Student Academic Performance Prediction Using Machine Learning Approach

Asim Iftikhar[1], Noor Ul Huda[1], Majid Kaleem[2], Fiza Maroof[1], Kashaf Suleman[3], Abiha Zehra[1]
[1]Institute of Business Management (IoBM), College of Computer Science & Info. Sys, Korangi Creek, Karachi, Pakistan
[2]Bahria University, Department of Software Engineering, 13 National Stadium Road, Karachi, Pakistan
[3]Institute of Business Administration (IBA), School of Business Studies, University Road, Karachi, Pakistan

E-mail: asim.iftikhar@iobm.edu.pk, noor.huda@iobm.edu.pk, majidkaleem.bukc@bahria.edu.pk, fizamaroof077@gmail.com, syeda.suleman.29900@iba.khi.edu.pk, abihazehra9501@gmail.com

Corresponding Author: Asim Iftikhar, asim.iftikhar@iobm.edu.pk

Received: 18-06-2024; **Accepted:** 25-07-2024; **Published**: 30-12-2024

Abstract: The academic performance of students in higher education has been a focal point of extensive research aimed at addressing issues such as academic underachievement, increased dropout rates, and delays in graduation, among other persistent challenges. Basically, student performance is the accomplishment of educational objectives, be they short-term or long-term. The study therefore describes the prediction of Student Performance aimed at improving the attainment of academic outcomes in higher education. By considering several performance factors like patterns of attendance, parent involvement, individual study habits, and preference in teaching methods, this system can portray a view of each student's academic performance in a holistic way. Support Vector Machine (SVM) and Artificial Neural Network (ANN) algorithms have implemented in this proposed model to classify and predict student academic performance. A comparative study was also undertaken to determine which of these two algorithms has better accuracy or precision. The results show that the SVM outperformed the ANN.

Keywords: Student Performance Prediction, Higher Education, Machine Learning, Academic Performance, Parental Involvement, Study Habits, Teaching Methods

1. Introduction

Nowadays, most colleges struggle to draw in prospective students due to very competitive academic markets [1]. Therefore, studying students' academic achievements is crucial for fostering their growth and raising the standard of higher education, both of which eventually improve an institution's standing. Academicians use a variety of metrics to assess students' performance, including final grades, grade point average (GPA), and potential employment opportunities [2], [3]. Attracting and keeping prospective students is a major problem for colleges and universities in today's fiercely competitive educational industry. As a result, studying students' academic performance is essential to encouraging their development and raising the standard of higher education as a whole.

Student academic performance is influenced by a diverse range of factors, including personal, social, psychological, and environmental variables [4], [5]. Machine learning can be leveraged by universities to predict which students might fail, identify those with poor performance, forecast exam results in specific subjects, and calculate graduation ratios. This strategic information enables universities to develop and enhance educational policies, supporting struggling students and guiding them towards suitable specializations [6], [7]. The implementation of Machine learning algorithms can significantly improve the academic performance of educational institutions. The results of academic performance prediction can be used to categorize students, enabling the university to provide them additional resources for tutoring or timely assistance [8]. Predicting a student's academic performance is essential to determine which group of students will perform well during the semester so that scholarships may be given to them and, more crucially, to identify which students may struggle during the semester so that remediation can be provided to them [5].

In this study a predictive model has been developed that offers a comprehensive view of student performance. We aim to explore how factors such as attendance patterns, parental involvement, individual study habits, and preferred teaching methods collectively to shape a student's academic journey. By understanding these multifaceted aspects, we intend to provide educators with a more holistic and personalized approach to student support, ultimately enhancing the educational experience and outcomes for students.

There are five (5) sections in this document. The research study's introduction is covered in the first section. Section 2 provides an explanation of related research activity. Section 3 provides an explanation of machine learning and implemented algorithms. Section 4 elaborates research Methodology. The Results findings and discussion are covered in Section 5, and the study is concluded in the final section.

2. Related Work

The student performance in a course was predicted by researchers in [7]. A method called data mining is used to find hidden patterns in vast volumes of existing data. These trends could be useful for forecasting and analysis. The collection of data mining applications in the subject of education is known as education data mining. The analysis of student and teacher data is the focus of these programs. The analysis could be applied to prediction or classification. Machine learning techniques including SVM, ID3, C4.5, and Nave Bayes are studied. The experimental investigation makes use of the UCI machinery student performance data set. Certain criteria, like as accuracy and error rate, are used to examine algorithms.

