

Student's Performance Prediction in Final Examination Using Real Dataset: A Deep Learning Approach

Abstract— The academic result is the most important thing in a student's career. This result depends on their academic performance and many other factors. Educational data mining can help both students and institutions develop their academic performance. For analysis of their performance, we can use new techniques Deep Learning, Convolution Neural Networks, Data Clustering, Optimization Algorithms, etc. In machine learning. Using Deep Learning, we will predict the student's performance yearly in the form of CGPA and compare that with the real CGPA. A real dataset can boost the prediction performance. We used a real dataset from the Institute of Science, Trade & Technology (ISTT).

Keywords— Deep Neural Network, CNN, Optimization algorithm, Data Mining, Student's Performance.

I. INTRODUCTION

Discovering Knowledge from a large number of Databases is known as Data mining. It extracts secret information from various data sources pertaining to several fields. Different techniques can be engaged in various fields of data mining together with weather forecasting, oil research, medical, business, EDM, marketing etc. [13]

In our paper we do Educational Data mining. Everybody knows that, Education is the claw of a nation. The Educational institutes and parents both parties are worried about their results. As this results, have a big impact on their career life. So, they to know that their child will do best in their exam or not?

Actually, Educational Data Mining mention the recipe to figure out exact information from large number of educational dataset. In our experiment, we collect the data from an educational institutions and apply several methods to produce meaningful and significant information. [1] Students academic performance, usually, is measured by year final examination results that are considered to point where the students able to reach in future. In the Comprehensive

evaluation method, examination outcome is pardoned by Cumulative Grade Point Average (CGPA) [14]. In order to derive knowledge from educational environments Data mining, statistics and machine learning are applied on EDM. Nowadays it is in need and gathering more care because of increase in the educational data of eLearning systems, and even developing usual education. Alarmed with sprouting methods for uncovering the distinctive types of data present in scholastic environments, it seeks to extract significant information in order to advance and praise learning processes from vast amounts of raw data [15]. According to the records of traditional database can provide answer to Problems such as “find the students who failed the examinations”, whereas EDM offers answers to additional problems like “predicting the students' performance pointing out the CGPA depending on some various factors such class attendance, test marks, personal facts and so on.

Over and again, many researchers started exploring various data mining techniques to create students' exploration models. It naturalizes the learning method, guesses the dropout rate, helps the grading scheme, etc. These patterns embody several techniques for prediction such as decision tree, regression, classification, clustering, and the like to proceed final output. But, the outcome of those predictions is not yet up to the mark. [1]

After looking into all the existing results, we are offering a Data Mining technique to resolve this problem in an additional authentic manner depending on 18 factors of every single student to show how these factors affect their academic performance. Our model will help us to know whether the particular student will do a good result or not. Most educational Institutes follow the CGPA system to measure students' performance. So we are planning to make a model which will predict their next year's result depending on some abstract and their previous semester's results and provide a CGPA. We collect those attributes from the students physically and through an online survey.

After Analysis all the previous attempt and to get a more accuracy we use a original dataset from the department of CSE at Institute of Science, Trade and Technology (ISTT). Here we take the data of 158 students for result analysis. we make the neural network with two hidden layers which is called deep neural network. Now we use Supervised machine learning algorithm CNN to train the deep neural network. We also used one hot encoding to convert our categorical data into numerical data. We refuted our entire dataset into three parts. One part is for training our neural network also we keep one for testing and last one is for verifying the result and we get the low number of percentage of error.

The main contribution of this papers that, we offer a new technique founded on Supervised Deep Neural Network which figure out the information from student's data. The empirical results display that the academic performance exceedingly depends on some factors both from academic and personal life.

The residual sections of the paper are embodied here. Related work with our experiment on predicting students' performance describes in section II. In section III, we bring in our proposed solution including some sub-sections. Section IV is responsible for the narration of the dataset. Section V is for the Implementation of our offered technique. We stated our Result in section VI & our future work is visible in the last section means VII.

II. RELATED WORKS

Fahim, Sajal and Jalal et. al. [1], propose a neural network which help to guess student's yearly accomplishment using neural network in the form of Cumulative Grade Point Average (CGPA) and compare that with original CGPA.

