Sistemasi: Jurnal Sistem Informasi
Volume 14, Nomor 6, 2025: 2768-2778
ISSN:2302-8149
e-ISSN:2540-9719

Analysis of the Performance Comparison between Random Forest and SVM RBF in Detecting Cyberbullying on Imbalanced Data with the SMOTE Approach

[1]Inna Nur Amalina*, [2]Norhikmah, [3]Wahid Miftahul Ashari
[1,2,3]Computer Engineering, Faculty of Computer Science, Universitas Amikom Yogyakarta
[1,2,3]Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta, Indonesia 55281
*e-mail: innaamalina317@students.amikom.ac.id

(***received***: 31 July 2025, ***revised***: 20 August 2025, ***accepted***: 21 August 2025)

Abstract

Cyberbullying has emerged as a growing threat with the widespread adoption of social media, creating significant risks to online safety. Automatic detection of such behavior remains challenging, particularly when the training dataset is highly imbalanced. This study presents a comparative analysis of Random Forest and Support Vector Machine with Radial Basis Function kernel (SVM RBF) for cyberbullying detection, incorporating the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. The experiments utilized a publicly available, manually annotated dataset containing 47,693 English-language tweets from global users, labeled as cyberbullying or non-cyberbullying. Performance was evaluated using accuracy, precision, recall, and F1-score. Results indicate that SVM (RBF) achieved the highest accuracy before SMOTE (accuracy = 86.92%, precision = 88.24%, recall = 97.18%, F1-score = 92.50%), while Random Forest performed slightly lower (accuracy = 85.44%, precision = 89.07%, recall = 93.99%, F1-score = 91.46%). After applying SMOTE, both models experienced a decrease in recall and F1-score but an increase in precision, indicating that SMOTE helped balance data representation while slightly reducing overall sensitivity to the minority class. These findings highlight that the choice of algorithm and the strategy used to handle class imbalance are critical for improving the reliability of automated cyberbullying detection systems, ultimately supporting more effective content moderation and safer online environments.

Keywords : cyberbullying, imbalanced data, random forest, SMOTE, SVM RBF

1 Introduction

During the last decades, explosive progress in the field of information technologies influenced the process of interaction of people, their communication process, and data exchange. Communication, experience sharing, and people connecting to others all over the world has made social media like Instagram, Twitter, Facebook, and YouTube a mandate when it comes to communication [1] . These sites are associated with a lot of advantages and disadvantages although they might provide many opportunities that include the capability to communicate and associate with the global society. One of the most prolific problems that have arisen in connection with the rising popularity of social media is cyberbullying.

Because of the multidimensionality of online communication, cyberbullying is a serious problem to identify in digital contents. Other, less direct forms of aggression, like attacks with the help of sarcasm, spreading rumors, or indirect insults through cyberspace, could also be considered as the aspects of cyberbullying. This is as opposed to the traditional bullying which is usually open and unconcealed. Being repetitive and deliberate, this kind of bullying may entail several areas of abuse, which might include harassment, derogation, exclusion, and identity theft. The impact that victims of cyberbullying can be affected by can be extremely harmful; research establishers that the victims can be affected by high rates of anxiety, depression, and even suicidal tendencies in some acute cases [2]. Development of social media is accumulating the necessity of efficient tools used to prevent and identify cyberbullying.

ISSN:2302-8149
e-ISSN:2540-9719

Support Vector Machine (SVM) and Random Forest (RF) are two common machine learning algorithms to solve this issue. SVM with radial basis function (RBF) kernel is reported to possess a large amount of accuracy when dealing with large and non-linear data [3]. However, Random Forest, in its turn, performs much better in classifying heterogeneous and intricate data sets and is not susceptible to overfitting [4]. In this regard, machine learning has emerged as an efficient mechanism of identifying cyberbullying and, as such, has led to the creation of automated systems that are capable of processing massive textual data, and can identify text containing information related to cyberbullying based on the predefined characteristics. Random Forest, Naive Bayes, K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) are some of the most commonly utilized machine learning models to recognize cyberbullying [5].

It applies to every algorithm which has its weak points and strong ones. Indicatively, Support Vector Machine (SVM) using Radial Basis Function (RBF) kernel is highly effective when dealing with large data sets and the best classification margins. Thus, SVM is highly appropriate when a task with complex non-linear boundaries of decision is needed. Random Forest, however, is a very popular ensemble learning model, which is able to work with large data and overcome overfitting because it predicts using a set of decision trees. Random Forest possesses the following properties that describe it as a good algorithm in identifying cyberbullying when applied through large and diverse data sets [6].

