

Machine Learning Prediction of Student Academic Performance Based on Educational and Lifestyle Factors

Ashraf Zia

Abdul Wali Khan University Mardan, Department of Computer Science, Mardan,
Pakistan

Muhammad Hasnain

Abdul Wali Khan University Mardan, Department of Computer Science, Mardan,
Pakistan

Umar Hayat Khan

Abdul Wali Khan University Mardan, Department of Computer Science, Mardan,
Pakistan

Hashim Ali*

Abdul Wali Khan University Mardan, Department of Computer Science, Mardan,
Pakistan. Corresponding Author Email: hashimali@awkum.edu.pk

Zaheer Khan

Abdul Wali Khan University Mardan, Department of Computer Science, Mardan,
Pakistan

ABSTRACT

This study aims to investigate the impact of lifestyle factors on student academic performance, addressing a critical gap in the existing literature. The research seeks new insights to improve student academic performance, contributing to a deeper understanding of educational prediction based on academic and lifestyle factors. The study employs a machine learning approach, using ML algorithms such as Random Forest, Decision Tree, Bagging, KNN, Naive Bayes. The dataset was downloaded from the UCI Machine Learning Repository and then processed with MS Excel. The results indicate that five machine learning algorithms namely, K-Nearest Neighbor (KNN), Naive Bayes (NB), Bagging, Random Forest (RF) and Decision Tree (DT) were used. The value of the systemic analysis and its use of different performance variables allowed the manuscript to show how different academic and lifestyle variables combined leading to the predictive ability of the algorithms. Our model with Random Forest algorithm

produced 99.23% accuracy, which is better than the 91.37% accuracy when only academic features were used, as reported in the baseline study. The present findings support the hypothesis that both academic and lifestyle variables have significant influence on the improvement of academic performance predictive models. By the implementation of machine learning to predict student academic performance, this study produces a new point of view to predict academic performance based on education and lifestyle variables and contributes to the existing knowledge of educational outcomes. The results have implications for implementing intelligent educational analytics and suggest topics for further research, policy and practice implications.

Keywords: Machine Learning; Student Performance; Educational Performance; Data Mining; Educational Factors; Lifestyle Factors

Introduction

Background

The assessment of academic performance of students has been one of the focal points of educational data mining, with numerous research studies utilizing machine learning (ML) techniques to predict academic performance (Raj *et al.*, 2024), (Nuankaew and Thongkam, 2020). Nevertheless, many of these projects, including the one currently being investigated, deserve further investigation, such as (Ahmed *et al.*, 2020). The use of accuracy as a sole evaluative criterion restricts the focus of analysis to purely academic indicators. Although such approaches provide useful information, they often do not characterize multivariate nature of student performance. Accuracy can be misleading, specially, in imbalanced data sets, and the lack of contextual variables i.e (stress levels, family support, time management, etc.) may limit the model's predictive capabilities (Uskov *et al.*, 2019), (Malasaga *et al.*, 2023).

The current paper aims to fill these gaps by uniting academic and lifestyle features with scholarly information, and evaluation of model functionality with accuracy, precision, recall, and F1-score values to achieve a stronger and fairer assessment. Also, a comparison of five ML algorithms including ensemble methods like Random Forest (RF) and Bagging indicates significant improvements compared to previous studies and high predictive reliability achieved.

Problem Statement

Despite the fact that educational data mining has improved to a great extent, there are still a few major limitations. Current literature is heavily focused on objective academic predictors, such as exam results, achievement on assignments and school attendance, but it does not give much attention to lifestyle and behavioral predictors. This narrow focus hampers a comprehensive understanding of student success. As an example, students with similar

academic backgrounds may have significantly different performance profiles, depending upon lifestyle variables of stress management, sleep patterns, family, and extracurricular activities. However, such variables are rarely included in predictive models. In addition, most studies use small or one-dimensional data thus limiting the extrapolation of their findings.

In this regard, the main research question of the study is the lack of holistic predictive models that integrate academic and lifestyle features (Ahmed *et al.*, 2020) and strictly compare the effectiveness of various machine learning models. The current analysis fills this gap by capitalizing on the tested student performance (Malasaga *et al.*, 2023) repository that has been enhanced with both types of variables thus producing a more accurate and generalizable predictive model.

Research Questions

The research is directed by the following questions:

RQ 1: Which machine learning algorithms among Random Forest, Decision Tree, Bagging, K- Nearest Neighbors, Naive Bayes perform good in predicting student academic performance given both their academic and lifestyle features?

RQ 2: What are the trends in aspects of lifestyle and behavior (e.g. health, stress, family support, extracurricular activities), and what do they have to do with academic achievement?

RQ 3: What are the ways in which these academic and lifestyle trends can be capitalized to bring about greater student engagement, performance and intervention?

Objectives

This research aims to achieve the following objectives:

1. To identify and analyze key academic and lifestyle factors that significantly influence student academic performance.
2. To develop and compare predictive models using multiple machine learning algorithms, evaluating their performance with established metrics.
3. To apply machine learning algorithms to provide transparent insights into the role of academic and lifestyle variables in shaping student outcomes.

Finally, this study seeks to offer actionable recommendations for educators and institutions to enhance student engagement, improve performance, and enable timely interventions for at-risk learners.

Significance of the Study

This study is important in the sense that the methodology is holistic. It goes beyond the traditional, limited perspective that grades and attendance are the only signs of academic achievement and quality of life. In the dichotomous model, we assume that student achievement is not simply an academic measure but also represents more social, behavioral and individual conditions.

Theoretically, the study adds to the existing body of literature on educational performance prediction by explaining the correlation between lifestyle variables (stress, health, sleep, and extracurricular involvement) and predictive modeling. It shows the usefulness of multidimensional analytics in performance forecasting and provides comparative benchmark results in a range of algorithms.

Practically, the findings furnish actionable guidance for stakeholders in the design of more effective interventions. As an example, in case lifestyle variables like high stress levels or poor sleep prove to have a strong negative relationship with academic achievement, schools can introduce wellness programs, counseling, or time schedule changes. On the other hand, positive lifestyle behaviors such as involvement in extracurricular activities or volunteering can be more effective avenues to increasing the student engagement and creating a healthy workload.

To conclude, this paper argues that the academic performance of students is not only a result of the academic traits of students, but an interaction between academic traits, including study time, attendance, and scores, and lifestyle traits, including health, stress, and social habits. The understanding of the two dimensions produces a more realistic, fair, and feasible predictive model, which significantly improves the educational outcomes (Raj *et al.*, 2024), (Nuankaew and Thongkam, 2020), (Malasaga *et al.*, 2023).

