LLM-Powered Data Dashboard

Sankalp Singh, Brandon Perez, Javid Mardanov, Muzhgan Aghazada, Niveditha Nerella, Laura Mo, Sai Lekkala, Kevin Patel, Shreyas Vendantham, Mayukha Reddivari, Rahil Mehta, Aiden Tian, Shubhro Biswas, Mihir Singh, Heng-I Chu, Arjun Khandelwal
TA: Siddharth Sancheti

INTRODUCTION

BACKGROUND

- Caterpillar equipment contains hundreds of sensors, producing large amounts of data
- Currently, engineers need to manually search through directories and then conduct data analysis on their own

PROJECT GOALS

- Create a data dashboard with an integrated chatbot, greatly reducing the time it takes to conduct data analysis tasks
- UI: ensure the dashboard is intuitive and user-friendly
- Ensure that the model can handle the conversion of GH5 files to CSV format

RESEARCH METHODOLOGY

Analyze Machine data

- Worked with Caterpillar machine data by analyzing payloads and familiarizing with time-series data

Explore LAMBDA

- Work with the open-sourced LAMBDA agent, adapt it to Caterpillar data format and features

Design and Integrate UI

- Built a Figma prototype and collected user feedback, created front-end UI with React.js, and combined with back-end

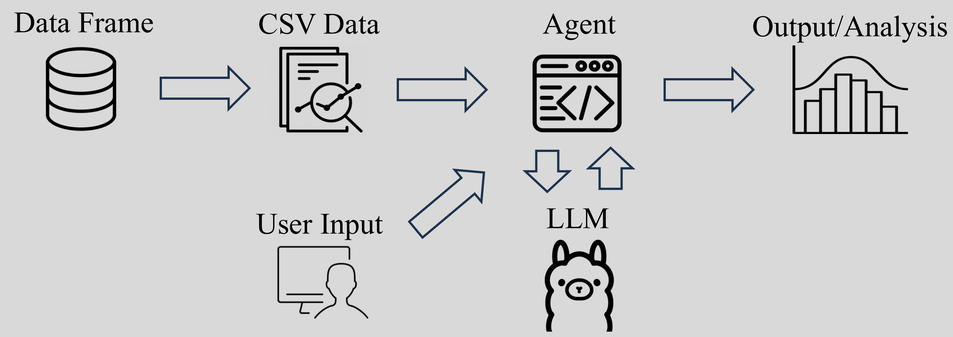
Fine-tune

- Prompt engineer and script the agent to familiarize it with Caterpillar jargon and needs

Testing

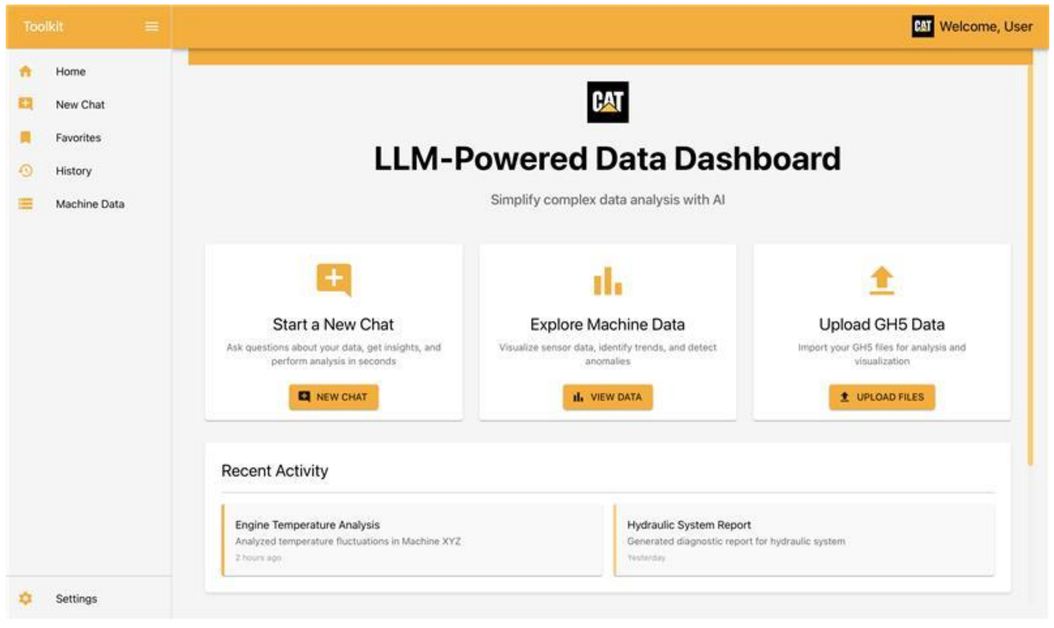
- Perform accuracy testing for LLM hallucinations, add any preventative measures or further prompt engineer to improve accuracy to 90%

WORKFLOW



UI AND FEATURES

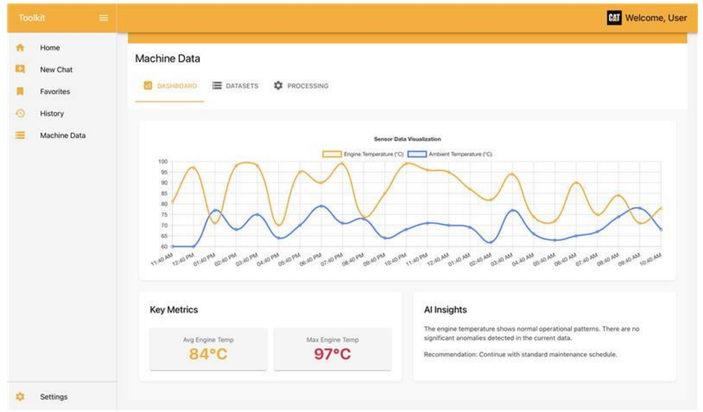
- **User Flexibility**: Ability to switch between LLMs, color themes, show/don't show code snippets, collapsing menus, comprehensive settings
- **Uploading/Accessing Data**: User can upload HDF5 data directly or access machine data through Caterpillar's database system
- **History**: Previous chat history and favorite chats
- **Instant Data Analysis**: Uploading to the agent's data system generates graphs, metrics, and insights



BEHIND THE PROGRAM

Programmer Agent: Generates code based on user input and knowledge base
Inspector Agent: Debugs programmer agent's code
Operational Loop: The two agents enter a feedback loop between each other until code is error-free, it hits the maximum attempts, or the user intervenes.

DATA DASHBOARD



CONCLUSION

- Improved efficiency of data analysis tasks for engineers
- Retrieved, processed, and visualized machine data
- Effectively summarize and answer employees' prompts

FUTURE GOALS

- Refining overall accuracy of the LLM's output
- Scaling data dashboard to other types of machine data

ACKNOWLEDGEMENTS

Thank you to Caterpillar, especially our mentors Jim Katter and Neo Mohsenvand, for your support and guidance throughout this project.
Thank you to the Data Mine staff and Dr. Ward, as well as our TAs: Siddharth Sancheti, Christina Joslin, Rayan Singh.
REFERENCES: LAMBDA — *A Large Model Based Data Agent*, https://www.polyu.edu.hk/ama/cmfai/lambda.html