Both clustering and classifying strategies were presented by the authors [9] in order to determine the early effects of student academic performance on GPA. In order to investigate the relationship between these factors and GPAs, the paper employs a dimensionality reduction mechanism by the T-SNE algorithm for the clustering technique with a variety of early-stage factors, including first-level courses, admission scores, general aptitude tests and academic achievement tests. The study offers trials on various machine learning models for the classification technique, which predicts student success at an early stage utilizing various characteristics such as course grades and admission test scores. We examine the models' quality using a variety of assessment indicators. The findings imply that early on, educational systems can reduce the likelihood of students failing.

The prediction structure based in the behavior classification-based e-learning performance (BCEP) has been presented by the researchers in [10]. By carefully choosing certain critical e-learning features, BCEP constructs a machine learning-based learning performance predictor at the very onset. Next the fusion feature is used: here the behavior data is used as an input to the behavior classification model, to ascertain the category feature values of the behavior of each type. The authors also suggest process-behavior classification (PBC) which is rooted in an online e-learning behavior classification process to deal with shortcomings in the existing model. They did this because they claim that the Open University Learning Analytics Dataset suggests superior performance of PBC over the other models. The researchers claim that this new approach of theirs gives a better option for evaluating e-learning categorization techniques in a quantitative way and is not merely a qualitative claim.

Authors in [11] presented a new classification system for blended courses after analyzing learning outcomes in terms of learning data across all of the mixed courses offered at a Chinese institution. This approach has the advantage that machine learning systems can easily use log data from a learning management system to algorithmically classify blended courses because the criteria used for classification are clearly defendable based on data about students' online behavior. According to the study's findings, prediction accuracy improved in each area once blended courses were categorized based on students' online behavior. The corresponding total accuracy rates were 38.2%, 48.4%, 42.3%, 42.4%, and 74.7%. These results suggested that rather than focusing on a single kind of online learning activity, the majority of students should be actively involved in a range of them, like watching online videos or turning in online assignments, in order to accurately predict students' learning outcomes in a blended course. In Course H, the prediction model's accuracy for grades was 80.6%, 85.3%, 63%, 54.8%, and 14.3%, respectively. The outcomes showed how well the suggested model could predict learning outcomes in blended learning environments. Lastly, we discovered that no single online learning activity significantly influenced the ability to predict students' final grades.

The use of CART, J48, K-Nearest Neighbor, Naïve Bayes, C5.0, Random Forest and Support Vector Machine for student performance prediction was examined by researchers in [12]. To compare, three datasets with different input parameters come from school, college, and e-learning platforms. Comparative findings of all algorithms used in this study on varying tuning parameters are presented in the paper. Three distinct datasets are used to evaluate these methods' performance. The findings indicate that Random Forest and C5.0 outperform other algorithms.

A group of researchers [13] is interested in increasing the efficiency of the academia by decreasing the failure rate of what the call are the “at-risk” students. They use various machine learning and deep learning methods to create predictive models to help enhance the performance of "at-risk" students. The performances of the predictive algorithms is compared and they always selected the one which is the best in terms of f-score, support, precision and accuracy for the different percentages of the time spans of the course. They claim that experimentally, the best results come from the models trained by random forest (RF).

The possibility of creating trustworthy models for forecasting student performance with artificial intelligence is examined by the authors in [14]. In particular, we suggest utilizing feature weighted support vector machines and ANN learning to predict academic achievement in two steps. For coarse-grained binary classification (pass, P1, or fail, P0), a feature weighted SVM is used, in which the significance of various characteristics to the result is determined using information gain ratios. The P1 and P0 classes are then fine-grained, multi-class trained independently using ANN learning after precise score levels are separated from D to A+. The efficacy of this hybridized approach has been demonstrated by the trials and our follow-up ablation investigation, which are carried out using student datasets from two secondary schools in Portugal.

3. Machine Learning

A branch of artificial intelligence (AI) called machine learning uses prior experience to autonomously comprehend and improve itself without the aid of any other software. This particular kind of learning focuses on teaching computer programs that can retrieve data and turn it into something useful for themselves. This learning process, which is founded on data or observations, such as teaching, examples, or firsthand experience, helps people make better judgments in the future. The main goal is to enable computers that can learn on their own without help and modify processes appropriately [15].

Machine learning (ML) has advanced significantly over the last three decades, from experimental interest to practical technology with widely distributed commercial applications [16]. For the building of useful speech recognition systems, computer vision systems, robot control, natural language processing and other artificial intelligence applications, machine learning has emerged as the method of choice. A program can use machine learning capabilities in a variety of contexts, including ML and ML applications software systems, methods, and libraries that provide ML features [17].

