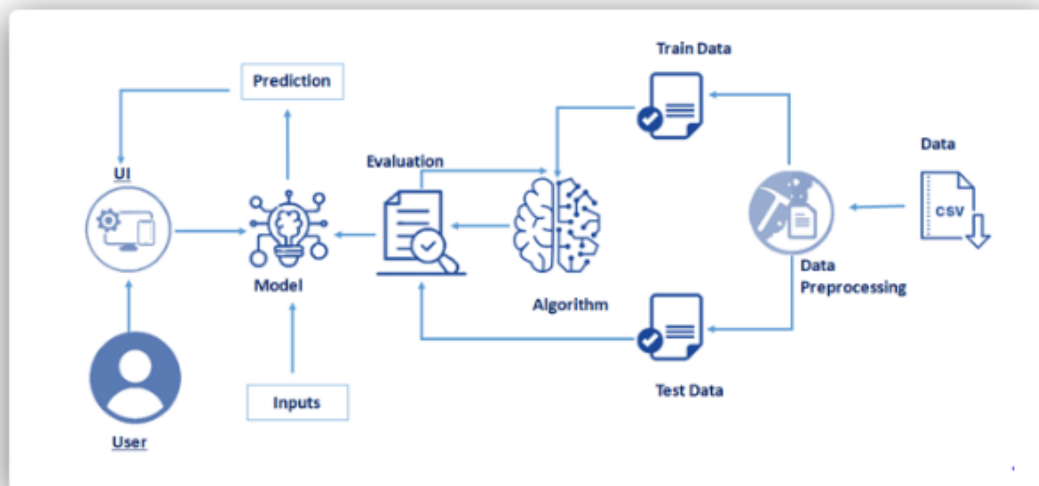# UniPredict - Project Details

Problem Statement:

University Admit Eligibility Predictor

Students are often worried about their chances of admission to University. The aim of this project is to help students in shortlisting universities with their profiles. The predicted output gives them a fair idea about their admission chances in a particular university. This analysis should also help students who are currently preparing or will be preparing to get a better idea.

Technical Architecture:

Our Project's GITHUB Repo:

https://github.com/IBM-EPBL/IBM-Project-42045-1660647960

Reference Project & Dataset:

https://www.kaggle.com/datasets/mohansacharya/graduate-admissions

Content

The dataset contains several parameters which are considered important during the application for Masters Programs.
The parameters included are :

1. GRE Scores (out of 340)
2. TOEFL Scores (out of 120)
3. University Rating (out of 5)
4. Statement of Purpose and Letter of Recommendation Strength (out of 5)
5. Undergraduate GPA (out of 10)
6. Research Experience (either 0 or 1)
7. Chance of Admit (ranging from 0 to 1)

Expected Inputs & Outputs for our Project:

- **Possibility**(%) of a Student with certain (1) *Cutoff Marks*, (2) *Rank*, (3) *Round*, (4) **Community** to get into a specific (5) **College** (out of 476 colleges) in his/her (6) **Preferred Branch.**

Our Findings, Assumptions & Theories to get this Possibility:

1	Cutoff Marks	Continuous	Independent	Range from 0 - 200
2	Rank	Discrete	Dependent on 1	Unique Values from 1 - 80K
3	Round	Discrete	Dependent on 2	1, 2, 3, 4
4	Community	Categorical	Independent	Ref. Info Brochure
5	College (code in dataset)	Categorical	??	476 Colleges - Ref. Info Brochure
6	Branch (code in dataset)	Categorical	Independent	Ref. Info Brochure
7	Allotted Category	Categorical	Independent	The Category in which the student got a seat.(either OC* or 4.)

* OC - Open Category

- **Possibility(%) = Correlation** of {2,3,6,7} & {5}
- Which ML Algorithm to generate a model for the whole dataset based on Correlation?
- **From Application's perspective:**
    - **Inputs:**
        - Cutoff Marks, Rank**, Round**, Branch, Community
        - College
    - **Output:**
        - Matching Strength(%) between two input sets
        - Higher if the student is able to get admission in the specified college with his/her input parameters.

** If not provided, will be predicted from dataset (Regression)

3. Kendall's Tau correlation of coefficient

Kendall's Tau is a non-parametric measure of relationships between columns of ranked data. The Tau correlation coefficient returns a value of 0 to 1, where:

0 is no relationship,
1 is a perfect relationship.

https://medium.com/analytics-vidhya/correlation-and-machine-learning-fee0ffc5faac

https://www.quora.com/What-machine-learning-algorithm-can-be-used-to-find-correlations-between-columns-of-a-table

https://ai.plainenglish.io/predicting-probabilities-with-python-b456334b85c6

Solution Architecture: