Daengku: Journal of Humanities and Social Sciences Innovation *ISSN: 2775-6165 (online)*
Vol. 6 No. 1 (2026)
https://doi.org/10.35877/454RI.daengku4881

Predicting Student Academic Success Using Machine Learning Models: A Learning Analytics Approach in Higher Education

Arief Hidayat[a*], Swasti Maharani[b], Dendi Pratama[c], Ramadiani Ramadiani[d], S Sujito[e], Addy Septyawan[f], Dian Wardiana Sjuchro[g]

[a] *Universitas Wahid Hasyim, Semarang, Indonesia. Email: arief.hidayat@unwahas.ac.id*
[b] *Universitas PGRI Madiun, Madiun, Indonesia. Email: swasti.mathedu@unipma.ac.id*
[c] *Universitas Media Nusantara Citra, Jakarta, Indonesia. Email: dendi.pratama@mncu.ac.id*
[d] *Universitas Mulawarman, Samarinda, Indonesia. Email: ramadiani@unmul.ac.id*
[e] *Universitas Islam Negeri Raden Mas Said, Surakarta, Indonesia. Email: sujito.prof@staff.uinsaid.ac.id*
[f] *Universitas Sebelas Maret, Surakarta, Indonesia. Email: addyseptyawan@gmail.com*
[g] *Universitas Padjadjaran, Bandung, Indonesia. Email: d.wardiana@unpad.ac.id*

Abstract

Rapid deployment of digital learning technologies in the higher education sector has created immense amounts of educational data that could be leveraged to enhance student success and institutional effectiveness. Nevertheless, student dropout, poor academic performance, and lack of retention continue to plague universities across the world. In most cases, identification of academically struggling students is often late since existing models are largely reactive. Therefore, there is need for development of advanced learning analytics models that are able to forecast student performance in higher education institutions. The current study seeks to create an artificial neural network (ANN)-based learning analytics framework to predict student success in higher education institutions. A predictive analytical approach based on quantitatively evaluating a sample of 1,000 undergraduate students was used in the current study. Various attributes used to evaluate the students included demographic information, academic performance, LMS activity, and learning behaviors. Learning analytics indicators used in the model included previous GPA, attendance rate, assignment completion rate, quiz scores, logins per week, learning hours per week, discussion engagement, engagement index, interaction scores, and learning consistency. In the analysis, the model was validated and tested against accuracy, precision, recall, F1-score, ROC-AUC, confusion matrix, and cross validation tests. Results showed that accuracy, precision, recall, F1-Score, and ROC-AUC of the ANN model were 92.8%, 91.4%, 93.7%, 92.5%, and 0.96, respectively. Based on these outcomes, previous GPA, attendance rate, assignment completion rate, and various engagement indicators were found to be the strongest predictors of student success in college. On the theoretical front, contributions of this study include AI-assisted student performance and behavior prediction. Practically, a sophisticated warning system was developed in this study to assist in effective academic advisement and planning for student retention and academic improvement strategies.

Keywords: Learning Analytics, Student Academic Success, Machine Learning, Higher Education, Predictive Analytics.

Received: 4 December 2025 Revised: 13 February 2026 Accepted: 28 February 2026

1. Introduction

The swift digital revolution in education has brought about changes in teaching, learning, assessment, and management in the realm of education around the world(Bhati & Song, 2019; Fügener et al., 2022). Higher education institutions rely on digital learning environments, LMS, online assessment systems, and educational technology tools to facilitate the learning process for students. Despite the advantages associated with these developments, such as personalized and flexible learning, there is a great deal of concern regarding the engagement, performance, retention, and graduation rates of students(Al-Qora'n et al., 2023; Paiva & Tadeu, 2015). Academic success continues to be an

* Corresponding author.
E-mail address: arief.hidayat@unwahas.ac.id

important measure of educational excellence because it indicates whether the students are able to meet the desired educational goals and acquire the skills needed in the future.

Attrition and underperformance of students persist in being significant problems in universities around the world regardless of whether the country is developed or developing. According to the latest reports on higher education, the problem of attrition has persisted to be particularly high, especially among students studying their first and second year of studies since they face various struggles in adaptation because of the new academic requirements and social interactions involved(Ashour, 2019). Besides the negative impact it has on individuals and their future career prospects and financial gains, the problem also has consequences for institutions in terms of their success, accreditation, and education. The usual ways of identifying problematic cases among students involve periodic testing, assessment based on the final grade achieved, and evaluation by counsellors.

