A Novel Approach For Detecting Tuberculosis Based On Observed Manifestations Using Supervised Machine Learning

Solomon Osarumwense Alile[1]
[1]Computer Science Department, Faculty of Physical Sciences,
University of Benin, Benin City, Nigeria
e-mail: solomon.alile@physci.uniben.edu
Tel.: +234-70364-44086

***Abstract*— Tuberculosis (TB) is a severe communicable respiratory tract infection caused by Mycobacterium Tuberculosis (MTB), thus infiltrating and disrupting the functionality of the lungs which is the major spot of infection. Tuberculosis is classified as a infectious disease transmitted via diminutive respiratory droplets let free into the air through coughs and sneezes from symptomatic patients which when inhaled by asymptomatic patients; it can be transmitted to these individuals inside the first few weeks, sometimes months leading to years before the manifestation of the infection begin to show in these infected patients with organs resident in their bodies such as kidneys, bladder and liver just name a few; being are at risk due to the spread of the infection. Owing to the widespread mode of TB infection, it was categorized as a global health pandemic by World Health Organization in 1993, which to date remains the world foremost infectious killer disease which has caused millions of untimely passings of infected persons. What's more, the manifestations of TB are hemoptysis (cough with sputum covered with blood), angina (chest pains), breathlessness, weakness, weight loss, fever, night sweats and loss of appetite just to name but a couple. Nevertheless, in recent past, several systems have been developed to detect this transmittable ailment, yet they delivered a ton of bogus negative during testing and couldn't distinguish Tuberculosis in view of its covering symptoms it imparts to other Respiratory Tract Infections (RTIs). Consequently, there was the need to proffer a solution for the quagmire of under-diagnosis and misdiagnosis of Tuberculosis which is much uncontrolled in Sub-Sahara Africa and South-East Asia respectively. Hence, in this paper, we proposed and developed a model to predict Tuberculosis and Respiratory Tract Infections using an AI technique called Bayesian Belief Network. The model was structured using Bayes Server and tested with data retrieved from Tuberculosis machine learning repository. The model had an overall prediction exactness of 99.98%; 99.84% and 99.96% sensitivity of Tuberculosis and Respiratory Tract infections in that order.**

***Keywords*—** Tuberculosis; Mycobacterium Tuberculosis; Prediction; Detection; Artificial Intelligence, Supervised Machine Learning; Bayesian Belief Network.

I. INTRODUCTION

Respiration is a procedure of inhaling oxygen to cells contained in tissues and exhaling carbon dioxide via the same medium in individuals which is regulated by a biological system resident in the human body called the respiratory system [1].

Respiratory system is an organic system which comprises of precise organs and frameworks employed for the sole aim of exchanging gases in the mode of oxygen (O) and carbon dioxide (CO_2) in humans, animals (mammals, birds, reptiles) and plants (trees, grasses and shrubs etc) [2]. Conversely, the respiratory system has organs resident within the system such as the trachea, the diaphragm and the lungs. Of the aforesaid organs of the respiratory system, the lungs are the most essential organ of all.

Lungs are liable for the retrieval of oxygen from the air we breathe thus conveying it to the cells in the blood vessels and expulsion of carbon dioxide when breathing out. In any case, regardless of the function of the lungs in everyday running of the respiratory system, it is susceptible to infections and diseases which can hamper the functionality of the system such as chronic bronchitis, pneumonia, asthma, sarcoidosis, lung cancer, cystic fibrosis/bronchiectasis, pneumonia, pleural effusion, tuberculosis and chronic obstructive pulmonary disease and tuberculosis infection. Of the previously mentioned ailments, Tuberculosis is the most dreadful of all.

Tuberculosis (TB) is a death-defying transmittable disease that generally affects the lungs and poses threat to other parts of the human body brought about by the Mycobacterium Tuberculosis (MTB). Moreover, Mycobacterium Tuberculosis (MTB) is transmitted from one individual to another via miniature respiratory droplets set free into the air through coughs and sneezes from symptomatic patients. The manifestations of this disease are hemoptysis (cough with sputum covered with blood), angina (chest pains), breathlessness, weakness, weight loss, fever, night sweats and loss of appetite just to name but a couple [3].

What's more, TB infections are categorized into two namely: Latent and Active Tuberculosis respectively with the sole aim of recognizing the conditions that comes with each form of the TB infection. Yet, men are majorly at risk of contracting and passing from TB infection compared to women; though TB can be predominantly ruthless in women specifically during their childbearing years (15 to 49 years) and during pregnancy. Then again, individuals of all ages are at risk of the TB infection. [4].

Latent TB is a type of TB infection that lingers in the human body in motionless state with no symptoms whatsoever evident. More so, this class of patients with this form of TB infection is classified as been asymptomatic. However, latent TB is not infectious but rather has potentials of transforming into active TB. Besides, patients with this kind of infection stand at approximately 2 billion of the world's population. In any case, it was further affirmed that 10% of the aforesaid estimated population sporadically advances to active TB if not diagnosed early and treated resulting in deaths of infected patients.

Active TB is a form of TB infection that can be due to the transformation from latent TB Infection; more so this condition makes one ill and majority of this kind of cases could lead to widespread of the TB infection in the very few weeks or sometimes years after contact with the MTB. Eventually, when the TB infection turns into active TB, the lungs are the most affected accounting for 90% of active TB infectivity to date [5].

In 2018, over 10 million confirmed cases of active TB were reported causing about 1.5 million passings; thus, classified as the foremost cause of death from contagious diseases in that year. Additionally, major cases of TB cases were recorded in different regions of the world, for instance, South-East Asia (44%), Africa (24%) and the Western Pacific (18%), with above half of confirmed cases being detected in eight nations to be specific: India, China, Indonesia, the Philippines, Pakistan, Nigeria and Bangladesh amounting to (27%), (9%), (8%), (6%), (6%), (4%) and (4%) cases respectively in the aforementioned countries attributable to under-analysis and misdiagnosis of TB infections [6].

Owing to the widespread mode of TB infection, it was categorized as an epidemic (global health emergency) by World Health organization in 1993, while it remains the world foremost infectious killer disease to date; even as efforts are being made to curb the disease by the year 2030.

Despite the epidemic nature of TB, clinical methods have been utilized in diagnosing this disease such as chest X-ray, multiple sputum cultures for acid-fast bacilli (AFB), Interferon-γ release assays, tuberculin skin tests, Nucleic acid amplification tests (NAATs) and adenosine deaminase testing. Conversely this usage of the aforementioned diagnostic methods exposes patients to radiation which has side effects such as vomiting, bleeding, fainting in the case of chest X-ray usage; acid fast bacilli test is deficient in making a distinction between MTB from NTM (Non-tuberculosis Mycobacterium); Interferon-γ release assays are capital-intensive and requires a ton of research facility resources to conduct the test; tuberculin skin tests lacks the ability to differentiate latent TB from active TB; nucleic acid amplification tests availability is limited and cost-intensive and adenosine deaminase testing requires chemical substance preparation; hence, it can only be conducted manually.

Subsequently, a lot of false negatives are produced as a result of usage of the above-listed diagnostic methods due to the overlapping symptoms Tuberculosis has with other respiratory tract infections leading to misdiagnosis of the aforesaid disease, with several of the aforesaid diagnostic methods quite invasive, risky and capital-intensive. Thus, there is the need to proffer a solution to assist in diagnosing Tuberculosis. Hence, the use of artificial intelligence (AI) is viewed as a non-invasive method which will assist in curbing the menace of misdiagnosis and untimely passings of patients recorded due to Tuberculosis infection.

Then again, several machine learning techniques have been utilized diagnosing Tuberculosis in the works of [7,8,9,10,11,12,13,14,15,16,17,18,19 and 20] but they generated a lot of false negative during testing and were unable to detect Tuberculosis due to the overlapping symptoms the disease shares with some respiratory tract infections.

In this paper, a managed AI technique called Bayesian Belief Network (BBN) was used in diagnosing Tuberculosis dependent on observed manifestations. BBN is a multifaceted probabilistic network that unites expert knowledge and observed datasets. It plans a course for conditions and intelligent outcomes relationship among factors and sets them up with probability that exhibits the level where one variable is most likely going to impact another. Be that as it may, BBN was our technique of choice because of its ability to make prescient surmising. The chosen approach uses Bayes theorem which is a statistical technique that guides high accuracy in predicting and detecting events and its occurrences.

One significant feature the proposed solution has over existing systems is its ability to diagnose Tuberculosis likewise the covering side effects this ailment has with other with Respiratory Tract Infections (RTIs) which will realize enhancement in the accompanying areas such as prediction, detection and diagnosis of Tuberculosis with possible overlapping symptoms identified with Respiratory Tract Infections.

In any case, the rest of the paper is sorted out as follows: Section II contains the related works on Tuberculosis and Respiratory Tract diseases diagnosis utilizing AI, Section III clarifies the chosen supervised AI method (Bayesian Belief Network) used in diagnosing Tuberculosis and Respiratory Tract Infections, Section IV contains the simulation, results and discussion and Section V finishes up research work with future directions.

II. RELATED WORK

Several studies have been conducted on diagnosing Tuberculosis (TB) using Artificial Intelligence. In [7], a specialized expert system for diagnosing lung diseases using Fuzzy logic (FL) as an approach was developed. The system diagnosed Lung diseases with high discovery accuracy. However, the system had the following accompanying quandaries such as complicatedness to identify respiratory tract diseases with overlapping symptoms as Tuberculosis. Besides, fuzzy expert systems have concern of real-time responsiveness, difficulty in making bi-directional inferences; fuzzy systems don't have the potentials of machine learning and neural network type pattern identification.

In [8], a dedicated expert system for forecasting results of Tuberculosis treatment using Fuzzy Logic was developed. What's more, the expert system predicted outcomes of Tuberculosis treatment with high recognition exactness. Nonetheless, the specialized system had the following dilemmas for instance, the inability of the specialized expert system to make bi-directional surmisings, problem of real-time sensitivity; fuzzy systems don't have the competence of machine learning and neural network type pattern recognition. In addition, the system failed to detect respiratory tract diseases with overlapping indications as Tuberculosis.

In [9], a novel system for diagnosing respiratory diseases using rules based on Fuzzy Logic was created. The system diagnosed only 3 respiratory diseases namely pneumonia, tuberculosis and normal influenza with high discovery precision. On the other hand, the system had the following accompanying issues for example, intricacy to recognize respiratory tract diseases with overlapping manifestations as Tuberculosis, complexity in making bi-directional deductions and issue of real-time responsiveness. All the same, fuzzy systems don't have the possibilities of machine learning and neural network type pattern classification.

In [10], a dedicated decision support expert system for recognizing Tuberculosis based on Fuzzy Logic was designed. The expert system identified Tuberculosis with high detection exactness aiding prompt diagnosis of the aforesaid disease. Then again, the system had the following inadequacies such as difficulty in responding in real-time and making bi-directional assumptions; failure to distinguish respiratory tract diseases with overlapping symptoms as Tuberculosis. Still, fuzzy systems don't have the skills of machine learning and neural network type pattern recognition.

