Jurnal Teknologi Informasi dan Pendidikan
Volume 17, No. 2, September 2024
https://doi.org/10.24036/jtip.v17i2.889

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Sentiment Analysis of Telkom University using the Long Short-Term Memory and Word2Vec Feature Expansion

Ahmad Alfarel[1*]✉, Hasmawati[1], Bunyamin[1]
[1]School of Computing, Telkom University, Bandung, Indonesia
*Corresponding Author: ahmadalfarel@student.telkomuniversity.ac.id

Article Information

Article history:

No. 889
Rec. August 13, 2024
Rev. December 17, 2024
Acc. December 19, 2024
Pub. December 23, 2024
Page. 468 – 481

Keywords:

- *Sentiment Analysis*
- *Word2Vec*
- *Confusion Matrix*
- *LSTM*
- *Feature Expansion*

ABSTRACT

One of Indonesia's top private universities is Telkom University, and branding is an important aspect of maintaining its reputation. In the digital era, social media has become the main platform for people to express their opinions on various topics, including educational institutions. This research aims to analyze public sentiment towards Telkom University on platform X (formerly Twitter) by using the Long Short-Term Memory (LSTM) method and Word2Vec Feature expansion. The data used consists of 6,627 tweets collected between November 2022 and November 2023. Sentiments were categorized into "Positive," "Negative," and "Neutral". The research stages include data collection, preprocessing, feature extraction using TF-IDF, and feature expansion with Word2Vec. The research results evaluated by calculating accuracy, F1-Score, Precision, and Recall with the help of a confusion matrix. There is a very severe data imbalance in Negative sentiment compared to other sentiments. By doing SMOTE oversampling, feature extraction, and also feature expansion combined with LSTM, the best results are obtained with 91% accuracy, 91% F-1 Score, 91% Precision, and 91% Recall. These results can help Telkom University understand public perception and manage its brand image more effectively.

How to Cite:

Alfarel, A., et al. (2024). Sentiment Analysis of Telkom University using the Long Short-Term Memory and Word2Vec Feature Expansion. Jurnal Teknologi Informasi Dan Pendidikan, 17(2), 468-481. https://doi.org/10.24036/jtip.v17i2.889

1. INTRODUCTION

One of Indonesia's top private universities is Telkom University. Being a top university means that branding matters. It is vital to understand what the general public thinks of Telkom University in order to uphold this. In this digital age, the public shares its knowledge and opinions about various topics on social media platforms [1]. There are several public opinions regarding Telkom University in social media, particularly about Application X. X, or formerly Twitter, providing a platform for the public to share knowledge, whether it be textual, audio, or video related to any organization, product, toy, or other topic. X can also be a useful tool for research because it contains a variety of emphases from the general public through text publishing. This explains why research on sentiment analysis has become rather popular [1].

Sentiment analysis is a text mining technique that assesses and extracts subjective data to help needy individuals [2]. Sentiment analysis can offer insightful information about how students and alumni feel about a specific university, like Telkom University, in the context of education. Researchers from the past have studied university sentiment. Research on public opinion of Telkom University based on postings shared on LinkedIn social media with positive, negative, and neutral categories was carried out in 2022 by Prakoso et al. [3]. The study also seeks to ascertain how well the Random Forest approach performs.

As Ryanto et al. demonstrated, sentiments voiced on social media can impact the public's image and perception of an educational institution [4]. Therefore, sentiment analysis allows us to obtain various information about the sentiments expressed on social media. In 2021, Wibowo et al.[5] talked about using Word2Vec and LSTM as techniques for sentiment analysis of hotel evaluations in the Traveloka application, using a total of 2500 review data. The model integrates effectively; LSTM helps manage the word order in the review, while Word2Vec offers a superior vector representation of words, enabling it to overcome the issue of language complexity with an accuracy of 85.96%. We have chosen to employ the LSTM and Word2Vec algorithms for sentiment analysis in the X application since it is evident that their combination performs sentiment analysis with a reasonably high degree of accuracy.

