A

Major Project

On

ADMISSION PREDICTION USING MULTIPLE LINEAR REGRESSION WITH PCA

(Submitted in partial fulfilment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

K.DHARANI (187R1A05E7)

K.DHRUVANESHWAR(187R1A05F0)

P.NEESHMA (197R5A0513**)**

A.SHRUTHI (197R5A0510)

Under the Guidance of

G.VINESH SHANKAR

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC,NBA,NIRF,Permanently Affiliated to JNTUH, Approved by AICTE, NewDelhi)

Recognized Under Section 2(f) & 12(B) of the UGCAct.1956,

Kandlakoya (V), Medchal Road, Hyderabad-501401.

2018-22

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled "**ADMISSION PREDICTION USING MULTIPLE LINEAR REGRESSION WITH PCA**"being submitted by **K.Dharani(187R1A05E7), K.Dhruvaneshwar(187R1A05F0) & P.Neeshma(197R5A0513), A.Shruthi(197R5A0510)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by him/her under our guidance and supervision during the year 2021-22.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Mr. G. Vinesh Shankar
Assistant Professor
INTERNAL GUIDE

Dr. A. Raji Reddy
DIRECTOR

Dr. K. Srujan Raju
HOD

EXTERNAL EXAMINER

Submitted for viva voice Examination held on _____

ACKNOWLEGDEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard to my guide. **Mr. G. VINESH SHANKAR,** Assistant Professor for his exemplary guidance, monitoring and constant encouragement through out the project work. The blessing, help and guidance given by him shall carry us a long way in the journey of life on which we are about to embark. We also take this opportunity to express a deep sense of gratitude to Project Review Committee (PRC) **Mr. J. Narasimha Rao, Dr. T. S. Mastan Rao, Dr. Suwarna Gothane, Mr. A. Uday Kiran, Mr. A. Kiran Kumar, Mrs. G. Latha** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju,** Head, Department of Computer Science and Engineering for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy,** Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy,** Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** who contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

K.DHARANI (187R1A05E7)

K.DHRUVANESHWAR(187R1A05F0)

P.NEESHMA (197R5A0513**)**

A.SHRUTHI (197R5A0510)

ABSTRACT

This project is titled as “Admission Prediction using multiple linear regression with pca”. In today's education world there are many numbers of students who want to pursue higher education after engineering or any graduate degree course in abroad universities like USA, UK etc. Students who want to do master’s program in America have to write GRE and TOEFL. Once they have attended the exams which are one of the crucial factors they have to consider. So our goal is to develop a model using machine learning which will tell the students their chance of admission into a respective university. These models should consider all the crucial factors which plays a vital role in student admission process and should have high accuracy. The predicted output gives them a fair idea about their admission chances in a particular university.

LIST OF FIGURES

LIST OF SCREENSHOTS

TABLE OF CONTENTS

1. INTRODUCTION

1. INTRODUCTION

1.1 PROJECT SCOPE

This project is titled as “Admission Prediction using multiple linear regression with pca”. This system predicts the probability of students getting admitted in a university. This project uses machine-learning methods. First, we use a machine learning algorithms for training dataset. We then compare a number of algorithms that gives fair idea about student admission process and should have high accuracy.

1.2 PROJECT PURPOSE

This has been developed for the purpose of the admission prediction model is to predict whether the student he/she gets admission or not in the universities. The algorithms are applied on the given dataset of admission prediction. Universities also can provide guidance to the students who are interested to take admissions to their courses. Also universities can confirm or reject the admission of students.

1.3 PROJECT FEATURES

The main features of this project are as input and learning machines play complementary roles in supervised learning. When features are discriminative, they place less demand on the learning machine in order to perform a task successfully. On the other hand, a powerful learning machine places less demand on features. In admission prediction system use only a few features such as GRE score of the student, TOFEL of the student, academic details, etc.

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

SYSTEM ANALYSIS

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside the system is done. A key question considered here is, “what must be done to solve the problem?” The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

2.1 PROBLEM DEFINITION

The system uses machine learning algorithms, It will help the students to identify the chances of their application to an university being accepted. Multiple machine learning classification algorithms were evaluated to develop the system. Accuracy and Graph outputs of all the models were compared and then the best model was chosen. Multiple Linear Regression model with PCA compared to other ones was the best suited model because it had the highest accuracy and the least error. Students can use the model to assess their chances of getting admission into a particular university.

2.2 EXISTING SYSTEM

The existing system is that it causes issue, Students admission is one of the most important activities in education industry. In existing system it contains algorithms like random forest regression , K-Nearest Neighbors’ etc , which contains of high errors , and less accuracy.

2.2.1 LIMITATIONS OF EXISTING SYSTEM

- Less efficient.
- The currently available resources is that they are very limited.
- They are not truly dependable taking into consideration of their accuracy and reliability.

To avoid all these limitations and make the working more accurately the system needs to be implemented efficiently.

2.3 PROPOSED SYSTEM

The aim of proposed system is to develop a system of improved facilities. The proposed system can overcome all the limitations of the existing system. The system provides higher efficiency for prediction. The existing system has several disadvantages and many more difficulties to work well. The proposed system eliminate or reduce these difficulties. The proposed system is that the system is to automate the process carried out in the organization with improved performance and realize the vision of admission. All the crucial factors plays a vital role in student admission prediction process and that have high accuracy.

2.3.1 ADVANTAGES OF THE PROPOSED SYSTEM

The system is very simple in design and to implement. The system requires very low system resources and the system will work in almost all configurations. It has got following features.

