A Brief Report on Sentiment Analysis

Siddhi Gharat[1] and Divyesh Patil[2]
Assistant Professor Department of IT[1]
Student, P.G. Department of IT[2]
Veer Wajekar ASC College, Phunde, Uran

Abstract: *In this paper, we present our preliminary experiments on tweets sentiment analysis. This experiment is designed to extract sentiment based on subjects that exist in tweets. It detects the sentiment that refers to the specific subject using Natural Language Processing techniques. To classify sentiment, our experiment consists of three main steps, which are subjectivity classification, semantic association, and polarity classification. The experiment utilizes sentiment lexicons by defining the grammatical relationship between sentiment lexicons and subject. Experimental results show that the proposed system is working better than current text sentiment analysis tools, as the structure of tweets is not same as regular text.*

Keywords*:* Sentiment Analysis; Natural Language Processing; Tweets

I. INTRODUCTION

Twitter, as a popular microblogging service, allows users to post tweets, status message with length up to 140 characters [2]. These tweets usually carry personal views or emotions towards the subject mentioned in the tweets. Sentiment analysis is a technique that extracts the users opinion and sentiment from tweets. It is an easier way to retrieve user views and opinions, compared to questionnaire or surveys. There have been studies on automated extraction of sentiment from the text. For example, Pang and Lee [10] have used movie review domains to experiment machine learning techniques (Naïve Bayes, maximum entropy classification, and Support Vector Machine (SVM)) in classifying sentiment. They achieved up to 82.9% of accuracy using SVM and unigram model. However, as the performance of sentiment classification is based on the context of documents, the machine learning approaches have difficulties in determining the sentiment of text if sentiment lexicons with contrast sentiment are found in the text. Subsequently, Pang and Lee [11] introduced the technique of minimum cuts in graphs before sentiment classification using machine learning method so that only subjective portion of the documents is used for text categorization [11]. They first classified the text as containing sentiment, and then classified the sentiment into positive or negative. It achieved better result compared to previous experiment, with an improved accuracy of 86.4%. Besides using machine learning techniques, Natural Language Processing (NLP) techniques have been introduced. NLP defines the sentiment expression of specific subject, and classify the polarity of the sentiment lexicons. NLP can identify the text fragment with subject and sentiment lexicons to carry out sentiment classification, instead of classifying the sentiment of whole text based on the specific subject [9]. Feature extraction algorithm is one of the NLP techniques. It can be used to extract subject-specific features, extract sentiment of each sentiment-bearing lexicons, and associate the extracted sentiment with specific subject. It achieved better result than machine learning algorithm, with accuracy up to 87% for online review article, and 91~93% of accuracy for reviewing general web page and news article [14]. This approach focused on general text, and it removed some difficult cases to obtain better result, such as sentences that were ambiguous, or sentences that have no sentiment. Previous machine learning and NLP studies in sentiment analysis for text may not be suitable for sentiment analysis for tweets, as the structure between tweets and text is different. Three main differences between sentiment analysis in tweets and previous research in text are:- 1. Length. The tweets are limited to 140 characters. In the experiment, Go et al. [4] found that average length of tweets is 14 words, and average length of sentence is 78 characters. Sentiment analysis in tweets and text are different, in the aspect of text sentiment analysis focused on review article that contained multiple sentences, where tweets are shorter from length. 2. Available data. The magnitude of data is different between tweets and normal text. In sentiment classification using machine

ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

learning techniques by Pang and Lee [10], the size of corpus for training and testing was 2053. However, Go et al. [5] managed to collect up to 15,000 tweets for Twitter sentiment analysis research. But now with the usage of Twitter API, we can collect thousands and millions tweets for training purpose. 3. Sentence structure. Tweets contain acronyms, abbreviations and elongated words, which result to a messy text. Besides the text features, other features such as URL, image, hashtags, punctuations and emoticons are included as well. Most of these features affect the accuracy of analysis process as they are not proper text that can be found in dictionary, or read and understand by machine. Some method should be sorted out, as machine cannot understand the informal languages. On the research by Blenn et al [1], a system that worked through a combination of grammatical analysis with traditional word frequency analysis was proposed. Grammatical analysis studied the structure of text, and associated the sentiment lexicons with subject by identifying the relationship between sentiment lexicons and subject. It was a significant improvement in sentiment analysis for short colloquial text as previous approach did not achieve high detection accuracy as it did. It did not need any supervised training, but managed to improve the accuracy of previous work by 40%. In this research, a system is proposed to carry out sentiment analysis on tweets based on specific topic. Several pre-processes steps have been carried out to clean the noise in tweets and present tweets in formal language. To determine the sentiment pf tweets, NLP is implemented to find out the subjective portion of tweets that associates to subject, and classify the sentiment of tweets. The tweets will be label as positive, negative or neutral. The remainder of the paper is structured as follows: Section II describes the framework of the proposed system and the dataset that is used for experiment. Section III discusses the experimental results when comparing the performance of the proposed system with current tools, and Section IV summarizes the findings of this paper.

