E-ISSN: 3089-8838

JUTEKOM

Jurnal Teknologi dan Ilmu Komputer

https://jurnal.nolsatu.co.id/jutekom 2026 Vol. 2 No. 2 /Hal.82-88

Evaluating Resampling Methods for Imbalanced Necrosis Classification on CT Scans

Akhmad Rezki Purnajaya[1✉], Masparudin[2]

[1,2] Department of Software Engineering, Faculty of Computer Science, Universal University, Batam, Indonesia

rezkipurnajaya@gmail.com

Abstract

Necrosis, or body tissue death, occurs when there is insufficient blood flow to the tissue, which can be caused by injury, radiation, or chemicals. One of the main challenges in the automated diagnosis of necrosis is data imbalance in medical datasets, where the number of pathological cases is far less than normal cases. To address this issue, this study implements and evaluates various data sampling techniques, including Random Undersampling (RUS), Random Oversampling (ROS), Combination of Over-Undersampling (COUS), Synthetic Minority Over-sampling Technique (SMOTE), and Tomek Link, then using a Support Vector Machine (SVM) as the classifier. The test results show that the best sampling technique is the Synthetic Minority Over-sampling Technique (SMOTE), which successfully achieved an accuracy of 100% and an Area Under Curve (AUC) of 100%, indicating its significant potential in improving the accuracy of necrosis diagnosis from CT scans.

Keywords: *Necrosis, CT Scan Classification, Imbalanced Data, SMOTE, Support Vector Machine (SVM)*

Abstrak

Nekrosis, atau kematian jaringan tubuh, terjadi ketika aliran darah ke jaringan tidak mencukupi, yang dapat disebabkan oleh cedera, radiasi, atau bahan kimia. Salah satu tantangan utama dalam diagnosis nekrosis secara otomatis adalah ketidakseimbangan data (data imbalance) pada dataset medis, di mana jumlah kasus patologis jauh lebih sedikit daripada kasus normal. Untuk mengatasi masalah ini, penelitian ini menerapkan dan mengevaluasi berbagai teknik sampling data, termasuk *Random Undersampling* (RUS), *Random Oversampling* (ROS), *Combination of Over-Undersampling* (COUS), *Synthetic Minority Over-sampling Technique* (SMOTE), dan *Tomek Link*, kemudian menggunakan *Support Vector Machine* (SVM) sebagai metode pengklasifikasian. Hasil pengujian menunjukkan bahwa teknik sampling terbaik adalah *Synthetic Minority Over-sampling Technique* (SMOTE), yang berhasil mencapai akurasi 100% dan *Area Under Curve* (AUC) 100%, yang menunjukkan potensi signifikannya dalam meningkatkan akurasi diagnosis nekrosis dari CT scan.

Kata kunci: Nekrosis, Klasifikasi CT Scan, Data Imbalance, SMOTE, *Support Vector Machine* (SVM)

1. Introduction

Necrosis, defined as the death of body tissue due to insufficient blood flow, presents a formidable diagnostic challenge in clinical medicine. This condition can be triggered by various factors, including physical injury, radiation, or chemical exposure. In the context of neuroradiology, which is the primary focus of this study, specific forms such as liquefactive and caseous necrosis are of particular concern. Liquefactive necrosis is often associated with cellular destruction and the formation of purulent matter, typically resulting from bacterial or fungal infections, while caseous necrosis is a characteristic form of coagulative necrosis commonly caused by mycobacteria and other foreign agents. The timely and accurate identification of necrotic tissue is paramount, as it serves as a critical indicator of underlying pathology that dictates the urgency and nature of clinical intervention [1].

Despite its central role in diagnostics, the interpretation of necrosis from Computed Tomography (CT) scans is fraught with ambiguity and limitations. Although CT is a primary and widely available imaging modality for investigating suspected necrotic infections, its

Diterima: 27-03-2026 | Revisi: 06-05-2026 | Diterbitkan: 08-05-2026 | doi: 10.65258/jutekom.v2.i2.56

diagnostic performance is inconsistent. Radiologists rely on a series of visual cues, including the presence of gas in soft tissues, multiple fluid collections, the absence or heterogeneity of tissue enhancement with intravenous contrast, and significant inflammatory changes below the fascia. However, these indicators lack definitive reliability. For instance, the presence of soft-tissue gas is a highly specific marker for necrotizing fasciitis, yet its diagnostic utility is severely hampered by low sensitivity [2].

