EE/CprE/SE 491 WEEKLY REPORT 7

10/24/2024 – 10/31/2024

Group number: 35

Project title: Universal Response Engine: LLMs for Good

Client &/Advisor: Ahmed Nazar and Mohamed Selim

Team Members/Role:

Abrahim Toutoungi - Stakeholder Liaison

Gabriel Carlson - Communications Manager

Halle Northway - Meeting Coordinator

Brianna Norman - Project Deliverables Manager

Ellery Sabado - Timeline Coordinator

Emma Zatkalik - Assignment Manager

Weekly Summary
We acquired the needed Virtual Machine resources and could start setting it up for our project. This week continued work on generating datasets and refining our fine-tuning methods.

Past Week accomplishments
- Continuing to experiment and explore with fine-tuning LLMs
- Cultivating a well-made collection of datasets

Pending Issues

- Compile group member's code into best-functioning for RAGs and fine-tuned model

Individual Contributions

Name	Individual Contributions	Hours this week	Hours cumulative
Abrahim Toutoungi	- Worked on generating and cleaning up a dataset - Accessed vm - Continued to work with RAG and fine-tuning - Checked out llama 3.2 - Looked at Emma's Google Colab	6	38
Gabriel Carlson	- Continued testing datasets for unsloth llama3 fine tuning - Cleaned up dataset used for training - Tested fine-tuning with different dataset	6	37

	- chunks, lengths, epochs, and batch size - Added test LLM to VM		
Halle Northway	- Accessed VM - Reviewing and rewriting fine-tuning LLM code - Reviewed datasets, began generating and creating a combined dataset	6	39
Brianna Norman	- Worked on setting up needed VM software - Reviewed some team member's code to determine best for use - Worked on generating and cleaning up a dataset	6	38
Ellery Sabado	- Collected more datasets and reformatting for Q&A format - Have a working Fine-Tuning model and tested JSON first aid dataset to the model and supervised the correct response (Ollama3.2:3B)	7	38

Emma Zatkalik	- Using Google Colab to finetune both llama3.1 8B and llama3.2 3B
	   - Uploaded those models to my huggingface account
	   - Uploaded the training data to weights and biases
	   - Used the mental health counseling dataset from huggingface
	- Accessed the VM via remote desktop
		6	39

Comments and extended discussion (optional)
N/A

Plans for upcoming week
- Combining Datasets for RAG and fine-tuning
- Continue setting up LLMs running locally on the VM
- Continue developing fine-tuned models with datasets
- Work on creating one unified LLM chain for VM
- Continue collecting performance data for running LLMs on student computers and collaborative vm

Summary of weekly advisor meeting

Did not have a meeting this week.