

SOPPY : A sentiment detection tool for personal online retailing

Nurliyana Jaafar Sidek¹, Mi-Hwa Song^{2*}

¹Data-Speaks (M) SDN BHD, Kuala Lumpur, Malaysia

²Division of Information and communication technology, Semyung University, Jechon, Korea
nurliyanajs@yahoo.com, *mhsong@semyung.ac.kr

Abstract

The best 'hub' to communicate with the citizen is using social media to marketing the business. However, there has several issued and the most common issue that face in critical is a capital issue. This issue is always highlight because most of automatic sentiment detection tool for Facebook or any other social media price is expensive and they lack of technical skills in order to control the tool. Therefore, in directly they have some obstacle to get faster product's feedback from customers. Thus, the personal online retailing need to struggle to stay in market because they need to compete with successful online company such as G-market. Sentiment analysis also known as opinion mining. Aim of this research is develop the tool that allow user to automatic detect the sentiment comment on social media account. RAD model methodology is chosen since its have several phases could produce more activities and output. Sopyy tool will be develop using Microsoft Visual. In order to generate an accurate sentiment detection, the functionality testing will be use to find the effectiveness of this Sopyy tool. This proposed automated Sopyy Tool would be able to provide a platform to measure the impact of the customer sentiment over the postings on their social media site. The results and findings from the impact measurement could then be use as a recommendation in the developing or reviewing to enhance the capability and the profit to their personal online retailing company.

Keywords: Sentiment Analysis, Natural Language Processing (NLP), Semantic Role Labeling (SRL), Social Media, Rapid Application Development (RAD) model

1. Introduction

Sentiment analysis is the analysis of the feeling (i.e. attitudes, emotions and opinions) behind the words using natural language processing tools. Sentiment analysis also known as opinion mining. This study focus at sentiment comment in social media for personal online retailing business. This tool allow user to track mentions for identified keywords in comments and to automatic classify the polarity and sentiment comment in the social media. This research will analyze the sentiment comment from social media, especially from Facebook. Many personal online retailing also use social media such as Facebook as the platform for online business. Therefore, this research will develop new solution to analyze the comment and user can easily manage the polarity of comment in Facebook by using this tool that call as Sopyy tool.

Manuscript Received: July. 9, 2017 / Revised: July. 21, 2017 / Accepted: August. 3, 2017

*Corresponding Author: mhsong@semyung.ac.kr

Tel: +82-43-645-1645, Fax: +82-43-649-1747

Division of Information and communication technology, Semyung University, Jechon, Korea

2. Related Works

This chapter presents the literature and related studies that related to the Sopyy tool project. The definition, concept and design of the existing tools are also discuss to know the differentiation between several exiting tools. This project will involve data analysis and Natural Language Processing (NLP) for text extraction and processing. Table 1 shows several type of NLP such as Semantic Role Labelling (SRL), Chunking (CHUNK), Part Of Speech (POS) Tagging and Name Entity Recognition (NER).

Table1. Table of Natural Language Processing (NLP)

| Type of NLP | Description |
|-------------------------------|---|
| Part Of Speech (POS) Tagging | Aim at each word with unique tag indicates it syntactic role such as plural, noun, verb, adverb, etc. |
| Chunking (CHUNK) | Aim at labelling segments of sentence with syntactic constituents such as noun or verb phrases (NP or VP). Each word is assign only one unique tag. |
| Name Entity Recognition (NER) | Classify sentences into categories such a "PERSON" or "LOCATION". Each word is assign a tag prefixed by an indicator of the beginning or the inside of an entity. |
| Semantic Role Labelling (SRL) | Aim at giving a semantic role to a syntactic constituent of a sentence. Use phrase structure and syntactic parsing. |

2.1 Literature Review

Nowadays, social media is use as their vehicle to do business and make the businessperson easier to move faster about their products or services into the market. Social media is a great "hub" to get closer with the citizens. However, there still have many problems happen through the online business. Table 2 shows several topic of past journals that quiet similar and related to this research.

