**Identify Sentiment Analysis **

## **Business Understanding:**

## Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material, and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations. Brands can use this data to measure the success of their products in an objective manner. In this challenge, we are provided with tweet data to predict sentiment on electronic products of netizens.

## **Problem Statement:**

Sentiment analysis remains one of the key problems that has seen extensive application of natural language processing. This time around, given the tweets from customers about various tech firms who manufacture and sell mobiles, computers, laptops, etc, the task is to identify if the tweets have a negative sentiment towards such companies or products.

**Data Understanding: (What type of data we have)**

**id (int) :** The id of the Tweet

**label(int) :** Whether the tweet is indicating -ve(0) or +ve(1) Sentiment

**tweet(object) :** The tweets posted by customers about various tech firms

**Data Preparation:**

This stage involves doing a little Exploratory Data Analysis(if any) and thinking about how our data will fit into the model we are going to build. Is the data in data types that are compatible with model or not? Is there any missing values in the given data?

If we have any missing values in the data, we should do Imputation. If categorical data is there, we should use Dummy variables for that column before passing that to the ML algorithm we are building.

**Modeling :**

Choose a model and turn the parameters before fitting it to our training set of data. Python's **scikit learn** library is a good place to get model algorithms.

**Evaluation :**

Withhold a test set of data to evaluate the model performance. We can use scikit-learn's Metrics class for measuring performace of our model.

**Deployment/Prototyping :**

Deployment and implementation are some of the key components of any data driven project.

**Solution:**

Team Members,

1. Sai kumar

2. Asim Ahmed

3.Krishna

4. Satish[[1]](#footnote-1)

1. [↑](#footnote-ref-1)