International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

Deciphering Voice of the Customer using Text Analytics and Sentiment Analysis: An Interpretable Review Rating Prediction using RoBERTa

Harish U.C.[1], Dhanya N.M.[2]
[1]MBA Department
City University Ajman
Ajman, United Arab Emirates
uc.harish@cuca.ae
[2]Department of Computer Science and Engineering,
Amrita School of Engineering, Amrita Vishwa Vidyapeetham
Coimbatore, India
dhanyahari07@gmail.com

Abstract— In this era of cut-throat competition among traditional and newer digital organizations, capturing, listening, and understanding customer voices are critical for success in the marketplace. The challenge to decipher the voice of the customer (VOC) has multiplied many times today, as now the number of customer reviews are present in multiple platforms and the data to be analyzed is huge. Sentiment analysis, and text analytics using machine learning, deep learning tools and transformer-based tools can be applied to gather meaningful insights from these data. This paper applies the traditional machine learning tools of the Naive Bayes classifier, Random Forest and AdaBoost to predict the customer review ratings. These results are compared with deep learning methods of CNN, RNN and Bi-LSTM and transformer-based approaches of BERT, DistilBERT and RoBERTa. The results show that RoBERTa has the highest accuracy among these methods. Paper also uses the explainable AI tool of LIME to understand the customer sentiments deeper in terms of why customers are giving a particular rating to the product. Business organizations will continue to use more and more AI tools to understand the customer feedback and the attempt in this paper is to learn how we can make predictions faster and more accurately.

Keywords-Sentiment analysis; voice of the customer; explainable artificial intelligence; machine learning; deep learning

I. INTRODUCTION

Business organizations look forward to listening to the voice of their consumers to design and improve the quality of the products and services as consumers are their source of revenue. An analysis of the voice of the customer can provide precise information about the customer’s requirements for a product or service [1]. It has always been a challenge to capture these voices and decipher them in the right manner so that the improvements of the product or service can lead to enhanced customer satisfaction. With the evolution of e-commerce sites such as Amazon as well as microblogging platforms, people can easily express their views about a product they purchased, customer service, etc. Reviews of products on e-commerce and social media platforms are valuable resources to understand the voice of the customers for business developers. These reviews also become a major source of information for consumers, and they will make their purchase decisions based on these reviews [2]. Today, these reviews come from all over the world and from a variety of people from various backgrounds. Customers write these reviews in a language which could be ambiguous for humans to understand [3].

Sentiment Analysis of review data from social media is becoming a very popular technique to understand the sentiment, preferences and expectations of customers [4]. It is critical that the comments provided by the customers are also analyzed along with the numerical ratings [5]. Such analysis using deep learning and NLP techniques will strengthen the business opportunity as it uncovers customer sentiments and helps in providing better marketing insights [6].

Sentiment analysis can be applied in various scenarios such as to gather public opinion about policy decisions of the government, election exit polls, reputation analysis etc. For example, sentiment analysis on Twitter data was done to understand the public opinion on demonetization policy in the Indian context by [7].

People will write the review in a very casual manner, and customers have the liberty to write reviews in their own style. The challenge lies in understanding the long-term dependency of the words in the review and inferring what other variables or features affect the review like time of review and title of the review. The predictions we make using a sentiment analysis

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

must be accurate and timely so that right decisions can be made and at the earliest possible time.

The need to process huge volumes of a variety of data makes sentiment analysis a challenging task. There are chances that the reviews are biased. Figure 1 shows a typical structure of reviews that can be found in online reviews. For example, "good facility and clean rooms" get a rating of 5.7/10, and "dirty rooms and unhygienic environment" gets a rating of 4.3/10. These are biased ratings and more accurate rating prediction is required based on the review. Hence predictions will be better if we are able to understand what aspect of the product or service leads to a particular rating which can provide better sense to the understanding of the customer sentiment.

This paper proposes to use machine learning and transformer-based approaches to predict customer sentiment. The data set on customer reviews of amazon echo dot customer reviews is analyzed using machine learning, deep learning tools and transformer-based tools to predict the customer ratings.

