Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

Credit Card Fraud Detection Using Support Vector Machine: A Study on Data Balancing Strategies

Lailan Sahrina Hasibuan*

[1] School of Data Science, Mathematics, and Informatics, Artificial Intelligence, IPB University, Bogor, Indonesia
Email: lailan.sahrina@apps.ipb.ac.id
Correspondence Author Email: lailan.sahrina@apps.ipb.ac.id
Submitted: **12/10/2025**; Accepted: **26/12/2025**; Published**: 26/12/2025**

Abstract−The rise in credit card transactions has been accompanied by an increase in fraudulent activities. One of the key challenges in detecting fraud is the distribution of the dataset, where fraudulent transactions are significantly outnumbered by normal ones. Despite their low occurrence, fraudulent transactions have a significant impact on the banking sector. Therefore, an effective model is needed to identify and estimate fraudulent transactions. This study aims to generate optimal training dataset from an imbalanced one using Adaptive Synthetic Sampling (ADASYN) to enhance the training process of Support Vector Machine (SVM) model. The dataset used consists of anonymized credit card transactions and labeled as either fraudulent or normal, sourced from the Kaggle dataset. It contains transactions made by European cardholders in September 2013, covering a two-day period with 492 fraud cases out of 284,807 transactions. Three datasets were derived from the original: raw, balanced, and support vector-based balanced. The SVM model training on these datasets resulted in sensitivities of 0.39, 0.64, and 0.70, respectively, while the precision values were 0.92, 0.72, and 0.01. The corresponding f-measure values were 0.55, 0.68, and 0.02. The best performance based on the f-measure was achieved using the balanced version of the raw dataset.

Keywords: Adaptive Synthetic Sampling (ADASYN); Support Vector-Based Dataset; Imbalanced Dataset; Credit Card Fraud Detection; Support Vector Machine (SVM)

1. INTRODUCTION

Credit cards are a mean of payment for financial transactions that do not require cash; therefore, they tend to be more practical. The ease of using credit cards has increased their use in society, including Indonesia. In Indonesia, the number of card instruments, transaction volume, and nominal transactions have continued to increase since 2020 until the latest data for 2022. The number of card instruments increased from 16.94 million units to 17.20 million units, transaction volume increased from 274 million transactions to 342 million transactions, and transaction value increased from 238 trillion to 323 trillion as stated in Asosiasi Kartu Kredit Indonesia at 2024 [1].

Unfortunately, the increase of credit card usage is followed by the increase of the number of crimes for credit card transactions. This term is known as a credit card fraudulent transaction: theft of authorization to use a person's credit card by a criminal group and use it to gain financial profit for them. Although the number of fraudulent transactions is much smaller than the number of normal transactions, financial losses can be broad. In the UK, total losses due to fraudulent transactions reached £580 million in the first six months of 2023. This amount decreased by 2% compared to that in 2022 during the same period. The improved security of credit card transactions used by UK banks has prevented losses of up to £651 million, as reported in UK Annual Fraud Report 2022 [2].

One of the primary challenges in developing a machine learning model for credit card fraud detection lies in the extreme imbalance between fraudulent and legitimate transaction data. In most real-world datasets, the proportion of fraudulent transactions is exceptionally small—often reaching a ratio as high as 1:580 compared to normal transactions. This significant disparity causes traditional learning algorithms to become biased toward the majority class, as they attempt to minimize overall classification error. As a result, the model tends to classify most transactions as legitimate, achieving seemingly high accuracy but failing to correctly identify the rare fraudulent cases that are of greatest interest. Ignoring this imbalance not only reduces the model's sensitivity to fraudulent activities but also undermines the reliability of the detection system in practical applications as stated in Kennedy et al at 2023 [3].

In general, the solutions proposed to address the class imbalance problem in fraud detection can be categorized into three main approaches: data-level, algorithm-level, and hybrid methods that integrate both strategies. The data-level approach focuses on manipulating the dataset itself to achieve a more balanced class distribution, either by oversampling the minority class, undersampling the majority class, or generating synthetic samples to equalize representation. The goal of this approach is to provide the learning algorithm with a more balanced training set, thereby reducing bias toward the majority class as stated in Kraiem et al at 2021 [4]. Meanwhile the algorithm-level approach modifies the learning process or the classifier's internal parameters to make it more sensitive to minority instances as stated in Liu et al at 2005 [5]. The hybrid approach combines both strategies, leveraging the strengths of resampling techniques and algorithmic adjustments to achieve higher robustness and better generalization in imbalanced data scenarios. Such a combination often yields superior performance, especially in complex domains like credit card fraud detection, where data imbalance is severe and dynamic as stated in Desuky and Hussain at 2021 [6].

Data-level approaches, which aim to address class imbalance before model training, can generally be categorized into three main strategies: undersampling, oversampling, and hybrid methods that combine the strengths of both. Undersampling techniques balance the dataset by reducing the number of samples in the majority class, either through random elimination or by using more advanced algorithms designed to retain the most informative instances.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

Common undersampling methods include Random Undersampling (RUS), Tomek Links, and Cluster Centroids, each aiming to remove redundant or less representative samples from the majority class. Although undersampling effectively reduces training time and helps mitigate bias toward the majority class, it also carries a significant drawback — the potential loss of critical information that could help the model better understand the majority class's characteristics. This information loss may lead to a decrease in overall model accuracy, particularly in cases where the dataset is already limited in size [4].

In contrast, oversampling methods address imbalance by increasing the number of samples in the minority class. This can be achieved either by duplicating existing minority samples or by generating new, synthetic instances that mimic the underlying data distribution. By creating synthetic samples rather than exact duplicates, these methods help reduce overfitting and improve the classifier's ability to generalize to unseen data [4].

Finally, hybrid approaches combine both undersampling and oversampling strategies to achieve a more balanced trade-off between data representativeness and training efficiency. In a typical hybrid framework, undersampling is first applied to remove redundant samples from the majority class, followed by oversampling to enrich the minority class with synthetic examples. This combination seeks to preserve essential information from both classes while achieving a more uniform class distribution [4].

