Sentiment Analysis for Customer Review: Case Study of GO-JEK Expansion

Alifia Revan Prananda[1)*], Irfandy Thalib[2)]

[1)2)] *Department of Electrical and Information Engineering, Faculty of Engineering, Universitas Gadjah Mada, Indonesia*

Jl. Grafika No. 2 Kampus UGM, Yogyakarta

[1)] alifia.revan@mail.ugm.ac.id, [2)] irfandythalib@mail.ugm.ac.id

Abstract

Background: Market prediction is an important thing that needs to be analyzed deeply. Business intelligence becomes an important analysis procedure for analyzing the market demand and satisfaction. Since business intelligence needs a deep analysis, sentiment analysis becomes a powerful algorithm for analyzing customer review regarding to the business intelligence analysis.

Objective: In this study, we perform a sentiment analysis for identifying the business intelligence analysis in GO-JEK.

Methods: We use Twitter posts collected from the Twint library which consists of 3111 tweets. Since the dataset did not provide a ground truth, we perform Microsoft Text Analytic for determining positive, neutral, and negative sentiment. Before applying Microsoft Text Analytic, we conduct a pre-processing step to remove the unwanted data such as duplicate tweets, image, website address, etc.

Results: According to the Microsoft Text Analytic, the results are 666 positive sentiment numbers, 2055 neutral sentiment numbers, and 127 negative sentiment numbers.

Conclusion: According to these results, we conclude that most GO-JEK customers are satisfied with the GO-JEK services. In this research, we also develop classification model to predict the sentiment analysis of new data. We use some classifier algorithms such as Decision Tree, Naïve Bayes, Support Vector Machine and Neural Network. In the result, the system shows   that the decision tree provides the best performance.

I. INTRODUCTION

Predicting market demand and customer satisfaction requires an extensive market research to predict the market demand and customer satisfaction. Agile software development (ASD) published for helping engineer in business management problem, has been modified by applying it in the business intelligence field for producing the fast analytics in data science. Recently, business intelligence known as BI has become an important analytical procedure for analyzing market demand and satisfaction. BI is related as a technology used to analyze the business data regarding to the decision making process based on prediction analysis and provided the data enhancement process followed by an analytical technology which is widely used in several aspects such as market intelligence, e-commerce, e-health etc. [1, 2]. Currently, researchers have been accomplishing studies and concluding that BI affords two things which are process and product. Process defines as a procedure that is used to produce the important information regarding the business process. Product defines as information about empowering organizations to predict the service, customer satisfaction, competitor, marketplace, technology, etc. These two important things become a necessary part in market research and have a powerful analysis result to predict the market demand and customer satisfaction [1, 3].

Regardless of enthusiasm in the advantages of BI, conducting the BI analysis process needs a deep analysis. Sentiment analysis is one of approaches that is widely used to support the BI analysis process. Sentiment analysis is defined as a text data analysis that provides a deep analysis about sentiments, opinions, even expressed emotions and it allows us to predict the chances regarding to the analysis result [4].

Sentiment analysis is widely used in several cases, particularly in the review analytics cases, for example, in the analysis of praise and complaint sentences. According to the previous study, praise sentences indicate a positive subset

---

[*] Corresponding author

which consists of adjectives, intensifiers, etc. Meanwhile, the complaint sentences showed a negative subset. However, in the real world, praise and complaint sentences have been delivered in the complex form. Hence, a deep analysis such as sentiment analysis may help to resolve this problem [5]. Other cases, sentiment analysis also successfully gives an effective recommendation in analyzing the customer review about airport services [6, 7, 8]. Sentiment analysis also has a powerful performance for giving recommendation on travel destination [9, 10, 11, 12, 13, 14]. Considering these facts, it can be concluded that sentiment analysis is suitable for giving recommendation and it may also give a suitable result in predictive analytics for business intelligence. Despite the powerful performance of sentiment analysis, the complexity of data also should be considered. Twitter posts are one of sentiment data that have high complexity and widely used for sentiment analysis. Twitter is a social media that is often used to share information, news and opinions. Previous study has reported the number of Twitter user is rapidly growth up in the last decade which turned into a popular media to express the public opinions [15].

