2023115062

Unit -4 Probabelistic Graphical Model

1) NAVIE BAYES ALGORITHM:

The Navie Bayes algorithm is a classification technically based on Bayes Theorem. It assumes that all features (attributes) are conditionally independent given the class label

$$P(c/x) = \frac{P(x/c)\ P(c)}{P(x)}$$

where

c -> class

x - feature ~~class~~ vector

since $P(x)$ is constant the decision is based on maximizing $P(x/c)\ P(c)$

Application:

* Spam Detection
* Medical Diagnosis
* Sentiment analysis

2) BAYESIAN BELEIF NETWORK: (BBN)

A Bayesian Belief Network (BBN) also

Known as a Bayesian Network, is a graphical model that represents probabilistic relationships among a set of random variables.

It is represented as a directed Acyclic Graph (DAG).

* Nodes represent random variables.
* Edges represents conditional dependencies.
* Each node has a conditional probability Table (CPT) that quantifies the effect of parent (cau

Example:

Consider medical diagnosis network:

Node 1 : Flu

Node 2 : Fever

Node 3 : Cough

Here Fever and Cough depends on Flu.

The network compactly encodes these dependencies.

Advantages :

* Captures casual relationship.
* Handles incomplete data.
* Supports both inference (predict unknowns) and Learning (update probabilities).

Application :

* Medical diagnosis (diseases & symptoms).
* Fault detection in engineering system.
* Decision support system.

3) Hidden Markov Model (HMM).

A Hidden Markov Model (HMM) is a statistical model for systems that evolve over time but where underlying states are hidden. Instead, we observe outcomes generated probabilistic from the hidden states.

Components a) Hidden states (S)
b) Observation (O)
c) Transition Probabilities (A)
d) Emission probabilities (B)

e) Initial state distribution (π)

Key problems solved by HMM:

* Evaluation: Compute probability of observation sequence given models.
* Decoding: Determine most likely sequence of hidden states (Viterbi algorithm).
* Learning: Estimate parameters.

Application:

* Speech recognition
* Natural Language Processing
* Bio Informatics.

4. BAYESIAN INFERENCE:

Bayesian Inference is a method of statistical inference where probability is used to represent uncertainity about parameter. Unlike frequentist methods, Bayesian methods update beliefs based on new evidence.

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)}$$

Where

H → hypothesis

D → Observed data

Advantages:

* Incorporates prior knowledge
* Naturally handles uncertainity.
* Produces full probability distributions.

PROBLEM BASED ON NAIVE BAYES

Problem:

A spam filter uses the Naive Bayes algorithm

Consider the word "offer" appearing in emails.

P(spam) = 0.4, P(Not spam) = 0.6, P(Offer|spam) = 0.8

P(offer|not spam) = 0.2

If an email contains the word "offer" classify it as spam or not spam.

Solution

$$P(\text{spam}|\text{offer}) = \frac{P(\text{offer}|\text{spam})\ P(\text{spam})}{P(\text{offer})}$$

P(offer) = P(offer|spam) + P(offer|Notspam) P(Not Spam).

$= (0.8)(0.4) + (0.2)(0.6)$

$= 0.32 + 0.12$

$= 0.44$

Now

$$P(\text{spam} / \text{offer}) = \frac{0.8 \times 0.4}{0.44} = 0.727$$

$$P(\text{Not spam} / \text{offer}) = \frac{0.2 \times 0.6}{0.44} = 0.273$$

Hence the email is classified as "spam".