
Comparative Study on Sentiment Analysis Approach for Online Shopping Review

Norhaslinda Kamaruddin¹, Siti Azizah Abas², Abdul Wahab³

^{1,2}Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40400 Shah Alam, Selangor, Malaysia

³Kulliyah of Information and Communication Technology, International Islamic University Malaysia, Gombak, Kuala Lumpur, Malaysia

norhaslinda@tmsk.uitm.edu.my¹, sitiazizahabas96@gmail.com², abdulwahab@iiu.edu.my³

Article History: Received: 10 November 2020; Revised: 12 January 2021; Accepted: 27 January 2021;

Published online: 05 April 2021

Abstract: The internet has revolutionized the way most people shop. Flexibility, convenience, products' variations, better price, and more privacy contribute to the exponential growth of online shopping platforms. However, due to the nature of online shopping, customers are not able to physically test the product before purchasing. They rely on the information given by the seller and previous customers' ratings to make their decision. Sometimes, the information that is given by sellers may be fraudulent, misleading, or over claim. Many researchers had found out that ratings and other customers' reviews can be manipulated and did not reflect on the actual customers' sentiment on the particular product. This research investigates how sentiment analysis can be used as an alternative solution to measure the positive, negative, and neutral feedback of the past reviews. It is to offer more comprehensive way to help the customers make an informed decision for the product that they wish to buy based on the totality of the reviews. This paper makes a comparative study on sentiment analysis methods on online shopping reviews. This can lead to the proposed theoretical framework of an alternative solution for better insight exploration. It is envisaged that this research would benefit the customer in making a better decision when doing online shopping and may act as a feedback mechanism for the seller to provide good products and services. A good product rating can influence many new buyers and increase business revenue and expansion.

Keywords: Customer Review, Sentiment Analysis, Text Analytics, Comparative Study, Customer Experience

1. Introduction

Shopping is an essential activity that affects customers' life in many aspects [1]. Customers regularly shop in physical retail stores where they can explicitly select and choose the products and goods that they want to purchase. Usually, the customers select the retail stores based on their familiarity, geographical location, and recommendation by other people. Therefore, word of mouth was the best advertising technique used by the resellers to promote their products in the past. Apart from that, people usually have little or no knowledge on the product and dependent on the advice of the salesperson [2]. Such a situation offers limited options to choose and make comparisons on the product that they want to purchase.

Now, the internet has revolutionized the way most people shop [3]. There is an exponential growth of online

shopping [4] due to its flexibility, convenience, products' variations, better pricing, and more privacy. This fact is also supported by a group of researchers who stated that online shopping provides so many options for customers in terms of price, details, and choices [5]. This will benefit the customer by saving their time of making a decision when selecting the optimal selection of a product and this selection is not just limited to the retail shops that are located in the vicinity of an individual but can cover a network of online reachability. A research conducted by Selyukh [6] revealed the top 5 reasons consumers prefer online shopping are as presented in Figure 1.

Figure 1 shows that two main factors of online shopping are the known availability of the product and can be done without any time restriction, with 88% survey respondents' agreement. This is different than a physical shop that has their strict operating hours, typically around 10 am to 10 pm. Any purchase cannot be done outside of this time range. Internet shopping offers the freedom to browse and traverse the site 24/7 as long as the web-site is running. This really makes people comfortable in selecting their preferred time to do their shopping. The next two factors that share 84% respondents' selection are the better product selection and the time to make the decision is shorter as the shopping website will list the products according to the specific categories. Moreover, some of the websites offer the service of search engines where the customer can apply a certain filter that restricts their choice based on the criteria they outlined. This method would surely shorten the list which shortens the time to make decision.