The Synthetic Minority Oversampling Technique (SMOTE) is a widely recognized data balancing method introduced by Nitesh V. Chawla in 2002 to address the problem of class imbalance in machine learning datasets. Unlike traditional oversampling, which merely duplicates minority class samples, SMOTE generates synthetic data points to increase the representation of the minority class without causing overfitting. The method works by calculating the distance between each minority class instance and its k nearest neighbors. New synthetic samples are then created along the line segments connecting each selected minority sample to one or more of its nearest neighbors. This interpolation process ensures that the generated samples represent plausible variations of the minority class, effectively expanding its decision region in the feature space [7].

ADASYN is an algorithm for balancing data introduced by He in 2008 [8]. This algorithm improves SMOTE in generating of synthetic data of minority class. SMOTE generates synthetic data with the same number for each data point in the minority class. Synthetic data generation in ADASYN was performed by learning the condition of each of minority data. The more major data that are around a minor data, the more synthetic data are generated for that minor data [7] [8]. According to its ability to learn each data specifically in the minority class, ADASYN is better than SMOTE in generating synthetic data.

Previous researches have been conducted to prevent fraudulent transactions. Li et al (2021) conducted a comparative study of SVM's ability to detect fraudulent transactions. SVM parameters were optimized using the Cukoo Search algorithm, Genetic Algorithm and Particle Swarm Optimization. Based on the accuracy, the SVM model optimized using PSO yielded the highest value of 98.6% [9]. Chang et al (2022) examined credit card transaction data using several machine learning techniques: Logistic Regression, K-NN, Decision Tree, and Randon Forest. This research also applied the SMOTE data-balancing technique to overcome the imbalance between fraudulent and normal transaction data. The Random Forest method is the best method according to the AUC value of 0.95 [10]. Chung et al (2023) applied ensemble learning from K-NN, Linear Discriminate Analysis, and Linear Regression to detect fraudulent transactions. The results show that the ensemble methods are better than single-machine learning methods [11]. Hasibuan et al (2023) compared grid search and Genetic Algorithm (GA) to optimized SVM model in order to identify fraudulent transaction. The dataset was balanced using ADASYN. Based on sensitifity, specificity and time consuming for SVM training, SVM model with grid search achived the best performance [12].

This study employs the ADASYN technique to address the issue of data imbalance commonly found in credit card transaction datasets by generating a balanced set of training samples. The balanced data are then utilized to train a Support Vector Machine (SVM) model for detecting fraudulent credit card transactions. SVM is a robust and widely recognized classifier known for its effectiveness in handling both linear and non-linear data distributions through the use of kernel functions that map data into higher-dimensional spaces [13][14]. Its ability to construct optimal hyperplanes that separate different classes makes it particularly suitable for fraud detection, where distinguishing between legitimate and fraudulent transactions is often challenging due to overlapping data patterns. In addition to training the SVM model on raw and balanced datasets, this study also investigates the impact of utilizing support vectors—data points that define the decision boundaries or hyperplanes in SVM as a new form of training dataset. This additional analysis aims to explore whether focusing on these critical instances can enhance model generalization and improve the overall detection performance in highly imbalanced conditions.

2. RESEARCH METHODOLOGY

2.1 Research Stages

The dataset used in this research is credit card transaction data obtained from the Kaggle open-source repository, which has been widely utilized in fraud detection studies due to its realistic representation of financial transactions. The dataset consists of credit card transactions carried out by European cardholders over a two-day period in September 2013. In total, it contains 284,807 transactions, of which 492 are classified as fraudulent, representing only

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

about 0.17% of the entire dataset. This extremely imbalanced distribution between the normal and fraudulent classes poses a significant challenge for machine learning models, as the minority class (fraudulent transactions) is heavily underrepresented. The dataset includes 31 features, where 28 of them are transformed features derived through Principal Component Analysis (PCA) to protect user confidentiality and prevent direct identification of the cardholders. The remaining features Time, Amount, and Class—are untransformed and represent the transaction timestamp, transaction value, and class label (fraudulent or normal), respectively [15]. The distribution of the two classes in this dataset is illustrated in Figure 1, clearly depicting the stark imbalance between normal and fraudulent transactions.

This dataset has become a benchmark dataset for numerous studies in the field of credit card fraud detection, enabling consistent evaluation and comparison of machine learning algorithms. For instance, researchers in [16] utilized the same dataset to evaluate the capability of self-supervised learning models in detecting fraudulent transactions without relying heavily on labeled data. Meanwhile, researchers in [10] investigated the performance of five different classifiers on this dataset to identify the most effective algorithm for fraud detection tasks. In another study, researchers in [12] applied Genetic Algorithm (GA) and Grid Search optimization techniques to enhance the performance of Support Vector Machine (SVM) models when trained on this dataset. The continued use of this dataset across various studies underscores its relevance, reliability, and importance as a standard reference dataset for evaluating fraud detection methods, particularly those focused on addressing issues related to data imbalance and model optimization.

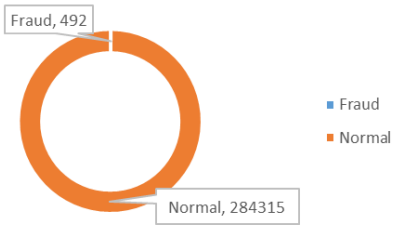


Figure 1. Data distribution

The method of this study is consist of praprocessing data, balancing data, model training, model testing and analysis. The model training consist of three different dataset and also resuting three different models. Figure 2 shows more detail about the method.

