(DEEMED TO BE UNIVERSITY)

ARTIFICIAL INTELLIGENCE
&
MACHINE LEARNING

SESSION NO : 17

Probability & Statistics

(Bayes' theorem and ML relevance)

Bayes' Theorem

Bayes' Theorem helps us update the probability of a hypothesis based on new evidence.

- It connects: Prior Probability (What we believed before)
- Likelihood (How likely the evidence is)
- Posterior Probability (What we believe after seeing the evidence)

◆ Formula:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Where:

- $P(A|B)$ = Posterior: Probability of A given B (updated belief)
- $P(B|A)$ = Likelihood: Probability of B assuming A is true
- $P(A)$ = Prior: Probability of A before seeing B
- $P(B)$ = Total probability of B

Bayes' Theorem

- **Example (Spam Filter using Bayes)**
- Suppose we want to detect if an email is spam based on the word "Free".

Prior: $P(\text{Spam}) = 0.3$

Likelihood: $P(\text{Free}|\text{Spam}) = 0.8$, $P(\text{Free}|\text{Not Spam}) = 0.1$

Evidence:

$$P(\text{Free}) = P(\text{Free}|\text{Spam})P(\text{Spam}) + P(\text{Free}|\text{Not Spam})P(\text{Not Spam})$$

$$= 0.8 \cdot 0.3 + 0.1 \cdot 0.7 = 0.31$$

Posterior:

$$P(\text{Spam}|\text{Free}) = \frac{0.8 \cdot 0.3}{0.31} \approx 0.77$$

If an email contains "Free", there's a **77% chance it's spam**.

Bayes' Theorem

Example: Weather and Umbrella

Prior (belief before evidence):

You wake up in the morning. Based on past experience, you think there's a **30% chance of rain today**.

Evidence (new data):

You look outside and see **dark clouds**.

Likelihood (how evidence fits the hypothesis):

- If it's going to rain, seeing dark clouds is **very likely** (say 80%).
- If it's not going to rain, seeing dark clouds is **rare** (say 20%).

Posterior (updated belief):

Using Bayes' theorem, after seeing clouds your belief changes:

Now, you estimate there's about a **66% chance of rain**.

So, you **update your belief**: you decide to **carry an umbrella**.

Machine Learning Relevance of Bayes’ Theorem

- Bayes’ theorem is fundamental in ML because it provides a **mathematical framework for reasoning under uncertainty.**

1. Naïve Bayes Classifier

A **direct application** of Bayes’ theorem.

Assumes features are conditionally independent given the class.

Formula:

$$P(C|x_1, x_2, \ldots, x_n) \propto P(C) \prod_{i=1}^{n} P(x_i|C)$$

Used in:

- Spam filtering
- Sentiment analysis
- Document classification

Machine Learning Relevance of Bayes' Theorem

2. Bayesian Inference (Learning Parameters)

In ML, model parameters (θ) are uncertain.

Bayes' theorem updates our belief about parameters after seeing data:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Used in:

- **Bayesian Neural Networks** (weights are probability distributions, not fixed values).
- **Bayesian Optimization** for hyperparameter tuning.

Machine Learning Relevance of Bayes' Theorem

3. Regularization as Priors

Many regularization techniques can be seen as **Bayesian priors:**

- L2 regularization ↔ Gaussian prior on parameters
- L1 regularization ↔ Laplace prior

This connection explains **why regularization prevents overfitting** → it encodes prior beliefs about parameter distributions.

Machine Learning Relevance of Bayes' Theorem

4. Uncertainty Estimation

Traditional ML models output point estimates (e.g., class label).

Bayesian models provide a **distribution** over predictions.

This is useful in:

- Medical diagnosis (confidence matters).
- Self-driving cars (uncertainty awareness prevents risky actions).
- Active learning (selecting uncertain samples for labeling).

---

5. Probabilistic Graphical Models

Bayesian Networks and Hidden Markov Models rely directly on Bayes' theorem for inference.

Widely used in **sequence modeling, speech recognition, recommendation systems**.

Machine Learning Relevance of Bayes’ Theorem

Bayes’ theorem in ML is important because :

- Enables **probabilistic classification** (Naïve Bayes).
- Supports **parameter learning and inference**.
- Justifies **regularization as prior knowledge**.
- Provides **uncertainty estimation** for safe and interpretable AI.
- Powers **graphical models** for structured prediction.

Where is it used in AI/ML?

1. Naive Bayes Classifier:
    - Classifies data using Bayes' Theorem + independence assumption
    - Fast, interpretable, works well for text classification, spam detection
2. Bayesian Networks:
    - Probabilistic graphical models
    - Represent complex conditional dependencies between variables
3. Model Updating:
    - Bayesian learning helps update beliefs as new data arrives

Terminal Questions

- State Bayes' theorem. Explain the terms **prior**, **likelihood**, **evidence**, and **posterior** with a real-life example.
- A disease affects 2% of the population. A test detects it correctly 95% of the time but gives a false positive 10% of the time.If a person tests positive, what is the probability they actually have the disease?
- Why is Bayes' theorem important in **Machine Learning**? Give at least two applications.
- What role does Bayes' theorem play in **uncertainty estimation** in ML models?

THANK YOU