Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 5 (2022) 840 - 846 ISSN Media Electronic: 2580-0760

Aspect-Based Sentiment Analysis on Twitter Using Logistic Regression with FastText Feature Expansion

Hanif Reangga Alhakiem[1], Erwin Budi Setiawan[2]
[1,2]Informatics, School of Computing, Telkom University
[1]hanzwew@student.telkomuniversity.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Social media has recently been widely used by users, especially Indonesians, as a place to express themselves in sentences, pictures, sounds, or videos. Twitter is one of the social media favored by people of diverse ages. Twitter is a social media that provides features like social media in general. However, Twitter has a unique feature where users can send or read text messages limited to only a few characters. Therefore, user tweets with topics related to a particular product can be utilized by companies to become input in the development of these products. This research was conducted using tweet data on the topic of Telkomsel, which is divided into two aspects, namely signal and service. Aspect-based sentiment analysis of Telkomsel was carried out using Logistic Regression with FastText feature expansion to reduce vocabulary mismatch in tweets so that the classification stage can be performed optimally. In addition, the Synthetic Minority Oversampling Technique (SMOTE) sampling method was applied to overcome data imbalance. The test results prove that feature expansion can improve F1-Score values for signal and service aspects. For the signal aspect, F1-Score increased by 3.33% from the baseline with a value of 96.48%. While for the service aspect, F1-Score increased by 12.91% from the baseline with a value of 95.57%.

Keywords: *aspect-based sentiment analysis, logistic regression, fasttext, feature expansion, twitter*

1. Introduction

Following the development of technology, social media users continue to increase and become more and more active in expressing opinions on issues currently discussed. A wide variety of social media provides a platform for people to express themselves and reveal their true selves conveniently. One of the most widely used social media is Twitter. Tweets written by users appear based on topics that are currently popular. Other users can also respond to the topic without any rules regarding expressing their sentiment, either positive, neutral, or negative[1]. These tweets can become valuable data that provide the possibility to be further analyzed and used as a baseline or reference for any company connected to the topic of the tweet.

Sentiment analysis is a variety of Natural Language Processing (NLP), text analysis, and Computational-Linguistics to analyze information obtained from its source[2]. Sentiment analysis was conducted to detect and classify sentiment based on specific topics contained in tweets that users. posted on Twitter. Sentiment analysis using aspect-based models is needed to obtain more detailed data from product comments which can then be processed as a review for related companies[3]. Sentiment analysis in this study focused on tweet data with topics about Telkomsel. The topic of Telkomsel was chosen because we often find tweets containing complaints about signals and services that have been provided. After all, Telkomsel is the largest cellular service provider with the most users in Indonesia.

Poornima. A. et al.[4], compared the use of several classification methods used in sentiment analysis. The classification methods used are Support Vector Machine (SVM), Logistic Regression, and Multinomial Naïve Bayes. The dataset used is tweets data obtained from Twitter. The three classification methods' accuracy results compared and placed the Logistic Regression method in the first position with an accuracy of 86.23%, SVM in the second position with an accuracy of 85.69%, while Multinomial Naïve Bayes in the third position with an accuracy of 83.54%.

R. Ahuja et al.[5], compared TF-IDF and N-Grams feature expansion methods applied in sentiment analysis using SS-Tweet data with K-Nearest Neighbor (KNN), Decision Tree, Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, and Random Forest classification methods. The test results prove that

Accepted: 16-08-2022 | Received in revised: 17-09-2022 | Published: 31-10-2022

TF-IDF feature expansion performs 3-4% better than N-Grams. In addition, the Logistic Regression method has the most optimal sentiment prediction results in measuring accuracy, recall, precision, and F-score using both TF-ID and N-Grams.

R. Velioglu et al.[6], compared the sentiment analysis with Uni-Gram and TF-IDF against emoticons in tweets using Naïve Bayes, Logistic Regression, Support Vector Machines (SVM), and Decision Tree methods. The best results were obtained by the Logistic Regression method with the F1-Score values of 62% for model-1 and 77% for model-2. Additional comparisons were made to compare the effect of FastText implementation on these classification methods. This comparison shows that FastText can increase the model's accuracy.