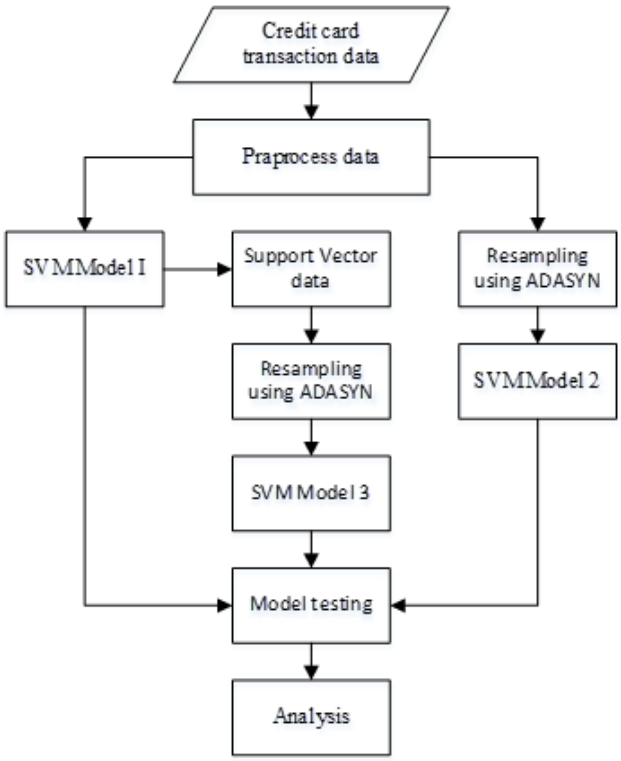


Figure 2. Research Methodology

2.2 Data Praprocessing

Data preprocessing is used to prepare data before being processed by SVM. Handling missing values and scaling data to the same range are preprocessing steps that must be performed, as they significantly affect the performance of

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

SVM. Figure 3 shows summary of the dataset. The dataset does not contain any missing values, but the data range needs to be scaled as stated in [17]. The scale method used was z-score standardization, the common scale method in R [18]. There are 31 features: time, V1-V28, amount and class. The "time" feature denotes the temporal difference, in seconds, between the transaction and the first recorded transaction in the dataset. The "amount" feature represents the transaction value, while "class" is the target variable encoded as 0 or 1 to indicate whether a transaction is normal or fraudulent.

```
      Time              V1                  V2                 V3                 V4                 V5                 V6
 Min.   :     0   Min.   :-56.40751   Min.   :-72.71573   Min.   :-48.3256   Min.   :-5.68317   Min.   :-113.74331   Min.   :-26.1605
 1st Qu.: 54202   1st Qu.: -0.92037   1st Qu.: -0.59855   1st Qu.: -0.8904   1st Qu.:-0.84864   1st Qu.:  -0.69160   1st Qu.: -0.7683
 Median : 84692   Median :  0.01811   Median :  0.06549   Median :  0.1799   Median :-0.01985   Median :  -0.05434   Median : -0.2742
 Mean   : 94814   Mean   :  0.00000   Mean   :  0.00000   Mean   :  0.0000   Mean   : 0.00000   Mean   :   0.00000   Mean   :  0.0000
 3rd Qu.:139321   3rd Qu.:  1.31564   3rd Qu.:  0.80372   3rd Qu.:  1.0272   3rd Qu.: 0.74334   3rd Qu.:   0.61193   3rd Qu.:  0.3986
 Max.   :172792   Max.   :  2.45493   Max.   : 22.05773   Max.   :  9.3826   Max.   :16.87534   Max.   :  34.80167   Max.   : 73.3016
       V7                 V8                  V9                 V10                V11                V12                V13
 Min.   :-43.5572   Min.   :-73.21672   Min.   :-13.43407   Min.   :-24.58826   Min.   :-4.79747   Min.   :-18.6837   Min.   :-5.79188
 1st Qu.: -0.5541   1st Qu.: -0.20863   1st Qu.: -0.64310   1st Qu.: -0.53543   1st Qu.:-0.76249   1st Qu.: -0.4056   1st Qu.:-0.64854
 Median :  0.0401   Median :  0.02236   Median : -0.05143   Median : -0.09292   Median :-0.03276   Median :  0.1400   Median :-0.01357
 Mean   :  0.0000   Mean   :  0.00000   Mean   :  0.00000   Mean   :  0.00000   Mean   : 0.00000   Mean   :  0.0000   Mean   : 0.00000
 3rd Qu.:  0.5704   3rd Qu.:  0.32735   3rd Qu.:  0.59714   3rd Qu.:  0.45392   3rd Qu.: 0.73959   3rd Qu.:  0.6182   3rd Qu.: 0.66251
 Max.   :120.5895   Max.   : 20.00721   Max.   : 15.59500   Max.   : 23.74514   Max.   :12.01891   Max.   :  7.8484   Max.   : 7.12688
      V14                V15                V16                V17                V18                V19                V20
 Min.   :-19.2143   Min.   :-4.49894   Min.   :-14.12985   Min.   :-25.16280   Min.   :-9.498746   Min.   :-7.213527   Min.   :-54.49772
 1st Qu.: -0.4256   1st Qu.:-0.58288   1st Qu.: -0.46804   1st Qu.: -0.48375   1st Qu.:-0.498850   1st Qu.:-0.456299   1st Qu.: -0.21172
 Median :  0.0506   Median : 0.04807   Median :  0.06641   Median : -0.06568   Median :-0.003636   Median : 0.003735   Median : -0.06248
 Mean   :  0.0000   Mean   : 0.00000   Mean   :  0.00000   Mean   :  0.00000   Mean   : 0.000000   Mean   : 0.000000   Mean   :  0.00000
 3rd Qu.:  0.4931   3rd Qu.: 0.64882   3rd Qu.:  0.52330   3rd Qu.:  0.39968   3rd Qu.: 0.500807   3rd Qu.: 0.458949   3rd Qu.:  0.13304
 Max.   : 10.5268   Max.   : 8.87774   Max.   : 17.31511   Max.   :  9.25353   Max.   : 5.041069   Max.   : 5.591971   Max.   : 39.42090
      V21                V22                V23                V24                V25                V26                V27
 Min.   :-34.83038   Min.   :-10.933144   Min.   :-44.80774   Min.   :-2.83663   Min.   :-10.29540   Min.   :-2.60455   Min.   :-22.565679
 1st Qu.: -0.22839   1st Qu.: -0.542350   1st Qu.: -0.16185   1st Qu.:-0.35459   1st Qu.: -0.31715   1st Qu.:-0.32698   1st Qu.: -0.070840
 Median : -0.02945   Median :  0.006782   Median : -0.01119   Median : 0.04098   Median :  0.01659   Median :-0.05214   Median :  0.001342
 Mean   :  0.00000   Mean   :  0.000000   Mean   :  0.00000   Mean   : 0.00000   Mean   :  0.00000   Mean   : 0.00000   Mean   :  0.000000
 3rd Qu.:  0.18638   3rd Qu.:  0.528554   3rd Qu.:  0.14764   3rd Qu.: 0.43953   3rd Qu.:  0.35072   3rd Qu.: 0.24095   3rd Qu.:  0.091045
 Max.   : 27.20284   Max.   : 10.503090   Max.   : 22.52841   Max.   : 4.58455   Max.   :  7.51959   Max.   : 3.51735   Max.   : 31.612198
      V28                Amount          Class
 Min.   :-15.43008   Min.   :    0.00   0:284315
 1st Qu.: -0.05296   1st Qu.:    5.60   1:   492
 Median :  0.01124   Median :   22.00
 Mean   :  0.00000   Mean   :   88.35
 3rd Qu.:  0.07828   3rd Qu.:   77.17
 Max.   : 33.84781   Max.   :25691.16
```

Figure 3. Statistics summary of dataset

2.3 Building The Model

The dataset was split into training and testing sets with an 80:20 ratio at first. The training set was then utilized to develop three SVM models with Radial Basis Function (RBF) as the kernel. The first model was trained on the raw dataset without any prior balancing adjustments. Grid search optimization was applied to determine the optimal model parameters, C and gamma. The selection of grid search optimization was based on previous research, [12] compared grid search and genetic algorithm optimization techniques for finding the optimal hyperparameters in SVM. The results showed that both grid search and genetic algorithm produced SVM models with comparable performance. However, the computational time for the genetic algorithm was significantly higher, taking up to 19 times longer than grid search.

The grid search optimization process was conducted to identify the most effective hyperparameter configuration for the SVM model trained on the raw dataset. Specifically, the range of values explored for the penalty parameter C included (1, 10, 30, 90, 270, 810), while the gamma parameter, which controls the influence of each training example in the RBF kernel, was tested across values of (0.1, 0.033, 0.011, 0.0037, and 0.00123). These parameter combinations were systematically evaluated to determine the optimal balance between model complexity and generalization capability. After obtaining the optimal parameters, the model was validated using the remaining 20% of the dataset, which was intentionally kept in its original imbalanced state to simulate real-world conditions where fraudulent transactions are significantly rarer than legitimate ones.

In the subsequent stage, the raw training dataset was balanced using the ADASYN technique, which generates synthetic samples for the minority class based on the distribution density of existing minority instances. In this study, the number of nearest neighbors used to generate synthetic samples was set to five. The resulting balanced dataset was then utilized to develop the second Support Vector Machine (SVM) model. To achieve optimal model performance, grid search optimization was applied to fine-tune the hyperparameters. The range of candidate values for the C parameter was defined as (0.1, 1, 10, 100, 1000), while the gamma parameter was explored within (0.0001, 0.001, 0.01, 0.1, 1). These combinations were systematically evaluated to identify the configuration that provided the best trade-off between bias and variance. Once the optimal parameters were determined, the model was evaluated using the remaining 20% of the dataset, which was deliberately kept imbalanced to reflect the natural distribution of credit card transaction data in real-world scenarios

In the final stage of experimentation, the support vectors obtained from the first SVM model were extracted to form a new subset of the dataset that represented the most critical boundary instances between classes. To address class imbalance within this subset, the Adaptive Synthetic Sampling (ADASYN) method was again applied, generating additional synthetic minority samples and resulting in a newly balanced training dataset. This balanced support vector dataset was subsequently used to develop the third SVM model. Similar to the previous stages, grid

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

search optimization was employed to determine the best hyperparameter configuration. The search range for the C parameter was set to (0.1, 1, 10, 100, 1000), while gamma values were explored within (0.0001, 0.001, 0.01, 0.1, 1). The model was then evaluated using the remaining 20% of the dataset, which remained unbalanced to simulate real-world fraud detection conditions. Table 1 summarizes the composition and number of records for each dataset used in the experiments.

Table 1. Comparision of number of data before and after balancing

Dataset	Total	Normal transaction	Fraud transaction
Training raw imbalanced	227845	227451	394
Training balanced	454902	227451	227451
Training SV balanced	17021	8484	8537
Testing	56962	56864	98

2.4 Model Evaluation

The performance of the developed models was evaluated using a confusion matrix, as presented in Table 2. This matrix provides a comprehensive overview of the model's classification results, including the number of correctly and incorrectly classified instances for each class. To assess the models more thoroughly, several commonly used evaluation metrics were calculated. These metrics are particularly relevant in fraud detection tasks, where class imbalance can significantly affect performance interpretation. Equations (1)–(5) present the formulas for the evaluation metrics used in this study.

Table 2. Confusion Matrix

Actual	Prediction	
	P	N
P	TP	FN
N	FP	TN

$$Accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \quad (1)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{FP}{(TN+FP)} \quad (4)$$

$$f-masure = \frac{2*Precision*Recall}{(Precision+Recall)} \quad (5)$$

The evaluation metrics employed in this study include accuracy, precision, sensitivity (also known as recall), specificity, and the F-measure. Accuracy measures the overall proportion of correctly classified transactions, providing a general indicator of model performance. However, in highly imbalanced datasets, accuracy alone can be misleading since a model can achieve high accuracy by predicting the majority class more frequently. Precision evaluates how many of the transactions classified as fraudulent are actually fraudulent, making it crucial in reducing false alarms. Sensitivity measures the model's ability to correctly identify fraudulent transactions, which is essential in minimizing missed fraud cases. Specificity, on the other hand, reflects the model's capability to correctly classify genuine transactions. Finally, the F-measure provides a harmonic mean of precision and sensitivity, offering a balanced assessment between detecting actual frauds and avoiding false positives.

In the context of fraud detection, sensitivity and F-measure are considered the most critical evaluation metrics. Since fraudulent transactions represent only a small fraction of the total dataset, the primary challenge lies in ensuring that these rare but significant cases are correctly identified. A model with low sensitivity may fail to detect many fraudulent transactions, resulting in substantial financial losses and security risks. Therefore, improving sensitivity directly enhances the model's practical value in real-world applications. Meanwhile, the F-measure offers a balanced perspective by simultaneously considering sensitivity and precision, providing insight into how well the model detects fraud without generating excessive false positives. This makes the F-measure particularly suitable for assessing model performance on imbalanced datasets, where trade-offs between false negatives and false positives must be carefully managed.

3. RESULT AND DISCUSSION

The experimental results highlight the significant influence of various data balancing strategies on the performance of the Support Vector Machine (SVM) model in detecting fraudulent credit card transactions. To evaluate this impact comprehensively, three distinct datasets were utilized: the raw unbalanced dataset, the balanced dataset generated

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

using the Adaptive Synthetic Sampling (ADASYN) method, and the support vector-based balanced dataset. Figure 4 presents the visualization of the datasets using the PCA-derived features V1 and V2 as the horizontal and vertical axes. The use of V1 and V2 aligns with the PCA-based feature transformation applied to the original dataset, as documented in the Kaggle data source. All datasets appear nearly identical, likely because the visualization uses only 2 out of the 30 available features. Information regarding the variation resulting from dimensionality reduction is not available in the source data.

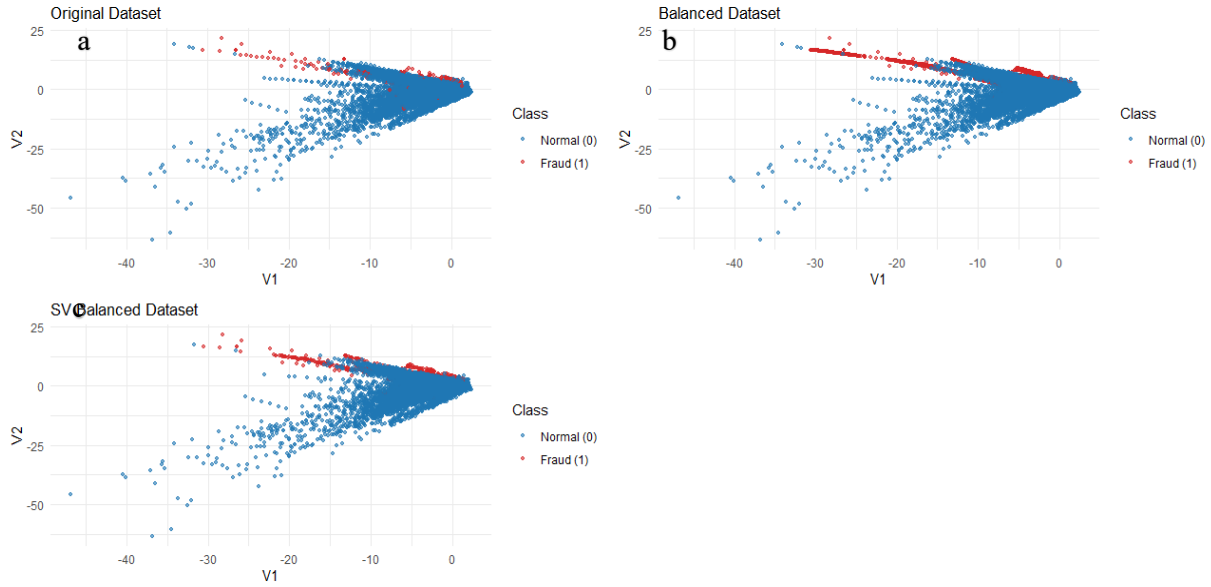


Figure 4. Dataset visualization. (a) The raw imbalanced dataset. (b) The balanced dataset. (c) The support-vector based balanced dataset

3.1 Performance of Different Models

In order to gain a deeper understanding of how data imbalance influences model performance, a grid search optimization process was conducted prior to model training. This procedure aimed to identify the most effective combination of hyperparameters for the Support Vector Machine (SVM), ensuring that the model achieved optimal classification performance under the given dataset conditions. The grid search systematically explored different values of the regularization parameter (C) and the kernel coefficient (gamma), which are crucial in controlling the trade-off between maximizing the margin and minimizing classification errors. After thorough experimentation using the raw (unbalanced) dataset, the optimal hyperparameter values were determined to be C = 10 and gamma = 0.1, indicating that a moderate level of regularization and kernel flexibility produced the best results for this dataset. The range of the domain search is [0.1, 1, 10, 100, 1000] for C and [0.0001, 0.001, 0.01, 0.1, 1] for gamma, as describe in the paper [12]. Table 3, illustrating the consistency and variations in hyperparameter tuning outcomes across different data balancing scenarios. The first Model was trained using raw dataset, the second model trained using balanced dataset, the third model trained using balanced support vector-based dataset. The optimal values of C and gamma were the same across all datasets, namely 10 and 0.1. This is presumed to occur because the datasets share similar characteristics, where fraudulent and normal data overlap in the area near the decision boundary.

Table 3. The best parametes of C and gamma

Parameters	The first model*	The second model**	The third model***
C	10	10	10
gamma	0.1	0.1	0.1

The Support Vector Machine (SVM) model was trained using the optimal hyperparameters obtained from the grid search process and subsequently evaluated on the testing dataset. The model achieved a sensitivity of 0.39, a precision of 0.92, and an f-measure of 0.55. The high precision value indicates that the majority of transactions predicted as fraudulent were indeed true fraud cases, suggesting that the classifier was highly conservative in labeling a transaction as fraud. However, the relatively low sensitivity demonstrates that the model failed to identify a large portion of actual fraudulent transactions, leading to a significant number of false negatives. This limitation is primarily attributed to the extreme class imbalance in the dataset, where fraudulent transactions represented only 0.17% of all records in the testing data. It is important to note that no balancing technique was applied to the testing dataset in order to preserve its real-world distribution, thus providing a realistic assessment of the model's detection capability. Despite the application of hyperparameter optimization, the model's performance remained constrained by the skewed data distribution. These results shows that, while optimization improves parameter tuning, it alone cannot adequately

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

address imbalance issues reinforcing the need for data-level techniques such as ADASYN to enhance sensitivity and achieve more reliable fraud detection results.

