JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7582

OPTIMIZING TRANSFORMER-BASED LEARNING MODEL WITH TABTRANSFORMER FOR PREDICTING ANTIBIOTIC SUSCEPTIBILITY FROM MICROBIOLOGY MEDICAL RECORDS

Feri Sulianta[1*]; Endang Amalia[1]; Rosalin Samiharjo[1]; Noval Eka Herdinata[1]

Informatics Department[1]
Widyatama Univeristy,Bandung, Indonesia[1]
www.widyatama.ac.id[1]
feri.sulianta@widyatama.ac.id*, endang.amalia@widyatama.ac.id, rosalin.samiharjo@widyatama.ac.id, noval.herdinata@widyatama.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract—*Antimicrobial Resistance (AMR) has become a growing threat due to the increase in infections that are unresponsive to conventional therapies. Therefore, the development and optimization of Transformer-based Deep Learning using TabTransformer was employed to model the complex interactions between categorical features. This model was trained to predict antibiotic susceptibility at the individual culture level using the Antibiotic Resistance Microbiology Dataset (ARMD). To address the challenge of highly imbalanced data, the methodology applied includes extensive feature engineering to create historical and clinical variables, as well as the use of Focal Loss during training. After optimization, the final model demonstrated excellent discriminatory ability, with an Area Under the ROC Curve (AUC-ROC) of 0.93 and balanced classification performance, yielding a macro average F1-score of 0.82. Interpretability analysis using SHAP confirmed that patient clinical history and prior drug exposure were the most dominant predictive factors. These findings suggest that the Transformer-based Deep Learning architecture using TabTransformer, combined with clinically relevant feature engineering, can produce a reliable and evidence-based predictive tool.*

Keywords: *Antibiotic Susceptibility Prediction, Antimicrobial Resistance (AMR), SHAP Interpretability Analysis, TabTransformer, Transformer-Based Deep Learning.*

Intisari—*Resistansi Antimikroba (AMR) menjadi ancaman dengan meningkatnya infeksi yang tidak responsif terhadap terapi konvensional. Sehingga mengembangkan dan mengoptimalkan Transformer-based Deep Learning menggunakan TabTransformer digunakan untuk memodelkan interaksi kompleks antar fitur-fitur kategorikal. Model ini dilatih untuk memprediksi suseptibilitas antibiotik pada level kultur individual menggunakan Antibiotic Resistance Microbiology Dataset (ARMD). Untuk mengatasi tantangan data yang sangat tidak seimbang, metodologi yang diterapkan mencakup rekayasa fitur ekstensif untuk menciptakan variabel historis dan klinis, serta penggunaan Focal Loss selama pelatihan. Setelah melalui optimasi, model final menunjukkan kemampuan diskriminatif yang sangat baik dengan Area Under the ROC Curve (AUC-ROC) sebesar 0.93 dan performa klasifikasi yang seimbang dengan macro average F1-score sebesar 0.82. Analisis interpretabilitas menggunakan SHAP mengkonfirmasi bahwa riwayat klinis pasien dan paparan obat sebelumnya adalah faktor prediktif yang paling dominan. Temuan ini menunjukkan bahwa arsitektur Transformer-based Deep Learning menggunakan TabTransformer dengan rekayasa fitur yang relevan secara klinis dapat menghasilkan alat prediksi yang andal dan berbasis bukti.*

Kata Kunci: *Antimicrobial Resistance (AMR), Analisis Interpretabilitas SHAP, Prediksi Sensitivitas Antibiotik, TabTransformer, Transformer-Based Deep Learning.*

VOL. 11. NO. 3 FEBRUARY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i3.7582

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

INTRODUCTION

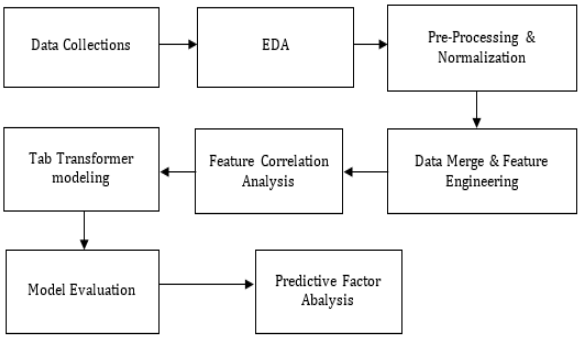
Antimicrobial resistance (AMR) has become a global threat to modern healthcare systems, with an increase in cases of bacterial infections that are unresponsive to conventional therapies[1]. This phenomenon is exacerbated by the irrational use of antibiotics and the lack of rapid diagnostic methods to guide targeted therapy[2]. There has been a 35% increase in multidrug-resistant (MDR) infection cases in the past decade [1]. Machine learning algorithms can be used to identify various underlying mechanisms of antimicrobial resistance, such as efflux pumps, target modification, and enzymatic inactivation, as well as to predict resistance in bacterial strains. After training on genomic data, several machine learning algorithms, including Support Vector Machines (SVM), Logistic Regression (LR) models, and Random Forests (RF), have demonstrated excellent accuracy in predicting antimicrobial resistance [3]. However, these models still rely on manually extracted features, which makes them less effective in handling data complexity [4]. Deep learning (DL) approaches have begun to show significant potential in predicting antibiotic resistance based on microbiological data, including mass spectrometry and genomic sequencing [4].