Kyndt et al.[2] used neural network to predict academic performance of the first year bachelor degree student, based on students' inspiration, way of learning, working memory retention and attention.

Ioannis E. Livieris, et al. [3] in their study predicted the performance of students in Mathematics using an Artificial Neural Network (ANN) classifier. They found that the modified spectral Perry trained artificial neural network performs better classification compared to other classifiers in this context.

S. Kotsiantis, et al. [4] investigated in distance learning using machine learning techniques for dropout prediction of students. Their study made an important contribution as it was a pioneer and helped to carve the path for educational data mining

San Pedro et. al [5] analysis a web based tutoring method on 3747 school students for mathematics course and try to guess whether a student will be present college (5 years later). Authors trained the students in middle school mathematics as weighted by the tutoring method are more likely to enroll 5 years' additional in college. On the other side, the students who proved chaos, inadvertency in the method have minor chance of college inscription. For the prediction they used logistic regression classifier.

Vihavainen et. al [6] Worked with some data from Computer Science students programming course of Helsinki University and tested to calculate whether a Student will pass or fail introductory mathematics course.

Bhardwaj and Pal [7] guess students' performance and found out living place has high impact on student's year final

outcome. They usage all the information from the Department of Computer Applications student's at Purvanchal University's

Al-Radaideh, et al [8] predict the final grade of students who taken the C++ course in Yarmouk University, Jordan. There are three various classification techniques namely, ID3, Naive Bayes, C4.5 and the are used. The outcome reported that Decision Tree model had better guess than other models. Nguyen and Peter [9] guess the performance of the students and likened the proficiency of two classifiers as Bayesian networks and Decision Tree using WEKA tool. They engaged two several groups of students of postgraduate and undergraduate level. The performance of Decision Tree was more correct around 3-12% comparing Bayesian networks. This experiment was beneficial for pointing the weak students for commanding and choosing good students for scholarship.

S. Anupama and Vijayalakshmi [10] Hope the performance of final exam of MCA students in accordance with their internal marks. This time they had used Decision tree algorithm. They liken the predicted results and original results which points out, that there was a meaningful uplift in results as the prediction support a lot to place good and weak both types of students to confirm good marks. They also likened their model with an algorithm named Decision tree & show that the developed system or method or model is better in terms of proficient and time taken to make the decision tree.

Smith et. al. [11] Bring forward in sight school safety pardon by student academic representation, attendance, and post-secondary aspirations.

Baris Cetin [12] Bring in to learning and age in guessing college students' academic acquisition. Improved higher order orthogonal iteration algorithm for student rending prediction.

[14] S. K. Yadav, S. Pal and B. Bharadwaj "Data mining applications: A comparative study for predicting student's performance," arXiv preprint arXiv:1202.4815, 2012[15] Scheuer et al., Educational data mining. In Encyclopedia of the sciences of learning pp. 1075–1079, Springer, 2012.

There are so many researchers had already done their experiment on this topic. In our paper we mention a model which is highly effective comparing the existing ones.

III. PROPSOED SOLUTION

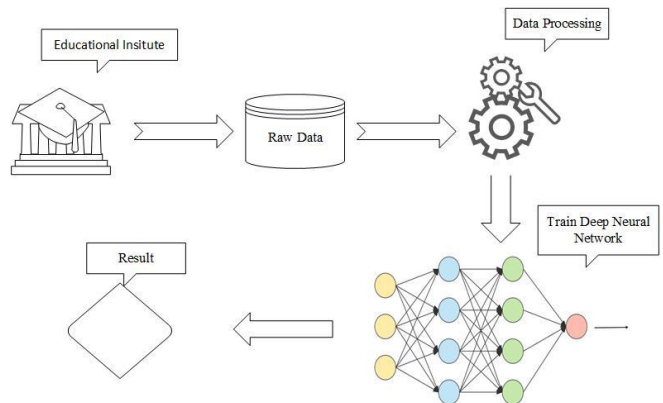


Fig. 1. System Overview

Fig. 1. shows the entire system overview. It describes the overall idea of how we made our model and how the performance will predict. First of all, we went to an educational institute and collect the data from students. After collecting the data, the most part is that we need to clean our data which means preprocessing the data, in this case we use one hot encoding to preprocess the data and convert categorical data into a numerical value. After preprocessing the data then we design a machine learning-based neural network which is a supervised neural network to help to train the data & figure out the essential information from the data. In this experiment we predict the student’s final year result in CGPA. We also used the CNN algorithm to train our neural network. In our model, we predict students’ performance and then attempt to find out the accuracy percentage comparing the original one. Every single part of our proposed system is described below in the subsections.

