

إبراهيم النخلي

Final Master Theses (3).docx

- Final Thesis Submission after Reduce Similarity (Word Format)
- Thesis_(2)_AI901_Dr_Amjad_Qatish_Third Trimester 2025 (Moodle PP)
- Midocean University

Document Details

Submission ID

trn:oid::1:3451869939

Submission Date

Dec 31, 2025, 7:25 PM GMT+4

Download Date

Jan 1, 2026, 2:15 AM GMT+4

File Name

xcod.832537143_011621__3__seseht_retsaM_laniF_إبراهيم_النخلي_57792

File Size

724.0 KB

37 Pages

6,170 Words

37,616 Characters

47% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups



43 AI-generated only 47%

Likely AI-generated text from a large-language model.



0 AI-generated text that was AI-paraphrased 0%

Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.





Analyzing and Improving Student Academic Performance using AI

Prepared By

Name: Ahmed Almoairfi

ID No.:251002473

Name: Ibrahim Alnakhli

ID No.:251002688

Submitted in fulfillment of the requirements for the degree Master of
Science in Artificial Intelligence

At

Collage of Informatics

Midocean University

UAE

Supervisor Name: Dr.Amjad Qtaish

Collage of Informatics, Midocean University, Comoros

[2025-03]

Acknowledgements

Special thanks to Dr.Amjad Qtaish

Abstract

AI is the new evolution in this generation. It has led to improvement in all fields. It has to be used to improve education and academic outcomes. Instead of simply allowing students to use AI generative models, users should have access to a comprehensive AI system that begins by creating a measurement process, monitors student performance, enhances the results, and suggests solutions. Data from studies were analyzed to evaluate students' work using AI and machine learning techniques. One investigation examined the work of 80 students enrolled in a genetics course, focusing on how their interpretations changed before and after receiving feedback from different sources, including AI systems, human instructors, or a combination of both. Students' interpretations were evaluated using a mixed-methods approach. Quantitative analysis relied on paired t-tests and descriptive statistics, while qualitative analysis involved manual coding to assess accuracy, depth, and conceptual understanding. In addition, a chatbot built on a retrieval-augmented generation (RAG) model within the CustomGPT.ai platform was employed to review students' answers to conceptual biology questions and deliver immediate feedback aimed at identifying conceptual misunderstandings. Other studies drew on datasets from both school and university students to develop machine learning models such as random forests, decision trees, support vector machines, k-nearest neighbors, Naive Bayes, and logistic regression. These datasets incorporated a range of academic and social factors, including attendance, study habits, health, travel time, and family background. Model performance was assessed using standard evaluation measures, including precision, recall, and the F1 score. In one case, an artificial neural network (ANN) trained with the stochastic gradient descent (SGD) algorithm and multiple input features showed performance comparable to that of traditional models. Another evaluation reported that a random forest model used to predict outcomes in Python programming courses achieved an R^2 value of 0.55, reflecting a moderate level of predictive accuracy.

List of Publications

Table of Contents

List of Abbreviations	vi
List of Tables	vii
List of Figures.....	vii
Chapter 1.....	1
Introduction.....	1
1.1 Problem Definition	1
1.2 Thesis Objectives.....	1
1.3 Thesis Contribution	2
1.4 Thesis Organization.....	3
Chapter 2.....	4
Overview	4
2.1 Introduction to AI-Driven Student Performance Prediction	4
2.2 Classical Machine Learning Models	4
2.3 Deep Learning Approaches	4
2.4 Hybrid, Optimized, and Explainable AI Models.....	5
2.5 Use of Real-World Educational Datasets.....	5
2.6 Quantitative Results Summary	5
2.7 Research Gaps and Opportunities	6
2.8 Conclusion.....	6
Chapter 3.....	7
Related Work.....	7
3.1 Introduction	7
3.2 History	7
3.3 Summarization of related work	10
Chapter 4.....	17

Proposed Approach.....	17
4.1 Data collection:.....	17
4.2 Data Processing:	18
4.3 Feature engineering and feature selection	18
4.4 Model selection and optimization:	19
4.5 Model Evaluation:	19
4.6 Visualization and Insights	19
Chapter 5.....	20
Experimental Results and Discussion	20
5.1 Experimental Setup	20
5.2 Baseline Model Performance	20
5.3 Feature Selection and Model Refinement	22
5.4 Visualization of Delay Factors	23
5.5 Discussion.....	25
Chapter 6.....	27
Conclusion and Future Work	27
6.1 Conclusion	27
6.2 Future Work.....	27
References	29

List of Abbreviations

- ANN: Artificial Neural Network
- API: Application Programming Interface
- AI: Artificial Intelligence
- CNN: Convolutional Neural Network
- DT: Decision Tree
- EDM: educational data mining
- GBR: Gradient Boosting Regressor
- GBM: Gradient Boosting Machine
- GridSearchCV: Grid Search with Cross-Validation
- GPA: Grade Point Average
- KNN: K-Nearest Neighbors Regressor
- LLMs: Large Language model-based
- LMS: Learning Management System
- LR: Logistic Regression
- LSTM: Long Short-Term Memory
- ITS: Intelligent Tutoring Systems
- MAE: Mean Absolute Error
- ML: Machine Learning
- NB: Naive Bayes
- PR: Precision-Recall
- RF: Random Forest
- R^2 : Coefficient of Determination
- RAG: Retrieval Augmented Generation
- RMSE: Root Mean Square Error
- RNN: Recurrent Neural Network
- SGD: Stochastic Gradient Descent
- SHAP: SHapley Additive exPlanations
- SVM: Support Vector Machine
- XAI: Explainable Artificial Intelligence
- XGBoost: Extreme Gradient Boosting

List of Tables

Table 1 problem definition and Thesis objective summary.....	1
Table 2 Quantitative Results Summary	6
Table 3 Summarization of related work.....	10
Table 4 Models hyperparameters.....	21
Table 5 ML models result.....	21
Table 6 ML models important features.....	22

List of Figures

Figure 1 Thesis organization graph.....	3
Figure 2 Proposed Approach flowchart	17
Figure 3 This visual illustrates the basis for discussing key predictive features	23
Figure 4 This visual explains the relationship between GPA and StudyTimeWeekly.	24
Figure 5 This visual explains the relationship between GPA and Absences.....	25

Chapter 1

Introduction

1.1 Problem Definition

Evolution in education has both pros and cons. Students encounter many challenges that mainly impact their ability to benefit fully from school. These include academic, social, economic, and study habit issues. This results in lower academic performance. Students must identify and address these issues early.

