The 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
22-25 September, 2021, Cracow, Poland

Customer Intent Prediction using Sentiment Analysis Techniques

Say Hong Lye [1], Phoey Lee Teh [2]
[1,2] Department of Computing and Information Systems, Sunway University, Bandar Sunway, Malaysia,
say.l@imail.sunway.edu.my, phoeyleet@sunway.edu.my https://university.sunway.edu.my/

Abstract— **Analysing the voice of the customer (VoC) through the customer intent has many applications ranging from personalised marketing to behaviour study. Individuals express their feelings in a language that is frequently accompanied by ambiguity and figure of speech, making it difficult even for humans to understand. Customer feedback is crucial as part of the customer experience (CX) management in customer retention and improves the sales strategy. Modern research has been using machine learning and word embedding technique for sentiment analysis, and it is focused on the predictive model without further context. In this study, the customer feedback comes in the form of Net Promoter Score (NPS) with a text box for written feedback. We analyse the data and demonstrate a hybrid representation that has resulted in the accuracy improvement of the sentiment classification task and predicting customer intent. The datasets were first trained using Word2Vec with the previous dataset and then fit into the Random Forest classifier, tested as the best configuration to prevent overfitting. The hybrid representation is compared against the baseline sentiment polarity tool through few experiments; the results have shown that the hybrid model has improved accuracy for the sentiment classification task. Lastly, we performed customer intent prediction by using the Power BI influencer module. The outcome of the result can be used as a reference for IT management in decision making.**

Keywords— Text Mining, Sentiment Analysis, Word Embedding, Random Forest, Classification, Customer Intent Prediction

I. INTRODUCTION

The customer experience (CX) refers to the overall experience a customer has with particular products or services, based on their engagement with and thoughts about the brand [1]. The sentiment elements of the CX relates to the features, speed, and availability of a service or product, and these sentiments can be positive or negative, for example, delight, regret, anger, outrage, joy or surprise [2]. The relationship between CX and customer commitment is crucial to retain customer loyalty. In the logistics industry, CX plays a crucial role in retaining the customer directly (Voice of Customer) and indirectly (Voice of Product) through an empirical study based on surveys [3]. The measurement of CX in this study is through Net Promoter Score (NPS) gathered from the customer feedback survey (CFS) in the IT department of the logistics company.

While the CFS form is the platform to capture customer feedback, there is currently no research to identify the customer intent by comparing NPS scoring. The analytic process for the sentiment analysis cannot be done just relying on the NPS scoring data.

Sentiment analysis is a standard method used to detect sentiment polarity scoring. The sentiment analysis of the short text of tonal estimates helps identify whether customer feedback has a negative or positive trend [4].

A. Models

In textual information classification, there are facts and opinions. The fact will provide the objective intent about entities, events, and attributes, while the opinion will address the subjective intent on human sentiment and feelings to their properties [4]. The author [4] highlights the sentiment analysis of the short text of tonal estimates that help identify the condition where the customer feedback has a negative or positive trend. Sentiment analysis had gone beyond the text mining of text categorisation and labelling capability. With the current research, sentiment analysis can analyse opinion expression, emoticons, emoji ideograms [5], and exclamation marks [6]. The author [7] shares various factors in influencing the sentiment level expressed in the textual comments. The example includes using the capital letter [8] tends to mark something for attention, emoticons to visualise the facial expression, and its potential to reverse the meaning of a statement [9]. Also, the test result of using exclamation marks pointed out that adding more exclamation marks on the negative comments will result in higher dissatisfaction [10]. The mentioned arguments validate that customer experience and sentiment feedback will influence overall customer satisfaction.

1) Word Embedding

Word embedding represents words within a continuous embedding space by exploiting the word similarity patterns to be learned and reused later; the representation is numeric. The data accuracy can be improved when the text segmentation technique of using word embedding is used [11][12]. The Word2Vec model is selected in this study because it is a simple method for finding phrases in the text and shows that learning a good vector representation of phrases is possible [13]; Word2vec is easy to implement with robust architecture and works well in capturing the semantic similarity. A recent study [14] pointed that Word2Vec provided an efficient implementation of CBOW and Skip-gram architectures for word vector representation. The study [14] shows more than 99% accuracy after the implementation of the Word2Vec model. While the other similar research presented the work experiments with the Word Embedding as the prediction model using different machine learning techniques to improve customers' shopping experience [11]. The results indicate that word embeddings gave the best predictor model and can be a valuable technique for many NLP applications. Both types of research indicate great results in capturing the words as text representation.

