Sentiment Analysis using Vader on Customer Review

Khaw Siang Lee
School of Computer Sciences
University of Science Malaysia
Penang, Malaysia
khawslee@student.usm.my

Thangeswariy
School of Computer Sciences
University of Science Malaysia
Penang, Malaysia
thangeswariy@student.usm.my

***Abstract*—Sentiment Analysis (SA) and opinion mining is an application of Natural Language Processing (NLP) and it has been booming in the business and social domain. Opinions and reviews are a central part of all human activities and it's the key influencers of our behaviors. This paper focuses on the opinion mining of customers by classifying the polarity of reviews in terms of positive (good), negative (bad), and neutral. It is important to capture the public opinion on products and services to strengthen customer support. Manual labeling of each review by humans is time-consuming and error prompting. Therefore, the VADER sentiment is used to label the reviews to provide fast insight into product option trends.**

Keywords—Sentiment Analysis, Opinion Mining, Reviews, VADER, SpaCY, Machine Learning

I. INTRODUCTION

Sentiment Analysis is one of the many areas of computational study that identifies and categorized opinions in text for emotion, mood recognition, ranking, and relevance to study the attitude of use within a particular domain [1]. Sentiment analysis is a combination of human intelligence and electronic intelligence in mining text and classifying user sentiments [2]. The rapid growth of user-generated content such as product reviews, customer satisfaction, and preferences turns to be an important source for business intelligence and marketing [3]. These customer reviews are vital to product manufacturers, service providers, and end-users to understand public opinions and make a concrete decision [3]. However, as customer reviews grow exponentially, it becomes a great challenge to obtain a comprehensive opinion through manual analysis. Therefore, efficient analysis and summarization of customer reviews enable an organization to identify the positive and negative opinions about a specific brand, product, or service [3]. Consequently, automatic analysis of customer reviews is widely preferred, and this is where text mining and computational intelligence which is known as sentiment analysis plays a major role [3]. There are three primary challenges within sentiment analysis which are aspect detection, opinion word detection, and sentiment orientation identification [3]. The opinion word detection is the opinion mining to produce a feature-based summary that is domain-specific such as topic modeling by grouping the same topic in one group [4].

In this paper, we will focus on sentiment orientation identification based on a customer's review of a food and beverage brand whereby the task is to classify the content as positive, negative, or neutral. It aims to automatically detect the subjective information contained in the text, identify the sentiment polarity and estimate the strength of sentiment polarity. The important aspect of this study is to identify the sentiment polarity of individual words which are known as words semantic orientation which evaluates the word and varies in the direction of positive or negative [5]. A positive semantic orientation indicates positive evaluation such as praising while a negative semantic orientation indicates negative evaluation such as complaints [5]. We are proposing a domain-dependent rule-based method for semantically classifying sentiment based on semantic orientation to calculate text sentiment. The rest of the paper is organized as follows: Section II Related Work, Section III Research Methodology, Section IV Results, and Discussion and Section V Conclusion.

II. BACKGROUND

This section briefly introduces some work related to semantic orientation identification using various techniques which constitutes an important basis for our work. Sentiment Analysis uses Natural Language Processing, Computational Linguistics, and text analysis for data mining [6]. There are two main approaches for sentiment extraction from review and classifying them as positive or negative which are the Lexicon Based Approach and Machine Learning [6].

The Lexicon technique uses the sentiment lexicon with information on words that are positive and negative respectively [6]. A sentiment lexicon has a list of lexical features which are labeled according to semantic orientation and then identify the polarity score of review based on the positive or negative indicators determined by the lexicon [6]. Lexicon is based on polarity determination which is mainly from two categories which are dictionary-based and corpus-based approaches [7]. The corpus-based approach is the probability of sentiment words together with sets of negative and positive words by searching for large amounts of texts [7]. On the other hand, the dictionary-based approach has been applied to hotel reviews, product reviews, and Twitter data [7]. There is various lexicon such as LIWC (Linguistic Inquiry and Word Count) and GI (General Inquirer) Maintaining the Integrity of the Specifications, however, the ANEW (Affective Norms for English Words), SentiWordNet, and SenticNet are associated with valence scores for sentiment intensity [6]. This Lexicon-based approach is very advantageous in terms of its being domain-independent, can be easily extended and improved.