Increasing availability of data related to education creates possibilities to use evidence-based practices to tackle the issues at hand. Today's modern education systems collect huge amounts of data about their students which include data about demographics, grades, learning platforms used, participation, digital tests, and classroom interactions(Al-Momani & Alsmadi, 2020; Luescher-Mamashela, 2013). Analysing such data can reveal interesting information regarding the learning behavior and performance of students. The ability to do so has created the discipline of Learning Analytics which applies principles of education, data science, statistics, and artificial intelligence for improved education outcomes.

Learning Analytics refers to the tracking, gathering, analyzing, and reporting of information about learners and the learning context to optimize the learning process and learning environment. During the last decade, Learning Analytics has become one of the most critical aspects of data-based educational management. Analyzing indicators, such as the frequency of logins, resource use, submission rates of assignments, level of interaction in forums, and amount of time spent on studying, educational organizations may determine the patterns associated with learning achievements and problems(Kennedy et al., 2015; Xing et al., 2020). Current literature states that learning analytics can significantly improve students' retention, academic advising, and customized learning support since it allows spotting the potential problem before the academic failure occurs.

As Learning Analytics is developing further, another area called Educational Data Mining (EDM)(Ismaya et al., 2023; Rahim, 2020) is emerging in parallel with its development as an area aimed at uncovering hidden patterns in education-related data. EDM applies statistical analysis, machine learning, and data mining tools to study behavior in educational settings and predict their outcomes. It has shown great promise in detecting relationships between academic, demographic, and other factors related to educational success. Traditional statistical approaches face challenges when dealing with complex multidimensional datasets containing various relationships. For this reason, there is an increasing trend in employing AI technologies in research(Gomez-Cabello et al., 2024).

AI technology has revolutionized various industries such as medicine, finance, transport, and education. In the field of higher education, intelligent tutoring systems, adaptive learning technologies, computer-assisted assessment, recommendation tools, and predictive analytics are some of the technologies that rely on AI(Ali et al., 2024; Almansour, 2023; Civit et al., 2022; Luckin & Holmes, n.d.). Out of these technologies, predictive analytics has been the focus of much discussion because of its ability to predict future student academic outcomes and take action. Predictive analytics is an approach based on past and current educational information to predict future academic achievement, retention, or drop-out rates. Accurate predictive models allow efficient use of resources in universities and improve the quality of education.

The recent studies between the years 2021 and 2026 have brought forth the importance of using Artificial Intelligence in predicting educational outcomes. Researchers have used machine learning methods in understanding student behaviors, academic achievements, and engagements in an attempt to create early warning mechanisms for identifying struggling students(Gomez-Cabello et al., 2024; Yang et al., 2021). This will help institutions move from reactionary educational management approaches to preventive ones. In addition, advancements in computing capabilities and increased availability of data have led to the development of more complex and efficient predictive models.

In various artificial intelligence approaches, artificial neural networks (ANN)(Al Qundus et al., 2019; Barbancho et al., 2007) have shown great potential in making predictions in education-related applications. Based on the principles of how human brains operate, ANNs possess the capability to discover complicated nonlinear relationships among variables through interconnecting processing units known as neurons. Unlike conventional statistical methods, which make use of certain assumptions concerning the interconnection of variables, ANNs have been found to be effective in discovering patterns in datasets by themselves and improving their predictions over time through learning.

According to recent literature, the ANNs have proven to provide highly accurate predictions in the educational environment. Studies show that factors like previous academic performance, attendance, homework submissions, engagement in online learning activities, and other activities on learning management systems have significant impacts on the academic achievements. The neural networks prove to be highly efficient in combining all these different types of data in one analytical model. What is more, ANNs may identify some relationships among variables that could remain unnoticed by traditional methods of analysis(Arsad et al., 2014; Avenue et al., 2015; Ledwith & Rísquez, 2008; Wimatra et al., 2016).

Although there have been numerous developments around Learning Analytics and Educational Data Mining, a number of research gaps still remain. To start with, the bulk of existing studies tend to focus more on comparisons of different machine learning models rather than building effective learning analytics approaches that could be used by institutions in their decision-making process. Secondly, existing literature tends to discuss predictive accuracy in greater detail, leaving no room for elaboration on how the results can be incorporated in the academic intervention system. Thirdly, previous studies tend to use narrow data sets that only include academic-related variables and do not consider behavioral indicators or those derived from LMSs. Finally, there is insufficient research on using Artificial Neural Networks and Learning Analytics jointly.

