I.J. Intelligent Systems and Applications, **2025, 5, 32-48**
Published Online on October 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2025.05.03

Machine Learning for Predicting Population Attitudes towards Tuberculosis Patients

Amadou Diabagaté
Faculty of Mathematics and Computer Science, Félix Houphouet-Boigny University, Côte d'Ivoire
E-mail: ahmadou.diabagate@gmail.com
ORCID iD: https://orcid.org/0000-0001-8466-4557

Yazid Hambally Yacouba*
National High School of Architecture and Urban Planning, University of Bondoukou, Côte d'Ivoire
E-mail: yazid.hambally@gmail.com
ORCID iD: https://orcid.org/0009-0005-2130-4083
*Corresponding author

Hafizatou Sani Yanoussa
Emy Polyclinique, Côte d'Ivoire
Email: hafiza.sani@yahoo.fr
ORCID iD: https://orcid.org/0009-0000-6233-9163

Adama Coulibaly
Training and Research Unit for Mathematics and Computer Science, Félix Houphouet Boigny University, Côte d'Ivoire
E-mail: couliba@yahoo.fr
ORCID iD: https://orcid.org/0000-0002-3029-9125

Abdellah Azmani
Faculty of Sciences and Technologies, Department of Computer Science, Tangier, Morocco
E-mail: abdellah.azmani@gmail.com
ORCID iD: https://orcid.org/0000-0003-4975-3807

Received: 22 April 2025; Revised: 17 June 2025; Accepted: 06 August 2025; Published: 08 October 2025

Abstract: Predicting attitudes towards people with tuberculosis is a solution for preserving public health and a means of strengthening social ties to improve resilience to health threats. The assessment of attitudes towards the sick in general is essential to understand the educational level of a given population and to measure its resilience in contributing to the health of all within the framework of community life. The case of tuberculosis is chosen in this study to highlight the need for a change in attitudes, particularly due to the preponderance of this disease in Africa. While it is clear that attitudes influence the organization of individuals and community life, it remains a challenge to put in place an effective mechanism for evaluating the metrics that contribute to determining the attitude towards people with tuberculosis. Knowledge of attitudes towards any disease is very important to understanding collective values on this disease, hence the need to predict attitudes in the case of tuberculosis in favor of health education for all social strata while targeting areas of training not yet explored or requiring capacity building among populations. Changing attitudes towards tuberculosis patients will contribute to preserving public health and will help reduce stigma, improve understanding of the disease and encourage supportive and preventive behaviors. Achieving these changes involves dismantling stereotypes, improving access to care, mobilizing the media and social networks, including people with TB in society and strengthening the commitment of public authorities. The approach adopted consists of assessing the state of attitude towards tuberculosis patients at a given time and in a specific space based on the characteristics of the different social strata living there. An analysis of several metrics provided by machine learning algorithms makes it possible to identify differences in attitudes and serve as a decision-making aid on the strategies to be implemented. This work also relies on the investigation and analysis of historical trends using machine learning algorithms to understand population attitudes towards tuberculosis patients.

Index Terms: Health Education, Machine Learning, Tuberculosis, Artificial Intelligence, Attitudes, Decision Support.

1. Introduction

Managing health care costs and other factors affecting community life or the operation of health centers can be done using predictive analysis. It is in this context that the evaluation of population attitudes towards people suffering from tuberculosis can improve the management of this disease and make it possible to effectively combat the prejudices linked to the disease of tuberculosis.

Today, Artificial Intelligence systems make it possible to tackle complex problems and even find creative solutions. This is why it is important to understand how these systems work in predicting people's attitudes towards tuberculosis patients.

How then can we apply Machine Learning technologies to solve the problem of predicting population attitudes towards people suffering from tuberculosis? Prediction is often used to return an educated guess or opinion.

First of all, Artificial Intelligence (AI) systems are only as good as the quality of the data provided to them. This data must be accurate and representative of the real world. If this is not the case, the system will learn from these inaccuracies and produce inaccurate results. AI systems learn by trying different things and seeing which work best. This means being patient while a system learns as there will be trial and error. However, AI systems still need supervision and direction, which requires thinking about ways they can help us.

Is it acceptable for machines to make decisions regarding the health, well-being and attitudes of populations towards the disease of tuberculosis? It is an ethical challenge which arouses real interest and the answer to which constitutes the main concern in predicting attitudes towards tuberculosis patients. Indeed, it is not only important at the individual level to have the elements to improve the functioning of health systems but also to contribute collectively to the adoption of a common vision in the service of well-being and the preservation of the health of populations. However, the data collected as well as the knowledge derived from these data can be compromising for the privacy of populations. There is therefore reason to make a responsible choice in terms of data ethics when using Artificial Intelligence.

This work will evaluate predictive models for decision-making on data collected about the population's attitude towards tuberculosis patients.

The theory used is the one based on the theory of reasoned action and the theory of planned behavior. This study combines both the use of demographic data and the classification of attitudes to drive changes at population level towards tuberculosis patients. Specific indicators include analysis of factors such as gender, marital status but also the classification of attitudes (wearing a respiratory protection mask, isolation, cohabitation).

In the specific context of this study, several limitations of proposed technology can be highlighted. Social attitudes are dynamic and can evolve over time in response to events such as a health crisis or an awareness campaign. Machine learning algorithms, once applied to a specific set of data, may have difficulty adapting quickly to new information or changes in social perceptions. If machine learning models are not regularly updated with new data, they risk becoming outdated and failing to reflect current attitudes toward TB, limiting their effectiveness in anticipating future trends. Machine learning algorithms can struggle to identify implicit or underlying attitudes in discussions about TB, such as when someone may express superficial support while conveying stereotypes or misconceptions about the disease. These nuances are difficult for algorithms to capture, especially when faced with indirect, ambiguous, or ironic discourse, and can lead to misinterpretations of attitudes, where data is classified as “positive” when it actually contains elements of stigma or lack of knowledge about the disease.

The data used to train models may not be inclusive, particularly for marginalized or vulnerable populations, which can make model predictions less reliable for these populations because they are not adequately represented in the training data.

This work will be structured as follows: section 2 is a presentation of the problematic. Then section 3 gives a state of the art of theories and methods of health education and section 4 is a literature review of artificial intelligence publications related to health education on tuberculosis. Section 5 then presents the data used for prediction and section 6 is a description of algorithms used for supervised learning. Sections 7 and 8 will deal respectively with experimentation and interpretation of the results and then the discussion. Finally, the presentation ends with section 9 which is the conclusion of the article.

2. Problematic

Tuberculosis (TB) poses a major threat to public health in Africa. It affects different social strata and health systems on several levels [1,2].

Sub-Saharan Africa is among the areas with the highest rates of TB in the world. The high prevalence is explained by co-infection with HIV, which results in the complication of its treatment and the increase in the mortality rate associated with tuberculosis [3,4]. Health centers are increasingly confronted with cases of multidrug-resistant tuberculosis (MDR-TB), which further complicates the control of the expansion of this disease within African societies. Africa records a significant proportion of TB-related deaths out of approximately 1.5 million deaths observed worldwide in 2020. The WHO Global TB Plan and the Stop TB Partnership are two major programs involved in TB control and prevention efforts [5]. Several research and development projects for new diagnostics, treatments and vaccines are underway, particularly in Africa [5].

Costs associated with tuberculosis care are high and significantly impact household income. Indeed, people suffering from tuberculosis suffer from a loss of productivity due to the often long periods of work stoppage imposed by the treatment of this disease, which reduces their income and that of their families [6]. African health systems are financially challenged by the costs of treating resistant forms of TB and this situation worsens the financial capacity of a generally underfunded health sector [6].

Discrimination, isolation and reluctance to seek treatment characterize the behavior of people with TB due to the social stigma they face [7]. Children in families of TB patients are often taken out of school to assist sick family members. In addition, these families are impacted economically and emotionally, thus worsening the cycle of poverty [7].

Knowledge of population attitudes towards tuberculosis patients is a key element in understanding collective values in order to develop areas of educational training not yet explored. Strengthening capacities of populations requires implementation of a formal attitude measurement approach.

Understanding the attitude of an individual in society in relation to a given issue meets the need to help in decision-making. While it is true that the answers to many concerns require the collection of information, it appears that community life involves the analysis of societal phenomena with a view to improving the common living environment.

Attitudes toward tuberculosis patients are influenced by complex social factors, such as family relationships, cultural beliefs, social norms and lived experiences. These human interactions are difficult to model comprehensively using quantitative or qualitative data alone; and algorithms risk oversimplifying these complex factors, leading to reductive or inaccurate conclusions about population attitudes.

