

Sentiment Analysis on Movie Review from Rotten Tomatoes Using Modified Balanced Random Forest Method and Word2Vec

Mohamad Rizki Nugraha^{*}, Mahendra Dwifabri P, Widi Astuti

Faculty of Informatics, Informatics, Telkom University, Bandung, Indonesia

Email: ¹mrizkinugraha@student.telkomuniversity.ac.id, ²mahendradp@telkomuniversity.ac.id, ³widiwdu@telkomuniversity.ac.id

Correspondence Author Email: mrnugraha2001@gmail.com

Submitted: **08/06/2023**; Accepted: **27/06/2023**; Published: **29/06/2023**

Abstract– The film industry is one of the impacts of the rapid development of technology. This causes the film industry to increase every year. In addition, technological developments also affect the public to make it easier to access various movie from various website. With many choices of movies, people need to know the quality of various movies by knowing the reviews of these movies from other people. However, the large number of audience reviews of a movie makes it difficult for people to categorize good movies or bad movies. The solution to the problem is to perform sentiment analysis on movie reviews. In this research, the classification method used is Modified Balanced Random Forest. This method was chosen because it can overcome imbalanced data and can increase accuracy and reduce time complexity. In this research also the feature extraction used is Word2Vec. This feature extraction was chosen because previous research explained that Word2Vec has the advantage of being able to show the contextual similarity of two words in the resulting vector. The best model produced from this research is a model built without using stemming in the preprocessing stage, using 300 dimensions in Word2Vec, and using the Modified Balanced Random Forest classification method which produces an f1-score of 84.15%.

Keywords: Sentiment Analysis; Movie Review; Modified Balanced Random Forest; Word2vec

1. INTRODUCTION

The film industry is a type of industry that is experiencing an increase every year, both from newly produced films and films that have been produced before [1]. This is due to the rapid development of technology. The rapid development of technology affects people to make it easier to find and watch various movies from various websites. With a large selection of movies, it is often confusing for someone to choose which movie they will watch [2]. People previously needed to know the quality of various kinds of movies by looking at reviews or other people's responses to the movie. With various reviews from the audience on a film, a sentiment analysis is needed to categorize the reviews into a sentiment [3].

The process that can automate the collection of opinions, emotions, and views from speech, text, and data sources by Natural Language Processing is called sentiment analysis [4]. The results of sentiment analysis can help in understanding audience satisfaction with a movie. Sentiment analysis evaluates the polarity of text in a sentence or data to establish which opinions expressed are positive, negative, or neutral [3]. The positive nature indicates that the comments given to the movie have a good value, it can be concluded that the reviewed movie is a great movie. Meanwhile, the negative nature indicates that the comments given to the movie have a less good value, it can be concluded that the reviewed movie is a bad movie. However, there is also a neutral trait which means it indicates that the movie is not too good and not too bad. Sentiment analysis can help benchmark the success of a movie by comparing it to similar films. The success of a movie can help the public so that they are not confused in choosing a movie that they will watch. This research can also determine the effectiveness of an algorithm and sentiment analysis techniques that are well used to process movie review data from various sources.

In sentiment analysis, there are several algorithms that can be used. In this research, the feature extraction used is Word2Vec. This feature extraction aims to map information or words in a sentence into a vector. This feature extraction was chosen because research [5] explains that Word2Vec has the advantage of being able to show the contextual similarity of two words in the resulting vector. In sentiment analysis, the imbalanced distribution of class data can affect classification performance. Therefore, Modified Balanced Random Forest is used as a classification method in this research. This classification method was chosen because the data used in this research is imbalanced, it is in accordance with the purpose of this method which is to handle imbalanced data. In addition, this method can also improve accuracy and reduce time complexity. In research [6] the feature selection used is Chi-square and Modified Balanced Random Forest as a classification method was successful in producing an accuracy of 81.75% and f1-score of 71.90%. This research uses some references from previous researches that have relevance to sentiment analysis, preprocessing stages, feature extraction, and classification methods.

Research [7] by Ardhan Fahmi Sabani in 2022, discusses sentiment analysis of movie reviews on the rotten tomatoes website but uses Support Vector Machine method and Word2Vec feature extraction. The preprocessing stages in the study are data type conversion, case folding, tokenization, stopword removal, stemming. In this study, testing was carried out on the dataset by selecting the kernel in the classification method, stemming and not stemming the dataset, using K-fold and not using K-fold, and undersampling. This research produces the best model using the RBF kernel and using K-fold which produces an average precision of 77.2%, recall 68.2%, F1-Score 70.2% and accuracy 79.0%. In this study, researchers want to use Word2Vec feature extraction which has been quite well used

previously on the same dataset in converting each word in the data into a vector but using a different classification method, namely Modified Balanced Random Forest.