In this study, a business intelligence based on customer review analytics was proposed to investigate and analyze the business solution in GO-JEK application case. We want to analyze the performance of sentiment analysis in evaluating customer satisfaction with GO-JEK services. We also want to know how the customer review to the GO-JEK services and want to evaluate what the services that are satisfied enough and what the services that need to be improved. In the business process, these analyses might become a suggestion for improving the services. Hence, conducting sentiment analysis to evaluate the GO-JEK is important for the business process.

GO-JEK was chosen because several countries were against GO-JEK to be operated in their countries such as Malaysia. Several opinions, reactions and perceptions are popped up about this case. Hence, we performed predictive analytics for business intelligence based on customer reviews in order to analyze how the GO-JEK customer responses about their services and to investigate the chance of GO-JEK to be operated in Malaysia. For the dataset, we use Twitter posts considering the trend of Twitter usage for expressing public opinions.

To address these research gaps, we presented a sentiment analysis on the customer review collected by Twitter posts. To handle sentiment classification task, we can use available commercial cloud services tools. Some cloud services that have been employed by other researchers are Amazon Web Service [16, 17], IBM Cloud [16], Google Cloud Platform [16] and Microsoft Text Analytics [16, 18, 19]. However, Arijit and William in [18] found that Microsoft Text Analytics was 10-20% better in identifying positive and negative sentiment when evaluated in various tweet datasets. Furthermore, Qaisi and Aljarah wrote that Microsoft could benefit from Twitter page by growing promotion offers to the customers to enhance their loyalty and satisfaction [19]. Therefore, according to the evidence stated by previous works, we performed Microsoft Text Analytics for making a ground truth data. Moreover, machine learning classifiers were approached for predicting the customer satisfaction analysis. We also analyzed the business process according to sentiment analysis result for making market decision and increasing the services.

I. METHODS

This research is conducted in four processes as depicted in Fig. 1. These all steps are described in the following subsection.

A. Dataset

In this study we used Twitter posts collected from the Twint library. Data collection was done by mining data using the following rules:
1. Twitter posts are filtered only in the countries that allows GO-JEK to operate such as Singapore, Thailand and Vietnam.
2. Twitter posts are collected since January 1, 2019.
3. All Twitter posts that contains of GOJEK tweeted by the Twitter users in Singapore, Thailand and Vietnam.
4. All Twitter posts that contains both "GOJEK" and "Singapore" keywords.
5. All Twitter posts that contains both "GOJEK" and "Thailand" keywords.
6. All Twitter posts that contains both "GOJEK" and "Vietnam" keywords.

According to rules above, we got the data as shown in the Table 1.

TABLE 1.
DATASET

Country	Singapore	Thailand	Vietnam
Number of tweets	1948	704	459

B. Pre-processing

Since the dataset were collected from the Twitter post which consist of many varied data, we should conduct a pre-processing step to normalize the data. In the first step, we performed the tweet cleaning process. We removed the duplicate tweet and cleaned the URL and images that exist in tweet.
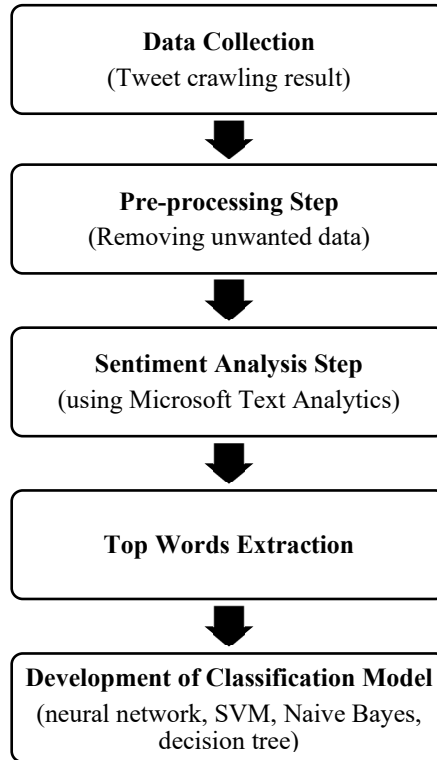
Data Collection
(Tweet crawling result)

↓

Pre-processing Step
(Removing unwanted data)

↓

Sentiment Analysis Step
(using Microsoft Text Analytics)

↓

Top Words Extraction

↓

Development of Classification Model
(neural network, SVM, Naive Bayes, decision tree)

Fig. 1.   Overall stages in this work.

