



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SENTIMENTAL ANALYSIS OF TWITTER

Karolinsudharani J , Karthikeyan M, Nandhini P, N. Anandkumar Assistant professor, Department of Information Technology
SNS College of Technology ,Coimbatore, Tamilnadu

ABSTRACT

Recently, Social media has arisen not only as a personal communication media, but also, as a media to communicate opinions about products and services or even political and general events among its users. Due to its widespread and popularity, massive amount of user reviews or opinions are produced and shared daily. Analysis of public information from social media could yield interesting results and insights into the world of public opinions about almost any product, service or personality. Social network data is one of the most accurate and effective indicators of public sentiment.

Twitter is one of the most widely used social media micro blogging sites. Mining user opinions from social media data is not a straight forward task; it can be accomplished in different ways. In this work, an open source approach is presented, throughout which, twitter Micro blogs data has been collected, pre-processed and visualized using open source tools to perform text mining and sentiment analysis for analyzing user contributed online reviews. Collecting customer task using conventional methods such as surveys.

The sentiment analysis of the customer opinions makes it easier for business to understand their competitive value in a changing market and to understand their customer views about their products and services, which also provide an insight into future marketing strategies and decision-making policies.

INTRODUCTION

Social media have become an emerging phenomenon due to the huge and rapid advances in information technology. People are using social media on daily basis to communicate their opinions with each other about wide variety of subjects, products and services, which has made it a rich resource for text mining and sentiment analysis. One of the major benefits of using a social listening tool is that this tool can quickly determine how people feel about your brand or product. Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis is widely applied to reviews and social media for a variety of applications, ranging from marketing to customer services.

Sentiment Analysis can be a subjective tool, and understandably businesses and people often vary the process of identifying the positive, negative or neutral sentiment from a comment. Sentiment Analysis aims to determine the attitude of the author of a specific piece of content with respect to the topic of interest. Comments and content can be referred to as positive, negative, neutral or have no sentiment at all. A basic task in sentiment analysis is classifying the polarity of a given text in the document, sentence, or feature/aspect level to determine whether the expressed opinion in a document, a sentence or an entity feature/aspect is positive, negative, or neutral. Advanced, "beyond polarity" sentiment classification looks, for instance, at emotional states such as "angry", "sad", and "happy". A different method for determining sentiment is the use of an scaling system whereby words commonly associated with having a negative, neutral or positive sentiment with them are given an associated number on a -10 to +10 scale (most negative up to most positive).

PROBLEM STATEMENT

The project involves in extracting the data set from the Twitter API and analyzing the data to find the sentimental score .Starting with Twitter is a great starting point for socialweb mining because of its inherent openness for public consumption, clean and well documented API, rich developer tooling, and broad appeal to users from every walk of life. Twitter data is particularly interesting because tweets happen at the “speed of thought” and are available for consumption as they happen in near real time, represent the broadest cross- section of society at an international level. In some point this might help to find a business need and execute a data mining process.

LITERATURE SURVEY

Sentiment analysis is a growing area of Natural Language Processing with research ranging from document level classification (Pang and Lee 2008) to learning the polarity of words and phrases (e.g. (Hatzivassiloglou and McKeown 1997; Esuli and Sebastiani 2006)). Given the character limitations on tweets, classifying the sentiment of Twitter messages is most similar to sentence level sentiment analysis

PROPOSED SYSTEM

To overcome the drawbacks of the methods we have reviewed above, we propose a new model for sentiment analysis. In this model we combine many techniques to reach our final goal of emotion extraction. The steps for the process are documented below.