In [11], a specialized expert system for Tuberculosis diagnosis which was based on Fuzzy Logic was developed. The dedicated expert system provided a decision support platform for medical professionals as regards Tuberculosis diagnosis with high identification precision. Conversely, the system had the following accompanying quagmires such as the inability to detect respiratory tract diseases with covering side effects as Tuberculosis, failure to respond in real-time and make bi-directional inferences. Yet, the developed fuzzy system does not have the capability of machine learning and neural network type pattern identification.

In [12], a proficient expert system that diagnosed Tuberculosis using Fuzzy Cluster Means (FCM) as an approach was developed. The specialized expert system classified Tuberculosis with high detection exactness. Even so, the system had the following accompanying inadequacies such as difficulty in distinguishing respiratory tract diseases with overlapping side effects as Tuberculosis. Also, Fuzzy C Means as a technique has concern of handling high datasets; yet, it is vulnerable to initialization and with ease gets trapped in the local optima.

In [13], a specialized diagnostic decision support expert system based on rules generated using Fuzzy Logic was created. The expert system diagnosed classes of Tuberculosis with high discovery precision. Nevertheless, the system had the following dilemmas for example, the intricacy in discovering respiratory tract diseases with overlapping symptoms as Tuberculosis, inability to make bi-directional conclusions and real-time responsiveness. In addition, the expert system lacks ability of machine learning and neural network type pattern classification; fuzzy systems are deficient in tackling uncertainties owing to unawareness, incompleteness and randomness.

In [14], an expert system and architecture for analyzing Lung infections based on rules created using Fuzzy Logic as an approach was proposed. The system detected lung diseases with high detection accuracy. However, the system had some accompanying issues such as complicatedness to identify respiratory tract diseases with overlapping side effects as lung diseases like Tuberculosis, the proposed system failed in making bi-directional deductions and as well lacks real-time responsiveness. Moreover, the expert system is short of the prospects of machine learning and neural network type pattern recognition.

In [15], a decision support expert system for recognizing Tuberculosis that employed Fuzzy logic was created. The system identified Tuberculosis with high discovery exactness. On the other hand, the system had the following accompanying inadequacies for instance: the expert system failed to detect respiratory tract diseases with overlapping side effects as Tuberculosis, powerlessness of the expert system to make bi-directional inferences and concern of real-time responsiveness. In addition, fuzzy expert systems don't have the competence of machine learning and neural network type pattern discovery.

In [16], a dedicated decision support expert system for diagnosing Tuberculosis bacterium class using Fuzzy logic as an approach was developed. The expert system diagnosed Tuberculosis causing bacterium with high detection accuracy. What's more, the expert system had the following accompanying inadequacies such as: the inability of the expert system to make bi-directional deductions, worry of real-time responsiveness and concern of identifying respiratory tract diseases with overlapping symptoms as Tuberculosis. Additionally, fuzzy expert systems lack the capability of machine learning and neural network type pattern detection.

In [17], a hybrid system for diagnosing Pulmonary Tuberculosis using a merger of Neural Networks and Fuzzy Logic called Adaptive Neuro Fuzzy Inference System (ANFIS) was developed. The expert system diagnosed Tuberculosis with high discovery precision. What's more, the system had the following drawbacks such as the difficulty in understanding the outcomes acquired from the learning process of the neural network; the system learning process is time-consuming and the system neural network outcome cannot be verified to see if it is credible because of its black box nature. Also, the system failed to detect respiratory tract diseases with covering manifestations as Tuberculosis.

In [18], Fuzzy logic as an approach was utilized in the development of an expert for diagnosing and treating Tuberculosis. The specialized system diagnosed Tuberculosis with high detection exactness. However, the system failed to identify respiratory tract diseases with overlapping symptoms as Tuberculosis. In addition, fuzzy systems don't have the competence of machine learning and neural network type pattern classification; issue of real time responsiveness and intricacy in making bi-directional inferences. Be that as it may, fuzzy systems are deficient in handling uncertainties owing to unawareness, incompleteness and randomness.

In [19], an intelligent system that diagnosed Tuberculosis using Adaptive Neuro Fuzzy Inference System (ANFIS) was developed. The system diagnosed Tuberculosis with 99.58% detection exactness. More so, the system had the following accompanying drawbacks such as inability of the system to recognize respiratory tract diseases with overlapping symptoms as Tuberculosis. Also, there exist the intricacies of comprehending the solutions derived from the learning process of the neural network. Yet, the learning process is time-consuming with the system neural network solution not easily verified to see its credibility because of its black box nature.

In [20], an Adaptive Neuro Fuzzy Inference System was developed with the solitary aim of diagnosing Tuberculosis. The hybrid system diagnosed Tuberculosis with high discovery precision. What's more, the system had the following downsides such as the difficulty in understanding the result attained from the learning process of the neural network; the learning process is lingering and the system neural network outcome cannot be established to see if it is credible. In addition, the system failed to detect respiratory tract diseases with overlapping symptoms as Tuberculosis.

III. METHODOLOGY

In this paper, the technique we intend to employ towards achieving the aim of this study which is diagnosing Tuberculosis, Respiratory Tract Infections as well as the overlapping symptoms they have in common; is the utilization of a non-invasive AI (Artificial Method) method called Machine learning.

Machine learning is a coalition of strategies for creating models that depicts or forecast utilizing data or past experience. Even so, there are a few techniques of AI namely Supervised Learning: it trains data and integrates required results (for instance, Bayesian Belief Networks, Neural Networks, Deep learning and so on.), Unsupervised Learning: it trains data and does leave out wanted end results (for example Grouping, Dimensionality Reduction), Semi-Supervised Learning: it trains data and hardly slots in any ideal outcome and Reinforcement Learning: it gains from series of activities (Temporal Difference Learning, Q-learning) [21].

In this paper, we expect to utilize an AI method called Bayesian Belief Network because of its prescient ability dependent on past experience and example data available to its during training and testing of observed datasets. Bayesian Belief Network (BBN) is directed acyclic graphical model that utilizes probability to show conditional dependencies that prevails among nodes on a graph [22]. It is a complex probabilistic network that blends expert information and investigative datasets. It structures out course of conditions and consistent outcomes associations among factors and encodes them with probability that indicates the amount wherein one variable is conceivable to affect another Furthermore, Bayesian Belief Network strives on the Bayes theorem which is relies on probability.

The Bayes theorem is represented in the mathematical equation below:

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)} \quad (1)$$

Where,

P(a) is the probability of event "a" happening without any information about event "b". It is called the "Prior".

P(a/b) is the conditional probability of event "a" happening given that event "b" has already occurred. It is otherwise called the "Posterior".

P(b/a) is the conditional probability of event "b" happening given that event "a" has already occurred. It is called the "Likelihood".

P(b) is the probability of event "b" happening without any information about event "a". It is called the "Marginal Likelihood".

The Naive Bayes classifiers are regularly spoken to as a type of directed acyclic graph (DAG). The Directed Acyclic Graph (DAG) comprises of vertices representing random variables and arrows connecting pairs of nodes.

Figure 1 shows a pictorial representation of a Bayesian Belief Network.

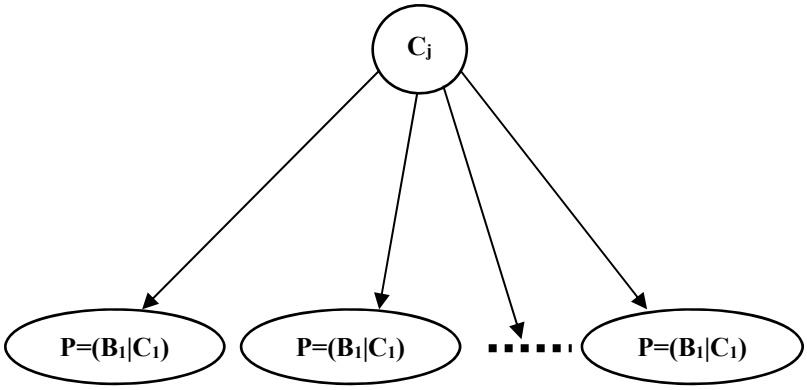


Figure 1. A Pictorial Representation of a Bayesian Belief Network

A couple of inclinations of this model are: it is exceptionally fast in making inductions, the subsequent probabilities are very simple to decode, the learning algorithm is clear and the model adequately solidifies with utility functions to make ideal surmisings. In this paper, we expect to recognize Tuberculosis, Respiratory Tract Infections and their symptoms utilizing a managed AI procedure called Bayesian Belief Network (BBN). A model comprising of 52 nodes where a few nodes speak to a type of ailment illness or elements that impact diagnosis of Tuberculosis, Respiratory Tract Infections and their side effects will be designed utilizing Bayes Server. A Tuberculosis dataset will be utilized to train and test the system. Utilizing the Pareto Principle, 80% of the dataset will be utilized to prepare the model while the remainder will be utilized in testing the model. The point of the model is to accomplish a high of identification precision with the utilization of the covering indications of Respiratory Tract Infections and Tuberculosis.

IV. SIMULATION, RESULTS AND DISCUSSION

The simulation was performed utilizing a Tuberculosis dataset in training, testing and predicting Tuberculosis which was retrieved from [23]. Besides, previews of used dataset, designed BBN model for predicting Tuberculosis, Respiratory Tract Infections (RTIs) and its manifestations, BBN model convergence chart, loglikelihood batch query chart, feature importance of nodes chart, in-sample anomaly detection chart, likelihood plots of ailments being cause of Respiratory Tract Infections and Tuberculosis, loglikelihood graph for detecting Tuberculosis and likelihood against loglikelihood plot for predicting Tuberculosis, Respiratory Tract Infections and symptoms were taken during the simulation process which appear beneath in figures 2, 3, 4, 5, 6, 7, 8,9,10, 11 and 12 respectively with the results discussed underneath the diagrams. On the other hand, the used dataset include mix of ailment afflictions and factors thought about in the recognition of Tuberculosis, Respiratory Tract Infections and symptoms amounting to 52 with each sickness and factor having a value which addresses the probability of such ailment malady and factor causing Tuberculosis and Respiratory Tract Infections.

The sicknesses and components are: Active Tuberculosis, Adult-Onset Asthma, Allergic Asthma, Angina, Asthma, Asthma-COPD Overlap, Breathlessness, Bronchitis, Chills, Chronic Obstructive Pulmonary Disease (COPD), Cough, Emphysema, Fatigue, Fever, Hemoptysis, Influenza, Laryngitis, Latent Tuberculosis, Loss of Appetite, Lower Respiratory Infections, Lung Abscess, Lung Cancer, Lung Nodules, Metastatic Lung Cancer, Migraine, Muscle Pain, Mycobaterium Tuberculosis Bacteria, Nausea, Night Sweats, Non-Allergic Asthma, Non-Small Cell Lung Cancer, Occupational Asthma, Pertussis, Phlegm, Pneumonia, Respiratory Tract Infections (RTIs), Runny Nose, Severe Angina, Sinus Infection, Small Cell Lung Cancer, Sneezing, Sore Throat, Stable Angina, Swine Flu, Tonsillitis, Tuberculosis, Unstable Angina, Upper Respiratory Infections, Variant Angina Pectoris, Weakness, Weight Loss and Wheezing respectively.

