**IMPLEMENTATION:**

**MODULES:**

* User
* Admin
* Data preprocess
* Model Execution

**MODULES DESCRIPTION:**

**User:**

The User can register the first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the customer. Once admin activated the User then user can login into our system. The dataset collected from US Bureau of Transport is not directly processed. Before process we need to clean the data. Once clean the data then user can test the departure delay performance based on selected models. The user can see the results in the browser. The all error scores displayed and graphical representation can be displayed.

**Admin:**

Admin can login with his credentials. Once he login he can activate the users. The activated user only login in our applications. We have studied from various sources to find out which parameters will be most appropriate to predict the departure and arrival delays. After several searches we conclude the dataset parameters are Day, Departure Delay, Airline, Flight Number, Destination Airport, Origin Airport, Day of Week, Taxi out. So this data we consider for further process.

**Data Preprocess:**

The admin provided data has been stored in the sqlite database. To process our methodology we need to perform data cleaning process. By using pandas data frame we can fill the missing values with its mean type. Once data cleaned the data will be displayed on the browser.

**Model Execution**

Machine learning models such as Logistic Regression, Decision Tree Regression, Bayesian Ridge, Random Forest Regression and Gradient Boosting Regression we predict result. The MSE is appropriate for our regression problems since it is differentiable, contributing to the stability of the algorithms. It also heavily punishes the bigger errors over smaller errors. MAE is a risk providing metric which tells the expected value of the absolute error loss. Explained Variance Score proportion with which our machine learning model explains the scattering of the dataset is measured by this technique. R2 Score Goodness of fit is indicated by this metric and hence it measures the probability of the model to predict unknown samples, through the proportion of explained variance. The best score can be 1.0 and the score can also be negative.