Student Academic Growth System using Machine Learning

[1]Juili Kadu, [2]Hrutwika Ambavane, [3]Deepali Surekar, [4]Prof. Sachin Chavan

[1, 2, 3] *Department of Computer Engineering,*
Mahatma Gandhi Mission College of
Engineering and Technology, Kamothe, Navi
Mumbai, Maharashtra, India.

Abstract—**In today's world, students' success plays a vital role in his education and career. Early prediction of their academic performance can help the students as well as teachers to find the problems and help future improvements. In this paper, we are trying to build a system that predicts the student's performance and failure risk using various ML Algorithms. Along with this we have also built a prediction model to find the depression level of students as due to pandemic a lot of students are suffering from mental health issues. This is implemented using a Django web-based application. By early detecting students' issues teachers can help them to upgrade their studying skills and improve their academics.**

Index Terms— **ML, Django, Students Performance.**

I. INTRODUCTION

In this era, education is considered to be one of the most important investments that a country makes for its own future. It is said that youth is foundation of future With the corona-virus outbreak life has disrupted around the globe in 2020, the educational systems have been affected in many ways. For providing every student proper education and keeping check on their mental is getting so much important. And here we have proposed a web-based system which perform prediction student performance. If teacher contain knowledge of students result before the exams, teacher can take the necessary steps to improve the students' performance [1].

In the life of student's grades are not the only one thing they have to focus. They may go through lot of mental stress as well. This may lead to frustration and student may not be able to do his studies properly. Moreover, they are likely to burden themselves during exams and as a result they score less marks. So, for this problem we have Depression Detection test for them. Because of their daily routine disturbance students go through stress and anxiety. Reading blogs will improve their positive outlook towards life. This will enhance overall development of a student.

Currently, there are many techniques being compare to evaluate students' performance. Machine learning is part of the Artificial Intelligence (AI), where computer can teach themselves to learn the data. While data mining is a technique to find pattern in large amount of data[3]. In the field of education Machine Learning is quite widely used. The useful information and patterns are often utilized in predicting students'

performance. As a result, it would assist the educators in providing an effective teaching approach. Besides, educators could also monitor their students' achievements. Students could improve their learning activities, allowing the administration to enhance the systems performance [2]. This paper assumes that this web-based application would be a prototype that would shape the future and there still remains much to do in terms of improvement of the existing models.

II. RELATED WORK

In Existing systems, the datasets used are using only marks as a parameter to predict the students' performance academically based on which the output is considered. The colleges need to keep a lot of paperwork to keep record of student's performance from which it is difficult to analyze the results. This system also has no idea about students' mental health which can affect their performance.

K.Nithya,M Phil Research Scholar1, Dr.V.Narayani, Assistant Professor2 did a comparative study of Machine Learning Approaches on Prediction System of Students Academic Performance. They compared three algorithms KNN, NB and SVM which gave the accuracy of 78, 81 and 85 percentage respectively based on their dataset.

Vedang Gharat1, Saurabh Rathod2, Chandan Saw3, Aditi Malkar4 made a STUDENT PERFORMANCE PREDICTION SYSTEM using C4.5 algorithm. The result showed that multilayer perceptron gives good prediction results than linear regression.[4]

Shubha Pawar1, Omkar Jadhav2, Shubham Kadam3, Pratik Jadhav4 Prof. Sneha Deshmukh5 did "Predicting Student's Performance For Early Prevention using Supervised Machine Learning Algorithm". They used SVM, Naïve Bayes, Logistic Regression, Decision Tree, KNN and Random Forest Classifier. Out of all the SVM gave the highest accuracy of 87.26% based

on the previous year data.[1]

III. PROPOSED WORK

The aim of this project is to improve the current trends in the higher education systems and to seek out which factors might help in creating successful students.[5].

