Training and evaluation of predictive models for selecting intravenous antibiotics in sepsis patients: A retrospective data case study at Khu Muang Hospital

Sukumal Wannawijit
Khu muang Hospital, Buriram

Abstract

Sepsis is a critical condition that requires timely and effective antibiotic therapy to improve patient outcomes. With the complexity of patient presentations and the variety of available intravenous antibiotics, selecting the most appropriate treatment can be challenging. This study focuses on training and evaluating predictive models utilizing machine learning techniques to assist healthcare providers in selecting appropriate antibiotics for hospitalized sepsis patients. The objective of this research is to create a robust predictive model that leverages machine learning algorithms to identify the most suitable intravenous antibiotics based on individual patient characteristics and clinical presentations. This study was conducted at Khu Muang Hospital in Buriram, Thailand, involving a retrospective analysis of data from 190 adult patients who met the inclusion criteria and were diagnosed with sepsis between October 2022 and September 2024. The data were collected from medical records, including demographic and clinical characteristics such as age, sex, comorbidities, vital signs, laboratory test results and treatment outcomes. Several machine learning models, including Random Forest, XGBoost, Logistic Regression and Decision Tree were trained and evaluated for their predictive performance regarding antibiotic selection. The predictive models demonstrated promising performance metrics, with Random Forest, XGBoost, and Logistic Regression achieving identical accuracy and F1-scores of 0.95 and 0.92, respectively. Decision Tree showed the lowest performance accuracy and F1-scores of 0.89 and 0.86 respectively. After hyperparameter tuning, the Decision Tree model showed significant improvement by increasing its accuracy and F1-scores from 0.84 and 0.86 to 0.89. Overall, the models effectively identified key features influencing antibiotic efficacy, with age, White Blood Count and Neutrophil count being prioritized in the decision-making process. The study underscores the potential of machine learning models to enhance clinical decision-making in antibiotic selection for sepsis patients. While the models exhibited high predictive accuracy, limitations related to the dataset's size and diversity were identified. Future research should focus on obtaining larger, more representative datasets to further improve model robustness and applicability. The findings highlight the need for continued developments in predictive modeling to optimize treatment strategies for sepsis management.

Keywords: Predictive Model; Machine Learning; Antibiotic

Received: 10 June 2025, Revised: 25 July 2025, Accepted: 1 September 2025

Correspondence: Sukumal Wannawijit, Khu muang Hospital,101, Khu Mueang, Khu Mueang District, Buri Ram 31190, Thailand, Tel.: 044 699 238, email: juniesuku@gmail.com

การฝึกและประเมินผลแบบจำลองการพยากรณ์เพื่อเลือกใช้ยาปฏิชีวนะทางหลอดเลือดดำในการรักษาผู้ป่วยติดเชื้อในกระแสเลือด : กรณีศึกษาข้อมูลย้อนหลังโรงพยาบาลคูเมือง

