

## Research Article

# Machine Learning Approaches to Acute Kidney Injury Prediction: Addressing the Class Imbalance Challenge

Evi Diana Omar<sup>1\*</sup>, Hasnah Mat<sup>1</sup>, Ainil Zafirah Abd Karim<sup>1</sup>, Zam-tira Seman<sup>1</sup>, Nabilah Hanis Zainuddin<sup>1</sup>, Anis-Syakira Jailani<sup>2</sup>, Ridwan Sanaudi<sup>1</sup>, Munerah Mohamad Zaman<sup>1</sup>, Najjah Tohar<sup>1</sup>, Yusrina M Yusoff<sup>1</sup>, Nik Noor Syamimi Ismail<sup>1</sup>, Asyraf Syahmi Mohd Noor<sup>1</sup>, Bak Leong Goh<sup>3</sup>, Mohd Azahadi Omar<sup>1</sup> and Muhd Zulfadli Hafiz Ismail<sup>1</sup>

<sup>1</sup>Sector for Biostatistics & Data Repository, National Institutes of Health, Ministry of Health Malaysia, Shah Alam, Selangor, Malaysia

<sup>2</sup>Research Grant Unit, National Institutes of Health, Ministry of Health Malaysia, Shah Alam, Selangor, Malaysia

<sup>3</sup>Hospital Serdang, Ministry of Health, Malaysia

## Abstract

### Introduction

Acute kidney injury (AKI) requiring dialysis after cardiac surgery is a critical condition associated with increased mortality and morbidity. Early prediction of AKI is essential for timely intervention and improved patient outcomes. However, the development of accurate predictive models is challenged by class imbalance in datasets, where non-AKI cases significantly outnumber AKI cases.

### Methods

This study explores various machine learning approaches to address this challenge, comparing nine algorithms: Logistic Regression, Gradient Boosted Trees, Support Vector Machines, Random Forest, Naïve Bayes, Generalized Linear Model, Fast Large Margin, Deep Learning and Decision Tree. Using clinical data from cardiac surgery patients in Malaysia (2011-2015), we evaluated model performance across three data configurations: original imbalanced

dataset (179 AKI vs. 1,562 non-AKI cases), random down sampling (134 cases per class) and Synthetic Minority Over-sampling Technique (SMOTE, 1,250 cases per class).

### Results

Results demonstrated that while models trained on imbalanced data achieved high overall accuracy (85-90%), they exhibited poor sensitivity (40-50%) for detecting AKI cases. Both balancing techniques significantly improved performance, with SMOTE showing the most substantial enhancements: sensitivity increased to 75-85%, precision improved to 80-85% and ROC-AUC scores reached 85-90%. Feature importance analysis revealed that balancing techniques uncovered more nuanced predictive factors, including creatinine levels, length of stay, insulin use and hypertension status.

### Conclusion

This study demonstrates that addressing class imbalance is crucial for developing reliable clinical decision-making tools for AKI prediction and provides a framework applicable to other rare but serious medical conditions.

**Keywords:** Acute Kidney Injury; Cardiac Surgery; Class Imbalance; Clinical Decision Support; Down Sampling; Feature Importance; Healthcare Analytics; Machine Learning; Predictive Modeling; SMOTE

## Introduction

Acute Kidney Injury (AKI) is a critical medical condition characterized by a sudden decline in kidney function, leading to increased serum creatinine levels or reduced urine output. This condition subsequently causes dysregulation of the body electrolytes and fluid balance, as well as abnormal retention of nitrogenous waste [1]. Paweena Susantitaphong et al. [2], conducted a meta-analysis of 154 studies examining the global incidence of AKI in adults and children. The pooled incidence rates of AKI were found to be 21.6% in adults and 33.7% in children. Meanwhile, the pooled AKI-associated mortality rates were 23.9% in adults and 13.8% in children. In addition, AKI can lead to long-term complications including an increased risk of developing Chronic Kidney Disease (CKD) with an adjusted Hazard Ratio (HR) reaching as high as 8.82 in certain populations.