Steps for implementing Machine Learning Algorithm:

1. Data Collection and Preprocessing:
   The first and foremost step is collection of data. Data maybe present on different spreadsheets in database. Data has to be collected and cleaned in order to build more accurate model. So, we remove noisy data, missing data and inconsistent data in this step.
2. Feature extraction:
   Finding the features that are most suitable for our model is an important step in machine learning. Based on the dataset we have extracted 14 features that give the higher accuracy in model building.
3. Prediction Model Generation:
   In this module the prediction model is made by using various several machine learning approaches. Among several machine learning approaches, this work uses K Nearest Neighbor, Naive Bayes and Support Vector Machine as prediction model generation.[2]

   Classification

   Classification algorithm is a data mining technique that helps us to map data into predefined category. It is a supervised learning technique which needs categorized training data so it can create rules for categorizing test data into pre-arranged category. [6] Its a 2-phase process. The first phase because the learning phase, where the classification rules are generated and training data is analyzed. The second phase because the classification phase, where test data is classified into predefined groups according to the generated rules.[5] Based on this classification we have classified the students' performance as "Failure, Poor, Good, Satisfactory and Excellent ."

   For depression detection we have used SVM Classifier model. On the basis of set of questionnaires, it predicts that the depression level ranging from No Depression, Mild Depression, Moderate Depression, moderately severe Depression and Severe Depression.

IV. METHODOLOGY

We are using three sorts of classification models so as to learn the predictive function which is required. The models are used for experimental analysis. They are selected on the basis of their frequent usage within the existing literature. The list of methods are as follows: [5]

1. Logistic Regression:
   Logistic regression is an example of supervised learning to predict the probability of a target variable. It is used to calculate or predict the probability of a binary (yes/no) event occurring but also can be used for ordinal classification like good, very good and excellent.
2. Support Vector Machine:
   Support Vector Machine is used for classification which is also a supervised learning method. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of every feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the 2

classes finely.

3. Random Forest Algorithm:
   A random forest represents a meta estimator that matches variety decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. Random Forest is a classifier that contains a variety of decision trees on various subsets of the given dataset and takes the average to enhance the predictive accuracy of that dataset.
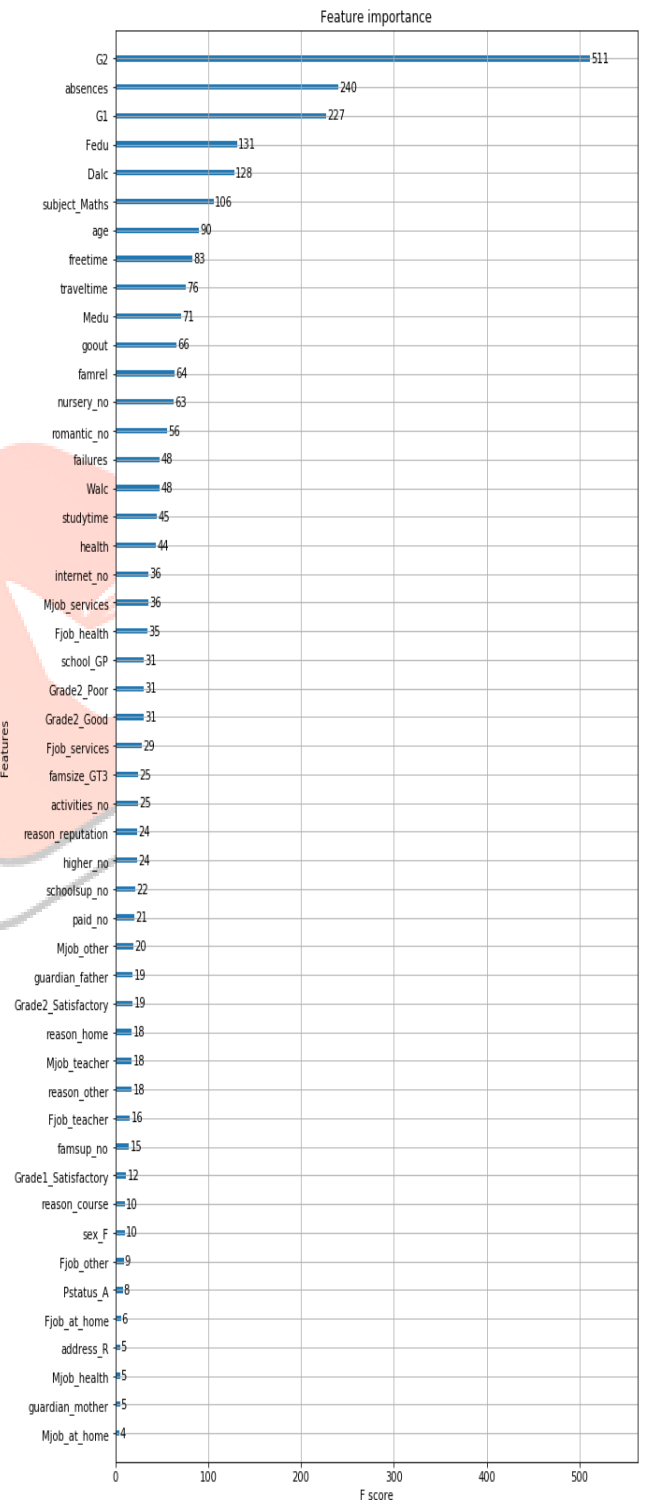

