Groq, Ollama Setup

1 Ollama allows you to run the LLM models locally

2 Groq allows to call LLMs in the cloud without any charges.

3 The inference (time it takes LLM to generate response) in Groq is very fast