Table 2. Table of related journal

| Title of research paper | Summary/ Description |
|--|---|
| Natural Language Processing (Almost) from Scratch [1] | This journal discuss about natural language processing or also known as NLP. This research state that there are various natural language processing (NLP) task including part-of-speech (POS) tagging, chunking (CHUNK), named entity recognition (NER), and semantic role labeling (SRL) [1] |
| Completed research about sentiment analysis of government social media towards an automated content analysis using semantic role labelling done by 2015 [2]. | Siti Salwa Hasbullah and Dr. Rita Zaharah Wan Chik state that "This research will investigate the many responses to these official social media accounts, focusing on the emotions expressed, whether it be people are making suggestions, complaining or just passing information" (p.1). Semantic Role labeling also known as SRL technique was involves in this project. |
| Opinion Mining and Sentiment Analysis done by Bo Pang and Lillian Lee, 2008 [3] | This journal discuss about how the opinion mining and sentiment analysis can give an impact in the decision making process. |

| | |
|---|---|
| This sentiment analysis on social media research have completed at 2012. There are several people that responsible in this project such as Federico Neri and other. | This journal discuss about how a sentiment analysis can help an organization and company to gather the unstructured data about their customers feedback, preferences and problem or complaints about their products or services through social media such as Facebook, Twitter, YouTube and Myspace . |
| Node's Prestige gives different weights of opinion and relates the opinions expressed by the individuals [4] | From the point of view of opinion mining, it allows the node's prestige to differentiate opinions of different individuals. |

2.2 Comparison between existing tools

There are several existing sentiment analysis tool. This entire tool only can extract the data content manually from social media into tool. This project purpose is to develop the automatic detection sentiment tool for personal online retailing account. Each tool have their advantage and disadvantage. The following paragraph will explain about the sentiment analysis tools.

2.2.1 TheySay technology

TheySay's technology is the real sentiment tool to determine what people are saying. This tool has simple and clear interface. TheySay can provide data visualization of deep insights into sentiment through a dashboard delivered through our browser or mobile device.

This tool also can integrated into existing CRM or Social Media monitoring platforms and workflows easily and quickly give us best-in-class text and sentiment analytics. This tool have extra features compare to other tool such as Trackur, Twirratr and Sopyy. However, because of this point, it become not practical for those who lack in applying software in their small business.

2.2.2 Trackur tool

Trackur is a sentiment analysis on social media tools that user- friendly. Tracktur are providing details searches, sentiment and influencer analytics to help users in planning and assessing campaigns or doing product research.

2.2.3 Twitrratr

Twirratr and Tweetfeel are some of the free sentiment analysis tools especially for social network, Twitter. Twitter post that consist of some input word on how the majority of Twitter users feel will be search and gather by these tools. Positive "tweets" will be determine by words like, "love", "kind", "great" etc. while negative "tweets" will show the vice versa. At the same time, neutral "tweets" does not belong in both positive and negative "tweets".

Table 3 shows the differences between the existing tools. For examples, TheySay [5], Trackur[6] and Twitrratr tool. Each of this tool have the different features such as usable for personal online retailing. The following paragraph will explain about others features of the sentiment analysis tool.

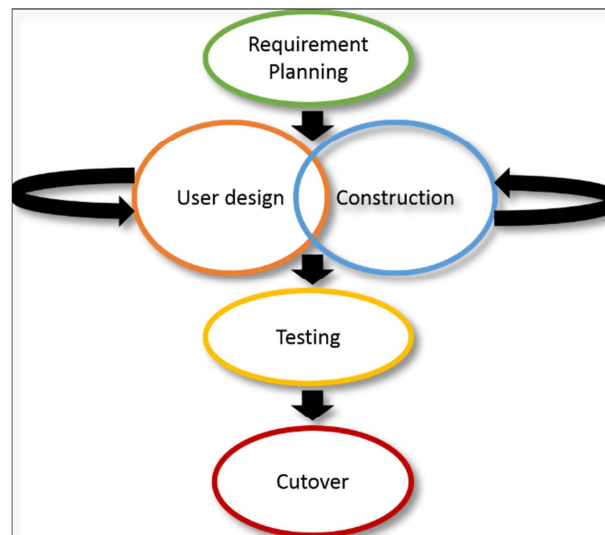
First, TheySay tool have user-friendly interface. Then, this tool can integrate the sentences and apply the rules based or heuristic. However, this tool fail to extract the text automatically and this tool not provide the data directory to store all the sentences. Then, Trackur is worst that TheySay tool because before use this tool, user need to sign up first. Next weakness of this tool are fail to extract the text, not provide own data directory and this tool do not have user-friendly interface. Next tool is Twirratr and this tool only available for Twitter user. This tool have no user-friendly interface and it do not have the data directory to store all the data.

