

Introduction to LLM Input / Output

<u>Instructor</u>

Dipanjan Sarkar

Head of Community & Principal Al Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author



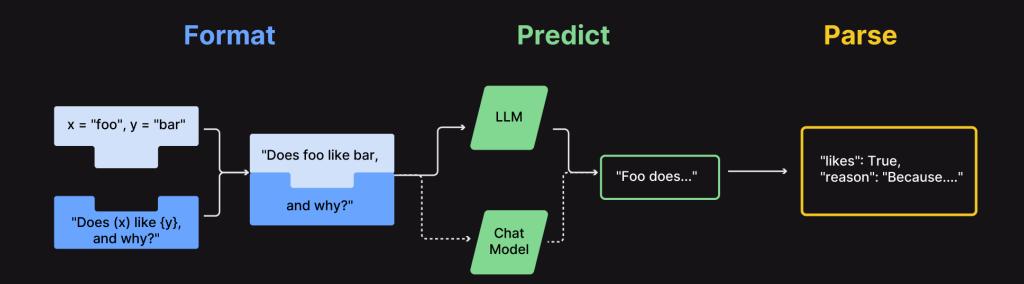
Outline

- LLM Input / Output Workflow
- LLMs and Chat Models
- Prompting with Prompt Templates
- Format LLM Response with Output Parsers
- Other Advanced Operations



LLM Input/Output with LangChain

Model I/O

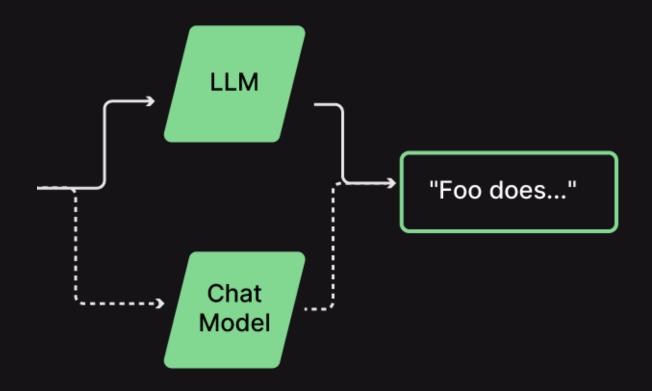




LLMs and Chat Models

- LangChain's LLM components has two different types of models
 - LLMs for general prompting
 - Chat Models for conversational prompting
- We can interface with these LLMs using Prompt Templates in LangChain

Predict

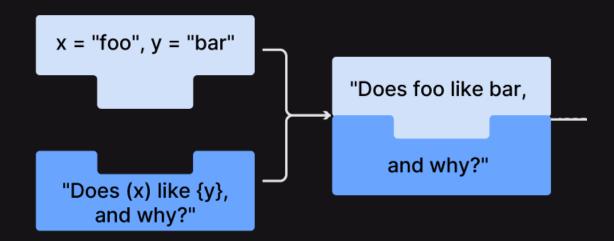




Prompting with Prompt Templates

- LangChain allows users to create
 Prompt Templates, which are sets
 of instructions and input variables.
- These input variables are usually populated at runtime in the Prompt Template and sent to the LLM

Format





Format LLM Response with Output Parsers

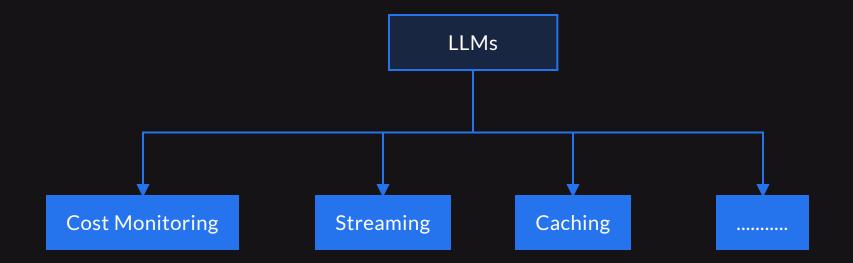
- LangChain allows to convert the raw LLM response into a more consumable format by using Output Parsers.
- There exists a variety of parsers including:
 - String parser
 - CSV parser
 - Pydantic parser
 - JSON parser







Other Advanced Operations



- LangChain offers additional capabilities when interfacing with LLMs.
 - Token cost monitoring
 - Streaming responses
 - Caching requests



Thank You