A systematic review combining data from multiple studies found a pooled sensitivity of only 48.6% for this criterion, indicating that gas is absent in more than half of confirmed cases, especially in the early stages of the disease when intervention is most critical. Other studies have reported similarly low sensitivities for various CT criteria, some as low as 39-43%, underscoring the inherent variability and uncertainty in visual interpretation. Furthermore, the morphological appearance of necrosis on CT can be highly variable, ranging from amorphous, cloud-like strands to well-defined, rim-enhancing masses, which can mimic other pathologies and mislead diagnosis. This diagnostic "weak spot" creates an urgent clinical need for automated and objective analytical tools that can augment radiological interpretation, reduce diagnostic delays, and improve patient outcomes by detecting subtle patterns that may be invisible to the human eye [2].

The development of such tools aligns with the current trajectory of artificial intelligence (AI) in neuroradiology. AI-driven systems are increasingly being integrated into clinical workflows to enhance diagnostic accuracy, accelerate imaging protocols, and improve overall efficiency. Current applications span the entire imaging cycle, from deep learning-based image reconstruction that improves image quality while reducing scan times, to computer-aided triage systems for acute conditions like intracranial hemorrhage, stroke, and spinal fractures. Research is also actively exploring AI for complex tasks such as volumetric analysis of brain tumors and the partial automation of radiology reports. The present study, which aims to develop an automated tool for the challenging task of necrosis detection, is therefore positioned at the forefront of this technological trend, addressing a well-documented limitation in current clinical practice where an AI-based solution could provide significant value [3].

While the clinical need for an automated necrosis detection tool is clear, its development faces a fundamental computational challenge that permeates all of medical AI: class imbalance. Medical imaging datasets are, by their nature, imbalanced. Pathological findings, which constitute the "minority class," are inherently less common in a given population than normal findings, which form the "majority class" [4]. This is particularly true for tasks involving the detection of specific or rare diseases, where the prevalence of the condition of interest may be very low [5].

Training a standard machine learning classifier on such imbalanced data leads to the development of a biased model. Because the learning algorithm is exposed to a vastly greater number of majority class examples, it optimizes its parameters to correctly classify these instances, often at the expense of the minority class [6]. This phenomenon can result in an "accuracy paradox," where a model exhibits a high overall classification accuracy that is profoundly misleading and clinically useless [4]. For example, a model trained on a dataset where only 1% of cases are positive for a disease could achieve 99% accuracy simply by learning to predict "negative" for every case, thus failing completely at its primary diagnostic purpose.

This issue is more than a technical inconvenience; it represents a core challenge to the safety, efficacy, and trustworthiness of clinical AI systems. Models that are not robust to class imbalance are prone to high false-negative rates for the cases requiring the most urgent clinical attention [7]. Such a failure mode would render an AI tool not just useless but potentially dangerous if integrated into a clinical workflow, as it could provide false reassurance and delay necessary treatment. Consequently, addressing class imbalance is not an optional preprocessing step but a mandatory prerequisite for the development of a clinically viable AI diagnostic tool. The systematic investigation of data sampling techniques, therefore, forms the scientific cornerstone of the present study, as it directly confronts this critical obstacle to building a reliable and effective model for necrosis detection.

2. Research Methodology

To obtain the results of this study, 5 data sampling techniques will be used for unbalanced data with the help of the RStudio program with the following stages, namely data collection, data initiation, dataset formation, inputting the SVM model, prediction and accuracy and will produce a performance table after data execution.

2.1. Data Collection

The dataset used in this study was sourced from a public repository on the Kaggle platform, titled "CT Head Scans for Necrosis Detection". It consists of a total of 52 grayscale CT scan images, each with a resolution of 225 x 225 pixels. The dataset is divided into two classes: a majority class of 30 non-necrotic (control) images and a minority class of 22 necrotic (pathological) images [8].

