

ORIGINAL ARTICLE

APPLYING MACHINE LEARNING TO ANALYZE TUBERCULOSIS SERVICE STANDARDS IN INDONESIA

Maria G. Catur Yuantari^{1*}, Yusthin M. Manglapy¹, Ririn Nurmandhani¹, Edi Jaya Kusuma², Haikal¹ and Dani Miarso³

¹Department of Public Health, Faculty of Health Science, Universitas Dian Nuswantoro, Semarang, Indonesia

²Department of Medical Record, Faculty of Health Science, Universitas Dian Nuswantoro, Semarang, Indonesia

³Public Health Office, Semarang, Indonesia

*Corresponding author: Maria Goretti Catur Yuantari

Email: mgcatur.yuantari@dsn.dinus.ac.id

ABSTRACT

Tuberculosis (TB) remains a significant public health challenge in Indonesia, with persistent disparities in service quality. Early identification of substandard TB services is critical to prevent transmission and accelerate elimination efforts. This study developed a machine-learning-based model to classify TB service compliance, identify key predictors, and provide evidence to guide targeted TB control policies. A retrospective design was applied using primary and secondary data from 1,722 TB patients in Semarang City, Indonesia. Three algorithms—Random Forest, K-Nearest Neighbors (KNN), and Boosting—were implemented with 5-fold stratified cross-validation, feature selection via Recursive Feature Elimination (RFE), hyperparameter tuning, and variable importance analysis. Model performance was evaluated based on accuracy, AUC, recall, and F1-score, with particular emphasis on the ability to detect the minority class ("non-compliant services"). The Random Forest model achieved the highest test accuracy (97.7%) with an average AUC of 0.655. However, the high accuracy should be interpreted cautiously, given the highly imbalanced class distribution (96.1% compliant vs 3.9% non-compliant), which limits the model's ability to detect minority cases. The analysis, therefore, emphasized recall and F1-score for the minority class as indicators of the model's practical usefulness for early detection. Five key predictors were identified, with anti-TB drug regimen and age contributing most to model performance. The model shows potential as an early warning tool for substandard TB services and can support risk-based interventions. Integration into national health information systems may strengthen decision-making and accelerate TB elimination.

Keywords: Machine learning, Tuberculosis, Indonesia

INTRODUCTION

Tuberculosis (TB) remains one of the deadliest communicable diseases worldwide. According to the World Health Organization (WHO), there were an estimated 10.8 million TB cases globally in 2023, resulting in approximately 1.25 million deaths¹. Indonesia, with the second-highest TB burden, faces challenges in reducing incidence and mortality. Achieving national elimination targets requires improved TB service quality². Within the framework of the Sustainable Development Goals (SDGs), Indonesia aims to reduce TB incidence to 65/100,000 and deaths by 90% from 2019 levels³. Achieving this target requires well-structured and focused efforts, particularly in improving the quality of TB services consistently across all healthcare facilities.

Tuberculosis control in Indonesia faces persistent challenges due to variations in the quality of TB service implementation across health facilities. Studies have reported disparities between public and private providers in diagnostic accuracy, treatment

management, and adherence monitoring. These inconsistencies may lead to delayed diagnosis, inappropriate treatment regimens, and weakened supervision, ultimately increasing the risk of treatment failure and continued community transmission^{4, 5, 6, 3, 7, 8, 9, 10, 11, 12, 13}. Non-compliant TB services not only negatively affect individual clinical outcomes but also increase the risk of community transmission and the emergence of drug-resistant TB cases^{14, 10}.

Current TB service monitoring relies on conventional manual supervision and periodic audits, which are reactive and resource-intensive. With over 30 public health centers and numerous private facilities, conventional methods fail to provide timely responses to service quality problems^{15, 4, 16, 17}. Delays in detecting facilities with substandard services may result in an accumulation of untreated cases and hinder the effectiveness of the city's TB control program^{18, 17}.

Non-standardized services increase the risk of treatment failure by up to 15-20% and prolong

the infectious period of patients^{11, 19}, ultimately amplifying the potential for community transmission. Therefore, a more proactive, accurate, and efficient approach is urgently needed to identify facilities or areas at high risk of providing substandard TB services.

