Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

Sentiment Analysis Based on Aspects of Telkomsel Users on Twitter Using FastText Feature Expansion and NBSVM Classification

Sukmawati Dwi Lestari, Erwin Budi Setiawan*

Fakultas Informatika, Informatika, Telkom University, Bandung Indonesia
Email: [1]sukmawatilestari@student.telkomuniversity.ac.id, [2*]erwinbudisetiawan@telkomuniversity.ac.id
Submitted: **26/08/2022**; Accepted: **29/08/2022**; Published: **30/08/2022**

Abstrak−Telkomsel merupakan layanan yang banyak digunakan masyarakat Indonesia. Keluhan dari pengguna yang mengacu pada aspek layanan dan sinyal Telkomsel sering dibuat dalam tweet Twitter dengan bahasa yang kasar atau baik. Hal ini dilakukan karena pengguna terus menuntut untuk mendapatkan layanan yang lebih baik. Oleh karena itu, diperlukan teknik analisis sentimen berbasis aspek untuk mengetahui pandangan seseorang terhadap setiap aspek, seperti aspek layanan dan sinyal Telkomsel. Analisis sentimen berbasis aspek merupakan solusi untuk mengetahui pendapat pengguna Telkomsel berdasarkan aspeknya. Dalam implementasinya, metode NBSVM digunakan sebagai model klasifikasi yang terbukti bekerja dengan baik dibandingkan dengan metode lain yaitu MNB dan SVM. Implementasi perluasan fitur FastText dapat mempengaruhi tingkat performansi model, dan hasil terbaik diperoleh pada fitur Top 1 pada aspek sinyal dan Top 5 pada aspek layanan dengan kombinasi korpus Twitter dan berita. Dalam penelitian ini, data yang digunakan tidak seimbang dan telah ditangani dengan menerapkan teknik SMOTE dan AdaBoost pada model perluasan fitur FastText. Berdasarkan hasil pengujian yang telah dilakukan, SMOTE dapat menangani ketidakseimbangan data dibandingkan dengan AdaBoost. Hasil performansi model perluasan fitur FastText setelah diterapkan SMOTE mendapatkan F1-Score 91,24% pada aspek signal dan F1-Score 88,75% pada aspek service.

Kata Kunci: Analisis Sentimen Berbasis Aspek; Ekspansi Fitur; FastText; NBSVM; *Handling Imbalanced Data*

Abstract−Telkomsel is a service that the people of Indonesia widely use. Complaints from users referring to Telkomsel's service and signal aspects are often made in Twitter tweets with harsh or good language. This is done because users continue to demand to get better service. Therefore, an aspect-based sentiment analysis technique is needed to determine a person's view of each aspect, such as Telkomsel's service and signal aspects. Aspect-based sentiment analysis is a solution to find out the opinions of Telkomsel users based on their aspects. In its implementation, the NBSVM method is used as a classification model that is proven to work well compared to other methods, namely MNB and SVM. The implementation of the expansion of the FastText feature can affect the level of performance model, and the best results are obtained in the Top 1 feature on the signal aspect and Top 5 on the service aspect with a combination of Twitter corpus and news. In this study, the data used is unbalanced and has been handled by applying SMOTE and AdaBoost techniques to the FastText feature expansion model. Based on the results of the tests that have been carried out, SMOTE can handle data imbalances compared to AdaBoost. The performance results of the FastText feature expansion model after SMOTE are applied to get F1-Score 91.24% in the signal aspect and F1-Score 88.75% in the service aspect.

Keywords: Sentiment Analysis Based on Aspect; Feature Expansion; FastText; NBSVM; Handling Imbalanced Data.

1. INTRODUCTION

Twitter is one of the most popular online media for people of all ages. According to data released by Statista [1], Indonesia is included in the category of ten countries with the most active Twitter users worldwide, 15.7 million users in July 2021 that can say. It can be understood the importance of Twitter's social media channel as one of the an alternative to marketing media. In addition, it can be used as a forum to accommodate user responses related to certain aspects, such as Telkomsel. PT Telkomsel, Tbk is one of the providers of mobile access service providers that has the largest users in Indonesia as many as 150 million [2]. A fairly good marketing strategy is the main factor in a large number of users of this operator. However, to become a better company and want to continue to grow, Telkomsel is able to accept criticism and suggestions from its users. Complaints from users who refer to aspects of Telkomsel's services and signals are often made in Twitter tweets with harsh or kind language. The written comments can be used as evaluation material to increase Telkomsel user satisfaction and a study is needed to analyze whether the comments are positive, negative or neutral.