2) Random Forest

Random Forest (RF) algorithm is a combined classifier algorithm [15]. With the acquiring of classification and regression tree (CART) as the classifier element, RF can tolerate noise further and improve higher classification performance. RF extends the basic idea of a single classification tree by growing many classification trees in the training phase and averaging to improve the predictive accuracy. To classify an instance, each tree in the forest generates its response (vote for a class). One significant advantage of RF over traditional decision trees is protecting against overfitting, making the model deliver high performance [16][17].

3) Sentiment Polarity Tool

Both VADER [18] and TextBlob [19] are frequently used and cited. VADER (Valence Aware Dictionary and sEntiment Reasoner) belongs to a type of sentiment analysis based on lexicons of sentiment-related words and rule-based. Each of the words in the lexicon is rated with a sentiment polarity of positive and negative. Multiple studies suggest that VADER provides higher accuracy in sentiment polarity scoring than other methods [20]. TextBlob is a python library for textual data processing. First, it provides the platform, to begin with, the standard natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more. The recent research pointed out that TexBlob can provide higher accuracy during the experiment in capturing the mobile branding sentiment polarity [19].

4) Word Embedding and Random Forest as Classification Model

The recent research concerning the pairing model of word embedding and RF is common with excellent results, but the focus is different from this study. Research paper [11] focuses on regression but not classification. This paper highlights the benefits of Word-Embedding, suggesting that with Word-Embedding, it will be the best predictor model. Fundamentally, classification is about predicting a label, and regression is about predicting a quantity. The author of the research paper [21] proposed a feature extraction using Word Embedding and a content-based classification model which uses machine learning to filter out unwanted messages. The performances of the classification algorithms are compared, and the RF method succeeded with an accuracy rate of 99.64%. For non-English research using Turkish text analysis, the author of the research paper [22] was able to measure higher accuracy of the results (56.48%) by using a word embedding and RF in comparison against the other four machine learning models like Support Vector Machine (SVM), Naïve Bayes (NB), k-nearest neighbours (KNN), and Decision Tree (DT). Using the similar approach of Word-Embedding and RF in the ontology field, the author [23] predicts the alignment between ontology concepts with more than 85% precision scoring findings that the hybrid approach is a good predictor model. Based on our literature review concerning the customer service industry, implementing machine learning techniques in detecting the sentiment polarity were focused on combining both Word-Embedding and Random Forest (RF) [11][23]. These studies suggested that the combination is a good predictor model, but there is no research in comparing the data accuracy from the NPS result with the submitted text. The NPS result only captures the customer scoring but not the customer intent in the context of satisfaction and dissatisfaction with the products and services they had subscribed to.

5) Customer Intent Prediction

Voice of Product (VoP) is the next step or the evolution of the products and what customer needs or wishes [24]. This can be summarised as customer intention. There are many predictive tools in the market. Still, without much programming intervention, Microsoft developed a module in Power BI reporting tool called Influencer Visual based on the ML.NET framework, a predictive machine learning model using the Gradient Boosting algorithm called MART [25]. The algorithm employed was an efficient implementation and show promising results in the mechanical prediction [26] and error compensation [27]. Based on the Google Scholar search, the application of ML.NET into the customer service industry for customer intent prediction is not available.

B. Research Objectives

There are three research objectives in this study. The first research objective (RO#1) compares the results between the accuracy of NPS scoring with sentiment polarity tool (TextBlob and VADER) and a hybrid model of Word-Embedding Random Forest. The accuracy result helps identify the best predictor model for customer intent prediction in the later stage and validate if the NPS is a suitable model for gauging VoC.

The second research objective (RO#2) is to identify the best method to perform customer intent prediction based on Result #1.

The third research objective (RO#3) is to compile the intent category to predict the customer intent concerning the IT products and services.

II. METHOD

In understanding the customer intent, further investigation to recognise the classification of customer sentiment is crucial. In this section, the researcher presents the methods conducted in this Sentiment Analysis study. This study has five methods detailed as follows and illustrated in Fig. 1.

A. Dataset Extraction

With the approval of the senior management of the IT department of the logistics company, the datasets from the year 2019 were extracted from their customer feedback system (CFS) database in Microsoft Excel data format with compliance and accordance with the data protection act and the data privacy law. A total of 27,897 rows of data was extracted.