In research [6] by Antika Putri Permata Wardani in 2022, discusses sentiment analysis using Modified Balanced Random Forest method and feature selection used is Chi-square of beauty product reviews. In this research, the preprocessing schemes carried out is data cleaning, case folding, normalization, stopword removal, stemming. In this study, the dataset was tested by stemming and not stemming the dataset, using feature selection and not using feature selection, and using K-fold and not using K-fold. The best results in this study are using stemming at the preprocessing stage, using selection features, and using K-fold in the performance test with a f1-score of 71.95% and an average accuracy of 81.75%. In this study, researchers want to use Modified Balanced Random Forest classification method which has been quite well used previously to overcome imbalanced datasets but the dataset used is not a movie review, does not use feature selection and uses different feature extraction, namely Word2Vec.

In the journal [8] by Yusuf Surya T in 2021, using Support Vector Machine and Word2Vec methods for sentiment analysis of movie reviews. This research produces the best accuracy value by using the lemmatization preprocessing stage, using 300 Word2Vec dimensions and using a linear Support Vector Machine classification model which produces the best F1-score of 78.74% and the best accuracy of 78.75%. In this study, researchers want to use Word2Vec which has been quite well used previously as a feature extraction, but researchers try to use a different preprocessing stage, namely stemming and use a different classification method, namely Modified Balanced Random Forest to get the best performance results expected.

In research [9] by Muhammad Asjad Adna Jihad in 2021, conducted sentiment analysis using Random Forest method and Word2Vec as a feature extraction of movie reviews. This study explains that the use of stemming in the preprocessing stage affects the final performance results where in this case stemming improves performance further. The best results in this study were obtained from datasets that applied the stemming process at the preprocessing stage, applied Adaptive Boosting to the base model, and used 300-dimensional Word2Vec skip-grams by producing the best accuracy of 75.76%. In this study, researchers want to use Word2Vec which has been quite well used previously as feature extraction, but researchers try to use a modified method of Random Forest, namely Modified Balanced Random Forest to get the best performance results expected.

In research [10] by Firdausi Nuzula Zamzami in 2021, describes sentiment analysis using Modified Balanced Random Forest Method with feature selection used is Mutual Information of movie reviews. This research produces F1-scores of 75% and accuracy of 79%. These results were obtained after the use of stemming in the preprocessing process which can improve the performance of the system, the use of Mutual Information feature selection which can improve the performance of the classification process and reduce less relevant features to be used in Modified Balanced Random Forest classification process, and the use of Modified Balanced Random Forest classification process which is able to increase the F1-score value of Random Forest by 27% for imbalanced English movie review datasets. In this study, researchers want to use Modified Balanced Random Forest classification method which has been quite well used previously to overcome the imbalanced movie review datasets but does not use feature selection and uses different feature extraction, namely Word2Vec.

2. RESEARCH METHODOLOGY

System model developed in this research is a sentiment analysis using Modified Balanced Random Forest and Word2Vec methods of movie reviews on the Rotten Tomatoes website. Figure 1 is an illustration of the system design flow.

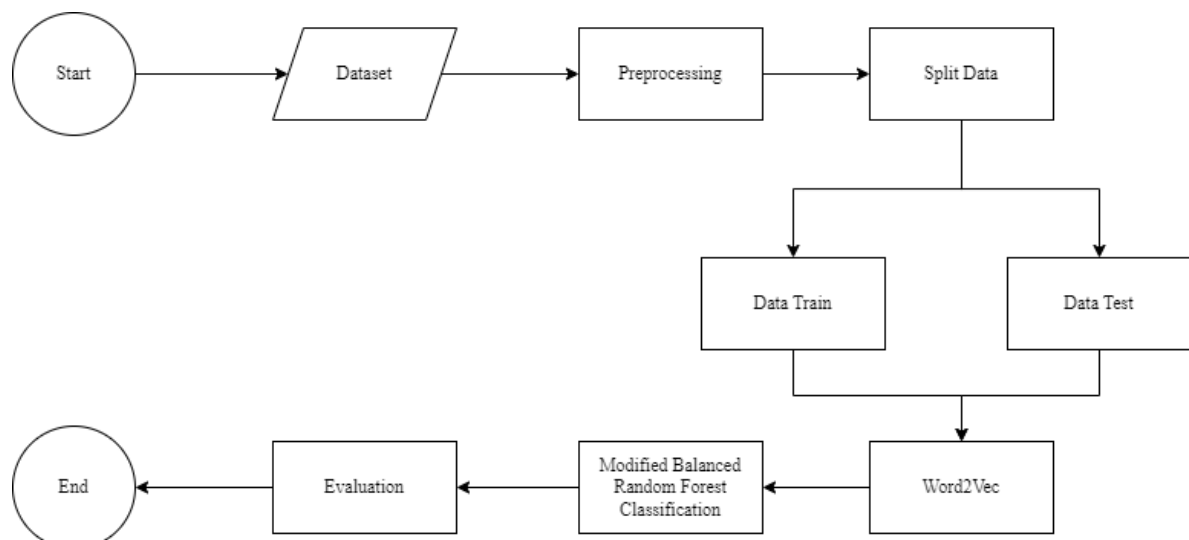


Figure 1. System Design Flow

2.1 Dataset

The datasets collected are English movie reviews data from the rotten tomatoes website obtained from Kaggle totaling 1,130,017 data. In this study, the only data used is movie review data that started from 2020 which amounted to 45,746 data. Sorting the data is done so that when running the program it does not take a long time. The dataset has been labeled with a positive class (Fresh) and a negative class (Rotten). Table 1 is an examples of datasets that have been labeled and Figure 2 is a diagram of the distribution labels on the dataset used.