A. Data Collection

Data collection is the methodical way to collect and standardize information from different origins to get a fulfilled and accurate image of an area of interest. Data collection enables a person, organization, or institution to answer relevant queries, assess results, and make predictions about coming possibilities and trends. In our system, first, we collect students’ relevant data for predicting the student’s yearly performance. Here we use 158 students’ data for our research purpose which is taken from the department of Computer Science & Engineering(CSE) of the Institute of Science Trade & Technology (ISTT). We collect these data from the students of ISTT via an online survey using Google docs and physically from the students. In this survey, we consider 18 factors where 17 are our input factors and the last one is our output factor. We consider it as not only academic data but also some other personal information like Living area, Family Education, Political involvement, Drug addiction, Social media interaction, Affairs, and so on.

B. Data Pre-Processing

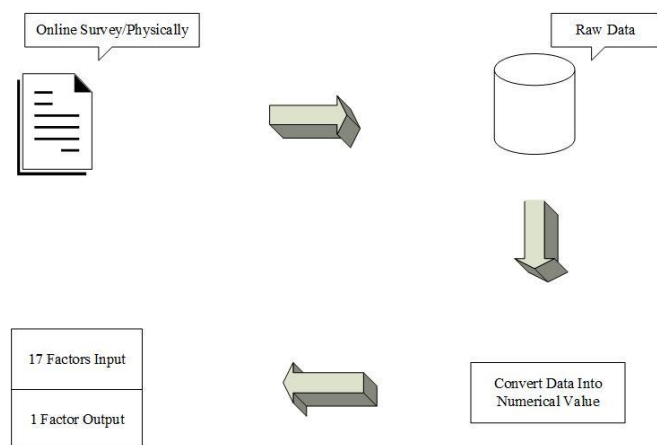


Fig. 2. Data Pre-Processing

After collecting the data, we preprocess the data. Now we have to do a filter to remove some unnecessary data. First of all, we convert all the categorical data into numerical value using one hot encoding. Then we divided the 18 factors into two parts. One part is the input part which uses 17 factors among these 18 factors and the remaining one factor is called target value. Figure 2. shows the steps of data preprocessing.

C. Deep Convolutional Neural Network

In living years, in the circumstances of research area neural network has been drawing rising intentness. Neural network has solved some of the riddle posed and makes some development for future predictions. The first thing is to get the prediction we build a deep neural network. Alter that we train the deep neural network with our preprocessed Data [1]. We used supervised neural network and train the neural network using CNN. Supervised learning is one of the Machine learning method which is use to performing a function from designated training data. In supervised learning method two kinds of data are provided which is known as inputs and outputs. After providing the data, system procedures the inputs and analogies its decision outputs against the required outputs. After this procedure, systems sometimes find some errors. These errors are once again propagating to the system. Then the System synthesize the weights which control the network. This method occurred repeatedly as the weights are frequently tweaked. The dataset we used in our system for the training purpose is known as the training set. In the training cycle of the neural network identical dataset is actioned repeatedly as the connection weights are always tasted. Sometimes training process continues for some days. This process is getting paused only when the system comprehensions some statistically aimed outcomes or expected accuracy. However, some networks never learn and the reason behind is, the input data does not carry the specific information. That is why we cannot find the expected output. Networks cannot manipulate necessary learning if there is less data. There should be sufficient of data so that some data can be caught backward as a testing phase. This network consists many layered including couple of nodes in each & every layer are efficient of deporting information [1].

IV. DESCRIPTION OF DATASET

Class Performance

One of the significant factors to measure students’ performance in class performance. In Our experimental analysis, we consider three divisions of class performance. If the students' level of class performance is not meet the expectations then we call it poor, if it is kind of satisfactory then average & if it meets the expectation level then we call it good.