In 2022, Mollah [1] used 27,056 data from the combination of 7 dataset to study the application of LSTM as a technique for sentiment analysis on Twitter social media. Because there are no emojis in the Word Embedding method that was employed during testing, the accuracy result obtained with 4,459 data is 68%. We may conclude that while LSTM is a pretty decent tool for sentiment analysis, there is still a need for improvement in terms of its word embedding techniques.

Research was done using sentiment analysis on social media X to compare and evaluate four universities in Michigan: Wayne State University, Michigan State University, Oakland University, and University of Michigan by Zohdy et al. in 2019 [6]. Naïve Bayes, Support Vector Machine, KNN, and Decision Tree were the methods used.

Pane and Ramdan [7] carried out additional research on sentiment analysis using Twitter social media in relation to the government's implementation of the Enforcing Restrictions on Community Activities (PPKM) policy in response to the COVID-19 pandemic. The 2,176 data are ready to be used for modeling after the preparation step. Using the LSTM model, a 94.3% accuracy rate was achieved when classifying sentiment as positive or negative. We may conclude that the LSTM model performs a good job of classifying the community's attitude on Twitter or what is now termed X.

Given the foregoing context, the purpose of this study is to categorize public sentiment on X social media regarding Telkom University into three categories: "Positive," "Negative," and "Neutral". LSTM and Word2Vec combined with TF-IDF for feature extraction, and also oversampling using SMOTE to overcome data imbalance in the dataset used.

2. RESEARCH METHOD

2.1. Research Stages

The study procedure has multiple stages, which include Data Crawling and Labelling, Data Preprocessing, Feature Extraction, Building Word2Vec Language Model, Feature Expansion, Building LSTM Model, and Evaluation, as illustrated by the flowchart in Figure 1.

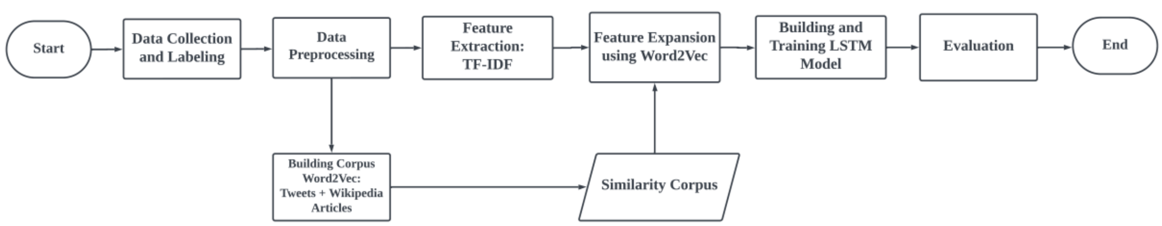

Figure 1. System Flowchart

2.2. Data Collection and Labeling

The dataset used is data derived from X social media using the keywords "Telkom University", "Tel-u", "telyu", and "Telkom University". Data taken in the period of November 2022 to November 2023. By using the Data Crawling method and the Application Programming Interface (API) of the X application, the crawling method uses the tweet-harvest library. Data from crawling results totaling 6,627 tweets will be stored in Comma Separated Value (CSV) format. Furthermore, the data will enter the labeling stage with 3 labels, namely Negative (-1), Positive (1), and Neutral (0) (Table 1). The labeling process is

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

carried out by 3 people, if there are different opinions on the label that will be given to the data, then a vote will be held to equalize the label of the data.

Table 1. Data Tweets

Full Text	Label
Imbas acara wisuda Telkom University, macetnya hingga 3km sebelum pintu keluar tol buahbatu https://t.co/qztMiWm557	-1
telyu! eh ini pbb masih mati lampu kah??	0
Memang butuh banyak pertimbangan Tapi aku lebih memilih yg pasti aja dan universitas telkom margusss jadi why not	1

2.3. Data Preprocessing

At this stage, the data that has been labeled will enter the preprocessing stage, tweets will be subjected to data cleaning, case folding, tokenization, and normalization before entering the next stage (Table 2).

