit	9.32	9
K9	Benefit	2.62	3

All criteria are included in the benefit attribute according to the table of classification criteria, therefore the most excellent value is the greatest value. Therefore, at this stage, the matrix normalization of all alternative values is converted to a percentage of the full value scale using the initial weight before calculating the sum using the AHP weight.

The normalized matrix multiplication with the AHP weight vector is added to the ranking procedure to get the final result. The finest option as a potential solution for beneficiaries of financing assistance is selected as having the most significant value.

Table 8: Ranking of Values from Largest to Lowest.

Recapitulation of the Reviewer Team's Assessment Results									
Alternative	K1	K2	K3	K4	K5	K6	K7	K8	K9	Total
27	3	29	8	27	9	9	3	7.2	2.4	97.60
30	3	29	8	23.62	9	9	3	9	2.4	96.03
48	2.7	24.36	7.2	26.32	8.55	8.55	2.7	8.55	2.7	91.64
23	3	26.1	7.2	27	9	9	3	3.6	2.4	90.30
...	...	...	...	...	...	...	...	...	...	
...	...	...	...	...	...	...	...	...	...	
87	1.8	20.3	5.2	16.2	5.4	6.3	2.1	5.85	1.8	64.95
88	1.8	20.3	5.2	16.2	5.4	6.3	2.1	5.85	1.8	64.95

3.3 Feature Selection

There are two types of feature selection, namely supervised and unsupervised. In this study, feature selection with a supervised method is used, which consists of a

Adiwijaya, et al.: Integration of the Hybrid Decision Support System and Machine Learning Algorithm ...

MENDEL
Soft Computing Journal

filter, wrapper, and intrinsic/embedded plans to reduce the number of features or input variables by selecting the features that are considered the most relevant and affect the model to be made. The following Table 9 of feature selection results.

Table 9: Feature Selection Results

Method	Description	Selection Result
Filter	Using Pearson Correlation and Variance Inflation Factor (VIF)	K4, K3, K1, K6
Wrapper	Using Recursive Feature Elimination (RFE) mechanism with Learning Random Forest Algorithm	K4, K3, K1, K9
Embedded	Using the Learning Logistic Regression Algorithm and using the L2 regularization feature as a penalty function to eliminate feature	K4, K3, K2, K6

Based on the feature selection table above, of the nine criteria or input variables used, the criteria that are considered the most relevant and influential on the model to be made are K3 (Profile of Technology and Innovation) and K4 (Stakeholder Support). This also proves that in predictive analysis, Each criterion's or variable's weight value, as determined by the AHP method's computation, does not necessarily reflect its importance to the input variable. In contrast to the feature selection procedure of the predictive analysis of the most pertinent and important K3 criteria, it is demonstrated that the K2 criteria (Profile of Regional Leading Products) are the most significant weight value (29%) in the computation of the AHP technique.

3.4 Imbalanced Data Test

In the classification process, we usually have various problems with data, both from data preprocessing, modeling, evaluation, and others. Sometimes, one thing that is not realized from the classification process is the number or proportion of existing labels/classes. It could be that the data we are dealing with is an Imbalanced Dataset. Therefore, dataset validation was also done in this study in order to aid create the desired model. The expected model is a model that can distinguish between proposals that are accepted/funded (rare class) and those that are not (abundant class). Figure 5 displays the outcomes of charting the dataset using the feature selection variable.

Plotting the dataset's findings reveals that one of the classes/labels (not passed) has a value that has a number that differs significantly from the graduated label class. So dataset validation is essential here to help form the desired model. Figures 6 and 7 are the results of resampling the dataset using Over-Sampling and Under-Sampling.

Considering the outcomes of Table 10, the dataset handled without imbalance has the best model eval-

Table 10: Dataset Evaluation Validation Table (Base Logistic Regression)

Evaluation Model	Without Imbalance Data	SMOTE (Over Sampling)	ENN (Under Sampling)
Accuracy	92.11	67.11	89.47
Precision	92.42	95.24	90.91
Recall	98.39	63.49	96.77
F1 Score	95.31	76.19	93.75

uation score compared to the dataset with imbalanced data handling using SMOTE (Over-Sampling) and ENN (Under-Sampling).

