

CS 514 - Applied Artificial Intelligence
Project -3 Bayesian Networks (NETICA)
Diagnosis Of Tuberculosis

Username: Snow White

Problem Specification: Diagnosis of Tuberculosis using Bayesian Networks in NETICA

Abstract:

The Bayesian Network is designed using NETICA and it can diagnose the presence of Tuberculosis for a given individual. In this network, based on the risk factors of an individual and the tests performed, the probability of having Tuberculosis is computed. This network can be used for probability reasoning for the presence or absence of Tuberculosis.

In this network, the parameters used for determining the risk for TB disease are - Age, HIV Positive, Previously Infected, Sick with other diseases. Similarly the parameters used for determining the risk for TB infection are - Age, From Country where TB is common, Living or working in High Risk settings, Exposed to Infected, Exposed to people with High Risk for TB. I have considered these variables, as they are the most common risk factors that determine the presence or absence of Tuberculosis.

The network performs reasoning, based on the prior probabilities, risk factor values and results of various tests, to find the probability of presence or absence of Tuberculosis, and probability whether it is latent TB Infection or TB Disease.

Nodes and dependencies in the network:

Tuberculosis is caused by *Mycobacterium tuberculosis*. TB bacteria can live in the body without making you sick. This is called latent TB infection.

In some people the TB bacteria remain inactive for a lifetime without causing disease. But in other people, especially people who have a weak immune system, the bacteria become active, multiply, and cause TB disease.

1. The people at high risk for TB Disease include people with HIV infection, people who became infected with TB bacteria in the last 2 years, Babies and young children, people who are sick with other diseases that weaken the immune system, Elderly people, People who were not treated correctly for TB in the past.

To represent these factors, there are nodes for Age, HIV Positive, Previously Infected, Sick with other diseases. These nodes have prior probabilities. There is a node for High Risk For TB Disease which depends on the above four nodes.

2. The people at high risk for latent TB Infection include people who have spent time with someone who has TB disease, people from a country where TB disease is common (most countries in Latin America, the Caribbean, Africa, Asia, Eastern Europe, and Russia), people who live or work in high-risk settings (for example: correctional facilities, long-term care facilities or nursing homes, and homeless shelters), Health-care workers who care for patients at increased risk for TB disease, Infants, children and adolescents exposed to adults who are at increased risk for latent tuberculosis infection or TB disease.

To represent these factors, there are nodes for Age, From Country where TB is common, Living or Working in High Risk settings, Exposed to Infected, Exposed to people with High Risk for TB. These nodes have prior probabilities.

There is a node for High Risk For TB Infection which depends on the above five nodes.

3. The tests for diagnosing the presence of TB Infection are Mantoux tuberculin skin test (TST) and TB blood test.

The presence of acid-fast-bacilli (AFB) on a sputum smear or other specimen often indicates TB disease.

A posterior-anterior chest radiograph is used to detect chest abnormalities. These abnormalities may suggest TB..

There are four nodes - TB skin test, TB Blood test, Presence of AFB in Sputum smear, Chest radiograph which are performed, if an individual is at high risk for having TB infection or TB disease. To represent this dependency, there are arrows from High Risk For TB Disease and High Risk For TB Infection to these four nodes that represent the tests.

Vaccination with BCG may cause a false positive reaction to a TB skin test. To indicate this, there is a node BCG Vaccine and there is an arrow from BCG Vaccine to TB skin test.

4. Since there are two types of Tuberculosis, there are two nodes one for latent TB Infection and other for TB disease. These nodes indicate the presence or absence