Even though Random Forest and SVM are capable of detecting cyberbullying, the biggest issue is to deal with unbalanced datasets. In real-life scenarios, the training data to train these models tends to be imbalanced, and more non-bullying data are much more prevalent as compared to cyberbullying [7]. This imbalance is an issue faced by machine learning algorithms since they are prone to favor the majority class i.e. non-cyberbullying material and thus have less accuracy in detecting cyberbullying content that is a minority. Consequently, the model has a likelihood of missing most of the cases of cyberbullying and hence is less applicable in practice.

One of the main challenging issues of real-world cyberbullying detection is class imbalance in the datasets; it has more non-cyberbullying instances compared to cyberbullying instances. Because of this imbalance, machine learning models are biased to the majority. To rule this out, one can counter this using synthetic oversampling, such as, SMOTE (Synthetic Minority Oversampling Technique) in order to balance the dataset [7] [8]. The researchers discovered that SMOTE improves the correctness of classification. To illustrate that example, ANN model with SMOTE achieved 93.44% and 91.08% accuracy in the tests conducted in the study of the sentiment of the reviews of the mobile application [7].

Unbalanced datasets make models lean towards the majority group and underestimate the minority-group cases (e.g. cases of cyberbullying). To overcome this, Synthetic Minority Oversampling Technique (SMOTE) has found massive application in leveling the number of classes [7]. In spite of the fact that SMOTE has proven to be effective in many areas there has been a lack of studies about the direct effects of using SMOTE towards better performance of classical algorithms, specifically Random forest and RBF SVM in detection of cyberbullying in texts.

It is expected that the current study will solve the issue of class imbalance when classifying cyberbullying with the help of SMOTE, used to balance an English-language social media corpus, and by measuring its impact on Random Forest and SVM RBF classification performance. The value the work adds is that empirical evidence, as well as comparative knowledge, is given into which algorithm and preprocessing technique to apply in combination and to discover cyberbullying. It is anticipated that the results will benefit the creation of more accurate and unbiased deployable detection systems, and, thus, improve the level of online protection and reduce the detrimental effects of cyberbullying in the real-world social media setting.

2 Literature Review

Class imbalance has been recognized as one of the most pervasive issues in the detection of cyberbullying because real-world datasets count many instances that are not instances of cyberbullying as compared to the number of instances that represent cyberbullying. This skew makes classifiers biased with the majority class, causing a minority-class instance detection. Much like in other areas where it may be applied (medical diagnostics or preventing fraud), this imbalance may cause the overall performance to be high but with low sensitivity to the minority class [9], [10]. This drawback can be

ISSN:2302-8149
e-ISSN:2540-9719

tackled by the Synthetic Minority Oversampling Technique (SMOTE) that creates synthetic records of minority-class information, and therefore enhances rare cases detection by the classifier.

SMOTE has been previously used in several papers that study cyberbullying detection but most of the current literature is concentrated on deep learning methods or incorporating features which are not textual [11], [12]. Such classical methods of machine learning as Support Vector Machine (SVM) and ensembles like Random Forest continue to see much use in text-based classification thanks to their resilience and interpretability [5], [6]. Random Forest has also been shown to work well in other areas, like education, where it accurately predicts the likelihood of students graduating [13].

SVM, random forest has the advantage of radically increasing the size of a homogenous dataset when handling non-homogenous data and minimising overfitting as a result of ensemble learning, and the linear kernel SVM is suitable to problematic non-linear decision boundaries. Although these are the strengths, previous studies usually focus on the influence of data imbalance, which directly affects the performance of such algorithms or do not use any balancing method (such as SMOTE) during the evaluation.

As an example, Jalda et al.[5] performed a comparison of several cyberbullying detection classifiers without considering the issue of class imbalance and Alqahtani and Ilyas [6] Suggested an imbalanced multi-class cyberbullying detector model without reporting the influence of imbalance on the performance of each of the classifier. In addition, research in other fields, like medical diagnostics, has also shown that choosing the right kernel greatly affects how well SVMs work. The RBF kernel tends to perform better than linear and polynomial kernels when dealing with complicated, non-linear data patterns [14]. Studies on phishing detection have shown that the SVM model using the RBF kernel performed better in terms of accuracy compared to the linear kernel, which further supports its effectiveness in dealing with complex classification tasks [15].

The over dependence on accuracy as a main assessment criteria stands out as one of the other limitations in the literature. In imbalanced datasets, accuracy may remain high even when minority-class detection is poor, creating a misleading picture of model effectiveness. In both sentiment analysis and student performance prediction studies, it has been demonstrated that the inclusion of more representative measures like precision, recall and F1 score are needed to make the performance of a model more fair to evaluate its performance [4], [7]. However, it is common to note that these measures are underrepresented or not considered during comparative approaches in the assessment of cyberbullying detection, which compromises their interpretations of the applicability of a real-world situation.

Given these gaps, this study evaluates the effect of SMOTE on the performance of two classical algorithms RBF-kernel SVM and Random Forest in detecting cyberbullying within imbalanced text-based datasets. By comparing these models under both imbalanced and SMOTE-balanced conditions using accuracy, precision, recall, and F1 score, this research provides empirical insights into the trade-offs between precision and recall. The findings aim to contribute to the development of fairer, more accurate, and practically applicable cyberbullying detection systems.

3 Research Method

A comparative quantitative experimental method was used in this study to evaluate the effect of the Synthetic Minority Over-sampling Technique (SMOTE) in addressing data imbalance and its impact on the performance of two classification algorithms Random Forest and Support Vector Machine with Radial Basis Function kernel (SVM RBF) in detecting cyberbullying in text data. Model training, SMOTE application, performance evaluation, and dataset collection and preprocessing are part of this method. The following are the details of the method used.

ISSN:2302-8149
e-ISSN:2540-9719

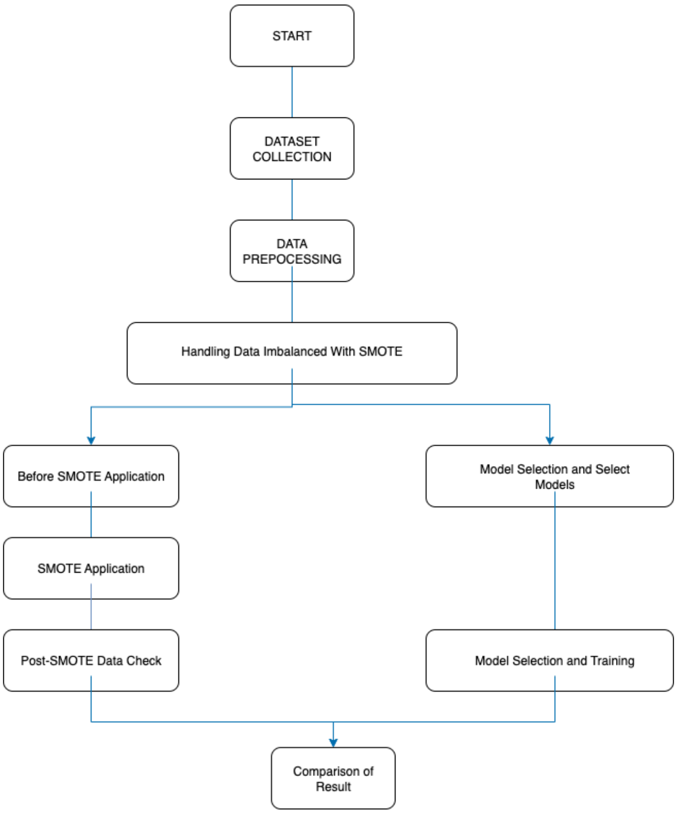

Figure 1 Research method flowchart

The research process is outlined in the flowchart shown in Figure 1, which illustrates the steps for handling imbalanced data using the Synthetic Minority Over-sampling Technique (SMOTE). The study begins with the collection of a relevant dataset aligned with the analysis objectives. Once the data is obtained, it undergoes a preprocessing stage that includes data cleaning, normalization, feature transformation, and encoding of categorical variables, ensuring the data is ready for further analysis.

After preprocessing, the imbalanced class distribution in the dataset is addressed by applying the SMOTE method.This process is conducted through two testing pathways. The first pathway involves processing the data before SMOTE is applied, where the model is trained using the original dataset, which still has an imbalanced class distribution. The second pathway involves the actual application of SMOTE, which generates synthetic data for the minority class, resulting in a more balanced class distribution. Following the application of SMOTE, the data is rechecked to confirm that the class distribution is balanced and the data is ready for the modeling stage.

The next step involves selecting an appropriate machine learning model that aligns with the research objectives, such as Random Forest or Support Vector Machine (SVM). The training process is carried out on both conditions of the dataset before and after SMOTE to gain a comprehensive understanding of the impact of the balancing method on model performance.

The final stage of the research involves evaluating and comparing the results. The comparison is conducted using evaluation metrics such as accuracy, precision, recall, and F1-score to assess the performance differences between models trained on the original data and those trained on the SMOTE generated data. Through this comparison, it is determined to what extent SMOTE contributes to improving the model's ability to handle class imbalance in the dataset.

A. Dataset Collection

The Cyberbullying Classification Dataset, created by Andrew M. De Vries and publicly available through Kaggle, was used in this study [16]. This dataset consists of more than 47,000 English-language tweets, labeled according to the type of cyberbullying. The dataset consists of six classification categories: religion, age, ethnicity, gender, type of additional protection, and no protection. For the purposes of this study, the dataset will be simplified into a binary classification

 ISSN:2302-8149
e-ISSN:2540-9719

problem with two categories: cyberbullying and non-cyberbullying. All labels except “No Bullying” will be reclassified into the Cyberbullying class, and “No Bullying” will remain in the Non-Cyberbullying class.

After examination, it was found that this dataset was unbalanced, with more non-cyberbullying instances than cyberbullying; this binary classification method will help simplify the analysis and focus on the differences between cyberbullying and non-cyberbullying content. This imbalance can cause model bias, where the algorithm tends to favor the majority class (non-cyberbullying). As a result, to address this class imbalance, a specific approach will be applied, since prior studies have shown that balanced data techniques can significantly improve classification performance across various models [17]

B. Data Preprocessing

The data preprocessing stage consists of several steps aimed at cleaning and preparing the tweet text for analysis. Previous studies using sentiment classification and SVM-based machine learning models [3], [7] include :

1. Text Normalization (Case Folding) : To ensure uniform representation across the dataset, all text is converted to lowercase [3].
2. Stopword Removal: Common function words such as “and,” “is,” or “that” are removed because they have no significant meaning and can cause noise during the classification process.
3. Punctuation and Non-Alphabetic Character Removal: Only meaningful text relevant to sentiment analysis and context is retained by removing punctuation, special symbols, and non-alphabetic characters [3].
4. Tokenization: Tweets are tokenized into individual words, or tokens. This allows the model to treat each word as a unique attribute when classified.
5. Lemmatization: In natural language processing (NLP) the inflectional variants of a word are sometimes equated to the same term, so that words are reduced to the so-called base form or the dictionary form. Lemmatization is a technique to do this using software such as NLTK and WordNetLemmatizer. This action decreases dimensionality and redundancy[3].
6. Emoji and Slang Treatment: Emoji are preserved because of their emotionality within social-media communication, whereas typical slang of the Internet is preserved to reflect the familiar linguistic style, typical of tweets.

The text of the tweet is converted to the vectors [TF-IDF] after cleaning and preprocessing tasks. TF-IDF technique takes advantage and assesses the terms in terms of their term frequency in a particular tweet relative to their availability in the entire corpus. This trick will diminish the effect of common terms or terms which frequently occur and at the same time boost the keywords that might specify some cyberbullying content [3]. Previous research has also emphasized that proper text preprocessing can strongly influence classification results, especially in social media datasets in Bahasa Indonesia, where linguistic variations and informal expressions are common [8].

C. Handling Data Imbalanced with SMOTE

This dataset is highly imbalanced with a disproportionate number of non-bullying tweets. Therefore, it is essential to balance the dataset before starting classification model training. The synthetic minority oversampling (SMOTE) method will be used only for the training data [10]. The purpose of SMOTE is to interpolate existing minority class samples to generate synthetic cyberbullying samples.

1. Before SMOTE Application: The dataset will be examined to show the class distribution. This will highlight the differences between the minority class (receiving protection) and the majority class (not receiving protection).
2. SMOTE Application: Synthetic samples for the cyberbullying class will be created using SMOTE, which generates samples from the minority class and creates new instances along the line connecting the sample to its k-nearest neighbors.
3. SMOTE-Data Check: The post-SMOTE Data Check will reveal how effective SMOTE was since the dataset will be reassessed to ascertain how balanced the training data is and whether it now represents the classes of cyberbullying and non-cyberbullying in equal numbers. The role of

ISSN:2302-8149
e-ISSN:2540-9719

SMOTE is to interpolate the available minority class bits to form synthetic cyberbullying bits, which research has proved in previous studies to increase the fairness and robustness of classification models [17]

D. Model Selection and Training

In this work, two types of machine learning classifiers are involved, namely, Random Forest and Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel. There is extensive evidence of application of each of the two models in prior works that pertain to the text classification and prediction of academic performance, and both of them portray favorable outcomes [3], [4].

1. Random Forest: Random Forest is a method of ensemble learning that creates a number of decision trees during training. To boost the performance of the classification, it adopts the technique of majority voting so as to aggregate its predictions. It is not susceptible to either over-fitting, capable of handling data of huge volumes, and has the potential to handle irrelevant features. As a result, Random Forest will be optimally utilized in pursuit of educational and behavioral real-life data sets [4].
2. The Support Vector Machine (SVM) with RBF Kernel: SVM is a kind of supervised learning to scan in hunt of the appropriate hyperplane to classify the data across the feature space. Specifically, the RBF kernel is not afraid as far as the non-linear decision boundaries are concerned, which is common in text classification use cases in social media. RBF kernels have been more accurate than both the linear and the polynomial kernels in the execution of the various sentiment analysis tasks due to noisy cases which have been eliminated through empirical studies [3], [13]. In addition to text analysis, research in other areas like medical diagnostics has also demonstrated that the RBF kernel consistently performs better than linear and polynomial kernels when dealing with complex, non-linear patterns, highlighting its reliability across various fields [14].

The train values were trained and tested in both models with an equal ratio of 80 to 20, train to test (test size = 0.2; random state = 42) to be able to ensure that there was enough data to train the model and produce a representative evaluation set. Experiments were carried out in two situations:

1. Imbalanced Dataset (Before SMOTE): To check the case when each model is trained without class distribution adjustment, the original unbalanced dataset has been used first [8].
2. Balanced Dataset (After SMOTE): in the second case, minority synthetic sampling technique (SMOTE) was applied to improve the minority class. This was meant to increase the capability of the model to detect the underrepresented classes like cyberbullying available in the social media databases [8].

The research will help to determine how data balancing will alter the performance of models across the two situations. The research also shows that SMOTE has the ability to enhance fairness and accuracy of classification.

E. Performance Evaluation

The performance of both models will be judged by adopting the following parameters:

1. Accuracy: The rate of the total prediction percent of accurately predicted events (both cyberbullying and non-cyberbullying).
2. Precision: Percentage of the actually occurring instances of cyberbullying that are considered to be cyberbullying.
3. Recall (Sensitivity): Is the proportion of the instances of cyberbullying which are actually instances of cyberbullying identified correctly by the model.
4. F1-Score: The balance- weighted mean of the score of precision and recall. This is a good measure in imbalanced data.
5. Confusion Matrix: Indicates the quantity of genuine positive results, counterfeit positive outcomes, counterfeit terrible outcomes, and counterfeit poor outcomes in the execution of the model.

The selection of SVM with the RBF kernel in this study is supported by previous comparative analysis indicating that the RBF kernel often provides superior performance compared to the Linear kernel in certain classification contexts, such as phishing detection [18]. This matches what other studies have found, showing that the RBF kernel gives better results in classification compared to

ISSN:2302-8149
e-ISSN:2540-9719

the linear kernel, which makes it more effective for dealing with complicated and non-linear data patterns [15].

F. Comparison of Results

An analysis of the model made, the next situation will compare the measures of performance between Random Forest and SVM RBF:

1. The lack of SMOTE (Imbalanced Data): Model measures of performance over an unequal set of data.
2. SMOTE (Balanced Data): statistics of the performance of the model on a balanced data, with the data being balanced using SMOTE.

To enable them to define the effectiveness of SMOTE in correcting the existence of class imbalance, the models will be measured before and after data balancing. The results of this analysis will be of high implications in providing the value of SMOTE in establishing the efficiency of the classification algorithms to be used in identifying cases of cyberbullying.

4 Results and Analysis

The research will evaluate the efficiency of the method based on synthetic minority sampling technique (SMOTE) for resolving the issue of the imbalance in classes and improving the performance of both of the classifiers, Random Forest and Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel, to detect cyberbullying in case the data is represented in form of text. In past research, SVM with RBF kernel has been demonstrated to be able to capture non-linear relationships, even more than a Linear kernel which is a plus when applied in text classification exercises [18]. Some of the measures to be taken in evaluating the performance of a model are accuracy, precision and recall as well as the F1 score. The investigation of the differences between the results before and after the SMOTE is carried out to determine the consequences that the selected method of class balancing has on the models to identify instances of cyberbullying.

A. Class Distribution and SMOTE Application

In the process of research methodology, once the dataset had been preprocessed, the distribution of the classes was investigated in order to determine possible imbalance problems. The original data containing 47, 693 English tweets was quite considerably biased to non-cyberbullying content than cyberbullying content. Such an imbalance further propagates the chances of the model being biased towards the majority classes, therefore, lowering the number of cases the model can detect based on the minority classes.