II. OVERVIEW OF FRAMEWORK

In this section, the outline of the proposed system is described. Tweets were extracted from a Twitter database for experiment. All tweets were manually labelled as positive, negative and neutral. This set of tweets was used to evaluate the performance of the proposed system, using metrics such as the accuracy and precision of the predictive result. To present the tweets in structured manner, some preprocessing have been done on the dataset before being further analyzed by the proposed system. Pre-processing ensures that the tweets will be prepared in formal language format that can be read and understood by machine. After pre-processing, sentiment of tweets can be determined through sentiment classification. There are three steps in sentiment classification: subjectivity classification, semantic association and polarity classification. Subjectivity classification was carried out to judge whether the tweets are subjective or objective, and subjective tweets went forward for semantic association to find out the sentiment lexicons that associates to subject. The sentiment classification predicted the tweets as positive, negative or neutral by classifying the sentiment of sentiment lexicons.

A. Dataset

A total of 1513 tweets were extracted from Twitter and manually labelled. These tweets contain keyword ‘Unifi’, which is a telecommunication service in Malaysia. It is also the subject in sentiment classification. There are 345 positive tweets, 641 negative tweets and 531 neutral tweets. Tweets were analyzed by the proposed system to obtain the predictive sentiment. For benchmarking, 1513 tweets were analyzed using Alchemy API1 and Weka2 . Alchemy API applies NLP techniques in sentiment analysis, while Weka is a tool which performs data mining using machine learning algorithms. The selected machine learning algorithms were Naïve Bayes, Decision Tree (J48) and Support Vector Machines. Raw tweets and pre-processed tweets were imported to Weka to check the importance and effect of preprocessing. In Weka, feature extraction algorithm is applied. 100 words with highest frequency were selected, and 10- folds cross-validation was used to train and test the data. Predictive results from Alchemy API, Weka and the proposed system were tabulated and compared to the manually labelled tweets to calculate the accuracy, precision, recall and F-measure

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
IJARSCT International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

B. Proposed System

The figure below shows the steps of the proposed system, where it begins from pre-processing to sentiment classification. The detailed steps of pre-processing are described in Section 1, and Section 2 explains the flow of sentiment classification.

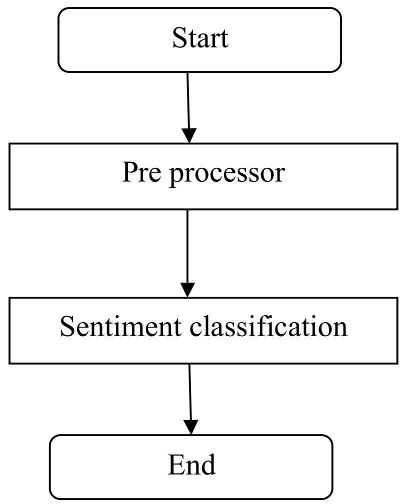

Figure 1: Flow Chart of Pre-process and Sentiment Classification.

1) Pre-process

Pre-processing aims to process and present the tweets in an organized format and increase the machine understanding on the text, as most tweets are in the form of unstructured text. It includes URLs and hashtags removal, special symbol replacement, repeated characters removal, abbreviations and acronyms expansion, and subject capitalization. URLs and image linkage are removed from the text as they do not carry meanings. Hashtags are removed from the text to avoid confusion, as the hash-tagged item may not refer to the subject. Special symbol are replaced as words to avoid confusion during text processing, for example, ‘>’ replaced by ‘greater’, ‘&’ replaced by ‘and’. For emoticons, according to the research on Automatic Sentiment Analysis of Twitter message, word-based analysis performed better than emoticon-based analysis [7]. Thus, emoticons are removed from tweets. Unstructured tweets are normalized by removing the repeated characters, or reducing the elongated words to normal form, for example, ‘goood’ to ‘good’. The contraction, acronyms and abbreviations are expanded as well. For example, “I’m not going to work 2mr” expands to “I am not going to work tomorrow”. Subject capitalization is done to prepare the subject in uniform pattern, and the machine can detect the subject in an easier way. The processed tweets will proceed to sentiment classification, in order to predict the sentiment of tweets.

