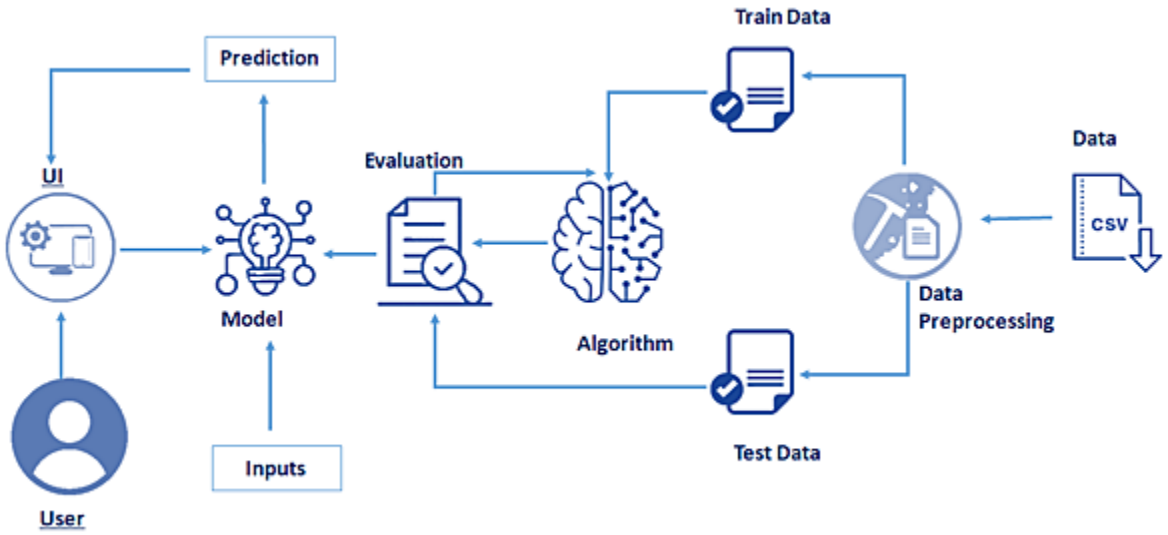
SESHADRI RAO GUDLAVALLERU

ENGINEERING COLLEGE

"UNIVERSITY ADMIT ELIGIBILITY PREDICTION USING IBM WATSON"



SUBMITTED BY:

SK.KAREEM

PIN NO : 20485A0330

Content

1 INTRODUCTION

Abstract:

The primary purpose is to discuss the prediction of student admission to university based on numerous factors and using logistic regression. Many prospective students apply for Master's programs. The admission decision depends on criteria within the particular college or degree program. The independent variables in this study will be measured statistically to predict graduate school admission. Exploration and data analysis, if successful, would allow predictive models to allow better prioritization of the applicants screening process to Master's degree programme which in turn provides the admission to the right candidates.

Overview:

Student admission for the Master's degree program consists of different criteria/scores which is taken into consideration before admitting the student to the degree program. This process is elaborative and requires lot of thought processing and analysis by the selection committee before choosing the right applicants to the Master's degree program. The purpose of this analysis is to demonstrate the top contributing scores which helps the student to get the admission into the Master's degree program. What factors contributes to successful admission to a Master's degree program? The analysis might seem straight forward but caution has to be exercised to consider the scores like GRE, TOEFL, university rating, SOP, LOR and CGPA and any outliers should not impact the decision making process.

