

LLM Apps

LLM App Types by Levels of Reasoning

LLM App Types by Levels of Reasoning

1. Single call.
2. Chain of multiple calls.
3. Router with no while loops.
4. Router with while loops (sometimes called agent).
5. Agent (autonomous).

Single Call LLM App

- One single call to the Foundation LLM model to provide the context and get the output.
- Example:
 - Provide the prompt template in the app.
 - The user completes the prompt.
 - The app makes a call to the LLM with the user prompt.
 - The LLM provides the output to the prompt.

Chain of multiple calls

- Chain multiple LLM calls and API calls. The steps are known.
- Example:
 - The user enter the search keyword.
 - The app calls the search engine API and gets the result.
 - The app builds the prompt template with the search result.
 - The user completes the prompt.
 - The app makes a call to the LLM with the user prompt.
 - The LLM provides the output to the prompt.

Router WITH NO while loops.

- The LLM decides the steps to take among several options (prompts to use, LLMs to use, tools to use, etc). With no while loops.
- Example:
 - The app can decide among several prompts (one for Math operations, another for translation, etc).
 - The app can decide among several LLMs to use (Claude when larger context window is needed, ChatGPT4 when better reasoning is needed, etc).
 - The app can decide among several tools to use (online search, calculator, online app, etc).
 - The process is sequential. DOES NOT HAVE while loops (“While the goal is not achieve, keep doing this”).

Router WITH while loops.

- Sometimes called state machine. Sometimes called agent.
- The LLM decides the steps to take among several options (prompts to use, LLMs to use, tools to use, etc). With while loops.
- Example:
 - The app can decide among several prompts (one for Math operations, another for translation, etc).
 - The app can decide among several LLMs to use (Claude when larger context window is needed, ChatGPT4 when better reasoning is needed, etc).
 - The app can decide among several tools to use (online search, calculator, online app, etc).
 - The process is sequential. DOES HAVE while loops (“While the goal is not achieve, keep doing this”).
 - On each loop, the LLM frequently does 3 steps: plan, take action, validate.

Agent (autonomous)

- The LLM decides the steps to take. The agent decides when to plan, act or validate.
- Popular example: AutoGPT.