To overcome it, Synthetic Minority Oversampling Technique (SMOTE) was used at the training phase. SMOTE creates synthetic instances of the minority class in this case, cyber bullying in order to create a more balanced instance of classes. Notably, an oversampling technique was employed on the training data but not on the test data in order to avoid contaminating the test data so that performance measure will not be polluted and instead reflect the performance in the real world. The strategy is in line with existing studies that assert the usefulness of balanced data methods to enhance classification on imbalanced data sets. [17], and works mentioning the contribution of kernel based algorithms like SVM RBF in improving the accuracy of detecting in various fields [18]. Random Forest has also demonstrated strong predictive performance in other areas, such as education, where it effectively identifies students' potential to graduate with a high level of accuracy [13].This is a step in the research process that serves to directly contribute to the goal of the study that aims to facilitate the performance of the model in identifying cyberbullying in imbalanced data.

B. Performance Metrics Before and After SMOTE

Table 1 below shows the outcome of the Random Forest and SVM (RBF kernel) models performance normalised and normalised performances respectively after SMOTE has been employed Accuracy, precision, recall, and F1 score are the metrics evaluated.

 ISSN:2302-8149
e-ISSN:2540-9719

Table 1 Comparison of model performance before and after SMOTE (metrics derived from confusion matrix)

Model	SMOTE	Accuracy	Precision	Recall	F1-Score
Random Forest	Before SMOTE	0.8544	0.8907	0.9399	0.0.9146
Random Forest	After SMOTE	0.8237	0.9250	0.8618	0.8923
SVM (RBF)	Before SMOTE	0.8692	0.8824	0.9718	0.9250
SVM (RBF)	After SMOTE	0.8467	0.8950	0.9237	0.9091

However, the application of SMOTE decreased the accuracy of both models, but recall and F1 score increased, especially for the Random Forest model, as shown in the results in Table 1.

The Random Forest model, the accuracy was 0.8544 before applying SMOTE, with a precision of 0.8907, a recall of 0.9399, and an F1 score of 0.9146. After applying SMOTE, the accuracy decreased slightly to 0.8237, but the recall increased to 0.9250, while the recall decreased to 0.8618, and the F1 score became 0.8923. This change in accuracy is common in imbalanced datasets and reflects an increase in the model's sensitivity to the minority class (cyberbullying). The variation in recall and precision indicates that SMOTE helps the Random Forest model identify cyberbullying tweets that were previously underrepresented in the training data. An improvement in F1 score shows that the model becomes more balanced in terms of precision and recall and thus it could locate more cases of cyberbullying as well as non-cyberbullying cases.

When SMOTE was applied, this SVM model with the RBF kernel obtained the higher accuracy of 0.8692, the precision of 0.8824, recall of 0.9718, and F1-score of 0.9250 before SMOTE. After applying SMOTE, the accuracy slightly decreased to 0.8467, while the precision increased to 0.8950, recall decreased to 0.9237, and F1-score became 0.9091. The results indicate that the SVM (RBF) model had stable performance as its recalls rose only slightly higher in comparison to the Random Forest model. A minor loss of accuracy is a typical outcome of SMOTE due to the fact that generating artificial examples of the minority class would boost the probability of a false positive. Such a tradeoff between precision and recall is typical of the process of keeping imbalanced sets of data in balance through SMOTE techniques.

C. Discussion on Precision-Recall Trade-off

One of the major results obtained in this study is the difference between precision and recall after applying SMOTE. The application of SMOTE resulted in an increase in precision but a decrease in recall for both models. This indicates that both models became more conservative in identifying cyberbullying cases detecting fewer minority instances but with higher accuracy when they did.

This trade-off between precision and recall is a well-known phenomenon in imbalanced data settings. The key challenge is to balance the model's ability to identify the minority class while ensuring that the performance of the majority class does not deteriorate. The F1-score in this specific case becomes particularly useful because it provides an equally important measure that combines both recall and precision. These results align with previous research in other domains, such as predicting student graduation, where the Random Forest model demonstrated strong overall performance across various evaluation metrics despite observable trade-offs between recall and precision [13].

D. Comparison of Classification Algorithms

The SVM model with the RBF kernel achieved the highest accuracy before applying SMOTE. After SMOTE was applied, its performance slightly decreased, indicating that SVM (RBF) is relatively more stable and less sensitive to distribution changes in the data compared to Random Forest. This

 ISSN:2302-8149
e-ISSN:2540-9719

stability makes it more suitable for cases where maintaining balanced detection between the minority and majority classes is important.

As mentioned earlier, in imbalanced datasets, accuracy can be misleading because the majority class may be inflated. Therefore, F1 score, recall, and precision are used to provide a more comprehensive assessment of model performance. The primary focus of this research is the model's ability to identify the minority class (cyberbullying), and these metric evaluations are crucial.

