MINISTRY OF EDUCATION
IMAM ABDULRAHMAN BIN
FAISAL UNIVERSITY
COLLEGE OF COMPUTER SCIENCE
& INFORMATION TECHNOLOGY

وزارة التعليم
جامعة الإمام
عبدالرحمن بن فيصل
كلية علوم الحاسب
وتقنية المعلومات

جامعة الإمام عبدالرحمن بن فيصل
IMAM ABDULRAHMAN BIN FAISAL UNIVERSITY

CIS 517: Data Mining and Warehousing
Term 1 – 2023/2024
Milestone 2

Enhancing Education with Data Mining: Analyzing Student Performance

EDUANALYTICS
ENHANCING EDUCATION WITH DATA MINING

- ✓ **Group No. 3**
- ✓ **Members of the Group:**

#	Name	ID	Role
1	Danah Yousef Aldossary	2200003424	Leader
2	Dana Khalid Alsharari	2200004812	Member
3	Sara Mohammed Alzharani	2200005923	Member
4	Fai Saleh Alsadlan	2200005122	Member
5	Fay Mohammed Balhareth	2200003560	Member

Project Submission
02/12/2023, 11:59 PM
15 / 15

أنصحكم تحطون أكواد الكلاسفيكيشنز لا تعتمدون اننا ماحطينا ولا بينقص بمزاجه 😽

Acknowledgment

We are first and foremost grateful to Allah for providing us with the knowledge necessary to complete this project successfully, and we also express our sincere gratitude to our academic supervisor, Dr. Irfan Ullah Abdur Rab, for his invaluable advice, steadfast support, and perceptive criticism during the whole research project. His knowledge and commitment have significantly influenced the course and quality of this project.

We also thank all team members who provided valuable insights and assistance at various project stages. Their collaborative spirit and diverse perspectives enriched the overall quality of the research and contributed to completing this project successfully. Furthermore, we express our gratitude to the participants of this study who shared their dataset and data. Their contribution is integral to the relevance and applicability of the findings. This research was only possible with the collective efforts of everyone mentioned above.

Abstract

In this study, we explore the potential of enhancing education by leveraging Python, a programming language known for its versatility and applicability. The focus is on simplifying the complexities of educational data through Python's accessible libraries, such as Pandas, NumPy, and Scikit-learn. By examining various aspects like student grades and attendance records, we aim to use Python's capabilities to uncover insights that can benefit students and teachers alike. We employ techniques like clustering to group students with similar performance, classification to predict how well a student might do, and association rule mining to find patterns in student behavior. Alongside these technical aspects, we also explore the ethical considerations and privacy concerns involved in using student data, ensuring our approach is responsible. The overarching goal is to provide valuable information to educators and school leaders, helping them identify students who might need extra support and making schools more supportive and effective for everyone involved. In essence, this research underscores how Python, when applied in education, can serve as a valuable tool for data-driven decision-making, contributing to continuous improvement in students' academic outcomes.

Table of Contents

Table of figures

1. Introduction

In the age of digital advancements, technology is playing an increasingly prominent role in shaping the teaching and learning process, significantly impacting education. A notable outcome of this transformation is the utilization of data mining to assess student success. Data mining, a sophisticated analytical tool, empowers educators and organizations to sift through vast amounts of data generated in educational settings, extracting crucial insights. This approach has the potential to revolutionize education by addressing concerns about student performance and learning outcomes. In an educational context, data mining involves extracting patterns, trends, and insights from extensive and intricate datasets, encompassing various features like student demographics, exam performance, attendance records, and even social interactions within the classroom. Consequently, it plays a pivotal role in aiding educators to better understand the factors influencing student achievements by employing complex algorithms and analytical tools. By delving into this data, educators can uncover hidden insights, anticipate future occurrences, and make informed, data-driven decisions. This data-driven approach holds the promise of enhancing educational performance in the long term. The significance of data mining in education is underscored by its potential to address ongoing issues of educational inequality. Teachers can implement personalized interventions to cater to the unique needs of individual students or groups, bridging gaps and fostering inclusive learning environments. Moreover, it helps instructors assess the effectiveness of their lesson plans and curriculum, enabling continuous improvement. Additionally, it proves invaluable in tailoring instructional strategies to each learner's specific needs and preferences, particularly in times of information overload. However, the real revolutionary potential lies in data mining's ability to unravel the complex network of factors influencing student accomplishments. Educators can identify patterns and correlations in historical data, allowing them to spot early signs of academic issues and take proactive measures. Data mining also facilitates the creation of predictive models, foreseeing student outcomes and enabling early intervention and support. Furthermore, it aids in optimizing resource allocation, ensuring that education funding is directed where it is most needed. This makes data mining an invaluable tool for enhancing educational effectiveness and efficiency.

