INTRODUCTION:

This project is based on the topic ‘Movie Success Prediction and Sentiment Study’.

ABSTRACT:

This project is all about predicting the success rate of a list of movies using IMDB/Kaggle data and based on the genre and other factors, sentiment for the movie is also analysed in this project.

TOOLS USED:

- Excel is used to extract data from Kaggle and import the data
- Python is used in this project to predict the success rate of the movie through regression method and sentiment study using VADER library.

STEPS INVOLVED IN BUILDING THE PROCESS:

1. Firstly, data is downloaded from Kaggle website and imported to Jupyter Notebook.

2. For predicting the success rate of the movie regression model is used, this model is used by importing train test split and Random Forest Regressor library.
3. The Regression Model is done through a scatter plot with X axis as actual revenue and Y axis as predicted revenue.
4. Next, to analyze the sentiment factor for the list of movies, the dataset is reviewed.
5. It provides the result of sentiment as positive and negative for the whole dataset from which a sample is taken to plot a graph.
6. Lastly, a bar graph is plotted with X axis as positive and negative and Y axis with number of sample taken.

CONCLUSION:

This project deals in success rate prediction of a list of movies and sentiment analysis of the movies. Through this project, I have learned to use different libraries in Python and learned how to analyse based on sentiment which was new to me and had a different procedure.