EE/CprE/SE 492 BI-WEEKLY REPORT 2

01/31/24 – 02/07/24

Group number: 49

Project title: Using Generative AI to Assess Learning

Client &/Advisor: Dr. Henry Duwe

Team Members/Role:
Alex Vongphandy - Backend and Prompt Developer
Akpobari - Backend and Prompt Development
Abram Demo - Frontend Developer
Drake Rippey - Frontend and Backend Developer

Weekly Summary

The overall objective of this week for the backend and prompt engineering team was to devise a solution to craft a coherent conversation between human ("student") and AI. The backend team also focused on selecting an Agent to handle these interactions along with crafting prompts to ensure workflow logic between the different tools (functions that the AI can call). We tested first with OpenAI Functions Agent and Pydantic Models but we did not receive the outcome we hoped for and there were numerous issues. We switched to a ReAct (Reasoning + Act) agent which allowed the AI to think first before making an action. This has given us better results and we are moving forward with this Agent. We modified our original prompt templates to reflect Bloom's Taxonomy, per client's request. We can also now trace all our AI and LLM prompts and calls in LangSmith which allows us to see visually the workflow and logic behind each AI output. The end-to-end connection team was focused on building the framework for the rest of the semester, this includes a working database that responds to get and post requests, and an LTI launch page that establishes a connection with Canvas. This work also included creating some of the first pages for the site that won't be run using chainlit. For now, this is just a webpage with a basic quiz for database testing purposes.

Past week accomplishments
- Agent Selection - Alex
  - Prototyped OpenAI Functions Agent vs ReAct Agent. Determined that ReAct Agent was much more responsive and usable.

- One example of why OpenAI Function Agent was not selected is the Agent would first formulate a follow-up question but analyze the student's response after. This logic was incorrect.
  - Crafted a base template for ReAct Agent to use to allow it to process student input.
- Agent Tools Engineering - Alex and Akpobari
  - Created originally one tool that analyzed and crafted a follow up question using a Pydantic model (a JSON schema)
    - Did not work well, explored not using Pydantic or a structured model and instead let the Agent do the processing rather than manually.
    - Split the one tool into two tools for separation and so that the Agent could dynamically select which to use.
    - Crafted the prompt templates required for the tools, one using Bloom's Taxonomy.
- Backend Troubleshooting - Alex
  - Langchain updated their whole repository which broke all our imports/packages.
    - Manually went through and found out which packages were not compatible and updated that.
- CI/CD Pipeline - Alex and Akpobari
  - Manually set a Git Runner on our local server to build and deploy our GitLab project so we can all work on it at the same time.
    - Was not working as intended as we are finding difficulties uploading our docker image into GitLab so we can develop tests on the backend
    - Prototyped Docker and was able to run our project in a container locally but not hosted on our server
- LTI Integration - Abram
  - Manually run a js based server that integrates with canvas and can appear as an LTI; this will allows us to get student information on the backend
    - This requires an HTTPS connection, which is more secure than HTTP and requires a little setup
    - There are a couple different security options to launch an LTI with, they include public and anonymous and we have not yet decided what we will do
- HTTPS Connection - Abram
  - Integrate an SSL certificate with the current program so that we can run Chainlit on HTTPS, was not able to get is running on VM, but was working locally
    - Canvas rejects all HTTP connections, so this is a mandatory step
    - HTTPS and chainlit seem to be competing for port control; information does seem to be sending, it's not helpful in any way
- Develop Frontend Pages - Abram and Drake
  - Develop the landing pages that professors will see when they are configuring an assessment to give to students

- ■ Some pages will include a sign up, log in, LTI configuration tutorial and assessment creation
- Python Canvas API - Abram
  - Understand and utilize the Python Canvas API
  - Successfully created an assignment as well as submitted a pdf as a submission for said assignment
    - ■ Based on how the LTI integration has been going, this may be a dead end and not be useful in the final product, but not sure yet.
- Database Remote Connection - Drake
  - Established all connections required to make MySQL database connectable from devices outside of the virtual machine
  - Created an external User to know when edits are being made by the external user or by someone on the VM
- Python Database Connection - Drake
  - Installed mysql-python connector that allows the database to be access and edited via our python program
  - Made a test case that shows the working connections with some hard coded variables

Pending issues

- ReAct Agent using some tools incorrectly - Alex
- ReAct Agent spams tools and processes them erroneously. - Alex, Akpobari
- After answering the correct answer, the conversation continues with the follow up question due to Pydantic boolean value not changing from false to true. - Akpobari, Alex
- Trouble setting up a SSL Certificate to run a HTTPS server on the backend that also runs on the same ports as Chainlit - Abram
- Unable to get the chainlit program to redirect to a new page on quiz "completion." - Drake

Individual contributions

NAME	Individual Contributions (Quick list of contributions. This should be short.)	Hours this week	HOURS cumulative
Alex Vongphandy	Selecting Agent, Prompt Engineering, Troubleshooting	8	24
Akpobari Godpower	Agent prompt engineering, Pydantic value prompt engineering,CI/CD	7	23
Abram Demo	LTI Launch, HTTPS configuration, frontend webpages	10	24
Drake Rippey	Database Remote Connection, Database Python Connection, Database Test Cases, Chainlit Redirecting/Styling	8	20

Plans for the upcoming week

- Agent - Alex
    - Need to verify if correct answer is being applied/provided to Agent
        - We manually have to input correct answer due to limitations found with GPT 3.5 and MIPS
        - Client recommended GPT-4 which we will test
    - Limit number of follow-up questions to 5 per client recommendation
    - Adjust prompt template so it responds more like an instructor instead of AI
- Agent Tools - Alex, Akpobari
    - Adjust prompt templates so that Agent uses the correct tool and follows workflow.
    - Create a Question Generation Tool for MIPS questions
- Memory - Alex, Akpobari
    - Need to add memory to our ReAct Agent so that it can remember the whole conversation and use that as context.
        - Need to figure out the limits/token limit of using this memory in our prompt.
- CI/CD - Akpobari
    - Deploy docker image of backend into GitLab repository
    - Create infrastructure to run tests
- HTTPS - Abram
    - get chainlit and https running together so that we can host LTI launch pages on our VM
- LTI submissions - Abram
    - Be able to submit a grade for a student after an assessment is completed
        - This can't be done using an API key
    - Submission includes a numeric score as well as a pdf (or something similar) of the chat history with the bot
- Frontend pages - Abram, Drake
    - Create more pages to keep building the frontend
        - Sign up, Login
        - LTI configuration tutorial and assessment configuration
    - Ability to dynamically create pages given information about an assessment
- Chainlit style - Drake
    - Customizing the chainlit pages to more fit our needs
    - Create custom pages that work going into and from a chainlit instance
- Database Integration - Drake
    - Merging all required database attributes to main branch
    - Test again to ensure it works.