As technology continues to evolve, the continuous adaptation and updating of data mining techniques are crucial to their effectiveness in the ever-changing educational landscape. Additionally, collaboration between researchers, educators, and policymakers is essential. This collaborative effort is vital for developing comprehensive policies and guidelines that protect student privacy while maximizing the potential of data mining in education.

2. Literatures Review

The second part of the report uses a combination of machine learning, deep learning, and instructional approaches to identify the findings, look for gaps, and provide a note that will help future researchers. A literature review produces and critically assesses earlier research on a specific topic. To provide a broad overview of the present state of knowledge on Enhancing Education using Data Mining: Analyzing Student Performance, relevant academic books, articles, and other sources will be examined and summarized.

In the paper created by Mustafa Yağcı's [3]. The study investigates the use of machine learning algorithms to predict undergraduate students' final test scores based on their midterm exam results. The study intends to assist decision-making processes and aid in the development of a learning analysis framework in higher education. The author assesses the efficacy of many machine learning methods, including logistic regression, Naïve Bayes, closest neighbor, random forests, support vector machines, and k-nearest neighbor, to predict students' final exam scores. The dataset utilized consists of the academic achievement grades of 1854 students who enrolled in Turkish Language I at a Turkish institution in the fall semester of 2019–2020. According to the study's findings, the suggested model had a 70–75% classification accuracy. This shows that the model has a moderate level of accuracy in predicting students' academic success. Three parameters were used in the prediction process: faculty, department, and midterm test grades. The significance of data-driven research in identifying and assisting students at a high risk of failing is also covered in this study. Earlier studies have used machine learning techniques to predict academic success grades, including multiple regression, neural networks, decision trees, and Naïve Bayes. According to recent studies, algorithms like random forests, genetic programming, and Naïve Bayes can reach high prediction accuracy. The literature study also emphasizes the necessity of raising educational standards by forecasting pupils' academic success. Student success and achieving are influenced by several factors and qualities, such as psychomotor abilities, course and pre-course performance, student involvement, motivation, and habits. The study shows that utilizing the midterm test scores as the source data, the suggested model, based on machine learning techniques, can accurately predict students' final exam grades.

The paper by Divya Thakur and Nitika Kapoor [4]. The authors emphasize that forecasting a student's academic success is a complicated and experimental issue in educational data mining, focusing on data mining strategies to predict student performance. They highlight the growing accessibility of educational datasets, especially in the context of virtual education, which has drawn more scholars to this field of study. This article outlines the goals of the study, which include examining current data mining techniques in the field of education and comparing, evaluating, and assessing their outcomes. To predict student performance, the authors use the Naive Bayes (NB) and Support Vector Machine (SVM) methods. They also discuss boosting and bagging as ensemble techniques to increase classification accuracy. The advantages of educational data mining, which closes the knowledge gap between computer science and education, are emphasized by the authors. Academic actors, such as instructors, administrators, and students, can gain from the appropriate information that data mining provides, resulting in highly qualified innovations in the field of education. The results obtained and the efficacy of the suggested algorithms are covered in the paper's conclusion. The authors point out that different techniques provide varied outcomes, emphasizing the significance of contrasting and assessing the efficacy of alternative algorithms.

In the research led by Nafuri [5], The authors propose classifying B40 students based on their institutional performance by employing a clustering-based approach. The paper discusses the necessity for additional research on unsupervised learning in the educational environment. The authors developed three unsupervised models—k-means, BIRCH, and DBSCAN—using the data from B40 pupils. They carried out feature selection and pre-processing operations to guarantee the caliber of the training data. The study found that the optimized k-means model on Model B (KMoB) performed the best among all the models. Five clusters of B40 students were formed by KMoB based on their performance, which can assist the government in reducing the proportion of HEI dropouts, increasing graduation rates, and enhancing the socioeconomic standing of students. Data preparation, modeling, and assessment are part of the proposed study's approach. The Ministry of Higher Education's Policy Planning and Research Division provided the dataset utilized in the study, which included 248,568 student records with various variables. Pre-processing operations were carried out on the data to clean and convert it, and feature selection was done in order to choose pertinent details for the clustering analysis. The experiment outcomes demonstrated that the BIRCH, DBSCAN, and k-means algorithms were The experimental findings demonstrated that the dataset was subjected to the k-means, BIRCH, and DBSCAN algorithms and that each model's performance was assessed. The highest-performing model among them was the optimized k-means model on Model B (KMoB), which showed how effectively it could cluster B40 students based on their performance. The study also emphasized the significance of early identification and intervention for improving pupils' academic performance. The clustering approach can assist and direct students in effectively mastering their lectures by recognizing their performance levels during their studies. This could lower the dropout rate and improve overall academic success.

