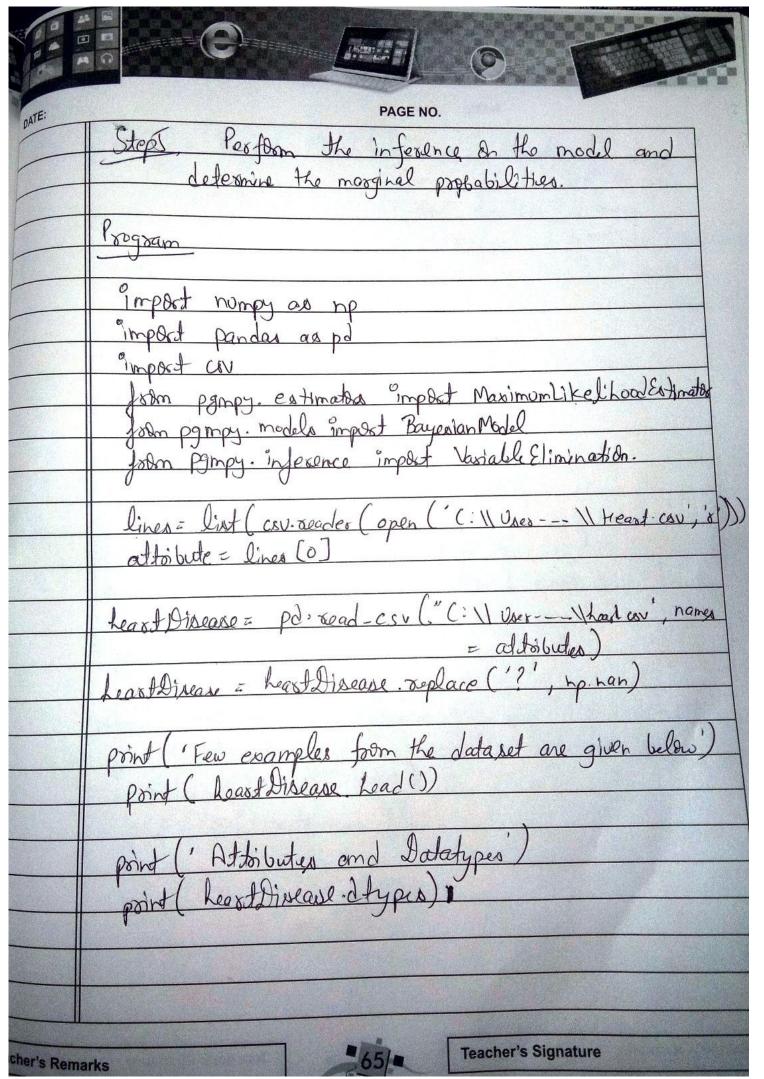# Lab Program - 7

Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Dataset. You can use Java/Python ML library classes/API

Theory- Steps to build the Bayesian networks-

Step1   Identify the variables which is set of attributes specified in the dataset (ex Medical Dataset)

Step2   Determine the domain of each variable that is set of values a variable may take.

Step3   Create a directed graph network of nodes where each node represents the attribute and edges represent parent child relationship. Edge represents that the child variable is conditionally dependent on the parent.

Step4   Determine the prior and conditional probability for each attribute.

Steps: Perform the inference on the model and determine the marginal probabilities.

Program

```python
import numpy as np
import pandas as pd
import csv
from pgmpy.estimators import MaximumLikelihoodEstimator
from pgmpy.models import BayesianModel
from pgmpy.inference import VariableElimination

lines = list(csv.reader(open('C:\\Uses---\\Heart.csv', 'r')))
attribute = lines[0]

heartDisease = pd.read_csv("C:\\User---\\heart.csv', names = attributes)
heartDisease = heartDisease.replace('?', np.nan)

print('Few examples from the dataset are given below')
print(heartDisease.head())

print('Attributes and Datatypes')
print(heartDisease.dtypes)
```

```python
model = BayesianModel([('age', 'trestbps'),
                       ('age', 'fbs'), ('sex', 'trestbps')
                       ('sex', 'trestbps'), ('exang',
                       'trestbps'), ('trestbps', 'heartdisease')
('fbs', 'heartdisease'), ('heartdisease', 'restecg'), ('heartdisease',
'thalach'), ('heartdisease', 'chol')])


print('Learning CPDs using Maximum Likelihood Estimators')
model.fit(heartDisease, estimator = MaximumLikelihoodEstimator)


print('Inferencing with Bayesian Network:')
HeartDisease_infer = VariableEliminator(model)


print('Probability of Heart Disease given Age=20')
q = HeartDisease_infer.query(variables = ['heartdisease'],
            evidence = {'age':28})
print(q['heartdisease'])


print('Probability of Heart Disease given chol (Cholesterol=100)')
q = HeartDisease_infer.query(variables = ['heartdisease'],
            evidence = {'chol': 100})
print(q['heartdisease']).
```

Output -

, Dataset are given below

	age	sex	cp	trestbps	chol	fbs	restreg	thalach	exang	oldpak	slpe
0	63	1	3	145	233	1	0	150	0	2.3	0
1	37	1	2	130	250	0	1	187	0	3.5	0
2	41	0	1	130	204	0	0	172	0	1.4	2
3	56	1	1	120	236	0	1	178	0	0.8	2
4	57	0	0	120	354	0	1	163	1	0.6	2

Infering the Bayesian Network:

1 Probability of HeartDisease Age = 20

Heart disease	phi (HeartDisease)
0	0.4058
1	0.5942

2 Probability of HeartDisease given chol (Cholesterol) = 100

Heart disease	phi (Heart disease)
0	0.0000
1	1.0000