2 LITERATURE SURVEY

2.1 Existing problem

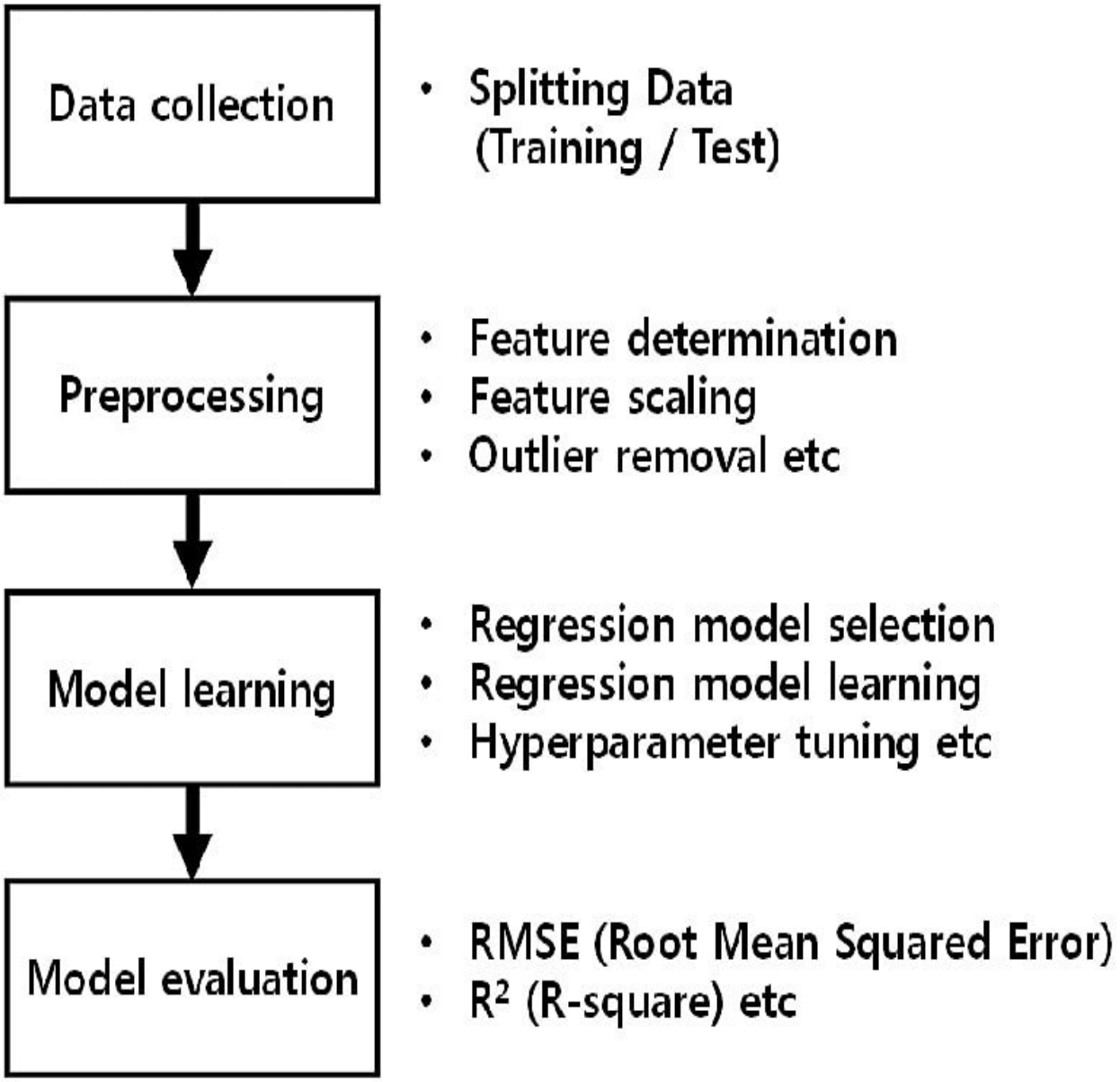
Every year the number of students seeking admission for the graduate studies is constantly increasing. As a result the competition gets tougher and the chances of admission becomes unpredictable. Given the growth of new programs and number of admission, a student is often unaware of the existence of such programs. As university education has become a basic part of most people's preparation for working life, admission to university is henceforth a topic of importance. How a student chooses a university, and conversely how a university chooses a student, determines the success of both sides in carrying through the education.

