Spring 2023 Research Literature and Techniques

(CSCI-595-01B-01W)

TITTLE: STUDENT LEARNING MANAGEMENT SYSTEM

Under the Supervision and Guidance of

Professor: Dr. Sang Suh.

Team Members:(Team No – 4)

S.NO	MEMBER NAME	CWID	CONTACT NO	EMAIL
1	Rukmini Sai Mohan Jayanth Srikantam	50298996	----	----

TABLE OF CONTENTS

ABSTRACT

I. INTRODUCTION 4

II. LITERATURE SURVEY 5

III. PROPOSED METHODOLOGY: 7

DATASET 10

IV.IMPLEMENTATION 12

V. RESULTS 12

VI. FUTURE WORK AND DISCUSSION 19

VII. REFERENCES 19

Abstract

Learning Management Systems (LMSs) are increasingly utilized for the administration, tracking, and reporting of educational activities. One such widely used LMS in higher education institutions around the world is Blackboard. This is due to its capabilities of aligning items of learning content, student-student and student-teacher interactions, and assessment tasks to specified goals and student learning outcomes. This study aimed to determine how certain **Key Performance Indicators (KPIs)** based on student interactions with Blackboard helped to forecast the learning outcomes of students. A mixed-methods study design was used which included analysis of four deep learning models for predicting student performance. They were analyzed using analysis approach to establish possible predictive KPIs associated with the electronic Blackboard report. Correlational analyses were performed to examine the extent to which these factors are linearly correlated with the performance indicators of students. This study suggests using the Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm to create a Student Learning Management System (SLMS). The approach enhances student achievement, gives each student individualized feedback, and helps students and teachers communicate. The suggested method makes use of machine learning techniques to give students personalized feedback and guidance, find out where they need more assistance, and raise their performance levels. The system also makes it easier for students and teachers to communicate, and it offers a simple way to access assignments and course materials. Results from experiments show how well the suggested strategy works to raise student achievement and offer individualized learning opportunities. The system has the power to transform how pupils learn and improve the standard of instruction. To understand the temporal dependencies in student data and forecast their performance based on their interactions with the system, researchers used the Bi-LSTM algorithm. The system also gives users access to communication tools, assignments, and course materials through a user-friendly interface.

I. Introduction

A software program called a **student learning management system (SLMS)** is created to assist or help teachers and students in controlling the learning process [1]. Teachers may manage assignments, assess student work, and give comments with the aid of the SLMS. However, it also makes it easier for students to access course materials, turn in assignments, get grades, and feedback, and interact with peers and teachers. Natural language processing (NLP) methods have recently attracted more attention to analyze student data and enhance SLMS. The Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm, a kind of recurrent neural network (RNN) that can deal with sequential data such as text, is one such method.

Based on their interactions with the SLMS, Bi-LSTM models can be used to forecast student performance. For instance, the algorithm can assess the written work of students and forecast their marks [7]. The programmed may also evaluate the language used in online forums and inform teachers of the problems their pupils are having. An SLMS that makes use of the Bi-LSTM model has the potential to enhance students' learning experiences by offering individualized feedback and support. The algorithm can pinpoint places where students need more assistance and offer tailored interventions to improve their comprehension of the subject matter by examining student data.

The **LMS ("Learning management system")** is one kind of software tool that is utilized in several educational institutions and organizations to deliver and provide online courses. The data of LMS can be a beneficial source of information to increase the education quality. The patterns can be easily identified which are useful to predict the outcomes of student learning by analyzing and investigating LMS data. machine and deep learning algorithms are available, but their prediction accuracy is not good enough. To increase prediction accuracy, we tried to combine or integrate multiple deep learning algorithms and then experiment with '**EDA-Student Academic Performance Dataset'**.As base algorithm the implementation project has chosen Convolution Neural Network (CNN) which is known for best features extraction or optimization algorithm and its proven performance in various fields such as image classification, medical disease prediction and many more. The initial modules will train CNN on students' performance dataset and then extract optimized features from CNN and then retrain with LSTM (Long Short-Term Memory) and BI-LSTM (Bidirectional – LSTM). After training, LSTM was shown not to increase accuracy, but BI-LSTM managed to get accuracy between 95 to 98%. BI-LSTM accuracy is higher than that of LSTM and CNN. LSTM will filter features in forward direction and never look backward to re-filter dropout features so its accuracy may not improve [5]. BI-LSTM will look both forward and backward directions for features filtration which helps in obtaining more enhanced features which result in better accuracy. filter dropout features so its accuracy may not improve.

The aim of creating a Student Learning Management System (SLMS) based on the Bi-LSTM model is to enhance or improve student learning and promote efficient communication between students and teachers.

Objectives:

- Enhance student performance: By offering personalized feedback and support, the SLMS seeks to enhance student performance. The Bi-LSTM model can detect student needs through analysis of student data and deliver customized interventions to assist students in better understanding the topic [8].
- Improved communication between students and teachers is the goal of the SLMS. Students may readily access course materials, turn in assignments, and interact with teachers and peers using the SLMS.
- Efficiency improvement: The SLMS tries to make the learning process more effective. Teachers can save time and concentrate on giving students more individualized support by automating some processes, such as grading and feedback.
- Give information: The SLMS aims to give information about student performance and behavior. The Bi-LSTM model can find trends in student data and give teachers information about potential problem areas for their students [9].

II. Literature Survey

To analyze student data and enhance the Student Learning Management System (SLMS), there has been an expansion in interest in applying Natural Language Processing (NLP) methods in educational research, including the Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm. Bi-LSTM models have been used in SLMSs in research, each with a different strategy and set of goals.

The application of Natural Language Processing (NLP) methods in educational research has gained significant interest in recent years, particularly in Student Learning Management Systems (SLMSs) [1]. One popular NLP algorithm used in SLMSs is the Bidirectional Long Short-Term Memory (Bi-LSTM) model. Several studies have investigated the use of Bi-LSTM models in SLMSs, each with different strategies and objectives. One study aimed to develop an SLMS that uses a Bi-LSTM model to predict student performance based on their interactions with the system [2]. The study demonstrated that the Bi-LSTM model could accurately predict student grades with over 80% accuracy using a dataset of student interactions with an online platform. The methodology could be applied to provide personalized feedback and support to students.

Another study utilized a Bi-LSTM model to analyze student comments in online forums and identify areas where students needed additional help [3]. The model accurately identified subjects that students struggled with, and it was recommended that the model be used to provide targeted interventions to improve student comprehension. Similarly, another study explored the use of a Bi-LSTM model to provide tailored recommendations and advice to students based on their learning histories [4]. The model accurately predicted the effectiveness of various learning resources for specific students and could be used to enhance their learning experience.

Another study used a Bi-LSTM model to analyze student essays and provide feedback on their writing skills [5]. The model was found to be effective in identifying errors in grammar and syntax and could provide personalized feedback to students.

In one study, the goal was to create an SLMS that makes use of a Bi-LSTM model to forecast student performance considering their interactions with the system. The study demonstrated that the Bi-LSTM model could predict student grades with an accuracy of over 80% using a dataset of student interactions with an online platform. According to the study, the methodology might be applied to give pupils personalized feedback and support [8].

Student comments in online forums were analyzed using a Bi-LSTM model to spot places where students needed more help. The study discovered that the model could correctly pinpoint the subjects that students were having trouble understanding, and it was recommended that the model might be used to deliver targeted interventions to enhance or increase student comprehension [9]. In a different study, student comments in online forums were analyzed using a Bi-LSTM model to spot places where students needed more help. The study discovered that the model could correctly pinpoint the subjects that students were having trouble understanding, and it was recommended that the model might be used to deliver targeted interventions to enhance student comprehension.