3.1. **Support Vector Mahine**

SVM, or Support Vector Machine, stands out as a unique form of supervised machine learning. It remains a cornerstone technique capable of effectively tackling classification problems and regression tasks even in the realm of big data [18]. This method, which is illustrated in Figure 2, greatly aids in the resolution of classification issues [19]. SVM uses an n-dimensional space to represent each data point, where n is the number of features. The value of each characteristic is associated with a coordinate in this space. Through the identification of a hyperplane, SVM distinguishes between two classes by effectively partitioning the data [20].

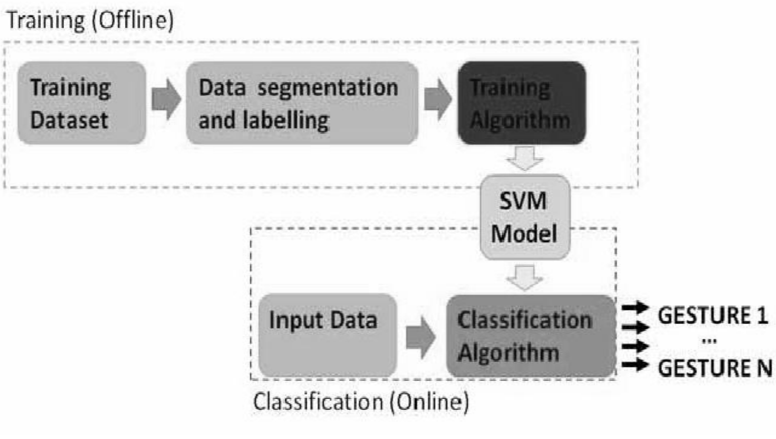


Figure.1. SVM algorithm block diagram *[19]*

3.2. **Artificial Neural Network**

Models known as ANN are based on biological networks, particularly the neural networks found in the brains of humans and animals. Among the neurological operations that the human brain carries out are speech recognition, pattern recognition, and face recognition, all of which it attempts to mimic [21]. +e networks continue the iteration process with forward and backward propagation after learning from instances. This process is carried out until the output satisfies the response that was given with the required level of precision. Each example has a sequence of inputs and responses, which are related outputs. The network makes adjustments using internal connections called weights [22]. There are three layers that make up a standard ANN (see Figure 2 [23]). For instance, an object (such as a cube) could be an input layer, and the output aim would be to precisely identify and recognize the cube. Through an iterative process, the hidden layer assists in completing the cube recognition job. Solid lines indicate the links (or channels) that link all three layers.

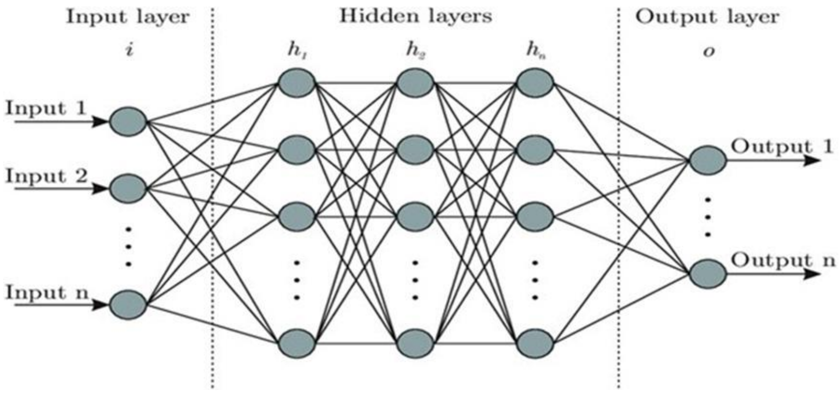


Figure.2. Artificial neural network architecture *[23]*

4. Methodology

Support Vector Machine (SVM) and Artificial Neural Network (ANN) algorithms have been used to forecast academic results and predict student performance under the umbrella of machine learning approaches.

4.1 **Research Design**

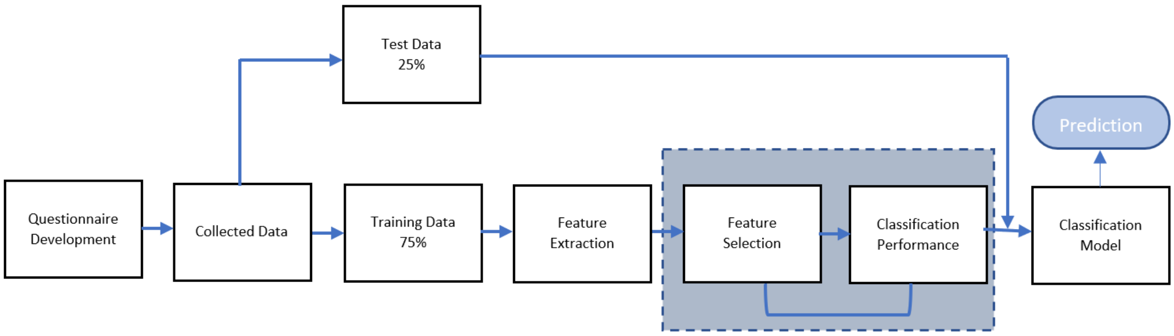
This research will employ experimental as well as simulation-based research design. A quantitative approach will be used to predict student academic performance using collected data through developed questionnaire. The study follows a step-by-step methodology as shown in Figure 3 comprising questionnaire development, data collection, pre-processing and analysis of results using ML algorithms.