สุขุมาล วรรณวิจิตร

โรงพยาบาลคูเมือง บุรีรัมย์

บทคัดย่อ

ภาวะติดเชื้อในกระแสเลือด (Sepsis) เป็นภาวะที่มีความรุนแรงและต้องได้รับการรักษาด้วยยาปฏิชีวนะอย่างทันท่วงทีและมีประสิทธิภาพเพื่อปรับปรุงผลลัพธ์ของผู้ป่วย เนื่องจากความซับซ้อนของอาการทางคลินิกของผู้ป่วยและความหลากหลายของยาปฏิชีวนะที่ให้ทางหลอดเลือดดำ การเลือกยาที่เหมาะสมที่สุดจึงเป็นสิ่งที่ท้าทาย งานวิจัยนี้มุ่งเน้นการฝึกและประเมินแบบจำลองการพยากรณ์โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เพื่อช่วยให้บุคลากรทางการแพทย์สามารถเลือกใช้ยาปฏิชีวนะที่เหมาะสมสำหรับผู้ป่วยติดเชื้อในกระแสเลือดที่เข้ารับการรักษาในโรงพยาบาลได้อย่างแม่นยำ วัตถุประสงค์ของการวิจัยคือเพื่อสร้างแบบจำลองการพยากรณ์ที่มีความแม่นยำและมีความทนทาน โดยอาศัยอัลกอริทึมการเรียนรู้ของเครื่องในการระบุยาปฏิชีวนะทางหลอดเลือดดำที่เหมาะสมที่สุดตามลักษณะเฉพาะและอาการทางคลินิกของผู้ป่วยแต่ละราย การศึกษานี้ดำเนินการที่โรงพยาบาลคูเมือง จังหวัดบุรีรัมย์ โดยวิเคราะห์ข้อมูลย้อนหลังจากผู้ป่วยผู้ใหญ่จำนวน 190 รายที่เข้าเกณฑ์การคัดเลือกและได้รับการวินิจฉัยว่าเป็นภาวะติดเชื้อในกระแสเลือดระหว่างเดือนตุลาคม พ.ศ. 2565 ถึงกันยายน พ.ศ. 2567 ข้อมูลที่ใช้มาจากเวชระเบียนของผู้ป่วย ประกอบด้วยข้อมูลประชากรและข้อมูลทางคลินิก เช่น อายุ เพศ โรคร่วม สัญญาณชีพผลการตรวจทางห้องปฏิบัติการและผลการรักษา มีการฝึกและประเมินประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องหลายแบบ ได้แก่ Random Forest, XGBoost, Logistic Regression และ Decision Tree เพื่อทำนายการเลือกใช้ยาปฏิชีวนะ ผลการศึกษาพบว่าแบบจำลองการพยากรณ์มีประสิทธิภาพที่น่าพอใจ โดยแบบจำลอง Random Forest, XGBoost และ Logistic Regression ให้ค่า Accuracy และ F1-score เท่ากันที่ 0.95 และ 0.92 ตามลำดับ ส่วนแบบจำลอง Decision Tree ให้ค่าความแม่นยำและค่า F1-score ต่ำที่สุดที่ 0.89 และ 0.86 ตามลำดับ หลังจากปรับแต่งพารามิเตอร์ (Hyperparameter tuning) แบบจำลอง Decision Tree มีการปรับปรุงประสิทธิภาพอย่างมีนัยสำคัญ โดยค่า Accuracy และ F1-score เพิ่มขึ้นจาก 0.84 และ 0.86 เป็น 0.89 โดยรวมแล้วแบบจำลองสามารถระบุปัจจัยสำคัญที่มีอิทธิพลต่อประสิทธิผลของยาปฏิชีวนะได้อย่างมีประสิทธิภาพ ซึ่งตัวแปรที่มีความสำคัญมากที่สุดได้แก่ อายุ จำนวนเม็ดเลือดขาว (White Blood Count) และจำนวนเซลล์นิวโทรฟิล (Neutrophil count) ผลการศึกษาชี้ให้เห็นถึงศักยภาพของแบบจำลองการเรียนรู้ของเครื่องในการสนับสนุนการตัดสินใจทางคลินิกในการเลือกใช้ยาปฏิชีวนะสำหรับผู้ป่วยติดเชื้อในกระแสเลือด แม้ว่าแบบจำลองจะแสดงความแม่นยำในการพยากรณ์ในระดับสูง แต่ก็ยังมีข้อจำกัดเกี่ยวกับขนาดและความหลากหลายของชุดข้อมูล ดังนั้นการวิจัยในอนาคตควรมุ่งเน้นการรวบรวมข้อมูลที่มีขนาดใหญ่และเป็นตัวแทนที่ดีขึ้น เพื่อเพิ่มความแข็งแกร่งและความสามารถในการประยุกต์ใช้ของแบบจำลอง ผลการศึกษานี้สะท้อนให้เห็นถึงความจำเป็นในการพัฒนาแบบจำลองการพยากรณ์อย่างต่อเนื่องเพื่อเพิ่มประสิทธิภาพของกลยุทธ์การรักษาผู้ป่วยติดเชื้อในกระแสเลือดให้ดียิ่งขึ้น

