ORIGINAL ARTICLE

CUSTOMER SENTIMENT ANALYSIS THROUGH SOCIAL MEDIA FEEDBACK: A CASE STUDY ON TELECOMMUNICATION COMPANY

Siti Nur Syamimi Mat Zain[1], Nor Azuana Ramli[1], Rose Adzreen Adnan[2]

[1]Centre for Mathematical Sciences, College of Computing & Applied Sciences, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.
[2]Credence Tech, Petaling Jaya, Selangor, Malaysia.

ABSTRACT – Customer sentiment analysis is an automated way of detecting sentiments in online interactions in order to assess customer opinions about a product, brand or service. It assists companies in gaining insights and efficiently responding to their customers. This study presents a machine learning approach to analyse how sentiment analysis detects positive and negative feedback about a telecommunication company's products. Customer feedback data were taken from Twitter through Streaming API (Application Programming Interface), where Tweets are retrieved in real time based on search terms, time, users and likes. Responses from the twitter API are parsed into tables and stored in a CSV file. Based on the analysis, it was found that there was no negative sentiment from the customers. The data were then split into training and testing to be tested on the three different supervised learning algorithms used in this study which are Support Vector Machine, Random Forest, and Naïve Bayes. Lasty, the performance of each model was compared to select the most accurate model and from the analysis, it can be concluded that Support Vector Machine gives the best performance in terms of accuracy, Mean Squared Error, Root Mean Squared Error and Area Under the ROC curve.

INTRODUCTION

In today's digital era, people use social media not just to make new connections or reunite with old friends but to share information and get a variety of information as well. According to the Statista Research Department, social media usage is one of the most popular online activities. Over 3.6 billion people used social media in 2020, with that number expected to rise over 4.41 billion by 2025. With these facts, companies are attempting to accommodate these growing trends to spread their corporate values, such as driving customer business behaviour, sustaining customer loyalty, increasing sales and revenue, increasing consumer satisfaction, and enhancing brand awareness and reputation [1].

With global Internet access, a massive quantity of data is generated, providing a potential way to learn about customer feedback on previously purchased and used items. Companies want to make use of this data and turn it into useful knowledge that will help them make better decisions, which could be done by analysing all accessible data. This type of analysis might help companies not in just determining what customers want, but also in designing new products and improving existing ones on the market.

Within this context, sentiment analysis techniques are a useful strategy to study. Sentiment analysis (SA), also known as opinion mining, is the study of people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions concerning things such as products, services, companies, persons, issues, events, and subjects, as well as their attributes. SA is a method of extracting, converting, and interpreting opinions from a text and classifying them as positive, negative, or natural sentiments using Natural Language Processing (NLP) [2].

SA encompasses psychology, sociology, natural language processing, and machine learning. More advanced forms of analytics have recently been enabled by the exponentially rising volumes of data and computational power. Hence, machine learning has become a common approach for SA [3]. The approach involved using a variety of machine learning classifiers and feature extractors. Naive Bayes, k-Nearest Neighbour (k-NN), decision tree and Support Vector Machines (SVM) are examples of machine learning classifiers while feature extractors including unigrams, bigrams, unigrams plus bigrams, and unigrams with part of speech tags.

Although with all these methods, it is not easy to obtain the data into an understandable format as the volume of data grows, extracting sentences, reading, analysing, summarizing, and organizing them becomes more challenging. Moreover, it is difficult to specifically assess what truly is negative, positive, and neutral on particular words used. This causes the accuracy of sentiment analysis to be reduced when the sentences are complex and too lengthy.

A positive and negative sentiment score is important, as it may assist the company to discover where they must improve and when they can continue business as usual. Thus, if the company unable to understand and measure their customer's emotions and behaviour in details, it might not only negatively impact existing customer relationships, but also drops the company's revenue. Therefore, in this paper, a case study is conducted to analyse if there is any negative feedback from telecommunication company's customers through a social media platform which is Twitter. Other

objectives of this study including to conduct customer sentiment analysis by using machine learning algorithms such as SVM, Random Forest and Naïve Bayes and comparison between several different machine learning algorithms will be conducted to determine the best algorithm. At the end of this study, it is important to achieve all these objectives to ensure that the telecommunication company can meet market demand and together provide the best product and service to the world, especially in Malaysia.

