Distinguishing Between Prior and Posterior Measures of Probability

What is Probability?

Def:- Probability is the chance of something happening

Ex…

If you flip a coin, you have two possible results: heads or tails

So, the chance (or probability) of getting heads is 1 out of 2, or 50%

Formula:

$$\text{Probability} = \frac{\text{Number of ways it can happen}}{\text{Total possible outcomes}}$$

For example:

- Dice has numbers 1 to 6.
- Probability of getting a 4 = $\frac{1}{6}$

Bayes Theorem Key Terms: Prior, Likelihood, Posterior, Evidence

Term	Meaning	Example
Prior	What you believe before seeing the data	Probability a person has a disease before a test
Likelihood	Probability of observing the data given a hypothesis	Probability of positive test if person has disease
Posterior	Updated belief after seeing the data	Probability person has disease given positive test
Evidence	Total probability of observing the data	Overall chance of getting a positive test

Bayes Theorem

Bayes' Theorem:

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

Application of bayes theorem

1) Medical Diagnosis

2) Spam Email Filtering

3) Recommendation Systems

4) Weather Forecasting

What is Prior Probability?

Def:-prior probability is like an educated guess or assumption we make about something before we get any new information

Ex…

A doctor initially believes a patient has a certain illness based on general knowledge of the population. If 5% of the population has a disease, then before examining the patient, the prior probability that they have it is 5%

What is Posterior Probability?

Def:- Posterior probability is the updated probability of an event happening after you have received new information or evidence

Ex…

imagine you initially thought there was a 60% chance it would rain today (prior probability). But then you saw dark clouds in the sky. Based on this new information (evidence), you update your belief, and now you might think there's an 80% chance of rain

Supervised vs Unsupervised Learning

Introduction to Machine Learning

Def:- Machine Learning is technique that design and developed algorithms that allow computers to learn from data

Types of Machine Learning

1) Supervised Learning

2) Unsupervised Learning

3) Reinforcement Learning

Supervised Learning

- **Definition**:- Supervised learning is a type of machine learning where the model is trained on a labeled dataset
- **Characteristics of Supervised Learning**

-Mainly used in classification and regression

-Common models include Linear Regression, Decision Trees, SVM, k-NN

Unsupervised Learning

- **Definition**:-Unsupervised learning is a type of machine learning where the model is trained on **unlabeled data**
- **Characteristics of Supervised Learning**

-Mainly used in Clustering and Dimensionality Reduction

-K-Means Clustering, Hierarchical Clustering, DBSCAN, PCA

Reinforcement Learning

- **Definition**:-Reinforcement learning (RL) is a type of machine learning where an agent learns to take actions in an environment to maximize cumulative rewards
- **Characteristics of Reinforcement Learning**

Game playing (e.g., AlphaGo), robotics, self-driving cars, automated trading

Examples of Supervised Learning

- Email Spam Detection
- Loan Prediction

Advantages of Supervised Learning

- High Accuracy
- Easy to Understand

Disadvantages of Supervised Learning

- Requires Labeled Data
- Expensive

Examples of Unsupervised Learning

- News Grouping
- Social Network Analysis
- Music Genre

Advantages of Unsupervised Learning

- No Labeled Data Needed

Disadvantages of Unsupervised Learning

- Hard to Interpret
- Lower Accuracy

Statistical Models for Structured and Unstructured Data

Introduction to Data

- What is Data?

  Data is is collection of RAW facts or information

- Types of Data

  Structured and Unstructured

Structured Data

Definition:-Structured data refers to any data that is organized in a predefined manner, typically stored in tabular formats like rows and columns

Ex…

Relational databases (Oracle, PostgreSQL)
Excel spreadsheets
CSV files

Characteristics of Structured Data

- Clearly defined information
- Organized in rows and columns
- Easily searchable

Unstructured Data

Definition:-Unstructured data is information that does not follow a predefined format or organizational structure

Ex…

Text documents (e.g., articles, emails)
Images (e.g., photos, scanned documents)
Audio (e.g., podcasts, voice messages)
Video (e.g., surveillance footage, movies)

Characteristics of Unstructured Data

- No predefined data model
- Difficult to search/analyze without processing

What are Statistical Models?

Definition:- A statistical model is a way to look at data and understand patterns or trends

Purpose:

- Helps make better decisions using data
- Makes guesses about what might happen next

Examples:

- Predicting house prices based on size and location
- Figuring out if an email is spam or not

Types of Statistical Models

- Descriptive Models

  Help us summarize and describe what has happened

- Inferential Models

  Help us make guesses or conclusions

- Predictive Models

  Help us predict what might happen in the future using past data

Models for Structured Data

- Regression models
- Classification models

Models for Unstructured Data

- Natural Language Processing (NLP)
- Image Processing

Tools & Libraries

- Scikit-learn
- TensorFlow
- PyTorch
- NLTK
- OpenCV