MODULE-3 Machine Learning: Concepts, Algorithms, and Applications

Individual Task

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

Introduction

Bayes’ Theorem is one of the most powerful concepts in probability and statistics because it allows us to update our belief about an event when new information becomes available. In real life, decisions are rarely made with complete certainty. Instead, we rely on probabilities that change as we gather evidence. Bayes’ Theorem provides a mathematical way to revise these probabilities logically and accurately. It is widely used in fields such as medicine, artificial intelligence, machine learning, finance, weather forecasting, and cybersecurity.

A particularly meaningful real-world application is **medical testing**, where doctors must determine whether a patient truly has a disease after receiving a test result. Many people assume that if a test is highly accurate and the result is positive, then the person definitely has the disease. However, probability theory shows that this assumption can be misleading. Bayes’ Theorem helps calculate the actual likelihood.

This report explains Bayes’ Theorem using a realistic medical screening example and demonstrates how probabilities change after receiving test results.

Bayes’ Theorem Formula

The mathematical expression for Bayes’ Theorem is:

P(A|B)=P(B|A)×P(A)P(B)P(A|B) = \frac{P(B|A)\times P(A)}{P(B)}P(A|B)=P(B)P(B|A)×P(A)

Where:

- **P(A|B)** → Posterior probability (probability of event A after evidence B is known)
- **P(B|A)** → Likelihood (probability of observing evidence B if A is true)
- **P(A)** → Prior probability (initial belief before evidence)
- **P(B)** → Total probability of evidence

In simple words:

Posterior = (Likelihood × Prior) ÷ Evidence

Real-Life Scenario: Disease Screening Test

Suppose a disease screening test is used to detect a rare illness in a population. The following information is known:

- Only **1% of people** actually have the disease.

  $P(Disease)=0.01P(Disease)=0.01P(Disease)=0.01$

- The test correctly detects the disease in **95% of infected people**.

  $P(Positive|Disease)=0.95P(Positive|Disease)=0.95P(Positive|Disease)=0.95$

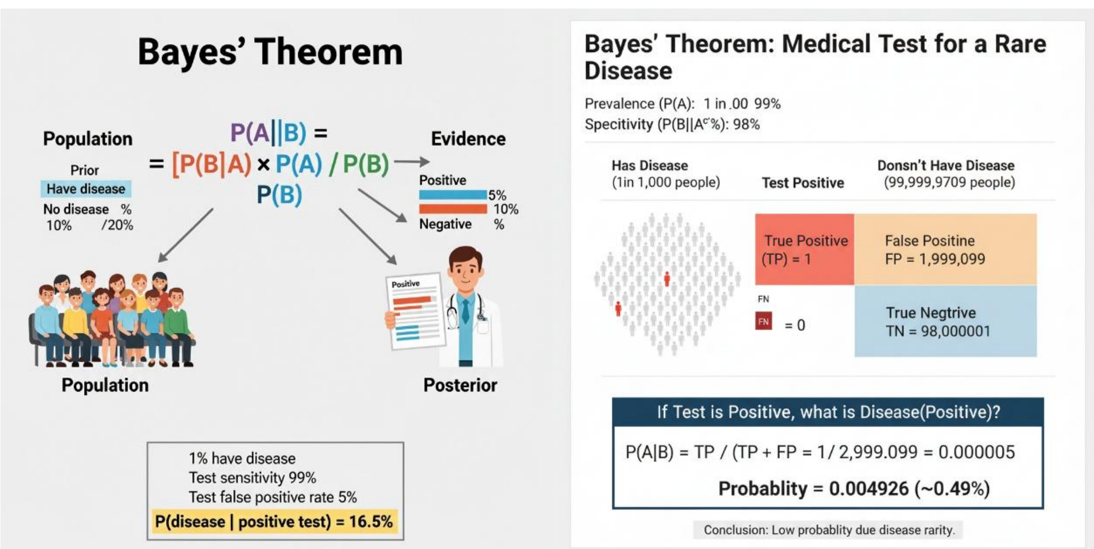
- The test gives a false positive result for **5% of healthy people**.

  $P(Positive|NoDisease)=0.05P(Positive|NoDisease)=0.05P(Positive|NoDisease)=0.05$

We want to calculate:

$P(Disease|Positive)P(Disease|Positive)P(Disease|Positive)$

This means: **If someone tests positive, what is the real probability that they actually have the disease?**



Step-by-Step Solution

Step 1 — Determine Prior Probabilities

The probability of not having the disease:

P(NoDisease)=1−0.01=0.99P(NoDisease)=1-0.01=0.99P(NoDisease)=1−0.01=0.99

Step 2 — Calculate Probability of Testing Positive

A positive test can occur in two ways:

1. True Positive → Person has disease and test detects it.
2. False Positive → Person is healthy but test still shows positive.

So,

P(Positive)=P(Positive|Disease)×P(Disease)+P(Positive|NoDisease)×P(NoDisease)P(Positive)= P(Positive|Disease)\times P(Disease) + P(Positive|NoDisease)\times P(NoDisease)P(Positive)=P(Positive|Disease)×P(Disease)+P(Positive|NoDisease)×P(NoDisease) =(0.95)(0.01)+(0.05)(0.99)= (0.95)(0.01) + (0.05)(0.99)=(0.95)(0.01)+(0.05)(0.99) =0.0095+0.0495= 0.0095 + 0.0495=0.0095+0.0495 =0.059= 0.059=0.059

Step 3 — Apply Bayes' Theorem

P(Disease|Positive)=P(Positive|Disease)×P(Disease)P(Positive)P(Disease|Positive)=\frac{P(Positive|Disease)\times P(Disease)}{P(Positive)}P(Disease|Positive)=P(Positive)P(Positive|Disease)×P(Disease) =0.95×0.010.059= \frac{0.95 \times 0.01}{0.059}=0.0590.95×0.01 =0.00950.059= \frac{0.0095}{0.059}=0.0590.0095 ≈0.161\approx 0.161≈0.161

Final Answer

P(Disease|Positive)≈0.161=16.1%P(Disease|Positive) \approx 0.161 = 16.1\%P(Disease|Positive)≈0.161=16.1%

Interpretation of Result

Even though the test is **95% accurate**, a positive result only means there is about a **16% chance** the person actually has the disease. This seems surprising at first, but it happens because the disease is very rare. Since 99% of people are healthy, even a small false positive rate produces many incorrect positive results.

This phenomenon is known as the **Base Rate Effect** — when the rarity of an event strongly affects probability outcomes.

Understanding with Realistic Numbers (Visualization Approach)

Imagine testing **10,000 people**:

- Diseased people = 1% of 10,000 = **100**
- Healthy people = **9,900**

Among diseased:

- True positives = 95% of 100 = **95**

Among healthy:

- False positives = 5% of 9,900 = **495**

Total positive results = 95 + 495 = **590**

Actual diseased among positives = 95

95590≈16.1%\frac{95}{590} \approx 16.1\%59095≈16.1%

So out of 590 positive results, only 95 actually have the disease.

Importance in Real Life

Medical Decision Making

Doctors do not rely solely on test results. They also consider:

- Patient symptoms
- Medical history
- Risk factors

Bayesian reasoning helps doctors decide whether additional testing is necessary.

Artificial Intelligence & Machine Learning

Many AI systems use Bayesian probability to make predictions and improve accuracy as new data arrives. For example:

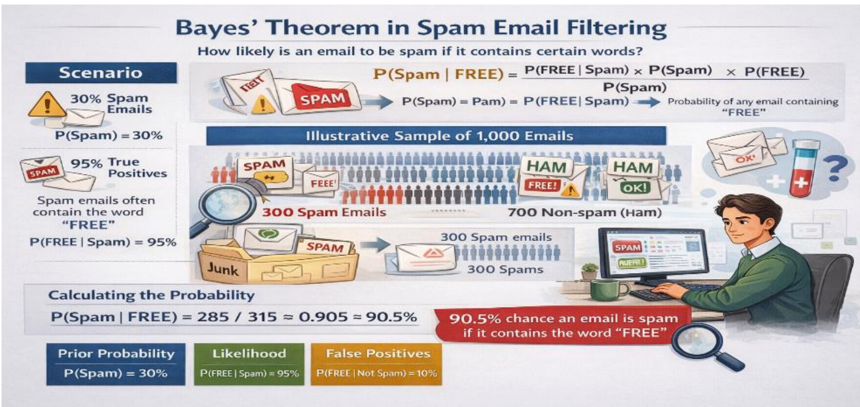
- Speech recognition
- Image classification
- Recommendation system.

Spam Email Filtering

Spam filters calculate the probability that an email is spam based on keywords, sender history, and previous spam data. Each new email updates the probability dynamically using Bayes' rule.

Fraud Detection

Banks estimate the probability of fraud for each transaction based on previous transaction patterns. If probability exceeds a threshold, the transaction is flagged.



Advantages of Bayesian Reasoning

- Updates probability with new evidence
- Works well with incomplete information
- Reduces decision errors
- Improves prediction accuracy
- Supports rational decision-making

Limitations

- Requires reliable prior probabilities
- Can be computationally intensive for large datasets
- Results depend heavily on input accuracy

Conclusion

The medical testing example clearly illustrates how Bayes' Theorem works in real life. Although the test in our scenario is highly accurate, the probability that a person actually has the disease after testing positive is only about **16.1%** because the disease itself is rare. This example highlights the importance of considering prior probabilities along with new evidence when making decisions.

Bayes' Theorem is not just a mathematical formula—it is a powerful reasoning tool used in healthcare, artificial intelligence, cybersecurity, business analytics, and many other fields.