Smart Multi-Criteria Analytics and Reasoning Technologies Volume 1, Issue 1 (2025) 22-32

Smart Multi-Criteria Analytics and Reasoning Technologies

Journal homepage: www.smart-journal.org
ISSN: 3104-3046

A Machine Learning and Multi-Criteria Decision-Making Framework for Student Grade Prediction

Tešić Duško*

Military Academy, University of Defence, Belgrade, Serbia

ARTICLE INFO

Article history:
Received 20 September 2025
Received in revised form 10 October 2025
Accepted 24 October 2025
Available online 25 October 2025

Keywords:
Machine Learning; MCDM; SAW; Random Forest; Linear Regresion

ABSTRACT

This paper presents an integrated methodological framework that combines machine learning (ML) algorithms with multi-criteria decision-making (MCDM) methods to predict student grades based on multiple input criteria. Unlike traditional approaches that focus on assigning grades based on static thresholds, the proposed system allows for numerical prediction of overall achievement before formal evaluation, thereby providing prediction of educational outcomes and supporting necessary decision-making. The research used regression models, including Random Forest and Linear Regression, to model the relationships between input attributes and target variables. The obtained criterion weights from the models were used in the Simple Additive Weighting (SAW) method, which allowed for information aggregation and generation of proportional results. These results were then transformed into classification classes while preserving ranking and proportionality. Special emphasis is placed on technical reproducibility, interpretability and stability of the classification throughout all processing stages. Pearson correlation between predicted and scaled values, as well as between their ranks, confirms that the transformation preserved the data structure with high accuracy. This shows that MCDM methods can serve not only for evaluation and ranking, but also as a valid tool for numerical prediction in educational systems. The proposed framework enables the development of scalable, transparent and automated systems for predicting student grades, with the potential for wider application in educational analytics, instructional planning and identification of educational needs.

1. Introduction

The development of modern education systems has stimulated the need for automated and transparent evaluation methods that enable reliable assessment and prediction of student results based on clearly defined criteria. Unlike traditional methods that rely on manual grading and fixed thresholds, data-driven predictive models allow quantitative assessment of student performance

* *Corresponding author.*
E-mail address: tesic.dusko@yahoo.com

https://doi.org/10.65069/smart1120253

before grades are formally assigned. The proposed approach is designed to complement, rather than replace, teachers' assessment practices, by providing analytical insight into learning outcomes [1].

Traditional methods of multi-criteria decision-making (MCDM) enable a transparent and structured evaluation of alternatives [2] based on predefined criteria [3]. However, their classical application often relies on subjectively determined weights of criteria [4], which can limit the objectivity and adaptability of the obtained results [5]. The modern development of machine learning (ML) methods enables empirical determination of the importance of individual criteria, thus providing a data-based basis for assigning weights in the MCDM approach [6], in addition to other purposes of these methods [7-9]. By using empirically derived weights from ML models into the MCDM framework, the difference between descriptive evaluation and predictive modeling is overcome, thus forming a robust and interpretable method applicable in educational, social and technical systems. In addition to the above, it is important to mention here the way in which these two approaches consider uncertainty and inaccuracy in the input data. While such uncertainty in MCDM is typically addressed by modeling ambiguity and imprecise human judgments using linguistic variables and membership functions [10-12], extensions based on some theories that handle uncertainty and imprecision well [13-14], machine learning approaches manage uncertainty using probabilistic models, ensemble techniques, and confidence estimation to quantify the reliability of predictions and handle data variability [15].

In this research, an integrated methodological framework was developed that combines ML algorithms with MCDM method, with the aim of generating numerical predictions of students' total success (total score), and consequently grade, based on several input attributes. Regression models, such as Random Forest and multidimensional Linear Regression, are used to model the relationship between the input criteria and the target variable, while Simple Additive Weighting (SAW) method is applied for weight aggregation and proportional mapping of results. This allows the model to predict numerical values and map them transparently into grade categories.

In contrast to classical evaluation approaches, this framework does not just provide ratings, but predicts their quantitative equivalents based on empirical patterns in the data. Such a prediction can serve as support for teachers, analysts and educational institutions in making necessary decisions, identifying educational needs and risks and improving the quality of teaching. Special emphasis is placed on technical reproducibility, transparency of transformation and preservation of classification stability through all stages of processing.

