Name: Omar Khaled Ahmed

WAZZUF RAG TASK

WUZZUF

Introduction:

Developing complete RAG system using open source models as well as fine-tuning the chosen LLM on the dataset provided, We will dive into:

- 2 different solution approaches as Colab notebooks one for RAG and one for Fine-tuning.

- Discussing the approach for each.

RAG Application: **Open in Colab**

In this notebook you will find complete rag app with llama-3 model

Methodologies:

- Exploratory data analysis for better understanding of the data.

- Preprocessing the necessary only. We kept the html parser for better visualization as markdown output.

- loading the model charts and quantizing model to work on low recourses VRam of google colab, in this case we choose **LLAMA-3-8b-instruct from META** as its considered SOTA model on the benchmark currently compared to the same sizes.

- choosing the embedding model **e5 from Microsoft** as it can work on both English and Arabic which makes it the perfect choice for a middle east company considering both languages are the main ones.

- Using Langchain library for its high compatibility and interpretability of code for chunking data, and Retrieval of the data using **FAISS** db as a vector store with the cosine similarity method.

• consider working with **qa chain** as a single inference only OR another option is with the **llm memory buffer** which keeps the chat history.

• finally simple GUI for chatting with your data and uploading documents you need to retrieve from.

• Some of the model outputs:

Question: Give me personalized career advices on how to become a business developer
Helpful Answer:

As a Business Development Executive, your primary goal is to identify and capitalize on new business opportunities. Here are some personalized career advice to help you succeed in this role:

1. **Develop a growth mindset**: As a Business Development Executive, you'll need to be adaptable, resilient, and open to new challenges. Focus on continuous learning, and be willing to take calculated risks to drive growth.
2. **Build a strong network**: Develop relationships with key stakeholders, including potential clients, partners, and industry experts. Attend conferences, join relevant associations, and engage in online communities to expand your network.
3. **Stay up-to-date with industry trends**: Continuously monitor industry developments, market shifts, and competitor activity. Stay informed about new technologies, innovations, and best practices to stay ahead of the curve.
4. **Develop a robust sales strategy**: Create a tailored sales strategy that aligns with your target market, competition, and company goals. Focus on building rapport, identifying pain points, and showcasing the value proposition.
5. **Foster a positive work environment**: As a Business Development Executive, you'll be responsible for managing a team. Focus on creating a positive, inclusive, and supportive work environment that fosters collaboration and innovation.
6. **Develop strong communication skills**: As a Business Development Executive, you'll need to communicate effectively with diverse stakeholders, including clients, partners, and colleagues. Develop strong written and verbal communication skills to articulate your vision, negotiate deals, and build trust.
7. **Set clear goals and targets**: Establish clear, measurable goals and targets for yourself and your team. Focus on achieving these targets, and continuously monitor progress to adjust your strategy.

Remember, as a Business Development Executive, your success depends on your ability to build relationships, identify opportunities, and drive growth. Focus on developing a growth mindset, staying informed, and building a strong network to achieve your goals. Good luck!

Question: Give me the job requirements for a Project Management job
Helpful Answer: Based on the provided job description, here are the job requirements for a Project Manager:

- Bachelor's degree in Engineering or related field
- Above 15 years of experience in project management
- Coordinate with different levels of stakeholders
- Manage a team of experts specialized in the Environmental industry
- Prepare a detailed execution and QHSE plan for the project
- Assure that the project deliverables have been submitted to client as per the agreed schedule
- Ensure project/client requirements are met and respected by the project team
- Being the focal point of the project
- Coordinate with the client representative in order to collect the required data
- Technical Expertise: Project Management

Please note that these requirements are based on the provided job description and may not be exhaustive. It's always best to verify the job requirements with the hiring manager or the company's HR department.

ChatWithYourData_Bot

| Conversation | Database | Chat History | Configure |

Enter text here...

1. **Gain relevant education and training**: As you're already a software engineer, you're off to a great start. Consider taking courses or earning a master's degree in machine learning or data science to gain the necessary knowledge and skills.
2. **Build your skills in AI and machine learning**: Focus on developing skills in AI and machine learning, such as data mining, data preparation, and machine learning algorithms. You can take online courses or work on personal projects to build your portfolio.
3. **Get hands-on experience with machine learning tools and technologies**: Familiarize yourself with popular machine learning tools and technologies, such as TensorFlow, PyTorch, or Scikit-learn. Practice building machine learning models and experimenting with different algorithms.
4. **Network with other machine learning engineers**: Attend conferences, join online communities, or participate in hackathons to connect with other machine learning engineers and learn from their experiences.
5. **Stay up-to-date with industry trends and advancements**: Follow industry leaders and researchers, and stay informed about the latest advancements in machine learning and AI.
6. **Consider earning a certification**: While not necessary, earning a certification in machine learning or AI can demonstrate your expertise and commitment to the field.

