International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

An Optimal Pruning Fuzzy Learning Model for Analysing Risk Factors of Tuberculosis

Karabi Saikia*

Submitted: 10/09/2022 **Accepted:** 20/12/2022

Abstract:The most prevalent pathogens disease that causes death is tuberculosis (TB). Approximately 1,6 million per year or 4,384 people die every day of Tuberculosis globally, surpasses death number of HIV and Malaria combined according to WHO, 2018. Computational intelligence has the ability to tackle complex, broad, and ambiguous problems. However, existing computational techniques such as Fuzzy logic and Neural Network suffer from performance degradation when fitting the model to a guidance data set. We use computational techniques to analyse the risk factors of Tuberculos is on patient dataset publicly available in WHO portal. We apply pruning fuzzy model to reduce the network size while maintaining the fitting accuracy. To categorise input data, fuzzy C-means classifier and fuzzy inference system are utilised. To improve prediction accuracy, we use an Adaptive Neuro-fuzzy Inference System to generate fuzzy rules. The findings indicate that the suggested technique has greater precision, which meets the needs of the physicians. As a result, the created system will benefit both regular people and medical professionals.

Keywords:*Computational intelligence, Fuzzy, Pruning Fuzzy, Fuzzy Neural Network, Tuberculosis*

1. Introduction

Bacillus Mycobacterium tuberculosis causes the chronic disease Tuberclosis (TB) [1]. Through the air, TB germs can spread from one individual to another. People who have lung TB cough, sneeze, or spit into the air, spreading the TB bacteria. Only a few percentages of these microbes must be inhaled in order to cause an infection [2]. An estimated 1.5 million people died from TB worldwide in 2020. TB is the second infectious killer in the globe, after COVID-19 (behind HIV/AIDS), and the thirteenth most frequent cause of death globally [3]. In 2020, the 30 countries with the greatest incidence of TB accounted for 86% of all new cases. Two thirds of the total is accounted for by eight countries, lead by India, with China, Indonesia, Philippines, Pakistan, Nigeria, Bangladesh, and South Africa following [4]. One of the Sustainable Development Goals of the United Nations calls for the eradication of tuberculosis by 2030 (SDGs) [5].66 million lives are anticipated to be saved by TB identification and treatment between 2000 and 2020 [6]. Although TB often damages the lungs and also harm the spine, brain, intestines, and kidneys. The location of the TB germs in the body affects the TB symptoms [7]. A lifetime risk of contracting TB for those who have been exposed to the TB bacteria is between 5 and 10%. A person's chance of getting sick is increased if they have a weakened immune system, which includes those who use cigarettes, are malnourished, have diabetes or HIV, or if they are obese [8]. A person may take several months to exhibit the symptoms of active TB disease, such as a cough, fever, night sweats, or weight loss. These delays receiving medical care and lead to the infection spreading to other persons [9]. Nearly all HIV-positive and 45% of HIV-negative TB patients will deteriorate without proper care [10].

**Department of Mathematics, Dhakuakhana College, Dhakuakhana, Assam-787055,India.*

** Corresponding Author Email: karabisaikia1980@gmail.com*

In order to overcome these issues Fuzzy Expert System is developed which diagnosis TB [11]. In oriental medicinal purposes, fuzzy expert system utilised to differentiate the syndromes, and a fuzzy expert system is employed for lung disorders [12]. Applications of artificial intelligence techniques (AITs) have been developed in a variety of fields, including medicine. These applications include disease diagnosis, therapy, patient recruitment, and risk prediction [13]. Due to the high degree of complexity and uncertainty in various areas, AITs such as fuzzy logic, artificial neural networks, genetic algorithms, artificial immune systems, and several others have been developed by many academics. A particularly suitable and practical foundation like fuzzy set theory and fuzzy logic is used for creating knowledge-based systems in medicine, such as the interpretation of collections of medical results, etc., [14]. A popular branch of mathematics called fuzzy logic (FL) [15] helps create the framework for interpreting behavioural patterns in individuals. Fuzzy Set Theory (FST), which establishes the values between "partially true" and "partially false," forms the basis of the system together with "true" and "false" values. The FST aims to provide an analytical explanation for life's uncertainties, such as the terms "warm" and "cool". The majority of the time, these rules is produced by a domain expert. Input along with output values in the FES model are clear values [16], as can be seen in Fig. 1. These crisp input/output values are fuzzyfied to provide fuzzy membership values and degrees. The fuzzy inference method is used to process these produced fuzzy values. The defuzzification unit receives the fuzzy output values that were also created using the rule-base and sends them on to produce the final crisp values [17].