B. Data Pre-Processing

Ensuring the data quality extracted from the CFS database helps transform the text into a more understandable format so that machine learning algorithms can perform better. Tokenisation, lemmatise, and stop words removal are the steps used for data pre-processing.

C. Data Benchmarking

After the data pre-processing, with a simple python script in capturing the word count and unique words; we have found 107,035 total numbers of words with 6,033 unique words in the dataset and a total of 21,359 rows of data with 20,023 (93.7%) positive and 1,336 (6.3%) negative feedback.

D. Sentiment Polarity Scoring

Word2Vec + RF (WE+RF): We deployed the Word2Vec and RF algorithm using a python script. First, we deploy the word embedding function using Gensim models of word2vec. Second, the words vector was trained with two types of data training, (1) Year 2018 Submitted feedback datasets from the CFS database from the IT department of the logistics company – 25,000 rows of data, (2) IMDB training dataset downloaded from Kaggle – 25,000 rows of data. The trained data passed into the Random Forest classifier configured at 100, 200, 400, 600, 800, and 1200 trees using the sklearn module through the python script. The result is presented in the Result section.

VADER (VD): A python script has been created using the NLTK SentimentIntensityAnalyzer library to capture the sentiment polarity scoring. The results were collected and presented in the Result section.

TextBlob (TB): A python script using the textblob library to capture the sentiment polarity scoring was developed. The results were collected and presented in the Result section.

E. Customer Intent Categorisation and Prediction

To analyse the customer intent, we manually labelled and categorised the customer intent category based on the ticket resolved under the support group and their function with the help of the human annotation from the IT department in the logistic company. The nine intent categories are Application, Architecture, Commercial, Facility, Infrastructure, Local Support, Service, Software Design, and Others. The motivation behind this categorisation is to retrieve the customer intent either in positive or negative sentiment through the automated sentiment polarity classification method towards a category, thus providing the service improvement options in supplement the CX.

With the best sentiment analysis method selected, the results are loaded into the Microsoft Power BI tool and is configured with Influencer Visual Module, which can help to visualise and assist in understanding the factors that drive the customer intent result. It analyses the data, ranks the factors that matter, and displays them as key influencers. This study visualised customer intent prediction with the average sentiment polarity result of the overall negative and positive. Follows by using key influencers visuals to show the area which requires improvement.

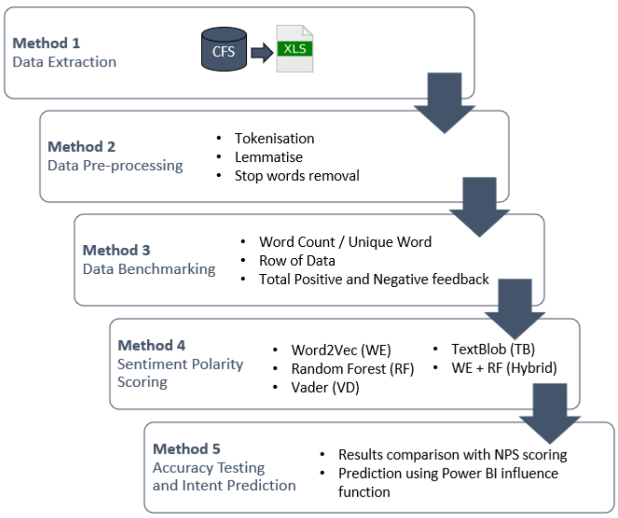

Figure 1. Methods used

III. RESULT

The result of the experiments was retrieved based on the evaluation of our method to predict the customer intent using the sentiment analysis method. Our preliminary study shows that the best method of a hybrid WE+RF has improved accuracy in sentiment polarity scoring. The WE+RF is trained with 100 trees and slowly increased to 1200 trees; this training uses the 21,359 processed data. Each training took less than 3 minutes. The result is stable after configured with 800 trees. The configuration of "n_estimators = 800" in the RF python code help in the overfitting issue.

A. *Sentiment Polarity Scoring*

Table I shows that all input method is lower than the baseline accuracy. The highest accuracy achieved was 85.50% using the VD polarity classification tool, follows by TB at 85.30%. The lowest value by just using RF with the configuration of 800 scores the lowest at 63.20%. The proposed model accuracy scoring is at 77.30%, which is lower than both VD and TB.