2. Methodology

The algorithm of the conducted research, which consists of four phases with multiple steps, is presented in Figure 1. For the purposes of this research, ML and MCDM approaches were used, which aim to predict the total score and grades of students. In the first stage, a data set with a sufficient number of alternatives, criteria and ratings is defined, after which the Random Forest algorithm is applied to determine the empirical weights of the criteria. In the second phase, the data is divided into training and test set in a ratio of 70:30, and then a multidimensional Linear Regression model is applied to predict the total score value on the test set. The third phase includes the application of the SAW method to 10 alternatives from the test set, whereby previously defined criteria weights are integrated into the aggregation function, and the obtained scores are proportionally scaled and transformed into grades. In the final, fourth stage, the model is validated by comparing the scaled values of the SAW method with regression predictions, as well as by analyzing the differences between the scores obtained by the SAW method and those predicted by the regression model, which checks the classification stability and methodological consistency of the proposed framework.

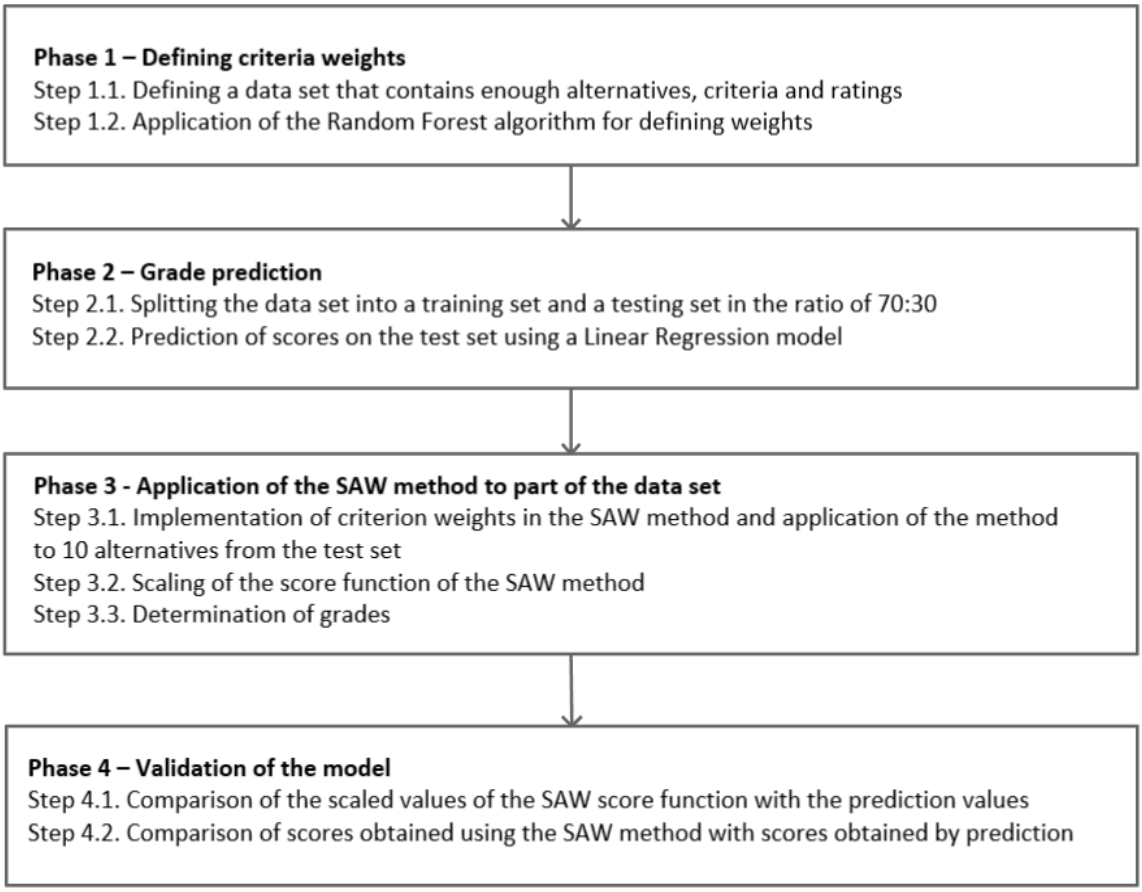

Fig. 1. Research algorithm

2.1 Random Forest

Random Forest is an ensemble machine learning method that combines multiple decision trees to improve prediction accuracy and reduce overfitting, instead of relying on a single tree [16-18]. This method builds multiple trees on different subsets of the data and uses averaging (for regression) or voting (for classification) to get the final result. Each tree is trained on a random sample of data and uses a random selection of branching criteria, which increases the diversity of the model [16]. The model automatically calculates the importance of criteria (Feature Importances), which makes it possible to see the influence of factors on the final decision. It is suitable for large datasets and noisy situations. In the context of MCDM, Random Forest provides empirically validated criterion weights that can be directly used in the MCDM methods. The algorithm for obtaining criteria weights from the data set is given in Figure 2.