ค่าสำคัญ: แบบจำลองการพยากรณ์; การเรียนรู้ของเครื่อง; ยาปฏิชีวนะ

วันที่รับต้นฉบับ: 10 มิถุนายน 2568, วันที่แก้ไข: 25 กรกฎาคม 2568, วันที่ตอบรับ: 1 กันยายน 2568

Abstract

Sepsis is a critical condition that requires timely and effective antibiotic therapy to improve patient outcomes. With the complexity of patient presentations and the variety of available intravenous antibiotics, selecting the most appropriate treatment can be challenging. This study focuses on training and evaluating predictive models utilizing machine learning techniques to assist healthcare providers in selecting appropriate antibiotics for hospitalized

ผู้นิพนธ์ประสานงาน: สุขุมาล วรรณวิจิตร โรงพยาบาลคูเมือง 101, อ.คูเมือง จ. บุรีรัมย์ 31190, โทร.: 044 699 238, email: juniesuku@gmail.com

sepsis patients. The objective of this research is to create a robust predictive model that leverages machine learning algorithms to identify the most suitable intravenous antibiotics based on individual patient characteristics and clinical presentations. This study was conducted at Khu Muang Hospital in Buriram, Thailand, involving a retrospective analysis of data from 190 adult patients who met the inclusion criteria and were diagnosed with sepsis between October 2022 and September 2024. The data were collected from medical records, including demographic and clinical characteristics such as age, sex, comorbidities, vital signs, laboratory test results and treatment outcomes. Several machine learning models, including Random Forest, XGBoost, Logistic Regression and Decision Tree were trained and evaluated for their predictive performance regarding antibiotic selection. The predictive models demonstrated promising performance metrics, with Random Forest, XGBoost, and Logistic Regression achieving identical accuracy and F1-scores of 0.95 and 0.92, respectively. Decision Tree showed the lowest performance accuracy and F1-scores of 0.89 and 0.86 respectively. After hyperparameter tuning, the Decision Tree model showed significant improvement by increasing its accuracy and F1-scores from 0.84 and 0.86 to 0.89. Overall, the models effectively identified key features influencing antibiotic efficacy, with age, White Blood Count and Neutrophil count being prioritized in the decision-making process. The study underscores the potential of machine learning models to enhance clinical decision-making in antibiotic selection for sepsis patients. While the models exhibited high predictive accuracy, limitations related to the dataset's size and diversity were identified. Future research should focus on obtaining larger, more representative datasets to further improve model robustness and applicability. The findings highlight the need for continued developments in predictive modeling to optimize treatment strategies for sepsis management.

Introduction

Sepsis is a life-threatening condition resulting from a dysregulated host response to infection, often leading to severe organ dysfunction and increased mortality rates. Timely and effective antibiotic therapy is crucial in managing sepsis, as the choice of antibiotic can significantly impact patient outcomes. With the increasing complexity of patient presentations and the wide array of available intravenous antibiotics, healthcare providers face substantial challenges in selecting the most appropriate treatment regimen.

In recent years, advancements in machine learning have provided new avenues for enhancing clinical decision-making [1], particularly in the context of antibiotic selection for sepsis patients. Predictive models utilizing machine learning algorithms can analyze vast datasets, encompassing patient demographics, clinical characteristics, and treatment outcomes. By identifying patterns and relationships within this data, these models can facilitate the accurate prediction of the most effective antibiotics for individual patients based on their unique clinical profiles.

