**DS Lab**

**Exp-1**

**Name: Shashwat Tripathi**

**Division: D20A**

**Roll no: 64**

Aim: To explain the concept,draw the diagram, mention a few applications of Bayesian networks. Implement it with different datasets and show the output. Test the same with different inputs

Theory:

Bayesian Belief Network in Artificial Intelligence

A Bayesian belief network (BBN) is a fundamental technology in computer science for managing probabilistic events and solving problems that involve uncertainty. It can be defined as follows:

"A Bayesian network is a probabilistic graphical model that represents a set of variables and their conditional dependencies using a directed acyclic graph (DAG)."

This network is also known by various other names, such as Bayes network, belief network, decision network, or Bayesian model.

Bayesian networks are probabilistic because they are constructed from a probability distribution and utilize probability theory for tasks like prediction and anomaly detection. In real-world scenarios, which are often probabilistic by nature, Bayesian networks help represent relationships between multiple events. They are used in various applications, including prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision-making under uncertainty.

Applications of Bayesian Belief Networks

1. Medical Diagnosis:

- Disease Diagnosis: BBNs model the probabilistic relationships between symptoms and diseases, assisting in diagnosing diseases by calculating the likelihood of different conditions given observed symptoms.

- Treatment Planning: They aid in selecting the best treatment options by evaluating the probable outcomes of different treatments.

2. Fault Diagnosis:

- Industrial Systems: Used for diagnosing faults in complex machinery and industrial systems by modeling the relationships between different components and their failure modes.

- IT Systems: Applied to diagnose issues in computer networks, software systems, and hardware by analyzing the dependencies between various system components.

3. Decision Support Systems:

- Business Decisions: BBNs help make informed business decisions by evaluating factors such as market trends, financial data, and risk assessments.

- Environmental Management: Used in environmental decision-making to assess the impact of different policies and actions on ecosystems.

4. Robotics and Artificial Intelligence:

- Robot Navigation: Employed in robotic systems for navigation and path planning by modeling uncertainties in sensor data and environmental conditions.

- Autonomous Vehicles: Used in self-driving cars for decision-making under uncertainty, such as obstacle detection and route planning.

5. Speech and Image Recognition:

- Natural Language Processing: Applied in speech recognition systems to model the probabilistic relationships between phonemes, words, and sentences.

- Computer Vision: Used in image recognition and classification tasks by modeling the dependencies between various visual features.

6. Genetics and Bioinformatics:

- Gene Expression Analysis: BBNs help understand the relationships between genes and their expression levels, aiding in the study of genetic diseases and personalized medicine.

- Protein Structure Prediction: Used to predict the 3D structure of proteins based on probabilistic relationships between amino acids.

7. Finance and Economics:

- Risk Assessment: Applied in financial risk modeling to evaluate the probabilities of different financial events, such as market crashes or credit defaults.

- Portfolio Management: Used to optimize investment portfolios by modeling uncertainties and dependencies between various assets.

8. Ecology and Environmental Science:

- Habitat Modeling: BBNs are used to model the probabilistic relationships between environmental variables and species distributions.

- Climate Change Impact: Applied in assessing the impact of climate change on ecosystems by evaluating various environmental factors and their interdependencies.

9. Forensic Science:

- Evidence Analysis: Used in forensic investigations to evaluate the probabilistic relationships between different pieces of evidence and potential suspects or scenarios.

10. Education and Learning Systems:

- Student Modeling: BBNs assist in modeling student knowledge and learning processes, providing personalized learning recommendations based on the probabilistic assessment of student performance.

- Intelligent Tutoring Systems: Used to create adaptive learning environments that tailor content and feedback to individual student needs.

Structure of a Bayesian Network

A Bayesian network consists of:

- Directed Acyclic Graph (DAG): This graph is composed of nodes and directed arcs (arrows). Each node represents a random variable, which can be either continuous or discrete. The directed arrows indicate causal relationships or conditional dependencies between the random variables. For example, if there is an arrow from node A to node B, A is considered the parent of B, and B is dependent on A. If there is no directed link between two nodes, they are considered independent.

- Conditional Probability Tables (CPTs): These tables specify the conditional probability distribution for each variable, given its parent variables.

Influence Diagram

An influence diagram is a generalized form of a Bayesian network used for representing and solving decision problems under uncertain conditions.

Joint Probability Distribution

For a set of variables \( X\_1, X\_2, \ldots, X\_n \), the joint probability distribution is the probability of different combinations of these variables. It can be expressed as:

\[ P(X\_1, X\_2, \ldots, X\_n) = P(X\_1 | X\_2, \ldots, X\_n) P(X\_2, \ldots, X\_n) \]