Apart from this, another approach for sentiment analysis is machine learning by contrasting algorithms using feature selection and learning from labeled training data [6]. Examples of classifiers are Naïve Bayes, Random Forest, and Support Vector Machine [6]. The sentiment classifier is trained on the labeled data from one domain, which is usually non-transferable to another domain [7]. To overcome this challenge, Lexicon-based approaches are recommended [6].

III. RELATED WORK

Various previous works have been done concerning sentiment analysis. With the current rapid growth in user-generated content, automatic sentiment analysis on customer reviews is essential [3]. In the paper by Ayoub, two steps have been applied to understand the semantic orientation of customer reviews on products [3]. The first is using a generalized set of heuristic methods for detecting the words with opinion, thereafter a bootstrapping unsupervised algorithm is applied to remove the incorrect opinions [3]. Then, a lexicon approach is used to determine the strength of polarity of the opinion word [3].

In another work by Aurangzeb, lexical-based semantic orientation for sentence-level semantic classification has been done by classifying the subjective and objective sentences using the SentiWordNet dictionary [8]. Then, with the usage of a lexical dictionary, the polarity of the subjective sentence is checked as positive, negative, or neutral [8]. SentiWordNet is a semi-supervised method to obtain polarity by applying scoring rules [8]. This proposed methodology obtained an accuracy of 87% [8].

Apart from that, the semantic orientation of the subjective terms for sentiment analysis can be made using a classification approach through a model known as S-HAL (Sentiment Hyperspace Analogue to Language) [5]. S-HAL generates a set of weighted features based on surrounding words and categorizes the semantic orientation information of words by feature space [5]. This approach produces fast and accurate semantic orientation identification without the usage of Internet search engines and produces an accuracy of 86% [5].

In addition to that, several works have been implemented using Valence Aware Dictionary and Sentiment Reasoner (VADER) which is also a lexicon [6]. It is a combination of sentiment lexicon, a list of lexical features which labels according to semantic orientation as either positive or negative [6]. In the paper by Chaitra, the comments on mobile unboxing videos have been analyzed using a hybrid combination of Naïve Bayes and VADER whereby the classifier produced an accuracy of 79.7% of accuracy in predicting the sentiment comment [9]. Furthermore, Gilbert applied VADER for the social domain and the effectiveness was compared to ANEW, LIWC, and support vector machines [10]. This resulted in VADER out-performing all the other models with an accuracy of 96% and said to be sensitive to sentiment as well as considered to be favorable to all other domains [10]. Shihab applied VADER sentiment on Twitter sentiment analysis to automatically classify semantic polarity by binary classification and the results were good in detection ternary when compared to Natural Language Toolkit [11]. On top of all these, sentiment analysis on movie reviews was made using VADER to overcome the challenges in identifying the polarity of each review and summarizing the opinions in them [6]. Hence, automated sentiment analysis to find the polarity score and classify the reviews as positive or negative was made using NLTK, Text blob, and VADER to compare the output result. VADER generated the highest precision, recall, F1-score, and accuracy in comparison to Textblob and NTLK [6].

In short, VADER is both a qualitative and quantitative method that empirically validates the fold standard sentiment lexicon [10]. VADER has several advantages on sentiment analysis work such as it does not require any training data, supports emoticons, works in an adverse domain, and has a relatively good speed performance [6]. VADER is able to identify the polarity indices using polarity_scores() function [6]. It will return the metric values of the negative, neutral, positive, and compound of a given sentence [6]. The compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1 and +1 where -1 indicates the most extreme negative and +1 indicates the most extreme positive [6].

IV. RESEARCH METHODOLOGY

In this section of the paper, the detailed description and methodology of the dataset, data cleaning, and sentiment analysis technique applied will be elaborated.

A. *Dataset Description*

The dataset set used in the project is obtained from an open-source known as Kaggle. This is a Ben & Jerry ice cream outlet's customer review dataset. The dataset consists of 8 attributes and 7943 instances.

B. *Data Cleaning*

Data cleaning is an essential step in the data mining process as the accuracy and integrity of result and model deployment is related to the quality of the data. In addition to that, data cleaning is crucial in eliminating invalid or error data to remove any noise or data cardinality. In the dataset of this project, there were invalid columns found in the title column and they were replaced with blanks. Next, the title and text column were combined into a new column called "com_title". Thereafter, the star ratings were converted into sentiments by defining 4 stars and above as positive, 2 stars and below as negative, and 3 stars as neutral. The positive is indicated by "pos", negative is indicated by "neg" and neutral is indicated by "neu".