Furthermore, cardiac surgery is reported to be one of the significant factors developing AKI. This is attributed to several factors such as hemodynamic, inflammatory and nephrotoxic factors [3]. These factors lead to renal vasoconstriction and ischemia, ultimately resulting in AKI. Various studies have highlighted the notably high prevalence of Acute Kidney Injury (AKI) following cardiac surgery, is. Up to 30% of patients who undergo cardiac surgery developed acute kidney injury with approximately 1% required dialysis. According to Cardinale D et al. the incidence rate of AKI following cardiac surgery ranges from 6% to 20%, with mortality rates between 3.7% to 34% [4]. Furthermore, Acute Kidney Injury (AKI) also complicates recovery from cardiac surgery in up to 30% of patients, injuries and impairs the function of the brain, lungs, gut and places patients at a 5-fold increased risk of death during hospitalization [5]. Conclusively, AKI

**\*Corresponding author:** Evi Diana Omar, Sector for Biostatistics & Data Repository National Institutes of Health (NIH), Jalan Setia Murni U13/52, Seksyen U13 Setia Alam, 40170 Shah Alam, Selangor, Malaysia, E-mail: evidiana@crc.moh.gov.my

**Citation:** Omar ED, Mat H, Karim AZA, Seman Z, Zainuddin NH, et al. (2025) Machine Learning Approaches to Acute Kidney Injury Prediction: Addressing the Class Imbalance Challenge. J Nephrol Renal Ther 11: 104.

**Received:** October 06, 2025; **Accepted:** October 14, 2025; **Published:** October 22, 2025

**Copyright:** © 2025 Omar ED, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

has a profound impact on patient morbidity and mortality. Therefore, early detection and timely intervention are essential to prevent the progression of AKI to more severe stages, reduce complications and improve overall patient outcomes including a lower mortality rate. Acute kidney injury can be predicted and managed using biomarkers, machine learning and fluid management [6]. Accurate prediction is crucial in healthcare as it helps to identify at-risk patients early and alleviates the burden on healthcare systems. The Cleveland Clinic Score has been widely used as a predictive tool to estimate the risk of Acute Kidney Injury (AKI) requiring dialysis after cardiac surgery. It incorporates multiple clinical variables, such as patient demographics, comorbidities and surgical details, to generate a risk score. This tool is adapted from the 2005 Thakar model, which is considered to be the best-validated and most predictive tool [7]. Despite its effectiveness, this model's reliance on conventional statistical approaches may limit its applicability in diverse populations or datasets with imbalanced cases—a common challenge in AKI datasets where non-AKI cases predominate. Furthermore, these conventional methods for monitoring renal function often fall short as they depend on late-stage indicators, potentially delaying necessary treatment and worsening patient outcomes.

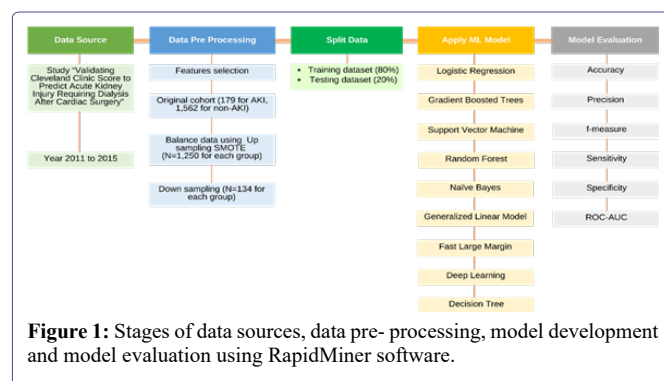
Recent advances in Machine Learning (ML) have emerged as powerful tools to improve the predictive accuracy of AKI outcomes. These advanced algorithms can analyse clinical variables with the ability of analysing and processing huge datasets in real time, enabling earlier identification of at-risk patients and more precise predictions of AKI. However, a persistent challenge in developing effective ML models for effective AKI prediction is data imbalance, which often results in biased model performance and reduced generalizability. Since the incidence rate of AKI is typically low, this creates a significant dataset class imbalance challenge for predictive modelling. This imbalance poses substantial methodological obstacles in developing accurate machine learning approaches. Researchers have explored multiple strategies to address the data imbalance challenges [8]. Machine learning techniques for handling data imbalance have evolved from traditional resampling methods to more sophisticated approaches. Early strategies involved oversampling minority classes and under-sampling majority classes, with hybrid techniques combining both methods [9]. SMOTE (Synthetic Minority Over-sampling Technique) innovatively creates artificial minority class samples, expanding limited dataset representation through synthetic data generation. Ensemble learning strategies like boosting and bagging enhance model performance by dynamically adjusting weights and creating balanced subsets [10-12]. Despite their effectiveness, these methods are often perceived as “black-boxed,” presenting challenges in understanding their intricate decision-making processes. Deep learning techniques have emerged as a powerful solution, offering advanced feature extraction and improved sensitivity for predicting rare events. These approaches demonstrate significant potential in managing complex, imbalanced datasets, particularly in domains like Acute Kidney Injury (AKI) prediction [13-14]. While the techniques mentioned have demonstrated potential benefits, significant gaps remain in understanding their effectiveness in predicting AKI. This study aims to explore various ML approaches for predicting AKI outcomes, particularly focusing on addressing the challenges posed by imbalanced AKI datasets. Imbalanced datasets often result in biased model performance and reduced generalizability in predictive algorithms. To address this, the study employs the Synthetic Minority Over-sampling Technique (SMOTE), an innovative method that