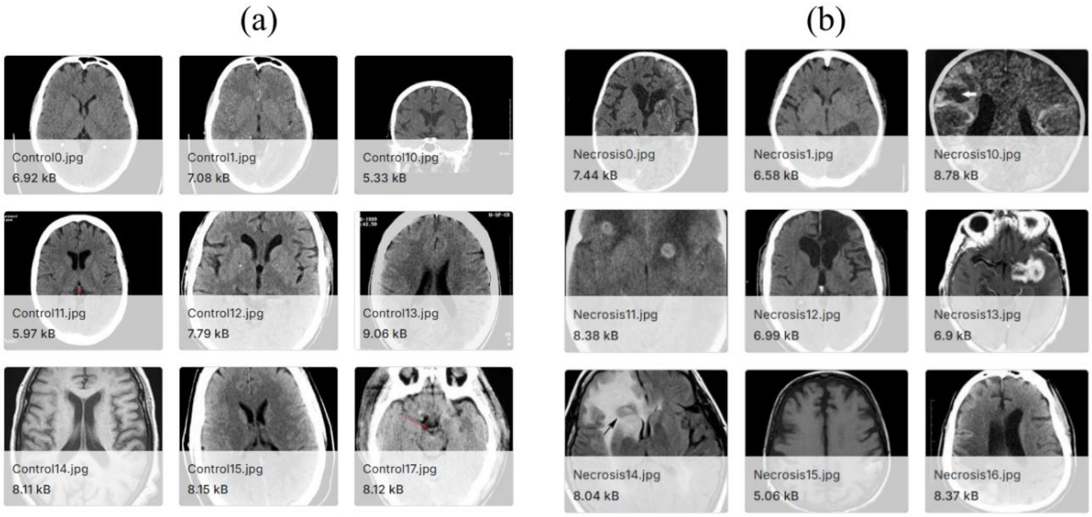

Figure 1. Representative samples of the CT head scan dataset: (a) Control group images (non-necrotic), and (b) Necrosis group images (pathological)

Visual representations of the dataset are presented in the Figure 1, where section (a) displays a sample of nine images from the control group, while section (b) illustrates nine examples from the necrosis group. These samples highlight the diversity of brain structures and the varying manifestations of necrotic tissue present within the dataset. The inclusion of these representative slices serves to demonstrate the visual complexity the model must navigate during the feature extraction and classification process. For the classification task, each image was reduced to 50 x 50 pixels flattened into a 2500-dimensional feature vector, which served as input to the machine learning model.

2.2. Data Sampling

To address the class imbalance inherent in the dataset, this study employed a data-level strategy by systematically evaluating several canonical resampling techniques. The fundamental goal was to modify the class distribution of the training data to mitigate the classifier's bias towards the majority class [4]. The evaluated techniques include:

- Random Undersampling (RUS): This technique balances class distribution by randomly removing instances from the majority class [8]. While it reduces computational load, it carries a significant risk of discarding valuable data, potentially leading to poor model generalization, especially with small datasets [6, 9]. This data sampling technique will be carried out by using the *ubUnder()* function from unbalanced package R.
- Random Oversampling (ROS): In contrast to RUS, ROS balances the dataset by randomly duplicating instances from the minority class [10]. This method avoids information loss but introduces a high risk of model overfitting, as the classifier may memorize specific duplicate samples rather than learning generalizable features [9]. This data sampling technique will be carried out by using the *ubOver()* function from unbalanced package R.
- Combination of Over-Under Sampling (COUS): This is a hybrid technique that combines the ROS and RUS algorithms. The combined algorithm works by simultaneously performing over-sampling on the minority class and under-sampling on the majority class until both classes have an equal number of samples. This approach seeks to find a balance, mitigating the risk of significant information loss from pure under-sampling while also

reducing the risk of overfitting associated with pure over-sampling [11]. This data sampling technique will be carried out by using the *ovun.sample()* function from ROSE package R.

- Synthetic Minority Oversampling Technique (SMOTE): As a widely recognized standard for handling imbalanced data, SMOTE generates new, synthetic minority class instances rather than duplicating existing ones [6]. It operates by selecting a minority instance, identifying its k-nearest minority neighbors, and creating a new sample along the line segments connecting them [12]. This approach effectively expands the decision region for the minority class, mitigating overfitting while avoiding the information loss associated with undersampling [4, 13]. This data sampling technique will be carried out by using the *ubSMOTE()* function from unbalanced package R.
- Tomek Links (T-Link): This is an advanced undersampling or data cleaning method. A Tomek Link consists of a pair of instances from different classes that are each other's nearest neighbors. The technique identifies these pairs and removes the majority class instance from the link, aiming to clarify the decision boundary between classes [10]. However, its performance can still degrade as the level of imbalance increases [9]. This data sampling technique will be carried out by using the *ubOver()* function from unbalanced package R.