Fig 1. Feature Importance

Model Performance

Our model got 84% using Logistic Regression and 85% accuracy using Random Forest Classifier and 83% accuracy

using Support Vector Machine on predicting student performance.

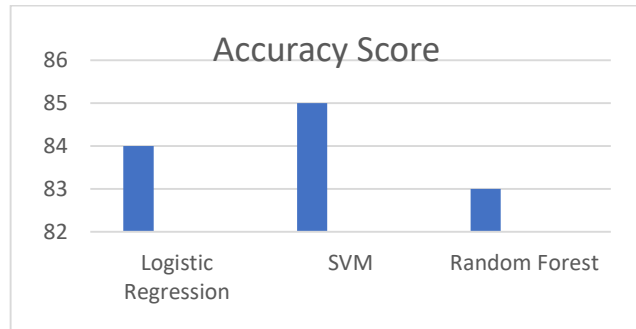| Model	Accuracy Score
Logistic Regression	0.84
Random Forest	0.85
SVM	0.83

Fig.2 Accuracy Score

V. IMPLEMENTATION DETAILS

A. Application Features

The system has following features-

- This system will allow only the authenticated users to use the web application.
- This system will help student and teacher to keep track on students' academic progress
- This system allow student to schedule their task
- The system provide facility to check depression level
- This system helps to boost mental health of student
- The system keeps track on number of people using app
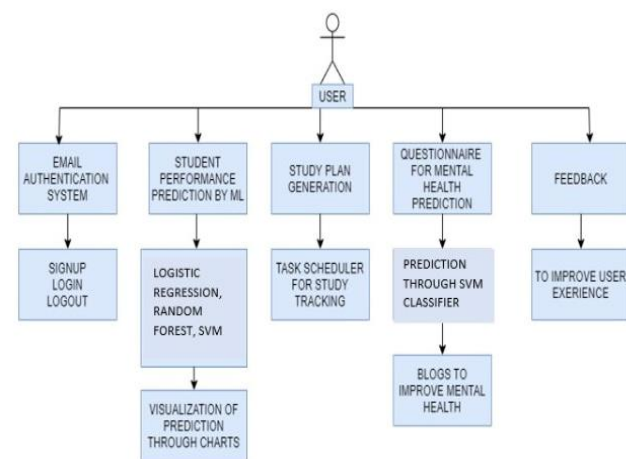- Helps teacher to personally look into each student academic growth

Fig.3 System workflow

B. Technologies Used

- Django
- SQLite3
- Python
- HTML
- CSS

Django – Django is free ,open-source Python based web framework for rapid development of web apps quickly and with less code. It is also called as batteries included framework because it provides: development WebServer,Forms,DatabaseSQLite3,Models,admin panel, User Authentication

SQLite3 – SQLite3 is a database engine return in a C language . It is an opensource platform. As it is not a standalone platform, we have used it with Django for storing our data

Python – Python is simple, high-level, interpreted, procedure and object-oriented language and free to use general purpose programming language. Python is simple and easy return in English like statements which implies more clarity and a compact code . Python program can be created using function and also using classes and object. Python compiler generate byte code which can be run on any system having python virtual machine. python is dynamically typed i.e. No data type to be specified.