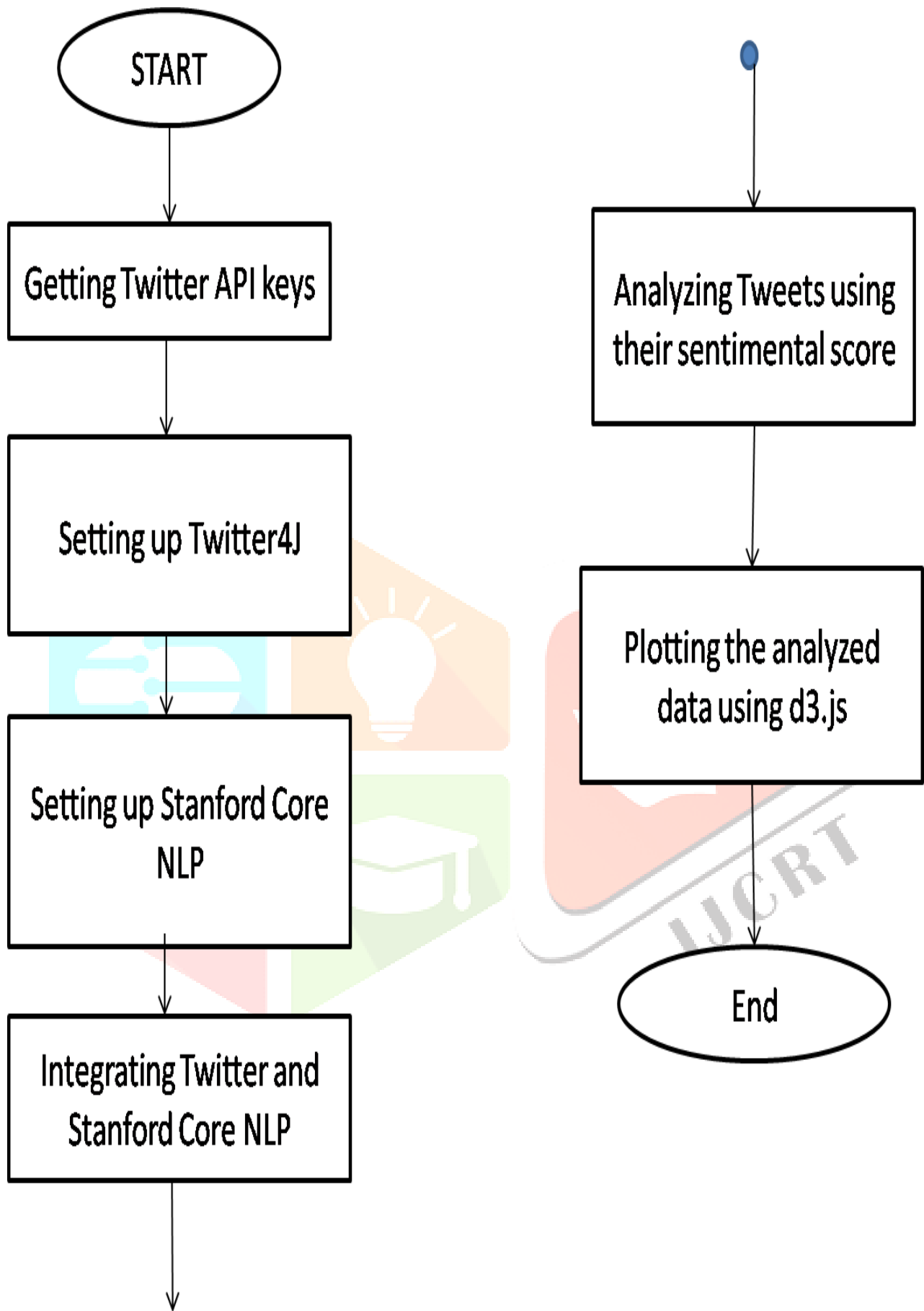
1. Retrieval of Data: Public Twitter data is mined using the existing Twitter APIs for data extraction. Tweets would be selected based on a few chosen keywords pertaining to the domain of our concern, i.e. product reviews. We have elected to use the Twitter API due to ease of data extraction.

2. Tweet Correction: As tweets are written for human perusal, they often contain slang, misspellings and other irrelevant data. Thus we correct the misspellings in the sentences and look to replace the slang in the sentences with words from standard English that may roughly relate to the slang in question. As slang itself can be used to display a wide variety of sentiment, often with greater emotional impact, this process is necessary so that slang words may be considered as part of the emotion expressed.