In a study conducted by Arun D. et al. [6], the focus is on using educational data mining to anticipate and prevent low-grade point averages among students. The researchers utilized a dataset from BMS College of Engineering and employed the WEKA platform for data analysis. To address classification and regression challenges, machine learning models were employed, incorporating ensemble methods with voting for subject-wise analysis. The authors opted for soft voting, where the average of predictions within specific algorithm groups is considered the final prediction. The ensemble model, categorized by certain factors, is applied to subject-wise analysis to predict students' performance in various subjects. Notably, the study found that the Random Forest Algorithm excelled in predicting GPAs through regression analysis. The authors underscore the significance of early performance prediction, advocating for timely support to students at risk of falling behind. They highlight the potential advantages of identifying struggling students at the outset rather than after results are disclosed. Overall, the research emphasizes the potential of machine learning models and ensemble methods in enhancing accuracy and early identification of students facing academic challenges.

In a paper prepared by Danping Duan [7], the authors explore the application of the XGBoost algorithm alongside LR, RF, and Lasso in predicting academic performance based on learning behavior data. Investigating the prediction of students' academic performance has evolved into a crucial area of research, with advancements in machine learning algorithms opening up new avenues for analysis. The objective is to identify influential characteristics, construct an effective prediction model, and provide early warnings for timely intervention in students' learning behaviors. The study utilizes data from 10,000 freshmen in an ordinary high school spanning from 2019 to 2020, with a gender distribution of 28.2% males and 72.7% females. The dataset is divided into an 8,000-student training set and a 2,000-student test set. The research focuses on the course "Fundamentals of Computer Application" and employs characteristic correlation analysis to select factors with high correlations as inputs for the algorithm model. A comprehensive comparison of LR, RF, Lasso, and XGBoost algorithms is conducted, with emphasis placed on overall accuracy, stability, and effectiveness. Notably, the XGBoost algorithm distinguishes itself by exhibiting superior performance, particularly under the influence of six robust feature correlations identified through correlation analysis and feature verification. Overall, these studies collectively underscore the potential of machine learning algorithms in predicting students' academic performance. Through the strategic utilization of diverse datasets and continuous algorithmic improvements, researchers have achieved promising outcomes in predicting academic results and gaining insights into the multifaceted factors influencing students' performance.

3. Proposed model description

The Decision Tree model is a versatile and interpretable algorithm chosen for its ability to capture non-linear relationships within the student performance dataset. Decision Trees are particularly well-suited for educational data analysis, as they allow us to understand the decision-making process behind student outcomes. The model will leverage features such as 'math score,' 'reading score,' 'writing score and other relevant factors to predict student performance categories.

The Support Vector Classifier (SVC) is a powerful classification algorithm selected for its effectiveness in handling complex relationships within the data. SVC works well for both linear and non-linear patterns, making it suitable for capturing nuanced interactions between various features impacting student performance. The model will be trained to distinguish between different classes of student performance based on a range of input features.

The Naive Bayes model is chosen for its simplicity and efficiency, particularly when dealing with categorical features and relatively small datasets. Despite its 'naive' assumption of feature independence, Naive Bayes often performs well in practice and can provide insights into the conditional probabilities of different factors influencing student outcomes. The model will be applied to predict student performance categories based on the likelihood of various input features.

All proposed models will undergo a rigorous training process on the preprocessed dataset, which includes handling outliers, addressing class imbalance, and encoding categorical variables. The models will be evaluated using appropriate performance metrics such as accuracy, precision, recall, and F1-score.

