Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

Aspect-Level Sentiment Analysis on Social Media Using Gated Recurrent Unit (GRU)

Ghani Kamil[1], Erwin Budi Setiawan[2,*]

School of Computing, Informatics, Telkom University, Bandung, Indonesia
Email: [1]ghanikamil@student.telkomuniversity.ac.id, [2,*]erwinbudisetiawan@telkomuniversity.ac.id
Correspondence Author Email: erwinbudisetiawan@telkomuniversity.ac.id
Submitted: **06/02/2023**; Accepted: **27/03/2023**; Published**: 31/03/2023**

Abstract− Twitter is one of the popular social media for sharing opinions, one of which is about movie reviews. There are many opinions related to movie reviews on Twitter social media so the assessment of a movie can vary. Therefore, sentiment analysis at the aspect level is needed to classify film reviews in order to provide optimal results as analytical material in making films that can increase audience satisfaction. This research was conducted by building a system using the Gated Recurrent Unit (GRU) method to perform sentiment analysis at the aspect level on movie reviews taken from Twitter. The aspects used in this research are plot, acting, and director. This research also conducted experiments by combining three techniques, which are feature extraction using TF-IDF, feature expansion with GloVe, and the application of SMOTE to improve model accuracy. The results show that each test scenario can improve the accuracy and F1-Score values of each aspect. The final value of each aspect is the accuracy value for the plot aspect is 70.40%(+7.62%) and F1-Score is 70.35%(+9.70%), the accuracy value is 93.75%(+6.28%) and F1-Score is 93.70%(+65.19%) for the acting aspect, and the accuracy value is 90.44%(+4.60%) and F1-Score is 90.17%(+122.80%) for the director aspect.

Keywords: Sentiment Analysis; Aspect; Gated Recurrent Unit; GloVe; SMOTE

1. INTRODUCTION

The use of social media is closely related to people's lives today. Social media allows each individual to communicate and interact with others through cyberspace, one of which is Twitter [1]. Based on Statista [2], the number of Twitter users in Indonesia is the fifth largest in the world with a total of 18.45 million users. There are various topics of discussion found on Twitter social media, one of which is movie reviews. Based on Statista [3], the theater or film entertainment market continued to grow from 2015 to generate 80.8 billion US dollars in 2020. This indicates that the film industry is growing from year to year. But with the increasing output, the quality of a movie is also evaluated by most netizens such as making reviews on social media. Twitter users can express their opinions in the form of tweets, such as giving a review of a movie. The reviews that Twitter users express in their tweets about a movie can be negative or positive [4]. However, due to the various opinions available, the assessment of reviews about a movie can be different [5]. The various opinions contain sentiment information that is in lines with this research such as plot, acting, and director aspects [6]. These opinions can be used as evaluations by filmmakers to increase audience satisfaction [7]. Therefore, research is needed to analyze the opinion whether it is a positive, negative, or neutral opinion.

Sentiment analysis is a means of determining whether an opinion is negative or positive or neutral from someone who can be realized in the written or oral form [8][9]. Sentiment analysis is comprised of three levels: sentence level, document level, and aspect level [10],[11]. Sentiment analysis at the aspect level is utilized to uncover sentiments expressed in reviews of a product. Considering multiple aspects of a product is crucial as a product typically has multiple facets that require evaluation. This cannot be done using the document level or sentence level because both levels cannot determine opinions specifically as in the aspect level [12],[13]. Since movie reviews have various aspects such as storyline, cast, visual effects, and so on, the implementation of aspect level can help in conducting this research.

This research was conducted based on previous research that has been done. In research [14], Rimdani conducted an aspect-based sentiment analysis on Twitter with applied the Gradient Boosting Decision Tree classification technique to signal and service aspects, as well as SMOTE and Random Undersampling. The results of this study indicate that SMOTE can produce better performance values than Random Undersampling. As a result, the final performance value obtained is 96.035% for the signal aspect, an increase of 25.260%, and 90.256% for the service aspect, an increase of 27.918%. In research [15], M. S. Mubarok et al. conducted experiments using the Naïve Bayes method in analyzing aspect-based sentiment and the dataset used is restaurant reviews. The aspects used in this study are food, service, price ambiance, and miscellaneous. With this method, the performance value of the f1-score is 78.12%. However, this research has not applied deep learning to the sentiment analysis process. Another study [16], N. Mohamed Ali et al. used several deep-learning methods namely CNN, LSTM, and CNN-LSTM. This research compares the performance value of each deep learning method used. The dataset used in this research is a movie review dataset. It is concluded that the CNN-LSTM method is better than the others with an accuracy value of 89.20%. However, the research is not based on the product review aspect. In research [17], J. S. Lee et al. uses the GRU method in analyzing sentiment. The dataset used is a dataset taken from product reviews on Chinese e-commerce. This research aims to compare the GRU method to the LSTM method in potential performance improvement. The accuracy

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

result obtained from this method on the Facebook dataset is 87%. However, this research has not implemented aspect-level sentiment analysis.