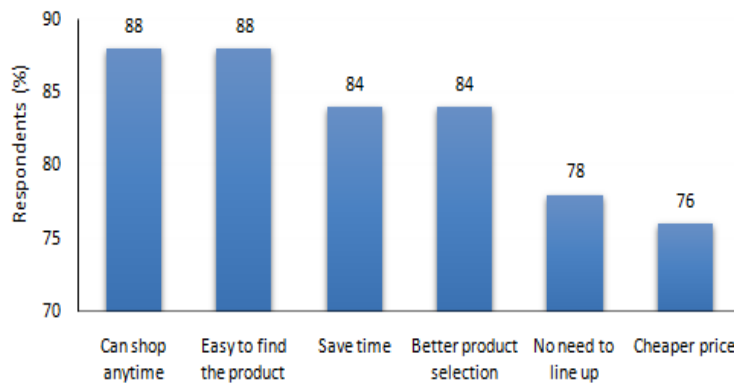


Figure 1. Top reasons for shopping online [6]

From the survey, it was also observed that 78% of people just hate to queue in the physical store. According to Kakava and Erasmus [7], queuing is regarded as a major inconvenience, unpleasant experience and a waste of time. Such activity can be avoided using online shopping. Customers can just select and drop the purchases in the cart and make payment using online banking. Subsequently, a cheaper price is the fifth reason why consumers prefer to do their shopping online. This is because due to fierce competition to stay relevant, companies have to offer lucrative deal to the customers to attract them to buy the products and goods. In summary, it can be inferred that online shopping is an alternative way to shop for busy people and for those who want to find a great product at a reasonable price.

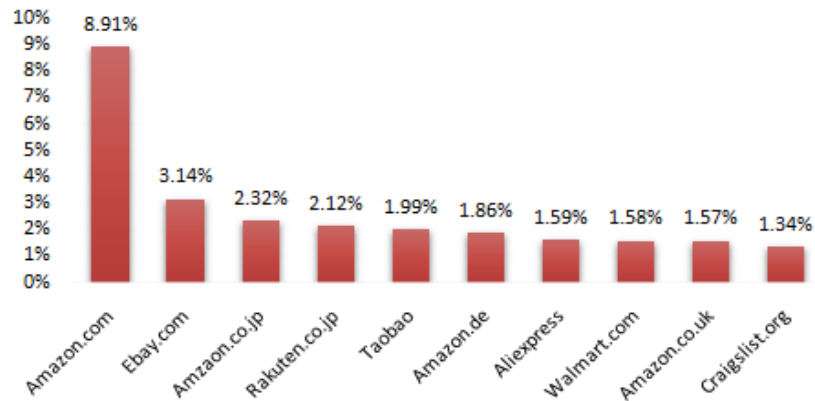


Figure 2. Top 10 worldwide online shopping platform [8]

Due to many demands from the customers, there are many online shopping platforms that have been developed. Based on Figure 2, the list of the top 10 e-commerce websites and online shopping platforms in the world is shown [8]. The data collected is of March 2020. The Amazon.com topped the chart with 8.91% of traffic shares with a difference of almost 6% of the worldwide market. The rest of the competitors are almost at par with market shares, ranging between 1% to 3%. Hence, it can be inferred that Amazon.com is the world's largest online shopping platform preferred by the worldwide customers.

Online shopping also has now become a trend in Malaysia because of its convenience and can be accessed anywhere at any time using an internet connection [9]. Table 1 shows the ranking of the top 5 websites in e-commerce and online shopping in Malaysia [8]. The March 2020 statistics reveal that the Malaysian online shopping traffic share is dominated by Shopee with approximately 30%, which amount to one-third of the Malaysian market. The superiority of Shopee traffic share is duplicated by Lazada with only half of the traffic share of 15%. Amazon.com although is the most preferred platform globally is only ranked in the fourth position for the Malaysian market. Such finding is also reported by [10] showing Shopee performance is far superior to the other online shopping platforms in Malaysian market.

Table 1. Top 5 online shopping platforms in Malaysia [8]

Online shopping platform	Traffic share (%)
1 Shopee (https://www.shopee.com.my)	29.65
2 Lazada (https://www.lazada.com.my)	15.52
3 Mudah.my (https://www.mudah.my)	7.96
4 Amazon.com (https://www.amazon.com)	3.10
5 Taobao (https://www.taobao.com)	2.93

Sitegiant [11] makes a comprehensive comparison of the services provided by the online shopping platforms. It is observed that the reasons of Shopee soaring popularity are free shipping for the buyer and no transaction fee

and gateway fee for the seller. The introduction of Shopee coin for login into the apps also captures the market's attention. Moreover, Shopee also provides daily service for dispute settlement and this can be settled within three to seven days once the transaction status is shown as delivered.

