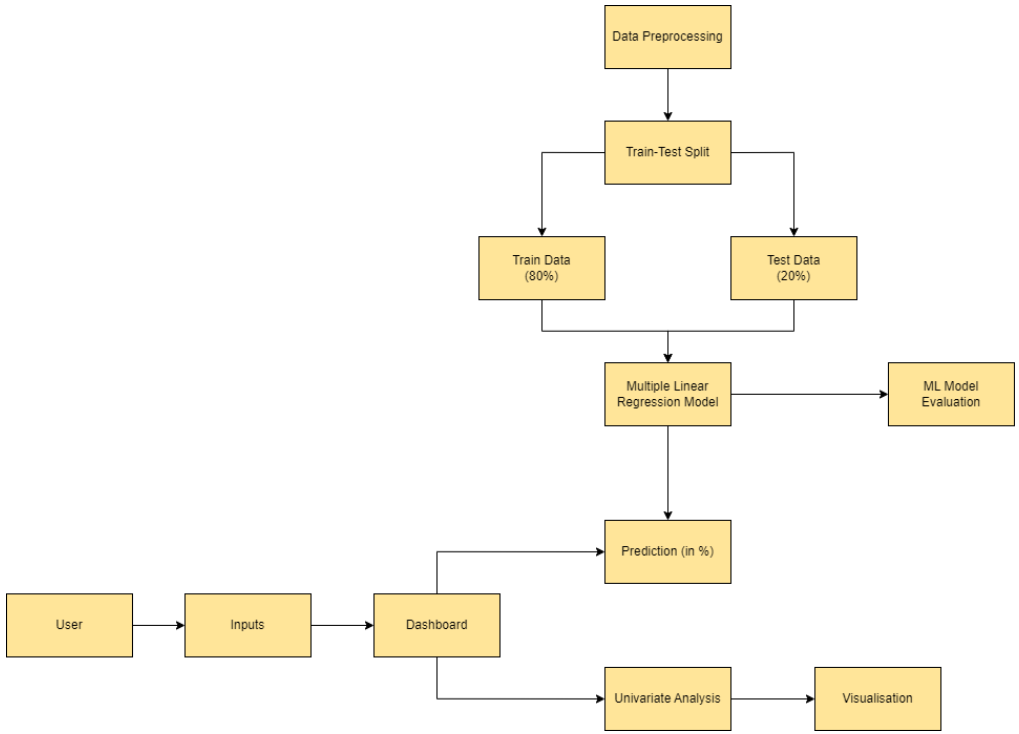
Project Design Phase-II
Data Flow Diagram & User Stories

Date	17 October 2022
Team ID	PNT2022TMID35368
Project Name	Project – University Admit Eligibility Predictor
Maximum Marks	4 Marks

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Developer	Exploratory Data Analysis	USN-1	Perform data cleaning if required and perform univariate, bivariate and multivariate analysis.	I can analyse the given dataset in the form of graphs.	Low	Sprint-1
Developer	Analysis of different regression models	USN-2	Compare the R2 scores of different fundamental regression models like Decision Trees, Random Forest, Multiple Linear Regression, Logistic Regression, etc and determine which model has the highest R^2 score.	Multiple Linear Regression was found to have the highest R2-Score of 0.819.	Medium	Sprint-1
End User	Web App Development	USN-3	Develop the web app using Streamlit. Predict the probability of acceptance given a test data for a candidate.	I am able to determine the percentage acceptance of a candidate to a particular university.	High	Sprint-2
Developer	Model integration using pickle file	USN-4	Persist the model with highest R^2 score as a pickle file and integrate it with the web app.	I was able to use the persisted model in the web app	Low	Sprint-2
Developer	Running Jupyter Notebook in Watson Studio	USN-5	Push the dataset CSV files and the Jupyter notebook as an asset in IBM cloud's deployment space.	Was able to execute the Jupyter notebook in IBM Watson studio.	Medium	Sprint-3
Developer/End User	Deploying Model in IBM Cloud	USN-6	Use IBM Watson's ML Service and deploy the multiple-linear regression model. Create a scrollable endpoint of the deployed model so that it can be accessed as an API.	Was able to make predictions using the deploy model.	High	Sprint-3
End User	Hosting	USN-7	Hosting the web app in Streamlit cloud platform and	We are able to access the application using the scoreable endpoint.	Medium	Sprint-4
End User	Integration	USN-8	Integrating the web app with the deployed model	We are able to test the deployed model via an API	Low	Sprint-4

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Developer	CI/CD	USN-9	Connect the respective Github repo and branch to Streamlit cloud platform and set up CI-CD to automatically deploy new changes that's pushed to the repo.	The deployment changes were being pushed to the cloud on a commit to the repo.	Low	Sprint-4
Developer	Model Performance Testing	USN-10	Evaluate the performance of the Multiple Linear Regression model using various model testing measures such as Accuracy, Recall, F1-score	We can tell how good the model is able to perform under various circumstances.	Medium	Sprint-4
Developer/User	Web Application Testing	USN-11	Test the performance of the web app and generate reports	We were able to test various UI components present in the web application	Medium	Sprint-4