The usage of a Bi-LSTM model to offer students tailored recommendations or advice based on their learning histories was investigated. According to the study, the model can be used to make tailored recommendations or advice to students to enhance their learning experience. The model was found to be able to accurately predict the efficiency of various learning resources for certain students [10].

This indicates that using Bi-LSTM models in SLMSs has the potential to enhance student learning by offering individualized feedback and support. The algorithm can pinpoint places where students need more assistance and offer tailored interventions to improve their comprehension of the subject matter by examining student data. The model can also give teachers information about how students behave and perform, which they can use to enhance the efficiency of their teaching methods.

III. Proposed methodology:

Bidirectional long-short term memory (Bidirectional-LSTM) is the process of making any neural network to have the sequence information in both directions backwards.
Features that can be implemented in this project using Bi-LSTM.

In the first stage, raw data is collected to initiate the data processing [15]. After the data is processed, cleaning of the dataset has been performed. Data cleaning is performed to make sure that the data utilized for analysis is consistent, accurate, as well as complete. Thereafter, the models have been implemented as per the requirement. Finally, the decisions have been made as per the visualizations [4].

1. Dataset collection: To generate a dataset for analysis or machine learning models, relevant data must be gathered from a variety of sources and organized.

2. Data pre-processing: To prepare raw data for analysis or machine learning models, it must be cleaned, transformed, and organized.

3**. Data Cleaning**:

a. ***Data Integration:*** Data integration is the process of merging information from several sources and formats into a single, cohesive format for use or analysis.

b. ***Data Transformation:*** Data transformation entails changing and altering data, such as normalization or aggregation, to make it acceptable for analysis or machine learning models.

c. ***Data Reduction or Dimension Reduction:*** Data reduction, also known as dimension reduction, is the process of lowering the amount of data while keeping the key characteristics needed for analysis or machine learning models.

4.***Training Models:*** To generate predictions or categorize new data, models must be trained, which entails employing algorithms to discover patterns and correlations in data.

5.**Identifying class labels**:

a. ***Representations or attributes that describe the data:*** The qualities and aspects of the data that are used to build models for analysis or machine learning are described by representations or attributes.

b. ***Extraction of the features from the dataset:*** In feature extraction, the dataset's most pertinent and instructive qualities are chosen and modified to provide a smaller collection of features that encapsulate the key aspects of the data. Techniques like feature engineering, feature scaling, and dimensionality reduction may be used in this process.

6. **Training the models to predict the classes/labels**:

a. ***Extraction of the Feature required for prediction:*** It is necessary to determine which features are most pertinent for the prediction job to extract the features needed for prediction. This may entail feature selection approaches, exploratory data analysis, and domain expertise.

Before being utilized for model training and prediction, the pertinent characteristics might be modified and preprocessed.

There are multiple steps in the suggested process for creating a Student Learning Management System (SLMS) that uses the Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm [11].

The first stage is gathering data. The SLMS will gather information from a variety of sources, including student use of the system, forums, and course materials. Pre-processing will be done on the data to get rid of extraneous data and noise.

Training the Bi-LSTM model is the next stage. To forecast student success based on their interactions with the system, the model will be trained on the pre-processed data. To increase or boost the model's precision, a validation set will be used for fine-tuning [3].

Creating a user interface for the SLMS is the third phase. Students and professors will have access to course materials, homework, and communication tools through the user interface. Teachers will be able to grade student work and offer feedback through the user interface.

The Bi-LSTM model is incorporated into the SLMS in the fourth stage. The approach will be included in the system to give students personalized feedback and help. In order to help students better understand the subject, the model will analyze student data to pinpoint areas in which they require further support.

The SLMS is put to the test in the fifth step. A sample of students will be used to measure the system's performance in raising student achievement and fostering student-teacher communication [12].

The system must be improved and optimized in the last step. The system will be improved and optimized with the help of student and teacher feedback. To increase accuracy and efficiency, the model will also be periodically retrained on fresh data.

The suggested process for creating an SLMS utilizing a Bi-LSTM algorithm entails data gathering, model training, and creation of the user interface, integration of the models, testing, and system improvement. The system uses data analysis, targeted feedback, and support to give each learner a personalized learning experience [4].

System Architecture:

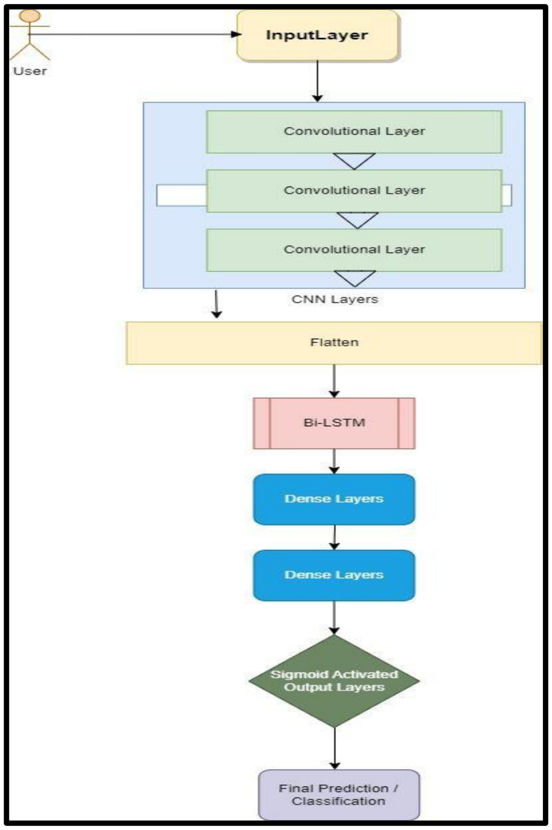

Figure 1: Architecture of the system

(Source: Created by the learner)

Figure1, displays the Architecture of the system in the form of a flowchart. Hence from the above figure, it can easily be inferred that the user imported the data by using the Input layer. After that, data has been passed through various convolutional layers as well as dense layers. Finally, after passing through the sigmoid-activated output layers, a final classification or prediction has been made [3].

While the pooling operation can minimize the number of parameters and prevent overfitting, the convolutional operation offers the advantages of local connection and weight sharing, which can simplify the model [1]. To obtain each character's contextual information from a distance, we employ Bi-LSTM. Despite the spacing between the characters, it can learn and capture the intricate connections between many characters inside sequential data. There are two layers of hidden nodes from two different LSTMs in the bidirectional design. The dependencies in separate directions are captured by the two LSTMs [12]. In contrast to the second hidden layer, where the direction of recurrent connections is reversed and activation is passed backward in the sequences, the first hidden layer has recurrent connections from the most recent words [2]. Because of this, in the LSTM layer, From the forward LSTM network, we may obtain the forward hidden state, and from the reverse LSTM network, the backward hidden state. The compositional semantics information in both directions of the character sequences is captured by the two-state.

- Deep neural networks have recently been shown to achieve highly competitive performance in many computer vision tasks due to their ability [13] to explore in a much larger hypothesis space.
- However, since most deep architectures like stacked RNNs tend to suffer from the vanishing-gradient and overfitting problems, their effects are still under research.

Dataset

The dataset used to train all algorithms is available at the following URL: https://www.kaggle.com/code/kianwee/eda-students-academic-performance-dataset/input?select=xAPI-Edu-Data.csv [17].

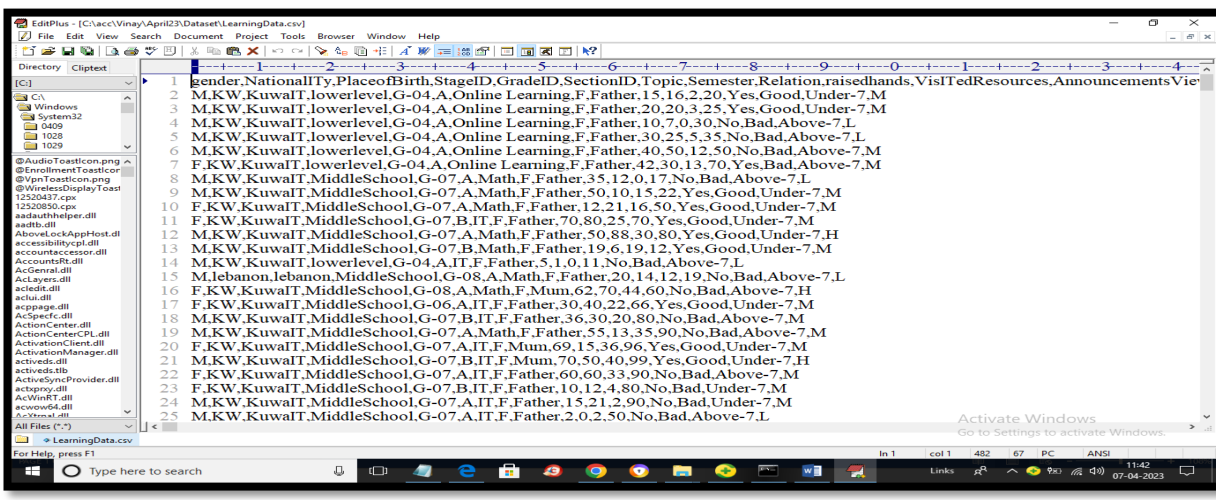

Figure 2: Dataset

The top row of the above dataset screen shows the dataset column names, while the following rows show the dataset values. Students' performances are listed in each row, and their outcomes are listed in the last column as L (low), M (Medium), and H (High). Therefore, the user will train and assess each algorithm's performance using a forementioned dataset in terms of accuracy, precision, recall, FSCORE, confusion matrix graph, and ROC [18].

Below are the code and output screens with blue-colored comments that the user created while using the JUPYTER notebook to code this project.

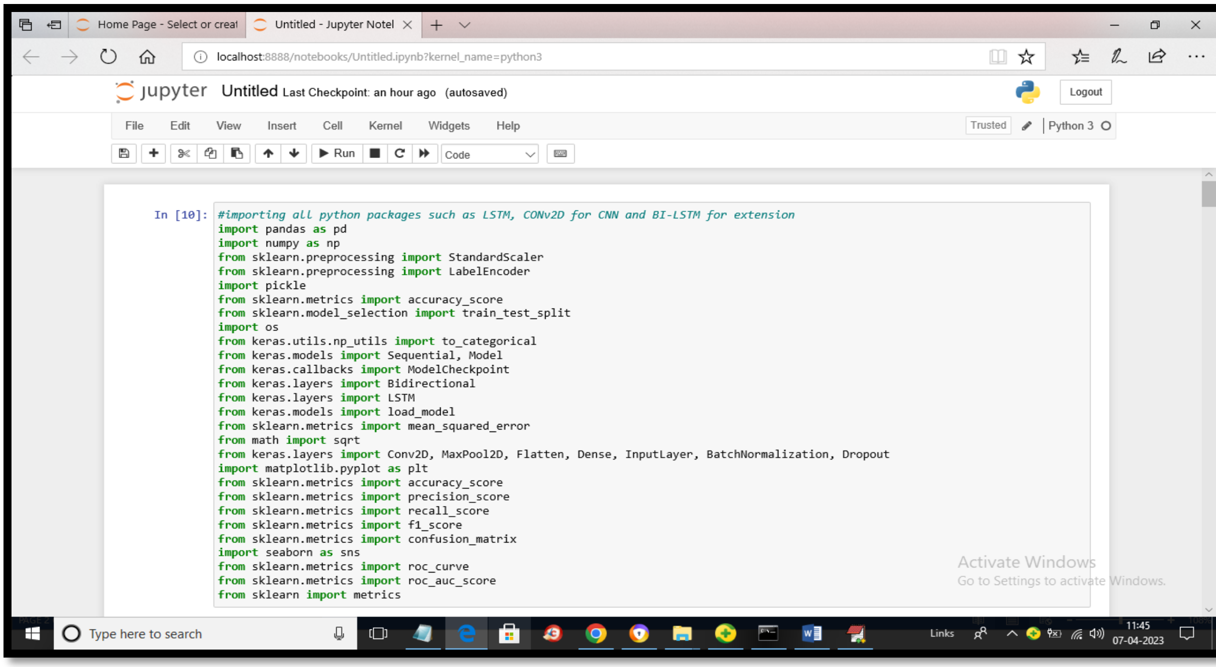

Figure 3: Importing the libraries.

In Figure 3, the user is importing the necessary Python classes and packages on the screen above.

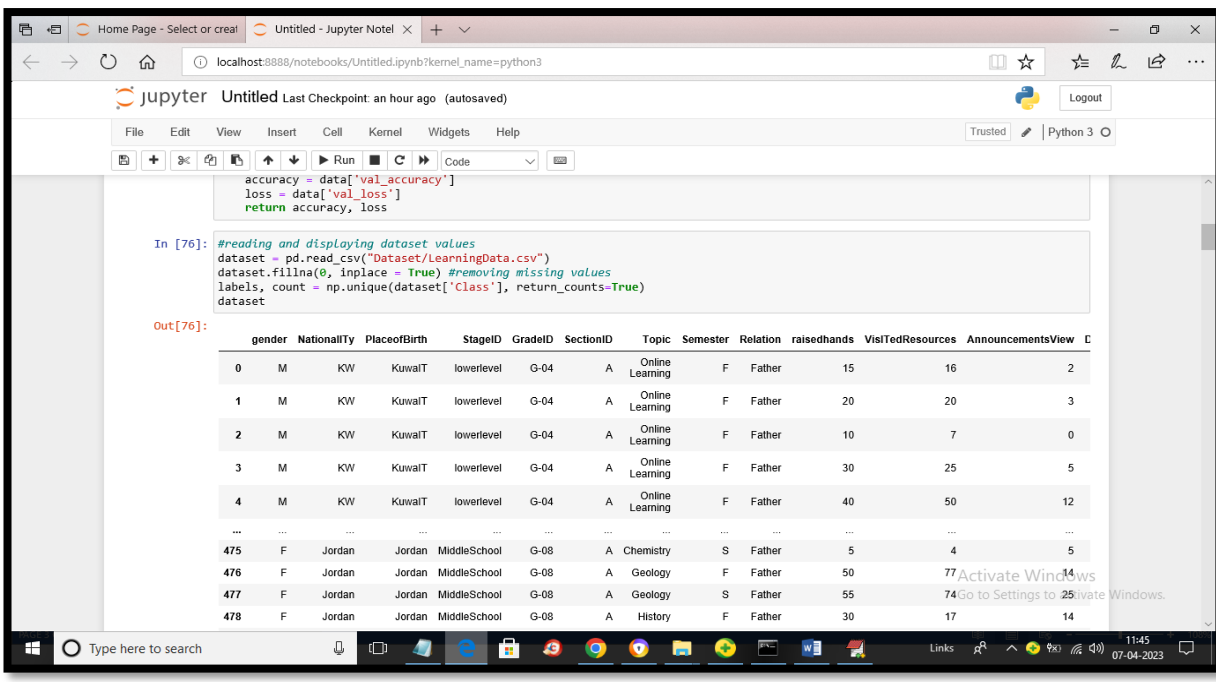

Figure 4: Importing the dataset.