1.2 Thesis Objectives

This thesis aims to design and evaluate a machine-learning-based framework for predicting students' performance and identifying the key factors that affect it.

The specific objectives are:

Explore and analyze a real-world student performance dataset to identify the most strongly correlated factors with student performance. Build and compare ML models and choose the best model for predicting student performance. Apply feature selection to improve the model accuracy. Evaluate the model performance using R^2 , MSE, MAE, RMSE, and Loss for RNN.

To visualize key trends and insights, particularly the role of Student Grade or student performance indicator, and the two main correlated factors.

Table 1 problem definition and Thesis objective summary

Problem Statement	Research Questions	Research Objective
Students need guidance to cross over the changes, identify their weaknesses, and resolve them.	How to analyze student performance using AI?	Create an ML model that analyzes and predicts students' performance
	How does artificial intelligence provide suggestions to detect student weakness?	Detect the student's weakness.

1.3 Thesis Contribution

Development of a Comprehensive Machine-Learning Framework for Student Performance Prediction

This thesis develops a full analytical framework that uses multiple machine-learning models—including Random Forest, Decision Tree, K-Nearest Neighbors, Gradient Boosting Regressor, XGBoost, and Recurrent Neural Networks—to predict student performance with high accuracy. A rigorous comparison across models demonstrates that XGBoost achieves the highest performance, reaching an R^2 score of 0.9614.

Identification of Key Factors Influencing Academic Performance

Select features that most affect student results, to allow the student to enhance their academic performance.

Use of Multiple Evaluation Metrics for Robust Comparison

The study employs a comprehensive set of evaluation metrics (MAE, MSE, RMSE, R^2 , and Loss for RNN) to ensure robust model assessment. The evaluation confirms the superiority of tree-based and ensemble models over distance-based models such as KNN.

Integration of Visual Analytics for Educational Insight

Correlation heatmaps and scatter plots are used to visualize the relationships between GPA and key predictors. These visualizations provide interpretable insights that support policymakers, instructors, and students in understanding learning behaviors and performance drivers.

Contribution to Student-Centered Educational AI

By analyzing weaknesses in study habits and personal factors, this thesis provides a foundation for future systems capable of delivering personalized academic guidance, early-warning alerts, and data-driven recommendations to support student improvement.

1.4 Thesis Organization

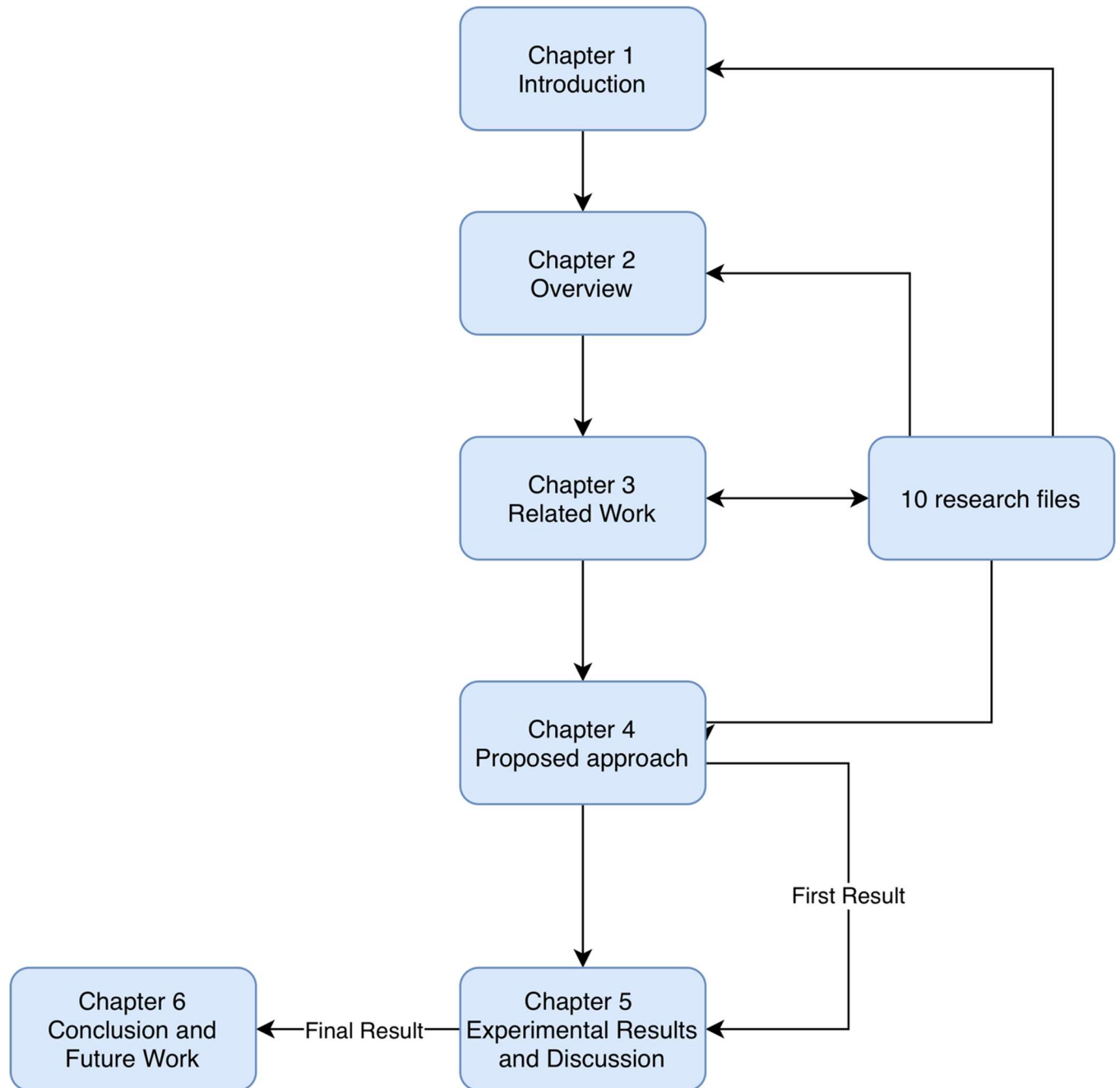


Figure 1 Thesis organization graph

Chapter 2

Overview

2.1 Introduction to AI-Driven Student Performance Prediction

AI is the new evolution in this generation. It has led to improvement in all fields. It must be used to improve education and academic outcomes. Instead of just letting students use AI generative models, users should have a complete AI system that starts by creating a measurement process, monitors student performance, enhances the result, and suggests solutions.

2.2 Classical Machine Learning Models

Traditional machine learning techniques have already been used in research to analyze educational data.