This approach reduces reliance on peak picking or baseline correction techniques, while improving accuracy by 10% compared to previous methods. The integration of transformer-based models has become an innovative solution for processing multidimensional data. The use of transformer-based models, as applied in predicting patient outcomes, enables simultaneous analysis of clinical and AST data with an error rate of < 2.5% for quinolones and cephalosporins [5]. Microbiological data play a crucial role in understanding antibiotic resistance patterns, which is critical for the development of effective treatment strategies[6]. Microbiological culture results and antibiotic sensitivity analysis provide the necessary information to determine appropriate therapy. For example, innovations in antibiotic sensitivity testing (AST) systems allow for rapid detection of bacterial resistance, thereby improving treatment effectiveness and reducing the risk of resistance spread. Systematic monitoring of microbiological data can reduce the spread of resistance genes in hospital environments and assist in better clinical decision-making[7].

This research aims to develop and evaluate a Transformer-based Deep Learning model using TabTransformer. The model is specifically designed to predict susceptibility outcomes (Resistant or Susceptible) at the individual culture level by leveraging complex microbiological electronic medical record (EMR) data. With the potential for integration into hospital business intelligence (BI) systems, this model is expected to become a practical decision support tool, providing evidence-based recommendations and contributing to the overall antibiotic stewardship program.

MATERIALS AND METHODS

There are several key stages to ensure the accuracy and effectiveness of the model in predicting antibiotic resistance based on microbiological medical record data using transformer-based deep learning techniques[8]. As described in Figure 1, the research process begins with the data collection phase, where relevant datasets are gathered from various sources to support the analysis of antibiotic resistance. Subsequently, an initial analysis is performed to understand the structure and characteristics of the data, followed by the merging of data from different sources into a unified dataset. The data then undergoes pre-processing and transformation, which includes data cleaning, handling missing values, and feature scaling to prepare it for use in the model. In the modeling phase, a transformer-based approach using TabTransformer is applied to build a prediction model capable of handling complex data. The model is evaluated using various metrics such as accuracy, precision, recall, F1-score, and AUC-ROC to ensure the quality of predictions and the model's generalization ability in the context of antibiotic resistance.



Source : (Koenigstein, 2025)

Figure 1. Reseach Framework

The Transformer architecture consists of an encoder and a decoder, each containing Nblocks. The input is a sequence of events, and the output is the predicted sequence of events [9]. Specifically, this model uses Scaled Dot-Product Attention, which is formulated as follows:

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

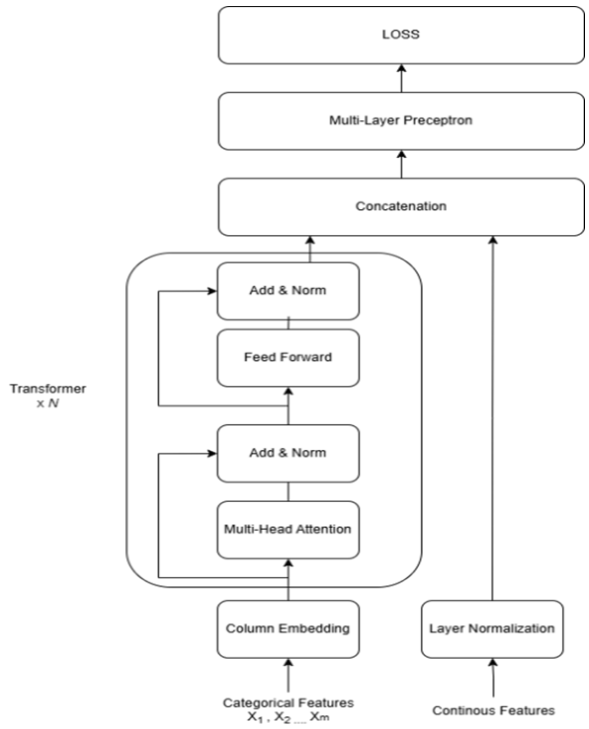
VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7582

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Tabel 1. Algorithms in Transformer Model

No	Layer	Algoritma
1	Selft-Attantion	Column-Aware Scaled Dot-Product
2	Multi-Head Attantion	Parallel Column-Aware Attention
3	Feed Forward Layer	Dense Layer + Gated Activation
4	Positional Encoding	Learnable Column Embedding
5	Masking	Feature Masking (untuk missing value)
6	Optimization	Adam Optimizer + Warmup Learning Rate

Source : (Sun [9], 2025)

