Journal of Engineering Sciences
Vol 11, Issue 6,June/2020
ISSN NO:0377-9254

SENTIMENT ANALYSIS USING PRODUCT REVIEW DATA

T. Aruna Jyothi[1], P. Sai Goutham[2], N. Heetesh Singh[3] and M. Sai Srujan[4]

[1]*Assistant professor, CSE, BE, Matrusri engineering college, Saidabad, Telangana State, India.*

[2]*Student, CSE, BE, Matrusri engineering college, Saidabad, TS, India.*

[3]*Student, CSE, BE, Matrusri engineering college, Saidabad, TS, India.*

[4]*Student, CSE, BE, Matrusri engineering college, Saidabad, TS, India.*

***Abstract*— Sentiment analysis or opinion mining is the major task of Natural Language Processing. Sentiment analysis has gained much attention in recent years. In this paper, we aim to tackle the problem of sentiment polarity categorization, which is one of the fundamental problems of sentiment analysis. A general process for sentiment polarity categorization is proposed with detailed process descriptions. Experiments for sentence-level categorization is performed with promising outcomes.**

***Keywords*— Sentiment analysis, we b scraping, text blob, Polarity, Review data, Positive, Negative, Neutral.**

1. INTRODUCTION

Sentiment is an attitude, thought, or judgment prompted by feeling. Sentiment analysis, which is also known as opinion mining, studies people's sentiments towards certain entities. From a user's perspective, people are able to post their own content through various social media, such as forums, micro-blogs, or online social networking sites. From a researcher's perspective, many social media sites release their application programming interfaces (APIs), prompting data collection and analysis by researchers and developers. However, those types of online data have several flaws that potentially hinder the process of sentiment analysis. The first flaw is that since people can freely post their own content, the quality of their opinions cannot be guaranteed. he second flaw is that ground truth of such online data is not always available Internet is a resourceful place with respect to sentiment information from a user's perspective, people are able to post their own content through various social media, such as forums, micro-blogs, or online social networking sites. "It is a quite boring movie…….. but the scenes were good enough. " The given line is a movie review that states that "it" (the movie) is quite boring but the scenes were good. Understanding such sentiments require multiple tasks. Hence, SENTIMENTAL ANALYSIS is a kind of text classification based on Sentimental Orientation (SO) of opinion they contain. Sentiment analysis of product reviews has recently become very popular in text mining and computational linguistics research.

2. SYSTEM DESIGN

All of the opinions and data that is gathered from your customers is based on the language they use when they talk about your product. So, if they are referring to your product or business in a positive, negative or neutral way, you will know about it through sentiment analysis.In order to find these opinions, data-miners use a method called Natural Language Processing (NLP). NLP is basically a system that is built to extract opinions from text and tell the difference between all the words, automatically. The project is carried out in following steps:

- Designing GUI
- Scraping the review data
- Data added to yml file
- Reviews to excel sheet
- Calculation of polarities.

Web scraping, web harvesting, or web data extraction is data scraping used for extracting data from websites. Web scraping software may access the World Wide Web directly using the HTTP, or through a web browser. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot. It is a form of copying, in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis.

YAML stands for YAML Ain't Markup Language, and this file format technology is used in

documents. These documents are saved in plain text format and are appended with the yml extension. Efficient data serialization was the main objective of the developers of the yml format, since it enables users to create yml files with content independent from any particular markup language. These yml files can also be read by any text editor developed for creating, opening and editing plain text files, be it text editing software for Microsoft Windows-based systems like Microsoft Notepad and Microsoft WordPad, or for Mac platforms like Apple TextEdit software. The data needed was right there in the HTML, and not a lot of transformation was required. Part of the frustration was that there were quite a lot of data points to grab from the website, for each row of data. Some might be in a nested format with parent element and child elements grouped together in different combinations, and we needed to format each of those fields. *TextBlob* is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.