TABLE I. SENTIMENT POLARITY SCORING AND ACCURACY RESULT

Sentiment	Methods				
	NPS	**RF (800)**	**WE+RF (800)**	**VD**	**TB**
Negative	1336	8166	4951	3334	3467
Positive	20023	13193	16408	18025	17892
Accuracy (%)		63.20	77.30	85.50	85.30
Total	21359	21359	21359	21359	21359

The differences show either the customer is merely rating the NPS or the proposed method is inaccurate. To analyse this situation, we use golden label by hiring an experienced IT staff in the logistic company to label random 100 data from the NPS and WE+RF manually. We observed that the customer NPS rating differs from their submitted text. The result of the golden label indicates that the accuracy of the polarity sentiment scoring tool does not understand the sarcasm and the meaning of the sentence. Both VD and TB are based on the sum of scoring. Hence, WE+RF shows a more accurate result, and the data output from WE+RF is used for the customer intent prediction - this met RO#1. While WE+RF is the best method to perform customer intent prediction, this met RO#2.

B. *Customer Intent Prediction*

The results processed by the Power BI tool are shown based on the positive (POS) (Fig. 2) and negative (NEG) (Fig. 3) sentiment. The predictive illustration based on the POS sentiment shows that the Commercial category has the likelihood of further increase of rating by 1.09 (9%). For NEG prediction, the Others category is having the likelihood of increased NEG rating by 1.60 (60%). Both POS and NEG results are crucial for IT management to improve the products and services - this met RO#3.

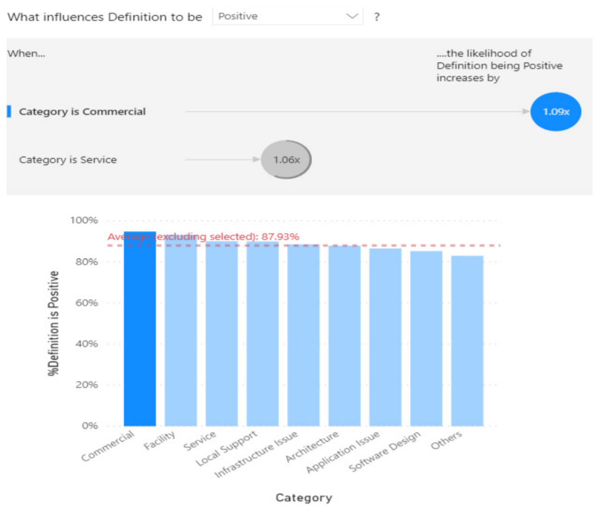

Figure 2. Positive (POS) result

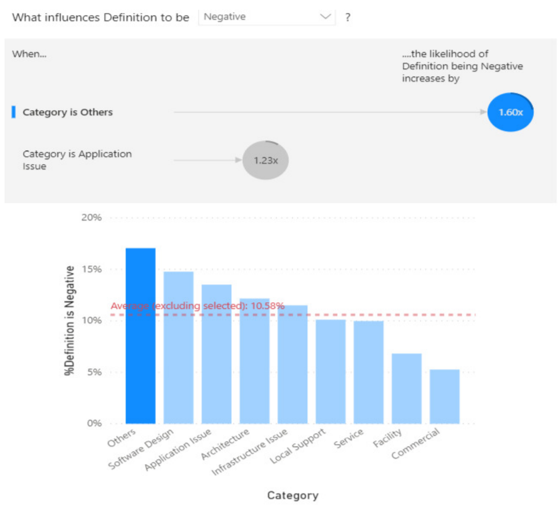

Figure 3. Positive (POS) result

C. *Conclusion*

Based on the experiments conducted, the results indicate that the combination hybrid model of WE+RF provides a more accurate result than the sentiment polarity tool. The sentiment polarity tool analyses the submitted text and converts it into words and compares it against its lexicon. It classified the word and assigned a scoring with negative and positive metric values. VD indicates the highest scoring, but it does not capture the customer intent. This method does not help in understanding the VoC.

The proposed WE+RF model shows the accurate result by verifying it through the golden label method through the test in this study. This shows that the feature and words similarity extraction through word embedding and the best overfitting method in machine learning using Random Forest supports this result. Power BI as the predictive tool able to provide customer intent prediction and the efficiency of the result can be compared with the future data extraction and analysis. All the research objectives were met through the test result.