generates artificial samples for minority classes, thereby improving dataset representation through synthetic data generation. We evaluate and compare the performance of various ML using nine machine learning algorithms; Logistic Regression, Gradient Boosted Trees, Support Vector Machines, Random Forest, Naïve Bayes, Generalized Linear Model, Fast Large Margin, Deep Learning and Decision Tree in predicting AKI requiring dialysis. Through this analysis, we utilize diverse clinical datasets to enhance the generalizability of the models and provide actionable insights for healthcare decision-making. By addressing these aspects, this study strives to refine predictive accuracy, facilitate informed clinical decisions and ultimately contribute to improved clinical decision-making and patient care. While all the techniques mentioned have shown potential benefits, significant gaps remain in understanding their effectiveness for AKI prediction.

This study aims to explore various ML approaches for predicting AKI outcomes with a specific focus on addressing the challenges posed by an imbalanced AKI dataset. We evaluate and compare the performance of various ML using nine machine learning algorithms; Logistic Regression, Gradient Boosted Trees, Support Vector Machines, Random Forest, Naïve Bayes, Generalized Linear Model, Fast Large Margin, Deep Learning and Decision Tree in predicting AKI requiring dialysis. Through this analysis, we utilize diverse clinical datasets to enhance the generalizability of the models and provide actionable insights for healthcare decision-making. By addressing these aspects, this study strives to refine predictive accuracy, facilitate informed clinical decisions and ultimately contribute to improved clinical decision-making and patient care.

## Methods

A rigorous methodology is employed, encompassing several key stages: data preprocessing, model training, evaluation and comparative analysis. Figure 1 shows the overall workflow of the study.



## Data source

Data was collected based on the Cleveland Clinic Score, which includes key variables relevant to AKI prediction. This study is grounded in the research titled “Validating Cleveland Clinic Score to Predict Acute Kidney Injury Requiring Dialysis After Cardiac Surgery [15].” The research involved comprehensive national-level data collection from patients who underwent cardiac surgery at a tertiary cardiothoracic center in Malaysia between 2011 and 2015. The methodology for this data collection process has been detailed in a prior publication [19]. The primary data source for this investigation was the Electronic Health Records (EHR), which contained extensive patient information, including clinical, surgical laboratory data. This

dataset includes clinical and demographic variables relevant to AKI prediction, with an original cohort comprising 179 patients with AKI and 1,562 patients without AKI. The Cleveland Clinic Score, a widely recognized predictive model for AKI requiring dialysis, served as the foundation for selecting the variables included in this study. Key variables utilized in the analysis encompassed gender, insulin use, history of Chronic Obstructive Pulmonary Disease (COPD), prior surgeries, Left Ventricular Ejection Fraction (LVEF), Intra-Aortic Balloon Pump (IABP) use, Congestive Heart Failure (CHF) diagnosis, type of surgery, emergency status, creatinine levels and the need for renal replacement therapy as the primary endpoint. Additionally, we incorporated other variables such as age, hypertension status, discharge status and length of stay due to their potential significance in predicting patient outcomes.

## Data preprocessing

In this study, the raw data was carefully explored and prepared based on its characteristics and structure. Initial steps included examining the dataset to identify any inconsistencies, errors, or outliers that could affect the analysis. The extent and pattern of missing values also were assessed and appropriate strategies, such as imputation and exclusion, were applied to minimize bias while preserving data integrity. This process encompasses both feature selection and data balancing techniques as essential components to prepare the dataset for effective machine learning model training.

## Feature selection

Identifying the most important factors for predicting AKI requiring dialysis requires feature selection to be applied in our method. The process is crucial for enhancing model efficiency by including only the most important predictive variables, identified through an analysis of their individual contributions and clear patterns in the data. Several key aspects were considered when selecting these factors. As correlation was assessed to determine how closely each factor relates to the outcome being predicted, uniqueness was somehow assessed to see how distinct each factor is from others. Stability was also important that we evaluated how consistent each factor remained across different groups of data. Additionally, the amount of missing information for each factor was taken into account, along with any factors that contained free-text data.