Machine learning offers predictive capabilities to strengthen TB surveillance and identify substandard services early²⁰⁻²¹. In the context of TB control, the adoption of machine learning has the potential to shift the paradigm from reactive monitoring to predictive surveillance, with the capability to detect facilities or areas at high risk of providing substandard TB services at an early stage^{22, 23, 21}. The strength of these algorithms lies in their ability to simultaneously process multiple variables and identify complex interactions among factors, making them an ideal tool for analyzing the multifactorial determinants of TB service quality²⁴⁻²⁵.

While machine learning is increasingly applied in healthcare, its use for evaluating and predicting TB service quality remains limited, particularly in developing countries. Existing studies have primarily focused on diagnosis or treatment outcomes^{26,25}, leaving the quality assessment of TB services underexplored. The absence of predictive tools to identify substandard care impedes targeted interventions and efficient resource allocation.

This study aims to develop and evaluate a machine learning-based model to classify TB service compliance with national standards and establish an early warning system to support TB control in Semarang City. By identifying key risk indicators, the model seeks to inform targeted public health interventions. Designed for integration into existing health information systems, it supports evidence-based decision-making, enhances resource allocation, and promotes equitable, high-quality TB services. This approach aligns with the WHO End TB Strategy, which advocates for patient- and people-centered care in advancing global TB elimination²⁷.

METHODS

Study Design

This study employed an observational retrospective design with a quantitative-predictive approach. The primary objective was to develop and evaluate a machine learning (ML)-based predictive model to classify the compliance of tuberculosis (TB) services with the applicable national standards.

Data Sources and Data Collection

This study utilized both primary and secondary data collected in 2023. Secondary data were

sourced from the Semarang City Health Office, including TB program reports and electronic systems such as SITB and SEMAR BETUL, which provide detailed information on TB cases and vulnerability distribution. Primary data were obtained through structured digital interviews with 1,722 TB patients from community health centers (Puskesmas) across Semarang City, using KoboToolbox. The unit of analysis was individual TB patients.

Variable Definitions

The outcome variable in this study was TB service compliance with national standards, categorized as “standard-compliant” and “non-compliant.” In this context, TB service compliance refers to the extent to which healthcare facilities adhere to key components of the Indonesian National TB Program guidelines, including appropriate diagnostic procedures, correct prescription of anti-TB drug regimens, accurate patient classification, and systematic adherence monitoring in accordance with national standards. Predictor variables included patient characteristics such as age, sex, history of diabetes mellitus, and HIV status, as well as service-related characteristics, including the type and ownership of health facilities, type of TB diagnosis, anatomical site classification, anti-TB drug regimen, and source of medication funding.

Data Preprocessing

All data were recoded and cleaned of missing values (No significant missing values were detected in the final dataset). Categorical variables were transformed into numeric format using one-hot encoding techniques. The data were not standardized because all models employed in this study are tree-based or non-parametric, which are not sensitive to scale differences.

Feature Selection

Recursive Feature Elimination (RFE) based on Random Forest importance was used to reduce 13 predictors to five key variables: Age, Anti-TB Drug Regimen, Facility Ownership, Diagnosis Type, and Anatomical Classification, based on their contribution to model accuracy and node purity.

Cross-Validation

A 5-fold stratified cross-validation was applied to all algorithms. This technique divided the training data into five subsets proportionate to the class distribution, then averaged the evaluation metrics. Cross-validation aimed to assess model performance robustly and minimize the risk of overfitting^{28,29}.

Hyperparameter Tuning

Manual hyperparameter tuning was applied to optimize model performance, with Random Forest achieving the best balance between

accuracy and minority class detection. Boosting showed high AUC but failed in minority recall.

Machine Learning Algorithms

Three classification algorithms were selected for their complementary strengths in handling tabular data and interpretability. Random Forest was preferred for its robustness with mixed data types, resistance to overfitting, and interpretable variable importance³⁰⁻³¹, optimized using 41 trees and three predictors per split. K-Nearest Neighbors (KNN), a non-parametric method capturing local data structures, performed best with seven neighbors, Euclidean distance, and rectangular weights³²⁻³³. Boosting (Gradient Boosting) was selected for its iterative learning and variable influence analysis, optimized using 14 trees and 0.1 shrinkage rate based on out-of-bag accuracy³⁴⁻³⁵.