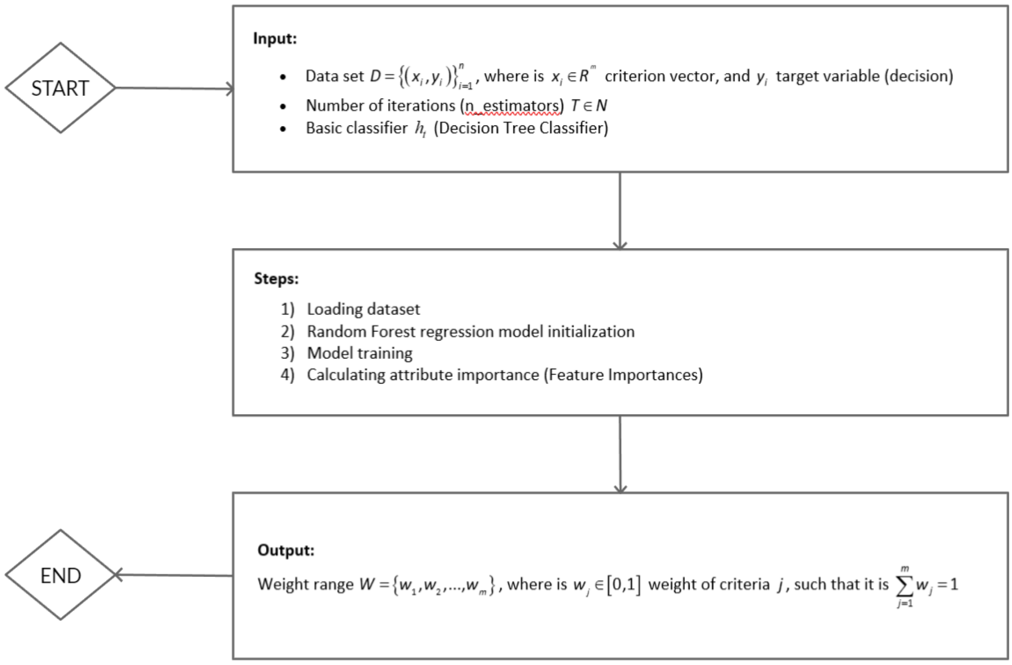

Fig. 2. Algorithm for determining the criteria weights

2.2 Linear Regression

Linear Regression is a basic method in machine learning that is used to predict numerical values based on known input criteria. The model tries to find the best possible linear relationship between the input and the goal by learning the coefficients that determine the impact of each criterion on the result [19, 20]. Its simplicity enables quick interpretation and analysis, so it is often used as a reference model in research. In the context of this study, Linear Regression was used to predict scores and grades, to compare the results obtained using the MCDM method, which were previously scaled using a linear transformation, with the output results of this method.

2.3 SAW method

The SAW method [21, 22] is among the simplest and most frequently applied approaches for solving MCDM problems [23, 24]. The essence of this method is that each considered alternative is evaluated collectively, based on the weighted sum of the value it achieves according to various criteria [25, 26]. Before calculation, criteria values are normalized, especially when they are expressed in different units. After that, each normalized value is multiplied by the corresponding weight that shows the importance of the given criterion in the overall assessment. The result is obtained by adding up the weighted values, and the alternative with the highest overall score is considered the most favorable. This method is simple to apply and clear in interpretation. The application of this method in combination with ML tools has been presented in various publications [27-29]. The steps of the SAW method are presented in Figure 3 [21, 22].