In Figure 4, screen reading and displaying dataset values.

IV. Implementation

There are various processes involved in the implementation of the Student Learning Management System (SLMS) using the Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm.

First, information is gathered from a variety or mixture of sources, including course materials, discussion forums, and student interactions with the system. The gathered data is then pre-processed to eliminate unnecessary data and background noise.

Second, based on how students engage with the system, the Bi-LSTM model is trained on the pre-processed data to forecast student success. A validation set is used to fine-tune the model and increase its accuracy.

Thirdly, a user interface for the SLMS is created. Both teachers and students have access to the interface's communication features and course materials. Teachers can evaluate student work and offer feedback using the interface [14].

Fourthly, the SLMS incorporates the Bi-LSTM model. The methodology examines student data to pinpoint areas where students need more assistance and offers focused interventions to enhance their comprehension of the subject. Additionally, or also, the model offers each student their own support system and feedback.

Fifth, the system is evaluated on a sample of students to determine whether it is helpful in raising student achievement and fostering student-teacher communication [15]. The system is improved and improved upon using student and teacher feedback.

Finally, the intended audience receives the SLMS. Students and teachers can use the system in their coursework and classroom activities. The performance of the system is tracked, and suggestions are used to make it more efficient.

Also, data gathering, model training, user interface development, model integration, testing, and system deployment are all steps in the implementation of the SLMS utilizing a Bi-LSTM algorithm. Through data analysis, individualized feedback, and support, the implementation aims to give each student a customized learning experience [5].

V. Results

The user has chosen Convolution Neural Network (CNN) as the basic method since it is well-known for its best feature extraction or optimization algorithm and has demonstrated its effectiveness in several different sectors, including image classification, disease prediction in humans, and many more. Because of this, the user trained CNN on the Students Performance dataset, extracted its optimized features, and then trained it again with LSTM and BI-LSTM

[16]. BI-LSTM was able to achieve accuracy levels between 95 and 98% after training but LSTM was unable to boost accuracy. More accurate than LSTM and CNN is BI-LSTM.

The accuracy of LSTM may not increase because it only filters feature in the forward path and never goes back to re-filter dropout features. The BI-LSTM will search in both forward and reverse directions for feature filtration to collect more enhanced features and hence improve accuracy.

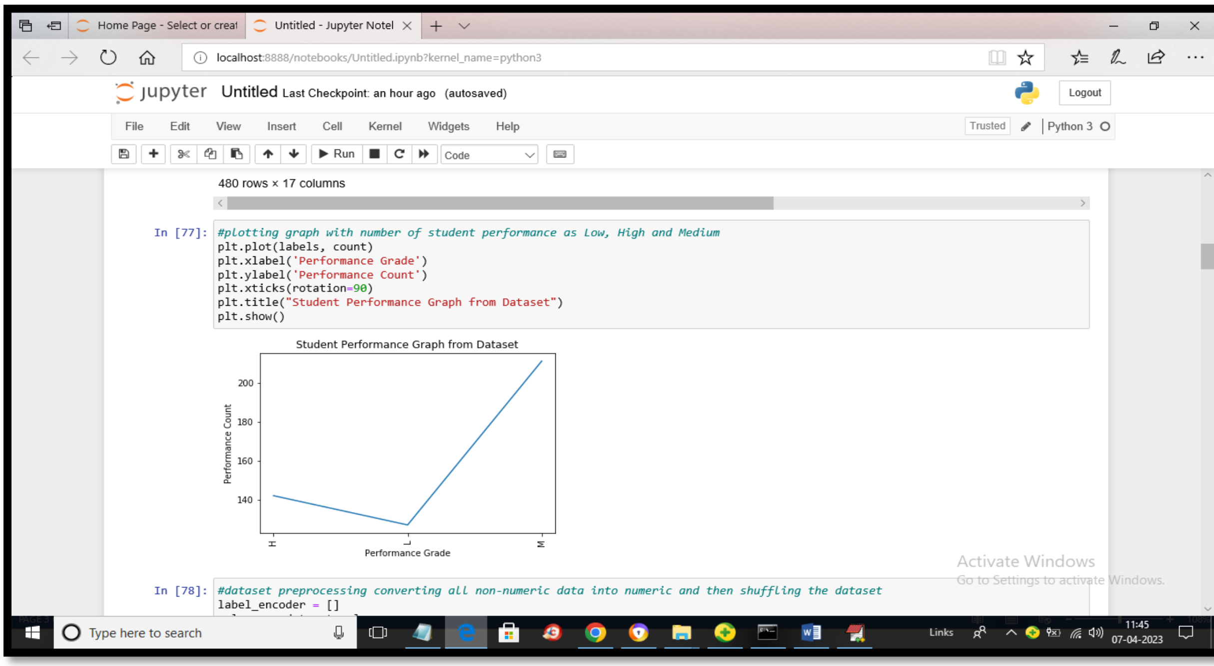

Figure 5: Number of students with various performance levels

The graph in figure 5, shows the number of students with various performance levels that were found in the dataset. The graph's x-axis shows the labels H, M, and L, while the y-axis shows the number of students who met that performance level.

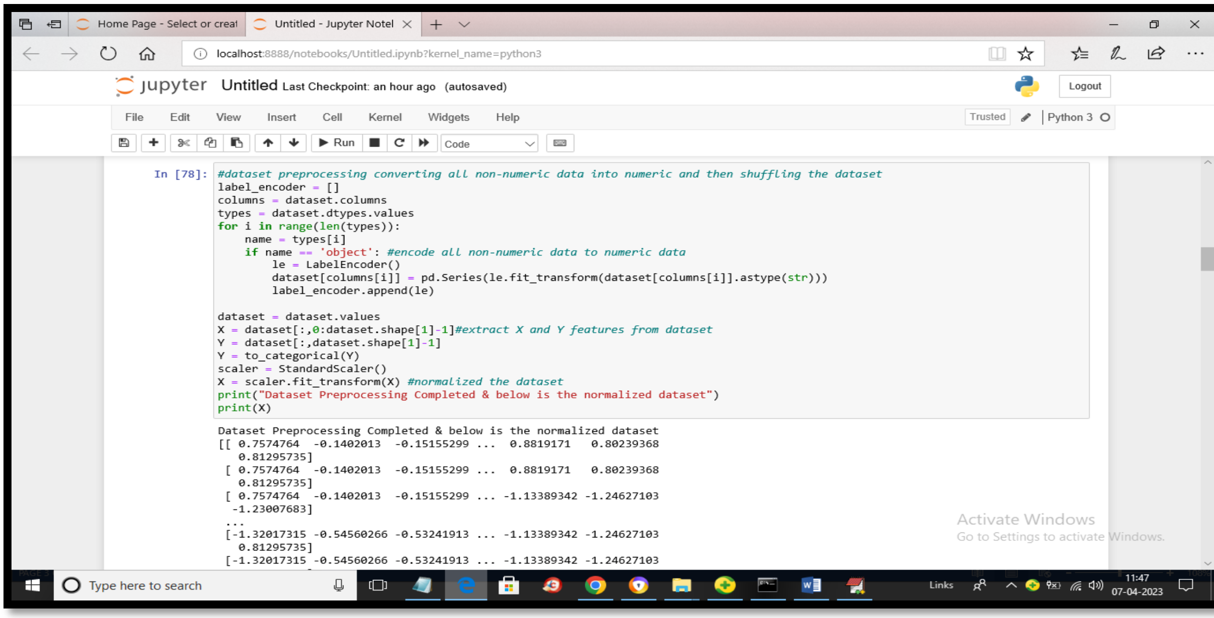

Figure 6: Preprocessing of the dataset

In figure 6, the user is preprocessing the dataset by using the label encoding technique to convert all non-numeric data into numeric data, normalizing the data, and then extracting the X training and Y target features from the dataset. After processing, the user is shown the dataset's numeric and normalized values [19].