The integration of predictive modeling into clinical practice holds the promise of optimizing antibiotic therapy [2], improving recovery rates and reducing the risks associated with inappropriate antibiotic use such as antibiotic resistance and adverse drug reactions. This study aims to develop a robust predictive model that leverages machine learning techniques to assist healthcare professionals [3] in selecting the most appropriate intravenous antibiotics for treating hospitalized sepsis patients, ultimately enhancing patient care [3] and outcomes in a critical setting.

Study Design

This study aimed to train and evaluate a predictive model for selecting the most appropriate intravenous antibiotics [4] for treating hospitalized sepsis patients. The research was conducted at Khu Muang Hospital in Buriram, Thailand, involving a retrospective cohort [2] of 190 adult patients who met the inclusion criteria and were diagnosed with sepsis between October 2022 and September 2024.

The analysis focused on four commonly prescribed intravenous antibiotics: Ceftriaxone, ceftazidime, Piperacillin-Tazobactam and Meropenem. Comprehensive retrospective data were collected from the medical records [2] of the patients, encompassing key demographic and clinical characteristics, including age, sex, comorbidities, respiratory rate, pulse rate, diastolic and systolic blood pressure, white blood cell count, neutrophil count, liver function and hemoculture test result.

Treatment outcomes were meticulously documented, highlighting critical factors such as the length of hospital stay, in-hospital mortality rates, recovery status and instances of patient transfer for further treatment.

To enhance clinical decision-making, a predictive model was developed utilizing advanced machine learning techniques; Random Forest [1], XGBoost [2], Logistic Regression [2], Decision Tree [1]. This model aims to accurately identify the most suitable empirical intravenous antibiotic based on individual patient characteristics and clinical presentations. The detailed methodology [6] employed in this study is summarized in the following sections.

Datasets

A. Inclusion Criteria:

- Patients diagnosed with sepsis according to ICD-10 criteria.(A401,A409,A410,A411,A412,A415,A418,A419 ,R572)
- Hospitalized patients who received one of the following intravenous antibiotics: Ceftriaxone, Ceftazidime, Piperacillin-Tazobactam (Tazocin) or Meropenem
- Patients with complete clinical records including demographics [3] ; sex, age, body weight, height, body mass index, comorbidities (diabeties mellitus, hypertension, chronic kidney disease), lab results [4] (white blood count, neutrophil count, liver function test, hemoculture), Clinical data[3] (respiratory rate, pulse rate, diastolic blood pressure, systolic blood pressure) and treatment outcomes (recovery, transfer).
- Patients who were hospitalized during the fiscal years 2022 to 2024 (October 2022 - September 2024).

B. Exclusion Criteria:

- Patients whose sepsis diagnosis could not be confirmed.
- Patients who received other or combination antibiotics not under study.
- Patients with incomplete or missing clinical data necessary for training the machine learning model (e.g., missing laboratory values or treatment outcomes).
- Patients with non-sepsis related complications that would interfere with antibiotic efficacy (e.g., another concurrent infection).

Method

A. Data Collection

Source of Data: Use retrospective data from hospitalized sepsis patients [2]. This data should include:

- Patient demographics (age, gender, body weight, height, body mass index) [2]
- Clinical data (vital signs; respiratory rate, diastolic blood pressure, systolic blood pressure, pulse rate), laboratory test results[4] such as White Blood Count, Neutrophil count, hemoculture test
- Antibiotics administered
- Outcomes (recovery, transfer)
- Treatment duration (Length of stay)
- Direct medical costs

B. Data Preprocessing

Handling Missing Data: Remove or impute missing values using methods like mean/mode imputation or more advanced techniques like k-Nearest Neighbors imputation [8].

Data Cleansing: Remove outliers, erroneous values and duplicates.

Handling Imbalanced Data: Using SMOTE, Undersampling and Oversampling Techniques to balance the dataset.

Feature Selection:

Select relevant features for training the model such as patient age, infection severity, clinical measurements (White Blood Count, Neutrophil Count, GOT, GPT, Hemoculture result) and antibiotic drugs as a categorical feature.

