**Session 1**

I have followed the tutorial: https://python.langchain.com/v0.2/docs/tutorials/llm\_chain/

The full tutorial list is below, my study plan is to follow these tutorials: https://python.langchain.com/v0.2/docs/tutorials/

Introduction page: https://python.langchain.com/v0.2/docs/introduction/

Optionally create a virtual environment guide is here: https://medium.com/@claudia.nikel/how-to-setup-a-jupyter-notebook-in-vs-code-w-virtual-env-kernels-install-packages-884cf643375e

Genel baslangic icin bunu takip ettim:

<https://www.gettingstarted.ai/everything-you-need-to-know-when-getting-started-with-langchain/>

**Session 2**

RAGSETUPI KUR LLMI KUR BAGLA RAGSTEUPA - ISTEDIGI BU

Langflow - SADECE ASINALIK ICIN

<https://www.langflow.org>

<https://blog.onesaitplatform.com/en/2023/11/27/getting-to-know-langflow/>

<https://cobusgreyling.medium.com/langflow-for-langchain-58c143ba9218> - u can use this as a guide - huggingface guide is here as well

<https://medium.com/logspace/langflow-micro-tutorials-pdf-parser-22e0772de330>

SADECE ASINALIK ICIN - RAGIN KULLANIM KONSEPTINI OTURTMAK ICIN

<https://selfrag.github.io>

Langgraph

<https://langchain-ai.github.io/langgraph/>

VECTOR DATABASE - SADECE ASINALIK ICIN

<https://www.pinecone.io/learn/vector-database/>

<https://www.youtube.com/watch?v=BrsocJb-fAo>

Youtube video:

<https://www.youtube.com/watch?v=BrsocJb-fAo>

Docs guide: <https://docs.google.com/document/d/1kghX2fpfLd2uul8vZIClh22QNAz1u-aVknlQTkSKi7M/edit>

Yeniden canlandirma guide:

Vscodedan LLM dosyasini ac,

cmmadn terminalde venvi reactivate et

Notebook dosyasina git ve yukarida > yapip python interpretordan venvimizi sec

**First Setup Openai add payment info:https://platform.openai.com/settings/organization/billing/overview**

**Openai api key:**

**sk-proj-250ZP0zkF1zIqrbZbeKS78-nzIXD1iP71cw-3SPXDgMw4mDUaCiYrn\_r7DeF3549xKneutvSBYT3BlbkFJS0zcciQlarKp6aZj9uSt2ewz4Hf1pHjs8Dp7ISHFtqZhuwufCaMbGK2CeD9F22GuVbOPbPJsQA**

**Find your keys here:https://platform.openai.com/api-keys**

Adam diyor ki its going to cost cents but you need it.

PINECONE API KEY IS HERE:

<https://app.pinecone.io/organizations/-O5bx7FhH8PcEjKS1xBc/projects/340e76fa-b81c-47f7-850a-cf7e68b864ec/keys>

Create a .env file with the following variables and install requirements/txt to ur venv

Using a language model like gpt to ask question about a video

Langchain allows us to access the model

chain = model | parser

This is the first instance of a chain usage, the whole idea behind lachain, it says ok run the model and take its output and feed it as an input into the parser

Tokens: the llm is basically a neural network where every word needs to be translate into vectorized set of numbers, this is called tokenization.

75% of token length equals the words length

We cant fit our huge database into a single prompt

Lets say we have 2 databases math and history class, if the student asks a question about the history class we dont need to send the whole 2 classes as context for tehh model, we only send history class so we need to find the portion of transaction thats required to answer to question and only feed that portion to the model as context

Recursivecharactersplitter; we gonna take the entire transcript and well specify how many characters we want in every chunk, then we will specify ho much of an overlap between the chunks

After having out chunks, we ned to determine which one is the relevant ones. We use the idea of embeddings.

Embedding is a special function that puts an idea, text, image, etc to a location in the 3d space, if another idea is very similar then their locations will bev ery close to each other in this 3d space, by comparing th embeddings of the question w the embeddings of the chinks we can comput a simimlarityy index for each of the chunks and take thee most similar ones

Vector store: a database for vectors, a database for embeddings, wecan use the database for storing all of our content/documents from the transcription, automatically generate and store the embeddings from all of these documents so that we dont have to generate theme or every question, and to perform optimized similarity search really quickly

Retriever, wth the input of vectorstore1 was formed

Once we invoke it it does everything above in the background and by default gives the most similar 4 chunks

setup = RunnableParallel(context=retriever1, question=RunnablePassthrough())

setup.invoke("What color is Patricia's car?")

Here: the retriever1 needs the question, context comes from the retriever, and the question comes from the previous part of the chain, so it is a passthrough (see the image)

So setup is ready for the model: context and question input is ready for it

Right now our vectorstore is in our memory, but we dont want it to be in our memory, we want it to be physically stored in a database, so we use pinecone

<https://www.youtube.com/watch?v=sVcwVQRHIc8>

<https://www.youtube.com/watch?v=JEBDfGqrAUA>

<https://www.youtube.com/watch?v=f-AXdiCyiT8>

Data pipeline llm flowlari olustur

Agent generation uzerine

OKUL ARASTIRMA

Hocanin karsisina bir konu cikmis ona odaklanmak istedigini soyledi,

<https://learning.cs.toronto.edu/theses.html>

<https://www.cs.toronto.edu/~ndjaitly/Jaitly_Navdeep_201411_PhD_thesis.pdf>

<https://www.cs.toronto.edu/~gdahl/papers/Dahl_George_E_201506_PhD_thesis.pdf>

Deep learning approaches to problems in speech recognition, computational chemistry, and natural language text processing

Exploring Deep Learning Methods for discovering features in speech signals

gENERATIV AIDA

Kisa bi calisma icin llm ya da

Fizik bazli modeller - daha vakit istiyor

Ne tur fiziik verisi goruntu mu ses mi yoksa multimodal mi

Fiziksel model icin doktora dusunebilirsin

Kendine bir senaryo ciz, phd mi masterda bi yokla fizikle generative ai ile ne kadar giderim

Fizik verisini isleyebilmek ozel nir kabiliyet data scientist olarak.

Tezini neye gore yazicaksin?

Llm kisminda senaryo: cok falza calisma var birsuru

Mert abi bana inaniyor ondan basarmam lazim, model yaptim diyince ne mutlu oldu