IV. Discussion and Conclusion

This study aims to explore the comparison of the data accuracy between NPS scoring with sentiment polarity tool and a hybrid model of Word-Embedding and Random Forest. The result shows accuracy improvement in actual sentiment polarity scoring with the best method, hybrid model using Word2Vec and Random Forest, to perform customer intent prediction. The attempted results of this research offer both research and managerial implications for the IT department in the logistic industry and beyond to now be detailed.

A. Research implication - Effects of NPS scoring on customer experience

Customer feedback on products and services are the essential attributes affecting the 5-Star rating for the IT department of the logistic company. The findings of this research contribute to the research in the CX area and supporting the existing literature, which found that NPS scoring is not a recommended method in gauging the VoC. The accuracy between NPS scoring and the submitted feedback have differences, and with the proposed model, we observed some of the data scored with 5-Star, but the feedback is with sarcasm, and the issue is not resolved.

B. Executive implication - Predict customer intent and prioritise improvement

The IT management team should consider the customer intent prediction result and use the data to prioritise the improvement on the suggested area to improve the performance of the products and services.

C. Conclusion

Text mining, primary on words and sentiment analysis, plays a crucial role in understanding the customer experience [28]. Based on the experiments, WE+RF with the configuration of 800 trees is more accurate in terms of the true meaning. It is suitable as the predictor model in helping this research achieve the customer intent prediction. The implication of this research has been discussed, contributing to both research and executive implication. The results emphasise the suggestion that it would be a disadvantage for a business to ignore the VoC in the form of submitted text in the CFS and solely depending on the NPS result. Concurrently, the result of the experiment shows that all the research objectives were met. Hence, the customer intent prediction can be achieved by using the proposed model.

References

[1] N. Ameen, A. Tarhini, A. Reppel, and A. Anand, “Customer experiences in the age of artificial intelligence,” Comput. Human Behav., vol. 114, no. June 2020, p. 106548, 2021, doi: 10.1016/j.chb.2020.106548.

[2] T. Keiningham et al., “The interplay of customer experience and commitment,” J. Serv. Mark., vol. 31, no. 2, pp. 148–160, 2017, doi: 10.1108/JSM-09-2016-0337.

[3] M. Querin, Francesco; Göbl, “An analysis on the impact of Logistics on Customer Service,” J. Appl. Leadersh. Manag., vol. 5, pp. 90–103, 2017.

[4] V. Sydorenko, S. Kravchenko, Y. Rychok, and K. Zeman, “Method of Classification of Tonal Estimations Time Series in Problems of Intellectual Analysis of Text Content,” Transp. Res. Procedia, vol. 44, no. 2019, pp. 102–109, 2020, doi: 10.1016/j.trpro.2020.02.015.

[5] K. Korovesis, G. Alexandridis, G. Caridakis, P. Polydoras, and P. Tsantilas, “Leveraging aspect-based sentiment prediction with textual features and document metadata,” ACM Int. Conf. Proceeding Ser., pp. 168–174, 2020, doi: 10.1145/3411408.3411433.

[6] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar, and M. Abdel-Basset, “Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network,” IEEE Access, vol. 7, pp. 23319–23328, 2019, doi: 10.1109/ACCESS.2019.2899260.

[7] P. L. Teh, I. Pak, P. Rayson, and S. Piao, “Exploring fine-grained sentiment values in online product reviews,” in ICOS 2015 - 2015 IEEE Conference on Open Systems, Jan. 2015, pp. 114–118, doi: 10.1109/ICOS.2015.7377288.

[8] I. Pak and P. L. Teh, “Value of expressions behind the letter capitalization in product reviews,” ACM Int. Conf. Proceeding Ser., pp. 147–152, 2018, doi: 10.1145/3185089.3185150.

[9] P. L. Teh, P. Rayson, I. Pak, S. Piao, and S. M. Yeng, “Reversing the polarity with emoticons,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, vol. 9612, pp. 453–458, doi: 10.1007/978-3-319-41754-7_48.

[10] P. L. Teh, P. Rayson, I. Pak, and S. Piao, “Sentiment analysis tools should take account of the number of exclamation marks!!!,” Dec. 2015, doi: 10.1145/2837185.2837216.