- Reducing time in activities.
- Higher accuracy.
- Improve the operational efficiency by improving the quality of the process.

2.4 FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put into the forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed an system is not a burden to the company. Three key considerations involved in the feasibility of the analysis are.

- Economic Feasibility
- Technical Feasibility
- Social Feasibility

2.4.1 ECONOMIC FEASIBILITY

The developing system must be justified by cost and benefit. Criteria to ensure that effort is concentrated on project, which will give best, return at the earliest. One of the factors, which affect the development of a new system, is the cost it would require.

The following are some of the important financial questions asked during preliminary investigation:

- The costs conduct a full system investigation.
- The cost of the hardware and software.
- The benefits in the form of reduced costs or fewer costly errors.

Since the system is developed as part of project work, there is no manual cost to spend for the proposed system. Also all the resources are already available, it give an indication of the system is economically possible for development.

2.4.2 TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

2.4.3 BEHAVIORAL FEASIBILITY

This includes the following questions:

- Is there sufficient support for the users?
- Will the proposed system cause harm?

The project would be beneficial because it satisfies the objectives when developed and installed. All behavioral aspects are considered carefully and conclude that the project is behaviorally feasible.

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 HARDWARE REQUIREMENTS:

Hardware interfaces specifies the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements.

- Processor : Intel Core i5.
- Hard disk : 1TB.
- RAM : 2GB.

2.5.2 SOFTWARE REQUIREMENTS:

Software Requirements specifies the logical characteristics of each of the interface and software components of the system. The following are some software requirements,

- Operating system : Windows 10
- Languages : Python
- Tool : PyQT

3. ARCHITECTURE

3. ARCHITECTURE

3.1 PROJECT ARCITECTURE

This project architecture describes how the application system is going to function.

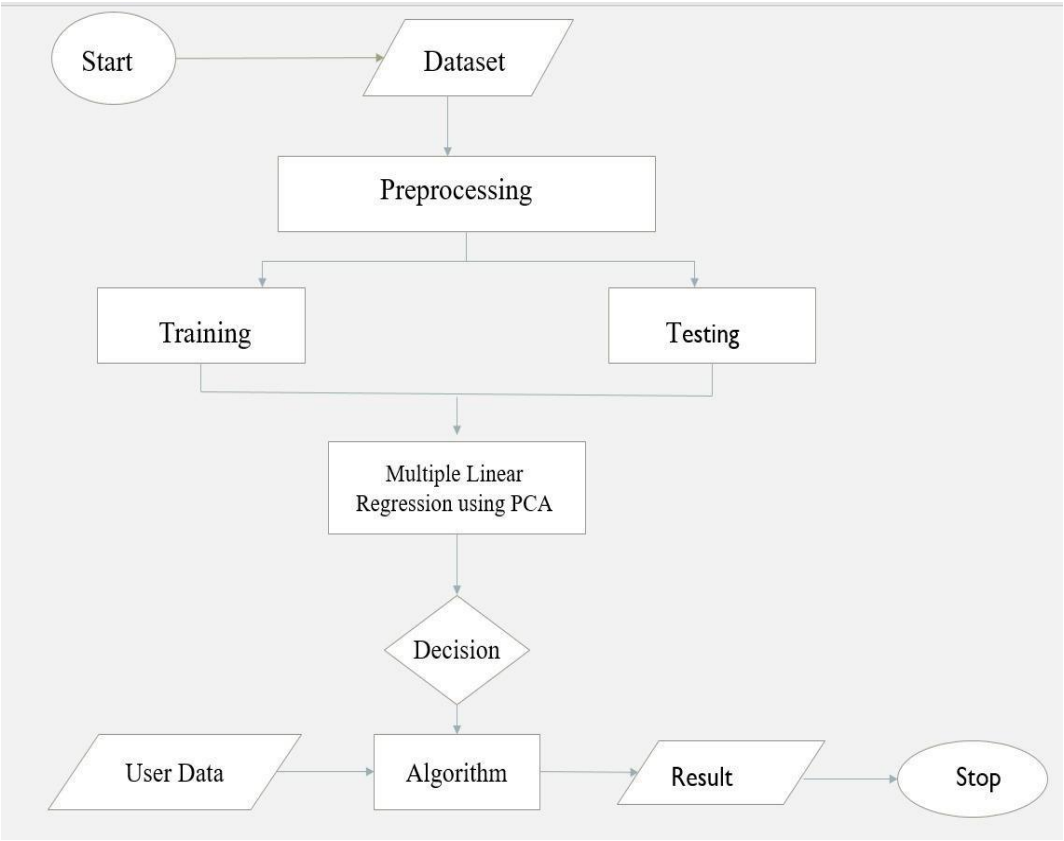


Figure 3.1: Project Architecture of admission prediction

3.2 MODULES DESCRIPTION

DATA GATHERING: Data gathering is the process of collecting and measuring information from countless different sources. In order to use the data we collect to develop practical machine learning solutions, it must be collected. The dataset has been collected from students of different colleges. The dataset collected consist of instances of students.