## Data balancing techniques

To address class imbalance, we apply the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic examples of AKI cases to balance the training dataset. This technique aims to improve model performance by ensuring that minority class instances are adequately represented during training. In this study, we implemented a combined approach to address class imbalance by utilizing both upsampling with SMOTE and downsampling. SMOTE generates synthetic samples of the minority class by interpolating between existing minority class samples. This approach yielded a balanced dataset with 1,250 samples in each class, enabling the models to generalize better without overfitting to the minority class. Meanwhile, downsampling reduces the majority class size to match the minority class, resulting in 134 samples per class. By employing these methods, this study aimed to evaluate the impact of class balancing techniques on model performance and identify the most effective approach for predicting acute kidney injury requiring dialysis.

## Data splitting

To ensure an unbiased evaluation of the machine learning models, the dataset was divided into training and testing sets. An 80:20 split was applied, where 80% of the data was used for training the models and the remaining 20% was reserved for testing. The choice of an 80:20 split was made to maximize the amount of data available for training while maintaining a reasonable testing set size to reduce the risk of overfitting. This method of splitting ensures that the evaluation of the models reflects their real-world performance, providing confidence in their predictive ability when applied to unseen data. The hyperparameters were automatically optimized using RapidMiner's AutoModel feature during the training phase to enhance the performance of each algorithm. Hyperparameter tuning is performed using grid search to optimize performance metrics such as accuracy, precision, recall Area Under The Receiver Operating Characteristic Curve (AUC-ROC).

## Analysis

In this study, we implement nine machine learning algorithms to evaluate their predictive capabilities for AKI namely Logistic Regression, Gradient Boosted Trees, Support Vector Machine (SVM), Random Forest, Naive Bayes, Generalized Linear Model (GLM), Fast Large Margin, Deep Learning and Decision Tree. The selection of these algorithms was based on several important factors relevant to the nature of the problem and the characteristics of the data. These algorithms represent a diverse range of methodologies, from simple linear models like Logistic Regression to complex non-linear models such as Deep Learning. This diversity allows for a comprehensive evaluation of different approaches to determine which model performs best in predicting acute kidney injury (AKI) outcomes. All selected algorithms are well-suited for binary classification tasks, which is essential for this study since the goal is to predict whether a patient will develop AKI requiring dialysis. Algorithms like Logistic Regression and SVM are particularly effective in binary classification scenarios due to their ability to model probabilities and decision boundaries effectively. Additionally, some algorithms offer high interpretability, such as Decision Trees and Logistic Regression, making it easier to understand how predictions are made. This interpretability is crucial in healthcare settings where understanding model decisions can impact clinical practices. Given that AKI requiring dialysis is a relatively rare outcome compared to non-AKI cases, algorithms like Random Forest and Gradient Boosted Trees can handle class imbalance effectively through techniques such as ensemble learning and boosting. The inclusion of various algorithms also allows for hyperparameter tuning and model optimization, significantly enhancing predictive performance. Techniques like grid search can be applied across these diverse models to find optimal parameters that improve metrics such as accuracy, precision, recall and area under the receiver operating characteristic curve (AUC-ROC). Furthermore, the selected models vary in their assumptions about data distribution and relationships among features. For instance, Naive Bayes assumes independence among predictors, while Random Forest can manage correlated features effectively. This robustness allows for better adaptability to different characteristics within the dataset.

## Model evaluation and comparative analysis

The evaluation phase tests our models on a separate validation dataset to assess their predictive capabilities in real-world scenarios. To extend these metrics, we also focus on other relevant indicators of



model performance, including the harmonic mean of precision and recall, true positive rate and true negative rate.

Feature Importance

In this study, we used correlation-based feature selection to determine the importance of different features. A higher correlation indicates that an attribute is more relevant to the prediction task. This technique is suitable for data with numerical or binomial labels. Correlation values range from -1 to +1 and measure how two attributes are related. A positive correlation means that larger values of one attribute are linked to larger values of another, while smaller values are linked to smaller values. On the other hand, a negative correlation indicates that larger values of one attribute are associated with smaller values of another. Correlation is calculated by summing the products of deviations from the mean for each attribute and normalizing this by the product of their standard deviations. This study adheres to ethical guidelines for data use and patient privacy. Patient identifiers were removed and data were anonymized to ensure confidentiality. The Medical Research and Ethics Committee, Ministry of Health, Malaysia approved this study.

Results

The study’s evaluation of nine machine learning algorithms revealed remarkable variations in performance across the three data configurations of balance and imbalance (Figure 2).

