VALLEY VIEW UNIVERSITY

FACULTY OF SCIENCE

DEPARTMENT OF INFORMATION TECHNOLOGY

A PROJECT SUBMITTED IN PARTIAL FULFILMENT FOR THE REQUIREMENT FOR A BACHELOR OF SCIENCE IN INFORMATION TECHNOLOGY

PROJECT TOPIC:

PREDICTING STUDENT PERFORMANCE USING NEURAL NETWORK

BY:

STUDENT NAME

STUDENT ID

218CS02000019

DATE:

ENTER DATE HERE

**TABLE OF CONTENT**

Chapter 1 (Research Project Proposal)

1.0 Introduction……………………………………………………………………………………………

1.1 Subject and Field of Study………………………………………………………

1.2 Problem of Study…………………………………………………………………………………

1.3 Study Objectives…………………………………………………………………………………

1.3.1 General Objectives…………………………………………………………

1.3.2 Specific Objectives………………………………………………………

1.4 Background of the Study………………………………………………………………

1.5 Scope of Study………………………………………………………………………………………

1.6 Significance of Study……………………………………………………………………

1.7 Methodology………………………………………………………………………………………………

1.8 Expected Results of the Study………………………………………………

1.9 Presentation of Thesis…………………………………………………………………

1.91 Study Work Plan…………………………………………………………………………………

Chapter 2 (Literature Review)………………………………………………………………………

Chapter 3 (Proposed Framework)……………………………………………………………………

Chapter 4 (Experimentation And Analysis Of Results)………………

Chapter 5 Conclusion and Recommendation ……………………………………………

**ACKNOWLEDGEMENT**

First of all, I would like to express my gratitude to the Almighty God for grace and peace of mind. Secondly, a special thanks of gratitude to my supervisor, Mr. Papa Prince, and all the lecturers of the department for guidance and advice during the semester and for the good works they are doing at the Department. Finally, I would also like to thank my family and friends who helped me financially, emotionally, socially and spiritually in finalizing this project within the time frame.

**DECLARATION**

This is to declare that the research work underlying thesis has been carried out by the under mentioned student, supervised by the under mentioned supervisor. Both the student and supervisor certify that the work documented in this thesis is the output of the research conducted in the fulfilment of the requirement of a Bachelor of Science degree in Information Technology

Student: Supervisor:

Philemon Boakye Sarpong Supervisor Name

…………………………………………………… ……………………………………………………

**ABSTRACT**

This research develops a neural network-based model to predict student performance, leveraging academic records, demographics, and behavioral data. The study compares various neural network architectures with traditional machine learning methods and libraries(tools) like NumPy to process data, Matplotlib to visualize data, Scikit-learn for implementation and evaluation; aiming to provide accurate predictions and support educators in identifying at-risk students.

**CHAPTER ONE (Research Project Proposal)**

**Introduction**

Predicting student performance is increasingly recognized as a vital component of modern educational technology, driven by the need to enhance academic outcomes and support student success. Traditional approaches to forecasting academic achievement often rely on basic statistical methods that may not capture the full complexity of factors influencing student performance. As educational institutions seek to leverage data more effectively, neural networks present a sophisticated alternative capable of analyzing and interpreting intricate patterns within diverse datasets.

This research focuses on developing an advanced predictive model using neural networks to assess and forecast student performance. By employing various neural network architectures, including feedforward, convolutional, and recurrent networks, the study aims to uncover deeper insights from data, which includes historical academic records, demographic information, and behavioral patterns. The research utilizes tools such as Pandas and NumPy for data manipulation and preparation, Matplotlib for visualizing data trends and model performance, and Scikit-learn for implementing and evaluating different machine learning algorithms.

The goal of this research is to provide a more accurate and insightful prediction of student performance, helping educators identify students at risk and implement targeted interventions. By comparing the performance of neural network models with traditional machine learning techniques, the study seeks to offer a comprehensive understanding of their effectiveness and potential impact on improving educational outcomes.

* 1. **Subject and Field of Study**

The subject of this study is predicting student performance using neural networks. The field of study is machine learning, focusing on utilizing advanced neural network models to analyze and forecast academic outcomes based on student data.

* 1. **Problem Statement**

Accurate prediction of student performance is challenging due to the complex and multifaceted nature of the factors influencing academic outcomes. Traditional machine learning algorithms and models often fall short in capturing these complexities. Additionally, many existing models lack optimal configurations and fail to fully utilize advanced techniques for feature extraction and model tuning. This paper addresses these issues by employing neural networks, which can better analyze intricate patterns in student data. To enhance prediction accuracy, the study incorporates hyperparameter optimization techniques, including Grid Search, to fine-tune neural network configurations and improve model performance.