Model Evaluation

Model performance was assessed using accuracy, precision, recall, F1-score, and AUC per class. Due to class imbalance, emphasis was placed on detecting the minority ('non-compliant') class, given its clinical and epidemiological relevance. A confusion matrix and ROC curves were used to examine misclassification and sensitivity-specificity trade-offs³⁶⁻³⁷. Variable importance analysis identified key contributors to model decisions.

Ethical Considerations

This study used anonymized aggregate data without any personally identifiable information. Data processing followed data privacy and security standards. Ethical approval was obtained from the Health Research Ethics Committee, Faculty of Health Sciences, Dian Nuswantoro University (No. 573/EA/KEPK-FKes-UDINUS/V/2023).

Handling Class Imbalance

In this study, the distribution of TB service compliance was highly imbalanced, with 96.1% of cases classified as "standard-compliant" and only 3.9% as "non-compliant." No resampling techniques (e.g., oversampling, undersampling, or SMOTE) were applied, as the study aimed to preserve the real-world distribution observed in routine health system data collected directly from TB patients at health facilities. Maintaining the original class distribution was considered important to ensure external validity and to preserve the model's representativeness of the target population. Nevertheless, the class imbalance was explicitly considered during model evaluation by emphasizing performance metrics sensitive to minority class detection, particularly recall and F1-score, rather than relying solely on overall accuracy.