The research scope in this study is focusing on the telecommunication company's customers who are using social media platforms as a medium to give their feedback. Majority of customers leave data points in the form of reviews, tweets, and comments to express their emotions towards the product or services. This study is to keep track of what customers were saying on Twitter about the company's products, and thus allow the company to understand customer sentiments towards their products. The emotions of customers are not always expressed in formal words, they might use emoticons, slang, short-form, images and others. Thus, the reviews, tweets and comments produced are in unstructured data and become a limitation in this study as it will reduce the accuracy of the sentiment analysis. In addition, data extraction from Twitter is also limited as the standard API allows retrieving tweets within the last 7 days only. Therefore, the extraction process to get more tweets needs to be done every week.

RELATED STUDIES

Sentiment analysis in the business world, is the act of recognising and indexing a piece of text based on its tone. The text might be in the form of tweets, comments, feedback, or even spontaneous rants with positive, negative, or neutral views. This automated sentiment analysis has been implemented in every business, with the investigation of public opinion toward brands, products, and service. Oktaviani et al. has proposed sentiment analysis of e-commerce applications in Traveloka data review on Google Play site to discover customer perceptions of service quality. Traveloka is a service for hotel reservation and purchasing transportation tickets. The study has gathered review data from September 2019 to November 2019. The machine learning used in classification was Naïve Bayes which provided 91.2% accuracy, and the best kappa accuracy is 59.56% [4].

Based on the previous study regarding sentiment analysis of tweets connected to three major Turkish telecommunications firms, Turk Telekom (TT), Turkcell, and Vodafone Turkey, Kündüm et al. proposed on supervised learning technique by analyzing Support Vector Machine, Decision Tree, Random Forest, Naïve Bayes, Multilayer Perceptron and k-NN. The study has gathered comments on Twitter data between 1 December 2019 until 1 September 2019. Pre-processing models such as removing punctuation marks, eliminating stop-words, removing tags, filtering URLs, and stemming are used to improve the performance of the system. Overall, the result suggested that the Random Forest model managed to give the best classification accuracy with over 80% for all the telecom operators [5].

Research by [6] proposed on sentiment analysis of tweets written in English belonging to Saudi Arabia telecommunication companies (Mobily, STC, and Zain). The purpose of the study is to extract sentiment, which may be positive, negative, or neutral, from Twitter data, which can then be used to assist define policies and provide better services. The study applied several machine learning algorithms such as Artificial Neural Networks (ANN), k-NN and Naive Bayesian. It is used for classification, and the best results were obtained by k-NN, with F-measure of 75.6%. Additionally, k-NN was also employed with various metrics including Euclidean distance and cosine similarity. The results achieved using cosine were quite better than those obtained with its counterpart. It is emphasised that increasing the value of k has a positive impact on the accuracy of certain algorithms. The k value, however, must not be so large that it contains noise points or points from a neighbouring class [7].

Meanwhile, in another study in Indonesia, sentiment analysis on social media was used by several data service operators to determine the level of public satisfaction with their data services for internet access. The data was gathered through official accounts from four providers by using Twitter's official API, from December 2017 to March 2018. Fitri et al. proposed this study by using the Naïve Bayes algorithm [8]. To get the best accuracy result, performance testing is done using k-fold Cross Validation, which involves performing the test four times with the value of k varying each time: 5-fold, 10-fold, 15-fold, and 20-fold. The greater the k-fold value, the more exact the accuracy rating. Therefore, the best k-fold value is 20-fold, which has a 94.17 % accuracy, whereas system performance testing yielded an average of 94.5% precision, 93.31% recall, F1-score of 93.15% and accuracy of 99.09%.