Based on the three articles above, it can be concluded that the classification method with consistent results is Logistic Regression. In addition, the selection of feature expansion to be applied in classification can affect performance results because feature expansion is used to reduce vocabulary mismatch. Data selection also has the potential to affect classification results, so we implement a sampling method using the SMOTE method to overcome data imbalance. Then to optimize the Logistic Regression classification process, we apply hyperparameter tuning so that classification can be carried out with the best parameters. This research was conducted because there is no other research that combines these methods. We are expecting that the implementation of Logistic Regression with FastText feature expansion assisted by SMOTE sampling method and hyperparameter tuning can achieve a great result.

The contents of this research paper are structured as follows. Section 2 discusses the method and description of the aspect-based sentiment analysis using Logistic Regression with FastText expansion features. Section 3 contains results and discussions regarding this research. Lastly, Section 4 contains conclusions from the research.

2. Research Methods

In this research, aspect-based sentiment analysis is carried out on Telkomsel data using the Logistic Regression classification method with FastText feature expansion, as shown in Figure 1. This research begins with collecting data from Twitter, then labeling, pre-processing, feature extraction with TF-IDF, implementation of FastText feature expansion, classification using Logistic Regression, and the last is model performance evaluation.

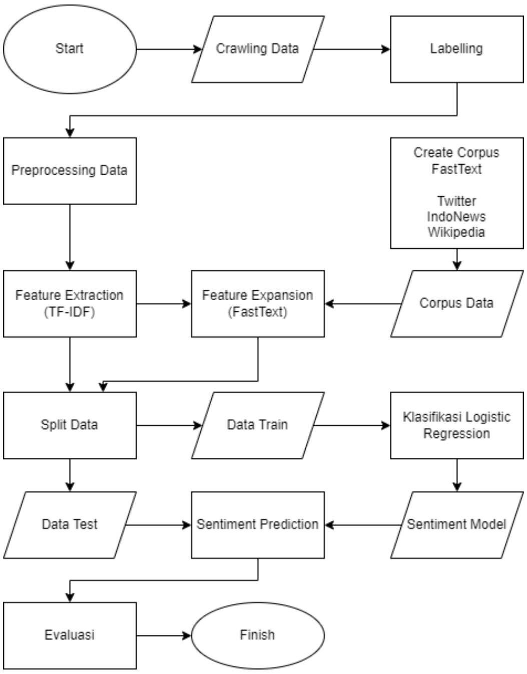


Figure 1. Aspect-Based Sentiment Analysis with Logistic Regression and FastText

2.1 Data Collection

Twitter data crawling is a process of extracting user data and tweets originating from Twitter based on specific keywords[7]. Data extraction is performed by implementing Application Programming Integration (API)[8]. The crawling process was carried out using the SNScrape library provided in the Python programming language. This research uses tweets data that discusses the topic of Telkomsel with the date ranging from September 01, 2021, to February 28, 2022. The keywords table can be seen in Table 1. Tweets are divided into two aspects, namely signal and service, with a total of 16,987 tweets.

Table 1. Telkomsel Data Collection Keywords

No.	Keywords
1	sinyal telkomsel
2	jaringan telkomsel
3	telkomsel lemot
4	kualitas telkomsel
5	telkomsel down
6	telkomsel roaming
7	telkomsel gangguan
8	telkomsel error
9	cs telkomsel
10	grapari
11	pelayanan telkomsel
12	mytelkomsel
13	pelayanan sms telkomsel

2.2 Data Labelling

The next phase is to label the tweet data that has been obtained. Sentiment labeling is done manually by a total of 7 people with the rule that each tweet is labeled by 3 people, so there are 3 opinions regarding the label. Therefore, the final label for each tweet was decided by

DOI: https://doi.org/10.29207/resti.v6i5.4429

the majority vote. Labeling is done based on signal and service aspects, with labels 1 as positive, 0 as neutral, and -1 as negative. An example of labeling is presented in Table 2. The comparison of the amount regarding data labeling for each aspect is presented in Table 3.