Figure 2 below shows a snapshot of the dataset utilized in training, testing and predicting Tuberculosis and Respiratory

Tract Infections.

Sore Throat	Stable Angina	Swine Flu	Tonsillitis	Tuberculosis	Unstable Angina	Upper Respiratory Infections	Variant Angina Pectoris	Weakness
-1.27	0.589	-2.95	0.161	-0.136	-0.179	-0.126	-0.278	-1.2
0.964	-1.14	-2.28	-0.336	2.49	-0.295	2.18	-0.527	0.208
0.93	0.925	-1.88	-2.85	0.288	-0.549	-0.249	1.18	-0.555
0.583	-0.482	-1.55	0.346	0.468	0.556	-1.37	-2.61	-0.362
0.00568	-0.455	-1.41	0.831	1.02	-0.777	-0.668	-2.57	0.798
0.202	-0.254	-1.35	0.477	-0.238	-0.893	-0.45	0.986	0.546
2.62	1.3	-0.924	0.526	0.284	0.0312	0.621	-0.831	-0.622
-1.6	0.519	-0.916	-0.856	0.575	0.905	-2.51	-0.756	0.589
0.492	-1.2	-0.905	2.28	1.96	-0.365	-0.179	-0.172	0.105
0.543	0.558	-0.893	-0.499	-0.967	2.92	1.81	1.21	-0.479
-1.94	0.783	-0.807	-0.187	-0.117	0.87	0.117	-0.192	-1.08
-0.602	0.588	-0.723	-1.55	0.293	-0.13	-0.894	-0.309	-1.55
-1.11	-0.667	-0.712	-0.0559	0.426	0.057	0.39	0.381	0.299
0.358	-1.51	-0.678	-1.72	-1.44	-1.11	-0.235	-1.05	0.738
-0.195	1.13	-0.635	0.139	-0.688	0.0762	-0.332	-0.122	-1.08
0.46	-0.571	-0.553	0.742	0.324	-0.561	-1.17	0.169	-0.797
-1.36	-1.03	-0.475	-0.0756	-1.87	-0.204	1.01	-1.91	0.855
-0.682	-1.17	-0.457	-0.0784	0.642	2.55	0.438	0.818	-0.201
0.398	-0.434	-0.42	-0.0176	-1.51	-0.0387	0.553	0.0869	-1.49
0.827	-0.0179	-0.351	-0.593	-0.535	0.95	-1.04	-0.158	2.06
1.55	-0.452	-0.292	-0.568	-0.746	-1.2	-1.47	-0.0996	-0.602
-0.218	-1.25	-0.264	1.9	1.51	0.853	1.25	-0.636	-0.961
-0.453	-0.231	-0.246	1.04	1.15	-0.897	-0.207	1.37	-0.0545
0.0107	-2.4	-0.213	-0.329	-0.495	-0.738	-0.135	0.46	-0.203

Figure 2. Snapshot of Dataset

The Bayesian Belief Network model was designed using Bayes-Server platform. The Bayesian Belief Network (BBN) for anticipating Tuberculosis (TB) was organized with the ultimate objective that the nodes on the network are connected reliant on the probability of an illness coming to fruition to another and factor influencing another factor. In our model for a case to be regarded as a Tuberculosis infection case, the infections and various factors taken into mindfulness in the diagnosis of Tuberculosis are: Active Tuberculosis, Adult-Onset Asthma, Allergic Asthma, Angina, Asthma, Asthma-COPD Overlap, Breathlessness, Bronchitis, Chills, Chronic Obstructive Pulmonary Disease (COPD), Cough, Emphysema, Fatigue, Fever, Hemoptysis, Influenza, Laryngitis, Latent Tuberculosis, Loss of Appetite, Lower Respiratory Infections, Migraine, Muscle Pain, Mycobaterium Tuberculosis Bacteria, Nausea, Night Sweats, Non-Allergic Asthma, Occupational Asthma, Pertussis, Phlegm, Pneumonia, Respiratory Tract Infections (RTIs), Runny Nose, Severe Angina, Sinus Infection, Sneezing, Sore Throat, Stable Angina, Swine Flu, Tonsillitis, Unstable Angina, Upper Respiratory Infections, Variant Angina Pectoris, Weakness, Weight Loss and Wheezing respectively.

Figure 3 shows the BBN model for detecting Tuberculosis (TB), Respiratory Tract Infections (RTIs) and their symptoms.

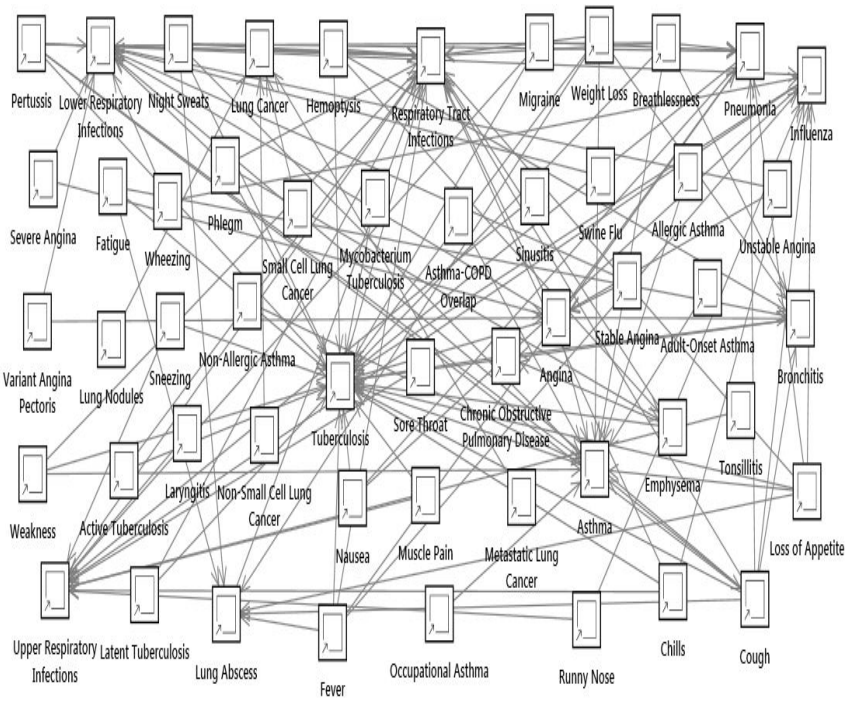


Figure 3. Bayesian Belief Network Model for Detecting Tuberculosis (TB), Respiratory Tract Infections (RTIs) and Its Symptoms.

So, to mathematically represent our model we have:

Tuberculosis

$$\%_{i=1}^{52} P(Disease_j) \mid Parents_j(Disease_j) \tag{2}$$

Where,

Disease: Node with a Disease Ailment

Parents ($Disease_i$) = Nodes that converge on Disease $Ailment_i$..

The dataset in figure 2 was used to train and test the model. Upon completion of training and testing the BBN model, the test data converged at time series 2. The log likelihood value for each case was recorded.

Figure 4 shows the BBN model convergence of Tuberculosis (TB), Respiratory Tract Infections (RTIs) and their symptoms at Iteration Count 2.

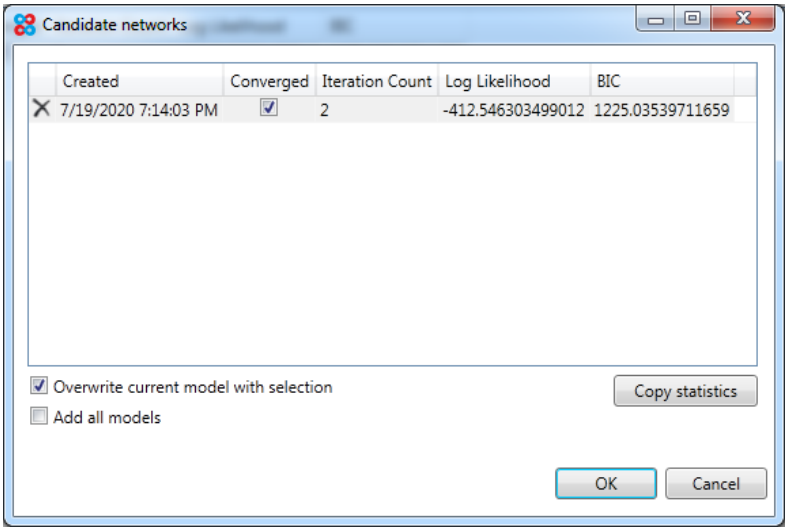
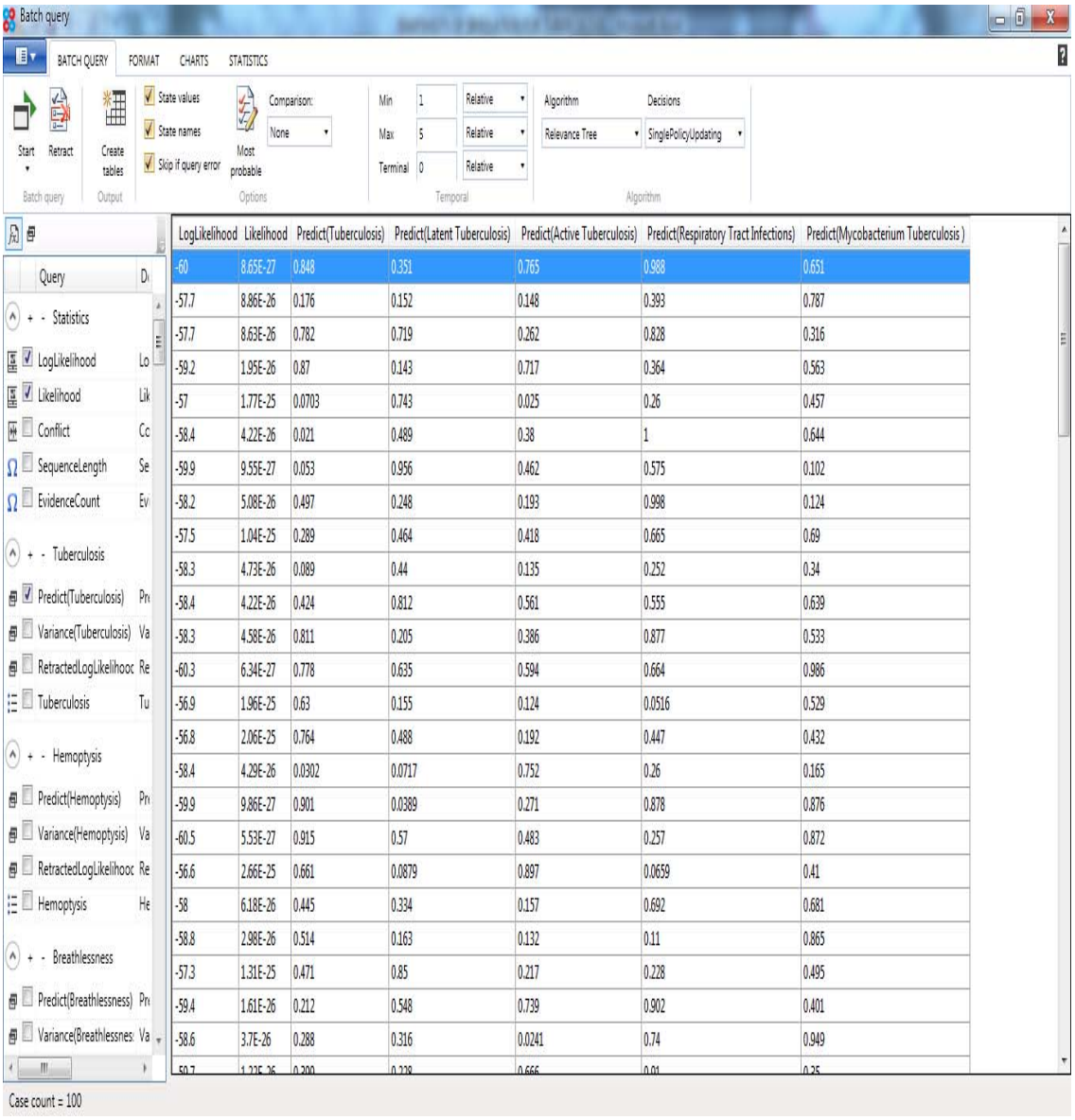