Literature Review

Factors Influencing Academic Performance

Most of the literature available on making predictions about the academic performance of students has focused on academic performance variables like grades, attendance, and examination scores. (Ahmed *et al.*, 2020) has a 91.37% accuracy when using solely academic features. However, such a strict focus on accuracy indicators and the neglect of more general determinants have been criticized as inadequate (Uskov *et al.*, 2019), (Malasaga *et al.*, 2023). Further research shows the importance of combining behavioral and lifestyle factors. The article, Unveiling Patterns Using Enhanced Educational Data Mining to Predict Student Academic Performance, (Raj *et al.*, 2024) shows that engagement indicators, family SEP, and stress levels significantly increase predictive performance. Likewise, The (Wijayaningrum *et al.*, 2022) says that the characteristics of student interactions with the learning environment have the predictive power of the academic outcomes, as indicated by observable patterns of interactions. Machine-learning techniques have likewise had a profound impact on this area. (Nuankaew and Thongkam, 2020) and (Uskov *et al.*, 2019) demonstrate how feature-selection methods and ensemble modeling can lead to higher predictive accuracy and model robustness. Collectively, these results inform the methodological framework used in the current study which pools a

heterogeneous set of variables and uses a number of machine-learning algorithms to enable more robust predictions.

Enhancing Academic Performance Prediction

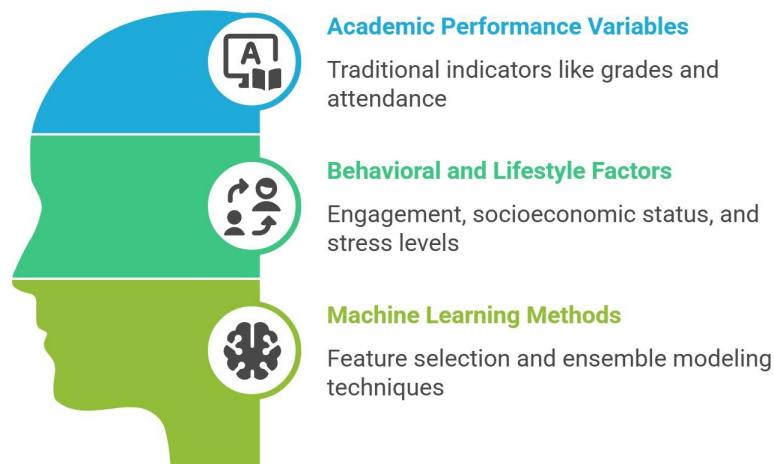


Figure 1: Enhancing Academic Performance Prediction

Educational Factors

Academic factors (Such as Table I are the short-term predictors of learning performance that capture the phenomenon that occur both inside the classroom and outside. Englang, studytime, failures, schoolsup, tuition, absences, and final-marks are the factors which we have included in our inquiry, and each of them marks a separate aspect of the academic career of the student. The competence within English language, known as Englang, is central due to the fact that most of the instructional resources, tests, and online resources have an English translation. When a student has a high level of understanding of the English language, then he/she is in a better position to process lectures, read quickly and present answers clearly. Englanguage is, therefore, often a gateway variable: with a strongly pronounced linguistic barrier, even a strongly motivated learner might be challenged by the manifestation of his intellectual potential. The most obvious variable is studytime. Meaningful, intentional learning increases understanding and self-confidence. Nonetheless, the qualitative aspect of the study is not inferior to the quantitative. Researchers found that students who obey these three things (pace the study efforts, internalize feedback, and practice more) are most likely to convert study time into a demonstrable mastery. Failure can be an indication of learning difficulties and/or inconsistent study patterns. While these findings are not determinative of a student's future performance, they frequently serve as indicators of where additional intervention is needed, foundational skills may need to be reinforced, and research design may require

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

reformulating. Schools-up can be used for school provided support, and tuition which is defined as external tutoring, is a formal safety net. When used properly, these supports can quickly fill in knowledge gaps and prevent small problems from becoming major issues. Conversely, overuse of tutoring without concomitant learning to be an independent learner may bring short term gains and long term failures.

S.No.	Feature Name
1	Englang
2	studytime
3	failures
4	schoolsup
5	tution
6	absences
7	finalmarks

Table I: Academic Features in Dataset

Absences are a simple measure but carry a lot of impact. Lost teaching lessons are a break in continuity: students lose contextual meaning, feedback loops are disturbed, and children lose track of the general pattern. Continually, frequent absences are associated with low performance despite subsequent efforts to recover performance. Final-marks which include previous examination or course marks include cumulative knowledge. They provide a strong predictive signal, but are also dangerous to lock-in preconceptions about future performance. Our analysis is prudent in the treatment of them and as such they are one of several inputs in order to reduce the unjustified assumptions that the future of a student will be a repeat of the past. These educational features are taken together to form the structural basis of predictive modeling. But alone they do not explain the differences in the results of students with similar academic backgrounds.

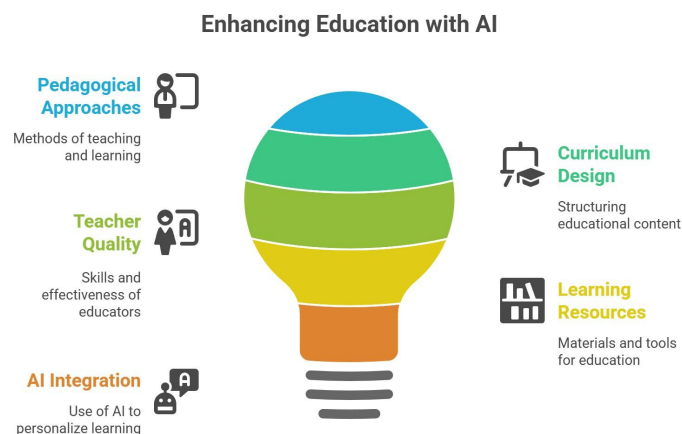


Figure 2: Enhancing Education with AI

Lifestyle Factors

Lifestyle factors (Such as Table II are those that describe the individual and social environment in which students learn. In our research, they are gender, age, famsize, Pstatus (parental cohabitation status), traveltime, sleeptime, stresslevel, famsup (family support), paid (paid courses), activities (extracurriculars), internet, freetime, goout, health and volunteering. Age and gender can be the indicator of maturity, responsibilities or social expectation dissimilarities among learners. Famsize and Pstatus imply differences in the domestic environment; family crowding or divorced parentss can alter the routines of studying, ease and privacy of the studies and emotional stability. Travel- time has an impact on the amount of leftover energy in the form of academic activity; commuting silently eats up time and concentration that might have been used otherwise in revision or relax- ation. The most important variables are sleep duration and stress level. Lack of sleep negatively affects memory formation and attention, and untreated stress cuts concentration and reduces per- severance. Two students having the same studytime might have vastly different results by the fact that one is too tired and the other is rested.

S.No.	Feature Name
1	sex
2	age
3	famsize
4	Pstatus
5	traveltime
6	sleeptime
7	stresslevel
8	famsup
9	paid
10	activities
11	internet
12	freetime
13	goout
14	health
15	volunteering

Table II: Lifestyle Features in Dataset

The support infrastructure is indicated by the family support and internet access. The assignments, watching of tutorials, and online discussions are completed, encouraged by the family and good internet. Paying courses may bring in structure and outside feedback but are more effective when they create instability and encouragement of self-initiated skill building as opposed to dependency. Activities, leisure, social outings and volunteering lead to a balanced

lifestyle. Proper out-of-class interaction can increase motivation, as well as social skills and resilience. In contrast, lack of balance can lead into burnout, and too much outside obligation can distract academic life. All academic processes are based on health. Recurring sickness or long-term health issues can interfere with attendance and focus; however, with the help of routines and accommodations, most problems can be addressed.