The general objective aims to educate about health to improve the living conditions of populations and the specific objectives seek to promote a change in attitude at the population level, to predict population attitudes towards people with tuberculosis and to help in decision-making through data analysis.

3. Theories and Methods of Health Education

Health education is part of an approach aimed at guiding practices and interventions to protect public health through several major theoretical frameworks.

Ottawa Charter for Health Promotion (1986) is the reference document for the implementation of public policies, health promotion, construction of favorable environments for the development of health services. public health and skills [8].

It makes it possible to predict behavior in the face of an illness on the basis of four fundamental criteria which are: susceptibility to illness, severity, obstacles and opportunities for health actions [9]. Theory of Reasoned Action and Theory of Planned Behavior explain that behavioral intention is not only the basis of health behaviors but this intention is also determined by subjective norms, attitudes and perceived control of behavior [10]. Transtheoretical Model of Change (Stages of Change) makes it possible to observe each of the different phases of change in an individual's behavior which are respectively precontemplation, contemplation, preparation, action and adaptation [11].

Socio-ecological model highlights several factors (social, economic, cultural, political and environmental) in the influence of health behaviors between individuals, taking into account several levels (interpersonal, intrapersonal, community, organizational and public) [12]. Self-determination theory places humans at the center of their own health behavior which is motivated by relational, psychological and autonomy needs [13]. In the empowerment model approach, the preservation of health and the environment is entirely based on the empowerment of people and communities [14].

4. Literature Review of Health Education on Tuberculosis Using Artificial Intelligence

Here is an overview of the integration of artificial intelligence into health education and tuberculosis control.

$$\mathcal{L}_n = \{(x_1, y_1), \dots, (x_n, y_n)\} \tag{1}$$

Saquib Alam and al. [15] propose initiatives to control tuberculosis through the use of Artificial Intelligence (AI) for more efficient and more inclusive health. Muhammad Faiz Mohd Hisham and al. [16] highlight an evaluation based on Artificial Intelligence software for the diagnosis and radiographic screening of tuberculosis. Lina Keutzer and al. [17] use mobile applications of Artificial Intelligence to improve the treatment of tuberculosis. Yejin Lee and al. [18] present a review of the use of Artificial Intelligence for strengthening tuberculosis control. Juliet Nabbuye Sekandi and al. [19] propose monitoring adherence to tuberculosis treatment in Africa through the development and validation of Artificial Intelligence algorithms.

Sukatemin and al. [20] present a literature review on the use of AI for early identification of TB recovery phases. Nijiati, Mayidili and al. [21] use AI in low-resource conditions for the diagnosis of tuberculosis by chest radiography. Ntwali Placide Nsengiyumva and al. [22] present an analysis of the cost and effectiveness of chest x-ray interpretation for triage of tuberculosis symptoms using Artificial Intelligence. S. Tamilselvi and al. [23] propose an Artificial Intelligence system based on biological sensors for the detection of tuberculosis.

Some works dealing more specifically with the evaluation rather than the prediction of the attitude towards tuberculosis patients have also been identified. The article [24] highlights the importance of taking into account psychological assessments of tuberculosis patients with regard to their disease. G. Garden uses a questionnaire to measure specific aspects of attitude including deeper feelings about mental health and recovery motivation. Jang and Jeon [25] develop and validate a predictive model to assess the quality of life of hospitalized pulmonary tuberculosis patients. The developed predictive model included several key factors, including TB symptoms, depression, social support and nutritional status. Jaramillo's study [26] used questionnaires to identify demographic, social, and psychological characteristics that predict stigma and prejudice toward people with TB. The study highlights that fear of contagion, lack of information, cultural beliefs, socioeconomic status, and history of discrimination strongly influence TB stigma. Loh, Zakaria, and Mohamad [27] reveal a moderate level of knowledge about TB in Malaysia among a sample of contacts of TB patients. The authors rely on statistical analyses to reduce stigma and promote positive or negative attitudes toward tuberculosis. West, Gadkowski, Østbye, Piedrahita, and Stout [28] conducted a cross-sectional survey of people at increased risk of tuberculosis in North Carolina based on questions of general knowledge of tuberculosis (modes of transmission, symptoms, treatment), personal beliefs (erroneous beliefs about contagion), and attitudes toward people with tuberculosis. The statistical analyses conducted concluded that further efforts are needed to improve knowledge about tuberculosis among populations at increased risk in North Carolina.

The study [29] by A. Y., Kidanemariam, B. Y., Tesfamariam, E. H., and Gulbetu, M. E. which focuses on the assessment of community attitudes and practices towards tuberculosis in the Nakfa sub-zone shows that there are still significant gaps in the in-depth understanding of the disease. The article [30] by Fan, Y., Zhang, S., Li, Y., Li, Y., Zhang, T., Liu, W., and Jiang, H. presents a reliable and valid tool to measure the attitudes and practices of students with tuberculosis in China based on a questionnaire which found that care-seeking behaviors were often inadequate but also highlighted the need for better health education. K., Spigt, M., Johanna, L., Noortje, D., Abera, S. F., and Dinant, G. J. [31] show that some prisoners in northern Ethiopia showed signs of stigmatization towards people with tuberculosis and emphasize that this has the effect of inhibiting the search for care for the disease.

The contribution of our work compared to the existing one consists in the prediction of one of the attitudes of cohabitation, isolation and wearing of respiratory protection masks towards tuberculosis patients. Several characteristics intervene in the determination of the attitude including sex, age, marital status, municipality of residence, persistent nature of the cough, existence of contact with a TB patient, consultation and behavior. In our literature review, we did not identify any scientific publication dealing specifically with the prediction of several attitudes towards people with tuberculosis based on well-defined variables and even less through the use of the machine learning algorithms that we used. Also, the need to implement a formal attitude measurement approach is essential to act precisely on the factors to be corrected in population health education. This study is very different from all existing studies and the model used can be automated for the prediction of attitude towards tuberculosis patients. The choice of methodological approaches and target populations in the context of predicting attitude towards people with tuberculosis also denotes the specificity of our study.

5. Data used for Prediction

A population of 507 individuals was randomly interviewed in five different municipalities of the city of Niamey, capital of Niger as part of a survey conducted in 2016 on tuberculosis disease.

The problem addressed is to determine the attitude of populations towards people with tuberculosis. Several criteria were defined as parameters for the prediction of one of the possible attitudes, namely cohabitation, isolation and wearing a respiratory protection mask.

This study is intended as a factual approach to predicting attitudes towards tuberculosis based on a specific number of parameters which are gender, age, marital status, municipality, persistent nature of the cough, existence of contact with a tuberculosis patient, consultation and behavior.

Table 1. Variables and response methods used for data collection

Variable	Response modalities
Sex	Female
	Male
Age	18 - 25 years old
	30 - 44 years old
	45 - 59 years old
	60 and more years old
Marital status	Single
	Married
	Divorce
	Widowed

municipality of residence	Niamey 1
	Niamey 2
	Niamey 3
	Niamey 4
	Niamey 5
Persistent cough or not	Yes
	No
Existence of contact with a TB patient	None
	Casual
	Permanent
Consultation	Yes
	No
Behavior	Abandonment
	Indifference
	Support
Attitude (towards the patient with tuberculosis)	Cohabitation
	Isolation
	Wearing of respiratory protection mask

The table shows the variables and response modalities used for data collection.

6. Description of Algorithms used for Supervised Learning

6.1. Decision Tree

A decision tree has an inverted tree structure composed of a root node that represents all of the data, internal decision nodes that represent the relationship between the values of one or more independent variables, and terminal nodes connected by branches that are the leaves of the tree and which represent the dependent variables. Decision trees are a powerful and intuitive tool that can solve both classification and regression problems by dividing data into subsets based on simple questions. Their intuitive nature makes them particularly useful in situations where transparency is crucial, notably thanks to their ability to manage non-linear data.

The decision tree method was popularized in the field of artificial intelligence and machine learning by several researchers such as Leo Breiman [32] but a key name associated with this technique is J. Ross Quinlan. He introduced the ID3 (Iterative Dichotomize 3) algorithm in the 1980s [33], one of the first widely used algorithms for building decision trees.

Quinlan later improved this algorithm [34], adding features like handling missing values and processing both numeric and categorical data, which became one of the most popular decision tree methods. Figure 1 illustrates a decision tree, as well as the partitioning of the input space that it implies.