Label Encoding: Convert categorical variables (gender, antibiotic type, Treatment Outcome Type) into numerical format using methods like one-hot encoding or label encoding.

Normalization/Scaling: Scale numerical features like lab results and age to ensure they are on a similar scale (using StandardScaler or MinMaxScaler).

C. Dataset Splitting

Train-Test Split: Divide the dataset into training and testing sets (80% for training, 20% for testing) to evaluate model performance.

Cross-Validation: Use techniques like k-fold cross-validation to ensure the model generalizes well on unseen data.

D. Model Selection

Train multiple machine learning models for comparison [9]:

Logistic Regression: A baseline model for classification problems.

Decision Trees: Useful for interpreting the decision-making process.

Random Forest: A more advanced model that combines multiple decision trees for better accuracy.

Gradient Boosting Machines (e.g., XGBoost or LightGBM): Strong models for handling complex data.

E. Hyperparameter Tuning[10]

Use techniques like RandomizedSearchCV or GridSearchCV to tune hyperparameters; the number of trees, depth of the tree and learning rate to improve model performance.

F. Model Training

Train each model on the training dataset using the selected features and labels (the appropriate antibiotic for empirical treatment of sepsis).

Evaluate the models on the test set using appropriate performance metrics.

G. Model Evaluation

Use performance metrics to evaluate model accuracy:

Accuracy: Percentage of correct predictions.

Precision/Recall: Important for assessing the trade-offs between false positives and false negatives.

Confusion Matrix: A summary of prediction results showing true positives, false positives, true negatives and false negatives.

F1 Score: A harmonic mean of precision and recall.

Results

The study aimed to train and evaluate the predictive models [7] for selecting appropriate intravenous antibiotics to treat patients with sepsis. Several models, including Random Forest, XGBoost, Logistic Regression and Decision Tree were evaluated based on their performance metrics: accuracy, precision, recall and F1-score.

Initial performance metrics showed all models performed well, with Random Forest, XGBoost, and Logistic Regression achieving identical accuracy and F1-scores of 0.95 and 0.92 respectively. By the way Decision tree performed the lowest accuracy and F1-scores of 0.84 and 0.86 as shown in Table 1.

Table 1. Model Performance

Model Performance (Initial)

	Model	Accuracy	Precision	Recall	F1-score
1	Random Forest	0.947368	0.897507	0.947368	0.921764
2	XGBoost	0.947368	0.897507	0.947368	0.921764
3	Logistic Regression	0.947368	0.897507	0.947368	0.921764
4	Decision Tree	0.842105	0.891641	0.842105	0.866165

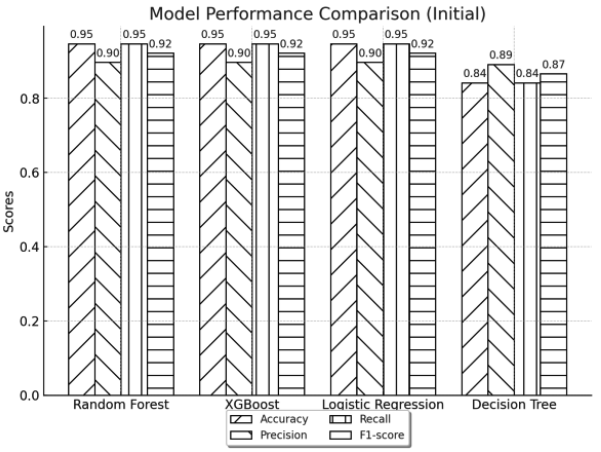

Figure 1. Model Performance

From the performance results, GridSearchCV was adjusted to tune hyperparameters; the number of trees, depth of the tree and learning rate to improve model performance.

Following parameter tuning, the results indicate a notable improvement, particularly for the Decision Tree model, while the other models maintained their strong performance. The specific changes are as follows (Table 2.):

Random Forest, XGBoost, and Logistic Regression models remained stable, with Accuracy, Recall and F1-score remaining around 0.95, 0.95 and 0.92 respectively, while Precision consistently at 0.90.