DATA-PREPROCESSING: Data preprocessing is a technique that is used to convert raw data into a clean dataset. The data is gathered from different sources is in raw format which is not feasible for the analysis. Pre-processing for this approach takes 4 simple yet effective steps:

- **Attribute Selection**: Some of the attributes in the initial dataset that was not pertinent (relevant) to the experiment goal were ignored. The attribute like serial no. is not required. The main attributes used for this study are GRE Scores, TOEFL Scores, CGPA, and University Ranking.
- **Cleaning missing values**: In some cases the dataset contain missing values. We could remove the entire line of data but what if you're inadvertently removing crucial information. One in every of the foremost common plan to handle the matter is to require a mean of all the values of the same column and have it to replace the missing data. The library used for the task is called Scikit Learn preprocessing. It contains a class called Imputer which will help us take care of the missing data.
- **Feature Scaling** : Feature Scaling is a technique to standardize the independent features present in the data in a fixed range. It is performed during the data pre-processing to handle highly varying magnitudes or values or units. If feature scaling is not done, then a machine learning algorithm tends to weigh greater values, higher and consider smaller values as the lower values, regardless of the unit of the values.
- **Training and Test data Splitting the Dataset into Training set and Test Set**: The next step is to split our dataset into two. Training set and a Test set. We will train our machine learning models on our training set and then we will test the models on our test set to examine how accurately it will predict. A general rule of the thumb is to assign 75% of the dataset to training set and therefore the remaining 25% to test set.

PROCESS: They are several ML models have to be developed using various machine learning algorithms for admission prediction to a particular university. Algorithms are :

- **MULTIPLE LINEAR REGRESSION** : Multiple linear regression is the extension of linear regression in the relationship between more than two variables. In simple linear regression, we have one predictor and one response variable. But in multiple regressions, we have more than one predictor variable and one response variable. There is the following general mathematical equation for multiple regression - y=b0+b1*x1+b2*x2+b3*x3+⋯bn*xn
  - y is a response variable.
  - b0, b1, b2...bn are the coefficients.
  - x1, x2, ...xn are the predictor variables

  In MLR the assumptions are Lack of Multicollinearity, Linearity.

- **RANDOM FOREST ALGORITHM**: Random forest is a Supervised Learning algorithm which uses ensemble learning method for classification and regression. Random forest is a bagging technique and not a boosting technique. The trees in random forests are run in parallel. There is no interaction between these trees while building the trees. It operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

  Disadvantages of random forest are Good at classification concept but not so good at Regression. As the inner calculations of the Random Forest Model are scarcely known we have very little control on how the model functions.

- **MULTIPLE LINEAR REGRESSION WITH PCA**: Multiple regression analysis is one of the most widely used methodologies for expressing the dependence of a response variable on several predictor variables. Due to the presence of multicollinearity, the standard errors of the parameter estimates could be quite high. The specific goals of principal component analysis are to reduce a large number of predictor variables to smaller no. of principal components and to provide a regression equation for an underlying process by using predictor variables. Principal components can be derived such that they are nearly uncorrelated or orthogonal. Thus the problem of multicollinearity among the variables can be solved by using PCA.

PRINCIPAL COMPONENT ANALYSIS PCA

- It is considered to be one of the most used unsupervised algorithms and can be seen as the most popular dimensionality reduction algorithm.
- It is a linear dimensionality reduction technique that can be utilized for extracting information from a high-dimensional space by projecting it into a lower-dimensional sub-space.
- It tries to preserve the essential parts that have more variation of the data and remove the non-essential parts with fewer variation. Dimensions are nothing but features that represent the data.

ADVANTAGES OF PCA

- It reduces computation time.
- It also helps remove redundant features, if any

APPLICATIONS OF PCA

- PCA is used as a dimensionality reduction technique in domains like facial recognition, computer vision and image compression.
- It is also used for finding patterns in data of high dimension in the field of finance, data mining, bioinformatics, psychology, etc.

INTERPRETATION: Multiple Linear Regression model with PCA compared to other ones was the best suited model because it had the highest accuracy and the least error. Students can use the model to assess their chances of getting admission into a particular university with an average accuracy of 83 percent. The predicted output gives them a fair idea about their admission chances in a particular university.

ERROR AND ACCURACY :

MSE : MLR (PCA) < RFR (PCA) < RFR < MLR

RMSE : MLR (PCA) < RFR (PCA) < RFR < MLR

ACCURACY : MLR < RFR < RFR (PCA) < MLR (PCA)

3.3 USECASE DIAGRAM

In the use case diagram, we have basically two actors who are the student and the administrator.