Despite the glowing recommendation of online shopping, it also suffers from a few drawbacks. First, the shopping experience is not as fulfilling as the physical exercise because the shopper could not physically test the product. They have to rely on the information given by the seller and sometimes the descriptions are misleading and over claim. There is also a growing case of a fraudulent seller as they sell fake products that they claimed as the original. There is no guarantee of the products' quality, the genuinity of the seller, and the safety of products' shipments. Mittal [12] confirmed that customers are unable to control the quality of the product, return policies, additional charges, and other issues.

Hence, to cater to the need for validation of the product, seller, and overall shopping user experience, a rating system is introduced. In the rating system, the customers are allowed to rate and review the product that they purchased. This is because the customers' opinions or reviews play a vital role in marketing the products and predicting sales for their business [13]. Recently, online reviews have received much attention because their visibility has been proven to play an important role during the purchase process [14]. A study by Fu et. al [15] has proven that the customers' online reviews search behavior is considerably affected by human association levels of recycled products because customers rely on safety perception reviews when buying a product. Moreover, the reviews may be in the form of text or image making it easier for customers to manipulate it.

Although many researchers had found that the provided reviews and ratings can help consumers in purchasing, it can still be manipulated and does not reflect on the consumers' sentiment of a specific product. Thus, the sentiment analysis technique could be an alternative solution to measure the positive, negative, and neutral feedback of the reviews. This can ensure that the reviews and feedbacks are given a more comprehensive way to help customers to make an informed decision for the product that they wish to buy. Hence, this paper attempts to study in more detail the suitability of sentiment analysis to be employed for insights exploration specifically for online shopping reviews.

This paper is organized in the following manner. Section 2 describes the fundamental of sentiment analysis. The comparative study on other researchers' works on customers' reviews is presented in Section 3. The proposed solution and discussion are provided in Section 4. This paper concludes with summary and future works in Section 5.

2. Sentiment Analysis

Sentiment analysis classification is a process analyzing a document to detect emotion within the text. Sentiment analysis is a type of data mining that measures the inclination of people's opinions through natural language processing (NLP), computational linguistics and text analysis, which are used to extract and analyze

subjective information from the Web - mostly from social media and similar sources. The analyzed data quantifies the general public's sentiments or reactions toward certain products, people, or ideas and reveal the contextual polarity of the information. Analysis of customers' experience is vital in online shopping to know the exact perceptions of how customers feel about the product. It is usually used to classify customers' experience towards the product; either it is positive, neutral, and negative. The customers' sentiment is basically influenced by the customers' satisfaction. Maks & Vosen [16] found that sentiment analysis tools perform better compared to reviewer ratings. This finding is also supported by Bhatt et. al that reported the classification of review with sentiment analysis provides accurate feedback to the user [17].

There are two types of sentiment analysis classification; which are, machine learning approach and lexicon-based approach. The summary is illustrated in Figure 3. The machine learning approach relies on natural language processing to classify and detect the emotion whereas the lexicon-based approach is used to classify a document based on the meaning of words and the current sentiment of the text.

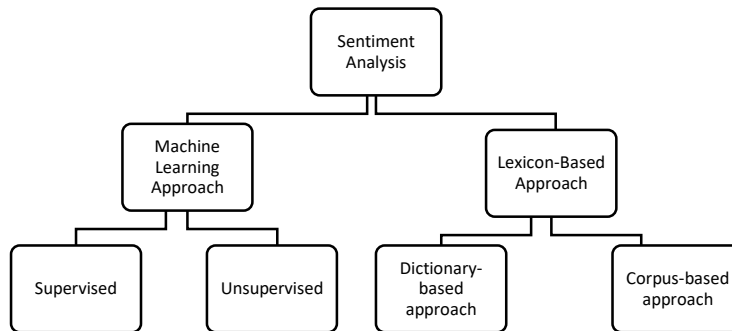


Figure 3. Sentiment analysis classification techniques

2.1. Machine Learning Approach

Machine learning approach uses linguistic feature or synthetic feature. It relies on machine learning algorithms to solve the sentiment analysis of text mining. The texts are split into keywords based on the sentences using a particular machine learning model. The text mining process selects keyword phrases for classification into classes based on a search term. For the text classification, it relies on an aspect of words or sentences to identify the class. The keywords, while not a part of the input dataset, provides important information to the ranking engine in the task of sentiment analysis.