[11] L. Jácome-Galarza, M. S. Viñán-Ludeña, J. Sinche-Freire, J. Charco-Aguirre, and W. Jaramillo-Sangurima, “Word-embedding-based model to improve the customer’ shopping experience,” RISTI - Rev. Iber. Sist. e Tecnol. Inf., vol. 2019, no. 19, pp. 567–580, 2019.

[12] E. Biswas, K. Vijay-Shanker, and L. Pollock, “Exploring word embedding techniques to improve sentiment analysis of software engineering texts,” IEEE Int. Work. Conf. Min. Softw. Repos., vol. 2019-May, no. May, pp. 68–78, 2019, doi: 10.1109/MSR.2019.00020.

[13] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed representations ofwords and phrases and their compositionality,” Adv. Neural Inf. Process. Syst., pp. 1–9, 2013.

[14] A. K. Sharma, S. Chaurasia, and D. K. Srivastava, “Sentimental Short Sentences Classification by Using CNN Deep Learning Model with Fine Tuned Word2Vec,” Procedia Comput. Sci., vol. 167, no. 2019, pp. 1139–1147, 2020, doi: 10.1016/j.procs.2020.03.416.

[15] L. Breiman, “Random Forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[16] K. Alkhatib and S. Abualigah, “Predictive Model for Cutting Customers Migration from banks: Based on machine learning classification algorithms,” 2020 11th Int. Conf. Inf. Commun. Syst. ICICS 2020, pp. 303–307, 2020, doi: 10.1109/ICICS49469.2020.239544.

[17] A. Borg, J. Ahlstrand, and M. Boldt, “Predicting e-mail response time in corporate customer support,” ICEIS 2020 - Proc. 22nd Int. Conf. Enterp. Inf. Syst., vol. 1, no. Iceis, pp. 305–314, 2020, doi: 10.5220/0009347303050314.

[18] B. Mathayomchan and V. Taecharungroj, "'How was your meal?' Examining customer experience using Google maps reviews," Int. J. Hosp. Manag., vol. 90, no. August, p. 102641, 2020, doi: 10.1016/j.ijhm.2020.102641.

[19] S. K. Sharma, M. Daga, and B. Gemini, "Twitter Sentiment Analysis for Brand Reputation of Smart Phone Companies in India," Lect. Notes Electr. Eng., vol. 605, no. September, pp. 841–852, 2020, doi: 10.1007/978-3-030-30577-2_75.

[20] M. A. Al-Shabi, "Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining," Int. J. Comput. Sci. Netw. Secur., vol. 20, no. 1, pp. 51–57, 2020.

[21] S. Ballı and O. Karasoy, "Development of content-based SMS classification application by using Word2Vec-based feature extraction," IET Softw., vol. 13, no. 4, pp. 295–304, 2019, doi: 10.1049/iet-sen.2018.5046.

[22] M. Bilgin and H. Köktaş, "Sentiment analysis with term weighting and word vectors," Int. Arab J. Inf. Technol., vol. 16, no. 5, pp. 953–959, 2019.

[23] N. Wiratunga, S. Massie, K.-Y. Hui, and R. Heaven, "Ontology alignment based on word embedding and random forest classification Ikechukwu Nkisi-Orji 1[0000−0001−9734−9978]," pp. 1–16, [Online]. Available: http://oaei.ontologymatching.org/.

[24] O. Y. Abramov, "'Voice Of the Product' To Supplement 'Voice Of the Customer,'" TRIZfest 2015, vol. 5, no. 5, pp. 1–9, 2015, [Online]. Available: https://www.researchgate.net/publication/286447524%0A'Voice.

[25] Z. Ahmed et al., "Machine learning at microsoft with ML.Net," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., no. June, pp. 2448–2458, 2019, doi: 10.1145/3292500.3330667.

[26] D. Ferreño et al., "Prediction of mechanical properties of rail pads under in-service conditions through machine learning algorithms," Adv. Eng. Softw., vol. 151, no. October 2020, 2021, doi: 10.1016/j.advengsoft.2020.102927.

[27] G. Ge, Z. Du, and J. Yang, "On-machine measurement-based compensation for machining of thin web parts," Procedia Manuf., vol. 48, pp. 844–851, 2020, doi: 10.1016/j.promfg.2020.05.122.

[28] F. V. Ordenes, B. Theodoulidis, J. Burton, T. Gruber, and M. Zaki, "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach," J. Serv. Res., vol. 17, no. 3, pp. 278–295, 2014, doi: 10.1177/1094670514524625.