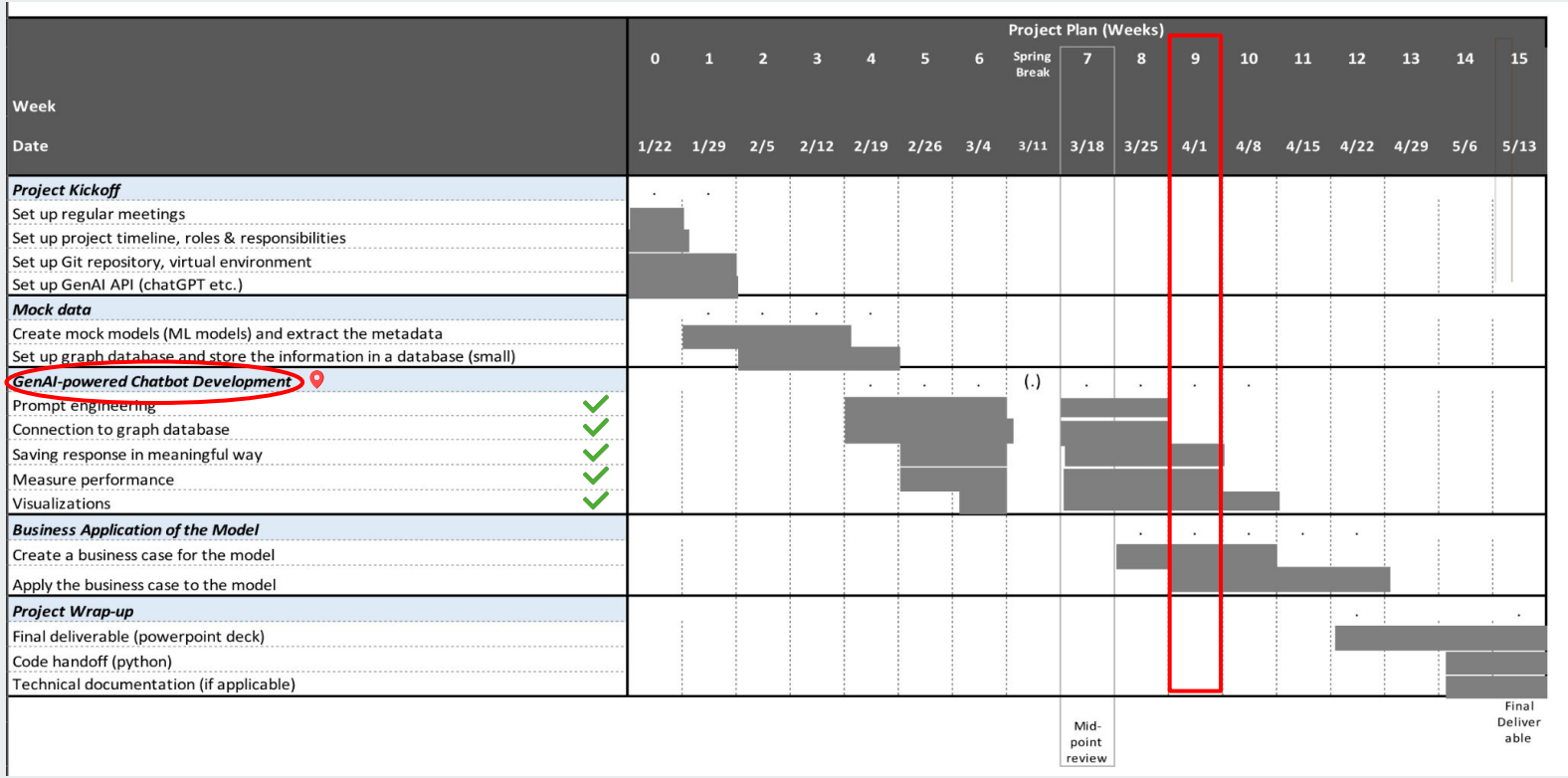
KPMG Capstone Weekly Update

2024.04.03

Schedule


Project Plan (Weeks)
Week
Date
0 1 2 3 4 5 6 Spring Break 7 8 9 10 11 12 13 14 15
1/22 1/29 2/5 2/12 2/19 2/26 3/4 3/11 3/18 3/25 4/1 4/8 4/15 4/22 4/29 5/6 5/13
Project Kickoff
Set up regular meetings
Set up project timeline, roles & responsibilities
Set up Git repository, virtual environment
Set up GenAI API (chatGPT etc.)
Mock data
Create mock models (ML models) and extract the metadata
Set up graph database and store the information in a database (small)
GenAI-powered Chatbot Development
Prompt engineering
Connection to graph database
Saving response in meaningful way
Measure performance
Visualizations
Business Application of the Model
Create a business case for the model
Apply the business case to the model
Project Wrap-up
Final deliverable (powerpoint deck)
Code handoff (python)
Technical documentation (if applicable)
Mid-point review
Final Deliverable

Update

1. Successfully implemented the 3-layer LLM model in StreamLit

2. UI/UX

    2.1 Instructions on prompts