HTML - HTML is abbreviation of Hyper Text Markup Language. Used for designing documents to be displayed on web browser. HTML describes structure of web page semantically and originally include cues for appearance of the document

CSS – CSS is abbreviation of Cascading Style Sheets is a stylesheet language that describes the presentation of a document written in HTML AND XML.

V. RESULTS

We have implemented our algorithms with the assistance of Python and integrated it with Django. We have made use of built-in python libraries and packages to implement our classification algorithms. We have made use of the subsequent libraries and packages:

1. NumPy
2. Pandas
3. Scikit-learn
4. Matplotlib
5. Django

A login window from where users as well as admin can login into their respective accounts and continue with the further access.
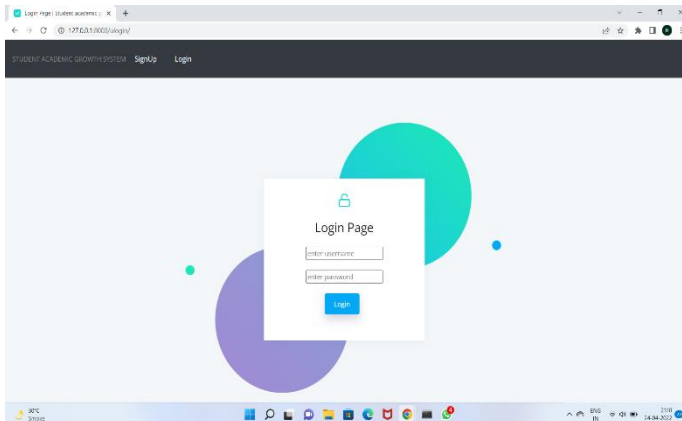
depression.

Fig. 4: Snapshot of the login window

In this web app, it shows all the app's main options for the navigation.
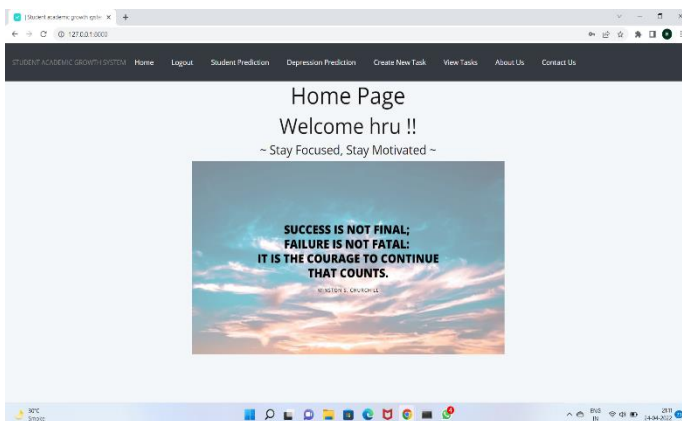

Fig. 5: Snapshot of the Dashboard

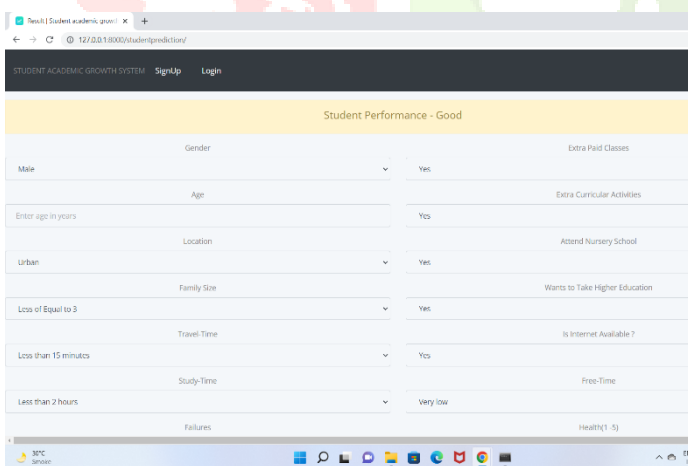This is the output page for student performance prediction.