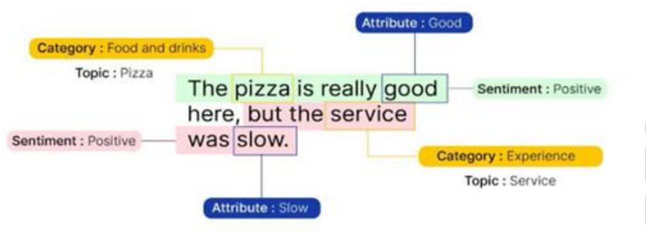

Figure 1. Structure of a review – an example

II. Related Works

There is an increasing number of works that carry out sentiment analysis using various machine learning, and deep learning tools. There are a number of tools which are used to carry out sentiment analysis and text analytics. These include lexicon based, machine learning based and deep learning-based methods [8, 9].

In the paper [10], have used a deep sequential model for review rating prediction deals with amazon's electronic dataset rating prediction. Deep learning mechanisms such as LSTM, GRNN and a combination of LSTM and GRNN is proposed and the combination gives higher accuracy. The goal of this research is to use a gated recurrent neural network to calculate vector representations of product reviews. The sentiment label of each document is further classified using this vector representation. A review's sentiment label is expressed as a ranking that runs from 1 to 5. Positive sentiment is denoted by higher ranking, whereas negative sentiment is denoted by lower value.

The paper [11] is predicting the review rating of open opinion using probability's classifier model. A case study of hotel recommendations is used as the application, where open comments by customers about the hotel reviews are taken as the data. Two categories of prediction are used here, either positive or negative. A decision tree and naive bayes classifier is used as the classification models. 400 customer reviews are taken as the dataset. The highest accuracy obtained is 94.37%.

The purpose of the work in [12] is to examine the advantages of PreTrained Models (PTM) in comparison to the current methods for app review analysis, particularly the tasks involving app review classification. The authors characterize the task of obtaining meaningful information from user comments, which might be connected to app needs, release planning, and software maintenance, as the app review classification. The information that was extracted may be used to identify various aspects of the application, including feature requests, aspect evaluations (e.g., feature strength, feature shortcoming, application performance), usability, portability, reliability, energy usage, problem reports, privacy concerns, and security concerns about the application.

The work [13] deals with the textual representation models for app review classification. The classical BoW model to the latest pre-trained model is evaluated in this paper. The results show that there is a significant accuracy for the classical Bag of Words model. The advantage of the pre-trained model is comparatively better accuracy, significant dimensionality reduction, and good semantic proximity in multilingual platforms.

[14] presents a comparison of several deep learning algorithms. The algorithms compared are CNN, RNN, BiLSTM, and various embedding schemes such as BERT, FastText, and Word2Vec. Word2Vec is showing the highest accuracy than any other models. The paper presents a benchmark comparison of a number of deep learning models, such as Convolutional Neural Networks, Recurrent Neural Networks, and Bi-directional Long Short-Term Memory, evaluated using a number of words embedding techniques, such as the Bi-directional Encoder Representations from Transformers (BERT) and its variants, FastText and Word2Vec.

[15] proposes a model for mining reviews from websites like Amazon, by automatically extracting reviews from the website. The model uses algorithms such as Naıve Bayes classifier, Logistic Regression and SentiWordNet algorithm to classify the review as positive and negative review. Naıve Bayes classification proves to be the most efficient among three algorithms for text classification of opinion mining.

[16] have proposed a methodology for constructing a domain specific lexicon paired with the algorithm for sentiment analysis. This methodology uses only three target labels of good, bad and neutral to extract sentiments and the precision showed to be high.

[17] uses online review extraction, bigram and trigram analysis, topic identification, to analyze online reviews posted for parcel and delivery service companies on social networking sites using text analytic tools.

[18] have used Naive Bayes classification methodology to online reviews of ecommerce platforms to assess the service

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

quality. The article states that traditional methods of customer surveys are no longer effective in terms of cost and time, and organizations need to use tools of sentiment analysis to analyze service quality.

The relationship between unstructured customer feedback and structured customer ratings is analyzed in [19] and concludes that one cannot solely rely on the ratings given by the customer to understand what the customer is actually intending to say.

[20] proposes a technique to improve classifiers for generating controlled text using the transformer-based Wasserstein autoencoder. The paper compares the results with classifiers trained on data generated by other synthetic data generators.