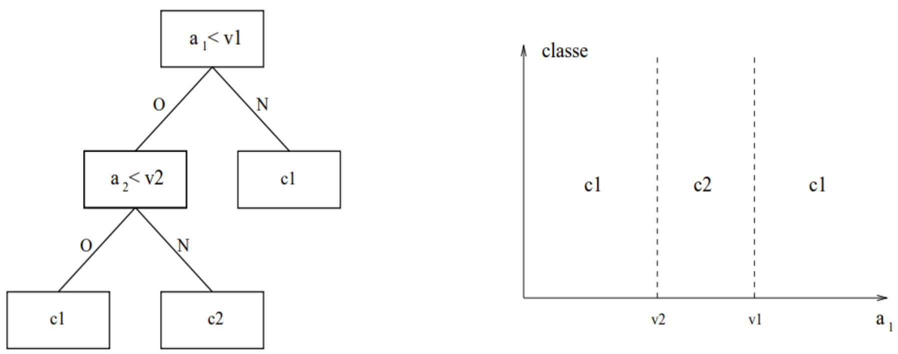

Fig.1. Decision tree and the partition it implies

Decision trees can be used to solve different types of problems in supervised machine learning, including classification problems (to classify data into different categories or classes) and regression problems (to predict continuous values).

After selecting the attributes, we move on to the predictions in the leaves of the decision trees, which are performed using formulas adapted to the nature of the target variable (classification or regression).

For a classification tree, the prediction corresponds to the majority class in the leaf. The predicted class in the leaf is:

$$\hat{y} = \arg\max_i(n_i) \tag{2}$$

where n_i represents the number of occurrences of class i among the observations in a terminal leaf of the decision tree.

For a regression tree, the prediction in a leaf is the average of the target values of the observations in that leaf. The predicted value is given by:

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{3}$$

where n is the number of observations in the leaf and y_i are the target values.

Finally, we have the performance evaluation which can be measured using different metrics, depending on whether it is classification or regression.

The performance of a classification model is evaluated by the error rate which is measured by the percentage of incorrect predictions.

$$\text{Error rate} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\lambda}(y_i \neq \hat{y}_i) \tag{4}$$

where:

- $\boldsymbol{\lambda}$ is the indicator which is equal to 1 if $y_i \neq \hat{y}$ and 0 otherwise.
- n is the number of observations in the leaf.
- $y_i$the actual value and $\hat{y}_i$ the corresponding predicted value.

Accuracy for regression models can be measured by the mean square error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5}$$

where n is the number of observations in the leaf, y_i the actual value and $\hat{y}_i$ the predicted value.

An alternative to the Mean Squared Error (MSE) is the Mean Absolute Error (MAE), which uses the sum of the absolute values of the errors rather than the sum of the squares. The formula is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6}$$

6.2. Random Forests

Random forests are a technique for constructing multiple decision trees by selecting a random subset of variables for each tree to determine the best correlation between the trees. For each set of variables, the allocation function determines how much training data to allocate while the combination function determines how the outputs of each tree should be combined into a single output [35].

Random forests are a new statistical learning method developed by Léo Breiman [35]. This new statistical learning technique is based on Bagging. Bagging consists of dividing the learning sample into several sub-samples, applying a prediction method to each sub-sample and aggregating the results obtained at the level of the different sub-samples [36]. Random forests are an improvement on Bagging and they differ from Bagging in that the number q of variables used at the level of each sub-sample is less than J (the total number of variables), which is not the case for Bagging where $q = J$. This reduction in the number of variables allows for better results [37].

Random forests are a family of methods, the method proposed by Leo Breimen called Random Forests-RI, which we will use, has been widely used in many applications because of its ease of use and exceptional performance [38].

In addition to building a predictor, the random forest algorithm calculates an estimate of its generalization error: the Out-Of-Bag (OOB) error. Let $(\mathrm{x}_i, \mathrm{y}_i)$ be an observation of the sample $\mathcal{L}_n$. Consider the set of trees built on the subsamples not containing this observation (the set of these subsamples constitute the OBB sample). We only use the predictions of the subsamples of the OOB sample to build a prediction $\hat{Y}_{iOOB}$ of y_i

In the case of a regression, the Out-Of-Bag error which is the error made by the predictions is obtained by the following formula:

$$E_{OOB} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_{iOOB} - y_i)^2 \tag{7}$$

In the case of a classification, it is obtained by:

$$E_{OOB} = \frac{1}{n}\sum_{i=1}^{n} T_i \ \ with\ T_i = \begin{cases} 1\ if\ \hat{Y}_{iOOB} \neq \mathrm{y}_i \\ 0\ if\ \hat{Y}_{iOOB} = \mathrm{y}_i \end{cases} \tag{8}$$

Consider a variable X_j , a subsample $\mathcal{L}_n^{\Theta_l}$ and the associated sample OOB_l. Let us calculate E_{OOB}^l the error committed on OOB_l by the tree built on $\mathcal{L}_n^{\Theta_l}$. We perform a random permutation of the values of the variable X_j in the sample OOB_l to obtain a perturbed sample $\widetilde{OOB_l}^j$. We calculate the error of the perturbed sample $\widetilde{OOB_l}^j$ which we denote by $E_{\widetilde{OOB_l}^j}^l$. The same operation is performed for all subsamples. The importance of the variable X_j which we denote by $Imp(X_j)$ is:

$$Imp(X_j) = \frac{1}{q}\sum_{l=1}^{q}(E_{\widetilde{OOB_l}^j}^l - E_{OOB}^l) \tag{9}$$

6.3. Gradient Boosting

Gradient boosting machines are another ensemble learning technique that combines multiple models to create a powerful predictive model. This technique builds multiple decision trees trained sequentially where each tree attempts to correct the errors of the previous one. The final predictions result from combining the predictions from all constructed trees, often by adding them together.

Boosting is a supervised learning method that consists of building a reliable prediction by aggregating the responses of base learners, i.e. estimators that are just better than chance. This family of machine learning algorithms sequentially builds base learners, also called weak learners. At each iteration, the new estimator favors its learning on the errors of the previous one and adds to it to finally obtain a strong learner. This method was particularly recognized with the Adaboost algorithm [39]. Even today, many challenges are won by similar methods such as XGBoost/LightGBM [40] known to be powerful on both regression and classification models.

The main idea behind this algorithm is to construct the new base learners to be maximally correlated with the negative gradient of the loss function, associated with the entire ensemble. The applied loss functions can be arbitrary, but to give a better intuition, if the error function is the classical squared-error loss, the learning procedure would result in a consecutive adjustment of the errors. In general, the choice of the loss function is up to the researcher, with both a large variety of loss functions derived so far and the possibility to implement one's own task-specific loss.

This high flexibility makes GBMs (Gradient Boosting Machines) highly customizable for any particular data-driven task. This introduces a great deal of freedom in model design, making the choice of the most appropriate loss function a matter of trial and error. However, boosting algorithms are relatively simple to implement, allowing experimentation with different model designs. Moreover, GBMs have shown considerable success not only in practical applications but also in various machine learning and data mining challenges [41-44].

Given the training dataset $D_n = \{(X_i, Y_i)_{1\leq i\leq n}\}$, containing n pairs (X, Y) of random variables, where, $X \in \mathbb{R}^d$ and $Y \in \{-1,1\}$ for a binary classification problem or $Y \in \mathbb{R}$ for a regression problem, the main objective of Gradient Boosting is to minimize the loss function on the training dataset. To do this, at each iteration m, a new mode $h_m(x)$ is trained to approximate the negative of the gradient of the loss function L with respect to the predictions of the global model at that step.

Gradient Boosting builds the prediction function progressively, by adding successive predictors to improve the performance of the model.

At each step m, $h_m(x)$ is approximated as follows:

$$h_m(x) \approx -\nabla_{F_{m-1}(x)} L\big(y, F_{m-1}(x)\big) \tag{10}$$

Here, $L\big(y, F_{m-1}(x)\big)$ is the loss function and $\nabla_{F_{m-1}(x)} L$ is the gradient of the loss function with respect to the model predictions at step $m-1$.
where:

- y is the true value.
- $F_{m-1}(x)$ is the predicted value at step $m-1$ based on the inputs x.