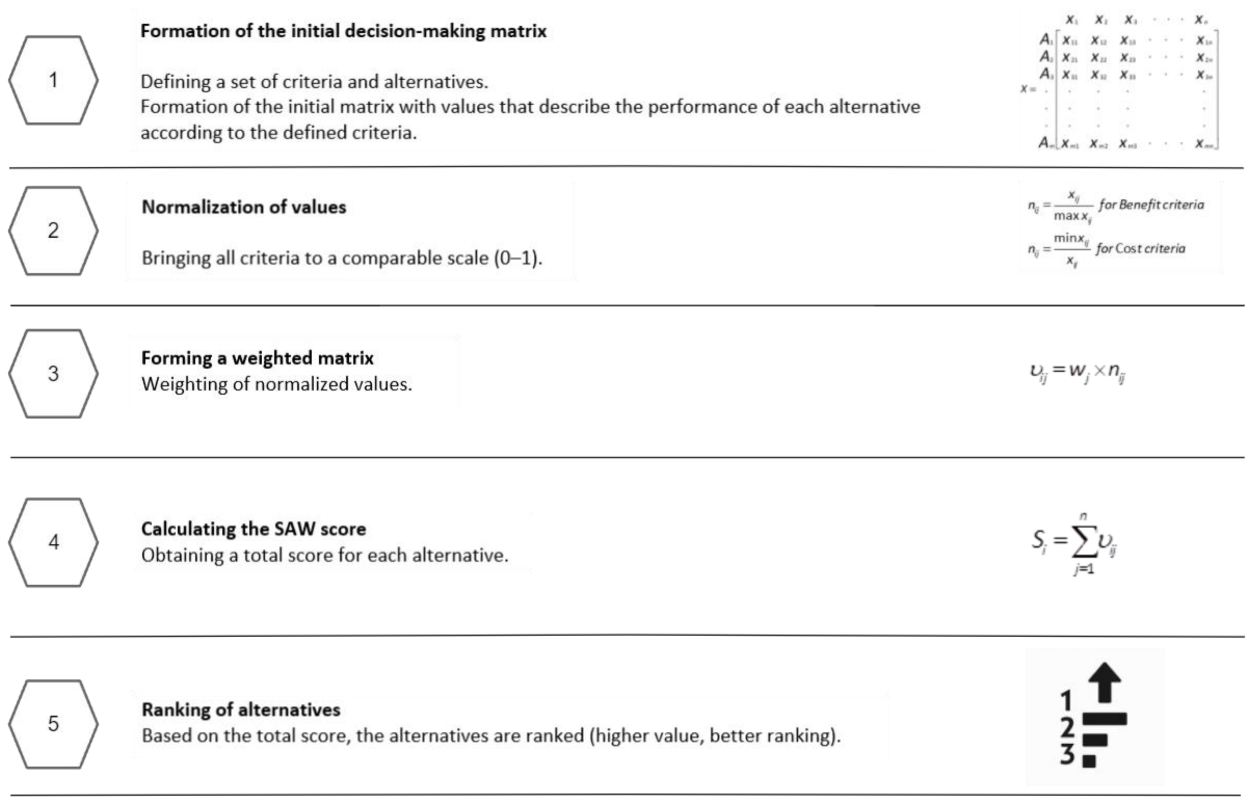

Fig. 3. The algorithm of the SAW method

After obtaining the final values of method, mapping of SAW values S_i is done as follows (Eq. 1-5):

$$y_{SAW} = aS_i + b \tag{1}$$

$$\overline{x} = \frac{1}{m}\sum_{i=1}^{m} S_i \tag{2}$$

$$\overline{y} = \frac{1}{m}\sum_{i=1}^{m} y_i \tag{3}$$

$$a = \frac{\sum_{i=1}^{m}(S_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{m}(S_i - \overline{x})^2} \tag{4}$$

$$b = \overline{y} - a\overline{x} \tag{5}$$

where is *a* - the slope coefficient that determines how much the SAW score is scaled to fit the target metric, and *b* is the free term (intercept) that determines the initial value of the transformation, i.e. where the line intersects the y-axis. The values $\overline{x}$ and $\overline{y}$ represent the average values of the score function of the SAW method and the values of the prediction score, respectively.

3. Results

Applying the research algorithm (Figure 1), in Phase 1, we first define a data set with more criteria and a larger number of alternatives, where the output data is given, in this case the total score and the student's grade.

3.1 Defining criteria weights

This phase consists of two steps. In Step 1.1., for the purposes of this research, the Student Performance Dataset was used [30] from Kaggle, the global data science and machine learning platform known for bringing together researchers, engineers and enthusiasts from around the world to share code, compete to solve problems and learn from each other. This dataset presents 1000000 rows of realistic student performance data, where the input data is integrated into three criteria: 1) Average weekly self study hours (0–40), 2) Attendance percentage (50–100), 3) Class participation (0–10, indicating how actively the student participates in class), while the output data represents Final performance score (0–100) and Grade (A, B, C, D, F) derived from total_score , according to the following rule A: ≥ 85, B: ≥ 70, C: ≥ 55, D: ≥ 40, F: < 40 [24].