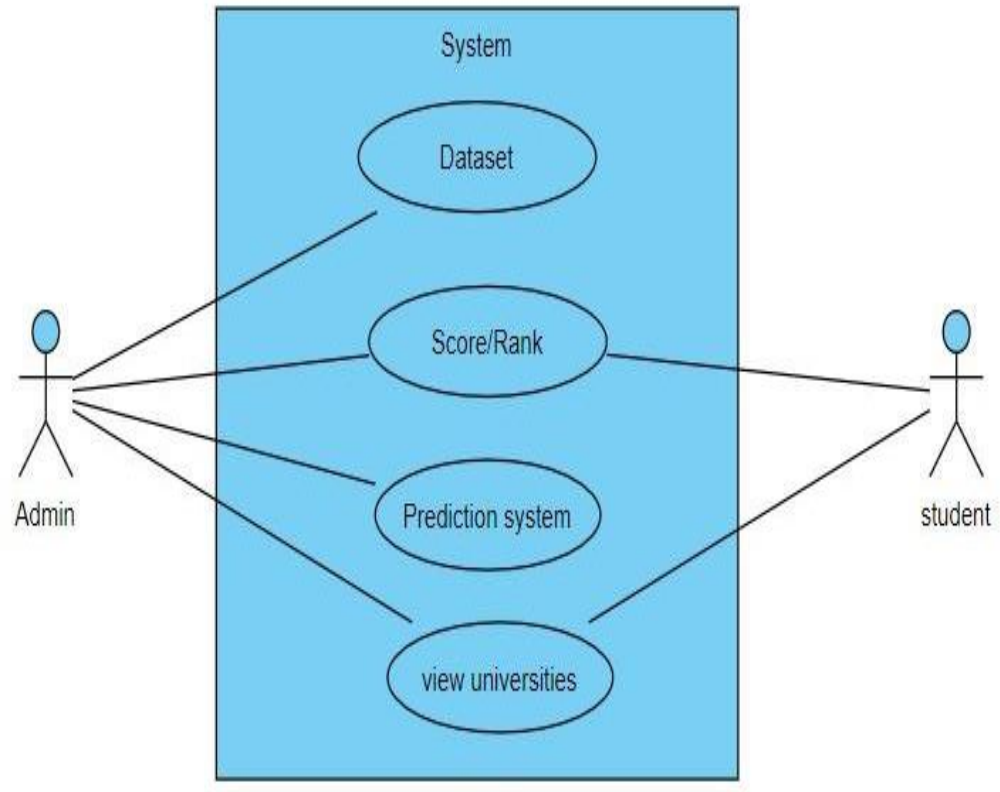


Figure 3.2: Use Case Diagram for admission prediction

3.4 SEQUENCE DIAGRAM

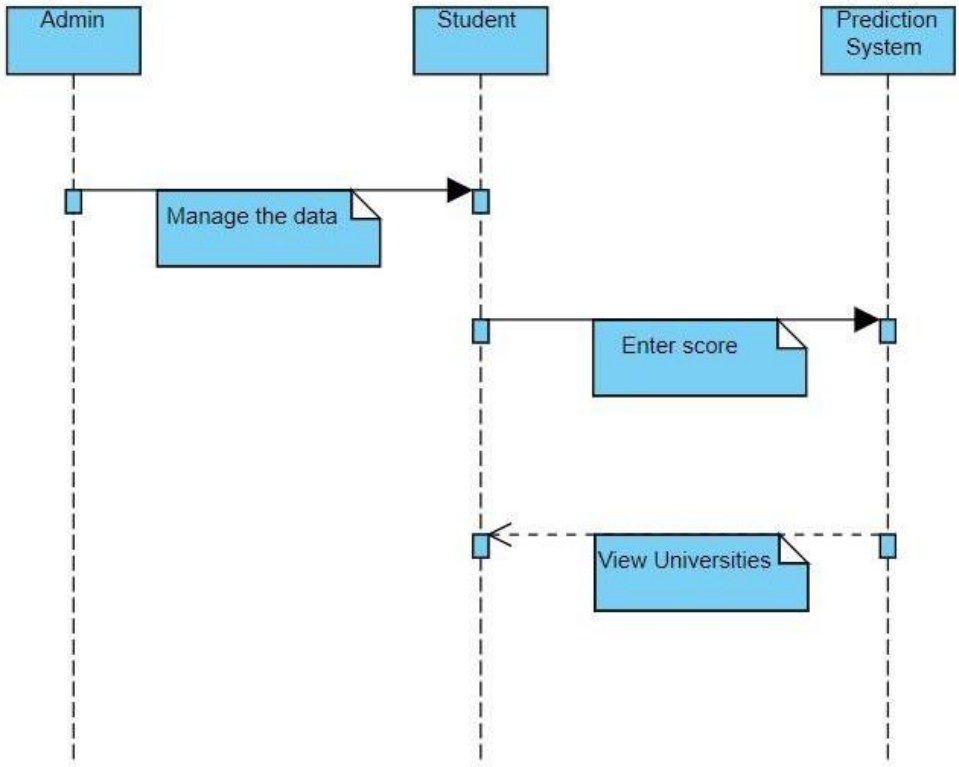
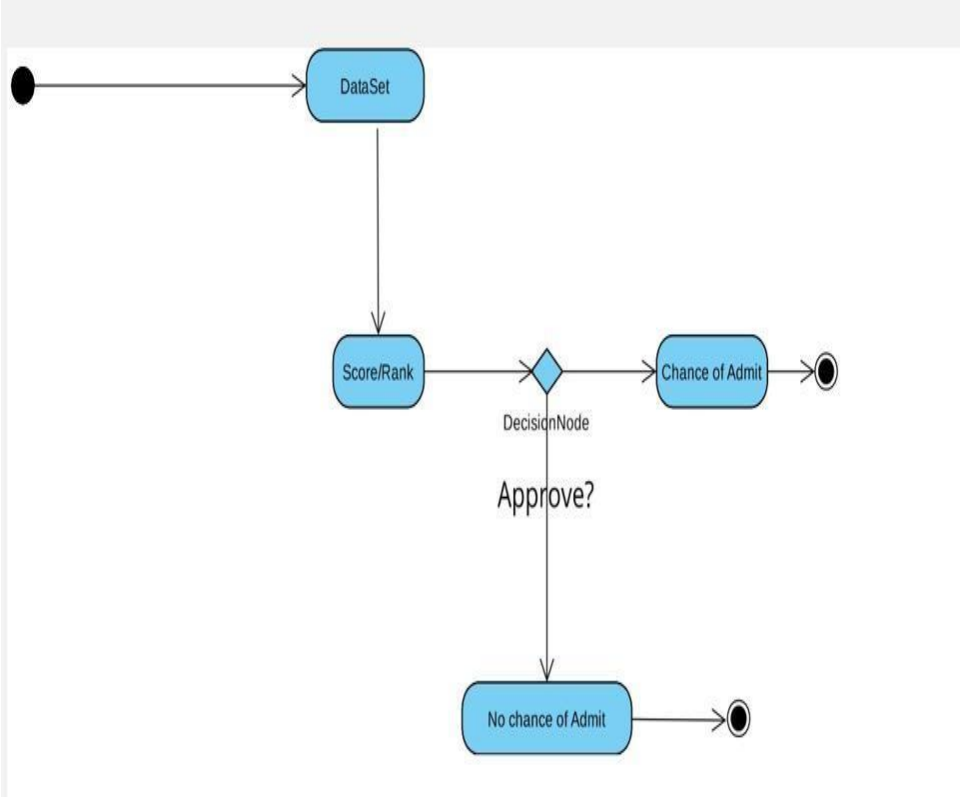