Sentiment analysis discusses the behavior of a person or society based on opinions related to certain topics, where these opinions will produce positive, negative, and neutral polarity values [3]. Bing Liu mentions [4], in general, sentiment analysis has three main levels, including Document Level, Sentence Level, and Aspect Level. In this study, Aspect-based [5] is one type of sentiment analysis technique that can be a solution to determine individual views on each aspect such as Telkomsel signals and services. In addition, this aspect-based sentiment analysis is carried out to identify the tendency of user loyalty to the product. The FastText feature expansion is used to help improve accuracy [6] which is based on the skip-gram model.

This research is based on various literature sources taken from previous studies related to research methods and objects. Felia et al. [7] conducted an aspect-based sentiment analysis using the Naïve Bayes and Chi-Square classification methods as selection features. mentions that adding n-gram and TF-IDF to feature selection helps to improve the F1-Score even more. This study obtained the best accuracy results, namely 80.18%, recall 72.49%, precision 77.25% and f1-score as much as 74.73%. The drawback of this study is the distribution of the data used is not balanced and there is no handling imbalanced data in this study. Furthermore, similar research discusses

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

sentiment analysis in the research of A.N Muhammad et al. [8], the results of the study stated that NBSVM method is claimed to have very good performance. In the test, the researcher divides the data into train data and test data with varying scales. A stronger level of accuracy and performance is obtained at a scale of 7:3 or 70% train data and 30% test data. 91% precision, 83% recall, and 87% F1-Score. Then the use of the NBSVM classification method is also found in the study of F. F. Zain et al [8] to classify film reviews. Model evaluation is done by comparing the performance of three methods, namely Nave Bayes, SVM, and NBSVM. Based on the evaluation results, the model gets accuracy as high as 88.8% when using NBSVM with unigram and bigram features, compared to using individual methods, namely NB and SVM.

Erwin *et al*.[6], implement feature expansion in Twitter sentiment analysis. In his research, the classification used is SVM, Logistic Regression and Nave Bayes. They mention that feature expansion is proven to help improve accuracy. The highest accuracy is obtained in the Logistic Regression classification. However, in this study have not implemented word embedding in feature expansion testing. The next research is done by Dimuthu Lakmal *et al*.[9] evaluate the types of word embedding, namely FastText, Word2vec, and GloVe for the Sinhalese language. The evaluation method used is intrinsic and extrinsic evaluation. From the two evaluation processes, FastText with 300 vector dimensions managed to get the best overall accuracy compared to Word2vec and GloVe.

This study aims to determine the effect of feature expansion with FastText and handling imbalanced data on aspect-based Twitter sentiment analysis. A comparison of aspect-based sentiment analysis methods in this study was carried out because each method's performance results can determine high and low aspects. In its design, this research will be built using various methods, namely, Nave Bayes – Support Vector Machine (NBSVM) will be used to classify the model. Feature extraction is carried out using the TF-IDF method based on a predetermined number of features. Then a performance comparison will be made before and after the implementation of the FastText feature expansion to find out the influential features. The number of data classes that is not balanced will affect the level of accuracy of the prediction results of a model. This problem can be handled with the SMOTE [10] technique and AdaBoost [11]. These methods are proven to get good results. Therefore, the authors decided to use all these methods to produce the best model. Based on the knowledge of the author and several studies, there has been no research similar using NBSVM as classification method, TF-IDF as feature extraction, FastText as feature expansion, SMOTE and AdaBoost as handling imbalanced data method, and Telkomsel tweets as the dataset.

2. RESEARCH METODHOLOGY

The system design in this research has several processes as described in Figure 1. The sentiment analysis process begins with collecting Twitter data using a crawling technique. Then the data that has been collected will be labeled positive, neutral, and negative sentiment. Furthermore, the dataset is processed at the preprocessing stage to become data that is more efficient and easy to process for the next process. Feature extraction is used to transform data. The implementation of feature expansion is done by building a FastText model first as a corpus similarity. Then there is a process of splitting the data into training data and test data. The model classification process with NBSVM is carried out on the training data and the test data is processed with the sentiment prediction model. The sentiment model that has been processed is then evaluated to determine its performance. The details of each step will be explained in the following sections.