Test Marks

Class test marks is one of the important element to predict the students’ performance. Generally, those students get a highest mark in their class test having a high possibility to do good in final year examination. We take the class test marks and divided into three parts which is poor, average and good. If the mark is below forty percent, then its poor. If the mark is between forty & eighty percent, then its average and if the mark is above eighty percent then its good.

Class Attendance

Another vital importance is Class attendance for student performance analysis. Those who attend classes regularly have a high chance to understand the teacher’s lectures comparing those who do not attend classes. We consider the

class attendance mark is 100%. If students' mark is less than 40% will count as poor, between 50 to 70% will count as average, and above 80% will count as good.

Due time assignment submission

We take assignment one of our factors for students' performance analysis. Most teachers give the assignment to students to check how much they understand the lecture sometimes teachers want to know how they apply their knowledge. We divided it into two categories. If the assignment mark is below 50 percent, then it is poor and if the mark is above 50 percent then it is good.

Lab Performance

One of the vital factors is lab performance for student performance analysis. Teachers can understand using this lab performance how much students get to understand the theoretical things. We consider three categories for lab performance first one is poor, the second one is average and the last one is good.

Previous semester result

This one is also an important factor. Students those who get the good result in previous semester they will inspire to make good result in future. Contrariwise those who not get good result may be fall into depression or disappointed. We consider this one for our experiment analysis.

Family Education

One of the universal truth is we all started our learning from our family. Educated parents are always wanted that their children also be an educated person hence they always try to inspire them. That's why family education is one important factor for predicting students' performance. We divided this into three categories which is poor, average and good.

Freelancer

Some students started freelancing when they are a student. In this case, they point to be able to give proper time to their study which will affect their performance. We divided it into two categories and they are Yes or No.

Relationship with faculty

We consider this one of our factors. Students who have a good relation with faculty have a chance to get career counseling from their faculty and those who haven't missed this opportunity. We divided it into three categories which are poor, average, and good.

Study Hours

The main factor for students' performance analysis. Students who study surely they have can gain knowledge daily even if they have the scope to learn a lot of modern things. We divide this Study hour into two categories excluding their class time hours. Those students who study less than 3 hours every week are in the average group & those who study more than 10 hours every week are in the good group.

Living area

To do students' performance analysis, we also give a little bit of concentration on the students living areas. Nowadays students who are living in town are a little bit more concerned than remote areas students. Contrariwise, students, who explore their higher education from his/her home and live with their parents they get extra care usually. Regarding this issue, we divided this factor into two categories. Students who are staying in the Mess/hall are in one group and others who are living at home are in different group.

Social Media Interaction

Nowadays, most students are involved in social media. Sometimes they do it in exceeds level which will be very harmful to their study and future career. The students consume most of their time on social media like Facebook, Twitter, WhatsApp, IMO, LinkedIn and so on they don't get sufficient time to study. That's why we consider it a factor of student performance analysis. Students who consume over 30 minutes & less than 2 hours per day on social media are in the average group and the rest of the students who consume more than 2 hours in a single day are in Exceed category.

Extra-Curricular Activity

Extra-Curricular activities are very important for every student. It helps to refresh and increase the important for every student. It helps to refresh and increase the power of thought as well as leadership and management skills. In this research paper, we consider students who are connected with extracurricular activities or different clubs in their university are in the "yes" category, and those who are not connected are in the "no" category.

Drug Addiction

Drug addiction is very harmful for both our mental and physical health. Students who take drugs cannot concentrate on studies. They are also involved in different kinds of violence in our society. Students are addicted to drugs in the yes group and not drug-addicted students are in the no group.

Financial support from Family

Those who get financial support from their families get enough time for study, and those who do not have to manage their expenses by doing jobs that will most of the time, they can't study well. We consider students who get financial support from their families to be in the "yes" category and those who are not in the "no" category.

Political Involvement

One of the things that is harmful is those who are involved in politics. It detracts our students from their studies. We consider students who are involved in the "yes" category and those who are not in the "no" category.

Affair

An affair is a relationship between boys and girls. Recent research shows that sometimes affairs can affect a student's academic performance. If a student (boy or girl) is engaged

with another student, then they have passed a lot of time doing hangouts. That’s why they don’t get proper time for study. So, the students engaged in affairs are in the yes category, and those who are not involved in affairs are in the no category.