Table 2. Data Labelling Example

Tweet	Signal	Service
@Telkomsel kecewa dengan pelayanan telkomsel	0	-1
Sumpah Telkomsel lemot banget plss lelah gua mati idupin data	-1	0
Akhirnya Jaringan @Telkomsel Udah Mulai Bagus	1	0

Table 3. Amount of Label From Each Aspect

Aspect	Positive	Neutral	Negative
Signal	471	6870	9646
Service	1216	13546	2225

2.3. Data Pre-processing

Pre-processing is the systematic process of transforming data into a more structured form so that it can be processed more efficiently. Pre-processing is done with the aim of decreasing the size and changing the data to make word processing easier[4]. There are five steps in this process. The first step is Data Cleaning to clean data from noise, such as the removal of URLs, mentions, symbols, emoticons, hashtags, and numbers. The second step is Case Folding to change all letters into lowercase letters. The third step is Tokenization which converts sentences into words. The fourth step is Stopwords Removal which removes words that do not have any meaning. The last step is Stemming, which converts words into their root words by removing prefixes, infixes, and suffixes.

2.4 TF-IDF Feature Extraction

The feature extraction stage may directly affect the classification accuracy, so this stage is one of the most critical stages. The TF-IDF feature extraction algorithm performs by calculating the weight of words in the text and then displaying it as a vector of defined features[9]. TF is the ratio of word occurrences divided by the total number of words[2]. Furthermore, TF-IDF is the multiplication of TF and IDF values, where TF is the frequency of word occurrence, while IDF is the rarity value of the word from all documents[2]. TF-IDF is represented as Formula (1) and Formula (2).

$$TF = \frac{No.\ of\ times\ word\ w\ occurred}{Total\ No.\ of\ words\ in\ the\ document} \quad (1)$$

$$IDF = \log\left[\frac{No.\ of\ documents}{No.\ of\ documents\ containing\ word\ w}\right] \quad (2)$$

2.5 FastText Feature Expansion

Feature expansion is done after the feature extraction process has been completed. The application of feature expansion in this research utilizes the FastText method with the aim of enriching the word division of the tweets that have been collected to avoid vocabulary mismatch. FastText is an open-source word embedding method developed by Facebook[10]. FastText has advantages in processing foreign languages because it can simplify language processing and reduce dependence on pre-processing data[10]. FastText categorizes each word into a series of vectors where each vector represents n-grams[11]. An additional data, namely IndoNews Corpus, is used to help the feature expansion process. IndoNews consists of Indonesian news media such as Liputan 6, CNN Indonesia, Detik.com, Republika, Kompas, and Tempo, with a total amount of 142,545 articles.

In this research, feature expansion is applied by utilizing the created corpus using FastText to generate semantically similar words. For example, the word "telkomsel" is given to search for other semantically similar words based on the IndoNews corpus. Table 4 displays 10 semantically similar words to "telkomsel" sorted by word similarity level. The first rank is the word with the highest semantic value, while the last rank is the word with the lowest semantic value.

Table 4. Top 10 Rank of Word Similar to "telkomsel"

Rank	Similar Word
1	telkomsel
2	telkomselku
3	msel
4	telkoms
5	telkomselnya
6	telkom
7	mytelkomsel
8	el
9	telk
10	nyesel

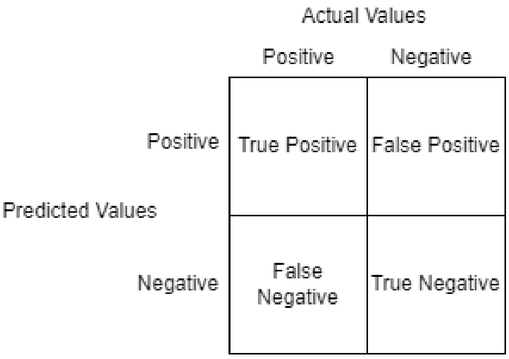
2.6 Logistic Regression

The classification method used in this research is Logistic Regression. This method is used with the aim of determining the sentiment class of the tweet data where positive sentiment is labeled 1, neutral is labeled 0, and negative is labeled -1. Logistic Regression is a method that is often used to predict results based on probability[4]. Logistic Regression is able to work well in overcoming linear relationships between variables[12]. Logistic Regression is a classification method that is very suitable for data with two labels, which are positive and negative, but this method can still be used for data that has more than two labels[13].

2.7 Performance Evaluation

The last stage in this research is the evaluation of the performance from the system that has been built. Performance measurement is carried out using an evaluation metric, namely F1-Score. F1-Score is a harmonic mean of precision and recall so that it can provide a better evaluation for imbalanced data[14].