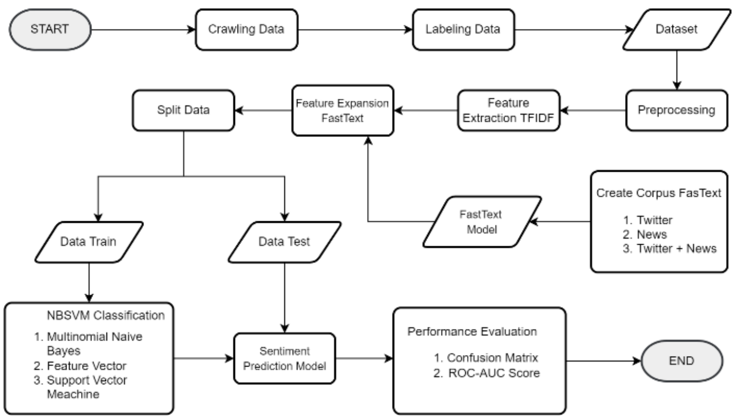

Figure 1 System Design of Sentiment Analysis Based on Aspect

2.1 Crawling Data

The dataset used in this study is Indonesian-language tweet data from Telkomsel users on Twitter. Data crawling is carried out with the snscrape scrapping module using Python 3. In the process, tweet data is retrieved from

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

September 2021 – February 2022 based on 12 keywords written in the Twitter search field. The keywords used are "Sinyal Telkomsel, Jaringan Telkomsel, Kualitas Lemot, Telkomsel Down, Telkomsel Roaming, Telkomsel Gangguan, Telkomsel Error, CS Telkomsel, GraPARI, Pelayanan Telkomsel, MyTelkomsel dan Pelayanan SMS Telkomsel". The total data obtained is 16992 tweets containing user comments.

2.2 Labeling Data

The labeling process was carried out manually by three annotators. If in the process you find a sentence that is difficult to assign a class to, a vote will be made for that sentence. Determination of annotations is determined based on a majority of the votes of the three annotators involved. On the sentiment label, if the text has a positive element, it is labeled "1", if it is negative, it is labeled "-1", and a "0" label is given if the text has a neutral element.The distribution of data obtained from the annotation results is that 9643 tweets are negative, 6878 are neutral, and 471 are positive on the Telkomsel signal aspect. In the service aspect, Telkomsel has a data distribution of 2225 negative tweets, 13547 neutral,and 1216 positive tweets. Illustrasion can be seen at Figure 2.

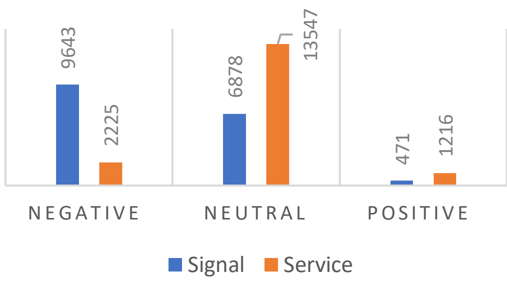

Figure 2 The Distribution of Data Sentiment Analysis Based on Aspect

2.3 Preprocessing Data

In this study, there are six steps used in preprocessing, including data cleaning, case folding, converting slang words (normalization), tokenization, stopword removal, and stemming. (1) Data cleaning is used to remove punctuation marks, numbers, special characters, mention "@", the hashtag "#", URLs, emoji, and so on [12]. (2) Casefolding is the process of converting text from uppercase and capital to lowercase [7]. (3) Convert slang words (Normalization) are used to convert slang words into standard words [13], such as the word "tdk" to "tidak". (4) Tokenization is a technique to change the text in a document in the form of tokens or words [14]. (5) Cleaning the text of words that are considered unimportant such as "dan", "di", "itu", etc. by using a stopword removal technique [7]. Then the process of (6) stemming is carried out with the aim of changing the affixes word into basic. Like the word "menari" is transformed into "tari" [14].

2.4 Feature Extraction TF-IDF

Feature extraction is performed to retrieve features on the dataset. Next will be organized to produce a vector or representation of the boolean (1,0) and assigned to each word [15]. Word weighting is done using the TF-IDF method. Term Frequency (TF) is the frequency with which the same word appears in a document. Inverse Document Frequency (IDF) as the weight of the term by looking at the occurrence of the word in the document [12]. for example there is the text "Saya pergi ke GraPARI" then the words are matched with vector form "saya", "pergi", "servis", "Telkomsel", "GraPARI", "ke". Then the result of the representation of the text is {1, 1, 0, 0, 1, 1}.

2.5 Handling Imbalanced Data

Data imbalance is one of the problems that may occur in data mining and machine learning. The data can be unbalanced if one class in the sample has more numbers than the other classes. SMOTE (Synthetic Minority Oversampling Technique) is one of the most commonly used imbalance data handling techniques. According to the term, SMOTE works with oversampling techniques to make synthetic samples in the minority class[13]. Another unbalanced data handling method is AdaBoost (Adaptive Boosting) which is a boosting technique by maintaining weight in a training data set. Take advantage of weak classifications to create stronger classifications [8]. This study needs SMOTE and AdaBoost techniques because the data used are unbalanced. It is hoped that using these two techniques can help reduce the data imbalance experienced and can help to improve the performance of the built model.

