Beyond Sentiment Classification: A Novel Approach for Utilizing Social Media Data for Business Intelligence

Ibrahim Said Ahmad[1], Azuraliza Abu Bakar[2], Mohd Ridzwan Yaakub[3], Mohammad Darwich[4]

Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia, Bangi, Malaysia

Abstract—**Extracting people's opinions from social media has attracted a large number of studies over the years. This is as a result of the growing popularity of social media. People share their sentiments and opinions via these social media platforms. Therefore, extracting and analyzing these sentiments is beneficial in many ways, for example, business intelligence. However, despite a large number of studies on extracting and analyzing social media data, only a fraction of these studies focuses on its practical application. In this study, we focus on the use of product reviews for identifying whether the reviews signify the intention of purchase or not. Therefore, we propose a novel lexicon-based approach for the classification of product reviews into those that signify the intention of purchase and those that do not signify the intention of purchase. We evaluated our proposed approach using a benchmark dataset based on accuracy, precision, and recall. The experimental results obtained prove the efficiency of our proposed approach to purchase intention identification.**

Keywords—*Purchase intention; sentiment analysis; lexicon; social media; product reviews*

I. INTRODUCTION

The internet and web technologies have experienced tremendous development in terms of how data is received, processed and managed over the last decade. The contemporary web provides users with the means to actively interact and modify the contents of the web through social networking platforms. People often share their opinion on these social networking platforms in the form of comments in a blog, debates, and arguments in discussion forums or status updates in social networking channels. The web 2.0 immensely contributed to this development. The web 2.0 provides features that enable users to actively interact and contribute to the web contents rather than merely reading the contents. These features make blogs, Facebook, Twitter and other social networking platforms possible. These platforms that enable people to share their opinion are referred to as social media. Extracting and analyzing the data generated on this social media data is referred to as sentiment analysis [1], [2].

Sentiment analysis research has attracted many studies over the years [3]–[5]. However, most of the research is on the accurate classification of the data into positive, neutral, and negative sentiments. Notwithstanding, several studies have emphasized the varying potentials of sentiment analysis research, in security, tourism, and business intelligence [6], [7].

In this paper, we focus on the potentials of sentiment analysis for business intelligence, specifically on the identification of people's intention to purchase a product, called purchase intention from product reviews.

Consequently, we proposed a lexicon-based approach to classify product reviews whether they signify intention of purchase or not. A lexicon-based approach was selected because it has been applied in other domains with good results. For example a study by [8] focused on cyber-harassment lexicon. In this paper, first we develop a purchase intention lexicon from product reviews, then use the lexicon to classify product reviews as to whether they signify purchase intention or not. We evaluated the approach using benchmark dataset. The experimental results show that it is possible to identify people's purchase intention from product reviews.

The remainder of this paper is organized as follows: Section 2 presents the literature review. Section 3 discusses the proposed method. Section 4 presents the results and discussion while Section 5 presents the conclusion and future works.

II. LITERATURE REVIEW

In the following sub-sections, we discuss the related literature on sentiment analysis and purchase intention identification.

A. Sentiment Analysis

Sentiment Analysis involves the classification of emotions in social media data into positive, negative or neutral sentiment. Sentiment Analysis is possible because of the huge amount of data available through web content, like twitter posts, discussion forums, product reviews, blogs, online markets and comments of web pages [9]–[11]. The task of sentiment analysis is usually achieved through the following steps:

1) Extract the desired content from twitter, blog, and forum.

2) Prepare the extracted data and furnish it by removing irrelevant pieces like symbols and repetition.

3) Detect the sentiment (if any) contained in the contents.

4) Classify the polarity of the contents into positive, negative or neutral.

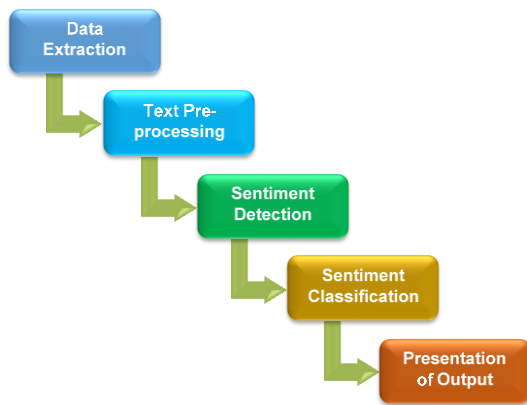*5)* Present the sentiment analysis result.