In Step 1.2, the Random Forest regression method is used to predict the numerical target value based on the input criteria. The target variable (total_score) is separated from the input attributes, thus forming the X (inputs) and Y (target) matrices for training the model. Then the data is split into training and test set, in a ratio of 70:30, in order to train the model on most of the data and then test it on the test set. A Random Forest Regressor is trained on the training set, which uses an ensemble of decision trees to model nonlinear relationships between criteria and target values. After training, the model is used to make predictions on the test set, and evaluation is done using metrics such as MSE (mean squared error) and R^2 (coefficient of determination), which show how well the model explains the variance in the data. Feature importances are then extracted from the trained Random Forest model, i.e. the relative contribution of each criterion in the prediction process. Those values are normalized so that their sum is 1, thus obtaining the proportional weights of the criteria that can be used in MCDM methods. In the Random Forest regression model, two key parameters were used: the number of trees (100) and a fixed value for randomness (42). The first parameter determines that the model consists of 100 decision trees, which achieves greater stability and accuracy of predictions through the aggregation of several independent decisions. The second parameter serves to control randomness in the training process, allowing model results to be reproduced identically at each run, which is important for consistency in research and evaluation. By applying the mentioned methodology, the following values are reached (Table 1):

Table 1
Random Forest Results

Indicators	Values
MSE	81.93
R^2	0.66
Average weekly self study hours	0.7628
Attendance percentage	0.1374
Class participation	0.0998

The Random Forest regression model shows a solid performance with an MSE value of 81.93, which means that it is wrong on average by about 9 points per example, while an R^2 of 0.66 indicates that it explains 66% of the variation in the target values, which is acceptable for educational data. The analysis of criteria weights reveals that average weekly hours of independent study are the most influential factor with 76.28% contribution, while class attendance (13.74%) and class participation (9.98%) are of secondary importance. This distribution shows that independent learning has a crucial role in success, while other factors contribute additionally, but with a significantly lower intensity. The weights can be directly used in the SAW method for proportional decision making, thus providing an empirically validated evaluation. In the next phase of the model, using Linear Regression, scores and grades are predicted for the test data.

3.2 Grade prediction

In this procedure, the method of multidimensional Linear Regression is used to predict the numerical target value (total score) based on the input criteria. The target variable (total_score) is separated from the input attributes, thus forming the X (inputs) and Y (target) matrices for training the model. Then the data is divided into training and test set, in a ratio of 70:30. The test data includes 30% of the input data used in the Random forest method. After training, the model is used for prediction on the test set. The obtained values are compared to the actual scores, and the evaluation is done using metrics such as MSE and R^2. Score values for 30,000 alternatives were obtained by the aforementioned prediction, and the first 10, which will be used in the MCDM method, are shown in Table 2. By applying this methodology, the values of MSE=80.83 and R^2=0.66 are obtained. E valuation of the model shows that an MSE of 80.83 has been achieved, which means that the model is wrong on average by about 8.99 points per alternative. This value indicates moderate precision, acceptable in the context of educational and social data with multiple influences. At the same time, the R^2 value shows that the model manages to explain 66% of the total variation in the target values, which is an indicator of stable linearity between the input criteria and the results.

Table 2
Values of the predicted total score and grade of 10 alternatives

Alternative	Average weekly self study hours	Attendance percentage	Class participation	Predicted total score	Rank	Grade
A1	1.3	80.1	2.1	59.3372	10	C
A2	6.7	100	5	69.1492	8	C
A3	1.9	78	5.7	60.4211	9	C
A4	10.5	94.1	2.8	76.0548	6	B
A5	7.4	78.9	5.7	70.4138	7	B
A6	12.9	87.1	3.6	80.4119	4	B
A7	11.2	82.6	3.2	77.3227	5	B
A8	17.4	74.7	5.5	88.581	3	A
A9	23.4	68.4	5.7	99.4798	2	A
A10	40	57	7.5	129.633	1	A

In the next phase, the SAW method is applied and the obtained values of the total score of the function are scaled and student grades were obtained.