3.5 Logistics Regression Equation Model

Based on the results of feature selection and data validation, it is known that for machine learning modeling, the dataset that will be used is a dataset without resampling with two explanatory variables, K3 (Profile Technology and Innovation) and K4 (Stakeholder Support) which have a significant influence on the response variable so that These two variables are included in the logistic regression equation, so that the explanatory variables K3 and K4 and the response variable $Y = (x)$ in the logistic regression model represent the status of the grant recipient. The following is the equation of the logistic regression model:

$$\hat{p} = \frac{e^{-17,84+1,53X_3+0,49X_4}}{1 + e^{-17,84+1,53X_3+0,49X_4}} \qquad (14)$$

Additionally, the classification results from the system (model) with the actual classification results are used to evaluate the classification of the logistic regression equation model using the confusion matrix table. The classification model's performance on a set of test data with known real values is described by the confusion matrix, which takes the form of a matrix table. A confusion matrix with four distinct combinations of expected values and actual values may be seen in the image below.

Table 11: Confution Matrix N=96.

	Predicted Positive	Predicted Negative
Actual Positive	76(TP)	4(FP)
Actual Negative	5(FN)	11(TN)

Table 12: Evaluation Metric

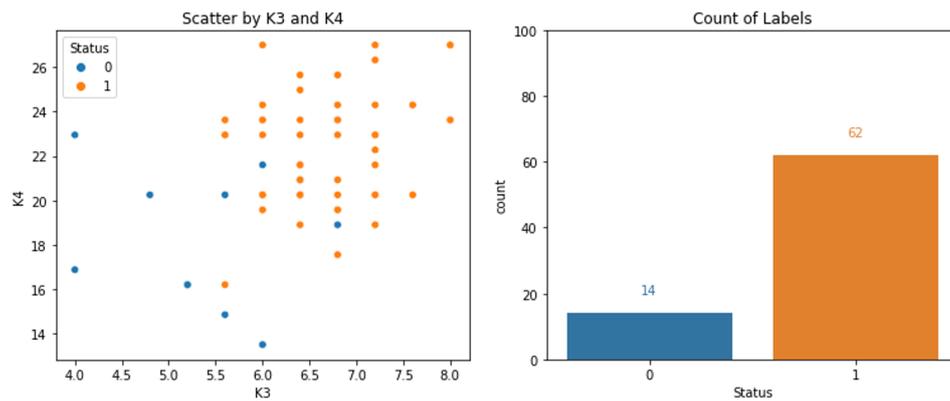| Metric	Score
Accuracy	90.62%
Precision	95%
Recall	93.83%
F1 Score	94.40%

Figure 5: Dataset Plotting Results Using Feature Selection Results Variables Without Imbalanced Data
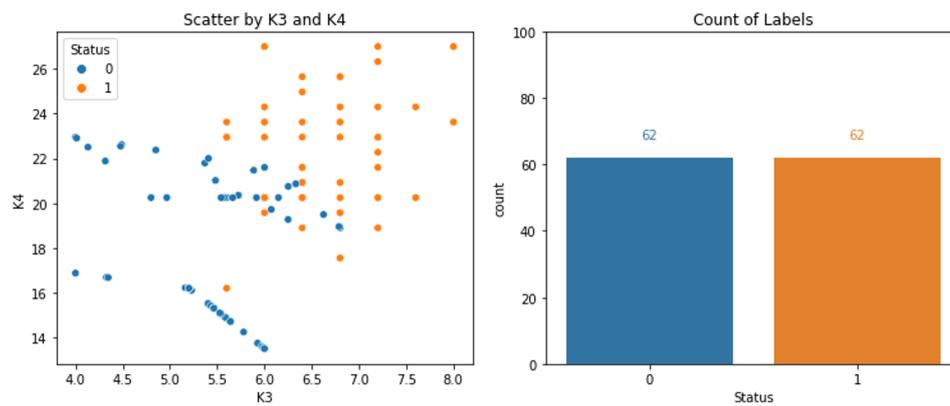

