

21i-2668 Manahil Kamran
21i-2669 Ali Arfa
21i-2676 Shitba Kashif DS-N

AI Project Report: “ MINDMATE”

Python Code:

Purpose:

The Python code aims to create a Flask web application for a psychology study chat. It utilizes language models and vector databases to retrieve relevant documents based on user queries.

Components:

Libraries:

```
from langchain_community.document_loaders import UnstructuredPDFLoader
from langchain_community.document_loaders import OnlinePDFLoader
from langchain_community.embeddings import OllamaEmbeddings
from langchain_text_splitters import RecursiveCharacterTextSplitter
from langchain_community.vectorstores import Chroma
from langchain.prompts import ChatPromptTemplate, PromptTemplate
from langchain_core.output_parsers import StrOutputParser
from langchain_community.chat_models import ChatOllama
from langchain_core.runnables import RunnablePassthrough
from langchain.retrievers.multi_query import MultiQueryRetriever
```

Document Loaders: The code imports UnstructuredPDFLoader and OnlinePDFLoader to load PDF documents for processing. Specifically, it focuses on the "Fundamentals-of-Psychological-Disorders.pdf" book, which aids in the study of psychological disorders.

```
local_path = "books\Fundamentals-of-Psychological-Disorders.pdf"

# Local PDF file uploads
if local_path:
    loader = UnstructuredPDFLoader(file_path=local_path)
    data = loader.load()
else:
    print("Upload a PDF file")

# Preview first page
data[0].page_content
```

Fundamentals of Psychological Disorders\n\nFundamentals of Psychological Disorders\n\nPDF Version of the Textbook - Fundamentals

21i-2668 Manahil Kamran
21i-2669 Ali Arfa
21i-2676 Shitba Kashif DS-N

Embeddings: OllamaEmbeddings are used to embed text data for similarity comparisons. The chosen embedding model, "nomic-embed-text," provides effective representations for the chatbot's understanding of psychological concepts.

```
# Add to vector database
vector_db = Chroma.from_documents(
    documents=chunks,
    embedding=OllamaEmbeddings(model="nomic-embed-text", show_progress=True),
    collection_name="local-rag"
)
```

OllamaEmbeddings: 100%|██████████| 129/129 [05:08<00:00, 2.39s/it]

Text Splitters: RecursiveCharacterTextSplitter is employed to split documents into smaller chunks for processing, enhancing efficiency and scalability.

```
# Split and chunk
text_splitter = RecursiveCharacterTextSplitter(chunk_size=7500, chunk_overlap=100)
chunks = text_splitter.split_documents(data)
```

Vector Stores: Chroma is used to create a vector database from the document chunks and embeddings. This enables efficient retrieval of relevant documents based on user queries.

Prompts: ChatPromptTemplate and PromptTemplate are used to structure prompts for the language model, facilitating coherent and context-aware responses.

```
QUERY_PROMPT = PromptTemplate(
    input_variables=["question"],
    template="""You are an AI language model assistant. Your task is to generate five
different versions of the given user question to retrieve relevant documents from
a vector database. By generating multiple perspectives on the user question, your
goal is to help the user overcome some of the limitations of the distance-based
similarity search. Provide these alternative questions separated by newlines.
Original question: {question}""",
)
```

Output Parsers: StrOutputParser and CommaSeparatedListOutputParser parse model outputs for display, ensuring readability and user-friendliness.

21i-2668 Manahil Kamran

21i-2669 Ali Arfa

21i-2676 Shitba Kashif DS-N

Chat Models: ChatOllama is utilized as the language model for generating responses. It was fine-tuned using LangChain and RAG techniques to optimize its performance for psychological queries.

```
# LLM from Ollama
#local_model = "mistral"
local_model='phi3'
llm = ChatOllama(model=local_model)
```

Retrievers: MultiQueryRetriever is employed to retrieve documents based on user queries, enhancing the chatbot's ability to provide relevant information.

```
retriever = MultiQueryRetriever.from_llm(
    vector_db.as_retriever(),
    llm,
    prompt=QUERY_PROMPT
)

# RAG prompt
template = """Answer the question based ONLY on the following context:
{context}
Question: {question}
"""

prompt = ChatPromptTemplate.from_template(template)
```

Model Selection:

After experimenting with various models like GPT-2, Llama 2 and 3, Mistral, and others, Phi3 emerged as the most suitable choice. Phi3 offers a balance of computational efficiency, speed, and accuracy, making it ideal for real-time chat interactions. Additionally, its compatibility with GPU acceleration further enhances performance.

