Chapter 2: Probability Fundamentals

Introduction

- Probability is the mathematical framework for expressing uncertainty.
- Forms the foundation for understanding data distributions, model generalization, and decision-making in machine learning.

Basic Probability Concepts

Sample Space

- The set of all possible outcomes of a random experiment.

Events

- A subset of the sample space.

Probability Measure

- A function that assigns probabilities to events.

Conditional Probability

- Probability of an event A given that another event B has occurred.
- Formula: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Independence

- Events A and B are independent if $P(A \cap B) = P(A) \times P(B)$

Bayes' Theorem

- Relates the conditional and marginal probabilities of events.

- Formula: $P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$

Random Variables

- A variable that can take different values randomly.

Distributions

Uniform Distribution

- All outcomes are equally likely.

Bernoulli Distribution

- A binary distribution with probability p for success.

Binomial Distribution

- The number of successes in n Bernoulli trials.

Poisson Distribution

- Models the number of events in a fixed interval of time or space.

Gaussian (Normal) Distribution

- A continuous distribution, defined by the mean μ and variance σ^2.

Exponential Distribution

- Models the time between events in a Poisson process.

Multivariate Gaussian Distribution

- Generalization of the Gaussian distribution to multiple dimensions.

Summary

- Understanding probability is crucial for machine learning.
- Different distributions model different kinds of data and uncertainty.