Yağcı (2022) applied Random Forest, Support Vector Machines, Logistic Regression, Naïve Bayes, and K-Nearest Neighbors to a dataset of 1,854 university students, achieving an accuracy of 70–75%. Midterm grades and attendance were identified as the most important indicators.

Also, Kalpana et al. (2020) used Decision Trees, Logistic Regression, and Random Forest to identify at-risk students using academic, behavioral, and demographic factors. The results confirmed the importance of study time, absences, and past academic performance as key factors.

Overall, studies using the tree model are superior to the linear model.

2.3 Deep Learning Approaches

The deep learning techniques employed have demonstrated strong predictive capabilities. Studies implementing neural network models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) have shown that hybrid structures like CNN–LSTM outperformed individual models by combining

spatial pattern extraction and chronological modeling, leading to improved early detection of academic risks.

These techniques illustrate the ability of deep learning to model complex learning patterns to a degree that surpasses traditional machine learning (ML).

2.4 Hybrid, Optimized, and Explainable AI Models

Recent AI research emphasizes hybrid models and structural optimizations, seeking improvements in both predictive accuracy and our ability to understand how models work. On the optimization front, techniques from genetic algorithms to Mayfly optimization are helping to stabilize networks and curb training errors. Parallel to this, Explainable AI (XAI) is rightfully gaining traction. In practice, researchers are applying interpretability tools—SHAP and Sobol methods, for example—to clarify the role of various academic and behavioral features. Another interesting direction is the use of large language models (LLMs) to power feedback systems that can offer surprisingly human-like, interpretative comments for formative assessment. The combined result of these approaches is AI that is not only more accurate but also more transparent and educationally valuable.

2.5 Use of Real-World Educational Datasets

The literature uses a variety of data sources, including:

Academic records from universities.

Data from schools.

Data from education systems.

Student surveys.

Classroom interaction.

Behavior and habits outside of school

2.6 Quantitative Results Summary

Table 2 Quantitative Results Summary

Study	Model	Key Findings
Yağcı (2022)	RF, SVM, LR, NB, K-NN	70–75% accuracy; midterm exam scores most influential
Kalpana et al. (2020)	RF, DT, LR	Effective identification of at-risk students using academic + behavioral variables
Misconception-analysis studies	AI Diagnostic Models	Identified conceptual errors; improved formative assessment
Learning analytics studies	ML + Analytics	Higher accuracy when engagement data is included
Li & Botelho (2024)	LLM-XAI	Human-aligned interpretable feedback
Pineda-Jaramillo et al. (2024)	SHAP / Sobol	Clear ranking of feature importance

2.7 Research Gaps and Opportunities

Despite notable progress, several research gaps remain: Most models rely on data from a single institution, limiting their generalizability.

There is a need for student-centered AI to improve personalization.

The limited number of models that utilize real-time data reduces the ability to intervene early.

There is a lack of research analyzing the quality of student interaction with AI systems, particularly. language learning models (LLMs).

2.8 Conclusion

The literature demonstrates the effectiveness of ML, DL, hybrid models, and XAI in predicting and understanding academic performance.

The results show that accuracy improves when behavioral data are combined with academic variables and when interpretability is incorporated into the modeling process.

Chapter 3

Related Work

3.1 Introduction

There has been great interest in educational outcomes since ancient times due to their impact on academic and practical paths. This has led researchers to seek the best ways to help students achieve the highest possible results, from initial learning through memorization and then writing, to the use of artificial intelligence. We will learn about the extent and developments the field has achieved through research efforts to improve student outcomes and predict student problems.

3.2 History

Artificial Intelligence analyzes student performance by defining their education needs and enhancing academic results. Also, AI measures student attitudes from learning, encourages their motivation, and uses adaptive learning techniques to guide them. This results in improving and enhancing learning performance. Data collection: • Mixed-methods approach with focus group discussions used. • Random sampling of 100 respondents based on criteria. (Mallillin, 2024).

The student analyzes student performance using a random forest algorithm to predict the learning outcomes from learning data from 395 students in a Python programming course. The model achieves great results. R^2 of 0.55, MAE of 3.61 points, and RMSE of 4.68 points, indicating accurate predictions. Early intervention and curriculum improvements can be achieved by including the scores of the 2nd unit test, the 4th unit test, and average homework scores in Key predictors influencing performance. Data collection: • The study collected learning data from 395 students. • These students were enrolled in a Python Programming course. • Chongqing University of Posts and Telecommunications offered the course. Machine Learning Model: The study developed a predictive model based on the random forest algorithm. Machine Learning Evaluation: • The model's performance

in predicting student performance was evaluated using several metrics. • The evaluation results indicated the following: R-squared (R^2) value: 0.55 Root Mean Squared Error (RMSE): 4.68 points Mean Absolute Error (MAE): 3.61 points. In summary, the research utilized a random forest model and assessed its accuracy using R^2 , RMSE, and MAE to determine its effectiveness in predicting student performance. (Zhu, 2024).

One promising way AI can help teachers is by digging into students' weaknesses through their written explanations in formative assessments, like multiple-choice questions with follow-ups. Tools like the Dewey chatbot, for instance, give instructors quick insights into how students are thinking—spotting common misunderstandings or outright misconceptions in real time.

This kind of analysis uncovers real gaps in knowledge, things like mistaken ideas about gene dosage or how transcription factors actually work. Once those are clear, teachers can adjust their lessons more precisely, focusing on the exact spots where kids are getting stuck, and ultimately boost overall learning.

Data collection:

Students explained not just why they picked an answer, but also why they thought the wrong options were incorrect.

An AI bot trained specifically in the discipline reviewed those explanations, flagging missing concepts or misapplied principles.

The questions came from the Biology Concepts Instrument (BCI) and were tested in a small-scale proof-of-concept study.

Machine Learning Model:

In this work, researchers used a custom AI chatbot called Dewey, built on a Retrieval-Augmented Generation (RAG) framework via the CustomGPT.ai platform. It delivers fast, practical feedback on student reasoning during formative assessments, helping overcome some of the shortcomings of plain multiple-choice testing.

Machine Learning Evaluation:

The CustomGPT.ai system—and Dewey in particular—was evaluated in a proof-of-concept setup. The focus was on how well it could break down student explanations and offer useful feedback, showing its potential to make formative assessment far

more insightful than standard multiple-choice approaches alone (Klymkowsky & Cooper, 2024).