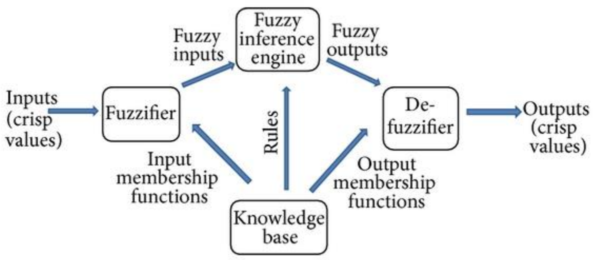


Figure 1. Fuzzy Expert System Model

2. Literature Review

In the past, a variety of methods utilising fuzzy mathematics have been developed in the field of medical research as well as modelling of diagnostic systems. Fuzzy relation equations, cluster decision making, conceptual algebra, grouping techniques, belief functions, fuzzy measurements, data theory, etc. are some examples of fuzzy mathematics techniques. Hernández-Julioet al.designed, applied, and validates a structure for the creation of fuzzy set theory-based decision support system that makes use of tables like huddle and dynamic [18]. Fuzzy inference system was built to find out breast cancer in order to validate the suggested framework. In classification issues, the usage of clusters and dynamic tables enables knowledge discovery through the identification of patterns. The likelihood of lowering input variables (feature extraction) is boosted by using these two techniques, which speeds up the decision-making process. The suggested framework might be a useful way to identify a design in the growth of classified fuzzy systems issues. The handling of Lfuzzy contexts with huge sizes can be accomplished by Alcalde&Burusco by using the Choquet integrals as helpful tools. Particularly, these operators are appropriate for the L-fuzzy context's aggregation of rows or columns. Finally, the study has established the effectiveness of the theory that has been devised to tackle the difficulty of making differential diagnoses for diseases that have a lot of symptoms [19].

On the basis of incidence of safety risk indicators amongst healthcare workers all through the COVID-19 pandemic outbreak, Rathore& Gupta determined the key hospital among five hospitals. This study suggested a decision making outline with the aid of combining fuzzy based approaches in systematic way in order to attain the aforementioned purpose. This study presented an ample collection of emotional, ergonomic, managerial, operational situation, equipment, method, and technological protection risk categories in a fuzzy based conclusion making framework [20]. For primary healthcare system, a Fuzzy Based Treatment Model (FBTM) is suggested by Akila&Balaganesh. In this method, the Fuzzy inference engine is given medical data derived since ontology model and rules in order to provide suitable treatment based on the known information[21]. Zhang et al. delved into how to identify secondary pulmonary (SPTB). A novel F3 model has been proposed by this study, where the primary F denotes the analysis of chest CT images using a four direction varying distance gray level cooccurrence matrix (FDVDGLCM), the second F denotes development of a five property feature set (FPFS) from the outcome of the FDVDGLCM, and the final F denotes the development of a fuzzy support vector machine (FSVM).The FSVM found to perform more effectively than standard SVM, and the suggested F3 model outperforms six cutting-edge SPTB recognition methods [22]. Methods like Data mining as well as fuzzy logic utilised in diabetes detection were main focal areas. It is the process of identifying patterns in large datasets by compounding various machine learning, database, and statistical techniques. Said systems are known as expert systems because fuzzy logic, which resembles human reasoning, can handle the uncertainty in medical diagnosis data. It was established that fuzzy inference systems are competent with high accuracy along with minimal difficulty because they assess the information from the existing data, which may be inaccurate, offer linguistic concepts amid significant approximation as their core to medical copy [23].

Auwalet al. applied image processing methods to evaluate the severity of tuberculosis (TB), a condition brought on by the bacterium Mycobacterium tuberculosis. The two main types of TB are latent and pulmonary. The severity level of TB, such as serene, modest, rigorous furthermore very severe, can be determined by the identification of TB utilising an upper body x-ray picture, nucleic acid enlargement test, culture, sensitivity [24]. Kukker& Sharma proposed reinforcement learning for correctly classifying pneumonia and tuberculosis (TB) using a database of X-ray images. End results demonstrate that in terms of accuracy and performance [25]. Doguet al. developed an agenda and devise a arithmetic model to help physicians to assess likelihood of antibiotic resistance. Intuitionistic fuzzy has used to represent levels of hesitation among the decision-makers in the model [26].

Reddy et al. suggested a novel method for categorising cardiac disease in 2020 that combines an adaptive genetic algorithm, fuzzy rule-based classification, and rough sets. Heart disease feature module for the model is found on rough sets. Genetic algorithm is used to optimise the rules that fuzzy classifiers generate [27]. Das et al.analysed medical diagnostics utilising computing techniques. Augmented that there were numerous corruptions in therapeutic judgement; instead of making an accurate diagnosis, the majority of practitioners chose to capture patients when they were already in a terrible stage of illness. This essay examines the drawbacks of Machine Intelligent Diagnostics, which can assist society in a normal capacity as a doctor and in times of crisis as a crisis-manager. In order to deal with the patients' emotions, a membership function was employed to determine how severe their symptoms were, and a fuzzy logic membership function was utilized to determine the likelihood that a disease would occur [28]. The study by Ontiveros-Robles et al.built General Type-2 Fuzzy Logic which was to develop General Type-2 Fuzzy Classifiers to handle ambiguous input. It has been proposed that the concept of embedded type 1 fuzzy membership role be used in the creation of broad type-2 fuzzy classifiers. Regarding the comparison of various fuzzy logic systems used to solve diagnosis problems, the study concluded that the suggested methodology could be enhanced in the future using optimization techniques [29].

A First Aid Guide by Dr.Flynxz based on Mamdani-Sugeno type Mantim Innocent et al. started the diagnosis in 2020. Three cutting-edge algorithms were introduced as part of the application of allergy treatment methods. Outcomes provided solid indication developed ES produces predictions and suggestions that are almost accurate. The outcomes also demonstrated that the fuzzy inference technique that was suggested may forecast illnesses for both positive and negative people [30]. Omisoreet al. proposed Genetic-Neuro-Fuzzy Inferential methodology to develop a resolution support platform may assist physicians in providing a precise, quick also cost-effective tuberculosis diagnosis. The cognitive and emotional filters used in this method were changed to take into account the environment that frequently affects medical professionals while using the standard and conventional ways of disease diagnosis. This study demonstrated how merging soft computing approaches can result in a system of medical diagnostics that is more precise and effective [31]. TinukeOmolewaOladeleet al. demonstrated of the useful appliance in the medical profession occurred in 2020. The doctors whose knowledge served as the foundation for the fuzzy expert system were interviewed by the authors [32].