Fig 1.   Sentiment Analysis Steps.

Fig. 1 represents the general steps involved in sentiment analysis. Several attempts have been made to achieve the task of sentiment analysis. The main challenge of sentiment analysis is sentiment detection and sentiment classification of the contents [10].

This classification is normally achieved in two main ways, i.e., supervised learning approach and unsupervised learning approach. Supervised learning approach train a sentiment classifier based on the training documents which are represented by the selected features. unsupervised learning approach divide features into three classes, ''positive'', ''negative'' and "neutral" based on sentiment lexicon and then count an overall positive/negative score for a document [12].

B. Supervised Learning Approach

Supervised learning approach of sentiment analysis, also known as the machine learning approach of sentiment analysis, involves the use of popular machine learning algorithms for sentiment analysis. This is achieved by training a machine learning algorithm with a labeled dataset and then using that trained algorithm for sentiment analysis [13]. The general methodology for the supervised learning approach is illustrated in Fig. 2 as given by [14]. It involves first creating a training dataset by manually annotating reviews into different sentiment classes, and then use that training set to train a machine learning algorithm so that it can be able to automatically classify new unclassified reviews based on the sentiment they carry.

C. Unsupervised Learning Approach

An unsupervised learning approach of sentiment analysis, also known as the lexicon-based approach involves the use of a list of words with known sentiment value called sentiment lexicon for sentiment analysis. Sentiment lexicon is usually manually developed. The general methodology for the lexicon-based approach of sentiment analysis is illustrated in Fig. 3 as given by [14]. Unsupervised learning approach requires no training set. A review is classified based on developed sentiment lexicon and devised rules.

D. Related Work on Purchase Intention Mining

Purchase intention identification is an important aspect in business intelligence. The Internet and the WWW have provided new avenues through which purchase intention can be identified. Purchase intention online perhaps finds its roots from web usage mining that first appeared in 2000 by [15]. They defined web usage mining as: *"The process of applying data mining techniques to the discovery of usage patterns Web data"*. These patterns are then used for various applications depending on the domain.

In relation to purchase-intention mining specifically, [16] proposed one of the earliest studies. They proposed a Hidden Markov Model (HMM) to predict and internet user's purchase intention based on his online activity data. That is, based on search history, pictures, sounds, and other activities on the web. The precision and recall they got is 51% and 73% respectively. However, with the advent of social media, the data generated on social media has become important for predicting purchase intention. Author in [17] proposed one of the earliest research on the use of social media reviews for purchase intention mining. They proposed a method to automatically identify 'wishes' from product reviews by extracting a list of specific list *wish* words from the reviews.
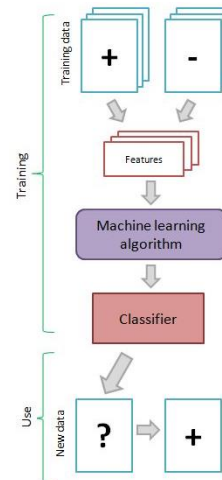

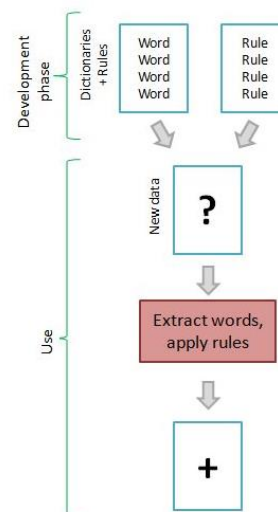Fig 2.   GeneralMethodology for Supervised Learning Sentiment Analysis Approach [14].

Fig 3.   General Methodology for unsupervised Learning Sentiment Analysis Approach [14].