Once $h_m(x)$ is found, the global model is updated. Then, the prediction function at iteration m is written as:

$$F_m(x) = F_{m-1}(x) + \eta . h_m(x) \tag{11}$$

where:

- F_m is the prediction at the m- th iteration.
- F_{m-1} is the prediction of the previous model.
- η is the learning rate, which adjusts the importance of each added tree.
- $h_m(x)$ is a simple model (often a shallow decision tree) that is trained, from the inputs x, at each step m to predict the residuals or errors of the model's prediction up to that iteration.

Thus, the final prediction function F^* for a Gradient Boosting model is obtained after the M iterations;

During training, the choice of the loss function is crucial, because it defines how the model adjusts its predictions to reduce the error. The loss functions used for Gradient Boosting are multiple. Here are some of them:

- Mean Squared Error (MSE)

MSE is used in standard regression problems and aims to minimize the sum of squared errors between the actual values y_i and the predictions $\hat{y}_i$.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{12}$$

- Mean Absolute Error (MAE)

Minimizing the MAE is the same as reducing the median of the errors. It is also used in regression problems.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{13}$$

- Log-likelihood (Log-Loss or Cross-Entropy Loss)

Used for classification problems, it measures the divergence between the probabilities predicted by the model and the real classes. The further the prediction is from reality, the higher the penalty.

$$\text{Log} - \text{Loss} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i)] \tag{14}$$

Where $\hat{p}_i$ is the predicted probability for the variable y_i.

- Loss Functions for Multiclass Classification Gradient Boosting

For multiclass classification problems, generalized log-loss or cross-entropy is used:

$$\text{Log} - \text{Loss multiclass} = -\sum_{i=1}^{N} \sum_{k=1}^{K} y_{i,k} \log(\hat{p}_{i,k}) \tag{15}$$

Where K is the number of classes and $\hat{p}_{i,k}$ is the predicted probability for class k.

6.4. Logistic Regression

Logistic regression is a model used for classification that models the probability that an observation belongs to a given class. A very important part of this process is normalizing the data. Logistic regression is affected by scale, so entities must be scaled to unit scale for optimal performance. Unit scale means having a mean of zero and a variance of 1 for the entities [45].

Multinomial or polytomous logistic regression generalizes binary logistic regression by treating problems where the dependent variable takes $K(K > 2)$ modalities [46].

The objective is to model the probability of an individual belonging to a category y_k (the associated class is C_k) of the dependent variable. This probability can be defined as follows:

$$\psi_k(\text{x}) = P(Y = y_k / X(\text{x})) \tag{16}$$

Such as:

$$\sum_{k=1}^{K} \psi_k(\text{x}) = 1 \tag{17}$$

The principle of multinomial logistic regression consists of modeling $(K - 1)$ probability ratio called odds. To do this, a category is taken as a reference and the logits are expressed in relation to this reference.

6.5. Naive Bayes

It is one of the most popular algorithms that helps classify items based on many data characteristics. Its main specificity is that it is naive because it assumes that all the predictors are independent of each other, which allows it to be used to do much more than simply classify elements by denying any correlation between the predictors. Naive Bayes creates a probability of class membership by looking at each of the predictors and making so few assumptions, which often makes it much more accurate when classifying data [47,48].

Given a feature vector $X = (x_1, x_2, \ldots, x_n)$ and a class variable C_k, Bayes Theorem states that:

$$P(C_k|X) = \frac{P(X|C_k)P(C_k)}{P(X)}, fot\ k = 1,2, \ldots, k \quad (18)$$

With $P(C_k|X)$ the posterior probability, $P(X|C_k)$ the likelihood, $P(C_k)$ the prior probability of class, and $P(X)$ the prior probability of predictor. We are interested in calculating the posterior probability from the likelihood and prior probabilities.

Using the chain rule, the likelihood $P(X|C_k)$ can be decomposed as:

$$P(X|C_k) = P(x_1, x_2, \ldots, x_n \mid C_k) = P(x_1 \mid x_2, \ldots, x_n, C_k)P(x_2 \mid x_3, \ldots, x_n, C_k) \ldots \mathrm{P}(x_{n-1} \mid x_n, C_k)P(x_n \mid C_k) \quad (19)$$

6.6. Support Vector Machine (SVM)

This method was first used for classification problems. Its main goal is to create an optimal margin that can separate the maximum number of data points. Sometimes the data is not linearly separable and it is not that simple to draw a line between them and in these cases, it is kernel tricks that are used to make them linearly separable. The objective is to find a hyperplane that minimizes the error and obtains a minimum margin interval with a maximum number of data points [49].

Support vector machines (SVM) are algorithms whose goal is to solve two-class discrimination problems. We call a two-class discrimination problem a problem in which we attempt to determine the class to which an individual belongs (individual is used here in the sense of constituting a set) among two possible choices. SVMs were developed by Cortes and Vapnik (1995) for binary classification [49].

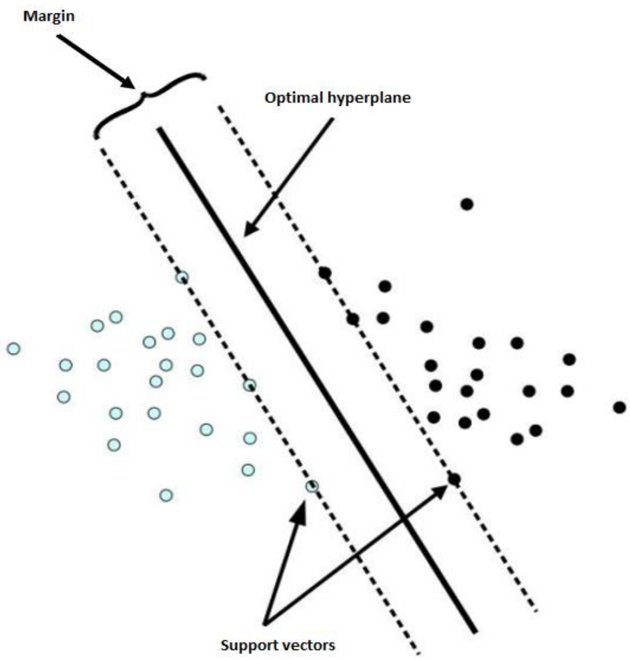

Fig.2. Classification (linear separable case)

The objective of SVM is to find a hyperplane that separates the data into distinct classes with the greatest possible margin, thus maximizing the separation between categories, which amounts to solving an optimization problem to find the hyperplane. optimal. Therefore, in order to find the canonical hyperplane that separates the data with the largest possible margin, it is enough to solve the following optimization problem:

$$\text{Minimize } \frac{1}{2}\|w\|^2, \text{ under the constraints } \ y_i(w \cdot x_i + b) \geq 1 \text{ pour tout } i \quad (20)$$

where:

- x_i is the input feature vector.
- y_i is the class (+1 ou -1) of observation x_i.
- w is the vector of weights determining the orientation of the hyperplane.
- b is the bias moving the hyperplane relative to the origin.

In a binary classification problem, the objective is first to find a hyperplane that separates the two classes. However, if the data are linearly separable, a hyperplane in an n-dimensional space is defined by the equation:

$$w \cdot x + b = 0$$

And for a hyperplane $w \cdot x + b = 0$, the margin M is given by the expression

$$M = \frac{2}{\|w\|} \tag{21}$$

While keeping the assumption that the data is linearly separable, finding a rule to classify them is very simple. So, simply use the following function (sometimes called the indicator function) to perform the classification:

$$Class(x) = sign(\, w \cdot x + b) = \begin{cases} -1 \ , & if\ w \cdot x + b < 0 \\ \ \ 0 \ , & if\ w \cdot x + b = 0 \\ +1 \ , & if\ w \cdot x + b > 0 \end{cases} \tag{22}$$

This function classifies the data relative to the side of the hyperplane on which it is located. Note that if a set of data is separated by a hyperplane, it will be perfectly classified by this function. Note that if a piece of data is directly on the hyperplane (which can happen when considering data that is not in the training set), it will be assigned to class 0, which means that it cannot be classified by the current model. In this case, it is possible to leave it unclassified, use another rule or randomly assign it to one of the two classes.

6.7. Neural Networks

Inspired by biological neurons, artificial neural networks are a type of machine learning that uses a structure like the human brain to break down massive data sets instead of using previous machine learning algorithms. They are particularly effective at processing patterns of inputs and outputs using a sequence of mathematical neurons arranged in layers. Neural networks are the basis of most of the recent revolutionary advances in artificial intelligence and break down data into much smaller pieces [50].

A neural network is a structure composed of several neurons connected to each other and exchanging information via connections between them. The information, assigned a weighting coefficient ω_{ij} is transmitted from a neuron i to a neuron j via the connection that links them.