C. Sentiment Analysis

After conducting pre-processing step, we worked an available premium API for text analytics namely Microsoft Text Analytics. This API worked by extracting a sentiment score from the given document ranged between 0 to 1 which 1 indicated the most positive sentiment and 0 indicated the most negative sentiment [20]. The analysis models used by the API were retrained using an extensive body of text and natural language technologies from Microsoft [20].

D. Top Words Extraction

This step was used to extract and analyze which features that have most number in negative and positive classes. We employed the combination of 1-gram and 2-gram token to split the tweet. This process split the sentences into one-word and two-words clause. The extracted word clauses represent the service. The top word clause in negative sentiment indicated that the services are not satisfy the customer and needs to be improved, while the top word clause in positive sentiment indicated that the service needs to be maintained.

E. Development of Classification Model

In this stage, we built a predictive model using machine learning that will be used by the businessman to classify new data. Before going through classification process, we balanced the data distribution of each class and the population that described in Table 2. After the process done, we split the sentence using tokenization technique and removed the stop words that exist in the tweet body. Furthermore, we employed the bag of words vector to convert

the tweet text into binary for the classification purpose. After that, we trained the data using four machine learning algorithms which are neural network, support vector machine, Naïve Bayes and decision tree.

Multinomial Naïve Bayes (MNB) extracted information without seeing the sentence many times in a document which make this method to be one of the best models to solve problems in text classification [21]. Furthermore, the benefit of using Naive Bayes was that this method only requires a small amount of data training to estimate the parameters in classification. It helped to determine exact probabilities for hypothesis, and it is robust to noise in input data [22]. Support vector machine (SVM) is a classifier algorithm that often used in both classification and regression. It worked by finding a hyper plane on the input, where the best hyper plane is the one which located in the middle between two sets of classes, look for the best hyper plane equivalent to maximize the margin or distance between two sets of objects from different classes [23]. Neural networks have been employed in handling document classification task. This algorithm reaches good results on complex domains and allows performing fast testing. However, the training process is very slow [24]. Neural network methods are generally better compared to other traditional machine learning algorithms, and it is effectively extracting the semantic features of text [25]. Tree-based classifiers like decision tree algorithm have also been studied for text analysis task, for example document categorization [26]. This algorithm is used effectively in various fields for classification purpose [27], and it is a very fast algorithm for prediction and learning task [28]. Furthermore, this algorithm has been proofed for its robustness in handling categorical data [29]. Considering the robustness of the four mentioned classifiers, we used these all classifier and evaluated each performance.

TABLE 2
DATA DISTRIBUTION FOR CLASSIFICATION PROCESS

Data	Total
Negative	127
Neutral	127
Positive	127

II. RESULTS

A. Pre-processing and Sentiment Analysis

We collected 3111 tweets from January 1, 2019 that contain keyword of "Gojek" and "Singapore", "Gojek" and "Vietnam", and keyword "Gojek" and "Thailand". After collecting the data, we perform the pre-processing step by removing any unwanted data. The pre-processing step successfully removed 163 which were duplicated tweets. In the next step, we perform Microsoft Text Analytics API to give label for each data. This process obtained 2055 neutral sentiments, 666 positive sentiments and 127 negative sentiments considering the number of sentiment score provided by Microsoft Text Analytics API. These results are illustrated in Fig. 2. An example of labelled tweet is drawn in Table 3. The sentiment score of each label provided by Microsoft Text Analytics API is drawn in Table 4.

TABLE 3
EXAMPLE OF LABELLED TWEET

Tweet	Score	Label
Fintech Viral GOJEK Cs Claimed Singapore Unicorn, BKPM Boss Apologize 30 July 2019 22:20 - CNBC Indonesia	0.5	neutral
Honestly I hate GOJEK drivers. They ALWAYS cancel on me. Little shits.	0.014	negative
In Indonesia grab is cheaper. While in Singapore GOJEK is cheaper. Interesting.	0.818	positive

TABLE 4
SENTIMENT SCORE FOR LABELLING PROCESS

Label	Score
Negative	0 to less than 0.5
Neutral	Exactly 0.5
Positive	More than 0.5 to 1

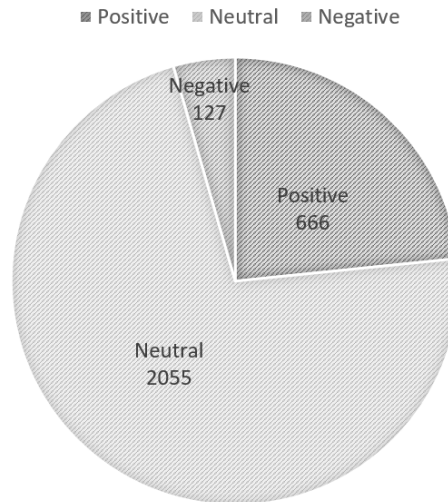■ Positive ■ Neutral ■ Negative

Fig. 2. Sentiment population.

