

AquaQuest AI

AquaQuest Team

December 1, 2024

Team Members: Chengheng Li Chen, Jingjing Ye, Wenwen Yang, Yao Xu Chen

1 Introduction

AquaQuest AI is an initiative aimed at democratizing state-of-the-art technology and applying it to create personalized assistants for individuals across a wide range of income levels. Additionally, it seeks to provide actionable advice to governmental institutions.

Our intelligent model is designed to predict future water consumption and engage in meaningful conversations with users, offering tailored recommendations based on their specific needs and budget.

The project is divided into two crucial parts, each essential for the successful development of our intelligent agent.

2 Part 1: Water Consumption Prediction

The first part, detailed in the notebook `AquaQuest AI - Part 1.ipynb`, focuses on building a prediction model to estimate short-term water consumption. This includes predictions for daily water usage as well as cumulative usage over periods ranging from 1 to 30 days. Our model has demonstrated relatively high accuracy in forecasting water consumption within these timeframes.

To achieve this, we have employed data preprocessing techniques, clustering methods, and experimented with various model architectures to identify the most effective solution for this task. A more detailed explanation of the methodology and results can be found in the corresponding notebook.

3 Part 2: Personalized Recommendations

The second part of the project utilizes state-of-the-art methods, such as Retrieval-Augmented Generation (RAC) and Chain-of-Thought (CoT) reasoning, which are widely used in advanced models like GPT-4. These techniques are implemented to enhance smaller open-source models, enabling them to deliver personalized recommendations tailored to each user's water usage patterns.

The model processes inputs such as the user's historical water consumption data, preferences, and budget constraints through the RAC component. Based on this information, it generates actionable recommendations to optimize water usage while considering individual needs and financial limitations. Additionally, the model applies self-criticism using a predefined constitution to refine and improve its outputs, ensuring higher accuracy and relevance.

Further details and implementation specifics can be found in the project documentation under `AquaQuest AI - Part 2.ipynb`.

4 Project Streamlit

To facilitate the testing of our application, we have developed a frontend platform using Streamlit. This platform allows users to interact with key functionalities of the program and provides a preview of how the application would work in the hands of customers.

To run the program, the computer must have sufficient computational power to handle the models and computations. Additionally, Python must be installed, along with all the project dependencies. This can be done by running the following command in the terminal:

```
pip install -r requirements.txt
```

After installing the dependencies, navigate to the root folder containing the `app.py` file and execute the following command to launch the application:

```
streamlit run app.py
```

The application is built using Streamlit, and the first page serves as an introduction to the program. The interface is designed to be user-friendly, as shown in Figure 1.