Table 2. Preprocessing

Preprocessing	Tweets
Original Tweets	Telkom University berhasil menjadi PTS Terbaik Berkelanjutan versi UI Green Metric World University Rankings 2022 #TelkomUniversity #KampusSwastaTerbaik #CreatingTheFuture #ContributetoTheNation #teluproud https://t.co/91bTKLk4z1
Data Cleaning	Telkom University berhasil menjadi PTS Terbaik Berkelanjutan versi Green Metric World University Rankings TelkomUniversity KampusSwastaTerbaik CreatingTheFuture ContributetoTheNation teluproud
Case Folding	telkom university berhasil menjadi pts terbaik berkelanjutan versi green metric world university rankings telkomuniversity kampusswastaterbaik creatingthefuture contributetothenation teluproud
Tokenization	['telkom', 'university', 'berhasil', 'menjadi', 'pts', 'terbaik', 'berkelanjutan', 'versi', 'green', 'metric', 'world', 'university', 'rankings', 'telkomuniversity', 'kampusswastaterbaik', 'creatingthefuture', 'contributetothenation', 'teluproud']
Normalization	['telkom', 'university', 'berhasil', 'menjadi', 'pts', 'terbaik', 'berkelanjutan', 'versi', 'green', 'metric', 'world', 'university', 'rankings', 'telkomuniversity', 'kampusswastaterbaik', 'creatingthefuture', 'contributetothenation', 'teluproud']

2.4. Feature Extraction

A document's maximum and minimum term frequency can be found via feature extraction [8], which helps to find other essential textual information. Term Frequency-Inverse Document Frequency (TF-IDF) is one technique that is frequently employed in the feature extraction stage.

Term frequency (TF) measures a word's frequency of occurrence in a document and indicates the word's significance to the content [9]. On the other hand, IDF quantifies the extent to which a word appears in a document. Equation 1 [8] shows how to calculate TF-IDF.

$$(\mathrm{T}F - IDF)_{ij} = (TF)_i \times \log{(IDF)_{ij}} \quad (1)$$

Where:

$(\mathrm{T}F - IDF)_{ij}$: TF-IDF Weight
$(\mathrm{T}F_i)$: Term Frequency
$(IDF)_{ij}$: Inverse Document Frequency

2.5. Word2Vec

Word2Vec is a word embedding technique developed by Mikolov et al. in 2013 to represent words including their meaning and context in a document, there are two approaches, namely Continuous Bag-of-Word (CBOW) and skip-gram algorithms [10]. Both strategies make use of the same network structure, but they depend on distinct input and output factors [11].

When calculating word vector representations, Word2Vec can offer an effective implementation of the CBOW and Skip-Gram designs. Several language processing activities can make use of this structure. Whereas the Skip-Gram architecture predicts terms that surround the target word, the CBOW architecture predicts target words depending on their context [12]. Figure 2 [12] displays the architectural layout of the two types of word2vec.

