**Jurnal Teknik Informatika (JUTIF)** Vol. 6, No. 2, April 2025, Page. 891-904
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

# Enhancing Prediction of Treatment Duration in New Tuberculosis Cases: A Comprehensive Approach with Ensemble Methods and Medication Adherence

**Rusdah*[1], Painem[2], and Dewi Kusumaningsih[3]**

[1]Master of Computer Science, Universitas Budi Luhur, Indonesia
[2,3]Information System, Universitas Budi Luhur, Indonesia

Email: [1]rusdah@budiluhur.ac.id

Received: Dec 28, 2024; Revised: Mar 1, 2025; Accepted: Mar 3, 2025; Published: Apr 26, 2025

**Abstract**

Tuberculosis (TB) remains a significant global health problem, with treatment duration varying among patients. TB patients have difficulty following a long-term treatment regimen. After the final diagnosis is determined, it is necessary to know the predicted duration of treatment for a patient. By increasing patient compliance with taking medication, the percentage of TB patients will increase, and this can reduce cases of multi-drug resistant patients and dropouts. This study aims to build a prediction model for the duration of treatment for new cases of Pulmonary TB patients by adding medication compliance parameters using the ensemble method. The research methodology uses CRISP-DM. This study begins with identifying problems and objectives, collecting data, preprocessing and analyzing data, modeling, evaluating, and validating models. The results showed that adding medication compliance parameters can improve model performance. However, the results of model exploration with feature selection techniques and various ensemble methods have not shown good performance. The medication adherence parameters used in this study are the number of medications swallowed in Phase I and Anti-Tuberculosis drug compliance in Phase I. These parameters had never been used in previous studies. The prediction model can be used as an early warning for a patient. If a patient is predicted to have a treatment duration of more than six months, then the patient will receive stricter drug intake supervision. Thus, this proposed model is expected to help achieve the target of eliminating Tuberculosis in 2030 to reduce the death rate by 90% compared to 2019.

***Keywords:*** *Ensemble Method, Machine Learning, Prediction Model of Treatment Period, Prognosis of TB, Treatment Duration of TB Patients*.



## 1. INTRODUCTION

Indonesia is among eight countries that account for 2/3 of Tuberculosis (TB) cases worldwide. Indonesia is second only to India, with 845,000 cases and 98,000 deaths, equivalent to 11 deaths/hour [1]. In the National Strategy for Tuberculosis Control in Indonesia for 2020-2024, the government has a target to reduce the number of deaths from TB to 27 per 100,000 population. One of the targets for Tuberculosis elimination by 2030 is to reduce the mortality rate by 90% compared to 2019 [2].

TB patients have difficulty following a long-term treatment regimen. Improving treatment outcomes requires a better understanding of adherence as a complex behavioral issue and the particular barriers to and facilitators of patient adherence [3]. Loss to follow-up (LTFU) is a significant barrier to the completion of anti-tuberculosis treatment and a major predictor of TB-associated deaths [4]. According to the National Guidelines for Tuberculosis Control in Indonesia, patient detection begins with detecting suspected TB, followed by physical and laboratory examinations, diagnosis, and type of TB patient [5]. Once a final diagnosis is made, it is necessary to know the predicted length of treatment

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

for a patient. The length of treatment that a person diagnosed with TB must undergo is at least six months. Suppose it is known early on that a patient's predicted length of treatment exceeds six months. In that case, the patient will need closer supervision from the family (Medication Monitoring Officer) or the Directly Observed Treatment Shortcourse (DOTS) at the health service where the patient is treated. Unfortunately, patient adherence to medications, particularly for conditions requiring prolonged treatment such as TB, is frequently less than ideal and can result in poor treatment outcomes [6]. Increasing a patient's medication adherence will increase the percentage of TB patients cured. On the other hand, it can reduce the number of Multi-Drug Resistance (MDR) patients and fail or Drop-Out cases [2]. The variables of education, age, income, knowledge, medication adherence, availability of medication supervisors, and counseling influence the treatment success rate [7]. Risk factors for loss to follow-up include the distance between the patient and the treatment center, side effects, gender, family support, motivation, knowledge, beliefs, economic status, and stigma, myths/beliefs. There is a need for involvement and support from multiple sectors to align policies, as well as the importance of patient-centered treatment and supervision to strengthen compliance in treatment [8].

Researchers have conducted several previous studies on developing prognosis models for TB patients' treatment length. The first study, conducted in 2022, used data on patients declared cured or completed treatment between 2010 and 2014. The results of the single classifier comparison showed that the best performance was C4.5. Unfortunately, the performance of the resulting model was still low, with 74% accuracy and 76% specificity [9].

On the other hand, a study to determine the determinants of the length of time to cure category I tuberculosis in Semarang Regency showed that the variable of treatment regularity was not associated with the length of time to cure [10]. This differs from the National Strategy for Tuberculosis Control in Indonesia 2020-2024 [2]. Thus, the second study was conducted by adding the parameter of drug compliance and using a multi-classifier, namely Gradient Boosting [11]. The results showed that the drug compliance parameter could improve the model's accuracy. However, the model's accuracy is 79.6% with an AUC of 62.5, which is still included in the poor classification category [12].