Table 3. Table of comparison between existing tools

| Features | TheySay | Trackur | Twirratr | Soppy tool |
|--------------------------------------|---------------------|-----------------|-------------------|------------|
| Usable for personal online retailing | Yes (many features) | Need to sign up | Twitter user only | Yes |
| Integration | Yes | Yes | Yes | Yes |
| Text extraction | No | No | Yes | Yes |
| User friendly | Yes | No | No | Yes |
| Rule based / Heuristic | Yes | Yes | Yes | Yes |
| Data directory | No | No | No | Yes |

2.3 Rapid Application Development (RAD) Model

Researcher of Soppy tool project has decided to use Rapid Application Development (RAD) model. James Martin introduce RAD model in 1980s. RAD model is use to take care the workflow of this project is going well from starting until the end of the project. By choosing this model, chances to missing the due date, budgets, and low quality of project can be avoid. Figure 1 shows the RAD model flow. "Research methodology is systematic way to solve the problem [7]". RAD model propose five phase and activities performed in each phases are described in followed paragraphs.

**Figure 1. RAD model**

2.3.1 Phase 1: Requirement Planning

Proper planning is need before developing Soppy tool. Through the real data collected, certain techniques been used such as data analysis, journal related work and comparison between existing tool before proceed next phases.

2.3.2 Phase 2: User Design (UD)

In this phase, all information, input, user interface of the tool and output need to be design. User design

phase is important to ensure the design correctly and validate the design achieve the quality of requirement project.

2.3.3 Phase 3: Construction

This phase is to generate and prepare documentation tool to operate the propose application. However, the main goal of construction phase is to complete details design into creating and translating the Sopyy tool into program code.

2.3.4 Phase 4: Testing

This research choose “functionality test” technique to test the tool. During functionality testing, check the core application functions, text input, menu functions, installation, and setup on localized machines, and other testing process [8]. Then, Sopyy tool will test core application functions, text input (classify word by word), main functions such as browse the file from data directory and then process the data by using rule-based.

2.3.5 Phase 5: Cutover

Cutover phase is the researcher will do some changes to upgrade current Sopyy tool function. Compared with traditional methods, the entire process is compressed. As a result, the new tool completely developed, delivered and place in operation much sooner.

3. SOPPY : Prototype Development

The prototype development is about the process that involves in development of Sopyy tool. The process of designing the project, which includes the system architecture, use case diagram, activity diagram, sequence diagram and interface of Sopyy tool. System development is the process of creating or altering systems, along with the processes, practices, models, and methodologies used to develop them [9].

3.1 Programming Language and Software Used in System

The Sopyy tool has use Visual Basic as programming language. Several rule-based use to complete development the Sopyy tool. Sopyy tool will cooperate with Facebook.

3.2 System Architecture diagram

Figure 2 show the overall process that happen while using the Sopyy tool. There also show that connection between Admin, Facebook user and Sopyy tool.

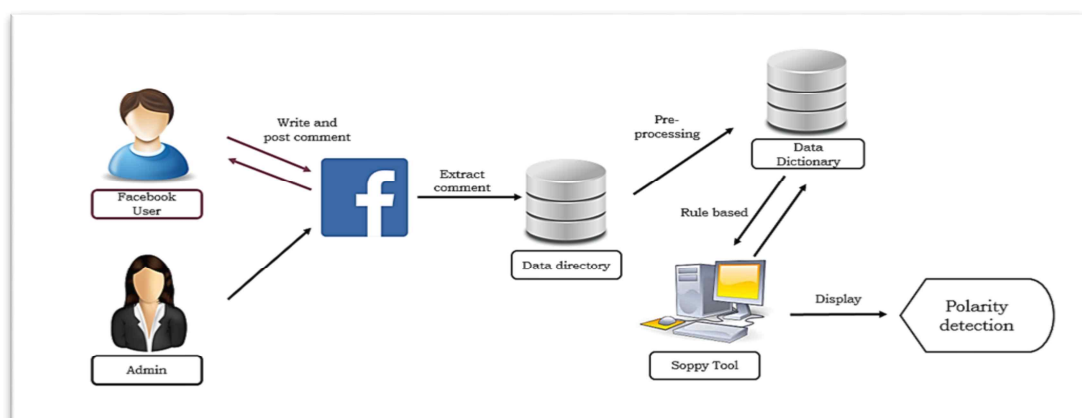


Figure 2. System Architecture of Sopyy tool

3.3 Extraction comment from Facebook to data directory

Admin will extract all comment from Facebook page and then locate all data into data directory. Admin can extract data and then store it date by date. All the extraction data is save in .txt file. This option make Admin easy to choose which data that need to be analyze by Sopyy tool.

3.4 Pre-processing from data directory to data dictionary

The Data Dictionary is actually the repository of the information about the data. Data dictionary is to save all the positive and negative words. Criteria of the polarity words is randomly choose by the researcher. Total of polarity words are 538 words; positive is 317 words and negative is 221 words. Data Dictionary defines each of the data elements and gives it a name for the easy access. Helps in the identification of the organizational data irregularity. Appendix 1 shows several polarity words that been used for this project. The sopyy tool will communicate with the data dictionary by using rule-based.

3.5 Rule-based

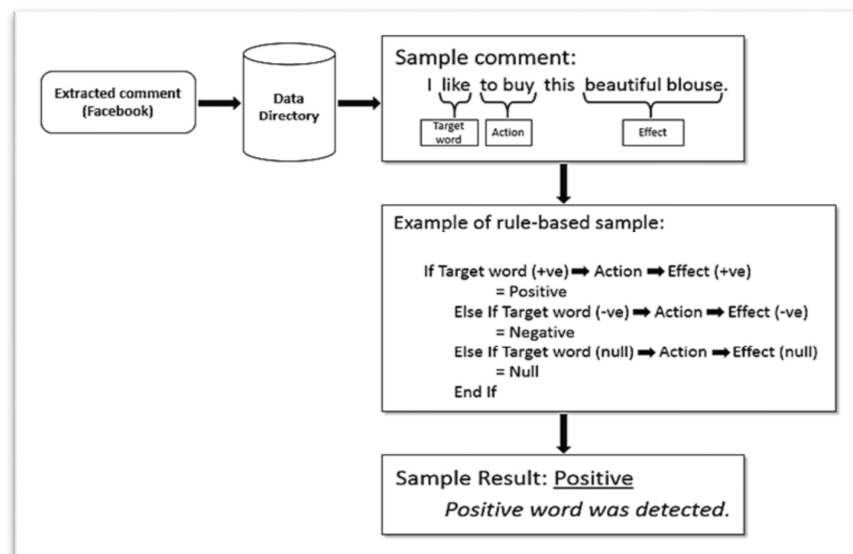


Figure 3. Example of Rule-based sample

This research used Microsoft Visual Basic 2008 Express Edition as programming language and apply some rule-based to develop the Sopyy tool. A rule-based language lets the researcher specify the rule declaratively, exact similar to the way that the rule are describe to the tool specified. A rule-based is the component that takes the declarative specification of the rules and decides at run time. Figure 3 shows the example of Rule-based sample for Sopyy tool.

3.6 System Design

This phase explain about the system design. System design is the process of defining the components, modules, interfaces, and data for a system to satisfy specified requirements [9]. The researcher need to validate the design in order to achieve the quality of requirement project. For example, Use Case diagram (Figure 4), Activity diagram (Figure 5) and Sequence diagram (Figure 6). This two diagram have been explain in following paragraphs.