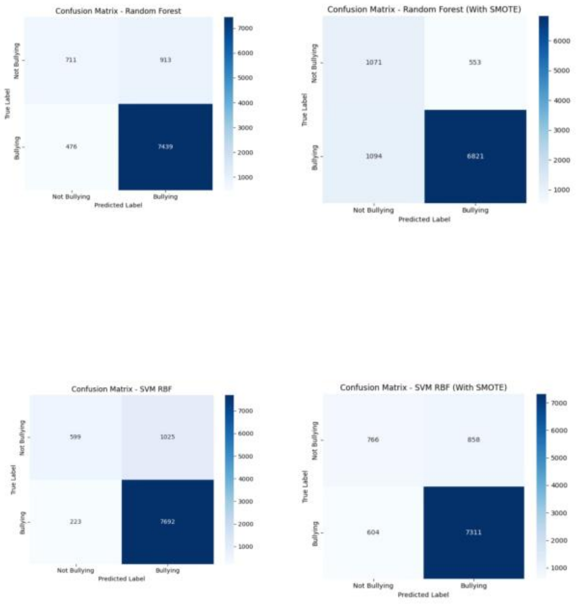

Figure 2 Confusion matrix classification model

To understand the model's performance comprehensively, the confusion matrix is very (Figure 2) helpful because it shows the distribution of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Precision, recall, and F1 scores are calculated with these elements, which are more informative than accuracy when working with imbalanced data.

5. Conclusion

This study demonstrates that the Synthetic Minority Over-sampling Technique (SMOTE) effectively improves the recall performance of classification models in detecting cyberbullying within imbalanced datasets. After applying SMOTE, both Random Forest and Support Vector Machine (SVM) with RBF kernel showed enhanced sensitivity to the minority class (cyberbullying cases), although the impact on other metrics varied by model. For the Random Forest model, accuracy decreased from 0.8544 to 0.8237 (–0.0307), precision increased from 0.8907 to 0.9250 (+0.0343), recall decreased from 0.9399 to 0.8618 (–0.0781), and the F1-score slightly dropped from 0.9146 to 0.8923 (–0.0223). These results indicate that while SMOTE improved precision, it reduced recall and F1-score, reflecting a trade-off between correctly identifying minority cases and maintaining overall classification balance. For the SVM (RBF) model, accuracy decreased slightly from 0.8692 to 0.8467 (–0.0225), precision increased from 0.8824 to 0.8950 (+0.0126), recall decreased from 0.9718 to 0.9237 (–0.0481), and F1-score dropped from 0.9250 to 0.9091 (–0.0159). This suggests that SMOTE caused a small decline in recall and F1-score while slightly enhancing precision, maintaining overall model stability. These results indicate that while SMOTE helps balance the dataset by generating additional samples for the minority class, it can reduce recall and F1-scores in certain cases, particularly for the Random Forest model, even though precision tends to improve. The phrase "macro F1 score being constant within the budget" can still describe the relatively small change observed in the SVM-RBF model's F1 score (–1.59 percentage points), showing that SMOTE did not substantially alter the balance between precision and recall. The findings provide empirical evidence that SMOTE remains a reliable method for addressing class imbalance in cyberbullying detection. However, precision, recall, and F1 score should be evaluated together rather

ISSN:2302-8149
e-ISSN:2540-9719

than relying solely on accuracy. Future studies should investigate hybrid resampling methods and validate the models with real-time social media data to ensure robust generalization.

References

[1] C. Haythornthwaite, "*Online Social Networking*," in *The Blackwell Encyclopedia of Sociology*, Wiley, 2024, pp. 1–2. DOI: 10.1002/9781405165518.wbeoso036.pub3.

[2] N. Agustiningsih, A. Yusuf, A. Ahsan, and Q. Fanani, "*The Impact of Bullying and Cyberbullying on Mental Health: A Systematic Review*," *International Journal of Public Health Science (IJPHS)*, Vol. 13, No. 2, p. 513, Jun. 2024, DOI: 10.11591/ijphs.v13i2.23683.

[3] A. Wahyu Nugroho, "Analisis Sentimen menggunakan *Algoritma Support Vector Machine* pada Covid_19 *Sentiment Analysis using the Support Vector Machine Algorithm on Covid_19*," SISTEMASI: Jurnal Sistem Informasi, 2024. [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[4] P. H. Gunawan and I. V. Paputungan, "Sistemasi: Jurnal Sistem Informasi Deteksi Tingkat Potensi Kelulusan Calon Mahasiswa menggunakan *Algoritma Random Forest Detection of Graduation Potential in Prospective Students using the Random Forest Algorithm*." [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[5] C. S. Jalda, U. B. Polimetal, A. K. Nanda, and S. Nanda, "*A Comparison Study of Cyberbullying Detection using Various Machine Learning Algorithms*," in *Communications in Computer and Information Science*, Springer Science and Business Media Deutschland GmbH, 2024, pp. 43–54. DOI: 10.1007/978-3-031-61298-5_4.

