# Abstract

Due to the pandemic spread of COVID-19, the only substitute of delivering education was the online teaching mode. A large number of schools, colleges and universities had to move on online teaching mode. In this regard, there was a need to investigate the quality of education by evaluating the performance of teacher. There is also need to investigate the quality of education in term of student assessment. In our project, we proposed a framework based on historical dataset to evaluate the performance of teacher and student by using different parameters. The ultimate goal of this study is to identify the most correlated parameters of teacher and student assessment. Now in this report, we are going to validate our proposed framework base on the different examples set. Different evaluation measures including the R2 Score and Root Mean Squared Error (RMSE) will be calculated to measure the performance of proposed framework. After the validation of proposed framework, the complete procedure of deployment will be a part of this report.

# Introduction

During the pandemic of COVID-19, the schools and universities around the world change their education style from physical to online mode. Many countries have launched policy and support to enhance online teaching from primary to university education. Online education has become the only substitute for many universities in this special circumstance. In this regard, there is need to analyze the quality and effectiveness of the online offered course. The delivery of online course through digital platforms, enables the authorities to store the historical data that can be used for different analysis tasks. The collection of historical data never easy in physical learning mode. Resultantly, the delivery of education via online courses opened the new opportunities. In particular the evolution of data analytics and machine learning allows these data to be collected, analyzed and possibly identify areas where improvements can be made.

To ensure the feasible and accurate calculation of teachers and students performance score we proposed a framework that was based on the training of different machine learning models by using historical data. Now there is need to validate the performance of trained models. In this regard, we chose a set of testing example from the dataset to make predictions from models and evaluate the performance of complete framework. To evaluate the prediction results, different evaluation measures including R2 Score, Mean Squared Error (MSE), RMSE, and Mean Absolute Error (MAE) was calculated. Lastly, the best performed trained model was deployed by developing the user interface for friendly and interactive use of proposed framework.

# Validation of Models

## Dataset Description

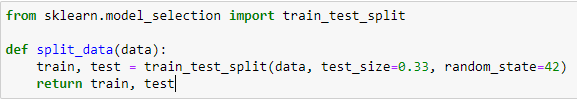
The dataset was prepared for the performance prediction of teachers and students by using python script. The students and teacher’s dataset were divided into four categories including the Bad, Average, Good and Excellent category of students and teachers. The python script generated the 3000 samples in students and teacher’s dataset for each category. The teacher’s dataset was based on four input variables with one target variable while the student dataset was based on # input variables with one target variable. The generated datasets were chosen for the performance prediction of teachers and students in the proposed project.

The original Dataset contains the following number of rows and Columns:

|  |  |  |
| --- | --- | --- |
| **Teachers Dataset** | **Number of Rows in Dataset:** | 12000 |
| **Number of Variables in Dataset:** | 5 |
| **Students Dataset** | **Number of Rows in Dataset:** | 12000 |
| **Number of Variables in Dataset:** | 14 |

## Train Test Split of Dataset

The dataset of teachers and students’ performance was split into training and testing set with the ratio of 70% and 30% respectively. The train test split function of Sklearn library was used to split the both datasets into training and testing set. The code sippet of splitting data into training and testing set is following:



## Test Set

After the split of both datasets, the training data was used for the training of the models (Chapter 2) while the testing data will be used here for the validation of trained models. The testing set have the 3960 number of samples in teacher and student dataset. The details of testing set are given in below table.

|  |  |  |
| --- | --- | --- |
| **Teachers Dataset** | Number of Samples in testing set | 3960 |
| Total attributes in testing set | 6 |
| Target Variable in testing set | Teacher Score |
| **Students Dataset** | Number of Samples in testing set | 3960 |
| Total attributes in testing set | 14 |
| Target Variable in testing set | Student Score |

## Target Variable Description

The attribute “Teacher-Score” and “Student-Score” were set as target variable in teachers and students’ test set respectively. The possible range of teachers-score is 10 to 100 while the possible range of students-score was 20 to 100. The target variable was used for the prediction of teachers and students’ performance based on the other attributes in both datasets.

