

Documentation for APL - assignment 4:

APPROACH:

- **Data Exploration:**

I started by exploring the dataset to get a feel for the features—things like GRE scores, SOP, LOR, etc. I checked for any missing data or outliers that could mess up my analysis.

- **Feature Scaling (Normalisation):**

I made sure all my features, like GRE, TOEFL, SOP, LOR, and CGPA, played nice together by scaling them to the same level. This way, no single feature would have too much say just because its numbers were bigger.

- **Linear Regression Model:**

I went with a simple linear regression model first because, well, it's straightforward. I put together a formula involving GRE, TOEFL, SOP, LOR, CGPA, Research, and a constant term. Then, I used the least squares method to find the best coefficients for my model.

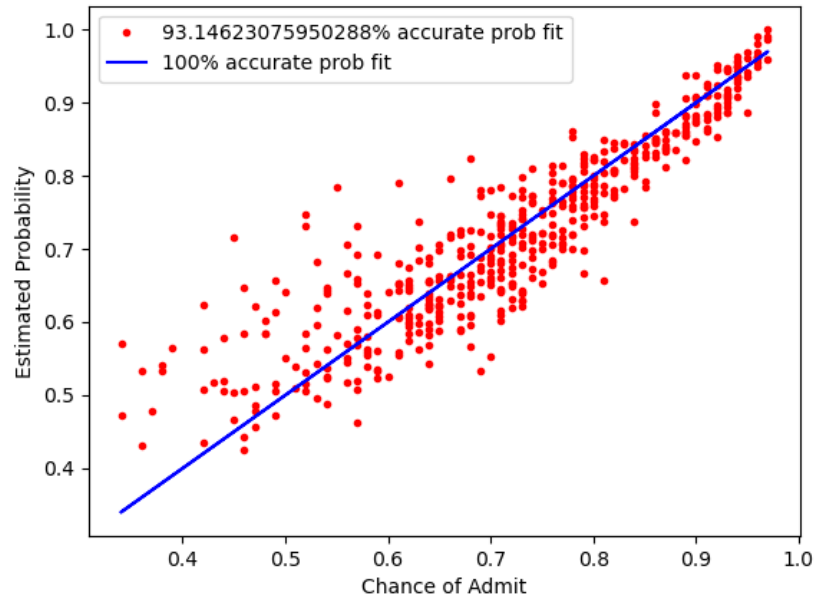
- **Coefficient Analysis:**

After crunching the numbers, I took a look at the coefficients. Basically, they told me how much each feature influenced the chances of getting admitted. I focused on the ones with the highest positive coefficients (magnitudes).

```
GRE coefficient      : 0.6318922049034613
TOEFL coefficient    : 0.3333566869703511
SOP coefficient      : 0.007930687278833844
LOR coefficient      : 0.08429371176209334
CGPA coefficient     : 1.1838505345773824
Research coefficient : 0.024307478582166402
```

- **Model Evaluation:**

To see if my model was any good, I calculated the percentage accuracy. I also whipped up a plot comparing my model's predictions with the real deal.



CONCLUSIONS:

- Concentrate more on your CGPA then GRE and TOEFL scores to get into top 5 ranked universities
- CGPA has the maximum impact on the chance of admit