A hybrid deep learning method is proposed by [21] and has concluded that RoBERTa-LSTM model benefits from the strengths of both RoBERTa and LSTM, where RoBERTa efficiently encodes the words into word embedding while LSTM excels in capturing the long-distance dependencies

III. METHODS AND MATERIALS

A. Dataset

Customer reviews about Amazon Echo Dot from November 2016 to November 2017 are collected from Kaggle with the following fields. A total of 53048 customer reviews are collected.

- A Unique identifier for each review
- The timestamp at which the data is crawled
- The URL of the review page
- Title for the review
- Review Text
- Review Color
- User verified or not
- Review date
- Review Useful count
- Declaration Text
- Rating

In the dataset, the following modifications are made as the two fields of 'Declaration Text' and 'Review Useful count' were not used for analysis as these fields had null data in most cases. The fields of 'Uniq Id', 'CrawlTimestamp' and 'Pageurl' were also not used for any analysis.

The Ratting field is pre-processed and converted into numeric rating for further analysis.

Figure 2 shows a sample dataset for Amazon Echo.

Uniq Id	Crawl Timestamp	Pageurl	Title	Review Text	Review Color	User Verified	Review Date	Review Useful Count	Configuration Text	Ratting	Declaration Text
63583450415a20094950528ffb4d955	2017-10-26T15:57:14Z	https://www.amazon.com/All-New-Amazon-Echo-Dot...	Five Stars	Love the Echo Dot.	Black	Verified Purchase	2017-07-03	NaN	Echo Dot	5.0 out of 5 stars	NaN
c8e5ca6b44bea1006c8bb85cdca3816	2017-10-26T15:57:14Z	https://www.amazon.com/All-New-Amazon-Echo-Dot...	Five Stars	Working just fine.	Black	Verified Purchase	2017-07-12	NaN	Echo Dot	5.0 out of 5 stars	NaN
f823996e2317dd65a6235011492b42	2017-10-26T15:57:14Z	https://www.amazon.com/All-New-Amazon-Echo-Dot...	Five Stars	I love my Echo Dot	Black	Verified Purchase	2017-08-01	NaN	Echo Dot	5.0 out of 5 stars	NaN
6c928e62707a1530c591b897b864d6	2017-10-26T15:57:14Z	https://www.amazon.com/All-New-Amazon-Echo-Dot...	Three Stars	Not great speakers	Black	Verified Purchase	2017-10-03	NaN	Echo Dot	3.0 out of 5 stars	NaN
275af85c81c1be55efd706f51a6c7cbe	2017-10-26T15:57:14Z	https://www.amazon.com/All-New-Amazon-Echo-Dot...	Five Stars	Great assistant !!	Black	Verified Purchase	2017-07-22	NaN	Echo Dot	5.0 out of 5 stars	NaN

Figure 2. Sample Dataset

Table 1 shows the category of customer reviews with the average rating per category. It is seen that reviews on Echo Dot are significantly larger than other combinations. Hence this paper analyses only Echo Dot reviews and ignores the other 9 reviews. Table 2 shows review colour summary and shows that more people prefer black model than the white model. Figure 3 shows the consumer ratings for the product.

TABLE I. CATEGORY OF CUSTOMER REVIEWS.

Category	*Count*	*Average Rating*
Ecodot	53039	4.296141
Echo Dot + Vaux Speaker	6	5.000000
Echo Dot + Sony XB10 Speaker	2	3.666667
Echo Dot + Philips Hue Smart Lighting Kit	1	5.000000

TABLE II. CATEGORY OF CUSTOMER REVIEWS

Review Color	*Count*	*Average Rating*
Black	39486	4.28766 8
White	13553	4.32069 6

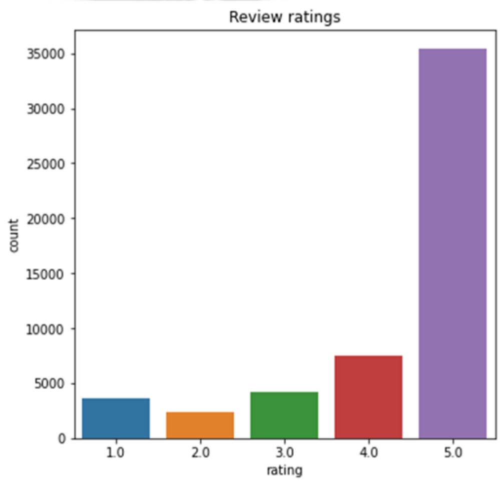

Figure 3. Count of review ratings

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

B. Architecture Diagram

This section presents the architecture diagram (Figure 4) that outlines the proposed methodology employed in this work