A machine learning model is an algorithm that performs a data analysis based on a set of input data. Machine learning algorithms use a set of parameters, called features, to "learn" what the inputs should be, and then perform an analysis on the training set to generate predictions for the future. Machine learning models typically use time series data from internet websites, such as Google Trends, Twitter, YouTube, and Yelp, in order to learn about future trends and search patterns.

The machine learning approach can be divided into two; namely, supervised and unsupervised approaches.

The supervised machine learning is usually carried out to map the input to labeled output. Example of supervised learning is Decision Tree, Naïve Bayes, Bayesian Network and Neural Network. On the contrary, unsupervised machine learning is usually used in clustering, representation learning, and density estimation. It is useful in the exploratory analysis because it can automatically identify structure in data. Example of unsupervised learning is clustering, partitioning, overlapping and probabilistic.

2.2. Lexicon-based Approach

Lexicon is a vocabulary of a language or subject. Senti-Lexicon or sentiment lexicon is a vocabulary of a word in terms of sentiment such as the word “impressed”, “problem”, “lack” and “good”. Senti-Lexicon algorithm proves to be scalable, simple, and achievable compared to machine learning. Challenges such as blind rejection, sarcasm involvement, nuanced phrases, fake comments, spam detection, vigilance to time, managing hidden features could be detected using senti-lexicon [18].

There are two types of lexicon-based approaches; which are, dictionary-based and corpus-based. The corpus-based approach can be defined in statistical and semantic. It helps to solve the problem in finding opinions in a specific context into a semantic orientation such as sentiment polarity which is negative and positive and sentiment strength of word or text. Dictionary-based approach is the most well-known and the most widely used approach. It measures the feeling of consumers within the text dictionary into a multiple dimension such as fear, sad, happy and others. This method classified using dictionary list to measure the word and keyword of the text.

3. Comparative Study On Other Researchers’ Works

For better understanding of the current situation, several state-of-the-art approaches are studied and summarize as presented in Table II. Because Amazon is the most popular online platform globally, we make the comparative study on the Amazon reviews using different approaches of sentiment analysis of machine learning, lexicon-based or even hybrid approaches. This is to give better understanding of the range of performance that can be yielded if such approach is applied.

Table 2. BCRs of different JPEG QF attacks on image LENA

Authors	Sentiment Analysis Approach	Data	Approach Detail	Result	Discussion
[19]	Supervised Machine Learning - Random Forest (RF) - Support Vector	Product review on Amazon	Combined supervised classification methods to find the performance in	two Accuracy RF = 81% SVM = 82.4 RF + SVM	Combination of classifiers produced an increment of 1% to 2.4% accuracy as compared to individual use of SVM and RF

	Machine (SVM)		large-scale reviews	= 83.4%	respectively.
[20]	Unsupervised Machine Learning - Latent Dirichlet Allocation (LDA) Supervised Machine Learning - Regression analysis	Coffee, online food and groceries from Amazon	LDA is used to determine the hidden meaning of all input while Regression analysis is used to examine the performance from LDA result.	LDA is a powerful tool to find hidden topic in underlying texts.	Combination of both supervised and unsupervised learning to capture the keywords.
[21]	Supervised Machine Learning - Naïve Bayes	Reviews on Amazon.com	Used Naïve Bayes to identify e-Commerce service quality.	Accuracy Naïve Bayes = 90%	Recorded 90% accuracy and recall and 90.5% precision.
[22]	Lexicon-based approach - Corpus-based approach -Dictionary-based approach Supervised Machine Learning Approach - Support Vector Machine - Decision Tree	Reviews on Amazon.com	Proposed both lexicon-based and supervised machine learning approaches. Then, combined to find out performance.	Accuracy Lexicon = 92.67% Lexicon + SVM = 91.33% Decision Tree = 82%	Lexicon-based method performed the best individually.
[23]	Supervised Machine Learning - K-Nearest Neighbor (kNN) - Decision Tree - Support Vector Machine (SVM) - Random Forest	Amazon Book	Using multiple supervised machine learning approaches to find accuracy for each approach	Average accuracy of 6 book: Random Forest = 90.15%	Random Forest classifier performs the best as compared to the other classifiers.