3.3 Application of the SAW method

In the first step of this phase, by applying the steps of the SAW method (Figure 3) on the initial matrix presented in Table 2 (without the part related to the prediction of grades) with the weight coefficients of the criteria presented in Table 1, the following values of the total score of the SAW function and the rank of the alternatives are obtained (Table 3):

Table 3
SAW method values

Alternative	Total score	Rank
A1	0.1628	10
A2	0.3317	7
A3	0.2193	9
A4	0.3668	6
A5	0.3254	8
A6	0.4136	4
A7	0.3697	5
A8	0.5076	3
A9	0.6161	2
A10	0.9409	1

In the next steps, the obtained score is scaled, so that the results obtained after the prediction can be compared with this, and in order to assign grades to the alternatives (students). In this procedure, a linear transformation of the SAW score is performed using Eqs. 1-5. Applying Eq. 2, the value is obtained $\bar{x} = 0.401120$, Eq.3 $\bar{y} = 78.758997$, Eq. 4 $a = 93.80907159303806$ and Eq.5 $b = 41.17615725807386$. Finally, using Eq. 1, the following scaled value of the total score of the SAW method according to the alternatives is obtained (Table 4):

Table 4
Scaled values of the total score SAW method, rank and grades

Alternative	Scaled total score	Rank	Grade
A1	56.44756	10	C
A2	72.29285	7	B
A3	61.74408	9	C
A4	75.58411	6	B
A5	71.69925	8	B
A6	79.97394	4	B
A7	75.85341	5	B
A8	88.79763	3	A
A9	98.96889	2	A
A10	129.4428	1	A

In the next phase, the validation of the obtained results is carried out through the comparison of the obtained results of the total scaled score of the SAW method and Linear Regression, as well as the comparison of the scores obtained using the SAW method and prediction.

3.4 Validation of the model

To validate the proposed approach, the results from Tables 2 and 3 will be compared. Comparative analysis of predicted and scaled total score SAW method values shows high consistency. In most cases, scores remain stable, and the differences between predicted and scaled values are minimal, generally below ±3 points. The only exception is alternative A2, where there is a change in grade from C to B, indicating an oscillation around the cut-off value between grades. This stability confirms that the scaling preserved the ranking and proportionality of the predictions, thus allowing reliable mapping into ratings and integration into MCDM methods such as SAW. The results show that the transformation is numerically precise and classification consistent.

By calculating the Pearson correlation coefficient [31] between the predicted and scaled values of the total score SAW method, a result is obtained that indicates an extremely strong positive linear relationship (0.9994), while the value of this coefficient for ranks is 0.9945. The values are almost perfectly matched, which confirms that the scaling did not distort the data structure, but preserved the proportionality and ranking among the alternatives. This high correlation means that the scaled values can be reliably used as a substitute for the original predictions in evaluation and classification procedures, without loss of information precision.

4. Conclusions

In this research, a multi-level analysis of predictive modeling, scaling and evaluation of educational data was carried out, with the aim of providing a proportional, reproducible and technically consistent classification of grades. Two prediction paradigms were tested, Random Forest regression and multidimensional Linear Regression, representing methodologically distinct approaches, with the application of clearly defined parameters and evaluation metrics. Beyond measuring accuracy, the approach also allows for a qualitative analysis of the contribution of individual criteria in the formation of the overall assessment. Although MCDM methods have traditionally been used for ranking and evaluating alternatives based on multiple criteria, the results of this research confirm that they can also be successfully applied for numerical prediction.

The results of the Random Forest model (MSE=81.93, R^2=0.66) and the Linear Regression (MSE=80.83, R^2=0.66) indicate a stable ability of both models to explain the variation of the target variable, with a moderate prediction error. The analysis of the importance of input attributes showed that the average weekly hours of independent study are the most dominant factor (76.28%), while attendance at classes (13.74%) and class participation (9.98%) have secondary importance. This distribution of weights allows direct integration into MCDM decision methods, thus providing a transparent and empirically validated evaluation.

After scaling the predictions into proportional values of the total score of the SAW method, classification stability was preserved, whereby the scores remained consistent in almost all cases. The Pearson correlation between predicted and scaled values was 0.9994, while the correlation between ranks was 0.9945, confirming that the transformation preserved order and proportionality with almost perfect precision. This stability is crucial for reliable classification and evaluation in automated decision systems, especially in educational and research contexts.

By using empirically based weights from models such as Random Forest, MCDM approaches allow the generation of aggregate scores that are then proportionally scaled to predictive values. Such a transformation allows MCDM logic to be used not only for decision making, but also for quantitative assessment of the target variable, while preserving the interpretability of the model. This extends the MCDM methodology beyond its classical domain, offering a transparent and reproducible alternative for predictive modeling in the context of educational, social and technical systems.

