CSCE 5290: NATURAL LANGUAGE PROCESSING
PROJECT PROPOSAL

GROUP – 10

Deeksha Thandra - 11545852

Avinash Chinta – 11523130

Sushmitha Dharmareddy – 11549862

Mavya Tekula – 11511499

Project Title:

Sentiment Analysis on Customer Reviews of Amazon Products.

Goals and Objectives:

- **Motivation:**

Now a days**,** we can observe that product reviews are making a significant impact due to the evolution of retail stores to e-commerce platforms. Lot of users are posting their reviews on different e-commerce sites and thereby helping the future users to get an idea about the quality of the products. The purpose of this project is to investigate how the companies can conduct sentiment analysis based on their reviews to gain more insights into customer experiences. Around 100 amazon review files will be analyzed using sentiment analysis along with NLP techniques to understand customer experiences using Amazon reviews and to get the feedback about the products.

- **Significance:**

Today's social media platforms and online networks enable businesses to gather honest feedback from clients all around the world. Customers' experiences with prices, value, quality, customer service, ease of shopping, and other

aspects of their online purchases are revealed in customer reviews. Since the customer reviews are unstructured, sentiment analysis will make it easier and more affordable to make sense of them.

- **Objectives:**

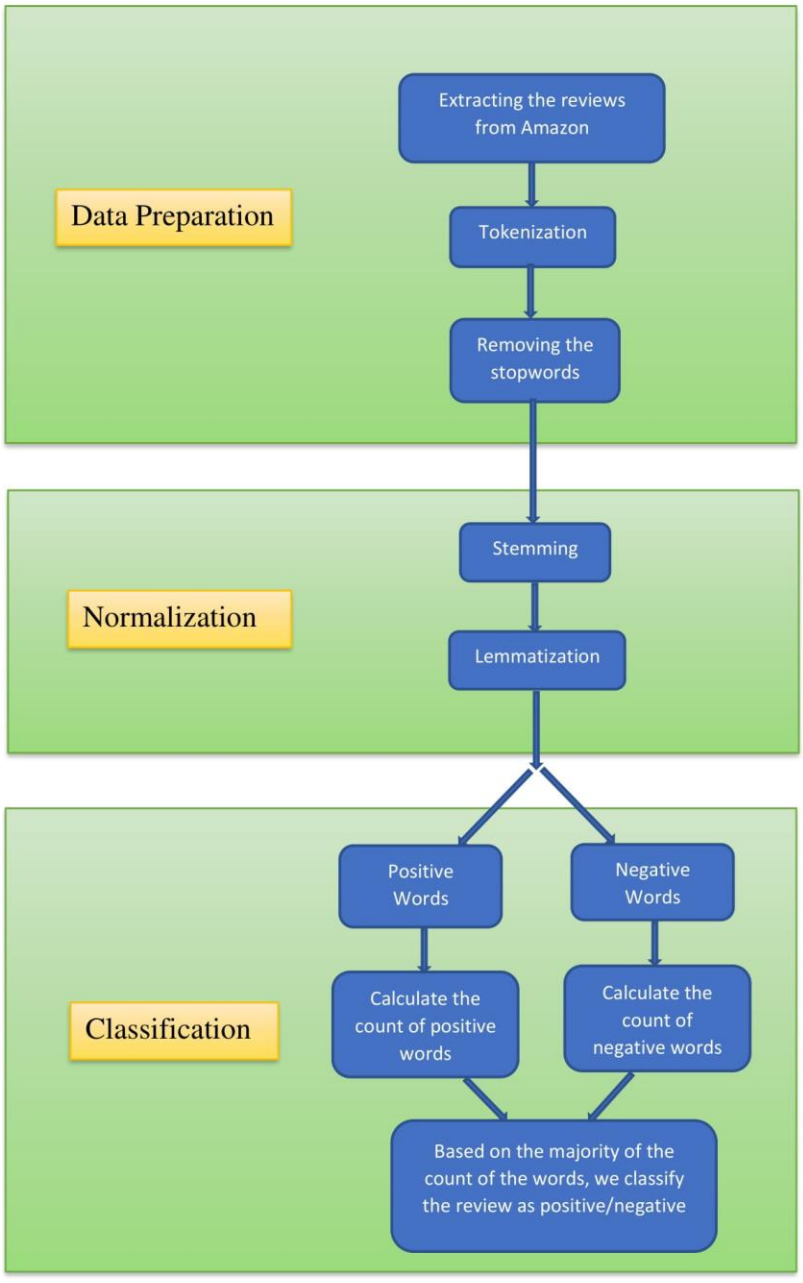
We will extract the data from all the 100 input files and then we'll implement the tokenization to split the text into individual tokens and then we are removing the stopwords using NLTK libraries.

We'll implement the normalization techniques like stemming and lemmatization which helps us to extract the root words from the generated tokens. Furthermore, we will use 'vader sentiment' library from python and the 'SentimentIntensityAnalyzer' module from above library to identify the inclination of people's opinions.

- **Features:**

Main feature of the project is to classify the reviews based on the opinions such as positive, negative and neutral. After normalizing and processing the data using the SentimentIntensityAnalyzer, we will extract the count of positive and negative tokens and classify the reviews as positive or negative feedbacks based on the count. Let's consider a amazon review for motor vehicle which says "This bike is a really nice bike, huffy did a great job!". By examining the above review, we count the positive and negative words, then we will predict if the above review is positive, negative or neutral.

Workflow diagram:



References:

1) https://techvidvan.com/tutorials/python-sentiment-analysis/
2) https://www.researchgate.net/publication/344869545_Sentiment_Analysis_on_Amazon_reviews
3) https://www.researchgate.net/publication/330871275_Natural_Language_Processing_Sentiment_Analysis_and_Clinical_Analytics

GitHub Link:

https://github.com/avinashchinta99/nlp-project-group-10