2) Sentiment Classification

The sentiment classification process is shown in Fig. 2.

a) Subjectivity Classification- Subjective classification differentiates the tweets into subjective or objective. The system scans the tweets word by word, and finds out the word that contains sentiment. If the word in the tweet carries positive or negative sentimentweightage, the tweet will be classified as subjective. Else, it will be objective, is which also neutral. For example, “Come and get internet package” or “Come and get new internet package” The first tweet does not have any word that carries sentiment score. It will be classified as Objective and tagged as Neutral. While in second tweet, “new” is a word that has sentiment score. The tweet will be classified as subjective, and proceed to next step for semantic association.

b) Semantic Association- In semantic association, the sentiment lexicons that associate to subject are being defined through grammatical relationships between subject and sentiment lexicons. As tweets are mostly short and straight

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, February 2025
Impact Factor: 7.67

forward, the grammar structure is simpler than normal text. Sentiment lexicons that mostly associate with subject will be adjectives or verbs. We can find direct opinion and comparisons in tweets [12]. In direct opinion, sentiment lexicons are describing one subject or more, with the help of preposition and conjunction. While in comparison opinion, there are at least two subjects, but the subjects are associated to the same sentiment lexicons without the existence of conjunction. An example of direct opinion "I love Unifi" is shown at Alg. 1. In the example shown, 'I' is the nominal subject of 'love', and 'Unifi' is the direct object of 'love'. It is a straight forward statement, and most tweets appear in this structure. As mentioned above, most of grammatical relationships show that verbs and adjectives are associated with subject. In this example, since the Subject 'Unifi' is direct object of 'love', we have to check the attribute of 'love'. From the POS tag result, we can see that 'I' is a personal pronoun, 'love' is a verb, and 'Unifi' is a noun. As 'love' is a verb which associates to the Subject, 'love' will be sent to polarity classification to determine whether it carries positive or negative sentiment.

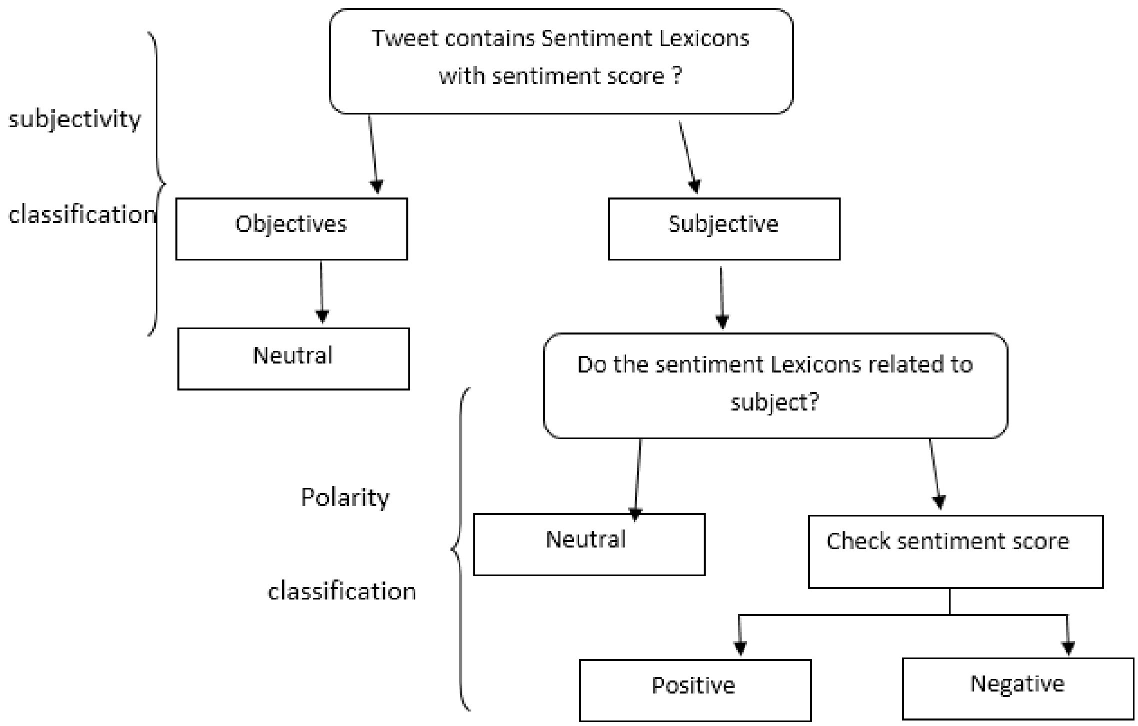

Figure 2 Sentiment Classification Process.

