Build a PDF QA chatbot using Azure OpenAI and Streamlit

Abstract:

In this 2-hour workshop, participants will explore the powerful capabilities of Retrieval Augmented Generation (RAG) with Azure OpenAI, Langchain, ChromaDB, and Streamlit. Designed for beginners with basic knowledge in generative AI and large language models, this interactive session will combine informative slides with hands-on code-along activities to demonstrate the integration and application of these cutting-edge technologies. Along the way, other tools and frameworks may be introduced to complement the learning experience. By the end of the workshop, attendees will gain a foundational understanding of how RAG can be applied in various use cases and improve user interactions.

Workshop Outline:

1. Introduction
   - Brief overview of the workshop agenda
   - Explanation of RAG and its relevance for beginners

2. Azure OpenAI
   - Introduction to Azure OpenAI and its capabilities
   - Simplified overview of integrating Azure OpenAI with RAG
   - Hands-on activity: Setting up Azure OpenAI for RAG (with step-by-step guidance)

3. Langchain
   - Introduction to Langchain and its role in text generation
   - Combining Langchain with RAG for advanced language models (explained in layman's terms)
   - Hands-on activity: Configuring and utilizing Langchain with RAG (with clear instructions)

4. ChromaDB
   - Overview of ChromaDB and its applications in RAG
   - Storing and retrieving data using ChromaDB (explained in simple terms)
   - Hands-on activity: Implementing ChromaDB for RAG data storage (with guidance)

5. Streamlit
   - Introduction to Streamlit and its benefits for creating interactive web applications

- Integrating RAG with Streamlit for user-friendly interfaces (explained in an approachable manner)

- Hands-on activity: Building a simple RAG-powered Streamlit application (with step-by-step instructions)

6. Real-world Applications

- Discussion of use cases and benefits for beginners

- How RAG can be applied in various scenarios and improve user interactions

7. Q&A and Closing

- Addressing questions from participants

- Recap of key takeaways

- Final thoughts and next steps for exploring RAG and associated technologies for beginners

Workshop Outline:

1. Introduction
    - Brief overview of the workshop agenda
    - Explanation of RAG and its relevance for beginners
    - Mention of other tools and frameworks to be introduced during the workshop
2. Azure OpenAI
    - Introduction to Azure OpenAI and its capabilities
    - Simplified overview of integrating Azure OpenAI with RAG
    - Hands-on activity: Setting up Azure OpenAI for RAG (with step-by-step guidan ce)
    - Introduction of any supplementary tools related to Azure OpenAI
3. Langchain
    - Introduction to Langchain and its role in text generation
    - Combining Langchain with RAG for advanced language models (explained in l ayman's terms)
    - Hands-on activity: Configuring and utilizing Langchain with RAG (with clear ins tructions)
    - Introduction of any additional frameworks related to Langchain
4. ChromaDB
    - Overview of ChromaDB and its applications in RAG
    - Storing and retrieving data using ChromaDB (explained in simple terms)
    - Hands-on activity: Implementing ChromaDB for RAG data storage (with guida nce)
    - Introduction of other relevant tools in the context of ChromaDB
5. Streamlit

    - Introduction to Streamlit and its benefits for creating interactive web applicati ons
    - Integrating RAG with Streamlit for user-friendly interfaces (explained in an app roachable manner)
    - Hands-on activity: Building a simple RAG-powered Streamlit application (with step-by-step instructions)
    - Introduction of any supplementary tools or frameworks related to Streamlit
6. Real-world Applications
    - Discussion of use cases and benefits for beginners
    - How RAG can be applied in various scenarios and improve user interactions
    - Emphasizing the focus on Azure OpenAI, Langchain, ChromaDB, and Streamlit
7. Q&A and Closing
    - Addressing questions from participants
    - Recap of key takeaways, including the main tools and frameworks covered
    - Final thoughts and next steps for exploring RAG and associated technologies f or beginners

To use Azure's OpenAI resources, you'll need to create an Azure Cognitive Services resource and obtain the necessary keys. Here's a step-by-step guide to help the attendees set up the required Azure OpenAI resources and get the API keys:

1. Sign in to the Azure Portal:

a. Visit the Azure Portal at https://portal.azure.com/.

b. Sign in with your Microsoft account credentials. If you don't have an account, you can sign up for a free trial: https://azure.microsoft.com/en-us/free/.

2. Create a Cognitive Services resource:

a. Click on the "Create a resource" button in the top-left corner of the Azure Portal.

b. In the search box, type "Cognitive Services" and select it from the list.

c. Click the "Create" button to create a new Cognitive Services resource.

d. Fill in the required details:

- Subscription: Choose your Azure subscription.

- Resource group: Create a new resource group or select an existing one.

- Region: Choose a region closest to your location or your users.

- Name: Provide a unique name for your Cognitive Services resource.

- Pricing tier: Choose a pricing tier based on your requirements (e.g., S0, F0).

e. Review the terms and click "Review + Create" to proceed.

f. Review your configuration and click "Create" to deploy the Cognitive Services resource.

3. Obtain the API keys:

a. Once the deployment is complete, navigate to the "Resource Management" section of the Azure Portal.

b. Locate and click on your newly created Cognitive Services resource.

c. In the resource's overview page, click on "Keys and Endpoint" in the left-hand menu.

d. You will see two API keys, either of which can be used to authenticate with the Cognitive Services API. Copy one of these keys and store it securely, as you'll need it to access the OpenAI resources.

Now that you have the Azure OpenAI resource and API keys, you can use them to authenticate and interact with Azure's Cognitive Services APIs, such as the Text Analytics API, Computer Vision API, and more.

Make sure to share these instructions with your workshop attendees, so they can set up their Azure OpenAI resources and obtain the required API keys before the workshop begins.

Title: "Engineers Guide to Spark in Azure"

In this session we'll explore the Data Engineering experience in Azure across Databricks and Microsoft Fabric, focusing on capabilities, performance, feature comparisons, developer experience, and use cases.