Although fuzzy systems were capable of offering the best answer to complicated problems, their primary flaw was wholly reliant on soul understanding. Fuzzy neural network, also known as a neuro fuzzy system, is a sophisticated learning tool that uses approximation methods from neural networks to determine the parameters of a fuzzy system (such as fuzzy sets or fuzzy rules). However, when building fuzzy neural networks (FNNs), there are two key issues related to the identification of fuzzy models: structure identification and parameter adjustment. Structure identification determines the input-output space division, antecedent and consequent variables of if-then rules, such as the number of rules and beginning locations of membership functions. During parameter modification, the parameters of the premises and consequences are established. When fitting a FNN to a training data set, it is better to lower the size of the network while keeping the target fitting accuracy [33].

Evidently, determining the proper balance between the quantity of restrictions and anticipated performance is challenging. In this paper, we use a novel pruning fuzzy neural network strategy to address this problem.

3. Methods and Materials Used

3.1 Pruning fuzzy neural network

Primary idea behind the hybrid model's design, emphasizes the approaches and training algorithms employed in the model and the components found in each of its layers. This study will be used to identify patients who are candidates for immunotherapy and cryotherapy treatment. It was originally developed for pattern categorization. In Figure 1, pruning fuzzy neural network model is shown, whereby Z neurons (uni-neuron) are linked to A neurons a fuzzy neurons created by the ANFIS (Adaptive Network-based Fuzzy Inference Systems), that come together to form a singular gangliocyte with a linear initiation function. The fuzzy inference system is represented by the model's top two layers. An aggregate neural network with one neuron is represented by the third layer.

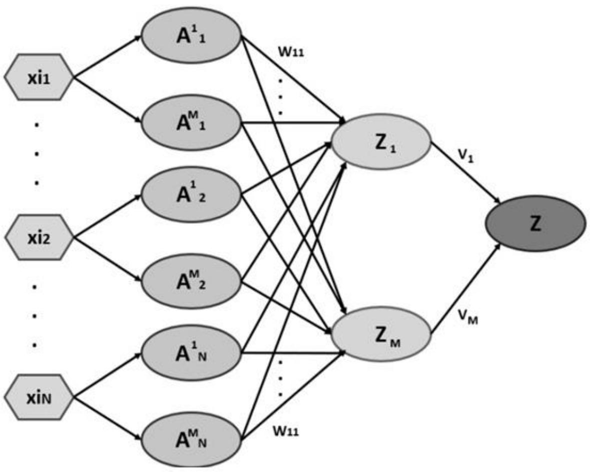


Figure 2. Pruning fuzzy neural network mode

3.2 Primary Layer

Fuzzy neural networks primary layer were employed usages principles for fuzzification that are reliant on techniques for describing the input space of ANFIS. Utilizing membership functions that are formed with equal spacing, Gaussian fuzzy neurons are produced in the primary layer. In range [1 & 1], these neurons' synaptic weight, as well as bias, are determined at random. In the second layer, these Gaussian neurons determine the significance of each neuron. Genfis1, a modification of the ANFIS strategy, generates membership functions having uniform spacing in order to partition the input space into equivalent depiction spaces. These areas provide you the chance to use different language forms to determine the logical interconnections of the input space. With the help of the genfis1 approach, it is possible to comprehend the constructed functionalities as language variables, like small, medium, and big, for instance. Despite the hidden neurons in the primary layer might rise exponentially out of control when there are too many neurons for the purpose, this method is effective for representing areas inside the solution space of input data. Each of the N dimensions of the issue is given a set of M membership functions. In systems with additional dimensions, there may be more neurons. The model creates options for the link between dimensions & membership functions at randomly when the connection among dimensions & membership functions generates more than 500 neurons. In which up to a maximum of 500 combinations, to overcome this issue. Because of this, algorithms use random combinations to minimize the complexity of neurons when the amount of dimensions is large.

3.3 Second Layer

The fuzzy logic-equipped neurons make up the next layer of the framework. Type III neurons are tasked with carrying out this responsibility because of their capability to formulate diffuse if/then rules that extract data from the data. These logical neurons use the fuzzy operator's t-norm and s-norm to cumulative the information from the neurons in the first layer. These are fuzzy set operators, and they make it possible to build representational values for the neurons in the network. Although s-norm serves as operator, neuron that employs t-norm is similarly regarded as a neuron. There's a logical operator that allows for the simultaneous application of an s-norm and a t-norm, despite the fact that two neurons are an efficient way to solve problems on their own. The logical neuron employed in this study is built on the operator known as uni-norm.

The uni-norm is articulated as

$$U(x,y)=\begin{cases} gT\left(\dfrac{x}{g},\dfrac{y}{g}\right) if \quad y \quad \in[0,g] \\ g+(1-g)S\left(\dfrac{x-g}{1-g},\dfrac{y-g}{1-g}\right) \quad ,if \quad y \quad \in(g,1] \\ \varphi(x,y), otherwise \end{cases}$$

And

$$\varphi(x,y)=\begin{cases} \max(x,y), & ,if\, g\in[0,0.5] \\ \min(x,y), & ,if\, g\in[0.5.1] \end{cases}$$

Whereas, T represents a t-norm and s-norm represented by S, also g represents individuality component with a value that may range anywhere from 0 to 1. In other words, uni-norms are able to make a seamless transition among an s-norm (when g equals zero) furthermore a t-norm (when g equals one). The t-norm operator, also known as invention, and the s-norm operator, also known as probabilistic sum, were both rethought. For the purpose of computing its output, the suggested uni-neuron carries out the following operations:

1. Every pair of values (ai, wi) combined into singular value represented by the equation bi = h (ai, wi)
2. Determine cohesive aggregated changed principles, denoted by U (b1, b2,……bn), whereas n as the total amount of inputs.