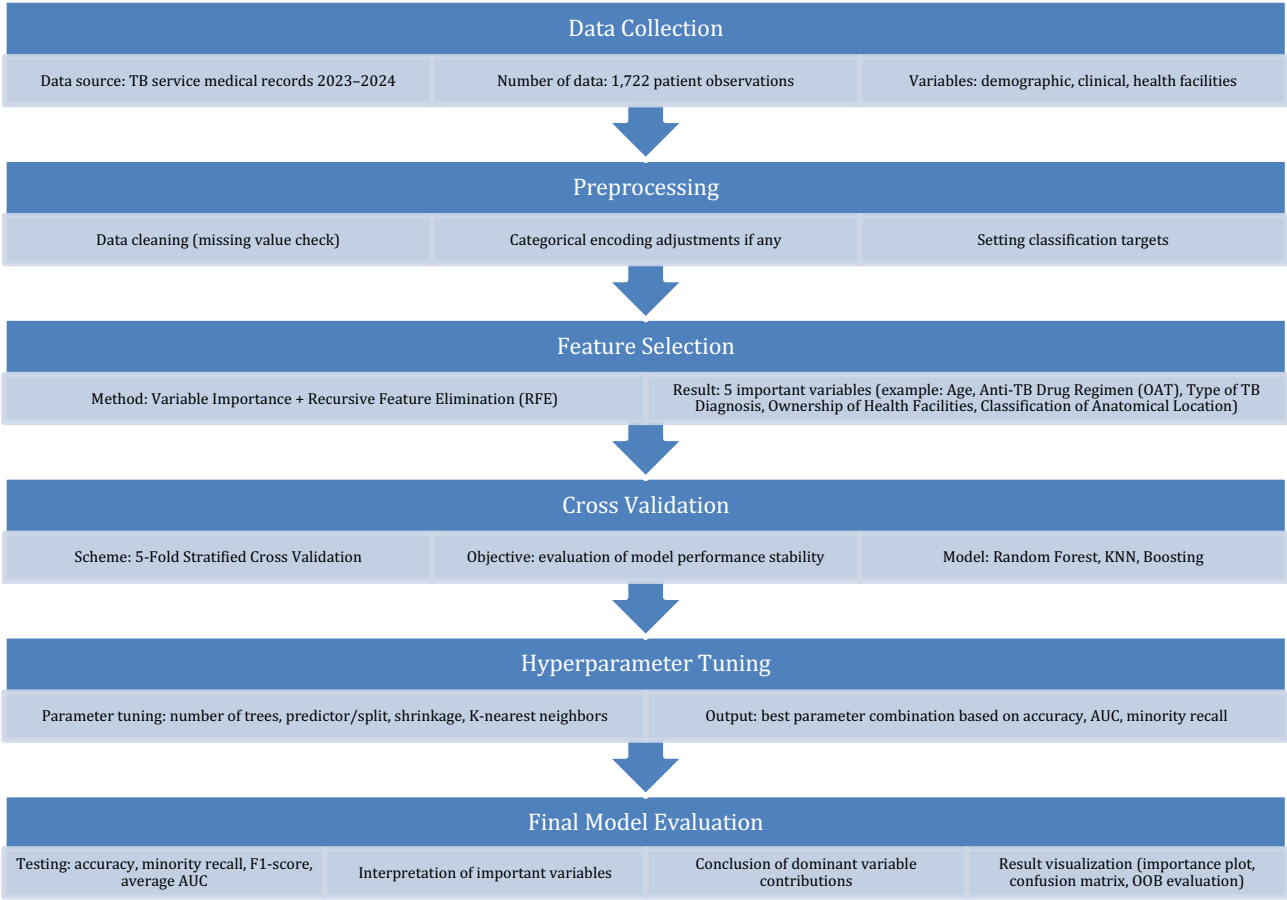


Figure 1. Data Analysis Workflow for the Classification of Tuberculosis Service Quality

RESULTS

Data Characteristics and Class Distribution

Analysis of 1,722 observations revealed a highly imbalanced class distribution, with 96.1% of cases classified as “standard-compliant” and only 3.9% as “non-compliant.” This distribution was consistent across all data subsets, indicating that the majority of TB services adhered to the established standards, although a small proportion of cases still requires special attention.

Feature Selection

To improve model efficiency and accuracy, Recursive Feature Elimination (RFE) was applied using Random Forest-based variable importance metrics. Features with low contributions—based on Mean Decrease in Accuracy (MDA) and Total Increase in Node Purity (TINP)—were iteratively removed.

From 13 initial features, 8 were eliminated, resulting in a final model using five key predictors: (1) Age, (2) Anti-TB Drug Regimen (OAT), (3) Health Facility Ownership, (4) Type of TB Diagnosis, and (5) Anatomical Classification. The refined model retained predictive performance and improved interpretability.

Cross-Validation

Cross-validation showed high accuracy across algorithms: Random Forest (96.2%), Boosting (95.0%), and KNN (94.5%). KNN achieved the highest discriminatory power (AUC = 0.716) despite overall classification challenges with the minority class. All models showed high accuracy; KNN had the highest AUC. Despite strong performance, minority class detection remained low.

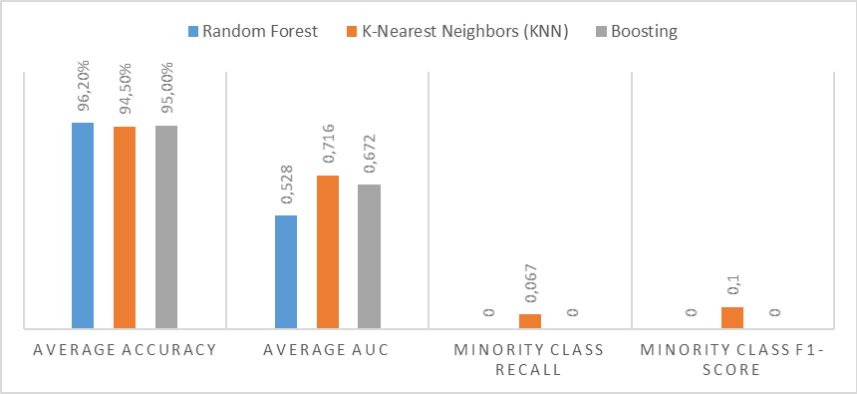


Figure 2. Summary of 5-Fold Cross-Validation Results (Test Data)

Hyperparameter Tuning

The results of hyperparameter tuning for the two algorithms are presented in the following table 1. Random Forest showed the most balanced performance. Although the highest accuracy (0.975) was achieved with 200 trees, the best recall for the minority class (0.231) was obtained using a lighter model with 50 trees and 2 predictors per split, suggesting improved sensitivity through simplified tuning. For KNN, minority class detection declined with larger k values and the use of Manhattan distance. The

best configuration—k = 3 with Euclidean distance—yielded a recall of 0.063 and F1-score of 0.100, substantially lower than Random Forest. Boosting was excluded from the main results due to consistently zero recall for the minority class, despite relatively high AUC values (up to 0.836), indicating strong overall discrimination but poor sensitivity to minority class patterns.