	- Naïve Bayes				
[24]	Supervised Machine Learning - Linear Support Vector Machine - Multinomial Naïve Bayes - Stochastic Gradient Descent - Random Forest - Logistic Regression - Decision Tree	Reviews on Amazon.com	Proposed supervised machine learning on a large dataset for each category of products from Amazon.com	Accuracy Support Vector Machine = 94.02%	SVM performed the best as compared to the other classifiers.
[25]	Supervised Machine Learning - Naïve Bayes - Support Vector Machine	Reviews from Amazon.com	To find accuracy for a large number of book reviews from Amazon	Accuracy Naïve Bayes SVM = 82%	Combination of SVM and Naïve Bayes recorded better accuracy as compared to individual classifier used.
[26]	Supervised Machine Learning - Naïve Bayes - Support Vector Machine	Product reviews from Amazon.com	Compared the performance in term of accuracy, precision, recall and F-measure	Accuracy Naïve Bayes SVM = 90% SVM = 83%	Naïve Bayes performed better than SVM.
[27]	Lexicon-based approach - SentiwordNet	Amazon products	Proposed to improve sentiWordNet from previous research with increment recorded by 11.86%	Accuracy SentiWordNet = 68.63%	Showed that sentiWordNet has lower accuracy than the other type of sentiment analysis

From the comparative table in Table II, it is observed that both machine learning and lexicon approach can perform well with accuracy ranging from 65% to 95% approximately. Both approaches are used by the researchers to find on the sentiments on the products on Amazon. Combination of the approach or hybrid

approach although not as popular as single approach, also shows potential to be developed further. From the perspective of machine learning approach, the most reported classifier with good performance is Support Vector Machine (SVM), Naïve Bayes, and Random Forest [19, 21, 23, 24, 25]. Although the data studied are different, these three classifiers are popular with the researchers. However, the unsupervised learning approach is not quite popular as the supervised approach. Hence, we are going to consider supervised approach only.

The lexicon-based approach focuses on the semantic of the words. Based on Table II, this approach gives mixed results as can be seen in [22] and [27]. Lexicon-based approach is most accurate if the dictionary used is comprehensive that can cater multiple meaning of a word depending on the context of its use. However, it may suffer if the words are ambiguous or vague that result in poor accuracy. Hence, for the purpose of this study, we would like to propose the combination of both approaches to compensate the weaknesses and strengths to gain the most benefit of the hybrid approach.

4. Methodology

The proposed research method is divided into two phases, namely, review sub-processing and sentiment analysis. The reviews from users are scrapped from the Shopee website on the comment section of the products. This is because in Malaysia, Shopee captures almost one third of the market share. The reviews are captured manually and stored in the database. Then, in the data preparation stage, the reviews are cleaned from the artefacts. The prefixes and suffixes are removed to get the root word. In addition, the stop words are also eliminated as it gives very minimal meaning if it is used alone.

Then, the relevant terms are extracted and stored into list of terms database. Once the list of terms completed, review analysis will be conducted to compare with emotional word dictionary to complete first phase on review sub-processing. This is to get the input to the sentiment analysis phase. Later in sentiment analysis phase, the result from the previous phase will be used to get the sentiments from the users’ comment on the product which to be categorized into positive, neutral or negative sentiment. The machine learning classifier will then be used to rank the product based on the sentiments that are categorized. The overall workflow is shown in Figure 4 and each phase will be explained further in the next section.

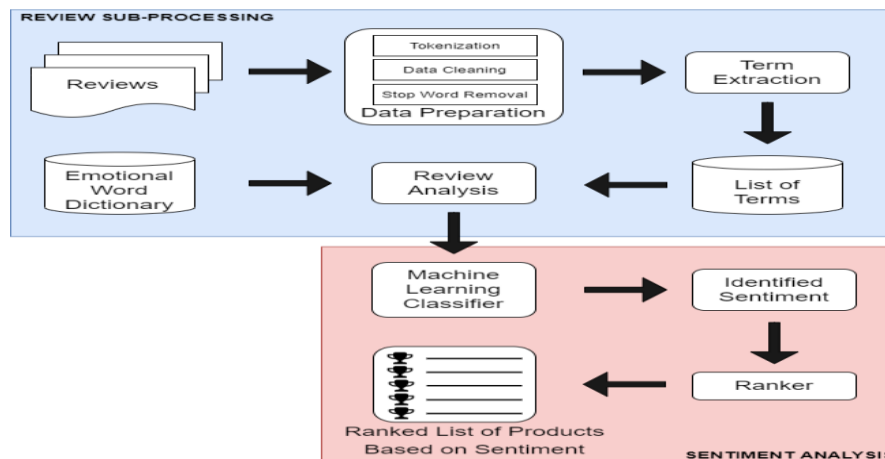


Figure 4. The proposed methodology for sentiment analysis for the online shopping reviews.