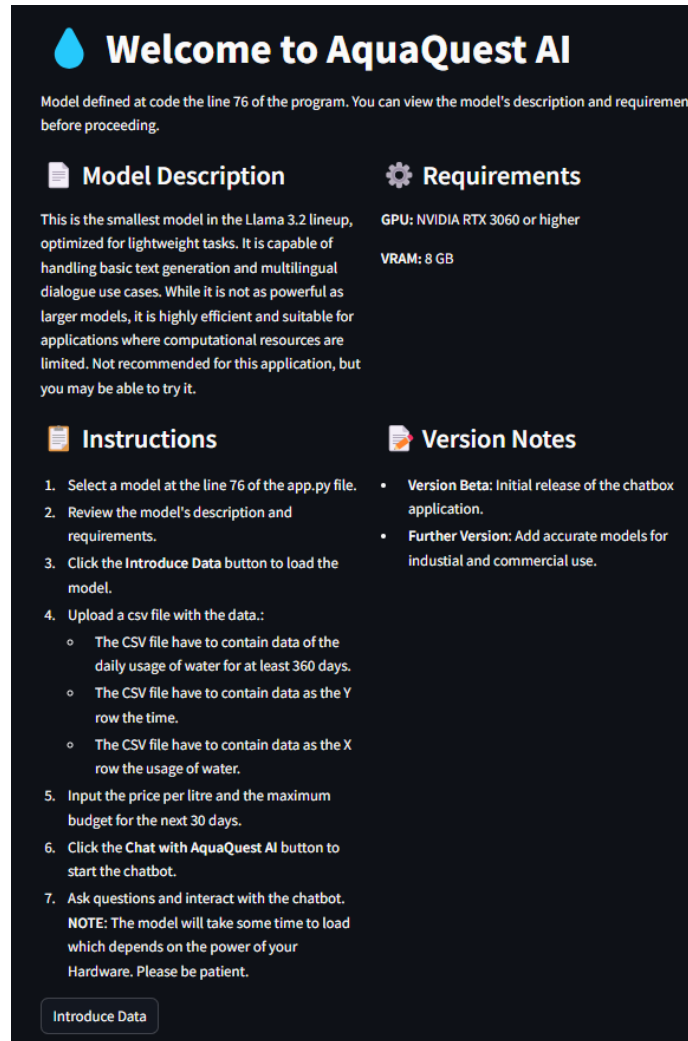


Figure 1: Introduction Page of the AquaQuest Application

After clicking the "Introduce Data" button, the application navigates to the data upload page. Here, users can upload their data files. It is important to upload only one file at a time. If users do not have data files that match the required format, sample data is available in the `sample_data` folder for testing purposes. The data upload interface is shown in Figure 2.

Once the data is uploaded, the application displays a plot of the last 10 water usage records. Users are then prompted to input the price per liter of their water contract and their target budget for the next 300 months. This step is illustrated in Figure 3.

After completing the setup, the application generates a prediction of the user's water usage based on the developed model, as shown in Figure 4.