2.3. SVM Classification

A Support Vector Machine (SVM) was selected as the classification algorithm. While Convolutional Neural Networks (CNNs) are dominant in image analysis, their effectiveness depends on large datasets to avoid severe overfitting [14, 15]. Given this study's small dataset (N=52) and high dimensionality (D=2500), an SVM presents a more strategic choice. SVMs are highly effective on small-to-medium-sized datasets and are robust in high-dimensional spaces, making them well-suited for "small N, large D" problems [15, 16].

The SVM learning mechanism, which defines a maximal margin hyperplane based on a small subset of critical data points (support vectors), is inherently resistant to the curse of dimensionality [4]. The combination of SMOTE and SVM is particularly synergistic. SVMs can produce suboptimal decision boundaries on imbalanced data [17], a weakness that SMOTE directly addresses by creating a more balanced and well-defined feature space. This preprocessing allows the SVM to identify a more robust and accurate hyperplane. The use of SVMs for brain CT image classification is also well-supported by previous research, which has demonstrated high classification accuracies with this approach [18]. The SVM model is created using the *svm()* function in the e1071 package. This function is used to train the SVM model with data. The arguments used in the *svm()* function are the class and dataset that have been processed based on each data sampling technique.

2.4. Performance Evaluation of Data Sampling

Predictions are made using the *predict()* function in R. The effectiveness of each data sampling technique in combination with the SVM classifier was quantitatively assessed using a set of standard performance metrics derived from the confusion matrix. The chosen metrics provide a comprehensive view of the model's performance beyond simple accuracy, which can be misleading in the context of imbalanced data. The evaluation metrics were [19]:

- Accuracy: The proportion of total predictions that were correct. The calculation is performed as shown in Equation (1).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{(Total\ Population)} \quad (1)$$

- Sensitivity (Recall): The proportion of actual positive cases (necrosis) that were correctly identified. This metric is crucial for evaluating a diagnostic tool's ability to detect the disease. This value is determined by the formula presented in (2).

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

- Specificity: The proportion of actual negative cases (non-necrotic) that were correctly identified. High specificity indicates a low rate of false alarms. It is computed according to the expression in (3).

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Fositives} \quad (3)$$

- Area Under the Curve (AUC): The AUC represents the area under the Receiver Operating Characteristic (ROC) curve. It provides a single scalar value summarizing the classifier's performance across all classification thresholds. An AUC of 1.0 indicates a perfect classifier, while an AUC of 0.5 suggests a performance no better than random chance. It is calculated as in (4).

$$AUC = \frac{Sensitivity + Specificity}{2} \quad (4)$$

3. Results and Discussion

Table 1 presents the achievement of 100% accuracy, sensitivity, specificity, and AUC using SMOTE-preprocessed data with an SVM classifier is a striking result that serves as the central finding of this study. This outcome is consistent with literature indicating that SMOTE can provide substantial, and at times dramatic, improvements in classifier performance on imbalanced datasets.

Table 1. Comparison of Imbalanced Data Performance and Data Sampling Techniques

Data Sampling Technique	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
Imbalanced	94.23	100	86.36	93.18
RUS	90.38	86.67	95.45	91.06
ROS	90.38	100	77.27	88.64
COUS	80.77	70.00	95.45	82.72
SMOTE	100	100	100	100
Tomek Link	90.38	89.67	95.45	91.06

The underlying mechanism for this success is SMOTE's ability to create larger and more coherent decision regions for the minority class. By generating synthetic samples within the convex hull of existing minority instances, SMOTE provides the classifier a more robust and stable target to learn from, effectively reducing the inherent bias toward the majority class. The perfect scores achieved suggest that for the specific feature space defined by the 50x50 pixel images in the Kaggle dataset, the necrosis samples likely form a relatively distinct and well-separated cluster. In such an "ideal" scenario, SMOTE's linear interpolation would be highly effective, generating synthetic samples that cleanly and unambiguously reinforce the decision boundary for the SVM without introducing noise or class overlap.

On the other hand, the table clearly shows that the Combination of Over-Undersampling (COUS) technique performed the worst among all methods tested. This technique yielded the lowest accuracy (80.77%) and the lowest AUC (82.72%). Even more concerning is its sensitivity value, which plummeted to 70.00%. This is a critical metric in medical diagnosis, as such a low value means the model failed to identify 30% of the actual (positive) necrosis cases. This result indicates that the combined oversampling and undersampling strategy in this case likely undermined the SVM's learning process. This could be due to the undersampling process removing too many informative minority samples, or an ineffective oversampling process, thereby creating a blurred and confusing decision boundary for the classifier. The results are summarized in Figure 2.