AI analyzes student performance by leveraging machine learning algorithms, natural language processing, and data analytics to identify patterns and trends that traditional assessment methods may overlook. It provides real-time feedback and personalized learning experiences, enabling early identification of struggling students for timely interventions. AI tools, such as adaptive learning platforms and AI-driven tutoring systems, collect and analyze performance data, enhancing educational outcomes and reducing the administrative burden on teachers, allowing them to focus more on instruction and student engagement. Data collection: • AI tools collect and analyze student performance data. • Data analytics provides personalized learning experiences. Machine Learning Model: Leverages algorithms to analyze student performance data. Machine Learning Evaluation: • Identifies patterns in student performance data. • Enables early intervention for struggling students. (Rajeena & Quraishi, 2024).

AI's real value in formative assessment lies in its ability to dissect student reasoning, particularly within multiple-choice frameworks. Platforms like Dewey go beyond simple grading; they pinpoint specific conceptual hurdles—such as confusion surrounding genetic expression or the mechanics of transcription factors—giving educators the immediate insights needed to pivot their teaching strategies. For this study, we utilized a pre-filtered, balanced dataset from the Learning Cup 2006, employing a cluster-linked sampling method. By integrating variables like travel time, family size, and health status, we trained Random Forest, Decision Tree, and Logistic Regression models to map the multifaceted nature of student performance. The evaluation methods used were based on machine learning and multivariate analysis strategies and were evaluated using performance tools, excluding failures due to high prevalence, G1/G2 quality charts, and gender-based education statistics. No precise details were provided regarding numerical evaluation metrics or data split ratios. (Kalpana et al., 2020).

In this study, we present an understandable AI system based on a large-scale language model (LLM) aimed at improving feedback in education. The system classifies the type of feedback and compares human-generated responses with AI-generated responses. The model improves the relevance of educational feedback by providing structured, clear, and personalized guidance to students. However, it is still a conceptual approach and further implementation in real classrooms is needed

to validate its effectiveness. The data was obtained (scraped) from the UCI Machine Learning Repository and covers two Portuguese secondary schools and includes a wide range of demographic, social, and related information such as grades, absences, study time, etc. Algorithms used: Decision Tree, Random Forest, Support Vector Machine / SVM, K-Nearest Neighbors / KNN, Naive Bayes, Logistic Regression. We split the dataset into training and test sets, and used cross-validation for reliability. The evaluation metrics we considered are accuracy, recall, precision, and F1 score. (Yağcı, 2022).

This systematic review of 22 studies examines the application of artificial intelligence in educational assessment. It highlights that AI is primarily used for formative and automated assessment. Despite its usefulness, the study suggests that current AI tools exhibit a notable lack of pedagogical logic, suggesting the need for deeper integration with principles of instructional design. (Li & Botelho, 2024).

A study combines AI-based predictions with real-world classroom learning analytics. Using a similar experimental design, it shows that incorporating this design increases student engagement, collaboration, and satisfaction in the classroom. This study provides practical evidence of how AI can be integrated into everyday learning and a strong example of the use of AI in the education sector. (González-Calatayud et al., 2021).

3.3 Summarization of related work

Table 3 Summarization of related work

Study	Key Findings/Results	Methodology	Limitations	Advantages

Artificial Intelligence (AI) Towards Students' Academic Performance, Mallillin,2024	AI enhance student. learning and measures student. Attitude	Quantitative research use to Measure and detect the related factors to enhance academic performance. Qualitative research is used to evaluate and analyze the AI Enhance the academic performance of students among the respondents.	The study uses a mixed-methods approach (focus groups + survey), but lacks a control or comparison group to assess the effectiveness of AI interventions versus no intervention	Targets specific learning needs for improved outcomes. And Enhances motivation and attitudes toward learning.
Prediction and Analysis of University Student Performance Based on Machine Learning, Xiaobo Zhu, 2024	The model provide a good result for student performance $R^2 = 0.55$, RMSE = 4.68 and MAE=3.61. analysis predict the learning difficulties and curriculum enhancement.	Predictive model based on random forest that analyze student performance specifically on python course.	Limited amount of data used in the research. Not all indicators have been used in this study such as classroom interaction frequency and homework completion time	Guidance for Early Warning The study provides valuable guidance for the early identification of students who may be experiencing learning difficulties. Curriculum Enhancement It offers insights that can be used for the continuous improvement and enhancement of educational curricula. Reference for Teaching Quality The findings serve as a useful reference for improving the overall quality of teaching. In summary, the study's advantages lie in its practical applications for identifying at-risk students, refining educational programs, and elevating teaching standards.

The end of multiple choice tests: using AI to enhance assessment, Klymkowsky, 2024	Dewey delivered a detailed analysis for students and define the miss understand concept by allowing student to explain why they choose this answer. Then the AI bot “Dewey analyze the student answers and identify misconceptions and missing knowledge	Develop a AI Bot	The current answer is unsure student understanding. Training Data: Dewey’s accuracy depends on its curated knowledge base. Adoption Barriers: Instructors must embrace formative assessment and redesign courses around feedback. Expansion Needed: Larger trials across diverse subjects and student populations.	The use of AI to analyze student explanations provides rapid and informative feedback on their understanding, allowing instructors to identify specific misconceptions and gaps in knowledge, particularly in complex topics like genetic principles and molecular processes. Incorporating AI-analyzed formative assessments enhances the depth of understanding by guiding instructors on how to address misunderstandings, ultimately leading to improved learning outcomes and a more accurate assessment of student comprehension compared to traditional multiple choice tests.
Data-Driven Insights: Using AI to Analyze Student Performance in English Language Courses, Adlin Jerusha, 2024	The study result in enhancing the performance by analyzing the the factors such as motivation and resource availability using AI. This indicates that understanding these interconnections can lead to enhance educational strategies.	The study use SPSS for quantitative analysis to study the increase in student performance in English course as a second language.by using both descriptive and inferential statistics to analyze survey data that collected from student according to their study habits, demographics and the others factors.	The study focus on English courses student only, so the generalization in this research is limited	AI enhances personalized course development for students. Identifies key factors influencing student success and performance.