This study will do several ways to solve the problem of low performance of the prediction model for the length of treatment of pulmonary TB patients, first, in terms of datasets. This study will focus on new TB patients, both compliant and non-compliant cases. As suggested by previous research [11]. Second, in terms of algorithms. The previous study used a gradient-boosting multi-classifier [11] that works on numerical data, while the dataset is categorical. So, this research will explore ensemble algorithms that can work on categorical data. Ensemble development will combine several single classifiers to create an ensemble model with the best performance.

Research has been conducted to build a prognosis model for the treatment period of tuberculosis patients using a single classifier [9]. The data is the data of patients who were declared cured or completed treatment in 2010-2014. The data was classified into four classes, namely 0-6 months, >6-9 months, >9-12 months and >12 months. The class composition of the dataset is unbalanced, where 74.5% of the data is patient data whose treatment period is more than six months to 9 months. The algorithms explored are C4.5, KNN, Naïve Bayes, and SVM. The comparison results show that the best performance is C4.5. Unfortunately, the performance of the resulting model was still low, with 74% accuracy and 76% specificity [9].

Meanwhile, a study to determine the determinants of the length of time to cure category I tuberculosis in Semarang Regency showed that the variable of treatment regularity was not associated with the length of time to cure [10]. On the other hand, a study found that factors of patient compliance in taking medication are patient knowledge, respondent attitude, distance to the health center, officer attitude, [13] and family motivation [13], [14]. The previous study was conducted by adding the medication adherence parameter (number of medications swallowed in stage I, Anti-Tuberculosis Drug

Compliance I, number of medications swallowed in stage II, and Anti-Tuberculosis Drug Compliance II) and using a multi-classifier, Gradient Boosting [11]. The results showed that the drug compliance parameter could improve the model's accuracy. However, the model's accuracy is 79.6% with an AUC of 62.5, which is still included in the poor classification category [12].

The ensemble method has performed better [15], [16]. Research related to the use of ensemble methods in the Tuberculosis disease domain has been conducted by several researchers. Prediction of the likelihood of death by TB to help TB prognosis and treatment decision-making process using balanced and unbalanced datasets from Brazil [17]. The best model for predicting the recovery class resulted from an ensemble model formed from Random Forest, Gradient Booster, and Multi-Layer Perceptron. The Gradient Boosting model was best for predicting the TB mortality class. The performance of four feature selection techniques was compared: Sequential Forward Selection, Sequential Forward Floating Selection, Sequential Backward Selection, and Sequential Backward Floating Selection [17]. In addition, research has been conducted for the prognosis of pathogen drug resistance in pulmonary tuberculosis with Catboost [18]. The dataset has drug-sensitive and drug-resistance classes. Research on predictive TB diagnosis has also proven that ensemble techniques can perform better than single classifiers [17], [19]–[23]. Likewise, data in the form of chest X-ray images that use ensemble learning to detect TB [21], multi-omic data using an ensemble of deep learning and machine learning models [24], and imbalanced datasets in clinical data [25]. The study aimed to prognosticate loss to follow-up (LTFU), exploring SVM, RF, XGBoost, and logistic regression. The result showed that the XGBoost model provides the highest predictive performance in identifying LTFU risk among tuberculosis patients [26].

This study aims to develop a prediction model for the duration of treatment for new cases of Pulmonary TB patients by adding medication compliance parameters using the ensemble method. The novelty is in predicting the length of treatment time for new cases of pulmonary TB patients, both for Smear-Negative and Smear-Positive Pulmonary Tuberculosis, with the addition of drug compliance parameters, using the development of ensemble methods. Ensemble methods have not been widely used for categorical datasets with compliance parameters. The medication adherence parameters used in this study are the number of medications swallowed in Phase I and Anti-Tuberculosis drug compliance in Phase I. The prediction model can be used as an early warning for a patient.

## 2. METHOD

This research uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, as presented in Figure 1. CRISP-DM has six stages; the first stage, Business Understanding, has already been completed. At this stage, a literature study was conducted to find research gaps in predicting the treatment period of pulmonary tuberculosis patients. After that, an interview was conducted with the Management of the Jakarta Respiratory Center at the Indonesian Tuberculosis Disease Eradication Association (JRC-PPTI). The aim is to identify the research problems and objectives. The output of this stage is the formulation of research problems and objectives. The following activities to be carried out are described below.

**Jurnal Teknik Informatika (JUTIF)** Vol. 6, No. 2, April 2025, Page. 891-904
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

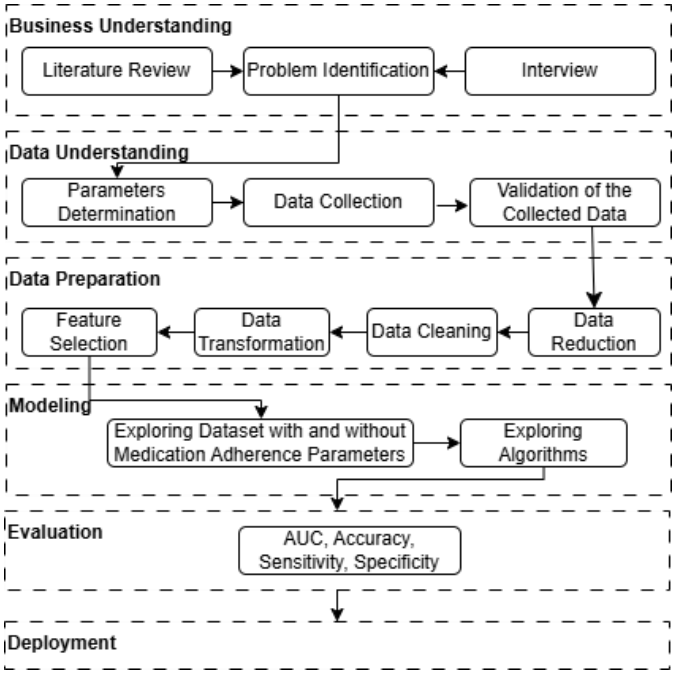

Figure 1. Research Methodology

### 2.1. Data Understanding

At this stage, the collection and validation of the collected data is carried out. Data was taken from the Jakarta Respiratory Center at *Perkumpulan Pemberantasan Tuberkulosis Indonesia* (JRC-PPTI) South Jakarta. The dataset used is the Medical Record Data of new cases of Pulmonary TB Patients, including the type of Smear-Positive Pulmonary Tuberculosis (SPPT) and Smear-Negative Pulmonary Tuberculosis (SNPT), with a minimum age category of 15 years, and have been declared cured (for SPPT) or completed treatment (for SNPT) and the addition of medication compliance parameters, both cases of compliant and non-compliant patients. These parameters were determined with guidance from a doctor. After the data were collected, a doctor validated the data.

### 2.2. Data Preparation

At this stage, the dataset will be prepared by conducting data reduction, cleaning, transformation, and feature selection using the wrapper method, filter method, and information gain. The output of this stage is a dataset ready for the modeling stage.

### 2.3. Modeling

In the modeling phase, datasets with and without medication adherence parameters were compared to prove which dataset performed better. Then, several multi-classifier algorithms (ensemble method) were explored to obtain the best-performing model. Those ensemble methods were Random Forest, Gradient Boosting, Voting, and Stacking using C4.5, Support Vector Machine (SVM), and Neural Network (NN). The output at this stage is the Prediction Model of the length of treatment of TB patients with the highest AUC value.

### 2.4. Evaluation

In this study, metrics are used to compare the models: accuracy, sensitivity, specificity, and AUC ROC [17], [27]. To understand these metrics, it is essential to define the composition of a confusion matrix: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy is

a performance metric that indicates how many samples were correctly classified about the whole: the ratio between the sum of TP and TN and the sum of all samples (1). Sensitivity (Recall) is an evaluation metric that describes how well a model correctly identifies positive classes (2). Specificity is an evaluation metric that shows how effective a model is in accurately classifying negative classes (3). The model with the highest sensitivity minimizes False Negatives. Receiver Operating Characteristic (ROC) curves are a popular way to visualize a binary classifier's tradeoffs between sensitivity and specificity. The ROC curve is plotted with sensitivity against its complement (1 − sensitivity) or False Positive Rate (FPR), where sensitivity is on the y-axis and FPR is on the x-axis. As the name suggests, the Area Under The Curve (AUC) ROC is the area underneath the entire ROC curve, representing the degree of separability between classes. The higher the AUC value, the better the model predicts class positive as class positive and class negative as class negative.

$$\boldsymbol{Accuracy} = \frac{(\boldsymbol{TP} + \boldsymbol{TN})}{(\boldsymbol{TP} + \boldsymbol{TN} + \boldsymbol{FP} + \boldsymbol{FN})} \quad (1)$$

$$\boldsymbol{Sensitivity} = \frac{\boldsymbol{TP}}{\boldsymbol{FN} + \boldsymbol{TP}} \quad (2)$$

$$\boldsymbol{Spesificity} = \frac{\boldsymbol{TN}}{\boldsymbol{TN} + \boldsymbol{FP}} \quad (3)$$

### 2.5. Deployment

A doctor validates the model at this stage to ensure the result is acceptable.

## 3. RESULT

### 3.1. Data Understanding

The medical record dataset was successfully collected from 2020 to 2023. The number of records in the medical record dataset is 779 records and 44 attributes. The attributes used in this study are under the doctor's instructions. The difference is that this study included patients with failed final results and loss to follow-up. For the medication adherence parameter, the attributes of Initial Date of Receipt of Phase I Medication, Final Date of Receipt of Phase I Medication, Number of Swallowed Medications in Phase I, Anti-Tuberculosis Drug Compliance Phase I, Initial Date of Receipt of Phase II Medication, Final Date of Receipt of Phase II Medication, Number of Swallowed Medications in Phase II, Anti-Tuberculosis Drug Compliance Phase II were used. After the data is collected, the doctor carries out data validation.

### 3.2. Data Preparation

After the doctor validates the data, pre-processing is carried out to produce a dataset ready to be modeled. This process includes data reduction, cleaning, transformation, and feature selection. The following is an explanation of each activity performed.

***Data Reduction.*** The reduction stage was performed by removing records and attributes (dimensionality reduction). First, 13 records of patients aged less than 15 years, four relapsed TB cases, and 23 former TB patients were deleted. Second, the attributes removed were Date of Registration, Village, Subdistrict, and Occupation because they were irrelevant to developing a prognosis model for the length of treatment of new Pulmonary TB patients. Furthermore, single-valued attributes were removed. The attributes MONTH 1 and Anti-Tuberculosis Drug 1 were removed because they only

contained the INITIAL FDC I. The attributes MONTH 3, Anti-Tuberculosis Drug 3, MONTH 4, and Anti-Tuberculosis Drug 4 were removed because they only contained CONTINUED FDC I. The result of this stage is a dataset with a single-valued attribute. The result of this stage was a dataset consisting of 34 attributes and 739 records. This deletion activity was carried out in consultation with the JRC-PPTI pulmonary doctor.

***Data Cleaning.*** At this stage, several missing values are handled. One patient data record does not have data on the Weight attribute. Then, it is filled with the average weight of all patients, which is 49.5 kgs. One missing value at the Sputum attribute in the 5th Month is filled with the value Not Performed. At the Cough Duration attribute, 50 missing value records are filled in with the value No. Five missing values are filled in in the Temperature attribute with a median of 37.

***Data Transformation.*** The data transformation activities carried out are a generalization of attribute values and the formation of new attributes. First, the Duration of Cough attribute was generalized into three values: No, <2 weeks, and >=2 weeks. This follows the Guidelines for Diagnosis and Management in Indonesia, which states that the symptom of Pulmonary TB is a cough with phlegm for two weeks or more [5]. Coughing can be followed by additional symptoms, namely sputum mixed with blood, coughing up blood, shortness of breath, weakness, decreased appetite, decreased weight, malaise, night sweats without physical activity, and fever for over a month [1]. In the Comorbidities attribute, data containing diseases other than diabetes mellitus, HIV, and COVID-19 were generalized to None. This was done because the comorbid diseases in Pulmonary TB are diabetes mellitus, HIV, and COVID-19. In general, the duration of treatment for uncomplicated and comorbid pulmonary TB is six months. In extra-pulmonary TB and TB with comorbidities, treatment can take more than six months [1]. In the Initial Diagnosis attribute, data containing BE, Bronchitis, Acute Bronchitis, Bronchopneumonia, Chest Pain, Dyspepsia, Pleural Effusion, Hemaptoe, Lymphadenitis, Pleuritis, Pneumonia, Pneumoniathorax, COPD, Lung Tumor are generalized to Non-TB. Miliary TB is generalized to Extra Pulmonary TB. BTA Positive TB and Pulmo TB are generalized into BTA Positive Pulmonary TB. The Formation of a new attribute, namely the Length of Month of Treatment, used the formula by reducing the end of treatment date attributed to the final diagnosis date by units of months. Then, based on the Length of Treatment Month attribute, the Length of Treatment attribute is formed by generalizing the contents of the Length of Treatment column into 0-6 months and >6 months (2). In contrast to previous studies that divided the Length of Treatment class into 4, namely 0-6 months, >6-9 months, >9-12 months, and >12 months [9].

After the data transformation process, there are two additional new attributes. So, at this stage, the total attributes are 36 attributes. Then, attribute reduction is carried out. The deleted attributes are the Date of Final Diagnosis and the date of End of Treatment because they have been transformed into the Length of Treatment Month. Then, the Length of Treatment Month attribute was deleted because it had been transformed into the Length of Treatment attribute. In addition, the attributes of Initial Date of Receipt of Stage I Medication, Final Date of Receipt of Stage I Medication, Initial Date of Receipt of Stage II Medication, and Final Date of Receipt of Stage II Medication were also deleted. So, at the end of this stage, the total attributes used were 22, namely Sex, Age, Duration of Cough, Phlegm, Blood, Fever >1 Month, Weight Loss, Loss of Appetite, Night Sweats, Shortness of Breath, Chest Pain, Comorbid, Ronchi, Temperature, Weight, Initial Diagnosis, Initial X-Ray, Initial Phlegm, Final Diagnosis, End of Intensive Phase, Number of Medications Swallowed Phase I, Anti-Tuberculosis Medication Compliance Phase I and Duration of Treatment.

***Feature Selection.*** At this stage, attribute selection is carried out to determine the correct number of attributes to produce the final diagnosis with the best accuracy. The attribute selection method used is the wrapper method. The wrapper method produces a selection of attributes that will produce the best accuracy when classified with a particular classifier. In this study, experiments were conducted with

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

multi-classifiers (ensemble method), Random Forest (Bagging), Gradient Boosting (Adaboost), Vote (C4.5, SVM, and NN), and Stacking (C4.5, SVM, and NN). In addition, filter and information gain methods were used. Validation is done with 10-fold cross-validation. Table 1 shows the experimental results of the feature selection techniques.

Table 1. The experimental results of the feature selection techniques

| Feature Selection Method | Number of attributes | Attributes |
|---|---|---|
| Wrapper Method – Baging (Random Forest) | 17 | Age, Weight, End of Intensive, Temperature, Anti-Tuberculosis Drug Compliance I, Number of Medications Swallowed in Phase I, Sex, Comorbid, Night Sweats, Phlegm, Initial X-ray, Initial Diagnosis, Fever >1 Month, Chest Pain, Shortness of Breath, Ronchi, Weight Loss |
| Wrapper Method – AdaBoost (Gradient Boosting) | 12 | Age, Weight, Cough with Phlegm, Night Sweats, Ronchi, Temperature, Initial Diagnosis, Gender, Shortness of Breath, Fever >1 Month, Comorbid, Chest Pain |
| Wrapper Method – Vote (C4.5, SVM, and NN) | 11 | Age, Weight, Anti-Tuberculosis Drug Compliance I, Temperature, Initial X-ray, Fever >1 Month, Gender, Ronchi, Cough with Blood, Duration of Cough, End of Intensive |
| Information Gain | 6 | Anti-Tuberculosis Drug Compliance I, Fever >1 Month, Number of Medications Swallowed in Phase I, End of Intensive, Cough with Phlegm, Initial Diagnosis |
| Filter Method | 21 | Number of Medications Swallowed in Phase I, Weight, Anti-Tuberculosis Drug Compliance I, age, Comorbidities, End of Intensive Care, Temperature, Initial X-ray, Fever >1 Month, Gender, Final Diagnosis, Initial Phlegm, Weight Loss, Coughing up Blood, Night Sweats, Shortness of Breath, Ronchi, Coughing Up Phlegm, Duration of Cough, Decreased Appetite, Chest Pain |

In Table 1, the attributes often selected in all feature selection techniques performed are Age, Weight, Initial Diagnosis, End-of-Treatment Sputum, End-of-Treatment X-ray, Number of Medications Swallowed in Phase I, and Anti-Tuberculosis Drug Compliance Phase I.

### 3.3. Modeling

The best accuracy level is tested at this research stage by comparing several classifiers and datasets with or without the drug compliance parameter. Some algorithms used for comparison are Random Forest (RF), Gradient Boosting (GB), Vote with C4.5, SVM, and NN classifiers, and Stacking with C4.5, SVM, and NN classifiers. The compliance parameters used are the number of medications swallowed in Phase I, Anti-Tuberculosis Drug Compliance Phase I, the number of medications swallowed in Phase II, and Anti-Tuberculosis Drug Compliance Phase II. Exploration of this data was

carried out to compare the level of accuracy obtained from several classifiers. The validation method used 10-fold cross-validation. The model accuracy and AUC comparison results are presented in Table 2 and visualized in Figure 2.

Table 2. Comparison of Model Accuracy and AUC using a Dataset with or without Medication Adherence Parameters

| Dataset | Accuracy | | | | AUC | | | |
|---|---|---|---|---|---|---|---|---|
| | RF | GB | Vote | Stacking | RF | GB | Vote | Stacking |
| **without** medication adherence parameters | 62,64 | 58,99 | 62,65 | 57,52 | 66,3 | 67,6 | 64,3 | 57,3 |
| **with** medication adherence parameters | **64,40** | **61,44** | **62,79** | **59,68** | **68,0** | **71,6** | **66,6** | **59,6** |

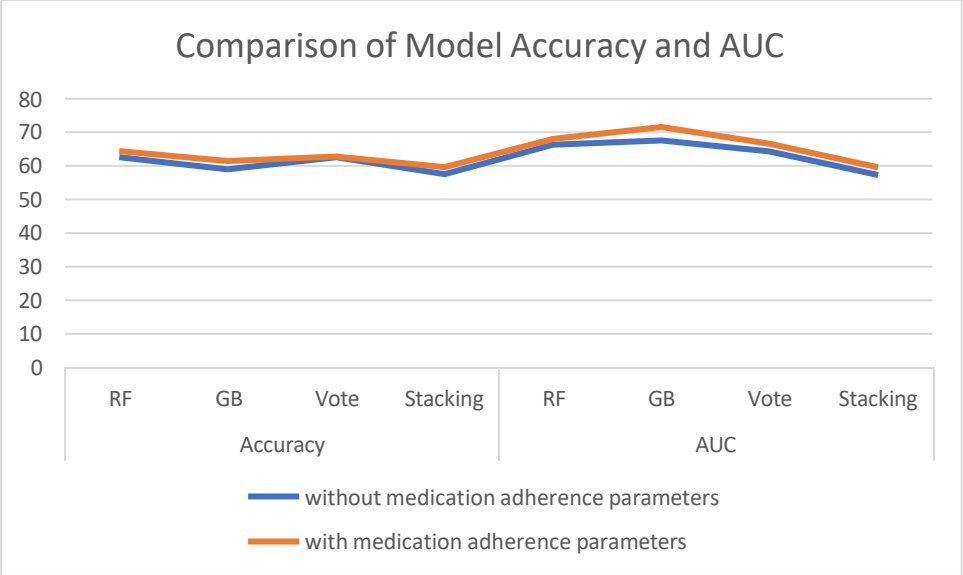


Figure 2. Comparison of Model Accuracy and AUC in terms of Dataset with or without medication adherence parameters

From Figure 2, it can be concluded that, in general, in the dataset with the addition of medication adherence parameters (the orange line), all classifiers have better performance in terms of accuracy and AUC. Thus, in the modeling phase, the dataset with the addition of the medication adherence parameter is used.

After experimenting with several feature selection methods on 739 medical records of tuberculosis patients, several datasets with several classifiers were explored. The five datasets explored were the dataset resulting from feature selection with wrapper method (Bagging, AdaBoost, and Vote), the dataset resulting from feature selection with filter method, and the dataset resulting from feature selection with information gain.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

Table 3. Comparison of Accuracy and AUC of Several Datasets

| Datasets | Accuracy | | | | AUC | | | |
|---|---|---|---|---|---|---|---|---|
| | RF | GB | Voting | Stacking | RF | GB | Voting | Stacking |
| FS-RF | **60.89** | 56.16 | 56.16 | 56.16 | **61.33** | 56.68 | 56.68 | 56.68 |
| FS-Adaboost | **55.34** | 51.01 | 51.01 | 51.01 | **55.23** | 49.99 | 49.99 | 49.99 |
| FS-Voting | **57.77** | 51.97 | 51.97 | 51.97 | **58.00** | 51.60 | 51.60 | 51.60 |
| FS-Filter | **59.67** | 57.11 | 57.11 | 57.11 | **59.66** | 57.88 | 57.88 | 57.88 |
| FS-Information Gain | 57.23 | **61.43** | **61.43** | 60.22 | 57.61 | **63.11** | **63.11** | 61.68 |

First, the experiments were conducted using the dataset of attribute selection results of the Wrapper method with Random Forest, AdaBoost, Voting, and Filter Methods. The Random Forest Model obtained the highest performance (see Table 3). The last experiment used the dataset of feature selection results with information gain. Voting and gradient boosting obtain the highest accuracy and AUC. However, the model with the best performance has not been obtained. AUC values below 70 indicate that the model is in the poor classification. Random Forest shows the best performance on the feature selection dataset with the wrapper and filter methods compared to other models. Gradient Boosting performs similarly to Voting (C4.5, SVM, and NN) on the feature selection dataset with information gain.

### 3.4. Evaluation

The model was evaluated using four test parameters: accuracy, sensitivity, specificity, and AUC. The model with the highest sensitivity minimizes False Negative. Sensitivity is the ability of the model to detect positively labeled data correctly, that is, 0-6 months. Meanwhile, specificity is the ability of the model to detect negative-labeled data correctly, which is more than six months.

Table 4 shows that the highest sensitivity value is Gradient Boosting and Voting in the feature selection dataset with information gain, which is 94.75%. This means that the model can correctly predict the class of treatment duration of 0-6 months in 94.75% of all cases of treatment duration of 0-6 months. However, the specificity value shows that the model can only predict the class of treatment duration >6 months, which is very low, at 31.46%.

Table 4. Model Evaluation

| Model | Test Size | AUC | Accuracy | Sensitivity | Specificity | Attributes |
|---|---|---|---|---|---|---|
| Gradient Boosting and Voting with Dataset FS Information Gain | 10-fold Cross Validation | 63,11 | 61,43 | 94,75 | 31,46 | Anti-Tuberculosis Drugs Compliance Phase I, Fever >1 Month, Number of Medications Swallowed in Phase I, End of Intensive, Cough with Phlegm, Initial Diagnosis |
| Stacking with dataset FS Information Gain | Split data 70-30 | 65,47 | 63,96 | 93,33 | 37,61 | Number of Medications Swallowed in Phase I, End of Intensive, Anti-Tuberculosis Drugs Compliance Phase I, Initial X-ray, Final Diagnosis, Initial Diagnosis |

### 3.5. Deployment

At this stage, the doctor validates the prediction results from the best model. The results of the doctor's analysis are related to low specificity because the dataset used has not specifically differentiated the drug dose for each patient. This study only used the number of swallowed medications in phase I. Therefore, the amount of medication swallowed is not sufficient to indicate patient compliance in swallowing medication. So, additional parameters are needed to determine whether a patient takes medication every day or intermittently.

Based on the attributes selected for building the model, as presented in Table 4, the doctor approved the attributes resulting from the stacking model using information gain. Those are the Number of Medications Swallowed in Phase I, End of Intensive, Anti-Tuberculosis Drugs Compliance Phase I, Initial X-ray, Final Diagnosis, Initial Diagnosis.

## 4. DISCUSSION

Previous research [9] used 596 record TB patients to predict the period of TB treatment with four classes, namely 0-6 Months, >6-9 Months, >9-12 Months, and >12 Months. C4.5 model obtained 74.33%. The problems related to the dataset are the unbalanced class and the inability to use medication adherence parameters. Therefore, the following research was conducted [11] using 500 TB patient records with two classes, 0-6 Months and >6 Months. Medication adherence parameters were the Number of Medications Swallowed in Phase I, Anti-Tuberculosis Drugs Compliance in Phase I, the Number of Medications Swallowed in Phase II, and Anti-Tuberculosis Drugs Compliance Phase I. The study used feature selection results with the filter method, namely the attributes of Comorbidities, Number of Medications Swallowed in Stage I, Sputum Test at the end of the treatment period, Number of Medications Swallowed in Stage II, and Duration of Treatment as class attributes. Gradient Boosting model got the best performance with 79,6% accuracy and 62,5 AUC. This performance is because the dataset in this study used 100% data from adherent patients (patients participating in the DOTS program). Thus, this study used a dataset that included patient data with the final results of failure and treatment discontinuation (loss to follow-up). Unfortunately, the result of this study shows that the model is still in poor classification based on the AUC and accuracy. Even though the model sensitivity is high, the specificity is very low. This means the model only performs better predicting the positive class (0-6 months).

The low performance of the model is likely due to a misunderstanding of the number of medication adherences that are not accompanied by information on whether the patient should take the medication every day or not (intermittent). So, collecting information related to each patient's dose is necessary. In addition, according to the research roadmap, future research will be made to improve model performance with parameter tuning on the ensemble method or by using a voting weighting technique [19], using ensemble deep learning [24], or analyzing factors associated with loss to follow up before and after treatment [4]. Previous research in the health domain has proven that educational background has no relation to medication adherence, but motivation has [28]. Meanwhile, the risk factors highlight the importance of targeted interventions, which could significantly improve treatment adherence and patient outcomes [4].

Despite the low performance of the model, this study can be a reference in developing a prediction model for the duration of treatment for new cases of pulmonary TB patients by adding medication adherence parameters using the ensemble method. This study proves that adding medication adherence parameters can improve model performance. This prediction model for the duration of treatment for new cases of TB patients can be developed into an early warning-based patient monitoring

system. Thus, it is expected to help achieve the target of eliminating Tuberculosis in 2030 to reduce the death rate by 90% compared to 2019.

## 5. CONCLUSIONS

The addition of medication adherence parameters has been proven to improve model performance. Unfortunately, the results of model exploration with feature selection datasets have not achieved the best performance. The results of the doctor's analysis are related to low specificity because the dataset used has not specifically differentiated the drug dose for each patient. The medication adherence parameters used in this study are the number of medications swallowed in Phase I and Anti-Tuberculosis Drug Compliance in Phase I. Therefore, the amount of medication swallowed is not sufficient to indicate patient compliance in swallowing medication. Besides, this research dataset includes patient data with the final results of failure and treatment discontinuation (loss to follow-up). This means that those patients are in the 0–6-month class. So, future research needs to add other attributes to describe a patient's medication compliance and improve the model's performance.

## CONFLICT OF INTEREST

There is no potential conflict of interest in this study.

## ACKNOWLEDGEMENT

The Indonesian Ministry of Research, Technology, and Higher Education supports this work and will fund it through the Beginner Lecturer Research Scheme in 2024.

## REFERENCES

[1] WHO, ‘Global Tuberculosis Report 2021’, 2021.
[2] Kemenkes RI, Strategi Nasional Penanggulangan Tuberkulosis di Indonesia 2020-2024. 2020.
[3] W. Xu, W. Lu, Y. Zhou, L. Zhu, H. Shen, and J. Wang, ‘Adherence to anti-tuberculosis treatment among pulmonary tuberculosis patients: A qualitative and quantitative study’, BMC Health Serv Res, vol. 9, 2009, doi: 10.1186/1472-6963-9-169.
[4] Y. Jiang et al., ‘Factors associated with loss to follow-up before and after treatment initiation among patients with tuberculosis: A 5-year observation in China’, Front Med (Lausanne), vol. 10, 2023, doi: 10.3389/fmed.2023.1136094.
[5] PDPI, Pedoman Diagnosis dan Penatalaksanaan di Indonesia, 2nd ed. 2021.
[6] E. E. Lutge, C. S. Wiysonge, S. E. Knight, D. Sinclair, and J. Volmink, ‘Incentives and enablers to improve adherence in tuberculosis’, Sep. 03, 2015, John Wiley and Sons Ltd. doi: 10.1002/14651858.CD007952.pub3.
[7] R. Andarwati, Masrah, and Z. I. Fauzi, ‘Analisis Faktor Keberhasilan Penyembuhan Tuberkulosis Paru’, Jurnal Ilmiah Pannmed, vol. 15, no. 3, pp. 337–344, 2020, doi: https://doi.org/10.36911/pannmed.v15i3.
[8] M. Masita and H. Andriani, ‘Analisis Determinan Kejadian Loss to Follow-up (Putus Berobat) pada Pasien Tuberkulosis Paru: Literature Review’, Media Publikasi Promosi Kesehatan Indonesia (MPPKI), vol. 6, no. 5, 2023.
[9] R. Rusdah and B. Agni Bregastantyo, ‘Model Prognosis Masa Pengobatan Pasien Tuberkulosis dengan Metode C4.5’, Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), vol. 10, no. 6, pp. 1197–1204, Dec. 2023, doi: 10.25126/jtiik.2023107393.
[10] I. Azizah, ‘Determinan Lama Waktu Kesembuhan Pada Kategori I Di RSUD Ungaran Kabupaten Semarang’, 2019. Accessed: Feb. 24, 2025. [Online]. Available: https://lib.unnes.ac.id/36449/
[11] Rusdah, B. A. Bregastantyo, and J. E. Riwurohi, ‘Prognosis Model of The Treatment Period of Tuberculosis Patients with Medication Compliance Parameters using The Gradient Boosting Algorithm’, in 2023 6th International Seminar on Research of Information Technology and

Intelligent Systems (ISRITI), Batam: IEEE, Dec. 2023, pp. 220–225. doi: 10.1109/ISRITI60336.2023.10467254.

[12] F. Gorunescu, 'Data Mining Techniques in Computer-Aided Diagnosis : Non-Invasive Cancer Detection', World Acad Sci Eng Technol, vol. 34, pp. 280–283, 2007, doi: doi.org/10.5281/zenodo.1059721.

[13] R. Amran, W. Abdulkadir, and M. Madania, 'Tingkat Kepatuhan Penggunaan Obat Anti Tuberkulosis Pada Pasien Di Puskesmas Tombulilato Kabupaten Bone Bolango', Indonesian Journal of Pharmaceutical Education, vol. 1, no. 1, pp. 57–66, 2021, doi: 10.37311/ijpe.v1i1.10123.

[14] M. S. Zamaa, D. P. Windasari, E. P. Pawerusi, and Z. Iskandar, 'Hubungan Dukungan Keluarga dengan Kepatuhan Minum Obat Pada Pasien Tuberkulosis', Jambura Nurisng Journal, vol. 5, no. 2, pp. 2656–4653, 2023, doi: 10.37311/jnj.v5i2.19547.

[15] M. Fadel and A. Wibowo, 'Application of Machine Learning in Predicting Employee Discipline Violations in Financial', Jurnal Teknik Informatika (JUTIF), vol. 5, no. 1, pp. 171–178, 2024, doi: 10.52436/1.jutif.2024.5.1.1229.

[16] N. N. Marpid, Y. I. Kurniawan, and S. P. Rahayu, 'Analisis Prediksi The Movie Database Rating Film dengan Menggunakan Ensemble Learning Menggunakan Metode Random Forest Regression', Jurnal Teknik Informatika (JUTIF), vol. 6, no. 1, pp. 1–10, Feb. 2025, doi: 10.52436/1.jutif.2025.6.1.1563.

[17] M. H. L. F. da Silva Barros et al., 'Benchmarking machine learning models to assist in the prognosis of tuberculosis', Informatics, vol. 8, no. 2, pp. 1–17, 2021, doi: 10.3390/informatics8020027.

[18] E. B. Postnikov, D. A. Esmedljaeva, and A. I. Lavrova, 'A CatBoost machine learning for prognosis of pathogen's drug resistance in pulmonary tuberculosis', LifeTech 2020 - 2020 IEEE 2nd Global Conference on Life Sciences and Technologies, no. LifeTech, pp. 86–87, 2020, doi: 10.1109/LifeTech48969.2020.1570619054.

[19] V. C. Osamor and A. F. Okezie, 'Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis', Sci Rep, vol. 11, no. 1, pp. 1–11, 2021, doi: 10.1038/s41598-021-94347-6.

[20] R. Rusdah, E. Winarko, and R. Wardoyo, 'Predicting The Suspect of New Pulmonary Tuberculosis Case using SVM , C5 . 0 and Modified Moran ' s I', International Journal of Computer Science and Network Security (IJCSNS), vol. 17, no. 12, pp. 164–171, 2017.

[21] M. Ayaz, F. Shaukat, and G. Raja, 'Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors', Phys Eng Sci Med, vol. 44, no. 1, pp. 183–194, 2021, doi: 10.1007/s13246-020-00966-0.

[22] K. M. Liao, C. F. Liu, C. J. Chen, J. Y. Feng, C. C. Shu, and Y. S. Ma, 'Using an Artificial Intelligence Approach to Predict the Adverse Effects and Prognosis of Tuberculosis', Diagnostics, vol. 13, no. 6, Mar. 2023, doi: 10.3390/diagnostics13061075.

[23] Y. Luo et al., 'Development of diagnostic algorithm using machine learning for distinguishing between active tuberculosis and latent tuberculosis infection', BMC Infect Dis, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12879-022-07954-7.

[24] O. B. Poirion, Z. Jing, K. Chaudhary, S. Huang, and L. X. Garmire, 'DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data', Genome Med, vol. 13, no. 1, pp. 1–15, 2021, doi: https://doi.org/10.1186/s13073-021-00930-x.

[25] B. Abdualgalil, S. Abraham, and W. M. Ismael, 'An Efficient Machine Learning Techniques as Soft Diagnostic for Tuberculosis Classification Based on Clinical Data', Journal of Scientific Research, vol. 66, no. 02, pp. 61–67, 2022, doi: 10.37398/jsr.2022.660209.

[26] J. Chen et al., 'Predictive machine learning models for anticipating loss to follow-up in tuberculosis patients throughout anti-TB treatment journey', Sci Rep, vol. 14, no. 1, p. 24685, Dec. 2024, doi: 10.1038/s41598-024-74942-z.

[27] T. F. Monaghan et al., 'Foundational statistical principles in medical research: Sensitivity, specificity, positive predictive value, and negative predictive value', Medicina (B Aires), vol. 57, no. 5, 2021, doi: 10.3390/medicina57050503.

[28] R. A. K. Wardhani, 'Hubungan Pengetahuan dan Motivasi dengan Kepatuhan Minum OAT pada

Vol. 6, No. 2, April 2025, Page. 891-904
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263

Penderita Tuberkulosis Paru di Puskesmas Mangunjaya Kabupaten Bekasi Tahun 2022', Fakultas Kedokteran, UIN Syarif Hidayatullah Jakarta, Jakarta, 2022.

**Jurnal Teknik Informatika (JUTIF)**
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 2, April 2025, Page. 891-904
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4263