However, several limitations should be noted. The models are trained on structured data with a limited number of criteria, which may affect generalization in more complex environments. The model was tested on only ten alternatives. Also, the classification boundaries between grades are not adaptive, but fixed, which can lead to oscillations in borderline cases. In future work, it would be useful to explore dynamic classification thresholds, integrating additional semantic and contextual factors, and applying advanced models such as Gradient Boosting and Explainable AI approaches to increase interpretability and accuracy.

Acknowledgement
This research was not funded by any grant.

Conflicts of Interest
The author declare no conflicts of interest.

References
[1] Verma, U., Garg, C., Bhushan, M., Samant, P., Kumar, A., & Negi, A. (2022, March). Prediction of students' academic performance using Machine Learning Techniques. In 2022 International Mobile and Embedded Technology Conference (MECON) (pp. 151-156). IEEE. https://doi.org/10.1109/MECON53876.2022.9751956
[2] Trung, D., Truong, N., Duc, D., & Bao, N. (2024). Data Normalization in RAWEC Method: Limitations and Remedies. *Yugoslav Journal of Operations Research, 35*(3), 467-482. https://doi.org/10.2298/YJOR240315020T
[3] Özekenci, E. K. (2025). A Multi-Criteria Framework for Economic Decision Support in Urban Sustainability: Comparative Insights from European Cities. *International Journal of Economic Sciences*, *14*(1), 162-181. https://doi.org/10.31181/ijes1412025188
[4] Seifi, N., Keshavarz, M., Kalhor, H., Shahrakipour, S., & Adibifar, A. (2025). Ranking of Criteria Affecting the Implementation Readiness of Internet of Things in industries Using TISM and Fuzzy TOPSIS Analysis. *Journal of Operations Intelligence*, *3*(1), 47-67. https://doi.org/10.31181/jopi31202533
[5] Hou, Y., Zhang, Y., & Liao, J. (2025). Performance Evaluation of Mechanized Construction of Overhead Transmission Lines Based on FUCOM-F-CM. *Engineering Research Express*, *7*(3), 035576. https://doi.org/10.1088/2631-8695/adf523
[6] Tešić, D., Božanić, D., & Puška, A. (2024). Smartphone Selection Using Structured User Reviews: A Hybrid Random Forest and Fuzzy DIBR II–WASPAS Approach. *Journal of Innovative Research in Mathematical and Computational Sciences*, *3*(2), 71-93. https://doi.org/10.62270/jirmcs.v3i2.37
[7] Turgay, S., & Aydin, A. (2025). Improving decision making under uncertainty with data analytics: Bayesian networks, reinforcement learning, and risk perception feedback for disaster management. *Journal of Decision Analytics and Intelligent Computing, 5*(1), 25–51. https://doi.org/10.31181/jdaic10009052025t
[8] Krones, F., Marikkar, U., Parsons, G., Szmul, A., & Mahdi, A. (2025). Review of multimodal machine learning approaches in healthcare. *Information Fusion*, *114*, 102690. https://doi.org/10.1016/j.inffus.2024.102690
[9] Zhang, H., Liu, Y., Zhang, C., & Li, N. (2025). Machine learning methods for weather forecasting: A survey. *Atmosphere*, *16*(1), 82. https://doi.org/10.3390/atmos16010082
[10] Božanić, D. I., & Pamučar, D. S. (2010). Evaluating locations for river crossing using fuzzy logic. *Military Technical Courier*, 58(1), 129-145. https://doi.org/10.5937/vojtehg1001129B
[11] Pamučar, D., Božanić, D., Đorović, B., & Milić, A. (2011). Modelling of the fuzzy logical system for offering support in making decisions within the engineering units of the Serbian army. *International journal of the physical sciences*, *6*(3), 592-609. https://doi.org/10.5897/IJPS11.012
[12] Pamučar, D., Božanić, D., & Đorović, B. (2011). *Fuzzy logic in decision making process in the Armed Forces of Serbia*. LAP Lambert Academic Publishing

[13] Pamučar, D., Đorović, B., Božanić, D., & Ćirović, G. (2012). Modification of the dynamic scale of marks in analytic hierarchy process (AHP) and analytic network approach (ANP) through application of fuzzy approach. *Scientific Research and Essays*, *7*(1), 24-37. https://doi.org/10.5897/SRE11.373

[14] Tešić, D., Božanić, D., & Milić, A. (2023). A multi-criteria decision-making model for pontoon bridge selection: An application of the DIBR II-NWBM-FF MAIRCA approach. *Journal of Engineering Management and Systems Engineering*, *2*(4), 212-223. https://doi.org/10.56578/jemse020403