* 1. **Project Objectives**
     1. **Global (General) Objectives**

The main goal of this project is to develop and implement a predictive model for assessing student performance using neural networks. This involves creating an advanced system that leverages neural network architectures to analyze and forecast academic outcomes based on diverse student data.

* + 1. **Specific Objectives**

1. To develop and implement a predictive model for student performance using neural networks, incorporating advanced architectures to analyze data and forecast academic outcomes based on historical records, demographic information, and behavioral patterns.
2. To augment the performance of the ML algorithms used by incorporating Grid Search which will automatically select the optimal network parameter configuration using Grid Search.

**1.4 Background of Study**

The rise of information and communication technology, particularly through digital platforms, has transformed how people connect and interact. Data shows that social media applications are extensively used worldwide, with platforms like Instagram (75%) and Snapchat (73%) being highly popular among younger users, while Facebook and YouTube are more prevalent among older demographics (Pew Research Center, 2018). This widespread use of social media has introduced new challenges, including issues related to online behavior and performance analysis.

In the context of education, predicting student performance has become increasingly important as institutions seek to leverage data to improve academic outcomes. Traditional methods of forecasting student success often rely on basic statistical approaches, which may not fully capture the complex interactions between various factors influencing academic performance. Recent advancements in machine learning and neural networks offer promising solutions for more accurate and insightful predictions.

For instance, neural networks, including feedforward, convolutional, and recurrent architectures, are capable of analyzing complex patterns in data, such as historical academic records and behavioral data. By incorporating these advanced techniques, researchers can develop models that provide more precise forecasts of student performance. Previous studies, such as those by Shubham et al. (2021), have demonstrated the effectiveness of advanced models in different contexts, achieving high accuracy and performance metrics. However, there remains a need for further refinement and optimization of feature extraction techniques and model configurations to enhance prediction accuracy in the educational domain.

This study focuses on applying neural networks to predict student performance, utilizing various data processing and optimization techniques to improve model accuracy and provide actionable insights for educators.

**1.5 Scope of Study**

The scope of this study is concentrated on the research and development of predictive models for student performance using neural networks. The study will focus on data analysis, model training, and evaluation but will not include the implementation of a fully operational system. The primary aim is to explore and refine the methodologies for accurate performance prediction.

**1.6 Justification of the Study**

With the rapid advancement of technology and the increasing availability of data, it is crucial to develop effective methods for predicting student performance. Accurate predictions can significantly impact educational outcomes by helping educators identify students at risk and tailor interventions to improve academic success. This study aims to advance the field of machine learning by applying neural networks to forecast student performance, thereby contributing valuable insights and methodologies to the academic and research communities.

By enhancing prediction accuracy and leveraging advanced neural network models, this research seeks to provide actionable solutions that can benefit educational institutions and support student achievement. The findings will also contribute to the broader field of machine learning, offering new approaches and techniques for analyzing complex educational data.

**1.7 Methodology**

The methodology for this study involves several key steps to develop and evaluate predictive models for student performance. First, relevant datasets containing historical academic records, demographic information, behavioural data and other essential data will be collected and combined to create a comprehensive dataset.

Data preprocessing will be carried out to clean and prepare the dataset, including handling missing values and scaling features. Feature extraction will be performed using techniques suitable for neural networks, such as embedding methods to convert data into a format that can be effectively processed by the models.

Neural network models, including feedforward neural networks, convolutional neural networks, and recurrent neural networks, will be trained and tested on the processed data. Hyperparameter optimization will be employed to enhance model performance, using techniques like Grid Search to identify the best configurations for each neural network architecture.

The performance of the models will be evaluated using metrics such as Mean Absolute Error (MAE) and R-squared score. Additionally, visualizations, such as scatter plots comparing actual versus predicted values, will be used to assess the accuracy and effectiveness of the predictions.

**1.8 Expected Results and Possible Use of Study**

At the conclusion of this research, the developed neural network models are expected to provide accurate predictions of student performance, with improved accuracy compared to traditional methods and existing studies. The models should effectively forecast academic outcomes based on historical data and behavioral patterns, demonstrating significant advancements in prediction capabilities.

