

SIT305 2.1p Thinking about Llama 2

Llama 2 is a free, large language model designed for chat and natural language generation tasks, such as text generation, coding assistance, and more. It focuses on improving the performance of smaller models, ranging from 7B to 70B parameters. Compared to the first generation, its context length has been increased, it has changed to grouped-query attention (GQA), it can be used by any organization, and it can better understand human language to complete tasks, which means it is more powerful and easier to use (Heller 2023).

Here are five scenarios where it can be used:

1. Natural language unit conversion

Llama2 can parse complex instructions entered by users (such as "convert 5 feet 3 inches to meters"), automatically identify numerical values and unit categories, and directly trigger conversion logic. For example, in this task, we have realized the unit conversion by selecting units and entering quantities, but the application of Llama2 can avoid the tedious operation of manually selecting units, which is suitable for scenarios that require real-time voice interaction.

2. Content generation and summary

Applications can use Llama 2 to generate or summarize content, helping users take notes on conference content or quickly understand the main points of a large amount of text. For example, when users need to find a large amount of literature for scientific research, Llama 2 can help users find relevant literature and summarize the main content, saving a lot of time.

3. Multi-language support

Llama 2 can convert languages in real time, such as querying in different languages, converting received information into the user's native language in real time, and quickly translating relevant text materials. This feature can also support the use of Llama 2 applications to be more global and available to people around the world.

4. Personalized scenario recommendation

Llama 2 can browse the user's history and recommend relevant content that the user may be interested in. If the user often uses unit conversion and changes the geographical location, it will analyze whether the user often travels abroad for work or travel and needs navigation and unit conversion in different countries. Then it will provide recommended conversion units and analyze the user's current situation, providing weather forecast and other help. People may be surprised that the app provides them with relevant information before they even mention it.

5. Intelligent error diagnosis and feedback

When a user uses a certain function but provides incorrect information, Llama 2 can analyze the user's thoughts and propose relevant solutions or explain relevant knowledge. We can still use the example of unit conversion. For example, the user wants to convert 100 degrees Celsius into kilometers, but the temperature unit obviously cannot be converted into a length unit. Then Llama 2 will tell the user that this is a wrong question, explain to the user that the two units cannot be converted into each other, and analyze the user's thoughts and recommend measures for the user: Maybe you want to convert 100 meters into kilometers?

In general, Llama 2 has a wide range of application scenarios, and it can significantly improve the functionality of applications and promote their globalization. Applying Llama 2 will make our future lives better.

Reference:

Martin Heller (12 September 2023)'[Understanding Llama 2 and Code Llama](#)', InfoWorld, accessed March 25, 2025.