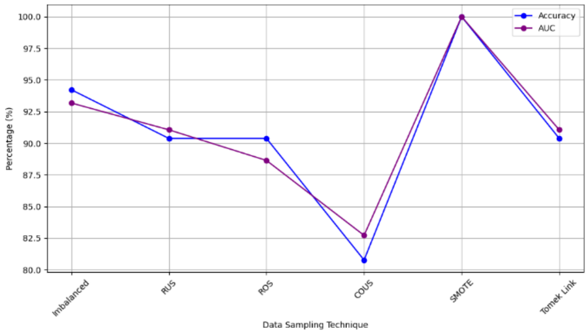

Figure 2. Comparison of Data Sampling Techniques Performance

One primary concern is SMOTE's behavior in high-dimensional spaces. Several rigorous evaluations have shown that SMOTE's effectiveness can diminish as the dimensionality of the data increases. The concept of "nearest neighbor," which is central to the SMOTE algorithm, becomes less meaningful in high-dimensional feature spaces due to the "curse of dimensionality," where all points tend to become equidistant from one another. In some high-dimensional contexts, such as microarray gene expression data, SMOTE has been found to be less effective than even simple random undersampling.

Another critical drawback is that the standard SMOTE algorithm generates synthetic samples without considering the distribution of the majority class. If the minority and majority classes have a high degree of overlap, SMOTE can inadvertently generate synthetic minority samples within dense regions of majority class instances. This introduces noise, blurs the decision boundary, and can actively harm classifier performance. The perfect result in this study suggests that significant class overlap was likely not a major issue in the chosen dataset, a critical point for discussion.

Furthermore, theoretical analyses have revealed that SMOTE can have unintended consequences on the statistical properties of the data in high-dimensional settings. It has been shown that SMOTE artificially reduces the minority class variance to approximately two-thirds of its original value leaving the mean unchanged. This underestimation of the true class variability can lead to a form of overfitting, where the classifier learns an oversimplified representation of the minority class that does not capture its full diversity. It can also lead to overgeneralization, where the model learns patterns from the synthetic data that do not accurately reflect the real-world distribution of the pathology.

Finally, the study can be framed appropriately: it successfully validates a promising computational pipeline (SMOTE + SVM) on a benchmark dataset, but the next critical phase of research must focus on bridging the "lab-to-clinic" gap. This demonstrates foresight and a deep understanding of the translational challenges in medical AI. The conclusion of the manuscript should therefore propose that future work must involve a rigorous validation of the proposed method on larger, multi-institutional, clinically-sourced datasets that reflect the true heterogeneity and complexity of the target patient population.

4. Conclusion

Based on the testing and analysis results, it can be concluded that the use of data sampling methods can effectively handle the problem of class imbalance in the task of necrosis classification on CT scans. The comparison between the imbalanced, RUS, ROS, COUS, SMOTE, and T-Link techniques showed significant performance variation, with respective accuracy results of (94%, 90%, 92%, 80%, 100%, 90%) and AUCs of (92.85%, 90.72%, 90.47%, 82.10%, 100%, 90.72%). Notably, the implementation of SMOTE yielded a substantial improvement over the baseline imbalanced model, increasing accuracy by 6% and the AUC by 7.15%. Of all the methods tested, SMOTE demonstrated the most superior and balanced performance, achieving 100% accuracy and 100% AUC.

This result underscores the great potential of SMOTE in creating a more representative training dataset, specifically by enhancing the model's ability to distinguish subtle pathological features within CT scan slices that characterize necrotic tissue. This, in turn, allows the SVM model to learn a more robust and accurate decision boundary. While these results are very promising, it is important to acknowledge that this performance was achieved on a public dataset that may not fully reflect the complexity and variability of real-world clinical data. Therefore, a crucial next step is to validate this approach on larger and more diverse datasets from various institutions to ensure its generalization and reliability before it can be considered for clinical application.

References

[1] W. Tonnus, C. Meyer, A. Paliege, A. Belavgeni, A. von Mässenhausen, S. R. Bornstein, dan A. Linkermann, "The pathological features of regulated necrosis," The Journal of Pathology, vol. 247, no. 5, hlm. 697–707, 2019.