4. Empirical Studies

4.1 Dataset description

The dataset, obtained from Royce Kimmons and consisting of 1000 instances [9], includes scores from three exams and various personal, social, and economic factors. The focus is on understanding how these factors interact and influence exam scores. This comprehensive dataset allows for a closer look at the relationships between academic performance, personal attributes, and socio-economic conditions.
The following Figure presents the initial 10 rows of the dataset:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group C	some college	standard	none	64	71	69
1	female	group C	some college	free/reduced	none	73	80	70
2	female	group A	bachelor's degree	standard	completed	63	71	77
3	female	group D	some high school	free/reduced	none	56	64	62
4	female	group D	some college	standard	none	70	84	84
5	male	group D	some high school	standard	none	68	70	65
6	female	group B	some college	standard	none	40	53	51
7	female	group E	some high school	free/reduced	completed	70	76	74
8	female	group B	high school	free/reduced	none	28	50	43
9	male	group E	some high school	free/reduced	completed	58	46	44

Figure 1 Dataset

The following Figure presents the first 10 rows of the dataset after adding 'overall score' (calculated as the mean of math score, reading score, and writing score), 'pass/fail' (set to 1 if the overall score is greater than or equal to 60 and 0 otherwise), and replacing categorical values with numerical values:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score	Overall Score	Pass/Fail
0	0	3	1	0	0	64	71	69	68	1
1	0	3	1	1	0	73	80	70	74	1
2	0	1	4	0	1	63	71	77	70	1
3	0	4	5	1	0	56	64	62	60	1
4	0	4	1	0	0	70	84	84	79	1
5	1	4	5	0	0	68	70	65	67	1
6	0	2	1	0	0	40	53	51	48	0
7	0	5	5	1	1	70	76	74	73	1
8	0	2	2	1	0	28	50	43	40	0
9	1	5	5	1	1	58	46	44	49	0

Figure 2 Replacing categorical values with numerical

The following Figure presents the description of the dataset:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score	Overall Score	Pass/Fail
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	0.506000	0.540000	0.403800	0.332000	0.331000	0.607529	0.640798	0.624071	0.623094	0.718000
std	0.500214	0.282365	0.319073	0.471167	0.470809	0.178564	0.175191	0.183570	0.169574	0.450198
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.250000	0.200000	0.000000	0.000000	0.482353	0.523810	0.500000	0.517647	0.000000
50%	1.000000	0.500000	0.400000	0.000000	0.000000	0.611765	0.642857	0.630952	0.635294	1.000000
75%	1.000000	0.750000	0.600000	1.000000	1.000000	0.741176	0.761905	0.750000	0.741176	1.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Figure 3 Dataset description

In order to assess the presence of outliers in the dataset, we employed a boxplot visualization. The boxplot provides a concise summary of the distribution of the data, highlighting the interquartile range (IQR) and identifying potential outliers beyond the whiskers. As shown in the figure 4, which includes boxplots with and without outliers, we proceeded to remove the outliers using Z-score normalization:

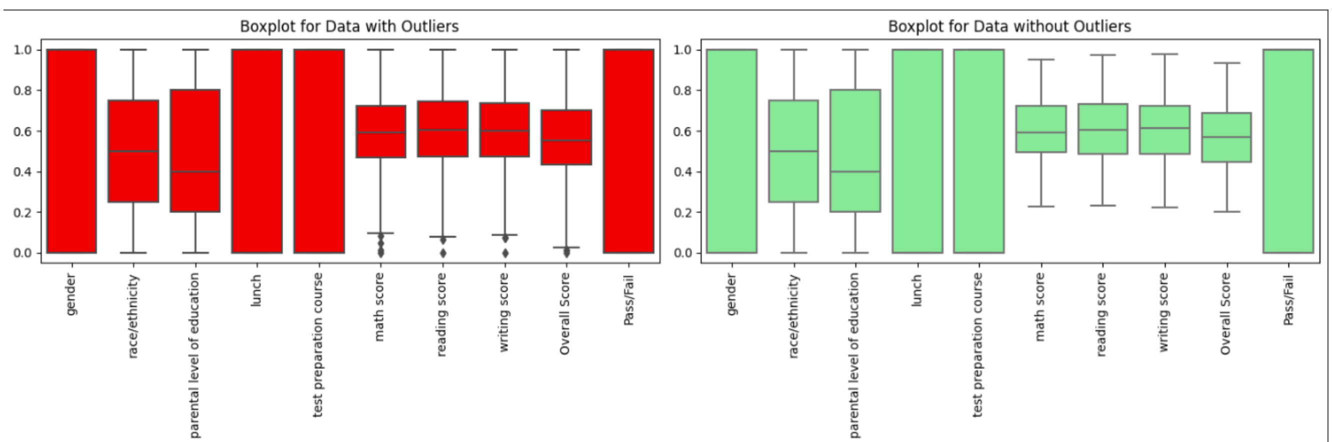

Figure 4 Remove the outlier

```
[155] print("Original Data Shape:", data.shape)
      print("Data Shape After Removing Outliers:", data_no_outliers.shape)

      Original Data Shape: (1000, 10)
      Data Shape After Removing Outliers: (909, 10)
```