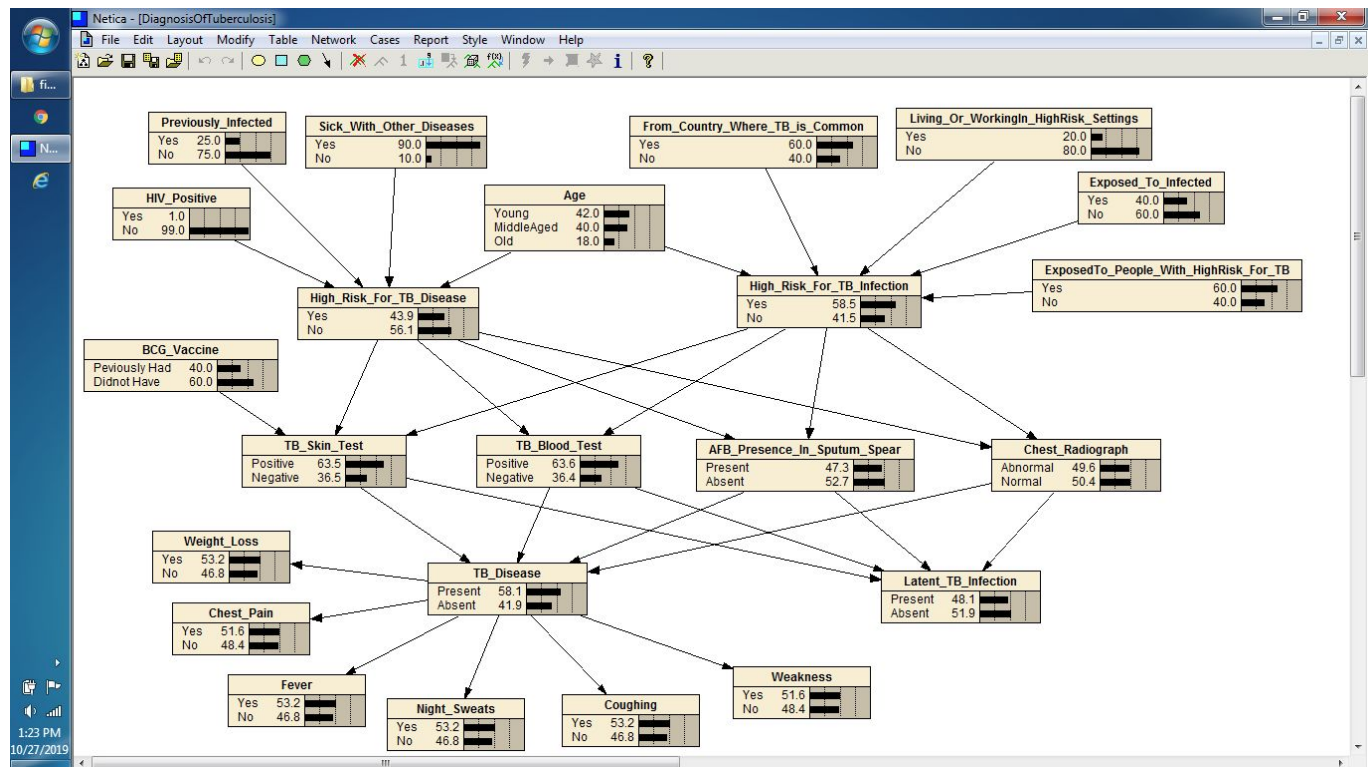
of latent TB Infection or TB disease. This depends on the four tests performed to diagnose TB. To represent this dependency, there are arrows from the four tests to latent TB Infection and TB disease.

5. The presence of TB disease causes few symptoms like Weight loss, Fever, Coughing, Chest pain, Night Sweats and Weakness. To represent this, there are arrows from TB disease to nodes representing Weight loss, Fever, Coughing, Chest pain and Weakness.