B. Top Words Extraction

In this stage, we extracted the top positive and negative sentiment. This extraction process was done by determining the highest number of sentiments in both positive and negative sentiment. Negative sentiment was extracted to determine which services that were often complained by customer. Positive sentiment was extracted to determine which services that should be maintained. Results of this step are illustrated in Table 5 and Table 6.

Table 5 shows the top three keywords indicating the negative sentiment. However, in the result, we found keywords that ambiguous, for example "via YouTube", "like", "di", therefore we dropped them.

TABLE 5
TOP NEGATIVE SENTIMENT

Keyword	Score
Cancel	9
Hate driver	3
'tidak bisa'	3

TABLE 6
TOP POSITIVE SENTIMENT

Keyword	Score
Ride hailing	50
Available consumer	37
Consumer commodities	33

C. Development of Classification Model

After giving label in each data and analyzing them, we used the labeled data for developing the classification model. This model was used to predict the new data. We applied machine learning approach to develop the best classification model. In this research work, we applied four machine learning algorithm which are neural network, support vector machine, Naïve Bayes and decision tree. To evaluate the performance of each model, we used precision, recall and f1-score calculation. These parameters were often used to validate the performance of classification model and have good performance in text classification cases.

In our experimental result, we achieved performance evaluation as shown in Fig. 3. The comparison results of each algorithm also shown in Fig. 3. From the performance evaluation result, neural network algorithm achieves precision of 0.52, recall of 0.51 and f1-score of 0.51. While support vector machine obtains precision of 0.54, recall of 0.53 and f1-score of 0.53. Naïve Bayes obtains precision of 0.52, recall of 0.52 and f1-score of 0.52. Finally, decision tree obtains precision of 0.55, recall of 0.55 and f1-score of 0.55.
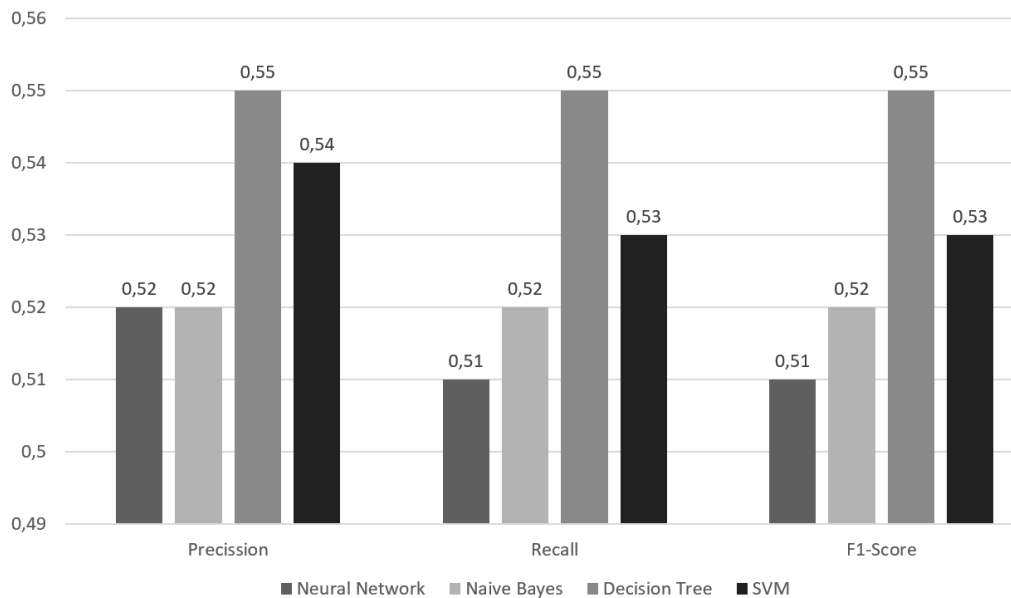
Fig. 3. Classification result of each model.