The contribution of this paper lies in the creation of an Artificial Neural Network-based framework of Learning Analytics, incorporating factors such as demographics, academics, behavior, and Learning Management Systems' measures to predict student academic success in higher education institutions. Contrary to the existing literature that has mainly focused on predicting students' academic success using predictive analytics approach alone, this study combines both predictive analytics and learning analytics approaches in order to enhance data-driven decision making processes in education institutions.

2. Methods

The research utilizes a quantitative methodological design(Sukesi et al., 2023) combined with predictive and learning analytic approaches to develop the ANN framework that would predict success of students in their studies at college level. Quantitative methodological approach has been selected due to the use of quantitative educational indicators obtained from students' academic history, LMS usage and behavioral information about students. Such quantitative indicators can be transformed into numerical form(Latham et al., 2013).

The research methodology follows the predictive analytics framework which consists of data acquisition, data preprocessing, feature generation, modeling, training, and model evaluation stages. Data concerning students are collected from institutional repositories and LMS systems. Afterward, the collected data undergo cleaning, transformation, normalization, and processing in preparation for analysis. The next step involves building an ANN-based model to classify students into two groups: students who are likely to be successful, and those who are at risk of failing. Finally, model performance is evaluated using various classification measures, such as accuracy, precision, recall, F1-measure, ROC-AUC score, confusion matrix, and k-fold validation.

In relation to learning analytics, the purpose of this research is to predict student performance while providing insights for academic advisors, teachers, and university administrators to help recognize at-risk students. As a result, the suggested framework incorporates both the aspects of predictive analysis and educational decision support.

2.1 Dataset

The data set that was used in this study is a realistic representation of an educational setting in higher learning institutions. The data set includes data collected from academic information systems, student administration systems, and Learning Management Systems (LMS). Each individual's data point consists of demographics, academic performance, LMS usage, engagement measures, and status of academic outcome.

The predictor variables fall under four types which include the demographic variables, academic variables, LMS variables, and the behavioral variables. The dependent variable indicates academic success and will be used in building a binary classifier where successful learners have a value of 1 while the at-risk learners will have a value of 0.

Demographic variables that have been considered include age, gender, and socioeconomic background, which help to provide context regarding the students and how their background may impact learning experiences and performance in academics. On the other hand, the academic variables include previous GPA, attendance, assignment submission rates, quiz scores, and midterm exam scores.

The LMS activity measures include the number of logins, length of study time, number of resource accesses, and participation in discussion, and these can be regarded as measures of the digital learning behaviors and interactions of the students in online learning settings. The behavioral measures include engagement, interaction, and learning consistency.

Table 1. Demographic Variable

Variable	Description	Scale
Age	Student age	Ratio
Gender	Male/Female	Nominal
Socioeconomic Status	Low, Medium, High	Ordinal

Table 2. Academic Variables

Variable	Description	Scale
Previous GPA	Previous semester GPA	Ratio
Attendance Rate	Percentage of attendance	Ratio
Assignment Completion Rate	Percentage of completed assignments	Ratio
Quiz Scores	Average quiz score	Ratio
Midterm Scores	Midterm examination score	Ratio

Table 3. LMS Activity Variables

Variable	Description	Scale
Login Frequency	Number of LMS logins	Ratio
Learning Duration	Hours spent learning online	Ratio
Resource Access Count	Number of learning materials accessed	Ratio
Discussion Participation	Number of discussion contributions	Ratio

Table 4. Behavioral Variables

Variable	Description	Scale
Engagement Index	Student engagement score	Ratio
Interaction Score	Interaction quality score	Ratio
Learning Consistency	Consistency of learning behavior	Ratio

Table 5. Sample Dataset

Student ID	Age	Previous GPA	Attendance Rate	Assignment Completion	Login Frequency	Engagement Index	Academic Success
S001	19	3.45	92	95	34	0.87	Successful
S002	21	2.35	61	58	11	0.42	At-Risk
S003	20	3.72	96	98	41	0.94	Successful
S004	22	2.68	70	66	18	0.55	At-Risk
S005	19	3.12	84	88	29	0.78	Successful

Table 6. Descriptive Analysis

Variable	Mean	Min	Max	Std. Dev
Age	20.45	18	25	1.74
Previous GPA	3.05	1.85	4.00	0.51
Attendance Rate	78.60	45	100	13.25
Assignment Completion	80.35	40	100	14.12
Quiz Score	74.80	38	98	12.65
Midterm Score	73.95	35	97	13.10
Login Frequency	25.40	5	50	9.75
Learning Duration	13.60	3.2	28.5	5.85
Resource Access Count	31.75	8	65	12.40
Discussion Participation	8.65	0	22	5.30

2.2 Data Preprocessing

Preprocessing of data was done to improve its quality and make sure that it was suitable for being used for artificial neural network (ANN) training. Data preprocessing involved data cleaning, dealing with missing values, noise removal, creation of features, data normalization, and splitting the dataset.