    2.2 Visualizations on the graph database

https://capstone-kpmg-chatbot.streamlit.app/

3. Automated the Testing Pipeline

Example Prompts

Give me the summary of the section Key Takeaways.

- Our model: correct answer
- Base model: AI hallucination

Does the model_925814632 and model_782951116 use the database sales?

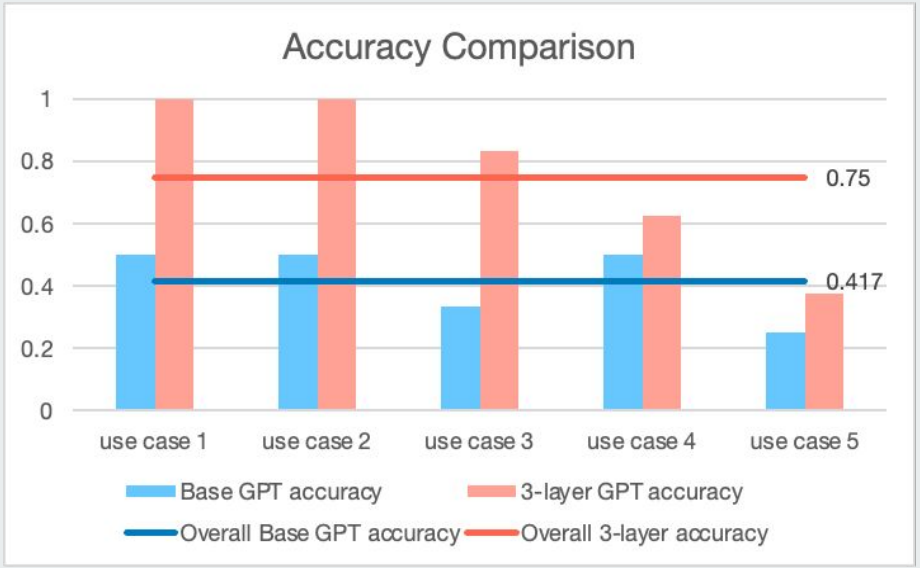
- Our model: need to implement follow up
- Base model: AI hallucination

Which models are predicting risk levels?

- Our model: correct answer
- Base model: incorrect answer

Automatic Testing Pipeline

1. Accuracy



2. Hugging Face: bge-m3

Future Plan

1. back-and-forth improvement, such as
    a. chatbot questioning (I don't understand → I don't understand, do you mean xxx?
    b. previous question references (Q: Can you provide information for **the previous model**?)
    c. question synonym or misspelling
        (To answer question like "who is the author/owner/maintainer of the model?")

2. Bad Performance on Use Case 4 & 5
    - Add more train cases?

3. Follow up questions

4. Visualization Tool on generated results