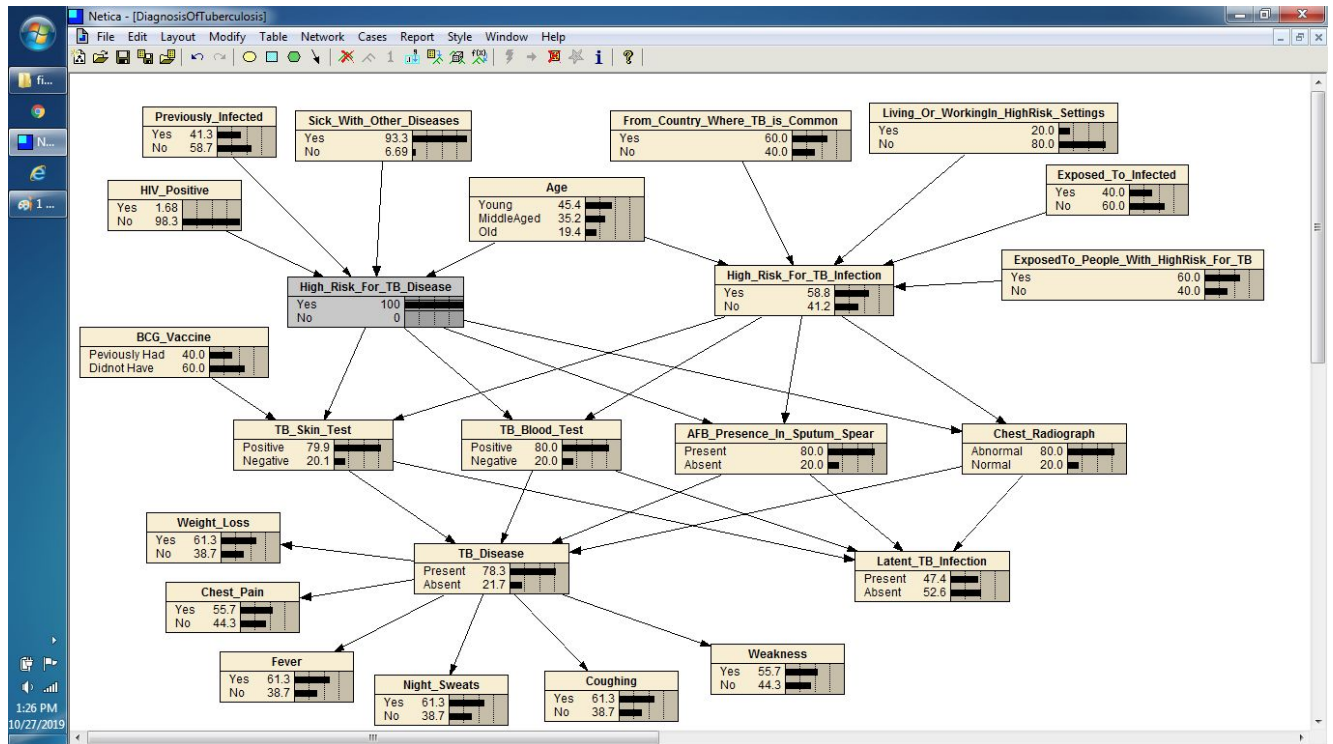
Note : Most of the probabilities are as close to reality as possible, while some have been formulated while hypothesizing the general trend.

Test Cases :