Figure 4. Bayesian Belief Network Model for Detecting Tuberculosis (TB), Respiratory Tract Infections (RTIs) and Their Symptoms Converging at Time Series 2.

Figures 4,5, 6, 7, 8, 9 ,10, 11 and 12 shows log likelihood batch query chart for predicting tuberculosis (TB), respiratory tract infections (RTIs) and its symptoms, feature importance chart of nodes in the BBN model, the in-sample anomaly detection chart, the likelihood plot of lung diseases, upper respiratory tract infections, lower respiratory tract infections and symptoms prompting respiratory tract infections case and its probabilities, the likelihood plot of mycobacterium tuberculosis bacteria, latent tuberculosis, active tuberculosis and its symptoms prompting tuberculosis infection case and its probabilities, the loglikelihood graph for detecting tuberculosis, respiratory tract infections and their symptoms and likelihood against loglikelihood for predicting tuberculosis, respiratory tract infections and their symptoms respectively. The results generated from the simulation demonstrated that the system had the option to anticipate 99% tuberculosis (TB), respiratory tract Infections (RTIs) and their symptoms on the dataset precisely and it had a loglikelihood of 62 on the test dataset.

The figure 5 below shows the loglikelihood batch query chart for predicting tuberculosis (TB) and respiratory tract infections with their symptoms.



LogLikelihood	Likelihood	Predict(Tuberculosis)	Predict(Latent Tuberculosis)	Predict(Active Tuberculosis)	Predict(Respiratory Tract Infections)	Predict(Mycobacterium Tuberculosis)
-60	8.65E-27	0.848	0.351	0.765	0.988	0.651
-57.7	8.86E-26	0.176	0.152	0.148	0.393	0.787
-57.7	8.63E-26	0.782	0.719	0.262	0.828	0.316
-59.2	1.95E-26	0.87	0.143	0.717	0.364	0.563
-57	1.77E-25	0.0703	0.743	0.025	0.26	0.457
-58.4	4.22E-26	0.021	0.489	0.38	1	0.644
-59.9	9.55E-27	0.053	0.956	0.462	0.575	0.102
-58.2	5.08E-26	0.497	0.248	0.193	0.998	0.124
-57.5	1.04E-25	0.289	0.464	0.418	0.665	0.69
-58.3	4.73E-26	0.089	0.44	0.135	0.252	0.34
-58.4	4.22E-26	0.424	0.812	0.561	0.555	0.639
-58.3	4.58E-26	0.811	0.205	0.386	0.877	0.533
-60.3	6.34E-27	0.778	0.635	0.594	0.664	0.986
-56.9	1.96E-25	0.63	0.155	0.124	0.0516	0.529
-56.8	2.06E-25	0.764	0.488	0.192	0.447	0.432
-58.4	4.29E-26	0.0302	0.0717	0.752	0.26	0.165
-59.9	9.86E-27	0.901	0.0389	0.271	0.878	0.876
-60.5	5.53E-27	0.915	0.57	0.483	0.257	0.872
-56.6	2.66E-25	0.661	0.0879	0.897	0.0659	0.41
-58	6.18E-26	0.445	0.334	0.157	0.692	0.681
-58.8	2.98E-26	0.514	0.163	0.132	0.11	0.865
-57.3	1.31E-25	0.471	0.85	0.217	0.228	0.495
-59.4	1.61E-26	0.212	0.548	0.739	0.902	0.401
-58.6	3.7E-26	0.288	0.316	0.0241	0.74	0.949

Figure 5. The Loglikelihood Batch Query Chart for Predicting Tuberculosis (TB), Respiratory Tract Infections (RTIs) with Their Symptoms.

This loglikelihood batch query chart shows the results of test data deployment. Here, 100 exploratory cases were conducted and the result produced from the test data showed the system ability to predict the likelihood of each case being Tuberculosis (TB) infection, a Respiratory Tract Infection and their symptoms with the loglikelihood and likelihood (probability values within 0 to 1) obtained from each of the 100 experimental cases and recorded as follows:

In Experiment 1: The probability of Predict(Tuberculosis) was 0.848, Predict(Latent Tuberculosis) was 0.351, Predict(Active Tuberculosis) was 0.765, Predict(Respiratory Tract Infections) was 0.988 and Predict(Mycobaterium Tuberculosis) was 0.651 in experiment 1 put side by side to 0.848000110,0.350700322,0.764910212, 0.988002030 and 0.650801013 respectively in the test data. What's more, Experiment 1 had 84.8%, 35.1%, 76.5%, 98.8% and 65.1% sensitivity of Tuberculosis, Latent Tuberculosis, Active Tuberculosis, Respiratory Tract Infections Mycobaterium Tuberculosis) and their symptoms after due completion of Experiment 1 correspondingly.

In Experiment 2: The probability of Predict(Tuberculosis) was 0.176, Predict(Latent Tuberculosis) was 0.152, Predict(Active Tuberculosis) was 0.148, Predict(Respiratory Tract Infections) was 0.393 and Predict(Mycobaterium Tuberculosis) was 0.787 in experiment 2 matches up to 0.175701102,0.151600212,0.148010143, 0.392601174 and 0.787002201 respectively in the test data. In addition, Experiment 2 had 17.6%, 15.2%, 14.8%, 39.3% and 78.7% sensitivity of Tuberculosis, Latent Tuberculosis, Active Tuberculosis, Respiratory Tract Infections Mycobaterium Tuberculosis) and their symptoms after termination of Experiment 2 respectively.

In Experiment 3: The probability of Predict(Tuberculosis) was 0.782, Predict(Latent Tuberculosis) was 0.719, Predict(Active Tuberculosis) was 0.262, Predict(Respiratory Tract Infections) was 0.828 and Predict(Mycobaterium Tuberculosis) was 0.316 in experiment 3 judged against 0.781802310,0.719000403,0.261602045, 0.828001046 and 0.316011005 respectively in the test data. Furthermore, Experiment 3 had 78.2%, 71.9%, 26.2%, 82.8% and 31.6% sensitivity of Tuberculosis, Latent Tuberculosis, Active Tuberculosis, Respiratory Tract Infections and Mycobaterium Tuberculosis and their symptoms after conclusion of Experiment 3 in that order.

Besides, this experiment continued up to Experiment number 100. Hence, the system results showed a 0.0001 value difference between the prediction results and original test data of 100% resulting to 99% prediction accuracy.

The figure 6 beneath shows the feature importance of the linked nodes in the designed BBN model in figure 3 above for predicting Respiratory Tract Infections (RTIs) with its symptoms.

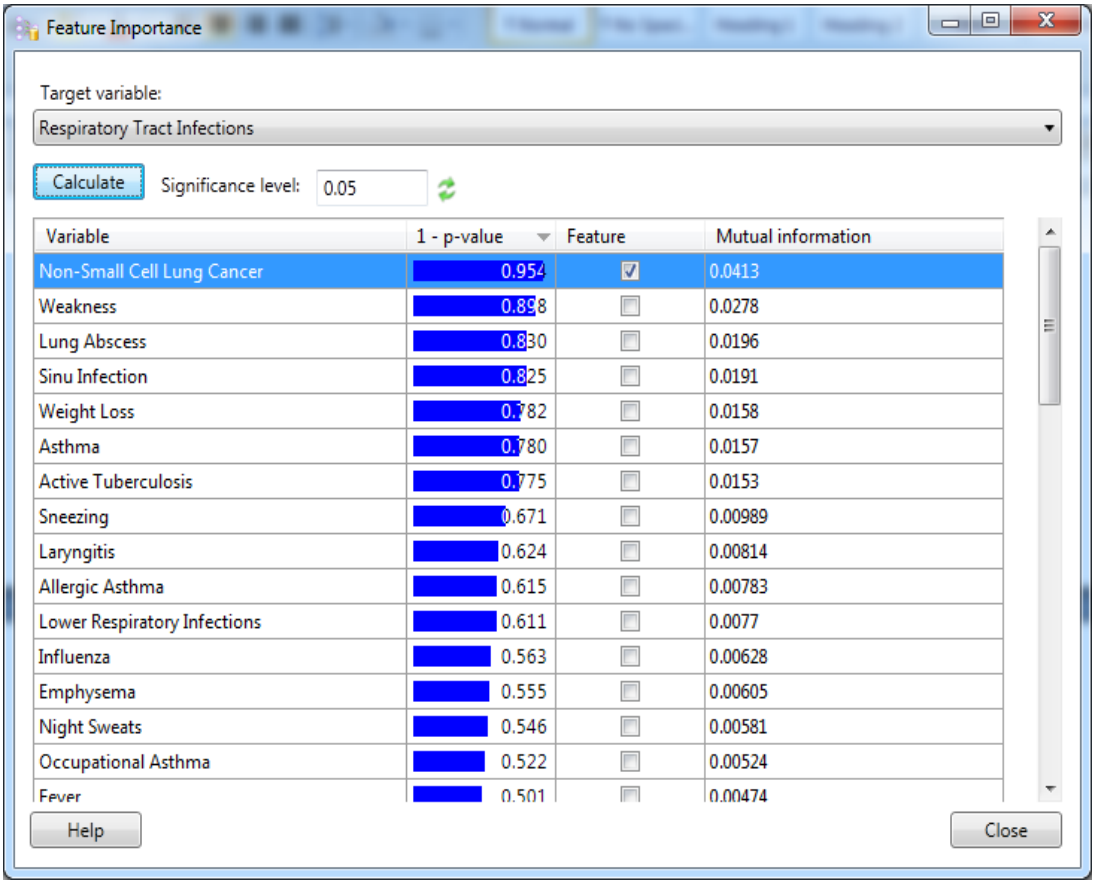


Figure 6. The Feature Importance Chart for Respiratory Tract Infections Node in The Designed BBN Model

The Feature Importance Chart shows probabilistic-value (p-value) of the variable (nodes), Feature and Mutual information in relation to the Respiratory Tract Infections node.