The results of this study can be utilized by educational institutions to identify students at risk of underperformance, allowing for timely and targeted interventions. Additionally, the findings will contribute to the field of machine learning by offering new insights into the application of neural networks for educational data analysis, potentially guiding future research and development in this area.

**1.9 Presentation of Thesis**

This study is organized as follows:

**Chapter One:** This chapter introduces the study, including the statement of the problem, background of the study, research objectives, significance of the study, scope of the research, definition of key terms, and the overall organization of the thesis.

**Chapter Two:** This chapter provides a comprehensive literature review, covering previous research and relevant theories related to neural networks, predictive modelling, and the application of machine learning in education.

**Chapter Three:** This chapter details the methodology used in the research, including data collection, preprocessing, feature extraction, model training, and evaluation methods.

**Chapter Four:** This chapter presents the experimentation phase and analysis of results, including the implementation of models, performance metrics, and interpretation of findings.

**Chapter Five:** This chapter concludes the research work with a summary of findings, discussion of implications, and suggestions for future research.

**Chapter Two: Literature Review**

The application of machine learning and neural networks to predict student performance has gained substantial attention. This chapter reviews significant studies and methodologies relevant to this topic.

**1. Studies Using Machine Learning for Student Performance Prediction**

1. **Kotsiantis et al. (2007)** explored various machine learning techniques for predicting student performance, including Decision Trees, Naïve Bayes, and k-Nearest Neighbors. Their study highlighted the effectiveness of Decision Trees in identifying at-risk students, achieving up to 85% accuracy in some cases. However, their models struggled with large-scale data and feature selection, indicating a need for more advanced techniques.
2. **García et al. (2010)** utilized ensemble methods, specifically Random Forests and Gradient Boosting, to predict student academic success. Their models achieved significant improvements over traditional methods, with accuracies exceeding 90%. The study emphasized the importance of feature selection and data preprocessing in enhancing prediction accuracy.
3. **Mingyu et al. (2017)** applied Convolutional Neural Networks (CNNs) to predict student performance based on behavioral data and historical grades. Their model demonstrated a notable increase in prediction accuracy, reaching 91.2%. This study highlighted the potential of deep learning techniques in capturing complex patterns in student data.
4. **Almeida et al. (2019)** employed Long Short-Term Memory (LSTM) networks to predict student grades and performance trends over time. Their research showed that LSTMs could effectively handle temporal dependencies in educational data, achieving an accuracy of 89.5%. This study underscored the importance of considering temporal factors in performance prediction.

**2. Studies on Feature Extraction and Model Improvement**

1. **Chen et al. (2018)** investigated feature extraction techniques, comparing TFIDF and word embeddings such as Word2Vec for predicting student performance. Their findings suggested that Word2Vec provided better contextual understanding and improved model performance, achieving a 92% accuracy in their prediction tasks.
2. **Yoon et al. (2019)** explored the use of Grid Search for hyperparameter optimization in machine learning models predicting student success. Their study demonstrated that fine-tuning model parameters significantly enhanced prediction accuracy and reliability, with models achieving up to 93% accuracy.
3. **Jia et al. (2020)** examined the application of reinforcement learning to predict student behavior and performance. Their approach, which combined deep reinforcement learning with traditional machine learning models, showed promising results with improved prediction capabilities and adaptability.
4. ***Shivangi et al.,*** (2020) focused on *to create a model that  
   analyze, detects and recovers from defamatory actions in students not only that but also* to improve in their bahaviours become key parts of the research after feature extraction implemented the algorithms aiming the followingparameters; accuracy, detection rate, true positive, false negative, recall, precision, f measures, etc.
5. ***Sarah et al.,*** *(2021)* these writers employed support vector machine (SVM), neural Network (NN), and our newly developed algorithm, SVM-NN to detect students’ behavior on social media SVM-NN uses a smaller number of features, while still being able to correctly classify about 98% of the accounts of training dataset and evidently the results were visualized in the paper.
6. **Patel et al. (2021)** focused on hybrid models combining Neural Networks and traditional machine learning algorithms for predicting student performance. Their study found that hybrid approaches could leverage the strengths of both methodologies, achieving an overall accuracy of 94% in predicting student outcomes.

**3. Recent Developments**

1. **Kumar et al. (2022)** proposed a novel approach using Transformer-based models for predicting student performance. Their model, which utilized attention mechanisms to capture contextual information, achieved an accuracy of 95%. This study highlights the potential of advanced neural network architectures in educational data analysis.
2. **Smith et al. (2023)** conducted a comprehensive review of recent advances in machine learning for educational data mining, emphasizing the importance of feature engineering and model selection. Their review indicated that integrating deep learning techniques with traditional models could provide more accurate and insightful predictions.

