ITCS451, Semester 1/2022

Bayesian Network Project

Group members

No.	Student ID	Student name	Section
1.	6388014	Waris Damkhum	1
2.	6388022	Pattanan Korkiattrakool	1
3.	6388094	Vipavee Ngamyingsurat	1
4.	6388097	Thanakij Pinyoboon	1

1. Description of the problem you are modeling.
    - Cancer is a condition in which the body's cells proliferate unchecked. Lung cancer is the name for cancer that first appears in the lungs. In addition to lymph nodes and other body organs including the brain, lung cancer can also start in the lungs. Lung cancer can potentially spread from other organs. Metastases are cancer cells that have spread from one organ to another. Small cell and non-small cell lung tumors are the two primary subtypes that are typically recognized (including adenocarcinoma and squamous cell carcinoma). These distinct forms of lung cancer develop differently and respond to various therapies. Compared to small cell lung cancer, non-small cell lung cancer is more prevalent. This dataset is the effectiveness of cancer prediction system helps the people to know their cancer risk with low cost and it also helps the people to take the appropriate decision based on their cancer risk status.

2. Did you use real data?

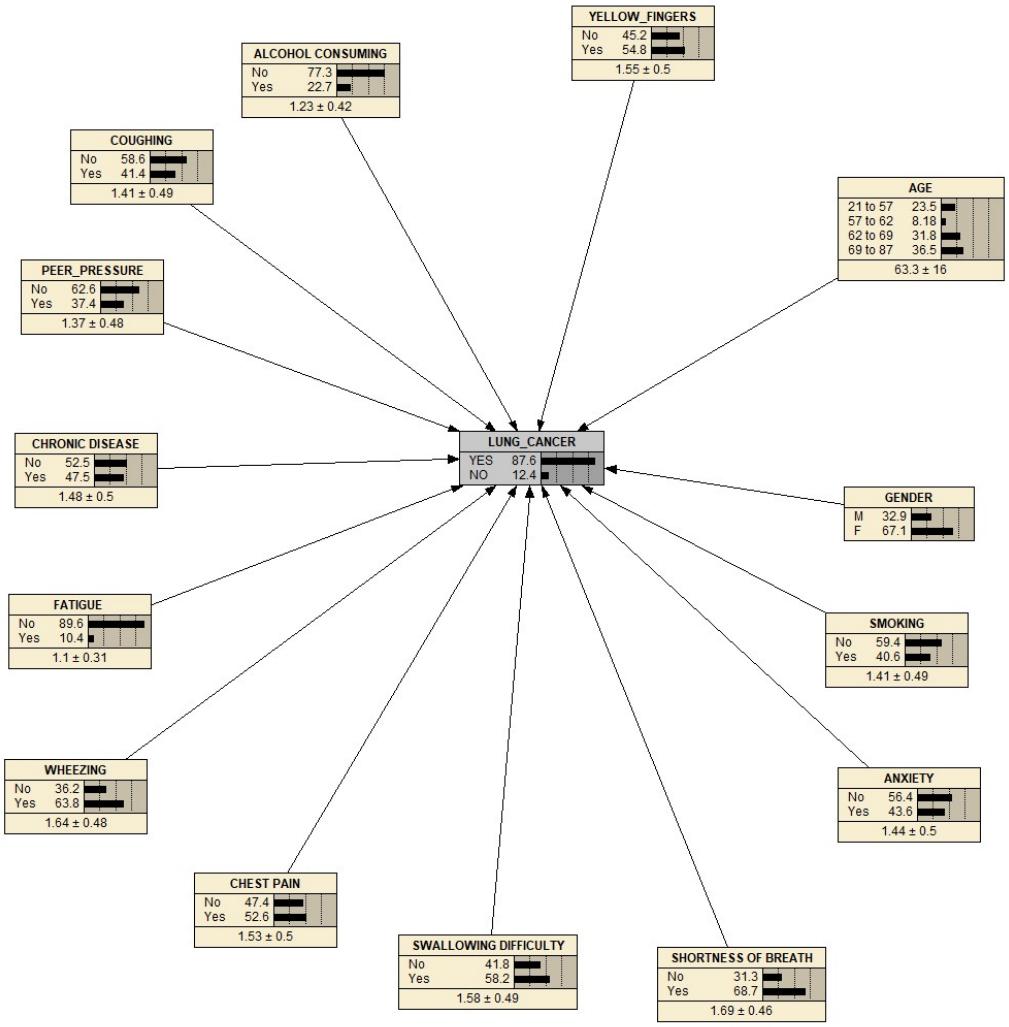
    ☑ Yes ☐ No

    If yes, describe your data in the following table.

Source of data	https://www.kaggle.com/datasets/mysarahmadbhat/lung-cancer
Type of data	☑ raw data ☑ published statistics
Modification to the data (If any)	Yes, modified. We have merged all ages into ranges from source data.