P is used to convert inputs as well as loads associated with them into values that have been individually altered (relevancy transformation). This function is said to satisfy the criteria of monotonicity in standards since the value that is converted must increase if the value that is being supplied increases. It also satisfies the requirements that no components of significance should be in play and that importance should be distributed normally. Finally, function p may ensure that the impact of $w_i$is consistent. The p function's formulation is as follows:

$$p(w,\ a)\ =\ wa\ +\ \overline{w}g,$$

The uni-neuron may be expressed as follows using the weighted aggregate described above:

$$z\ =\ \mathrm{UNI}(w;a)\ =\ U_{i=1}^{n}p(w_i,a_i).$$

The foundation of the fuzzy inference method is rules fuzzy that may be produced by these neurons. To create a crucial knowledge group for diverse difficulties, they extract the information from the test's database. This rule-base may be used to create expert systems that can help practitioners in a variety of fields, such as physicians who treat illnesses and need the information to choose the best therapies for patients.

3.4 Third Layer

The model's third layer has a neural network, and the ultimate reactions are produced by this network. A neural network for aggregation that consists of only one neuron. These neurons have an initiation purpose that might potentially deliver outcomes based on the database that was used during the training. Formula generating the model's output is:

$$y=\sum_{i=1}^{L_p} f_{linear}(z_i,v_i)$$

Whereas, z_0 = 1, v_0 is the partiality, z_j, and v_j, j = 1, ...,l_p is the output produced by every fuzzy neuron in the second layer and the load that is associated with it, accordingly. In this scenario, the activation function is linear, which indicates that the correct weight should be given to the fuzzy inputs and the weights of the additional layer.

The training is intimately connected to the procedures that generate or update the fundamental parameters that are necessary for the network to perform its intended tasks. Extreme learning machine, abbreviated as ELM, is used actively in the research that has been published Moore and Penrose's pseudo-inverse notions may be used to produce the weights analytically. The neurons in the initial layer as well as the neurons in the second layer both have weights as well as biases that were generated randomly. Weights connecting the second layer of the partial least squares-based neural network of aggregation to it. This method differs from others that accomplish the performance of the current standards using backpropagation because it uses a different algorithm. It has been shown by Huang that an improvement in the processing speed of intelligent networks may be achieved by generating final weights in the network using the pseudo-inverse. The intelligent framework is becoming autonomous from the network's recurrent update and provides replies through a higher amount of accuracy. Since it is not essential to adjust the other parameters. It is not essential to change any of the additional variables since the weights are obtained in a single step.

Because we have now specified the model, we are able to express y as Z multiplied by v, whereas y is the vector of outputs and the vector of loads v for the output layer. It has been decided that Z will:

$$Z=\begin{bmatrix} f(w_1,a_1+b_1) & \dots & f(w_m,a_1+b_1) \\ f(w_1,a_2+b_2) & \dots & f(w_m,a_2+b_2) \\ f(w_1,a_n+b_1) & \dots & f(w_m,a_n+b_1) \end{bmatrix}_{N\times 1}$$

The Single Layer Feed forward Network (SLFN) output hidden neurons are represented by Z matrix column. These outputs are measured in relation to the input.

$$a=[a_1,a_2,......,a_N]_{m\times N}^{T}$$

Hidden layer weights are randomly initialized by the ELM, w_k. The outputs layer's weights are then produced using the pseudo inverse in accordance with the formula:

$$v\ =\ Z^{+}y$$

Methods that make utilize ANFIS framework to create participation purposes with equal distance have a connection for exponential neuron formation, which is connected to the connection between membership functions & problem dimensions. These methods are used to generate membership functions with equal spacing. The amount of neurons present in fuzzy neural networks' initial layer is identical to that of the network's secondary layer. In order to lessen the impact of this logarithmic link and make pseudo-inverse determining more straightforward, the models of re-sampling make use of the regularization strategy. It is important to keep in mind that these results were obtained by following a set of hazy criteria, despite the fact that they performed poorly when compared to the data utilized in the study on TB disease. The technique presented does not involve choosing the parameters that will be employed in the pruning of useless data. The method for selecting the neurons that are the most important of the F-scores first suggested pruning structures learned using ELM, the idea of pruning is significant in the model. The F-scores approach has two unique features: its denominator represents the total of the deviations within each set of resources, and Its numerator illustrates the differentiation that may be made between positive and undesirable sets. A higher F-score reflects the discriminative power of the support better than a lower one does. The challenge that arises when determining which hidden neurons are most relevant is that it often results in the classic problem of having to choose between competing resources. The F-scores measure is used in this way in order to assess the capability of the neurons in the next layer to differentiate between various categories of patterns. Any fuzzy rules that have an F-score that is lower than the limit that has been established for elimination are regarded as being irrelevant to the matter at issue. The model that is being used to calculate this threshold does so only based on the training data, eliminating the requirement for procedures for evaluating or the computationally intensive cross-validation approaches that are used in the methods that are outlined in the proposal. The process of pruning is completed in a single step, and then loads of the network output layer are modified to increase the processing speed while maintaining the same level of predicted accuracy. The F-score is a straightforward way of measuring that actually creates input data for a high-dimensional space and thereafter selects the characteristics that are most relevant to the classes while still contributing to an evaluation of the discriminative power of the feature set's variables. This is done in order to calculate the score.