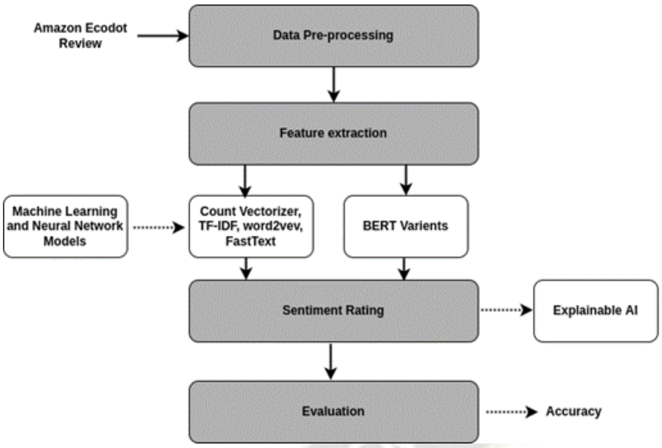

Figure 4. Architecture diagram of proposed methodology

C. Data pre-processing

Pre-processing is done to remove punctuations and convert the reviews into lower cases. The stop words were removed from the review text. Then, tokenization is done to split sentences into separate individual words. After this, lemmatization is done to convert words into its root formats. Zero padding is done to make the reviews into equal length as the models will accept only inputs of equal lengths.

D. Feature extraction

The text is converted to numerical form by using bag of words (count vectorizer) and Tf-idf (Term frequency - inverse document frequency) methods and the results are compared. The following deep learning-based word embedding methods are also used to obtain better accuracy.

Word2vec: It is a pre-trained model that understands the relationship between the words in a corpus, and returns a vector representation for each word in the text

FastText: It is an extension of word2vec which is based on n-grams, where the words are divided into smaller representations for better understanding.

E. Machine Learning Models

Naive Bayes classifier: This is a popular probabilistic machine learning model used for classification. The multinomial Naive Bayes algorithm is used in this paper as it is mostly used for text classification problems. The algorithm calculates the probability of each tag and provides the tag with the highest probability. Random Forest: Random Forest classifier constructs a number of decision trees and provides the output which is the class selected by most of the trees

Ada boost: AdaBoost which is similar to Random Forest is another classification algorithm that can be used with other algorithms to improve performance.

F. Deep Learning Models

Three well-known neural network algorithms are used for sentiment analysis such as CNN, RNN and Bi-LSTM. RNN is a neural network model which is good for modeling sequential data. An LSTM model is used for this experiment where long-term dependencies are captured efficiently. An LSTM layer of 256 units and a dropout of 0.5 and a learning rate of 0.01 is used to train and test the sentiment prediction

CNN which is a popular method for images is a good model for text processing also. It is designed to learn spatial hierarchies. It mainly consists of a 1D convolution layer, pooling, and fully connected layers. The first two layers perform the feature extraction and the fully connected layer maps these features to output. Here 256 filters with a window size of 3 is used along with ReLU activation function. The drop out selected is 0.5.

Bi LSTM consists of LSTM units that work in both directions. One LSTM takes inputs in the forward direction and another in the backward direction. It usually learns features faster than LSTM. A layer of BiLSTM with 256 units is created with a dropout of 0.5.

G. Explainable AI

In order to explain using the LIME explainer the input around the neighborhood is considered to see how the model behaves. The data points are then weighted by the proximity to the original sample. For example, if the explanation is for the sentence "I love this product" we will perturb the sentence and get a prediction for the following sentences. "I love product", "I this product", "I love". Even if the classifier uses word embeddings, the sentences can be used with these embeddings and LIME still works and gives the explanation in terms of "love" and "movie" [22].