Table 1. Results of Hyperparameter Tuning

Algorithm	Main Parameters	Accuracy	Recall (Minority)	AUC	F1-Score (Minority)
Random Forest	Trees=50, Pred=2, Train=50%	0.972	0.231	0.703	- (na)
Random Forest	Trees=100, Pred=3, Train=70%	0.966	0.143	0.782	-
Random Forest	Trees=200, Pred=4, Train=80%	0.975	0.100	0.760	-
KNN	K=3, Euclidean, Rectangular	0.948	0.063	0.566	0.100
KNN	K=5, Euclidean, Rectangular	0.909	0.000	0.723	0.000
KNN	K=3, Manhattan, Rectangular	0.956	0.000	0.579	0.000

Model Interpretability (Variable Importance)

To assess each predictor's contribution to model performance, a variable importance analysis was conducted using the Random Forest algorithm. Mean Decrease in Accuracy and Total Increase in Node Purity (Figure 3) were used to identify features with the greatest influence on model splitting decisions. Anti-TB Drug Regimen, Age, and Diagnosis Type were the most influential features in the model.

Final Model Evaluation

A final evaluation was conducted using the best-performing Random Forest model based on hyperparameter tuning, with a configuration of 50 decision trees and 2 predictors per split, using 50% of the training data. This model was applied to the test dataset comprising 344 observations. The final model achieved 97.7% accuracy, though recall for the minority class remained zero, reflecting imbalanced class issues. Although the model achieved high overall accuracy, this performance is largely influenced by the dominant majority class. Consequently, the model showed limited ability to correctly classify the minority “non-compliant” cases, as reflected by the low recall and F1-score.

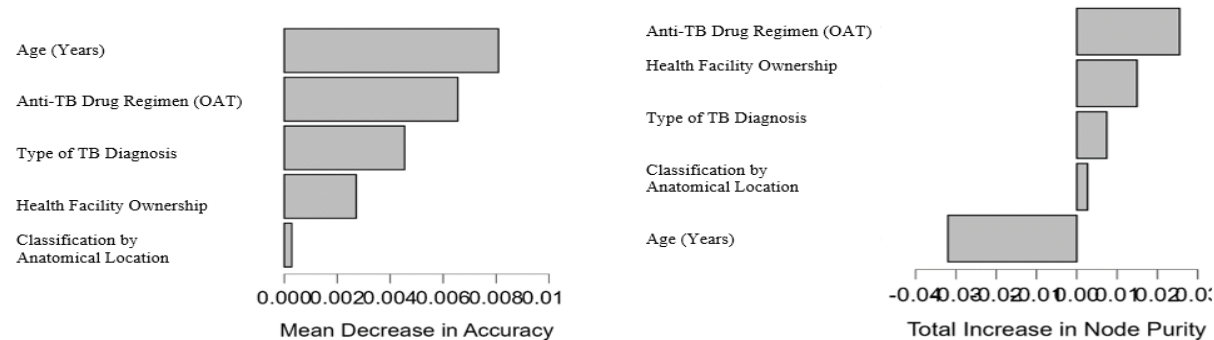


Figure 3. Visualization of variable importance based on Mean Decrease in Accuracy and Total Increase in Node Purity

Table 2. Final Model Performance Evaluation

Metric	Value
Test Accuracy	0,977
Average AUC	0,655
Recall (Minority Class)	0,000
F1-Score (Minority Class)	NaN (0)
Out-of-Bag Accuracy	0,133
Model Configuration	50 trees, 2 predictors per split

Despite high overall performance, the model failed to classify the 'non-compliant' (minority) class in the test data, reflected by undefined (NaN) recall and F1-scores due to zero positive predictions and only 8 minority cases out of 344. Nonetheless, the AUC for this class reached 0.660, indicating moderate discriminatory ability.