Step 1 - Questionnaire Development: The process begins with the development of a detailed questionnaire. This questionnaire will be designed to gather relevant information about the students, including their academic background, study habits, preferred Teaching Methods, Parental Involvement, personal circumstances, and other factors that may influence their performance.

Step 2 - Data Collection: Once the questionnaire is developed, data will be collected from the targeted student population. This step involves distributing the questionnaire, collecting responses, and organizing the data for further processing.

Step 3 - Feature Extraction and Selection: Data will be transformed into a more meaningful representation by identifying and extracting most relevant information, that leads to improved model accuracy, faster training time and better interpretability by reducing data complexity.

Step 4 - Data Classification: For classification tasks, Support Vector Machine SVM and ANN algorithms will be



employed. These algorithms help in categorizing students based on various features, enabling us to identify patterns and groupings within the data.

Step 5 - Performance Prediction: Using the same algorithms algorithm, student performance will be predicted. The SVM and ANN models will be trained on the pre-processed data, and then used to forecast outcomes based on new or unseen data.

Figure.3. Research Framework

4.2 **Data Collection Procedure**

The well-designed questionnaire will be used to gather important information from students at various universities. The dataset includes details such as demographics, academic history, attendance, parental involvement, class participation, study hours, and study habits. After collecting the data, it will be cleaned and processed to handle any missing or unusual values. Data will be collected from different universities, with contributions coming in both soft copy form and through Learning Management Systems, as well as directly from the questionnaires. To ensure diversity and credibility, the survey was distributed to Academic degree awarding institutions based in Pakistan. Out of the 2,110 responses received, 110 were rejected due to being incomplete or containing inaccurate answers. The remaining responses formed a robust dataset that enabled our machine learning model to predict student performance effectively. The dataset contains 2 classes as Performing Well and Not Performing Well as shown in Table 1.

Performance Criteria	**Class**	**Category**
Grades in (C, D and F category)	0	Not Performing Well
Grades in (A and B category)	1	Performing Well

Table.1. Dataset Classes Related to Student Performance

Tables 2 provide a glimpse of the responses received. The results presented in these tables indicate that the data contains variability in responses, with numerical values representing answers to all questions. This variability underscores the diverse range of input from students across different universities, enhancing the robustness and reliability of our dataset for accurate student performance prediction.

Attendance	Parental Involvement	Study Habits	Preferred Teaching Methods	Final Grade	performance
9	2	4	1	5	0
28	3	1	3	4	0
8	3	1	1	4	0
2	2	3	3	6	0
23	3	4	2	2	1
7	1	3	1	2	1
9	3	4	1	5	0
9	2	4	2	6	0
2	1	2	1	6	0
3	2	4	2	2	1
26	2	3	3	2	1
7	2	3	3	1	1
28	3	2	3	4	0
18	2	4	2	5	1

TABLE.2. SAMPLE DATASET

5. Results and Findings

The primary objective of this study is to evaluate and compare the performance of SVM and ANN models in classifying a given dataset. By analyzing each model's precision, recall, F1 score, and accuracy, we aim to identify the strengths and weaknesses of these algorithms in different classification scenarios.

5.1 **SVM Results**

Table 3 represents the classification performance metrics SVM model depicting an exceptionally performed result with respect to all parameters that were calculated. The classification yielded an accuracy of 98% for both classes demonstrating both efficacy and reliability of SVM as the classifier for the task of classification. Based on the SVM model, we get a precision of 0.99 for class 0, which represents the accuracy of our all predicted class 0 instances (99% of the instances predicted as class 0 was predicted correctly). Similarly, class 0 recall was 0.98, so we covered 98% of the real class 0 instances. For class 0 the F1-score was 0.99. As we have an F1-score that balances recall and precision we can see that we have a good performance. The model also performed consistently for class 1 having precision 0.97, recall 0.98 and F1-score 0.98.

For this dataset, the actual instances of class 0 and class 1 are 253 and 148 respectively as per the support column indicators.