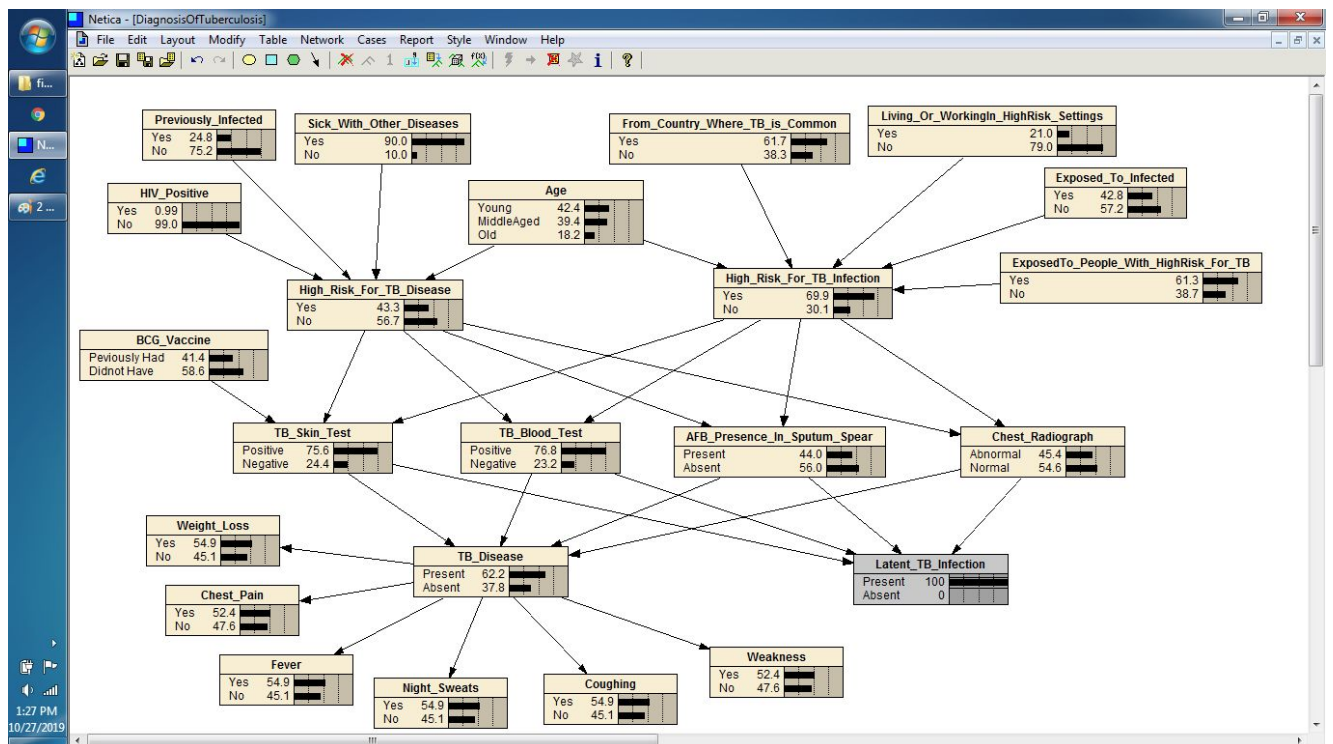
- ### 1. Probability of latent TB infection and TB disease without any evidence



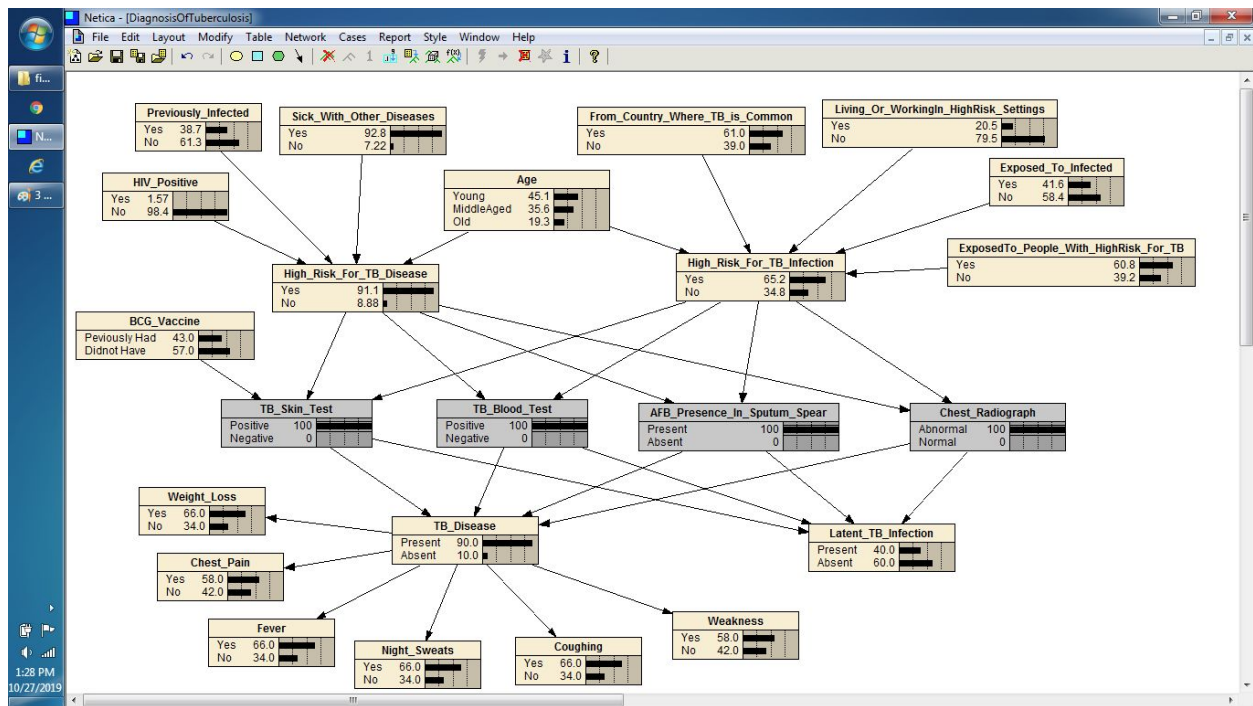
- ## 2. Probability of TB disease with evidence of high risk for TB disease