Figure 3.3: Sequence Diagram for admission prediction

3.5 ACTIVITY DIAGRAM

It describes about flow of activity states.



\

Figure 3.4: Activity Diagram for admission prediction

3.6 CLASS DIAGRAM

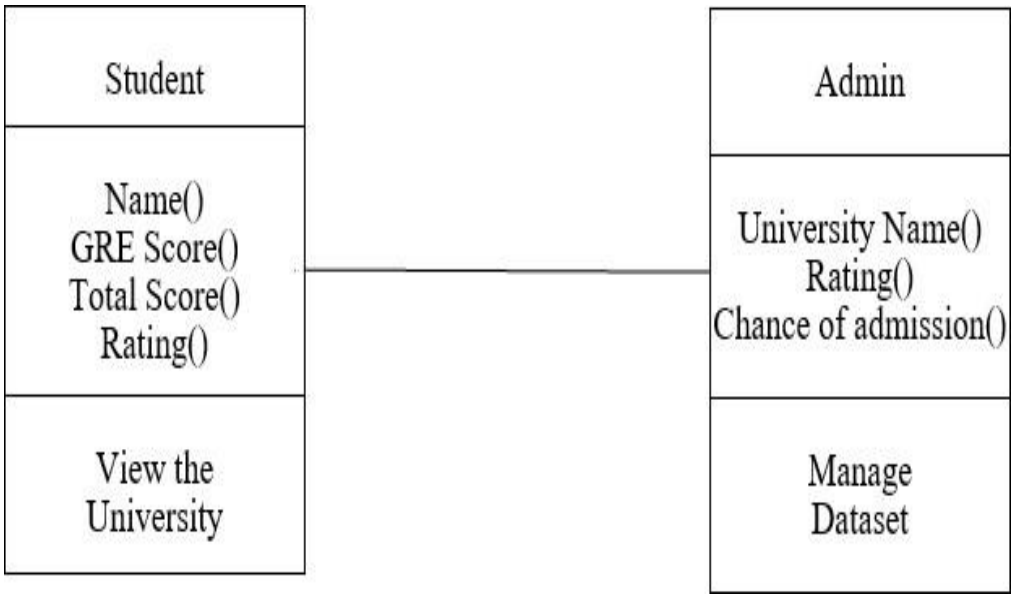


Figure 3.5: Class Diagram for admission prediction

4. IMPLEMENTATION

4. IMPLEMENTATION

4.1 SAMPLE CODE

Admission_prediction .ipynb:

run_commands.txt

```
!pip install sklearn
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from google.colab import drive
drive.mount('/content/drive')
df = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/Admission_Predict.csv")
print(df)
df.isna().sum()
x = df.iloc[:,0:-1].values
y = df.iloc[:,-1].values
from sklearn.preprocessing import StandardScaler
sc= StandardScaler()
x = sc.fit_transform(x)
print(x)
from sklearn.decomposition import PCA
pca = PCA(n_components=None)
x = pca.fit_transform(x)
print(pca.explained_variance_ratio_)
plt.plot(np.cumsum(pca.explained_variance_ratio_))
from sklearn.model_selection import train_test_split
x_train,x_test, y_train,y_test = train_test_split(x,y,test_size=0.25,random_state=50)
from sklearn.linear_model import LinearRegression
lr = LinearRegression()
lr.fit(x_train,y_train)
y_pred =lr.predict(x_test)
print(y_pred)
from sklearn.metrics import r2_score,mean_absolute_error,mean_squared_error
r2 = r2_score(y_test,y_pred)
mae=mean_absolute_error(y_test,y_pred)
mse=mean_squared_error(y_test,y_pred)
rmse=np.sqrt(mse)
print('r2_score=',r2)
print('mae=',mae)
print('mse=',mse)
print('rmse=',rmse)
```

```
plt.scatter(y_test, y_pred, color =
'red') plt.scatter(y_test, y_test,
color = 'blue') plt.xlabel('Actual
vales') plt.ylabel('Predicted
values') plt.title('Actual vs
Prediceted values') plt.show()
```