[2] R. M. Kwee dan T. C. Kwee, "Diagnostic performance of MRI and CT in diagnosing necrotizing soft tissue infection: a systematic review," Skeletal Radiology, vol. 51, no. 4, hlm. 727–736, 2022.

[3] Y. W. Lui, et al., "Artificial intelligence in neuroradiology: current status and future directions," American Journal of Neuroradiology, vol. 41, no. 8, hlm. E52–E59, 2020.

[4] A. Fernández, S. Garcia, M. Galar, R. C. Prati, B. Krawczyk, dan F. Herrera, Learning from Imbalanced Data Sets. Cham: Springer International Publishing, 2018.

[5] R. Ghasemi, et al., "Detection and classification of brain tumor using a hybrid learning model in CT scan images," Scientific Reports, vol. 15, no. 1, hlm. 35085, 2025.

[6] A. Fernández, S. Garcia, F. Herrera, dan N. V. Chawla, "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," Journal of Artificial Intelligence Research, vol. 61, hlm. 863–905, 2018.

[7] M. Naseriparsa, A. Al-Shammari, M. Sheng, Y. Zhang, dan R. Zhou, "RSMOTE: improving classification performance over imbalanced medical datasets," Health Information Science and Systems, vol. 8, no. 1, hlm. 22, 2020.

[8] C. Saben, "CT Head Scans (jpg files)," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/clarksaben/ct-head-scans. [Accessed: July 1, 2025].

[9] D. Singh dan A. Sharma, "A Comparative Analysis of Oversampling and Undersampling Techniques for Image Data Classification Across Varying Imbalance Levels," dalam International Conference on Data Science and Applications, Singapore: Springer Nature Singapore, 2024, hlm. 371–385.

[10] T. Elhassan dan M. Aljurf, "Classification of imbalance data using tomek link (t-link) combined with random under-sampling (rus) as a data reduction method," Global J Technol Optim S, vol. 1, no. S1, 2016.

[11] A. R. Purnajaya dan F. D. Hanggara, "Using Data Sampling Technique for Improving Classification of Covid-19 and Lung Diseases," dalam The 2nd Universitas Kuningan International Conference on System, Engineering, and Technology (UNISET 2021), EAI, 2022, hlm. 167–175.

[12] G. A. Pradipta, R. Wardoyo, A. Musdholifah, I. N. H. Sanjaya, dan M. Ismail, "SMOTE for handling imbalanced data problem: A review," dalam 2021 Sixth International Conference on Informatics and Computing (ICIC), IEEE, 2021, hlm. 1–8.

[13] R. S. Abdulsadig dan E. Rodriguez-Villegas, "A comparative study in class imbalance mitigation when working with physiological signals," Frontiers in Digital Health, vol. 6, hlm. 1377165, 2024.

[14] H. Hasan, H. Z. Shafri, dan M. Habshi, "A comparison between support vector machine (SVM) and convolutional neural network (CNN) models for

hyperspectral image classification," dalam IOP Conference Series: Earth and Environmental Science, vol. 357, no. 1, IOP Publishing, 2019, hlm. 012035.

[15] N. A. Almansour, et al., "Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study," Computers in Biology and Medicine, vol. 109, hlm. 101–111, 2019.

[16] V. Kecman, "Support vector machines–an introduction," dalam Support vector machines: theory and applications, Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, hlm. 1–47.

[17] Afif, A. (2023). Analisis Metode Klasifikasi Data Naïve Bayes dan SVM Dalam Menentukan Keunikan Hotel. Jurnal Teknologi Komputer Dan Informasi, 11(1), 38–46. https://doi.org/10.52072/jutekinf.v11i1.526.

[18] I. A. Illan, et al., "Support vector machine failure in imbalanced datasets," dalam International Work-Conference on the Interplay Between Natural and Artificial Computation, Cham: Springer International Publishing, 2019.

[19] A. Padma dan R. Sukanesh, "SVM based classification of soft tissues in brain CT images using wavelet based dominant gray level run length texture features," Middle-East Journal of Scientific Research, vol. 13, no. 7, hlm. 883–888, 2013.

[20] A. R. Purnajaya dan F. D. Hanggara, "Perbandingan Performa Teknik Sampling Data untuk Klasifikasi Pasien Terinfeksi Covid-19 Menggunakan Rontgen Dada," JAIC, vol. 5, no. 1, hlm. 37–42, Jun. 2021.