

# Solving Bayesian reasoning tasks with ChatGPT and Gemini

Renato A. Krohling

UFES- Federal University of Espírito Santo  
Labcin - Nature-Inspired Computing Lab &  
PPGI - Graduate Program in Informatics  
Av. Fernando Ferrari, 514, Prédio CT X  
CEP 29060-270, Vitória, ES  
e-mail: krohling.renato@gmail.com

**Abstract:** This paper expands on the exploration of Bayesian problem-solving capabilities in large language models (LLMs), specifically ChatGPT, and Gemini. Building upon our prior study, where ChatGPT excelled in solving 10 Bayesian problems, we extend the scope by introducing four additional tasks to both ChatGPT and Gemini in order to compare performance. The results demonstrate ChatGPT and Gemini consistent accuracy in tackling all four reasoning problems presented. The obtained results suggest the potential of LLM like ChatGPT and Gemini for effectively handling Bayesian reasoning tasks relevant to science and engineering fields.

**Keywords:** Bayesian Reasoning, generative models, ChatGPT, Gemini.

## 1. Introduction

ChatGPT is a language model developed by OpenAI, built on the GPT (Generative Pre-trained Transformer) architecture. It is designed to generate human-like text based on the input it receives (ChatGPT, 2024). GPT models are a type of artificial intelligence that have been trained on a massive amount of numeric/text data to understand and generate coherent and contextually relevant responses. ChatGPT specifically refers to a version of the GPT model that has been fine-tuned to perform well in generating conversational responses.

Gemini (previously known as Bard) is the name of the tool using large language model (LLM) created by Google AI. This complex model is trained on a huge dataset of text and code, which allows to understand and process information, generate different creative text formats, and answer questions in an informative way (Gemini, 2024).

In our previous study (Krohling, 2023), we investigated the capacity of ChatGPT to solve 10 probabilistic reasoning problems. In this paper, we aim to extent our test suite by applying chatGPT and also Gemini to 4 additional Bayesian reasoning problems. Indeed, it is worth mentioning that people often fail in situations where the Bayes's formula is not known. Experiments conducted with children (Zhu & Gigerenzer, 2006) indicate that the Bayesian reasoning abilities depend on the representation of information.

In this paper, we show that the generative models can effectively solve Bayesian reasoning problems. Next, we present the 4 Bayesian reasoning tasks (Woike, Hertwig, and Gigerenzer, 2006) along with the respective solutions provided by ChatGPT and Gemini.

## 2. Bayesian reasoning problems solved by ChatGPT and Gemini

Next, we present the Bayesian inference tasks.

### Task 1. Cab task (Kahneman & Tversky, 1972)

Two cab companies operate in a given city, the Blue and the Green (according to the color of cab they run). Eighty-five percent of the cabs in the city are Blue, and the remaining 15% are Green. A cab was involved in a hit-and-run accident at night. A witness later identified the cab as a Green cab. The court tested the witness' ability to distinguish between Blue and Green cabs under nighttime visibility conditions. It found that the witness was able to identify each color correctly about 80% of the time, but confused it with the other color about 20% of the time. What do you think are the chances that the errant cab was indeed Green, as the witness claimed?

### ChatGPT answer

*To determine the chances that the errant cab was indeed Green, we can use Bayes' Theorem. Let's denote the following:*

*G: The event that the cab is Green.*

*B: The event that the cab is Blue.*

*W: The event that the witness identifies the cab as Green.*

*We want to find  $P(G|W)$ , the probability that the cab is Green given that the witness identified it as Green.*

*Bayes' Theorem states:*

$$P(G|W) = P(W|G) \cdot P(G) / P(W)$$

*Where:*

*$P(W|G)$  is the probability that the witness identifies the cab as Green given that it is Green.*

*$P(G)$  is the prior probability that the cab is Green.*

*$P(W)$  is the total probability of the witness identifying the cab as Green.*

*The probability that the witness identifies the cab as Green given that it is Green ( $P(W|G)$ ) is the accuracy of the witness, which is given as 80% or 0.8.*

*The prior probability that the cab is Green ( $P(G)$ ) is given as 15% or 0.15.*

*The total probability of the witness identifying the cab as Green ( $P(W)$ ) can be expressed as the sum of two mutually exclusive events: the witness correctly identifying a Green cab and the witness mistakenly identifying a Blue cab as Green.*

$$P(W) = P(W|G) \cdot P(G) + P(W|B) \cdot P(B)$$

*The probability that the witness identifies a Blue cab as Green ( $P(W|B)$ ) is the complement of the accuracy, which is  $1 - 0.8 = 0.2$ . The prior probability that the cab is Blue ( $P(B)$ ) is given as 85% or 0.85.*

Now, we can plug in the values into Bayes' Theorem:

$$P(G|W) = 0.8 \cdot 0.15 / (0.8 \cdot 0.15 + 0.2 \cdot 0.85)$$

$$P(G|W) \approx 0.4138$$

So, the probability that the errant cab was indeed Green, given the witness's claim, is approximately 0.4138 or 41.38%.