2.6 Feature Expansion with FastText

FastText is a linear-based model word embedding released by Facebook's AI research team [16] which utilizes sub-words using a skip-gram model vector with n-gram characters [17] . According to Mikolov et al. (2013) stated that basically, FastText has a structure similar to the CBOW model. An image of the FastText architectural model

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

can be seen in Figure 3. Besides that, the potential benefits of FastText can result in better word insertion to handle words that are not recognized (out of vocabulary). Feature expansion is the process of developing short text using semantics to look like a large text document [6]. In the expansion of this feature, FastText word embedding is used to deal with word mismatch issues.

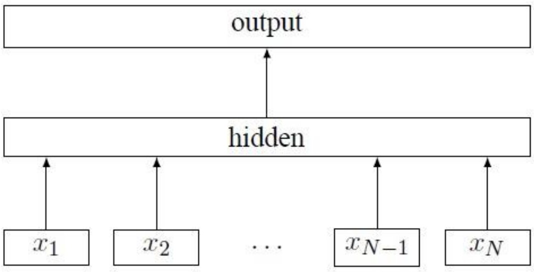

Figure 3 FastText Model Architecture by Mikolov et al.(2013)

In the process, a corpus is needed that contains a collection of words that have value as a dictionary to identify the similarity of words. In this study, the corpus was built from the FastText model using the Gensim library in python with parameters window size 5, min_count size 3, workers size 4, sg size 0, hs size 0. FastText corpus data comes from Twitter data, news, and a combination of Twitter and Twitter data. news. The total data for the construction of the FastText corpus can be seen in Table I.

Table 1 Corpus FastText

Corpus	Total
Twitter	16992
News	168110
Twitter + News	185110

2.7 Naïve Bayes-Support Vector Machine (NBSVM)

The NBSVM classification method is a method formed from the combination of Naïve Bayes and Support Vector Machine [18]. NBSVM is claimed to have excellent performance for handling snippets of documents or full-length documents and has a good level of accuracy[8]. In general, the idea of the NBSVM classification model is a linear classifier to predict the value of k, as in equation (1). Then there is the log-count ratio to calculate the vectors p and 1, which is stated in equation (2).

$$y = sign\left(w^T x^{(k)} + b\right) \quad (1)$$

$$r = \log\left(\frac{p/\|p\|_1}{q/\|q\|_1}\right) \quad (2)$$

In the Multinomial Naïve Bayes classification model, the linear form of the classifier is obtained when the log-ratio (2) that is owned can be transformed into the form (1). The statement produces equation (3). Where is the variable $x^{(k)} = f^{(k)}$, $w^{(T)} = r^{(T)}$ dan $b = \log(N_+ / N\_)$ represent the positives and negatives of the training case.

$$y = sign\left(\log\frac{N_+}{N_-} + r^T f\right) \quad (3)$$

Next, for the SVM classification model, the loss function is minimized with the variable $x^{(k)} = \hat{f}^{(k)}$, variable w, and b. The function loss is a value to measure the distance between the output model and the target. The equation of the minimalist loss function can be seen in formula (4).

$$L(w,b) = w^T w + C\sum \max\left(0,1 - y^{(i)}\left(w^T F^{(i)} + b\right)\right)^2 \quad (4)$$

The Support Vector Machine model with Multinomial Naïve Bayes features uses the ratio of the probabilities of the feature vectors, where the variable $x^{(k)} = \hat{r} \circ \hat{f}^{(k)}$. NBSVM works very well for all documents because of the interpolation of Multinomial Naïve Bayes and SVM which is a regulation of both. This interpolation can provide more optimal accuracy results by using formula (5).

$$w' = (1-\beta)\bar{w} + \beta w \quad (5)$$

The variable w represents the frequency data, $\bar{w}$ that is $\|w\|\ 1/\ |V|$ which states the average magnitude of w.V is an array of words in the vocabulary, and for the interpolation parameter expressed in $\beta \in [0,1]$.

2.8 Performance Evaluation

An evaluation was conducted to measure the performance of the system against classification by calculating accuracy, precision, recall, F1-measurement and ROC AUC score. The confusion matrix is an important

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

component used to develop performance by considering the values of True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) [19]. The calculation of the ROC-AUC score evaluates by considering the True Positive Rate (TPR) and False Positive Rate (FPR) values[20].