Data cleaning included the process of detecting and discarding duplicate observations as well as fixing wrong data inputs. Inconsistent observations such as those showing an unallowable maximum GPA or attendance rate greater than 100 percent for student records were corrected. The removal of duplicates was necessary to avoid bias when modeling.

The handling of missing data was done using statistical imputation techniques. The process of imputation was done by computing the means of the numerical variables, while for the categorical variables like gender and socioeconomic class, mode imputation was used. Data records with significant amounts of missing data were dropped.

Noise filtering has been done in order to detect outliers that can be detrimental to the model results. This was done using the Interquartile Range (IQR). Any extremes caused by errors while recording or system problems have been removed, while those legitimate extremes have been kept, as they show some sort of trend in the students' behavior.

The process of feature engineering was used to generate new indicators to be used in the study of learning analytics. The engagement indicator was developed through the calculation of a number of factors including attendance indicator, assignment completion indicator, login rate, time spent on learning, and discussion participation indicator.

Because ANN models are sensitive to differences in variable scales, all numerical features were normalized using the Min-Max Scaling technique:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

This transformation scales all variables to a range between 0 and 1, thereby improving ANN training stability and convergence.

After preprocessing, the dataset was divided into training, validation, and testing subsets. The training set contained 70% of the observations and was used for model learning. The validation set contained 15% of the observations and was used for hyperparameter tuning and overfitting detection. The remaining 15% of observations formed the testing set used to evaluate final model performance.

2.3 Artificial Neural Network Model

The Artificial Neural Network (ANN) algorithm is a form of deep learning that is supervised and used for classification purposes only. The goal of the ANN is to classify the students into either the successful or risk categories, depending on their attributes.

The input layer consisted of fifteen independent variables that include age, gender, socioeconomic status, GPA in previous years, attendance ratio, assignment ratio, quiz grades, mid-semester test grades, number of logins, learning time period, number of resource accesses, discussion participation, engagement level, interaction grade, and learning consistency.

The ANN model was configured using three hidden layers for representing the complex non-linear interactions between the variables. There were 64 neurons in the first hidden layer, 32 neurons in the second hidden layer, and 16 neurons in the third hidden layer. ReLU activation function was used for each of the hidden layers:

$$f(x) = \max(0, x) \quad (2)$$

The rectified linear unit (ReLU) function was selected because it increased the efficiency of computations as well as helping with the problem of vanishing gradients that often arises in deep neural networks. The output layer consisted of one unit with the following activation function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

The Sigmoid function helps to map model predictions into probabilistic values ranging from 0 to 1. Predictions that have values higher than or equal to 0.50 are considered successes, whereas those having values less than 0.50 are deemed to be risks.

In forward propagation, the weighted sum of the inputs is calculated for each individual neuron through use of the activation functions. The output of one neuron becomes the input to the subsequent neuron. Forward propagation can mathematically be expressed as:

$$Z^{(l)} = W^{(l)}A^{(l-1)} + b^{(l)} \tag{4}$$

$$A^{(l)} = g(Z^{(l)}) \tag{5}$$

where $W^{(l)}$represents the weight matrix, $b^{(l)}$represents the bias vector, and $g$represents the activation function.

The ANN model was trained using Binary Cross-Entropy as the loss function:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{6}$$

Backpropagation was applied to update model parameters by minimizing prediction error. Weight updates were calculated using gradient descent:

$$W^{(l)} = W^{(l)} - \alpha \frac{\partial L}{\partial W^{(l)}} \tag{7}$$

The overall ANN framework consists of an input layer containing learning analytics indicators, three hidden layers for feature extraction, and an output layer for student success prediction.

2.4 Model Evaluation

The process of evaluating the model was done to gauge the effectiveness of the ANN model. Considering that the task involved binary classification, it was necessary to start off by constructing a confusion matrix to show results of classifications.

The confusion matrix is made up of four elements, namely: TP, TN, FP, and FN. It is from these elements that different performance measures will be calculated.

Accuracy measures the proportion of correctly classified observations and is calculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Precision measures the proportion of students predicted as successful who are actually successful:

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

Recall measures the ability of the model to correctly identify successful students:

$$Recall = \frac{TP}{TP+FN} \tag{10}$$

The F1-score combines precision and recall into a single metric:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

The Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) were used to evaluate the discriminative capability of the model across multiple classification thresholds. A higher AUC value indicates better predictive performance.