Precision and Recall values need to be calculated first to be able to produce F1-Score values. Precision, Recall, and F1-Score measurements are carried out by utilizing the confusion matrix.



		Actual Values	
		Positive	Negative
Predicted Values	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Figure 2. Confusion Matrix

Within the confusion matrix presented in Figure 2, there are True Positive, False Positive, True Negative, and False Negative values. True Positive (TP) is a result that is predicted as positive and is actually positive, False Positive (FP) is a result that is predicted as positive but is actually negative, True Negative (TN) is a result that is predicted as negative and is actually negative, False Negative (FN) is a result that is predicted as negative but is actually positive.

Precision is the ratio of the accurate positive prediction results to the total positive predictions. Formula 3 is used to calculate the Precision value.

$$Precision = \frac{TP}{TP + FP} \qquad (3)$$

Recall is the ratio of the accurate prediction results to the positive value of the overall prediction. Formula 4 is used to calculate the Recall value.

$$Recall = \frac{TP}{TP + FN} \qquad (4)$$

F1-Score is a weighted average of the precision and recall values. Formula 5 is used to calculate the F1-Score value.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \qquad (5)$$

3. Results and Discussions

This research is focused on F1-Score results from the classification using Logistic Regression with FastText feature expansion. F1-Score is selected to be a comparative value considering that the data used are imbalanced. There are 4 test scenarios that will be carried out on 2 categories of aspects, namely signal and service. The first scenario compares the F1-Score obtained from Logistic Regression classification using the ratio of training and test data by applying TF-IDF of 70:30, 80:20, and 90:10. The best data ratio and classification results obtained from this scenario will be used as a baseline to be compared with the results of other scenarios. The second scenario test is carried out using SMOTE. SMOTE is capable of overcoming unbalanced data[15]. The third scenario tested the classification results by applying FastText feature expansion. Each corpus data feature will be tested in the form of Top 1, Top 5, Top 10, Top 20, and Top 25. The best results are selected to be used in the following scenario. The last scenario applied hyperparameter tuning to improve the test results. Classification testing is done five times for each scenario, and then the F1-Score results will be averaged to get optimal results.

3.1 Results

The first scenario was conducted to find the baseline by comparing the best classification results from the three combinations of training and test data ratios. The data ratios used are 70:30, 80:20, and 90:10.

Table 5. Best Data Ratio and Baseline

Aspect	Data Ratio	F1-Score (%)
Signal	90:10	93.34
	80:20	**93.37**
	70:30	92.81
Service	90:10	**84.64**
	80:20	83.73
	70:30	83.76

Table 5 shows that the best F1-Score value is obtained from the 80:20 data ratio, which is 93.37% for the signal aspect. While for the service aspect, the best result is obtained from the 90:10 data ratio, which is 84.64%. These two results will be the baseline to compare with the following test scenario.

The second scenario applied SMOTE to overcome the imbalanced data. As seen in Table 6, the implementation of SMOTE was able to increase F1-Score by 2.70% to 95.89% for the signal aspect and an increase of 9.90% to 93.02% for the service aspect.

Table 6. Performance of Logistic Regression with Data Sampling

Aspect	F1-Score (%)
Signal	95.89 (+2.70)
Service	93.02 (+9.90)

The third scenario implemented the FastText feature expansion on all three corpuses. Table 7 and Table 8 show the F1-Score values from this test.

Table 7. Performance of Logistic Regression with FastText Feature Expansion for Signal Aspect

Features	F1-Score (%)		
	Tweet Corpus	IndoNews Corpus	Tweet+IndoNews Corpus
Top-1	95.61 (+2.40)	95.89 (+2.70)	95.89 (+2.70)
Top-5	95.53 (+2.31)	95.95 (+2.76)	94.87 (+1.61)
Top-10	93.65 (+0.30)	95.91 (+2.72)	92.76 (-0.65)
Top-20	89.41 (-4.24)	**96.02 (+2.84)**	89.38 (-4.27)
Top-25	87.74 (-6.03)	95.99 (+2.81)	88.73 (-4.97)

DOI: https://doi.org/10.29207/resti.v6i5.4429

Table 8. Performance of Logistic Regression with FastText Feature Expansion for Service Aspect