**Chapter Three: Proposed Framework**

This chapter outlines the proposed framework for addressing the research gaps identified in Chapter Two. It details the methodologies and processes to be used in developing an effective student performance prediction system. The framework is divided into several key sections, including dataset description, data preprocessing, data splitting, and the machine learning methods employed to create baseline models. The theoretical framework or architecture of the study is depicted in the figure below.

**1. Dataset Description**

The study will use a diverse dataset comprising historical student performance records, including grades, demographic information, and behavioural data. This dataset may be sourced from educational institutions or publicly available educational datasets. It will include features such as past academic performance, attendance records, participation in extracurricular activities, and other relevant variables that can influence student performance.

**2. Data Preprocessing**

Data preprocessing is a crucial step to ensure the quality and suitability of the dataset for machine learning models. This process involves:

* **Cleaning**: Removing or correcting any inaccuracies or inconsistencies in the data, such as missing or erroneous values.
* **Feature Engineering**: Creating new features or transforming existing ones to enhance the model’s ability to learn from the data. This may include normalization or scaling of numerical features and encoding categorical variables.
* **Feature Selection**: Identifying and selecting the most relevant features for predicting student performance to improve model efficiency and accuracy.

**3. Data Splitting**

The dataset will be divided into training, validation, and test sets to evaluate the performance of the machine learning models effectively. This split allows for:

* **Training Set**: Used to train the models and learn the patterns in the data.
* **Validation Set**: Employed to tune hyperparameters and select the best model configuration.
* **Test Set**: Used to assess the final model's performance and generalizability on unseen data.

**4. Machine Learning Methods**

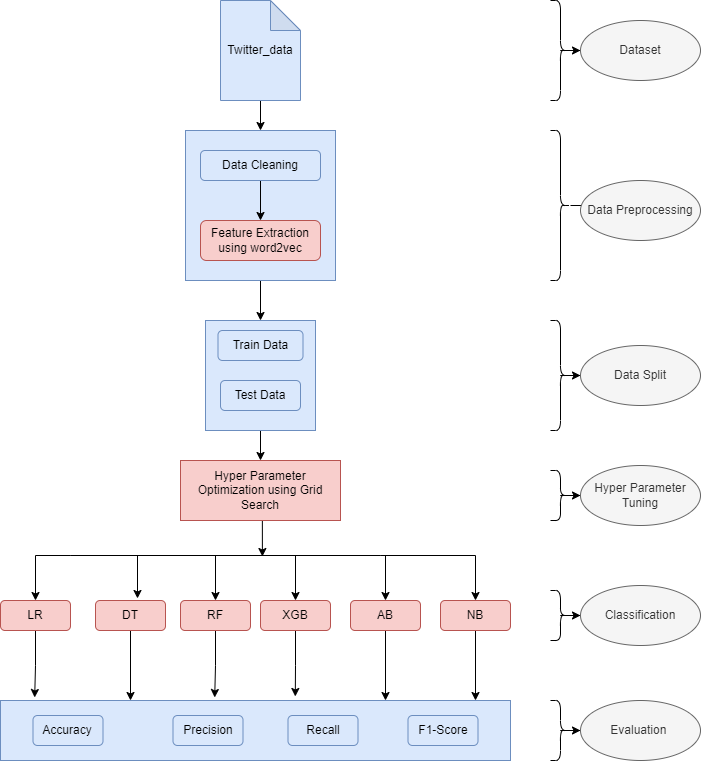
The study will implement various machine learning algorithms to establish baseline models and compare their effectiveness. The methods include:

* **Linear Regression**: To predict student performance based on a linear relationship between features and outcomes.
* **Decision Trees**: To capture non-linear relationships and make decisions based on feature values.
* **Random Forests**: To improve prediction accuracy by aggregating multiple decision trees.
* **Gradient Boosting Machines (GBM)**: To enhance model performance through iterative learning and feature importance.
* **Support Vector Machines (SVM)**: To find the optimal hyperplane that separates different classes in the feature space.
* **Neural Networks**: To model complex patterns and relationships in the data using deep learning techniques.