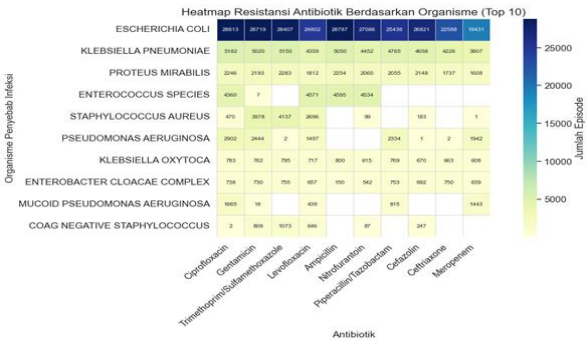


Source : (Koenigstein [8], 2025)
Figure 2. The TabTransformer - model architecture

In the Transformer architecture, Q(Queries), K(Keys), and V(Values) are matrices that represent the input, and $d_k$is the dimension of the keys. Each standard Transformer block consists of a Multi-Head Attention layer, followed by a Feed-Forward Network (FFN), with residual connections and layer normalization applied to each sub-layer. However, due to the tabular nature of the dataset used in this research, which has a mixed structure, the standard Transformer model is not directly applicable without modification. This is due to the differences in characteristics between text data and tabular data. In tabular data, the order of rows or columns generally does not have sequential meaning, but each column has a unique semantic meaning, and the interaction between features becomes critical [10]. To address this issue, this research adopts the TabTransformer approach, a neural network architecture developed to handle tabular data by leveraging a modified self-attention mechanism tailored to the columnar structure of the data. In the TabTransformer, as shown in Figure 2, each categorical feature is converted into a vector embedding through a lookup table, while numerical features are normalized before further processing.Each layer in the TabTransformer contains critical subcomponents, with many essential algorithms performing their respective functions, as outlined in Table 1.

Data Source

The Antibiotic Resistance Microbiology dataset consists of eleven main interconnected tables based on patient identifiers and examination procedures. These eleven tables provide comprehensive information, including clinical and demographic data, antibiotic resistance results, and patient antibiotic exposure history. All data has undergone a de-identification process to protect patient privacy, including the removal of direct identifiers, time jittering, age grouping, and binary encoding of gender without explicit labels. This standardized data structure supports comprehensive epidemiological and predictive analysis related to antimicrobial resistance [11]. The dataset consists of eleven tables that are linked through unique identifiers (anon_id for patients and order_proc_id_coded for culture procedure processes). The data is longitudinal in nature, enabling temporal analysis in antibiotic resistance prediction based on deep learning models. The dataset includes over 2,241,050 microbiological cultures from 283,715 patients. Figure 3 illustrates the antibiotic resistance patterns for the 10 most common infection-causing organisms.



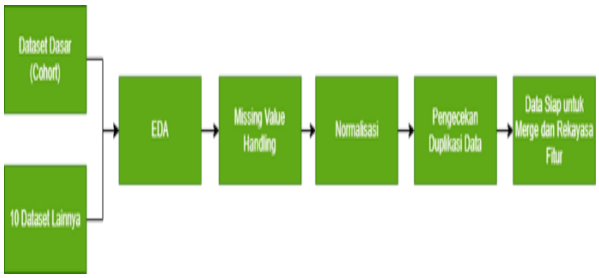
Source: (Research Results, 2025)
Figure 3. Top 10 Antibiotic & Organism

VOL. 11. NO. 3 FEBRUARY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i3.7582
JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

Pre-processing

Pre-processing improves machine learning model accuracy by eliminating noise and redundancy in the data [12]. Key steps in data pre-processing include:

a. **Exploratory Data Analysis (EDA):** The initial step to understand data patterns, trends, anomalies, and structure through statistical methods and visualizations [13].
b. **Missing Value Handling:** Critical for accurate model building, especially in clinical research. A hybrid imputation strategy was used based on feature meaning and data type.
c. **Normalization:** Ensures features with different scales don't affect model accuracy. Numeric features were standardized using StandardScaler, setting a mean of 0 and standard deviation of 1 [14].
d. **Duplicate Data Check:** Removing duplicate rows prevents bias during model training, as redundancy from merging data can reduce quality [15].



Source: (Research Results, 2025)

Figur 4. Data Preprocessing Workflow

The initial stage focused on integrating and preparing core data, namely microbiology_cultures_cohort.csv, along with ten supporting tables containing demographic, clinical, medical history, and laboratory results. The dataset had a "long" format, where each unique culture was identified by order_proc_id_coded and split into multiple rows, each representing the susceptibility test results for a specific antibiotic against the isolated organism. This format was intentional and not a result of duplication, allowing for antibiotic-specific analysis. To focus on binary classification, the target variable susceptibility was filtered to retain only definitive 'Susceptible' or 'Resistant' results. Ambiguous categories like 'Intermediate', 'Inconclusive', and 'Synergism' were excluded, reducing the dataset from 2,241,050 rows to 1,554,329 relevant rows for modeling.

The class distribution in the filtered dataset, revealing a major methodological challenge: severe class imbalance. The *Susceptible* class is significantly dominant (N ≈ 1.29 million) compared to the *Resistant* class (N ≈ 0.26 million), with a ratio of approximately 4.8:1. This imbalance guided the selection of evaluation metrics (e.g., F1-score and PR AUC) and modeling strategies, such as using Focal Loss to address data imbalance in later stages, avoiding bias towards the majority class. This preparation phase concluded with the creation of the target variable.