Connection.php :

```
<?php
session_s
tart();
$db=mysqli_connect("localhost","root","","kdproject");
define('SERVER_PATH',$_SERVER['DOCUMENT_ROOT'].'http://localhost/trans
port/'); define('SITE_PATH','http://localhost/transport/');

define('PRODUCT_IMAGE_SERVER_PATH',SERVER_PATH.'admin/media/images/');
define('PRODUCT_IMAGE_SITE_PATH',SITE_PATH.'admin/media/images/');
if (mysqli_connect_errno()) {printf("Connect failed: %s\n", mysqli_connect_error());}
?>
```

index.php :

```
 <!doctype html>
<?php include 'connection.php'; ?>
<html class="no-js" lang="">
<head>
<meta charset="utf-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<title>Admission Prediction System</title>
<meta name="description" content="A Major Project By KD">
<meta name="viewport" content="width=device-width, initial-scale=1">
<link rel="apple-touch-icon" href="https://www.clipartmax.com
/png/middle/101-1019829_internet-multiple-tab-tabs-web-page-webpage-
icon-web-page-tab-png.png">
<link rel="shortcut icon" href="https://www.clipartmax.com
/png/middle/101-1019829_internet-multiple-tab-tabs-web-page-webpage-icon-web-page-tab-
  png.png">
<link rel="stylesheet" href="https://cdn.jsdelivr.net/npm/normalize.css@8.0.0/normalize.min.css">
<link rel="stylesheet"
  href="https://cdn.jsdelivr.net/npm/bootstrap@4.1.3/dist/css/bootstrap.min.css">
<link rel="stylesheet" href="https://cdn.jsdelivr.net/npm/font-awesome@4.7.0/css/font-
  awesome.min.css">
<link rel="stylesheet" href="https://cdn.jsdelivr.net/gh/lykmapipo/themify-icons@0.1.2/css/themify-
  icons.css">
<link rel="stylesheet" href="https://cdn.jsdelivr.net/npm/pixeden-stroke-7-icon@1.2.3/pe-icon-7-
  stroke/dist/pe-icon-7-stroke.min.css">
```

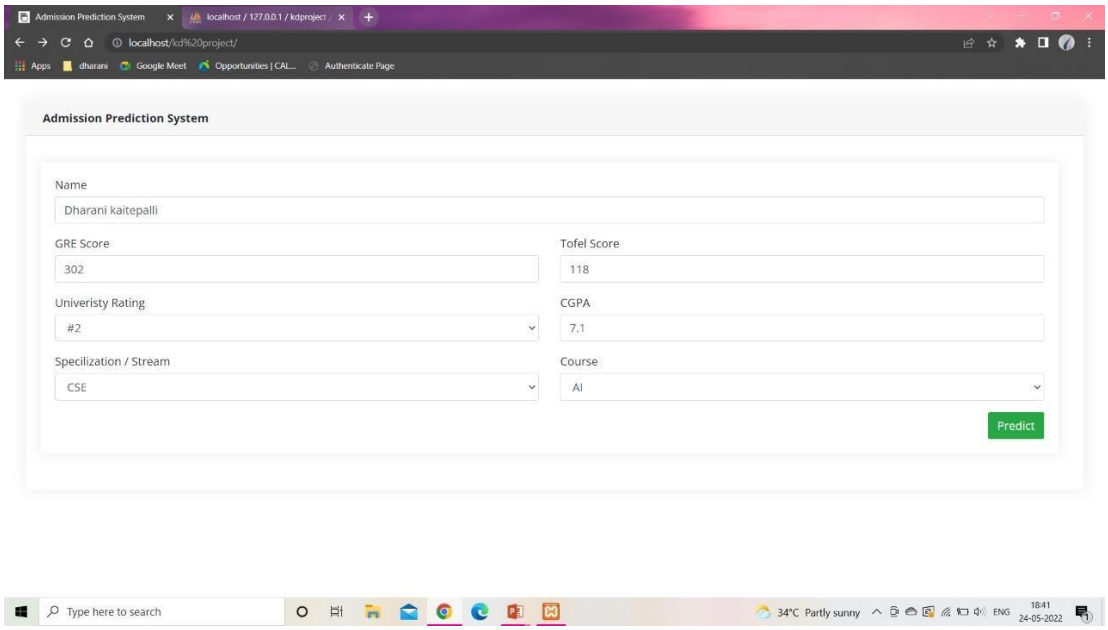
```
<link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/flag-icon-css/3.2.0/
css/flag- icon.min.css">
<link rel="stylesheet" href="assets/css/cs-skin-elastic.css">
<link rel="stylesheet" href="assets/css/style.css">
<link href='https://fonts.googleapis.com/css?family=Open+Sans:400,600,700,800'
rel='stylesheet' type='text/css'>
</head>
<body>
 <?php function get_safe_value($con,$str)
{
 if($str!='')
 {
 $str=trim($str);
 return mysqli_real_escape_string($con,$str);
 }
  }

  if(isset($_POST['submit']))
  {

 $name=get_safe_value($db,$_POST['name']);
 $grescore=get_safe_value($db,$_POST['grescore']);
 $tofelscore=get_safe_value($db,$_POST['tofelscore']);
 $rating=get_safe_value($db,$_POST['rating']);
 $cgpa=get_safe_value($db,$_POST['cgpa']);
 $stream=get_safe_value($db,$_POST['stream']);
 $course=get_safe_value($db,$_POST['course']);

   }

 ?>
 <div class="content">
 <div class="row">
 <div class="col-lg-12"><div class="card">
 <div class="card-header pl-3">
 <strong class="card-title pl-2">Admission Prediction System</strong>