For an imbalanced dataset, the accuracy ranged approximately between 85-90%. The sensitivity was hovering around 40-50% to detect minority class (AKI patients). Conversely, the specificity was extremely high, reaching around 95-98%. The precision remained moderate, ranging between 60-70%, while the ROC-AUC score settled at approximately 75-80%. The random down-sampling approach yielded notable improvements in model performance. The overall accuracy is slightly lower at 80-85%. The sensitivity improved substantially, increasing to 60-70%, which indicated a better ability to identify AKI cases. The specificity experienced a modest decline to 85-90%, while the precision showed an encouraging increase to 75-80%. The ROC-AUC score improved to 82-85%, demonstrating enhanced discriminative capabilities. The Synthetic Minority Over-Sampling Technique (SMOTE) demonstrated the most significant improvements in model performance. The accuracy stabilized around 82- 87%. The sensitivity experienced a substantial increase to 75-85% in detecting AKI cases. The specificity was maintained at 80-90%, while the precision significantly improved to 80-85%. The ROC-AUC score was enhanced to 85-90%, indicating a notable improvement in the model’s discriminative power.

The feature importance analysis provided critical insights into the predictive modelling of Acute Kidney Injury (AKI), revealing profound differences between the imbalanced and balanced datasets (figure 3). The balancing techniques (both downsampling and SMOTE) reveal more nuanced insights, giving greater importance to attributes like creatinine levels, length of stay, insulin use and hypertension status. Specifically, insulin use and hypertension status show a significant increase in weight, indicating their stronger role in AKI prediction when both classes are equally represented.

Discussion

This study provides a critical examination of machine learning approaches to Acute Kidney Injury (AKI) prediction, addressing the

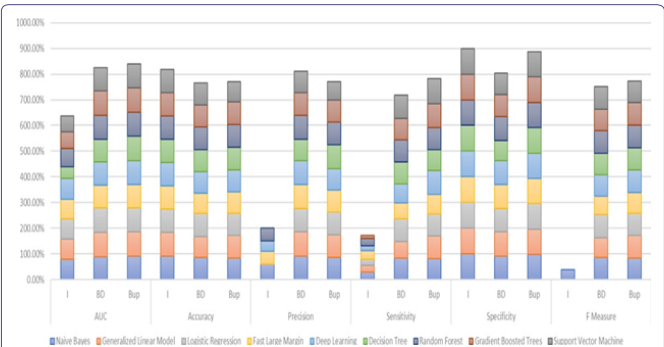


Figure 2: Performance of machine-learning algorithms for AKI prediction between Imbalance and Balanced Data.

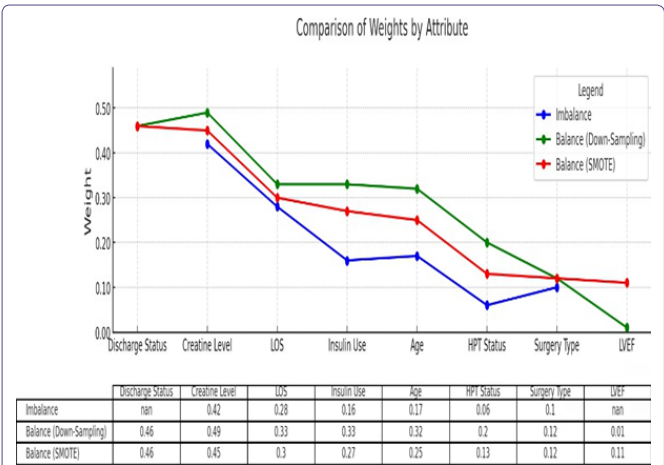


Figure 3: Feature Importance for AKI prediction between Imbalance and Balanced Data.

challenge of class imbalance in clinical datasets. Our findings underscore the significant impact of data balancing techniques on predictive model performance. The most striking observation from our analysis is the improvement in model performance when addressing class imbalance. In the original imbalanced dataset, models demonstrated a concerning limitation: while maintaining a high overall accuracy (85-90%), they exhibited poor sensitivity (40-50%) for detecting AKI cases. This demonstrates the effect of untreated imbalance issues, which lead to higher misclassification of critical AKI events due to the disproportionately fewer positive examples of AKI cases during training. This highlights a critical weakness in traditional machine learning approaches when applied to medical prediction tasks with uneven class distributions [16-18]. Both down-sampling and Synthetic Minority Over-sampling Technique (SMOTE) techniques showed substantial improvements as compared to an imbalanced dataset. By exposing the model to a more representative AKI case, the minority-class detection was then enhanced. As such, the performance of the models offers a more reliable representation of how the model will perform in a real-world clinical situation. Comparing both down-sampling and SMOTE, SMOTE emerged as the more promising approach. The sensitivity increased dramatically to 75-85%, a clinically significant enhancement that directly translates to improved early detection capabilities [19]. The Receiver Operating Characteristic Area Under the Curve (ROC-AUC) score improvement from 0.75-0.80 to 0.85-0.90 further validates the effectiveness of this balancing technique. Plausible explanation to the superior performance by SMOTE compared

