**Predicting Student Success in Online Learning Environments using Machine Learning**

**Abstract**

Recent years have witnessed an increased interest in online education, both massive open online courses (MOOCs) and small private online courses (SPOCs). This significant interest in online education has raised many challenges related to student engagement, performance, and retention assessments. With the increased demands and challenges in online education, several researchers have investigated ways to predict student outcomes, such as performance and dropout in online courses. This paper presents a comprehensive review of state-of-the-art studies that examine online learners’ data to predict their outcomes using machine and deep learning techniques. The contribution of this study is to identify and categorize the features of online courses used for learners’ outcome prediction, determine the prediction outputs, determine the strategies and feature extraction methodologies used to predict the outcomes, describe the metrics used for evaluation, provide a taxonomy to analyze related studies, and provide a summary of the challenges and limitations in the field.

**Existing system**

In Existing system, using machine learning algorithms like Naïve Bayes, Logistic Regression to predict student success in online learning environments.

**Disadvantages:**

1. Less Accuracy
2. More time taking process

**Proposed system**

In this proposed system, we are employing OULAD dataset to predict grade and score using various machine learning algorithms like Random Forest and gradient boosting. Each algorithm performance is evaluated using accuracy, precision, recall and FSCORE.

Random Forest classifier is used to predict grade and Random Forest Regressor is used to predict score. Similarly Gradient Boosting classifier to predict Grade and Gradient Boosting Regressor to predict score. Regressor algorithm performance is evaluated using RMSE (root mean square error). RMSE refers to difference between true value and predicted value so the lower the difference the better is the value.

**Advantages:**

1. High Accuracy
2. Takes less time

**HARDWARE & SOFTWARE REQUIREMENTS:**

**HARDWARE REQUIRMENTS:**

* processor :   intel i3(min)
* Hard Disk  :   40 GB.
* Floppy Drive :   1.44 Mb.

**SOFTWARE REQUIRMENTS:**

* Operating system : Windows 10 (min)
* Coding Language  : python 3.7.0