Figure 5 Dataset size after remove the outliers

In order to address the issue of class imbalance in the dataset, we utilized a RandomOverSampler technique. This method involves oversampling the minority class to achieve a more balanced distribution of classes. As shown in the figure 6, which includes visualizations before and after oversampling, we applied RandomOverSampler to mitigate the class imbalance and enhance the representation of the minority class in the dataset:

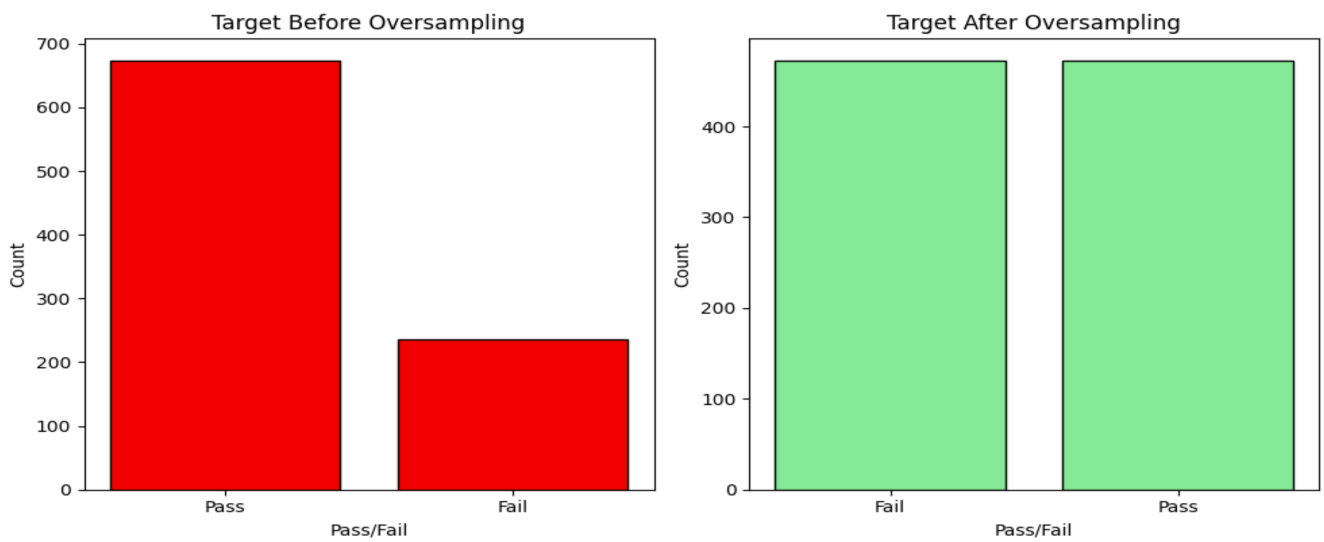

Figure 6 Over Sampling

Given the absence of any missing values, in figure 7 is the description of the dataset after addressing outliers and handling class imbalance:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score	Overall Score	Pass/Fail
count	925.000000	925.000000	925.000000	925.000000	925.000000	925.000000	925.000000	925.000000	925.000000	925.000000
mean	0.509189	0.534595	0.398703	0.322162	0.321081	0.614143	0.645997	0.629279	0.628769	0.739459
std	0.500186	0.283508	0.317591	0.467558	0.467144	0.152357	0.149015	0.157702	0.142310	0.439167
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.258824	0.297619	0.261905	0.305882	0.000000
25%	0.000000	0.250000	0.200000	0.000000	0.000000	0.505882	0.535714	0.511905	0.517647	0.000000
50%	1.000000	0.500000	0.400000	0.000000	0.000000	0.611765	0.654762	0.630952	0.635294	1.000000
75%	1.000000	0.750000	0.600000	1.000000	1.000000	0.729412	0.750000	0.738095	0.729412	1.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	0.941176	0.988095	0.988095	0.952941	1.000000

Figure 7 Handle imbalance

4.2 Experimental setup

Following the preprocessing stage, we will now investigate the experimental study's specifics. The integrated development environment (IDE) for the study was Jupyter Notebook, which was built with Python and the Scikit-learn module. The primary objective of this study is to predict whether a student's overall grade qualifies for a pass or not. For analysis, a dataset of 909 rows of data was obtained.

To begin the investigation, we separated the dataset into a training set and a testing set using Scikit-learn's "train_test_split" method. The training dataset accounted for 70% of the data, with the remaining 30% constituting the testing dataset. Following that, we built prediction models using various classification techniques, as described in Section 3. On the dataset, the chosen methods are Decision Tree, Support Vector Machines (SVM), and Naive Bayes were used independently.