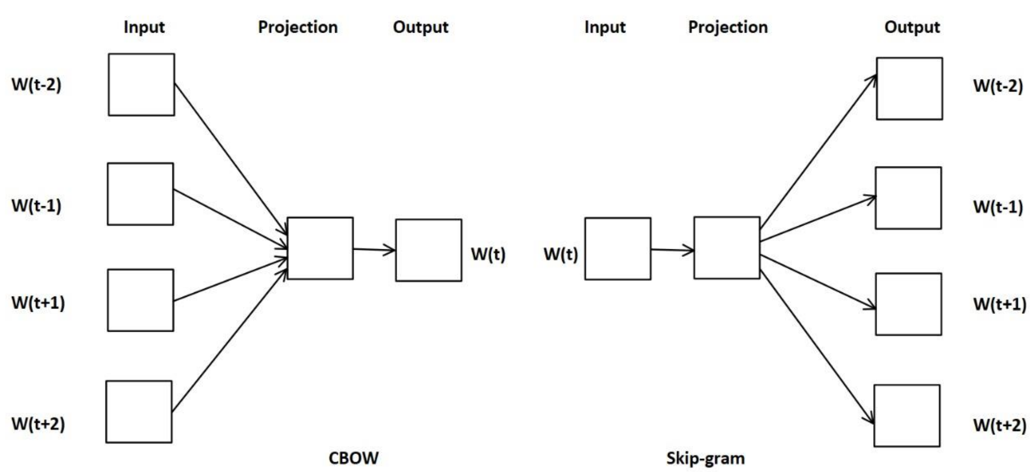

Figure 2. Word2Vec Architecture

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

2.6. Feature Expansion

Feature Expansion is used to add new words related to the feature itself [13]. In this research, the author uses word2vec to perform Feature Expansion and create a language model using Wikipedia language model data of Indonesian articles, then the language model trained using tweets data that has gone through the previous preprocess stage. The language model that has been created will be combined with the results of vectorized TF-IDF at the Feature Extraction stage. As a result of the vectorizer performed at the feature extraction stage, the vector values at each iteration will be multiplied by the word2vec embedding of the feature. This will expand the vector with embedding information.

2.7. Long Short-Term Memory

LSTM is an improved variation of Recurrent Neural Network (RNN) that has internal memory and multiplication. LSTM is designed to overcome the problem of exploding/disappearing gradients when learning long-term information [14].

LSTM consists of memory cells, input gates, output gates, and forget gate [14]. The four parts have their respective functions. An essential component of LSTM is the memory cell, which is used to store and preserve data during training. Input gates control the amount of new data that is stored in the memory cell, while output gates control the amount of data that is extracted from the memory cell to generate output. Finally, forget gates control the amount of previously stored data that needs to be removed. The architecture of LSTM can be seen in Figure 3 [14].

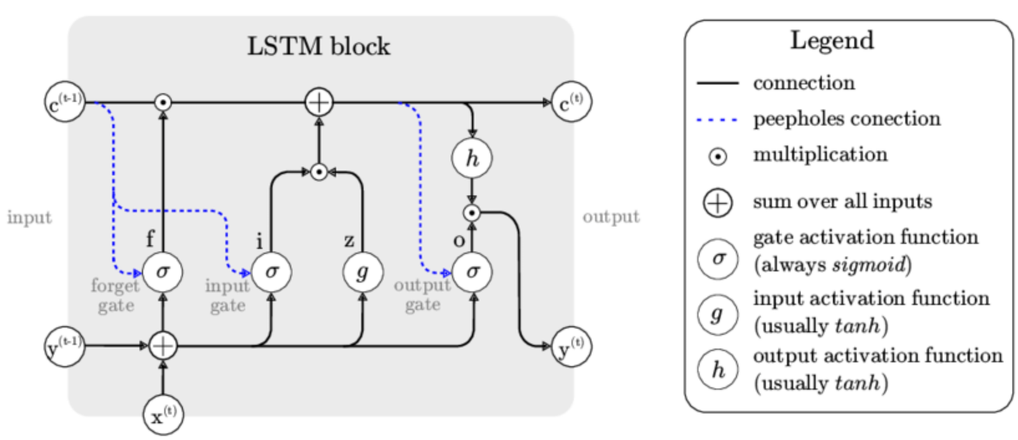

Figure 3. LSTM Architecture

Some steps must be taken in LSTM. The main step in LSTM is to determine whether the incoming input can pass through the cell state or not. This is the role of the forget gate, if the resulting decision is 1 then the data will be stored, while if the decision result is 0 then the information will be deleted. The way to calculate the value of the forget gate (f_t) can be seen in equation (2) [15].

$$f_t = \sigma(w_f[y_{t-1}, x_t] + b_f) \tag{2}$$

Where:

f_t : Forget gate
σ : Sigmoid
w_f : Weight
x_t : Input
y_{t-1} : Hidden state
b_f : Bias

Furthermore, it determines the new information stored in the cell state. The input gate plays a role in determining which part will be updated. Then the g layer or usually tanh makes the new vector, $\tilde{C}_t$ can be added to the cell state. After the determination, the two can be combined for the state update. The way to determine the input gate value (i_t) and the new candidate value ($\tilde{C}_t$) can be seen in equations (3) (4).

$$i_t = \sigma(w_i[y_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh\,(w_c[y_{t-1}, x_t] + b_c) \tag{4}$$

Where:

i_t : Input gate
$\tilde{C}_t$: Candidate values
σ : Sigmoid
b_i : Bias
x_t : Input
w_i : Weight
y_{t-1} : Hidden state

The next step is to update the old cell state to the new cell state. Perform multiplication between the old cell state and the forget gate (f_t) then add $i_t * \tilde{C}_t$. It can be seen more clearly in equation (5).