METHOD & MATERIALS

First step in the customer sentiment analysis process starts with collecting data from social media which in this case study, the social media platform that has been chosen is Twitter. Customer feedback data were taken using a streaming API that is provided by Twitter platforms using a hashtag. Data from Twitter was selected in this study due to its higher popularity compared to other platforms and the data available is free to the public. The next step in the analysis is pre-processing the data. This pre-processing stage assists in the conversion of noise from high-dimensional features to low-dimensional spaces, enabling the extraction of as much accurate data from the text as feasible. There are many phases involved in data pre-processing which are tokenization, stopwords, normalization, POS tagging, stemming and lemmatization. Then, feature extraction was performed.

Several approaches that can be used are CountVectorizer and TFIDF (term frequency-inverse document frequency). Count-Vectorization is a method for counting all the corpus words and turning them into document representations. This approach merely counts the number of times a word appears. Meanwhile, the TF-IDF method is used for weighting of each regularly used word. The goal of this method is to compute the Term Frequency (TF) and Inverse Document

Frequency (IDF) of each token (word) in the corpus. The TF-IDF method, on the other hand, is used to determine how frequently a word appears.

Once tweets are converted to numbers, one of the best approaches to adding labels to unlabeled tweets is to use the TextBlob library in Python. Three new columns were added to the Twitter dataset namely Subjectivity, Polarity and Sentiment. Sentiment score calculations were made on columns containing text data, using the getSubjectivity, getPolarity and getSentiment functions. Sentiment labels were classified according to sentiment score, where, if the sentiment score is less than 0, then it is classified as Negative, and if the sentiment score is more than 0, then it is categorized as Positive, otherwise it is Neutral. Based on the sentiment label, the frequency of each sentiment can be calculated and will be illustrated in pie charts and bar charts.

As a part of the machine learning process, the next step is to train the Twitter data using a machine learning algorithm. Before being assigned, the data must be labelled first. The labelled sentiment data will then be used to measure the accuracy of the model by comparing the correctly classified sentences with the falsely classified data which in turn helps to tune the performance of the machine learning model. This sentiment data will be split into training and test dataset and will be trained using machine learning algorithms. Three supervised machine learning algorithms that were applied in this study are SVM, Random Forest, and Naive Bayes. Lastly, models were compared based on their accuracy, precision, recall and F1 score, Mean Square Error (MSE) and Root Mean Square Error (RMSE) in order to select the best algorithm. Figure 1 shows the flow of this research:
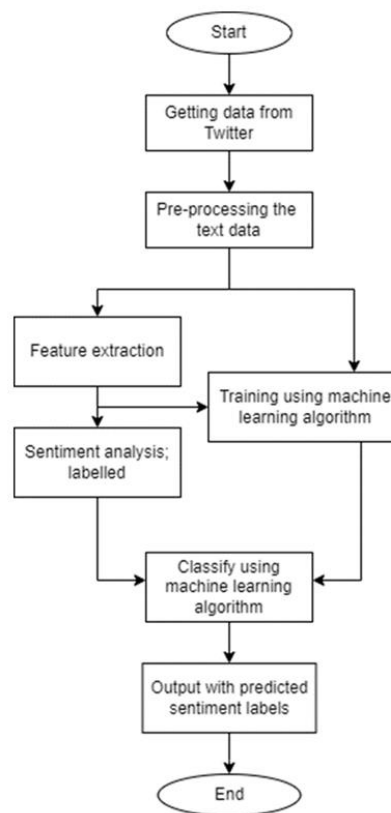
Figure 1. Process flowchart of the research framework.

Development of Predictive Modelling

In developing the predictive model, this study will build several supervised machine learning algorithms in python. Three types of machine learning algorithms involved are Support Vector Machine (SVM), Random Forest and Naïve Bayes. These models were selected based on the previous literature where their performance always stood out compared to other models.

1) Support Vector Machine (SVM)

SVM are learning algorithms that learn to interpret patterns from data. It is employed in the solving of classification issues with large datasets. The SVM model focuses on identifying hyperplanes that divide data points into as many different classes as possible, so that new data are classified according to which part of the gap they land on. Generally, hyperplanes are decision boundaries that assist in data classification. It also searches for the optimal surface for separating positive, negative, and neutral training data. There are several hyperplanes to choose from to separate the two classes of data points. The purpose is to locate the plane with the largest margin, or the distance between two classes' data points. Increasing the maximum margin distance provides more robustness, so that future data points can be classified more accurately.