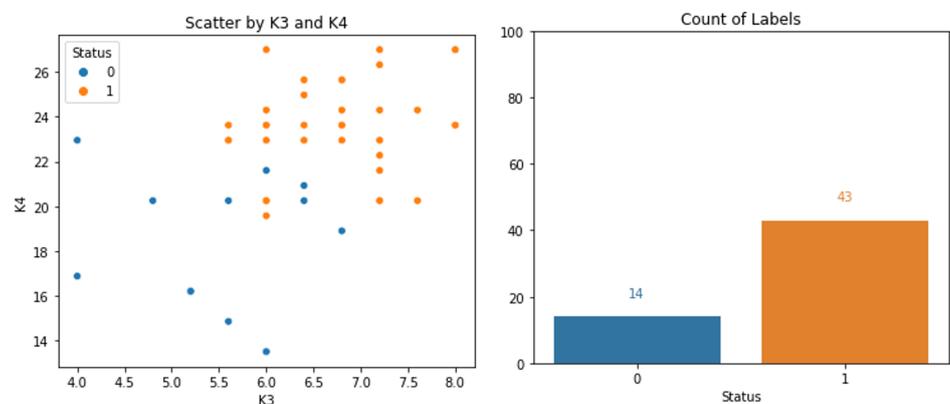Figure 6: SMOTE (Oversampling) Data Imbalance Plotting Results

Figure 7: ENN (Under Sampling) Data Imbalance Plotting Results

4   Conclusions

According to the results of a study, each evaluation criterion's weight value is determined by the AHP computation and has a value of K1 3%, K2 29%, K3 8%, K4 27%, K5 9%, K6 9%, K7 3%, K8 9% and K9 3%, and the value of the consistency ratio is $0.065749 < 0.1$, then the preference value of the assessment criteria is consistent and does not call for the evaluation to be revised. The SAW approach may be employed in the subsequent mathematical step to use the weight value,

or eigenvalue. The normalized matrix multiplication with the AHP weight vector is added to the ranking procedure to get the final result. The finest option as a potential solution for beneficiaries of financing assistance is selected as having the most significant value.

In the predictive analysis of the feature selection process, which aims to select variables relevant to the grantee's response variables, the relevant variables are K3 (Profile Technology and Innovation) and K4 (Stakeholder Support). These results prove that in predictive analysis, the weight value of each of the criteria or vari-

ables derived by the AHP method's computation does not automatically represent the value of relevance to the input variable. It is proven that the K2 criteria (Profile of Regional Leading Products) are the most considerable weight value (29%) in the AHP method's computation. The K3 criteria are the most pertinent and important factors in comparison to the feature selection procedure.

In the data validation process, the best model is without imbalanced data handling compared to imbalanced data handling using Over Sampling and Under Sampling, resulting in 92.11%Accuracy, 92.42%Precision, 98.39%Recall and 95.31%F1 Score. Therefore, a logistic regression equation is obtained to predict the response variable y (recipient of funding) $Y = -17.84 + 1.53K3 + 0.49K4$ with evaluation values using a confusion matrix, namely the Accuracy value of 90.62%, Precision 95%, Recall 93.83%and F1 Score 94.40%. Judging from the regression equation, the K3 value is greater than the K4 value. The K3 value indicates the slope of X (Profile of Technology and Innovation), and K4 indicates the slope of X (Stakeholders Support). In this case, it can be concluded that the Technology and Innovation Profile percentage is more influential than Stakeholder Support.

It is hoped that the following research can use serial data to enrich the dataset used and be tested to find weights using AHP with group preferences rather than individuals. The subsequent research can compare several machine learning algorithms for predictive analysis to get the best model from several algorithms used. It is also hoped that the following output can be application-based[1].

References

[1] AL NAWAISEH, A. J., ALBTOUSH, A., AL-MSIEDEEN, R., AND AL NAWAISEH, S. J. Evaluate Database Management System Quality By Analytic Hierarchy Process (AHP) and Simple Additive Weighting (SAW) Methodology. *Mendel 28*, 2 (2022), 67–75.