3. Bayes net description.
    a. Bayes net description
        - This model will show all possible symptoms of lung cancer by calculating all environments that cause lung cancer such as alcohol consumption, smoking, chest pain, age, gender, and furthermore.

    b. Bayes net screenshot

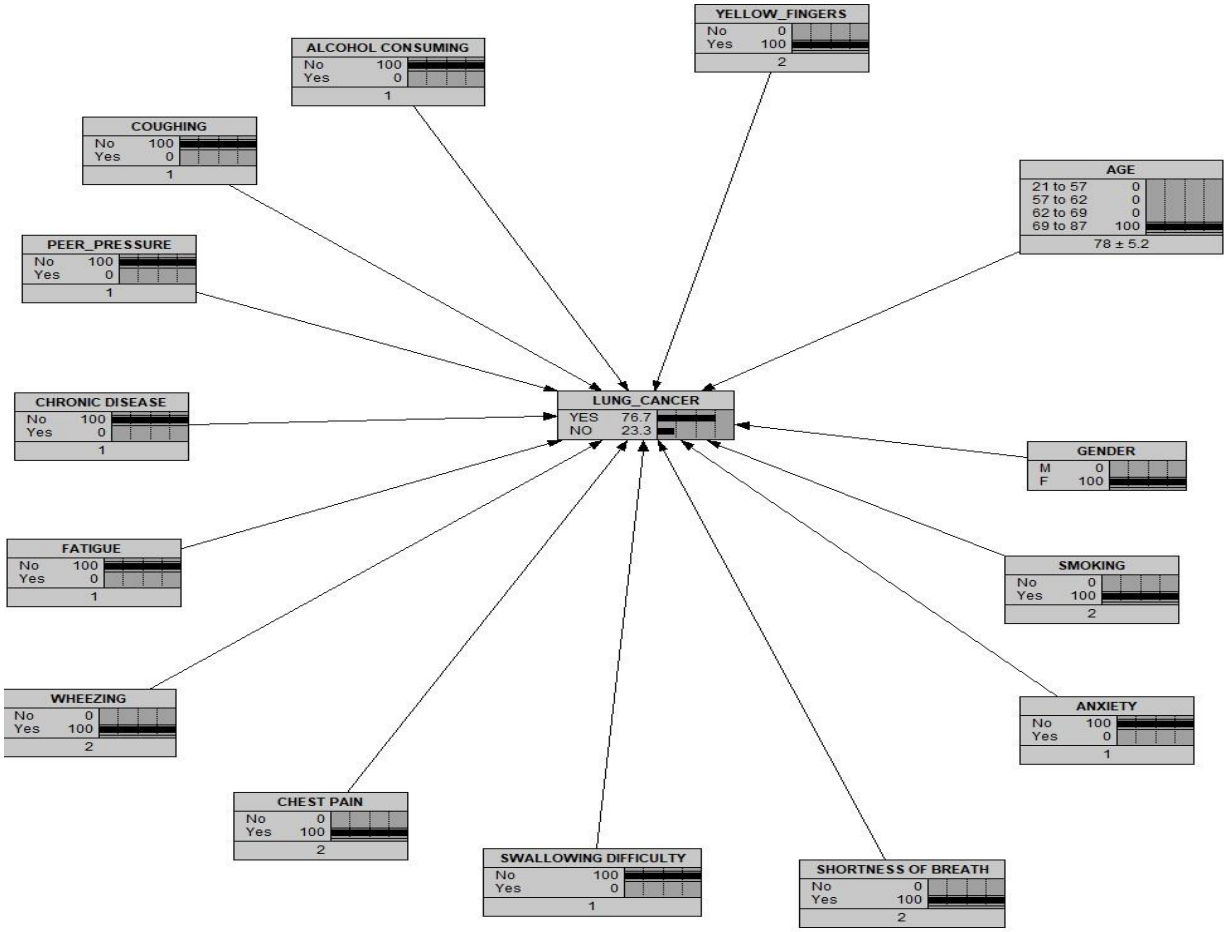


4. Explanation of your network

No.	Node Name	Node Description
1.	Lung_cancer	Have this specific symptom Yes = 2, No = 1
2.	Fatigue	Have this specific symptom Yes = 2, No = 1
3.	Age	In range of: 21 to 57 years old 57 to 62 years old 62 to 69 years old 69 to 87 years old
4.	Wheezing	Have this specific symptom Yes = 2, No = 1
5.	Alcohol Consuming	Have this specific symptom Yes = 2, No = 1
6.	Coughing	Have this specific symptom Yes = 2, No = 1
7.	Chronic Disease	Have this specific symptom Yes = 2, No = 1
8.	Peer_Pressure	Have this specific symptom Yes = 2, No = 1
9.	Anxiety	Have this specific symptom Yes = 2, No = 1
10.	Yellow_Fingers	Have this specific symptom Yes = 2, No = 1
11.	Smoking	Have this specific symptom Yes = 2, No = 1
12.	Gender	M(male), F(female)
13.	Shortness of breath	Have this specific symptom Yes = 2, No = 1

14.	Swallowing Difficulty	Have this specific symptom Yes = 2, No = 1
15.	Chest Pain	Have this specific symptom Yes = 2, No = 1

5. Two examples of inference: Enter some evidence and show the updated probabilities of the main hypothesis node(s). Explain what each example is showing.
    a. Example1
        - If the female subject had lung cancer between the ages of 69 to 87 with a yellow finger, they had always smoked, had shortness of breath, had chest pain, and had wheezing. Patients, however, do not experience anxiety, difficulty swallowing, fatigue, chronic disease, peer pressure, coughing, or alcohol consumption. The subject will have 76.7% lung cancer.



b. Example2

- If the male subject had lung cancer between the ages of 57 to 62, had always smoked, had difficulty swallowing, had shortness of breath, had fatigue, had a chronic disease, and had wheezing, Patients, however, do not experience anxiety, chest pain, peer pressure, coughing, a yellow finger, or alcohol consumption. The subject will have 53.4% lung cancer.