to undersampling is the fact that SMOTE enhances the minority class by synthesizing plausible new data points, as opposed to discarding a portion of majority class data, leading to risk of information loss in undersampling. This leads to a better model learning, as there is more comprehensive representation of AKI cases, enhancing the discriminative ability [20,21]. The feature importance analysis revealed a profound shift in predictive modelling when addressing class imbalance. In the original imbalanced dataset, prediction relied heavily on a few dominant clinical markers. After implementing balancing, we observed several key changes: The model's dependence shifted from relying primarily on prevalent indicators in non-AKI cases (the majority class) to incorporating a more comprehensive range of predictive factors. This shift highlights two important scientific insights. Firstly, it highlights that class imbalance can significantly distort which features appear predictive in standard analyses, leading to incomplete or skewed clinical inferences. In the imbalanced dataset, predictions largely hinged on a small subset of dominant clinical variables, potentially overshadowing critical but underrepresented signals that are more common among AKI patients. Secondly, it demonstrates that the balancing of minority-class examples allows the algorithm to more robustly capture the multifactorial nature of AKI, thereby unmasking additional risk signals. Models gain a broader view of patient characteristics, enabling them to unearth nuanced risk factors that might otherwise remain undetected. In essence, this transformation enabled a more holistic approach to risk assessment, rather than focusing exclusively on dominant indicators. Recent literature supports these insights, particularly in the context of rare medical event detection [22,23].

Our findings have significant implications for both AKI prediction and broader clinical care. The balanced models demonstrate superior early detection capabilities for AKI, potentially enabling more timely medical interventions and improved patient outcomes [24]. Furthermore, these models provide deeper insights into AKI risk factors compared to traditional approaches, facilitating more comprehensive patient risk assessment. This is due to the fact that by counteracting the bias introduced by skewed representation of non-AKI examples, balancing expands the algorithm's exposure to the more diverse presentations of AKI. This leads to a better understanding of the interplay among comorbidities, treatment factors and patient outcomes, reflecting the real-world clinical situations more accurately. The methodology developed in this study extends beyond AKI prediction and could be applied to improve predictive modelling for other rare but serious medical conditions. Our results strongly suggest that addressing class imbalance is crucial for developing reliable clinical decision-making tools, particularly when dealing with conditions that occur infrequently but carry significant health risks [24].

## Limitations

Our study has three key limitations. First, since we only used data from cardiac surgery patients, our findings may not apply equally well to other patient groups or medical settings. Second, while using synthetic data helped balance our dataset, this approach might introduce biases and may not fully capture the complexities of real AKI cases. This limitation raises important questions about the robustness of the model when applied to real-world clinical scenarios. Finally, although our results show significant improvements, they need validation across different medical settings and patient populations to confirm their broader applicability.

## Conclusion

This study represents a significant step forward in addressing the challenge of class imbalance in medical predictive modelling. By demonstrating the substantial improvements achievable through careful balancing techniques, we provide a framework for more accurate and clinically relevant machine learning approaches to AKI prediction.

## Acknowledgement

The author would like to thank the Director-General of Health Malaysia for granting permission to publish this report. The author is also very thankful for the reviewers' insightful comments and constructive suggestions. Additionally, we extend our appreciation to Dr Goh Kheng Wee for providing valuable data that significantly contributed to the completion of this research project.

## Conflicts of Interest

Nothing to disclose.

## Funds

Nothing to disclose.

## References

1. Tamargo C, Hanounh M, Cervantes CE (2024) Treatment of acute kidney injury: a review of current approaches and emerging innovations. *J Clin Med* 13: 2455.
2. Susantitaphong P, Cruz DN, Cerda J, Abulfaraj M, Alqahtani F, et al. (2013) Acute Kidney Injury Advisory Group of the American Society of Nephrology. World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol* 8: 1482-1493.
3. Rosner MH, Okusa MD (2005) Acute kidney injury associated with cardiac surgery. *Clin J Am Soc Nephrol* 1: 19-32.
4. Cardinale D, Cosentino N, Marenzi G, Cipolla CM (2019) High-volume hydration for the prevention of acute kidney injury after cardiac surgery. *J Thorac Dis* 11(Suppl 9): S1135-S1138.
5. O'Neal JB, Shaw AD, Billings FT (2016) Acute kidney injury following cardiac surgery: current understanding and future directions. *Crit Care* 20: 187.
6. Pickkers P, Darmon M, Hoste E, Joannidis M, Legrand M, et al. (2021) Acute kidney injury in the critically ill: An updated review on pathophysiology and management. *Intensive Care Med* 47:835-850.
7. Thakar CV, Arrigain S, Worley S, Yared JP, Paganini EP (2005) A clinical score to predict acute renal failure after cardiac surgery. *J Am Soc Nephrol* 16: 162-168.
8. Salmi M, Atif D, Oliva D, Abraham A, Ventura S (2024) Handling imbalanced medical datasets: review of a decade of research. *Artif Intell Rev* 57: 273.
9. Khushi M, Shaikat K, Alam TM, Hameed IA, Uddin S, et al. (2021) A comparative performance analysis of data resampling methods on imbalance medical data. *IEEE Access* 9: 109960-109975.
10. Noble RA, Lucas BJ, Selby NM (2020) Long-term outcomes in patients with acute kidney injury. *Clin J Am Soc Nephrol* 15: 423-429.
11. Zhao H, Wang R, Lei Y, Liao WH, Cao H, et al. (2022) Severity level diagnosis of Parkinson's disease by ensemble k-nearest neighbor under imbalanced data. *Expert Syst Appl* 189: 116113.
12. Bhattacharya M, Jurkovicz C, Shatkay H (2017) Assessing chronic kidney disease from office visit records using hierarchical meta-classification of an imbalanced dataset. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE 663-670.

