SENTIMENT ANALYSIS FRAMEWORK FOR SOCIAL MEDIA

GROUP NO : 14

Riddhi Prajapati - 15104015

Pratiksha Patil - 16204011

Aafreen Shaikh - 16204009

Guide- Prof. Sunil A. Sushir

Abstract

Social media in today's world is a collection of people's feedback, emotions, reviews and even personal experiences which result in proper knowledge of the given content.The time spent by uses is almost two or more hours looking for papers that asks search engine to optimize the result. Sentiment analysis is a method which analyses emotions or views of an individual regarding to a topic. Here we design a framework for sentiment analysis in case of social media data. The available data is analyzed in terms of positive or negative and an clear summary is obtained. The input data that needs to be analyzed is given as input to the NLP (Natural Language Processing) and processed.


Introduction

- Sentiment data is unstructured data that represents opinions, emotions, and attributes contained in sources such as :
  - Social media posts
  - Blogs
  - Online product reviews
  - Customer support interactions.
- Sentiment analysis aims to determine the attitude or the emotional communication of the public with respect to some topic.
- R programming language will be used for data analysis along with NLP algorithm and Corpus approach.

Corpus

- Corpus is a collection of text document over which we apply text mining or NLP.
- Corpus may have written language / Spoken language. Corpus analysis give us semantic information.
- There are 3 types of corpus used for sentiment analysis - Positive, Negative, Neutral.
- Corpus is used after tokenization for spelling correction or data cleaning and grammar checking.

R Programming language

R Programming language used for -

- Statistical interface
- Data Analysis

R language is basically made for statistics.

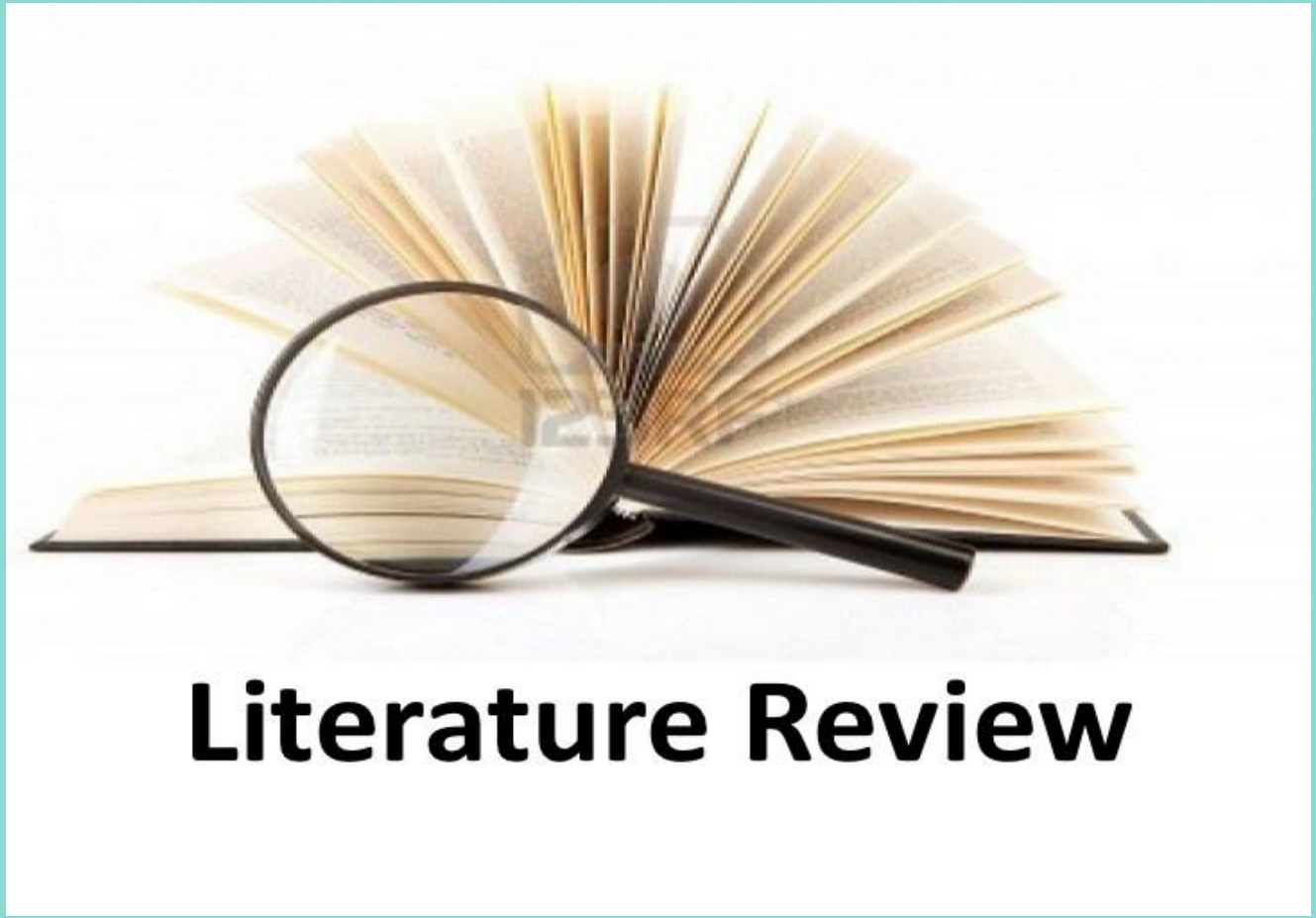
R can perform visualization and prediction.

