

## Phase 1: Problem Definition and Design Thinking

### Phase 1: Problem Definition and Design Thinking

<b>PROJECT TITLE</b>	<b>SENTIMENT ANALYSIS FOR MARKETING</b>
<b>TEAM ID</b>	<b>5277</b>
<b>TEAM NAME</b>	<b>Proj_204223_Team_1</b>
<b>STUDENT NAME</b>	<b>Sneha.K</b>
<b>COLLEGE CODE - NAME</b>	<b>9238 - MANGAYARKARASI COLLEGE OF ENGINEERING PARAVAI, MADURAI.</b>
<b>GITHUB REPOSITORY LINK</b>	<b><a href="https://github.com/Sneha1743/IBM-NaanMudhalvan-AI.git">https://github.com/Sneha1743/IBM-NaanMudhalvan-AI.git</a></b>

## 1. Project Definition

### 1.1 Project Overview

Sentiment Analysis involves classifying a text into various sentiments, such as positive or negative, Happy, Sad or Neutral, also the sentiment or emotional tone expressed in textual data, such as customer reviews, social media posts, survey responses, and more. Here's an overview of sentiment analysis for marketing:

### 2.DESIGN THINKING :

#### 2.1.DATA COLLECTION

##### 2.1.1 DATASET CONTAINING CUSTOMER REVIEW

Customer sentiment analysis is the process of using machine learning (ML) to discover customer intent and opinion about a brand from customer feedback given in reviews, forums, surveys, and so on. Sentiment

analysis of customer experience data gives businesses deep insight into motivations behind purchase decisions, the patterns in changing brand sentiment based on timelines or events, and market-gap analysis that can help in product and service improvement.

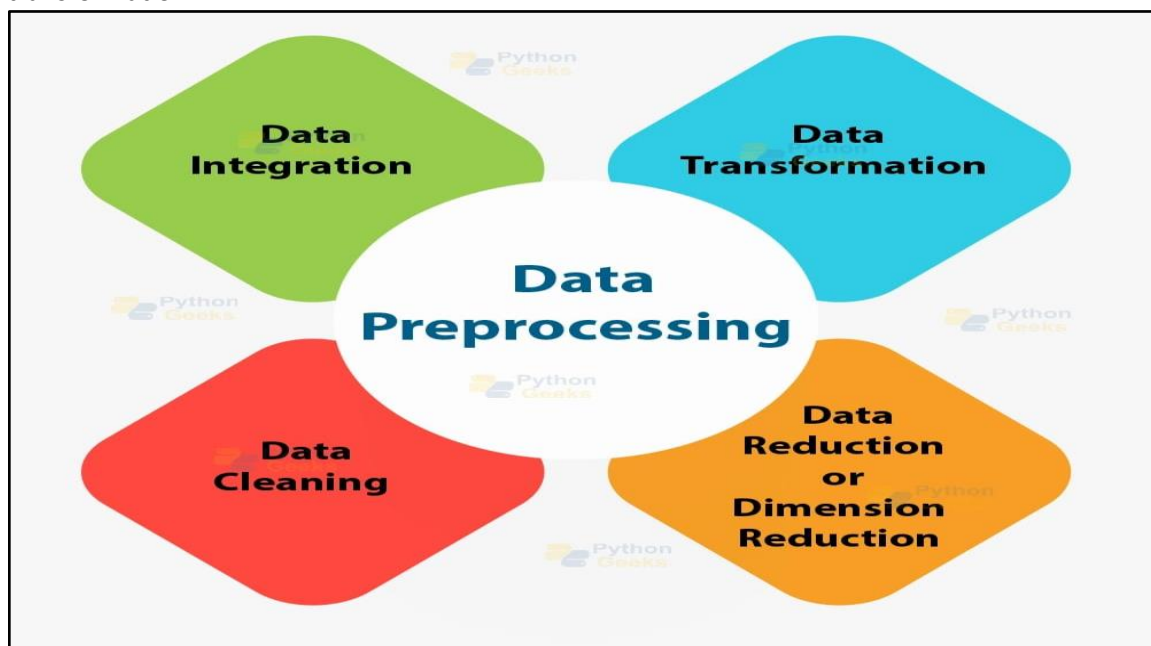
### 2.1.2 SENTIMENTS ABOUT COMPETITOR PRODUCT

Evaluating sentiment about a competitor's product typically involves assessing the opinions and emotions expressed by consumers or reviewers. The sentiment can range from positive, indicating satisfaction and praise for certain aspects, to negative, highlighting. The sentiment analysis process may involve analyzing customer reviews, social media discussions, or expert opinions to gain insights into how consumers perceive and react to the competitor's product. These insights can be invaluable for businesses to identify areas for improvement in their own products.

## 2.2.DATA PREPROCESSING

### 2.2.1DATA CLEAN :

This involves identifying and correcting errors or inconsistencies in the data, such as missing values, outliers, and duplicates. Various techniques can be used for data cleaning, such as imputation, removal, and transformation.