To overcome the limitations identified in the model trained on the raw dataset, the training data was reprocessed using the Adaptive Synthetic Sampling (ADASYN) technique to achieve a more balanced class distribution. The number of nearest neighbors for ADASYN was set to five, allowing the algorithm to generate synthetic minority samples effectively and reduce class bias. Following data balancing, grid search optimization was performed once again to determine the most suitable hyperparameters, resulting in C = 10 and gamma = 0.1 as the optimal values. The model trained on this balanced dataset exhibited a substantial improvement in performance, achieving a sensitivity of 0.64, precision of 0.72, and F-measure of 0.68. Compared to the model using the raw dataset, the marked increase in sensitivity demonstrates the model's enhanced capability to detect fraudulent transactions. Although a slight reduction in precision indicates a modest increase in false positive predictions, the overall improvement in F-measure confirms that data balancing using ADASYN significantly strengthens the model's ability to identify fraud more effectively.

To further enhance the fraud detection performance, a third SVM model was developed using a dataset derived from the support vectors of the initial model. This approach was designed to focus the training process on the most critical boundary data points that define the decision margins between fraudulent and non-fraudulent transactions. By utilizing support vectors as the foundation for training, the model was expected to improve generalization and reduce the influence of redundant or less informative samples. Additionally, this method sought to mitigate the adverse effects of class imbalance by emphasizing instances that contribute most significantly to the classifier's decision boundaries.

In binary classification, Support Vector Machine (SVM) employs a hyperplane to separate data based on their respective classes. During the training process, SVM determines the optimal hyperplane that minimizes classification errors while ensuring proper data separation. From the constructed hyperplane, SVM establishes the maximum margin that separates the data, with a tolerance for classification errors regulated by the parameter C. By incorporating this margin, the initial hyperplane transforms into a two-dimensional decision boundary. The margin of the hyperplane is defined by a subset of training samples, known as support vectors, which are selected by SVM during the training process. These support vectors, positioned along the class boundary, play a crucial role in defining the decision boundary and influencing the classification of new samples. Therefore, in this study, dataset generation for training is conducted using support vector data to enhance the model's learning process. Figure 5 shows illustration of SVM hyperplane. In this figure, four samples—(a), (b), (c), and (d)—are shown in relation to the decision boundary. Sample (a) is correctly classified and located well outside the margin, indicating high confidence in its classification. Sample (b) is correctly classified and lies exactly on the margin, serving as a crucial boundary point. Sample (c) is also correctly classified but positioned within the margin, meaning it is correctly labeled yet not sufficiently distant from the decision boundary. Sample (d), however, is misclassified, lying on the wrong side of the hyperplane. Samples (b), (c), and (d) are identified as support vectors, as they directly influence the position and orientation of the decision boundary in the soft-margin SVM formulation.

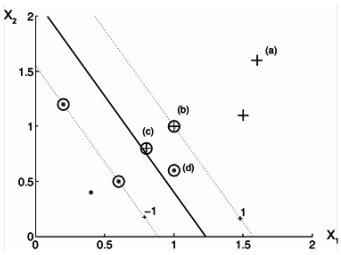


Figure 5. Illustration of soft margin hyperplane.

As in the previous training process, parameter optimization is also performed in this step. After balancing the support vector-based dataset, grid search optimization was applied, yielding C = 10 and gamma = 0.1 as the optimal parameters. The model trained on this dataset achieved a sensitivity of 0.70, a precision of 0.01, and an f-measure of 0.02 after tested to the test dataset. Despite the improvement in sensitivity, the drastic drop in precision indicates a significant increase in false positives, leading to poor overall performance. These results suggest that while support vector-based balancing enhances sensitivity, it may not be effective for maintaining precision in fraud detection. Consequently, the balanced raw dataset remains the best-performing approach based on the f-measure. Table 4 shows matrix confusion for each model, where class 1 indicates fraud transaction while 0 is normal. The first model was trained using raw dataset, the second model was trained using balanced dataset, the third model was trained using balanced support vector-based dataset.

Table 4. Matrix confusion of the models

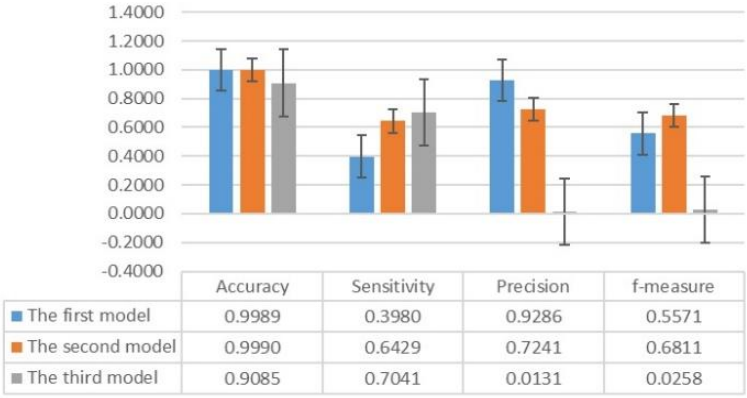
		Prediction					
		The first model		The second model		The third model	
		1	0	1	0	1	0
Actual	1	39	59	63	35	69	29
	0	3	56861	24	56840	5185	51679

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

A comparative analysis of the three SVM models clearly illustrates the significant influence of different data balancing strategies on fraud detection performance. The model trained on the raw dataset achieved a high precision score of 0.92, indicating that the majority of transactions identified as fraudulent were indeed true fraud cases. However, this model exhibited very low sensitivity (0.39), meaning that it failed to detect a substantial portion of actual fraudulent transactions. This imbalance between precision and sensitivity reflects a common challenge in imbalanced classification problems, where the model becomes biased toward the majority (non-fraudulent) class due to the overwhelming number of normal transactions. As a result, the raw dataset model demonstrated poor overall fraud detection capability and limited effectiveness for real-world applications where missing fraud cases can have severe financial consequences.