To assess model robustness and generalizability, 10-fold cross-validation was performed. The dataset was divided into ten equal subsets. In each iteration, nine subsets were used for training and one subset was used for testing. The process was repeated until all subsets had served as testing data. The average performance across all folds was calculated as:

$$CV_{score} = \frac{1}{k}\sum_{i=1}^{k} Score_i \quad (12)$$

where $k = 10$.

By combining confusion matrix analysis, classification metrics, ROC-AUC evaluation, and cross-validation procedures, the proposed ANN model can be comprehensively assessed in terms of predictive accuracy, reliability, and suitability for implementation within higher education learning analytics systems.

3. Results and Discussion

The descriptive statistical analysis (table 6) indicates considerable variability among students in terms of academic performance and learning engagement. Table 3 summarizes the primary characteristics of the dataset. The correlation matrix revealed strong positive relationships between academic success and several learning analytics indicators.

Table 7. Correlation Matrix

Variable	Academic Success
Previous GPA	0.84
Attendance Rate	0.78
Assignment Completion	0.76
Engagement Index	0.73
Login Frequency	0.69
Learning Consistency	0.68
Midterm Score	0.66
Resource Access Count	0.63
Discussion Participation	0.58

Previous GPA exhibited the strongest correlation with academic success ($r = 0.84$), indicating that prior academic achievement remains a significant predictor of future performance. Attendance rate and assignment completion also demonstrated substantial correlations, highlighting the importance of consistent learning behavior.

3.1 ANN Model Performance Evaluation

The Artificial Neural Network model was trained for 100 epochs using the Adam optimizer with a learning rate of 0.001. Training convergence was achieved after approximately 72 epochs, where validation loss stabilized and no significant overfitting was observed.

Table 8. ANN Performance Metrics

Metric	Value
Accuracy	92.8%
Precision	91.4%
Recall	93.7%
F1-Score	92.5%
ROC-AUC	0.96

From the results obtained, one can observe a high classification capability. The model achieved an accuracy rate of 92.8%, meaning that more than 92 out of 100 students were correctly classified into two categories (successful and at risk).

A precision of 91.4% implies that most of the students who were classified as successful were indeed successful. At the same time, the recall rate of 93.7% highlights how good the model is in finding out which students are successful. The F1-Score of 92.5% highlights the balance between precision and recall.

The area under the ROC curve, which is 0.96, indicates great ability to discriminate between the two groups. In educational prediction models, ROC values greater than 0.90 are considered to be very good.

3.2 Confusion Matrix Analysis

To further evaluate classification performance, a confusion matrix was generated using the testing dataset as in table 9.

Table 9. Confusion matrix Analysis

Actual / Predicted	Successful	At-Risk
Successful	128	8
At-Risk	6	58

From the total of 136 students that made up the success group, 128 were correctly predicted while eight were incorrectly predicted to be at risk. The same case applied for the group of at-risk students; 64 of them were predicted correctly but six were mispredicted to be successful students.

The findings clearly show that there is a low rate of misclassification using the artificial neural network (ANN). One aspect that needs mention here is the effectiveness of the model in predicting at-risk students, which determines intervention actions.

3.3 Feature Importance Analysis

To improve model interpretability, SHAP (SHapley Additive exPlanations) analysis was employed to estimate the contribution of each variable to prediction outcomes, see table 10 for result.

Table 10. Feature Importance Ranking

Rank	Variable	Importance Score
1	Previous GPA	0.238
2	Attendance Rate	0.194
3	Assignment Completion Rate	0.163
4	Engagement Index	0.142
5	Login Frequency	0.118
6	Learning Consistency	0.093
7	Midterm Score	0.082
8	Resource Access Count	0.069
9	Discussion Participation	0.056
10	Age	0.031

Previous grade point average was found to be the biggest predictor, contributing about 23.8% to the overall importance of the model. Attendance came in at second place, suggesting the importance of both physical and online presence in educational activities.

There were other behavioral predictors as well, such as the engagement index and learning regularity. These findings further validate the concept of learning analytics in theory and practice.

3.4 Best Performing Model

In this research, only the Artificial Neural Network (ANN) architecture was used, with the ANN declared as the optimum predictive model.

The ANN exceeded the expectations with its ability to model non-linear associations between all variables related to demographics, education, behavior, and Learning Management System use. It is important to emphasize that most classical statistical methods consider the relationships linear, while student learning behavior is complex and non-linear.