III. Discussion

According to the sentiment result depicted in Fig. 2, we conclude that the most sentiment number is neutral sentiment. Furthermore, Table 5 indicates some GO-JEK services that need to be improved, for example their driver's performance. On the other hand, regarding to the positive sentiment, GO-JEK services related with 'ride hailing', 'available consumer' and 'consumer commodities' should be maintained to provide more customer satisfaction as shown in Table 5. However, in the top words of negative sentiment, there is 'tidak bisa', which is a non-English sentence. This is caused by Indonesian users tweeting from Singapore about Gojek. For example, tweet from user with username @norikadwight: "@gojekindonesia sore, saya ada notif untuk paylater nya gojek. Saya mau bayar tapi tidak bisa karna saya tidak bisa akses akun, saya posisi di singapore sekarang" and from @deeadrianto: "@gojekindonesia gojek sudah masuk ke Singapore. Tapi kenapa bagi pengguna gojek indonesia apps gojek tidak bisa langsung dipakai di sana, diharuskan memasukkan DBS card?". In the future, it is needed to remove the non-English tweets, for example by checking the user's region in the users metadata.

According to the experimental result of sentiment analysis described in the section C, we conclude that Microsoft Text Analytics was very useful to give data label since unlabeled data were much published. This tool is proper to be used in further research which does not have a ground truth data. However, we did not perform a manual rechecking for the classification result from Microsoft Text Analytics.

Regarding to the classification result, it can be concluded that decision tree algorithm obtains the highest score in precision, recall and f1-score with 0.55, 0.55 and 0.55, respectively. The performance results illustrated in Fig. 3 seems small caused by the limitation of the dataset that we had. For further research, increasing the number of data is needed to develop a better model of classification. Prediction system would be a robust system if their model has been extracted varied data. Hence, a big number of data might improve the performance result. This procedure might improve the performance evaluation of classification model.

IV. Conclusions

Business intelligence is an important analysis in business process. In this research work, we proposed a business intelligence analysis by using sentiment analysis method based on customer reviews. We choose GO-JEK case study in our research. Our proposed sentiment analysis scheme consists of pre-processing, sentiment analysis, top words extraction, and development of classification model. In the experimental result, we achieve positive sentiment number of 666, neutral sentiment number of 2055 and negative sentiment number of 127. According to these results, we conclude that most of GO-JEK customers were satisfied with the GO-JEK services considering the number of neutral

sentiments. In this research work, we also developed a classification algorithm to predict the sentiment analysis of new data. We used neural network, Naïve Bayes, decision tree and support vector machine as classifier. Our experimental research concludes that decision tree is the best classifier. For the further research, it is necessary to conduct balancing data method to improve the performance of classification model. Furthermore, in the keyword extraction stage, we still found some keywords that did not represent the services of the company. Therefore, we suggest for the future research to get the list of services provided by the company in order to tackle the ambiguous result in top words extraction.

REFERENCES

[1] P. Wazurkar, R. S. Bhadoria and D. Bajpai, "Predictive analytics in data science for business intelligence solutions," in *7th International Conference on Communication Systems and Network Technologies (CSNT)*, Nagpur, 2017.

[2] G. Shmueli and O. Koppius, "Predictive analytics in information systems research," *MIS Quarterly,* vol. 35, no. 3, pp. 553-572, 2011.

[3] R. Vedder, M. Vanecek, C. Guynes and J. Cappel, "CEO and CIO Perspectives on Competitive Intelligence," *Communications of the ACM,* vol. 42, no. 8, pp. 108-116, 1999.

[4] L. Zhang and B. Liu, "Sentiment Analysis and Opinion Mining," in *Encyclopedia of Machine Learning and Data Mining*, Boston, Springer, 2017.

[5] Ganesan, Kavita and G. Zhou, "Linguistic Understanding of Complaints and Praises in User Reviews," in *NAACL-HLT*, 2016.

[6] J. Suter, "Airport image: An exploratory study of McCarran International Airport," *Journal of Air Transport Management,* vol. 67, pp. 72-84, 2018.

[7] N. Haghighat, "Airline service quality evaluation: A review of concepts and models," *Economics, Management, and Sustainability,* vol. 2, no. 2, pp. 31-47, 2017.

[8] V. Bogicevic, W. Yang, A. Bilgihan and Bujisic, "Airport service quality drivers of passenger satisfaction," *Tourism Rev.,* vol. 68, no. 4, p. 3–18, 2013.