Machine Learning & NLP (Natural Language Processing)

- Both ML & NLP are subsets of Artificial Intelligence
- ML deals with developing systems which improve previous performance whereas NLP deals with developing systems that understand computer language and does human interpretation.
- We obtain better results when NLP is used along with ML.

Objectives

- To implement an algorithm for automatic classification of text into positive, negative or neutral.
- Sentiment analysis to determine the attitude of the mass is positive, negative or neutral towards the subject of interest.
- Graphical representation of the sentiment in form of Word Cloud and also frequency plotting.


Literature Review

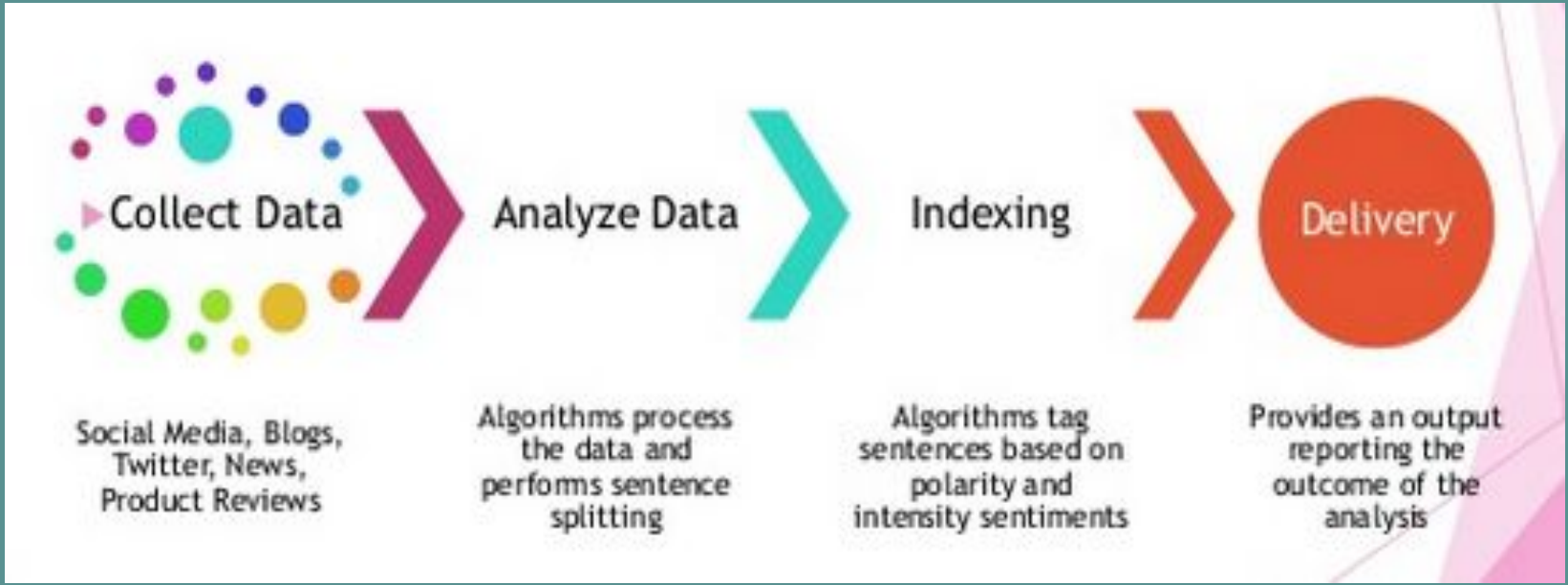
Sr.No	Title Of Paper	Idea	Tool/Algorithm	Issues	Overcome
1	Onto-based Sentiment classification using Machine Learning Techniques.	The basic idea of this paper is performing Sentiment Analysis on customers social media data which has direct impact on business world.	Sentiment Analysis, Machine Learning, NLP, Statistics, R-programming Language	In preprocessing stage, various rules and regulations are defined to standardize the text that make text mining process efficient.	The machine learning algorithms are applied for better results.
2	Text Mining: Techniques, Applications and Issues	Text mining techniques are used to analyze the interesting and relevant information effectively and efficiently from large amount of unstructured data.	Classification, Knowledge Discovery, Information Extraction ,k-means clustering	-	To develop specific algorithms to overcome the issues.