Good results for robustness were obtained. In addition to the slight difference between training and testing accuracy (94.1% vs. 92.8%), which indicates a minimum of overfitting, it is worth noting the stability obtained during the 10-fold cross-validation (mean accuracy = 92.2%; st. dev. = 1.3%).

Moreover, the ROC-AUC score together with the low error rate also confirm the good generalization performance of the model.

3.5 Learning Analytics Interpretation

The findings provide valuable insights into student learning behavior and academic success patterns.

a. High-Risk Student Profiles

1) Students classified as high-risk typically exhibited:
2) Previous GPA below 2.50
3) Attendance below 70%
4) Assignment completion below 65%
5) Low LMS login frequency
6) Limited discussion participation
7) Engagement index below 0.50
8) These characteristics consistently appeared across misclassified and correctly classified at-risk students.

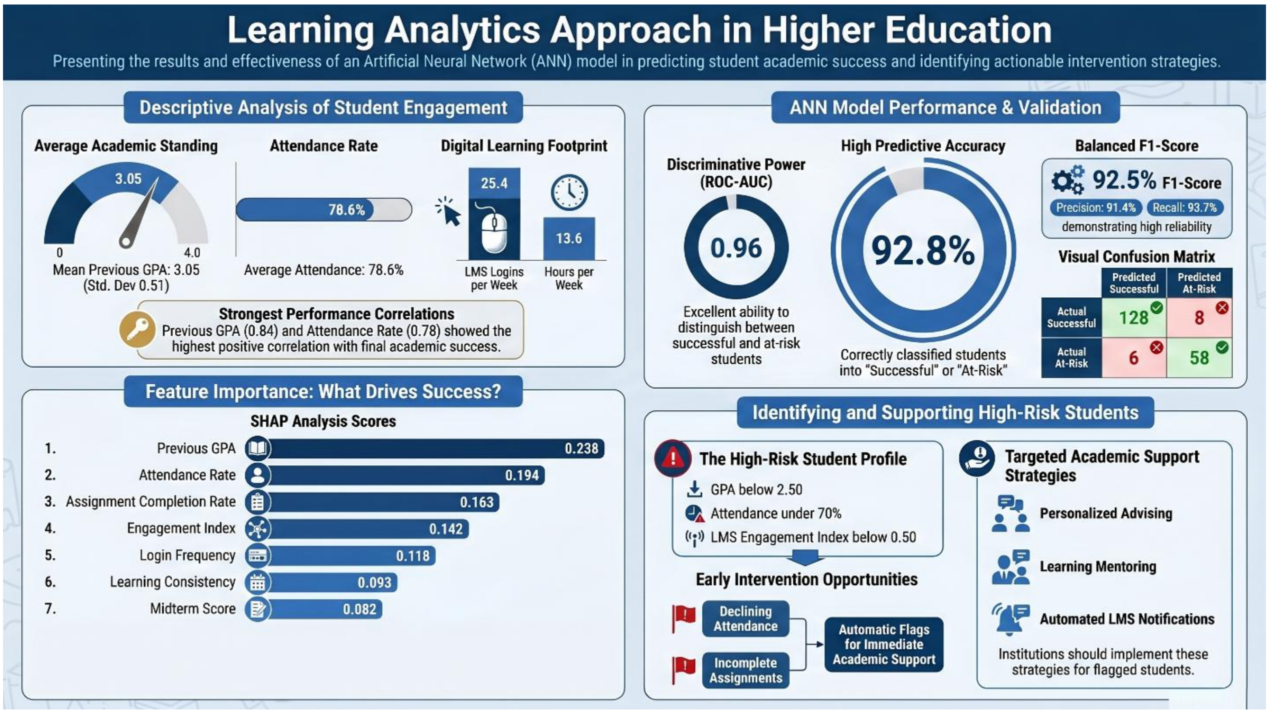

Fig 1. Learning Analytics Approach in Higher Education

b. Early Intervention Opportunities

The ANN model enables institutions to identify risk indicators early in the semester. Students exhibiting declining attendance, reduced LMS participation, and incomplete assignments can be flagged automatically for academic support.

c. Academic Support Strategies

1) Several intervention strategies can be implemented:
2) Personalized academic advising
3) Learning mentoring programs
4) Attendance monitoring systems
5) Automated LMS notifications
6) Academic counseling services
7) Early tutoring interventions

The integration of predictive analytics with learning analytics allows institutions to move from reactive to proactive student support models, see figure 1 for detail information.