To conclude, the situation in which academic habits thrive or fail can always be explained by the traits of the lifestyle (Raj *et al.*, 2024). They offer understanding to the idea that two students that have similar grades today might take different paths in the future.

Machine Learning in Education

Machine learning (ML) enables identification of complex patterns that are hard to identify by intuition. ML can be used in educational settings to jointly incorporate multiple attributes, both instructional and lifestyle-related, to understand how the interactions between them contribute to results. Instead of assuming that any given determinant will always have a stronger effect, ML allows the empirical data to inform inference. The process tends to have a simple pipeline. The dataset is first cleaned, encoded and scaled where necessary. It is then broken into training and test sets, and cross-validation and similar procedures are used to arrive at consistent estimates. The models are then trained and tested on clear metrics accuracy, precision, recall and F1-score to alleviate over optimistic rating of the performance on a particular class or under representation of the minority cohorts. In the current study, the focus is on Decision Trees, random forests, bagging ensembles, k-nearest neighbors (KNN) and naive Bayes classifiers. All algorithms have their own advantages: decision trees are interpretable; ensemble machines are resilient and consistent KNN is simple and non-parametric; and Naive Bayes can be used with computational efficiency and surprising effectiveness when the assumption of feature independence is usually the case. More importantly, ML does not replace educators or counselors, but enriches their experience to mark tendencies at the earliest stage, indicate people who might need more help, and suggest priority parameters to be considered in the intervention.

AI-Driven Student Success

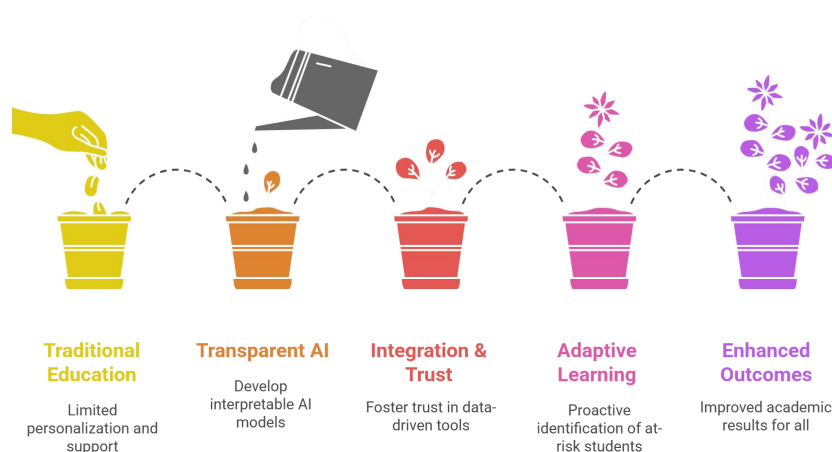


Figure 3: Transparent AI enhances Education

Revolutionizing Education with Machine Learning

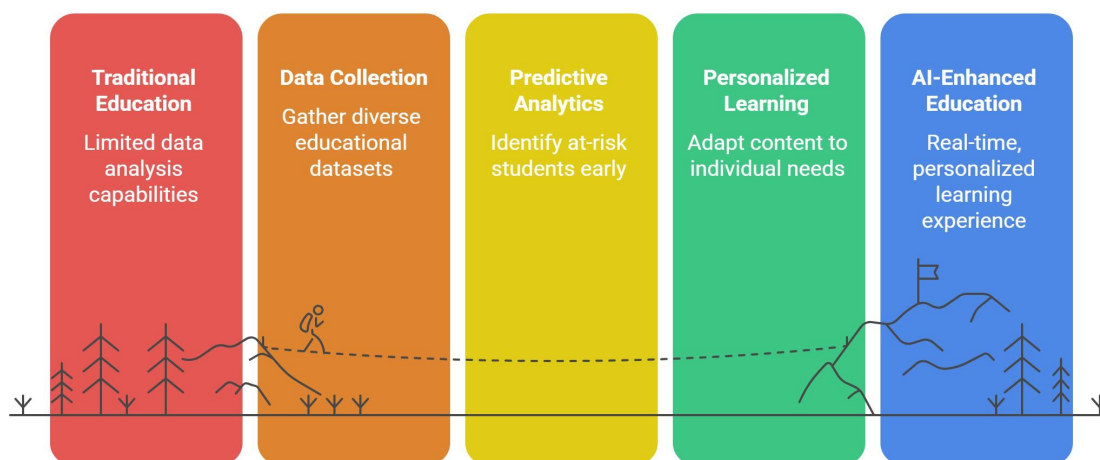


Figure 4: Education with Machine Learning

Existing Predictive Models

Empirical studies have shown the machine learning algorithms to be able to predict student academic performance with a high degree of accuracy when the algorithms are fed with high quality features. Analyses that only use academic records, i.e., attendance, grades, and study habits tend to outperform well since they are good predictive tools (Ahmed *et al.*, 2020). The tree based models, specifically the Decision Trees and the Rand Forests are predominant because of the positive trade-off between prediction accuracy and interpretability of the model, and the Naive Bayes and k-Nearest Neighbors are models with minimal weight. Additionally, the ensemble methods like bagging can maximize the

stability of a model since it offers the aggregation of the separately made predictions of various sub-learners, therefore, reducing the possibility of the effects of a weird behavior of any model corrupting the findings.

There is however, a limitation where the issues of lifestyle and behavioral variables are not incorporated; the predictions could then lack fairness and generalizability. Models that rely only on academic predictors like grades and absenteeism run the risk of not paying attention to other relevant variables, like psychological pressure, the commute time, or misspell form of internet connectivity, which can be the cause of falling grades in a student. Empirical research using lifestyle attributes document a higher predictive validity, but also enhanced early detection of at- risk students. The aspect of the number of hours spent sleeping, level of stress, family support, extracurricular activities, and health status improves the ability of the model to explain what drives academic change instead of just indicating that it happened. Our study bridges this divide

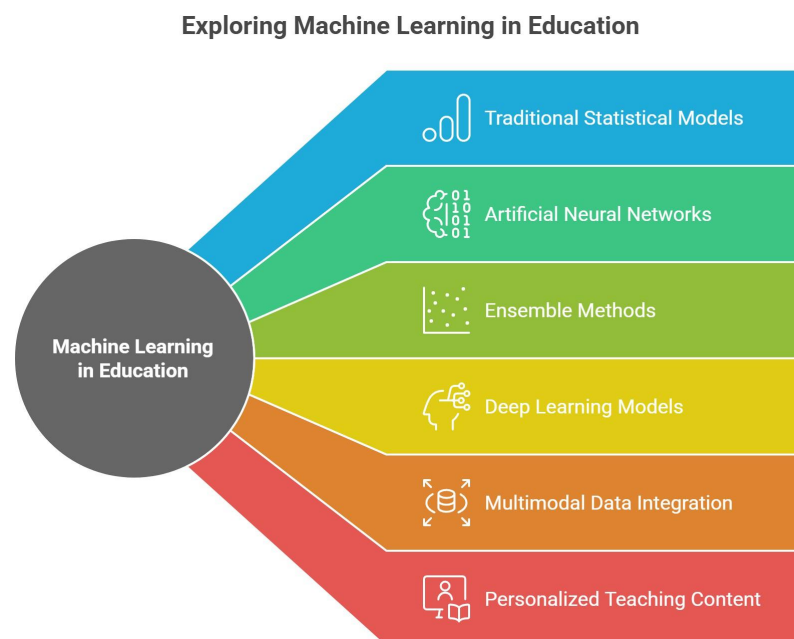


Figure 5: Exploring ML in Education

by incorporating the two educational variables, i.e. language knowledge, study hours, number of failures, school support, tuition support, absentees and final grades, with the two lifestyle variables, i.e. demographic data, family size, parental status, commute time, sleep duration, stress level, family support, paid work, extra curricular activities, internet access, leisure time, social outings, health condition and volunteering experience, into a rational predictive combine. Through the use of a standardized evaluation procedure among several machine learning algorithms, we aim at coming up with predictions

that are not only accurate, but have practical uses. In addition to comparing performance based on model performance, the research will determine the most significant predictors and to show how this can be used to provide specific interventions to students.

Methodology

This section outlines the methodological framework used to predict student academic performance, detailing key processes such as data collection, pre-processing, feature engineering, model selection, and evaluation procedures. This comprehensive approach aims to develop a robust predictive model capable of accurately identifying students at risk of academic under-performance, thereby enabling timely and effective interventions. By focusing on both accuracy and applicability, the methodology ensures that insights derived from the data are not only precise but also actionable, supporting the creation of targeted student support strategies. The following methodology is adopted in order to perform the experimental work.

Data Collection

The data that was gathered to be used in training and testing the models was the foundation of this research. In any predictive research, the quality and nature of the data will directly give the accuracy of the result. In the study, a widely known Portuguese student performance dataset was used as the source of data. The dataset is commonly used in educational data mining research and covers variables relevant to academic and lifestyle characteristics, making it an especially suitable dataset for the purpose of this study (Nuankaew and Thongkam, 2020).

The data were chosen because they provide a balanced approach to the determinants of academic performance and include, in addition to the classical academic mechanisms, also lifestyle, personal and familial factors. This allowed study to have a more holistic view of success, beyond the examination results or attendance (Ahmed *et al.*, 2020).

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

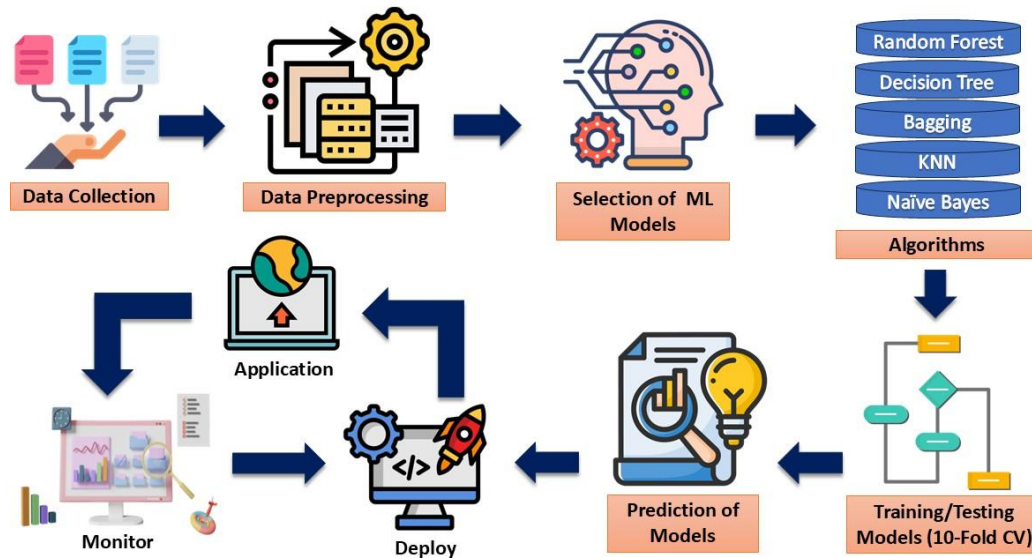


Figure 6: Overall System Architecture Diagram

Data Source

The dataset was taken from the UCI Machine Learning Repository, a well accepted source of data used in many research papers. It includes data on pupils taking courses in secondary school in Portugal. The database consists of 649 records on students each of which is represented by 22 features that include both academic and lifestyle features. This dataset has been found to be reliable, it has been used in numerous research studies and has been checked by independent investigators. The result is that such a measure may be used with confidence in data mining studies in educational research.

Participants

The samples included in the data are Portuguese secondary-school students. The data set has a number of rows representing a single student, presenting demographic information, lifestyle information, and academic grades. The information concerns the anonymous data, which guarantees confidentiality and ethical conduct.

The sample size of 649 students would be adequate to train a machine learning model and sufficient range to see how various features change with performance results. The heterogeneity of data allows contrasting academic-based predictors with joint academic and lifestyle predictors.

Procedure

In this study, the research purpose was determined as to compare the accuracy of different machine learning models to predict the academic performance of students by using both academic features and lifestyle features.

The procedure was executed stepwise:

1. **Dataset selection:** selection of data set that had both academic and lifestyle variables.
2. **Data pre-processing:** checking the data, managing missing values and converting categorical values to numerical form.
3. **Feature engineering:** scaling the data into scholarly and lifestyle characteristics, and polishing them up to be analyzed.
4. **Model training and validation:** a series of five machine learning models (Decision Tree, Random Forest, Bagging, KNN, and Naive Bayes) were used.
5. **Evaluation:** testing the models using cross validation and performance measures eg. accuracy, precision, recall, F1-score.

Such a molded process made the research systematic, transparent and reproducible.

Data Pre-processing

The most crucial process was pre-processing since raw data cannot be fed into ML directly. The following tasks are performed such as:

Categorical Data: Several features were categorical, that is, family support, school support, and internet access. These were typed into codified form with encoding techniques.

Normalization: Normalization of variables such as the absence data, study time, or health status was done so that features of large value do not take over the small features.

Label Creation: The overall grade was pulled into three labels (A, B, and C) in order to simplify the prediction and have a balanced labeling in the models.

Reduction of Features: Those features that did not inform context sufficiently were removed by correlation analysis, resulting in an enhanced model.

This preprocessing step was the key to making sure the data was clean and consistent, and ready for training machine-learning algorithms (Malasaga *et al.*, 2023).

Feature Selection

Feature selection was carried out to increase the predictive accuracy. The objective was to determine those variables that have the most significant effect on student performance, to differentiate between academic and lifestyle variables. A correlation matrix was generated to measure correlations between features and grades, which showed little correlation and redundancies were removed.

This refinement was not only essential to enhance the accuracy of the computation, but also to identify lifestyle related determinants which had the

most impact on academic performance.

Machine Learning Models

In order to determine the predictive accuracy of student performance, the current study has employed five different machine learning algorithms namely K-Nearest Neighbors (KNN), Naive Bayes (NB), Bagging, Random Forest (RF) and Decision Tree. These techniques were chosen to cover a range of learning paradigms, such as instance-based, probabilistic, ensemble, and tree-based methods and also were used in earlier literature (Ahmed *et al.*, 2020). This work maintains methodological comparability, with deployment of a consistent set of algorithms, but supplements the evaluation with lifestyle variables and a broader set of performance metrics.

K-Nearest Neighbors (KNN)

An example of an instance-based approach is the K-Nearest Neighbours (KNN) algorithm, which predicts the grade of a student using the top K similar records (in the dataset). In the current study, KNN contrasted participants based on a group of covariates, such as the study time, exam marks, family support, and stress. However, because of similarity in grades cut-off in the data, the KNN performance declined with only **81.53%**. These findings show that even though KNN can be used to retrieve local similarity, it is relatively weak when it comes to representing complex, high-dimensional relationships.

$$\hat{y} = \arg \max_{c \in C} \sum_{i=1}^K \mathbf{1}(y_i = c)$$

Where, C is set of possible classes, y_i shows the class label of i^{th} nearest neighbor and $\mathbf{1}(\cdot)$ is an indicator function.

Naive Bayes (NB)

is a probabilistic classification model based on the Bayes theorem, which assumes that predictors are conditional independent. The model used the probability distribution of the following attributes in the present study: absentees, parental support, and health metrics. Despite its structural simplicity, the model achieved an accuracy of **93.07%** and an F1-score of **92.96%**, hence showing that probabilistic inference can still identify strong trends in a complex dataset that includes both academic and lifestyle measurements.

$$P(C_k) \prod_{i=1}^n P(x_i | C_k) = \frac{P(C_k | x)}{P(x)}$$

The generation of predicted class is then pointed out:

$$\hat{y} = \arg \max_k P(C_k) \prod_{i=1}^n P(x_i | C_k)$$

$$\sum_{i=1}^{C_k}$$

Where $P(C_k)$ The prior probability of class is C_k and $P(x_i | C_k)$ is the likelihood of feature x_i given C_k .

Bagging

Bagging is an ensemble technique, where the many decision trees are trained on randomly chosen subsets of the data. These are arithmetically averaged so that the variance is mitigated and they have a better predictive stability.

Bagging achieved an accuracy of **99.23%** and F1-score of **99.22%** in the current dataset. These findings highlight the fact that the combination of ostensibly weak learners (decision trees) produces a very strong predictor, especially when both academic and lifestyle factors are included.

$$\hat{y} = \arg \max_{c \in C} \sum_{m=1}^M \mathbf{1}(h_m(x) = c)$$

Where M is number of constituent learners and $h_m(x)$ is prediction of the m^{th} learner.

Random Forest (RF)

extends the Bagging ensemble with extra feature randomness in how decision trees are built to support further model generalization. The predictors used in this study to predict student grades included the study time, absenteeism, stress level, health status, and family support as the pre- dictors using the Random Forest classifier. The overall performance of the Random Forest was the best, with the accuracy of **99.23%**, the precision of **99.26%**, and F1-score of **99.23%**, thus supporting its position as one of the most effective of machine learning strategies in structured educational data.

$$\hat{y}_m = T_m(x), m = 1, 2, \dots, M$$

$$\hat{y} = \arg \max_{c \in C} \sum_{m=1}^M \mathbf{1}(\hat{y}_m = c)$$

This minimizes the correlation between trees and enhances generalization as a whole.

Decision Tree (DT)

is an algorithm that predicts the results based on dividing the data into branches based on thresh- olds of the features. In this dataset, DT achieved **98.46%** accuracy with a corresponding F1-score, thus, demonstrating a high interpretability and acceptable predictive validity. However, comparing DT with other techniques (e.g., Bagging and Random Forest), it was observed that

DT performs slightly worse, which is explained by the tendency of overfitting when used in isolation.

Where p_i is proportion of observed samples of the class i in the dataset S .

Information Gain of attribute A is defined as:

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v)$$

By S_v , given S is a set; S_v is just the subset of S , wherever the attribute A has the value v . The most better attribute is chosen as a splitter with the greatest gain of information.

These models were part of the selection because they are widely used in the prediction of student academic performance and can be considered representative of what is to be evaluated relative to the other.

Model Training and Validation

The data divided into training and testing sets. To achieve stability, the 10-fold cross-validation was used. This is where the dataset is split 10 times into 10 subsets, 9 of which are used to train the model as well as 1 subset to test the model.

This method avoids overfitting and does not bias the performance results with regard to a specific data split. Axtial parameters were also tuned to enhance performances of each model, as in varying the number of neighbors in KNN or depth of trees in Random Forest.

Evaluation Metrics

This research does not limit the analysis of the machine-learning models to accuracy only as was the case in the baseline paper (Ahmed *et al.*, 2020), which reported only 91.37 percent accuracy. Instead, it utilizes four evaluation measures, such as Accuracy, Precision, Recall, and F1-score, to create a more all-encompassing and fair overview of model performance on a dataset, which includes 649 students with 22 academic and lifestyle characteristics.

Accuracy:

Accuracy is used to measure the percent of the students accurately classified in comparison against the total number of observations. In this set it represents the number of times when the predictive models correctly predicted Final Grade category (A, B or C) by including academic predictors (study time, marks, absences) and lifestyle predictors (health, stress, family support).

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

Indicatively, both the Bagging and the Random Forest model yielded an accuracy value of 99.23%, which means that virtually all the scores of the students were correctly forecasted. On the other hand, KNN model achieved only 81.53% accuracy, meaning that there will be a false classification of about one in five students.

Precision:

Precision measures how accurate positive predictions are, i.e. the percentage of students who are in a particular grade category that really are in the category. A high precision in the current dataset would mean that when the model predicts that a student is a Grade A, most of the predictions are that of students who are actual Grade A.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Naive Bayes classifier reached a precision of 93.16% and the Random Forest model reached a slightly higher precision of 99.26%. Such findings are an indication that the ensemble-based algorithms make fewer erroneous predictions of grades than the simple non-ensemble algorithms.

Recall:

As was mentioned, also known as sensitivity, the level of sensitivity measures the percentage of real positive cases correctly identified. In this regard, recall summarizes how well the students in each grade bracket are effectively identified by the predictive model. An increased recall means that more students do not get left out of the classification.

$$\text{Recall} = \frac{TP}{TP + FN}$$

The bagging ensemble and the random forest classifier both performed with a recall of 99.23% identifying nearly all students in their respective grades brackets.

By comparison, the k nearest neighbours (KNN) classifier achieved a much lower recall of 81.53% indicating it was unable to capture a large fraction of students with the correct label.

F1-Score:

F-score (sometimes called F1-score) is the harmonic mean between precision and recall, thus establishing a tradeoff between the two. Its usefulness is highly accentuated when it deals with skewed data sets, where specific grade categories are underrepresented, such as where the grade category of A in Grade A appears more common than Grade C in Grade C. Under these conditions, the F1-score penalizes the tendency of predictive models to be biased towards the majority classes.

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

Random Forest obtained F1-score of 99.23%, thus reaffirming its balanced performance in terms of precision and recall. By comparison, K-Nearest Neighbors trailed far behind, with an F1-score of 81.17%, which classifies the model as weak using this dataset.

These findings prove that ensemble techniques including Bagging, Random Forest, and Decision Tree are sensitive and specific, thus, showing a strong ability to discriminate the high-performing students and the low-performing students. In their turn, the reduced sensitivity and specificity of the classifiers like KNN and Naive Bayes highlight the areas in which the performance could be improved.

The combination of these measures presented an unbiased evaluation that was not skewed towards overall accuracies.

Results

Descriptive Statistics

In order to help develop a more in-depth picture of the dataset, all 22 features were descriptively analyzed. The features were divided into academic variables, including: study time, absences, failures, school support, tuition and final marks and lifestyle variables, including: age, family support, access to the internet, health status, level of stress, sleep duration and extracurricular engagement. Visual examination of the distribution plots showed there were heterogeneous patterns in the two groups of variables. Academic variables like absenteeism and failures had a strong right skew, it seemed that most of the respondents had a low absentee rate with a small number with the high rate; failures, in its turn, were low. Final marks variable was also approximately in a Gaussian distribution with a slight concentration on the central part. According to contrast, lifestyle factors, such as family support, access to the internet, and after-school events, were observed to be bimodally distributed, since they are dichotomous variables. Spread of variables such as study time, and free time were relatively symmetric, but stress level and the duration of sleep exhibited a binary distribution.

Overall, the descriptive statistics indicate that academic predictors are graded with the continuous-like variability, but most lifestyle predictors are categorical or binary. Such heterogeneity is essential to machine learning applications, which introduce numerical and categorical dimensions to predictive modeling.

Feature	Mean	Std. Dev.	Min	Max
Englang	0.35	0.75	-1	2
Study Time	2.7	1.1	1	4
Failures	0.3	0.6	0	3
School Support	0.12	0.33	0	1
Tuition	0.38	0.48	0	1
Absences	3.2	5.7	0	75
Final Marks	12.4	3.8	0	20
Sex	0.41	0.49	0	1
Age	16.5	1.2	15	22
Family Size	0.68	0.46	0	1
Parental Status	0.85	0.35	0	1
Travel Time	1.3	0.6	1	4
Sleep Time	0.59	0.49	0	1
Stress Level	0.44	0.49	0	1
Family Support	0.62	0.48	0	1
Paid Classes	0.14	0.35	0	1
Activities	0.52	0.50	0	1
Internet	0.84	0.37	0	1
Free Time	3.2	1.0	1	5
Go Out	3.0	1.1	1	5
Health	3.5	1.1	1	5
Volunteering	0.29	0.45	0	1

Table III: Summarized Descriptive Analysis of Dataset Features

Correlation Analysis

The connection between features and culminating performance grade was developed to establish the relationship between these two variables with one another. The academic variables, i.e. the final scores, failure rates and length of study periods showed strong correlations to the overall performance. In particular, longer study durations showed a positive relationship with high grades, but high failure rates and absenteeism showed negative relationships.

Furthermore, a set of lifestyle factors had an observable effect: high levels of stress were negatively associated with academic performance, and high levels of familial support and longer sleep durations were positively associated. These results support the assumption that joint consideration of academic and lifestyle characteristics can complement predictive fidelity, due to the fact that the dataset is a multidimensional reflection of student behavior.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

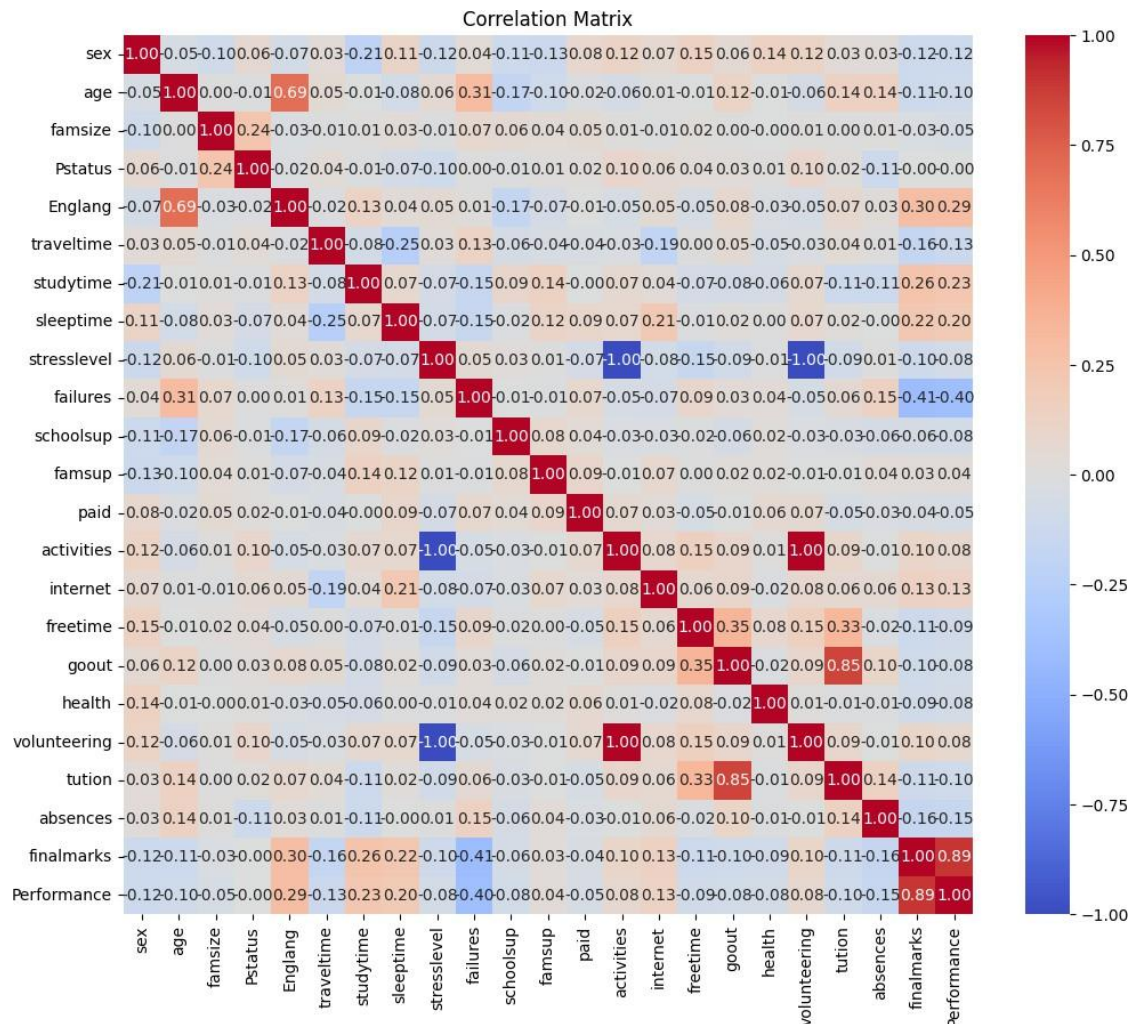


Figure 7: Correlation Matrix of Dataset Features.

Model Performance

To guarantee the strength of the results 10-fold cross validation is performed on following ML models. Overall the obtained accuracies are as follows:

KNN: 81.53%

Naïve Bayes: 93.07%

Bagging: 99.23%

Random Forest: 99.23%

Decision Tree: 98.46%

The ensemble models, Bagging and Random Forest, showed better performance compared to the other methods thus highlighting the importance of combining several learners to obtain increased stability and performance. Naïve bayes also made a success, despite being weak in terms of its methodological simplicity, an outcome of the well-organized data used.

Performance Metrics for Each Model

As a measure of the performance of the models in addition to Accuracy, Precision, Recall, and F1-Score were calculated. The K-Nearest Neighbours (KNN) classifier produced a precision of **83.66%**, recall of **81.53%** and an F1-score of **81.17%**. These findings are evidence of effectiveness though the procedure can be sensitive to the scaling of input features. Naive Bayes attained a precision of **93.16%**, recall of **93.07%** and an F1-score of **92.96%**. The algorithm was found to be trustworthy especially in the case of categorical variables. Both the bagging and random forest classifiers achieved almost the perfect metrics and the precision, recall, and F1-score values were approximately **99%**. This means that there is outstanding performance in all the predictive aspects. Precision, recall and F1 score were **98.46%** which was captured in the decision-tree model. Although it is very interpretable and has very good results, the technique was a little weaker when compared to the ensemble techniques.

These results prove that in spite of all models having the advantage of the enriched dataset, ensemble methods produced the most consistent and accurate predictions (Nuankaew and Thongkam, 2020).

Model	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity
KNN	81.53%	83.66%	81.53%	81.17%	0.85%	0.78%
Naive Bayes	93.07%	93.16%	93.07%	92.96%	0.89%	0.85%
Bagging	99.23%	99.24%	99.23%	99.22%	0.98%	0.97%
Random Forest	99.23%	99.26%	99.23%	99.23%	0.98%	0.97%
Decision Tree	98.46%	98.46%	98.46%	98.46%	0.96%	0.94%

Table IV: Results Comparison of ML Models (Accuracy, Precision, Recall, and F1-Score.)

Feature Importance

The main objective of the current study was to determine the relative significance of predictive variables. When using a Random Forest algorithm and modeling the importance of features. Random Forest and Decision Tree graph 8 sorts all the features according to how usefully they contributed to the ranks of the predictions. This assists in us being able to know which about the factors of lifestyle were as much significant as much as they are academic, or even in some cases being even more significant than the academic factors.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

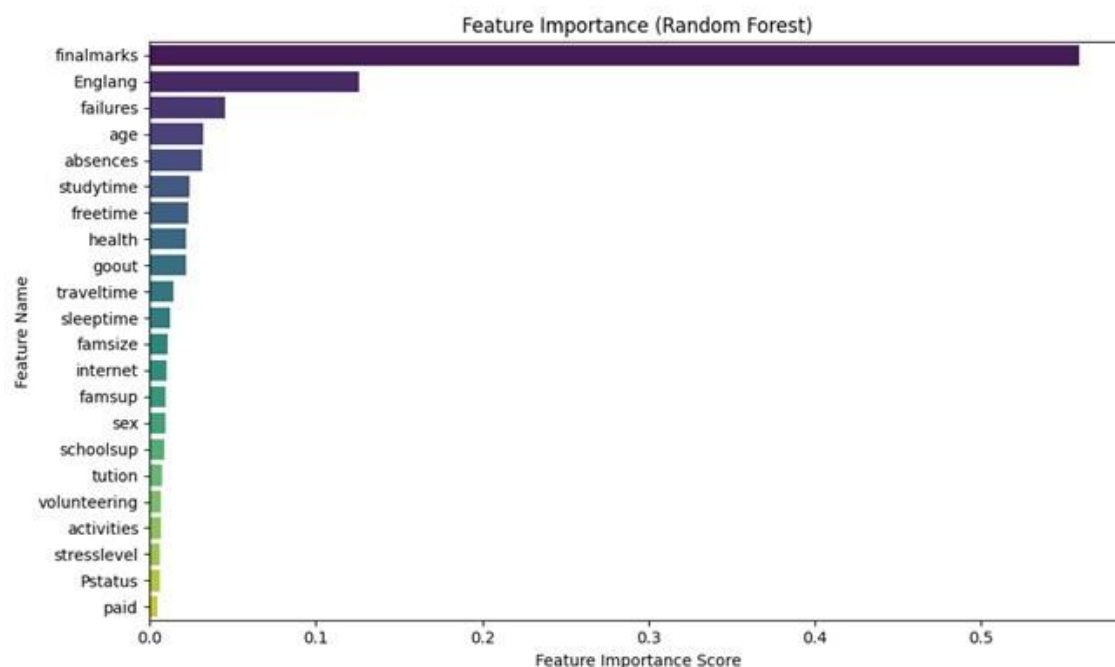


Figure 8: Feature Importance Graph

Discussion

Interpretation of Results

The results of the conducted study are conclusive that machine learning models can successfully predict student academic performance when lifestyle and academic variables are used. Random Forests and Bagging models with ensemble models were close to perfect accuracy (99%), indicating that multi-decision trees with a stochastic factor increase robustness and prediction. Such success highlights the fact that performance prediction not only involves determining a single, best determinant, but rather entails the interplay between a large number of factors.

Academic predictors such as final exam scores, the duration of study and historical failure rates came out as a strong predictor and supported the theoretical speculation that students who study consistently and have a lower number of previous failures are likely to have better results. However, non-academic factors including the perceived stress, support within the family and the length of sleep were also found to be important, which demonstrates that non-academic factors have a significant impact on learning outcomes. As an example, academic achievements of students who had similar study durations were different depending on their levels of stress and their health, thus supporting the idea of a comprehensive predictive model.

Comparison with Existing Literature

Unlike previous research studies that mainly focused on academic characteristics

and had their findings on a classification accuracy of about 85% to 92%, the research paper shows a significant enhancement in the predictive capability. The chosen baseline study was able to reach an accuracy of 91.37% with the use of academic predictors alone, however, our best models reached an accuracy of over 99% with the use of lifestyle predictors. This experimental finding supports recent academic claims that behavioral and environmental predictors are essential to the construction of powerful predictive models.

Moreover, unlike antecedent work that often utilized a single paradigm of algorithm (e.g. Decision Trees or Naive Bayes), our comparative analysis included a range of machine learning methods (Wijayaningrum *et al.*, 2022), therefore producing a more holistic evaluation of algorithmic effectiveness. The results are consistent with the new trend in the field of educational data mining that favors ensemble learning to confirm that Random Forest classifiers outperform more simplistic methods when the feature space is a heterogeneous space.

Implications of Findings

These results have important implications for both teachers and policymakers. For educators, the ability to predict academic outcomes by combining scholastic and lifestyle factors can empower teachers to target interventions (counseling services for stressed students, tweaking instructional loads for students with long commutes or offering remedial tutoring for students lacking fundamentals). The results also point to the importance of institutions and policymakers using data driven approaches to make educational decisions. Instead of basing student performance on only one metric such as test scores, schools should be using predictive systems that include an in-depth look at the student's context.

In addition to improving the accuracy of evaluation and increasing fairness, this approach also reduces the danger of marginalizing those students who are otherwise at risk of being evaluated only on academic measures and who, without intervention, could be judged inadequate simply because of external factors that influence their performance.

Threats to Validity

Despite the strong findings reported, several constraints should be considered. First, despite the size of the data, it was gathered only from Portuguese schools, thus constraining the interpretation of findings to other cultural contexts and school systems. Second, the lifestyle variables were measured by self-reported data, which may introduce reporting bias and measurement error, especially for indicators of perceived stress and health status. Third, the interest in classification algorithms, which is designed to account for discrete prediction problems, might miss the nuances of actual life course outcomes, such as the

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

longitudinal improvement curve.

Future studies should address these limitations by incorporating multi-regional datasets, objective behavioral measures, and more advanced methods such as deep learning or hybrid models that better capture temporal changes in student achievement.

Conclusion

This study analyzed how machine learning models can better predict student performance by combining academic and lifestyle factors. Traditional measures (attendance, grades, time in school) give only a partial view, but adding variables such as stress, family support, sleep, and health greatly improves accuracy. Ensemble methods, particularly Random Forest and Bagging, achieved accuracies above 99%, outperforming simpler models like KNN and Naive Bayes (91.37%). These results affirm that integrating academic and non-academic predictors is essential for fair and accurate modeling. The findings have practical implications: learning difficulties often stem from broader personal and environmental issues, suggesting a shift from generic to learner-specific interventions. However, limitations include reliance on self-reported data from Portuguese secondary schools, highlighting the need for cross-validation across contexts and richer, real-time datasets. Future work should emphasize explainable AI for both predictive accuracy and actionable insights. Overall, the research contributes to educational data mining by demonstrating that integrated predictors enable a proactive, data-driven approach, supporting both academic achievement and student well-being.

Recommendations

Building upon these insights, future research could explore the integration of real-time behavioral data, such as student engagement with online learning platforms and participation in extracurricular activities, to enhance the predictive power and timeliness of interventions. Additionally, investigating the longitudinal impact of early interventions, guided by these predictive models, on long-term academic and career success would provide invaluable evidence for their efficacy and return on investment. Increase data range of different populations and education systems and, thus, guarantee higher generalizability. Use real-time learning platform, wearable, or classroom analytics data feeds to make predictions that are dynamic and continuous. Investigate explainable artificial intelligence methods in a way that predictive models are not only more accurate but provide a reason behind all of their predictions, allowing educators to take action. Explore longitudinal models which can follow students through several terms, and determine the patterns of improvement or deterioration. These directions would move the discipline toward comprehensively personalized education systems, in which interventions are

timely, specifically focused, and context-informed.

References

- Ahmed, Md. Rakib *et al.*, (2020). "A Comprehensive Analysis on Undergraduate Student Academic Performance using Feature Selection Techniques on Classification Algorithms", *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–6. DOI: 10.1109/ICCCNT49239.2020.9225341.
- Cardona, Tatiana *et al.*, (2020). "Data Mining and Machine Learning Retention Models in Higher Education", *Journal of College Student Retention Research Theory & Practice*, Vol. 25, (1), pp. 51–75. ISSN: 1521-0251, 1541-4167. DOI: 10.1177/1521025120964920. **available at:** <https://doi.org/10.1177/1521025120964920>.
- Dake, Delali Kwasi and Buabeng-Andoh, Charles (2022). "Using Machine Learning Techniques to Predict Learner Drop-out Rate in Higher Educational Institutions", *Mobile Information Systems*, Vol. 2022, pp. 1–9. ISSN: 1574-017X, 1875-905X. DOI: 10.1155/2022/2670562. **available at:** <https://doi.org/10.1155/2022/2670562>.
- Gillani, Nabeel *et al.*, (2023). "Unpacking the "Black Box" of AI in Education", *arXiv (Cornell University)*, DOI: 10.48550/arxiv.2301.01602. **available at:** <https://arxiv.org/abs/2301.01602>.
- Guo, Ying *et al.*, (2022). "Identifying Critical LMS Features for Predicting At-risk Students", *arXiv (Cornell University)*, DOI: 10.48550/arxiv.2204.13700. **available at:** <https://arxiv.org/abs/2204.13700>.
- Maghsudi, Setareh *et al.*, (2021). "Personalized Education in the Artificial Intelligence Era: What to Expect Next", *IEEE Signal Processing Magazine*, Vol. 38, (3), pp. 37–50. ISSN: 1053-5888, 1558-0792. DOI: 10.1109/msp.2021.3055032. **available at:** <https://doi.org/10.1109/msp.2021.3055032>.
- Malasaga, Elisa V. *et al.*, (2023). "Analyzing Machine Learning Algorithm Performance in Predicting Student Academic Performance in Data Structures and Algorithms Based on Lifestyles", *2023 IEEE 15th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, pp. 1–4. DOI: 10.1109/HNICEM60674.2023.10589068.
- Mao, Hui *et al.*, (2025). "What factors enhance students' achievement? A machine learning and interpretable methods approach", *PLoS ONE*, Vol. 20, (5). ISSN: 1932-6203. DOI: 10.1371/journal.pone.0323345. **available at:** <https://doi.org/10.1371/journal.pone.0323345>.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

- Nag, Prashant Kumar, Bhagat, Amit, and Priya, R. Vishnu (2025). *Expanding AI's Role in Health-care Applications: A Systematic Review of Emotional and Cognitive Analysis Techniques*, DOI: 10.1109/access.2025.3562131. **available at:** <https://doi.org/10.1109/access.2025.3562131>.
- Nuankaew, Wongpanya and Thongkam, Jaree (2020). "Improving Student Academic Performance Prediction Models using Feature Selection", *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI- CON)*, pp. 392–395. DOI: 10.1109/ECTI-CON49241.2020.9158286.
- Qadir, Hafiz Muhammad *et al.*, (2025). "Optimizing learning outcomes: a deep dive into hybrid AI models for adaptive educational feedback", *Journal Of Big Data*, Vol. 12, (1). ISSN: 2196-1115. DOI: 10.1186/s40537-025-01187-6. **available at:** <https://doi.org/10.1186/s40537-025-01187-6>.
- Raj, Roop *et al.*, (2024). "Unveiling Patterns Using Enhanced Educational Data Mining for forecasting Student Academic Performance", *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, pp. 1–7. DOI: 10.1109/ACCAI61061.2024.10602439.
- Thakkar, Anoushka, Gupta, Ankita, and Sousa, Avinash De (2024). *Artificial intelligence in positive mental health: a narrative review*, DOI: 10.3389/fdgth.2024.1280235. **available at:** <https://doi.org/10.3389/fdgth.2024.1280235>.
- Ujkani, Berat, Minkovska, Daniela, and Hinov, Nikolay (2024). "Course Success Prediction and Early Identification of At-Risk Students Using Explainable Artificial Intelligence", *Electronics*, Vol. 13, (21), pp. 4157–4157. ISSN: 2079-9292. DOI: 10.3390/electronics13214157. **available at:** <https://doi.org/10.3390/electronics13214157>.
- Uskov, Vladimir L. *et al.*, (2019). "Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education", *2019 IEEE Global Engineering Education Conference (EDUCON)*, pp. 1370–1376. DOI: 10.1109/EDUCON.2019.8725237.
- Wijayaningrum, Vivi Nur *et al.*, (2022). "Prediction of Student Academic Performance in Practicum Courses Based on Activity Logs and Student Background", *2022 International Conference on Electrical and Information Technology (IEIT)*, pp. 366–371. DOI: 10.1109/IEIT56384.2022.9967888.