3. SYSTEM DESIGN

Web scraping a web page involves fetching it and extracting from it. Fetching is the downloading of a page (which a browser does when you view the page). Therefore, web crawling is a main component of web scraping, to fetch pages for later processing. Once fetched, then extraction can take place. The content of a page may be parsed, searched, reformatted, its data copied into a spreadsheet, and so on. Web scrapers typically take something out of a page, to make use of it for another purpose somewhere else. An example would be to find and copy names and phone numbers, or companies and their URLs, to a list (contact scraping).

The GUI window is created using tkinter module which is present as inbuilt in python 3.6. The title of window is Sentiment analysis and a bold text is displayed as SENTIMENT ANALYSIS USING PRODUCT REVIEW DATA with an image of three different types of sentiments. It also has text and textbox where a product's URL is entered. The model scrapes the review data from the website when a user clicks on submit button.

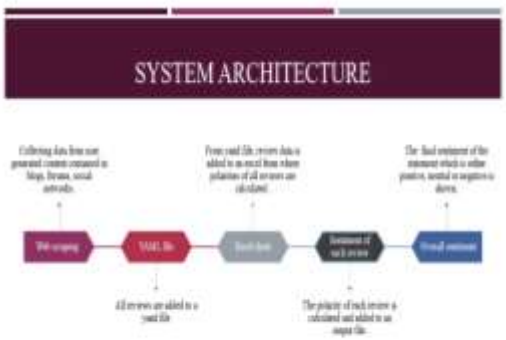

Fig.1: System architecture

4. REQUIREMENT SPECIFICATIONS

Tkinter:

What Is A Graphical User Interface (GUI)?

Graphical User Interface (GUI) is nothing but a desktop application which helps you to interact with the computers. They are used to perform different tasks in the desktops, laptops and other electronic devices.

- **GUI** apps like **Text-Editors** are used to create, read, update and delete different types of files.
- **GUI** apps like Sudoku, Chess and Solitaire are games which you can play.
- **GUI** apps like **Google Chrome, Firefox and Microsoft Edge** are used to browse through the **Internet**.

They are some different types of **GUI** apps which we daily use on the laptops or desktops. We are going to learn how to create those type of apps.

What Is Tkinter?

Tkinter is actually an inbuilt **Python** module used to create simple **GUI** apps. It is the most commonly used module for **GUI** apps in the **Python**.

We don't need to worry about installation of the **Tkinter** module as it comes with **Python** default. We can create our own GUIs.

Text blob:

The TextBlob's sentiment property returns a Sentiment object. The polarity indicates sentiment with a value from -1.0 (negative) to 1.0 (positive) with 0.0 being neutral. The subjectivity is a value from 0.0 (objective) to 1.0 (subjective).