The p-value indicates the likelihood (probability) of the nodes inciting Respiratory Tract Infections.

The Feature box is checked if that precise node is actively involved and assists in the detection of Respiratory Tract Infection case.

The Mutual information shows the bond with nodes directly linked to one another (i.e. in this case the direct association of the nodes with the Respiratory Tract Infections node) and assigned a value.

The Significance Level indicates the margin of error encountered in the detection of Respiratory Tract Infections.

The figure 7 below shows the feature importance of the associated nodes in the designed BBN model for predicting tuberculosis and its symptoms in figure 3.

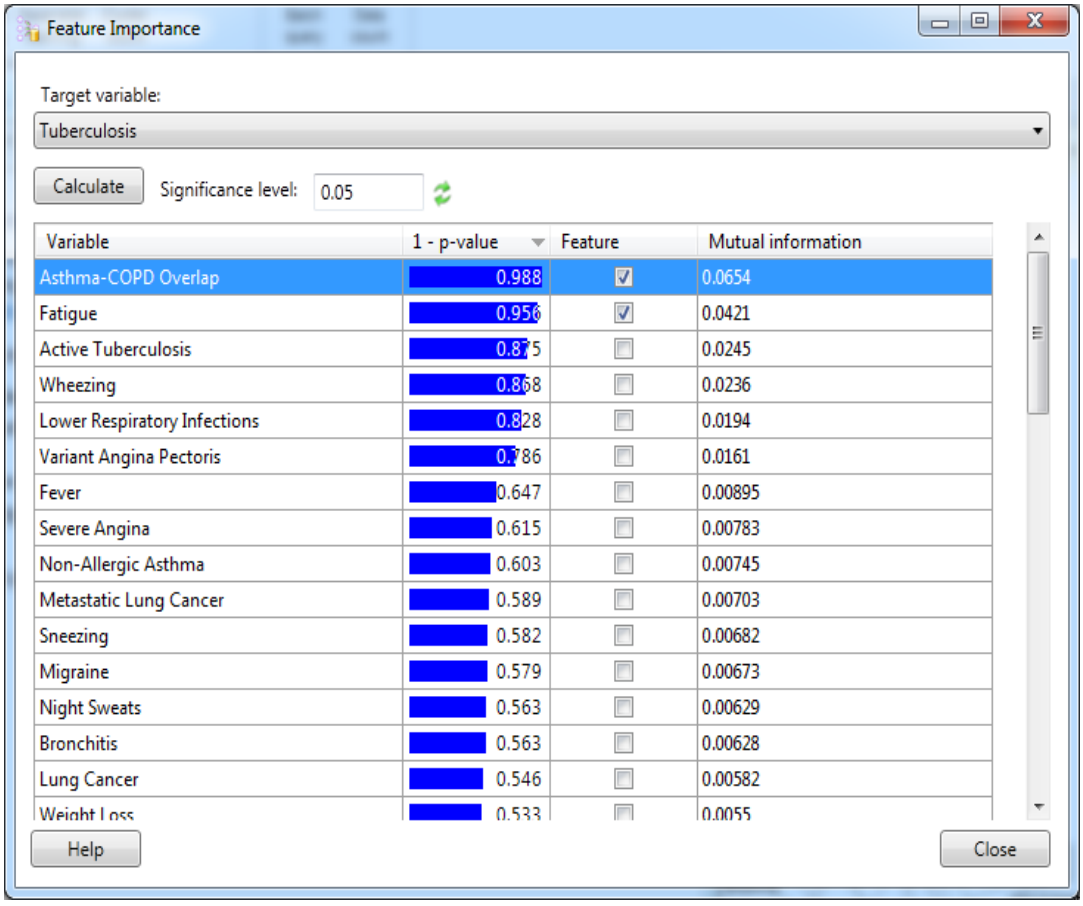


Figure 7. The Feature Importance Chart for Tuberculosis Node in The Developed BBN Model

The Feature Importance Chart shows probabilistic-value (p-value) of the variable (nodes), Feature and Mutual information in relation to the Tuberculosis node.

The p-value specifies the likelihood (probability) of the nodes prompting Tuberculosis infection.

The Feature box is checked if that precise node is actively involved and aids the recognition of Tuberculosis infection case.

The Mutual information shows the connection with nodes directly linked to one another (i.e. in this case the direct association of the nodes with the Tuberculosis node) and allotted a value.

The Significance Level point to the margin of error encountered during the course of detecting of Tuberculosis.

The figure 8 below shows the in-sample anomaly detection chart of the developed BBN model in figure 3.

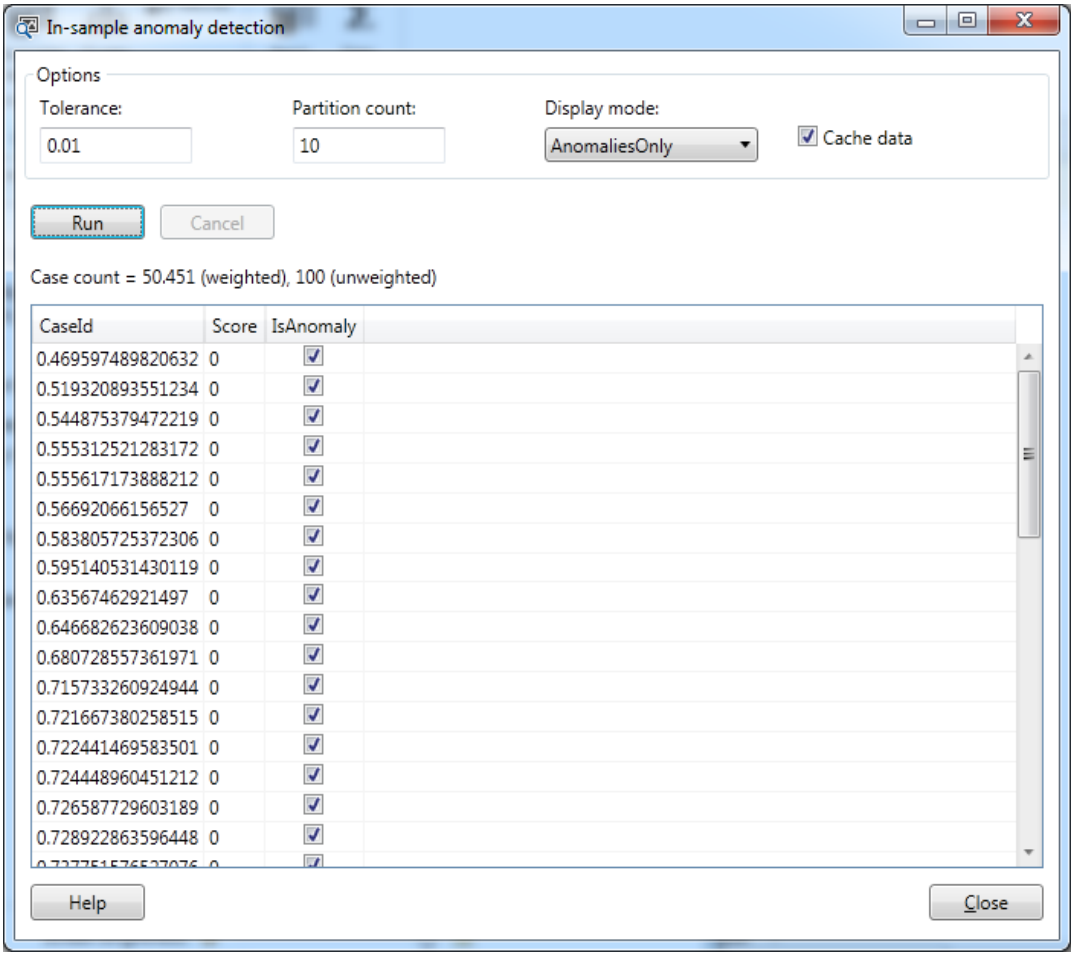


Figure 8. The In-sample Anomaly Detection Chart of BBN Model

The In-sample Anomaly Detection Chart shows 100 experimental tests of distinguishing tuberculosis and respiratory tract infections. Each case is doled out an ID(Identification value) which are estimations of the

Predict(Tuberculosis), Predict(Latent Tuberculosis), Predict(Active Tuberculosis), Predict(Respiratory Tract Infections) and Predict(Mycobaterium Tuberculosis Bacteria) in figure 5.

The IsAnomaly checkbox is checked to identify that each case is a confirmed instance of Tuberculosis, Respiratory Tract Infections with their side effects. The 100 cases includes Tuberculosis, Respiratory Tract Infections with their side effects having a case tally estimation 50.451 (weighted) which implies the impact of the cases prompting Tuberculosis and Respiratory Tract Infections. On the other hand, the 100 (unweighted) implies the quantity of cases in the pool of information accessible to the system for recognition of Tuberculosis, Respiratory Tract Infections and their manifestations in the dataset pool. The tolerance is the margin of error that could be encountered as regards to the detection of the Tuberculosis, Respiratory Tract Infections and their symptoms.

The figure 9 below shows the likelihood plot of lung diseases, upper respiratory tract infections, lower respiratory tract infections and symptoms prompting respiratory tract infections of the designed BBN model in figure 3.