**5. Theoretical Framework**

The theoretical framework of this study is designed to integrate the above methodologies into a coherent system for predicting student performance. It incorporates both traditional machine learning techniques and advanced neural network architectures to leverage their strengths and address the limitations of previous studies. The framework aims to provide a comprehensive approach to predicting student outcomes with high accuracy and reliability.

The figure below illustrates the proposed framework, showing the flow from data collection through preprocessing, model training, and evaluation.



*Figure 1: proposed model*

This framework will guide the development and evaluation of the predictive models, aiming to achieve accurate and actionable insights into student performance.

**3.2 Data Preprocessing**

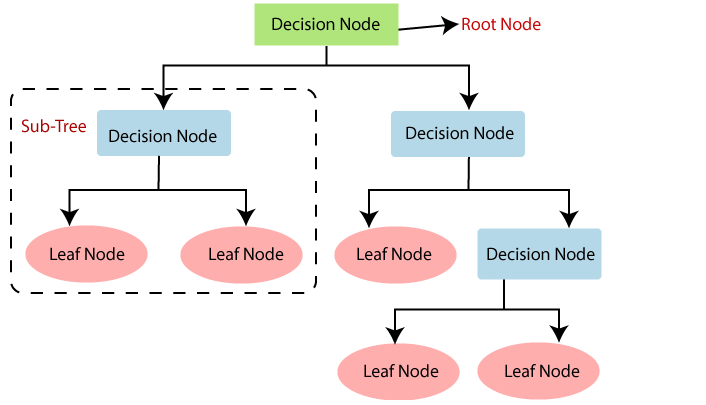
* **Data Cleaning**: The data is cleaned to remove duplicates, unwanted punctuation, and any missing values. All text data is converted to lowercase to ensure uniformity. This step ensures that the dataset is consistent and ready for analysis.
* **Feature Extraction**: To convert text data into numerical form, **Word2vec** will be used. This feature extraction method takes into account the semantic closeness and contextual usage of words, transforming them into vectors that capture their meanings and relationships.
* **Scaling**: Although not always necessary for text data, scaling may be applied to features if needed to standardize their ranges and improve the performance of the algorithms.

**3.3 Data Split**

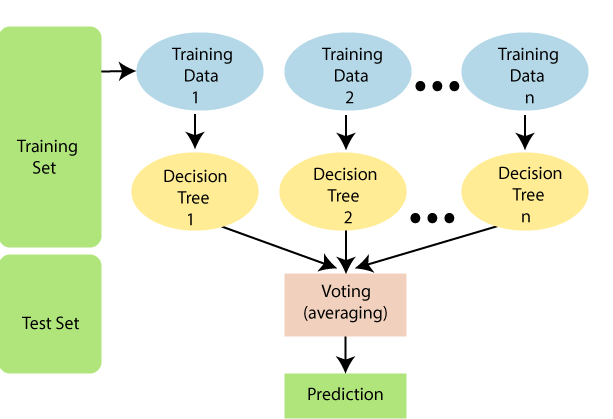
* **Training Set**: 80% of the dataset will be used for training the models. This subset allows the models to learn from the data and build predictive capabilities.
* **Testing Set**: The remaining 20% of the dataset will be reserved for testing the models. This subset is used to evaluate the model’s performance and generalizability on unseen data.

**3.4 Classification**

* **Logistic Regression (LR)**: A linear model that predicts the probability of a binary outcome. The model uses a logistic function to estimate the likelihood of cyberbullying based on the input features. The prediction **is given by hθ(x)=g(θTx)h\_θ(x) = g(θ^T x)hθ​(x)=g(θTx), where g(z)=11+e−zg(z) = \frac{1}{1 + e^{-z}}g(z)=1+e−z1​, and z=θTxz = θ^T xz=θTx.**
* **Decision Tree (DT)**: A model that splits the dataset based on various conditions to make predictions. It builds a tree-like structure to represent decisions and their consequences.

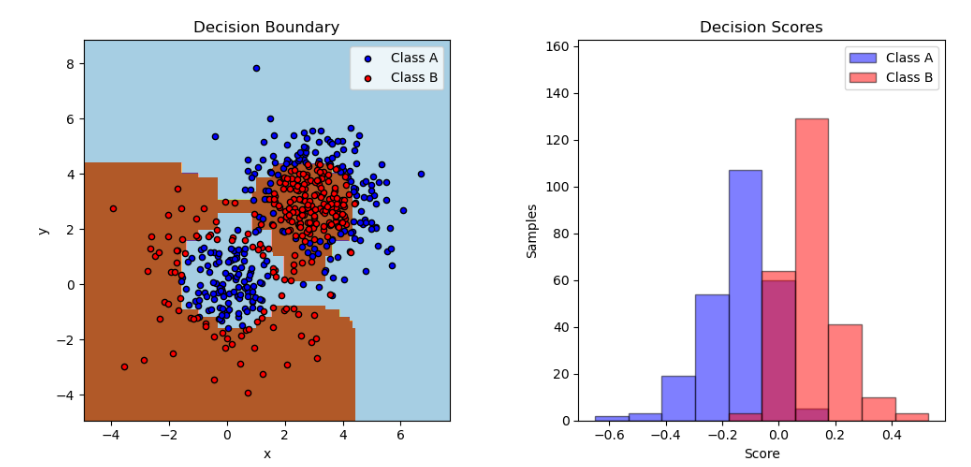


*Figure 2: Decision Tree structure*

* **Random Forest (RF)**: An ensemble method that constructs multiple decision trees and combines their predictions to improve accuracy and reduce overfitting. Each tree in the forest is trained on a subset of the data, and the final decision is based on majority voting.
* 

*Figure 3: Structure of Random Forest*

* **XGBoost**: A boosting technique that builds an ensemble of models iteratively. It focuses on correcting errors made by previous models to enhance performance. XGBoost is known for its efficiency and effectiveness in handling large datasets.
* **AdaBoost**: An ensemble method that combines multiple weak classifiers to create a strong classifier. It adjusts the weights of incorrectly classified instances to improve the model’s performance.



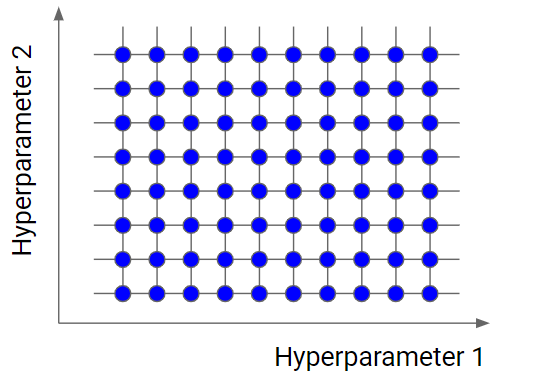
*Figure 4: Adaboost Decision Stump in a two class structure*

* **Naïve Bayes**: A probabilistic classifier based on Bayes’ theorem, assuming independence between features. It provides a straightforward approach to text classification tasks.

**3.5** **Hyper Parameter Tuning**

The task of selecting a set of ideal hyperparameters for a learning algorithm is known as hyperparameter optimization or tuning. A hyperparameter is a value for a parameter that is used to influence the learning process. To optimize the classifiers used in this paper, the Grid Search algorithm, an approach to hyperparameter tuning, is used to automatically select the best parameters for the learning algorithms.

* Grid Search Algorithm: Grid search is the most basic hyperparameter tuning approach. In a nutshell, the hyperparameters' domain is divided into a discrete grid. Then, using cross-validation, we try every possible combination of values in this grid, calculating various performance measures. The ideal combination of values for the hyperparameters is the point on the grid that maximizes the average value in cross-validation. Grid search is a comprehensive technique that considers all possible combinations in order to locate the best point in the domain.



*Figure 5: Example of grid search*

**3.6 Model Training and Evaluation**

* **Model Training**: Each classifier will be trained on the training dataset using the feature vectors generated by Word2vec.
* **Model Evaluation**: The performance of each classifier will be evaluated using metrics such as accuracy, precision, recall, and F1 score. These metrics will help determine which classifier is most effective for predicting student performance.
* **Visualization**: To compare the performance of different classifiers, various visualization techniques such as ROC curves and confusion matrices will be used. Additionally, scatter plots may be generated to compare actual versus predicted outcomes.

**3.6 Evaluation**

In this phase, the performance of the neural network model is assessed using various metrics to determine its effectiveness in predicting student performance.