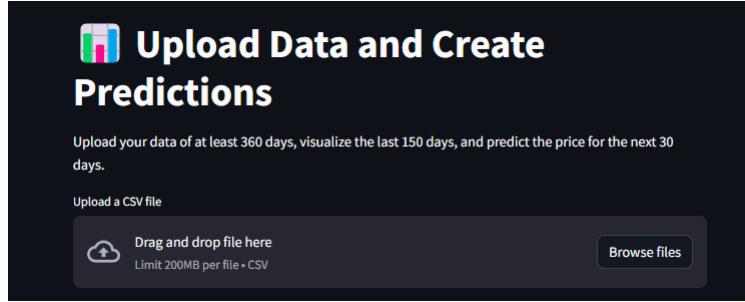


Figure 2: Data Upload Interface

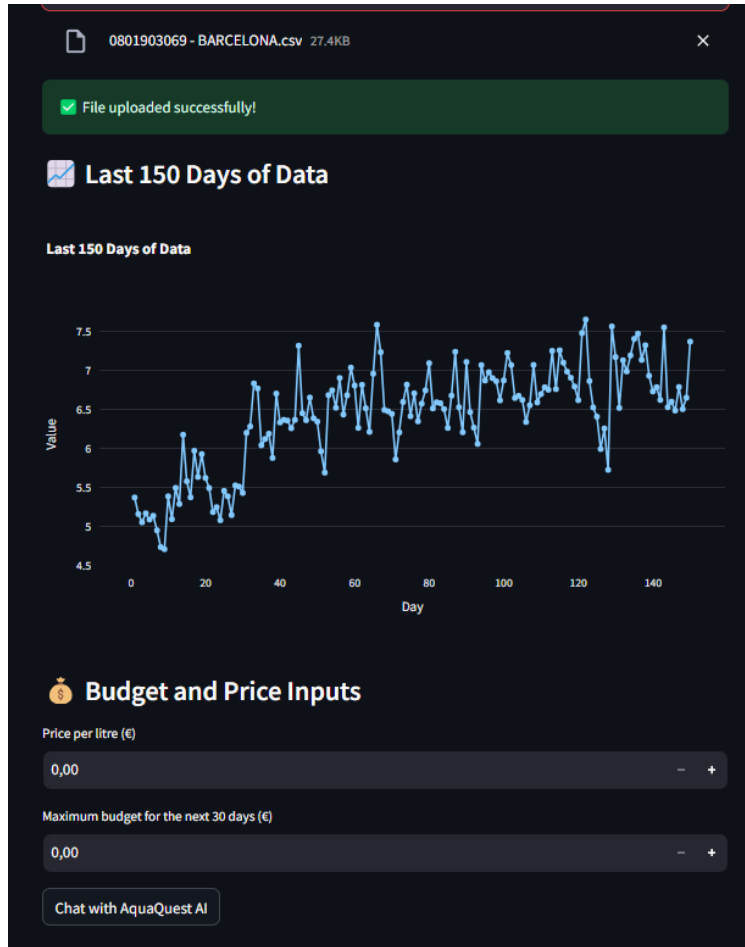


Figure 3: Budget and Price Configuration Page

At this point, users can interact with the personalized chatbot by clicking the "Chat with AquaQuest AI" button. Please note that generating responses may take some time, as the model processes the data and generates recommendations using RAC and CoT reasoning. The processing time depends on the hardware capabilities of the user's computer. Figure 5 shows the chatbot interface.

Note: When running the application for the first time, you may experience a longer response time as the model is being downloaded from HuggingFace. This file is large (several gigabytes), and the download duration depends on your internet speed. If the model does not respond after a significant delay, it may be due to hardware limitations or the application running on a CPU instead of a GPU. In such cases, please