There are several types of neural networks. We will present the multilayer perceptron that we will use in our study for the following reasons:

- it is very widespread and gives good results [51];
- it is the only model implemented by almost all software and statistical tools;
- this is the only neural model implemented by Mahout.

Multilayer perceptrons are organized in successive layers and each of them includes neurons that are not connected to each other. Below, Figure 3 represents a multilayer perceptron with two hidden layers:

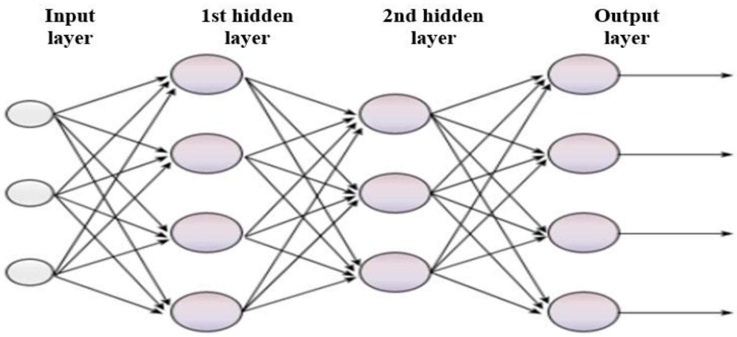

Fig.3. Multilayer perceptron

Consider the following notations:

- O: the set of possible states of neurons.
- o_i: the state of neuron i, where $o_i \in O$.
- f_i: the transfer function associated with neuron i.
- s_i: the input of neuron i, $s_i \in \mathbb{R}$

- ω_{ij}: the weight of the connection from neuron j to neuron i ; $\omega_{ij} \in W$ the set of weights of the multilayer perceptron

When neuron i receives information from n_i neurons, it performs the following operation to define its state:

$$o_i = f_i(s_i) \text{ avec } s_i = \left(\sum_{j=1}^{n_i} \omega_{ij} * o_j\right) - \omega_{i0} \tag{23}$$

ω_{i0}: is an additional weight called bias coefficient linked to the bias which we will note o_0 such that $o_0 = -1$. Note that $\omega_{i0} \in W$

Consider that the first layer which is the input layer has n neurons and the last layer which is the output layer has K neurons. The states of the neurons of the first layer and the last layer can respectively be represented by the vectors $\mathrm{x} = (x_1, \ldots, x_n)$ and $\mathrm{y} = (y_1, \ldots, y_K)$. Thus, the multilayer perceptron can be modeled by the following function F:

$$\begin{gathered}\mathbb{R}^n \rightarrow \mathbb{R}^K \\ \mathrm{x} \mapsto \mathrm{y} = \mathrm{F}(W, \mathrm{x})\end{gathered} \tag{24}$$

Like random forests, the multilayer perceptron can be used for classification and regression problems. We present its properties within the framework of the classification which is the aim of our study.

For classification, the multilayer perceptron is applied by carrying out coding of the dependent variable making it possible to represent the different classes. Let x be an input, the class associated with x is represented by a vector $\mathrm{y} = (y_1, \ldots, y_n)$ whose components can only take two values a or b. The vector which encodes the class C_k has all its coordinates equal to b except the k^{th} which takes the value a. The multilayer perceptron approximates the mathematical expectation of the conditional random variable Y/x. The approximation of the mathematical expectation of the conditional variable Y/x for the k^{th} component is given by:

$$E(Y/\mathrm{x})_k = a * P(C_k/\mathrm{x}) + \mathrm{b} * (1 - P(C_k/\mathrm{x})) \tag{25}$$

7. Experimentation and Interpretation of Results

7.1. Methodology

To achieve this goal, we will follow a roadmap which is based on a widely used method for planning data science projects. This roadmap includes six steps briefly described as follows:

- The first step is understanding the problem to be solved and the benefits it will bring.
- understanding the data before even being able to use it as well as understanding the types of data to know how to process them during predictions.
- Data preparation to ensure that the data can be processed by the prediction models and which sometimes requires the transformation of categorical data into numerical data because most models only accept numerical data.
- Modeling to make predictions by dividing the data into test data and training data, using the training data to train the model and then using the test data to evaluate the model's success score.
- Evaluation is the next step and it involves testing to see if the model provides a good result using the test data. There are different evaluation methods to predict the success of a model. Sometimes, the evaluation will help to decide that the model does not have a good result and, in this case, we return to the modeling stage. Many times, in this process, going back and forth between the modeling and evaluation stages is necessary.
- The deployment of the best model identified to solve the problem posed. The evaluation metrics are useful but there is also a human aspect which is the opinion of subject matter experts which must be taken into account before making a major final evaluation.

The health behavior model used to predict attitudes towards people with tuberculosis is the theory of reasoned action and planned behavior. This approach helps to efficiently extract information to make an informed estimate which then leads to calculated decisions and impactful actions adding added value for populations and decision-makers.

The data used on all algorithms was divided into two sets. The training data sets and the test data sets are made up of 70% and 30% of the data collected for each of the algorithms used, respectively. We are in a case of supervised machine learning insofar as our objective is to build a model which maps input data to output data. Also, the nature of our dependent variable is categorical, which leads us to a classification problem. Popular machine learning algorithms such as decision trees, naive bayesians, neural networks, support vector machines, and K-nearest neighbors can be used to solve regression and classification problems.

After training a supervised machine learning model, it is important to evaluate it to see how well it fits the problem at hand. This is the goal of the evaluation phase in the machine learning process to obtain an unbiased assessment of our model performance by training the model with a different dataset than the one we use for the assess.

Each algorithm has strengths and weaknesses hence the need to try several algorithms to determine which one is best suited to a specific set of data.

The accuracy of the predictions also depends on the hyperparameters used. We will then do the hyperparameter tuning. In a machine learning model, hyperparameters are parameters that must be adjusted for optimal performance. This process involves systematically trying different combinations of these parameters to maximize model accuracy and generalization to new data. This is crucial because it ensures that the model adapts to the specific characteristics of the data set in order to avoid overfitting or underfitting and ultimately to provide the model with predictive capabilities.

7.2. Analysis and Interpretation of Results

Table 2. Results of the algorithms used

	Accuracy (%)	f1_score (%)	Precision (%)	Recall (%)
Random Forest	62.09	39.38	39.92	39.87
Decision Tree	59.47	38.42	38.78	38.55
Gradient Boosting	58.82	38	38	38
Logistic regression	55.55	36.44	36.08	36.82
Naive Bayes	10.45	11.81	50.77	40.09
SVM	70.58	45.90	46.48	46.12
Neural network	64.05	42.26	41.82	42.72

Assessing attitudes towards tuberculosis through machine learning algorithms involves several steps, from data collection to interpretation of results. Here are some specific metrics and indicators used in this context, which can help better understand and assess the impact of Artificial Intelligence models for tuberculosis surveillance, prevention and treatment.

Accuracy measures the proportion of correct predictions made by the algorithm, whether for classifying attitudes or predicting behaviors related to tuberculosis prevention. Accurate accuracy indicates that the algorithm is doing a good job in identifying individuals' actual attitudes toward tuberculosis.

Precision and Recall are two important indicators in a context where attitudes towards tuberculosis are either positive or negative. Precision measures the proportion of correctly identified positive cases among positive predictions, while recall evaluates the proportion of positive attitudes identified among all truly positive attitudes. These two measurements make it possible to judge the ability of the algorithm to correctly deduce attitudes without neglecting certain ones.

The metric (F1-Score) combines both precision and recall into a single value. It is useful when the balance between these two criteria is essential, for example when we want to minimize both false positives and false negatives in attitudes analysis.

The SVM, Neural Network and Random Forest algorithms respectively give the three best accuracy, f1_score and recall while the Naive Bayes, SVM and Neural Network algorithms respectively provide the best precision.

SVM is the best algorithm with an accuracy value of 70.58%, an f1_score value of 45.90% and a recall value of 46.12%. Only the Naive Bayes algorithm obtained a better precision value of 50.77% compared to that of the SVM algorithm which is 46.48%.

However, it should be noted that the second-best performance is achieved by the Neural Network algorithm with an accuracy value of 64.05%, an f1_score value of 42.26% and a recall value of 42.72%.

The third best ranking is achieved with the Random Forest algorithm which obtains an accuracy value of 62.09%, an f1_score value of 39.38% and a recall value of 39.92%.