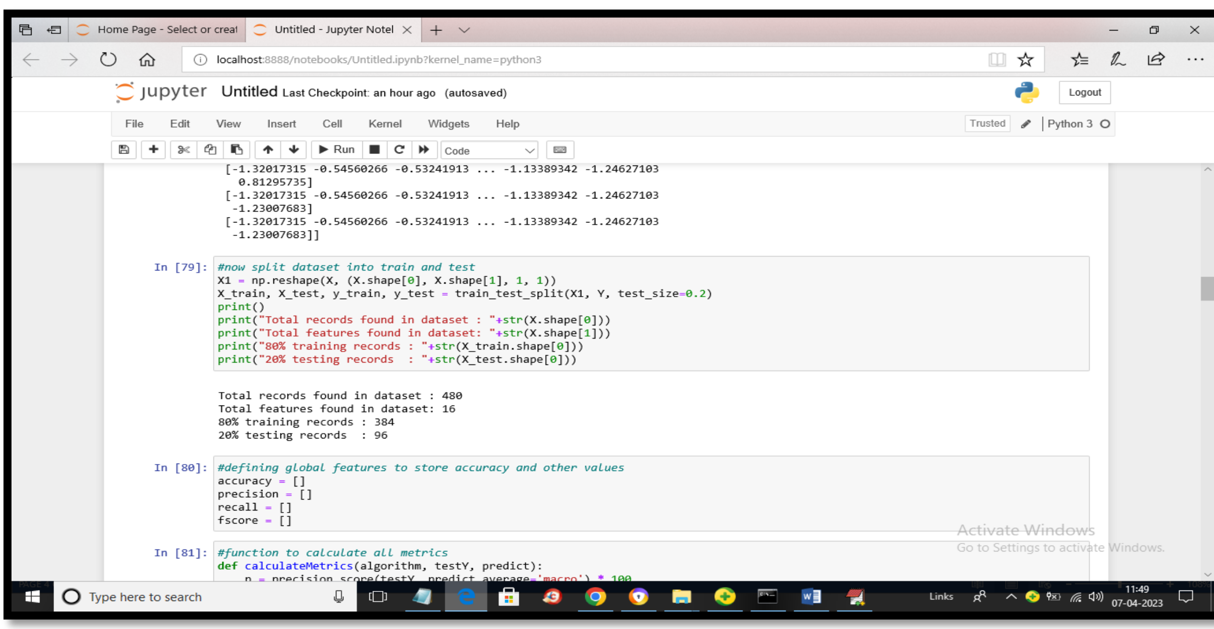

Figure 7: Splitting the dataset

The user has divided the dataset into train and test in figure 7, utilizing 80% of the dataset for training and 20% for testing.

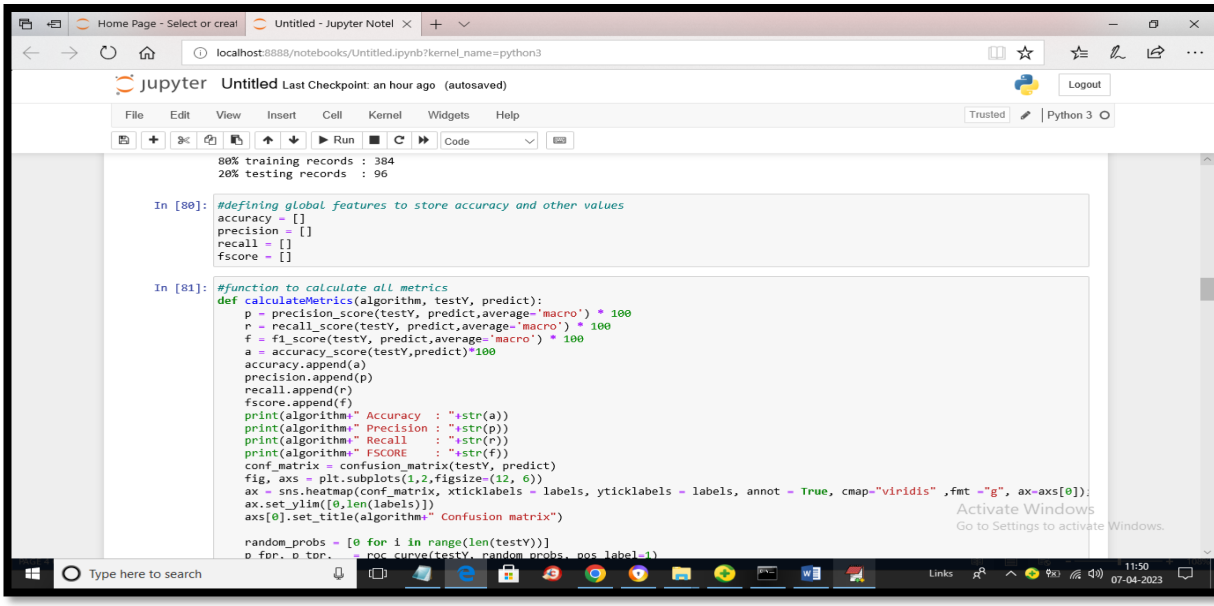

Figure 8: Get the accuracy.

The accuracy, precision, recall, and FSCORE values are defined in figure 8, and the confusion matrix and ROC graph are also presented.

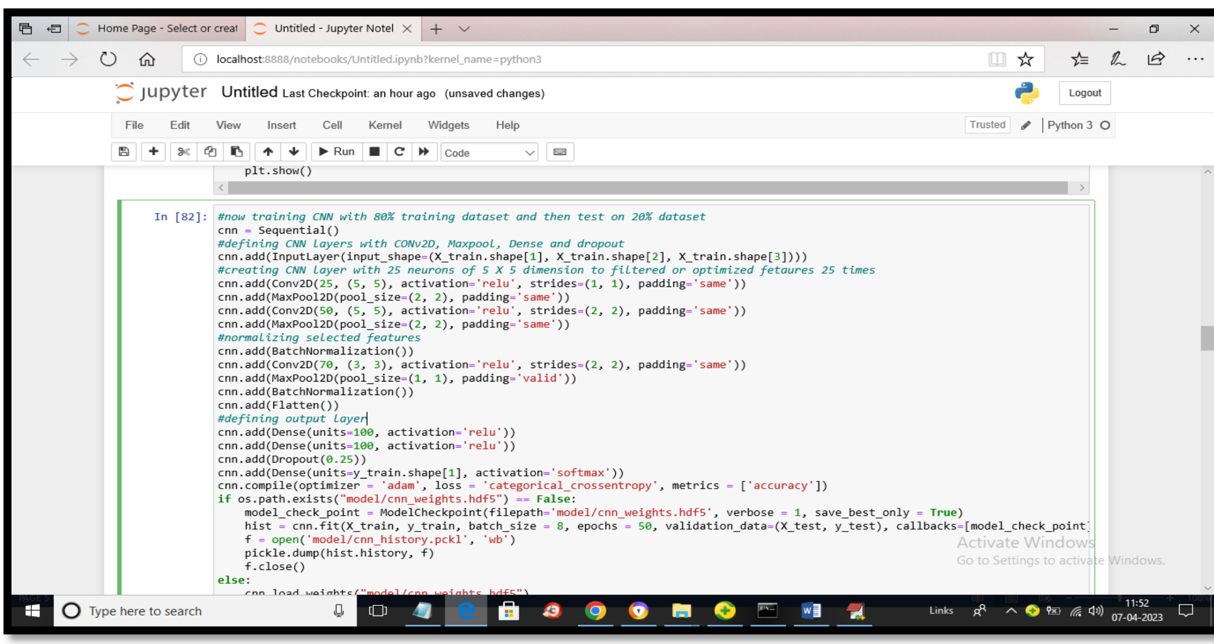

Figure 9: CNN algorithm

Defining the CNN algorithm with various layers in the previous screen, this block will be executed to produce the accuracy and other performance metrics shown in figure 10.