	Precision	Recall	F1-Score	Support
0	0.99	0.98	0.99	253
1	0.97	0.98	0.98	148
			Accuracy	98%

Table. 3. SVM Classification Results

The confusion matrix effectively demonstrates the classification performance of the SVM model as shown in figure 4. It consists of four sections: the top left (249) indicates true positives for class 0, showing that 249 instances were accurately classified as class 0. The top right (4) represents false positives, where 4 instances were incorrectly identified as class 0 instead of class 1. The bottom left (3) denotes false negatives, meaning 3 instances of class 1 were wrongly classified as class 0. Lastly, the bottom right (145) shows true positives for class 1, with 145 instances correctly identified as class 1. These results reflect the model's high accuracy, as it successfully classified most instances with minimal errors, highlighting its effectiveness and reliability in distinguishing between the two classes.

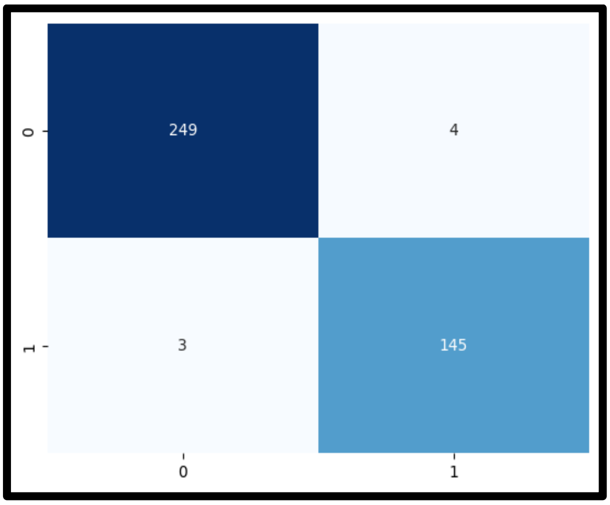


Figure. 4. SVM Confusion Matrix

Figure 5 illustrates the Precision-Recall (PR) curve for a Support Vector Machine (SVM) classifier. This curve assesses the trade-off between precision and recall across a range of decision thresholds. Precision (y-axis) indicates the proportion of true positive predictions out of all positive predictions made by the model. High precision means fewer false positives. Recall (x-axis) reflects the proportion of true positive predictions out of all actual positive instances. High recall indicates fewer false negatives. At the leftmost point (recall = 0), precision is very high (~1.0). This indicates that the classifier exhibits a conservative approach, issuing positive predictions only when it has a high level of confidence in its assessment. As recall increases, precision slightly decreases, indicating that the model starts predicting more positives (reducing false negatives), but also introduces some false positives. At the rightmost point (recall = 1.0), precision drops sharply. This indicates that when all instances are classified as positive (maximizing recall), the number of false positives increases significantly.

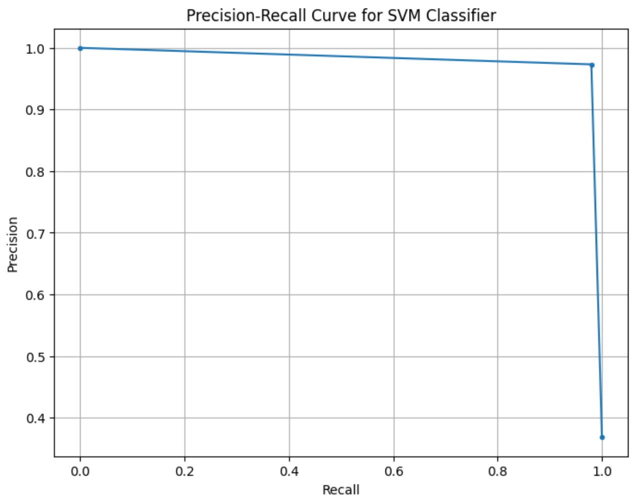


Figure. 5. SVM precision-recall curve

5.2 **ANN Results**

Table 4 depicts classification results of the ANN model used in classifying data that were in the two classes represented by the label "0", label "1"" Additionally, these are the performance metrics: Precision, Recall, F1-Score, and Support Precision in category "0" on the model was 0.99, which indicated that after the conclusion of the model, 99% of items labelled "0" were actually "0". The Recall for this category is 0.98, indicating that the model correctly identified 98% of all actual "0" items. The F1-Score, which is the harmonic mean of Precision and Recall, is 0.98 for this category, reflecting a balanced high performance. The Support, representing the actual number of items in this category, is 253. For category "1", the Precision is 0.96, meaning that 96% of the items it labeled as "1" were correct. The Recall is 0.98, showing that the model identified 98% of all actual "1" items correctly. The F1-Score for this category is 0.97, indicating strong performance, although slightly lower than category "0". The Support for this category is 148, showing fewer items compared to category "0".