 </div>
 <div class="card-body">
 <div class="tab-content pl-3 pt-2" id="nav-tabContent">
 <div class="tab-pane fade show active" id="custom-nav-home"
  role="tabpanel" aria-labelledby="custom-nav-home-tab">
  <div class="row">
  <div class="row col-lg-12">
  <div class="card mt-2 col-lg-12" ><div class="card-body">
  <form action="next.php" method="get">
  <div class="form-group">
  <label class="control-label mb-1">Name</label>
```

```
<input id="name" name="name" type="text" class="form control"></div>
<div class="row"><div class="form-group col-md-6">
<label  class="control-label mb-1">GRE Score</label>
<input id="grescore" name="grescore"
type="number" min="1" max="340" class="form-control" >
</div>
<div class="form-group col-md-6">
<label  class="control-label mb-1">Tofel Score</label>
<input id="tofelscore" name="tofelscore" type="number" min="1"
max="320" class="form-control" >
</div></div>
<div class="row">
<div class="form-group col-md-6">
<label  class="control-label mb-1">Univeristy Rating</label>
<select name="rating" id="rating" class="form-control">
 <option value="0">Select Rating</option>
 <option value="1">#1</option>
 <option value="2">#2</option>
 <option value="3">#3</option>
 <option value="4">#4</option>
 <option value="5">#5</option>
 </select>
 </div><div class="form-group col-md-6">
 <label  class="control-label mb-1">CGPA</label>
 <input id="cgpa" name="cgpa" type="text" class="form-control" >
 </div>
 </div>
 <div class="row">
 <div class="form-group col-md-6">
 <label  class="control-label mb-1">Specilization / Stream</label>
 <select name="stream" id="stream" class="form-control">
 <option value="0">Select stream</option>
 <option value="1">CSE</option>
 <option value="2">ECE</option>
 <option value="3">MECH</option>
 </select>
 </div>
 <div class="form-group col-md-6">
 <label  class="control-label mb-1">Course</label>
 <select name="course" id="course" class="form-control">
 <option value="0">Select Course</option>
 <option value="1">AI</option>
 <option value="2">ML</option>
 <option value="3">ROBOTICS</option>
 </select></div></div><div>
 <button type="submit" class="btn btn-success float-right">
   Predict</button>
```

```
</div>
</form>
</div>
</div>
</div>
</div>
</div>
</div>
</div>
</div>
</div>
 </div>
 <script src="https://cdn.jsdelivr.net/npm/jquery@2.
 2.4/dist/jquery.min.js"></script>
 <script src="https://cdn.jsdelivr.net/npm/popper.js@1.14.
 4/dist/umd/popper.min.js"></script>
 <script src="https://cdn.jsdelivr.net/npm/bootstrap@4.1.
 3/dist/js/bootstrap.min.js"></script>
 <script src="https://cdn.jsdelivr.net/npm/jquery-match-
height@0.7.2/dist/jquery.matchHeight.min.js"></script>
 <script src="assets/js/main.js"></script>
 </body>
 </html>
```