## Evaluation Measures

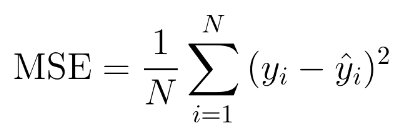
For the validation of the models, we chose some evaluation measures to measure the performance of our trained model. Evaluation measures were selected based on the nature of proposed problem. As the evaluation score prediction of teachers and students is a regression problem, that’s why we chose all the evaluation measures that are used for regression problems. The detail description and formula of calculating each evaluation measure is mentioned below:

**R2 Score -** The R2 score is a very important metric that is used to evaluate the performance of a regression-based machine learning model. It is pronounced as R squared and is also known as the coefficient of determination. It works by measuring the amount of variance in the predictions explained by the dataset.

**R2 = (var(mean) - var(line)) / var(mean)**

**Mean Squared Error (MSE) –** The Mean Squared Error (MSE) is perhaps the simplest and most common loss function, often taught in introductory Machine Learning courses. To calculate the MSE, you take the difference between your model’s predictions and the actual value, square it, and average it out across the whole dataset.

The MSE will never be negative, since we are always squaring the errors. The MSE is formally defined by the following equation:

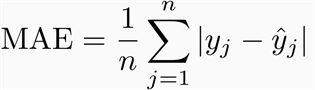


**Root Mean Squared Error (RMSE) -** RMSE is the standard deviation of the errors which occur when a prediction is made on a dataset. This is the same as MSE (Mean Squared Error) but the root of the value is considered while determining the accuracy of the model.



**Mean Absolute Error (MAE) -** The Mean Absolute Error (MAE) is only slightly different in definition from the MSE, but interestingly provides almost exactly opposite properties! To calculate the MAE, you take the difference between your model’s predictions and the actual value, apply the absolute value to that difference, and then average it out across the whole dataset.

The MAE, like the MSE, will never be negative since in this case we are always taking the absolute value of the errors. The MAE is formally defined by the following equation:



## Validation

All models were tested on test set of the dataset to evaluate the performance of trained models. The input variables of students and teachers test set were passed to each model and make prediction on each sample. Every model returns a list of predations base on its learning equal to the length of testing examples (rows). The predicted scores were labeled as y’ while the actual scores were labeled as y. The score of all evaluation measures were calculated with y (actual scores) and y’ (predicted scores) by using the relative formulas (described above).

**Linear Regressor Validation:** The input variables of all the testing examples were passed to the Linear Regressor model and it return the predicted scores list (y’) equal to the length of testing examples (3960). The actual scores/target variables (y) and predicted scores (y’) were passed to the evaluation measure function that calculate the R2 score, MSE, RMSE, MAE by using the equation 1-4 respectively.Linear Regressor Model showed the 0.917555, 7.727267, 59.71066, and 6.362078 values of R2 Score, RMSE, MSE, MAE for teachers’ dataset respectively. The Linear Regression model also showed the 0.930227, 6.193569, 38.36029, and 5.268767 values of R2 Score, RMSE, MSE, MAE for students’ dataset respectively.

**Random Forest Regressor Validation:** The input variables of all the testing examples were passed to the Random ForestRegressor model and it return the predicted scores list (y’) equal to the length of testing examples (3960). The actual scores/target variables (y) and predicted scores (y’) were passed to the evaluation measure function that calculate the R2 score, MSE, RMSE, MAE by using the equation 1-4 respectively.Random ForestRegressor Model showed the 0.92693, 7.274672, 52.92085, and 6.0097 values of R2 Score, RMSE, MSE, MAE for teachers’ dataset respectively. The Random ForestRegression model also showed the 0.936384, 5.913984, 34.97521, and 5.094838 values of R2 Score, RMSE, MSE, MAE for students’ dataset respectively.

