# Project Report

By Sarvesh Pai

1714098 LY-IT

### Purpose

To create a Sentimental Analysis Model to correctly identify and determine the sentiment of sentences extracted from amazon reviews, yelp reviews and imdb reviews. This has been performed on Python using **NLP techniques and Support Vector Machine**.

### Hypothesis

Opinion mining (sometimes known as sentiment analysis or emotion AI) refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. So my hypothesis is to determine which words could possibly be responsible to bring out a positive or negative sentiment.

### Motivation

There are so many websites and sources where sentences and documents written can be found . Determining the motive and sentiment of those sentences without physically looking through every sentence can not only save time but also prove to be immensely helpful for analysis and decision making.

### Algorithms Used

NLP algorithms like Lemmatization and Stemming, N-gram Analysis and ML Algorithm Support Vector Machine.

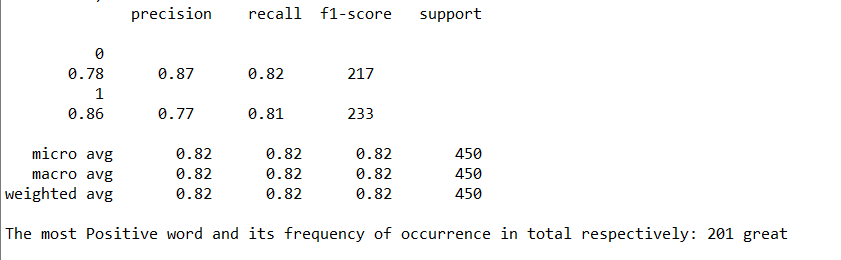
### How I used it?

I went through the essential pipeline of NLP: Data Acquisition, Text Preprocessing, Text Tokenization, Text Cleaning, Vectorization and then finally ML- Algorithm Modelling.

Data was acquired from Kaggle. I used nltk library for lemmatization after tokenization and stop word removal. This data then used to create a corpus which was then used for n-gram analysis.

This data was then fitted into a SVM to train the model and find how accurate the model was.

### Output



Where 0 stands for Negative and 1 stands for Positive.

### Future Works

There is a considerable application of this project after I have made its predictions (currently at 86%) more accurate and equipped enough to handle more complex sentences.

### Conclusion

Successful development of a simple sentimental analysis model using NLP