Bayesian Intelligence: From Data to Decisions - A Case Study

Tarikka B.[1]; Saanvi G.[2]; P. Sri Lekha[3]; Dr. Sivasakti Balan D. P.[4]; R. J. Thayumanaswamy[5]

[1;2][11E]; [3][Guide]; [4][Principal]; [5][CEO]

[1;2;3;4;5]K. R. M. Public School

Publication Date: 2025/12/27

Abstract: Bayes' Theorem is a mathematical formula in probability theory that calculates the conditional probability of an event based on prior knowledge of related conditions. Bayes' Theorem is used in many areas like weather forecasting, spam filtering, medical diagnosis, and more. Other examples include credit risk assessment, quality control, search engines, recommendation systems, stock market prediction, and AImachine learning. This research highlights how Bayes theorem helps in interpreting medical test results, classifying emails as spam or legitimate and to identify suspects or causes in Forensics more accurately.

Keywords: *Bayes Theorem, Probability, Theoretical Frame, Smoking, Covid 19.*

How to Cite: Tarikka B.; Saanvi G.; P. Sri Lekha; Dr. Sivasakti Balan D. P.; R. J. Thayumanaswamy (2025) Bayesian Intelligence: From Data to Decisions - A Case Study. *International Journal of Innovative Science and Research Technology,* 10(12), 1811-1814. https://doi.org/10.38124/ijisrt/25dec1306

I. INTRODUCTION

Bayes' theorem is named after Reverend Thomas Bayes, an English statistician, philosopher, and Presbyterian minister. Around 1760, Thomas Bayes worked on a mathematical problem: How can we calculate the probability of a cause, given the effect we observe?His manuscript was published after his death in 1763 by Richard Price.Today, it forms the foundation of modern probability and statistics.

➢ *Objective of the Study:*

The main objective of this study is to understand Bayes' Theorem and examine its practical applications in real-life situations, especially in modern technologies like spam detection, medical diagnosis, and machine learning. The primary objective of this research involving Bayes theorem is to explore and refine its ability to systematically calculate conditional probabilities and update beliefs based on new evidence. This research aims to improve decision-making processes in uncertain situations across various fields.

➢ *Scope of the Study:*

This study focuses on understanding the theoretical foundation of Bayes' theorem, its historical development, and its importance in modern probability theory. The study also covers simple numerical examples to illustrate the working of the theorem and discuss its advantages, limitations and relevance in today's data driven world. The research is confined to conceptual understanding,practical applications and case based explanations rather than advanced mathematical derivations.The scope includes analyzing how Bayes' theorem is applied in real-life decision-making, especially in fields such as medicine (diagnostic testing), spam filtering and Forensics.

II. METHODOLOGY

The methodology that we used to analyze the research is a qualitative approach. Secondary data has been used for this research. We have also used some of the case studies from medical and email spam detection.

III. REVIEW OF LITERATURE

Bayes' Theorem was first introduced by Thomas Bayes in 1763 to find the probability of an event based on prior information and new evidence. Later mathematicians like Laplace developed it further and gave it a clear mathematical form. Modern textbooks explain Bayes' theorem as an important tool in probability and statistics for updating beliefs when new data is obtained. Many researchers have shown that it is widely used in medical diagnosis, forensics, machine learning, and decision making, where accurate prediction based on evidence is required. Thus, the existing literature proves that Bayes' Theorem has both strong theoretical importance and practical applications. Thomas Bayes is credited with discovering the theorem who wrote a paper titled " An essay towards solving a problem in the Doctrine

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1306

of chances". After bayes died in 1761, his friend Richard Price found the unpublished manuscript and published and presented it to the royal society in 1763. Later Pierre Simon Laplace generalized, developed and popularized it.

➢ *Theoretical Framework:*

$$P(E1/A) = \frac{P(E1)\,P(A/E1)}{P(E1)\,P(A/E1) + P(E2)\,P(A/E2)}$$

- E1 : First possible reason for event A .
- E2 : Second possible reason for event A .
- A : Event that has happened.
- P(E1) : Chance that E1 happens.
- P(E2) : Chance that E2 happens.
- P(A/E1) : Chance that A happens if E1 happened.
- P(A/E2) : Chance that A happens if E2 happened.
- P(E1/A) : Chance that E1 is the reason after A happens.
- Denominator: Total chance that A happens.(Total probability)