### 2.2.2 PREPROCESS :

Text preprocessing is an essential step in natural language processing (NLP) that involves cleaning and transforming unstructured text data to prepare it for analysis. It includes tokenization, stemming, lemmatization, stop-word removal, and part-of-speech tagging.

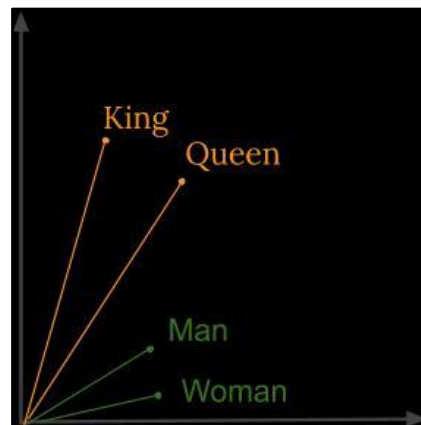
## 2.3 SENTIMENT ANALYSIS TECHNIQUES :

Sentiment analysis techniques, also known as opinion mining, are natural language processing methods used to determine and quantify the emotional tone or sentiment expressed in text data. These

techniques employ a variety of approaches, including machine learning models such as Naive Bayes, Support Vector Machines, or deep learning methods and transformers to classify text as positive, negative, or neutral. These techniques find extensive applications in understanding customer feedback, social media monitoring, and market research, among others, to extract valuable insights from textual data.

### 2.3.1 WORD EMBEDDINGS :

we humans understand the words like king and queen, man and woman, tiger and tigress have a certain type of relation between them but how can a computer figure this out? Word embeddings are a form of word representation that bridges the human understanding of language to that of a machine.



### 2.3.2 TRANSFORMER MODELS :

Transformer models are a class of neural network architectures designed for natural language processing tasks.

They revolutionized the field by introducing the self-attention mechanism, enabling parallelization and capturing long-range dependencies.

They power state-of-the-art language models like BERT, GPT, and T5, achieving impressive results in various NLP tasks.

Transformers excel in tasks like machine translation, sentiment analysis, and text generation.

Their architecture has been adapted beyond NLP, making them a cornerstone of modern deep learning.

### 2.4 FEATURE EXTRACTION :

#### 2.4.1 EXTRACT FEATURES

In sentiment analysis, we detect tweets that have negative sentiment, i.e, racist, sexist or general hate speech. Here, tweets with a label '1' denotes a negative tweet, while '0' denotes the absence of hate speech in the tweet.

For textual data, we need to perform feature extraction in order to train a supervised machine learning model. In this notebook, our main goal is to explore the different feature extraction methods, and not to optimise the performance of the model.

We extract features using this :

#### 1. Bag of Words Features

Bag of Words model is used to preprocess the text by converting it into a bag of words, which keeps a count of the total occurrences of most frequently used words. This model can be visualized using a table, which contains the count of words corresponding to the word itself.

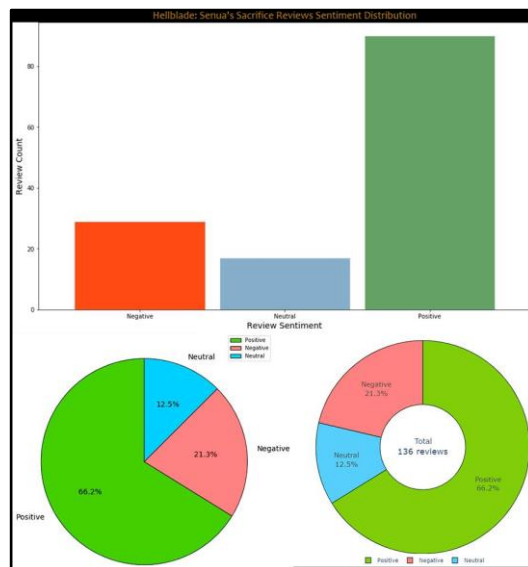
### 2.4.2 SENTIMENT FROM TEXT DATA :

Sentiment analysis is the process of analyzing digital text to determine if the emotional tone of the message is positive, negative, or neutral. Today, companies have large volumes of text data like emails, customer support chat transcripts, social media comments, and reviews.

### 2.5 VISUALIZATION :

#### 2.5.1 SENTIMENT DISTRIBUTION

Sentiment analysis is a technique that uses natural language processing and machine learning to extract and quantify the emotional tone of text or speech. It can help you understand how your customers, users, or stakeholders feel about your products, services, or topics.



#### 2.5.2 ANALYZE TRENDS :

Sentiment analysis is the task of automatically identifying and extracting the opinions, emotions, and attitudes expressed in natural language texts. It has many applications in domains such as social media, customer service, e-commerce, and marketing.

### 2.6 INSIGHTS GENERATION :

Targeted Sentiment Analysis (TSA) is a contemporary AI method for generating valuable insights from user reviews. Such insights may aid consumers in their decision-making, or help companies when they strive to understand customer satisfaction and guide marketing campaigns.