In research [18], R. Ahuja et al. conducted sentiment analysis by comparing two feature extraction methods namely TF-IDF and N-Gram on the SS-Tweet dataset. The study used six classification models such as the Decision Tree, SVM, KNN, Logistic Regression, Random Forest, and Naïve Bayes. The results of the study found that the use of TF-IDF has a better performance than N-Gram. In research [19], Aditya and Erwin applied Glove feature expansion on an aspect-based sentiment analysis on Twitter using the Random Forest classification. The accuracy value obtained for the service aspect is 80.12%, while the accuracy value for the signal aspect is 80.37%. In addition, GloVe feature expansion is able to increase signal accuracy by 13.15 percent and service accuracy by 5.37%. M. A. Raihan and E. B. Setiawan in their research [20], conducted aspect-based sentiment analysis on Twitter social media using FastText and SVM. The aspects used in the study are signal and service. The research also uses SMOTE data sampling techniques to overcome unbalanced data. The results of this study obtained the highest performance value after the application of SMOTE with an F1-Score value for the signal aspect of 95.93% and the service aspect with an F1-Score value of 94.53%.

Based on previous research, most of these studies have not conducted aspect-level sentiment analysis using GRU with feature extraction, GloVe, and SMOTE methods. The use of feature extraction, feature expansion and SMOTE can improve the performance value of the model built to be better. Therefore, In this study, an aspect-level sentiment analysis was carried out using the GRU classification.

The purpose of this research is to apply the GRU classification method to aspect-level sentiment analysis on movie review data and calculate the accuracy value generated by the model. The results of this study provide the performance value of the GRU classification model built. In this study, experiments were also conducted by adding feature extraction with TF-IDF, feature expansion using Global Vectors (GloVe) trained based on corpus datasets, and adding SMOTE to improve model performance. The problem limitation of this research is that the data used is movie review data taken from Indonesian Twitter. The aspects used in this research are plot, acting, and director.

2. RESEARCH METHODOLOGY

2.1 Research Stages

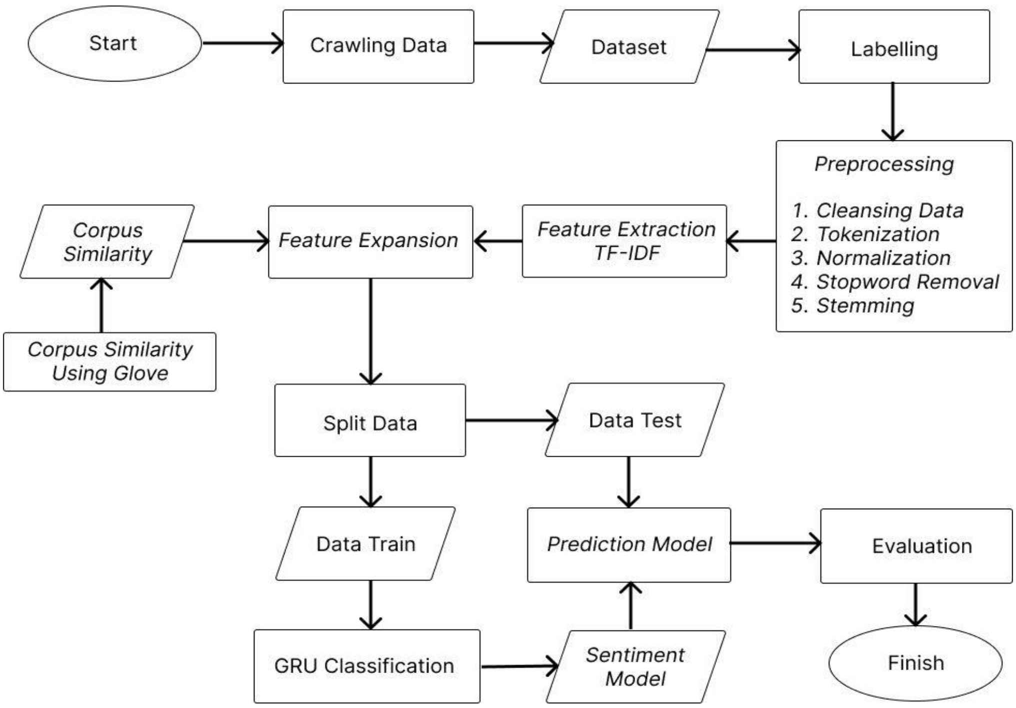

Figure 1. Aspect-Level Sentiment Analysis with Gated Recurrent Unit (GRU)

Figure 1 shows the stages of building the GRU classification model used to conduct sentiment analysis on social media.