2.2 Proposed solution

In today's fast-paced world, every technological innovation influences the importance of higher education. Being the top destination for higher education universities of United States enroll thousands of international students every year. For students across the world who wish to pursue graduate studies in the USA, choosing a suitable college and universities and earning an admit is a challenge. Although, many Internet resources and forums are available, they do not offer satisfactory suggestions, as most of them are based on assumptions from college rankings and not the actual statistical relations. From a student's perspective cost of application and amount of dedication and determination to the process is also high. Hence, to guide the students in an efficient manner, the University Recommendation and Admission Prediction System has to be developed, based on the input of the students' academic data.

3 THEORITICAL ANALYSIS

3.1 Block diagram



3.2 Hardware / Software Requirements

Recommended System Requirements

1. Processors:Intel® Core™ i5 processor 4300M at 2.60 GHz or2.59 GHz (1 socket, 2 cores, 2 threads per core), 8 GB ofDRAMIntel® Xeon® processor E5-2698 v3 at 2.30 GHz (2sockets, 16 cores each, 1 thread per core), 64 GB ofDRAMIntel® Xeon Phi™ processor 7210 at 1.30 GHz (1 socket,64 cores, 4 threads per core), 32 GB of DRAM, 16 GB ofMCDRAM (flat mode enabled)
2. Disk space: 2 to 3 GB
3. Operating systems: Windows® 10, macOS*, and Linux*

Minimum System Requirements

1. Processors: Intel Atom® processor or Intel® Core™ i3processor
2. Disk space: 1 GB
3. Operating systems: Windows* 7 or later, macOS, and Linux • Python* versions: 3.9

Software requirements:

anaconda navigator:

Anaconda is an open-source distribution for python and R. It is used for data science, machine learning, deep learning, etc. With the availability of more than 300 libraries for data science, it becomes fairly optimal for any programmer to work on anaconda for data science.

Pycharm:

PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.

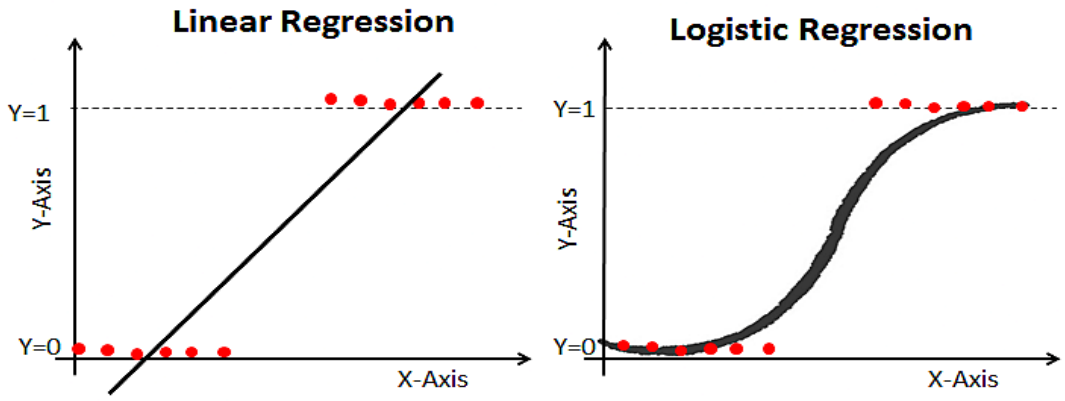
4 EXPERIMENTAL INVESTIGATIONS

Logistic Regression:

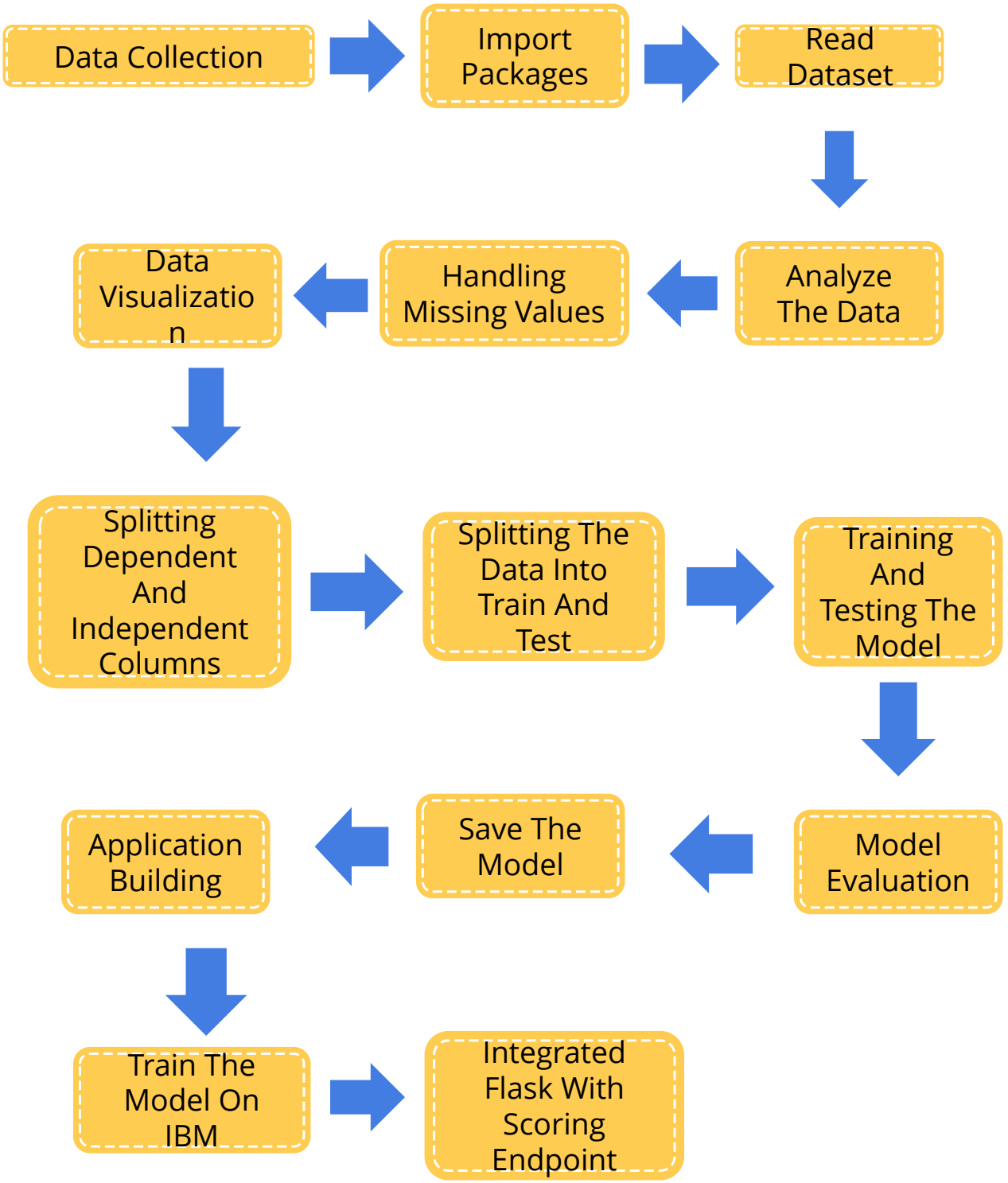
This type of statistical model (also known as *logit model*) is often used for classification and predictive analytics. Logistic regression estimates the probability of an event occurring, such as voted or didn't vote, based on a given dataset of independent variables. Since the outcome is a probability, the dependent variable is bounded between 0 and 1. In logistic regression, a logit transformation is applied on the odds—that is, the probability of success divided by the probability of failure.

Linear Regression:

Linear regression models are used to identify the relationship between a continuous dependent variable and one or more independent variables. When there is only one independent variable and one dependent variable, it is known as simple linear regression, but as the number of independent variables increases, it is referred to as multiple linear regression. For each type of linear regression, it seeks to plot a line of best fit through a set of data points, which is typically calculated using the least squares method.