The *susceptibility* column was transformed using LabelEncoder, with 'Resistant' labeled as 0 and 'Susceptible' labeled as 1. After forming the base dataset, the next step was to enrich the data with clinically relevant features through feature engineering, a crucial step in the machine learning pipeline for electronic health record (EHR) data, shown to improve predictive performance [16]. The goal was to transform raw data into predictive signals understandable by the model. Feature engineering was performed by aggregating and transforming data from various source tables into a single summary row for each culture order.

Feature engineering specifically addressed two main challenges in the ARMD dataset. The first challenge is the complex longitudinal data structure, where a single patient (anon_id) may have multiple culture events (order_proc_id_coded) over time. The data cannot be directly used by the model, so the main feature engineering strategy focused on transforming this longitudinal data into static historical features [17].

The second challenge is the significant class imbalance. The target variable distribution analysis showed a much higher number of *Susceptible* cases compared to *Resistant* cases. This poses a challenge as predictive models tend to be biased towards the majority class and perform poorly on the minority class, which in this case is *Resistant*. While this imbalance was mainly addressed during the modeling stage (e.g., using Focal Loss), awareness of this issue also guided the feature engineering process to create strong signals for the minority class [18]. This process generated 189 features covering various dimensions, including demographics, comorbidities, procedure history, drug exposure, lab results, vital signs, and engineered historical features.

Final Transformation

After the integration and feature engineering phase, a comprehensive dataset rich in information is created, ready for final transformation before modeling. The TabTransformer architecture employs a unique approach for processing categorical features, differing from typical models. This model requires categorical inputs in integer format (e.g., 0, 1, 2),

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

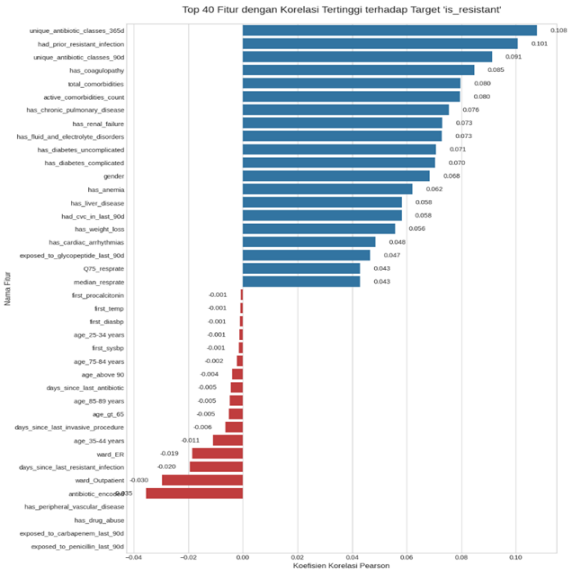
VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7582

which are then mapped to internal embedding layers. Therefore, all categorical features, including key ones like organism and antibiotic, are transformed using Label Encoding. This step is essential for the model to learn dense, meaningful vector representations (contextual embeddings) for each category. This approach enables the model to effectively capture complex relationships between organisms and antibiotics. Given the large data volume, the encoding process is efficiently executed using the Dask library. The transformation results in a final dataset with 1,554,329 rows and 189 feature columns. For flexibility in analysis and modeling, this dataset is saved in three file formats (parquet, csv, and pkl), serving as the single data source for all subsequent stages.

Before proceeding to modeling, a correlation analysis is conducted to validate the predictive signals of the engineered features. This analysis measures the strength of the linear relationship between each numeric feature and the target variable, typically using Pearson correlation coefficient, a standard metric for assessing linear associations [19]. While low correlations are common in complex medical datasets, this step remains crucial as a sanity check to ensure the engineered features are statistically relevant before further modeling.

The results of this analysis are visualized in Figure 6, highlighting features with the strongest positive and negative correlations. Key insights include that the strongest positively correlated features, such as *unique_antibiotic_classes_365d* (0.108) and *had_prior_resistant_infection* (0.103), are clinically logical predictors of resistance. This confirms that diverse antibiotic exposure history and prior resistant infections are major risk factors. Additionally, various comorbidities (*has_coagulopathy*, *has_chronic_pulmonary_disease*, *has_renal_failure*) also show positive correlations, suggesting that patients with more complex health conditions are more likely to experience resistant infections.

The strongest negatively correlated features are *antibiotic_encoded* (-0.035) and *ward_Outpatient* (-0.030). The negative correlation with *ward_Outpatient* is logical, as community-acquired infections (outpatients) are generally more susceptible compared to hospital-acquired infections (Carestia, M. et al., 2023). These results provide initial confidence that the engineered features carry valid predictive signals and are ready for modeling.



Source: (Research Results, 2025)

Figure 5. Feature Correlation with Susceptibility

Train-Test Split

The dataset is divided into two parts: the training set (80%) and the test set (20%). This split ensures that the model is trained and evaluated objectively, as shown in Figure 12. The train-test split aims to ensure that the model not only learns patterns from the training data but also provides an unbiased estimate of its performance on unseen data. As explained by Géron (2022), using a separate test set is the standard method for evaluating generalization error, or the model's error rate on new data. In this study, the dataset is split with an 80-20 ratio, applying stratification to maintain class proportions.