Year Final Result

All the conversations we have had till now are input factors. We measure students’ performance based on these factors, which is called the year’s final result. This one is our target factor and we have already stored the final year CGPA of the student. After doing the prediction, we get the predicted result, which will be compared to the real results and will find the accuracy percentage. The overall view of these eighteen factors, including value, is given in Table I.

TABLE I
DATA SET

Factors	Value	Factors	Value
Class Performance	Poor / Average / Good	Study Hours	Average / Good
Test Marks	Poor / Average / Good	Living Area	Hostel / Mess / Home
Class Attendance	Poor / Average / Good	Social Media Interaction	Average / Exceed
Due Time assignment Submission	Poor / Good	Extra-curricular Activity	Yes / No
Lab Performance	Poor / Average / Good	Drug Addiction	Yes / No
Previous Semester Result	SCGPA	Financial Support from Family	Yes / No
Family Education	Poor / Average / Good	Political Involvement	Yes / No
Freelancer	Yes / No	Affair	Yes / No
Relationship with faculty	Poor / Average / Good	Year Final Result	CGPA

V. IMPLEMENTATION

In the implementation part of our system, first we need to import all the libraries named Pandas, numpy, matplotlib.pyplot, and sklearn. Then we import the necessary modules, i.e., train_test_split, mean_squared_error & sqrt. We also import the ML library Keras. After that, we take the dataset which we have preprocessed earlier and use it in the Jupyter Notebook. Then we create a deep neural network with two hidden layers. Now we train our deep neural network and a training algorithm can be added so that the network can be trained. In order to train, we use the CNN algorithm. In our exploratory analysis, we divided the dataset into three parts. The first section is the training dataset, which contains 98% of the data. The second part is the testing dataset, where we use only 1%. And most importantly, another 1% of the data is used to validate the predicted results. After taking all these steps, the train function is used to train the network. After running the neural network, we expect a result which will be positive in terms of accuracy. The following Figure 3. shows the implementation procedure.

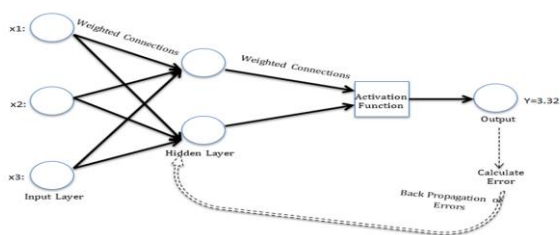


Fig. 3. Deep Neural Network Architecture

VI. RESULT & DISCUSSIONS

After running the deep neural network, we have the predicted result in our hand. Fig. 4. shows the original and predicted result.

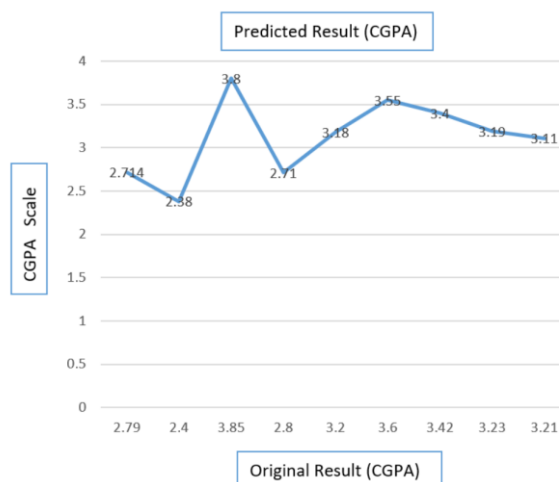


Fig. 4. Original result VS Predicted result.

In our experiment, we used nine students' predictions. In Table II, we show these students' real results and predicted results. Based on these results, we find the accuracy percentage. In table II we show the highest is 99.5% when a student’s original result is 3.42 and we get the lowest accuracy of 97% when the original result is 2.8. After observing these results, we can say that our system is better because for both cases, we got an accuracy that is very close to the original results.