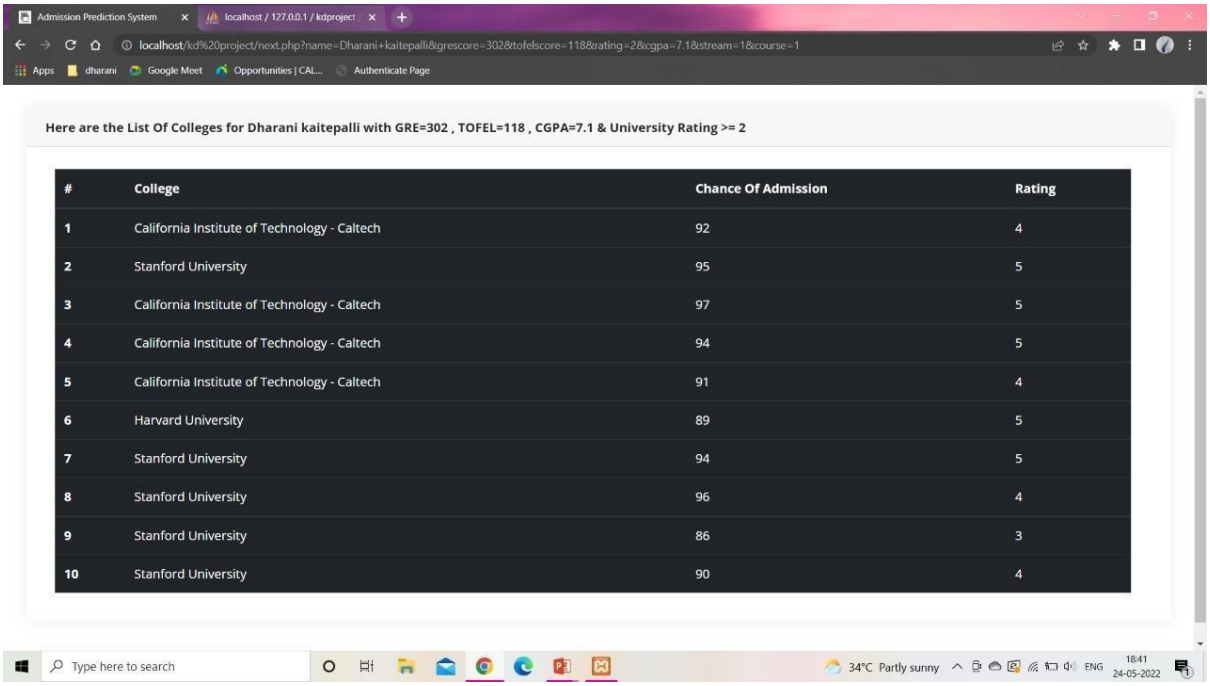
5. SCREENSHOTS

5.1 INPUT OF ADMISSION PREDICTION



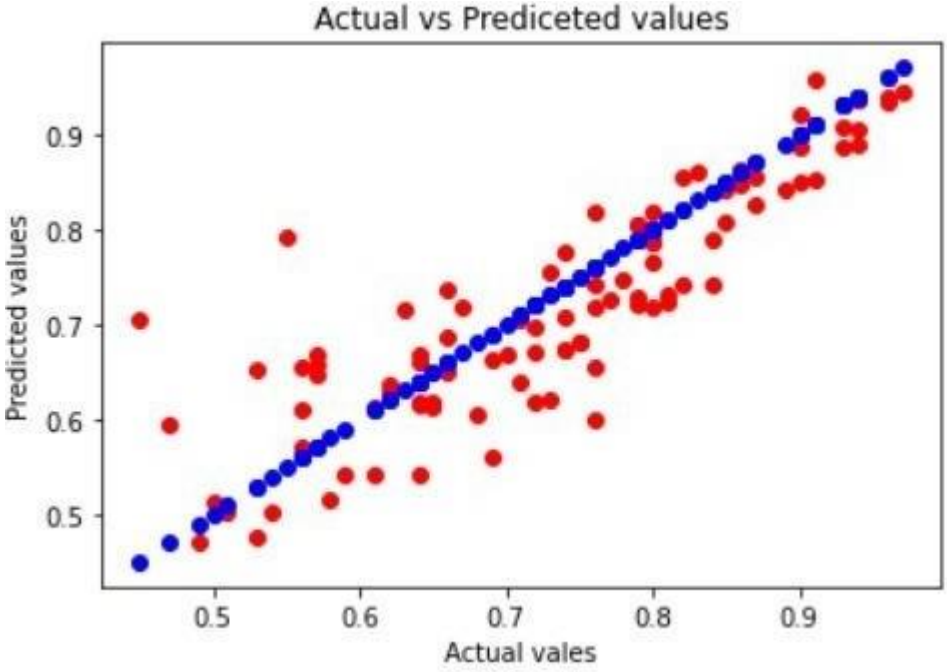
Screenshot 5.1: Input of admission prediction

5.2 RESULT OF ADMISSION PREDICTION



Screenshot 5.2: Result of admission prediction

5.3 ESTIMATED GRAPH RESULT



Screenshot 5.3: Estimated graph result

6. TESTING

6. TESTING

6.1 INTRODUCTION

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of is ensuring that the Software of system meets its requirements and user expectations and does not fail in an unacceptable manner.There are various types of test. Each test type addresses a specific testing requirement.

6.2 TYPES OF TESTING

6.2.1 UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

6.2.2 INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine the actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, is the combination components is correct and consistent. Integration testing is specifically aimed at exposing problems that arise from the combination of components.

6.2.3 FUNCTIONAL TESTING

The Functional tests will provide systematic of the functions tested are available as demonstrations that specified by the business and system documentation, and user technical requirements, manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be functions.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised

Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes.

7. CONCLUSION

7. CONCLUSION & FUTURE SCOPE

7.1 PROJECT CONCLUSION

The project titled as “Admission prediction using multiple linear regression with pca” the main goal of this project is to create a Machine Learning model which could be used by students who want to pursue their education in the US. Students can use the model to assess their chances of getting admission into a particular university. Multiple linear regression with pca proved to best-fit for development of the system when compared with other Linear regression model. The model can be used by the students for evaluating their chances of getting shortlisted in a particular university with an average accuracy of 85%.

The system is very flexible and versatile. The application has been tested with live data and has provided a successful result. Hence the software has proved to work efficiently.

7.2 FUTURE SCOPE

In future we can use Ongoing work includes extending more data related to additional universities , courses can be added to the system facilities to enhance the project and also to add some security features. It is possible to test other classification algorithms if they have high accuracy score than the current algorithm, the framework can be easily modified to support the new algorithm by changing the server code in the Node Red. Finally students can have an open source machine Learning model which will help the students to know their chance of admission into a particular university with high accuracy. Identified classes of invalid input must be functions

8. BIBILOGRAPHY

8. BIBILOGRAPHY

8.1 REFERENCES

[1] L. Breiman, Accuracy Predictors, Machine Learning, 24(2):123-140,1996.

[2] Abdul Fatah S; M, A. H. (2012). System for Predicting College Admission, pp. 107–113.

[3] Jamison, J. (2017). Applying Machine Learning to Predict Davidson College's Admissions Yield, pp. 765–766.

[4] Jared Cirelli, Andrea M. Konkol, Faisal Aqlan, "Predictive Analytics Models for Student Admission and Enrolment.

8.2 WEBSITES

[1] Data Cleaning and Analytics, Machine Learning https://archieve.ics.uci.edu/ml/index.php

[2] Data Visualizaton, Machine Learning https://www.analyticsvidhya.com/blog/2017/09/common-machin e-learning-algorithms/

[3] Jupyter Notebook, Implementing the Algorithms, Machine Learning, https://jupyter-notebook.readthedocs.io/en/stable/

8.3 GITHUB LINK

[1] https://github.com/dharani966/major-project