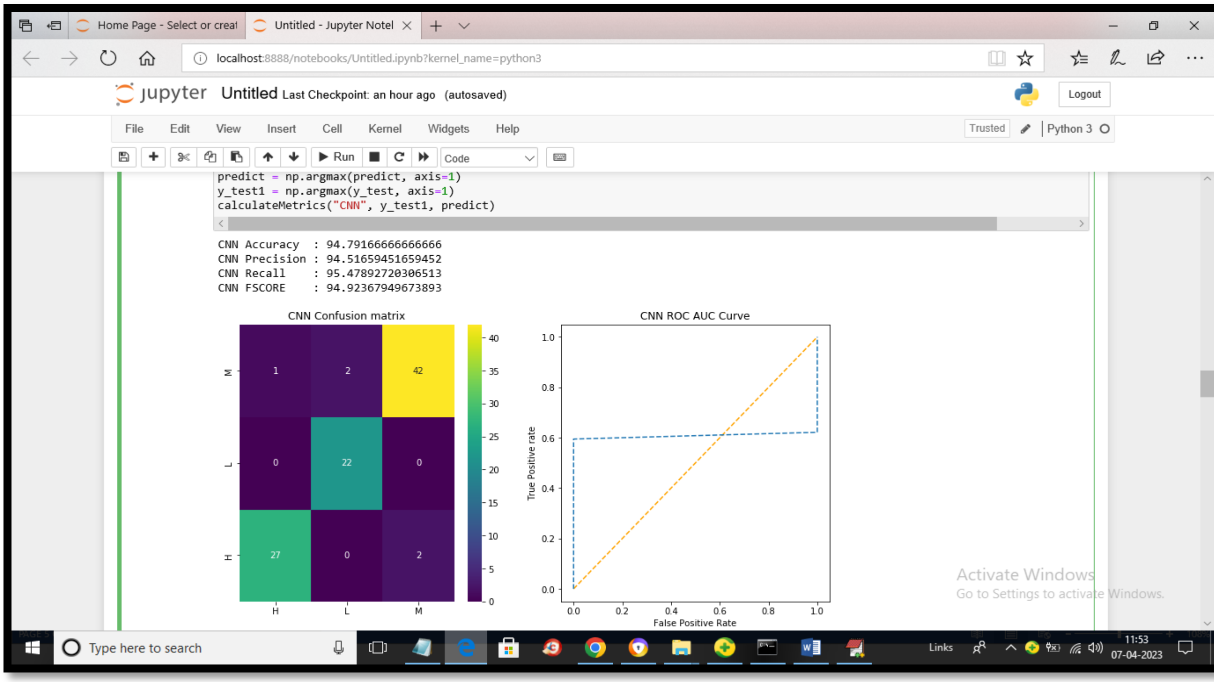

Figure 10: Accuracy and confusion matrix

In the figure 10, CNN integrated BI-LSTM got 97% accuracy which is more than CNN and LSTM.

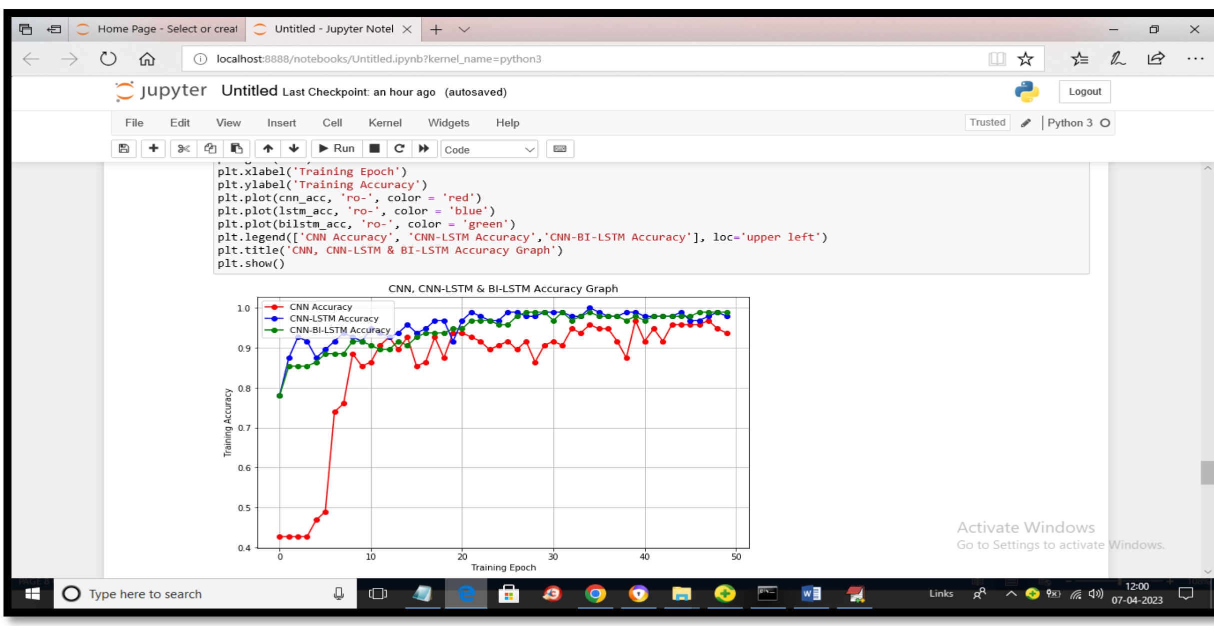

Figure 11: Training data accuracy

In the training accuracy graph in figure 11, the Y-axis shows training accuracy, and the X-axis shows training epoch. The red line represents CNN, the blue line CNN-LSTM, and the green line CNN-BI-LSTM, which has higher accuracy across all the algorithms.