Features	F1-Score (%)		
	Tweet Corpus	IndoNews Corpus	Tweet+IndoNews Corpus
Top-1	92.36 (+9.12)	93.11 (+10.01)	93.00 (+9.88)
Top-5	91.63 (+8.26)	**93.15 (+10.05)**	90.90 (+7.40)
Top-10	87.78 (+3.71)	93.02 (+9.90)	85.31 (+0.79)
Top-20	81.50 (-3.71)	93.00 (+9.88)	79.92 (-5.58)
Top-25	79.47 (-6.11)	93.06 (+9.95)	77.66 (-8.25)

Table 7 and Table 8 present the F1-Score value after implementing the feature expansion. The best value on the signal aspect is obtained from the Top-20 feature of IndoNews corpus valued at 96.02%, with an increase of 2.84%. Meanwhile, the best value in the service aspect is obtained from the Top-5 feature of the IndoNews corpus, valued at 93.15%, with an increase of 10.05%. Figure 3 visualizes the comparison of the F1-Score for the signal aspect on each corpus and each Top-n feature. Figure 4 visualizes the comparison of the F1-Score for the service aspect on each corpus and each Top-n feature.

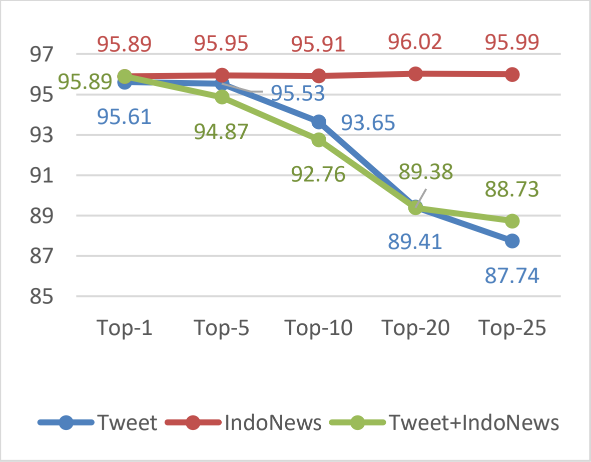


Figure 3. F1-Score Comparison After Feature Expansion for Signal Aspect

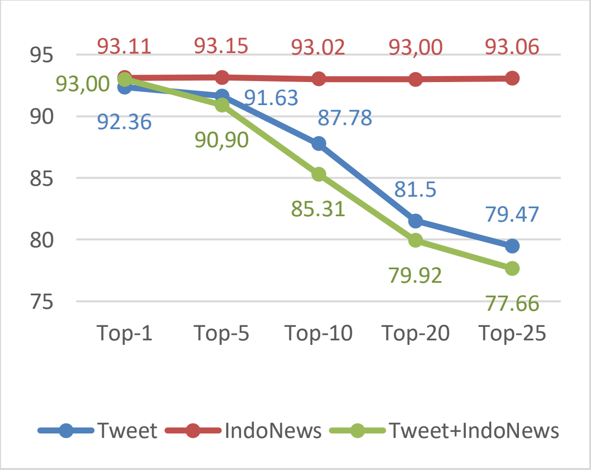


Figure 4. F1-Score Comparison After Feature Expansion for Service Aspect

The last scenario implemented hyperparameter tuning with C and Penalty as the used parameters. The best C and Penalty parameter values are searched using the GridSearchCV method provided in the Python library. The parameters used are listed in Table 9.

Table 9. Parameter for Hyperparameter Tuning

Parameter	Values
C	[-4, 4, 20]
Penalty	[‘L1’, ‘L2’]

Table 10. Best Parameter for Hyperparameter Tuning

Aspect	C	Penalty	F1-Score (%)
Signal	11.29	L2	96.48 (+3.33)
Service	1438.45	L2	95.57 (+12.91)

Table 10 shows the best parameters for each aspect. The C=11.29 and Penalty="L2" values are the best parameters for the signal aspect. While the C=1438.45 and Penalty="L2" values are the best parameters for the service aspect. The F1-Score value for the signal aspect improved by 3.33%, which amounted to 96.48%, and the F1-Score value for the service aspect improved by 12.91%, which amounted to 95.57%.

3.2 Discussion

According to the four test scenarios that have been executed, the application of FastText feature expansion is, in fact, able to improve the F1-Score value of the classification using Logistic Regression. In addition, the use of SMOTE and hyperparamater tuning significantly improved the test results. Table 11 shows the comparison of the results from all four test scenarios.

Table 11. Experiment Result Comparison

Aspect	Scenario	F1-Score (%)
Signal	Baseline	93.37
	Baseline + SMOTE	95.89 (+2.70)
	Baseline + SMOTE + Feature Expansion	96.02 (+2.84)
	Baseline + SMOTE + Feature Expansion + Hyperparameter Tuning	**96.48 (+3.33)**
Service	Baseline	84.64
	Baseline + SMOTE	93.02 (+9.90)
	Baseline + SMOTE + Feature Expansion	93.15 (+10.05)
	Baseline + SMOTE + Feature Expansion + Hyperparameter Tuning	**95.57 (+12.91)**

Table 11 presents the F1-Score values as well as the improvement obtained after comparing with the baseline. The baseline search is the first scenario tested, with results for the signal aspect of 93.37% and 84.64% for the service aspect. Then the implementation of SMOTE data sampling improved the results to 95.89% for the signal aspect and 93.02% for the service aspect. Entering the third scenario, which is the implementation of feature expansion, the results are increased to 96.02% for the signal aspect and 93.15% for the service aspect.

The last scenario, that is, the implementation of hyperparameter tuning, was able to improve the results to 96.48% for the signal aspect and 95.57% for the service aspect. There is a pretty considerable increase over the baseline from each aspect in each test scenario, with a total of 3.33% increase for the signal aspect and 12.91% for the service aspect.

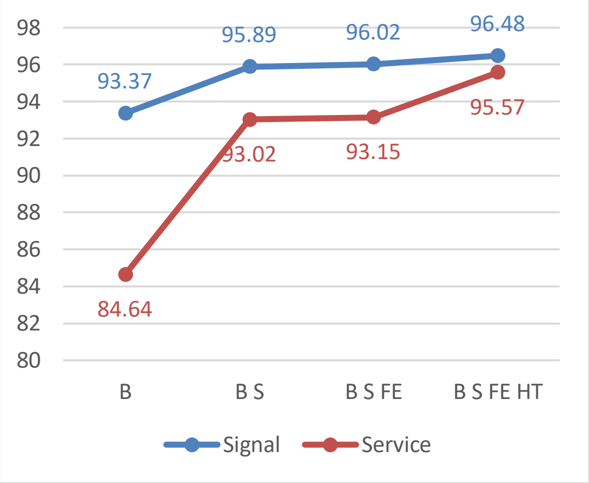


Figure 5. F1-Score Improvement on Each Scenario of Each Aspect

Figure 5 illustrates the results from the table, with the blue line being the signal aspect and the red line being the service aspect. The first scenario is denoted by the letter B, which refers to Baseline. The second scenario is denoted by the letters B S, which refers to Baseline + SMOTE. The third scenario is denoted by the letters B S FE, which refers to Baseline + SMOTE + Feature Expansion. Finally, the fourth scenario is denoted by the letters B S FE HT, which refers to Baseline + SMOTE + Feature Expansion + Hyperparameter Tuning. The line graph illustrates a significant improvement in each aspect, particularly the service aspect.

4. Conclusion

In this research, aspect-based sentiment analysis using Logistic Regression has been performed by applying the FastText feature expansion. The data used is in the form of tweets on Telkomsel with two categories of aspects, namely signal, and service. The amount of data used is 16,987 tweets. In addition, we use IndoNews data in the feature expansion process to help find optimal results. Model testing was conducted based on four scenarios by combining the SMOTE sampling method, Logistic Regression classification, FastText feature expansion, and hyperparameter tuning. The sampling method using SMOTE is able to improve classification results significantly. The application of hyperparameter tuning at the end of the test improved the results quite remarkably. The best result for the signal aspect is obtained from the Top 20 features in the IndoNews corpus, with the F1-Score value equal to 96.48%. Meanwhile, the best result for the service aspect is obtained from the Top 5 features in the IndoNews corpus with the F1-Score value equal to 95.57%. The advice for further research is probably to expand the number of aspects used and combine classification methods with different feature expansions.

