INDIVIDUAL TASK – MODULE 3

MACHINE LEARNING : CONCEPTS , ALGORITHMS AND APPLICATIONS

Bayes' Theorem in Real-life :Choose a real-world scenario (like medical testing or email spam filtering) and apply Bayes' theorem to calculate probabilities.

Bayes' Theorem in Real Life

Introduction

Bayes' Theorem is a fundamental concept in probability and Machine Learning. It is used to calculate the probability of an event after receiving new evidence. It helps in updating prior beliefs when new information becomes available.

Bayes' Theorem

Definition

Bayes' Theorem is a probability rule used to find the probability of an event after getting new information.It:

Calculates conditional probability

Updates old probability into new probability

Helps when we know but want to find

Formula

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

Meaning of Terms:

P(A|B) → Posterior probability

Probability of event A occurring given that B has occurred

P(B|A) → Likelihood

Probability of event B occurring given that A has occurred

P(A) → Prior probability

Initial probability of event A before evidence

P(B) → Evidence

Total probability of event B occurring

Expanded Form of Evidence

$P(B) = P(B|A)P(A) + P(B|A')P(A')$

Full Bayes' Theorem

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

Applications of Bayes' Theorem

1. Medical Test Diagnosis

Doctors use medical tests to detect diseases such as:

Cancer

Diabetes

COVID-19

HIV

Even if a test is highly accurate, a positive result does not always mean the disease is confirmed. This happens because the disease may be rare and false positives can occur.

Bayes' Theorem helps calculate the actual probability that a person truly has the disease after testing positive.

Simple Example (10,000 People)

Suppose:

Total people tested = 10,000

Disease rate = 1%

So:

Diseased people = 100

Healthy people = 9,900

Test Sensitivity = 99%

→ Out of 100 diseased:

99 test positive

1 test negative

Test Specificity = 95%

→ Out of 9,900 healthy:

9,405 test negative

495 test positive (false positives)

Total Positive Tests

True positives = 99

False positives = 495

Total positives = 99 + 495 = 594

Probability of Having Disease After Positive Test

P(Disease|Positive) = \frac{99}{594}

= 0.166 = 16.6\%

Conclusion:

Even with a 99% accurate test, the actual chance of having the disease after testing positive is only 16.6%.

This shows why Bayes' Theorem is important.

2. Spam Email Detection

Email services like Gmail, Yahoo, and Outlook must decide whether an email is:

Spam

Not Spam

Bayes' Theorem checks the probability of an email being spam based on words like:

"Congratulations"

"Free"

"Win money"

"Click here"

Machine Learning uses the Naïve Bayes Classifier, which is based on Bayes' Theorem.

Bayes' Theorem in Machine Learning

Bayes' Theorem is widely used in classification problems such as:

Sentiment analysis

Predicting student results

Image classification

Document classification

The Naïve Bayes algorithm is one of the simplest and most effective ML classification algorithms.

Real-Life Interpretation

Even if a system or test is very accurate:

Rare events can lead to misleading results

False positives can be high

Actual probability may be much lower than expected

Bayes' Theorem helps us find the true probability after considering all factors.