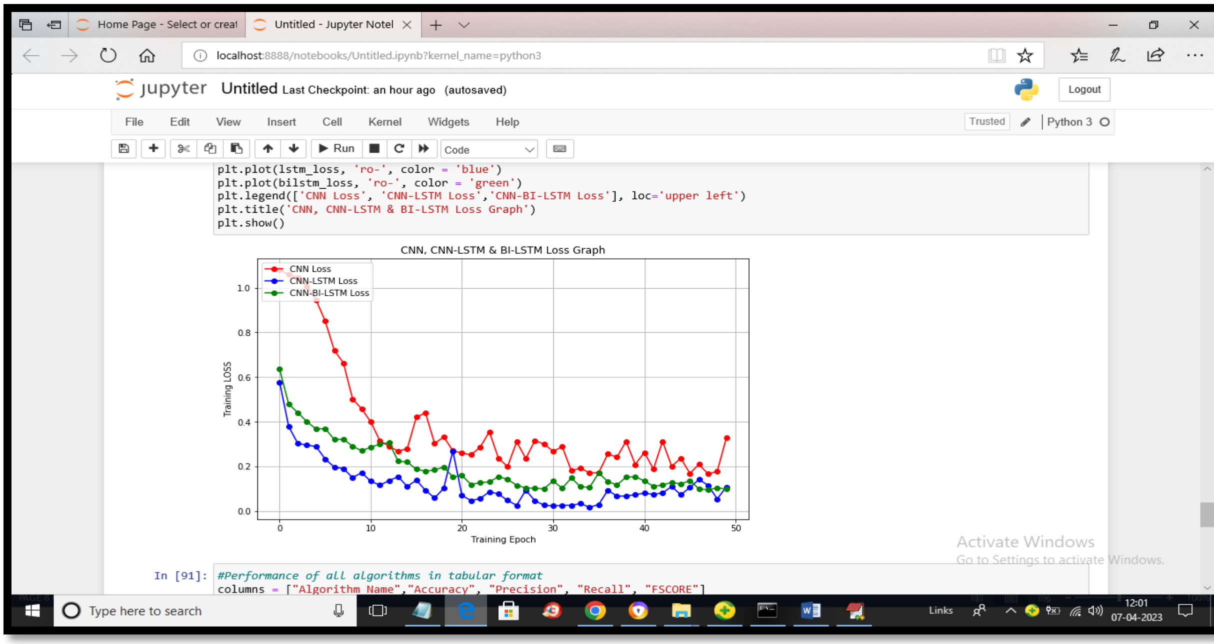

Figure 12: Loss graph

In figure 12, the user can see the loss graph for each algorithm.

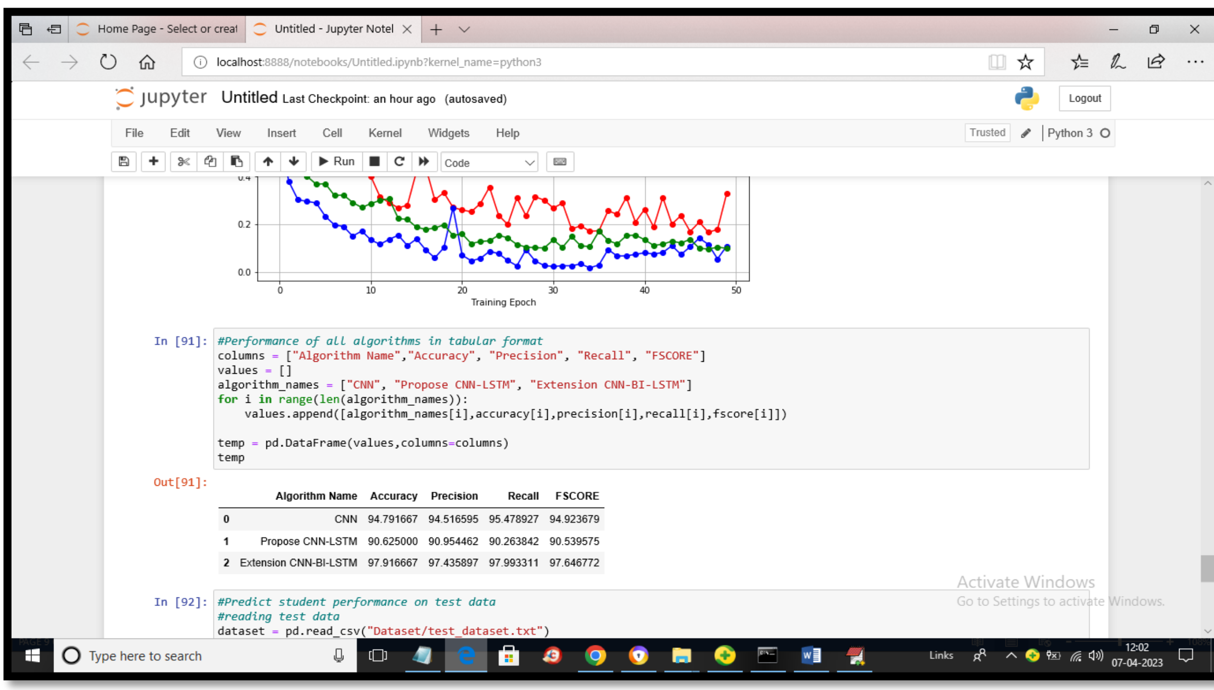

Figure 13: Each algorithm in tabular format

In figure 13, the user can see the performance of each algorithm in tabular format.

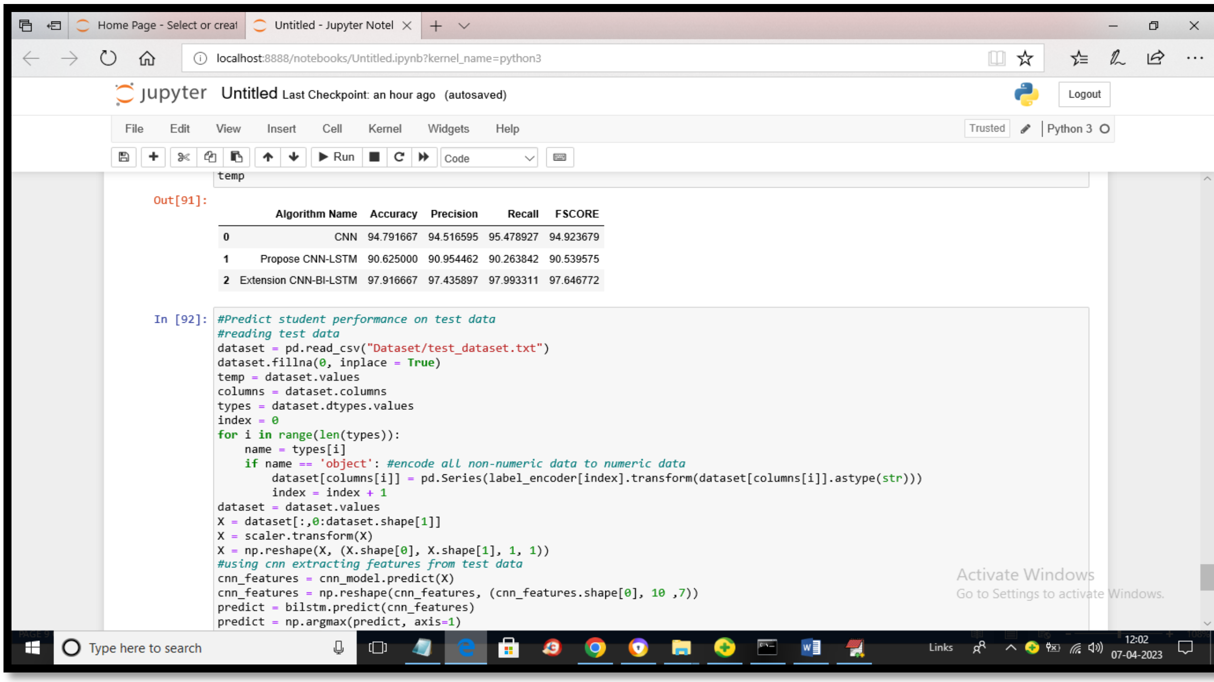

Figure 14: Test data as an input value

In figure 14, the user is giving test data as input and then BI-LSTM will give the below prediction output.