Figure 4 show the use case diagram of Sopyy tool. Use case diagram is dynamic in nature; there should be some internal or external factors for making in interaction [10]. Use case diagram represents the user's

interaction with the system that shows the relationship between internal and external agent. These internal and external agents, also known as actors. The actors for Soppo tool are Facebook users and the researcher (admin of Soppo tool). The diagram used to model the system and capture a particular functionality of a system.

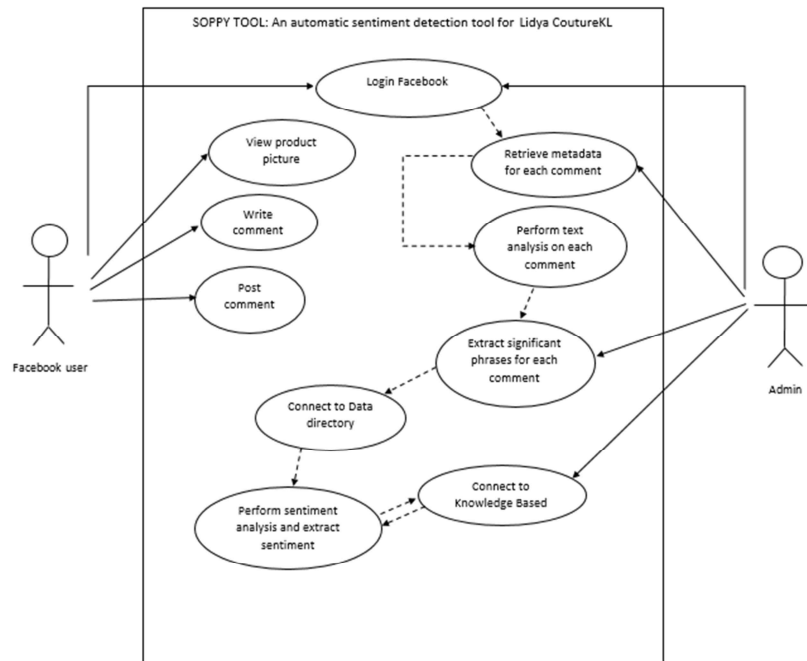


Figure 4. Use case diagram of Soppo tool

Activity diagram is fundamentally a flow chart to represent the flow from one activity to another activity. The activity can be describe as an operation of the system. Therefore, the control flow is describe from one operation to another. Figure 5 show the activity diagram of Soppo tool. There are some activity during the process in Soppo tool.