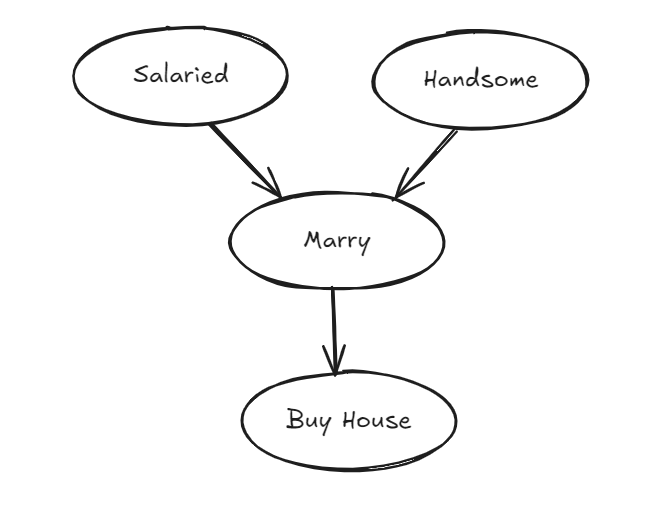
\[ = P(X\_1 | X\_2, \ldots, X\_n) P(X\_2 | X\_3, \ldots, X\_n) \ldots P(X\_{n-1} | X\_n) P(X\_n) \]

In general, for each variable \( X\_i \), the joint probability can be written as:

\[ P(X\_i | X\_{i-1}, \ldots, X\_1) = P(X\_i | \text{Parents}(X\_i)) \]

This formula illustrates how the probability of each variable depends on its parents in the network.

**Example:** We have the following Directed Acyclic Graph

****

**Code:**

from pgmpy.models import BayesianNetwork

from pgmpy.factors.discrete import TabularCPD

from pgmpy.inference import VariableElimination

# Define the Bayesian Network structure

model = BayesianNetwork([

('Salaried', 'Marry'),

('Handsome', 'Marry'),

('Marry', 'Buy House')

])

# Define the CPDs (Conditional Probability Distributions)

cpd\_salaried = TabularCPD(variable='Salaried', variable\_card=2, values=[[0.2], [0.8]])

cpd\_handsome = TabularCPD(variable='Handsome', variable\_card=2, values=[[0.35], [0.65]])

cpd\_marry = TabularCPD(variable='Marry', variable\_card=2,

values=[[0.05, 0.32, 0.47, 1.0],

[0.95, 0.68, 0.53, 0.0]],

evidence=['Salaried', 'Handsome'],

evidence\_card=[2, 2])

cpd\_buy\_house = TabularCPD(variable='Buy House', variable\_card=2,

values=[[0.08, 0.79],

[0.92, 0.21]],

evidence=['Marry'],

evidence\_card=[2])

# Add the CPDs to the model

model.add\_cpds(cpd\_salaried, cpd\_handsome, cpd\_marry, cpd\_buy\_house)

# Check if the model is valid

assert model.check\_model()

# Perform inference

inference = VariableElimination(model)

# Calculate the probability P(Buy House, Marry, Salaried, ~Handsome)

query = inference.query(variables=['Buy House', 'Marry'],

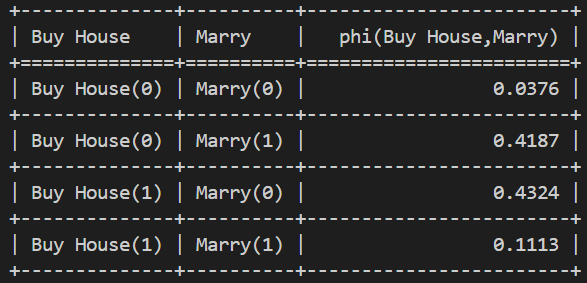
evidence={'Salaried': 1, 'Handsome': 0},

joint=True)

print(query)

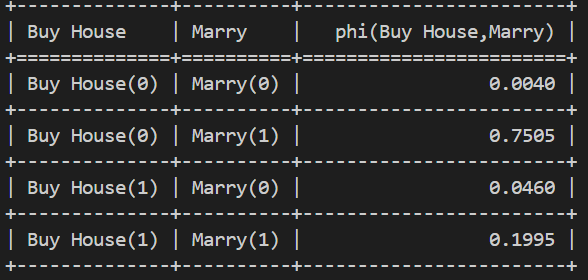
**Example 1:** Probability of Buy House when Married, Salaried, not Handsome P(Buy House, Marry, Salaried, ~Handsome)

**Output:**

****

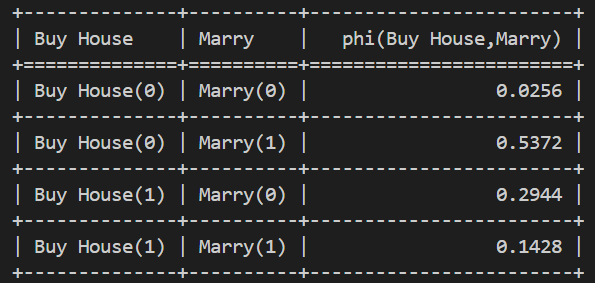
**Example 2:** Probability of Buy House when Married, not Salaried, not Handsome P(Buy House, Marry, ~Salaried, ~Handsome)

**Output:**

****

**Example 3:** Probability of Buy House when Married, not Salaried, Handsome P(Buy House, Marry, ~Salaried, Handsome)

**Output:**

****

**Conclusion:** Therefore we have studied the bayesian belief network and implemented the same in python for different data sets.