H. BERT Variants

Pretrained transformer-based models of BERT, DistilBERT and RoBERTa are applied to increase the prediction accuracy [23, 24]. BERT is a pre-trained Language model for many NLP tasks [25]. BERT can still be improved based on the data in which it is trained and the duration of training with the proposed optimized and robust version of BERT called RoBERTa (Robustly Optimized BERT-Pretraining Approach) [26]. Increasing the training data generally improves the performance of the model. RoBERTa is trained on a vast dataset over 160GB of text [27]. The dataset consists of book corpus, english wikipedia, CC-News, open WebText and stories [28, 29]. The models are evaluated based on the predicted value against the actual rating using accuracy score. To relate the features of the device with the ratings, Wordcloud and bigram analysis is carried out. The results of these are used to understand the product features which leads to 'good' or 'bad' ratings.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

IV. RESULTS AND DISCUSSION

A. Word Cloud

The figures 5 - 9 shows the word cloud for the customer ratings of the data set. Word cloud shows the most used words in the reviews for the ratings of 5 to rating of 1. 'Love', 'use', 'music', 'fun', and 'speaker' are some of the most used words for a rating of 5 and rating 4 which relates to the positive aspects of the product. For lower ratings of 1 and 2, some of the words seen are 'wifi', 'connect and 'time'

Figure 5. Word cloud for rating 5.

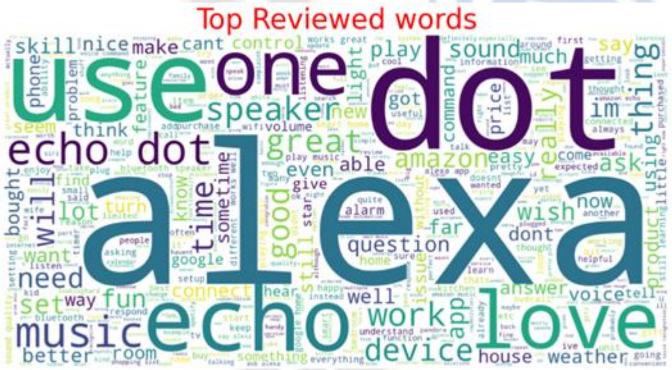

Figure 6. Word cloud for rating 4.

Figure 7. Word cloud for rating 3.

Figure 8. Word cloud for rating 2.

Figure 9. Word cloud for rating 1.

B. Bigram Analysis

Result of bigram analysis is shown in figure 10 below. Bigram analysis is done by considering reviews associated with rating 5 and 4 as good reviews and the reviews associated with ratings 1 & 2 as bad reviews. Reviews for rating 3 are treated as neutral. The good reviews which show why and how people use echo are 'love echo', 'works great', 'play music', 'great product', 'easy set', 'bluetooth speaker', 'smart home', 'shopping list' and others. The bad reviews which show the difficulties in using the product are 'play music', 'google home', 'doesn't, work', 'don't know', 'waste money', 'sound quality', 'customer service' and others.

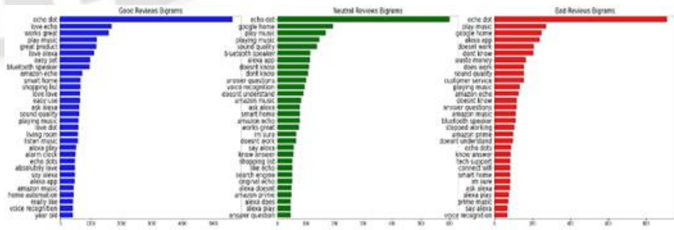

Figure 10. Bigram Analysis

C. Ratings of the Features

Some of the important features of echo dot are considered and the ratings which are given for these features are shown in

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
DOI: https://doi.org/10.17762/ijritcc.v10i12.5840
***Article Received**: 02 October 2022 **Revised**: 11 November 2022 **Accepted**: 30 November 2022*

Table 3. It can be seen that the average rating is relatively low for wifi, Alexa app and tech support.

TABLE III. RATINGS OF FEATURES OF ECHO DOT

Rating	Total Reviews	5	4	3	2	1	Average rating
Wifi	1295	435	223	162	165	310	3.238
Bluetooth	2345	1397	460	221	134	133	4.217
Speaker	5673	3319	1181	587	302	284	4.225
Alexa app	743	291	139	107	64	142	3.502
Voice	2834	1507	571	349	220	187	4.055
Tech support	99	12	5	10	17	55	2.01

D. *Rating prediction models*

The following Table 4 shows the accuracy percentages of the traditional machine learning approaches of bayes classifier, random forest, adaboost and deep learning approaches such as CNN, RNN and Bi LSTM and transformer-based approaches of BERT, Distil BERT and RoBERTa. RoBERTa has the highest accuracy percentage of 92.45%.