	Precision	**Recall**	**F1-Score**	**Support**
0	0.99	0.98	0.98	253
1	0.96	0.98	0.97	148
Accuracy				98%

Table. 4. ANN Classification Results

The confusion matrix shown in Figure 6 illustrates the performance of the Artificial Neural Network (ANN) model in classifying data into two categories, labeled as "0" and "1". The matrix is divided into four sections: The top-left cell (247) represents the number of true positives for category "0", meaning that the model correctly identified 247 items as "0". The top-right cell (6) indicates the number of false positives for category "0", meaning that the model incorrectly classified 6 items as "0" when they belong to category "1". The bottom-left cell (3) shows the number of false negatives for category "1", meaning that the model incorrectly classified 3 items as "1" when they belong to category "0". The bottom-right cell (145) represents the number of true positives for category "1", meaning that the model correctly identified 145 items as "1". This confusion matrix highlights the high accuracy of the ANN model, with a large number of correct classifications (247 for "0" and 145 for "1") and a small number of misclassifications (6 false positives and 3 false negatives). This visual representation supports the earlier reported accuracy of 98%, demonstrating the model's effectiveness in distinguishing between the two categories.

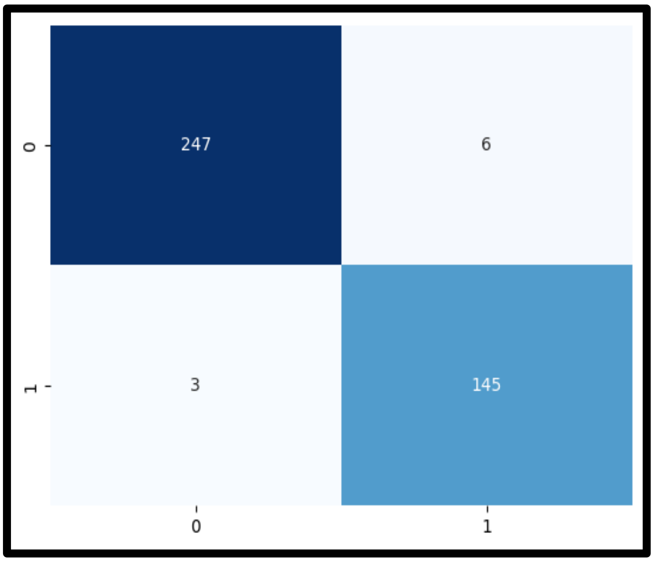


Figure. 6. ANN Confusion Matrix

In Figure 7, At the leftmost point, the classifier makes fewer positive predictions, achieving high precision (close to 1.0) at the cost of very low recall (near 0.0). This indicates only the most confident predictions are being considered. As recall increases (more positive predictions), precision gradually decreases. This reflects the trade-off: as the model tries to identify more true positives, it risks more false positives. At the far right, recall reaches its maximum value (all positives are identified), but precision drops sharply. This happens because the model predicts every instance as positive, increasing the number of false positives.

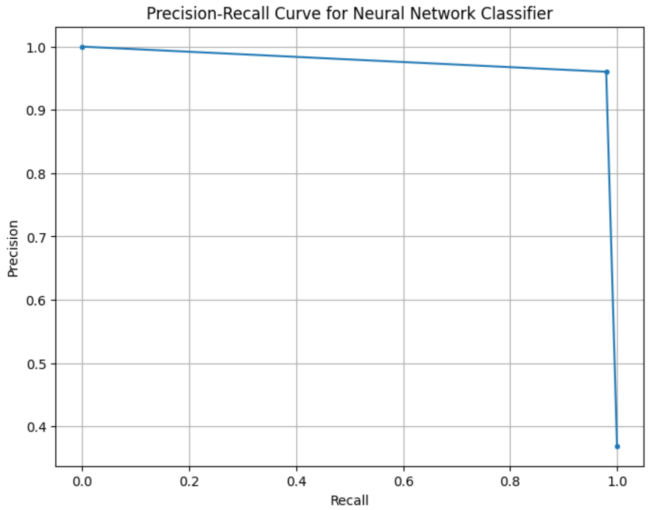


Figure. 7. ANN precision-recall curve

Figure 8 illustrates those significant features which were extracted after feature engineering. Their contribution and significance in predicting student performance is shown by the length of the respective bars. Their importance were extracted by performing multicollinearity. This technique selects and combines multiple predictor variables for greater degree of accuracy.

