

NAME: Padmashri Adgonda Malgonnavar
HAWKID: padgondamalgonnavar@hawk.iit.edu
CWID: A20457612

NAME: Chinmay Thakare
HAWKID: cthakare@hawk.iit.edu
CWID: A20454398

PROJECT: GRADUATE SCHOOL ADMIT PREDICTION

INTRODUCTION

Objective:

The number of people opting for a masters grows year on year and there is an increase in competition to get a seat at the best ranked university increase simultaneously. Graduate admissions can be a mapping problem between students and universities where students always face a dilemma deciding universities of their choice while applying to a master's programs. Universities qualify students based on certain criteria, one of the most important being the test scores of a student. Students applying to any university in the United State of America, must have appeared for Graduate Record Examination (GRE)/Graduate Management Admission Test (GMAT) and International English Language Testing System (IELTS)/Test of English as a Foreign Language (TOEFL) and have a grading system in a recognized undergraduate school. Based on these scores, a student might qualify to be considered for a seat in the university. During the application process, an average student considers applying to at least 5 universities as the selection criteria for each university varies.

In this project, we will be using the graduate admission dataset which is in a csv format to predict the chances of a student getting admitted to a university based on various academic and non academic scores. Given a set of standardized scores like the GRE, TOEFL, Letter of recommendation, Statement of purpose the university may also consider other factors such as background of the student, research papers published, work experience. Since some of these factors are not of a statistical nature, we cannot determine the acceptance / rejection from a university based on them. However, making use of the statistical data available, we are interested in predicting the probability of a student being admitted based on certain parameters and design a model for this prediction.

Dataset information:

The dataset has about 500 rows and consists of following 8 columns:

- 1) GRE scores (out of 340)
- 2) TOEFL scores (out of 120)

NAME: Padmashri Adgonda Malgonnavar
HAWKID: padgondamalgonnavar@hawk.iit.edu
CWID: A20457612

NAME: Chinmay Thakare
HAWKID: cthakare@hawk.iit.edu
CWID: A20454398

- 3) University rating (out of 5)
- 4) SOP (out of 5)
- 5) LOR (out of 5)
- 6) GPA of undergraduate (out of 10)
- 7) Research paper published 1 or 0)
- 8) Chance of admit (ranging from 0 to 1)

The top three parameters that help in determining the chances of admit are students GRE scores, TOEFL scores and their CGPA. Students must mandatorily input these scores when applying to any university in the USA. The data is placed in google drive and the link of the same is provided below:

https://drive.google.com/open?id=1MGk6uBx3ITI0QHSsizcZ7t_u9892Gn_5

A	B	C	D	E	F	G	H	I
Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
1	337	118	4	4.5	4.5	9.65	1	0.92
2	324	107	4	4	4.5	8.87	1	0.76
3	316	104	3	3	3.5	8	1	0.72
4	322	110	3	3.5	2.5	8.67	1	0.8
5	314	103	2	2	3	8.21	0	0.65
6	330	115	5	4.5	3	9.34	1	0.9
7	321	109	3	3	4	8.2	1	0.75
8	308	101	2	3	4	7.9	0	0.68
9	302	102	1	2	1.5	8	0	0.5
10	323	108	3	3.5	3	8.6	0	0.45

Figure: Admission predict dataset.

Hypothesis:

1. We will check the correlation between various factors available and the chances of admissions.
2. We will test different models with different number of variables to determine which model would give us the best result in predicting the chances of admissions.
3. We will check if having a research paper published affects the chances of admissions.