3. RESULTS AND DISCUSSION

3.1 Experiment Result

This study has three test scenarios. In the first scenario, a baseline test was carried out on the model to find the optimal value based on the comparison ratio of training data and test data, namely 90:1, 80:20, and 70:30 in each aspect. The second scenario tests TF-IDF feature extraction with the basic model. The third scenario compares the FastText feature expansion model with or without using SMOTE and AdaBoost imbalance data handling techniques

3.1.1 Results of Baseline Test

The baseline test was conducted five times with the NBSVM classification method created by Luis Rei based on the research of Sida Wang, et.al (2012) and has been carried out by Gregoir Mesnil. Results based on the performance that has been written in Table II, the best average F1-score is produced by the 80:20 ratio in the signal aspect. While the optimal value of F1-score in the service aspect is obtained at a ratio of 90:10. The best F1-score value from the second ratio that has been obtained, can be used as a basis for testing the next scenario.

Table 2. Baseline Test

Aspect	Ratio	F1-Score (%)	ROC-AUC (%)
Signal	70:30:00	72,15	82,51
	80:20:00	**72,71**	82,98
	90:10:00	72,4	82,62
Service	70:30:00	60,63	71,44
	80:20:00	61,06	71,5
	90:10:00	**63.71**	71,72

3.1.2 Results of the Effect of TF-IDF on the Baseline Model

Table 3. Find Maximum Feature

Aspect	Evaluation	Max Feature (%)		
		Max 1000	Max 5000	Max 10000
Signal	F1-Score	**73,46**	72	72,31
	ROC AUC	**83,37**	82,45	82,7
Service	F1-Score	**66,2**	62,25	64,5
	ROC AUC	**73,25**	72,2	71,77

The second scenario aims to determine the effect of TF-IDF weighting on the baseline. The application of TF-IDF requires a parameter in the form of a max feature. The researcher has experimented five times to get the best average max feature value. The number of features used is 1000, 5000, and 10000. Based on Table 3, max feature 1000 gets the highest performance among other features in signal and service aspects. It can be seen that the baseline using the max feature parameter is the most useful feature of the other three features to be applied to the TF-IDF weighting. Furthermore, a baseline comparison was made before and after the implementation of TF-IDF feature extraction. In testing this model, three classification methods are used, namely NBSVM, and from these methods, namely MNB, and SVM. Results Based on what is shown in Table 4 for the signal of an aspect, the NBSVM classification method has the best value of the other methods. The F1-Score value obtained is 74.29%, and the ROC-AUC value is 83.68% from the baseline model with TF-IDF weighting. In addition, in terms of the results of the model table, there was a significant increase in the percentage value after applying TF-IDF for all classification methods.

Table 4. Compare Baseline and TF-IDF Signal

Model	Signal Aspect			
	Evaluation	NBSM (%)	MNB (%)	Linear SVM (%)
Baseline	F1-Score	**72,71**	65.88	71.07
	ROC-AUC	**82,98**	79.70	81.56
Baseline + TF-IDF	F1-Score	**74.29 (+ 1.57)**	67.72 (+1.84)	72.38 (+1.31)
	ROC-AUC	**83.68 (+0.02)**	80.20 (+0.5)	82.44 (+0.88)

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

Table 5. Compare Baseline and TF-IDF Service

Model	Service Aspect			
	Evaluation	**NBSM (%)**	**MNB (%)**	**Linear SVM (%)**
Baseline	F1-Score	**63,71**	61.73	61.28
	ROC-AUC	**71,72**	73.33	68.00
Baseline + TF-IDF	F1-Score	**66.67 (+2,96)**	65.18 (+3,45)	65.09 (+ 3.81)
	ROC-AUC	**73.36 (+1,62)**	72.05 (-1,28)	71.08 (+3.08)

The test results on the service aspect shown in Table 5 show that NBSVM obtained the highest performance with F1-Score 66.67% and ROC-AUC 73.36%. However, in the service aspect, the model's performance value after applying TF-IDF weighting experienced a significant decrease by the MNB methods -1,28 for ROC- AUC. Based on the results of previous tests, it can be seen that the combined NBSVM method has the highest performance value among the other two methods. Because this method is the best, it is decided to only use the NBSVM method for further testing.

3.1.3 Applying Feature Expansion with or without Handling Imbalanced Data Technique

Table 6. NBSVM with Feature Expansion FastText Signal

Feature	Signal Aspect with NBSVM and FastText (%)					
	News Corpus		Twittr Corpus		Twitter + News Corpus	
	F1-Score	**ROC-AUC**	**F1-Score**	**ROC-AUC**	**F1-Score**	**ROC-AUC**
Top 1	74,64 (+0,85)	83,81 (+0,15)	74,24 (+0,45)	83,64 (-0,02)	**75,03 (+2,32)**	**83,96 (+0,28)**
Top 5	74,73 (+0,94)	83,85 (+0,19)	74,10 (+0,31)	83,58 (-0,08)	74,56 (+0,77)	83,84 (+0,18)
Top 10	74,12 (+0,33)	83,54 (-0,12)	73,86 (+0,07)	83,34 (-0,32)	73,13 (-0,66)	83,65 (-0,01)
Top 20	74,7 (+0,91)	83,63 (-0,03)	73,13 (-0,66)	82,87 (-0,79)	74,36 (+0,57)	83,72 (+0,06)