There are some grammatical patterns to check the sentiment lexicons that are associated to subject:- 1. Adjective that describes the subject (good Unifi finally upgrade the service) 2. Adjective or verb that complements the subject (happy when Unifi is recovered) 3. Adverb that describes the subject (Unifi speed is fine) 4. Adjective that has a preposition between adjective and subject (fast like Unifi) 5. Adjective that present in superlative mode and has a preposition between adjective and subject (let us face it Unifi is not the best but it is better than M) 6. Adjectives, verb or noun that complements the subject (50% of my drafts are about Unifi to be honest) 7. Adjectives or verb that associates to the subject (Unifi forever no lag) 8. Adjective or verb will reverse the sentiment if negate word exists (I do not want to uninstall Unifi) Meanwhile, comparison opinion has different sentence structure and grammatical relationship. An example of comparison opinion "Unifi is better than M" is shown in Alg. 2.

Sentence I love Unfi
Pos Tagging
I/PRP
love/VBP
Unfi/NNP
Parse
(ROOT (S (NP (PRP I)) (VP (VBP love) (NP (NNP Unfi)))))
Typed dependencies
nsubj(love-2, I-1)
root(ROOT-0, love-2)
dobj(love-2, Unfi-3)

Algorithm 1 Dependencies type and POS tag of Direct Opinion.

Sentence Unifi is better than M
POS Tagging
Unifi/NNP
is/VBZ
better/JJR
than/IN M/NNP
Parse (ROOT (S (NP (NNP Unifi)) (VP (VBZ is) (ADJP (ADJP (JJR better)) (PP (IN than) (NP (NNP M)))))))
Typed dependencies, **collapsed**
nsubj(better-3, Unifi-1) cop(better-3, is-2) root(ROOT-0, better-3) prep_than(better-3, M-5)

Algorithm 2 Dependencies type and POS tag of Comparison Opinion.

In this example, ‘Unifi’ is the nominal Subject of comparative adjective ‘better’, and there is a preposition between ‘better’ and another noun ‘M’. We can find a grammatical pattern of adjective is complementing subject (‘Unifi is better’), and superlative adjective that has a preposition between it and subject (‘better than M’). c) Polarity Classification In polarity classification, subjective tweets are classified as positive or negative. The sentiment of tweets is classified based on the sentiment lexicons that associate to subject. For example, ‘I love Unifi’, the verb ‘love’ is the sentiment lexicons. While checking with SentiWordNet, ‘love’ has a positive score of 0.625. Hence, we can conclude that ‘Unifi’ has Positive sentiment, thus classify the tweet as Positive. For comparative opinion, the position of the subject is very important [13]. For instance, in the tweet ‘Unifi is better than M’, adjective ‘better’ is found, but there are 2 Subjects – ‘Unifi’ and ‘M’. The subject that exists after the comparative adjective carries a contrast sentiment with the Subject that appears before. In this case, as ‘better’ carries a Positive score of 0.825, ‘Unifi’ will be classified as positive, and ‘M’ will be classified as negative

III. RESULTS AND DISCUSSIONS

The results are tabulated in confusion matrix. It records the predicted result and the actual result. Confusion matrix is of size $\ell \times \ell$, where ℓ is the number of different label values [6]. In this research, there are three label: positive, negative and neutral. While comparing the predictive results to labelled results, we can get the score for precision, recall and F-measure. Table 1 summarizes the performance of proposed system and Alchemy API. Compares to Alchemy API, the proposed system scores higher in overall, where it hits 59.85% in accuracy, 53.65% in precision and 0.48 in F-measure. For Alchemy API, it only scored 58.87% accuracy, 40.82% precision and 0.43 in F-measure. Alchemy API may not perform well as it is designed for normal text sentiment analysis, and the structure of tweets is different from text. Performances of machine learning algorithms are summarized in Table 2. Among three selected machine learning algorithm, SVM scores 58.67% in accuracy, 60.44% in precision, 0.48 in F-measure while using raw tweets as training and test data, and scores 64.95% accuracy, 66.54% precision, 0.57 F-measure when pre-processed tweets were imported for training and test data. Its results outperform Naïve Bayes and Decision Tree classifier, regardless of the tweets are pre-processed or not. The results from Weka prove that NLP-based preprocessing does help in improving the performance of classifier, as all classifiers that are trained and tested with pre-processed tweets score higher compared to classifiers that use raw tweets as corpus. Generally, the scores were increased by an average of 2.09% in accuracy,

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

IJARSCT International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

5.23% in precision and 0.10 in F-measure. Table 3 compares the performance of the proposed system, Alchemy API and SVM. For overall assessment, our proposed system outperforms Alchemy API, but not as well as SVM. SVM performs better if tweets were pre-processed before training and testing. At such, the proposed system still needs to be improved in order to achieve a better performance.