➢ *Applications:*

- Spam filtering: It classifies emails as spam based on content and sender info. Words like "free", "offer", "win" have higher chances in spam mails. Bayes rule updates the probability of a message being spam based on words appearing.
- Medical diagnosis: It predicts disease likelihood based on symptoms and test results. A test result is 90% accurate. Diseases are rare in the population. Using Bayes' theorem, doctors calculate the actual chance a person has the disease after a positive test. This prevents wrong treatment.
- Weather forecasting: It helps in predicting weather events using past historical data and current climatic conditions.Weather departments combine previous data + new satellite images to update probability of rain.
- AI and Other Uses Machine Learning: Powers image classification, recommendation systems, and self-driving car navigation.
- Finance: Assesses investment risks and market forecasts.Assesses investment risks and market forecasts using economic indicators.Banks evaluate transaction risks or flag fraud by combining patterns with economic data.
- Forensics: Evaluates DNA evidence strength in criminal cases.Courts quantify evidence strength by matching crime scene samples to suspect profiles.Updating probability of suspect being guilty after new evidence appears.

➢ *Advantages:*

- It starts with a prior belief. As data comes in, we can update to a posterior probability. Perfect for dynamic Systems like medical diagnosis and spam filtering.
- Gives a probability distribution instead of a single yes or no answer. You know how confident the conclusion is.
- Prior can come from past studies, expert opinion, or even a neutral guess. The theorem merges it with current data, making use of everything you have.
- It is the foundation of many algorithms like text classification, spam detection, casual modelling, risk assessment, recommendation systems, robotics and finance risk models etc.
- It is flexible across many domains like medicine, law, search engines, finance etc.
- The formula of Bayes theorem is intuitive.
- Unlike frequenting methods that need large Samples, Bayesian approach can still give meaningful results when data is scarce.
- It helps pick actions that maximize expected payoff.

➢ *Case Study -1*

- *Cancer Diagnosis*

A 45-year-old man is a chain smoker and has been addicted to smoking for many years. He visits a doctor complaining of persistent cough, chest pain, breathlessness, and fatigue.

The doctor observes that these symptoms are commonly associated with lung cancer, but they can also occur due to other lung diseases like bronchitis or infection.To confirm the diagnosis, the doctor recommends a lung cancer screening test (for example, CT scan or biopsy).

However, no medical test is 100% accurate. Sometimes, a test may give:

✓ A positive result even when the patient does not have cancer (false positive), or

✓ A negative result even when the patient has cancer (false negative).

Hence, the doctor uses Bayes' Theorem to calculate the actual probability that the patient has lung cancer given that the test result is positive.

P(C) = 0.05 = 5%

P(C') = 0.95 = 95%

P(+|C) = 0.90

P(+|C') = 0.10

By applying these values in the formula,

P(C|+) = 0.321(approx)

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1306

P(C|+) = 32%

- *Interpretation:*

Even though the patient is a smoker and has symptoms of lung cancer and the test result is positive, the actual probability that he had lung cancer is only 32%.

Bayes theorem is crucial in cancer diagnosis because it prevents panic due to false positives, avoids unnecessary treatment, helps doctor decide whether further treatment is needed or not and improves clinical decision making.

➢ *Case Study 2*

- *Covid 19*

A woman visits a doctor showing symptoms of COVID-19 such as fever, cough, and fatigue. The doctor recommends a COVID diagnostic test to determine whether she is infected. Since medical tests are not 100% accurate, Bayes' Theorem is used to calculate the probability that the woman actually has COVID given that her test result is positive.

C = event that the woman has covid

T+= event that the COVID test result is positive.

- *Probability that the Test Result is Positive*

P(C) = 0.05 = 5%

P(T+|C) = 0.95 = 95%

P(T+|C) = 0.10

By applying these values in the formula,

P(C | T+) = 0.33(approx) = 33%

- *Interpretation:*

Even though the woman shows symptoms of COVID and the test result is positive , the actual probability that she has COVID is only 33%.

➢ *Case Study 3 -Spam Filtering*

Suppose out of all emails,40% of emails are spam and 60% of emails are non-spam.

P(spam)=0.4

P(non spam)=0.6

The word 'Free' appears in 50% of spam emails and in 10% of non spam emails.

P(Free/Spam)=0.5

P(Free/non spam)=0.1

Applying bayes theorem,

$$P(spam/ Free)=\frac{P(Free/ spam)\ P(spam)}{P(Free)}$$

P(Free) = (0.5 x 0.4) + (0.1 x 0.6)

P(Free) = 0.20 + 0.06 = 0.26

$$P(Spam \mid Free) = \frac{0.5 \times 0.4}{0.26} = \frac{0.20}{0.26} = 0.77$$

P(Spam | Free) = 77%

- *Interpretation:*

An email containing the word "free" has a 77% of being spam. So the spam filter will classify it as spam.