21i-2668 Manahil Kamran

21i-2669 Ali Arfa

21i-2676 Shitba Kashif DS-N

```
C:\Users\Manahil Kamran\AppData\Local\Ollama>
Microsoft Windows [Version 10.0.19045.4291]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\system32>ollama run phi3
pulling manifest
pulling 4fed7364ee3e... 97% ▒ 2.3 GB/2.3 GB 2.1 MB/s
Error: Post "http://127.0.0.1:11434/api/show": dial tcp 127.0.0.1:11434: connect: No connection could be made
the target machine actively refused it.

C:\Windows\system32>ollama run phi3
pulling manifest
pulling 4fed7364ee3e... 100% ▒ 2.3 GB
pulling c608dc615584... 100% ▒ 149 B
pulling fa8235e5b48f... 100% ▒ 1.1 KB
pulling d47ab88b61ba... 100% ▒ 140 B
pulling f7eda1da5a81... 100% ▒ 485 B
verifying sha256 digest
writing manifest
removing any unused layers
success
>>>
>>> hi
Hello! How can I assist you today? If you have any questions or need information, feel free to ask.

>>> /bye

C:\Windows\system32>ollama list
NAME                                ID                                SIZE    MODIFIED
nomic-embed-text:latest            0a109f422b47                    274 MB  20 hours ago
phi3:latest                         a2c89ceaed85                    2.3 GB  2 minutes ago
```

Model Setup:

The `model_load()` function sets up the language model, vector database, retriever, and chat chain. It loads the "Fundamentals-of-Psychological-Disorders.pdf" book, splits it into chunks, creates embeddings, initializes the ChatOllama model, defines a query prompt, creates a retriever, and constructs a processing chain.

```
chain = (
    {"context": retriever, "question": RunnablePassthrough()}
    | prompt
    | llm
    | StrOutputParser()
)
```

Input/Output (question and answers):

```
chain.invoke("What is Epidemiology?")

OllamaEmbeddings: 100%|██████████| 1/1 [00:06<00:00, 6.40s/it]
OllamaEmbeddings: 100%|██████████| 1/1 [00:02<00:00, 2.09s/it]
OllamaEmbeddings: 100%|██████████| 1/1 [00:02<00:00, 2.08s/it]
OllamaEmbeddings: 100%|██████████| 1/1 [00:02<00:00, 2.10s/it]
OllamaEmbeddings: 100%|██████████| 1/1 [00:02<00:00, 2.12s/it]

"Clinical assessment refers to the process of gathering information about a patient through various methods such as observation,
```

21i-2668 Manahil Kamran
21i-2669 Ali Arfa
21i-2676 Shitba Kashif DS-N

Question:

```
chain.invoke("What is Epidemiology?")
```

Answer:

```
"Clinical assessment refers to the process of gathering information about a patient through various methods such as observation, psychological tests, neurological evaluations, and interviews. Reliability in this context relates to the consistency with which these methods produce stable and consistent results over time, while validity ensures that the assessments accurately measure what they are intended to measure. Standardization is crucial as it provides a uniform approach to administering tests and evaluations, thereby minimizing variability in how each patient's data is collected.\n\nSection 3.1 Review Questions\n\n1. Clinical assessment involves multiple methods that have their own strengths (e.g., objectivity of psychological testing) and limitations (e.g., subjective interpretation of interviews). Reliability ensures consistent results across repeated tests, whereas validity guarantees the assessment measures accurately. Standardization allows for uniform administration, critical to producing comparable data among patients.\n\n2. A syndrome in clinical diagnosis refers to a group of symptoms that consistently occur together and often have predictable patterns and courses specific to certain disorders."
```

Flask Routes:

- /model and /: Render the chat interface template for user interaction.
- /query: Handle POST requests containing user queries, invoke the chat chain, and return model responses.
- /rem: Provide an endpoint to delete the vector database collection, enabling maintenance and optimization.

Challenges Faced:

- Ensuring seamless integration of various components for effective chatbot functionality.
- Managing model selection and optimization for performance and accuracy.
- Addressing scalability concerns, especially when dealing with large datasets and complex queries.

GUI Code:

Purpose:

The HTML/JavaScript code provides the frontend interface for the psychology study chat. It allows users to input queries, submit them to the backend, and display responses from the language model.

21i-2668 Manahil Kamran
21i-2669 Ali Arfa
21i-2676 Shitba Kashif DS-N
Components:

HTML Structure: Defines the layout of the chat interface, including input form and message display area.

CSS Styling: Styles the chat interface with colors, borders, and spacing, ensuring a visually appealing and user-friendly experience.

JavaScript: Handles form submission, fetches model responses from the backend, and displays messages in the chat history, facilitating smooth user interactions.