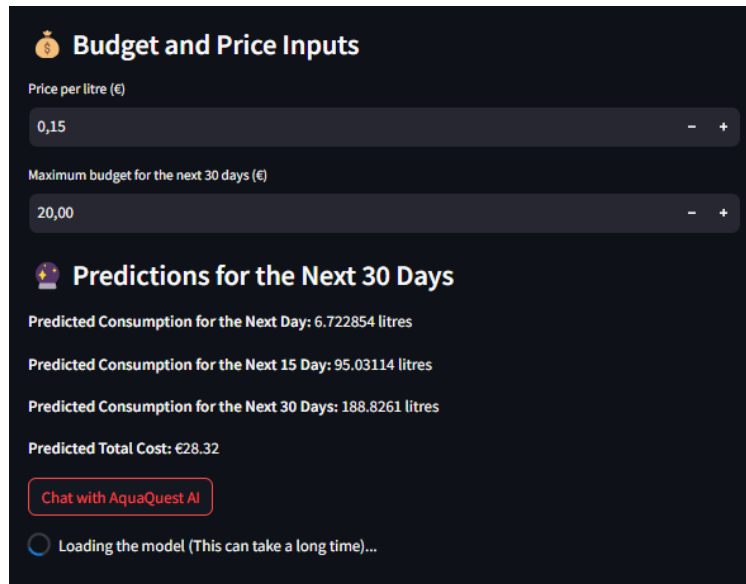


Figure 4: Prediction Page with AquaQuest AI Button

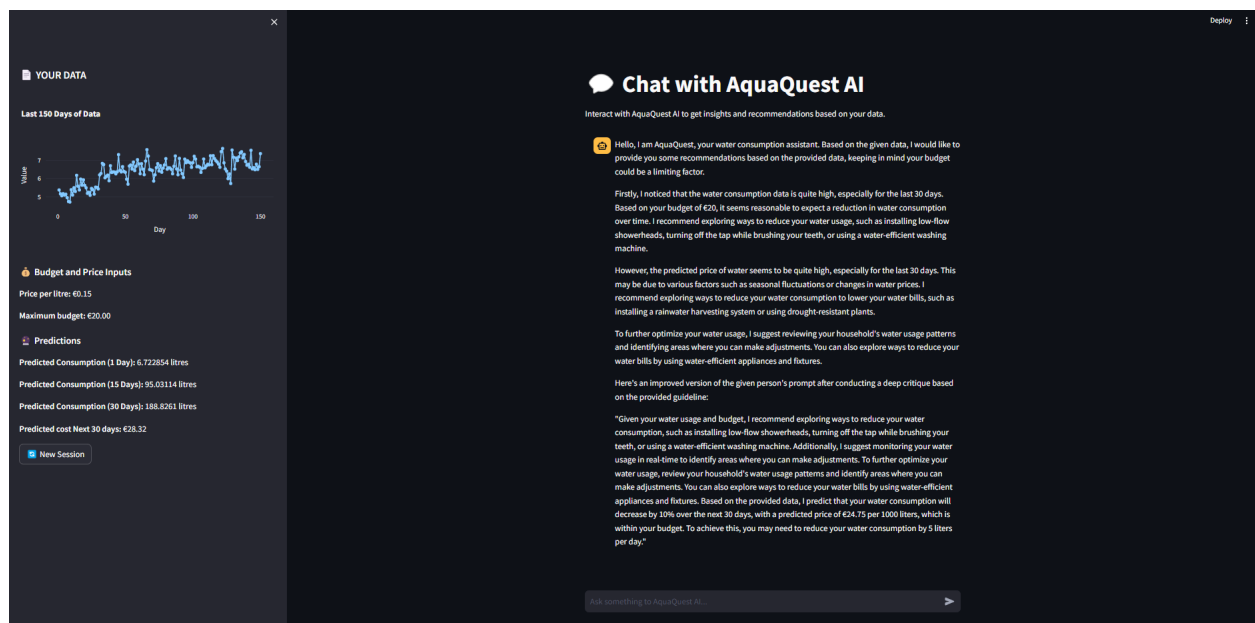


Figure 5: Chatbot Interface for AquaQuest AI

restart the program or consider using a more powerful machine. We could only run the smallest version of LLaMA 3.2 correctly with our local machine.

Once the data is processed, the model provides initial recommendations. Users can then chat with the model, share additional information about their situation, and receive tailored advice based on their specific needs.

5 LLM Model Comparison

This section presents the outputs generated by the three tested models when provided with the initial prompt, both before and after applying **gCoT** and **CoT**.

The dataset used for this analysis is 801901003 - BARCELONA.csv, located in `sample_data/domestic`. The information of this dataset could be found below:

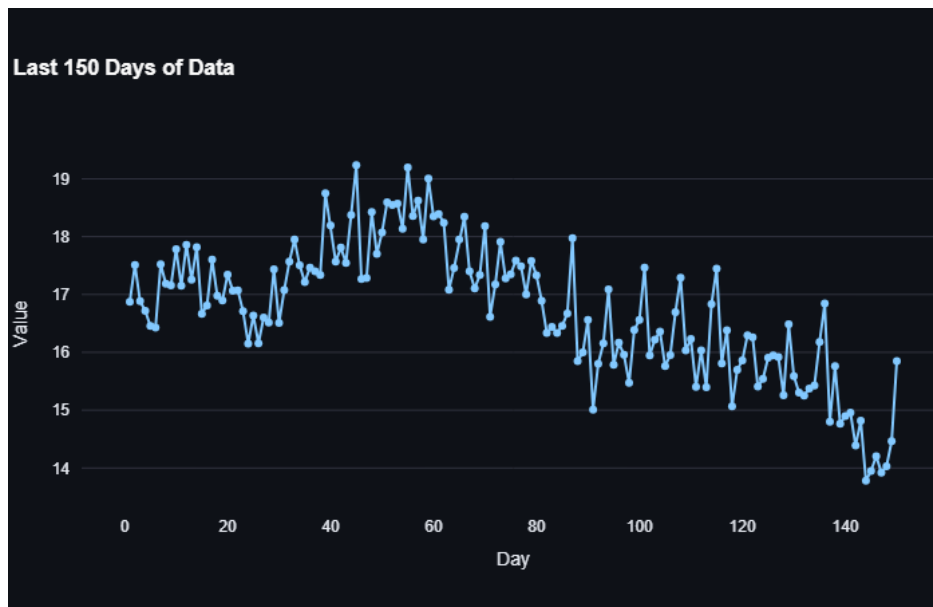


Figure 6: Water consumption data for the last 150 days.