The third scenario is to compare the FastText feature expansion model before and after using SMOTE and Adaboost. First, the application of feature expansion in the previous classification model is executed five times for each top-n using a different corpus. The corpus used is the news corpus, the Twitter corpus, and the combined news corpus with Twitter. This corpus contains word similarities that will be used to find the values of Top 1, Top 5, Top 10, and Top 20 to represent the number of words that have the same meaning in the dataset. Based on the results shown in Table 6, the highest F1-score and ROC-AUC values were generated by Top 1 using a combination of Twitter and news corpus, which were 75.03% and 83.96%, respectively, in the signal aspect. When viewed from the second test, the classification model experienced a significant increase in performance after implementing the expansion of the FastText feature both from the F1-Score and ROC-AUC values. However, there is also a decrease in some of its Top-n values.

Furthermore, the same test will be carried out on the service aspect. The results of the service aspect are shown in Table 7. It can be seen that the Top 5 features using a combination of Twitter corpus and news get the best performance. The resulting F1-Score value is 66.70% and ROC-AUC 73.82%. This model has improved performance after implementing the FastText feature expansion of 2.99% for F1-Score and 2.1% for ROC-AUC from the previous test.

Table 7. NBSVM with Feature Expansion FastText Service

Feature	Service Aspect with NBSVM and FastText (%)					
	News Corpus		Twitter Corpus		Twitter + News Corpus	
	F1-Score	**ROC-AUC**	**F1-Score**	**ROC-AUC**	**F1-Score**	**ROC-AUC**
Top 1	66,44 (+2,73)	73,85 (+2,13)	65,74 (+2,03)	73,46 (+1,74)	65,66 (+1,95)	73,23 (+1,51)
Top 5	65,87 (+2,16)	73,16 (+1,44)	64,46 (+2,93)	72,2 (-1,16)	**66,70 (+2,99)**	**73,82 (+2,1)**
Top 10	64,87 (+1,16)	72,45 (+0,73)	64,14 (+0,43)	72,17 (+0,45)	65,92 (+2,21)	73,55 (+1,83)
Top 20	65,99 (+2,28)	73,45 (+1,73)	62,87 (-0,84)	70,87 (-0,85)	65,44 (+1,73)	72,88 (+1,16)

As explained in section 2 poin 2.5, the data used in this study were not balanced. Therefore, in this test, an experimental feature expansion model was carried out by applying SMOTE and AdaBoost techniques to the signal and service aspects. In the signal aspect, the results shown in Table 8 know that the FastText feature expansion model with SMOTE implementation has a higher value than the feature expansion with AdaBoost. The F1-Score value is 91.24% and ROC-AUC 93.47% for the feature expansion model with SMOTE, and the F1-Score value is

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

69.93% and ROC-AUC 81.07% for the FastText feature expansion with AdaBoost. The F1-Score value is 91.24% and ROC-AUC 93.47% for the feature expansion model with SMOTE, and the F1-Score value is 69.93% and ROC-AUC 81.07% for the FastText feature expansion with AdaBoost.

Table 8. Feature Expansion Using Handling Imbalance Data Technique Signal

Signal Aspect (%)		
Model	**Evaluation**	**Twitter + News Corpus Top 1**
Feature Expansion FastText +SMOTE	F1-Score	**91,24 (+16,2)**
	ROC-AUC	**93,47 (+9,5)**
Feature Expansion FastText +AdaBoost	F1-Score	69,93
	ROC-AUC	81,07

Then, in service aspect the results of testing the SMOTE and AdaBoost techniques applied to the FastText feature expansion model can be seen in Table 9. Results Based on what was shown, the SMOTE model obtained better results than the AdaBoost technique. The performance produced by the SMOTE model is F1-Score 88.75% and ROC-AUC 87.64%. Next, the results obtained from the AdaBoost model are 56.39% for the F1-Score and 65.59% for the ROC-AUC.

Table 9 Feature Expansion Using Handling Imbalanced Data Techniques Service

Service Aspect (%)		
Model	**Evaluation**	**Twitter + News Corpus Top 5**
Feature Expansion FastText +SMOTE	F1-Score	**88,75 (+22,05)**
	ROC-AUC	**87,64 (+14,62)**
Feature Expansion FastText +AdaBoost	F1-Score	56,39
	ROC-AUC	65,59

3.2 Discussion

Based on the results of the three test scenarios that have been carried out and are stated in Table 2-9, it shows that the implementation of the FastText feature expansion on the sentiment model with TF-IDF weighting affects improving performance. This is because every data in the model has been transformed into vector form and given weight in the TF-IDF feature extraction process. Then the word in the vector will be replaced with the word that has the greatest similarity value in the corpus. Then from the results of scenario three, it is known that the performance value increases when applying the SMOTE technique to the feature expansion model. This can happen because the SMOTE technique works by utilizing sampling data to create an artificial sample in the minority class so that the data has a balanced amount. Regarding this statement, it can be said that the balance of data can affect the good and bad of a model. Furthermore, Figure 5 states that the performance movement of each experiment on the signal aspect has increased. Starting from the application of TF-IDF in the baseline model, the F1-Score value is 1.57%, then the application of FastText feature expansion is 2.32%, and the application of SMOTE technique to the feature expansion model is 16.2%.

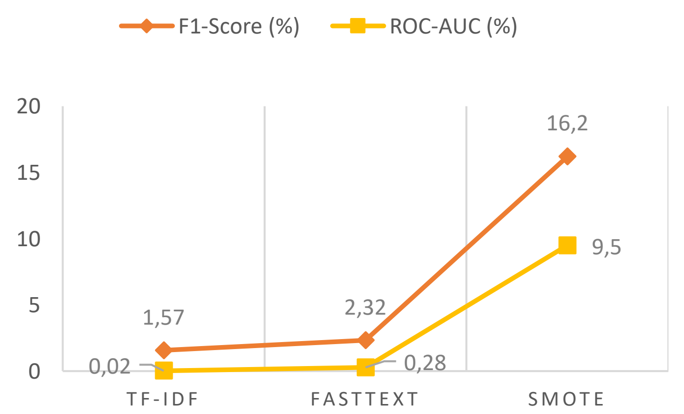

Figure 4. Illustration of Improvement Performance Signal Aspect

Likewise, in the service aspect, as stated in Figure 6, the difference in the F1-Score value is 2.96% when the model applies TF-IDF. Then, 2.99% when applying the FastText feature expansion and 22.02% when the feature expansion model applies the SMOTE technique. Based on the experiment result on illustration of the improvement performance from both aspects, the performance value of the model has increased significantly when the model implements the SMOTE technique.

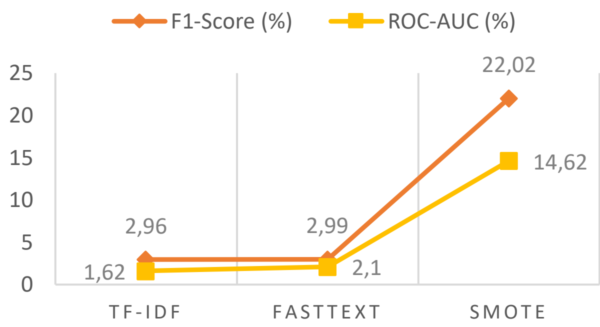

Figure 5. Illustration of Improvement Performance Service Aspect

4. CONCLUSION

This study has conducted an aspect-based sentiment analysis on Telkomsel users using Indonesian-language Twitter data for as many as 16992 tweets. The classification model is made by determining the baseline and the maximum number of features and weighting on the TF-IDF. Then the model is classified by the NBSVM method, and the expansion of the FastText feature is applied to the signal and service aspects. In the classification process, the model is compared using three classification methods, namely NBSVM, MNB, and SVM. NBSVM gets the highest performance value in the signal and service aspects of the three methods before implementing the feature expansion. The expansion of the FastText feature in the model is proven to help improve performance and get the highest results in Top 1 for the signal aspect and Top 5 for the service aspect with a combination of Twitter and news corpus. Uneven distribution of data when labeling makes the data unbalanced. The application of SMOTE is very influential in this aspect-based sentiment analysis research because it has a good performance in balancing data and can increase accuracy significantly compared to the AdaBoost technique. The results obtained are F1-Score 91.24% and ROC-AUC 93.47% in the signal aspect. Then for the service aspect, F1-Score is 88.75%, and ROC-AUC is 87.64%. Based on the experiments conducted in this study, it can be stated that the amount of data distribution can affect the good and bad of the model. In addition, from the results of this study, it can be concluded that Telkomsel users tend to give their sentiments on the signal aspect compared to the service aspect.