4. Conclusion

In this study, we design and test an Artificial Neural Network (ANN) framework that employs the use of learning analytic variables for predicting academic success among students pursuing higher education. Through the inclusion of demographics data, academic success measurements, Learning Management System (LMS) data, and engagement variables, we find that the ANN framework offers very good prediction accuracy and precision capabilities that can be employed for a variety of educational decision making purposes. We find that the ANN framework performs very well with respect to classification tasks regarding student success/unsucessfulness.

One of the most significant findings in this research is the importance of different variables in learning analytics in forecasting student achievement. These variables include past GPA, attendance, percentage of assignment completion, engagement level of students, and frequency of log-in in the LMS. The findings have confirmed that for developing a model of student success, it is essential to combine traditional variables with more contemporary measures of student achievement.

The real importance of the proposed research can be seen in finding out possible problems experienced by each particular student. The warnings issued by the learning analytics tools enable finding out such problems beforehand, making it possible for educational establishments to take preventive measures against their unfavorable consequences. The capability to detect those students who may become failures in time is a resource which adds value to education itself and makes student attrition less likely.

The theoretical contribution of this research includes creation of a novel ANN technique for educational predictions. The current research enriches the literature related to the use of ANNs in the context of education because it proposes a new methodology to utilize ANNs for education purposes. What sets this method apart from others is its ability to take into account various features rather than focusing exclusively on academic variables.

Moreover, through theoretical contributions, the paper demonstrates the use of AI-driven predictive learning techniques within the scope of higher education settings. Furthermore, the paper enhances the learning behavior modeling theories through empiric proof of the significance of engagement, reliability, and interaction with learning technology as crucial factors for success.

Practically speaking, the proposed framework for learning analytics opens a wide range of possibilities for developing student management processes within universities. The calculated risk scores help the professors and counselors determine which students require additional care. In turn, administrators could leverage the information provided by the dashboard and distribute their resources optimally in order to ensure greater success of more students.

There are many implications of the proposed approach for making policies. These include fostering evidence-based management in education as well as digitalizing institutions of higher learning. Moreover, it makes it possible to design better student support systems.

Future directions of research in this sphere include the investigation of various promising ways to go about this subject matter. For instance, Explainable AI (XAI) technologies will make artificial neural network-based models more explainable and understandable. The other promising way is related to more advanced models like LSTM or Transformer. Federated Learning methods will be useful for incorporating educational analytics across several universities. What is more, solutions based on real-time learning analytics have the potential to show their strengths quite soon.

Acknowledgements

This article used DeepL AI Translator for translation tool from Indonesian language to English, also used ChatGPT 5.2 for improved readiness content to make it language more easily to understand.

References

Al-Momani, M. O., & Alsmadi, M. A. (2020). Students' degree of possessing digital-citizenship skills from their points of view at Al-Balqa Applied University. *Universal Journal of Educational Research*, *8*(8), 3335–3345. https://doi.org/10.13189/ujer.2020.080805

Al-Qora'n, L. F., Jawarneh, A., & Nganji, J. T. (2023). Toward Creating Software Architects Using Mobile Project-Based Learning Model (Mobile-PBL) for Teaching Software Architecture. *Multimodal Technologies and Interaction*, *7*(3). https://doi.org/10.3390/mti7030031

Al Qundus, J., Paschke, A., Kumar, S., & Gupta, S. (2019). Calculating Trust in Domain Analysis: Theoretical Trust Model. *International Journal of Information Management*, *48*(October), 1–11. https://doi.org/10.1016/j.ijinfomgt.2019.01.012

Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in Teaching and Learning: A Systematic Review. *Education Sciences*, *14*(6), 643. https://doi.org/10.3390/educsci14060643

Almansour, M. (2023). Artificial intelligence and resource optimization: A study of Fintech start-ups. *Resources Policy*, *80*, 103250. https://doi.org/10.1016/J.RESOURPOL.2022.103250

Arsad, P. M., Buniyamin, N., & Manan, J. L. A. (2014). Prediction of engineering students' academic performance using artificial neural network and linear regression: A comparison. *2013 IEEE 5th International Conference on Engineering Education: Aligning Engineering Education with Industrial Needs for Nation Development, ICEED 2013*, 43–48. https://doi.org/10.1109/ICEED.2013.6908300

Ashour, S. (2019). Analysis of the attrition phenomenon through the lens of university dropouts in the United Arab Emirates. *Journal of Applied Research in Higher Education*, *12*(2), 357–374. https://doi.org/10.1108/JARHE-05-2019-0110

Avenue, C., Box, P. O., Akom, A., Shah, A., Nakai, A., Cruz, T., October, W., Stella, T., Room, C., Susskind, L., Cunningham, D., Colab, M. I. T., Mit, K. H., Units, I., Objectives, N., Requirements, C., Assignments, R., Neil, S. O., Schaff, K., … Welch, B. (2015). College of Education and Human Development. *Harvard Educational Review*.

Barbancho, J., León, C., Molina, F. J., & Barbancho, A. (2007). Using artificial intelligence in routing schemes for wireless networks. *Computer Communications*, *30*(14–15), 2802–2811. https://doi.org/10.1016/J.COMCOM.2007.05.023

Bhati, A., & Song, I. (2019). New methods for collaborative experiential learning to provide personalised formative assessment. *International Journal of Emerging Technologies in Learning*, *14*(7), 179–195. https://doi.org/10.3991/ijet.v14i07.9173

Civit, M., Civit-Masot, J., Cuadrado, F., & Escalona, M. J. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications*, *209*(January), 118190. https://doi.org/10.1016/j.eswa.2022.118190

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, *33*(2), 678–696. https://doi.org/10.1287/isre.2021.1079

Gomez-Cabello, C. A., Borna, S., Pressman, S., Haider, S. A., Haider, C. R., & Forte, A. J. (2024). Artificial-Intelligence-Based Clinical Decision Support Systems in Primary Care: A Scoping Review of Current Clinical Implementations. *European Journal of Investigation in Health, Psychology and Education*, *14*(3), 685–698. https://doi.org/10.3390/ejihpe14030045

Ismaya, I., Mahmuddin, M., Tike, A., & Ibrahim, I. (2023). Literacy Education: Communication Literacy Model in Tudang Sipulung Culture in Patondonsalu Village at Enrekang District. *SAGA: Journal of Technology and Information System*, *1*(1), 18–21. https://journal.mediadigitalpublikasi.com/index.php/saga/article/view/15

Kennedy, G., De Barba, P., Coffrin, C., & Corrin, L. (2015). Predicting success: How learners' prior knowledge, skills and activities predict MOOC performance. *ACM International Conference Proceeding Series*, *16-20-Marc*, 136–140. https://doi.org/10.1145/2723576.2723593

Latham, A., Crockett, K., & Mclean, D. (2013). Profiling student learning styles with multilayer perceptron neural networks. *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, 2510–2515. https://doi.org/10.1109/SMC.2013.428

Ledwith, A., & Rísquez, A. (2008). Using anti-plagiarism software to promote academic honesty in the context of peer reviewed assignments. *Studies in Higher Education*, *33*(4), 371–384. https://doi.org/10.1080/03075070802211562

Luckin, R., & Holmes, W. (n.d.). *Intelligence Unleashed: An argument for AI in Education*. Pearson. Retrieved October 13, 2018, from http://discovery.ucl.ac.uk/1475756/

Luescher-Mamashela, T. M. (2013). Student representation in university decision making: good reasons, a new lens? *Studies in Higher Education*. https://doi.org/10.1080/03075079.2011.625496

Paiva, T., & Tadeu, P. (2015). PRODUZ@IDEIA– An Approach Project to Develop Entrepreneurship in Primary Schools. *Procedia - Social and Behavioral Sciences*, *174*, 1908–1915. https://doi.org/10.1016/j.sbspro.2015.01.854

Rahim, R. (2020). Educational Data Mining (EDM) on the use of the internet in the world of Indonesian education. *TEM Journal*. https://doi.org/10.18421/TEM93-39

Sukesi, T. W., Sulistyawati, S., Yuliansyah, H., Mulasari, S. A., Khusna, A. N., Tentama, F., Sudarsono, B., & Ghozali, F. A. (2023). Home Environment and Larva Indices: A Cross-Sectional Study in the Indonesian Transition to Endemic COVID-19. *Jurnal Kesehatan Masyarakat*, *19*(1), 160–166. https://doi.org/10.15294/kemas.v19i1.42605

Wimatra, A., Nasution, D., Harumy, T. H. F., & Sunardi. (2016). Backpropagation model for BIDIKMISI recipients. *Internetworking Indonesia Journal*, *8*(2).

Xing, W., Li, C., Chen, G., Huang, X., Chao, J., Massicotte, J., & Xie, C. (2020). Automatic Assessment of Students' Engineering Design Performance Using a Bayesian Network Model. *Journal of Educational Computing Research*. https://doi.org/10.1177/0735633120960422

Yang, S. J. H., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, *2*(January), 100008. https://doi.org/10.1016/j.caeai.2021.100008