ChatWithYourData_Bot

Conversation   Database   Chat History   Configure

Enter text here…

3. **Get hands-on experience with machine learning tools and technologies**: Familiarize yourself with popular machine learning tools and technologies, such as TensorFlow, PyTorch, or Scikit-learn. Practice building machine learning models and experimenting with different algorithms.
4. **Network with other machine learning engineers**: Attend conferences, join online communities, or participate in hackathons to connect with other machine learning engineers and learn from their experiences.
5. **Stay up-to-date with industry trends and advancements**: Follow industry leaders and researchers, and stay informed about the latest advancements in machine learning and AI.
6. **Consider earning a certification**: While not necessary, earning a certification in machine learning or AI can demonstrate your expertise and commitment to the field.
7. **Highlight your transferable skills**: As a software engineer, you likely have strong analytical, problem-solving, and communication skills. Highlight these transferable skills when applying for machine learning engineer positions or discussing your career goals with others.

Remember, becoming a machine learning engineer requires dedication, persistence, and continuous learning. Stay focused, and you'll be well on your way to a successful career in this exciting field. Good luck!

ChatWithYourData_Bot

Conversation **Database** Chat History Configure

DB query:

Give me personalized career advices on how to become a machine learning engineer

Result of DB lookup:

page_content='job_title: Machine Learning Engineer' metadata={'source': '/content/

page_content='pick up new software systems and data analysis tools. Experience with

page_content='requirements:

- Bachelor's or master's degree in computer science, Engineering, Data Science,
- Proven experience in software testing, with a focus on AI or machine learning
- Strong understanding of AI concepts.
- Strong analytical, problem-solving, and communication skills.
- Ability to work effectively in a fast-paced, collaborative environment.
- Certification in AI or software testing (e.g., ISTQB, AI Certification) is a

career_level: Experienced (Non-Manager)
description_length: 138
requirements_length: 55' metadata={'source': '/content/sampled_jobs.csv', 'row': 2

page_content='Administrator).
- Experience with microservices architecture.
- Knowledge of serverless computing.
- Familiarity with agile methodologies.
' metadata={'source': '/content/sampled_jobs.csv', 'row': 346}

r in another tab. Show diff

ChatWithYourData_Bot

Conversation Database **Chat History** Configure

Current Chat History variable

('Give me personalized career advices on how to become a machine learning engineer

- Bachelor's or master's degree in computer science, Engineering, Data Science,
- Proven experience in software testing, with a focus on AI or machine learning
- Strong understanding of AI concepts.
- Strong analytical, problem-solving, and communication skills.
- Ability to work effectively in a fast-paced, collaborative environment.
- Certification in AI or software testing (e.g., ISTQB, AI Certification) is a

\ncareer_level: Experienced (Non-Manager)\ndescription_length: 138\nrequirements_l

- Experience with microservices architecture.
- Knowledge of serverless computing.
- Familiarity with agile methodologies.

\n\nQuestion: Give me personalized career advices on how to become a machine learn
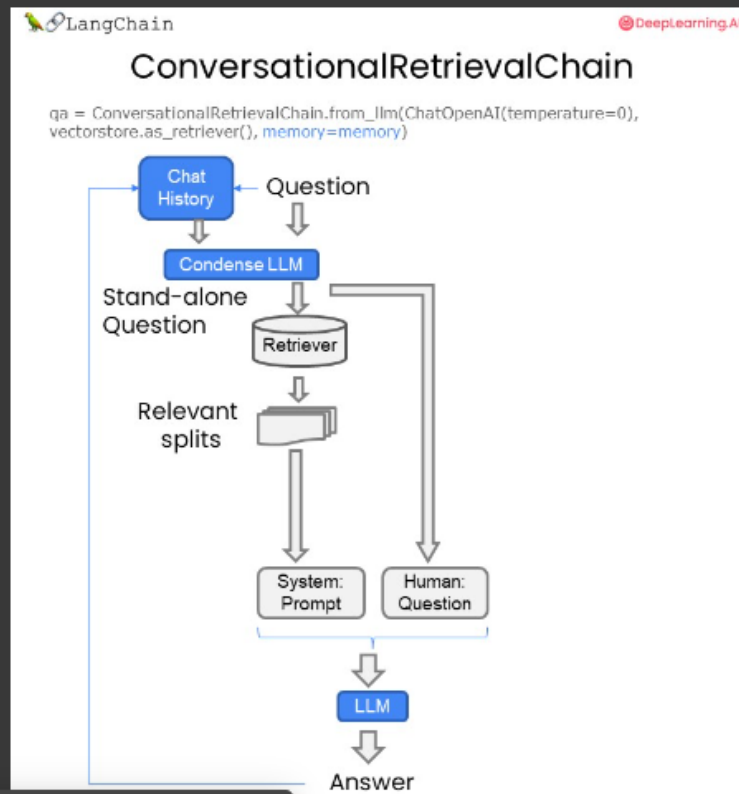
ChatWithYourData_Bot