Reference

[1] M. Rezwanul, A. Ali, and A. Rahman, “Sentiment Analysis on Twitter Data using KNN and SVM,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 6, pp. 19–25, 2017, doi: 10.14569/ijacsa.2017.080603.

[2] S. Thavareesan and S. Mahesan, “Sentiment Analysis in Tamil Texts: A Study on Machine Learning Techniques and Feature Representation,” *2019 IEEE 14th Int. Conf. Ind. Inf. Syst. Eng. Innov. Ind. 4.0, ICIIS 2019 - Proc.*, pp. 320–325, 2019, doi: 10.1109/ICIIS47346.2019.9063341.

[3] N. Zainuddin, A. Selamat, and R. Ibrahim, “Hybrid sentiment classification on twitter aspect-based sentiment analysis,” *Appl. Intell.*, vol. 48, no. 5, pp. 1218–1232, 2018, doi: 10.1007/s10489-017-1098-6.

[4] A. Poornima and K. S. Priya, “A Comparative Sentiment Analysis of Sentence Embedding Using Machine Learning Techniques,” *2020 6th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2020*, pp. 493–496, 2020, doi: 10.1109/ICACCS48705.2020.9074312.

[5] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, “The impact of features extraction on the sentiment analysis,” *Procedia Comput. Sci.*, vol. 152, pp. 341–348, 2019, doi: 10.1016/j.procs.2019.05.008.

[6] R. Velioglu, T. Yildiz, and S. Yildirim, “Sentiment Analysis Using Learning Approaches over Emojis for Turkish Tweets,” *UBMK 2018 - 3rd Int. Conf. Comput. Sci. Eng.*, pp. 303–307, 2018, doi: 10.1109/UBMK.2018.8566260.

[7] T. D. Dikiyanti, A. M. Rukmi, and M. I. Irawan, “Sentiment analysis and topic modeling of BPJS Kesehatan based on twitter crawling data using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation algorithm,” *J. Phys. Conf. Ser.*, vol. 1821, no. 1, 2021, doi: 10.1088/1742-6596/1821/1/012054.

[8] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, “Feature expansion for sentiment analysis in twitter,” *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 2018-Octob, pp. 509–513, 2018, doi: 10.1109/EECSI.2018.8752851.

[9] R. Dzisevic and D. Sesok, “Text Classification using Different Feature Extraction Approaches,” *2019 Open Conf. Electr. Electron. Inf. Sci. eStream 2019 - Proc.*, pp. 1–4, 2019, doi: 10.1109/eStream.2019.8732167.

[10] S. Shumaly, M. Yazdinejad, and Y. Guo, “Persian sentiment analysis of an online store independent of pre-processing using convolutional neural network with fastText embeddings,” *PeerJ Comput. Sci.*, vol. 7, pp. 1–22, 2021, doi: 10.7717/peerj-cs.422.

[11] B. Athiwaratkun, A. G. Wilson, and A. Anandkumar, “Probabilistic fasttext for multi-sense word embeddings,” *ACL 2018 - 56th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.*, vol. 1, pp. 1–11, 2018, doi: 10.18653/v1/p18-1001.

[12] A. De Caigny, K. Coussement, and K. W. De Bock, “A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees,” *Eur. J. Oper. Res.*, vol. 269, no. 2, pp. 760–772, 2018, doi: 10.1016/j.ejor.2018.02.009.

[13] P. Lauren, G. Qu, J. Yang, P. Watta, G. Bin Huang, and A. Lendasse, “Generating Word Embeddings from an Extreme Learning Machine for Sentiment Analysis and Sequence Labeling Tasks,” *Cognit. Comput.*, vol. 10, no. 4, pp. 625–638, 2018, doi: 10.1007/s12559-018-9548-y.

[14] M. Ibrahim, M. Torki, and N. El-Makky, “Imbalanced Toxic Comments Classification Using Data Augmentation and Deep Learning,” *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 875–878, 2019, doi: 10.1109/ICMLA.2018.00141.

[15] A. Fernández, S. García, F. Herrera, and N. V. Chawla, “SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary,” *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, 2018, doi: 10.1613/jair.1.11192.

DOI: https://doi.org/10.29207/resti.v6i5.4429