**K-Nearest Neighbor Validation:** The input variables of all the testing examples were passed to the K-Nearest Neighbormodel and it return the predicted scores list (y’) equal to the length of testing examples (3960). The actual scores/target variables (y) and predicted scores (y’) were passed to the evaluation measure function that calculate the R2 score, MSE, RMSE, MAE by using the equation 1-4 respectively. K-Nearest NeighborModel showed the 0.928033, 7.219567, 52.12214, and 5.948182 values of R2 Score, RMSE, MSE, MAE for teachers’ dataset respectively. The K-Nearest Neighbormodel also showed the 0.928191, 6.283286, 39.47969, and 5.295505 values of R2 Score, RMSE, MSE, MAE for students’ dataset respectively.

**Decision tree Regressor Validation:** The input variables of all the testing examples were passed to the Decision treeRegressor model and it return the predicted scores list (y’) equal to the length of testing examples (3960). The actual scores/target variables (y) and predicted scores (y’) were passed to the evaluation measure function that calculate the R2 score, MSE, RMSE, MAE by using the equation 1-4 respectively.Decision treeRegressor Model showed the 0.877154, 9.432427, 88.97068, and 705246 values of R2 Score, RMSE, MSE, MAE for teachers’ dataset respectively. The Decision treeRegression model also showed the 0.872907, 8.359081, 69.87424and 6.832828 values of R2 Score, RMSE, MSE, MAE for students’ dataset respectively.

## Validation Results

After the validation of all the models using test set of teachers and student data, the models were compared base on the evaluation scores. Random Forest Regressor Model Showed the 0.92693 and 0.936384 highest R2 Score for teachers and student test data respectively. The Random Forest model also showed the lowest error for teachers and student test data. The comparative table of evaluation measures for teacher and student test set is available below:

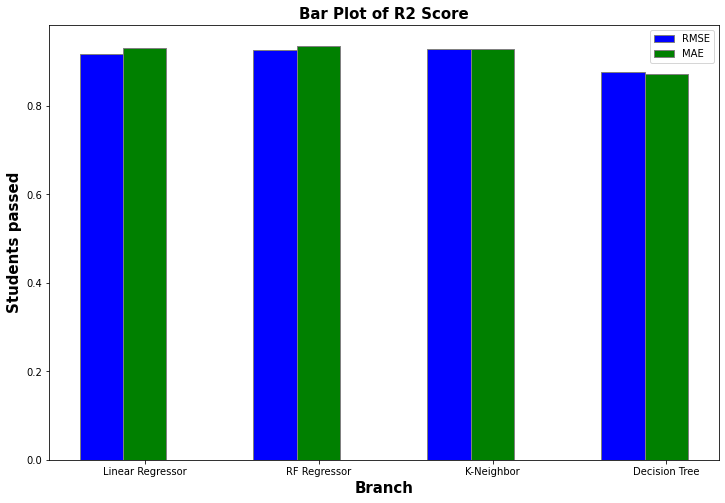
Table : Evaluation Measures report of teachers test set.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | R2 Score | RMSE | MSE | MAE |
| Linear Regressor | 0.917555 | 7.727267 | 59.71066 | 6.362078 |
| Random Forest Regressor | 0.92693 | 7.274672 | 52.92085 | 6.009769 |
| K-Nearest Neighbor | 0.928033 | 7.219567 | 52.12214 | 5.948182 |
| Decision tree Regressor | 0.877154 | 9.432427 | 88.97068 | 7.5246 |

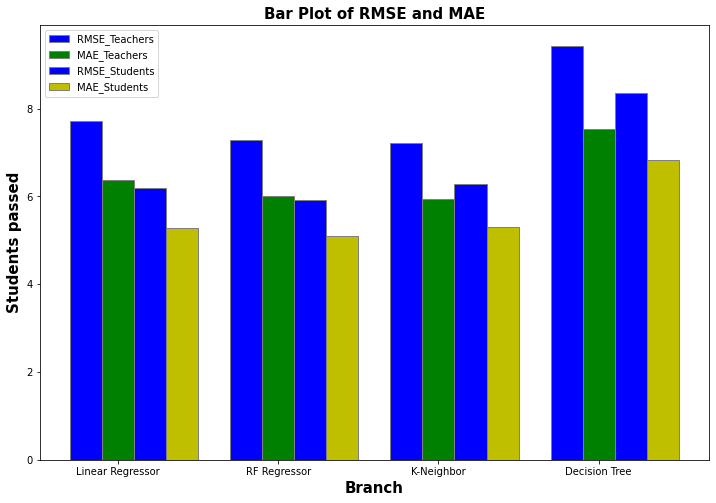
Table : Evaluation measure report of student test set.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | R2 Score | RMSE | MSE | MAE |
| Linear Regressor | 0.930227 | 6.193569 | 38.36029 | 5.268767 |
| Random Forest Regressor | 0.936384 | 5.913984 | 34.97521 | 5.094838 |
| K-Nearest Neighbor | 0.928191 | 6.283286 | 39.47969 | 5.295505 |
| Decision tree Regressor | 0.872907 | 8.359081 | 69.87424 | 6.832828 |

The below bar plot is also showed that Random Forest Regressor showed the best R2 Score for teachers and students’ dataset.



The below bar plot is also showed that Random Forest Regressor showed the lowest RMSE and MAE error Score for teachers and students’ dataset.



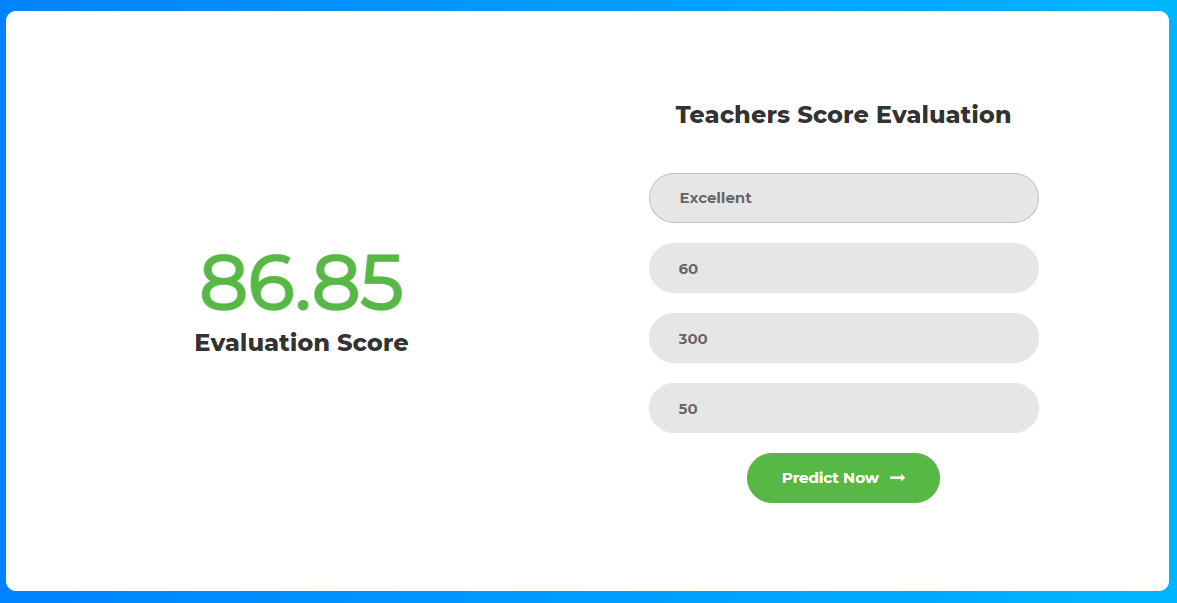
By the study of comparative table of evaluation measures and bar plot of evaluation measures, it is clear that Random Forest Regressor outer perform by all models in term of R2 Score and other evaluation measures. The best result conclude that Random Forest is the best model for deployment in real world environment to make predictions on real world data.

