#!/usr/bin/env python3

import os

import re

import json

import time

import requests

import numpy as np

from random import choice

from dotenv import load\_dotenv

import chromadb

from sentence\_transformers import SentenceTransformer

from sklearn.metrics.pairwise import cosine\_similarity

from openai import OpenAI

import openai  # for exception classes

from langchain.chat\_models import ChatOpenAI

from langchain.chains import LLMChain

from langchain.prompts import PromptTemplate

from langchain.memory import ConversationBufferMemory

from langchain.schema import SystemMessage, HumanMessage

# ─── 0) Paths & env ───────────────────────────────────────────────────────────

BASE\_DIR    = os.path.abspath(os.path.join(os.path.dirname(\_\_file\_\_), ".."))

DATA\_DIR    = os.path.join(BASE\_DIR, "data")

DB\_DIR      = os.path.join(BASE\_DIR, "quran\_db")

TAGGED\_JSON = os.path.join(DATA\_DIR, "quran\_chunks\_tagged.json")

load\_dotenv(os.path.join(BASE\_DIR, ".env"))

OPENAI\_API\_KEY = os.getenv("OPENAI\_API\_KEY")

HADITH\_API\_KEY = os.getenv("HADITH\_API\_KEY")

# ─── 1) LLM & memory ───────────────────────────────────────────────────────────

llm = ChatOpenAI(

    model\_name="gpt-4.1",

    temperature=0.5,

    openai\_api\_key=OPENAI\_API\_KEY

)

memory = ConversationBufferMemory(

    memory\_key="history",

    input\_key="question",

    return\_messages=False

)

# ─── 2) ChromaDB setup ────────────────────────────────────────────────────────

emb\_fn = chromadb.utils.embedding\_functions.SentenceTransformerEmbeddingFunction(

    model\_name="all-MiniLM-L6-v2"

)

client\_db = chromadb.PersistentClient(path=DB\_DIR)

collection = client\_db.get\_or\_create\_collection(

    name="quran",

    embedding\_function=emb\_fn

)

# ─── 3) Load verse→tags map ────────────────────────────────────────────────────

with open(TAGGED\_JSON, "r", encoding="utf-8") as f:

    tagged = json.load(f)

verse\_tags = {}

for v in tagged:

    t = v.get("tags", [])

    if isinstance(t, list):

        tags = t

    else:

        tags = [s.strip() for s in t.split(",") if s.strip()]

    verse\_tags[v["id"]] = set(tags)

# ─── 4) Tag detection helpers ──────────────────────────────────────────────────

EMOTION\_PATTERNS = {

    "anxiety":    [r"\banxiou?s?\b", r"\bworry(ing)?\b", r"\bpanic\b"],

    "hope":       [r"\bhope\b", r"\boptimis(m|tic)\b"],

    "gratitude":  [r"\bthank(s|ful)\b", r"\bgratitude\b"],

    "fear":       [r"\bfear\b", r"\bscared\b"],

    "grief":      [r"\bgrief\b", r"\bloss\b", r"\bsorrow\b"],

    "self-worth": [r"\bworthless\b", r"\bself-worth\b", r"\bestee?m\b"],

    "patience":   [r"\bpatient\b", r"\bpatience\b"],

    "strength":   [r"\bstrong\b", r"\bstrength\b"],

    "community":  [r"\bcommunity\b", r"\bbelong\b"],

    "optimism":   [r"\boptimism\b"],

    "trust":      [r"\btrust\b", r"\brely\b"],

}

def detect\_tag\_regex(text: str) -> str | None:

    low = text.lower()

    for tag, patterns in EMOTION\_PATTERNS.items():

        for pat in patterns:

            if re.search(pat, low):

                return tag

    return None

# Prepare local model for fallback fuzzy detection

local\_model = SentenceTransformer("all-MiniLM-L6-v2")

tag\_list = list(EMOTION\_PATTERNS.keys())

tag\_embs = local\_model.encode(tag\_list, convert\_to\_numpy=True)

def detect\_tag\_fuzzy(text: str, threshold: float = 0.4) -> str | None:

    vec = local\_model.encode([text], convert\_to\_numpy=True)

    sims = cosine\_similarity(vec, tag\_embs)[0]

    best\_idx, best\_score = int(np.argmax(sims)), float(np.max(sims))

    return tag\_list[best\_idx] if best\_score >= threshold else None

# OpenAI fallback

openai\_client = OpenAI(api\_key=OPENAI\_API\_KEY)

def detect\_tag\_openai(text: str) -> list[str]:

    prompt = f"""

Pick all relevant topics (comma-separated, lowercase) from this list:

{', '.join(tag\_list)}

User message: "{text}"

Tags:"""

    resp = openai\_client.chat.completions.create(

        model="gpt-4.1",

        messages=[{"role":"user", "content": prompt}],

        temperature=0

    )

    return [t.strip() for t in resp.choices[0].message.content.split(",") if t.strip()]

def detect\_tag(text: str) -> list[str]:

    tag = detect\_tag\_regex(text) or detect\_tag\_fuzzy(text)

    return [tag] if tag else detect\_tag\_openai(text)

# ─── 5) Softening input ─────────────────────────────────────────────────────────

def soften\_input(text: str) -> str:

    reps = {

        "hopeless": "emotionally drained",

        "depressed": "feeling low",

        "suicidal": "very overwhelmed",

        "worthless": "struggling with self-worth",

        "useless": "feeling unmotivated",

        "anxious": "feeling nervous",

        "panic": "feeling very stressed",

        "lonely": "feeling isolated",

        "sad": "feeling heavy-hearted",

    }

    for w, r in reps.items():

        text = re.sub(w, r, text, flags=re.IGNORECASE)

    return text

# ─── 6) Retrieve Quran verses ─────────────────────────────────────────────────

def retrieve\_ayahs(query: str, tags: list[str] | None = None, n\_results: int = 5) -> str:

    result = collection.query(query\_texts=[query], n\_results=n\_results \* 2)

    docs, metas = result["documents"][0], result["metadatas"][0]

    pairs = list(zip(docs, metas))

    if tags:

        pairs = [

            (d, m) for d, m in pairs

            if any(t in verse\_tags.get(m["verse\_id"], set()) for t in tags)

        ]

    pairs = pairs[:n\_results]

    if not pairs:

        return ""

    return "\n".join(f"{m['verse\_id']} ({m['surah']}): {d}" for d, m in pairs)

# ─── 7) Retrieve Hadiths ───────────────────────────────────────────────────────

HADITH\_URL = "https://hadithapi.com/api/hadiths"

def retrieve\_hadiths(query: str, n\_results: int = 1) -> str:

    params = {

        "apiKey":        os.getenv("HADITH\_API\_KEY"),

        "hadithEnglish": query,

        "status":        "Sahih",

        "paginate":      n\_results

    }

    resp = requests.get(HADITH\_URL, params=params, timeout=10)

    data = resp.json()

    # Normalize into a list of hadith dicts

    if isinstance(data, list):

        hadiths = data

    elif isinstance(data, dict):

        if isinstance(data.get("data"), list):

            hadiths = data["data"]

        elif isinstance(data.get("hadiths"), list):

            hadiths = data["hadiths"]

        else:

            hadiths = []

    else:

        hadiths = []

    # Build output strings

    out = []

    for h in hadiths[:n\_results]:

        text = h.get("hadithEnglish", "").strip()

        # book may be dict or string

        book = ""

        b = h.get("book")

        if isinstance(b, dict):

            book = b.get("nameEnglish") or b.get("slug", "")

        elif isinstance(b, str):

            book = b

        num = h.get("hadithNumber", "")

        ref = f"{book} (Hadith {num})" if book and num else ""

        if text:

            out.append(f"“{text}”{(' — ' + ref) if ref else ''}")

    return "\n".join(out)