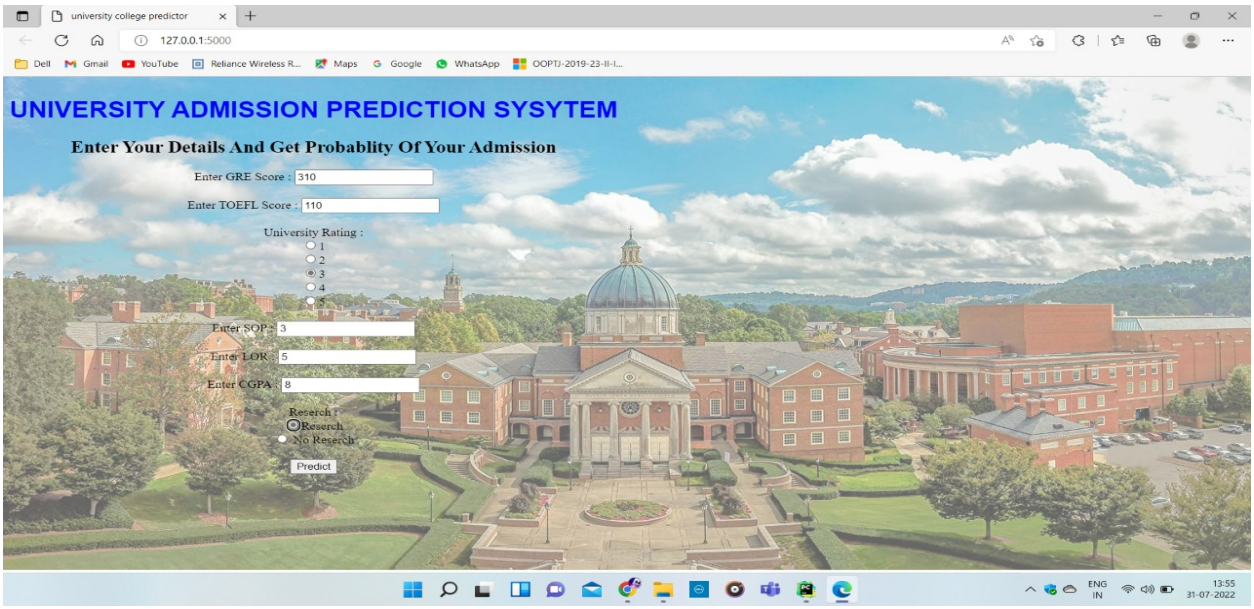


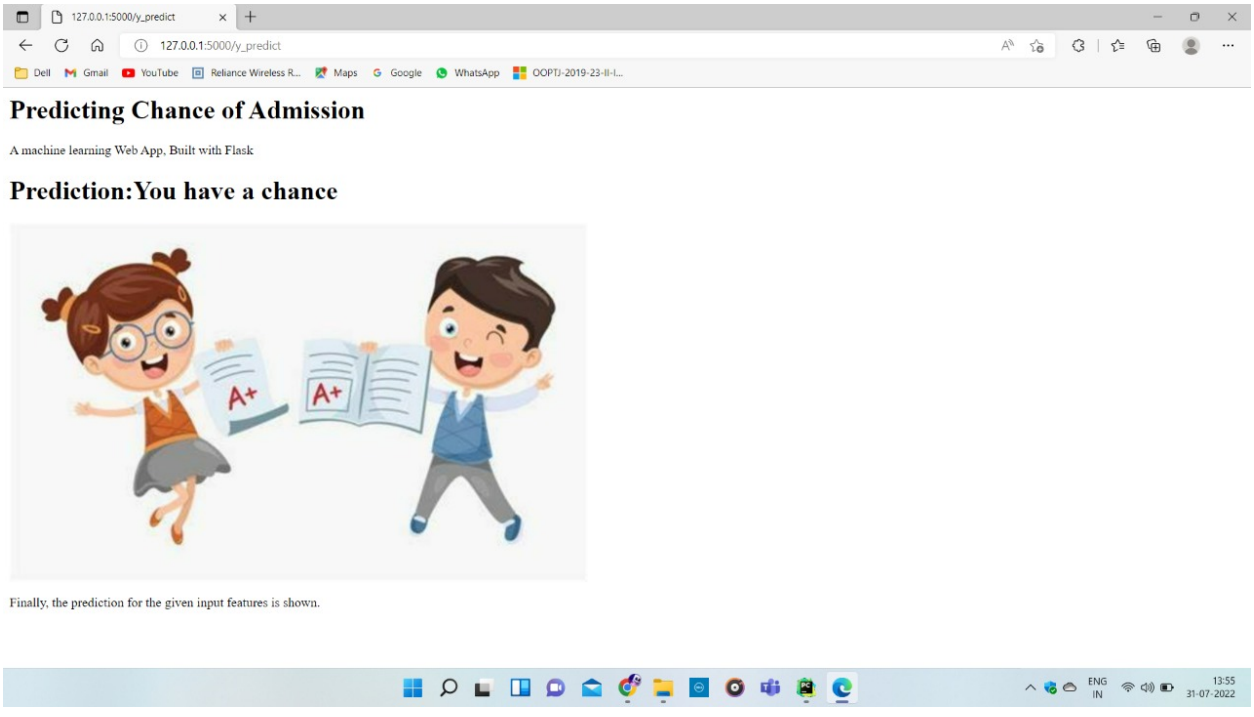
5 FLOW CHART


Data Collection
Import Packages
Read Dataset
Analyze The Data
Handling Missing Values
Data Visualization
Splitting Dependent And Independent Columns
Splitting The Data Into Train And Test
Training And Testing The Model