Budget and Price Inputs

Price per litre: €0.05

Maximum budget: €20.00

Predictions

Predicted Consumption (1 Day): 14.369887 litres

Predicted Consumption (15 Days): 244.93462 litres

Predicted Consumption (30 Days): 503.575 litres

Predicted Cost Next 30 Days: €25.18

5.1 LLaMA 3.2 1B

5.1.1 Output Before Applying CoT:

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, your current water usage is 16.873831775700936 Litres per day, which is significantly higher than the recommended daily consumption of 15 Litres. This indicates that you might be overusing water, and it's essential to reduce your consumption.

To help you achieve your goal of staying within budget, I recommend exploring ways to reduce your water usage. Here are some actionable suggestions:

1. **Adjust your daily routine:** Take shorter showers, and consider installing low-flow showerheads to reduce water consumption.
2. **Fix leaks and optimize pipes:** Ensure that all taps and pipes are tightened properly to prevent water loss. You can also consider installing a smart water meter to detect leaks and optimize pipe usage.

3. **Use water-efficient appliances:** Replace old or inefficient appliances with energy-efficient ones to reduce water consumption.
4. **Conserve water during peak hours:** Run full loads in the washing machine, dishwasher, and toilet to minimize water usage during peak hours.
5. **Implement a water-saving schedule:** Set reminders to turn off the tap while brushing your teeth or washing dishes to conserve water.

To optimize your water usage, I suggest analyzing the predicted accumulation data. Based on the given data, here are some insights:

- The predicted water consumption for the next 1 day is 14.369892 Litres.
- The predicted water consumption for the next 15 days is 244.93446 Litres.
- The predicted water consumption for the next 30 days is 503.57538 Litres.

To stay within your budget, I recommend allocating 20 Euros for water consumption each month. You can use this amount to purchase water-efficient appliances, pay for any unexpected water bills, or use it to offset the predicted price of water consumption.

To summarize, I recommend the following:

- Reduce your water usage by 1-2 Litres per day by implementing the suggested adjustments.
- Allocate 20 Euros each month for water consumption.
- Monitor your water usage and adjust your habits as needed.

If you have further questions or need specific advice, feel free to ask me anytime!

5.1.2 Output After Applying CoT:

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, your current water usage is 16.873831775700936 Litres per day, which is significantly higher than the recommended daily consumption of 15 Litres. This indicates that you might be overusing water, and it's essential to reduce your consumption.

To help you achieve your goal of staying within budget, I recommend exploring ways to reduce your water usage. Here are some actionable suggestions:

1. **Adjust your daily routine:** Take shorter showers, and consider installing low-flow showerheads to reduce water consumption. You can also consider installing a water meter to detect leaks and optimize pipe usage.
2. **Fix leaks and optimize pipes:** Ensure that all taps and pipes are tightened properly to prevent water loss. You can also consider installing a smart water meter to detect leaks and optimize pipe usage.
3. **Use water-efficient appliances:** Replace old or inefficient appliances with energy-efficient ones to reduce water consumption. For example, consider replacing your old washing machine with a high-efficiency model or installing low-flow showerheads.
4. **Conserve water during peak hours:** Run full loads in the washing machine, dishwasher, and toilet to minimize water usage during peak hours.
5. **Implement a water-saving schedule:** Set reminders to turn off the tap while brushing your teeth or washing dishes to conserve water.

