Study Notes: Fundamental Probability Rules & Bayes' Rule

Course: Machine Learning / Probability & Statistics

Purpose: Conceptual clarity, quick revision, and exam preparation.

1. Basic Probability Concepts

Sample Space (S): The set of all possible outcomes of a random experiment.

Event (A): A subset of the sample space.

2. Fundamental Rules of Probability

2.1 Addition Rule (Union of Events)

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Used when calculating the probability that A or B (or both) occur.

If A and B are mutually exclusive:

$P(A \cap B) = 0 \rightarrow P(A \cup B) = P(A) + P(B)$

2.2 Multiplication Rule (Joint Probability)

$P(A \cap B) = P(A) \times P(B|A) = P(B) \times P(A|B)$

Used to compute the probability that A and B both occur.

If A and B are independent: $P(A \cap B) = P(A) \times P(B)$

2.3 Conditional Probability

$P(A \mid B) = P(A \cap B) / P(B)$, if $P(B) > 0$

Represents the probability of A given B has occurred.

Conditional probability is not symmetric: $P(A \mid B) \neq P(B \mid A)$

3. Bayes' Rule (Bayes' Theorem)

Formula: $P(A \mid B) = [P(B \mid A) \times P(A)] / P(B)$

Interpretation:

• Prior Probability (P(A)) – Initial belief before evidence

• Likelihood (P(B|A)) – Probability of observing B if A is true

• Evidence (P(B)) – Total probability of the observed data

• Posterior Probability (P(A|B)) – Updated belief after seeing evidence

Example: Medical Diagnosis

A disease affects 2% of the population. A test detects it 98% of the time if present and gives a false positive 3% of the time.

Given: $P(D) = 0.02$, $P(+ \mid D) = 0.98$, $P(+ \mid \neg D) = 0.03$

Bayes' Rule:

$P(D \mid +) = [0.98 \times 0.02] / [(0.98 \times 0.02) + (0.03 \times 0.98)] = 0.0196 / 0.049 \approx 0.4$

Even with a positive test, the actual probability of having the disease is only 40%.

4. Applications in Machine Learning

Concept	Role
Naive Bayes Classifier	Uses Bayes' Rule to predict class labels
Spam Filtering	P(Spam | "free", "offer", etc.)
Bayesian Networks	Graphical models with conditional dependencies
Probabilistic Inference	Update predictions based on data

5. Summary Table

Rule	Formula	Use
Union	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$	"A or B"
Intersection	$P(A \cap B) = P(A) \times P(B\|A)$	"A and B"
Conditional	$P(A \mid B) = P(A \cap B) / P(B)$	"Given B"
Bayes' Rule	$P(A \mid B) = [P(B \mid A) \times P(A)] / P(B)$	Reverse inference