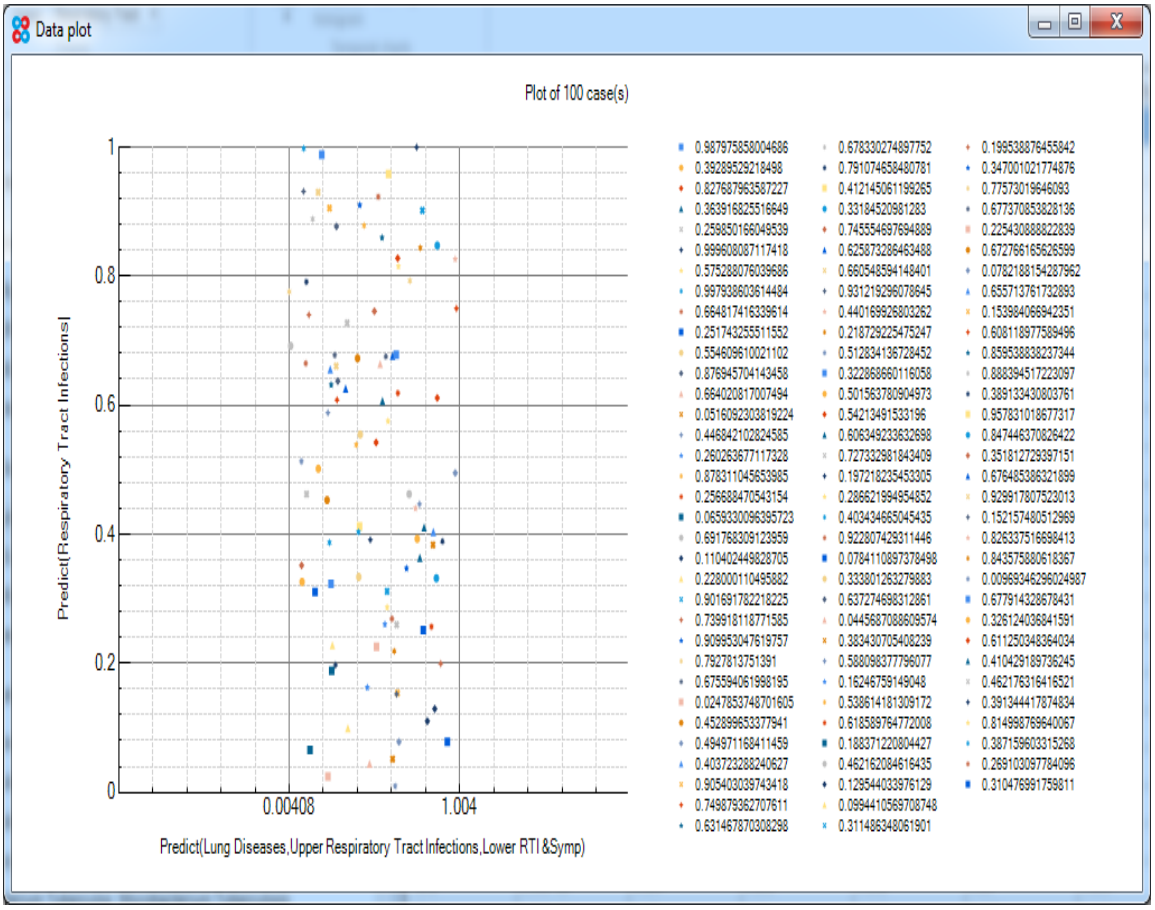


Figure 9. The Likelihood Plot of Lung Diseases, Upper Respiratory Tract Infections, Lower Respiratory Tract Infections and Symptoms Prompting Respiratory Tract Infections

The likelihood plot shows the probability of lung diseases, upper respiratory tract infections, lower respiratory tract infections and symptoms prompting respiratory tract infections case. In this plot, 100 experiments were undertaken with each shaded point in the chart named a case and allotted a probabilistic value which is in the scope of 0 to1 on Y-axis and 0.00408 to 1.004 on X-axis and positioned on the privilege of the plot. The variable marked "Predict(Respiratory Tract Infections)" is situated on the Y-axis is plotted against another variable named "Predict(Lung Diseases, Upper Respiratory Tract Infections, Lower Respiratory Tract Infections and Symptoms)" stationed on the X-axis. Nonetheless, from this plot, there are five analytical classes of Respiratory Tract Infections cases which our system had the choice to distinguish; they are asymptomatic, mild, moderate, severe, and critical classes in the aforesaid order.

Asymptomatic Class: This class likelihood ranges from 0 to 0.2 on Y-axis and 0.00408 to 1.004 on X-axis. This territory has 16 colored points (cases). This infers that the 16 colored points in this region represents 16 cases of no Respiratory Tract Infections by any stretch of the imagination; hence this class of patients is viewed as being Asymptomatic.

Mild Class: This class likelihood ranges from 0.2 to 0.4 on Y-axis and 0.00408 to 1.004 on X-axis. This region has 25 tinted points (cases). This suggests that the 25 colored points here corresponds to 25 occurrences of symptomatic patients with Respiratory Tract Infections with the seriousness level identified as being Mild.

Moderate Class: This class ranges from 0.4 to 0.6 on Y-axis and 0.00408 to 1.004 on X-axis. This zone has 16 tinted points (cases).This infers that the 16 toned points in this area signifies 16 examples of symptomatic patients with Respiratory Tract Infections with the seriousness level classified as being Moderate.

Severe Case: This class ranges from 0.6 to 0.8 on Y-axis and 0.00408 to 1.004 on X-axis. This locale has 24 toned points (cases). This implies that the 24 hued points in this area symbolizes 24 occurrences of symptomatic patients with Respiratory Tract Infections with the seriousness level recognized as being Severe.

Critical Class: This level spans from 0.8 to 1 on Y-axis and 0.00408 to 1.004 on X-axis. This region has 19 toned points (cases). This infers that the 19 toned points around this zone connotes 19 occurrences of symptomatic patients with Respiratory Tract Infections with the severity level identified as being Critical.

All 100 cases in figure 9 had likelihood values less than or equal to 1; with the most essential probability estimation of Lung Diseases, Upper Respiratory Tract Infections, Lower Respiratory Tract Infections and symptoms prompting Respiratory Tract Infections reported to be 0.999608087117418 which is very well below 1.

Of the 100 experimental cases, the system anticipated 100 cases of Lung Diseases, Upper Respiratory Tract Infections, Lower Respiratory Tract Infections and symptoms prompting Respiratory Tract Infections case extending from asymptomatic, mild, moderate, severe, and critical classes precisely from the test data with 99.96% sensitivity of Respiratory Tract Infections.

The figure 10 below shows the likelihood plot of mycobaterium tuberculosis, latent tuberculosis, active tuberculosis and its symptoms prompting tuberculosis infection case.

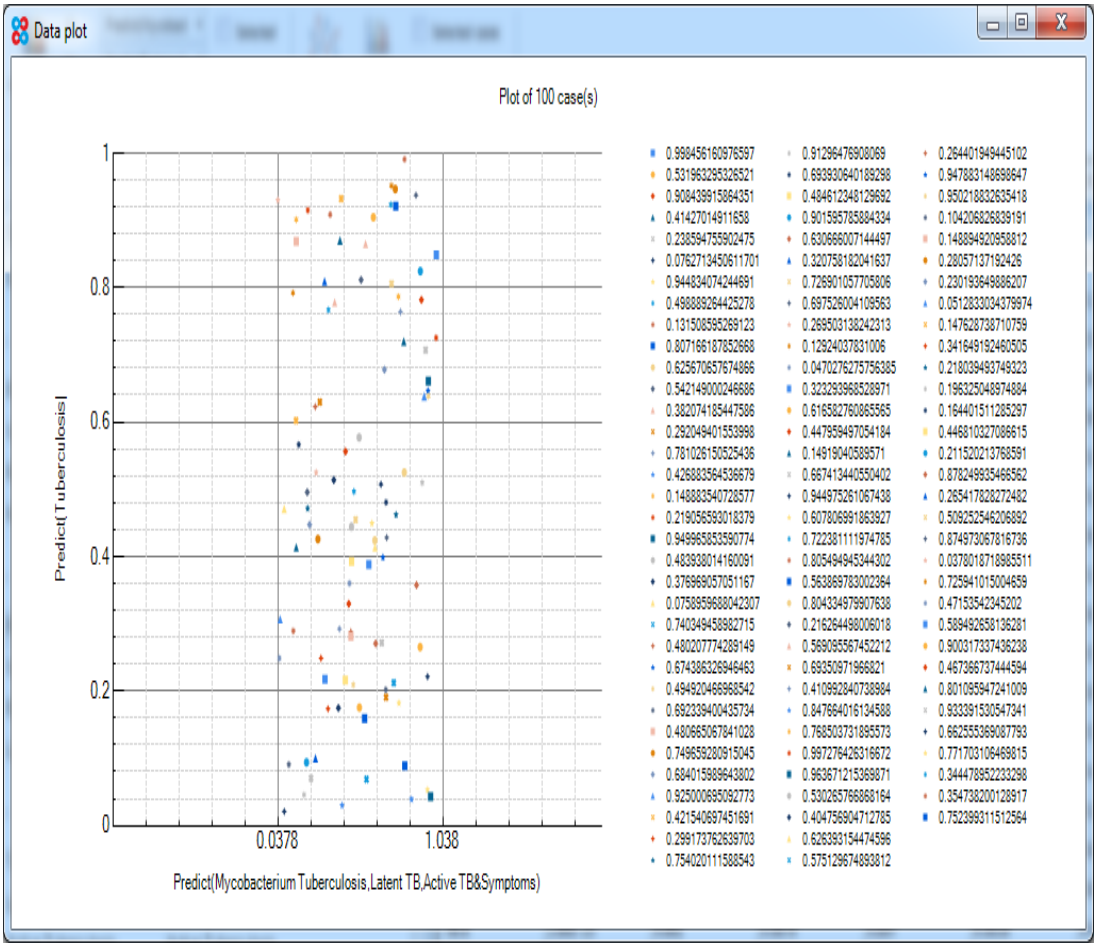


Figure 10. The Likelihood Plot of Mycobaterium Tuberculosis, Latent Tuberculosis, Active Tuberculosis and Its Symptoms Prompting Tuberculosis Infection Case

The likelihood plot shows the probability of Mycobaterium Tuberculosis, Latent Tuberculosis, Active Tuberculosis and its symptoms prompting Tuberculosis infection case. In this plot, 100 experiments were undertaken with each shaded point in the chart named a case and allotted a probabilistic value which is in the scope of 0 to1 on Y-axis and 0.0378 to 1.038 on X-axis and positioned on the privilege of the plot. The variable marked "Predict(Tuberculosis)" is situated on the Y-axis is plotted against another variable named "Predict(Mycobaterium Tuberculosis, Latent Tuberculosis, Active Tuberculosis and Symptoms)" stationed on the X-axis. Nonetheless, from this plot, there are five diagnostic classes of Tuberculosis infections cases which our system had the choice to distinguish; they are asymptomatic, mild, moderate, severe, and critical classes in the aforesaid order.

Asymptomatic Class: This class likelihood ranges from 0 to 0.2 on Y-axis and 0.0378 to 1.038 on X-axis. This territory has 18 colored points (cases). This infers that the 18 colored points in this region represents 18 cases of no Tuberculosis infection whatsoever; hence this class of patients is viewed as being Asymptomatic.

Mild Class: This class likelihood ranges from 0.2 to 0.4 on Y-axis and 0.0378 to 1.038 on X-axis. This region has 21 tinted points (cases). This suggests that the 21 colored points here corresponds to 21 occurrences of symptomatic patients having Tuberculosis infection; with the seriousness level identified as being Mild.

Moderate Class: This class ranges from 0.4 to 0.6 on Y-axis and 0.0378 to 1.038 on X-axis. This zone has 24 tinted points (cases).This infers that the 24 toned points in this area signifies 24 examples of symptomatic patients having Tuberculosis Infection; with the seriousness level classified as being Moderate.

Severe Case: This class ranges from 0.6 to 0.8 on Y-axis and 0.0378 to 1.038 on X-axis. This locale has 17 toned points (cases). This implies that the 17 hued points in this area symbolizes 17 occurrences of symptomatic patients having Tuberculosis Infection; with the seriousness level recognized as being Severe.

Critical Class: This level spans from 0.8 to 1 on Y-axis and 0.0378 to 1.038 on X-axis. This region has 20 toned points (cases). This infers that the 20 toned points around this zone connotes 20 occurrences of symptomatic patients having Tuberculosis Infection; with the severity level identified as being Critical.