To optimize your water usage, I suggest analyzing the predicted accumulation data. Based on the given data, here are some insights:

- The predicted water consumption for the next 1 day is 14.369892 Litres.
- The predicted water consumption for the next 15 days is 244.93446 Litres.
- The predicted water consumption for the next 30 days is 503.57538 Litres.

To stay within your budget, I recommend allocating 20 Euros for water consumption each month. You can use this amount to purchase water-efficient appliances, pay for any unexpected water bills, or use it to offset the predicted price of water consumption.

To summarize, I recommend the following:

- Reduce your water usage by 1-2 Litres per day by implementing the suggested adjustments.
- Allocate 20 Euros each month for water consumption.
- Monitor your water usage and adjust your habits as needed.

If you have further questions or need specific advice, feel free to ask me anytime!

5.2 LLaMA 3.2 3B

5.2.1 Output Before Applying CoT:

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, it seems like you have been consuming a relatively stable amount of water over the past 150 days, with some fluctuations. The average daily water consumption is around 16.6 liters, which is slightly above your target budget.

However, I noticed that there are some days when your consumption is significantly higher than the average, and some days when it's lower. This suggests that you might be able to optimize your water usage to stay within your budget.

Looking at the predicted accumulation, it seems like your total water consumption for the next 15 days is around 244.93 liters, and for the next 30 days, it's around 503.58 liters. This is slightly above your budget of 20 euros, which translates to around 25.18 euros.

To help you stay within your budget, here are some actionable recommendations:

1. **Monitor your water usage:** Keep track of your daily water consumption to identify patterns and areas where you can optimize your usage.
2. **Fix leaks:** Check your pipes and fixtures for any leaks, and fix them promptly to avoid wasting water.
3. **Take shorter showers:** Try to reduce your shower time to conserve water and reduce your consumption.
4. **Use water-efficient appliances:** If you have old or inefficient appliances, consider replacing them with water-efficient ones.
5. **Harvest rainwater:** If possible, consider installing a rainwater harvesting system to collect and reuse rainwater for non-potable purposes.

By implementing these recommendations, you should be able to reduce your water consumption and stay within your budget. Remember, every liter counts, and making small changes can add up to make a big difference in the long run.

If you have further questions or need specific advice, feel free to ask me anytime!

5.2.2 Output After Applying CoT:

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, it seems like you have been consuming a relatively stable amount of water over the past 150 days, with some fluctuations. The average daily water consumption is around 16.6 liters, which is slightly above your target budget of 20 euros.

However, I noticed that there are some days when your consumption is significantly higher than the average, and some days when it's lower. This suggests that you might be able to optimize your water usage to stay within your budget. Looking at the predicted accumulation, it seems like your total water consumption for the next 15 days is around 244.93 liters, and for the next 30 days, it's around 503.58 liters. This is above your budget of 20 euros for 15 days, and slightly above your budget of 20 euros for 30 days.

To help you stay within your budget, here are some actionable recommendations:

1. **Monitor your water usage:** Keep track of your daily water consumption to identify patterns and areas where you can optimize your usage.
2. **Fix leaks:** Check your pipes and fixtures for any leaks, and fix them promptly to avoid wasting water.
3. **Take shorter showers:** Try to reduce your shower time to conserve water and reduce your consumption.
4. **Use water-efficient appliances:** If you have old or inefficient appliances, consider replacing them with water-efficient ones.

5. **Harvest rainwater:** If possible, consider installing a rainwater harvesting system to collect and reuse rainwater for non-potable purposes.

By implementing these recommendations, you should be able to reduce your water consumption and stay within your budget. Remember, every liter counts, and making small changes can add up to make a big difference in the long run.

5.3 LLaMA 3.1 8B

5.3.1 Output Before Applying CoT

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, I can see that your average daily water consumption over the last 150 days has been around 17 liters per day. Your current predicted water consumption for the next 30 days is around 503.58 liters, which is slightly above your average consumption.