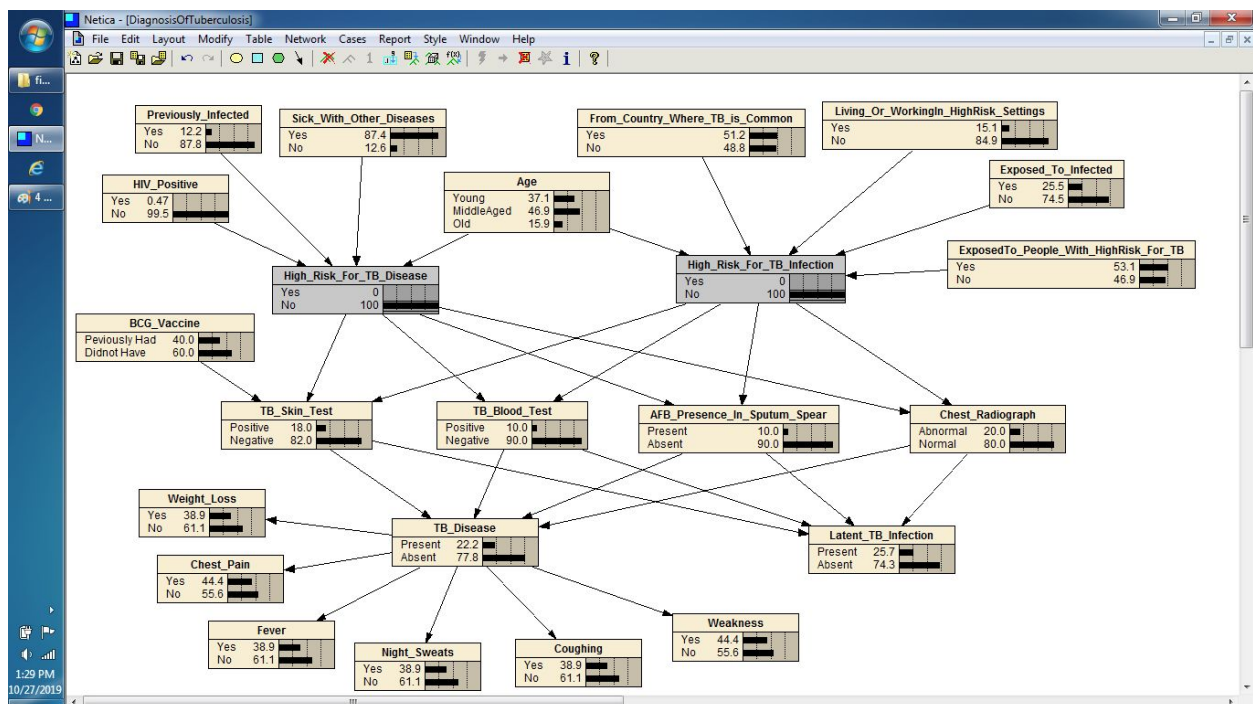
3. Marginal probability of each variable when TB Infection is present



4. Probability of TB disease with evidence of positive results in all tests.



5. Probability of TB disease and TB Infection, with evidence of no risk for latent TB infection or TB disease



Instructions to run :

1. Download and place "DiagnosisOfTuberculosis.neta" in any directory and open it through NETICA application.
2. Compile the project.
3. After the network is initialized, modify and set the values and see how the network adapts to the change in the values.

References :

1. <https://www.cdc.gov/tb/default.htm>