Variable importance analysis identified five key features:

1. Anti-TB Drug Regimen (MDA = 0.012; Node Purity = 0.076),
2. Age (MDA = 0.013; Node Purity = -0.009),
3. Type of TB Diagnosis,
4. Health Facility Ownership,
5. Anatomical Classification.

Overall, the model reliably classifies compliant TB services but remains limited in detecting non-compliance.

DISCUSSION

Interpretation of Key Findings

Despite high overall accuracy, the models consistently underperformed in detecting the minority class—a common issue in imbalanced health datasets^{43, 44}. Cross-validation and hyperparameter tuning yielded low or zero recall and F1-scores for minority cases, reflecting a well-known limitation in which machine learning models tend to favor majority-class performance at the expense of minority-class detection⁴⁰⁻⁴³. The models’ inability to accurately detect the minority class highlights a broader challenge in healthcare datasets, where rare events are often underrepresented. In the context of TB service monitoring, non-compliant services represent a relatively small but critical subset of cases. Although the model struggled to directly classify these cases, the identified predictors still

provide valuable signals for risk-based supervision⁴⁰

Despite its limited sensitivity to minority cases, the model remains useful as an early risk-screening tool. By identifying patterns associated with service compliance, health authorities can prioritize supervision for facilities with similar profiles, enabling more efficient allocation of monitoring resources⁴⁴. Cross-validation and hyperparameter tuning yielded low or zero recall and F1-scores for minority cases, reflecting a well-known limitation where models favor majority class performance at the expense of minority class detection.

The five key predictors—Anti-TB Drug Regimen (OAT), facility ownership, type of diagnosis, anatomical classification, and patient age—underscore the influence of structural, clinical, and demographic factors on service quality. Patient age may influence TB service outcomes because older patients often present with comorbidities, delayed diagnosis, or more complex clinical management, which may affect adherence to treatment standards and monitoring procedures⁴⁵. The type of TB diagnosis also reflects differences in diagnostic pathways and facility capacity, as variations in the use of bacteriological confirmation or clinical diagnosis may indicate disparities in adherence to national diagnostic guidelines⁴⁶. In addition, the anatomical classification of TB (pulmonary versus extrapulmonary) can influence treatment complexity and follow-up requirements, potentially affecting the consistency with which TB service standards are implemented in healthcare facilities⁴⁶.

Anti-TB Drug Regimen and facility ownership were especially prominent, aligning with prior studies on treatment success and supervision. These findings are consistent with the WHO Global Tuberculosis Report 2022, which highlights ongoing post-COVID-19 challenges, including the need to strengthen DOTS strategies⁴⁷. Additionally, Parwati et al. (2021) emphasize that combining direct observation with patient empowerment models, such as the Health Belief Model and Motivational Interviewing, can further enhance treatment outcomes⁴⁸.

Health facility ownership also emerged as an important determinant, supporting studies by Lestari and Dewi which showed variations in compliance with TB management standards between public and private facilities in Indonesia^{46, 47}. Although Partial Dependence Plots or SHAP values were not produced in this study, the variable importance plots still provide relevant insights for evidence-based policy formulation. Improving service quality requires stronger supervision and the integration of predictive tools into public health systems^{48, 49}. The Random Forest model showed stable performance and reliable OOB accuracy. While

hyperparameter tuning preserved overall accuracy, consistent with Hossain & Timmer (2021) and Pfob et al. (2022), challenges in detecting the minority class remain, indicating that tuning alone cannot fully resolve class imbalance bias^{50, 51}.

These findings demonstrate the feasibility of automating TB service quality evaluation using data-driven predictive models. The proposed machine learning approach supports integration into health information systems for precision public health and risk-based monitoring, aligning with efforts to enhance surveillance sensitivity (Ahmad et al.) and advance IoT-based risk frameworks for TB elimination^{52, 53}.

Relevance to Previous Studies

Results confirm Random Forest reliability for healthcare data with heterogeneous predictors⁵⁷. Identified predictors align with previous studies highlighting socioeconomic and health system factors influencing TB outcomes^{58, 59}. The persistent emergence of the Anti-TB Drug Regimen (OAT) as a main predictor aligns with the findings of^{57, 58} which demonstrate that patient adherence to OAT regimens is key to achieving treatment success, preventing drug resistance, and maintaining TB service quality through strategies such as DOTS, PMO supervision, and patient-centered support.