Several specific measures and indicators have been used to assess attitudes towards tuberculosis, including demographic data and attitude classification. Analyzing attitudes based on variables such as age, gender, and education level makes it possible to predict how these demographic groups react to tuberculosis using supervised learning models (SVM, etc.).

Also, the classification of attitudes (wearing a respiratory protection mask, isolation, cohabitation) makes it possible to predict population attitudes using supervised models such as random forests or neural networks on the data collected. This study provides an overview of different measures to help make decisions on strategies to implement in favor of changing attitudes towards tuberculosis patients.

To assess the impact of interventions or awareness campaigns on changing attitudes, it is essential to compare attitudes before and after exposure to the intervention.

A key indicator of success may be the reduction of negative or stigmatizing attitudes towards people with tuberculosis.

Another way to measure success is to monitor the evolution of preventive behaviors and supervised learning models can predict these behaviors based on attitudes towards tuberculosis.

By analyzing attitudes and behaviors, health authorities can target populations who show reluctance or negative attitudes towards prevention actions, thus making it possible to better adapt communication strategies.

8. Discussion

Stigma of tuberculosis patients is a recognized phenomenon in many parts of the world, and it constitutes a major obstacle to the care and treatment of this infectious disease. Several studies have examined public attitudes towards people with tuberculosis, highlighting negative perceptions that can hamper control and prevention efforts. However, few studies have explored how specific demographic factors (such as age, gender, education level, etc.) influence these attitudes, and how these elements could be integrated into public health strategies to change behaviors and reduce stigma. The study [52] found that younger and better educated people tended to have a better understanding of tuberculosis and prevention practices. People in urban areas were also better informed than those in rural areas. The study concludes that while the majority of participants recognized the importance of preventive measures, some myths and stereotypes persisted, including erroneous beliefs about the transmission of tuberculosis, which could lead to the stigmatization of those who were ill. The study [52] highlighted the impact of education on perceptions of tuberculosis. Individuals with a higher level of education were more likely to understand the mechanisms of transmission and treatment of the disease, which decreased stigma. However, this study [52] also noted that knowledge of the disease was not always sufficient to change behaviors, and that deeper cultural factors were involved. Thus, the majority of existing studies have focused on unidimensional demographic variables and have mainly measured attitudes through questionnaires.

The study by Yuan and al. [53] highlights a significant gender difference in how perceived stigma affects the acceptance of preventive treatment for TB. The results suggest that women, due to higher stigma, are less likely to accept preventive treatment, highlighting the need to adapt public health strategies to take into account gender dimensions. The study observed that young adults, particularly those from rural backgrounds, were more likely to associate the disease with poverty or unsanitary lifestyle, which reinforced stigma. This finding is supported by the work of Yuan and al. [53] who found that stigma varies considerably by gender and area of residence. Men, for example, showed greater reluctance to discuss the disease due to social pressure related to masculinity, while urban residents were more open to discussion and less likely to stigmatize tuberculosis patients. The authors recommend that awareness-raising and education interventions should specifically target women to reduce stigma and promote greater acceptance of preventive treatments. In addition, efforts should be made to address social and cultural concerns that contribute to stigma, while building trust in the effectiveness of treatments. Poorer populations and marginalized groups such as informal workers and people living in substandard housing conditions were less likely to be diagnosed or receive full treatment. These socio-economic inequalities have exacerbated the spread of the disease in vulnerable communities [54]. However, there is a lack of cross-analysis of different demographic factors and their combined influences on behaviors towards tuberculosis patients [55].

Sociodemographic factors, such as education level, economic status, and cultural beliefs, played an important role in shaping attitudes toward tuberculosis [55]. People with low education levels and those living in low socioeconomic conditions were more likely to hold stigmatizing attitudes toward tuberculosis patients. The study [55] showed that stereotypes and fear associated with tuberculosis are more pronounced in certain age groups and among people with low education levels.

Our study sought to overcome the limitations of using unidimensional demographic variables by using a more holistic approach and combining the following demographic factors: age, sex, education level, marital status, area of residence, as well as specific attitudes (mask wearing, isolation, cohabitation). Using a survey of 507 participants in Niamey, Niger, the study categorized attitudes into three categories: support for wearing respiratory protection masks, attitudes toward isolating patients, and acceptance of cohabitation with tuberculosis patients.

The analysis revealed interesting results:

- Young adults, particularly men, were less likely to accept isolation of tuberculosis patients, although they showed greater support for mask-wearing. Older women, on the other hand, showed higher support for isolation of patients and were more likely to accept cohabitation with patients.
- People with a higher level of education were significantly more likely to accept isolation and mask-wearing, suggesting that continued education and information campaigns are essential to reduce stigma. However, even among the most educated, cohabitation with patients remained a sensitive topic.

The results suggest that combining demographic factors in the analysis of attitudes allows us to better understand the different dimensions of tuberculosis stigma. These results highlight the importance of specific targets for public health interventions:

- Young adults, particularly men, should be a key target for campaigns that encourage not only mask wearing but also acceptance of isolation of patients. Educating men on prevention practices and reducing stigma could be reinforced through media campaigns.
- The results also show the need for educational programs and providing information on the mechanisms of tuberculosis transmission and treatment.
- Support for cohabitation with patients is another essential aspect. The development of psychosocial support programs for families of tuberculosis patients could help overcome the obstacles to the acceptance of cohabitation and isolation.

9. Conclusions

Most previous scientific publications on attitudes towards tuberculosis patients focus on social and cultural factors such as stigma, discrimination and perceptions of the disease. However, they have not often included a detailed analysis of demographic data (age, gender, education level, marital status, area of residence, etc.) alongside specific behavioral factors. This publication, which includes these demographic data as well as observable behaviors (mask wearing, isolation and cohabitation), allows a more detailed understanding of different social groups and their specific responses to the disease.

Although machine learning algorithms offer significant potential for analyzing and predicting attitudes toward tuberculosis patients, they face notable limitations, including data bias, interpretation difficulties, dynamics of social attitudes and ethical limitations. Comparing previous studies with this study shows the evolution of research in the area of tuberculosis stigma. The integration of demographic factors combined with specific attitudes makes it possible to better understand behaviors and identify effective levers to modify the attitudes and behaviors of populations. This new vision paves the way for more targeted and effective public health strategies to reduce tuberculosis stigma and promote lasting behavior changes.

This study revealed certain forms of stigmatization, isolation and abandonment towards tuberculosis patients. Awareness raising must be intensified in order to further educate populations and promote a change in behavior among the population of Niamey in particular and that of Niger in general in relation to tuberculosis patients.

This study allowed us to know the attitudes and behavioral practices of the population when faced with a person suffering from tuberculosis. It will constitute a basic tool on which subsequent studies can be based.

The algorithms used in the context of attitude prediction have given encouraging results and have made it possible to identify several algorithms. This approach to comparing algorithms made it possible to highlight the power of the SVM algorithm in front of other algorithms.

This work also made it possible to propose a formal process for analyzing the attitude of populations relative to people suffering from tuberculosis disease by defining the characteristics of cohabitation, isolation and the wearing of respiratory protection masks. This approach will make it possible to bring about changes in favor of a social environment ensuring the well-being of each individual.

Indicators such as precision, precision, recall, and sentiment analysis, combined with specific attitude and behavior measures, not only allow us to evaluate the performance of machine learning algorithms but also to clarify the objectives for success in tuberculosis interventions. These tools help to monitor changing perceptions and guide public health strategies.

On the prospects for future research work, the prospecting of other artificial intelligence algorithms with the aim of improving the attitude prediction rate and the determination of other variables can contribute significantly to improve attitudes prediction. Among the perspectives, it is worth highlighting that factors such as cultural preferences (religious or traditional beliefs) and decision-making behavior were not taken into account and can help refine the prediction of attitudes.