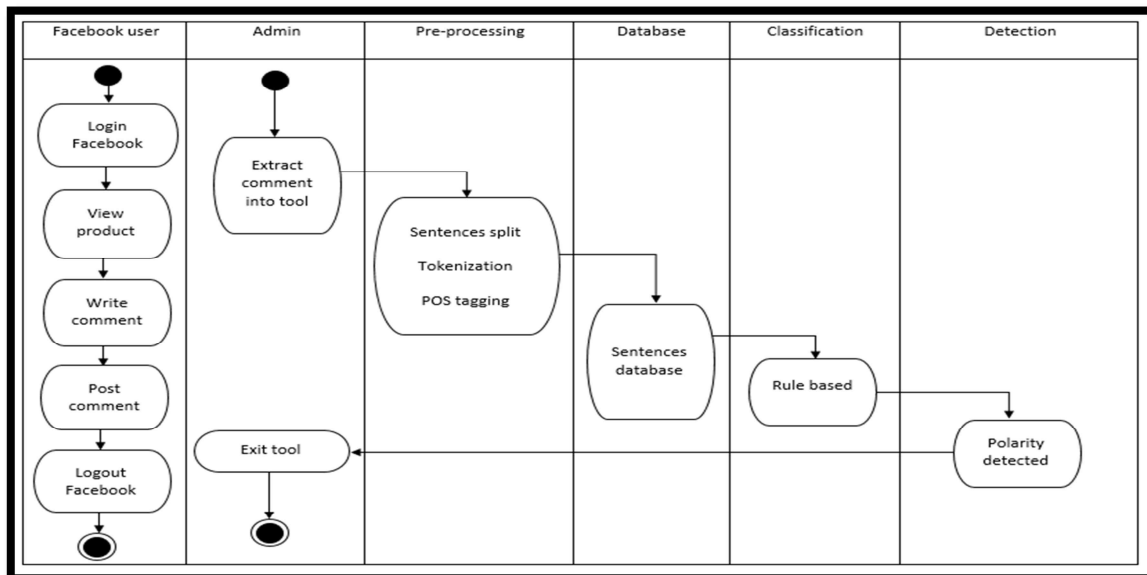


Figure 5. Activity diagram of Soppo tool

Figure 6 shows a Sequence diagram of Soppo tool. Sequence diagram are interaction diagrams that detail how operations are carry out [11] and shows how procedures operate with one another and in what order. Other than that, a sequence diagram also shows object interactions arranged in time sequence.

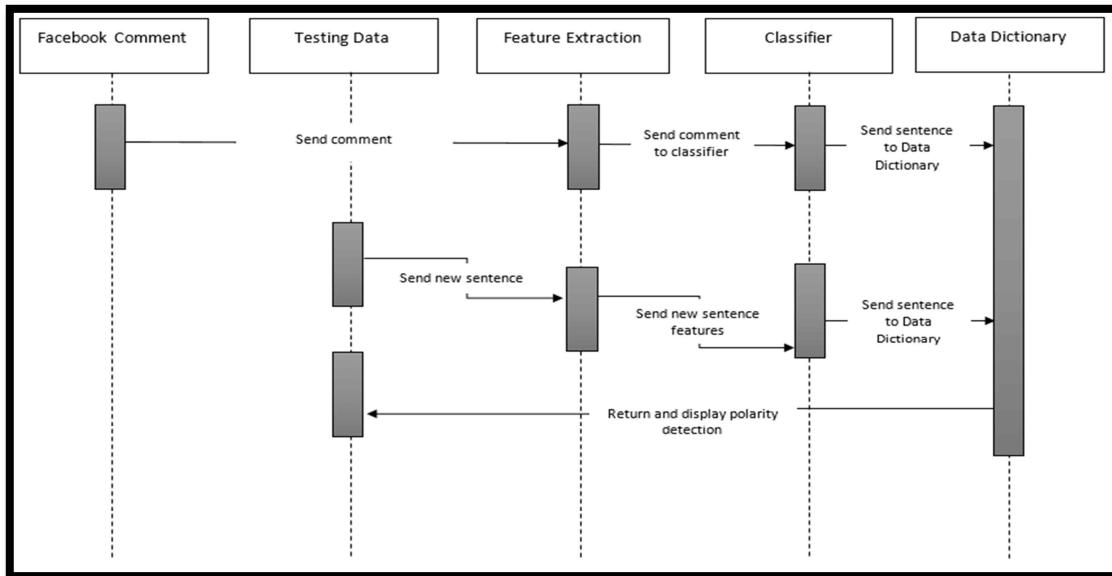


Figure 6. Sequence diagram of Soppo tool

4. System Implementation

Systems implementation is the process of defining how the information system should be built (i.e., physical system design), ensuring that the information system is operational and used, ensuring that the information system meets quality standard (i.e., quality assurance). The researcher will explain more detail about this topic in next paragraph.

Soppo tool choose the Visual basic as the programming language to develop the Soppo tool. Figure 7 show the main interface of Soppo tool. User can start using this tool after they click at the button as shown in Figure 7.