### **Gemini answer**

This situation is a classic application of Bayes' theorem to calculate the probability of the event (cab being green) given the evidence (witness identification). Here's how to solve it:

1. Define events:

Let  $G$  be the event that the cab involved in the accident was Green.

Let  $B$  be the event that the cab involved in the accident was Blue.

Let  $W$  be the event that the witness identified the cab as Green.

2. Given information:

$P(G) = 0.15$  (probability of a Green cab, as 15% are Green)

$P(B) = 1 - P(G) = 0.85$  (probability of a Blue cab, as 85% are Blue)

$P(W|G) = 0.8$  (probability of witness correctly identifying Green, 80% accuracy)

$P(W|B) = 0.2$  (probability of witness misidentifying Blue as Green, 20% error)

3. What we need to find:

$P(G|W)$  = probability of the cab being Green given the witness identified it as Green

4. Apply Bayes' theorem:

$$P(G|W) = (P(W|G) * P(G)) / (P(W|G) * P(G) + P(W|B) * P(B))$$

5. Substitute the values:

$$P(G|W) = (0.8 * 0.15) / (0.8 * 0.15 + 0.2 * 0.85)$$

$$P(G|W) = 0.12 / 0.29$$

$$P(G|W) \approx 0.4137 \text{ (approximately 41.37\%)}$$

### **Task 2. Weather forecast** (Dohmen et al., 2009)

Imagine you are on vacation in an area where the weather is mostly sunny and you ask yourself how tomorrow's weather will be. Suppose that, in the area you are in, on average 90 out of 100 days are sunny, while it rains on 10 out of 100 days. The weather forecast for tomorrow predicts rain. On average, the weather forecast is correct on 80 out of 100 days. What do you think is the probability, in percent, that it is going to rain tomorrow?

### **ChatGPT answer**

In this scenario, let's denote the following events:

*R: It will rain tomorrow.*

*S: It will be sunny tomorrow.*

*F: The weather forecast predicts rain.*

*We are interested in finding the probability that it will rain tomorrow given that the forecast predicts rain, i.e.,*

*$P(R|F)$ .*

*We can use Bayes' Theorem:*

$$P(R|F) = P(F|R) \cdot P(R) / P(F)$$

*Where:*

*$P(F|R)$  is the probability that the forecast predicts rain given that it will rain. This is given as the accuracy of the forecast, which is 80% or 0.8.*

*$P(R)$  is the prior probability that it will rain tomorrow, given as 10% or 0.1.*

*$P(F)$  is the total probability of the forecast predicting rain, which can happen either because it will rain and the forecast is correct or because it won't rain, but the forecast is incorrect.*

$$P(F) = P(F|R) \cdot P(R) + P(F|S) \cdot P(S)$$

*We know that  $P(F|S)$  is the probability that the forecast predicts rain given that it will be sunny. This is the complement of the accuracy, so it's  $1 - 0.8 = 0.2$ . The prior Now, let's plug in the values:*

$$P(F) = 0.8 \cdot 0.1 + 0.2 \cdot 0.9$$

$$P(F) = 0.08 + 0.18$$

$$P(F) = 0.26$$

*Now, we can use Bayes' Theorem to find  $P(R|F)$ :*

$$P(R|F) = 0.8 \cdot 0.1 / 0.26$$

$$P(R|F) = 0.08 / 0.26$$

*Now, let's perform the division:*

$$P(R|F) \approx 0.3077$$

*So, the probability that it will rain tomorrow, given that the forecast predicts rain, is approximately 30.77%.*

### **Gemini answer**

*The probability of rain tomorrow is approximately 30.8%. Here's the breakdown:*

*1. Define events:*

*Let  $R$  be the event that it rains tomorrow.*

Let  $F$  be the event that the forecast predicts rain.

2. Given information:

$P(\text{not } R) = 0.9$  (probability of a sunny day, as it rains 10% of the time)

$P(F|R) = 0.8$  (probability of an accurate forecast for rain)

$P(F|\text{not } R) = 0.2$  (probability of an inaccurate forecast for no rain)

3. What we need to find:

$P(R|F)$  = probability of rain tomorrow given the forecast predicts rain

4. Apply Bayes' theorem:

$$P(R|F) = (P(F|R) * P(R)) / (P(F|R) * P(R) + P(F|\text{not } R) * P(\text{not } R))$$

5. Substitute the values:

$$P(R|F) = (0.8 * 0.1) / (0.8 * 0.1 + 0.2 * 0.9)$$

$$P(R|F) = 0.08 / 0.26$$

$$P(R|F) \approx 0.3077$$
 (approximately 30.8%).

### **Task 3. Quality control** (Hoffrage et al., 2015)

The Skiwell Manufacturing Company gets material from two suppliers. Supplier A's materials make up for 30% of what is used, with supplier B providing the rest. Past records indicate that 15% of supplier A's materials are defective and 10% of B's materials are defective. Since it is impossible to tell which supplier the material came from once they are in the inventory, the manager wants to know: What is the probability that material comes from supplier A given that it has been identified as defective?

#### **ChatGPT answer**

Let's denote the following events:

$A$ : Material comes from supplier A.

$B$ : Material comes from supplier B.

$D$ : Material is defective.

We want to find  $P(A|D)$ , the probability that the material comes from supplier A given that it has been identified as defective.

We can use Bayes' Theorem for this:

$$P(A|D) = P(D|A) \cdot P(A) / P(D)$$

Where:

$P(D|A)$  is the probability that the material is defective given that it comes from supplier A, given as 15% or 0.15.

$P(A)$  is the prior probability that the material comes from supplier A, given as 30% or 0.3.

$P(D)$  is the total probability of the material being defective, which can happen either because it comes from supplier A and is defective or because it comes from supplier B and is defective.

$$P(D) = P(D|A) \cdot P(A) + P(D|B) \cdot P(B)$$

We know that  $P(D|B)$  is the probability that the material is defective given that it comes from supplier B, given as 10% or 0.1. The prior probability that the material comes from supplier B  $P(B)$  is the complement of  $P(A)$ , so it's  $1 - 0.3 = 0.7$ .

Now, let's plug in the values into the formula:

$$P(A|D) = 0.15 \cdot 0.3 / (0.15 \cdot 0.3 + 0.1 \cdot 0.7)$$

$$P(A|D) = .045 // (0.045 + 0.07)$$

$$P(A|D) = 0.045 / 0.115$$

Now, let's perform the division:

$$P(A|D) \approx 0.3913$$

So, the probability that the material comes from supplier A given that it has been identified as defective is approximately 39.13%.

### **Gemini answer**

Let's calculate the probability that the defective material comes from supplier A, given the information provided:

Define the probabilities:

$$p(A): \text{Probability of material coming from supplier A (30\%)} = 0.3$$

$$p(\text{defect}|A): \text{Probability of defective material from supplier A (15\%)} = 0.15$$

$$p(B): \text{Probability of material coming from supplier B (1 - p(A))} = 0.7$$

$$p(\text{defect}|B): \text{Probability of defective material from supplier B (10\%)} = 0.1$$

Calculate the overall probability of defective material:

$$p(\text{defect}): \text{Probability of any defective material} = p(\text{defect}|A) \cdot p(A) + p(\text{defect}|B) \cdot p(B) = 0.15 \cdot 0.3 + 0.1 \cdot 0.7 = 0.045 + 0.07 = 0.115$$

Apply Bayes' theorem:

$$p(A|\text{defect}): \text{Probability of material coming from supplier A given it is defective} = p(\text{defect}|A) \cdot p(A) / p(\text{defect}) = 0.15 \cdot 0.3 / 0.115 = 0.045 / 0.115 \approx 0.3913$$

Therefore, the probability that the defective material comes from supplier A is approximately 39.13%.

#### **Task 4. Caries detection** (Nadanovsky et al., 2018)

Imagine a population survey for interproximal caries detection using bite-wing radiographs. The information below refers to asymptomatic adults who took part in this survey. The probability that 1 of these adults has interproximal caries requiring restorative treatment, confirmed by tooth separation and direct visual and tactile examination, was 8%. The probability of having a positive bite-wing radiograph among adults who have a confirmed interproximal caries is 90%. The probability of having a positive bite-wing radiograph among adults without interproximal caries is 8%. Question: Imagine an adult who had a positive bite-wing radiograph in the survey. What is the probability that he actually has interproximal caries?