Before training the model, hyperparameters need to be set. In terms of developing a robust and accurate model, hyperparameters are needed as it assists in finding the right combination of bias and variance, as well as preventing the model from overfitting. Bias-variance problem is commonly encountered in supervised learning algorithms. As a result of a bias error, underfitting will occur because it misses critical interactions between features and classes. Variance error, on the other hand, leads to overfitting because it is very sensitive to noise and changes in the training set.

Therefore, the parameters C, kernel, and gamma must be tuned to get the best accuracy in the SVM model. The regularization parameter C determines the exchange between the slack variable's penalty (misclassification) and the margin width. Meanwhile, kernel can turn data points that are linearly inseparable into data points that are linearly separable. It is used as a metric in the equation. The input is the original feature, and the output is a measure of the equation in the new feature space. One of the kernel functions that is often used is the radial base function (RBF). For the Gamma parameter, it needs to be optimized simultaneously with the C parameter if using the RBF kernel. However, if the kernel is linear, only the C parameter needs to be optimized. Therefore, since these parameters are so important to a model's accuracy, these grid-search approach is employed to identify the best values.

2) Random Forest

The next prediction model for customer sentiment analysis is Random Forest. Random forest is a classification technique that uses ensemble learning. It classifies based on the outputs of numerous decision trees formed during the training, with the forest output being the targeted output from each decision tree. Random Forest tends to reduce numerous decision trees since trees are known to overfit data due to their low bias and high variance. In classification trees, the most important features are reflected in the tree's nodes, with leaves indicating class labels and the most critical traits being high up in the tree. The importance of a feature is determined by the Gini Index, where the smaller the decrease in accuracy by modifying the value of a feature randomly, the lower the importance of the feature [9].

In Random Forest algorithm, parameters are implemented to improve the model's prediction accuracy as well as make it easier to train. Several parameters which will be used in this study are n_estimators, max_depth, Random_state and max_features. n_estimator is the number of trees to be built before selecting the maximum rating or average predictions, max_depth of a tree is specified as the longest path between the root node and the leaf node, Random_state makes it easy to replicate as if given with the same training parameters and data, it will always produce the same results, and lastly, max_features is the maximum number of features considered for splitting a node.

However, all values for this parameter are not set because in the analysis of this study, the grid search method will be used in the Random Forest model to find the optimal value for each parameter using the best_params function.

3) Naïve Bayes

The third methodology in this study for customer sentiment analysis is by using Naïve Bayes. Naïve Bayes is a statistical classifier that is used to predict the likelihood of a class's membership. At the classification stage, the document category value is determined based on the terms that occur in the classified document. Bayes' theorem provides a method of calculating the posterior probabilities of $P(c|x)$ from $P(c)$, $P(x)$ and $P(x|c)$ as below:

$$P(c|x) = \frac{P(x|c)\, P(c)}{P(x)} \tag{1}$$

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \ldots \times P(x_n|c) \times P(c) \tag{2}$$

where $P(c|x)$ is the posterior probability of class, c is the target given predictor, x is the attribute, $P(c)$ is the class's prior probability, $P(x|c)$ is the likelihood which is the probability of predictor given class, and $P(x)$ is the predictor's prior probability.

For this study, a Multinomial Naïve Bayes will be chosen as it tends to be the basic solution for the sentiment analysis task. The Bayes theorem is used in this method to predict text tags. It will compute each tag's likelihood for a given sample and then extract the tag with the highest probability.