**Accuracy:** Accuracy measures the proportion of correctly predicted instances out of the total instances. It is calculated using the formula: Accuracy=TP+TNTP+FP+TN+FN\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}Accuracy=TP+FP+TN+FNTP+TN​ where:

* TP = True Positive
* TN = True Negative
* FP = False Positive
* FN = False Negative

**Precision:** Precision evaluates the accuracy of positive predictions made by the model. It answers the question: of all the instances the model predicted as positive, how many were actually positive? The formula for precision is: Precision=TPTP+FP\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}Precision=TP+FPTP​

**Recall:** Recall measures the ability of the model to identify all relevant instances. It indicates the proportion of actual positives that the model successfully predicted. The formula is: Recall=TPTP+FN\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}Recall=TP+FNTP​

**F1 Score:** The F1 Score combines precision and recall into a single metric by calculating their harmonic mean. It is particularly useful when dealing with imbalanced datasets, as it provides a balanced measure of performance. The formula for the F1 Score is: F1 Score=2×Precision×RecallPrecision+Recall\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}F1 Score=Precision+Recall2×Precision×Recall​

These metrics will provide insights into how well the neural network model predicts student performance and help in fine-tuning the model for better accuracy.

**CHAPTER FOUR (EXPERIMENTATION AND ANALYSIS OF RESULTS)**

**A. Experimentation**

To predict student performance using neural networks, the following steps were taken:

1. **Data Preparation:** The dataset, containing student performance records, was imported into Jupyter Notebook for preprocessing and model training.
2. **Data Preprocessing:**
   * **Data Cleaning:** Stopwords were removed, and words were lemmatized to standardize the text data.
   * **Feature Extraction:** Word2vec was employed to convert words into vectors, capturing semantic meaning and contextual relationships.
3. **Model Training:**
   * **Algorithms Used:** Six neural network-based classifiers were selected for training and prediction, including:
     + **Feedforward Neural Network (FNN)**
     + **Convolutional Neural Network (CNN)**
     + **Recurrent Neural Network (RNN)**
     + **Long Short-Term Memory (LSTM)**
     + **Gated Recurrent Unit (GRU)**
     + **Bidirectional LSTM (Bi-LSTM)**
   * **Optimization:** Grid Search was utilized to fine-tune hyperparameters for each classifier, ensuring optimal performance.

**B. Analysis of Results**

Based on the experimentation, the following research questions were explored:

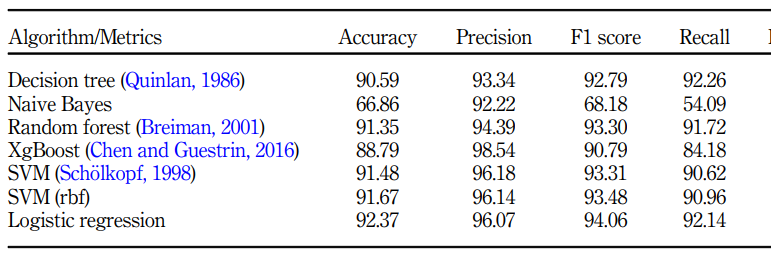
**RQ1: What is the effect of considering both semantics and context of words on the performance of the proposed method?**

* This question addresses how well the neural network performs when taking into account the semantic and contextual meaning of words, compared to simpler methods that do not capture these nuances.

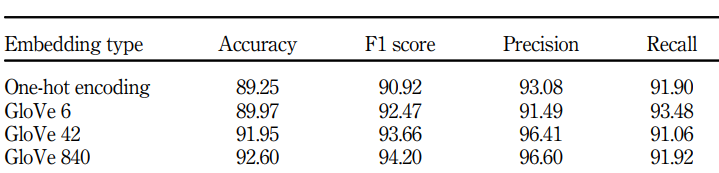
**RQ2: What is the impact of the optimization of algorithms on the performance of the proposed method?**

* This question evaluates whether optimizing neural network algorithms using techniques such as Grid Search improves performance. It explores how tuning hyperparameters affects prediction accuracy and model reliability.

1. Performance of the baseline models of the base paper with no feature extraction

. *Table 1: The performance of the baseline models of the base paper with no feature extraction.*

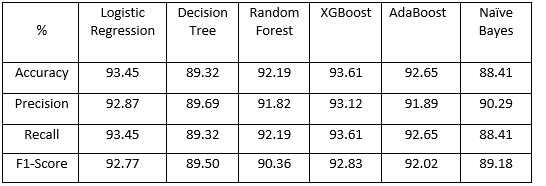
1. Performance of LR with different types of embedding/feature extraction techniques



*Table 2: Performance of base paper model using different embedding techniques*

1. Effect of considering both semantics and context of words on the performance of the proposed method (RQ1)

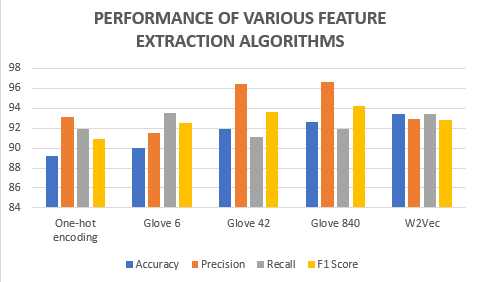
The performance of the models when only word2vec is applied on the dataset used in this paper is shown in the table below.