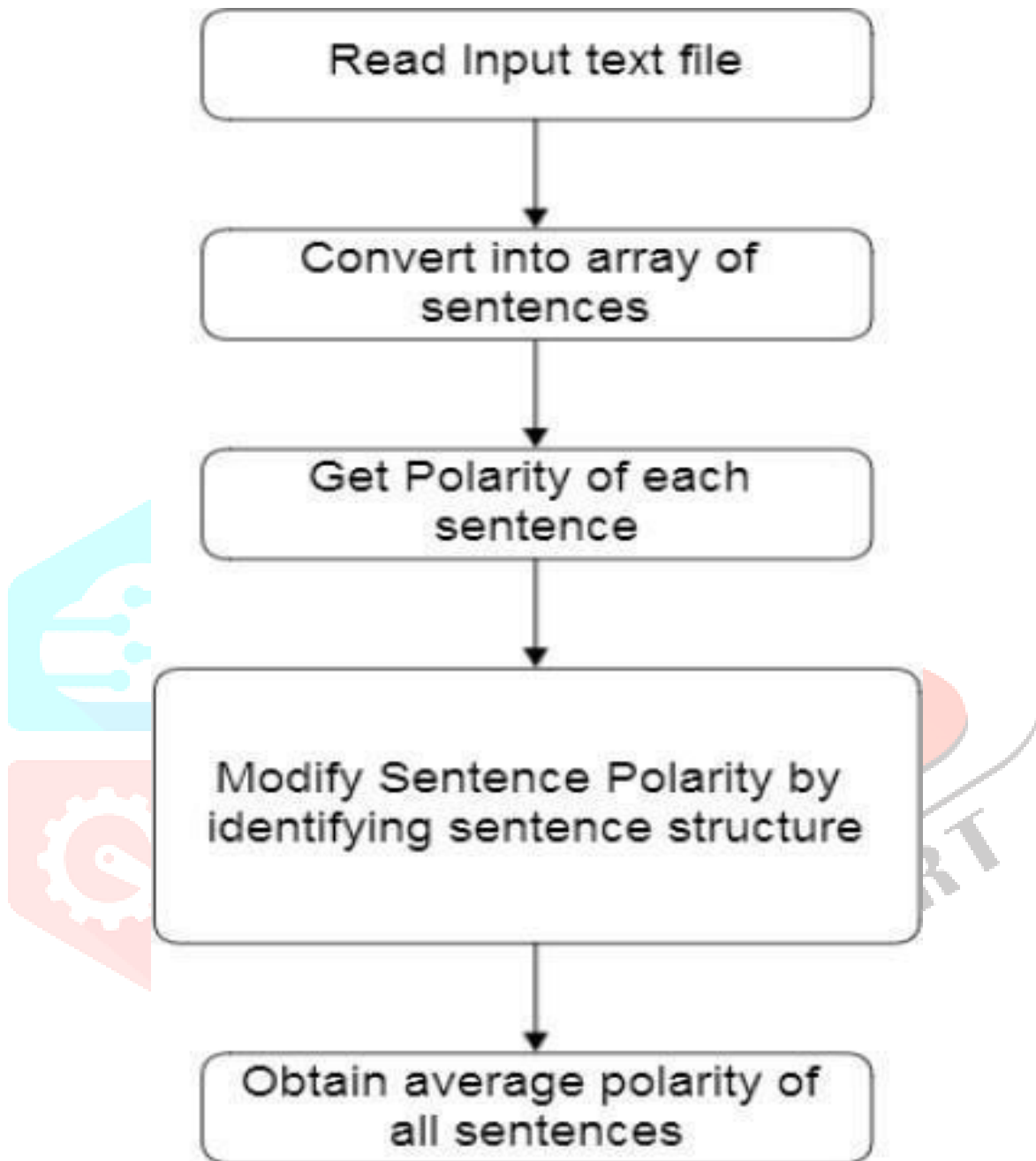


HUMAN LANGUAGES SUPPORTED

The basic distribution provides model files for the analysis of well-edited English, but the engine is compatible with models for other languages. Packaged models for Chinese, French, German, and Spanish are provided. A jar that contains all of our English models, which includes various variant models, and in particular has one optimized for working with uncased English are provided.(e.g., mostly or all either uppercase or lowercase). Stanford NLP models for Arabic are also usable inside CoreNLP.



CALCULATION OF SENTIMENTAL SCORE/ POLARITY



Graph plotted using d3.js for the topic Paralympics