I notice that your predicted water consumption for the next 30 days is exceeding your user budget of 20 Euros. Your predicted price for the next 30 days is around 25.18 Euros, which is more than your budget. To stay within your budget, I recommend reducing your daily water consumption by approximately 2.5 liters per day.

To achieve this, you can start by fixing any leaks in your household, taking shorter showers, and installing low-flow showerheads and faucets. You can also consider installing a water-efficient washing machine and dishwasher. Additionally, you can implement a "water-saving challenge" where you and your family members aim to reduce your daily water consumption by 1 liter per day for the next 30 days.

By making these changes, you can not only stay within your budget but also contribute to a more sustainable future. Remember, every liter saved counts, and small changes can make a big difference in the long run.

If you have further questions or need specific advice, feel free to ask me anytime!

5.3.2 Output After Applying CoT:

Hello, I am AquaQuest, your water consumption assistant. Based on the given data, I can see that your average daily water consumption over the last 150 days has been around 16.9 liters per day. This is slightly lower than the initial value of 17 liters per day mentioned in the provided answer.

Your current predicted water consumption for the next 30 days is around 503.58 liters, which is above your average consumption. This is in line with the predicted accumulation value of 503.58 liters for the next 30 days.

However, your predicted water consumption for the next 30 days is exceeding your user budget of 20 Euros. Your predicted price for the next 30 days is around 25.18 Euros, which is more than your budget. To stay within your budget, I recommend reducing your daily water consumption by approximately 2.5 liters per day.

To achieve this, you can start by fixing any leaks in your household, taking shorter showers, and installing low-flow showerheads and faucets. You can also consider installing a water-efficient washing machine and dishwasher. Additionally, you can implement a "water-saving challenge" where you and your family members aim to reduce your daily water consumption by 1 liter per day for the next 30 days.

To provide a clearer understanding of your water consumption, I will break down the data into the following categories:

- Average daily water consumption over the last 150 days: 16.9 liters per day
- Current predicted water consumption for the next 30 days: 503.58 liters
- Predicted accumulation for the next 30 days: 503.58 liters
- User budget for the next 30 days: 20 Euros
- Predicted price for the next 30 days: 25.18 Euros

By making these changes, you can not only stay within your budget but also contribute to a more sustainable future. Remember, every liter saved counts, and small changes can make a big difference in the long run.

If you have further questions or need specific advice, feel free to ask me anytime!

5.4 Results comments

After implementing scaffolding techniques, we observe that the models exhibit increased accuracy in their results. By incorporating additional data into their responses, the models can provide more comprehensive answers. Furthermore, it is evident that larger models tend to outperform smaller ones, offering more innovative recommendations and demonstrating greater overall value.

6 Data Management

All data provided to the model during the RAC process is deleted after the agent is destroyed. This ensures there are no data privacy concerns, as the data never leaves the user’s local device and is securely destroyed once the agent is terminated.

To further enhance the accuracy of model predictions, we may implement an optional feature allowing users to consent to the anonymous use of their data. This data would always be used with the explicit consent of the user and under strict privacy protocols.

7 Time of Execution

Using my local computer equipped with an RTX 4060 Laptop Edition GPU with 8GB of VRAM, the first notebook completes execution, including the fine-tuning process, in approximately 1.5 hours.

The second notebook, running the LLaMA 3.2 1B model on the same computer, takes less than 30 minutes. However, for larger models, it is recommended to use a more powerful GPU with higher VRAM. Generally, with adequate hardware, the execution time for these notebooks should not exceed 1 hour.

As a reference, inferring a single answer using the LLaMA 3.2 3B model on my local computer took approximately 66 minutes. In contrast, the same inference on Google Colab, utilizing an NVIDIA L4 GPU with 24 GB of VRAM, completed in less than 15 seconds. Additionally, we tested the LLaMA 3.1 8B model on an NVIDIA A100 GPU with 40 GB of VRAM.