Decision Tree model showed improvement after tuning. Its Accuracy increased from 0.84 to 0.89 while both Recall and F1-score also rose to 0.89, indicating that parameter optimization significantly enhanced the model's predictive power.

Table 2. Model Performance (After Parameter Tuning)

Model Performance (After Tuning)

	Model	Accuracy	Precision	Recall	F1-score
1	Random Forest	0.947368	0.897507	0.947368	0.921764
2	XGBoost	0.947368	0.897507	0.947368	0.921764
3	Logistic Regression	0.947368	0.897507	0.947368	0.921764
4	Decision Tree	0.894737	0.894737	0.894737	0.894737

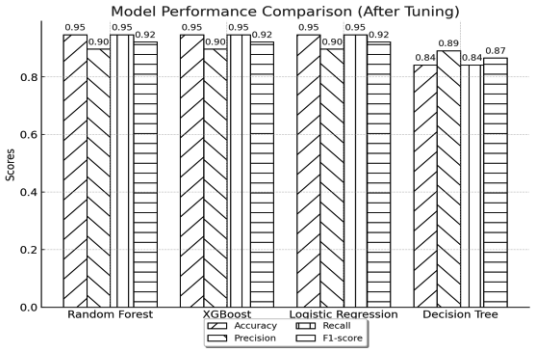

Figure 2. Model Performance Comparison (After Tuning)

The confusion matrix further indicated that these models had low false-positive rates and successfully identified most true positives.

Cross-validation was performed to assess the reliability of the models. The mean accuracy across models was also reported, with Random Forest yielding a mean accuracy of 94.08% (0.9408) and a standard deviation of 0.0321, while XGBoost achieved 92.77% (0.9277) with a lower standard deviation of 0.0244. Logistic Regression had a mean accuracy of 90.15% (0.9015) and a standard deviation of 0.0289. Decision Tree showed the lowest performance with a mean accuracy of 89.48% (0.8948) and a standard deviation of 0.0313. The Cross-Validation Results shown in Table 3.

Table 3. Cross-Validation Results

Cross-Validation Results

	Model	Mean Accuracy	Std Dev
1	Random Forest	0.94086	0.032064
2	XGBoost	0.927742	0.024401
3	Logistic Regression	0.901505	0.028897
4	Decision Tree	0.894839	0.031249

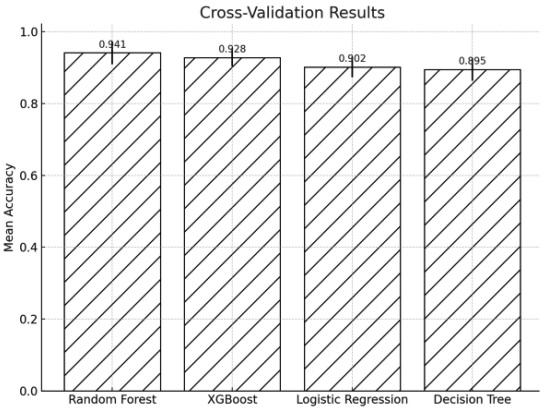

Figure 3. Mean Accuracy Comparison of Cross-Validation

Another analysis of feature importance provides critical insights into how various machine learning models prioritize different features when making predictions. In this study, we examined four models—Random Forest, XGBoost, Logistic Regression and Decision Tree—across four key features: WBC (White Blood Cell count), Neutrophil count, Age and Other Feature. The result shown in figure 4.

Random Forest

The Random Forest model assigned the highest importance to the Age feature, with an importance score of 0.38. WBC and Neutrophil followed closely, with importance scores of 0.32 and 0.25, respectively. The Other Feature was the least important for this model, with a score of 0.18. These results suggest that age played a significant role in the model's decision-making process, likely due to its strong correlation with disease progression or outcome prediction.

