**Project Proposal**

**Title:** Movie Review Sentiment Analysis

**Team Members:** 1. Viswa Teja Ravipati

2. Sai Vivek Uppala

**GitHub**: <https://github.com/tejaravipati/NLP-Project>

**Goals and Objectives:**

* **Motivation:**

The Success of a movie is based on its quality. That quality is known from the opinions or reviews(positive/neutral/negative) of the audience who watched the movie. Sentiment analysis is the most famous applications of Natural Language Processing and is most useful for identifying customers/audience opinion. Producers and moviemakers through social media and online platforms can find out reviews, views and thoughts from the viewers.

A textual movie review tells us about the strong and weak points of the movie and deeper analysis of a movie review can tell us if the movie in general meets the expectations of the reviewer. Sentiment analysis helps us to understand the state of the audience whether they are positive or neutral or negative about the movie.

* **Significance:**

The significance of this project is to use sentiment analysis on a set of movie reviews given by the audience and to predict overall response of the movie whether the audience liked the movie or not.

* **Objectives:**

The main objective of this project is to predict the movie review based on the sentiment for several movie reviews obtained from the Internet or from the text movie review dataset.

* **Features:**

1. Data Pre-processing: It includes Removing HTML tags, Removing special characters, converting to lowercase, Tokenization, Removing stopwords and stemming.
2. Feature Selection: Creating features from bag of words by making use of scikit learn. The bag of words learns a vocabulary from all the documents. And the models the eac document by learning howmany tims each words appears. TF-IDF method is also used to know how important a word is in the document.
3. Modelling and Classification: In this we make use of Random Forest and SVM Classifier for opinion classification.
4. Evaluation: Evaluating the the model by finding the accuracy

**References:**

1. <https://www.kaggle.com/c/word2vec-nlp-tutorial/overview/part-2-word-vectors>
2. <https://towardsdatascience.com/sentiment-analysis-a-how-to-guide-with-movie-reviews-9ae335e6bcb2>
3. <https://medium.com/nerd-for-tech/analysing-movie-reviews-using-sentimental-analysis-77e28e463b1b>
4. <https://en.wikipedia.org/wiki/Sentiment_analysis>