4.1. Reviews Sub-processing

The users' comments from the website are extracted using web scraping method. After scrapping process, the comments will go through data preparation stage. In this stage, the reviews will do tokenization process where the comments in strings are break into tokens. When the tokenization process is done, the comments must be cleaned to avoid some unnecessary errors in the data. Then, all the stop words will be removed since it can hinder the analysis process. Once done, all the relevant terms will be extracted and stored in list of terms database. Next, after the relevant terms are identified, the review analysis of the products can be done with the help of emotional word dictionary to complete the first phase of the method.

4.2. Sentiment Analysis

Once the review analysis is done, the sentiment analysis from the reviews is started. In this phase, the machine learning classifier is used to identify the sentiment into positive, neutral or negative feedbacks. From the identified sentiments, the ranking process is next to be executed to rank the list of products based on sentiments to help the user to get better insights of the products instead of long text in the comment section.

4.3. Expected Output

The findings of this study are beneficial to the consumers and online reseller by providing an effective analysis review which plays an important role in online shopping platform. This research makes it easier for the customer in making a better decision when doing online shopping in the future. This research also would be a beneficial input for online shopping sellers. It may act as a feedback mechanism for the seller to provide good products and services. It is observed that an improvement in product rating can influence many new buyers and can increase revenue and business expansion.

5. Conclusion and Future Work

This paper investigates how sentiment analysis can be used as an alternative solution to measure the positive, negative, and neutral feedback of the past reviews. From the comparative study, it is found that the sentiment analysis approach can yield performance between 65% to 95% accuracy regardless of the approach used, either; machine learning, lexicon-based or hybrid approaches. In this work, the hybrid approach of using both lexicon-based approach in review sub-processing and machine learning approach in sentiment analysis is proposed to compensate the weaknesses and strengths of the approaches. It is hoped that such approach can help customers to make informed decision to buy the products through online shopping platform and can be extended to the seller as value-added features that can attract confidence and trust of the customers to their products. This approach can also be extended in many other domains such as sentiment mining in education [28], extracting requirements [29] and job matching system [30].

References

1. Maggioni, S. Sands, R. Kachouie, and Y. Tsarenko, Shopping for well-being: The role of consumer decision-making styles. *Journal of Business Research*, 105, 21–32, 2019.
2. S. Ramaswamy, Shopping Then and Now: Five ways retail has changed and how businesses can adapt. *Google Think Insights*, 1(1), 1- 4, 2013.
3. P. Yannopoulos, The Impact of the Internet on marketing strategy, *International Journal of E-Business Research*, 6(4), 38–51, 2011.
4. N. Karthikeyan and S. Selva Brunda, Online consumer inspection on using E-shopping service of E-commerce. *International Journal of Recent Technology and Engineering*, 8, 438–442, 2019.
5. Y. Maker, A. Callahan, B. Mcsherry, J. M. Paterson, and L. Brophy, Improving access and support for consumers with cognitive disabilities. *A Guide for Retailers*, 1(1), 1–37, 2019.
 - a. Selyukh, What Americans Told Us About Online Shopping Says A Lot About Amazon. *National Public Radio*, 1–17, 2018.
6. N. Z. Kakava and A. C. Erasmus, Waiting as determinant of store image and customer satisfaction: A literature review. *Journal of Family Ecology and Consumer Sciences*, 40(1), 99–113, 2012.
7. SimilarWeb, Similar Web, March 2020. Accessed on May. 30, 2020. [Online]. Available: <https://www.similarweb.com/website/flytap.com>
8. Ismail, Whoopie, it's Shopee. *News Strait Times*, 1-2, Aug. 28, 2017
9. Statista, Top 10 e-commerce sites in Malaysia as of 1st quarter 2020, by monthly traffic(in 1,000 clicks), April 2020. Accessed on Sep. 15, 2020. [Online]. Available: <https://www.statista.com/statistics/869640/malaysia-top-10-e-commerce-sites/>
10. Sitegiant, Sitegiant.my, April 2020. Accessed on Jul, 22, 2020. [Online]. Available: <https://www.sitegiant.my>.
11. T. Mittal, E-commerce: It's impact on consumer behavior. *Global Journal of Management and Business Studies*, 3(2), 131-138, 2017.
12. S. P. Sheela, Sentiment analysis and prediction of online reviews with empty ratings. *International Journal of Applied Engineering Research*, 13(14), 11532–11539, 2018.
13. N. Korfiatis, E. García-Bariocanal and S. Sánchez-Alonso, Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11(3), 205–217, 2012.
14. H. Fu, G. Manogaran, K. Wu, M. Cao, S. Jiang, and A. Yang, Intelligent decision-making of online shopping behavior based on internet of things. *International Journal of Information Management*, 50(1),515–525, 2019.