C. *Sentiment Analysis using Lexicon (VADER)*

Upon completing the data cleaning steps, VADER was applied to analyze the sentences and convert the obtained compound score to sentiments. This is done by setting the polarity as less than -0.05 as negative, more than 0.05 as positive, and the rest as neutral. The prediction of the sentiment based on assigned polarity was done on a newly combined column named "com_titletext" and the outcome of the prediction is stored in a new column called "com_score".

V. RESULTS AND DISCUSSION

In this section, the results obtained after data cleaning and VADER application for sentiment analysis will be discussed in detail.

The outcome of the sentiment prediction is plotted in the form of a confusion matrix. The confusion matrix is a table describing the performance classification of the model in the form of a 3 x 3 matrix in this project [12]. Fig.1 shows the output in the form of the confusion matrix while Table 1 shows the detailed classification metrics of the sentiment analysis.

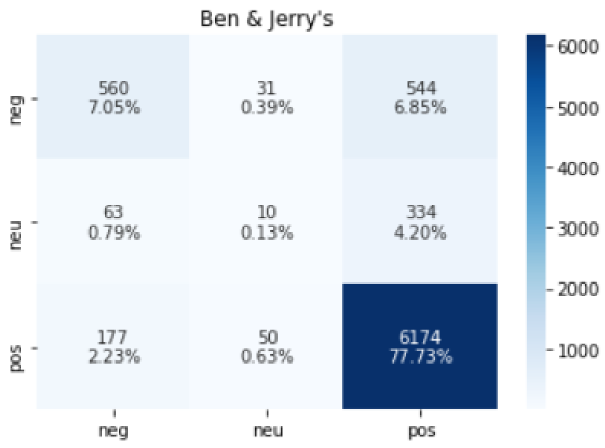

Fig. 1 The confusion matrix from Vader sentiment analysis.

Table 1 The classification report of sentiment analysis.

Classification Report	**neg**	**neu**	**pos**
Precision	70%	11%	88%
Recall	49%	2.5%	97%
F1-Score	58%	4%	92%
Support	1135	407	6401
Accuracy	85%		
Macro Avg	56%	50%	51%
Weighted Avg	81%	85%	82%

Based on the above confusion matrix and classification report, the VADER-based sentiment analysis generated an accuracy of 85%. The precision of the negative and positive classification is relatively good as this indicates the accuracy of correctly labeled sentiment is true [12]. From this analysis, it can be derived that 89% of users have a positive sentiment on Ben and Jerry's product while 10% of the users have a negative sentiment while 1% have a neutral review being predicted. In short, the customers of Ben & Jerry's are majorly happy with the products and services while only a small minority section of customers are unhappy with the products and services. The summary of this sentiment analysis is plotted in the bar chart and pie chart in Fig.2 and Fig.3 respectively.

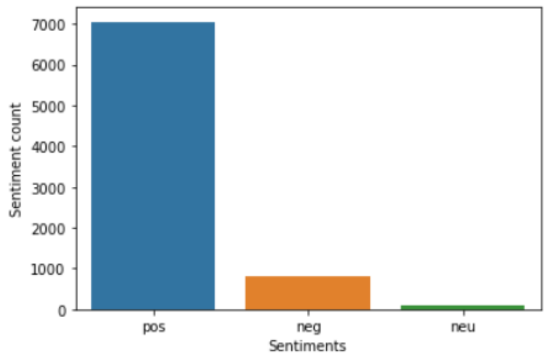

Fig. 2 The sentiment count bar chart using VADER.

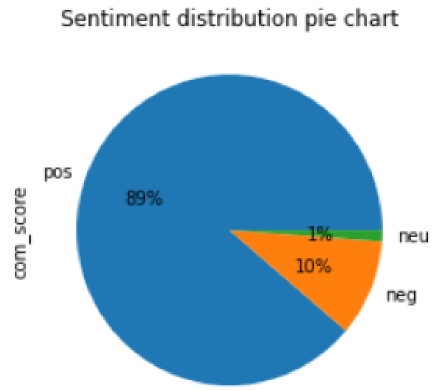

Fig. 3 The sentiment distribution using VADER.