# Deployment

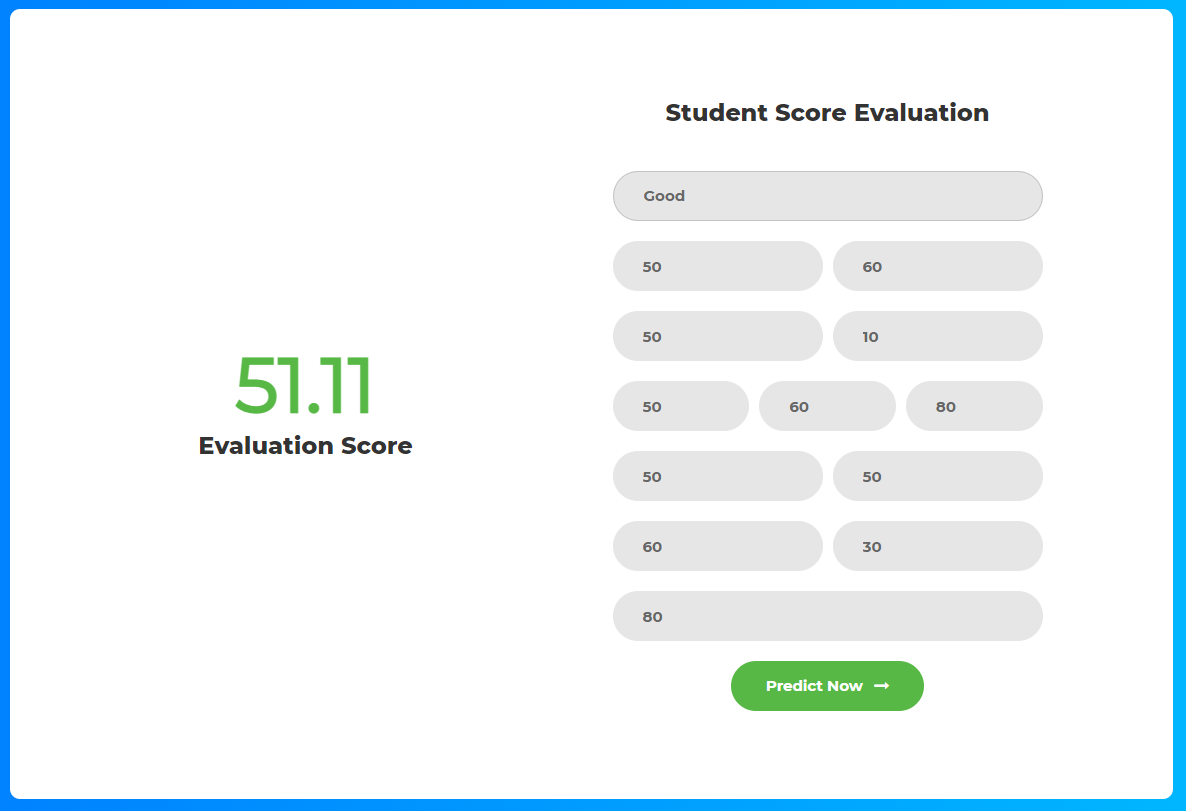
After the validation of all models, results conclude theta Random Forest Regressor is the best model for the evaluation score prediction of teachers and students. We made an interface for getting the data from user and made prediction on the collected data by Random Forest Regression. The Deployment of the best model is base on two phases: front-end phase and backend phase. The detail working of each phase is described below:

## Frontend Phase

In the frontend phase, we developed the 2 screens for getting the teachers and students input variable data respectively. The teacher interface screen is based on the 5 input fields followed by the submit button while the student interface was base on the 14 input fields based on the input variables of student dataset followed by the submit button. In the teachers Interface Screen, one input field is the dropdown filed that take categorical values from dropdown while rest of the fields are numerical fields that take integral values. In the last of the screen a submit button is place to submit the user entered data at server side. The overview of teacher’s screen is below:



In the student interface screen, one field is dropdown field that take categorical values from user while rest of the fields take the numerical values. There is also a submit button to send the user entered value at the server side. The snippet of Student Screen is presented below:



## Backend Phase

The backend phase base on several functions with the two main functions that calculate the teachers and students’ evaluation score. Both functions received the form data from user interface via hyper parameters. After converting the categorical data into numerical representation using one hot encoding, both functions send the converted to Random Forest Regressor model that calculate the evaluation encode bae of the input variables. After the calculation of evaluation score, both functions return the score towards the frontend interface in JSON format. The interface is capable to receive the score and show to the user as show in above figures.

## Configurations

The frontend and backend phases were developed by using the Django Framework. The Django Framework of Python is the secure and stable framework for the development of user friendly, secure and attractive web application. The MVC (Model, view, Controller) architecture was followed to developed the whole application. After the complete development of application, the application was deployed on the local server with 127.0.0.1 IP address. Minimum 4GB Ram and 500GB are required to run the application on local server.

After the complete deployment of application, the application was tested by entering the values in the form and calculated the teachers and students’ evaluation scores. After, clicking of the submit button, proposed model, and developed application perform well and showed the evaluation scores for both teachers and students also as shown in above figures.

# Discussion

In the proposed study, we trained the different machine learning models for different features of the students and teachers. The identification of most correlated features for the better MSE of the machine learning model was a biggest challenge. Sometimes, teachers are unnecessarily set the paper or assignment as very easy and very hard. So, the impact of that exam and assignment can’t directly translate the final grade of the student. Like the assignment 1, assignment 2 and assignment 3 in our study are less correlated with the final output score of the students. The calculation of correlation score required the mapping of each feature against the final score that is very difficult and time consuming for large amount of data. Secondly the historical data for the teacher and student score evaluation is no longer available publicly. The generation of the student and teacher historical data was the main challenge that face during the implementation of the proposed study. The quality assurance of collected data on the basis of quality parameters is further challenging.

Here, we generated our own historical data for the teacher and student final grade prediction. The proposed study contributes in the final assessment of the student score and teacher score for online offered courses. Although, we use correlation calculation technique for the selection of features, but our dataset was limited to few features. The dataset can be updated by adding more features of teachers and students that may or may not be correlated with the final score. However, future studies can update the dataset with more features to more accurately predict the final grades. Moreover, for the large set of features data, the feature selection approach can be used for the selection on best features for ML models after correlation process. For the limited features, the proposed study performs well and the MSE scores showed that the model is robust enough to deploy in real world environment.

# Future Directions and Scope

In this project, we trained the different machine learning for the performance assessment of teachers and student in online teaching mode. The proposed results showed the result that are robust enough for the assessment of teachers and students in online teaching model. In online teaching mode, teachers and student related data like online teaching time of teachers and duration of attending class by student can be collected easily. The rest of the performance data like assignment marks and exam marks of students can be collected from the LMS of concerned organization department. After compiling the data of students and teachers, the evaluation score of each teacher and student can be calculated one by one as show in below figures.

The automatic calculation of finding the evaluation score for teachers and students will reduce the effort of generating complex criteria of calculating score for teachers and students. The proposed project will also reduce the human effort to calculate the score manually for each teacher and student by selected criteria. Most prominently, the teacher’s evaluation usually calculated by taking the feedback from students and the totally score of teachers is depended on the feedback of students. This calculated score usually don’t cater the personal interest of students in bad and good feedback. The proposed project also contributing the other measures to calculate the teachers and students score that can be collected automatically and reduce. Hence the proposed project will reduce the dependency of teacher and students’ personal interests on giving tine of feedback and calculate the performance score totally based on performance measure.