[9] Q. Ye, Z. Z. and L. B., "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Systems with Applications,* vol. 36, p. 6527–6535, 2009.

[10] C. V., W. R and S. Purwanto, "Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film," *Journal of Intelligent Systems,* vol. 1, no. 1, pp. 55-59, 2015.

[11] S. Tan and J. Zhang, "An empirical study of sentiment analysis for chinese documents," *Expert Systems with Applications,* vol. 34, no. 4, p. 2622–2629, 2015.

[12] S. Lidya, O. Sitompul and S. Efendi, "Sentimen analysis pada teks bahasa Indonesia menggunakan Support Vector Machine (SVM) dan KNearest Neighbor (K-NN)," in *Seminar Nasional Teknologi Informasi dan Komunikasi*, 2015.

[13] A. Balahur, "Sentiment Analysis in Social Media Texts," in *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Atlanta, Georgia, 2013.

[14] C. Bhadane, H. Dalal and H. Doshi, "Sentiment analysis: Measuring opinions," *Procedia - Procedia Computer Science,* vol. 45, p. 808–814, 2015.

[15] A. Lestari and D. Karolita, "Summarizing Netizens' Sentiments Towards the 1st Indonesian Presidential Debate using Lexicon Sentiment Analysis," in *9th Annual Basic Science International Conference 2019 (BaSIC 2019)*, 2019.

[16] A. Koneru, N. B. N. S. R. Bhavani, K. P. Rao, G. S. Prakash, I. P. Kumar and V. V. Kumar, "Sentiment Analysis on top five Cloud Service Providers in the Market," in *2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018)*, 2018.

[17] D. A. Ivor, D. Do, R. Ge and S. I. Ahamed, "A Reference Architecture for Social Media Intelligence Applications in the Cloud," in *IEEE 39th Annual International Computers, Software & Applications Conference*, 2015.

[18] C. Arijit and W. Perrizo, "Investor Classification and Sentiment Analysis," in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2016.

[19] L. M. Qaisi and I. Aljarah, "A Twitter Sentiment Analysis for Cloud Providers: A Case Study of Azure vs. AWS," in *7th International Conference on Computer Science and Information Technology (CSTT)*, 2016.

[20] Microsoft Azure, "What is Text Analytics API?," [Online]. Available: https://docs.microsoft.com/en-us/azure/cognitive-services/text-analytics/overview. [Accessed 11 December 2019].

[21] A. I, S. D, N. F and H. A. N., "Utilizing hashtags for sentiment analysis of tweets in the political domain," in *9th Int. Conf. on Machine Learning and Computing - ICMLC 2017*, 2017.

[22] J.Ren, S.D.Lee, X.Chen, B.Kao, R.Cheng and D.Cheung, "Naive Bayes Classification of Uncertain Data," in *Ninth IEEE International Conference on Data Mining (ICDM '09)*, 2009.

[23] M. O. Pratama, W. Satyawan, R. Jannati, B. Pamungkas, Raspiani, M. E. Syahputra and I. Neforawati, "The sentiment analysis of Indonesia commuter line using machine learning based on twitter data," in *International Conference of Computer and Informatics Engineering (IC2IE)*, 2019.

[24] G. Kaur and K. Bajaj, "News Classification and Its Techniques: A Review," *IOSR Journal of Computer Engineering (IOSR-JCE),* vol. 18, no. 1, pp. 22-26, 2016.

[25] C. Hong, L. Ziwei, Y. Cuitin, L. Jie and Y. Air, "A Shallow Neural Network based Short Text Classifier for Medical Community Question Answering System," in *Proceedings of 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems*, 2018.

[26] X. B., G. X., Y. Ye and J. Cheng, "An Improved Random Forest Classifier for Text Categorization," *JCP,* vol. 7, p. 2913–2920, 2012.

[27] S. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst. Man Cybern.,* vol. 21, p. 660–674, 1991.

[28] C. Giovanelli, X. Liu, S. Sierla, V. Vyatkin and R. Ichise, "Towards an aggregator that exploits big data to bid on frequency containment reserve market," in *43rd Annual Conference of the IEEE Industrial Electronics Society (IECON 2017)*, Beijing, China, 2017.

[29] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes and D. Brown, "Text Classification Algorithms: A Survey," *Information,* vol. 10, no. 4, p. 150, 2019.