5. EXPERIMENTAL RESUTS

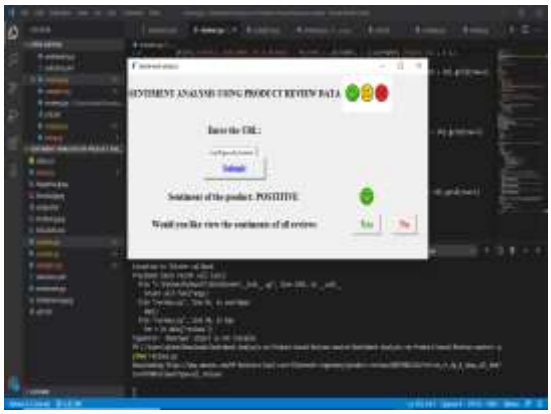

Fig.2: Positive review

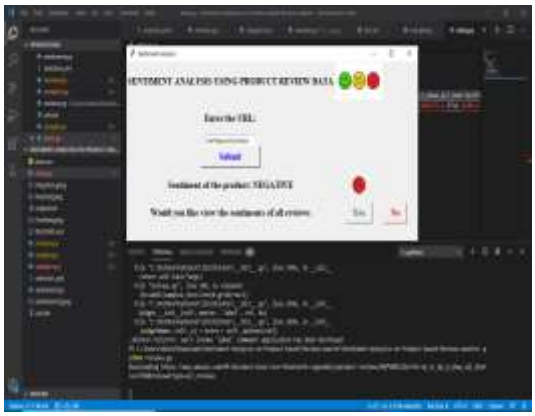

Fig.3: Negative review

Fig.4: Neutral review

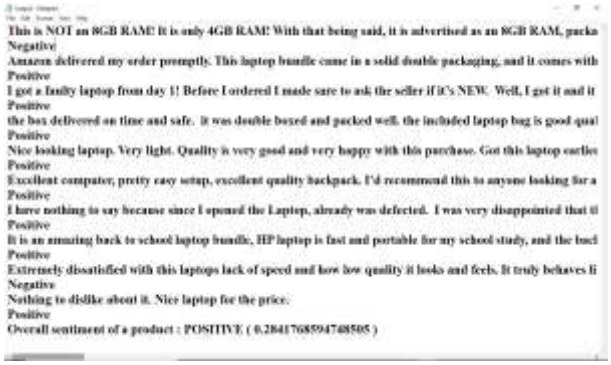

Fig.5: Individual sentiments

6. CONCLUSION

Sentiment analysis deals with the classification of texts based on the sentiments they contain. This article focuses on a typical sentiment analysis model consisting of three core steps, namely data scraping, review analysis and sentiment classification, and describes representative techniques involved in those steps. Sentiment analysis is an emerging research area in text mining and computational linguistics, and has attracted considerable research attention in the past few years. Future research shall explore sophisticated methods for opinion and product feature extraction, as well as new classification models that can address the ordered labels property in rating inference. Applications that utilize results from sentiment analysis is also expected to emerge in the near future. Sentiment analysis handles the evaluation of opinions of various kinds, broadly categorized into positive or negative opinions. Reviews show that various attributes as well as classification protocols fuse effectively for overcoming individual shortcoming and benefit from one another's advantages. In the end, sentiment classification performance is enhanced. Further work is required for improving performance metrics. The primary obstacle is utilizing other languages, handling negations as well as providing opinion summaries on the basis of product attributes, handling implicit product attributes.

7. REFERENCES

1. Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." Foundations and trends in information retrieval 2.1-2 (2008): 1-135.

2. Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. "Recognizing contextual polarity in

phrase-level sentiment analysis." Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics, 2005.

3. Whitelaw, Casey, Navendu Garg, and Shlomo Argamon. "Using appraisal groups for sentiment analysis." Proceedings of the 14th ACM international conference on Information and knowledge management. ACM, 2005.

4. Liu, Bing. "Sentiment analysis and subjectivity." Handbook of natural language processing 2 (2010): 627-666.

5. Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. "Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis." Computational linguistics 35.3 (2009): 399-433. 6. Liu, Bing. "Sentiment analysis and opinion mining." Synthesis Lectures on Human Language Technologies 5.1 (2012): 1-167.

7. Liu, Bing. "Sentiment analysis: A multi-faceted problem." IEEE Intelligent Systems 25.3 (2010): 76-80.

8. Liu, Bing, and Lei Zhang. "A survey of opinion mining and sentiment analysis." Mining Text Data. Springer US, 2012. 415-463.

9. Wilson, Theresa, et al. "OpinionFinder: A system for subjectivity analysis." Proceedings of HLT/EMNLP on Interactive Demonstrations. Association for Computational Linguistics, 2005. K Paramesha .et.al. 250 10. Ding, Xiaowen, Bing Liu, and Philip S. Yu. "A holistic lexicon-based approach to opinion mining." Proceedings of the 2008 International Conference on Web Search and Data Mining. ACM, 2008.

11. Benamara, Farah, et al. "Towards Context-Based Subjectivity Analysis." IJCNLP. 2011.

12. Nasukawa, Tetsuya, and Jeonghee Yi. "Sentiment analysis: Capturing favorability using natural language processing." Proceedings of the 2nd international conference on Knowledge capture. ACM, 2003.

13. Godbole, Namrata, Manja Srinivasaiah, and Steven Skiena. "Large-Scale Sentiment Analysis for News and Blogs." ICWSM 7 (2007).