	Proposed System Alchemy API	Proposed System Alchemy API
Accuracy	59.85%	58.87%
Precision	53.65%	40.82%
F-measure	0.48	0.43

Table 1 Result from Proposed system and Alcemy API

	Naïve Bayes		Decision Tree		Support Vector Machine (SVM)	
	Raw Tweets	Preprocessed Tweets	Raw Tweets	Preprocessed Tweets	Raw Tweets	Preprocessed Tweets
Accuracy	55.04%	60.58%	49.70%	57.60%	58.67%	64.95%
Precision	47.47%	53.59%	44.05%	53.65%	60.44%	66.54%
F-measure	0.44	0.55	0.35	0.43	0.48	0.57

Table 2 Result from Weka using Machine Learning Algorithm

	Proposed System	Alchemy API	SVM (Pre-processed tweets)
Accuracy	59.85%	58.87%	64.95%
Precision	53.65%	40.82%	66.54%
F-measure	0.48	0.43	0.57

Table 3 Summarized Result for 3 different Sentiment Analysis tools.

IV. CONCLUSION AND FUTURE WORK

There is lots of research on analyzing sentiment from tweets, as Twitter is a very popular social media platform. . In this paper, we present the preliminary results of our proposed system that corporates NLP technique to extract subject from tweets, and classify the polarity of tweets by analyzing sentiment lexicons that are associated to subject.

From the experiments, the proposed system performs better compared to Alchemy API, but still need to be improved as SVM is doing better. For future works, the focus will be on how to enhance the accuracy of sentiment analysis. It is also noticeable that due to the highly appeared misspelled words and slangs in tweets, it is not easy to extract the sentiment lexicons if it is not pre-processed to formal language. In pre-processing, the unstructured tweets should convert to formal sentence, which is still not that efficient, as more training data is needed.

REFERENCES

[1] N. Blenn, K. Charalampidou, and C. Doerr, “Context-Sensitive sentiment classification of short colloquial text”, in NETWORKING 2012, pp. 97-108, Springer Berlin Heidelberg, 2012.

[2] S. W. Davenport, S. M. Bergman, J. Z. Bergman, J. Z., and M. E. Fearrington, “Twitter versus Facebook: Exploring the role of narcissism in the motives and usage of different social media platforms.”, Computers in Human Behavior 32, pp. 212-220, 2014.

[3] A. Esuli, and F. Sebastiani, “Sentiwordnet: A publicly available lexical resource for opinion mining”, in Proceedings of LREC, vol. 6, pp. 417-422, May 2006.

[4] A. Go, R. Bhayani, and L. Huang, L., “Twitter sentiment classification using distant supervision”, CS224N Project Report, Stanford, 1-12, 2009.

[5] A. Go, L. Huang, and R. Bhayani, R, “Twitter sentiment analysis”. Entropy 17, 2009.

[6] R. Kohavi, and F. Provost, “Glossary of terms” Machine Learning 30(2-3), pp 271-274, 1998.

[7] A. C. Lima, and L. N. de Castro, "Automatic sentiment analysis of Twitter messages", in Computational Aspects of Social Networks (CASoN), 2012 Fourth International Conference, pp. 52-57, IEEE, 2012.

[8] M. C. De Marneffe, B. MacCartney, and C. D. Manning, "Generating typed dependency parses from structure parses", in Proceedings of LREC, vol. 6, pp. 449-454, 2006.

[9] T. Nasukawa, and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing", in Proceedings of the 2nd international conference on Knowledge capture, pp. 70-77, ACM, October 2003

[10] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up?: sentiment classification using Machine Learning techniques", in Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol 10, pp. 79-86, Association for Computational Linguistics, July 2002.

[11] B. Pang, and L. Lee, "A sentimental education: Sentiment analysis using Subjectivity summarization based on minimum cuts", in Proceedings of the 42nd annual meeting on Association for Computational Linguistics, pp. 271, Association for Computational Linguistics, July 2004.

[12] A. Westerski, "Sentiment Analysis: Introduction and the State of the Art overview", Universidad Politecnica de Madrid, Spain, pp 211- 218, 2007.

[13] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis", Computational linguistics 35, no 3, pp 399-433, 2009 [14] J. Yi, T. Nasukawa, R. Bunescu, and W. Niblack, "Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques", in Data Mining 2003, ICDM 2003. Third IEEE International Conference, pp. 427-434, IEEE, 2003.