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{5}$$

Where:

i_t : Input gate
$\tilde{C}_t$: Candidate values
C_{t-1} : Old cell state
f_t : Forget gate
C_t : New Cell state

Choosing the output that the output gate will produce is the final stage. To ascertain which portion of the cell state will be output, run the sigmoid known as the output gate (o_t). Next, pass the cell state through tanh and multiply the result by the sigmoid's output (o_t). The calculation can be seen in Equation (6) (7).

$$o_t = \sigma(w_o[y_{t-1}, x_t] + b_o) \quad (6)$$

$$y_t = o_t * \tanh(C_t) \quad (7)$$

Where:

o_t	: Output gate	x_t	: Input
σ	: Sigmoid	y_{t-1}	: Hidden state
w_o	: Weight	b_o	: Bias
y_t	: New hidden state	C_t	: Cell state

2.8. Matrix Evaluation

The system is evaluated using a confusion matrix. Confusion matrix are widely used for scientific model evaluation in several fields, such as computer vision and natural language processing [16].

True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) are the four potential classification outcomes. When the positive prediction findings match the actual data, there are True Positive (TP) situations. When a positive prediction result is made but a negative result happens, this is known as a false positive (FP). A True Negative (TN) is when the outcomes of the negative prediction coincide with the actual events. The final type of result is False Negative (FN), which happens when a prediction is made and the actual result is positive [17]. For a better understanding can be seen in Table 3 [17].

Table 3. Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

We can determine the values of accuracy, precision, recall, and F1 Measure based on the classification results of the confusion matrix. This can be calculated using the formula below:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$F1-Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (11)$$

3. RESULTS AND DISCUSSION

This section contains the results of the research that has been done, in the form of experiments that have been carried out using the LSTM, Word2Vec, Feature Extraction, and Feature Expansion methods.

3.1. Dataset

The data generated after the crawling and labeling process is 6627, but after the preprocessing stage in accordance with Table 2 which includes data cleaning, tokenization, and normalization. The data was reduced to 5052 data. The distribution of data that has gone through the preprocessing stage can be seen in Table 4.

Table 4. Dataset after Preprocessing

Sentiments	Label	Amount of Data
Positive	1	1459
Neutral	0	3481
Negative	-1	112

Based on Table 4, it can be seen that there is a very uneven distribution of data, this is due to the tweets taken, one of which uses the keyword "telyu" where most of the data taken from the "telyufess" account, which is an unofficial account from Telkom university which contains outpourings about someone. This causes a lot of Neutral sentiment, namely 3481 data. Due to imbalanced data, this research will conduct an oversampling experiment on existing data. The Synthetic Minority Over-sampling Technique (SMOTE) is an oversampling method by creating new synthetic samples from minority classes that are combined with nearby samples [18]. The results of the dataset that has gone through the oversampling stage using SMOTE, the distribution of data can be seen in Table 5.

Table 5. Dataset after SMOTE

Sentiments	**Label**	**Amount of Data**
Positive	1	3,481
Neutral	0	3,481
Negative	-1	3,481

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

3.2. Classification Results

We will conduct multiple tests and discuss the outcomes in this part regarding LSTM testing. Initially, split data, oversampling, feature extraction, and Word2Vec feature expansion are performed. Table 6 displays the outcomes of the tests that were carried out.

