CS510 - Project - Development Track

Team Member: Palveet Kaur Saluja(psaluja2@illinois.edu), Rudrik Sanmukhlal Patel(rudrikp2@illinois.edu)

Co-ordinator: Palveet

Title: CodeClash : LLM Code Generator and Comparator

Functions and Users

We are planning to develop a tool that takes input of a descriptive coding problem in natural language from users and uses LLMs like ChatGPT and Gemini to generate the code for the problem. It gives users a chance to edit and run the code in the same screen and evaluates the run time and efficiency of the code generated by the two LLMs. The Users are going to be anybody who wants to learn coding or knows coding already.

Significance

Right now users have to go to each LLM to get the code and the code's are uneditable on the LLM window. This tool will save time and will give a comparison between the codes generated by ChatGPT and Gemini. This will be a code repository of frequently asked coding problems and their correct solutions. Also, we can use this repository to analyze which LLM actually gives accurate results for a given problem.

Approach

We are using APIs from OpenAI and Google for prompting, Flask for backend, React and Tailwind for frontend. Also we have provided support for code compilation in 40+ languages using Vscode's editor called 'Monaco'. This gives a neat visualization and code edit option. In the UI we give users fields such as question description, input & output constraints etc. in order to create strong and detailed prompts so that the LLM provides accurate code.

Evaluation

Our evaluation will be based on code correctness and efficiency of the code. The evaluation will be done by the users and their satisfaction of the generated code by the two LLMs. Since we are giving an editable window for the user to edit the code generated by the LLM they can edit it there itself as per their liking.

Timeline

We have 2 weeks in hand right now:

Week 1: Understanding the API get and send request architecture and figuring out other requirements. Starting the frontend and backend development.
Week 2: Completing the implementation and user testing. Fix any bugs that are found.

Task Division:
Backend: Rudrik
Frontend: Palveet
Since we are just two people in the group, we are putting efforts in both the areas