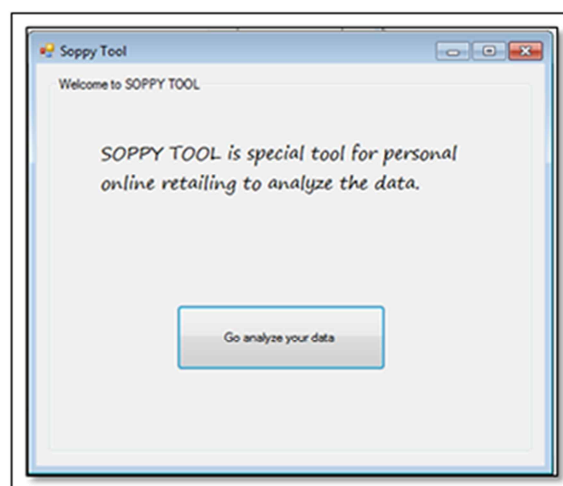


Figure 7. Main Soppo tool interface

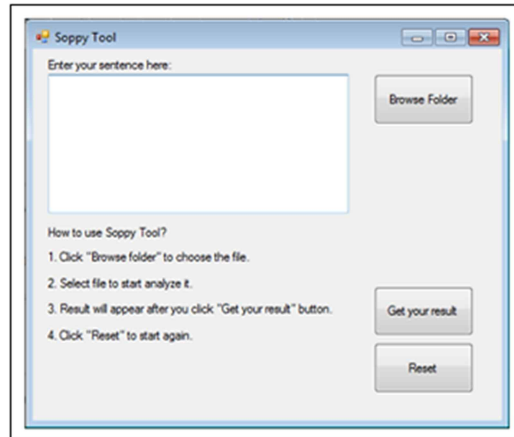


Figure 8. Process interface of Soppo tool

Then, Figure 8 show the process interface that the user can start analyze the data and get the result. Figure 8 show that user also can click the Reset button if they want analyze new data.

Next step is Admin just click the “Get your result” button to proceed last step. Then, rule based will start detection of sentiment. The result either POSITIVE or NEGATIVE will appear after the process was complete. Then, the result will appear in pop-up box as shown in Figure 9. Figure 9 shows the positive result because Soppo tool detect the positive word in that sentence. All the polarity words been stored in data dictionary or also known as knowledge based. User just need to click “Reset” button to use this tool again.

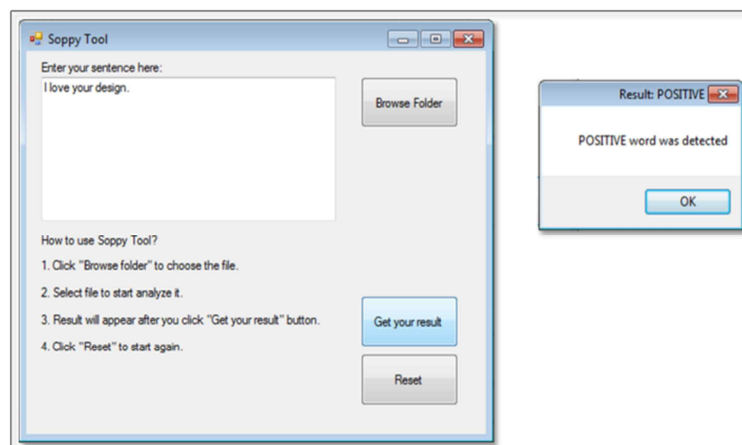


Figure 9. Result interface of Soppo tool

5. Testing and Results

Testing phase is important because to ensure the quality of the product. Customer gain their confidence in product depend on the quality of product. There are several type of testing. Such as Black Box testing, White Box testing, Unit testing, Functional testing, System testing, etc. [12]. Soppo tool research have chosen the Functional testing to test the functional of Soppo tool [13]. This type of testing focus on the output is as per requirement or not. Then, testing phase is required for an effective performance of the Soppo tool. Functionality testing involves several step such as identify that the tool expected to perform and determine the outcome based on the function’s specifications. Then, this testing execute the test case and lastly to compare the actual and expected outcomes.

5.1 Functionality testing Method

The Sopyy tool have chosen the Functional testing to test the functional of Sopyy tool. This type of testing focus on the output is as per requirement or not. Then, testing phase is required for an effective performance of the Sopyy tool. Table 4 shows the functionality testing.