Fig. 6: Snapshot of student prediction

This is the depression detection page where a set of questionnaires is given to the user and the output is predicted ranging from no depression to sever
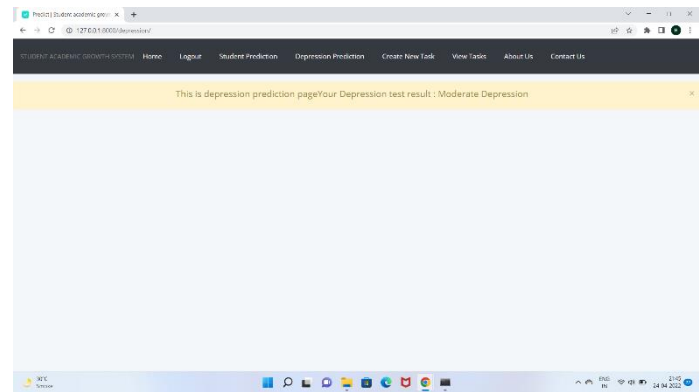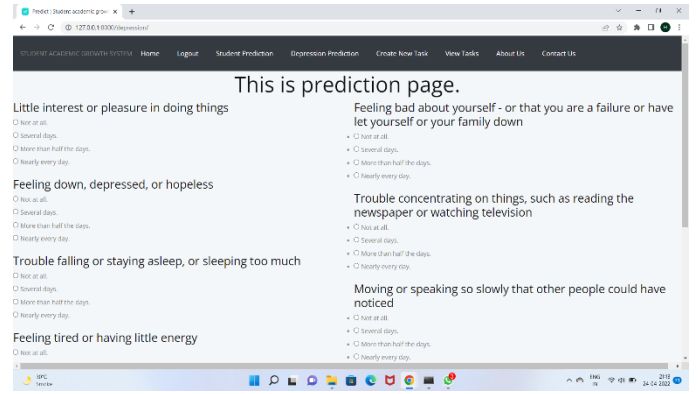

Fig. 7: Snapshot of depression prediction

VII. CONCLUSION

Prediction of students' academic progress plays a key role for teachers and students as well as the institutions. Early detection of student success and failures help to find out in which areas the student is lagging behind. Here we have used random forest, logistic regression and support vector machine algorithm to classify the student performance. Also, we have used the SVM classifier for depression detection in students based on their habits. So, the student prediction and depression detection can be a warning for parents as well as teachers to know about the child's progress.

In the future if we collect more data, we will get more accurate result by applying different algorithm using machine learning. Furthermore, after detecting mental health, we can provide live counselling and is also possible to get a better idea of illness from facial expression and analysis using AI. Also, this research can help us to explore in many fields.

VIII. REFERENCES

[1] Shubha Pawar1, Omkar Jadhav2, Shubham Kadam3, Pratik Jadhav4 Prof. Sneha Deshmukh5," Predicting Student's Performance for Early Prevention using Supervised Machine Learning Algorithm" International Research Journal of Engineering and Technology (IRJET) Volume: 08 Issue: 05 | May 2021.

[2] K.Nithya,M Phil Research Scholar1, Dr.V.Narayani, Assistant Professor2, "A Comparative Study of Machine Learning Approaches on Prediction System of Students Academic Performance" International Research Journal of Engineering and Technology (IRJET) VOLUME: 07 ISSUE: 05 | MAY 2020

[3] I. No, S. Anam, and S. Gupta, "A research Review on Comparative Analysis of Data Mining Tools, Techniques and Parameters," Int. J. Adv. Res. i, vol. 8, no. 7, pp. 523– 529, 2017.

[4] Vedang Gharat1, Saurabh Rathod2, Chandan Saw3, Aditi Malkar4, "Student Performance Prediction System", Volume: 07 Issue: 05 | May 2020

[5] Durgesh Ugale1, Jeet Pawar2, Sachin Yadav3, Dr. Chandrashekhar Raut4, "STUDENT PERFORMANCE PREDICTION USING DATA MINING TECHNIQUES", Volume: 07 Issue: 05 | May 2020

[6] K. V. J.K. Jothi Kalpana, "Intellectual performance analysis of students by using data mining techniques", International Journal of Innovative Research in Science, Engineering and Technology, vol. 3, 2014.