Performance Evaluation

After a predictive model using each machine learning algorithm is developed and a test dataset is generated, the model will then be assessed to determine its performance. The prediction consists of a True Positive Reviews (TP), False Positive Reviews (FP), True Negative Reviews (TN), and False Negative Reviews (FN). TP indicates the number of samples that the model correctly classified as positive class, TN indicates the number of samples that the model correctly classified as negative class, FP indicates the number of samples that the model incorrectly classified as positive class, and FN indicates the number of samples that the model incorrectly classified as negative class. All values of TP, FP, TN, and FN can be found using the confusion_matrix function, which can be clearly seen in the confusion matrix boxes in Figure 2.
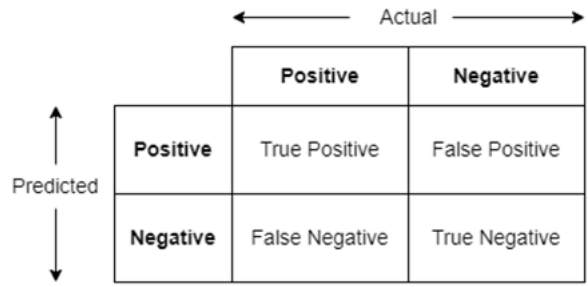
Figure 2. Confusion matrix.

From the confusion matrix, the performance of the model can be measured using accuracy, precision, recall dan F1-measure and through Python, all the measurements can be obtained using classification_report function. However, in evaluating machine learning models, accuracy is not the only performance parameter need to be considered. Performance of the models will also be measured using RMSE (Root Mean Square Error) and AUC (Area Under the Curve). RMSE is a measure to confirm how close an estimate or forecast is to the actual value. The model can be considered as a good model if it has the lowest RMSE value.

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(Predicted - Actual)^2}{N}} \tag{3}$$

As for AUC, it is a metric for a classifier's ability to distinguishes between positive and negative classes. If AUC value is equal to 1, the classifier is capable of correctly differentiate between all Positive and Negative class points. On the other hand, if AUC value is 0, the classifier expects all Negatives to be Positives and all Positives to be Negatives.

RESULTS AND DISCUSSION

All the results obtained in this study will be discussed under this section. First, the discussion starts with data analysis.

Data Extraction from Twitter

Data collection involved gathering of relevant tweets regarding on a particular topic. In this study, data was gathered from Twitter, using the hashtag on Twitter's streaming API, for a specified time of analysis. The tweets were extracted weekly from 15th March 2022 to 20th April 2022. The number of tweets extracted is limited by Twitter's rate limit. During this time, 33 tweets from the hashtag were extracted. This extracted text format was then converted to a CSV file. The raw data extracted from Twitter contains 32 rows and 5 attributes, consisting of Unnamed: 0, Time, User, Tweets and Likes.

Data Pre-processing

Pre-processing of data is an important process in machine learning where it has prepared the tweets for classification in sentiment analysis. Due to the widespread use of machine learning models, the best model performance cannot be guaranteed simply by using big data. This is because labelling has affected the performance of the model. Therefore, the data was pre-processed to obtain high quality data. The processes that have been involved were conversion to lowercase, removing URLs (Uniform Resource Locators), removing stop words, removing short words (eliminate words with fewer than three letters that had no significance), tokenization (breaks raw text into small chunks) and lemmatization (step-by-step approach of determining the fundamental form of a word).

Proportion of Sentiment

After analyzing the twitter data through pre-processing, each word in the tweet was converted to a numerical representation using the bag of words method. With this method, polarity values can be obtained. A sentiment value can be obtained using the function getSentiment, which then determines whether it is positive, negative, or neutral. Based on the pie chart output from Figure 3, the proportion of sentiment consists of positive and neutral tweets only. There were 40.6% positive and 59.4% neutral. This means that most of the telecommunication company's customers are very satisfied with the products and services provided.
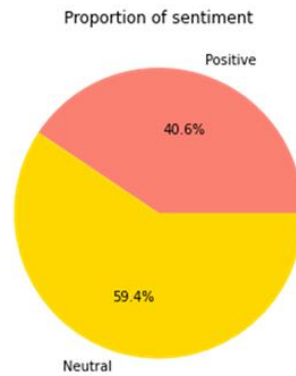
Proportion of sentiment

Figure 3. Proportion of sentiment.