Table 1. Dataset Examples

Label	Sentence
Fresh (Positive)	“Thought-provoking, continually riveting, and absolutely unforgettable - and surprisingly designed around a very simple, tightly budgeted, special-effects-free premise.”
Rotten (Negative)	“It has some gags - some are even quite funny - but not nearly enough.”

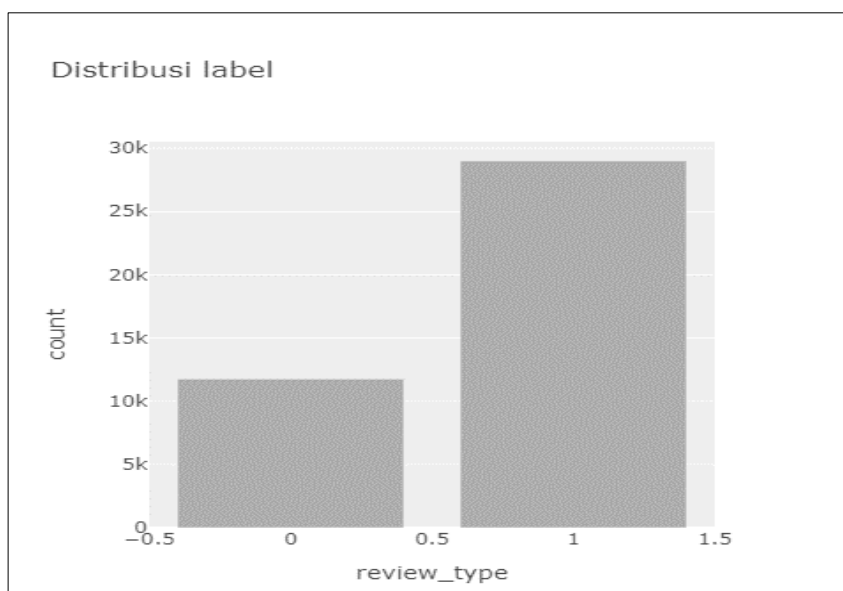


Figure 2. Label Distribution

2.2 Preprocessing

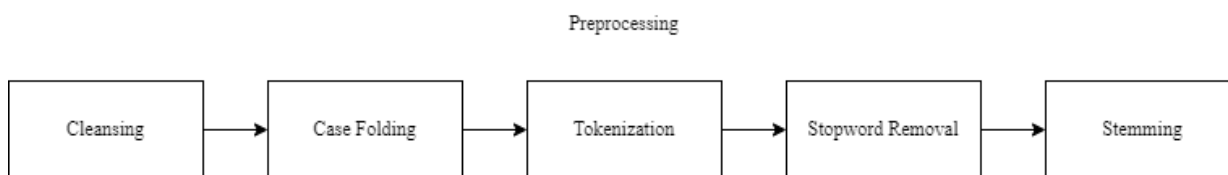


Figure 3. Preprocessing Stages

The stage after collecting and preparing the data is the preprocessing stage. Preprocessing is a stage where data will be cleaned and corrected first before being classified. Figure 3 is the preprocessing stages conducted in this research. Preprocessing stages carried out are Cleansing, Case Folding, Tokenization, Stopword Removal, and Stemming. In addition, the preprocessing stage also changes the data type in the review_type column. The data type change is changing the Fresh (positive class) to 1 and the Rotten (negative class) to 0.

2.2.1 Cleansing

Cleansing is one of the preprocessing stages carried out before classification. Cleansing aims to remove numbers, punctuation marks, and symbols [11]. Table 2 is result of cleansing stage.

Table 2. Cleansing Result

Text	Cleansing Result
“Thought-provoking, continually riveting, and absolutely unforgettable - and surprisingly designed around a very simple, tightly budgeted, special-effects-free premise.”	“Thought provoking continually riveting and absolutely unforgettable and surprisingly designed around a very simple tightly budgeted special effects free premise”

2.2.2 Case Folding



Case Folding is a stage in preprocessing that aims to change the letter in each sentence from capital letters into non-capital letters [12]. Table 3 is result of case folding stage.

Table 3. Case Folding Result

Cleansing Result	Case Folding Result
“Thought provoking continually riveting and absolutely unforgettable and surprisingly designed around a very simple tightly budgeted special effects free premise”	“thought provoking continually riveting and absolutely unforgettable and surprisingly designed around a very simple tightly budgeted special effects free premise”

2.2.3 Tokenization

Tokenization is a stage in preprocessing that aims to divide each sentence into a number of tokens by using delimiters or spaces that match the system requirements [11]. Table 4 is result of tokenization stage.

Table 4. Tokenization Result

Case Folding Result	Tokenization Result
“thought provoking continually riveting and absolutely unforgettable and surprisingly designed around a very simple tightly budgeted special effects free premise”	“[thought], [provoking], [continually], [riveting], [and], [absolutely], [unforgettable], [and], [surprisingly], [designed], [around], [a], [very], [simple], [tightly], [budgeted], [special], [effects], [free], [premise]”

2.2.4 Stopword Removal

Stopword Removal is a stage in preprocessing that aims to remove words that are not important or meaningless words during the classification process [13]. Table 5 is result of stopword removal stage.

Table 5. Stopword Removal Result

Tokenization Result	Stopword Removal Result
“[thought], [provoking], [continually], [riveting], [and], [absolutely], [unforgettable], [and], [surprisingly], [designed], [around], [a], [very], [simple], [tightly], [budgeted], [special], [effects], [free], [premise]”	“[thought], [provoking], [continually], [riveting], [absolutely], [unforgettable], [surprisingly], [designed], [around], [simple], [tightly], [budgeted], [special], [effects], [free], [premise]”

2.2.5 Stemming

Stemming is a stage in preprocessing that aims to convert words that have affixes into basic words [7]. Table 6 is result of stemming stage.

Table 6. Stemming Result

Stopword Removal Result	Stemming Result
“[thought], [provoking], [continually], [riveting], [absolutely], [unforgettable], [surprisingly], [designed], [around], [simple], [tightly], [budgeted], [special], [effects], [free], [premise]”	“thought provok continu rivet absolut unforgett surprisingli design around simpl tightli budget special effect free premis”

After going through several preprocessing stages such as Cleansing, Case Folding, Tokenization, Stopword Removal, and Stemming, an example of clean sentence results can be shown in Table 7.

Table 7. Preprocessing Stages Result

Text	Preprocessed Text
“Thought-provoking, continually riveting, and absolutely unforgettable - and surprisingly designed around a very simple, tightly budgeted, special-effects-free premise.”	“thought provok continu rivet absolut unforgett surprisingli design around simpl tightli budget special effect free premis”

2.3 Split Data

After doing the preprocessing stage, the next stage is split data. Split data in this research is done by dividing the data into 20% for test data and 80% for train data. Table 8 is the result of the split data.

Table 8. Split Data Result

Split Data	Data Train	Data test
Total Data	32589	8148
Fresh (Positive Class)	23192	5794
Rotten (Negative Class)	9397	2354

2.4 Word2Vec

Word2Vec is one of the word embedding methods that aims to present each word in a sentence into a vector [14]. In this study, Word2Vec is used as a feature selection. Word2Vec maps words in a continuous vector space, words with comparable semantic properties are processed and mapped in nearby vector spaces [15]. There are two types of models in Word2Vec, which are known as Skip-Gram and Continuous Bag of Words (CBOW). In this research, Word2Vec model used is Skip-Gram model. Skip-Gram model operates by generating predictions of adjacent words in a sentence. Skip-Gram model will increase the mean log probability of a word [16]. Skip-Gram model equation can be shown below.

$$\frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j} | w_t) \tag{1}$$

Description:

$p(w_{t+j} | w_t)$ = Word probability in the current word

w_t = Center word

w_{t+j} = Word after word centre

c = Size of the training context

Skip-Gram model architecture in Figure 4 is an implementation of the equation above.

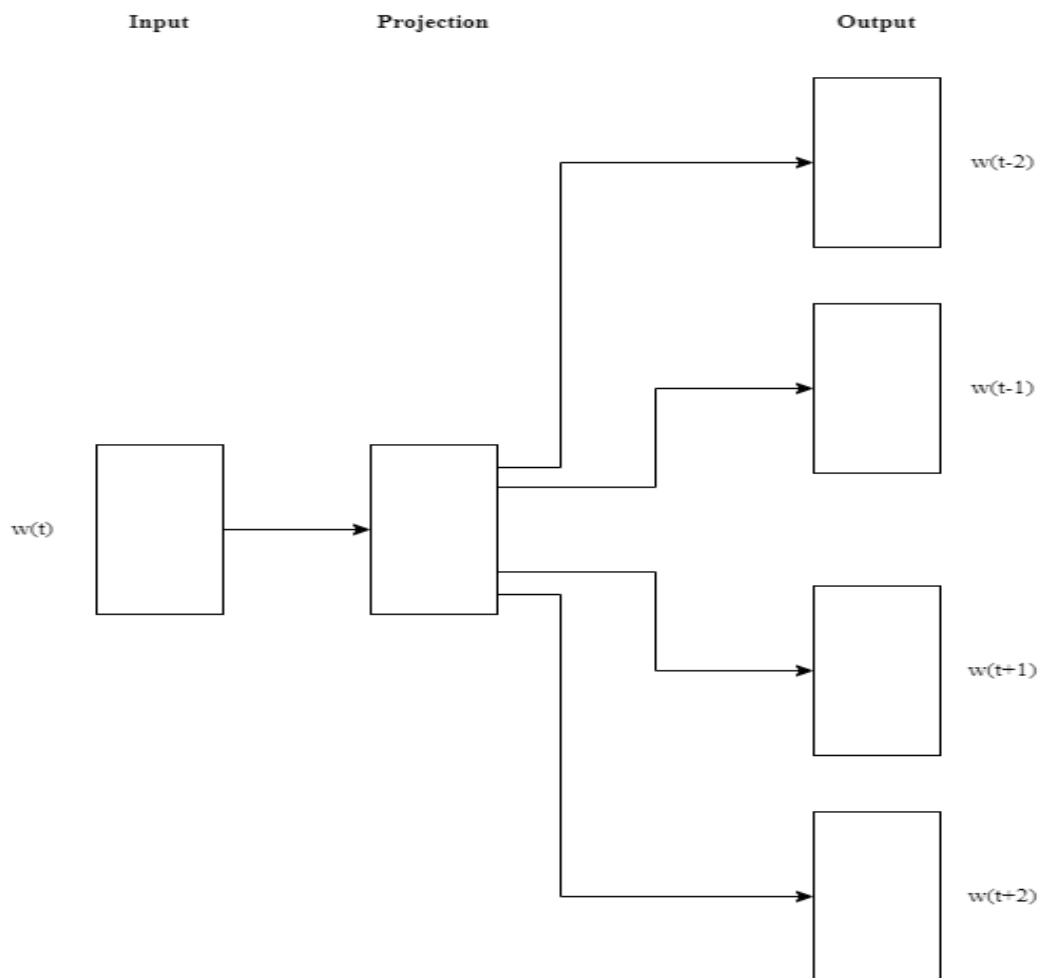


Figure 4. Skip-Gram Architecture [17]

2.5 Modified Balanced Random Forest

After performing the feature extraction stage using Word2Vec, the next stage is classification. Classification method used in this study is Modified Balanced Random Forest. Modified Balanced Random Forest is a modified method of Random Forest and Balanced Random Forest that can overcome imbalanced data. This method modifies Balanced Random Forest algorithm process by reducing most of the data [18]. This method can also improve accuracy and reduce time complexity. To address the issue of imbalanced dataset in this study, this method includes an oversampling stage in its process. Figure 5 is the model of Modified Balanced Random Forest.

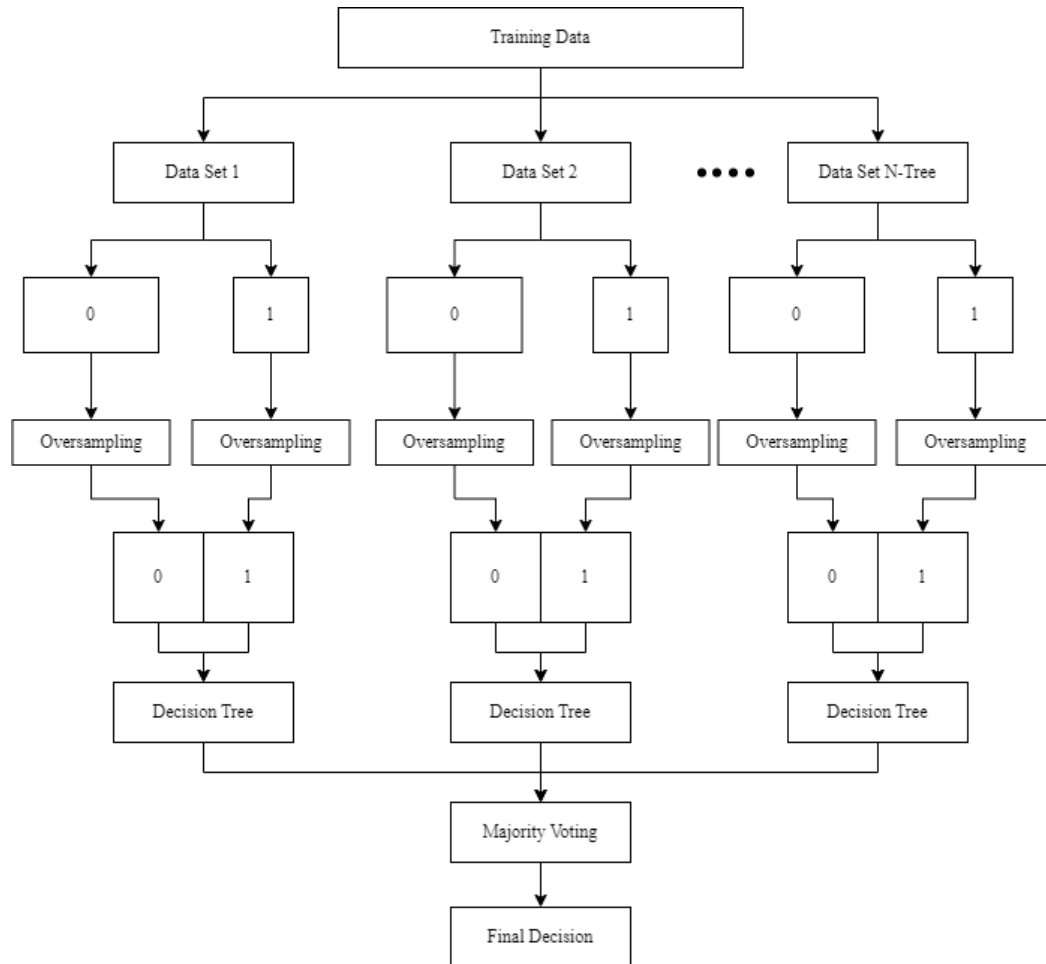


Figure 5. Modified Balanced Random Forest Model [19]

2.6 Evaluation

The evaluation stage in this research is carried out using a confusion matrix which aims to determine classification results in form of predicted class and actual class. Performance of classification model is generally evaluated using the data matrix [20]. Table 9 is a confusion matrix table.

Table 9. Confusion Matrix

Confusion Matrix		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Description:

TP = positive predicted data and positive actual data

FN = negative predicted data and positive actual data

FP = positive predicted data and negative actual data

TN = negative predicted data and negative actual data

To measure performance of classification process, f1-score, precision and recall are calculated. The following are the formulas for calculating evaluation:

F1-Score is a measure of combining recall and precision into a single average value for comparison [21]. F1-score can be expressed using the following formula:

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

Precision is quantifies the level of correctness when predicting a specific class [21]. Precision can be expressed using the following formula:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Recall can be defined as the capacity of a predictive model to correctly identify and choose examples belonging to a specific class [21]. Recall can be expressed using the following formula:

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

3. RESULT AND DISCUSSION

In this study, after going through several stages of preprocessing the data that was successfully obtained was 40737 data. Then the data was split with a ratio of 80:20 which resulted in 32589 train data and 8148 test data. The data that has been split then enters the next stage, namely performing feature extraction with Word2Vec. In feature extraction, each word in the data will be converted into a vector. After going through feature extraction stage, next stage is to carry out the classification process using Modified Balanced Random Forest method. There are three scenarios carried out in this research. First scenario is testing to find out the performance results when using stemming and without using stemming at preprocessing stage. This first scenario was conducted to determine effect of using stemming in preprocessing stage on performance results. Second scenario is to compare use of dimension 100 and dimension 300 in Word2Vec feature extraction. This second scenario is conducted to determine effect of dimension selection in Word2Vec feature extraction on performance results. Third scenario is to compare performance results between Modified Balanced Random Forest and Random Forest methods. This third scenario was conducted to determine the best classification method used in sentiment analysis of movie reviews. Table 10 are several of the scenarios conducted in this study.

Table 10. Experiment Scenario

Scenario	Experiment
1	Compared use of stemming in preprocessing stage
2	Compared use of dimension 100 and dimension 300 in Word2Vec feature extraction
3	Compared the performance results between the Modified Balanced Random Forest and Random Forest methods

3.1 Effect of Stemming

In first scenario, testing is carried out to determine effect of stemming process on performance of system model built. This test is conducted using Word2Vec data of dimension 300 and the Modified Balanced Random Forest method. Table 11 is the result of testing scenario 1.

Table 11. Scenario 1 Test Results

Preprocessing	Precision	Recall	F1-Score
With Stemming	74.54%	93.11%	82.80%
Without Stemming	77.22%	92.45%	84.15%

Based on Table 11, recall in testing without the use of stemming in preprocessing decreased recall by 0.66% compared to testing the use of stemming in preprocessing which was able to produce a recall of 93.11%. While the highest precision and f1-score values are obtained by testing without using stemming at preprocessing stage with a precision of 77.22% and f1-score of 84.15%, both values have increased precision by 2.68% and f1-score by 1.35% compared to testing using stemming at the preprocessing stage which only produces a precision of 74.547% and f1-score of 82.80%. The stemming process can affect performance results, same as the results in paper [7] which produced f1-score of 70.2%. These results were produced by testing without using the stemming process and using the same dataset but using different classification methods. The decrease in performance in testing using stemming occurs because stemming will convert words with affixes into basic words without considering the grammatical structure. Therefore, stemming can produce invalid or meaningless words that can cause the word to be ambiguous. So that preprocessing stage without using stemming produces the best performance results in using the same dataset in the system model built and the other system models that use different classification methods.

3.2 Effect of Dimension on Word2Vec

In second scenario, testing is carried out to determine effect of Word2vec on performance of system model built by comparing use of dimension 100 and dimension 300 in Word2vec. This test is conducted using data without the use of stemming in preprocessing and the Modified Balanced Random Forest method. Table 12 is the result of testing scenario 2.

Table 12. Scenario 2 Test Results

Word2Vec Dimension	Precision	Recall	F1-Score
100	73.26%	92.47%	81.75%

300	77.22%	92.45%	84.15%
-----	---------------	--------	---------------

Based on Table 12, testing using Word2Vec 300 dimensions can produce better performance compared to testing using Word2Vec 100 dimensions. Testing using dimension 300 can produce f1-score of 84.15% while testing using dimension 100 can only produce f1-score of 81.75%. This is the same as testing the dimensions of word2vec in paper [17] which produces the best performance results by using a dimension of 300 compared to using a dimension of 100. This is because by using a longer vector representation, the model can better understand the relationship between words in vector space. Also, the dataset used in this research is quite large, so the selection of higher dimensions is good enough to produce better performance.

3.3 Effect of Classification Method

In third scenario, testing is carried out to determine comparison of performance results between the Modified Balanced Random Forest and Random Forest methods. This test is conducted using data without use of stemming in preprocessing and use of dimension 300 in Word2vec. Table 13 is the result of testing scenario 3.

Table 13. Scenario 3 Test Results

Classifier	Precision	Recall	F1-Score
Random Forest	74.86%	94.97%	83.72%
Modified Balanced Random Forest	77.22%	92.45%	84.15%

Based on Table 13, testing using Modified Balanced Random Forest method can produce better performance results compared to testing using Random Forest. F1-score that can be generated using Modified Balanced Random Forest is 84.15%, this value has increased not too significantly by 0.43% compared to using Random Forest which only produces f1-score of 83.72%. This is the same as in paper [10] which compares the Modified Balanced Random Forest method and Random Forest method. In the test, testing using Modified Balanced Random Forest method has better performance results by producing f1-score of 74% compared to testing using the Random Forest method. This is because Modified Balanced Random Forest method is made to overcome problems with imbalanced data, and in the method can also be applied internally random undersampling or random oversampling processes so that the data in the majority and minority classes are balanced. Therefore, testing using Modified Balanced Random Forest can produce better performance.

4. CONCLUSION

Based on the results of the research that has been conducted on Sentiment Analysis Using Modified Balanced Random Forest and Word2Vec Methods for Movie Reviews on Rotten Tomatoes, by conducting three test scenarios to get the best performance results. First scenario is a test to compare use of stemming in preprocessing stage. Second scenario is a test to compare use of dimension 100 and dimension 300 in Word2Vec feature extraction. Third scenario is a test to compare performance results between Modified Balanced Random Forest and Random Forest methods. From results of the scenario testing that has been done, it can be concluded that use of stemming in preprocessing stage, selection of dimensions in Word2Vec, and use of different classification methods can affect performance results. In first scenario, testing without using stemming in preprocessing gets better performance results compared to testing using stemming in preprocessing. In second scenario, testing using 300 dimensions in Word2Vec gets better performance results compared to testing using 100 dimensions in Word2Vec. In third scenario, testing using Modified Balanced Random Forest method gets better performance results compared to testing using Random Forest method. So that the best model produced from the three scenarios is a model built without using stemming in preprocessing, using 300 dimensions in Word2Vec, and using Modified Balanced Random Forest classification method which produces an f1-score of 84.15%. Some suggestions for further research are, replacing the stemming process to lemmatization in preprocessing stage, trying to do the normalization stage, and selecting a different Word2Vec pre-trained model.

REFERENCES

- [1] A. R. Yosafat and Y. Kurnia, "Aplikasi Prediksi Rating Film dengan Perbandingan Metode Naïve Bayes dan KNN Berbasis Website Menggunakan Framework Codeigniter," 2019. [Online]. Available: <https://jurnal.buddhidharma.ac.id/index.php/alogor/index>
- [2] S. A. Pratomo, S. Al Faraby, and M. D. Purbolaksano, "Analisis Sentimen Pengaruh Kombinasi Ekstraksi Fitur TF-IDF dan Lexicon Pada Ulasan Film Menggunakan Metode KNN," *eProceedings Eng.*, vol. 8, no. 5, pp. 10116–10126, 2021.
- [3] A. Z. Amrullah, A. Sofyan Anas, and M. A. J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *Jurnal*, vol. 2, no. 1, pp. 40–44, 2020, doi: 10.30812/bite.v2i1.804.
- [4] A. Andreyestha and A. Subekti, "Analisa Sentiment Pada Ulasan Film Dengan Optimasi Ensemble Learning," *J. Inform.*, vol. 7, no. 1, pp. 15–23, 2020, doi: 10.31311/ji.v7i1.6171.
- [5] D. I. Af'idah, Dairoh, S. F. Handayani, and R. W. Pratiwi, "Pengaruh Parameter Word2Vec terhadap Performa Deep Learning pada Klasifikasi Sentimen," *J. Inform. Jurnal Pengemb. IT*, vol. 6, no. 3, pp. 156–161, 2021.
- [6] A. Putri, P. Wardani, and M. D. Purbolaksano, "Sentiment Analysis on Beauty Product Review Using Modified Balanced



- Random Forest Method and Chi-Square,” vol. 4, no. 1, pp. 1–7, 2022, doi: 10.47065/josh.v4i1.2047.
- [7] A. Fahmi Sabani, Adiwijaya, and W. Astuti, “Analisis Sentimen Review Film pada Website Rotten Tomatoes Menggunakan Metode SVM Dengan Mengimplementasikan Fitur Extraction Word2Vec,” *e-Proceeding Eng.*, vol. 9, no. 3, p. 1800, 2022.
- [8] M. D. P. Y. Surya, S. Al Faraby, “Analisis Sentimen Terhadap Ulasan Film Menggunakan Word2Vec dan SVM,” vol. 8, no. 4, pp. 4136–4144, 2021.
- [9] M. A. A. Jihad, Adiwijaya, and W. Astuti, “Analisis sentimen terhadap ulasan film menggunakan algoritma random forest,” *e-Proceeding Eng.*, vol. 8, no. 5, pp. 10153–10165, 2021.
- [10] F. N. Zamzami, A. Adiwijaya, and M. D. P., “Analisis Sentimen Terhadap Review Film Menggunakan Metode Modified Balanced Random Forest dan Mutual Information,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 415, 2021, doi: 10.30865/mib.v5i2.2844.
- [11] R. S. Murti and S. Al-faraby, “Analisis Sentimen pada Ulasan Film Menggunakan Word2Vec dan Long Short-Term Memory (LSTM) Pendahuluan Studi Terkait,” *Telkom Univ.*, 2019.
- [12] M. B. Hamzah, “Classification of Movie Review Sentiment Analysis Using Chi-Square and Multinomial Naïve Bayes with Adaptive Boosting,” *J. Adv. Inf. Syst. Technol.*, vol. 3, no. 1, pp. 67–74, 2021.
- [13] C. A. Putri, “Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 2, pp. 181–193, 2020, doi: 10.35957/jatisi.v6i2.206.
- [14] D. T. Hermanto, A. Setyanto, E. T. Luthfi, and U. A. Yogyakarta, “Algoritma LSTM-CNN untuk Sentimen Klasifikasi dengan Word2vec pada Media Online,” *Citec J.*, vol. 8, pp. 64–77, 2021.
- [15] A. Salama and S. Al Faraby, “Klasifikasi Topik Ayat Al- Qur ’ an Terjemahan Berbahasa Inggris Menggunakan Metode Support Vector Machine Berbasis Vector Space Model dan Word2Vec,” vol. 6, no. 2, pp. 9133–9142, 2019.
- [16] I. G. Bagus, A. Bayu, P. Yuda, and M. Dwifabri, “Analisis Sentimen Review Film Berbahasa Inggris Menggunakan Word2Vec dan Naïve Bayes,” 2022.
- [17] S. Rizal, Adiwijaya, and M. D. Purbolaksono, “Sentiment Analysis on Movie Review from Rotten Tomatoes Using Word2Vec and Naive Bayes,” *2022 1st Int. Conf. Softw. Eng. Inf. Technol. ICoSEIT 2022*, pp. 180–185, 2022, doi: 10.1109/ICoSEIT55604.2022.10030009.
- [18] A. M. D. Purbolaksono, M. I. Tantowi, A. I. Hidayat, “Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes,” *J. RESTI*, vol. 1, no. 10, pp. 393–399, 2021.
- [19] Z. Agusta and K. Adiwijaya, “Modified balanced random forest for improving imbalanced data prediction,” *Int. J. Adv. Intell. Informatics*, vol. 5, pp. 58–65, 2019, doi: 10.26555/ijain.v5i1.255.
- [20] I. Prayoga and M. D. P., “Sentiment Analysis on Indonesian Movie Review Using KNN Method With the Implementation of Chi-Square Feature Selection,” *J. Media Inform. Budidarma*, vol. 7, pp. 369–375, 2023, doi: 10.30865/mib.v7i1.5522.
- [21] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, “An improved method to construct basic probability assignment based on the confusion matrix for classification problem,” *Inf. Sci. (Ny.)*, 2016, doi: 10.1016/j.ins.2016.01.033.