Each model was trained and evaluated using relevant performance indicators, allowing for a thorough comparison of the three techniques. This comparison research provides significant insights into each algorithm's strengths and drawbacks in terms of prediction accuracy for forecasting student achievement using machine learning approaches.

4.3 Performance measure

The following performance measures have been used [8]:

Accuracy: The overall correctness of the model's predictions.

$$\frac{(TP + TN)}{(TP + FP + FN + TN)}$$

Precision: The accuracy of the positive predictions made by the model.

$$\frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})}$$

Recall: The ability of the model to capture all positive instances.

$$\frac{TP}{(TP + FN)}$$

F1 Score: A balanced measure combining precision and recall.

$$\frac{2(precision\ x\ recall)}{(precision\ +\ recall)}$$

5. Result and discussion

In the evaluation of the Decision Tree (DT) model, we achieve perfect predictions with an accuracy of 100%. For class 0, precision, recall, and the F1-score are all 100%, indicating flawless performance. Similarly, class 1 demonstrates impeccable precision, recall, and F1-score, contributing to the overall success of the model, Classification Report shown below:

Classification Report for Decision Tree:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	72
1	1.00	1.00	1.00	201
accuracy			1.00	273
macro avg	1.00	1.00	1.00	273
weighted avg	1.00	1.00	1.00	273

Figure 8 DT Result

Moving on to the Support Vector Machines (SVM) model, we observe an accuracy of 96%. The precision for class 0 is 87%, indicating 87% accuracy in predicting instances of class 0. The recall for class 0 is 99%, showing the model's effectiveness in identifying 99% of actual instances of class 0. The F1-score for class 0, representing a balanced measure of the model's performance, is 92%. For class 1, the precision is 99%, the recall is 95%, and the F1-score is 97%. The macro-average and weighted average F1-scores stand at 95%, emphasizing the robust predictive capabilities of the SVM model, Classification Report shown below:

Classification Report for SVM:

	precision	recall	f1-score	support
0	0.87	0.99	0.92	72
1	0.99	0.95	0.97	201
accuracy			0.96	273
macro avg	0.93	0.97	0.95	273
weighted avg	0.96	0.96	0.96	273

Figure 9 SVM Result

In the case of Naive Bayes, an accuracy of 74% is achieved. The precision for class 0 is 100%, indicating that all predicted instances of class 0 are correct. However, the recall for class 0 is just 1%, resulting in a nuanced performance for class 0. For class 1, the precision is 74%, the recall is 100%, and the F1-score is 85%. The macro-average F1-score, reflecting the average performance across both classes, is 44%, highlighting the model's limitations in dealing with class imbalances, Classification Report shown below:

Classification Report For Naive Bayes:

	precision	recall	f1-score	support
0	1.00	0.01	0.03	72
1	0.74	1.00	0.85	201
accuracy			0.74	273
macro avg	0.87	0.51	0.44	273
weighted avg	0.81	0.74	0.63	273

Figure 10 NB Result

Decision Tree model stands out as the optimal choice, achieving flawless accuracy, precision, recall, and F1-scores in both classes. While the SVM performs well at 96% accuracy, the Decision Tree's perfect forecasts make it the preferred option for predicting student performance in this context. The Naive Bayes model, though accurate for class 0, faces imbalances and achieves a lower macro-average F1-score. Overall, the Decision Tree excels in forecasting student success.

A Confusion Matrix is a tabular representation of a classification algorithm's performance in machine learning. It breaks into true positive (TP), true negative (TN), false positive (FP), and false negative (FN) forecasts in great depth. These components allow for the computation of several performance measures including as accuracy, precision, recall, and F1-score. The matrix provides insight into how well a model performs, particularly in terms of correctly and wrongly predicted examples across different classes. [8] Below shows the Confusion Matrix for each algorithm that was used:

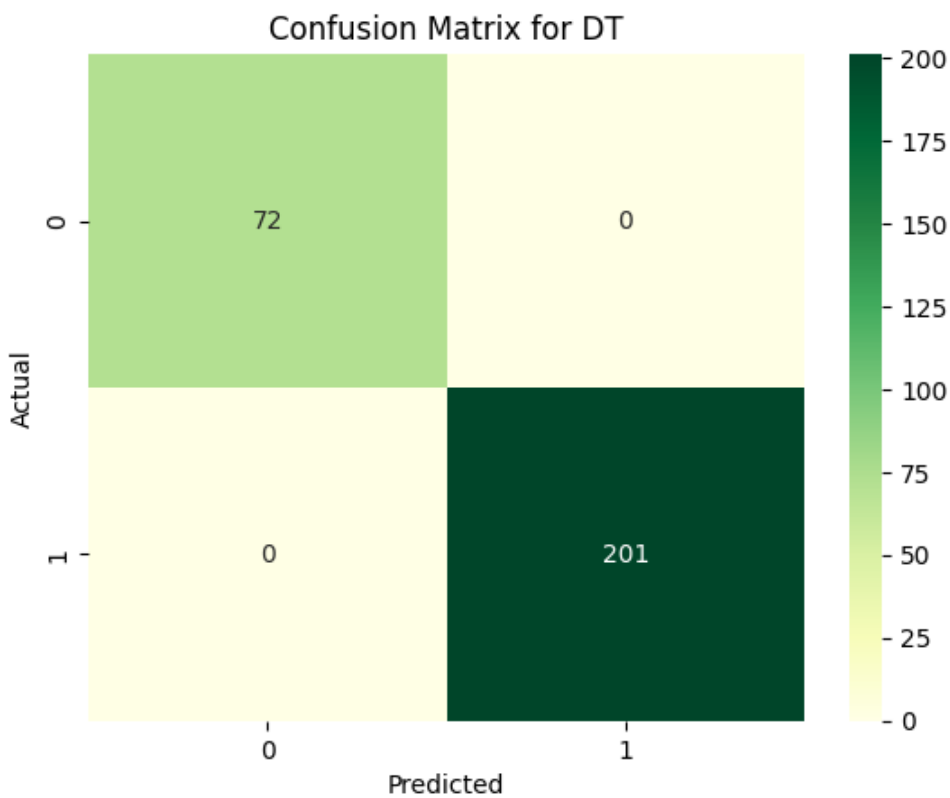

Figure 11 Confusion Matrix Of DT

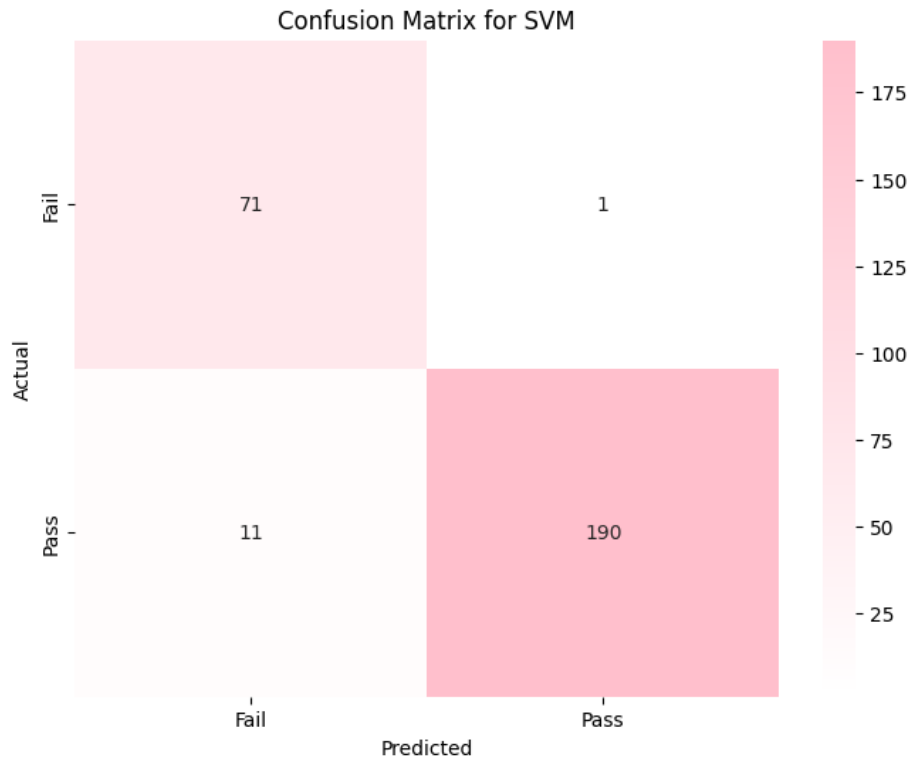

Figure 12 Confusion Matrix Of SVM

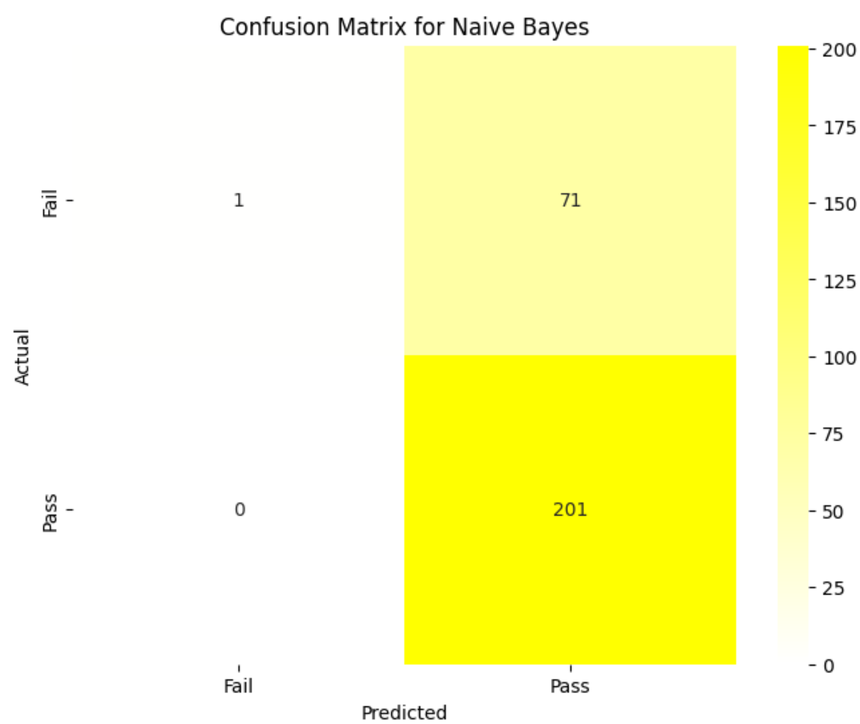

Figure 13 Confusion Matrix Of NB

6. Conclusion

In the end, we used three different machine learning models, including Decision Tree (DT), Support Vector Machines (SVM), and Naive Bayes, to predict student achievement based on a precisely produced dataset. The findings indicate the Decision Tree model's supremacy, as it not only attained faultless accuracy but also displayed remarkable precision, recall, and F1-scores for both pass and fail classes. This exceptional result establishes the Decision Tree as the most trustworthy option for forecasting student outcomes in this specific situation. Although the SVM model achieved 96% accuracy, the Decision Tree's perfect predictions made it the better choice for this dataset. Naive Bayes, on the other hand, struggled with imbalances, as seen by a lower macro-average F1-score. Finally, the Decision Tree model outperforms other educational prediction models in terms of forecasting student achievement.

References

[1] Carnegie Mellon University, "What is EDM? - DataLab - Carnegie Mellon University," *Cmu.edu*, 2013. https://www.cmu.edu/datalab/getting-started/what-is-edm.html (accessed Nov. 24, 2023).