Model Development using Machine Learning Algorithms

1) Support Vector Machine (SVM)

The first prediction was conducted using a SVM to identify hyperplanes that divide data points into as many different classes as possible, so that new data are classified according to the part of the gap in which they land. In the implementation of SVM, the kernel has been set as linear, while the C and gamma parameters were obtained through 'GridSearchCV'. 'GridSearchCV' will fit all the parameter values on the train set, which was then evaluated and yields the best parameter combination. As a result of the fitting on the test set for the SVM model, the best parameter combinations are 10 and 1, for C and gamma, respectively. In classification, a lower value of C will encourage a larger margin size, hence it avoids misclassification. SVM executes supervised learning and provides a sample of data in a high-dimensional space [10].

At the beginning of the study, it was observed that the SVM model accurately predicted 60% of the customer sentiments based on the text in tweets, while correctly predicted 56% neutral sentiments and 100% positive sentiments. However, as the best combination parameters are tuned with GridSearchCV, the accuracy increases where the SVM model accurately predicts 80% of customer sentiment based on the text in the tweet. Moreover, the MSE and RMSE results show that the model has an error rate of 0.2 and 0.4472, respectively. This indicates that the SVM model has a good prediction due to the low MSE and RMSE values.

2) Random Forest

The analysis was then conducted using the Random Forest model which was classified based on the output of various decision trees formed during training, with forest output being the target output from each decision tree. For this study, the model used GridSearchCV function, a library function in sklearn model_selection in which to select the best hyperparameters that provided the most accurate prediction. Initially, the range for each parameter max_depth, max_features and n_estimator was defined and became the input for GridSearchCV. While n_jobs, the number of cores to be used in model evaluation has been set as 1. After fitting GridSearchCV, the value of the best_params_ attribute for the Random Forest model will be returned. For this model, the best parameter combinations are 4, 2, and 3 for parameter max_depth, max_features, and n_estimator, respectively. The value will then be re-run in the RandomForestClassifier for evaluation purposes.

From the initial evaluation of model performance, Random Forest accurately predicted 60% of the customer sentiments based on the text in tweets. The model also correctly predicted 56% neutral sentiments and 100% positive sentiments. However, the performance accuracy of the model increased by 10% after the hyperparameter was tuned. Thus, Random Forest managed to predict 70% of customer sentiment accurately. Meanwhile, the results of the MSE and RMSE evaluations show that the Random Forest model has error rates of 0.3 and 0.5477, respectively.

3) Naïve Bayes

The next analysis was conducted by using the Naïve Bayes model in which it employs Bayes theorem to predict the probability of a given collection of characteristics as part of any label. Naïve Bayes discovers the class probabilities assigned to a text by using the probability of a combination of words as well as classes [11]. It is ideally used when the size of the training data is small, especially in this study where the Twitter data collected was only 32 rows. In the implementation of Naïve Bayes, the Multinomial Naïve Bayes approach was chosen since it is widely used to solve sentiment analysis problems. Multinomial works with integers that have been generated as frequencies for each word, hence it is good at classifying text. The model has also used GridSearchCV function in finding the best hyperparameters to improve its accuracy. Alpha parameter was used to control the formation of the model. By using GridSearchCV, alpha parameters ranging from 1.0 to 5.0 were fitted into the train set. Thus give the alpha values as 1.0.

Based on an initial evaluation of the Naïve Bayes model, it was observed that the model accurately predicted 70% of the customer sentiments based on the text in tweets. The model also correctly predicted 62% neutral sentiments and 100% positive sentiments. However, once the parameter was fitted, it was observed that the accuracy values obtained did not change from the previous evaluation despite been carried out several times. Thus, Naive Bayes model maintains 70% of accuracy. Finally, for the MSE and RMSE evaluation results show that Naive Bayes model has error rates of 0.3 and 0.5477, respectively, equivalent to the result of the Random Forest model.

Comparison of Model Performance

To select the best machine learning model, there are several conditions that need to be fulfilled. Among the conditions is to choose a model that has high Area Under Curve (AUC) value, low MSE, low RMSE, and high accuracy value. For AUC, the higher it is, the better the model's ability in classifying between positive and neutral classes. For MSE, the lower it is, the better the model evaluates how near predictions are to the actual value. Meanwhile, for accuracy, higher accuracy means the model can predict all labels correctly.