All 100 cases in figure 10 had likelihood values less than or equal to 1; with the most essential probability estimation of Mycobaterium Tuberculosis, Latent Tuberculosis, Active Tuberculosis and after-effects of diseases prompting Tuberculosis Infection reported to be 0.998456160976597 which is very well below 1.

Of the 100 experimental cases, the system anticipated 100 cases of Mycobaterium Tuberculosis, Latent Tuberculosis, Active Tuberculosis and Symptoms Prompting Tuberculosis Infection case extending from asymptomatic, mild, moderate, severe, and critical classes precisely from the test data with 99.84% sensitivity of Tuberculosis Infection.

Having attained the severity levels of Tuberculosis and Respiratory Tract Infections ranging from asymptomatic, mild, moderate, severe, and critical, we intend to plot the chart for the loglikelihood graph for detecting Tuberculosis and Respiratory Tract Infections; likelihood against loglikelihood graph for foreseeing Tuberculosis and Respiratory Tract Infections and what's more discover the loglikelihood value for distinguishing Tuberculosis and Respiratory Tract Infections; and prediction exactness of the BBN model which will be discussed in figure 11 and 12 beneath.

Figure 11 shows the loglikelihood chart for detecting Tuberculosis and Respiratory Tract Infections.

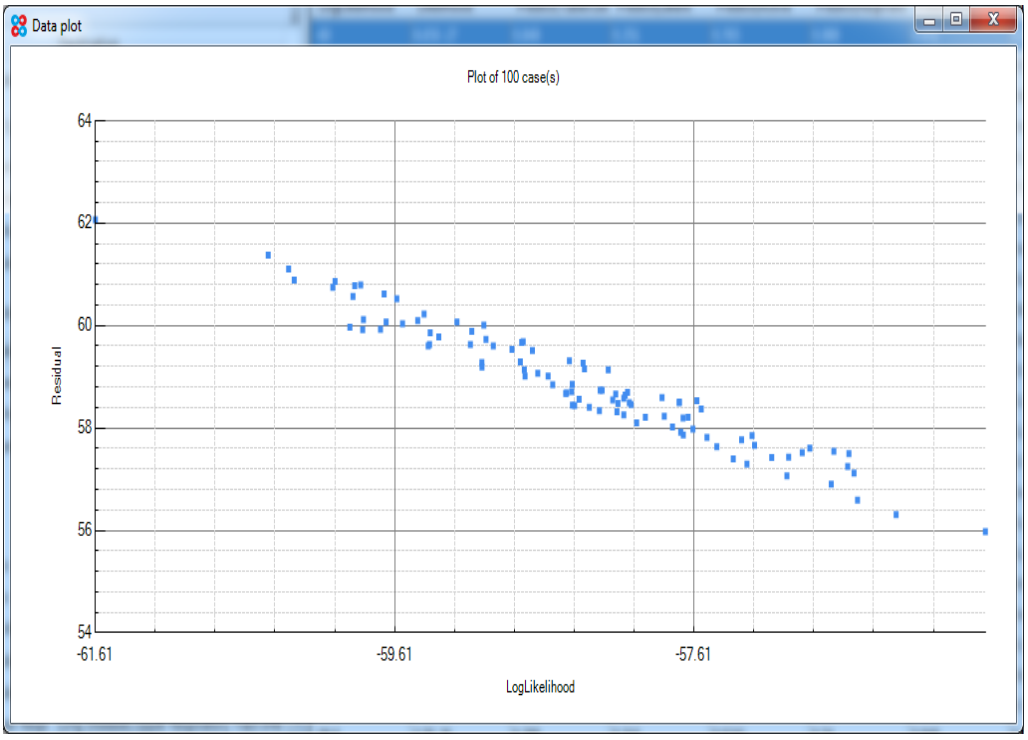


Figure 11. Loglikelihood Chart for Detecting Tuberculosis and Respiratory Tract Infections.

This loglikelihood graph for detecting Tuberculosis and Respiratory Tract Infections shows residual values on the Y-axis plotted against the loglikelihood values on the X-axis which are independent variables. Be that as it may, a residual value is a degree of how much a regression line perpendicularly fails a data spot. In essence, regression lines are the principal assault of tons of data. The lines are ordered as midpoints; conversely, a few data spots will match the line and others will fall short of hitting the spot. In this graph, it shows that 100 experimental cases achieved estimation of 62,61.35,61.28,61.24,61.22,61.20,60.19,60.16,60.13,60.11,60.09,60…..56.24 and 56 respectively.

On the other hand, residual values should be equally and intuitively isolated around the level lines. Taking a standpoint on the system' experimental results regards acquired from the even lines on the outline, it might be seen that where the uppermost residual value and the loglikelihood independent factor accomplished congregates at -61.61 on horizontal line with 64 being the most vital value that can be reached on the vertical line. The residual value achieved is 62 and loglikelihood independent esteem is -61.61, the differentiation between the two characteristics is 0.39 which is the complexity between the estimations of the prediction outcomes and unique test data of 100% in figure 5.

Consequently, the attained Loglikelihood value for recognizing Tuberculosis, Respiratory Tract Infections and their manifestations is 62.

Figure 12 shows the likelihood against loglikelihood for predicting Tuberculosis, Respiratory Tract Infections with its symptoms of the designed BBN model in figure 3.

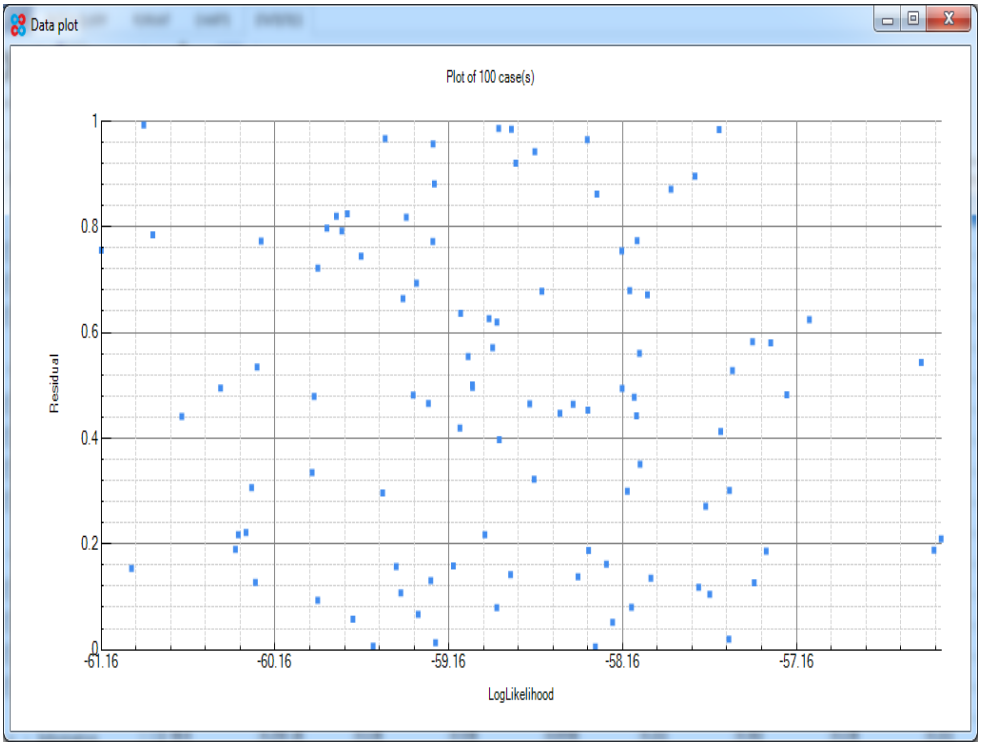


Figure 12. The Likelihood against the Loglikelihood Graph for Predicting Tuberculosis, Respiratory Tract Infections with Their Manifestations

This likelihood against loglikelihood plot for Tuberculosis, Respiratory Tract Infections and their manifestations shows the residual (likelihood) on the Y-axis plotted against the Loglikelihood on the X-axis both of which are independent variables. However, the likelihood of events (Predicting Tuberculosis and Respiratory Tract Infections) occurring are probabilistic values placed between 0 and 1. In this plot, 100 exploratory cases were conducted which achieved the estimations of 0.9998, 0.9997, 0.9996, 0.9994, 0.9992, 0.9990, 0.8998, 0.8995,…..0.0001,0 individually. Moreover, residual (likelihood) values ought to be reliably and randomly stretch around the level lines. Essentially, viewing the system' exploratory results obtained from the level lines on the diagram, it will in general be seen that the residual probability value attained is 0.9998 and loglikelihood independent value is –61.16.

Consequently, in this system, the most essential likelihood esteem that can be attained is 1. With 1, being the 100 % residual (probability) rate mark, to get our exactness precision rate, we divide attained likelihood probability esteem by highest likelihood value that can be attained and increase by most imperative residual mark, that is 0.9998/1*100% = 99.98% forecast exactness rate on the test data.

Besides, the likelihood graphs in figure 9, 10 showcased all classes of severity status of lung diseases, upper respiratory tract infections, lower respiratory tract Infections and symptoms prompting respiratory tract infections, mycobaterium tuberculosis bacteria, latent tuberculosis, active tuberculosis and its symptoms prompting tuberculosis infection ranging from asymptomatic, mild, moderate, severe, and critical classes independently with their probabilities while figure 11 demonstrated the system loglikelihood estimation of 62 for distinguishing tuberculosis, respiratory tract infections and their manifestations while the likelihood against loglikelihood anticipation plot of tuberculosis, respiratory tract infections and their manifestations in figure 12 ascertained a 99.98% forecast precision of the system.

Nevertheless, the likelihood given evidence of Lung Diseases, Upper Respiratory Tract Infections, Lower Respiratory Tract Infections and Symptoms Prompting Respiratory Tract Infections is denoted as:

P(Respiratory Tract Infections| Active Tuberculosis, Adult-Onset Asthma, Allergic Asthma, Angina, Asthma, Asthma-COPD Overlap, Breathlessness, Bronchitis, Chills, Chronic Obstructive Pulmonary Disease (COPD), Cough, Emphysema, Fatigue, Fever, Hemoptysis, Influenza, Laryngitis, Latent Tuberculosis, Loss of Appetite, Lower Respiratory Infections, Lung Abscess, Lung Cancer, Lung Nodules, Metastatic Lung Cancer, Migraine, Muscle Pain, Mycobaterium Tuberculosis Bacteria, Nausea, Night Sweats, Non-Allergic Asthma, Non-Small Cell Lung Cancer, Occupational Asthma, Pertussis, Phlegm, Pneumonia, Respiratory Tract Infections (RTIs), Runny Nose, Severe Angina, Sinus Infection, Small Cell Lung Cancer, Sneezing, Sore Throat, Stable Angina, Swine Flu, Tonsillitis, Tuberculosis, Unstable Angina, Upper Respiratory Infections, Variant Angina Pectoris, Weakness, Weight Loss and Wheezing) = 0.99960807117418.

