

Shadow AI

Overview:

Shadow AI is an agent which is powered by the LLM (Large Language Model) in our case LLAMA 2 (specifically, llama-2-7b-chat model). The Agent is designed to help his boss in an efficient way for

- Drafting Emails
- Searching a Document and providing relevant information and tasks which are asked.
- Generating the presentation content
- Boss Mode: Combining all the crucial information and providing it to the boss. (Jus like Jarvis)

The agent specifically provides a combination of all the stuff listed above and specifically lists them. The agent is also able generate the relevant images with the help of stable-diffusion

Key Technology:

- AWS Sagemaker (Our agent extensively uses the AWS sagemaker)
 - For running the llama-2-7b-chat model
 - For running the stable-diffusion-v2-1-base
- AWS Lamda
 - For communication with our models
- AWS API Gateway
 - For providing the Api access
- AWS S3
 - For storing the objects and the frontend
- Python 3

Details:

We are using AWS's sagemaker service for running the LLM model and the

stable diffusion model. Using AWS's sagemaker we can run the model with ease. The models are then interacting with the help of AWS's Lamda function which is also authenticated.

All the object is stored at AWS's S3 bucket and then used on the front-end

The API Details:
Method: POST, [URL](#)
Body:
{
 "menu":{
 "email": 1,
 "document_search": 0,
 "presentation": 1,
 "search": 1
 },
 "prompt": "Information about AI"
}

The API then properly algin's the data and provide it to the boss.

Thanking You

Neeraj Rana, Aimen Sadat & Ayesha Kalsoom