Apart from that, Fig. 4 shows the top 10 ice cream brands by customer review breakdown which indicates that customers mostly have positive reviews. Based on Fig. 5 and Fig.6, the highest positive review comes from the brand 16_bj and the lowest positive review comes from 46_bj. The positive reviews correspond to the 4 and 5-star ratings. In addition to that, we can see in Fig.7 and Fig.8 that the brand 24_bj has the highest negative review with the count of 100 only which is very much smaller than the highest positive review, and the 1 and 2-star rating corresponds to the negative reviews. On top of that, Fig.9 and Fig. 10 shows the neutral review corresponds to the 3-star rating.

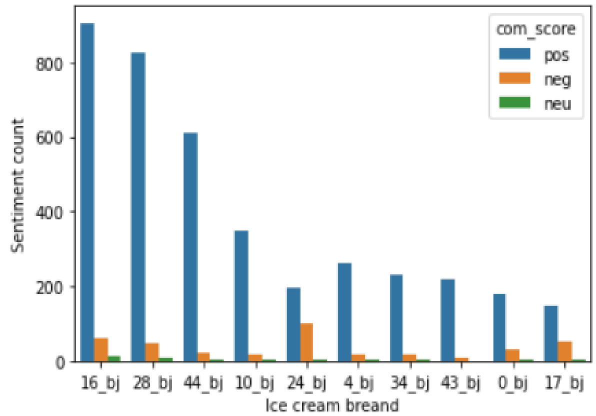

Fig. 4 Top 10 brands by the number of reviews.

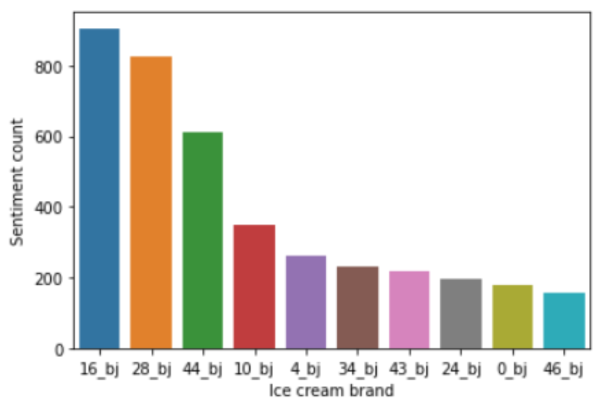

Fig. 5 Top 10 brands with positive reviews.

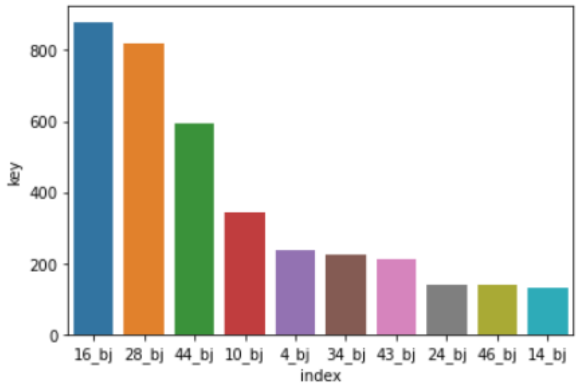

Fig. 6 Top 10 brands with 4 and 5 stars rating.

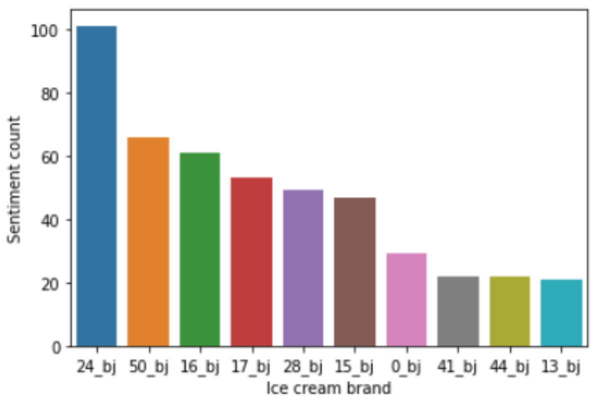

Fig. 7 Top 10 brands with negative reviews.

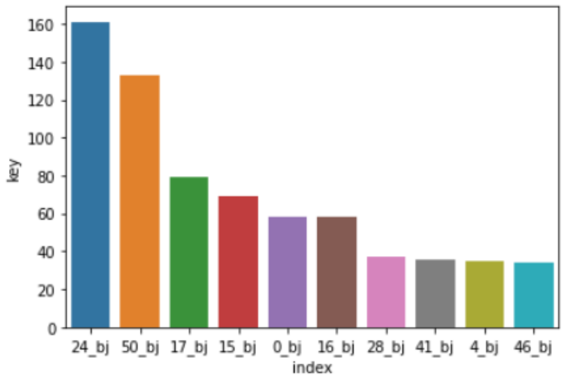

Fig. 8 Top 10 brands with 1 and 2 stars rating.