Implications for TB Prevention and Control

This study advances TB control by identifying key predictors—anti-TB drug regimen, facility ownership, and patient age—that inform targeted interventions at the primary care level. Aligned with the WHO End TB Strategy, the model highlights private facilities as higher-risk settings and supports early detection of vulnerable patients to improve prevention, contact tracing, and treatment outcomes²⁷.

Contribution to the Public Health System

By leveraging predictive analytics, this study enhances public health information systems and offers actionable insights for integrating machine learning models—such as Random Forest—into HMIS. These models can strengthen early warning systems and inform data-driven TB control efforts, guiding policymakers in implementing cross-sectoral strategies involving local authorities, professional bodies, and private healthcare regulators.

Limitations and Future Research Directions

Key limitations include lack of data balancing techniques leading to poor minority class detection, and absence of advanced interpretability methods. External validation is needed for generalizability assessment.

Practical Recommendations

Key recommendations include:

- Strengthen supervision and audits, especially in high-risk private facilities.
- Use the classification model as a decision-support tool at local TB program levels.
- Train data managers to apply key quality indicators effectively.
- Improve interoperability for integration into the national TB surveillance system.

CONCLUSIONS

This study developed a machine learning model to classify the compliance of TB services with national standards in Semarang City. Random Forest achieved the most stable overall performance, although all models showed limited ability to identify the small number of non-compliant cases due to the highly imbalanced dataset. Despite these challenges, the analysis consistently identified five significant predictors: the anti-TB drug regimen, facility ownership, diagnosis type, anatomical classification, and patient age. These findings demonstrate that machine learning can complement routine supervision by identifying facilities that may require closer attention. Although further refinement is needed to improve sensitivity to the minority class, the model offers an early foundation for incorporating predictive analytics into TB information systems. With stronger and more balanced data, this approach has the potential to support more timely and targeted TB control efforts.

Acknowledgements

This research was supported by the KEMENDIKBUDRISTEK Kedaireka 2023 and Health Department of Semarang City.

REFERENCES

1. World Health Organization. Global tuberculosis report 2024 [Internet]. 2024. Available from: <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024>
2. Kementerian Kesehatan Republik Indonesia. Profil Kesehatan Indonesia 2023 [Internet]. Jakarta; 2024. Available from: https://kemkes.go.id/app_asset/file_content_download/172231123666a86244b83fd8.51637104.pdf
3. Kementerian Kesehatan Republik Indonesia. Strategi nasional penanggulangan tuberkulosis di Indonesia 2020-2024. Vol. 135, Pertemuan

Konsolidasi Nasional Penyusunan STRANAS TB. 2020.

4. Handayani S, Isworo S. Evaluation of Tuberculosis program implementation in Primary Health Care, Semarang, Indonesia. *Int J Public Heal Asia Pacific*. 2024;1-11.
5. Rahayu SR, Susilastuti MS, Saefurrohim MZ, Azam M, Indrawati F, Supriyono M, et al. Lost to Follow-Up among Tuberculosis Patients during the Public-Private Mix Era in Rural Area of Indonesia. *Ethiop J Health Sci*. 2023;33(1).
6. Dinas Kesehatan Kota Semarang. Profil Kesehatan Kota Semarang 2023 [Internet]. Semarang; 2023. Available from: <https://pustakadata.semarangkota.go.id/upload/pdf/451-profil-kesehatan-2022.pdf>
7. Kementerian Kesehatan Republik Indonesia. Pedoman Nasional Pelayanan Kedokteran: Tatalaksana Tuberkulosis. Jakarta; 2020.
8. Ahmad RA, Matthys F, Dwiwardiani B, Rintiswati N, de Vlas SJ, Mahendradhata Y, et al. Diagnostic work-up and loss of tuberculosis suspects in Jogjakarta, Indonesia. *BMC Public Health*. 2012;12:1-6.
9. Lestari BW, McAllister S, Hadisoemarto PF, Afifah N, Jani ID, Murray M, et al. Patient pathways and delays to diagnosis and treatment of tuberculosis in an urban setting in Indonesia. *Lancet Reg Heal Pacific*. 2020;5.
10. Lestari BW, Nijman G, Larasmanah A, Soeroto AY, Santoso P, Alisjahbana B, et al. Management of drug-resistant tuberculosis in Indonesia: a four-year cascade of care analysis. *Lancet Reg Heal Asia*. 2024;22.
11. Sugiyono RI, Naysilla AM, Susanto NH, Handayani D, Burhan E, Karuniawati A, et al. Treatment outcomes of pulmonary TB in adults in Indonesia. *IJTLD open*. 2025;2(3):145-52.
12. Surya A, Setyaningsih B, Suryani Nasution H, Gita Parwati C, Yuzwar YE, Osberg M, et al. Quality tuberculosis Care in Indonesia: using patient pathway analysis to optimize public-private collaboration. *J Infect Dis*. 2017;216(suppl_7):S724-32.