Training the Model

The training process for the TabTransformer model is carefully designed to optimize predictive performance while addressing the main challenge of class imbalance in the dataset. Every component of the training workflow, from the loss function to the callback mechanisms, is fine-tuned. Table 2 outlines the hyperparameters used during the model training.

Table 2. Hyperprameter Setting

Hyperparameter	Description	Value
embedding_dim	Embedding vector dimensions for categorical features.	64
depth	Number of Transformer block stacks.	6
heads	Number of heads in Multi-Head Attention.	8
attn_dropout	Dropout rate in the Attention layer.	0.1

Hyperparameter	Description	Value
ff_dropout	Dropout rate in the Feed-Forward layer.	0.1
mlp_hidden_factors	Scaling factor for the layer size in the MLP head.	[4, 4, 2]
mlp_dropout	Dropout rate in the MLP head.	0.2
learning_rate	Initial learning rate for the Adam optimizer.	3e-4 (0.0003)
batch_size	Number of samples per batch during training.	256
epochs	Maximum number of epochs (controlled by EarlyStopping).	50

Source: (Research Results, 2025)

The BinaryFocalCrossentropy loss function with a gamma parameter of 2.0 is used. Unlike standard cross-entropy, which assigns equal weight to all samples, Focal Loss dynamically reduces the contribution of easily classified samples, forcing the model to focus on the harder-to-classify one [20]. This approach is particularly effective for addressing severe class imbalance. For model parameter optimization, the Adam optimizer is chosen for its computational efficiency and ability to compute adaptive learning rates for each parameter [21]. The initial learning rate is set to 3e-4. To ensure stable convergence, a dynamic learning rate strategy is employed using the ReduceLROnPlateau callback. This callback monitors the validation_loss metric and reduces the current learning rate by a factor of 0.5 if no improvement in performance is detected over three epochs.

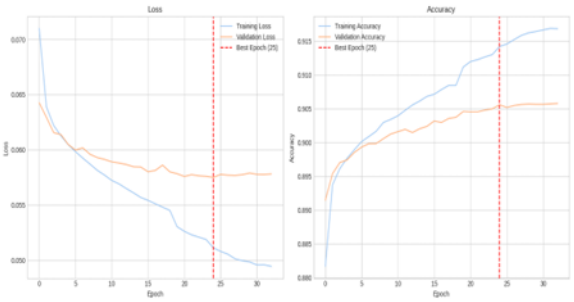
To prevent overfitting and enhance training efficiency, two additional callbacks are implemented. First, EarlyStopping is activated to halt training prematurely if the validation_loss does not improve over eight epochs. The argument *restore_best_weights=True* ensures that the model's weights are reverted to those from the epoch with the best performance [21]. Simultaneously, the ModelCheckpoint callback is used to permanently save the best model version (*best_model.keras*) whenever a new minimum validation_loss is achieved. The model is trained with a batch size of 256 for up to 50 epochs, with the actual training duration determined by EarlyStopping.

Model Evaluation Metrics

Model performance is evaluated using a comprehensive set of metrics to provide a holistic view, particularly given the imbalanced nature of the data where accuracy alone is insufficient. The primary metrics for assessing the model's ability to identify susceptibility classes are Precision, Recall, and F1-Score. Additionally, the model's overall discriminative ability across all probability thresholds is evaluated using the Area Under the Curve (AUC) for both the ROC (Receiver Operating Characteristic) curve and the Precision-Recall (PR) curve.

RESULTS AND DISCUSSION

The model performance is objectively measured through a comprehensive set of metrics, including the classification report, confusion matrix, and analysis of the Area Under the Curve (AUC) for both the ROC (Receiver Operating Characteristic) and Precision-Recall curves to assess the model's discriminative power. Additionally, an in-depth interpretative analysis is conducted using SHAP (SHapley Additive exPlanations) to identify and discuss the most dominant predictive factors. The purpose of using two evaluation methods, AUC-ROC and Precision-Recall curves, is to provide a more comprehensive assessment of the model's performance, especially with imbalanced data. AUC-ROC measures class distinction, while Precision-Recall focuses on detecting the minority class. Together, they offer a balanced view of the model's ability to handle both classes effectively The training process is dynamically evaluated at each epoch to monitor convergence and prevent overfitting. Figure 6 presents the loss and accuracy graphs for the model on both the training and validation data throughout the training process.



