**Prediction of Student Performance in**

**University Admission Systems**

**Introduction:**

Today, all higher education institutions, especially computer and engineering colleges, face challenges in the admissions process. Each university should strive for an admissions system based on valid and reliable admissions criteria that select candidates likely to succeed in its programs. In addition, each university should use the best possible techniques for predicting applicants’ future academic performance before admitting them. This would support university decision makers as they set efficient admissions criteria. However, most higher education institutions face challenges when they analyze their large educational databases to predict students’ performance. This is because they use only conventional statistical methods rather than new and efficient predictive techniques such as Educational Data Mining (EDM), which is the most popular technique to evaluate and predict students’ performance. EDM is the process of extracting useful information and patterns from a huge educational database, which can then be used to predict students’ performance. As a result of better information, student performance can be more effectively improved through more effective strategic programs.

**About abstract:**

An admissions system based on valid and reliable admissions criteria is very important to select candidates likely to perform well academically at institutions of higher education. This study focuses on ways to support universities in admissions decision making using data mining techniques to predict applicants’ academic performance at university. The results demonstrate that applicants’ early university performance can be predicted before admission based on certain pre-admission criteria (high school grade average, Scholastic Achievement Admission Test score, and General Aptitude Test score). The results also show that Scholastic Achievement Admission Test score is the pre-admission criterion that most accurately predicts future student performance

**Existing Methods:** In Existing methods, we are using Decision Trees, Support Vector Machines, and Naïve Bayes will help classifying admitted applicant who will accepted and not accepted. Algorithms has less performance and accuracy.

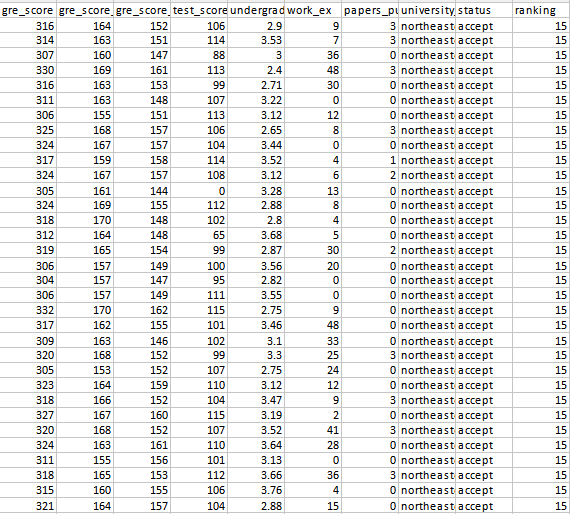
**Proposed Method:**

In this paper, we are proposing ANN algorithms which will increase the model performance and over and underfitting

**Dataset Location:** <https://www.kaggle.com/c/datasciencebowl/data>

The dataset consists of 480 student records and 16 features. The features are classified into three major categories: (1) Demographic features such as gender and nationality. (2) Academic background features such as educational stage, grade Level and section. (3) Behavioral features such as raised hand on class, opening resources, answering survey by parents, and school satisfaction.

The dataset consists of 305 males and 175 females. The students come from different origins such as 179 students are from Kuwait, 172 students are from Jordan, 28 students from Palestine, 22 students are from Iraq, 17 students from Lebanon, 12 students from Tunis, 11 students from Saudi Arabia, 9 students from Egypt, 7 students from Syria, 6 students from USA, Iran and Libya, 4 students from Morocco and one student from Venezuela



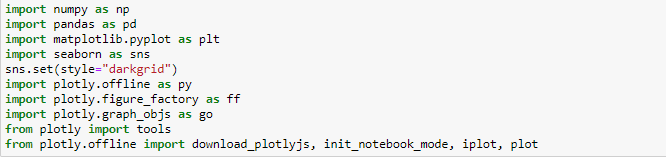
**Project Development Modules:**

1. **Data collection:** Collect the data from Kaggle
2. **Data Preparation:** Load the dataset and normalize the data, cleaning
3. **Data Split:** Split the data in train and test in the ration 8:2
4. **Model:** Machine Learning (RF, SVM,DT) algorithms. Train the model with training data
5. **Evaluation:** Test the model with testing data.
6. Predict on classes using trained model