[2] ANGELINA, SURBAKTI, R. D., SIMAMORA, R. S., CENDANA, E., SITANGGANG, D., BANJARNAHOR, J., AND TURNIP, M. Application Selection Lending Houses Subsidized by the Method of AHP and SAW. In *Journal of Physics: Conference Series* (sep 2019), vol. 1230, Institute of Physics Publishing.

[3] AZHAR, N. A., RADZI, N. A. M., AND WAN AHMAD, W. S. H. M. Multi-criteria Decision Making: A Systematic Review. *(Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering) 14*, 8 (nov 2021), 779–801.

[4] AZZEDDINE, D., JABRI, S., YOUSSE, B., AND TAOUFIQ, G. The selection of the relevant association rules using the electre method with multiple criteria. *IAES International Journal of Artificial Intelligence 9*, 4 (dec 2020), 638–645.

[5] BÜYÜKÖZKAN, G., GÖÇER, F., AND KARABULUT, Y. A new group decision making approach with IF AHP and IF VIKOR for selecting hazardous waste carriers. *Measurement: Journal of the International Measurement Confederation 134* (feb 2019), 66–82.

[6] CAHYAPRATAMA, A., AND SARNO, R. Application of Analytic Hierarchy Process (AHP) and Simple Additive Weighting (SAW) methods in singer selection process. In *2018 International Conference on Information and Communications Technology, ICOIACT 2018* (apr 2018), vol. 2018-Janua, Institute of Electrical and Electronics Engineers Inc., pp. 234–239.

[7] CHEN, Y. Application of analytic hierarchy process (AHP) and simple additive weighting (SAW) methods in mapping flood-prone areas. In *Frontiers in Artificial Intelligence and Applications* (oct 2021), vol. 341, IOS Press BV, pp. 435–441.

[8] DURSUN, M., AND OGUNCLU, O. Agile Supplier Evaluation Using Hierarchical TOPSIS Method. *WSEAS Transactions on Information Science and Applications 18* (2021), 12–19.

[9] FISHBURN, P. C. Methods of Estimating Additive Utilities Author (s): Peter C . Fishburn Source : Management Science , Vol . 13 , No . 7 , Series A , Sciences (Mar ., 1967), pp . 435-453 Published by : INFORMS Stable URL : http://www.jstor.org/stable/2627687. *Management Science 13*, 7 (1967), 435–453.

[10] HU, J., CHEN, C., AND ZHU, K. Application of Data Mining Combined with Fuzzy Multicriteria Decision-Making in Credit Risk Assessment from Legal Service Companies. *Mathematical Problems in Engineering 2021* (2021).

[11] INDONESIA. Regulation of the President of the Republic of Indonesia Number 81 of 2010 Concerning the Grand Design of Bureaucratic Reform 2010-2025, 2010.

[12] INDONESIA. Regulation of the Minister of Research, Technology, and Higher Education of the Republic of Indonesia Number 38 of 2019 concerning National Research Priorities for 2020-2024, 2019.

[13] INDONESIA. Technical Guidelines for Government Assistance in the Form of Desa Berinovasi, 2021.

[14] INTERNATIONAL FEDERATION OF ACCOUNTANTS, AND CHARTERED INSTITUTE OF PUBLIC FINANCE AND ACCOUNTANCY. International framework : good governance in the public sector. Tech. rep., IFAC dan CIPFA, 2014.

[15] JAIN, V., SANGAIAH, A. K., SAKHUJA, S., THODUKA, N., AND AGGARWAL, R. Supplier selection using fuzzy AHP and TOPSIS: a case study in the Indian automotive industry. *Neural Computing and Applications 29*, 7 (apr 2018), 555–564.

---

[1]The code for the experiments can be found at https://github.com/indrarusyadi/LR_Desa_Berinovasi

[16] Jena, R., Pradhan, B., Beydoun, G., Nizamuddin, Ardiansyah, Sofyan, H., and Affan, M. Integrated model for earthquake risk assessment using neural network and analytic hierarchy process: Aceh province, Indonesia. *Geoscience Frontiers 11*, 2 (mar 2020), 613–634.

[17] Kim, J. H., and Ahn, B. S. The hierarchical VIKOR method with incomplete information: Supplier selection problem. *Sustainability (Switzerland) 12*, 22 (nov 2020), 1–14.

[18] Lopes, A. P., and Rodriguez-Lopez, N. A decision support tool for supplier evaluation and selection. *Sustainability (Switzerland) 13*, 22 (nov 2021).

[19] Macieira, M., Mendonça, P., Miranda Guedes, J., and Tereso, A. Evaluating the efficiency of membrane's refurbishment solutions to perform vertical extensions in old buildings using a multicriteria decision-support model. *Architectural Engineering and Design Management 18*, 1 (2022), 1–25.

[20] Melvin, Sutrisno, T., and Herwindiati, D. E. Decision support system for election and evaluation of assistant lecturer using analytical hierarchy process and simple additive weighting: Case study faculty of information technology Tarumanagara university. In *IOP Conference Series: Materials Science and Engineering* (dec 2020), vol. 1007, IOP Publishing Ltd.

[21] Mohammed, M. A., Abdulkareem, K. H., Al-Waisy, A. S., Mostafa, S. A., Al-Fahdawi, S., Dinar, A. M., Alhakami, W., Baz, A., Al-Mhiqani, M. N., Alhakami, H., Arbaiy, N., Maashi, M. S., Mutlag, A. A., Garcia-Zapirain, B., and De La Torre Diez, I. Benchmarking Methodology for Selection of Optimal COVID-19 Diagnostic Model Based on Entropy and TOPSIS Methods. *IEEE Access 8* (2020), 99115–99131.

[22] Noviyanto, F., Tarmuji, A., and Hardianto, H. Food Crops Planting Recommendation Using Analytic Hierarchy Process (AHP) And Simple Additive Weighting (SAW) Methods. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH 9* (2020), 2.

[23] Rodrigues, D., Godina, R., and da Cruz, P. E. Key performance indicators selection through an analytic network process model for tooling and die industry. *Sustainability (Switzerland) 13*, 24 (dec 2021).

[24] Saaty, T. L. Applications of analytical hierarchies. *Mathematics and Computers in Simulation 21*, 1 (mar 1979), 1–20.

[25] Saaty, T. L. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research 48*, 1 (sep 1990), 9–26.

[26] Saputra, M., Sitompul, O. S., and Sihombing, P. Comparison AHP and SAW to promotion of head major department SMK Muhammadiyah 04 Medan. In *Journal of Physics: Conference Series* (apr 2018), vol. 1007, Institute of Physics Publishing.

[27] Wang, C. N., Pan, C. F., Nguyen, V. T., and Husain, S. T. Sustainable supplier selection model in supply chains during the COVID-19 pandemic. *Computers, Materials and Continua 70*, 2 (2022), 3005–3019.

[28] Wang, C. N., Tsai, H. T., Ho, T. P., Nguyen, V. T., and Huang, Y. F. Multi-criteria decision making (MCDM) model for supplier evaluation and selection for oil production projects in vietnam. *Processes 8*, 2 (feb 2020).

[29] Wijayanto, S., Napitupulu, D., Adiyarta, K., and Windarto, A. P. Decision Support System of New Student Admission Using Analytical Hierarchy Process and Simple Additive Weighting Methods. In *Journal of Physics: Conference Series* (sep 2019), vol. 1255, Institute of Physics Publishing.

[30] World Bank. *Governance and development.* World Bank, 1992.

[31] Xu, J., Li, L., and Ren, M. A Hybrid ANP Method for Evaluation of Government Data Sustainability. *Sustainability (Switzerland) 14*, 2 (jan 2022).

[32] Yu, L., Zhou, C., Wang, Y., Cao, Y., and Peres, D. J. Coupling Data-and Knowledge-Driven Methods for Landslide Susceptibility Mapping in Human-Modified Environments: A Case Study from Wanzhou County, Three Gorges Reservoir Area, China. *Remote Sensing 14*, 3 (feb 2022).

[33] Zhao, K., Dai, Y., Ji, Y., and Jia, Z. Decision-Making Model to Portfolio Selection Using Analytic Hierarchy Process (AHP) with Expert Knowledge. *IEEE Access 9* (2021), 76875–76893.