specified i^{th} feature vector (the z-vector representing uni-neurons in the case of the FNN), the total no of occurrences N, the no of positive occurrences n+, the no of negative cases n+, and the numeral of negative instances n+, ith feature's F score value is described as pursue:

$$F(i)=\frac{\left(\overline{x}_i^{(+)}-\overline{x}_i\right)^2+\left(\overline{x}_i^{(-)}-\overline{x}_i\right)^2}{\frac{1}{n_+-1}\sum_{k=1}^{n_+}\left(x_{k,i}^{(+)}-\overline{x}_i^{(+)}\right)^2+\frac{1}{n_--1}\sum_{k=1}^{n_-}\left(x_{k,i}^{(-)}-\overline{x}_i^{(-)}\right)^2}$$

Where, the average of $\overline{x}_i^{(+)}$ positive, $\overline{x}_i^{(-)}$ negative, and $\overline{x}_i$ total samples, correspondingly. The characteristic value that corresponds to the kth position in the i-faith attribute vector. Because of this, each neuron in the second layer z has been assigned an F score and newly formed neurons is made up of completely the neurons that have an f-score calculation with a value that is more important and the mean of the f-scores is comparatively less significant.

Relevance of the L neurons as of the primary layer is assessed, and the f-scores approach apprises fifty percent of the neurons that are considered to be the most important by using the mean values of the f-score in this particular instance. This evaluation is performed without making use of parameters in a single step. $L_p$refers to these pruned neurons as such throughout this body of work.

Data collected from patients with TB disease help the intelligent models understand the aspects involved in the study. The database's usage aided in the first assessments of methods for creating intelligent systems to enable the definition of remedies for effectively treating curls. To help with the prediction of treatment effectiveness, existing models were first utilized in the literature. Several models may be used with the Weka software to accurately identify various aspects of the patient profile. The outcomes of the will be compared to those of other classifiers that are often used in the research that has been published. The essential actions that need to be taken in order to carry out the categorization of designs. There are two variables to consider:

1. The no of functions associated with membership, M
2. The sort of neuron known as a uni-neuron, which is used in fuzzy logic

Algorithm: Recognition of TB along with secondary diseases using a Fuzzy Neural Network (FNN) trainings.

1. Determine M, total membership functions.
2. Find out M neurons using ANFIS feature in the initial layer.
3. Produce L fuzzy neurons using the centre and values of ANFIS and Gaussian membership functions.
4. Randomly describe the loads and partiality of the neurons.
5. Weld primary layer of L fuzzy neurons together to create L fuzzy logical neurons with random loads in addition to partiality on the network's secondary level.
6. Usage F scores to identify maximum important neurons in the issue (Lp).
7. Do this for all K inputs.
8. Using logical neurons, compute the mapping zk (xk).
9. With Equation, approximation loads of the output layer (7).
10. With Equation, compute the result y. (5).

4. Results

4.1Dataset description

Thedata set used for this work is been downloaded from the following link https://www.who.int/teams/global-tuberculosis-programme/data. This includes WHO-generated estimates of TB mortality and incidence. According to age group, sex, and risk factor, WHO TB incidence data are broken down [0.6 Mb].

4.2 Pruning fuzzy architecture

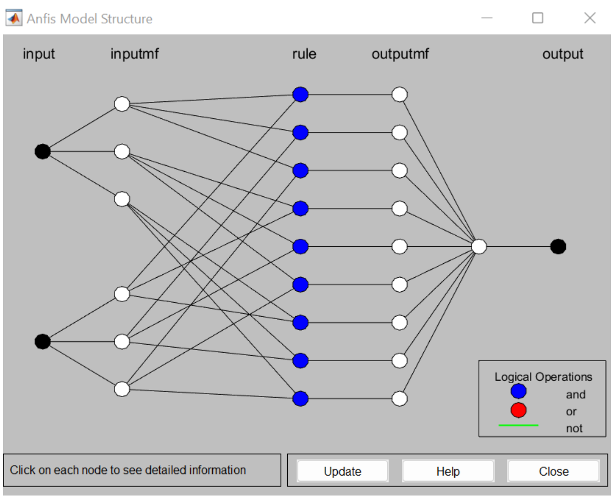


Figure 3: ANFIS Model Structure

4.3 Pruning fuzzy parameters

Total data pairs checked	**0**
Total fuzzy rules	**9**
Total linear parameters	**9**
Total nonlinear parameters	**18**
Total parameters	**27**
Total Nodes	**35**
Total training data pairs	**726**

4.4 Training data set

There are 1064 elements in the data collection. In these components, the dataset is split into training and testing, with 70% of the dataset is used. 70% of data trained with triangle membership function.

Start training ANFIS ...