Table 1. Comparison of AUC, MSE, RMSE, and accuracy.

Model	AUC	MSE	RMSE	Accuracy
Support Vector Machine	0.8	0.2	0.4472	0.8
Random Forest	0.7	0.3	0.5477	0.6
Naïve Bayes	0.7	0.3	0.5477	0.7

Table 1 above shows a comparison of the performance results of the models. From the table, it has been observed that based on the three models, the SVM model meets all the conditions in the selection of the best model. This is due to the SVM model has the highest AUC value of 0.8, the lowest MSE value of 0.2, the lowest RMSE value of 0.4472, and the highest Accuracy value of 0.8.

Comparison with Previous Studies

Analysing the results of this study has shown significant findings in terms of performance comparison between SVM, Random Forest and Naive Bayes in customer sentiment analysis through social media feedback. Furthermore, the process of developing the machine learning model itself has showcased the Jupyter Notebook's capabilities in data pre-processing and prediction of the model. Previous literature reviews have provided preliminary conclusions about the performance of each machine learning model.

One of the studies conducted by [12] has concluded Random Forest is the best model for sentiment analysis for customer reviews in Telecom Operators in Turkey. The input data used were Turkish customer comments from Twitter which were collected with the label of each name operator. In the data collection method, the authors used the Selenium browser, in contrast to this current study since customer data was collected using the Twitter API. However, the Twitter API has its own limitations because data retrieval is limited. Specification on the algorithm used was SVM, K-Nearest Neighbors, Naïve Bayes, Decision Tree, Random Forest, and Multilayer Perceptron as an artificial neural network. The result of the research shows that Random Forest model outperforms all models.

Case study also had been conducted by [13] on Twitter data to obtain views from the public. This study only focuses on the SVM as a classification algorithm and unigram as an extraction feature. The results from the study showed that the algorithm achieved an accuracy of 87%. Nevertheless, the authors suggested to use different algorithms to improve the performance of the study. The performance results are almost similar to this study when using the SVM with a linear kernel.

SVM is significant in customer sentiment analysis, as many studies employed this algorithm due to its performance. However, in a previous study on the analysis of customer data sentiment that has been made by [14] discovered that the classification algorithm performs differently depending on the data. This is because the performance of the algorithm was found to be different when tested on two different data, namely reviews on products from the Amazon site and comments on IMDB movies. Yet when the two data are combined into single dataset, SVM algorithm is superior to the k-NN and Naive Bayes algorithms. Overall, evaluation between current and previous studies showed similarity of results in prediction, however, this similarity is essentially due to the fitting of the dataset to the machine learning model.

CONCLUSION

This paper focused on the analysis of customer sentiment through social media feedback using machine learning. Data on telecommunication company's customer feedback was extracted from Twitter. The tweets were fed to machine learning models to be trained and then checked the accuracy of each model. This sentiment analysis process consists of several steps such as data collection using Twitter API, text pre-processing, feature extraction, sentiment classification, training, and testing of the model. All data extraction processes from Twitter as well as sentiment analysis were conducted in Python through the Jupyter Notebook platform.

Continuing with the process of analyzing the sentiments of the tweets, the customer tweets collected were pre-processed through conversion to lowercase, removing URLs, removing stop words, removing short words, tokenization, and lemmatization. Text pre-processing is important because raw text taken from Twitter may contain unwanted or irrelevant symbols or words that would interfere the accuracy of the analysis. Using the bag of words method, which is CountVectorizer tool in the scikit-learn library in Python, each word in the tweets was transformed to a numeric representation, in turn counting how many times each word appears. Therefore, a sentiment classification was made to determine whether each tweet had a positive or neutral sentiment score. The results of the analysis have confirmed that there is no negative feedback from the telecommunication company's customers.

Three supervised machine learning models were developed based on the next objective of the study, which is to apply machine learning in improving the accuracy of sentiment analysis results. The models were SVM, Naive Bayes, and

Random Forest. The performance of each of these models have been successfully compared in terms of accuracy, AUC, MSE, and RMSE. By comparing the overall machine learning model, SVM provides the best sentiment prediction with lowest MSE and RMSE.