When the dataset was balanced using the Adaptive Synthetic Sampling (ADASYN) technique, a notable improvement in performance was observed. The ADASYN-balanced model achieved a significantly higher sensitivity of 0.64, meaning it became much more capable of identifying fraudulent transactions. Although precision decreased slightly to 0.72, the trade-off led to a marked improvement in the model's overall performance, as reflected by the highest f-measure (0.68) among all tested configurations. This outcome confirms that data-level balancing can effectively enhance a classifier's ability to recognize minority class patterns by providing a more representative and balanced training distribution.

In contrast, the support vector-based balancing approach produced the highest sensitivity of 0.70 but suffered from an extremely low precision value of 0.01, resulting in an overall f-measure of only 0.02. This outcome suggests that while the model became highly sensitive to detecting potential fraud, it also generated an excessive number of false positives, misclassifying many normal transactions as fraudulent. Such behavior indicates overfitting to minority samples and highlights the limitations of this technique in achieving balanced performance.



	Accuracy	Sensitivity	Precision	f-measure
The first model	0.9989	0.3980	0.9286	0.5571
The second model	0.9990	0.6429	0.7241	0.6811
The third model	0.9085	0.7041	0.0131	0.0258

Figure 6. Comparison of models performance

These findings collectively demonstrate that while data balancing techniques are essential for improving sensitivity in fraud detection, they also introduce inherent trade-offs that must be carefully managed. Among the three approaches, the ADASYN-balanced dataset provided the most effective compromise between detecting actual frauds and minimizing false alarms, thus yielding the most reliable and practical performance. Figure 6 illustrates a comparative visualization of the precision, sensitivity, and f-measure values across the three SVM models, emphasizing the superiority of the ADASYN-balanced approach in this study. The first model: model trained using raw dataset, the second model was trained using balanced dataset, the third model was trained using balanced support vector-based dataset

3.2 Discussion

To gain a deeper understanding of the effectiveness of the data balancing strategy, the results of this study were compared with previous research on classification under imbalanced datasets. Prior studies have consistently emphasized that addressing class imbalance is a critical step in improving the detection performance of minority classes, particularly in fraud detection problems where fraudulent transactions represent a very small proportion of the total data. The findings of this study are in line with existing literature, confirming the inherent trade-off between sensitivity and precision when different balancing techniques are applied.

Consistent with earlier research, the model trained on the raw dataset in this study demonstrated high precision (0.92) but low sensitivity (0.39), indicating that while the model was accurate in labeling detected fraud cases, it failed to identify many actual fraudulent transactions. Similar observations were reported by previous studies, such as [19], where machine learning models trained on highly imbalanced data achieved high precision but struggled to recognize minority instances due to the dominance of the majority class. This pattern reinforces the well-known challenge that classifiers naturally tend to bias toward the majority class when the data distribution is severely skewed.

The results obtained after applying ADASYN further support previous findings regarding the benefits of oversampling-based balancing techniques. In this study, the use of ADASYN led to a notable increase in sensitivity (0.64) with a slight reduction in precision (0.72), resulting in a higher f-measure (0.68) compared to the model trained

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

on raw data. Similar improvements have been observed in prior works [19] [20] [21] [22], which reported that resampling techniques such as oversampling, undersampling, and hybrid methods effectively improve the model's ability to detect minority classes, albeit at the expense of a minor decrease in precision. Moreover, several recent studies have indicated that both sensitivity and precision can be improved simultaneously through the adoption of ensemble learning approaches, such as Bagging and AdaBoost, which integrate multiple classifiers to enhance generalization and robustness [19].

Overall, the findings of this study reaffirm the critical importance of data balancing in the development of robust and effective fraud detection systems. In highly imbalanced datasets, such as those typically found in credit card transactions where fraudulent cases represent only a tiny fraction of the total data, conventional machine learning algorithms often struggle to recognize minority class patterns. The results clearly demonstrate that applying the Adaptive Synthetic Sampling (ADASYN) technique substantially enhanced the performance of the Support Vector Machine (SVM) model when compared to both the raw dataset and the support vector–based balanced dataset. This improvement underscores that appropriate data-level handling remains a fundamental step in optimizing classification outcomes for imbalanced data scenarios.

4. CONCLUSION

This study explored the influence of various data balancing strategies on the performance of Support Vector Machine (SVM) models for credit card fraud detection. Three types of datasets were employed: an imbalanced raw dataset, a dataset balanced using Adaptive Synthetic Sampling (ADASYN), and a support vector-based balanced dataset. The SVM model training on these datasets resulted in sensitivities of 0.39, 0.64, and 0.70, respectively, while the precision values were 0.92, 0.72, and 0.01. The corresponding f-measure values were 0.55, 0.68, and 0.02. The experimental results revealed that data imbalance has a substantial impact on model performance. The SVM model trained on the raw dataset achieved high precision but very low sensitivity, indicating that while the model was highly reliable in identifying legitimate transactions, it failed to detect a large proportion of fraudulent cases. This limitation underscores the inherent challenge of applying traditional SVM directly to highly imbalanced financial datasets. When the training dataset was balanced using ADASYN, the SVM model's sensitivity increased significantly, resulting in a more effective fraud detection capability. The ADASYN method succeeded in improving the model's ability to identify minority-class instances by synthetically generating realistic fraud samples based on data distribution. Although a slight reduction in precision occurred due to the introduction of some false positives, the overall performance, as reflected by the f-measure, improved notably—demonstrating a better trade-off between precision and recall. Further investigation using the support vector-based balancing approach produced even higher sensitivity but caused a substantial decline in precision, suggesting that the model overfitted to minority-class patterns. This extreme imbalance between sensitivity and precision renders the support vector-based method less practical for real-world deployment, where minimizing false alarms is as critical as detecting fraud itself. Overall, the ADASYN-balanced dataset provided the best compromise between sensitivity and precision, yielding the highest f-measure among the three experimental configurations. These findings affirm that addressing data imbalance is essential for developing robust and reliable fraud detection systems. Future research should focus on hybrid or adaptive strategies that integrate data-level balancing methods such as ADASYN with algorithm-level approaches like ensemble learning or cost-sensitive SVM. Such integration could further enhance both sensitivity and precision, leading to more accurate and dependable fraud detection in real-world financial environments.