Sr.No	Title Of Paper	Idea	Tool/Algorithm	Issues	Overcome
3	A Framework for Sentiment Analysis with Opinion Mining of Hotel Reviews	The basic idea of this paper is taking textual data of hotel reviews. Sentiment Analysis is been done based on customer perspective.	Opinion mining, Sentiment Analysis, Machine Learning algorithm , Natural Language Processing(NLP)	Issues arise that some comments may be wrongly viewed as neutral i.e some words which are not included in dictionary are classified under neutral.	Here lot of research is done in fine tuning the feature extraction algorithm.
4	Business reviews classification using sentiment Analysis	The basic idea of this paper is applying Sentiment analysis by applying preprocessing steps , remove stop words, apply stemming, handling negotiations.	Naïve Bayes for text classification, Feature Extraction vector using TF idf.	The accuracy of the system decreased with 14% when compared to first approach result. Here also Sentiment ambiguity occurs.	The performance can be improved by the usage of bigrams or trigrams, word chunks, or part-of-speech as features in order to distinguish the same word features that use different POS.

Problem Definition

A major benefit of Social media is that we can see the good and bad things people say about a particular brand or personality. The bigger your company gets difficult it becomes to keep a handle on how everyone feels about your brand. For large companies with thousands of daily mentions on social media, new sites and blogs, it is extremely difficult to do this manually.

HOW TO DO IT ?


Collect Data
Analyze Data
Indexing
Delivery
Social Media, Blogs, Twitter, News, Product Reviews
Algorithms process the data and performs sentence splitting
Algorithms tag sentences based on polarity and intensity sentiments
Provides an output reporting the outcome of the analysis

Summary

People have always had an interest in what people think, or what their opinion is. With the interesting numbers of people are using websites and services to express their opinion. With social media, it is becoming easier to automate what public opinion is on a given topic, news story,product, or brand. Opinion that are analyzed from such services can be valuable. Data sets that are gathered can be analyzed and presented in such a way that it becomes easy to identify if the online mood is positive,negative or neutral. This allows individuals or business to be alert when a negative opinion arises, alternatively, positive sentiment can be identified thereby allowing the identification of product advocates or to see which parts of a business strategy are working.

References

[1] Walaa Medhat, Ahmed Hassan, HodaKorashy, Sentiment analysis algorithms and applications: A survey, "Onto-based sentiment classification using Machine Learning Techniques", 2014.

[2] R.Ezhilarasi, R.I. Minu, Automatic Emotion Recognition and Classification, International Conference on modelling, optimization and computing, 2012.

[3] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," Proceedings of the ACL-02 conference on Empirical methods in natural language processing, 2002.

[4] S. M. Weiss, N. Indurkhya, T. Zhang, and F. Damerau, Text mining: predictive methods for analyzing unstructured information. Springer Science and Business Media, 2010.


Thank
You