➢ *Case Study 4 Forensics*

A crime has occurred in a city. Some DNA evidences are collected from the crime scene by the forensics team. A suspect is tested and the DNA test result matches the crime scene DNA. The court now asks an important question: " What is the probability that the suspect is actually guilty, given that the DNA matches?

Bayes' theorem can help us answer this question scientifically?

- *Given Datas:*

G = Event that the suspect is guilty

I = Event that the suspect is guilty

M = Event that the DNA test matches

P(G) = 1/10,000

P(M | G) = 0.99

P (M | I) = 0.001

P(I) = 1 – P(G) = 9999/10,000

Applying these values in the Bayes theorem we get,

$$P(G \mid M)= \frac{P(M \mid G)\ P(G)}{P(M|G)\ P(G) + P(M|I)\ P(I)}$$

P(G | M) = 0.09 or 9%

- *Interpretation:*

Even though the DNA test matched, the probability that the suspect is actually guilty is only about 9%. This shows that DNA evidence must not be used alone, prior probability is crucial and Bayes' Theorem prevents wrong convictions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1306

IV. RESULT

From all the above case studies we can see how Bayes' theorem plays an important role in medical diagnosis, spam filtering and forensics. From case study 1 and 2 we can say that Bayes theorem' helps doctors estimate the true probability of cancer yt combining test results with prior risk factors, making diagnosis more accurate and reduces misdiagnosis by incorporating prior probabilities. From case study 3 we can say that bayes theorem has been highly effective in accurately classifying emails as spam or legitimate. From case study 4 its clearly visible that Bayes' theorem helps in transforming forensic evidences into reliable truth by combining data with probability.

V. LIMITATIONS

- The prior is often based on belief, past data or guesswork. If it is not proper, the posterior can be twisted, especially with small data.
- As the number of variables increase, calculating the denominator becomes difficult. It requires approximation methods which adds complexity and runtime.
- If features are dependent, the model can give poor probability estimates despite decent accuracy.
- With sparse data, the prior dominates. If prior is weak or wrong, results stay unreliable until enough Upevidence accumulates.
- People sometimes confuse P(A/B) with P(B/A). The theorem itself is correct, but human intuition often misleads it.
- Calculating P(B) over all possible A can be infeasible When A has many categories or is continuous.
- If the model is misspecified, Bayesian updating can become over confident in a wrong conclusion because it never questions the model structure.
- When prior and observed data clash strongly, the posterior may still be pulled toward the prior, leading to delayed Learning.

VI. CONCLUSION

In conclusion, Bayes' Theorem plays a vital role in understanding and managing uncertainty in real-life situations. By combining prior knowledge with new evidence, it provides a systematic and logical method for updating probabilities and making informed decisions. This makes Bayes' Theorem not just a theoretical concept in probability, but a powerful tool with wide practical applications.Thus, Bayes' Theorem forms a bridge between mathematics and real-world problem solving. Its ability to continuously refine predictions as new information becomes available highlights its importance in both scientific research and everyday life. Therefore, the study of Bayes' Theorem is essential for developing analytical thinking and understanding how mathematics influences decision – making in the modern world .

RECOMMENDATIONS

Bayes' theorem plays a key role in artificial intelligence and machine learning for handling uncertainty . It helps in data science and big data analytics by updating predictions with new data . It is important in cyber security and fraud detection to identify threats and spam . It is essential for robotics and autonomous systems . Therefore the future scope of Bayes' theorem is vast and promising.

REFERENCES

[1]. Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances. Philosophical Transactions of the Royal Society of London, 53, 370–418.
[2]. Feller, W. (1968). An Introduction to Probability Theory and Its Applications (Vol. 1). Wiley, New York.
[3]. Ross, S. M. (2014). Introduction to Probability Models (11^{th} ed.). Academic Press.
[4]. DeGroot, M. H., & Schervish, M. J. (2012). Probability and Statistics (4^{th} ed.). Pearson Education.
[5]. Jaynes, E. T. (2003). Probability Theory: The Logic of Science. Cambridge University Press.
[6]. Gelman, A., et al. (2013). Bayesian Data Analysis (3^{rd} ed.). Chapman & Hall/CRC.
[7]. Grinstead, C. M., & Snell, J. L. (2012). Introduction to Probability. American Mathematical Society.
[8]. Devore, J. L. (2015). Probability and Statistics for Engineering and the Sciences (9^{th} ed.). Cengage Learning.
[9]. Kreyszig, E. (2011). Advanced Engineering Mathematics (10^{th} ed.). Wiley.
[10]. Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2017). Probability and Statistics for Engineers and Scientists (9^{th} ed.). Pearson.