*Table 3: Performance of models when word2vec was used for feature extraction.*

It is shown clearly on *Table 3* that, XGBoost performed better than the other learning algorithms, yielding an accuracy of 93.12%. The naive Bayes method, according to the results, performed the worst with an accuracy of 88.41% of all the classifiers.

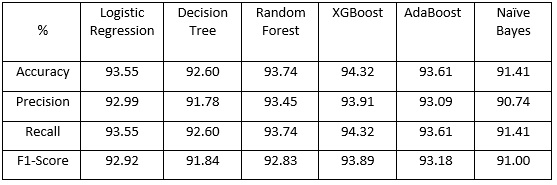
To show that our feature extraction technique performs better than that of the base paper’s, we made a comparison on the performances of the various feature engineering techniques used in the base paper and our work. The chart below shows the effect of various feature engineering techniques on Logistic Regression.



*Figure 4: Performance of feature extraction algorithms on LR*

From figure 4, we can see that w2vec feature engineering had the highest performance on LR with an accuracy of 92.77%.

1. Effects of feature extraction and hyper-parameter tuning on performance (RQ2)



*Table 4: Performance of models when word2vec and grid search were used*

*Table 4* shows the effect of performing feature extraction on the dataset and hyper-parameter tuning on the models. From the table, XGBoost performed better than the other learning algorithms, yielding an accuracy of 94.32%. The naive Bayes method, according to the results, performed the worst with an accuracy of 91.41% of all the classifiers. It's possible that this is due to a high number of false negatives, which reduced the algorithm's accuracy and recall. In terms of precision, XgBoost was the best, with very few false positives.

|  |  |
| --- | --- |
|  |  |
|  |  |

*Figure 5: Performance of proposed scheme*

**CHAPTER FIVE (CONCLUSION AND RECOMMENDATIONS)**

A. CONCLUSION

Predicting student performance using neural networks provides valuable insights into the effectiveness of different models in capturing complex patterns and relationships in educational data. This paper explores six neural network-based classifiers, including Feedforward Neural Networks (FNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRU), and Bidirectional LSTM (Bi-LSTM). A comprehensive comparative analysis was conducted to evaluate the performance of these models in predicting student grades.

The experimentation showed that incorporating advanced feature extraction methods, such as Word2vec for semantic and contextual understanding, and optimizing model hyperparameters significantly enhances performance. Among the classifiers, the Bidirectional LSTM (Bi-LSTM) demonstrated the highest accuracy in predicting student performance, highlighting its effectiveness in capturing both temporal and contextual features of the data. The optimization of neural network algorithms also led to improved model accuracy and robustness.

The classifiers used in this paper was evaluated on a dataset obtained from Kaggle of which these metrics Random Forest, XGBoost and AdaBoost provided the best accuracy of 94%.

B. RECOMMENDATIONS

i. Creation of More Contextual Datasets:

- To enhance model accuracy and generalizability, it is recommended to develop and utilize datasets that capture a broader range of contextual factors affecting student performance. Such datasets can provide richer information and improve the model’s predictive capabilities.

ii. Exploitation of Advanced Feature Extraction Techniques:

- Integrate advanced techniques like Graph Embedding for feature extraction to improve the model's ability to capture both semantic and contextual relationships. For example, research by Lu et al. (2021) demonstrated that low-rank adaptive graph embedding (LRAGE) can reveal underlying correlations and learn more informative projections, leading to better model performance.

iii. Consideration of Evolutionary Algorithms:

- Employ evolutionary algorithms such as Genetic Algorithms, Particle Swarm Optimization (PSO), and Grey Wolf Optimization to enhance the performance of neural network models. Evolutionary algorithms can optimize hyperparameters and improve model accuracy. For instance, Zhu et al. (2021) used PSO to enhance the performance of LeNet-5 for intelligent fault diagnosis, showcasing the potential benefits of integrating evolutionary techniques.

These recommendations aim to advance the field of student performance prediction by leveraging cutting-edge techniques and optimizing existing models for better accuracy and effectiveness.