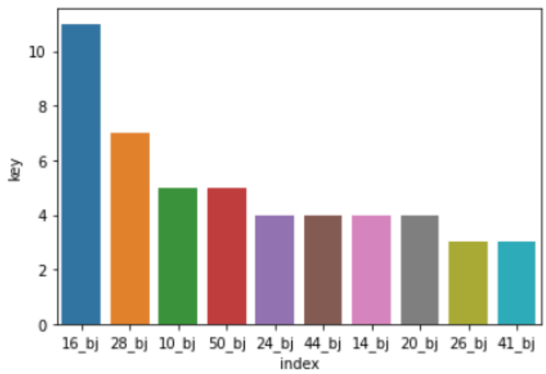

Fig. 9 Top 10 brands with neutral reviews.

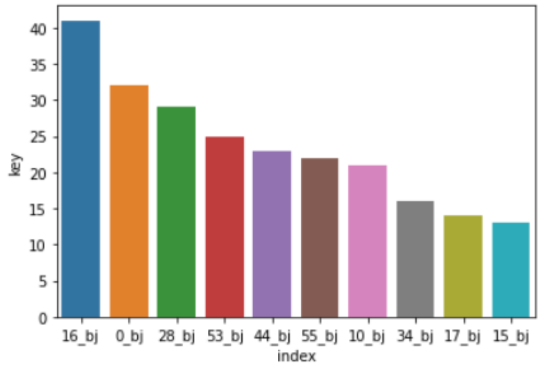

Fig. 10 Top 10 brands with 3 stars rating.

VI. CONCLUSION

The sentiment analysis using VADER has been successful in predicting the sentences with positive, negative, and neutral orientations. This has enabled us to determine the user reviews by classifying each product based on the sentiment polarity. Therefore, Ben & Jerry's can clearly distinguish the product with positive reviews, negative reviews, and neutral reviews, as well as the overall customer's reviews, which is relatively good. In short, this lexicon-based classifier manages to detect the sentiment of customer reviews with the accuracy of 85% which is almost in line with the results obtained based on previous work done as discussed in the related work section.

VII. REFERENCES

[1] T. M. S. Akshi Kumar, "Sentiment Analysis: A Perspective on its Past, Present and Future," *I.J. Intelligent Systems and Applications*, pp. 1-14, 2012.

[2] V. G. Amandeep Kaur, "A Survey on Sentiment Analysis and Opinion Mining Techniques," *JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE*, vol. 5, no. 4, 2013 .

[3] M. S. F. d. J. Ayoub Bagheri, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," *Knowledge Based System*, vol. 52, pp. 201-213, 2013.

[4] B. L. H. X. P. J. Zhongwu Zhai, "Constrained LDA for Grouping Product Features in Opinion Mining," in *Proceedings of 15th Pacific–Asia Conference, Advances inKnowledge Discovery and Data Mining*, 2011.

[5] Q. P. Y. C. Tao Xu, "Identifying the semantic orientation of terms using S-HAL for sentiment analysis," *Knowledge-Based Systems*, vol. 3, pp. 279-289, 2012.

[6] N. K. d. N. J. Venkateswarlu Bonta, "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis," *Asian Journal of Computer Science and Technology*, vol. 8, no. 2249-0701, pp. 1-6, 2019.

[7] M. B. Anton Borg, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert System with Applications* , vol. 162, 2020.

[8] M. Z. A. S. A. F. M. K. M. Q. F. Aurangzeb khan, "Sentiment Classification through Semantic Orientation Using SentiWordNet," *Life Science Journal*, vol. 10, p. 11, 2014.

[9] C. V. D., "Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments," *International Journal of Electrical & Computer Engineering*, vol. 9, no. 5, pp. 4452-4459, 2019.

[10] C. J. H. E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media tex," in *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*, 2016.

[11] S. E. a. J. Yang, "Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment," in *Proceedings of the*

International MultiConference of Engineers and Computer Scientists, Hong Kong, 2019.

[12] M. B. K. Badriya Murdhi Alenzi, "Application of Sentiment Lexicons on Movies Transcripts to Detect Violence in Videos," *International Journal of Advanced Computer Science and Applications,* vol. 10, no. 2, 2019.