2.2 Crawling Data

This research uses a dataset in the form of downloads from tweets on Twitter social media. Data retrieval is done through the use of the snsscrape library in python. The dataset taken is in the form of movie review tweets from various Twitter users. The total number of datasets used in this study is 17247 movie review tweet data for each aspect. The crawling results will be saved in CSV format.

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

2.3 Labeling

In this research, data labeling is carried out so that the classification process runs well. Each data obtained is labeled manually. Example of labeling on the movie review dataset for each aspect can be seen in Table 1.

Table 1. Example of data labeling for each aspect

Tweet	Plot	Acting	Director
the desperate hour lakewood 2022 beberapa waktu lalu ada yang bersalah dengan 98 cerita hanya lewat suara telepon dan ber setting ruangan operator 911 ini juga lewat telepon sekitar 94 dan 89 ber setting outdoor terutama hutan tegang banget asli	1	1	1
plot utama adalah dua orang deserter pursuit atau pemburu wamil junho hoyeol ditugaskan utk memburu para wamil yg kabur dari wajib militer dua karakter ini sangat berbeda junho pendiam dan hoyeol humoris tapi keduanya sama2 menentang kekerasan dlm pendidikan wamil	1	1	0

Based on Table 1, the labels used for each aspect use a number range from -1 to 1, with a label of 1 for data that is considered positive, -1 for data that is considered negative, and 0 for data that is considered neutral. Based on the labeling results on the data, the distribution of the amount of data for each class is obtained as in Table 2.

Table 2. Number of Sentiment Labels for Each Aspect

Category Aspect	Positive	Neutral	Negative
Plot	6999	6566	3682
Acting	2907	13770	570
Director	1907	14753	587

Table 2 shows the amount of data distribution for each class in each aspect. There is a total of 17247 data for each aspect.

2.4 Preprocessing Data

The data obtained from Twitter still contains characters that are not needed in the classification process of symbols, emoticons, mentions, and URLs, causing the data to be unstructured, therefore a process is carried out to correct the data so that it can be processed. The stages in data preprocessing are as follows:

a. **Cleansing Data**: This process is done to remove unnecessary characters such as symbols, numbers, emotes, mentions, and URLs that are usually present in a tweet.
b. **Tokenization**: This process is carried out to convert the sentences in the tweet into words and stored them into tokens.
c. **Normalization**: This process is done to standardize each word in each existing data.
d. **Stopword Removal**: This process is done to remove words that appear in large numbers but have no meaningful meaning.
e. **Stemming**: This process is done to remove suffixes and affixes from each word in each existing data.

2.5 TF-IDF Feature Extraction

Term Frequency-Inverse Document Frequency (TF-IDF) is an algorithm that determines the relevance of words in a document. Based on the number of words in a document, Term Frequency (TF) calculates how frequently a term appears. The formula for the frequency of occurrence of words in the document is shown in formula (1) [21].

$$TF = \frac{number\ of\ times\ word\ w\ occured}{total\ number\ of\ words\ in\ the\ document} \quad (1)$$

Based on formula (1), the number of times word *w* occurrences in document is divided by the total number of words in the document. Meanwhile, Inverse Document Frequency (IDF) determines a term's importance based on its distribution across the corpus of documents [18]. The formula for IDF is shown in formula (2) [21].

$$IDF = log \frac{Number\ of\ documents}{number\ of\ documents\ containing\ word\ w} \quad (2)$$

Based on formula (2), IDF is obtained by calculating the log of the number of documents divided by the number of documents containing the word *w*.

Overall the TF-IDF formula can be shown in formula (3).

$$TF - IDF = TF \times IDF \quad (3)$$

At this stage, feature extraction is carried out on the dataset that has been preprocessed. Feature extraction in this study was carried out using Term Frequency-Inverse Document Frequency (TF-IDF) vectorization with a maximum number of features of 1000, 5000, and 10000.