To conclude, sentiment analysis has been performed to detect positive and negative feedback about a company's products from the Twitter platform. This study's methodological approach may support the company in making a paradigm change toward data-driven decision-making. This study will be extremely useful to other business as well which have an online presence on Twitter and wants to maximize the chances accessible to them. Even so, there are some improvements that can be done in the future such as collecting data from other social media platforms such as Facebook, Instagram, Tik Tok, web, and forums, and perform a customer sentiment analysis study on Malay-language texts.

ACKNOWLEDGEMENT

REFERENCES

[1]   M. H. Saragih, and A. S. Girsang, "Sentiment analysis of customer engagement on social media in transport online," *Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology (SIET 2017)*, 2018-January, pp. 24–29, 2018. https://doi.org/10.1109/SIET.2017.8304103

[2]   Z. Drus & H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," *Procedia Computer Science*, vol. 161, pp. 707–714, 2019. https://doi.org/10.1016/j.procs.2019.11.174

[3]   A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artificial Intelligence Review*, Vol. 54, Issue 7, 2021. https://doi.org/10.1007/s10462-021-09973-3

[4]   V. Oktaviani, B. Warsito, H. Yasin, R. Santoso & Suparti. "Sentiment analysis of e-commerce application in Traveloka data review on Google Play site using Naïve Bayes classifier and association method," *Journal of Physics: Conference Series*, Vol. Vol. 1943, 2021. https://doi.org/10.1088/1742-6596/1943/1/012147

[5]   D. Kündüm, Z.H. Kilimci, M. Uysal & O. Uysal, "Evaluation of Customer Satisfaction about Telecom Operators in Turkey by Analyzing Sentiments of Customers through Twitter," Data Science and Applications, vol. 3, no.2, pp. 15–20, 2020.

[6]   Ali Mustafa Qamar, Suliman A. Alsuhibany and Syed Sohail Ahmed, "Sentiment Classification of Twitter Data Belonging to Saudi Arabian Telecommunication Companies" *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 8, no.1, 2017. http://dx.doi.org/10.14569/IJACSA.2017.080150

[7]   A. Rane and A. Kumar, "Sentiment Classification System of Twitter Data for US Airline Service Analysis," *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, pp. 769-773, 2018, doi: 10.1109/COMPSAC.2018.00114.

[8]   F. S. Fitri, M. N. S. Si and C. Setianingsih, "Sentiment Analysis on the Level of Customer Satisfaction to Data Cellular Services Using the Naive Bayes Classifier Algorithm," *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*, pp. 201-206, 2018, doi: 10.1109/IOTAIS.2018.8600870.

[9]   A. Nayak, "Comparative study of Naïve Bayes, Support Vector Machine and Random Forest Classifiers in Sentiment Analysis of Twitter feeds," *International Journal of Advanced Studies in Computer Science and Engineering*, vol. 5, no.1, pp. 14–17, 2016.

[10]   R. M. Alzoman and M. J. F. Alenazi, "A comparative study of traffic classification techniques for smart city networks," *Sensors*, vol. 21, no.14, pp. 1–17, 2021. https://doi.org/10.3390/s21144677

[11]   M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, Issue 0123456789, 2022. https://doi.org/10.1007/s10462-022-10144-1

[12]   D. Kündüm, Z.H. Kilimci, M. Uysal, and O. Uysal, "Evaluation of Customer Satisfaction about Telecom Operators in Turkey by Analyzing Sentiments of Customers through Twitter," *Data Science and Applications*, vol. 3, no.2, pp. 15–20, 2020.

[13]   Al-Otaibi, S., Alnassar, A., Alshahrani, A., Al-Mubarak, A., Albugami, S., Almutiri, N., & Albugami, A., "Customer satisfaction measurement using sentiment analysis," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, pp. 106–117, 2018. https://doi.org/10.14569/IJACSA.2018.090216

[14]   O. Grljević and Z. Bošnjak, "Sentiment analysis of customer data," *Strategic Management*, vol. 23, no. 3, pp. 38–49, 2018. https://doi.org/10.5937/straman1803038g