References

[1] Glanz Karen, Barbara Rimer, Kasisomayajula Viswanath, "Health Behavior and Health Education: Theory, Research, and Practice", John Wiley & Sons, ISBN 978-0-7879-9614-7, 2008.
[2] Don Nutbeam, Elizabeth Harris, Marilyn Wise, "Theory in a nutshell: a practical guide to health promotion theories, Sydney", McGraw Hill, ISBN: 9780070278431, 2010.
[3] World Health Organization, "Global Tuberculosis Report 2020", Retrieved from WHO, ISBN: 978-92-4-001313-1, 2020.
[4] UNAIDS, "Global HIV & AIDS statistics — 2020 fact sheet", Retrieved from UNAIDS, 2020.
[5] Stop TB Partnership, "The Global Plan to End TB 2018-2022", Retrieved from Stop TB Partnership, 2018.
[6] Knut Lönnroth, Ernesto Jaramillo, Brian Williams, Christopher Dye, Mario Raviglione, "Drivers of tuberculosis epidemics: The role of risk factors and social determinants", Social Science & Medicine, Vol.68, No.12, pp.2240-2246, 2009. DOI:10.1016/j.socscimed.2009.03.041.
[7] Courtwright Andrew, Turner Abigail Norris, "Tuberculosis and stigmatization: Pathways and interventions", Public Health Reports, Vol.125, No.4, pp. 34-42. DOI: 10.1177/00333549101250s407.
[8] World Health Organization, "Ottawa Charter for Health Promotion", Retrieved from WHO, 1986.
[9] Irwin Rosenstock, "Historical Origins of the Health Belief Model", Health Education Monographs, Vol.2, No.4, pp.328-335. DOI:10.1177/109019817400200403.
[10] Ajzen Icek, "The theory of planned behavior", Organizational Behavior and Human Decision Processes, Vol.50, No.2, pp.179–211. DOI:10.1016/0749-5978(91)90020-T.
[11] Prochaska James, DiClemente Carlo, "Stages and processes of self-change of smoking: Toward an integrative model of change", Journal of Consulting and Clinical Psychology, Vol.51, No.3, pp.390-395, 1983. DOI:10.1037//0022-006x.51.3.390.
[12] McLeroy Kenneth, Bibeau Daniel, Steckler Allan, Glanz Karen, "An Ecological Perspective on Health Promotion Programs", Health Education Quarterly, Vol.15, No.4, pp.351-377, 1988. DOI:10.1177/109019818801500401.
[13] Edward Deci, Richard Ryan, "Intrinsic Motivation and Self-Determination in Human Behavior", Berlin: Springer Science & Business Media, 1985. DOI:10.1007/978-1-4899-2271-7.
[14] Wallerstein Nina, "Powerlessness, empowerment, and health: Implications for health promotion programs", American Journal of Health Promotion, Vol.6, No.3, pp. 197–205, 1992. DOI:10.4278/0890-1171-6.3.197.

[15] Alam Saquib, "From Waves to Wisdom: Harnessing AI For Tuberculosis Education Through Community Radio", Media and Al: Navigating The Future of Communication, p.205-219.

[16] Muhammad Faiz Mohd Hisham, Noor Aliza Lodz, Eida Nurhadzira Muhammad, Filza Noor Asari, Mohd Ihsani Mahmood, Zamzurina Abu Bakar, "Evaluation of 2 Artificial Intelligence Software for Chest X-Ray Screening and Pulmonary Tuberculosis Diagnosis: Protocol for a Retrospective Case-Control Study", JMIR Res Protoc, Vol.12, 2023. DOI: 10.2196/36121.

[17] Keutzer Lina, Wicha Sebastian, Simonsson Ulrik, "Mobile Health Apps for Improvement of Tuberculosis Treatment: Descriptive Review", JMIR Mhealth Uhealth, Vol.8, No.4, 2020. DOI: 10.2196/17246.

[18] Lee Yejin, Raviglione Mario, Flahault Antoine, "Use of Digital Technology to Enhance Tuberculosis Control: Scoping Review", J Med Internet Res, Vol.22, No2, 2020. DOI: 10.2196/15727.

[19] Sekandi Juliet Nabbuye, Shi Weili, Zhu Ronghang, Kaggwa Patrick, Mwebaze Ernest, Li Sheng, "Application of Artificial Intelligence to the Monitoring of Medication Adherence for Tuberculosis Treatment in Africa: Algorithm Development and Validation", Journal of Medical Internet Research, Vol.2, No.1, 2023. DOI: 10.2196/40167.

[20] Sukatemin Sukatemin, and Yuli Peristiowati, "Transformation Concept of Artificial Intelligencein the Early Identification of Tuberculosis Recovery Phase: A Systematic Literature Review: Early Identification of Tuberculosis Recovery", Indonesian Journal of Health Sciences Research and Development (IJHSRD), Vol.5, No.1, pp.30-41, 2023. DOI:10.36566/ijhsrd/Vol5.Iss1/144.

[21] Mayidili Nijiati, Ziqi Zhang, Abudoukeyoumujiang Abulizi, Hengyuan Miao, Aikebaierjiang Tuluhong, Shenwen Quan, Lin Guo, Tao Xu and Xiaoguang Zou, "Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings", Journal of X-Ray Science and Technology, Vol.29, No.5, pp.785-796, 2021. DOI:10.3233/XST-210894.

[22] Ntwali Placide Nsengiyumva, Hamidah Hussain, Olivia Oxlade, Arman Majidulla, Ahsana Nazish, Aamir Khan, Dick Menzies, Faiz Ahmad Khan, Kevin Schwartzman, "Triage of Persons With Tuberculosis Symptoms Using Artificial Intelligence-Based Chest Radiograph Interpretation: A Cost-Effectiveness Analysis", Open Forum Infect Dis, Vol.8, No.12, 2021. DOI: 10.1093/ofid/ofab567.

[23] Tamilselvi Saravanakumar, Saravana Kumar, Lavanya Selvaraj, Bindhu Jayaprakash, Kaviyavarshini Naagarrajan, "Artificial intelligence for a bio-sensored detection of tuberculosis", Network Modeling Analysis in Health Informatics and Bioinformatics, Vol.11, No.51, 2022. DOI:10.1007/s13721-022-00396-w.

[24] Calden George, "A method for evaluating the attitudes of tuberculous patients", American Review of Tuberculosis, Vol.67, No.6, pp. 722-731, 1953. DOI: 10.1016/s0041-3879(54)80016-x.

[25] Jang Kwang Sim, Jeon Gyeong Suk, "Prediction Model for Health-Related Quality of Life in Hospitalized Patients with Pulmonary Tuberculosis", Journal of Korean Academy of Nursing, Vol.47, No.1, pp.60-70, 2017. DOI: 10.4040/jkan.2017.47.1.60.

[26] Jaramillo Ernesto, "Tuberculosis and Stigma: Predictors of Prejudice Against People with Tuberculosis", Journal of Health Psychology, Vol.4, No.1, pp.71-79, 1999. DOI:10.1177/135910539900400101.

[27] Loh Shin Yee, Zakaria Rosmani, Mohamad Noraini, "Knowledge, Attitude, and Stigma on Tuberculosis and the Associated Factors for Attitude Among Tuberculosis Contacts in Malaysia", Medeniyet Medical Journal, Vol.38, No.1, pp.45-53, 2023. DOI: 10.4274/MMJ.galenos.2023.14478.

[28] Elizabeth West, Gadkowski Lara Beth, Truls Ostbye, Carla Piedrahita, Jason Stout, "Tuberculosis knowledge, attitudes, and beliefs among North Carolinians at increased risk of infection", North Carolina Medical Journal, Vol.69, No.1, pp.14-20, 2008. DOI:10.18043/ncm.69.1.14.

[29] Abiel Yehdego Kidanemariam, Betiel Yihdego Kidanemariam, Eyasu Tesfamariam, Mengisteab Embaye Gulbetu, "Community knowledge, attitude and practice towards tuberculosis in Nakfa subzone: cross-sectional study 2021", Journal of Environmental Science and Public Health, Vol.7, No.2, pp.44-56, 2023. DOI: 10.26502/jesph.96120186.

[30] Yahui Fan, Shaoru Zhang, Yan Li, Yuelu Li, Tianhua Zhang, Weiping Liu, Hualin Jiang, "Development and psychometric testing of the Knowledge, Attitudes and Practices (KAP) questionnaire among student Tuberculosis (TB) Patients (STBP-KAPQ) in China", BioMed Central Infectious Diseases, Vol.18, No.1, pp.213, 2018. DOI:10.1186/s12879-018-3122-9.

[31] Kelemework Adane, Mark Spigt, Laturnus Johanna, Dorscheidt Noortje, Semaw Ferede Abera, Geert-Jan Dinant, "Tuberculosis knowledge, attitudes, and practices among northern Ethiopian prisoners: Implications for TB control efforts", PLoS One, Vol.12, No.3, 2017. DOI: 10.1371/journal.pone.0174692.

[32] Leo Breiman, Jerome Friedman, Richard Olshen, Charles Stone, "Classification and Regression Trees", Chapman and Hall/CRC, 1984. DOI:10.1201/9781315139470.