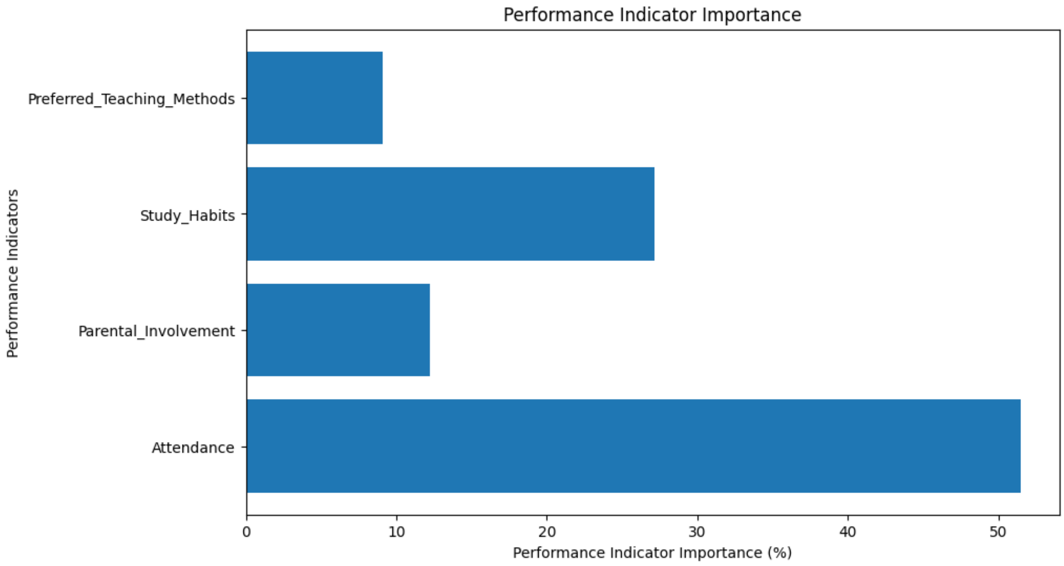


Figure. 8. Performance Indicator Importance

5.3 **Comparative Analysis**

Both the Support Vector Machine (SVM) and Artificial Neural Network (ANN) models demonstrated high effectiveness in predicting student performance, each achieving an accuracy of 98%. However, upon closer inspection of the performance metrics, SVM slightly outperforms ANN. Specifically, SVM has higher precision for students classified as "Performing Well" (0.97 compared to ANN's 0.96), indicating fewer false positives in this category. Additionally, SVM achieves higher F1-Scores for both "Not Performing Well" (0.99 vs. ANN's 0.98) and "Performing Well" (0.98 vs. ANN's 0.97), suggesting a better balance between precision and recall. While both models are highly capable, SVM's superior precision and F1-Score make it a slightly more reliable choice for accurately classifying student performance. However, considerations such as computational efficiency, ease of implementation, and resource availability might influence the final choice between these two models. Further tuning and evaluation might also enhance their respective performances.

6. Conclusion and Future Directions

The primary aim of this research is to create a predictive model that captures comprehensively the performance of students considering several influencing factors-attendance patterns, parental involvement, study habits, and preferred teaching methods. To ensure better accuracy so that corrective measures can be initiated well in advance, the study undertakes classification and prediction of student performance using SVM and ANN algorithms. It would help policy-makers and educators in making rational choices as it throws some light on which particular machine learning algorithm stands the best chance of success relative to others. It overcame the limitations of traditional assessment methods because it included a much wider array of variables with an impact on student success.

Future studies can be extended to this research project in several aspects. Firstly, the Automated tool can be developed for Academic institutions and serve as a performance predictive system in detecting at-risk students and optimizing teaching methods for improved retention rates and higher quality educational outcomes. Secondly, more diverse variables can be added, along with extracurricular activities, mental health indicators, and socioeconomic factors. This will further enhance the model's predictive accuracy regarding student performance. Finally, Deep Learning algorithms could be introduced that would ensure enhanced responsiveness of the system with a great deal of adaptability.

Acknowledgments

The authors would like to thank their mentors whose consistent help and directions made this research possible. The authors would like to thank Institute of Business Management (IoBM) for facilitating such a positive research environment.

References

[1] D. T. Ha, P. T. T. Loan, C. N. Giap, and N. T. L. Huong, “An empirical study for student academic performance prediction using machine learning techniques,” *Int. J. Comput. Sci. Inf. Secur. IJCSIS*, vol. 18, no. 3, pp. 75–82, 2020.

[2] R. Ghorbani and R. Ghousi, “Comparing different resampling methods in predicting students’ performance using machine learning techniques,” *IEEE Access*, vol. 8, pp. 67899–67911, 2020.