ACKNOWLEDGMENT

The author wishes to extend sincere appreciation to IPB University for the facilities, technical resources, and academic support that have greatly contributed to the successful completion of this research. The author also gratefully acknowledges the continuous encouragement and understanding provided by family and colleagues, whose support has been instrumental throughout the research and manuscript preparation process.

REFERENCES

[1] Asosiasi Kartu Kredit Indonesia, "Credit Card Growth." [Online]. Available: https://www.akki.or.id/index.php/credit-card-growth. [Accessed: 31-Jan-2024].
[2] UK Finance, "Annual Fraud Report," 2022.
[3] R. K. L. Kennedy, Z. Salekshahrezaee, F. Villanustre, and T. M. Khoshgoftaar, "Iterative Cleaning and Learning of Big Highly-Imbalanced Fraud Data Using Unsupervised Learning," *J. Big Data*, vol. 10, no. 1, 2023. doi: 10.1186/s40537-023-00750-3.
[4] M. S. Kraiem, F. Sánchez-Hernández, and M. N. Moreno-García, "Selecting The Suitable Resampling Strategy for Imbalanced Data Classification Regarding Dataset Properties. An Approach Based On Association Models," *Appl. Sci.*, vol. 11, no. 18, 2021. doi: 10.3390/app11188546.
[5] Y. H. Liu and Y. T. Chen, "Total Margin Based Adaptive Fuzzy Support Vector Machines for Multiview Face Recognition," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2, pp. 1704–1711, 2005. doi: 10.1109/icsmc.2005.1571394.
[6] A. S. Desuky and S. Hussain, "An Improved Hybrid Approach for Handling Class Imbalance Problem," *Arab. J. Sci. Eng.*,

Building of Informatics, Technology and Science (BITS)
Volume 7, No 3, December 2025 Page: 1900−1909
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i3.8514

vol. 46, no. 4, pp. 3853–3864, 2021. doi: 10.1007/s13369-021-05347-7.
[7] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research.*, 2002. doi: https://doi.org/10.1613/jair.953.
[8] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning," *Proc. Int. Jt. Conf. Neural Networks*, no. 3, pp. 1322–1328, 2008. doi: 10.1109/IJCNN.2008.4633969.
[9] C. Li, N. Ding, Y. Zhai, and H. Dong, "Comparative Study on Credit Card Fraud Detection Based On Different Support Vector Machines," *Intell. Data Anal.*, vol. 25, no. 1, pp. 105–119, 2021. doi: 10.3233/IDA-195011.
[10] V. Chang, L. M. T. Doan, A. Di Stefano, Z. Sun, and G. Fortino, "Digital Payment Fraud Detection Methods In Digital Ages and Industry 4.0," *Comput. Electr. Eng.*, vol. 100, p. 107734, May 2022. doi: 10.1016/J.COMPELECENG.2022.107734.
[11] J. Chung and K. Lee, "Credit Card Fraud Detection: An Improved Strategy for High Recall Using KNN, LDA, and Linear Regression," *Sensors*, vol. 23, no. 18, 2023. doi: 10.3390/s23187788.
[12] L. S. Hasibuan and F. A. Jannah, "Deteksi Penipuan Kartu Kredit Menggunakan Support Vector Machine Dengan Optimasi Grid Search Dan Genetic Algorithm," *Building of Informatics, Technology and Science*, vol. 6, no. 1, pp. 344–353, 2023. doi: 10.47065/bits.v6i1.5355.
[13] N. G. Ramadhan, "Comparative Analysis of ADASYN-SVM and SMOTE-SVM Methods on The Detection of Type 2 Diabetes Mellitus," *Sci. J. Informatics*, vol. 8, no. 2, pp. 276–282, 2021. doi: 10.15294/sji.v8i2.32484.
[14] K. Shing Lim, L. Hong Lee, and Y.-W. Sim, "A Review of Machine Learning Algorithms for Fraud Detection in Credit Card Transaction," *Int. J. Comput. Sci. Netw. Secur.*, vol. 21, no. 9, pp. 31–40, 2021. doi: 10.22937/IJCSNS.2021.21.9.4.
[15] "Credit Card Fraud Detection." [Online]. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud/data. [Accessed: 16-Jun-2024].
[16] C. T. Chen, C. Lee, S. H. Huang, and W. C. Peng, "Credit Card Fraud Detection via Intelligent Sampling and Self-supervised Learning," *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 2, pp. 1–29, 2024. doi: 10.1145/3641283.
[17] D. Singh and B. Singh, "Investigating The Impact Of Data Normalization On Classification Performance," *Appl. Soft Comput.*, vol. 97, p. 105524, Dec. 2020. doi: 10.1016/J.ASOC.2019.105524.
[18] K. Hornik, A. Weingessel, F. Leisch, and M. D. M. Davidmeyerr-projectorg, *Package 'e1071.'* 2021.
[19] B. Juba and H. S. Le, "Precision-Recall Versus Accuracy And The Role Of Large Data Sets," *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 4039–4048, 2019. doi: 10.1609/aaai.v33i01.33014039.
[20] A. Balla, M. H. Habaebi, E. A. A. Elsheikh, M. R. Islam, and F. M. Suliman, "The Effect of Dataset Imbalance on the Performance of SCADA Intrusion Detection Systems," *Sensors*, vol. 23, no. 2, 2023. doi: 10.3390/s23020758.
[21] S. J. Yen and Y. S. Lee, "Cluster-based Under-sampling Approaches for Imbalanced Data Distributions," *Expert Syst. Appl.*, vol. 36, no. 3 PART 1, pp. 5718–5727, 2009. doi: 10.1016/j.eswa.2008.06.108.
[22] S. Bagui and K. Li, "Resampling imbalanced data for network intrusion detection datasets," *J. Big Data*, vol. 8, no. 1, 2021. doi: 10.1186/s40537-020-00390-x