Choose File | No file chosen | **Load DB** | Loaded File: /content/sampled_jobs.csv

Clear History | Clears chat history. Can use to start a new topic

LangChain   DeepLearning.AI

ConversationalRetrievalChain

```
qa = ConversationalRetrievalChain.from_llm(ChatOpenAI(temperature=0),
vectorstore.as_retriever(), memory=memory)
```

Chat History ← Question

↓

Condense LLM

Stand-alone Question

↓

Retriever

Relevant splits

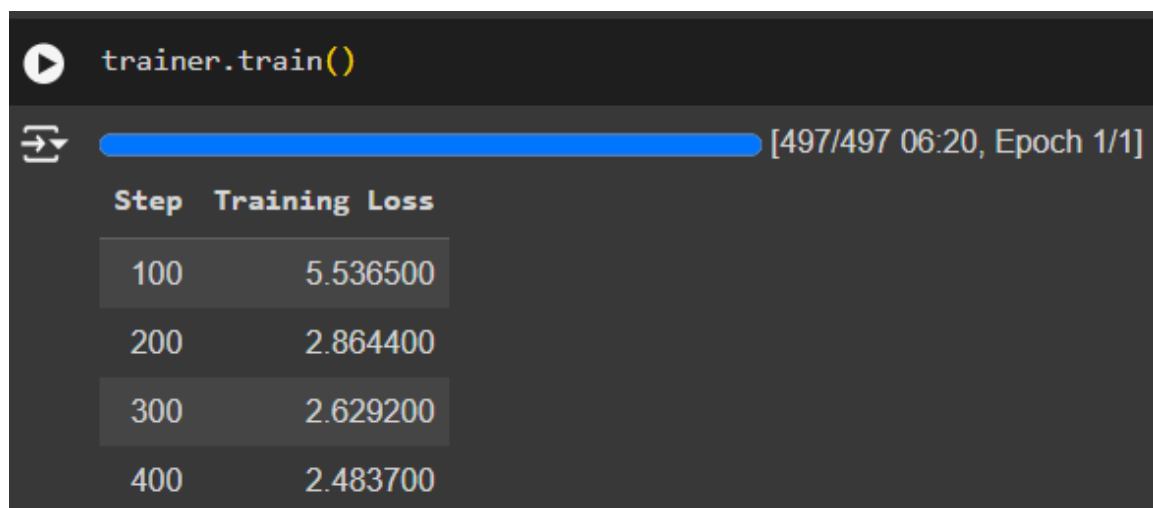↓

System: Prompt   Human: Question

↓

LLM

↓

Answer

Finetuning LLMs: Open in Colab

In this notebook you will find finetuning llama-3 model on the dataset after preprocessing.

Methodologies:

• Data preparation preprocessing for tokenizing the data.

• Model configuration and fine tuning using PEFT with LoRA.

• Save the model and testing the ouput.


```
trainer.train()
```

[497/497 06:20, Epoch 1/1]

Step	Training Loss
100	5.536500
200	2.864400
300	2.629200
400	2.483700

Apologies:

1. Time consumption and deadline due to personal reasons and working hours pressure.
2. Lack of complete deployement environment and recourses which lead to difficulties setting up the docker, fast-api environment and frontend with vue.js for the UI of the chatbot
3. The dismiss of model evaluation scores which can be used using different approaches as setting up anpther LLM to evaluate the model, ROUGE scores for comparing the model output with the original one, relevance, coherence.
4. The unavailability for writing a well organized read me file and a documentation for this task due to time.

Note: Future work and more time would lead to correct and complete the full application desired isa. 😊