

Individual insight report - Magnus

Limit questions, not only total prompt length when searching through a long document

I tried to extract useful information from a 500-page contract using 50 questions. If I applied all 50 questions to a subset of pages, it would answer well, but only for the first 4-5 questions and cut off the rest of the answers. Therefore, I not only had to split the pages but also questions into subsets. In this task it was also difficult to get the LLM to provide a blank output. It was easier if I just told it to output a very specific sentence when it did not find the answer like "No answer found" and then later go through and delete these entries.

GPT 3.5 VS 4 cost

I didn't appreciate the difference until I switched to 4. After using 3.5 for several weeks of debugging at almost no cost, I switched to 4 and blew through almost my entire budget in an hour.

Changing APIs

We used Llama_index and it was updated after we built it into our application. Had to change several lines of code. Could be an issue if in a real product.

Indent prompts

Can be very useful to create a hierarchy that the LLM can follow, putting overall considerations first and indenting details.

Use 3.5 to draw out errors in your prompt – Switch to 4 if you need better performance

If you can make it work with 3.5, you can probably make it work with 4 and its much less expensive. So use 3.5 to understand which errors the LLM is likely to make and mitigate them. The errors may just be a result of a prompt that is not correctly formatted. GPT3.5 may not be able to see past the errors as well and may force you to create a more robust prompt, also reducing the chance that GPT4 will make those errors.

The basics – Use branches

Its been really useful to learn how to use tools like github VS code, how to work with libraries and learning to debug. I had very limited programming experience coming in so it was good to work with it in practice. By the end of the project we learned that we should extensively use branches and use GitHub in a smarter way than just pushing and pulling on the main branch.

Leading a team – Clear goals

This was especially hard when I was the one on the project who knew the least about coding. Didn't go smoothly in the start as i had unrealistic expectations about how fast we could do things did not realize how detailed plans have to be for everyone to have the same understanding. Don't have too high expectations in the sprint. Have a very clear goal going into the sprint and make sure everyone is on the same page.