Table 4. Table of Functionality Testing

| Case | Task | Test objective | Expected result | Result |
|------|--|---|--|--------|
| 1 | Main interface | To test the functioning of "Go analyze your data" button. | After clicking the button, interface of process in Sopyy tool will appear. | PASS |
| 2 | File selection | To test the functioning of "Browse Folder" button. | This button is to browse the data directory in other interface. | PASS |
| 3 | Extraction from data directory into Sopyy tool | To confirm the extraction data process is success. | Extraction data will appear at the provided space at process interface (refer Figure 7). | PASS |
| 4 | Detection word in Sopyy tool | To display the result of data analysis. | The result either POSITIVE or NEGATIVE will appear after the process was complete. | PASS |
| 5 | Connect to knowledge based | To detect word from Sopyy tool with the knowledge based. | Certain word detected by Sopyy tool. | PASS |

6. Conclusion

The Sopyy tool was successful achieved the objective of this project. The Sopyy tool features including data extraction process between data directory and the Sopyy tool, creating own knowledge based and detection word process. The future enhancement is to add several features such as import API to automatic communicate with Sopyy tool to be done in order to make this tool become more effective. In addition, the researcher plan to develop the Sopyy tool for other social media such as Instagram.

The Sopyy tool features including data extraction process between data directory and the Sopyy tool, creating own knowledge based and detection word process. This tool come with data directory and knowledge based. Therefore, it make admin of Sopyy tool feel free to add, edit or delete any data that have been store in the directory. Then, they also can classified data into folder any folder that they requested.

In addition, this research plan to develop the Sopyy tool for other social media such as Instagram. This is because while completing this project, there also have many personal online retailing that used Instagram as they platform to do marketing.

References

- [1] R. Collobert, JasonWeston, L. Bottou, M. Karlen, K. Kavukcuoglu and P. Kuksa, "Natural Language Processing (Almost) from Scratch," pp. 2-5, 2011.

-
- [2] S. S. Hasbullah and R. Z. Wan Chik, "SENTIMENT ANALYSIS OF GOVERNMENT SOCIAL MEDIA TOWARDS AN AUTOMATED CONTENT ANALYSIS USING SEMANTIC ROLE LABELING," pp. 1,4, 2015.
- [3] Bo Pang and L. Lee, "Opinion mining and sentiment analysis," vol. Vol. 2 , 2008.
- [4] S. Roy, S. Dhar, A. Paul, S. Bhattacharjee, A. Das and D. Choudhury, "CURRENT TRENDS OF OPINION MINING AND SENTIMENT ANALYSIS IN SOCIAL NETWORKS," IJRET: International Journal of Research in Engineering and Technology, p. 2, 2013.
- [5] TheySay, "Live Demo: TheySay," 2014. [Online]. Available: <http://apidemo.theysay.io/>.
- [6] Trackur , "Home: Trackur," 2007-2014. [Online]. Available: <http://www.trackur.com>.
- [7] S.Rajasekar; P.Philominathan; V.Chinnathambi, "RESEARCH METHODOLOGY," RESEARCH METHODOLOGY, 2013.
- [8] "Testing throughout the testing lifecycle: istqbexamcertification," 2015. [Online]. Available: <http://istqbexamcertification.com/what-is-functionality-testing-in-software/>.
- [9] www.mitre.org, "publication: System Engineering Guide: System Design and Development," The MITRE Corporation, 1997-2016. [Online]. Available:<https://www.mitre.org/publications/systems-engineering-guide/se-lifecycle-building-blocks/system-design-and-development>. [Accessed 2016].
- [10] M. Mohtashim, "UML Tutorial: UML- Use case Diagram," 2016. [Online]. Available: <http://www.tutorialspoint.com>.
- [11] F. Massimo, "Teaching/courses/seoc/2011-2012: THE UNIVERSITY of EDINBURGH," 2004-2011. [Online]. Available: http://www.inf.ed.ac.uk/teaching/courses/seoc/2011_2012/notes/SEOC08_notes.pdf.
- [12] Vijay, "Type of Software Testing: software testing help," 2015. [Online]. Available: <http://www.softwaretestinghelp.com/types-of-software-testing/>.
- [13] "Functionaly testing: Software Testing Fundamental," 2016. [Online]. Available: <http://softwaretestingfundamentals.com/test-case/>.