Leveraging Artificial Intelligence for Student Performance Monitoring. <i>International Journal of Research Publication and Reviews</i>, Rajeena,M., & Quraishi, 2024	The study focus on detecting the positive effect of AI on students outcomes and how can AI provide a real time feedback. The research found that how can AI access just handle simple management task to analyze student performance and enhance student outcome.	The study review existing researches in AI tools used in effectiveness in monitoring and evaluating student performance. Which include examining adaptive learning platforms, AI-driven tutoring systems. That will lead in detecting struggling students which will lead to enhance education outcome.	The study discuss the ethical issues and challenges related in Using AI. including privacy issues which is counted as main limitation in using AI in analyzing student performance.	Provides real-time feedback and personalized learning experiences. Enables early intervention for struggling students.
Predicting Student Academic Performance Using Neural Networks: Analyzing the Impact of Transfer Functions, Momentum and Learning Rate, Agarwal, M., & Agarwal, 2024	The study get great result in prediction student performance. The prediction accuracy is 98% using artificial neural network. As a comparative analysis of various Machine Learning Algorithms the result the ANN achieve Better result than SVM, KNN, NB, DT and others ML algorithms indicate that ANN has the potential to improve student performance.	The study utilizes Artificial Neural Networks (ANN) as the primary method of analyzing student performance. The study used SVM DT, KNN,LR,NB as comparative research to show the superiority of ANN with a 97% accuracy rate.	The dataset of students is not large which is less than 700 student. That can not lead to generalize the result. The research omit some factors like students' social interactions, interpersonal skills and academic engagement which might indicate a gap in the understanding all variables that affect student outline	Why is it important? This research used artificial intelligence specifically neural networks to predict student grades with very high accuracy (98%). - Can identify struggling students early, before failure occurs. - Tested various configurations like learning rate and momentum to optimize performance. - Compared many algorithms like SVM and Decision Trees, and neural networks came out on top. - Used real data from 689 students, giving credibility to the findings.

Student Performance Analysis using Machine Learning. International Journal of Innovative Technology and Exploring Engineering, Kalpana et al., 2020	Random Forest model to predict student outcomes using assessment data	Machine Learning (Random Forest)	No live classroom application; lacks qualitative evaluation	Why is it useful? The goal is to help teachers detect students who need support early, based on their grades and classroom behavior. - Provides early warning system for at-risk students. - Combines insights from psychology, education, and computer science. - Improves communication between teacher, student, and administration. - Includes user-friendly interface and auto-generated reports (PDFs and graphs). - Helps institutions understand student levels and improve academic programs
--	--	---	--	---

Educational data mining: Prediction of students' academic performance using machine learning algorithms, Yağcı, 2022	Midterm scores predict final grades with 70–75% accuracy	ML (SVM, RF, NB, KNN)	Simple features; lacks full behavioral data	What does it do? Builds a model to predict final exam grades based on midterm scores and department/faculty data. - Used real data from over 1800 students, ensuring robust and realistic outcomes. - Tested multiple algorithms (Random Forest, Nave Bayes, etc.) and identified the most accurate. - Helps detect at-risk students early through predictive insights. - Supports smart learning environments driven by data analytics.
Bridging Human and AI Explanations to Support Students' Revisions of Open-Ended Responses. Proceedings of the 17th International Conference on Educational Data Mining (EDM 2024),	Develops explainable feedback using LLMs; taxonomy of feedback types	LLMs + Text Analysis	Still theoretical; requires real classroom deployment	Intelligent Feedback Systems (Li & Botelho, 2024) Development of explainable systems that provide human-like, precise feedback using large language models like GPT Classification of feedback types according to specific educational objectives Analysis of student-AI interactions to improve educational tools Providing individualized feedback in large classrooms, enhancing educational equity

AI for Student Assessment - Systematic Review, González-Calatayud et al., 2021	AI mostly used in formative assessment and auto-grading	Review of 22 studies	Lack of educational/pedagogical grounding	Why is it important? This isn't an experimental study but a review of 22 previous studies on AI-based student assessment. - Highlights the use of AI in formative assessment and instant feedback. - Confirms that AI tools can automate grading and save teacher's time. - Discusses differences between AI-enabled and traditional education. - Stresses the need to train teachers on using AI pedagogically, not just technically. - Offers useful recommendations for schools and universities aiming to adopt AI in assessment.
--	---	----------------------	---	---

Chapter 4

Proposed Approach

This chapter focuses on developing an AI model that analyzes student performance using a predefined dataset, predicts outcomes based on selected features, and identifies which features have a greater impact on outcomes than others. To achieve this, multiple models will be used, and their results will be compared to determine the model with the highest accuracy.

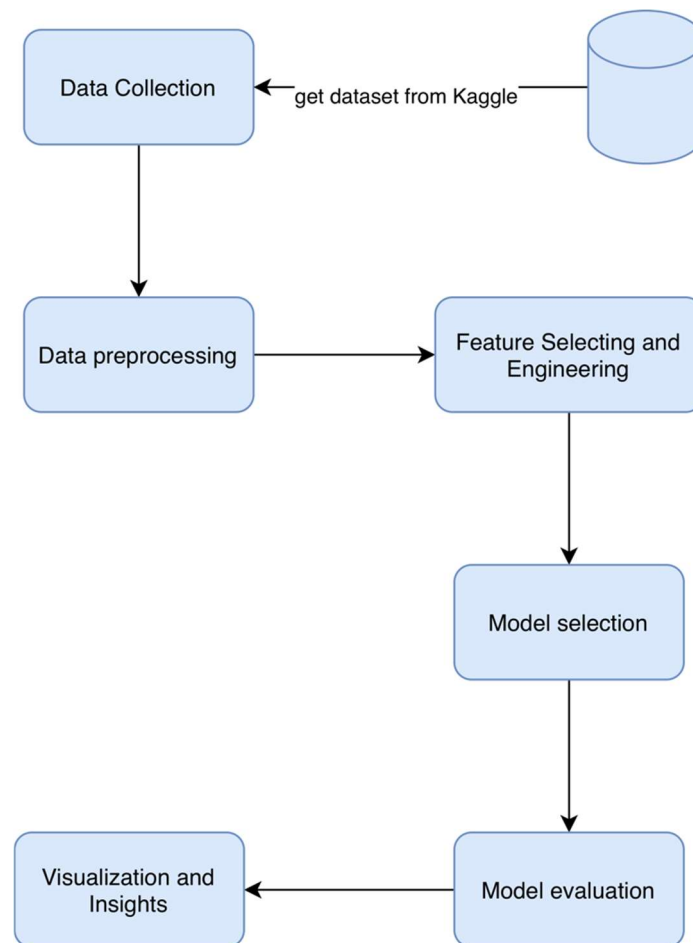


Figure 2 Proposed Approach flowchart

4.1 Data collection:

The data source is a dataset containing records for 2392 students, with GPA and 13 additional attributes, excluding student ID. The attributes are arranged as follows:

Demographic Details, which are: Age, Gender, Ethnicity, ParentalEducation.

Study Habit: (StudyTimeWeekly, Absences, Tutoring)

Parental Involvement (ParentalSupport)

Extracurricular Activities (Extracurricular, Sports, Music, Volunteering), and finally the target attribute (GPA and GradeClass). One of these targets will be ignored as they reflect each other.