Source: (Research Results, 2025)
Figure 6. Grafik Loss dan Grafik Accuracy in Training Process

The loss graph (left) shows a consistently decreasing training loss curve from start to finish, indicating that the model effectively learned patterns from the training data. Similarly, the validation loss curve also experiences a smooth decline, reaching its lowest point at epoch 25, marked by the vertical line, before leveling off. The stability of the validation curve, which does not increase after reaching its minimum, strongly suggests that the EarlyStopping callback successfully halted training at the optimal convergence point, effectively preventing overfitting. This positive trend is confirmed in the

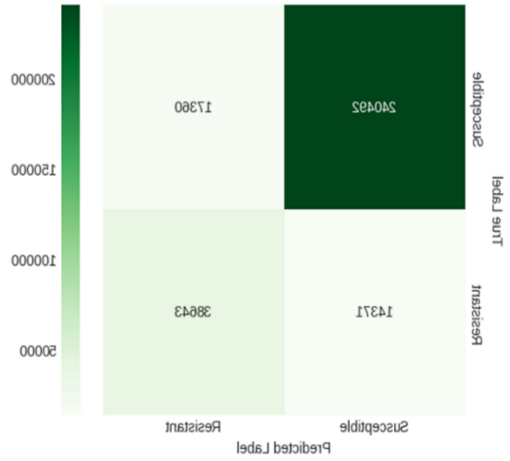
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7582

accuracy graph (right), where both the training accuracy and validation accuracy curves rise synchronously and plateau at a high level (around 90.6% for validation). The small and stable gap between the two accuracy curves further supports the conclusion that the model demonstrates good generalization capability. The curve dynamics indicate a stable and effective training process. Evaluation on the test data shows that the final optimized model performs strongly and balanced. The quantitative results of the evaluation are presented in the Classification Report in Figure 7, with performance visualization in the Confusion Matrix in Figure 8.

	precision	recall	f1-score	support
Susceptible	0.94	0.93	0.94	257852
Resistant	0.69	0.73	0.71	53014
accuracy			0.90	310866
macro avg	0.82	0.83	0.82	310866
weighted avg	0.90	0.90	0.90	310866

Source: (Research Results, 2025)
Figure 7. Classification Report

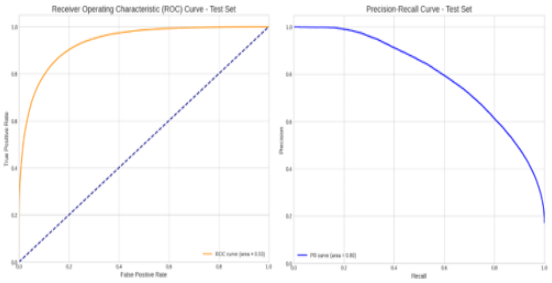


Source: (Research Results, 2025)
Figure 8. Confusion Matrix

Test Data Evaluation

Evaluation on the test data shows that the developed final model performs strongly and balanced. As summarized in the Classification Report in Figure 14, the model achieves an overall accuracy rate of 90%, indicating high predictive validity across the test samples. However, given the inherent class imbalance in medical data, this metric alone is not sufficiently representative. Therefore, the macro average F1-score is used as the primary indicator of balanced performance, with the model achieving a score of 0.82. This score reflects the model's ability to maintain a balance between macro average precision (0.82) and macro average recall (0.83).

The model is not simply accurate by predicting the majority class but is also effectively capable of fairly identifying cases from both classes. This balance confirms that the resulting model is a reliable and not significantly biased tool, a fundamental prerequisite for clinical applications. The model's discriminative ability, or its capacity to distinguish between classes, is further evaluated through the ROC and Precision-Recall curves presented in Figure 10.



Source: (Research Results, 2025)
Figure 9. AUC-ROC and Precision-Recall Curve

The model demonstrates excellent class separation ability, as evidenced by an Area Under the ROC Curve (AUC-ROC) score of 0.93. The AUC score measures how well the model differentiates between the two classes, with a score of 0.5 representing random guessing and 1.0 indicating perfect separation. With a score of 0.93, the model shows near-perfect discriminative ability, meaning it achieves a very high True Positive Rate (Detection Rate for Resistant cases) while maintaining a very low False Positive Rate (Error Rate for Susceptible cases).

Given the imbalanced nature of the dataset, Precision-Recall (PR) curve analysis is crucial. The model achieves an Area Under the Precision-Recall Curve (PR AUC) score of 0.80, highlighting its ability to maintain high precision even while maximizing recall for the minority Resistant class. This indicates the model's robustness in providing practically useful predictions, where detecting rare Resistant cases is the main challenge.

Model Performance Interpretation

The quantitative results presented in the previous chapter show that the optimized TabTransformer model has achieved strong and balanced performance. The macro average F1-score of 0.82 confirms that the model excels not only in predicting the majority class (Susceptible) but also has solid capabilities in handling the minority class (Resistant). This balanced performance is critical, as

the goal of the research is to create a reliable tool for clinical data environments that are inherently imbalanced.