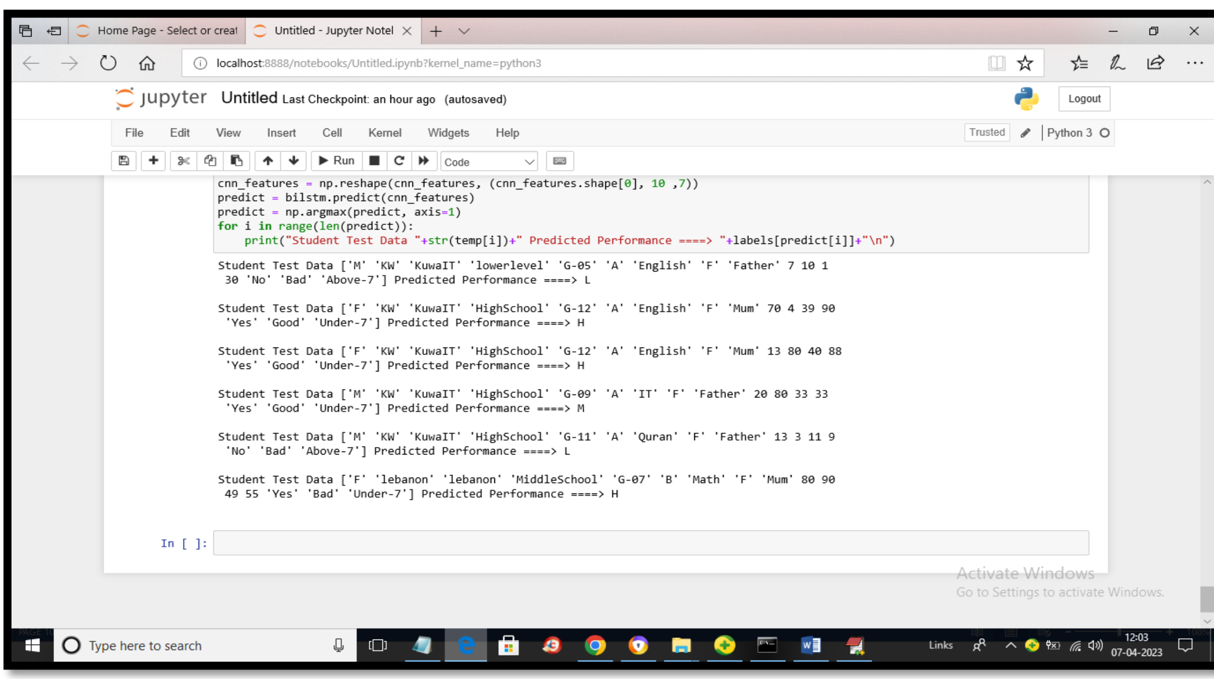

Figure 15: Output

Academic information about the student is shown in figure 15 as square brackets, and the expected output is shown as L, M, or H after the arrow symbol =□.

VI. Future Work and Discussion

By adding additional data sources and utilizing cutting-edge machine learning techniques, the Student Learning Management System (SLMS) that uses a Bidirectional Long Short-Term Memory (Bi-LSTM) algorithm could be improved in future studies [6]. The system might also be enhanced to handle more languages and academic disciplines, opening it up to a wider range of pupils. To further enhance the precision and efficacy of the model, more sophisticated methods for studying student data could be investigated.

To give students a more complete learning experience, the system might also be coupled with other tools and platforms [20]. Finally, it will be essential to continuously assess and evaluate the system's performance to maintain its efficacy and relevance to the needs of teachers and students.

VII. References

[1]Yang, Z., Zhou, L. and Jing, Z., 2022. A Novel Affective Analysis System Modeling Method Integrating Affective Cognitive Model and Bi-LSTM Neural Network. *Computational Intelligence and Neuroscience*, *2022*.

[2]Ngwira, B., Gobin-Rahimbux, B. and Sahib, N.G., 2023. A Deep-Learning Framework for Analysing Students' Review in Higher Education. *Computational Intelligence and Neuroscience*, *2023*.

[3]Jha, S. and Sonawane, P., 2022, March. Smart Student Grievance Redressal System with Foul Language Detection. In *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 187-192). IEEE.

[4]Yang, X., Zhou, Z. and Xiao, Y., 2021. Research on Students' Adaptive Learning System Based on Deep Learning Model. *Scientific Programming*, *2021*, pp.1-13.

[5]Zeberga, K., Attique, M., Shah, B., Ali, F., Jembre, Y.Z. and Chung, T.S., 2022. A novel text mining approach for mental health prediction using Bi-LSTM and BERT model. *Computational Intelligence and Neuroscience*, *2022*.

[6]Liu, J., Xie, B. and Shi, Y., 2021. *Semantic repeatability screening mechanism of intelligent learning platform based on Bi-LSTM* (pp. 200-209).

[7]Jang, B., Kim, M., Harerimana, G., Kang, S.U. and Kim, J.W., 2020. Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism. *Applied Sciences*, *10*(17), p.5841.

[8]Almars, A.M., 2022. Attention-based Bi-LSTM model for Arabic depression classification. *CMC-COMPUTERS MATERIALS & CONTINUA*, *71*(2), pp.3091-106.

[9]Özkaya, U., Öztürk, Ş., Melgani, F. and Seyfi, L., 2021. Residual CNN+ Bi-LSTM model to analyze GPR B scan images. *Automation in Construction*, *123*, p.103525.

[10]Rahul, J. and Sharma, L.D., 2022. Automatic cardiac arrhythmia classification based on hybrid 1-D CNN and Bi-LSTM model. *Biocybernetics and Biomedical Engineering*, *42*(1), pp.312-324.

[11]Yadav, S., Ekbal, A., Saha, S., Kumar, A. and Bhattacharyya, P., 2019. Feature assisted stacked attentive shortest dependency path based Bi-LSTM model for protein–protein interaction. *Knowledge-Based Systems*, *166*, pp.18-29.

[12]Kamyab, M., Liu, G. and Adjeisah, M., 2021. Attention-based CNN and Bi-LSTM model based on TF-IDF and glove word embedding for sentiment analysis. *Applied Sciences*, *11*(23), p.11255.

[13]Palanisamy, K., Singhania, D. and Yao, A., 2020. Rethinking CNN models for audio classification. *arXiv preprint arXiv:2007.11154*.

[14]Sharma, P., Berwal, Y.P.S. and Ghai, W., 2020. Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. *Information Processing in Agriculture*, *7*(4), pp.566-574.

[15]Rhanoui, M., Mikram, M., Yousfi, S. and Barzali, S., 2019. A CNN-BiLSTM model for document-level sentiment analysis. *Machine Learning and Knowledge Extraction*, *1*(3), pp.832-847.

[16]Dai, Y., Zhou, Q., Leng, M., Yang, X. and Wang, Y., 2022. Improving the Bi-LSTM model with XGBoost and attention mechanism: A combined approach for short-term power load prediction. *Applied Soft Computing*, *130*, p.109632.

[17]Wang, S., Chen, Y., Ming, H., Huang, H., Mi, L. and Shi, Z., 2020. Improved Danmaku emotion analysis and its application based on Bi-lstm model. *IEEE Access*, *8*, pp.114123-114134.

[18]Kuo, K. and Carpuat, M., 2020, July. Evaluating a Bi-LSTM model for metaphor detection in TOEFL essays. In *Proceedings of the Second Workshop on Figurative Language Processing* (pp. 192-196).

[19]Zhang, G., Li, X., Wang, X., Zhang, Z., Hu, G., Li, Y. and Zhang, R., 2022. Research on the Prediction Problem of Satellite Mission Schedulability Based on Bi-LSTM Model. *Aerospace*, *9*(11), p.676.

[20]Budak, Ü., Cömert, Z., Rashid, Z.N., Şengür, A. and Çıbuk, M., 2019. Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images. *Applied Soft Computing*, *85*, p.105765.