Model Evaluation
Save The Model
Application Building
Train The Model On IBM
Integrated Flask With Scoring Endpoint

6 RESULT

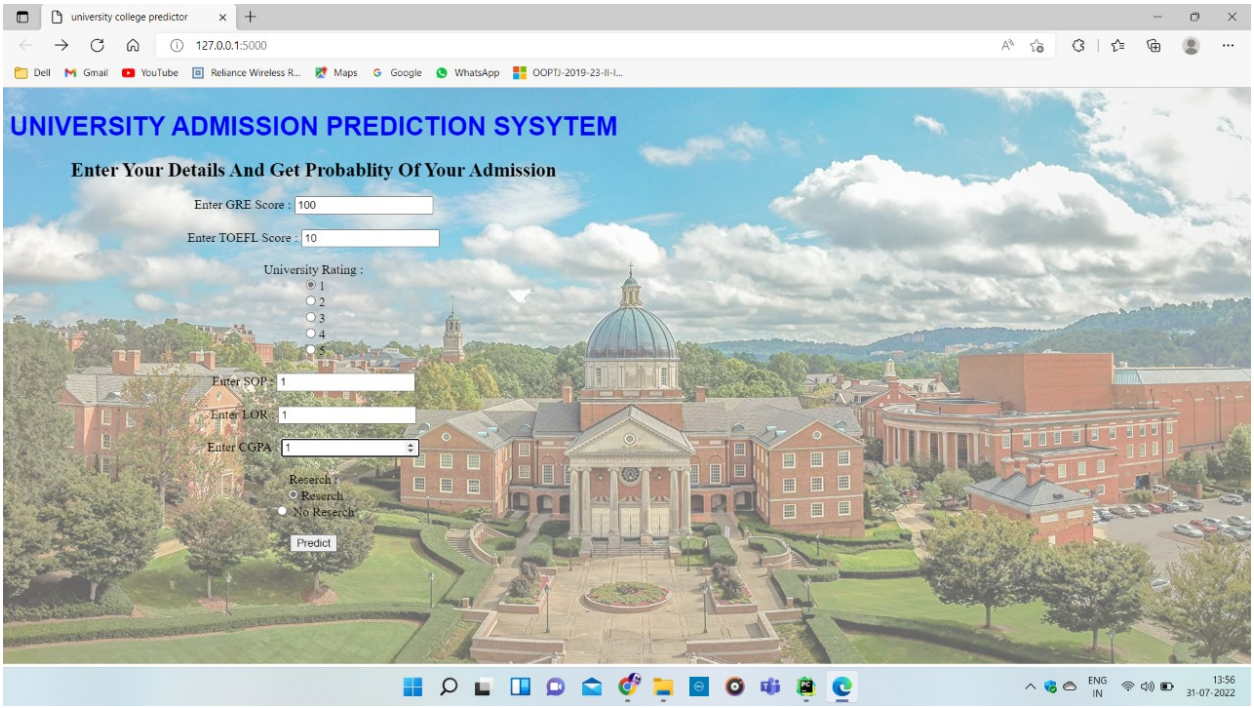
For Getting Admission In Reputed Universty For Master's Degree. We will Enter Our Results Of All The Fields Enter Here And Check The Probablity Here:

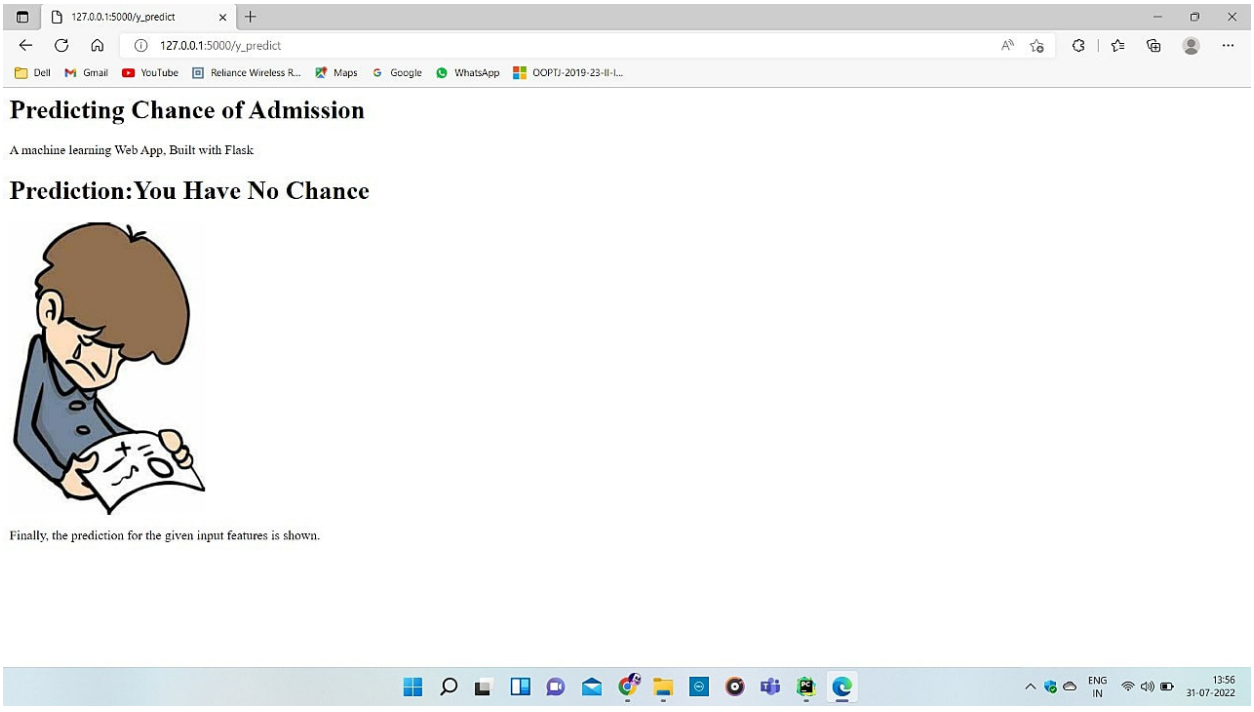




→ If Student Getting Good Results Prediction Shows Your Eligible

→ If Student Is Not Eligible It Shows You Are Not Eligible Based On Your Results Only.





7 ADVANTAGES & DISADVANTAGES OF A PROJECT

ADVANTAGES :

There is an endless number of advantages of ML. We can take a look at the one. which are really helpful.

The advantages of Machine Learning tell us how using ML would benefit us. So, let's have a look at the advantages of Machine Learning

→ **Automation of Everything**

→ **Wide Range of Applications**

→ **Scope of Improvement**

→ **Efficient Handling of Data**

→ **Best for Education**

DISADVANTAGES :

Similar to the advantages of Machine Learning, we should also know the disadvantages of Machine Learning. If you don't know the cons, you won't know the risks of ML. So, let's have a look at these disadvantages:

→ **Possibility of High Error**

→ **Algorithm Selection**

→ **Data Acquisition**

→ **Time and Space**

→ **Internet Issues**

8 CONCLUSION

The subject of this examination was to determine if the below variables contribute to the admission of student to Master's degree program.

GRE Score
TOEFL Score
University Rating
SOP
LOR
CGPA

The results of this examination appear to indicate that it greatly contributes to the response variable 'Chance of Admit'. Higher the **GRE, TOEFL** score then higher the admit chances. The model predicts 87.5% accuracy and can be used for predicting the admit chances based on the above factors. This model will be helpful for the universities to predict the admission and ease their process of selection and timelines.

As part of the hypothesis, the model proved that admission to Master's degree program is dependent on **GRE, TOEFL** and other scores.

This model would likely be greatly improved by the gathering of additional data of students from different universities which has similar selection criteria to choose the candidates for Master's program.

9 FUTURE SCOPE

University Admission depends on many factors, among them GPA, GRE and TOEFL are most important. We have used these three features in this project but more features can be implemented to get more accurate result. Other features could be Statement of Purpose (SOP), Letter of Recommendation (LOR), industry experience, internship experience, papers published, journals published etc. Also, as an extension to this work, recommendation of university with respect to research interest can be made with further study.
As mentioned before the data we actually fed to model is far less than the total data set as we first filter it based on user's intended major . So, we are working

on insufficient dataset. Even though the accuracy is ranged between 70 % - 85%. the result will never be satisfied to the user. So, we can work on huge data set for that we can scrape the data from thegradcafe.com and edulix.com. We can also add more features to our system like University ranking, University News feed.

This project is focused for the students intending to pursue master and PhD degrees, we can use dataset of undergraduate college admission and use it for the students intending to pursue bachelor degree.

10 BIBILOGRAPHY

Installation of Anaconda Navigator:

https://www.youtube.com/embed/5mDYijMfSzs

Installation Of Pycharm Professionals:

https://www.youtube.com/embed/z73PyNDgVyQ

Installation Of Python Packages:

https://www.youtube.com/embed/akj3_wTploU

Data Collection:

https://www.kaggle.com/datasets/rishal005/admission-predict

Data Pre-processing:

https://thesmartbridge.com/documents/spsaimldocs/Datapreprocessing.pdf

Handling Null Values:

https://towardsdatascience.com/7-ways-to-handle-missing-values-in-machine-learning-1a6326adf79e

Data Visualization:

https://www.youtube.com/embed/TLdXM0A7SR8

Splitting Dependent And Independent Columns:

https://www.youtube.com/embed/A_V6daPQZIU

Splitting The Data Into Train And Test:

https://www.youtube.com/embed/xgDs0scjuuQ

Training And Testing The Model:

https://www.youtube.com/embed/yIYKR4sgzI8

Model Evaluation:

https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226

Flask Frame Work Refarance:

https://www.youtube.com/embed/lj4I_CvBnt0

Flask Refarance To Run:

https://www.youtube.com/embed/UbCWoMf80PY

Train The Model On IBM:

Account Creation:

https://www.youtube.com/embed/4y_zD-0Q3F8

Train Model On IBM Watson:

https://www.youtube.com/embed/TysuP3KgSzc

Integrate Flask With Scoring Endpoint:

https://www.youtube.com/embed/ST1ZYLmYw2U