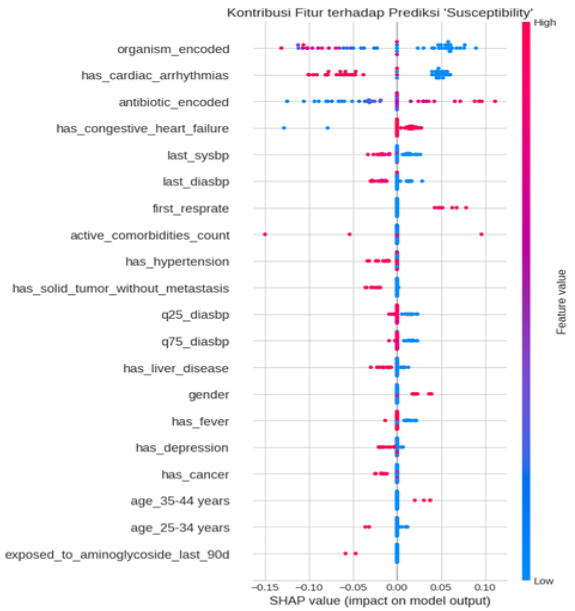
A deeper analysis of each individual class provides a comprehensive view of the model's performance profile. For the majority class, Susceptible, the model shows very high and reliable performance. With precision (0.94) and recall (0.94), the model almost never errs in identifying Susceptible cases and rarely misclassifies Resistant cases as Susceptible. Reliability in this class is essential to ensure that patients who should be treated with certain antibiotics are not misclassified.

On the other hand, for the minority class, Resistant, the model strikes a pragmatic balance between precision (0.69) and recall (0.73). A recall rate of 73% implies good sensitivity in detecting most Resistant cases. Meanwhile, a precision rate of 69% indicates a manageable false positive rate, with most of the model's Resistant predictions being correct. This combination of strong performance in one class and solid balance in the other shows that the developed model is a robust tool, capable of handling both prediction scenarios with clear and measurable capability.

The high AUC-ROC (0.93) and PR AUC (0.80) further confirm the fundamental strength of the model. The near-perfect AUC-ROC score indicates that the model has good discriminative ability in separating the two class populations. Meanwhile, the PR AUC score, well above the random baseline, proves that the model's performance remains strong even under the pressure of severe class imbalance. The developed model has successfully maximized the available predictive signals to generate reliable and clinically meaningful predictions.

Analysis of Key Predictive Factors

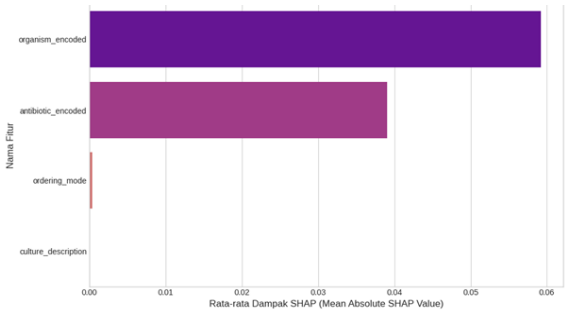
To understand the model's internal decision-making process and identify the most influential factors, an interpretability analysis was conducted using SHAP (SHapley Additive exPlanations). The SHAP analysis results in Figure 11 indicate that *organism_encoded* is the most dominant predictive factor in the model. The feature's values show a broad range of impact along the SHAP axis, signifying that the type of organism consistently plays a key role in determining whether a case is likely to be Resistant or Susceptible. The next most influential factors are combinations of comorbidity and microbiological features, namely *has_cardiac_arrhythmias* and *antibiotic_encoded*.



Source: (Research Results, 2025)

Figure 10. SHAP (SHapley Additive exPlanations

The finding that heart conditions such as arrhythmias and congestive heart failure (*has_congestive_heart_failure*) are important predictors offers significant clinical insights. The presence of these comorbidities (indicated by red/high feature values) consistently contributes negative SHAP values, meaning they push the prediction away from Susceptible or towards Resistant.



Source: (Research Results, 2025)

Figure 11. SHAP Categorical Features (SHapley Additive exPlanations)

It is evident that *organism_encoded* and *antibiotic_encoded* have a far greater impact compared to other categorical features, such as *ordering_mode* and *culture_description* (Fig. 11). This is clinically logical, as the identity of the microorganism and the type of antibiotic tested are the foundation of susceptibility testing, while the context of sample collection is secondary information. Overall, this SHAP analysis validates that the model does not act as a "black box" but

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7582

bases its predictions on clinically relevant factors, with a strong emphasis on microbial identity and the patient's heart health status. The SHAP analysis demonstrates that the model's strong performance is attributed to its ability to integrate information from three distinct domains: the patient's long-term history, specific microbiological context, and acute clinical status. The ability to weigh and prioritize features from these diverse sources is key to the reliability of the developed predictive model.

CONCLUSION

Transformer-based deep learning model, is developed and optimized a based on TabTransformer, to predict antibiotic susceptibility in response to the global challenge of Antimicrobial Resistance (AMR). A key innovation is the integration of methodological approach, including the engineering of clinical and historical features, as well as using Focal Loss to address data imbalance, the resulting model demonstrated strong and balanced performance. Evaluation results on test data revealed excellent discriminative ability, with an AUC-ROC of 0.93 and a macro average F1-score of 0.82. Further interpretability analysis confirmed that the model's predictions were based on clinically relevant factors, with patient history being the most dominant predictor. The study concludes that the Transformer-based deep learning architecture, utilizing TabTransformer combined with rich contextual data, can serve as a reliable and evidence-based decision support tool, especially when implemented in a hybrid system alongside deterministic clinical rules, to guide more prudent antibiotic use.