The dataset is a public resource on Kaggle “Predicting Students Grades”.
<https://www.kaggle.com/code/annastasy/predicting-students-grades/notebook>

4.2 Data Processing:

The predefined dataset doesn't require an emissive process. Because it's clear of null values, missing attributes, and duplicates, all values are “numerical”; even classification, nominal, and ordinal values are converted already, and the keys of each value are provided.

The selected target value is GPA, so the GradeClass attribute is deleted, and StudentID is removed because it may negatively affect the prediction results.

The data was split into two groups: training data (80 % of the dataset) and a testing set (20 %).

4.3 Feature engineering and feature selection

Based on the correlation head map, the features that affect GPA results are:

Absences, StudyTimeweekly, Tutoring, ParentalSupport, ExtraCurricular, and music

Thus, the model mainly depends on the selected feature for training.

4.4 Model selection and optimization:

In this research, we use 5 models, compare their results, and select the most accurate. The chosen models are (Random Forest Regressor, Decision Tree Regressor, K-Nearest Neighbors Regressor, XGBoost, Gradient Boosting Regressor, and RNN)

The results weren't far from each other, except for the KNN.

The most accurate were XGBoost and KNN.

All models are regression models that predict the expected GPA value from 0 to 4.

By using the GridSearchCV tool in scikit-learn, which uses Cross-Validation to find the best.

4.5 Model Evaluation:

The neural network model was evaluated twice: once with PCA and once without.

The evaluation results are R^2 and the Loss value.

Other models' evaluation result is R^2 , MSE, RMSE, MAE

The best result across all models is achieved by: XGBoost.

4.6 Visualization and Insights

Some visualizations are created to describe the model's behavior and the correlations among features.

Correlation Heatmaps:

It shows the relation between features. We can examine the relationship between GPA and other features and identify which features are more closely associated with GPA.

Scatter Plots:

It shows the relation between GPA and the most important features, each one in a separate plot.

Chapter 5

Experimental Results and Discussion

5.1 Experimental Setup

The Experiment is implemented in Python using the Python AI and data analysis libraries, including scikit-learn, seaborn, pandas, and PyTorch.

The dataset is publicly available on Kaggle. This contains 2392 records. And 13 features after eliminating unused features. The dataset is already preprocessed and doesn't require additional encoding or null value removal.

5.2 Baseline Model Performance

GridSearchCV was used with 5 models using all features

The selected models are (DT – RF – KNN – GBR – XGBoost)

The following parameter was fixed on all models:

fold cross-validation (CV): 5

Metric to optimize(scoring): neg_root_mean_squared_error

The input for each model in the table below:

Table 4 Models hyperparameters

DT Regressor	RF Regressor	GBR	KNN	XGBoost	RNN
max_depth: [None, 10, 20, 30]	n_estimators: [100, 200, 300],	n_estimators: [50, 100, 200],	n_neighbors: [3, 5, 7, 10],	n_estimators: [100, 200, 300],	Epocs: 5000
min_samples_split: [290, 295, 150]	max_depth: [None, 10, 20, 30],	learning_rate: [0.01, 0.05, 0.1, 0.2],	weights: ['uniform', 'distance'],	learning_rate: [0.01, 0.1, 0.2],	Learning rate: 0.002
min_samples_leaf: [1, 2, 4],	min_samples_split: [2, 5, 10],	max_depth: [3, 5, 10, 20, 30],	metric: ['euclidean','manhattan', 'minkowski'],	max_depth: [3, 6, 10],	batch_size: 128
	min_samples_leaf: [1, 2, 4],	min_samples_split: [2, 5, 10, 50, 100],		subsample: [0.8,1.0],	
		min_samples_leaf: [1, 2, 5, 10],		Colsample_bytree :[0.8,1.0],	

The resulting performance metrics on the test set were:

Table 5 ML models result

	DT Regressor	RF Regressor	GBR	KNN	XGBoost	RNN
MAE	0.2246	0.1668	0.1409	0.3464	0.1383	Loss: 0.0391
MSE	0.0956	0.0491	0.0339	0.1941	0.0325	
RMSE	0.3092	0.2217	0.1841	0.4406	0.1804	
R ² Score	0.8865	0.9416	0.9597	0.7694	0.9614	0.9537

The above result shows that KNN has the lowest accuracy among the models. Other models are close to each other. XGBoost is the best one; GBR and RNN are very close to it.

5.3 Feature Selection and Model Refinement

RNN value after PCA

The result with PCA:

When the PCA is 8, the R^2 is:

R^2 is 0.6553.

The result of using PCA led to a deterioration in the accuracy.

The result of the RNN after selecting the top 4 features that are common in most of the models: $R^2 = 0.91$

Based on the above result, three models are close to each other (Random Forest, RNN, XGBoost, GBR).

The most affected features in the model:

Table 6 ML models important features

Random Forest Regressor	Gradient Boosting Regressor	Xgboost	RNN
Tutoring	Tutoring	Tutoring	Absences
Absences	Absences	ParentalSupport	StudyTimeWeekly
Age	Extracurricular	Extracurricular	ParentalSupport
Extracurricular	Age	Music	Tutoring
ParentalSupport	ParentalSupport	Age	Extracurricular

The common ones among them are:

(Tutoring – ParentalSupport - ExtraCurricular)

After those features, “Absences” comes next.

The result after filtering by the main features selected is:

RNN: $R^2=0.91$

XGBoost: $R^2= 0.9107$, MAE=0.2173, MSE=0.0752, RMSE= 0.2742

GBR: $R^2= 0.9095$, MAE=0.2188, MSE=0.0762, RMSE=0.2761

Selecting only four features doesn't substantially reduce accuracy. If training performance is more critical than this accuracy percentage, as in training on a large dataset, the model will perform better by using only a subset of features.