TABLE II
RESULT COMPARISON

Original Result as CGPA	Predicted Result as CGPA	Accuracy
2.79	2.714	97.5%
2.40	2.38	99.5%
3.85	3.8	98.9%
2.8	2.71	97%
3.2	3.18	99.5%
3.6	3.55	98.7%
3.42	3.40	99.5%
3.23	3.19	99%
3.21	3.11	98.2%

In Table III, it showed the comparison of accuracy percentage among different techniques. Here we can see four techniques and the lowest accuracy belongs to the artificial neural network, which is 77.8%, and the highest accuracy belongs to the deep neural network, which is 97%. For neural networks, the accuracy is 90% and for recurrent neural networks, the accuracy is 85.4%. So, according to the above analysis, our model deep neural network is the one with the highest accuracy. So we can say that our model is best for predicting students' performance.

TABLE III
Comparison of Accuracy among Different Techniques

Techniques	Accuracy Percentage
Artificial Neural Network	93%
Recurrent Neural Network	85.40%
Deep Neural Network	97%
Naive Bayes Simple	89.40%
Decision Tree	74%
Data Clustering	74%

VII. FUTURE WORK

We will attempt to incorporate a sizable amount of data from various sorts of instructions into our model in the near future.

REFERENCES

[1] S. Halder, F. Sikder, J. Uddin, "predicting student's yearly performance using neural network: A case study of BSMRSTU", Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on At: Dhaka, Bangladesh Volume: 13-14 May 2016

[2] Kyndt et al. used neural network to predict academic performance of the first year bachelor degree student, based on students' motivation, approaches to learning, working memory capacity and attention.

[3] Ioannis E. Livieris, et al. in their study predicted the performance of students in Mathematics using an Artificial Neural Network (ANN) classifier. They found that the modified spectral Perry trained artificial neural network performs better classification compared to other classifiers in this context.

[4] S. Kotsiantis, et al. investigated in distance learning using machine learning techniques for dropout prediction of students. Their study made an important contribution as it was a pioneer and helped to carve the path for educational data mining

[5] San Pedro et. al analysis a web based tutoring system for mathematics from 3747 school students and predicted whether a student will (5 years later) attend college. Authors learned the students are successful in middle school mathematics as measured by the tutoring system are more likely to enroll 5 years later in college. On the other hand, the students who showed confusion, carelessness in the system have lower probability of college enrollment.

[6] Vihavainen et. al worked with a snapshot data from Computer Science students programming course of Helsinki University and tried to predict whether a Student will fail introductory mathematics course.

[7] Bhardwaj and Pal predicts students' performance and found out living location has high influence on student's final grade. They used Purvanchal University's Department of Computer Applications student's data and used Bayesian Classifier for predicting.

[8] Al-Radaideh, et al guess the final grade of students who studied the C++ course in Yarmouk University, Jordan. Three different classification methods namely ID3, C4.5 and the Naive Bayes are used. The results indicated that Decision Tree model had better prediction than other models.

[9] Nguyen and Peter forecast the performance of the students and compared the efficiency of two classifiers namely Decision Tree and Bayesian networks using WEKA tool. They used two different groups of students of undergraduate and postgraduate level. The performance of Decision Tree was 3-12% more accurate than Bayesian networks. This research was helpful for identifying the weak students for guiding and selecting good students for scholarship.

[10] S. Anupama and Vijayalakshmi expect the performance of final exam of MCA students according to their internal marks. They used C4.5 Decision tree algorithm. They compare the predicted results and actual results which indicates, that there was a significant improvement in results as the

prediction helped a lot to identify weak and good students and help them to score better marks. They also compared the model with ID3 Decision Tree algorithm and prove that the developed model is better in terms of efficient and time taken to build the decision tree.

[11] Smith et. al. proposes visible school security measures by student academic performance, attendance, and post-secondary aspirations.

[12] Baris Cetin introduce an approaches to learning and age in predicting college students' academic achievement. Enhanced higher order orthogonal iteration algorithm for student performance prediction has been proposed by Prema et. at [13].

[13] Han, pei and Kamber, Data Mining Concepts and Techniques, The Morgan Kaufmann series in data management systems, 3 rd edition, 2011.

[14] S. K. Yadav, B. Bharadwaj, and S. Pal, "Data mining applications: A comparative study for predicting student's performance," arXiv preprint arXiv:1202.4815, 2012

[15] Scheuer et al., Educational data mining. In Encyclopedia of the sciences of learningpp. 1075–1079, Springer, 2012.