13. Potharaju SP, Sreedevi M (2016) Ensembled rule-based classification algorithms for predicting imbalanced kidney disease data. *J Eng Sci Technol Rev* 9: 201–207.
14. Chaudhary K, Vaid A, Duffy Á, Paranjpe I, Jaladanki S, et al. (2020) Utilization of deep learning for subphenotype identification in sepsis-associated acute kidney injury. *Clin J Am Soc Nephrol* 15: 1557-1565.
15. Alfieri F, Ancona A, Tripepi G, Crosetto D, Randazzo V, et al. (2021) A deep-learning model to continuously predict severe acute kidney injury based on urine output changes in critically ill patients. *J Nephrol* 34: 1875-1886.
16. Goh KW, Teo WL, Ang YK, Kasmin FA, Diana OE, et al. (2017) Validating Cleveland Clinic score to predict acute kidney injury requiring dialysis after cardiac surgery. *Kidney Int Rep* 2: S18.
17. Fernando KRM, Tsoko CP (2021) Dynamically weighted balanced loss: class imbalanced learning and confidence calibration of deep neural networks. *IEEE Trans Neural Netw Learn Syst* 33: 2940-2951.
18. Liu N, Li X, Qi E, Xu M, Li L, Gao B (2020) A novel ensemble learning paradigm for medical diagnosis with imbalanced data. *IEEE Access* 8: 171263-171280.
19. Du G, Zhang J, Ma F, Zhao M, Lin Y, et al. (2021) Towards graph-based class-imbalance learning for hospital readmission. *Expert Syst Appl* 176:114791.
20. Phetlasy S, Ohzahata S, Wu C, Kato T (2019) Applying SMOTE for a sequential classifiers combination method to improve the performance of intrusion detection system. In: 2019 IEEE International Conference on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech).IEEE 255-258.
21. Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, et al. (2017) Learning from class-imbalanced data: review of methods and applications. *Expert Syst Appl* 73: 220-239.
22. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 16: 321-357.
23. Zhao Y, Wong ZS Y, Tsui KL (2018) A framework of rebalancing imbalanced healthcare data for rare events' classification: a case of look-alike sound-alike mix-up incident detection. *J Healthc Eng* 6275435.
24. Li J, Liu LS, Fong S, Wong RK, Mohammed S, et al. (2017) Adaptive swarm balancing algorithms for rare-event prediction in imbalanced healthcare data. *PLoS One* 12: e0180830.