15. Maks and P. Vossen, Sentiment analysis of reviews: Should we analyze writer intentions or reader perceptions? Paper Presented at the International Conference Recent Advances in Natural Language Processing (RANLP 2013), Hissar, Bulgaria, 415–419, 2013.
 - a. Bhatt, A. Patel, H. Chheda, and K. Gawande. Amazon review classification and sentiment analysis. *International Journal of Computer Science and Information Technologies*, 6, 5107–5110, 2015.
16. D. Mumtaz and B. Ahuja, Sentiment analysis of movie review data using senti-lexicon algorithm. Paper Presented at the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATCCT 2016), 592–597, 2017.
17. Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, Random Forest and Support Vector Machine-Based hybrid approach to sentiment analysis. *Procedia Computer Science*, 127, 511–520, 2018.
18. Y. Heng, Z. Gao, Y. Jiang, and X. Chen, Exploring hidden factors behind online food shopping from Amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42, 161–168, 2020.
19. K. Sari, A. Alamsyah, and S. Wibowo, Measuring e-Commerce service quality from online customer review using sentiment analysis. *Journal of Physics: Conference Series*, 971, 1–6, 2018.
 - a. A. Prakoso, W. Yananta and A. F. Setyawan, A lexicon-based sentiment analysis for Amazon web review. Paper Presented at the 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life (iSemantic 2018), 503–508, 2018.
20. K. S. Srujan, S. S. Nikhil, H. R. Rao, K. Karthik, B. S. Harish and H. M. K. Kumar, Classification of Amazon book reviews based on sentiment analysis. *Advances in Intelligent Systems and Computing*, 672, 401–411, 2018.
21. T. U. Haque, N. N. Saber and F. M. Shah, Sentiment analysis on large scale Amazon product reviews. Paper Presented at the 2018 IEEE International Conference on Innovative Research and Development, (ICIRD 2018), 1–6, 2018.
22. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar and M. Dey, A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews. Paper Presented at the 2020 International Conference on Contemporary Computing and Applications (IC3A 2020), 217–220, 2020
23. Vanaja and M. Belwal, Aspect-level sentiment analysis on e-commerce data. Paper Presented at the 2018 International Conference on Inventive Research in Computing Application (ICIRCA), 1275–1279, 2018
24. Hamouda and M. Rohaim, Reviews classification using sentiwordnet lexicon. *World Congress on*

- Computer Science and Information Technology (OJCSIT), 2, 120–123, 2011.
25. M. Bakri, N. Kamaruddin, M. Hamiz, P. Marlia, A. H. S. Nurhasmiza, A. B. Shahirah, Z. Othman, N. A. S. Nilam, M. S. O. Shahir, S. A. H. Nurul and L. C. L. Effie, Insights extraction on cross-cultural interaction through astronomy online labs using data analytics. *Indonesian Journal of Electrical Engineering and Computer Science*, 16(1), 508-515, 2019.
 26. N. Kamaruddin, A. Wahab, M. Bakri and M. Hamiz, Science lab repository requirements elicitation based on text analytics. Paper Presented at the International Conference on Soft Computing in Data Science (SCDS 2019), 351-360, 2019.
 27. N. Kamaruddin, A. Wahab and R. A. M. Lawi, Jobseeker-industry matching system using automated keyword selection and visualization approach. *Indonesian Journal of Electrical Engineering and Computer Science*, 13(3), 1124-1129, 2019.