[15] Petkovski, I., & Vranić, P. (2025). PCA - Enhanced regression approach for predicting internet use based on formal education. *Journal of Decision Analytics and Intelligent Computing*, *5*(1), 87–98. https://doi.org/10.31181/jdaic10014062025p

[16] Breiman, L. Random Forests. *Machine Learning*, 45, 5–32 (2001). https://doi.org/10.1023/A:101093340432

[17] Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024, 69-79. https://doi.org/10.58496/BJML/2024/007

[18] Iranzad, R., & Liu, X. (2025). A review of random forest-based feature selection methods for data science education and applications. *International Journal of Data Science and Analytics*, 20(2), 197-211. https://doi.org/10.1007/s41060-024-00509-w

[19] Fishburn, P. C. (1967). Additive Utilities with Incomplete Product Sets: Applications to Priorities and Assignments. *Operations Research*, *15*, 537–542, https://doi.org/10.1287/opre.15.3.537

[20] Roustaei, N. (2024). Application and interpretation of linear-regression analysis. *Medical Hypothesis, Discovery and Innovation in Ophthalmology*, *13*(3), 151. https://doi.org/10.51329/mehdiophthal1506

[21] Taherdoost, H. (2023). Analysis of Simple Additive Weighting Method (SAW) as a MultiAttribute Decision-Making Technique: A Step-by-Step Guide. *Journal of Management Science & Engineering Research*, *6*(1), 21–24. https://doi.org/10.30564/jmser.v6i1.5400

[22] Karakuş, C. B. (2024). Assessment of ecotourism potentiality based on GIS-based fuzzy logarithm methodology of additive weights (F-LMAW) method for sustainable natural resource management. *Environment, development and sustainability*, *26*(10), 27001-27055. https://doi.org/10.1007/s10668-024-05283-0

[23] Ginting, S. H. N. ., & Sridewi, N. (2024). Implementation of Decision Support System for New Employee Selection at PT Triotech Solution Indonesia using SAW Method. *Jurnal Minfo Polgan*, *13*(1), 856-862. https://doi.org/10.33395/jmp.v13i1.13842

[24] Grybaitė, V., & Burinskienė, A. (2024). Assessment of Circular Economy Development in the EU Countries Based on SAW Method. *Sustainability*, *16*(21), 9582. https://doi.org/10.3390/su16219582

[25] Radulescu, C. Z., & Radulescu, M. (2024). A Hybrid Group Multi-Criteria Approach Based on SAW, TOPSIS, VIKOR, and COPRAS Methods for Complex IoT Selection Problems. *Electronics, 13*(4), 789. https://doi.org/10.3390/electronics13040789

[26] Jones, L., Barnett, A., & Vagenas, D. (2025). Linear regression reporting practices for health researchers, a cross-sectional meta-research study. *PloS one*, 20(3), e0305150. https://doi.org/10.1371/journal.pone.0305150

[27] Nabiollahi, K., Kebonye, N. M., Molani, F., Tahari-Mehrjardi, M. H., Taghizadeh-Mehrjardi, R., Shokati, H., & Scholten, T. (2024). Assessment of Land Suitability Potential Using Ensemble Approaches of Advanced Multi-Criteria Decision Models and Machine Learning for Wheat Cultivation. *Remote Sensing*, *16*(14). https://doi.org/10.3390/rs16142566

[28] Impollonia, G., Croci, M., & Amaducci, S. (2024). Upscaling and downscaling approaches for early season rice yield prediction using Sentinel-2 and machine learning for precision nitrogen fertilisation. *Computers and Electronics in Agriculture*, 227, 109603. https://doi.org/10.1016/j.compag.2024.109603

[29] Chowdhury, S. J., Mahi, M. I., Saimon, S. A., Urme, A. N., & Nabil, R. H. (2023, January). An integrated approach of MCDM methods and machine learning algorithms for employees' churn prediction. In 2023 3rd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) (pp. 68-73). IEEE. https://doi.org/10.1109/ICREST57604.2023.10070079

[30] Qureshi, N. (n.d.). Student Performance Dataset [Dataset]. Kaggle. Retrieved August 27, 2025, from https://www.kaggle.com/datasets/nabeelqureshitiii/student-performance-dataset/code

[31] Rodgers, J.L., & Nicewander, W.A. (1988). Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, *42*(1), 59. https://doi.org/10.2307/2685263