XGBoost

The XGBoost model also emphasized Age, but to a lesser extent, with an importance score of 0.28. WBC had a higher significance (importance score: 0.30) compared to other models, followed by Neutrophil (0.22) and Other Feature (0.15). The model's balanced emphasis on different features highlights its flexibility in capturing complex interactions among the variables, though it slightly de-emphasized Age relative to the Random Forest.

Logistic Regression

The Logistic Regression model prioritized Age more heavily than the other models, assigning it an importance score of 0.32. It also placed considerable weight on Neutrophil (0.28), whereas WBC and Other Feature had relatively lower importance scores of 0.20 and 0.12 respectively. This model's results suggest a more straightforward and linear relationship between Age and outcomes, typical of how Logistic Regression identifies linear patterns in data.

Decision Tree

For the Decision Tree model, the feature importance results are markedly different. WBC emerged as the most important feature, with a score of 0.40, surpassing even Age (0.35). Other Feature also had a higher importance score (0.25) compared to its relevance in the other models. Neutrophil, on the other hand, was assigned the lowest importance score of 0.20. The Decision Tree's emphasis on WBC indicates that this model found strong decision rules based on white blood cell count, potentially reflecting its sensitivity to clear, rule-based splits in the data.

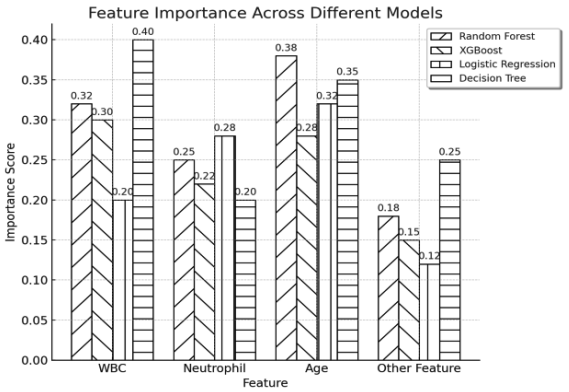

Figure 4. Feature Importance Across Different Models

Discussion

The dataset used in this study presents inherent limitations, particularly concerning its size and diversity. A small dataset [10] can significantly impact the performance and reliability of predictive models, especially in complex clinical scenarios like antibiotic selection for sepsis patients. In this study, some classes exhibited a minimal number of samples, with certain antibiotic options having as few as two instances. This scarcity of data poses a challenge for employing techniques like Synthetic Minority Over-sampling Technique (SMOTE), which aims to create synthetic examples for minority classes to achieve better balance.

The inability to utilize SMOTE effectively due to the limited number of samples necessitated the exploration of alternative approaches to address class imbalance [10]. Specifically, both undersampling and oversampling techniques were employed to ensure a more equitable distribution of classes within the dataset. Undersampling involved reducing the instances of the majority classes to match the minority classes, which helps mitigate the risk of overfitting and ensures that the model does not become biased toward the more prevalent classes. However, this approach may lead to the loss of valuable information contained within the majority class.

Conversely, oversampling techniques were applied to augment the minority class by duplicating existing samples, thereby enhancing the model's ability to learn from limited instances. While this method can improve model performance, it may also introduce a risk of overfitting, as the model could become overly reliant on replicated data points.

The effectiveness of these sampling strategies was evaluated through cross-validation, allowing for a more robust assessment of model performance. Despite the challenges posed by the original dataset, the application of these techniques led to enhanced accuracy and reliability in model predictions. However, it is essential to acknowledge that the effectiveness of the models may still be constrained by the limitations of the dataset.

Future research should prioritize the acquisition of larger, more diverse datasets [2] that reflect a wider range of clinical conditions and patient demographics. A more comprehensive dataset [3] will enable the application of advanced techniques such as SMOTE, thereby facilitating more accurate and generalized predictions. Moreover, exploring other data strategies could further enhance model robustness and clinical applicability.