2.6 Global Vectors (GloVe)

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

Global Vectors (GloVe) is a method of unsupervised learning used to derive vector representations of words [22]. The implementation of GloVe makes it possible to evaluate the performance of various classification algorithms that are often used in the field of Natural Language Processing and also those that are not commonly used in text data analysis [23]. GloVe can be used as a simpler and more efficient alternative to word2vec [23].

At this stage, the Global Vectors GloVe model is built for use in feature expansion. The GloVe model used is built based on movie review corpus data. Each word in the data is used to build the corpus. The resulting corpus is then trained using the GloVe model.

2.7 Feature Expansion

In this stage, feature expansion is performed using the GloVe model that has been built in the previous stage. Feature expansion is done to identify words that are missing in the representation of tweet data and distributed with semantically related words. Feature expansion is done by looking for word similitude in the GloVe model. The similitude results will then be inserted into the data to replace each 0 value according to its index.

2.8 Gated Recurrent Unit (GRU) Classification

Gated Recurrent Unit (GRU) is a neural network algorithm that uses a gate system that was first introduced by Kyunghun Cho in 2014 [24]. In this study, the GRU model is used to perform classification. The GRU model used consists of 3 bidirectional layers and 3 hidden layers and 1 output layer.

2.9 Performance Evaluation

At this stage, the performance evaluation process of the system that has been built using the confusion matrix is carried out. The confusion matrix will calculate the accuracy, recall, precision, and F1-Score values of the system that has been built, as in Table 3.

Table 3. Confusion Matrix [21]

	Positive Prediction	Negative Prediction
Actual Positive	TP	FN
Actual Negative	FP	TN

Based on the Table 3:
a. **True Positive**: The number of events that the system predicted was positive and actually positive.
b. **False Positive**: The number of events that the system predicted was positive but actually negative.
c. **True Negative**: The number of events that the system predicted was negative and actually negative.
d. **False Negative**: The number of events that the system predicted was negative but actually positive.

3. RESULT AND DISCUSSION

This research is conducted with several test scenarios to achieve the research objectives. Each scenario will use the GRU classification model. The scenarios used in this study include:
a. The first scenario is a test to find the baseline using GRU classification.
b. The second scenario is the addition of feature extraction using TF-IDF.
c. The third scenario is adding feature expansion with GloVe.
d. The fourth scenario applies the SMOTE algorithm for handling unbalanced data.

The test results for each scenario are in the form of average accuracy and F1-Score values after 5 times execution.

3.1 Evaluation Results

In the first scenario, tests were conducted to find a baseline model using GRU classification. In this test, the dataset used is the entire movie review dataset that has been preprocessed previously. The ratio of training data and test data used in this test is 90:10, 80:20, and 70:30. Tests were conducted for each aspect category. The test results for this scenario are shown in Table 4. The highest accuracy value for the plot aspect of 65.41% was obtained with a data ratio of 80:20. The highest accuracy value for the acting aspect is 88.21% obtained with a data ratio of 80:20. While the highest accuracy value for the director aspect is 86.46% obtained with a data ratio of 70:30. Based on these results, the best test size ratio for each aspect is used as a baseline for the next scenario.

Table 4. GRU baseline test results

Category Aspect	90:10		80:20		70:30	
	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)
Plot	64.67	69.21	**65.41**	**64.13**	65.03	63.96
Acting	86.78	58.12	**88.21**	**56.72**	87.31	57.11

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

Director	**86.46**	**40.46**	86.16	49.42	85.89	49.03

In the second scenario, testing was carried out by adding feature extraction using TF-IDF to the baseline model. The max number of features tested in this feature extraction is 1000, 5000, and 10000. The test results for this scenario can be seen in Table 5. Based on Table 5, the greatest accuracy value is generated for the plot aspect of 67.34% and F1-Score of 65.69% for the max feature of 1000. In the acting aspect, the highest accuracy value is 89.36% and F1-Score is 51.68% for the max feature of 1000. while for the director aspect, the highest accuracy value is 87.69% and F1-Score is 51.68% for the max feature of 1000.