1 0.299107

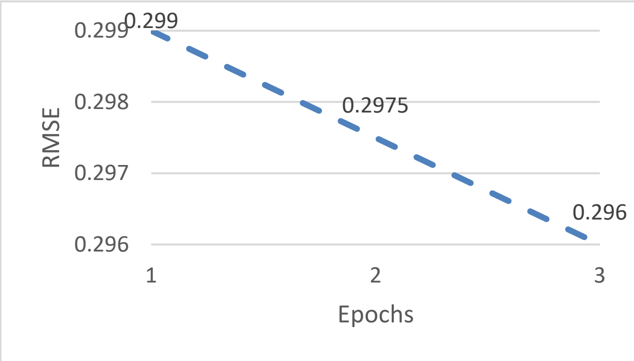
2 0.297504

Reached the selected epoch number at epoch 2, ANFIS training had been finished.

Minimal training RMSE = 0.297504

Start training ANFIS ...

1 0.297504
2 0.295885
ANFIS training was finished at epoch 2, which was the designated epoch number.
Minimal training RMSE = 0.295885



The above graph shows that pruning fuzzy training was completed at 0.2975 at epoch2.

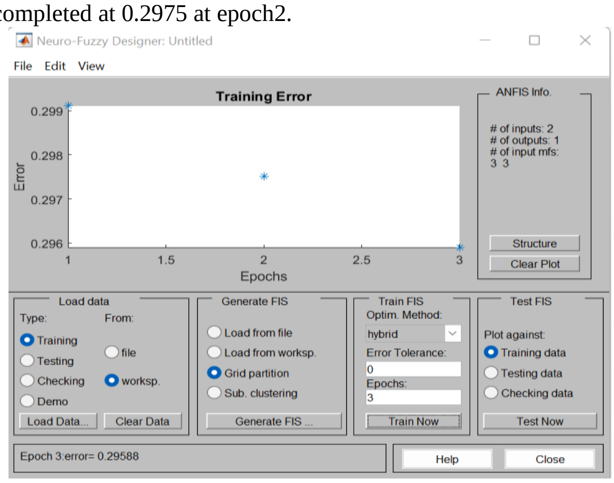


Figure 4. Training the Dataset in Neuro-Fuzzy

4.5 Testing data set

The data set consists of 1064 elements. In these elements, 30 percent of data elements are used for testing.

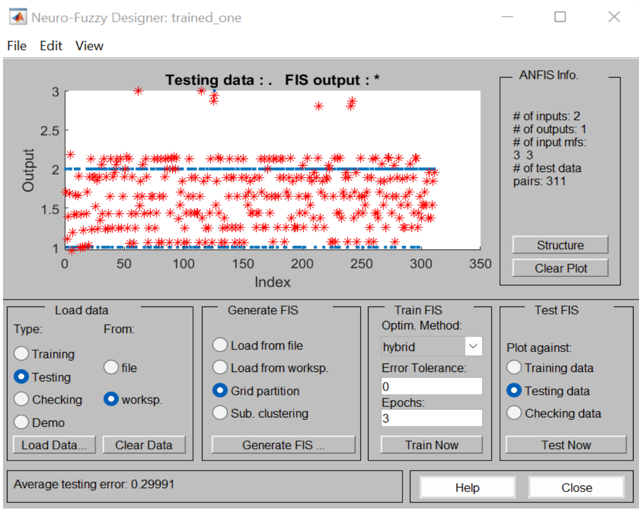


Figure 5 Testing Dataset in Neuro-Fuzzy

4.6 Groups for risk factor

The following are the group of diseases, we have given numbers for all the diseases to find the relative risk factor.

1. alc=Harmful use of alcohol;
2. dia=Diabetes;
3. hiv=HIV;
4. smk=Smoking;
5. und=Undernourishment;
6. all=All (not disaggregated by specific risk factors)

4.7Data preparation

$$\text{Relative risk factor for group 'i'} = \frac{M_i}{\sum_{\substack{j=1 \\ j\neq 1}}^{n} M_j}$$

Where M_i = Mortality of group 'i'
M_j = Mortality of group 'j'
n = Total number of groups

4.8 Risk judgment

From the below table R is the relative risk, if R is less than 0 or greater than 0.2 then the risk is low. If R is less than or equals 0.2 or greater than or equals 0.5 then the risk is moderate. If R is less than or equals 0.5 or greater than or equals 1 then the risk is high.

S.no	Rule	Risk Judgement
1	0<R<0.2	1
2	0.2<=R<=0.5	2
3	0.5<=R<=1	3

1→Low Risk
2→Moderate Risk
3→High Risk

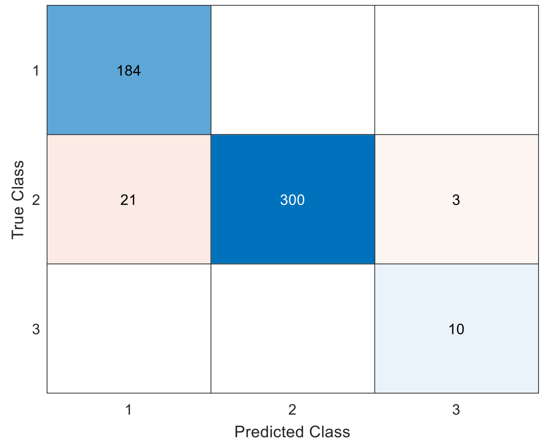


Figure 6. Confusion Matrix of Pruning Fuzzy

Figure 6 shows the confusion matrix of pruning fuzzy to determine the risk factor. This confusion matrix displays 2 classes where 1 is the predicted class and 0 is the true class. It is observed that 184 times,300 times, and 10 times the proposed pruning fuzzy predicted accurately. There are some misclassifications where the model had misinterpreted the risk factor.