TABLE IV. ACCURACY PERCENTAGE OF THE RATING PREDICTION

Method	Review Text	Title	Combined
Count vectorizer+bayes classifier	73.62	75.55	78.94
Count vectorizer + random Forest	74.81	79.47	81.19
CountVectorizer+ Adaboost	71.03	77.03	77.23
Tf-idf + NB	68.51	77	78.13
Tf-idf + random forest	69.23	72.15	79.69
Tf-idf + ada boost	73.56	76.98	80.15
word2vec+CNN	78.26	79.84	80.46
word2vec+RNN	79.42	80.96	82.76
word2vec+BiLSTM	81.93	83.63	85.56
FastText+CNN	80.49	80.97	81.29
FastText+RNN	82.19	85.19	86.17
FastText+BiLSTM	82.82	85.73	86.27
BERT	85.21	86.21	89.21
Distil BERT	83.45	86.31	89.37
RoBERTa	86.21	89.24	92.45

E. *LIME text explainer*

Results of the LIME text explainer for rating 5 is given below where in the highlighted text as seen in figure 11 is contributing to the probability of the customer ratings. This analysis explains the aspect or the feature that impacts the rating for the product. Examples for each of the five ratings with the probability values are given below

Figure 11. Example for rating 5.

V. CONCLUSION

In this paper, analysis is done on customer reviews to predict the sentiment of the customers regarding the product using machine learning tools of bayes classifier, random forest, adaboost and deep learning tools of CNN, RNN, BiLSTM and transformer-based approaches of BERT, Distil BERT and RoBERTa. RoBERTa gave an accuracy percentage of 92.45 and is the highest among all the tools. Today, when the customers are able to review and post their reviews very quickly on various online platforms, it is critical that the companies are able to understand and respond faster to those product reviews. In such a scenario, accuracy of prediction is also very critical. This requires efficiently mining such data [30] which is complex as it is unstructured and is voluminous. The paper helps in usage of NLP tools to process the customer reviews which are an important source of customer review information and can be used to predict the customer usage of various features of the product. This can help a company to modify its features in its existing products and also in its upcoming other products. These predictions on customer reviews need to be improved in terms of accuracy and relevancy as some of the challenges are not addressed in this work. One challenge is about reviews done in languages other than English which is a critical issue to be addressed. Another challenge is to verify the authenticity of the reviews. An analysis to see whether review is done by a verified customer, reviewer's previous history can be included in future work.

REFERENCES

[1] Aguwa, C. C., Monplaisir, L., Turgut, O. Voice of the customer: Customer satisfaction ratio based analysis. Expert 20 Systems with Applications 2012; 39(11), 10112-10119.

[2] Jin, J., Liu, Y., Ji, P., Kwong, C. K. Review on recent advances in information mining from big consumer opinion data for product design. Journal of Computing and Information Science in Engineering 2019; 19(1).

[3] Lye, S. H., Teh, P. L. Customer Intent Prediction using Sentiment Analysis Techniques. In 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) 2021; Vol. 1, pp. 185-190. IEEE. 26

[4] Ligthart, A., Catal, C., Tekinerdogan, B. Systematic reviews in sentiment analysis: a tertiary study. Artificial Intelligence Review 2021; 54(7), 4997-5053.

[5] O¨ zdaˇgoˇglu, G., Kapucugil-Ikiz, A., Celik, A. F. Topic modelling-based decision framework for analysing digital voice of the customer. Total Quality Management and Business Excellence 2018; 29(13-14), 1545-1562.

[6] Ramaswamy, S., DeClerck, N.Customer perception analysis using deep learning and NLP. Procedia Computer Science 2018; 140, 170-178.

[7] Dhanya, N. M., Harish, U. C. Sentiment analysis of twitter data on demonetization using machine learning techniques. In Computational vision and bio inspired computing 2018; pp. 227-237. Springer, Cham.