- Advances In Industrial Biotechnology | ISSN: 2639-5665
- Advances In Microbiology Research | ISSN: 2689-694X
- Archives Of Surgery And Surgical Education | ISSN: 2689-3126
- Archives Of Urology
- Archives Of Zoological Studies | ISSN: 2640-7779
- Current Trends Medical And Biological Engineering
- International Journal Of Case Reports And Therapeutic Studies | ISSN: 2689-310X
- Journal Of Addiction & Addictive Disorders | ISSN: 2578-7276
- Journal Of Agronomy & Agricultural Science | ISSN: 2689-8292
- Journal Of AIDS Clinical Research & STDs | ISSN: 2572-7370
- Journal Of Alcoholism Drug Abuse & Substance Dependence | ISSN: 2572-9594
- Journal Of Allergy Disorders & Therapy | ISSN: 2470-749X
- Journal Of Alternative Complementary & Integrative Medicine | ISSN: 2470-7562
- Journal Of Alzheimers & Neurodegenerative Diseases | ISSN: 2572-9608
- Journal Of Anesthesia & Clinical Care | ISSN: 2378-8879
- Journal Of Angiology & Vascular Surgery | ISSN: 2572-7397
- Journal Of Animal Research & Veterinary Science | ISSN: 2639-3751
- Journal Of Aquaculture & Fisheries | ISSN: 2576-5523
- Journal Of Atmospheric & Earth Sciences | ISSN: 2689-8780
- Journal Of Biotech Research & Biochemistry
- Journal Of Brain & Neuroscience Research
- Journal Of Cancer Biology & Treatment | ISSN: 2470-7546
- Journal Of Cardiology Study & Research | ISSN: 2640-768X
- Journal Of Cell Biology & Cell Metabolism | ISSN: 2381-1943
- Journal Of Clinical Dermatology & Therapy | ISSN: 2378-8771
- Journal Of Clinical Immunology & Immunotherapy | ISSN: 2378-8844
- Journal Of Clinical Studies & Medical Case Reports | ISSN: 2378-8801
- Journal Of Community Medicine & Public Health Care | ISSN: 2381-1978
- Journal Of Cytology & Tissue Biology | ISSN: 2378-9107
- Journal Of Dairy Research & Technology | ISSN: 2688-9315
- Journal Of Dentistry Oral Health & Cosmesis | ISSN: 2473-6783
- Journal Of Diabetes & Metabolic Disorders | ISSN: 2381-201X
- Journal Of Emergency Medicine Trauma & Surgical Care | ISSN: 2378-8798
- Journal Of Environmental Science Current Research | ISSN: 2643-5020
- Journal Of Food Science & Nutrition | ISSN: 2470-1076
- Journal Of Forensic Legal & Investigative Sciences | ISSN: 2473-733X
- Journal Of Gastroenterology & Hepatology Research | ISSN: 2574-2566
- Journal Of Genetics & Genomic Sciences | ISSN: 2574-2485
- Journal Of Gerontology & Geriatric Medicine | ISSN: 2381-8662
- Journal Of Hematology Blood Transfusion & Disorders | ISSN: 2572-2999
- Journal Of Hospice & Palliative Medical Care
- Journal Of Human Endocrinology | ISSN: 2572-9640
- Journal Of Infectious & Non Infectious Diseases | ISSN: 2381-8654
- Journal Of Internal Medicine & Primary Healthcare | ISSN: 2574-2493
- Journal Of Light & Laser Current Trends
- Journal Of Medicine Study & Research | ISSN: 2639-5657
- Journal Of Modern Chemical Sciences
- Journal Of Nanotechnology Nanomedicine & Nanobiotechnology | ISSN: 2381-2044
- Journal Of Neonatology & Clinical Pediatrics | ISSN: 2378-878X
- Journal Of Nephrology & Renal Therapy | ISSN: 2473-7313
- Journal Of Non Invasive Vascular Investigation | ISSN: 2572-7400
- Journal Of Nuclear Medicine Radiology & Radiation Therapy | ISSN: 2572-7419
- Journal Of Obesity & Weight Loss | ISSN: 2473-7372
- Journal Of Ophthalmology & Clinical Research | ISSN: 2378-8887
- Journal Of Orthopedic Research & Physiotherapy | ISSN: 2381-2052
- Journal Of Otolaryngology Head & Neck Surgery | ISSN: 2573-010X
- Journal Of Pathology Clinical & Medical Research
- Journal Of Pharmacology Pharmaceutics & Pharmacovigilance | ISSN: 2639-5649
- Journal Of Physical Medicine Rehabilitation & Disabilities | ISSN: 2381-8670
- Journal Of Plant Science Current Research | ISSN: 2639-3743
- Journal Of Practical & Professional Nursing | ISSN: 2639-5681
- Journal Of Protein Research & Bioinformatics
- Journal Of Psychiatry Depression & Anxiety | ISSN: 2573-0150
- Journal Of Pulmonary Medicine & Respiratory Research | ISSN: 2573-0177
- Journal Of Reproductive Medicine Gynaecology & Obstetrics | ISSN: 2574-2574
- Journal Of Stem Cells Research Development & Therapy | ISSN: 2381-2060
- Journal Of Surgery Current Trends & Innovations | ISSN: 2578-7284
- Journal Of Toxicology Current Research | ISSN: 2639-3735
- Journal Of Translational Science And Research
- Journal Of Vaccines Research & Vaccination | ISSN: 2573-0193
- Journal Of Virology & Antivirals
- Sports Medicine And Injury Care Journal | ISSN: 2689-8829
- Trends In Anatomy & Physiology | ISSN: 2640-7752

Submit Your Manuscript: <https://www.heraldopenaccess.us/submit-manuscript>