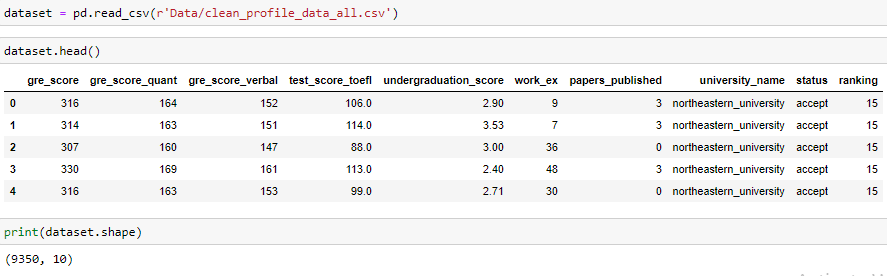
**Extension or improvements in project:** Increase the accuracy of model or implement of using CNN model

**Conclusion of project:** We will build the machine learning algorithm model to predict whether student performance(accepted or not)

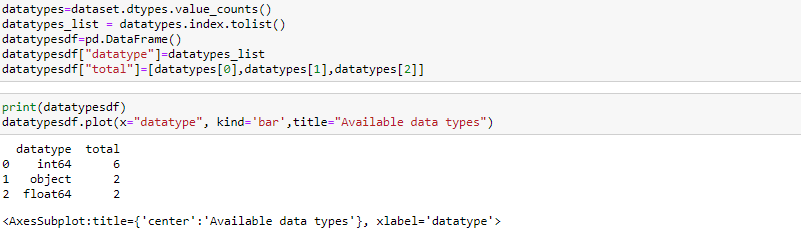
Importing required packages:



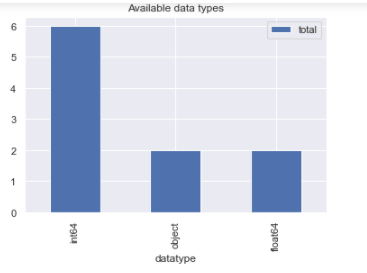
Read the data

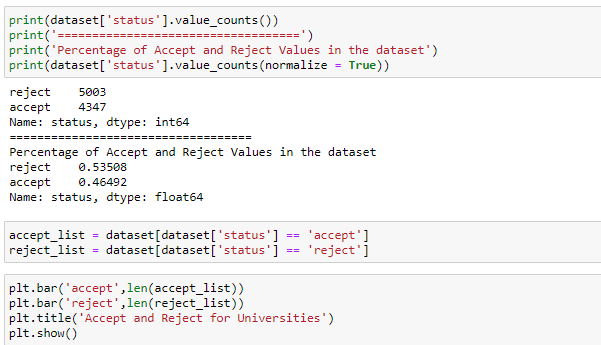


Preprocessing

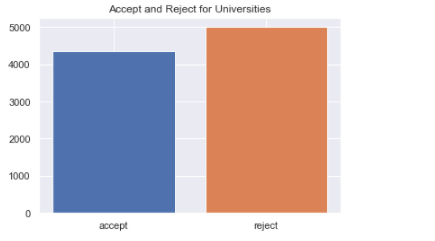


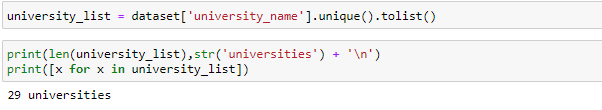
Visualization





Visualization

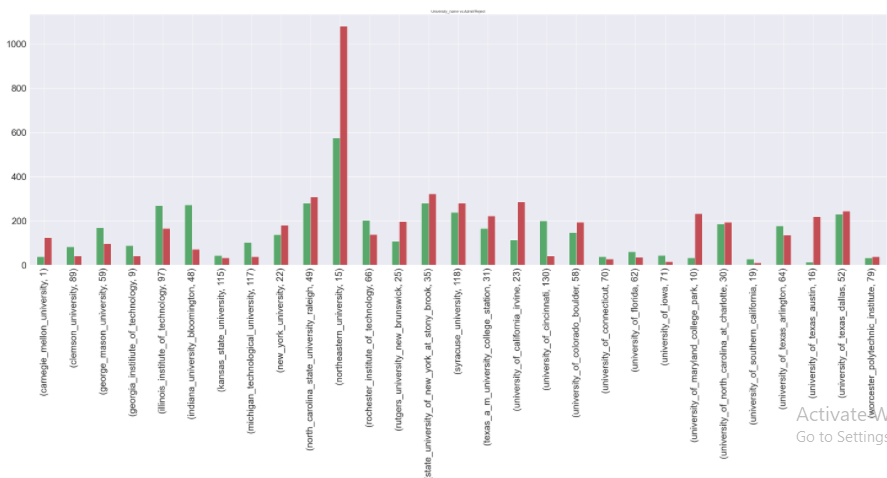


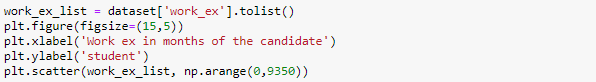


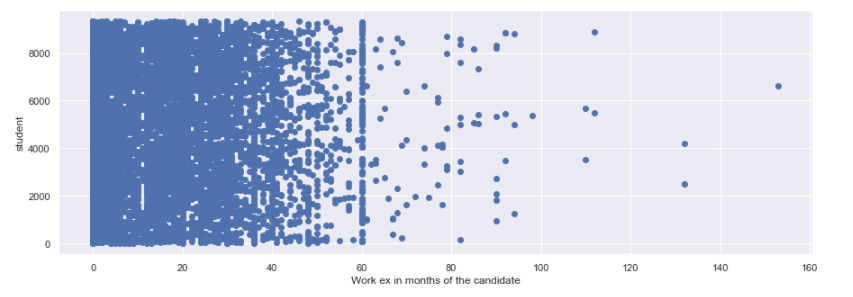


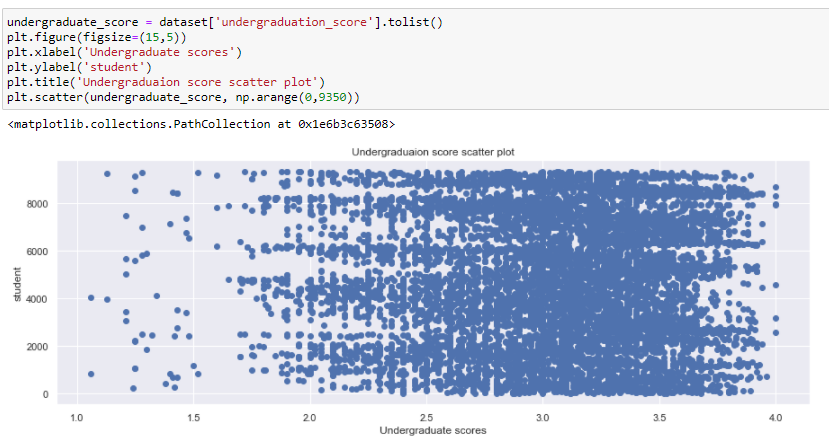
Visualize data for better understanding



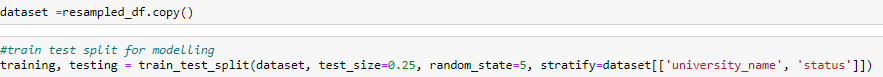




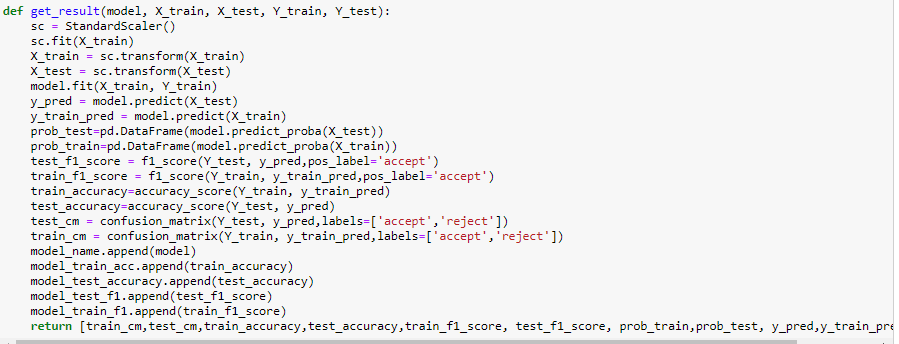




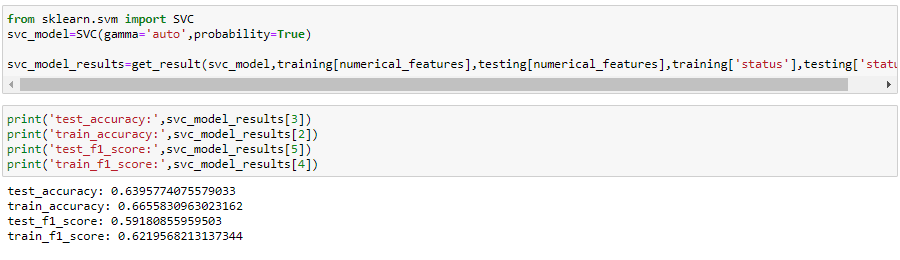
Splitting the dataset



Applying machine learning algorithms



SVM:



Visualization:

