Journal of Optimization in Industrial Engineering
Vol. 16, Issue 2, Summer & Autumn 2023, 213-219
DOI: 10.22094/JOIE.2023.1988443.2079

Research Article

Short-Term Tuberculosis Incidence Rate Prediction for Europe using Machine learning Algorithms

Jamilu Yahaya Maipan-uku [a] , Nadire Cavus [b] , Boran Sekeroglu [c]

[a] *Computer Information Systems, Near East University, N.Cyprus, Nicosia, 99138, Mersin 10, Turkey*
[b] *Department of Computer Science, Ibrahim Badamasi Babangida University, Lapai, Nigeria*
[c] *Computer Engineering, Near East University, N.Cyprus, Nicosia, 99138, Mersin 10 Turkey*
Received 10 June 2023; Revised 30 August 2023; Accepted 28 September 2023

Abstract

Tuberculosis (TB) remains a significant public health concern in Europe, necessitating effective disease management and resource allocation. Predicting short-term TB incidence rates using machine learning algorithms offers a data-driven approach to aid policymakers and healthcare professionals in making informed decisions. Machine learning (ML) algorithms are essential for prediction tasks due to their ability to establish a relationship for data sequences. In this study, three machine learning algorithms, namely, Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN), are implemented to predict the tuberculosis incidence rates and to compare the efficacy of ML algorithms for tuberculosis incidence rates prediction for 2025, among Europe. Even though all models achieved considerable results, DT obtained superior prediction rates for the future TB incidence rate with MSE, MAE, and R^2 of 0.000555, 0.01506, and 0.96430 while RF 0.000882, 0.01781, and 0.94329, and ANN 0.000767, 0.02315, and 0.95066. The prediction results showed that a significant decrease in TB incidence rates is expected for 2025 form 49,752 in 2019 to 38,509 in 2025, except Finland and Malta.

Keywords: *Tuberculosis Incidence Rates; Europe; Machine Learning; Decision Tree; Random Forest; ANN*

1. Introduction

Tuberculosis remains a crucial community health threat and one of the major communicable diseases that cause death worldwide [1]. Diseases attributed to tuberculosis include lung and heart disease, chronic respiratory disease, HIV, cancer, and diabetes. In accordance with the WHO Framework Convention on Tobacco Control (WHO-FCTC), MPOWER (Monitoring tobacco use and prevention measures, Protecting the public from tobacco smoke, offering tobacco cessation, Warning about the dangers of tobacco, enforcing bans on tobacco advertising, promotion and sponsorship, raising tobacco taxes) measures were introduced [2] as six cost-effective, high-impact ways to assist countries in reducing tobacco demand. However, according to WHO-Global Tobacco Epidemic 2021 report, only two countries in the world (Brazil and Turkey) have implemented all MPOWER measures at a comprehensive level. As a result, although the predominance of smoking among those aged 15 years and older has decreased from 22.7% to 17.5%, the total number of smokers remains high. As a result, the number of deaths caused by TB has amplified from 1.2 million people in 2019 to 1.3 million [3].

However, systems for different kinds of tuberculosis infections were introduced in diverse developed and developing countries for recordings of TB-related cases, and the registration proportions do not provide precise statistical information. As a result, it provided a broad picture of the incidence and mortality rates of various kinds of tuberculosis around the globe [4].

According to the WHO regional office for European Tuberculosis Fact sheet 2021, about 246,000 people are estimated to become ill with TB, and about 20,000 people lost their lives due to TB in the region. Thus, there was a fast decline in TB cases of about 5% per year in the region. However, the region recorded the maximum rate of Multi-Drug-Resistant TB (MDR-TB) worldwide [3].

Different types of tuberculosis have diverse occurrence rates in different regions, making public education and screening research inefficient. Predicting occurrence rates within diverse countries might offer valuable information on prevalence and public consciousness. This would result in savings in pecuniary and human resources, which require substantial funding. In addition, predictive results would offer researchers information on the progression and effectiveness of tuberculosis prevention interventions. However, such predictions are not as easy to accomplish for humans as compared to a machine due to their processing times, their ability to establish connections among data, and their continuous deployment capability. Therefore, ML models have been used in almost all areas of our life to aid mortal experts to resolve hitches or make decisions [2-3]. Furthermore, ML has been used and explored in various areas of health sciences such as biotechnology and especially in tuberculosis prediction to assist tuberculosis researchers and the public in further prevention and education.

Kumar et al., (2019) applies machine learning algorithms to predict tuberculosis cases in an urban area of India. It highlights the potential of machine learning for accurate

* Corresponding author Email address: jymaipanuku@ibbu.edu.ng

disease prediction. Similarly, Tang et al., 2023 utilize machine learning techniques for tuberculosis incidence prediction based on meteorological factors and air pollutants. The result of the study shows that machine learning model can accurately predict the incidence trend of tuberculosis in Changde City. The following stream of researchers also worked and shows the significant of using Machine learning methods for tuberculosis incidence prediction. (Akshita & Srabanti, 2019; Zheng et al., 2023; Ali et al., 2021; Abdualgalil et al., 2022) They explores the potential of machine learning in enhancing disease prediction accuracy (*see table 1*).

In all the prior studies, the high prediction was obtained majorly using classification [5-11], and some research reflects regression [12-13], and none of the studies centers on tuberculosis incidence prediction. This study tends to adopt regression methods since the objective is to quantify the number of tuberculosis incidence (continuous) in nearby future, and it could only be achieved using regression models. This study aims to employ machine learning models for tuberculosis incidence prediction among the European continent. In essence, using machine learning to predict short-term TB incidence rates empowers healthcare systems with proactive strategies to curb the disease's spread, allocate resources efficiently, and ultimately improve public health outcomes. the contribution of predicting short-term TB incidence rates using machine learning lies in its ability to empower healthcare systems with actionable insights for early detection, efficient resource allocation, targeted interventions, and evidence-based decision-making. This approach ultimately aids in reducing the burden of tuberculosis and improving public health outcomes.

Table 1
Summary of Related Literature

Article(s)	Model (s) used	Evaluation Metrics	Validation Techniques	Learning Algorithm	Dataset	Obtained Results
Tang et al., 2023	Regression	RMSE, MAE, and MAPE	Train/ Test	SVR, RF, and BPNN	2010 – 2021 meteorological factors and air pollutants data of Changde city	High accuracy and error minimization by BPNN
Lewis et al., 2021	Classification	Accuracy	X2CT-GAN		LIDC-IDRI	Improves TB Identification by 7.50%
Mohidem et al., 2021	Regression	MSE. R^2		MLR, ANN	2013-2017 3325TB sociodemographic data	ANN proves to be more efficient in TB cases prediction
Rashidi et al., 2021	Classification	Sensitivity, AUC, Accuracy and Specificity	Train/ test split	kNN, LR, SVM, DNN, RF, NB, & XGBoost GBM	Gulab Devi Chest Hospital Lahore, Pakistan	Improved in sensitivity by 90.5% and 100.0% - 74.6% specificities
Ali et al., 2021	Classification	Sensitivity, Accuracy and Specificity	Train/ test split	DT, ANN, RF, LR, SVM, KNN, & LASSO	PMDT site, eXpert site and additional eight (8) districts headquarter hospitals	Recommended the following models to be used for patients' classifications. Accuracy, Sensitivity, Specificity, SVM and RF
Jamal et al., 2020	Classification	Accuracy, AUC	Train/ test split	Naïve bayes, kNN, SVM, & ANN	PDB, MUBII-TB-DB, GMTV	Indicated the relevance of ML/AI in Predicting Mycobacterium tuberculosis
Lai et al., 2020	Classification	Accuracy, Sensitivity, & Specificity	Train/ test split	ANN, SVM, & RF	Talpel Medical University-Wanfang Hospital	Shows ANN as the best Predicting algorithm in compare with RF & SVM
Asad et al., 2020	Classification	Accuracy, Precision, Recall & F-measure	Train/ test split	ANN, SVM, kNN, RF, & J48	OCICB under NIAID	Shows ANN as the best Predicting algorithm
Van et al., 2019	Regression	RMSE	Jackknife	Bayesian, multiple linear regression	MFX-RIF	Limited sample strategies can be implemented to facilitate Therapeutic Drug Monitoring of moxifloxacin in TB patients

The rest of the paper is organized as follows: Section 2 introduces materials and methods used in this study, while Section 3 presents and discusses the obtained results. Finally, Section 4 concludes the remarks of the study.

2. Materials and Methods

This section briefly introduces the basic characteristics of the considered dataset and machine learning algorithms (models) implemented in this study.

2.1. Dataset

The tuberculosis incidence data were collected from the WHO 2021 datasheet [3] through European Centre for Disease Prevention and Control data repository [14]. It includes incidence rates of tuberculosis infections across 31 European countries between 1995 – 2019. Table 2 shows the list of European countries included in the dataset.

Table 2
List of European countries included in the dataset

Country	Country	Country
Austria	France	Netherlands
Belgium	Croatia	Norway
Bulgaria	Hungary	Poland
Cyprus	Ireland	Portugal
Czechia	Iceland	Romania
Germany	Italy	Sweden
Denmark	Liechtenstein	Slovenia
Estonia	Lithuania	Slovakia
Greece	Luxembourg	United Kingdom
Spain	Latvia	
Finland	Malta	

The nearest neighbor imputation [15] data technique was used to reduce the missing values. Missing values of Latvia (2018 and 2019) were replaced by values of 2017, while that of Liechtenstein for 2019 was replaced by the value of 2018. Each attribute was normalized using minimum-maximum scaler data between 0 and 1.

2.2. Machine learning algorithms

In this study, Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN) are implemented to predict the TB incidence and perform a comparative analysis. The following subsections introduce the considered algorithms briefly.

Decision Tree. A DT is frequently used for classification and regression problems [16]. Instances and attributes are used to create a hierarchical tree form that includes root and decision nodes. The construction of a decision tree requires very little computational time. However, several decision trees can be constructed from a set of features by assigning the root and following nodes. Therefore, different algorithms, such as Gini and entropy, have been used to construct the most effective and accurate decision trees for classification problems. In regression applications, the mean squared error (MSE) is commonly used to determine the sequence of the nodes by providing the degree of impurity.

Random Forest. Random forests [16] are a tree-based ensemble method and aim to achieve more accurate results by optimizing the mean regression of the created individual trees. However, the number of trees created during the training significantly affects the algorithm's accuracy.

Artificial Neural Networks. Artificial Neural Networks aim to simulate human perception and decision-making abilities. It is frequently used for regression, optimization, and classification tasks [17-18]. Neuron interconnections are updated in each iteration using the back-propagated error which is calculated by considering the actual and target outputs of the neural network. Gradient-descent is the most common algorithm for error minimization.

2.3. Experimental design

This section presents the properties of experiments, the used parameters of the algorithms, and the considered evaluation metrics.

Experiments. The training of models was performed using the hold-out method. As depicted in Fig. 2, the dataset was randomly split into training and testing sets with 70% and 30%, respectively. The final year of the dataset (2019) was used as a target output since the remaining years were considered as the training data.

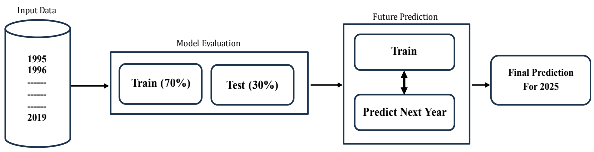

Fig 1. Conceptual framework of the study

The comparative evaluation was performed to determine a superior algorithm, and this model performed the prediction of further years.

The estimation of 2025 was performed based on the sequential test of the superior model by adding the years into the test set. Both the country-based and continental estimations were performed.

As mentioned above, MSE was considered as the tree creation criteria in DT. Random Forest was trained with several parameters, and with 1600 estimators, it achieved its highest prediction rates. The architecture of ANN was determined after varied experiments, and its' superior architecture was created using four hidden layers with 100, 100, 50, and 100 neurons, respectively. The MSE

was used as a loss function, and the network was trained using Adam optimizer.
Each neuron was activated using the Sigmoid activation function, and a 20% dropout was applied to avoid over-fitting.

Evaluation Metrics. The comparative evaluation of the considered machine learning algorithms was performed using three evaluation metrics, namely Mean Squared Error (MSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2 Score), which are commonly used for regression problems [19].
MSE is more sensitive to the outliers by assigning higher weights to errors. Therefore, significant errors could be more observable by the researchers. On the other hand, the error frequencies significantly affect the MSE results and cause the increment of MSE. The formula of MSE is shown in Equation 1.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(O_i - \hat{O}_i)^2 \quad (1)$$

where N, O_i and $\hat{O}_i$ are the number of samples, target output, and actual output, respectively.
MAE measures the general error by using the mean of absolute errors, and the direction of the error does not affect the results, similar to MSE. The formula of MAE is shown in Equation 2.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|O_i - \hat{O}_i| \quad (2)$$

The R^2 score measures the correlation level of predicted and target outputs. The scaled evaluation results of the models' abilities provide a more robust evaluation between them. Equation 3 shows the formula of the R^2 score.

$$R^2 = 1 - \frac{\sum O_i - \hat{O}_i}{\sum O_i - \overline{O_i}} \quad (3)$$

where $\overline{O_i}$ is the mean of target data.

3. Result and Discussion

This section presents the comparative evaluation results of the considered machine learning algorithms for TB incidence rate predictions and the prediction of incidence rates for 2025.

3.1 Comparative evaluation

A decision tree, artificial neural network, and random forest were trained to determine the superior algorithm predicting incidence rates.
Random forest with the superior parameters obtained 0.00088, 0.0178, and 0.943 MSE, MAE, and R^2 scores. Even though the achieved error and prediction levels are considerable, the RF could not reach the scores obtained by the ANN and DT. The ANN produced 0.00076, 0.0231, 0.950 MSE, MAE, and R^2 scores, respectively, and outperformed the RF model in MSE and R^2 scores. However, the MAE results showed that the RF performed predictions with small errors on most of the data since the relatively higher errors than ANN occurred in some data points. This caused the lower MAE and higher MSE results in RF.

The DT achieved superior results compared to the other considered ML algorithms. It achieved 0.00055, 0.01506, and 0.964 MSE, MAE, and R^2 scores and outperformed other algorithms both in error minimization and scaled prediction results. Table 3 presents the obtained results in detail. Bold values indicate superior results. In addition, Figure 2 demonstrates the comparison of the R^2 scores obtained by the models.

Table 3.
Results obtained in this study.

ML model	MSE	MAE	R^2 Score
DT	**0.000555**	**0.01506**	**0.96430**
RF	0.000882	0.01781	0.94329
ANN	0.000767	0.02315	0.95066

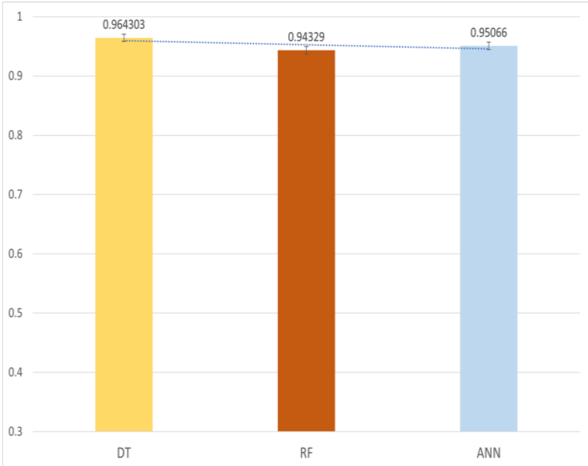

Fig. 2. The visualization of the obtained R^2 scores by the considered models.

3.2 Future estimation of tuberculosis incidence rates

Based on the obtained results in comparative evaluation experiments, the future estimation of TB incidence rates was performed using the decision tree.
The initial training of the DT was followed by the prediction of the following year and adding the predicted value to the test set in order to predict several years. As a result, the final prediction was performed in 2025, which comprises six years between 2019 and 2025. Figure 3 shows the basic steps of the future TB incidence rate prediction process.
The total TB incidence rate in the European continent was 110,315 in 1996; however, it was decreased to 49,752 in 2019. There has been a steady decrease except for two years (1998 and 2002). Therefore, the prediction of the incidence rates by DT are also realized in this direction,

showing that the total TB incidence rate of the European continent would be 38,509 in 2025. Figure 4 shows the TB incidence rate graph, including the estimated 2025 rates. Red portions indicate the increase in the total incidence rates.

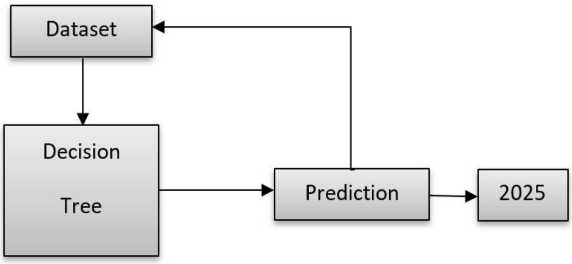

Fig. 3. Block diagram of basic steps for the future estimation process.

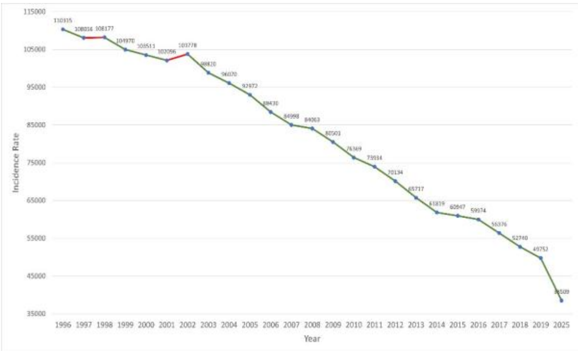

Fig. 4. The European continent TB incidence rates graphs with the prediction of 2025.

It is also predicted that only two European counties, Malta and Finland, would increase the TB incidence rates in 2025. It is estimated that the TB incidence rate of Malta will have a slight increase to 101 in 2025. However, the obtained prediction results showed that the increase in Finland would be more significant, and the TB incidence rate would reach 266 in 2025, from 225. All of the other European countries are predicted to have a decrease in TB incidence rates. Figure 5 shows Finland and Malta's TB incidence rate graphs, including 2025 predictions.

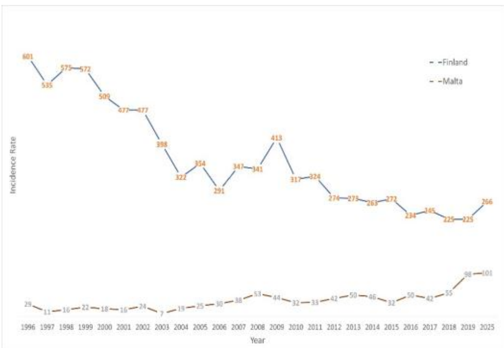

Fig. 5. Finland and Malta TB incidence rates graphs with the prediction of 2025.

3.3 Discussion

Machine learning algorithms have the ability to make high-accurate predictions with the relationships they create between linear or non-linear data. However, the determination of a superior algorithm even in a particular application is a challenging task and requires comparative studies on the relevant subject.

This study applied three ML algorithms, and tuberculosis incidence rates were estimated for 2025. Although all of the considered algorithms achieved reasonable predictive levels, Decision Tree outperformed ANN and RF in all evaluation metrics and achieved an R^2 score of 0.964 on test data during the training phase. Therefore, future prediction is also performed using the superior algorithm, DT. The regular decreasing trend of the data and the simplicity of DT enabled it to make predictions at a higher rate than other algorithms. Figure 6 presents the regression lines on normalized data for each ML algorithm.

The results obtained were consistent with the data between 1996-2019 and predicted that the decrease in TB incidence rates would continue. In this context, it will be possible to review or increase the precautions to reduce the incidence rates estimated to be reached in 2025.

Although the achieved prediction rates on the test set are above 0.96, it is required to optimize the obtained results and further reduce the error in order to make longer term predictions. Consideration of more ML algorithms may contribute to the improvement of results, and it would be possible to predict individual and total tuberculosis incidence rates for all countries in the world over the next 20 years.

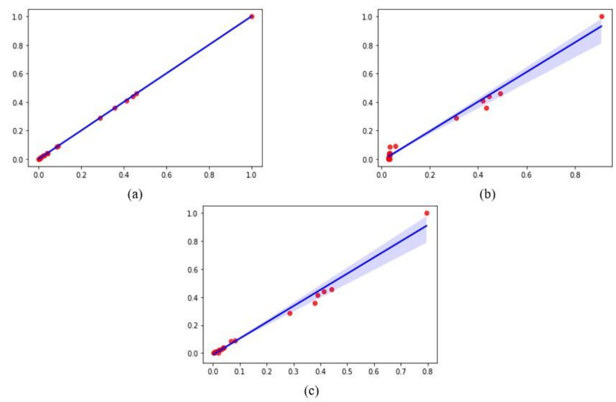

Fig. 6. Regression line plots (a) Decision Tree, (b) Artificial Neural Networks, and (c) Random Forest

4. Conclusion

Different types of tuberculosis have diverse occurrence rates in different countries within European Region. Therefore, predicting occurrence rates within diverse countries might offer valuable information on prevalence and public consciousness. This would, in turn, result in funds savings in pecuniary and human resources, which require substantial funding. In addition, predictive incidences would offer WHO information on the progression and effectiveness of tuberculosis prevention interventions. However, such predictions are not as easy for humans as a machine because of machine processing times, the ability to establish connections among data, and their continuous deployment capability.

This research study modeled three benchmark machine learning algorithms, Decision Tree, Random Forest, and Artificial Neural Network (ANN), to compare the effectiveness of these algorithms in tuberculosis incidence prediction using the tuberculosis dataset of 1995-2019. The comparative study showed that the machine learning algorithms are capable of predicting the tuberculosis incidence rates of Europe; however, the decision tree is the superior algorithm of this study in predicting the rates more accurately.

The prediction results showed that there would be a significant decrease in the number of TB incidence rates of the European continent in 2025, except for Malta and Finland.

Predicting short-term tuberculosis (TB) incidence rates in Europe using machine learning algorithms is a promising approach, but it comes with several limitations that need to be acknowledged such as: Data Quality and Availability; Identification of the most relevant features and engineering predictors; Model complexity; Data imbalance, and External Factors. Despite these limitations, a well-designed and carefully validated predictive model can still offer valuable insights into short-term TB incidence rates, helping public health officials and policymakers make more informed decisions to combat the disease effectively.

Our future work will include the estimation of totally eliminating years of tuberculosis in Europe, addressing the issue of data quality, model interpretability, and employing additional benchmark machine learning algorithms to increase the prediction ability.

Reference

World Health Organization: Global tuberculosis report 2015. 20th edn. World Health Organization (2015). https://apps.who.int/iris/handle/10665/191102

World Health Organization: WHO Report on the global tobacco epidemic: Addressing new and emerging products (2021).

World Health Organization: Global tuberculosis report 2021 (2021).

Şekeroğlu, B., & Emirzade, E. (2018, July). A computer aided diagnosis system for lung cancer detection using support vector machine. In *Third international workshop on pattern recognition* (Vol. 10828, pp. 143-150). SPIE. https://doi.org/10.1117/12.2502010.

Ali, M. H., Khan, D. M., Jamal, K., Ahmad, Z., Manzoor, S., & Khan, Z. (2021). Prediction of multidrug-resistant tuberculosis using machine learning algorithms in swat, Pakistan. *Journal of healthcare engineering*. 2567080 (2021). https://doi.org/10.1155/2021/2567080

Asad, M., Mahmood, A., & Usman, M. (2020). A machine learning-based framework for Predicting Treatment Failure in tuberculosis: A case study of six countries. *Tuberculosis*, *123*, 101944. https://doi.org/10.1016/j.tube.2020.101944.

Jamal, S., Khubaib, M., Gangwar, R., Grover, S., Grover, A., & Hasnain, S. E. (2020). Artificial Intelligence and Machine learning based prediction of resistant and susceptible mutations in Mycobacterium tuberculosis. *Scientific reports*, *10*(1), 5487. https://doi.org/10.1038/s41598-020-62368-2

Kouchaki, S., Yang, Y., Walker, T. M., Sarah Walker, A., Wilson, D. J., Peto, T. E., ... & Clifton, D. A. (2019). Application of machine learning techniques to tuberculosis drug resistance analysis. *Bioinformatics*, *35*(13), 2276-2282. doi:10.1093/bioinformatics/bty949.

Lewis, A., Mahmoodi, E., Zhou, Y., Coffee, M., & Sizikova, E. (2021). Improving tuberculosis (TB) prediction using synthetically generated computed tomography (CT) images. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3265-3273).

Lai, N. H., Shen, W. C., Lee, C. N., Chang, J. C., Hsu, M. C., Kuo, L. N., ... & Chen, H. Y. (2020). Comparison of the predictive outcomes for anti-tuberculosis drug-induced hepatotoxicity by different machine learning techniques. *Computer Methods and Programs in Biomedicine*, *188*, 105307. https://doi.org/10.1016/j.cmpb.2019.105307.

Rashidi, H. H., Dang, L. T., Albahra, S., Ravindran, R., & Khan, I. H. (2021). Automated machine learning for endemic active tuberculosis prediction from multiplex serological data. *Scientific Reports*, *11*(1), 17900. https://doi.org/10.1038/s41598-021-97453-7.

van den Elsen, S. H., Sturkenboom, M. G., Akkerman, O. W., Manika, K., Kioumis, I. P., van der Werf, T. S., ... & Alffenaar, J. W. C. (2019). Limited sampling strategies using linear regression and the Bayesian approach for therapeutic drug monitoring of moxifloxacin in tuberculosis patients. *Antimicrobial Agents and Chemotherapy*, *63*(7), 10-1128. https://doi.org/10.1128/AAC.00384-19.

Mohidem, N. A., Osman, M., Muharam, F. M., Elias, S. M., Shaharudin, R., & Hashim, Z. (2021). Prediction of tuberculosis cases based on sociodemographic and environmental factors in gombak, Selangor, Malaysia: A comparative assessment of multiple linear regression and artificial neural network models. *International Journal of Mycobacteriology*. https://www.ijmyco.org/text.asp?2021/10/4/442/332352.

ECDC, 2019. Surveillance Atlas of Infectious Diseases. https://atlas.ecdc.europa.eu/public/index.aspx?Dataset=27&HealthTopic=54, last accessed 2022/01/10.

Kuhn, M., Johnson, K. (2018). Applied Predictive Modeling. *Springer, Heidelberg*. 2nd edn,

Hussain, S. A., Cavus, N., & Sekeroglu, B. (2021). Hybrid machine learning model for body fat percentage prediction based on support vector regression and emotional artificial neural networks. *Applied Sciences*, *11*(21), 9797. https://doi.org/10.3390/app11219797.

Sekeroglu, B., Hasan, S. S., & Abdullah, S. M. (2020). Comparison of machine learning algorithms for classification problems. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 2 1* (pp. 491-499). Springer International Publishing. https://doi.org/10.1007/978-3-030-17798-0_39.

Cavus, N., Mohammed, Y. B., & Yakubu, M. N. (2021). Determinants of learning management systems during COVID-19 pandemic for sustainable education. *Sustainability*, *13*(9), 5189. https://doi.org/10.3390/su13095189.

Botchkarev, A. (2019). A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, *14*, 045-076. doi:10.28945/4184.

Tang, N., Yuan, M., Chen, Z., Ma, J., Sun, R., Yang, Y., He, Q., Guo, X., Hu, S., & Zhou, J. (2023). Machine Learning Prediction Model of Tuberculosis Incidence Based on Meteorological Factors and Air Pollutants. *International journal of environmental research and public health*, *20*(5), 3910. https://doi.org/10.3390/ijerph20053910.

Tiwari, Akshita and Maji, Srabanti, Machine Learning Techniques for Tuberculosis Prediction (March 14, 2019). International Conference on Advances in Engineering Science Management & Technology (ICAESMT) - 2019, Uttaranchal University, Dehradun, India, Available at SSRN: https://ssrn.com/abstract=3404486 or http://dx.doi.org/10.2139/ssrn.3404486

Zheng, Y., Zhang, X., Wang, X., Wang, K., & Cui, Y. (2021). Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China. *BMJ open*, *11*(1), e041040.

Ali, M. H., Khan, D. M., Jamal, K., Ahmad, Z., Manzoor, S., & Khan, Z. (2021). Prediction of multidrug-resistant tuberculosis using machine learning algorithms in swat, Pakistan. *Journal of healthcare engineering*.

Abdualgalil, B., Abraham, S., Ismael, W. M., & George, D. (2022, January). Modeling and Forecasting Tuberculosis Cases Using Machine Learning and Deep Learning Approaches: A Comparative Study. In *International Conference on Data Management, Analytics & Innovation* (pp. 157-171). Singapore: Springer Nature Singapore.