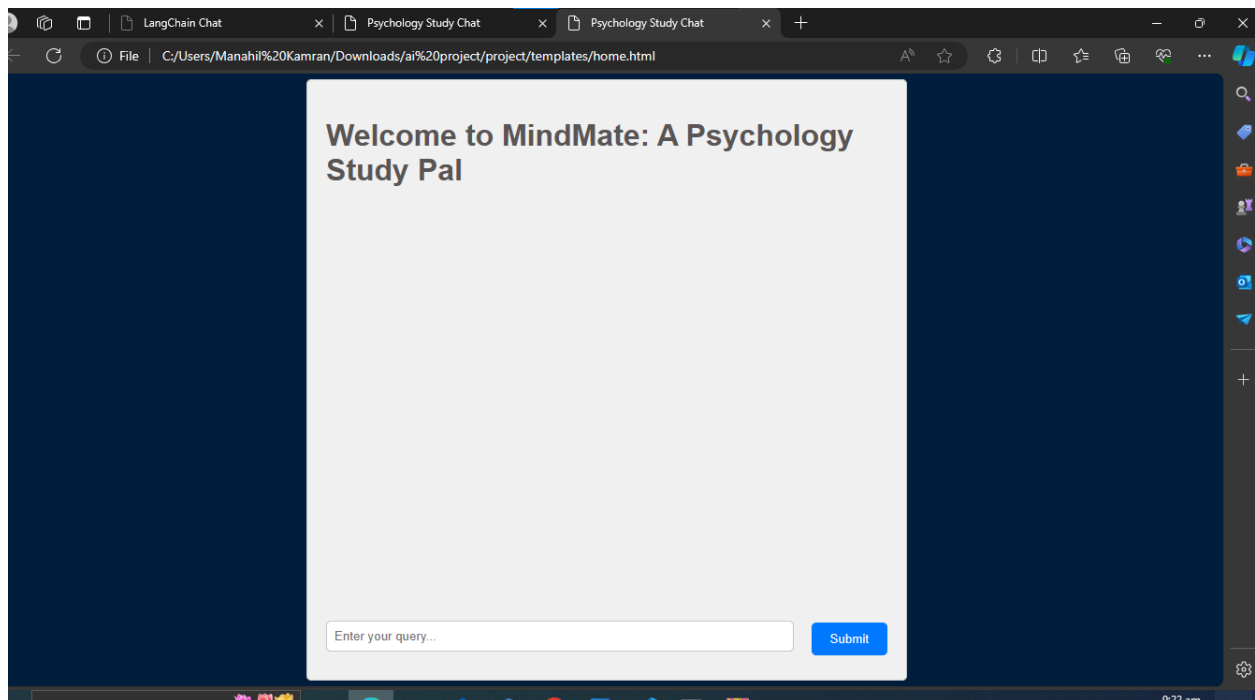
Features:

Input Form: Allows users to enter queries for the chatbot.

Submit Button: Submits queries to the backend for processing.

Model Button: Allows users to reload the model for updates or maintenance.

Chat History Display: Dynamically displays user queries and model responses in the chat area, providing a comprehensive conversation log.



Executed:

Welcome to Personal therapy Chat

You: hi

Bot: Hello! I'd be glad to provide information about specific learning disorders and related topics within psychology. If you have any questions or need clarification on these areas, feel free to ask! Here is a brief overview of the concepts mentioned in your initial prompt: 1. Rationalizing Behaviors: People sometimes justify their actions using various reasoning strategies, such as shifting blame onto external factors (external locus of control) or justifying themselves based on perceived necessity (just-world fallacy). However, it's essential to understand the underlying reasons for behaviors genuinely. 2. Intellectualization: This defense mechanism involves focusing excessively on intellectual aspects while disregarding emotions related to an event or situation. For example, someone dealing with grief might immerse themselves in research about loss and statistics rather than processing their emotions directly. 3. Autism Spectrum Disorder (ASD): ASD is a neurodevelopmental condition characterized by two core domains: social communication difficulties and restricted/repetitive patterns of behavior, interests, or activities. The manifestation of ASD varies in severity across the spectrum, including challenges with understanding others' perspectives, reading emotions, sensory sensitivities, and more. For detailed information on these topics, I recommend consulting academic resources like textbooks, research articles, or online educational platforms (e.g., Opentext). These materials can provide valuable insights into the complexities of human behavior and psychological disorders while avoiding oversimplification or generalizations about individuals.

You: what is the authors name

Bot: The author of "Fundamentals of Psychological Disorders" referenced in Section 3.2.1, Diagnostic and Statistical Manual of Mental Disorders (DSM) history, is quoted as the American Psychiatric Association (APA). However, individual chapters or sections within this text may be written by various authors. To identify a specific author's name, you would need to look at the chapter or section in question and consult the book's index for author names. If we were discussing the DSM-5 itself as an edited volume, then the primary editor could often be cited alongside APA due to its role with the American Psychiatric Association publishing it. However, without a specific chapter or section reference, the most accurate information would attribute the content to APA and their publication process rather than individual authors. For more precise authorship details on particular chapters within "Fundamentals of Psychological Disorders," you'd need access to the full text for an index listing contributors.

Enter your query:

Challenges Faced:

- Ensuring proper alignment and styling of the chat interface elements for a visually appealing design.
- Implementing form submission handling and asynchronous communication with the backend to maintain responsiveness and interactivity.
- Managing dynamic content updates and ensuring a smooth user experience, especially during periods of high traffic or complex queries.