Introduction

- This project is mainly focused on predicting the sentiment of the review passed to the trained model.
- So the work focused on this project is to pass a review to the model using front-end designed for taking the input, and the prediction of sentiment will be printed out

Dataset used

- The model is trained using IMDB movie dataset which contains 50000 highly polarised movie reviews along with their sentiment for sentiment analysis.

Proposed Methodology

- The project is divided into 4 stages
    - Data Cleaning and Preprocessing
    - Creating Bag-of-words model
    - Training the classification mode
        - Naive Bayes
        - Logistic Regression
    - Testing the model

- Using Naive bayes model for sentiment analysis gave accuracy around 76% and precision of 82% and for the same data using Logistic Regression model gave accuracy of 87% and precision of 86%.

- So the project is divided into two sections

    - model trainer.py
    - mainWindow.py

- Model trainer is focussing on first three stages of project i.e cleaning, pre-processing, creating Bag-of-Words model and training of model.
- MainWindow focuses on front-end of the project and is used to take input of the review from the user and passed to model for prediction