# ─── 8) Main + follow-up chains ────────────────────────────────────────────────

prompt = PromptTemplate(

    input\_variables=["history", "context", "hadith\_context", "question"],

    template="""

You are “ImanTherapist,” an AI trained in Cognitive Behavioral Therapy (CBT)

and well-versed in the Qur’an and authentic Ḥadīth. You help users identify

their thoughts, challenge cognitive distortions, guide them toward Islamic

spiritual resources, and suggest practical cognitive and behavioral exercises.

Always:

1. Reflect their feelings empathetically.

2. Point out common CBT patterns (e.g. “That feels like all-or-nothing thinking”).

3. Offer a Qurʾān verse or Ḥadīth that addresses the concern, with reference.

4. Give a gentle spiritual & CBT-aligned action step.

Conversation so far:

{history}

Relevant Quran Verses:

{context}

Relevant Hadith:

{hadith\_context}

User's message:

{question}

Your reply (2-7 sentences, empathetic, Quran‐and‐Hadith‐based):

"""

)

main\_chain = LLMChain(llm=llm, prompt=prompt, memory=memory)

followup\_prompt = PromptTemplate(

    input\_variables=["user\_input", "bot\_response"],

    template="""

Based on the user’s last message and the bot’s reply, ask one brief, open-ended CBT-style question to deepen reflection.

User said: {user\_input}

Bot replied: {bot\_response}

Question:

"""

)

followup\_chain = LLMChain(llm=llm, prompt=followup\_prompt)

closers = [

    "How does that feel for you?",

    "Does that help you reflect further?",

    "What do you notice as you think about this?"

]

# ─── 9) Chat loop with catch & regenerate ─────────────────────────────────────

FALLBACK\_SIG = "unable to provide the help that you need"

def is\_greeting(text: str) -> bool:

    return bool(re.search(r"\b(hi|hello|salam|assalamu alaikum|hey)\b", text.lower()))

if \_\_name\_\_ == "\_\_main\_\_":

    print("CBT + Quran & Hadith Chatbot (type 'exit' to quit)\n")

    while True:

        user = input("You: ").strip()

        if user.lower() in ("exit", "quit"):

            print("Bot: Goodbye. May Allah grant you ease.")

            break

        if is\_greeting(user):

            print("Bot: Wa alaikum assalam! How are you feeling today?")

            continue

        safe = soften\_input(user)

        tags = detect\_tag(safe)

        ctx  = retrieve\_ayahs(safe, tags=tags, n\_results=3)

        # pick first tag or fallback to full input

        hquery = tags[0] if tags else safe

        hctx   = retrieve\_hadiths(hquery, n\_results=1)

        reply = main\_chain.predict(

            context=ctx,

            hadith\_context=hctx,

            question=safe

        )

        if FALLBACK\_SIG in reply.lower():

            hist = memory.load\_memory\_variables({})["history"]

            override\_sys = (

                "You are a compassionate CBT therapist and Islamic counselor. "

                "Do NOT use any generic crisis disclaimer. "

                "Always provide supportive, Quran‐and‐Hadith‐based CBT advice."

            )

            override\_usr = (

                f"Conversation so far:\n{hist}\n\n"

                f"Relevant Quran Verses:\n{ctx}\n\n"

                f"Relevant Hadith:\n{hctx}\n\n"

                f"User's message:\n{safe}\n\n"

                "Your reply (2–3 sentences):"

            )

            msgs = [

                SystemMessage(content=override\_sys),

                HumanMessage(content=override\_usr)

            ]

            reply = llm.invoke(msgs).content

        print(f"Bot: {reply}")

        time.sleep(0.5)

        # follow-up

        follow = followup\_chain.predict(user\_input=user, bot\_response=reply)

        print("Bot:", choice(closers), follow, "\n")