REFERENCE

[1] Y. Li, X. Cui, X. Yang, G. Liu, and J. Zhang, "Artificial intelligence in predicting pathogenic microorganisms' antimicrobial resistance: challenges, progress, and prospects," *Front Cell Infect Microbiol*, vol. 14, Nov. 2024, doi: 10.3389/fcimb.2024.1482186.

[2] F. Farhat, M. T. Athar, S. Ahmad, D. Ø. Madsen, and S. S. Sohail, "Antimicrobial resistance and machine learning: past, present, and future," *Front Microbiol*, vol. Volume 14-2023, 2023, doi: 10.3389/fmicb.2023.1179312.

[3] A. Jain, G. R. Dabburu, B. Samanta, N. Singhal, and M. Kumar, "An explainable machine learning pipeline for prediction of antimicrobial resistance in *Pseudomonas aeruginosa*," *Bioinformatics Advances*, vol. 5, no. 1, Dec. 2024, doi: 10.1093/bioadv/vbaf190.

[4] X. A. López-Cortés, J. M. Manríquez-Troncoso, R. Hernández-García, and D. Peralta, "MSDeepAMR: antimicrobial resistance prediction based on deep neural networks and transfer learning," *Front Microbiol*, vol. 15, Apr. 2024, doi: 10.3389/fmicb.2024.1361795.

[5] J. S. Inda-Díaz *et al.*, "Prediction of antibiotic resistance at the patient level using deep learning," *bioRxiv*, p. 2023.05.09.539832, Jan. 2024, doi: 10.1101/2023.05.09.539832.

[6] S. Cannella *et al.*, "Evaluation of EUCAST Rapid Antimicrobial Susceptibility Testing for Gram-Negative ESKAPEEc Pathogens in Blood Cultures, with a Focus on Carbapenemase-Producing Klebsiella pneumoniae in a University Hospital in Palermo, Italy," *Antibiotics*, vol. 14, no. 12, p. 1251, Dec. 2025, doi: 10.3390/antibiotics14121251.

[7] A. Elbehiry *et al.*, "Detection of antimicrobial resistance via state-of-the-art technologies versus conventional methods," *Front Microbiol*, vol. Volume 16-2025, 2025, doi: 10.3389/fmicb.2025.1549044.

[8] N. Koenigstein, *Transformers in Action*. Manning Online, 2025.

[9] Y. Sun *et al.*, "Efficient Attention Mechanisms for Large Language Models: A Survey," Aug. 2025.

[10] G. Badaro, M. Saeed, and P. Papotti, "Transformers for Tabular Data Representation: A Survey of Models and Applications," *Trans Assoc Comput Linguist*, vol. 11, pp. 227–249, Mar. 2023, doi: 10.1162/tacl_a_00544.

[11] F. , A. F. , & M. M. Nateghi Haredasht, "Antibiotic Resistance Microbiology Dataset (ARMD): A de-identified resource for studying antimicrobial resistance using electronic health records. ," 2025.

[12] D. Rifaldi, Abdul Fadlil, and Herman, "Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet 'Mental Health,'" *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 161–171, Apr. 2023, doi: 10.51454/decode.v3i2.131.

[13] P. Kamath B., G. Sharma, A. Bongale, D. Dharrao, and M. Seitshiro, "Exploratory Data Analysis and Water Potability Classification using Supervised Machine Learning Algorithms," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp.

20898–20903, Apr. 2025, doi: 10.48084/etasr.8904.

[14] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[15] J. , P. J. , & K. M. Han, *Data mining: Concepts and techniques*, 4th ed. Morgan Kaufmann, 2022.

[16] F. Xie *et al.*, "Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies," *J Biomed Inform*, vol. 126, p. 103980, Feb. 2022, doi: 10.1016/j.jbi.2021.103980.

[17] J. Xu, X. Xi, J. Chen, V. S. Sheng, J. Ma, and Z. Cui, "A Survey of Deep Learning for Electronic Health Records," *Applied Sciences*, vol. 12, no. 22, p. 11709, Nov. 2022, doi: 10.3390/app122211709.

[18] G. S. Hida and A. C. Alves Do Nascimento, "Overview of machine learning in class imbalance scenarios: Trends, challenges, and approaches," *Expert Syst Appl*, vol. 298, p. 129592, Mar. 2026, doi: 10.1016/j.eswa.2025.129592.

[19] Y. Shi, P. Wei, K. Feng, D.-C. Feng, and M. Beer, "A survey on machine learning approaches for uncertainty quantification of engineering systems," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, p. 11, Jun. 2025, doi: 10.1007/s44379-024-00011-x.

[20] X. Qian, S. Gao, W. Deng, and W. Wang, "Improving Oriented Object Detection by Scene Classification and Task-Aligned Focal Loss," *Mathematics*, vol. 12, no. 9, p. 1343, Apr. 2024, doi: 10.3390/math12091343.

[21] Y. Shao *et al.*, "An Improved BGE-Adam Optimization Algorithm Based on Entropy Weighting and Adaptive Gradient Strategy," *Symmetry (Basel)*, vol. 16, no. 5, p. 623, May 2024, doi: 10.3390/sym16050623.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri