# NLP Project 1

<https://drive.google.com/file/d/1QEPJLixxDMEE-_v4iBcoj58cptqAL0zy/view?usp=sharing>

The primary objective of this project is to build a machine learning classification model that can predict the sentiment of IMDb movie reviews. The dataset contains a collection of movie reviews, and each review is labeled as either positive or negative.

We need to import the necessary libraries and after that we need to load our data in jupyter notebook.

We need to check for the null values and as the data shows there are no null values.

We check for the sentiments and there are 50% positive sentiments and 50% negative sentiments.

We need to use the preprocess the text. In the preprocess text we lowercase the text and tokenize the sentence and use the lemmatize the text.

After preprocessing we use the Tf-Idf vectorizer after vectorizer we train our data and use the machine learning algorithms like logistic regression, naïve\_bayes, svm and random forrest.

The results of the algorithm are as follows-

Logistic regression-0.8826

Naïve\_bayes-0.8499

Svm-0.8821

Random forrest-0.8439

The result clearly shows that the logistic regression is providing the best results.

The Lstm model results are as follows –

Lstm-0.8871