5. Conclusion

The most lethal infectious disease TB kills 1.5 million people annually. Despite the fact that tuberculosis can be found anywhere in the world, most patients live in low as well as middle income countries. Eight nations account for around half of all TB cases: Bangladesh, China, India, Indonesia, Nigeria, Pakistan, the Philippines, South Africa, and Indonesia. According to the World Health Organization, tuberculosis kills more people compared to HIV and malaria combined. The use of computational knowledge to assist with this challenge is in great demand. The risk factor is estimated in this study as a ratio of mortality with TB illness plus secondary diseases. The risk element for system design is the emphasis of the pruning fuzzy neural network that we have proposed. We identified the superfluous risk factors that are less likely to result in TB in the patient dataset using some of the computations. TB illnesses are categorized into categories based on secondary disorders such as Diabetes, harmful use of alcohol, HIV, smoking, and undernourishment. We have assigned numbers to each secondary condition in order to determine the proportional risk factor. Following that, the risk factor for each group is assessed, with a value ranging between (0 to 1). The danger is minimal if R is less than 0 or larger than 0.2. The risk is considerable if R is less than or equal to 0.2 or more than or equal to 0.5. The danger is considerable if R is less than or equal to 0.5 or more than or equal to 1. This severity level is used to determine whether a patient has a group of disorders in addition to TB and to compute the mortality ratio. The evaluation of the performance metrics shown that the proposed system had better learning capacity, generalization capacity and higher accuracy.

References

[1] N. Awoke, B. Dulo and F. Wudneh, "Total Delay in Treatment of Tuberculosis and Associated Factors among New Pulmonary TB Patients in Selected Health Facilities of Gedeo Zone, Southern Ethiopia, 2017/18", Interdisciplinary Perspectives on Infectious Diseases, vol. 2019, pp. 1-14, 2019. Available: 10.1155/2019/2154240.

[2] B. Patterson and R. Wood, "Is cough really necessary for TB transmission?", Tuberculosis, vol. 117, pp. 31-35, 2019. Available: 10.1016/j.tube.2019.05.003.

[3] A. Kizny Gordon, B. Marais, T. M. Walker, and V. Sintchenko, “Clinical and public health utility of mycobacterium tuberculosis whole genome sequencing,” International Journal of Infectious Diseases, vol. 113, pp. S40-S42, 2021.

[4] M.-R. Lee, H.-L. Huang, L.-C. Chen, H.-C. Yang, J.-C. Ko, M.-H. Cheng, I.-W. Chong, L.-N. Lee, J.-Y. Wang, and G. Dimopoulos, “Seroprevalence of Aspergillus IGG and disease prevalence of chronic pulmonary aspergillosis in a country with intermediate burden of tuberculosis: A prospective observational study,” Clinical Microbiology and Infection, vol. 26, no. 8, pp.1091-e1, 2020.

[5] R. Perrigot and K. ElomGbetchi, “Social franchise chains operating in African countries: Are their social goals aligned with the 2030 United Nations Sustainable Development Goals?,” International Journal of Emerging Markets, 2022.

[6] P. J. Dodd, C. M. Yuen, S. M. Jayasooriya, M. M. van der Zalm, and J. A. Seddon, “Quantifying the global number of tuberculosis survivors: A modelling study,” The Lancet Infectious Diseases, vol. 21, no. 7, pp. 984–992, 2021.

[7] A. Sharma, A. Sharma, R. Malhotra, P. Singh, R. K. Chakrabortty, S. Mahajan, and A. K. Pandit, “An accurate artificial intelligence system for the detection of pulmonary and extra pulmonary tuberculosis,” Tuberculosis, vol. 131, p. 102143, 2021.

[8] K. D. Dale, M. Karmakar, K. J. Snow, D. Menzies, J. M. Trauer, and J. T. Denholm, “Quantifying the rates of late reactivation tuberculosis: A systematic review,” The Lancet Infectious Diseases, vol. 21, no. 10, pp. e303-e317, 2021.

[9] Dale, K.D., Karmakar, M., Snow, K.J., Menzies, D., Trauer, J.M. and Denholm, J.T., 2021. Quantifying the rates of late reactivation tuberculosis: a systematic review. The Lancet Infectious Diseases, 21(10),.

[10] B. Long, S. Y. Liang, A. Koyfman, and M. Gottlieb, “Tuberculosis: A focused review for the Emergency Medicine Clinician,” The American Journal of Emergency Medicine, vol. 38, no. 5, pp. 1014–1022, 2020.

[11] L. Chelkeba, G. Fekadu, G. Tesfaye, F. Belayneh, T. Melaku, and Z. Mekonnen, “Effects of time of initiation of antiretroviral therapy in the treatment of patients with HIV/TB co-infection: A systemic review and meta-analysis,” Annals of Medicine and Surgery, vol. 55, pp. 148–158, 2020.

[12] M. I. Fale and Y. G. Abdulsalam, “Dr. Flynxz – a first aid Mamdani-SUGENO-type fuzzy expert system for differential symptoms-based diagnosis,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 4, pp. 1138–1149, 2022.

[13] S. Varela-Santos and P. Melin, “A new modular neural network approach with fuzzy response integration for lung disease classification based on multiple objective feature optimization in chest X-ray images,” Expert Systems with Applications, vol. 168, p. 114361, 2021.

[14] B. M. Yahya, F. S. Yahya, and R. G. Thannoun, “Covid-19 prediction analysis using artificial intelligence procedures and GIS Spatial Analyst: A case study for Iraq,” Applied Geomatics, vol. 13, no. 3, pp. 481–491, 2021.

[15] Z. Khalid, G. Khan, and M. A. Arbab, “Extrinsically evolved system for breast cancer detection,” Evolutionary Intelligence, 2022.

[16] L. J. Muhammad and E. A. Algehyne, “Fuzzy based expert system for diagnosis of coronary artery disease in Nigeria,” Health and Technology, vol. 11, no. 2, pp. 319–329, 2021.

[17] R. Kunhimangalam, S. Ovallath, and P. K. Joseph, “A novel fuzzy expert system for the identification of severity of carpal tunnel syndrome,” BioMed Research International, vol. 2013, pp. 1–12, 2013.

[18] D. Rasi and S. N. Deepa, “Energy optimization of internet of things in wireless sensor network models using type-2 Fuzzy Neural Systems,” International Journal of Communication Systems, vol. 34, no. 17, 2021.

[19] Hernández-Julio, Y.F., Hernández, H.M., Guzmán, J.D.C., Nieto-Bernal, W., Díaz, R.R.G. and Ferraz, P.P., 2019, February. Fuzzy knowledge discovery and decision-making through clustering and dynamic tables: Application in medicine. In International Conference on Information Technology & Systems (pp. 122-130). Springer, Cham.

[20] C. Alcalde and A. Burusco, "Reduction of the size of L-fuzzy contexts. A tool for differential diagnoses of diseases", International Journal of General Systems, vol. 48, no. 7, pp. 692-712, 2019. Available: 10.1080/03081079.2019.1620740.

[21] B. Rathore and R. Gupta, "A fuzzy based hybrid decision-making framework to examine the safety risk factors of healthcare workers during COVID-19 outbreak", Journal of Decision Systems, vol. 31, no. 1-2, pp. 68-101, 2020. Available: 10.1080/12460125.2020.1862988.

[22] Akila, D. and Balaganesh, D., 2021, April. Fuzzy Based Treatment Model for Critical Healthcare System. In 2021 2nd International

Conference on Intelligent Engineering and Management (ICIEM) (pp. 193-198). IEEE.
[23] Y. Zhang, W. Wang, X. Zhang and S. Wang, "Secondary Pulmonary Tuberculosis Recognition by 4-Direction Varying-Distance GLCM and Fuzzy SVM", Mobile Networks and Applications, 2022. Available: 10.1007/s11036-021-01901-7.
[24] H. Thakkar, V. Shah, H. Yagnik and M. Shah, "Comparative anatomization of data mining and fuzzy logic techniques used in diabetes prognosis", Clinical eHealth, vol. 4, pp. 12-23, 2021. Available: 10.1016/j.ceh.2020.11.001.
[25] N. Auwal, I. Goni, D. Ali, U. Ngene and I. Manga, "Image Processing Approach to Determine the Severity Level of Tuberculosis", Current Journal of Applied Science and Technology, vol. 37, no. 3, pp. 1-8, 2019.
[26] A. Kukker and R. Sharma, "Modified Fuzzy Q Learning Based Classifier for Pneumonia and Tuberculosis", IRBM, vol. 42, no. 5, pp. 369-377, 2021. Available: 10.1016/j.irbm.2020.10.005.
[27] E. Dogu, Y. Albayrak and E. Tuncay, "Multidrug-resistant tuberculosis risk factors assessment with intuitionistic fuzzy cognitive maps", Journal of Intelligent & Fuzzy Systems, vol. 38, no. 1, pp. 1083-1095, 2020. Available: 10.3233/jifs-179470.
[28] G. Reddy, M. Reddy, K. Lakshmanna, D. Rajput, R. Kaluri and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis", Evolutionary Intelligence, vol. 13, no. 2, pp. 185-196, 2019. Available: 10.1007/s12065-019-00327-1.
[29] S. Das and M. Sanyal, "Machine intelligent diagnostic system (MIDs): an instance of medical diagnosis of tuberculosis", Neural Computing and Applications, vol. 32, no. 19, pp. 15585-15595, 2020. Available: 10.1007/s00521-020-04894-8.
[30] E. Ontiveros-Robles and P. Melin, "Toward a development of general type-2 fuzzy classifiers applied in diagnosis problems through embedded type-1 fuzzy classifiers", Soft Computing, vol. 24, no. 1, pp. 83-99, 2019. Available: 10.1007/s00500-019-04157-2.
[31] M. Fale and Y. Abdulsalam, "Dr. Flynxz – A First Aid Mamdani-Sugeno-type fuzzy expert system for differential symptoms-based diagnosis", Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 4, pp. 1138-1149, 2022. Available: 10.1016/j.jksuci.2020.04.016.
[32] M. Omisore, O. Samuel and E. Atajeromavwo, "A Genetic-Neuro-Fuzzy inferential model for diagnosis of tuberculosis", Applied Computing and Informatics, vol. 13, no. 1, pp. 27-37, 2017. Available: 10.1016/j.aci.2015.06.001.
[33] T. Oladele, R. Ogundokun, J. Awotunde, M. Adebiyi and J. Adeniyi, "Diagmal: A Malaria Coactive Neuro-Fuzzy Expert System", Computational Science and Its Applications – ICCSA 2020, pp. 428-441, 2020. Available: 10.1007/978-3-030-58817-5_32 [Accessed 20 September 2022].
[34] J. Du and M. J. Er, "A fast pruning algorithm for an Efficient Adaptive Fuzzy Neural Network," IEEE ICCA 2010, 2010, pp. 1030-1035, doi: 10.1109/ICCA.2010.5524417.