Table 6. Classification Results

Method	Split	Imbalanced Data				Balanced Data			
		Accuracy	F1-Score	Precision	Recall	Accuracy	F1-Score	Precision	Recall
Word2Vec + LSTM	60:40	80%	79%	78%	80%	66%	66%	68%	66%
	70:30	82%	81%	80%	81%	69%	69%	69%	69%
	80:20	**82%**	**81%**	**80%**	**82%**	71%	71%	71%	71%
	90:10	82%	80%	80%	82%	71%	70%	73%	71%
Word2Vec Feature Expansion + LSTM	60:40	82%	81%	80%	82%	**83%(+1%)**	**82%(+1%)**	**82%(+2%)**	**83%(+2%)**
	70:30	82%	81%	80%	82%	**84%(+2%)**	**84%(+3%)**	**84%(+4%)**	**84%(+2%)**
	80:20	83%	81%	80%	83%	**91%(+8%)**	**91%(+10%)**	**91%(+11%)**	**91%(+8%)**
	90:10	**85%**	**84%**	**83%**	**85%**	**88%(+3%)**	**88%(+4%)**	**88%(+5%)**	**88%(+2%)**

By using Word2Vec as word embedding, and LSTM to perform classification, the results obtained are quite satisfying with the greatest accuracy at 82%, the accuracy is obtained from split data 90:10, 80:20, 90:10, and also not oversampling using SMOTE. The final result shows that the combination of using word2vec and LSTM experiences a considerable decrease after oversampling. The resulting evaluation matrix after oversampling is 66% for the lowest, and 71% for the highest. This decrease in accuracy also occurs in research [18].

3.2.1. Feature Expansion

Before feature expansion, we must go through the Feature Extraction stage first, in this research, the feature extraction technique used is TF-IDF. The dataset from the preprocessing results will be converted into a vector matrix using TfIdfVectorizer, with parameters max features = 300, smooth_idf = False, and ngram with a range of 1-4. Creating language model is very important at this stage, we create language model with corpus data from Wikipedia, and the finished language model will be combined with the dataset from application X which has been previously crawled and has gone through the preprocessing stage. The word2vec language model is created with vector_size = 300, Skip-Gram, window = 5, and the minimum occurrence of a word is 5 occurrences.

The vector value in TF-IDF will be multiplied by the vector value in the Word2Vec language model. The results obtained from Feature Expansion can be seen in Table 6, where there is no reduction in all evaluation matrices used, instead, it tends to increase and stabilize. Before Feature Expansion, the best classification results obtained with 60:40-90:10 split data experiments and without oversampling were 82% accuracy, 81% F-1 Score, 80% Precision, and 82% Recall, while the best classification results after Feature Expansion and without oversampling were 85% accuracy, 84% F-1 Score, 83% Precision, and 85% Recall. There is an additional value of 2%-3% in each evaluation matrix.

Next are the classification results using Feature Expansion and data balancing using SMOTE. The best overall result is 91% in Accuracy, F-1 Score, Precision, and Recall, when compared to the results before SMOTE, there is an additional of 6%-8%. When compared to the best results before Feature Expansion, the addition of 9%-11%. This proves that Feature Expansion affects improving the performance of text classification [13].

3.2.2. Oversampling Results

Due to the small number of negative sentiments obtained when crawling with the keywords "Telkom University", "Tel-u", "telyu", and "Telkom University" in the November 2022-November 2023 timeframe in Application X, it causes severe imbalanced data, which can be seen in Table 4. When testing with split data 80:20, negative sentiment at the time of testing amounted to only 27 data, although at the time of imbalanced data the model created provided a pretty good matrix evaluation, but if we look at Table 7, the results of negative sentiment testing get 0% in each matrix evaluation, this is due to imbalanced data problems [19]. Therefore, oversampling is done to overcome this problem, SMOTE which is one of the best methods available at this time [20]. It can be seen in Table 7 that SMOTE has an effect on improving the performance of the LSTM model [21]. The evaluation matrix for other sentiments has also grown. Previously, there was only 0% negative sentiment, due to oversampling, this number has climbed to 96%, 93%, and 100%. In neutral sentiment, there is no significant difference in the evaluation results because the amount of data before and after oversampling is not much different, while in positive sentiment there is a significant increase due to the difference of 407 data before and after oversampling. However, this contradicts the research results from [22] with undersampling results that are more suitable for the LSTM model.