The likelihood given evidence of Mycobaterium Tuberculosis Bacteria, Latent Tuberculosis, Active Tuberculosis and Its Symptoms Prompting Tuberculosis Infection is denoted as:

P(Tuberculosis| Active Tuberculosis, Adult-Onset Asthma, Allergic Asthma, Angina, Asthma, Asthma-COPD Overlap, Breathlessness, Bronchitis, Chills, Chronic Obstructive Pulmonary Disease (COPD), Cough, Emphysema, Fatigue, Fever, Hemoptysis, Influenza, Laryngitis, Latent Tuberculosis, Loss of Appetite, Lower Respiratory Infections, Migraine, Muscle Pain, Mycobaterium Tuberculosis Bacteria, Nausea, Night Sweats, Non-Allergic Asthma, Occupational Asthma, Pertussis, Phlegm, Pneumonia, Respiratory Tract Infections (RTIs), Runny Nose, Severe Angina, Sinus Infection, Sneezing, Sore Throat, Stable Angina, Swine Flu, Tonsillitis, Unstable Angina, Upper Respiratory Infections, Variant Angina Pectoris, Weakness, Weight Loss and Wheezing) = 0.998456160976597.

From the test, it will in general be viewed that our model has a higher residual loglikelihood value which is 62; on the whole, a prediction precision of 99.98%; 99.84% and 99.96% sensitivity of Tuberculosis and Respiratory Tract Infections in that order.

In the end, comparing the 99.98% prediction exactness of our model with the works conducted by [7, 8,9,10,11,12,14,15, 16,17, 18, 19 and 20] which has less than 99.58 % prediction exactness individually; it is clear our model has a superior prediction precision than the aforementioned systems. The higher forecast exactness achieved by our model could be a direct result of the range of the dataset used in training and testing the model similarly as its ability to foresee the covering symptoms of tuberculosis and respiratory tract infections, which assisted in the high recognition exactness of the aforementioned diseases.

V. CONCLUSION AND FUTURE SCOPE

Tuberculosis (TB) is a life threatening communicable disease caused by Mycobacterium Tuberculosis (MTB) which usually influences the lungs, hence interrupting its function of regulating the inhaling of oxygen and exhaling of carbon dioxide. Nonetheless, the Mycobacterium Tuberculosis (MTB) is transmitted from one individual to another via aerosols (miniature respiratory droplets) expelled from the human body into the air through coughs and sneezes from symptomatic patients. The TB Infection has potentials of affecting other organs of the human body such as liver, kidney just to a name a few which are at risk due to its spread.

Owing to the widespread of the TB infection, In time past, several systems have been utilized in diagnosing Tuberculosis and Respiratory Tract Infections with the solitary aim of curbing the untimely deaths of patients due to late diagnosis and misdiagnosis of the aforesaid diseases an area health and Information Technology experts are making remarkable attempts to enhance.

In this paper, we used an AI strategy called Bayesian Belief Network to predict Tuberculosis, Respiratory Tract Infections and their manifestations. The model had 52 nodes with each node addressing a select ailment and factors that influence diagnosis of Tuberculosis, Respiratory Tract Infections and their side effects. The model was trained and tested and had a general prediction accuracy of 99.98% and a loglikelihood estimation of 62 in predicting Tuberculosis, Respiratory Tract Infections with their side effects; 99.84% and 99.96% sensitivity of Tuberculosis and Respiratory Tract Infections in a that order.

We can surmise that the acquired results exhibited the learning capacity of the BBN model, despite the fact that the data utilized can be improved significantly. In this paper, we concentrated on the symptoms of Tuberculosis and Respiratory Tract Infections which supported the model' inference mechanism for early diagnosis of the aforementioned maladies.

For future works, there is need to incorporate more data inclined by the infection in other to improve the prescient and derive optimal outcomes which will be utilized and lead to improvement in the accompanying areas: forecast and classification of Tuberculosis and Respiratory Tract Infections.

REFERENCES

[1] Nilsson, Goran E. (2010). "Respiratory Physiology of Vertebrates",Cambridge University Press, pp 1-5, 2010. ISBN 978-0-521-70302-4.

[2] N.A. Campbell (1990): "Biology",2nd edition. Redwood City, Calif.: Benjamin/Cummings Publishing. Company. pp. 834–835, 1990. ISBN 0-8053-1800-3.

[3] World Health Organization, "Global Tuberculosis Report 2017", WHO/HTM/TB/2017.23, pp. 1-8, 2017.

[4] World Health Organization, "Tuberculosis in Women Report 2018" WHO Global TB Programme, WHO/CDS/TB/2018.36 REv.1, pp. 1-2, 2019.

[5] World Health Organization, "Global Tuberculosis Report 2018".WHO/CDS/TB/2018.20, pp 7, 2018. ISBN 978-92-4-156564-6

[6] World health Organization, "Global Tuberculosis Report 2019". WHO/CDS/TB/2019.15, pp 1-3, 2019. ISBN 978-92-4-156571-4

[7] N.H. Phuong , H.H. Dang, N.R. Prasad, "Developing Supporting Expert System for Lung-Disease Diagnosis Using Fuzzy Logic". Biomedical SoftComputing and Human Sciences, Vo1.5, Issue 2, pp.37-44, 2000.

[8] S.R.N. Kalhori, X-J Zeng, "Fuzzy Logic Approach to Predict the Outcome of Tuberculosis Treatment Course Destination", Proceedings of the World Congress on Engineering and Computer Science 2009 Vol.II, WCECS 2009, San Francisco, USA, pp.1-5, October 20-22, 2009. ISBN: 978-988-18210-2-7.

[9] A.K. Ali, X.D Zhi, S.K. Ali, "Novel Respiratory Diseases Diagnosis by Using Fuzzy Logic". Global Journal of Computer Science and Technology Vol. 10 Issue 13 (Ver. 1.0), pp.61-65, October 2010.

[10] N.Walia, S.K.Tiwari, R.Malhotra, "Design and Identification of Tuberculosis using Fuzzy Based Decision Support System". Advances in Computer Science and Information Technology, Vol.2, Issue 8, pp.57-62, 2015. ISSN: 2393-9915.

[11] X.Y. Djam, Y.H. Kimbi, "A Decision Support System for Tuberculosis Diagnosis". The Pacific Journal of Science and Technology, Vol.12, Issue 2, pp.410-425, November 2011 (Fall).

[12] A.A.Imianvan, J.C.Obi, "Fuzzy Cluster Means Expert System for the Diagnosis of Tuberculosis". Global Journal of Computer Science & Technology, Vol.11, Issue 1.0, April 2011.

[13] K.Soundararajan, S.Sureshkumar, C.Anusuya, "Diagnostics Decision Support System for Tuberculosis using Fuzzy Logic". IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 2, Issue 3, pp.684-689, June 2012. ISSN: 2249-9555.

[14] S.K.Tiwari, N.Walia, H.Singh, A.Sharma, "Effective Analysis of Lung Infection using Fuzzy Rules". International Journal of Bio-Science and Bio-Technology, Vol.7, Issue 6, pp.85-96, 2015.

[15] N.Walia, S.K.Tiwari, R.Malhotra, "Design and Identification of Tuberculosis using Fuzzy Based Decision Support System". Advances in Computer Science and Information Technology (ACSIT), Vol.2, Issue 8, pp. 57-62, April-June, 2015. Print ISSN: 2393-9907; Online ISSN: 2393-9915.

[16] N.Walia, H.Singh, S.K. Tiwari, A.Sharma, "A Decision Support System for Tuberculosis Diagnosability". International Journal on Soft Computing (IJSC), Vol.6, Issue 3, pp.1-14, August 2015.

[17] A.K.Shrivastava, A.Rajak, N.Singhal, "Modeling Pulmonary Tuberculosis using Adaptive Neuro Fuzzy Inference System". International Journal of Innovative Research in Computer Science & Technology (IJIRCST), Vol.4, Issue 1, January 2016. ISSN: 2347-5552.

[18] A.Angbera, M.B. Esiefarienrhe, I.Agaji, "Efficient Fuzzy-Based System for the Diagnosis and Treatment of Tuberculosis (EFBSDTTB)".International Journal of Computer Applications Technology and Research, Vol.5, Issue 2,pp.34-48, 2016. ISSN: 2319–8656.

[19] I.Goni, C.U. Ngene, I.Manga, A.Nata'ala, J.C.Sunday, "Intelligent System for Diagnosing Tuberculosis using Adaptive Neuro-Fuzzy". Asian Journal of Research in Computer Science, Vol.2, Issue 1, pp.1-9, 2018; Article no.AJRCOS.44210.

[20] J.M.Gumpy, I.Goni,M.Isa, "Neuro-Fuzzy Approach for Diagnosisng And Control of Tuberculosis". The International Journal of Computational Science, Information Technology and Control Engineering (IJCSITCE) Vol.5, Issue 1, pp.1-10, January 2018.

[21] A. Simon, M.S. Deo, S. Venkatesan, D.R.R. Babu, "Machine Learning and its Applications". International Journal of Electrical Sciences & Engineering (IJESE), Vol. 1, Issue 1, pp. 22-24, 2015.

[22] I. Ben-Gal, F. Ruggeri, F. Faltin, R. Kenett, "Bayesian Networks". Encyclopedia of Statistics in Quality and Reliability. John Wiley and Sons, Ltd. www.eng.tau.ac.il/bengal/BN.pdf/. pp. 1-6, 2007.

[23] S. Zhen, Montgomery, "Tuberculosis Dataset", Kaggle Open Machine Learning Datasets", 2020.

Authors Profile

Mr. Solomon Osarumwense Alile obtained a Diploma in Data Processing degree from University of Benin in 2004, B.Sc. degree in Computer Science in 2012 from Lagos State University (LASU), Ojo, Lagos and M.Sc in Computer Science from University of Benin in 2019. He is a Cisco Certified Network Associate (Routing and Switching) and System Engineer. He is a corporate member of Nigerian Institution of Professional Engineers and Scientists (NIPES), and member of the following societies namely: World Academy of Science, Engineering and Technology (WASET), International Computer Science and Engineering Society (ICSES), Institute For Engineering Research and Publication (IFERP), International Association of Engineers (IAENG), International Association of Engineers Society of Computer Science (ISCS), International Association of Engineers Society of Wireless Networks (ISWN), International Association of Engineers Society of Scientific Computing (ISSC), International Association of Engineers Society of Internet Computingand Web Services (ISICWS), International Association of Engineers Society of Information System Engineering (ISISE), International Association of Engineers Society of Data Mining (ISDM), International Association of Engineers Society of Artificial Intelligence (ISAI), and International Association of Engineers Society of Software Engineering (ISSE). He has published more than 12 research papers in reputed international journals which covers the area of cybersecurity and application of artificial intelligence in health diagnosis, all of which are available online. His skills are Machine Learning, Supervised Learning, Computer Networking, Routing and Switching and Network Security. His area of interest includes Information Technology, IoT(Internet of Things), Machine Learning and Cybersecurity. He is currently conducting research works in the area of Cybersecurity and development of AI Based Systems.