Team ID: 99626

Project Name: Sentiment Analysis

Project Type: UDP

Literature Review Report

170020107041 (Tirth Patel), 170020107049 (Ravi Sahani)

Abstract:

- In the age of artificial intelligence and machine learning, competition is between best and best. So inorder to gain control over market, it is essential to understand market condition especially during covid-19 situation.
- For that sentiment of market is very important and sentiment of market is what consumer think of certain product.
- Sentiment analysis is the interpretation and classification of emotions (positive, negative and neutral) within text data using text analysis techniques. Sentiment analysis tools allow businesses to identify customer sentiment toward products, brands or services in online feedback.
- Apply the following Machine Learning algorithms to arrive at the best result:
    - K-Means Clustering
    - Support Vector Machine
    - Logistic Regression
    - K Nearest Neighbours

Keywords:

Sentiment, NLP, Python, Jupyter Notebook, Twitter, API

Tools and Technology:

Frontend : HTML, CSS, Javascript

Backend : Python, Jupyter Notebook

Libraries : Numpy, Pandas, Scikitlearn, Matplotlib, Seaborn

Study of current Existing System:

- Sentiment analysis has been handled as a Natural Language Processing task at many levels of granularity.
- Starting from being a document level classification, it has been handled at the sentence level and more recently at the phrase level.
- Microblog data like Twitter, on which users post real time reactions to and opinions about “everything”, poses newer and different challenges.

Limitation of Existing System:

- They build models using Naive Bayes, Max Ent and Support Vector Machines (SVM), and they report SVM outperforms other classifiers. In terms of feature space, they try a Unigram, Bigram model in conjunction with parts-of-speech (POS) features.
- The data they use for training and testing is collected by search queries and is therefore biased.
- Our data will be a random sample of streaming tweets unlike data collected by using specific queries.
- The size of our hand-labeled data will allow us to perform cross validation experiments and check forth variance in performance of the classifier across folds.

Usability of project:

- Real time data analysis makes it possible for business organizations to keep track of their services and generates opportunities to promote, advertise and improve from time to time. Through comprehensive analysis, businesses gain valuable insights towards their customers.
- It can help enhance customer experience towards brands and business. Insights on what is being done right from positive sentiment analytic reports and what needs improvement.
- It can help revitalize a business' brand. Companies can quantify people's perception about their products or services, marketing strategies and their customer experience.
- When used right, it can help companies develop engaging marketing strategies to improve people's perception about their brands.
- Can provide a competitive advantage- A trove of business opportunities and product possibilities.
- Gives predictions of consumer trends so companies can develop strategies to gain an advantage. Gives an insight on what customers are looking for, what they like or what they want to see.

Innovative ideas incorporated in system:

- Tweet Sentiments : Returns the Sentiments of tweets.
- Text Processing : Performs sentiment analysis, parts of speech tagging and chunking, phrase extraction and named entity recognition.
- ML Analyzer : Preforms text classification, article summarization, sentiment analysis, extracts stock symbol, extracts person names, language detection, and extracts locations.