Table 7. Classification Results

Method	Label	Imbalanced Data				Balanced Data			
		Amount	F1-Score	Precision	Recall	Amount	F1-Score	Precision	Recall

Word2Vec Feature Expansion + LSTM	-1	27	0%	0%	0%	696	**96%**	**93%**	**100%**
	0	680	85%	**93%**	88%	682	**88%**	92%	**88%**
	1	304	78%	69%	73%	711	**89%**	**89%**	**89%**

3.3. Discussion

In this study, we successfully implemented Long Short-Term Memory (LSTM) network combined with Word2Vec for sentiment analysis of tweets about Telkom University. Analyzing a dataset of 6,627 tweets, this research categorizes sentiment into “Positive”, “Negative”, and “Neutral”. The preprocessing steps, feature extraction using TF-IDF, and feature expansion through Word2Vec significantly improved the model performance, achieving 91% accuracy, up from 82%.

This study demonstrates a notable improvement in accuracy due to the integration of Word2Vec and the application of the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance [6]. While earlier studies achieved similar methodologies, the current research highlights the effectiveness of combining multiple techniques (TF-IDF, Word2Vec, and SMOTE) to enhance sentiment classification performance. This comprehensive approach not only improves accuracy but also provides deeper insights into public sentiment dynamics, marking a significant contribution to the field of sentiment analysis in social media contexts [3][4].

4. CONCLUSION

This research successfully implements Long Short-Term Memory (LSTM) and Word2Vec for sentiment analysis toward Telkom University on X social media platform. By analyzing 6,627 tweets, this research classifies sentiment into three categories: "Positive," "Negative," and "Neutral" with promising accuracy. The preprocessing, feature extraction, oversampling, and feature expansion steps proved to improve the performance of the LSTM model with 91% accuracy. The results show that the combination of LSTM and Word2Vec is a powerful approach to performing sentiment analysis. From this research, it can be seen that the negative sentiment towards Telkom University on the X social media platform in the November 2022-November 2023 timeframe with the keywords "Telkom University", "Tel-u", "telyu", and "Telkom University" is very little when compared to other sentiments such as neutral and positive as can be seen in Table 4. These results can help Telkom University understand public sentiment, manage brand image more effectively, and make better strategic decisions.

Future research can explore other word embedding techniques such as fast-text, conduct deeper exploration of the effect of stopword removal and stemming on the language model, or look for data sources that are more balanced and do not show too far imbalanced data in each class on other popular social media platforms such as Instagram.

REFERENCES

[1] M. Parvez Mollah, "An LSTM model for Twitter Sentiment Analysis," *arXiv e-prints*, p. arXiv-2212, 2022.

[2] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 5, no. 2, pp. 697–711, Sep. 2021, doi: 10.30645/J-SAKTI.V5I2.369.

[3] I. B. Prakoso, D. Richasdy, and M. D. Purbolaksono, "Sentiment Analysis of Telkom University as the Best BPU in Indonesia Using the Random Forest Method," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, pp. 2050–2055, Oct. 2022, doi: 10.30865/mib.v6i4.4567.

[4] S. A. Ryanto, D. Richasdy, and W. Astuti, "Partner Sentiment Analysis for Telkom University on Twitter Social Media Using Decision Tree (CART) Algorithm," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, pp. 1940–1948, Oct. 2022, doi: 10.30865/mib.v6i4.4533.

[5] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews," *Procedia Comput Sci*, vol. 179, pp. 728–735, Jan. 2021, doi: 10.1016/J.PROCS.2021.01.061.

[6] R. Al Bashaireh, V. Sabeeh, and M. Zohdy, "Towards a new indicator for evaluating universities based on twitter sentiment analysis," *Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019*, pp. 1398–1404, Dec. 2019, doi: 10.1109/CSCI49370.2019.00261.

[7] S. F. Pane and J. Ramdan, "Pemodelan Machine Learning : Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Menggunakan Data Twitter," *Jurnal Sistem Cerdas*, vol. 5, no. 1, pp. 12–20, May 2022, doi: 10.37396/JSC.V5I1.191.