#### **ChatGPT answer**

*Let's denote the following events:*

*C: The event that an adult has interproximal caries requiring restorative treatment.*

*+: The event that an adult has a positive bite-wing radiograph.*

*We want to find  $P(C|+)$ , the probability that an adult has interproximal caries given that they have a positive bite-wing radiograph.*

*We can use Bayes' Theorem for this:*

$$P(C|+) = P(+|C) \cdot P(C) / P(+)$$

*Where:*

*$P(+|C)$  is the probability of having a positive bite-wing radiograph given that an adult has interproximal caries. This is given as 90% or 0.9.*

*$P(C)$  is the prior probability that an adult has interproximal caries, given as 8% or 0.08.*

*$P(+)$  is the total probability of having a positive bite-wing radiograph, which can happen either because an adult has interproximal caries and the radiograph is positive or because an adult does not have interproximal caries, but the radiograph is positive.*

$$P(+)=P(+|C) \cdot P(C)+P(+|\sim C) \cdot P(\sim C)$$

*We know that  $P(+|\sim C)$  is the probability of having a positive bite-wing radiograph given that an adult does not have interproximal caries, given as 8% or 0.08. The prior probability that an adult does not have interproximal caries  $P(\sim C)$  is the complement of  $P(C)$ , so it's  $1-0.08=0.92$ .*

*Now, let's plug in the values into the formula:*

$$P(C|+)=0.9 \cdot 0.08 / (0.9 \cdot 0.08 + 0.08 \cdot 0.92)$$

$$P(C|+)= 0.072 / (0.072 + 0.0736)$$

$$P(C|+)=.072 / 0.1456$$

*Now, let's perform the division:*

$$P(C|+) \approx 0.4945$$

So, the probability that an adult actually has interproximal caries, given that they have a positive bite-wing radiograph, is approximately 49.45%.

### **Gemini answer**

Let's calculate the probability that an adult with a positive bite-wing radiograph actually has interproximal caries.

We can use Bayes' theorem to approach this problem. This theorem helps determine the probability of an event (having caries) given that another event (positive bite-wing radiograph) has already occurred.

Here's what we know from the information provided:

Probability of having interproximal caries ( $P(\text{caries})$ ) = 8% = 0.08

Probability of a positive bite-wing radiograph among with caries ( $P(\text{positive} | \text{caries})$ ) = 90% = 0.9

Probability of a positive bite-wing radiograph among without caries ( $P(\text{positive} | \text{no caries})$ ) = 8% = 0.08

We need to find the probability of caries given a positive test ( $P(\text{caries} | \text{positive})$ ).

Applying Bayes' theorem:

$$P(\text{caries} | \text{positive}) = (P(\text{positive} | \text{caries}) * P(\text{caries})) / (P(\text{positive} | \text{caries}) * P(\text{caries}) + P(\text{positive} | \text{no caries}) * (1 - P(\text{caries})))$$

Calculation:

$$P(\text{caries} | \text{positive}) = (0.9 * 0.08) / ((0.9 * 0.08) + (0.08 * (1 - 0.08))) = 0.4945$$

Therefore, the probability that an adult with a positive bite-wing radiograph actually has interproximal caries is approximately 49.45%.

### **3. Concluding remarks**

This paper show ChatGPT and Gemini ability to solve all four Bayesian reasoning tasks presented. Since both tools can solve Bayesian reasoning problems given text-based data, their use, as suggested by Krohling (2023), may prove beneficial for enhancing the learning process of pupils/undergraduate students in engineering and other scientific disciplines taking probability and statistics courses.

### **References**

Zhu, L. & Gigerenzer, G. Children can solve Bayesian problems: the role of representation in mental computation, *Cognition*, January 2006.

Krohling, R. A. Bayesian artificial brain with ChatGPT, 2023, available at [arxiv.org/abs/2308.14732](https://arxiv.org/abs/2308.14732)

ChatGPT, openAI official website, available from <https://platform.openai.com/apps>, May 2024.

Gemini, Google AI Generative Chat Tools , available from <https://gemini.google.com/app>, May 2024



Wolke, J. K., Hertwig, R., & Gigerenzer, G. Heterogeneity of rules in Bayesian reasoning: A toolbox analysis, *Cognitive Psychology*, V. 143, June 2023, 101564

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430–454.

Dohmen, T., Falk, A., Huffman, D., Felix, M., & Sunde, U. (2009). The Non-Use of Bayes Rule: Representative Evidence on Bounded Rationality. Technical Report, Maastricht University, Research Centre for Education and the Labour Market, ROA-RM-2009/1, February 2009.

Hoffrage, U., Hafenbrädl, S., & Bouquet, C. (2015). Natural frequencies facilitate diagnostic inferences of managers. *Frontiers in Psychology*, 6, 642.

Nadanovsky, P., dos Santos, A. P. P., Lira-Junior, R., & de Oliveira, B. H. (2018). Clinical accuracy data presented as natural frequencies improve dentists' caries diagnostic inference: Evidence from a randomized controlled trial. *The Journal of the American Dental Association*, 149(1), 18–24.