REFERENCES

[1] Statista Research Department, “Countries with the most Twitter users 2021,” *Statista*, 2021. https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/#professional (accessed Oct. 31, 2021).
[2] Daon001, “Inilah rincian jumlah pelanggan prabayar masing-masing operator,” *Kominfo*, 2018. https://kominfo.go.id/content/detail/13131/inilah-rincian-jumlah-pelanggan-prabayar-masing-masing-operator/0/sorotan_media (accessed Oct. 31, 2021).
[3] A. Mittal and S. Patidar, “Sentiment analysis on twitter data: A survey,” *ACM Int. Conf. Proceeding Ser.*, pp. 91–95, 2019, doi: 10.1145/3348445.3348466.
[4] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publisher, 2012.
[5] R. Feldman, “Techniques and applications for sentiment analysis,” *Commun. ACM*, vol. 56, no. 4, pp. 82–89, 2013, doi: 10.1145/2436256.2436274.
[6] 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.
[7] F. Novitasari and M. D. Purbolaksono, “Sentiment Analysis Aspect Level on Beauty Product Reviews,” no. January, pp. 18–30, 2021.
[8] A. N. Muhammad, S. Bukhori, and P. Pandunata, “Sentiment Analysis of Positive and Negative of YouTube Comments Using Naïve Bayes-Support Vector Machine (NBSVM) Classifier,” *Proc. - 2019 Int. Conf. Comput. Sci. Inf. Technol. Electr. Eng. ICOMITEE 2019*, vol. 1, pp. 199–205, 2019, doi: 10.1109/ICOMITEE.2019.8920923.
[9] D. Lakmal, S. Ranathunga, S. Peramuna, and I. Herath, “Word embedding evaluation for Sinhala,” *Lr. 2020 - 12th Int. Conf. Lang. Resour. Eval. Conf. Proc.*, no. May, pp. 1874–1881, 2020.
[10] J. Li, H. Li, and J. L. Yu, “Application of Random-SMOTE on imbalanced data mining,” *Proc. - 2011 4th Int. Conf. Bus. Intell. Financ. Eng. BIFE 2011*, pp. 130–133, 2011, doi: 10.1109/BIFE.2011.25.
[11] W. Wang and D. Sun, “The improved AdaBoost algorithms for imbalanced data classification,” *Inf. Sci. (Ny).*, vol. 563, pp. 358–374, 2021, doi: 10.1016/j.ins.2021.03.042.
[12] W. Etaiwi and G. Naymat, “The Impact of applying Different Preprocessing Steps on Review Spam Detection,” *Procedia Comput. Sci.*, vol. 113, pp. 273–279, 2017, doi: 10.1016/j.procs.2017.08.368.
[13] F. S. Alnaz and W. Maharani, “Analisis Emosi Melalui Media Sosial Twitter Dengan Menggunakan Metode Naïve Bayes dan Perbandingan Fitur N-gram dan TF-IDF,” pp. 1–14, 2021.

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 3, No. 4, August 2022, Page 469−477
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v3i4.2202

[14] D. H. K. Al-Khafaji and A. T. Habeeb, "Efficient Algorithms for Preprocessing and Stemming of Tweets in a Sentiment Analysis System," *IOSR J. Comput. Eng.*, vol. 19, no. 3, pp. 44–50, 2017, doi: 10.9790/0661-1903024450.
[15] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison between BM25 and TF-IDF," *2019 Int. Conf. Adv. Sci. Eng. ICOASE 2019*, pp. 124–128, 2019, doi: 10.1109/ICOASE.2019.8723825.
[16] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of tricks for efficient text classification," *15th Conf. Eur. Chapter Assoc. Comput. Linguist. EACL 2017 - Proc. Conf.*, vol. 2, pp. 427–431, 2017, doi: 10.18653/v1/e17-2068.
[17] D. Roy, D. Ganguly, S. Bhatia, S. Bedathur, and M. Mitra, "Using word embeddings for information retrieval: How collection and term normalization choices affect performance," *Int. Conf. Inf. Knowl. Manag. Proc.*, pp. 1835–1838, 2018, doi: 10.1145/3269206.3269277.
[18] S. Wang and C. D. Manning, "Baselines and bigrams: Simple, good sentiment and topic classification," *50th Annu. Meet. Assoc. Comput. Linguist. ACL 2012 - Proc. Conf.*, vol. 2, no. July, pp. 90–94, 2012.
[19] S. Narkhede, "Understanding Confusion Matrix," *towardsdatascience.com*, 2018. https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62 (accessed Dec. 05, 2021).
[20] D. Karani, "What is ROC-AUC and when not to use it," *towardsdatascience.com*, 2022. https://towardsdatascience.com/read-this-before-using-roc-auc-as-a-metric-c84c2d5af621 (accessed Jul. 06, 2022).