[8] Nandwani, P., Verma, R. A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining 2021; 11(1), 1-19.

[9] Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., Baz, M. Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning.Sensors 2022; 22(11), 4157.

[10] Verma, S., Saini, M., Sharan, A. Deep sequential model for review rating prediction. In 2017 Tenth International Conference on Contemporary Computing (IC3),2017; pp. 1-6). IEEE.

[11] Songpan, W. The analysis and prediction of customer review rating using opinion mining. In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) 2017; pp. 71- 77. IEEE.

[12] Hadi, M. A., Fard, F. H. Evaluating pre-trained models for user feedback analysis in software engineering: A study on classification of app-reviews, 2021; arXiv:2104.05861.

[13] Araujo, A., Golo, M., Viana, B., Sanches, F., Romero, R., Marcacini, R. From bag-of-words to pre-trained neural language models: Improving Automatic Classification of app reviews for requirements engineering. Anais Do Encontro Nacional De Inteligˆencia Artificial e Computacional (ENIAC 2020) 2020; https://doi.org/10.5753/eniac.2020.12144

[14] Balakrishnan, V., Shi, Z., Law, C. L., Lim, R., Teh, L. L., Fan, Y. A deep learning approach in predicting products' sentiment ratings: a comparative analysis. The Journal of Supercomputing 2022; 78(5), 7206-7226.

[15] Kumar, K. S., Desai, J., Majumdar, J. Opinion mining and sentiment analysis on online customer review. In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) 2016; pp. 1-4.

[16] Grabner, D., Zanker, M., Fliedl, G., Fuchs, M. Classification of customer reviews based on sentiment analysis. In ENTER 2012; pp. 460-470.

[17] Rajendran, S. Improving the performance of global courier and delivery services industry by analyzing the voice of customers and employees using text analytics. International Journal of Logistics Research and Applications 2021; 24(5), 473-493.

[18] Sari, P. K., Alamsyah, A., Wibowo, S. Measuring e-Commerce service quality from online customer review using sentiment analysis. In Journal of Physics: Conference Series 2018; Vol. 971, No. 1, p. 012053. IOP Publishing.

[19] Gallagher, C., Furey, E., Curran, K. The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying. International Journal of Data Warehousing and Mining (IJDWM) 2019; 15(4), 21-47.

[20] Harikrishnan, C., Dhanya, N. M. Improving Text Classifiers Through Controlled Text Generation Using Transformer Wasserstein Autoencoder. In Inventive Communication and Computational Technologies 2022; pp. 97-105. Springer, Singapore.

[21] Tan, K. L., Lee, C. P., Anbananthen, K. S. M., Lim, K. M. RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network. IEEE Access 2022; 10, 21517-21525.

[22] Ribeiro, M., Singh, S., Guestrin, C. "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. 2016; https://doi.org/10.18653/v1/n16-3020

[23] Jain, P. K., Quamer, W., Saravanan, V., Pamula, R. Employing BERT-DCNN with sentic knowledge base for social media sentiment analysis.Journal of Ambient Intelligence and Humanized Computing 2022; pp1-13.

[24] Mingyu, J., Jiawei, Z., Ning, W. AFR-BERT: Attention-based mechanism feature relevance fusion multimodal sentiment analysis model. PloS one 2022; 17(9), e0273936.

[25] Devlin, J., Chang, M., Lee, K., Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. 2018; https://doi.org/10.48550/arXiv.1810.04805

[26] Liao, W., Zeng, B., Yin, X., Wei, P. An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. Applied Intelligence 2021; 51(6), 3522-3533.

[27] Rajapaksha, P., Farahbakhsh, R., Crespi, N. BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits. IEEE Access 2021; 9, 154704-154716.

[28] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettle6 moyer, L., Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. 2019; arXiv. https://doi.org/10.48550/arXiv.1907.11692 8

[29] Catelli, R., Pelosi, S., Esposito, M. Lexicon-based vs. Bert-based sentiment analysis: A comparative study in Italian. Electronics 2022; 11(3), 374.

[30] Li, W., Hu, B., Ke, S., Xiao, X., Deng, Y., Du, S. . The Application of NLP Technology In Customer Voice Analysis. In 2022 4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP) 2022; pp. 894-897. IEEE.