[2] "Better Learning via Educational Data Mining and Analytics," *Kdnuggets.com*, 2018. https://www.kdnuggets.com/2012/11/better-learning-via-educational-data-mining-analytics.html (accessed Nov. 24, 2023).

[3] Mustafa Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, Mar. 2022, doi: https://doi.org/10.1186/s40561-022-00192-z.

[4] D. Thakur and N. Kapoor, "Predicting Student's Performance using Data Mining Algorithm," Mar. 2022, doi: https://doi.org/10.1109/icacta54488.2022.9753265.

[5] A. Fikri, Nor Samsiah Sani, Nur, A. Hadi, and M. Aliff, "Clustering Analysis for Classifying Student Academic Performance in Higher Education," *Applied sciences*, vol. 12, no. 19, pp. 9467–9467, Sep. 2022, doi: https://doi.org/10.3390/app12199467.

[6] D. K. Arun, V Namratha, B. V. Ramyashree, Y. Jain, and Antara Roy Choudhury, "Student Academic Performance Prediction using Educational Data Mining," Jan. 2021, doi: https://doi.org/10.1109/iccci50826.2021.9457021.

[7] D. Duan, Dai Chun-Ping, and R. Tu, "Research on the Prediction of Students' Academic Performance Based on XGBoost," Dec. 2021, doi: https://doi.org/10.1109/eitt53287.2021.00068.

[8] N. Beheshti, "Guide to Confusion Matrices & Classification Performance Metrics," *Medium*, Feb. 10, 2022. https://towardsdatascience.com/guide-to-confusion-matrices-classification-performance-metrics-a0ebfc08408e (accessed Nov. 24, 2023).

[9] "Exam Scores," roycekimmons.com. http://roycekimmons.com/tools/generated_data/exams