[33] Quinlan Ross, "Induction of Decision Trees", Machine Learning, Vol.1, No.1, pp.81-106, 1986. DOI:10.1007/BF00116251.

[34] Salzberg Steven, "Book Review: C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993", Machine Learning, Vol.16, pp.235-240, 1994. DOI:10.1023/A:1022645310020.

[35] Breiman Leo, "Random Forests", Machine Learning, Vol.45, pp.5-32, 2001. DOI:10.1023/A:1010933404324.

[36] Malekipirbazari Milad, Aksakalli Vural, "Risk Assessment in Social Lending via Random Forests", Expert Systems with Applications, Vol.42, pp.4621-4631, 2015. DOI:10.1016/j.eswa.2015.02.001.

[37] James Gareth., Witten Daniela., Hastie Trevor, Tibshirani Robert, "An Introduction to Statistical Learning", Springer, New York, Vol.112, pp.3-7, 2013. DOI:10.1007/978-1-4614-7138-7.

[38] Robin Genuer, "Forêts aléatoires : aspects théoriques, sélection de variables et applications", Université Paris Sud - Paris XI, 2010.

[39] Freund Yoav, Robert Schapire, "Experiments with a New Boosting Algorithm", International Conference on Machine Learning, pp.148-156, 1996.

[40] Jean-Marc Fellous, Guillermo Sapiro, Andrew Rossi, Helen Mayberg, Michele Ferrante, "Explainable Artificial Intelligence for Neuroscience: Behavioral Neurostimulation", Front Neurosci, pp.13-1346, 2019. DOI: 10.3389/fnins.2019.01346.

[41] Alessandro Bissacco, Ming-Hsuan Yang, Stefano Soatto, "Fast human pose estimation using appearance and motion via multi-dimensional boosting regression", IEEE conference on computer vision and pattern recognition, pp.1-8, 2007.

[42] Daniiel McKenney, Michael Hutchiinson, Pia Papadopol, Keviin Lawrence, John Pedlar, Kathy Campbell, Ewa Milewska, Ron Hopkinson, David Price, Tim Owen, "Customized spatial climate models for North America", Bulletin of the American Meteorological Society, Vol.92, No.12, pp.1611-1622, 2011. DOI: 10.1175/2011BAMS3132.1.

[43] Simon Pittman, Kerry Brown, "Multi-scale approach for predicting fish species distributions across coral reef seascapes", PloS one, Vol.6, No.5, 2011. DOI:10.1371/journal.pone.0020583.
[44] Johnson Rie, Zhang Tong, "Learning nonlinear functions using regularized greedy forest", IEEE transactions on pattern analysis and machine intelligence, Vol.36, No.5, pp.942-954, 2014. DOI:10.1109/TPAMI.2013.159.
[45] Chen Tianqi, Guestrin Carlos, "XGBoost: A Scalable Tree Boosting System", Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, pp.785-794, 2016. DOI:10.1145/2939672.2939785.
[46] Mahmoud Shehadeh, Nader Ebrahimi, Abel Ochigbo, "Predicting the Type of Nanostructure Using Data Mining Techniques and Multinomial Logistic Regression", Procedia Computer Science, Vol.12, pp.392-397, 2012. DOI:10.1016/j.procs.2012.09.092.
[47] Menard Scott, "Applied Logistic Regression Analysis, 2nd ed", SAGE Publications Inc, 2002. DOI:10.4135/9781412983433.
[48] Shenglei Chen, Geoffrey Webb, Linyuan Liu, Xin Ma, "A novel selective naïve Bayes algorithm", Knowledge-Based Systems, Vol.192, 2020. DOI: 10.1016/j.knosys.2019.105361.
[49] Cortes Corinna, Vapnik Vladimir, "Support-vector networks", Machine Learning, Vol.20, pp.273–297, 1995. DOI:10.1007/BF00994018.
[50] Bernhard Chölkopf, Alexander Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond", ISBN electronic:9780262256933, 2001. DOI:10.7551/mitpress/4175.001.0001.
[51] Thenmozhi Ma'am, "Forecasting stock index returns using neural networks", Delhi Business Review, Vol.7, No.2, pp.59-69, 2006.
[52] Shuaihu Ni, Gang Chen, Jia Wang, Yuhong Li, Hui Zhang, Yan Qu, Yanlin Zhao, Xiaofeng Luo, "Assessment of public literacy in TB prevention and control in the National 13th Five-Year plan for Tuberculosis Prevention and Control (2016–2020) in China", BMC Health Services Research, Vol.25, No.1, pp.50-60, 2025. DOI:10.1186/s12913-024-12155-w.
[53] Yemin Yuan, Jin Jin, Xiuli Bi, Hong Geng, Shixue Li, Chengchao Zhou, "Gender-Specific Association Between Perceived Stigma Toward Tuberculosis and Acceptance of Preventive Treatment Among College Students With Latent Tuberculosis Infection: Cross-Sectional Analysis", Journal of Medical Internet Research, 2023. DOI: 10.2196/43972.
[54] Fentabil Getnet, Tom Forzy, Latera Tesfaye, Awoke Misganaw, Solomon Tessema Memirie, Shewayiref Geremew, Tezera Moshago Berheto, Naod Wendrad, Bantalem Yeshanew Yihun, Mizan Kiros Mirutse, Fasil Tsegaye, Mesay Hailu Dangisso, Stéphane Verguet, "Inequalities in tuberculosis control in Ethiopia: A district-level distributional modelling analysis", Tropical Medicine and Health, Vol.30, No.1, pp.31-42, 2025. DOI:10.1111/tmi.14066.
[55] Goutam Kumar Dutta, Mostafizur Rahman, Dipika Shankar Bhattacharyya, Kazi Robiul Alom, "Social Stigmatisation among Tuberculosis Patients and Community People in Bangladesh: An Exploratory Study", Preventive Medicine: Research & Reviews, 2024. DOI:10.4103/PMRR.PMRR_196_24.

Authors' Profiles

Dr. Amadou Diabagate received his Ph.D in computer science (artificial intelligence) at the Faculty of Science and Technology of Abdelmalek Essaadi University in Morocco (Tangier) in 2016. Since 2018, he has been Assistant Professor in computer science at the Faculty of Mathematics and Computer Science of Félix Houphouet-Boigny University in Abidjan, Côte d'Ivoire. He is a member of the artificial intelligence and big data research team at the Faculty of Mathematics and Computer Science. His research focuses on artificial intelligence, data science and big data. He is the President of the NGO IDTDS (Intelligent Digital Transformation and Data Strategies) which works to promote artificial intelligence and digital transformation.

Dr. Yazid Hambally Yacouba received his doctorate in December 2022 at Faculty of Mathematics and Computer Science of Félix Houphouët-Boigny University in Abidjan, Côte d'Ivoire. He is currently a higher education assistant since October 2, 2023 at the National High School of Architecture and Urban Planning Bondoukou University, Côte d'Ivoire. His current research interests include issues related to Big Data, Software engineering, IA and Data Science. He is the Head of the IT Service of the International Organization Conseil de l'Entente since 2016.

Dr. Hafizatou Sani Yanoussa received her doctorate in General Medicine in November 2016 at Abdou Moumouni University in Niamey, Niger. She also holds a specialty diploma in pulmonology obtained in October 2020 at the Faculty of Medicine of Félix Houphouët-Boigny University in Abidjan, Côte d'Ivoire. She currently works at Emy Polyclinique.

Prof. Adama Coulibaly works as a professor at the Faculty of Mathematics and Computer Science of Félix Houphouet-Boigny University. He obtained his doctorate in mathematics in July 1994 at Blaise-Pascal University in Clermont-Ferrand. He is currently Director of the Mathematics Research Institute (IRMA).

Prof. Abdellah AZMANI received his Ph.D in Industrial Computing at the University of Science and Technology of Lille in 1991. He worked as a professor at the Ecole Centrale de Lille and at the Institute of Computer and Industrial Engineering from Lens. He is a member of the Laboratory of Automatics and Informatics of Lille (LAIL). He is a professor at Faculty of Sciences and Technology of Tangier, Morocco. He has contributed to many scientific researches.

How to cite this paper: Amadou Diabagaté, Yazid Hambally Yacouba, Hafizatou Sani Yanoussa, Adama Coulibaly, Abdellah Azmani, "Machine Learning for Predicting Population Attitudes towards Tuberculosis Patients", International Journal of Intelligent Systems and Applications(IJISA), Vol.17, No.5, pp.32-48, 2025. DOI:10.5815/ijisa.2025.05.03