5.4 Visualization of Delay Factors

Correlation Heatmap for the dataset:

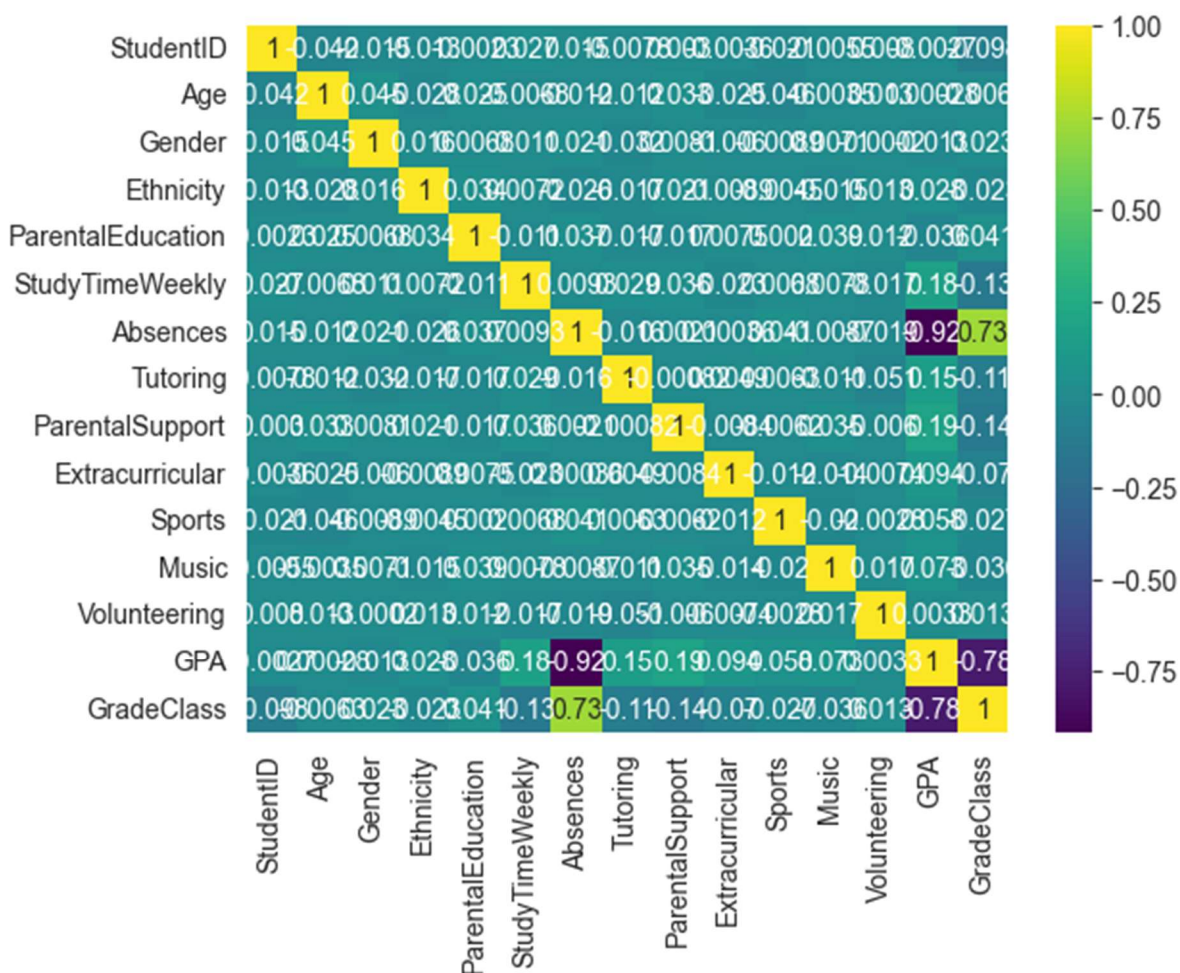


Figure 3 This visual illustrates the basis for discussing key predictive features

The Plot shows that there is a strong relationship between GPA and the following features: Absences, StudyTimeWeekly, Tutoring, ParentalSupport, ExtraCurricular, and music.

Scatter Plot:

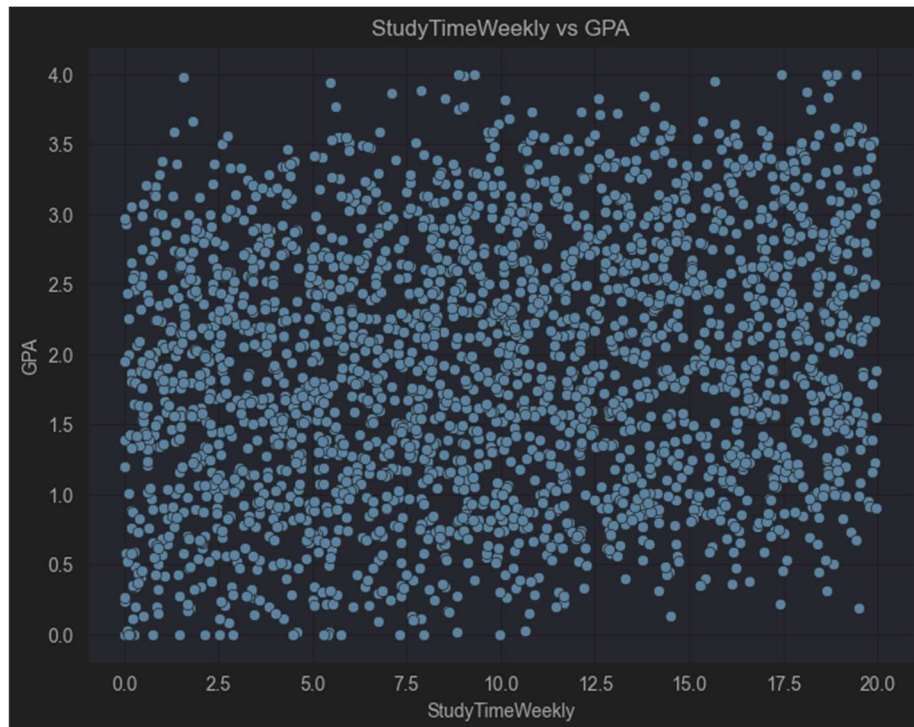


Figure 4 This visual explains the relationship between GPA and StudyTimeWeekly.

The Plot shows the relationship between GPA and StudyTimeWeekly. There is no strong relation between the two columns, and the plot shows the points distributed randomly.