[3] S. Priya, T. Ankit, and D. Divyansh, “Student performance prediction using machine learning,” in *Advances in parallel computing technologies and applications*, IOS Press, 2021, pp. 167–174. Accessed: Dec. 17, 2024. [Online]. Available: https://ebooks.iospress.nl/volumearticle/58051

[4] S. Priya, T. Ankit, and D. Divyansh, “Student performance prediction using machine learning,” in *Advances in parallel computing technologies and applications*, IOS Press, 2021, pp. 167–174. Accessed: Dec. 18, 2024. [Online]. Available: https://ebooks.iospress.nl/volumearticle/58051

[5] A. Acharya and D. Sinha, “Early prediction of students performance using machine learning techniques,” *Int. J. Comput. Appl.*, vol. 107, no. 1, pp. 37–43, 2014.

[6] A. Asselman, M. Khaldi, and S. Aammou, “Enhancing the prediction of student performance based on the machine learning XGBoost algorithm,” *Interact. Learn. Environ.*, vol. 31, no. 6, pp. 3360–3379, Aug. 2023, doi: 10.1080/10494820.2021.1928235.

[7] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, “Classification and prediction of student performance data using various machine learning algorithms,” *Mater. Today Proc.*, vol. 80, pp. 3782–3785, 2023.

[8] Q. Zhou, W. Quan, Y. Zhong, W. Xiao, C. Mou, and Y. Wang, “Predicting high-risk students using Internet access logs,” *Knowl. Inf. Syst.*, vol. 55, pp. 393–413, 2018.

[9] E. Alhazmi and A. Sheneamer, “Early predicting of students performance in higher education,” *IEEE Access*, vol. 11, pp. 27579–27589, 2023.

[10] F. Qiu *et al.*, "Predicting students' performance in e-learning using learning process and behaviour data," *Sci. Rep.*, vol. 12, no. 1, p. 453, 2022.

[11] Y. Luo, X. Han, and C. Zhang, "Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses," *Asia Pac. Educ. Rev.*, vol. 25, no. 2, pp. 267–285, 2024.

[12] M. T. Sathe and A. C. Adamuthe, "Comparative study of supervised algorithms for prediction of students' performance," *Int. J. Mod. Educ. Comput. Sci.*, vol. 13, no. 1, p. 1, 2021.

[13] M. Adnan *et al.*, "Predicting at-risk students at different percentages of course length for early intervention using machine learning models," *Ieee Access*, vol. 9, pp. 7519–7539, 2021.

[14] C. Huang, J. Zhou, J. Chen, J. Yang, K. Clawson, and Y. Peng, "A feature weighted support vector machine and artificial neural network algorithm for academic course performance prediction," *Neural Comput. Appl.*, vol. 35, no. 16, pp. 11517–11529, 2023.

[15] B. Van Liebergen, "Machine learning: a revolution in risk management and compliance?," *J. Financ. Transform.*, vol. 45, pp. 60–67, 2017.

[16] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.

[17] Z. Wan, X. Xia, D. Lo, and G. C. Murphy, "How does machine learning change software development practices?," *IEEE Trans. Softw. Eng.*, vol. 47, no. 9, pp. 1857–1871, 2019.

[18] S. Suthaharan, "Support Vector Machine: Machine Learning Models and Algorithms for Big Data Classfication," *Integr. Ser. Inf. Syst.*, vol. 36, 2016.

[19] S. Benatti, B. Milosevic, E. Farella, E. Gruppioni, and L. Benini, "A prosthetic hand body area controller based on efficient pattern recognition control strategies," *Sensors*, vol. 17, no. 4, p. 869, 2017.

[20] A. Iftikhar *et al.*, "Risk classification in global software development using a machine learning approach: A result comparison of support vector machine and K-nearest neighbor algorithms," *J. Inf. Technol. Res. JITR*, vol. 15, no. 1, pp. 1–21, 2022.

[21] S. Jafarzadeh-Ghoushchi, "Optimization of transportation system based on combined model using artificial neural networks and response surface methodology," *Int. J. Tech. Res. Appl.*, vol. 23, pp. 69–76, 2015.

[22] T. F. Awolusi, O. L. Oke, O. O. Akinkurolere, A. O. Sojobi, and O. G. Aluko, "Performance comparison of neural network training algorithms in the modeling properties of steel fiber reinforced concrete," *Heliyon*, vol. 5, no. 1, 2019, Accessed: Dec. 19, 2024. [Online]. Available: https://www.cell.com/heliyon/fulltext/S2405-8440(18)34603-6

[23] F. Bre, J. M. Gimenez, and V. D. Fachinotti, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks," *Energy Build.*, vol. 158, pp. 1429–1441, 2018.