

Introduction To LangServe

Instructor

Dipanjan Sarkar

Head of Community & Principal AI Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author



Course Outline

- Understanding LangServe
- Hosted LangServe

Understanding LangServe



```
#!/usr/bin/env python
from fastapi import FastAPI
from langchain.chat_models import ChatOpenAI
from langserve import add_routes

app = FastAPI(
    title="LangChain Server",
    version="1.0",
    description="A simple api server using Langchain's Runnable interfaces",
)

add_routes(
    app,
    ChatOpenAI(),
    path="/openai",
)

if __name__ == "__main__":
    import uvicorn
    uvicorn.run(app, host="localhost", port=8000)
```

- LangServe helps developers deploy LangChain runnable and chains as a REST API
- Built with FastAPI, it gives API documentation and a playground for the LangChain apps.
- LangServe enables LLMs as a service.

Understanding LangServe

```
#!/usr/bin/env python
from fastapi import FastAPI
from langchain.chat_models import ChatOpenAI
from langserve import add_routes

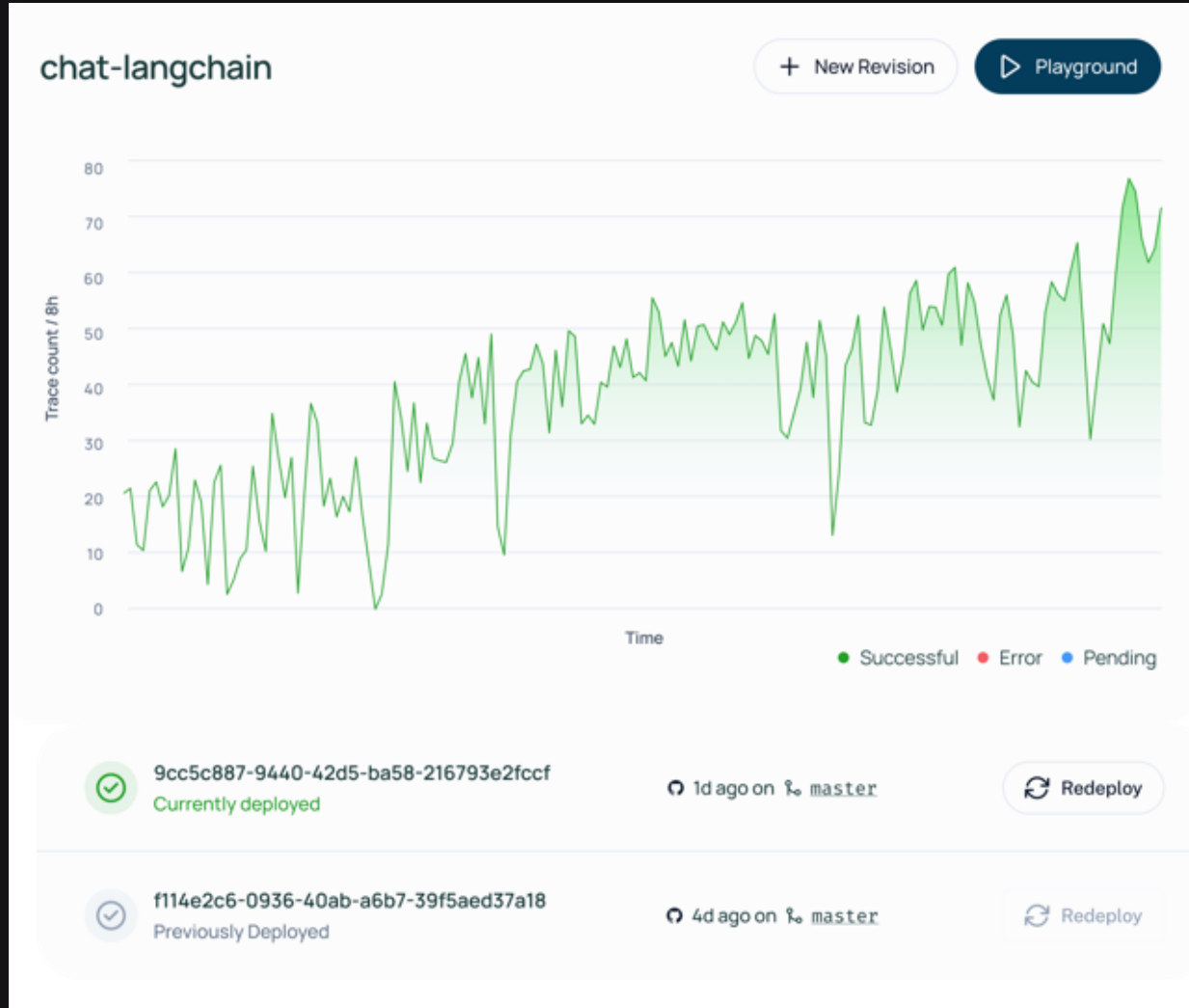
app = FastAPI(
    title="LangChain Server",
    version="1.0",
    description="A simple api server using Langchain's Runnable interfaces",
)

add_routes(
    app,
    ChatOpenAI(),
    path="/openai",
)

if __name__ == "__main__":
    import uvicorn
    uvicorn.run(app, host="localhost", port=8000)
```

- LangServe is integrated with FastAPI and uses Pydantic for data validation.
- LangServe also provides a client that can be used to call into LLM chains deployed on a server as an API

Hosted LangServe



- A hosted version of LangServe is available for one-click deployment of LangChain applications.
- Connect to the GitHub repository and deploy the app in a single click, with LangSmith tracing and observability automatically built in on every request.

Thank You