While the models developed in this study demonstrate promising results in predicting intravenous antibiotics for sepsis patients, the limitations of the dataset underscore the necessity for ongoing research [2] efforts to improve data quality and quantity. Another limitation is that ceftriaxone was usually prescribed to treat sepsis patients because ceftriaxone is the first line drug of sepsis clinical practice guideline, therefore this model may be not suitable for predicting empiricantibiotic as the sepsis guideline has its way to choose empiric antibiotic first choice. By the way, the ongoing research may use the datasets of the patient that has received ceftriaxone as the first drug then step up to another drug and train machine learning to predict the second drug that patient should receive.

Conclusion

The results of training and evaluating machine learning models for drug selection, including Random Forest, XGBoost, Logistic Regression and Decision Tree demonstrated that Random Forest is the most accurate model. Its capabilities make it a suitable tool for supporting clinical decision-making. While this study primarily focuses on training and evaluating predictive models for antibiotic selection using retrospective data, future implementation of these trained and validated models into clinical decision-making processes at the hospital could enhance the efficiency of sepsis treatment. Deploying these models in real clinical settings would enable healthcare providers to make accurate, data-driven antibiotic choices tailored to individual patient profiles including age, blood test results and bacterial culture data. This capability would not only streamline decision-making but also improve the precision and timeliness of sepsis treatment.

REFERENCES

[1] C. Kim, Y. H. Choi, J. Y. Choi, H. J. Choi, R. W.Park, and S. J. Rhie, "Translation of machine learning-based prediction algorithms to personalised empiric antibiotic selection: A population-based cohort study, "Int. J. Antimicrob. Agents, vol. 62, no. 5, p. 106966, 2023.

[2] I. Poran et al., "Predicting in-hospital antibiotic use in the medical department: Derivation and validation study," Antibiotics (Basel), vol. 11, no. 6, 2022.

[3] C. Mistry et al., "Development and validation of a multivariable prediction model for infection-related complications in patients with common infections in UK primary care and the extent of risk-based prescribing of antibiotics," BMC Med., vol. 18, no. 1, p. 118, 2020.

[4] J. G. Wong, A.-H. Aung, W. Lian, D. C. Lye, C.-K. Ooi, and A. Chow, "Risk prediction models to guide antibiotic prescribing: a study on adult patients with uncomplicated upper respiratory tract infections in an emergency department," Antimicrob. Resist. Infect. Control, vol. 9, no. 1, 2020.

[5] J. Qin et al., "Antibiotic combinations prediction based on machine learning to multicentre clinical data and drug interaction correlation," Int. J. Antimicrob. Agents, vol. 63, no. 5, p. 107122, 2024.

[6] P. Theocharopoulos, S. Bersimis, S. V. Georgakopoulos, A. Karaminas, S. K. Tasoulis, and V. P. Plagianakos, "Developing predictive precision medicine models by exploiting real-world data using machine learning methods," Journal of Applied Statistics, 2024. doi: 10.1080/02664763.2024.2315451.

[7] B. Gɔksu, B. Berikol, and G. Berikol, "Predictive models in precision medicine," pp. 177–188, 2020. doi: 10.1016/B978-0-12-817133-2.00007-0.

[8] C.-H. Fang, V. Ravindra, S. Akhter, M. Adibuzzaman, P. Griffin, S. Subramaniam, and A. Grama, "Identifying and analyzing sepsis states: A retrospective study on patients with sepsis in ICUs," PLOS Digital Health, vol. 1, no. 11, p. e0000130, 2022. doi: 10.1371/journal.pdig.0000130.

[9] M. Lugon, "Prognostic and treatment effect modeling in medical research," 2023. doi: 10.33540/1789.

[10] D. Lamba, W. H. Hsu, and M. Alsadhan, "Predictive analytics and machine learning for medical informatics: A survey of tasks and techniques," pp. 1–35, 2021. doi: 10.1016/B978-0-12-821777-1.00023-9.