It is important to focus on a specific domain when identifying purchase intention from social media data. This is because different words are used to convey opinions in different domains. Author in [18] proposed a domain-dependent model for identifying user consumption identification from social media data using CNN. They reported an accuracy of 92.54% which is an improvement on previous studies. Author in [19] proposed an approach that uses linguistic features along with statistical features for purchase intention classification. They reported that their proposed approach achieves a significant improvement compared to BOW based features model using Quora post. The best result they obtained is 93% based on AUC. Author in [20] proposed a framework based on the fuzzy set model and association rule mining to predict purchase intention from business companies fan page reviews. They illustrated the effectiveness of their approach using theoretical experiments. Author in [21] proposed an approach based on Recurrent Neural Network (RNN) for purchase intention identification. Their dataset was semi-automatically created from tweets. The RNN model achieved an F-measure of 83% which is better than other models based on linear regression, decision tree random forest, and naive Bayesian algorithms.

III. PROPOSED MODEL

A. Dataset

Purchase intention mining from social media is an emerging field of research, therefore sufficient literature in the field is yet to be established. Therefore, there is very little benchmark dataset for purchase intention mining from social media. On product reviews specifically, we are able to find one dataset by [22]. The dataset consists of 7,522 instances, divided as 6,016 for training, 752 for development and 754 for testing. We used this dataset to evaluate our proposed approach.

B. Purchase Intention Classification

In this step, a purchase intention mining approach was proposed. The approach is a classification-based task through which a product review is classified as to whether it signifies purchase intention or not. An unsupervised learning approach was adopted. Therefore, a purchase intention lexicon for movie reviews was developed.

The lexicon was generated using a set of seed words. This was then expanded using therasus.com to include synonyms in a recursive manner exponentially. The synonyms of the synonyms are also included until the list cannot be expanded anymore. The seeds were identified from previous studies. The seed words are *must buy, cannot wait, looking forward, keep an eye on, and must have*. Furthermore, a purchase intention mining approach from product reviews is proposed. The approach uses the lexicon to determine whether a review indicates purchase intention or not. The process is illustrated in Fig. 4.

To determine whether a review indicates purchase intention or not, the developed lexicon is used. A review is classified as indicating purchase intention if it contains any phrase in the seed words, while it does not indicate purchase intention if it does not contain any word in the seed words. The step involved in the classification is given as:

1) Search through a product reviews dataset.

2) For each review, if it contains a phrase from the product review purchase intention lexicon, classify that review as purchase intention review, else, as a review with not purchase intention.

3) Aggregate the total number of reviews that signify purchase intention and reviews that do not signify purchase intention.

C. Evaluation

In sentiment analysis and other classification problems, accuracy, precision, and recall are commonly used in the evaluation of the classification. Therefore, we also used accuracy, precision, and recall in evaluating our proposed approach. In order to understand how these metrics are computed, a knowledge of the confusion matrix is needed. The confusion matrix is a table that shows the performance of a classification task for which the actual values are known. In our case, there are two possible classes, whether a review contains purchase intention or not. Therefore, the 2 * 2 matrix used is shown in Table I.

The number of reviews that are correctly classified as to whether they signify purchase intention or not will be placed in TP and TN respectively, while the number of reviews that are wrongly classified as to whether they signify purchase intention or not will be placed under FP and FN, respectively.

1) Accuracy: Accuracy is a simple evaluation measure calculated as the ratio correctly predicted values to the total values. The equation is given by Equation 1:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

2) Precision: Precision is calculated as the ratio of correctly predicted positive values to the total predicted positive values. Precision tells us how much of the classified data is classified correctly. Precision is given by Equation 2:

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

3) Recall: A recall is the ratio of correctly predicted positive values to all values in the actual class. Recall tells us the amount of the correctly classified data; it is given by Equation 3:
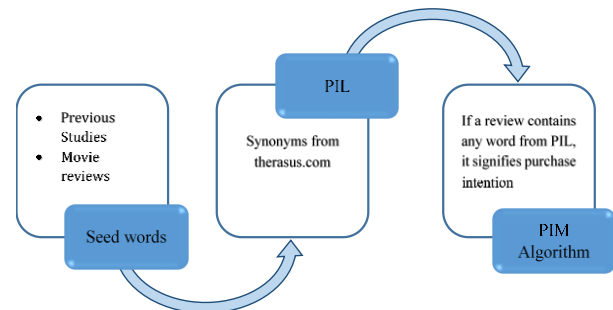
$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

Fig 4.   Product Review Purchase Intention Approach.