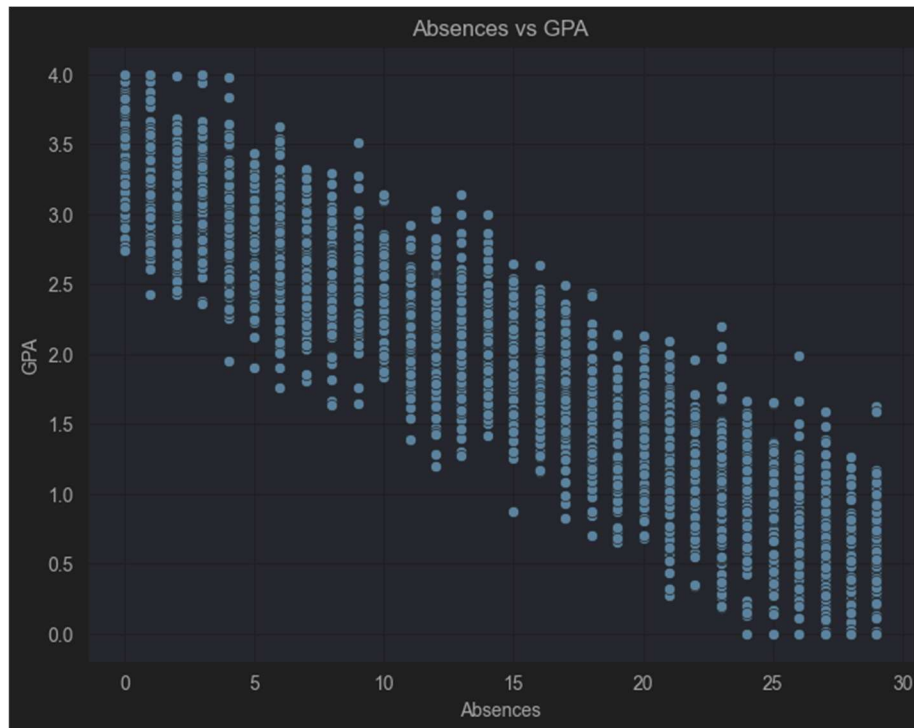


Figure 5 This visual explains the relationship between GPA and Absences

There is a strong negative relation between Absences and GPA; as when the number of absences increases, the GPA will decrease and vice versa.

5.5 Discussion

The experimental results support the conclusion that study habits play an essential role in student performance. The study examines features and their relationship to GPA as an indicator of student performance. We found that “Tutoral” is the most effective feature, and that “study habits” features, in general, are the most effective for GPA.

The results were obtained from different AI models and show only a limited difference, as shown above.

As a result, XGBoost was the most accurate model, with an R^2 of 0.9614. GBR and RNN results are close to XGBoost. In contrast, the KNN accuracy result is very weak compared to the previous models.

The feature selection and filtering results were less accurate than those from the complete feature model. It is only useful when training a model with a massive amount of data.

The visualization shown in the previous section will support the analysis and help yield better results, describing the relationship between features and the target “GPA”.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

To guide students in their academic path, Machine learning models can analyze student performance and provide suggestions. In previous studies, a different Machine learning model has been used. For example, the Random Forest Algorithm, which is used to analyze student performance as in (Xiaobo Zhu 2024), is used to analyze student performance in a Python programming course. An AI chatbot built on a RAG has been used by (Klymkowsky 2024) to provide actionable feedback on student thinking in the context of formative assessments. ANN has been used by (Agarwal, M., & Agarwal, B. B., 2024) and compared with SVM, Decision Trees, and other Machine learning models, and provides a better result. The limitation of the findings was mostly about the amount of data, which needs to be increased.

Key findings from the study include:

In this research, the amount of data was 2392 records, which can be increased further as an enhancement.

The used model in this research includes (Random Forest, Decision Tree, K Nearest Neighbors, XGBoost, GradientBoostingRegressor, and Recurrent Neural Network).

The best model was XGBoost with 0.9614 accuracy.

The Study habits appear as the most affecting factor on student performance.

Also, the visualizations show the relation between student GPA and other features, especially in the heatmap.

6.2 Future Work

Increase the training data:

The study could be improved by incorporating a larger dataset of students from different countries to enable generalization of the results. Additionally, the features used in the prediction can be expanded to capture other factors that affect academic performance. Also, ensure that you include samples from different study levels and social classes.

Provide and suggest Guides to enhance the detected weak points:

The next step is to provide a roadmap for each student to address the weaknesses identified that may affect their GPA and overall academic performance.

Include behavior indicators:

In future studies, students' psychological status should be incorporated into training to improve outcomes.

Develop an adaptive learning system:

Implement an adaptive personalized learning system. Start at the student's level, increase difficulty gradually, and select the most appropriate tool for the student, especially when the student has a disability.

References

- Mallillin, L. L. D. (2024). Artificial Intelligence (AI) Towards Students' Academic Performance. *Innovare Journal of Education*, 16–21. <https://doi.org/10.22159/ijoe.2024v12i4.51665>
- Xiaobo Zhu. (2024). Prediction and Analysis of University Student Performance Based on Machine Learning. *Academic Journal of Computing & Information Science*, 7(8). <https://doi.org/10.25236/ajcis.2024.070811>
- Klymkowsky, M. W., & Cooper, M. M. (2024). The end of multiple choice tests: using AI to enhance assessment. <https://doi.org/10.48550/arxiv.2406.07481>
- Jerusha J, A., & Rajakumari, R. (2024). *Data-Driven Insights: Using AI to Analyze Student Performance in English Language Courses*. 533–539. <https://doi.org/10.1109/icdici62993.2024.10810827>
- Rajeena, M., & Quraishi, A. H. (2024). Leveraging Artificial Intelligence for Student Performance Monitoring. *International Journal of Research Publication and Reviews*, 5(5), 9642–9645. <https://doi.org/10.55248/gengpi.5.0524.1364>
- Agarwal, M., & Agarwal, B. B. (2024). Predicting Student Academic Performance Using Neural Networks: Analyzing the Impact of Transfer Functions, Momentum and Learning Rate. *International Journal of Experimental Research and Review*, 40(Spl Volume), 56–72. <https://doi.org/10.52756/ijerr.2024.v40spl.005>
- **Kalpana, P., Arumarana, E., Hanif, S., & Deebak, T. (2020). *Student Performance Analysis using Machine Learning*. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(6), 211–215. <https://doi.org/10.35940/ijitee.F3585.049620>**

- **Yağcı, M.** (2022). *Educational data mining: Prediction of students' academic performance using machine learning algorithms*. Smart Learning Environments, 9, Article 11. <https://doi.org/10.1186/s40561-022-00192-z>
- **Li, Y., & Botelho, A. F.** (2024). *Bridging Human and AI Explanations to Support Students' Revisions of Open-Ended Responses*. Proceedings of the 17th International Conference on Educational Data Mining (EDM 2024), 100–109.
- **González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R.** (2021). *Artificial Intelligence for Student Assessment: A Systematic Review*. Applied Sciences, 11(12), 5467. <https://doi.org/10.3390/app11125467>
- **Ouyang, F., Zhou, M., & Hwang, G.-J.** (2022). *Exploring the relationships among students' self-efficacy, learning engagement, and academic achievement in online learning: A structural equation modeling approach*. Smart Learning Environments, 9, Article 2. <https://doi.org/10.1186/s40561-022-00192-z>
- Student Performance Analysis and Prediction dataset from Kaggle
<https://www.kaggle.com/code/annastasy/predicting-students-grades/input>