Table 5. GRU test results with TF-IDF Feature Extraction

Category Aspect	Max Feature					
	1000		5000		10000	
	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)
Plot	**67.34 (+2.95)**	**65.69(+2.43)**	66.85 (+2.20)	65.79(+2.58)	65.84 (+0.65)	64.14 (+0.01)
Acting	**89.36 (+1.30)**	**64.23(+13.24)**	88.34 (+0.14)	64.99 (+14.58)	88.47 (+0.29)	63.87 (+12.60)
Director	**87.69 (+1.42)**	**51.68(+27.73)**	87.03 (+0.65)	53.63 (+32.55)	86.93 (+0.54)	52.31 (+29.28)

In the third scenario, feature expansion using GloVe is added to the results of the previous scenario. The GloVe model used in this feature expansion is a model trained using a corpus of movie review tweet data. In the testing process, the number of similarities used are top 1, top 5, and top 10. The test results using the GloVe expansion feature for each aspect can be seen in Table 6. Based on these results, the highest accuracy value for the plot aspect is generated by the top 1 of 67.69% and the F1-Score value of 65.21%. In the acting aspect, the highest accuracy value is generated by the top 1 of 89.93% and the F1-Score value of 64.99%. As for the director aspect, the highest accuracy value is generated by the top 1 of 88.03% and the F1-Score value of 48.09%. The results of this scenario testing are then used in the next scenario.

Table 6. Baseline test results with TF-IDF and GloVe

Category Aspect	Top Similiarity					
	1		5		10	
	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)	Akurasi (%)	F1-Score (%)
Plot	**67.69 (+3.49)**	**65.21 (+1.68)**	67.05 (+2.51)	64.61 (+0.75)	67.03 (+2.48)	65.20 (+1.67)
Acting	**89.93 (+1.95)**	**64.99 (+14.58)**	89.11 (+1.02)	64.70 (+14.07)	89.13 (+1.04)	64.71 (+14.08)
Director	**88.03 (+1.82)**	**48.09 (+18.85)**	87.56 (+1.27)	48.26 (+19.27)	87.61 (+1.33)	48.53 (+19.95)

In the fourth scenario, the SMOTE algorithm was applied to overcome unbalanced data. The results of applying the SMOTE algorithm can be seen in Table 7.

Table 7. Baseline + TF-IDF + GloVe + SMOTE test results

Category Aspect	SMOTE	
	Accuracy (%)	F1-Score (%)
Plot	70.40 (+7.62)	**70.35 (+9.70)**
Acting	93.75 (+6.28)	**93.70 (+65.19)**
Director	**90.44 (+4.60)**	**90.17 (+122.8)**

In Table 7, the application of the SMOTE algorithm significantly improves the accuracy and F1-Score values for each aspect. In the plot aspect, the accuracy value increases to 70.40%, an increase of 7.62%. Meanwhile, the F1-Score value becomes 70.35%, an increase of 9.70%. In the acting aspect, the accuracy value increased to 93.75%, an increase of 6.28%. Meanwhile, the F1-Score value became 93.70%, an increase of 65.19%. In the director aspect, the accuracy value increased to 90.44%, an increase of 4.60%. Meanwhile, the F1-Score value became 90.17%, an increase of 122.8%.

3.2 Discussion

In this research, tests have been conducted on four scenarios, namely the search for a baseline model, the addition of TF-IDF feature extraction, GloVe feature expansion, and the application of SMOTE to the GRU classification model. In the plot aspect, the best baseline model is obtained by using a test size of 80:20. Then in scenario 2, an increase in accuracy of 2.95% and F1-Score of 2.43% was obtained after applying TF-IDF with a max feature value of 1000. In scenario 3, the increase in accuracy and F1-Score values is obtained with the top 1 in the tweet corpus. The increase in accuracy and F1-Score values is 3.43% and 1.68%, respectively. While in scenario 4, the increase in accuracy value is 7.62% and F1-Score is 9.70%. Figure 2 shows the rate of increase in model performance in the plot aspect for each scenario.

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

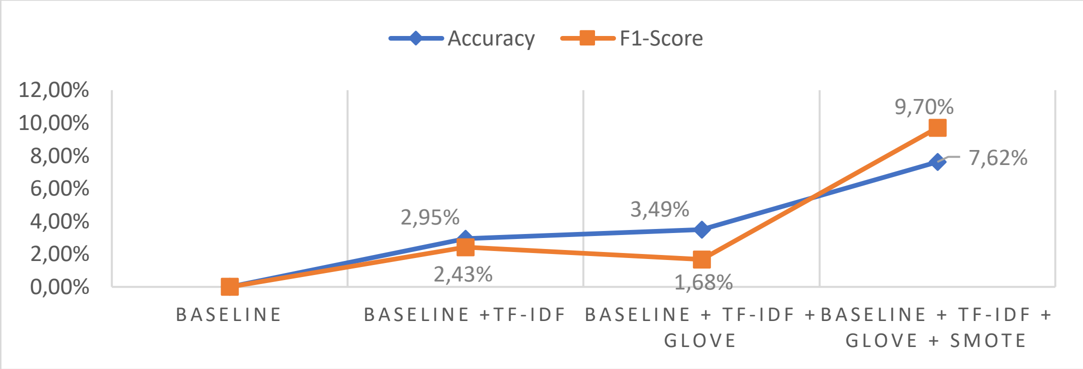

Figure 2. Increase in Accuracy and F1-Score on plot aspects

In the acting aspect, the best baseline model is obtained by using test size 80:20. In scenario 2, an increase in accuracy of 1.30% and F1-Score of 13.24% was obtained after applying TF-IDF with a max feature value of 1000. In scenario 3, the increase in accuracy and F1-Score values was obtained with the top 1 in the tweet corpus. The increase in accuracy and F1-Score value is 1.95% and 14.58% respectively. While in scenario 4, the increase in accuracy value is 6.28% and F1-Score is 65.19%. Figure 3 shows the rate of increase in model performance in the plot aspect for each scenario.

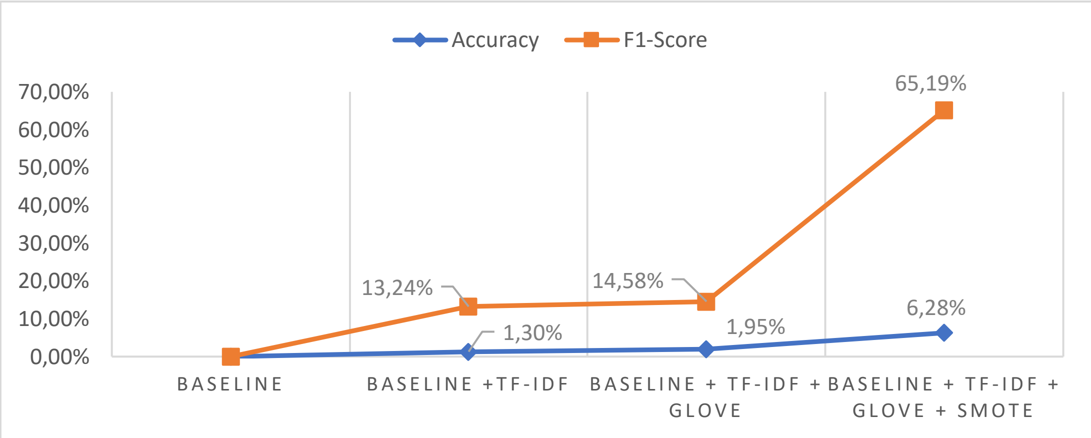

Figure 3. Increase in Accuracy and F1-Score on acting aspects

In the director aspect, the best baseline model is obtained by using a test size of 70:30. In scenario 2, an increase in accuracy of 1.42% and F1-Score of 27.73% was obtained after applying TF-IDF with a max feature value of 1000. In scenario 3, the increase in accuracy and F1-Score was obtained with the top 1 in the tweet corpus. The increase in accuracy and F1-Score values is 1.82% and 18.85%, respectively. While in scenario 4, the increase in accuracy value is 4.60% and F1-Score is 122.8%. Figure 4 shows the rate of increase in model performance in the plot aspect for each scenario.

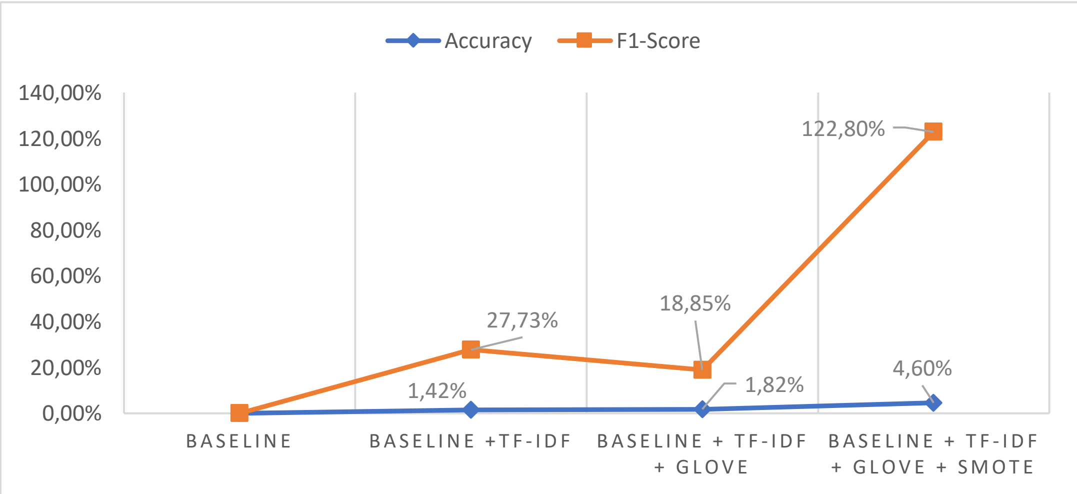

Figure 4. Increase in Accuracy and F1-Score on director aspects

Building of Informatics, Technology and Science (BITS)
Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

Based on the test results, shows that each scenario improves the model performance for all aspects. The use of feature extraction with TF-IDF, feature expansion with GloVe, and SMOTE can improve accuracy and F1-Score. SMOTE has a significant impact in improving accuracy and F1-Score because it can balance each class in the data used so as to prevent poor model performance due to data imbalance.

4. CONCLUSION

In this research, sentiment analysis of movie reviews using the GRU classification model has been conducted. This research also conducted experiments by adding feature extraction with TF-IDF, feature expansion with GloVe, and applying SMOTE. The analysis was conducted to determine the accuracy of the application of the GRU classification model to each aspect of the movie review such as plot, acting, and director. The data used in this study consisted of 17,247 movie reviews obtained from Twitter. Based on the four scenarios that have been carried out, each aspect has increased the accuracy value and F1-Score. The final accuracy value for the plot aspect is 70.40% and F1-Score is 70.35%, the final accuracy value for the acting aspect is 93.75% and F1-Score is 93.70%, and the final accuracy value for the director aspect is 90.44% and F1-Score is 90.17%. The results obtained show that each scenario can increase the accuracy value of each aspect. The use of feature expansion with GloVe also contributes to increasing the accuracy value of the baseline model. Further research can be done using other classification models and using other word embedding models.

REFERENCES

[1] T. Siswanto, “Optimalisasi Sosial Media Sebagai Media Pemasaran Usaha Kecil Menengah,” Liquidity: Jurnal Riset Akuntansi dan Manajemen, vol. 2, no. 1, pp. 80–86, Jul. 2013, doi: 10.32546/LQ.V2I1.134.
[2] S. Dixon, “Countries with most Twitter users 2022,” Statista, Nov. 22, 2022. https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ (accessed Feb. 06, 2023).
[3] J. Stoll, “Box office and home video revenue worldwide 2020,” Statista, Mar. 24, 2021. https://www.statista.com/statistics/1194522/box-office-home-and-mobile-video-entertainment-revenue-worldwide/ (accessed Feb. 06, 2023).
[4] S. Pandya and P. Mehta, “A Review On Sentiment Analysis Methodologies, Practices And Applications,” INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, vol. 9, p. 2, 2020, Accessed: Jan. 30, 2023. [Online]. Available: www.ijstr.org
[5] N. Shafirra, N. A. Shafirra, and I. Irhamah, “Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Speech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN),” Jurnal Sains dan Seni ITS, vol. 9, no. 1, pp. D95–D101, Jun. 2020, doi: 10.12962/j23373520.v9i1.51825.
[6] V. Parkhe and B. Biswas, “Aspect based sentiment analysis of movie reviews: Finding the polarity directing aspects,” Proceedings - 2014 International Conference on Soft Computing and Machine Intelligence, ISCMI 2014, pp. 28–32, Apr. 2014, doi: 10.1109/ISCMI.2014.16.
[7] C. Nanda, M. Dua, and G. Nanda, “Sentiment Analysis of Movie Reviews in Hindi Language Using Machine Learning,” Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing, ICCSP 2018, pp. 1069–1072, Nov. 2018, doi: 10.1109/ICCSP.2018.8524223.
[8] N. Muchammad Shiddieqy Hadna, P. Insap Santosa, and W. Wahyu Winarno, “STUDI LITERATUR TENTANG PERBANDINGAN METODE UNTUK PROSES ANALISIS SENTIMEN DI TWITTER,” Seminar Nasional Teknologi Informasi dan Komunikasi, 2016.
[9] A. Alsaeedi and M. Z. Khan, “A Study on Sentiment Analysis Techniques of Twitter Data,” Article in International Journal of Advanced Computer Science and Applications, vol. 10, no. 2, 2019, doi: 10.14569/IJACSA.2019.0100248.
[10] M. Cindo, D. P. Rini, and E. Ermatita, “Literatur Review: Metode Klasifikasi Pada Sentimen Analisis,” Seminar Nasional Teknologi Komputer & Sains (SAINTEKS), vol. 1, no. 1, pp. 66–70, Feb. 2019, Accessed: Jan. 30, 2023. [Online]. Available: http://seminar-id.com/prosiding/index.php/sainteks/article/view/124
[11] R. Feldman, “Techniques and applications for sentiment analysis,” Commun ACM, vol. 56, no. 4, pp. 82–89, Apr. 2013, doi: 10.1145/2436256.2436274.
[12] S. M. Jiménez-Zafra, M. T. Martín-Valdivia, E. Martínez-Cámara, and L. A. Ureña-López, “Combining resources to improve unsupervised sentiment analysis at aspect-level,” http://dx.doi.org/10.1177/0165551515593686, vol. 42, no. 2, pp. 213–229, Jul. 2015, doi: 10.1177/0165551515593686.
[13] S. Vanaja and M. Belwal, “Aspect-Level Sentiment Analysis on E-Commerce Data,” Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018, pp. 1275–1279, Dec. 2018, doi: 10.1109/ICIRCA.2018.8597286.
[14] R. A. Annisa and E. B. Setiawan, “Aspect Based Sentiment Analysis on Twitter Using Word2Vec Feature Expansion Method and Gradient Boosting Decision Tree Classification Method,” 2022 1st International Conference on Software Engineering and Information Technology (ICoSEIT), pp. 273–278, Nov. 2022, doi: 10.1109/ICOSEIT55604.2022.10030010.
[15] M. S. Mubarok, A. Adiwijaya, and M. D. Aldhi, “Aspect-based sentiment analysis to review products using Naïve Bayes,” AIP Conf Proc, vol. 1867, no. 1, p. 020060, Aug. 2017, doi: 10.1063/1.4994463.

Volume 4, No 4, Maret 2023 Page: 1837−1844
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i4.3105

[16] N. Mohamed Ali, M. Mostafa Abd El Hamid, and A. Youssif, "SENTIMENT ANALYSIS FOR MOVIES REVIEWS DATASET USING DEEP LEARNING MODELS," International Journal of Data Mining & Knowledge Management Process (IJDKP), vol. 9, no. 3, 2019, doi: 10.5121/ijdkp.2019.9302.
[17] J. S. Lee, D. Zuba, and Y. Pang, "Sentiment analysis of Chinese product reviews using gated recurrent unit," Proceedings - 5th IEEE International Conference on Big Data Service and Applications, BigDataService 2019, Workshop on Big Data in Water Resources, Environment, and Hydraulic Engineering and Workshop on Medical, Healthcare, Using Big Data Technologies, pp. 173–181, Apr. 2019, doi: 10.1109/BIGDATASERVICE.2019.00030.
[18] 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, Jan. 2019, doi: 10.1016/J.PROCS.2019.05.008.
[19] A. M. Zakaria and E. B. Setiawan, "Aspect-Based Analysis of Telkomsel User Sentiment on Twitter Using the Random Forest Classification Method and Glove Feature Expansion," Jurnal Teknologi dan Sistem Komputer, vol. 0, no. 0, Sep. 2022, doi: 10.14710/JTSISKOM.2022.14558.
[20] M. A. Raihan and E. B. Setiawan, "Aspect Based Sentiment Analysis with FastText Feature Expansion and Support Vector Machine Method on Twitter," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 6, no. 4, pp. 591–598, Aug. 2022, doi: 10.29207/RESTI.V6I4.4187.
[21] S. Thavareesan and S. Mahesan, "Sentiment Analysis in Tamil Texts: A Study on Machine Learning Techniques and Feature Representation," 2019 IEEE 14th International Conference on Industrial and Information Systems: Engineering for Innovations for Industry 4.0, ICIIS 2019 - Proceedings, pp. 320–325, Dec. 2019, doi: 10.1109/ICIIS47346.2019.9063341.
[22] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," Stanford, 2014. https://nlp.stanford.edu/projects/glove/ (accessed Feb. 06, 2023).
[23] P. Cichosz, "A case study in text mining of discussion forum posts: Classification with bag of words and global vectors," International Journal of Applied Mathematics and Computer Science, vol. 28, no. 4, pp. 787–801, 2018, doi: 10.2478/AMCS-2018-0060.
[24] L. Zaman, S. Sumpeno, and M. Hariadi, "Analisis Kinerja LSTM dan GRU sebagai Model Generatif untuk Tari Remo," Jurnal Nasional Teknik Elektro dan Teknologi Informasi, vol. 8, no. 2, 2019, Accessed: Jan. 30, 2023. [Online]. Available: https://journal.ugm.ac.id/v3/JNTETI/article/view/2595/717