13. Fuady A, Houweling TAJ, Mansyur M, Burhan E, Richardus JH. Cost of seeking care for tuberculosis since the implementation of universal health coverage in Indonesia. *BMC Health Serv Res.* 2020;20:1-10.
14. Udwadia Z, Furin J. Quality of drug-resistant tuberculosis care: Gaps and solutions. *J Clin Tuberc other Mycobact Dis.* 2019;16:100101.
15. Iswandari HD, Sugiharto S, Harsono H. The evaluation of pulmonary tuberculosis (TB) control program at Bandarharjo Community Health Center Semarang city in 2021. In: *AIP Conference Proceedings.* AIP Publishing; 2023.
16. Tuharea R, Suparwati A, Sariatmi A. Analysis on Factors Associated with the Implementation of the Pulmonary Tuberculosis Care Finding in Tuberculosis Control Program at Primary Healthcare Center in Semarang. *J Manaj Kesehat Indones.* 2(2).
17. Paskaria C, Sunjaya DK, Parwati I, Bestari MB. Digital Health to Strengthen District-Based Public-Private Mix Tuberculosis Control in Purwakarta District, Indonesia: A Qualitative Research. *Int J Environ Res Public Health.* 2022;19(19):12097
18. Reviono R, Setianingsih W, Damayanti KE, Ekasari R. The dynamic of tuberculosis case finding in the era of the public-private mix strategy for tuberculosis control in Central Java, Indonesia. *Glob Health Action.* 2017;10(1):1353777.
19. Fuady A, Houweling TAJ, Mansyur M, Burhan E, Richardus JH. Catastrophic costs due to tuberculosis worsen treatment outcomes: a prospective cohort study in Indonesia. *Trans R Soc Trop Med Hyg.* 2020;114(9):666-73.
20. Hansun S, Argha A, Bakhshayeshi I, Wicaksana A, Alinejad-Rokny H, Fox GJ, et al. Diagnostic Performance of Artificial Intelligence-Based Methods for Tuberculosis Detection: Systematic Review. *J Med Internet Res.* 2025;27:e69068.
21. KB HM, Jose SA, Jirawattanapanit A, Mathew K. A comprehensive study on tuberculosis prediction models: Integrating machine learning into epidemiological analysis. *J Theor Biol.* 2025;597:111988.
22. Gadicherla S, Krishnappa L, Madhuri B, Mitra SG, Ramaprasad A, Seevan R, et al. Envisioning a learning surveillance system for tuberculosis. *PLoS One.* 2020;15(12):e0243610.
23. Chen J, Jiang Y, Li Z, Zhang M, Liu L, Li A, et al. Predictive machine learning models for anticipating loss to follow-up in tuberculosis patients throughout anti-TB treatment journey. *Sci Rep.* 2024;14(1):24685.
24. Peng A zhou, Kong XH, Liu S tao, Zhang H fen, Xie L ling, Ma L juan, et al. Explainable machine learning for early predicting treatment failure risk among patients with TB-diabetes comorbidity. *Sci Rep.* 2024;14(1):6814.
25. Ahamed Fayaz S, Babu L, Paridayal L, Vasantha M, Paramasivam P, Sundarakumar K, et al. Machine learning algorithms to predict treatment success for patients with pulmonary tuberculosis. *PLoS One.* 2024;19(10):e0309151.