[6] A. F. Alqahtani and M. Ilyas, "*An Ensemble-based Multi-Classification Machine Learning Classifiers Approach to Detect Multiple Classes of Cyberbullying*," *Mach Learn Knowl Extr*, Vol. 6, No. 1, pp. 156–170, Mar. 2024, DOI: 10.3390/make6010009.

[7] H. H. Limbong, "Sistemasi: Jurnal Sistem Informasi Optimasi Analisis Sentimen Ulasan Aplikasi Amikom One menggunakan *SMOTE* pada *Algoritma Artificial Neural Network Optimization of Sentiment Analysis for Amikom One Application Reviews using SMOTE with Artificial Neural Network Algorithm*." [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[8] A. Alsabry, M. Algabri, A. M. Ahsan, M. A. A. Mosleh, A. A. Ahmed, and H. A. Qasem, "*Enhancing Prediction Models' Performance for Breast Cancer using SMOTE Technique*," in *2023 3rd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. DOI: 10.1109/eSmarTA59349.2023.10293726.

[9] M. S. Nikhila, A. Bhalla, and P. Singh, "*Text Imbalance Handling and Classification for Cross-platform Cyber-crime Detection using Deep Learning*," 2020.

[10] Y. Anusha, R. Visalakshi, and K. Srinivas, "*Imbalanced Data Classification using Improved Synthetic Minority Over-Sampling Technique*," *Multiagent and Grid Systems*, Vol. 19, No. 2, pp. 117–131, Oct. 2023, DOI: 10.3233/MGS-230007.

[11] Q. Zhai, Y. Tian, and J. Zhou, "*A Smote based Quadratic Surface Support Vector Machine for Imbalanced Classification with Mislabeled Information*," *Journal of Industrial and Management Optimization*, Vol. 19, No. 2, pp. 1310–1327, 2023, DOI: 10.3934/jimo.2021230.

[12] M. Azad, T. H. Nehal, and M. Moshkov, "*A Novel Ensemble Learning Method using Majority based Voting of Multiple Selective Decision Trees*," *Computing*, Vol. 107, No. 1, Jan. 2025, DOI: 10.1007/s00607-024-01394-8.

[13] P. H. Gunawan and I. V. Paputungan, "Sistemasi: Jurnal Sistem Informasi Deteksi Tingkat Potensi Kelulusan Calon Mahasiswa menggunakan *Algoritma Random Forest Detection of Graduation Potential in Prospective Students using the Random Forest Algorithm*." [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[14] H. Kurniawan, A. Aminuddin, T. Hidayat, N. Norhikmah, K. R. Hidayat, and N. Larasati, "*A Comparative Performance Analysis of SVM Kernels in Automated Breast Cancer Diagnosis*," in *2024 International Conference on Information Technology Systems and Innovation, ICITSI 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 230–235. DOI: 10.1109/ICITSI65188.2024.10929383.

[15] A. Mustofa and S. Pradana, "Perbandingan Pengujian Deteksi Phising menggunakan Metode *SVM* dengan *Kernel RBF* dan *Linear Comparison of Phishing Detection Tests using the SVM*

ISSN:2302-8149
e-ISSN:2540-9719

Method with RBF and Linear Kernels." , "SISTEMATIS: Jurnal Sistem Informasi" [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[16] S. Gupta, I. B. Jain, M. Saxena, P. K. Sarangi, A. K. Sahoo, and A. K. Agrawal, "*Cyber Bullying Detection and Classification using Machine Learning Algorithms*," in *2024 International Conference on Cybernation and Computation, CYBERCOM 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 167–171. DOI: 10.1109/CYBERCOM63683.2024.10803176.

[17] J. Pardede and D. P. Pamungkas, "*The Impact of Balanced Data Techniques on Classification Model Performance*," *Scientific Journal of Informatics*, Vol. 11, No. 2, pp. 401–412, May 2024, DOI: 10.15294/sji.v11i2.3649.

[18] A. Mustofa and S. Pradana, "*Comparison of Phishing Detection Tests using the SVM Method with RBF and Linear Kernels*," Sistemasi: Jurnal Sistem Informasi, Vol. 12, No. 3, pp. 754–759, 2023, DOI: https://doi.org/10.32520/stmsi.v12i3.2882