[8] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison between BM25 and TF-IDF," *2019 International Conference on Advanced Science and Engineering, ICOASE 2019*, pp. 124–128, Apr. 2019, doi: 10.1109/ICOASE.2019.8723825.

[9] D. Atika, A. Ari Aldino, S. Informasi, J. Pagar Alam No, L. Ratu, and K. Kedaton, "Term Frequency-Inverse Document Frequency Support Vector Machine Untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Mental Pada Media Sosial Twitter," *Jurnal Teknologi dan Sistem Informasi*, vol. 3, no. 4, pp. 86–97, 2022, Accessed: Aug. 01, 2024. [Online]. Available: https://jim.teknokrat.ac.id/index.php/sisteminformasi/article/view/2054

[10] B. Jang, I. Kim, and J. W. Kim, "Word2vec convolutional neural networks for classification of news articles and tweets," *PLoS One*, vol. 14, no. 8, p. e0220976, Aug. 2019, doi: 10.1371/JOURNAL.PONE.0220976.

[11] G. Di Gennaro, A. Buonanno, and F. A. N. Palmieri, "Considerations about learning Word2Vec," *Journal of Supercomputing*, vol. 77, no. 11, pp. 12320–12335, Nov. 2021, doi: 10.1007/S11227-021-03743-2/FIGURES/7.

[12] D. Jatnika, M. A. Bijaksana, and A. A. Suryani, "Word2Vec Model Analysis for Semantic Similarities in English Words," *Procedia Comput Sci*, vol. 157, pp. 160–167, Jan. 2019, doi: 10.1016/J.PROCS.2019.08.153.
[13] A. R. Royyan and E. B. Setiawan, "Feature Expansion Word2Vec for Sentiment Analysis of Public Policy in Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 78–84, Feb. 2022, doi: 10.29207/RESTI.V6I1.3525.
[14] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif Intell Rev*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/S10462-020-09838-1/TABLES/1.
[15] S. F. Handayani *et al.*, "Analisis Sentimen pada Data Ulasan Twitter dengan Long-Short Term Memory," *JTERA (Jurnal Teknologi Rekayasa)*, vol. 7, no. 1, pp. 39–46, Jun. 2022, doi: 10.31544/jtera.v7.i1.2022.39-46.
[16] D. Krstinić, M. Braović, L. Šerić, and D. Božić-Štulić, "Multi-label classifier performance evaluation with confusion matrix," *Computer Science & Information Technology*, vol. 1, 2020.
[17] Ž. Vujović, "Classification model evaluation metrics," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599–606, 2021.
[18] M. Rahayu, A. Luthfiarta, L. Cahyaningrum, and A. N. Azzahra, "Pengaruh Oversampling dan Cross Validation Pada Model Machine Learning Untuk Sentimen Analisis Kebijakan Luaran Kelulusan Mahasiswa," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 1, pp. 163–172, Jan. 2024, doi: 10.30865/MIB.V8I1.7012.
[19] H. Kaur, H. S. Pannu, and A. K. Malhi, "A Systematic Review on Imbalanced Data Challenges in Machine Learning," *ACM Computing Surveys (CSUR)*, vol. 52, no. 4, Aug. 2019, doi: 10.1145/3343440.
[20] P. Vuttipittayamongkol, E. Elyan, and A. Petrovski, "On the class overlap problem in imbalanced data classification," *Knowl Based Syst*, vol. 212, p. 106631, Jan. 2021, doi: 10.1016/J.KNOSYS.2020.106631.
[21] K. E. Turner, A. Thompson, I. Harris, M. Ferguson, and F. Sohel, "Deep learning based classification of sheep behaviour from accelerometer data with imbalance," *Information Processing in Agriculture*, vol. 10, no. 3, pp. 377–390, Sep. 2023, doi: 10.1016/J.INPA.2022.04.001.
[22] S. Riyanto, I. S. Sitanggang, T. Djatna, and T. D. Atikah, "Comparative Analysis using Various Performance Metrics in Imbalanced Data for Multi-class Text Classification," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, p. 2023, Accessed: Aug. 08, 2024. [Online]. Available: http://gcancer.org/pdr