TABLE I.   CONFUSION MATRIX

		Classified Values	
		Positive (PI)	*Negative (Not PI)*
Actual Values	Positive (PI)	True-positive (TP)	False-negative (FP)
	Negative (Not PI)	False-positive (FP)	True-negative (TN)

IV. RESULTS AND DISCUSSION

In this section, we present the experimental results along with a discussion on the results. We used the dataset by [22] to investigate the efficiency of our approach in classifying product reviews based on those that signify purchase intention and those that do not. However, because we are able to get only one benchmark dataset, we randomly divided the dataset into three parts. Then run the experiment on each of the parts, and finally on the entire dataset. This is to be able to study the performance of the proposed approach on different parts of the dataset and deduce a more reliable conclusion. The results of the experiment are presented in Table II.

From Table II, we can see that the proposed is effective in identifying purchase intention from product reviews with an accuracy of 90%, precision of 92%, and a recall of 85%. Similarly, in the three samples of the data created, the accuracy is between 89% and 91%, while the precision is between 88%-100%. However, the recall is low compared to the accuracy and precision, ranging between 75% and 92%. When using the entire dataset, the recall is still the lowest. Recall tells us how many of the reviews that actually signify purchase intention are correctly classified. Therefore, a recall of 85% means that 15% of the reviews that signify purchase intention were not classified correctly. On the other hand, precision tells us the number of reviews classified as signifying purchase intention are actually signifying purchase intention. Therefore, a precision of 92% indicates that only 8% of the reviews that are classified as signifying purchase intention are wrong. Finally, the accuracy tells how accurate our approach is in classifying the reviews are signifying purchase intention or not. An accuracy of 90% means that our approach can classify 90% of the reviews correctly. The results are further illustrated in Fig. 5.

TABLE II.   PERFORMANCE OF PROPOSED APPROACH

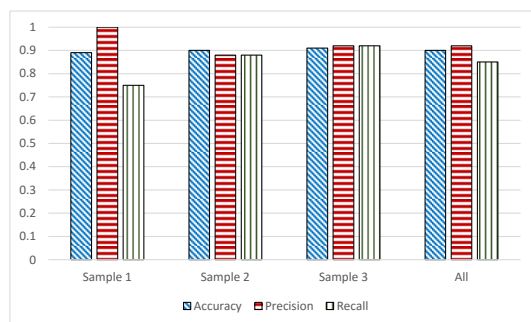| Data	Accuracy	Precision	Recall
Sample 1	0.89	1	0.75
Sample 2	0.90	0.88	0.88
Sample 3	0.91	0.92	0.92
All	0.90	0.92	0.85

Fig 5.   Performance of Proposed Approach.

V. CONCLUSION AND FUTURE WORKS

In this paper, we proposed an approach to automatically classify product reviews into two categories, whether the reviews indicate purchase intention, or whether the reviews do not indicate purchase intention. The approach proposed is a lexicon-based approach, in which a domain-specific purchase intention lexicon was developed and used in the classification. Based on the accuracy, precision, and recall, our approach achieved promising results and hence affirms the notion that reviews contained in social media can be used for business intelligence.

In the future, we intend to use the proposed approach in predicting business performance and the success of a business based on the amount of purchase intention from the reviews.

REFERENCES

[1]   J. Awwalu, A. A. Bakar, and M. R. Yaakub, "Hybrid N-gram model using Naïve Bayes for classification of political sentiments on Twitter," Neural Comput. Appl., pp. 1–14, May 2019, doi: 10.1007/s00521-019-04248-z.

[2]   M. R. Yaakub, Y. Li, and J. Zhang, "Integration of Sentiment Analysis into Customer Relational Model: The Importance of Feature Ontology and Synonym," Procedia Technol., vol. 11, no. Iceei, pp. 495–501, 2014, doi: 10.1016/j.protcy.2013.12.220.

[3]   T. Al-Moslmi, N. Omar, S. Abdullah, and M. Albared, "Approaches to Cross-Domain Sentiment Analysis: A Systematic Literature Review," IEEE Access, vol. 5, pp. 16173–16192, 2017, doi: 10.1109/ACCESS.2017.2690342.

[4]   T. Al-Moslmi, M. Albared, A. Al-Shabi, N. Omar, and S. Abdullah, "Arabic senti-lexicon: Constructing publicly available language resources for Arabic sentiment analysis," J. Inf. Sci., vol. 44, no. 3, pp. 345–362, Jun. 2018, doi: 10.1177/0165551516683908.

[5]   S. R. Ahmad, A. A. Bakar, and M. R. Yaakub, "Metaheuristic algorithms for feature selection in sentiment analysis," in 2015 Science and Information Conference (SAI), 2015, pp. 222–226, doi: 10.1109/SAI.2015.7237148.

[6]   M. Garcia-Herranz, E. Moro, M. Cebrian, N. A. Christakis, and J. H. Fowler, "Using friends as sensors to detect global-scale contagious outbreaks," PLoS One, vol. 9, no. 4, p. e92413, Apr. 2014, doi: 10.1371/journal.pone.0092413.

[7]   A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with Twitter: What 140 characters reveal about political sentiment," in Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010, pp. 178–185, doi: 10.1074/jbc.M501708200.

[8]   M. Rezvan, K. Thirunarayan, S. Shekarpour, V. L. Shalin, L. Balasuriya, and A. Sheth, "A quality type-aware annotated corpus and lexicon for harassment research," in WebSci 2018 - Proceedings of the 10th ACM Conference on Web Science, 2018, pp. 33–36, doi: 10.1145/3201064.3201103.

[9]   B. Liu, Sentiment analysis and opinion mining, vol. 5, no. 1. Morgan & Claypool Publishers, 2012.

[10]   W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Eng. J., vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/J.ASEJ.2014.04.011.

[11]   D. Zimbra, A. Abbasi, D. Zeng, and H. Chen, "The state-of-the-art in twitter sentiment analysis: A review and benchmark evaluation," ACM Trans. Manag. Inf. Syst., vol. 9, no. 2, 2018, doi: 10.1145/3185045.

[12]   S. Wang, D. Li, X. Song, Y. Wei, and H. Li, "A feature selection method based on improved fisher ' s discriminant ratio for text

sentiment classification," Expert Syst. Appl., vol. 38, no. 7, pp. 8696–8702, 2011, doi: 10.1016/j.eswa.2011.01.077.

[13] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," Expert Syst. Appl., vol. 57, pp. 117–126, Sep. 2016, doi: 10.1016/J.ESWA.2016.03.028.

[14] M. Taboada, "Sentiment analysis: an overview from linguistics," Annu. Rev. Linguist., vol. 2, pp. 325–347, 2016.

[15] J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan, "Web usage mining: Discovery and applications of usage patterns from web data," ACM SIGKDD Explor. Newsl., vol. 1, no. 2, p. 23, Jan. 2000, doi: 10.1145/846183.846188.

[16] F. Wu, I. H. Chiu, and J. R. Lin, "Prediction of the intention of purchase of the user surfing on the web using hidden Markov model," in 2005 International Conference on Services Systems and Services Management, Proceedings of ICSSSM'05, 2005, vol. 1, pp. 387–390, doi: 10.1109/ICSSSM.2005.1499501.

[17] J. Ramanand, K. Bhavsar, and N. Pedanekar, "Wishful Thinking Finding suggestions and 'buy' wishes from product reviews," in NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, 2010, pp. 54–61.

[18] X. Ding, T. Liu, J. Duan, and N. J.Y., "Mining user consumption intention from social media using domain adaptive convolutional neural network," in Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.

[19] V. Gupta, D. Varshney, H. Jhamtani, D. Kedia, and S. Karwa, "Identifying purchase intent from social posts," in Eighth International AAAI Conference on Weblogs and Social Media, 2014.

[20] L.-J. Kao and Yo-Ping Huang, "Predicting purchase intention according to fan page users' sentiment," in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 831–835.

[21] R. Haque, A. Ramadurai, M. Hasanuzzaman, and A. Way, "Mining Purchase Intent in Twitter," Comput. y Sist., vol. 23, no. 9, 2019, doi: 10.13053/CyS-23-3-3254.

[22] X. Ding, B. Cai, T. Liu, and Q. Shi, "Domain adaptation via tree kernel based maximum mean discrepancy for user consumption intention identification," in IJCAI International Joint Conference on Artificial Intelligence, 2018, vol. 2018-July, pp. 4026–4032.