Deployment and usage of Large Language Models in Small Tech Companies

STS Research Problem: How are schools currently using LLMs in teaching and what kind of experiments are they running?

A Thesis Prospectus

In STS 4500

Presented to

The Faculty of the

School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science in Computer Science

By

Jason Ton

October 27, 2023

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Kent Wayland, Department of Engineering and Society

Briana Morrison, Department of Computer Science

Introduction

A Large Language Model (LLM), refers to a type of artificial intelligence model that is designed to understand, generate, and interact with human language at a large scale. These models are typically trained on vast amounts of text data. These are great tools when used appropriately and honestly and can assist with many different tasks such as writing, explanation, etc. Many people and organizations are questioning the effect of ChatGPT in school environments and how it may affect academic performance and integrity. Hopefully here we will be able to separate the hype from the reality and take a look at how we can use these LLMs to accelerate our society's development. This issue is important because if used correctly, it can transform the way we work, learn, and overall function as humans. However, if not, it can devastatingly impact many of our sociotechnical systems.

For my technical research problem, I will look at the effect that LLMs had at my internship over the summer LinQuest. Another intern and I were tasked with figuring out whether an LLM could be deployed at Linquest's Lab. However, we ran into many obstacles, and I will discuss more about the deployment of a LLM and the obstacles that come with that in my technical section.

For my STS problem I will look into how schools are currently using LLMs in teaching and what kind of experiments they are running. When looking at the STS problem, I will also look into differences between how schools are handling the LLM situation. It's clear that every school will have a different reaction and I will see what different schools do and why they

decided to handle it the way they did. Mutual shaping is the most important part of my research. I will be trying to see how these faculty are planning to change the landscape of the education system with LLMs.

Technical Research Problem: Deployment and usage of Large Language Models in Small Tech Companies

ChatGPT is an amazing tool, it has the ability to enhance software developers and engineer's work and productivity. The company I interned at, Linquest, wanted to have an LLM assistance tool to assist the Machine Learning team at Linquest, and to provide general support for software engineers. However, for several companies, like LinQuest, one major drawback of ChatGPT is that its creator, OpenAI, may collect and record chat information. Linquest specifically was a company that worked with proprietary satellite and software solutions so if that fell into the hands of another company, Linquest's information would be practically worthless. There are also many other concerns with using a 3rd party tool which can be found in the Ethical Regulatory Framework (agency, accountability, privacy, and transparency) (Piñeiro-Martín 2023) which is a framework software engineers strive to follow when developing new technology. So, my team was tasked with figuring out if there was a large-scale LLM that could be deployed given our resources so that we could keep our data proprietary while also giving us functionality similar to ChatGPT. We also wanted to use this tool for tasks such as data analytics or aggregation. LLMs are projected to have the ability to partially automate up to around ⅔ of professions (Mearian 2023), so its important businesses take advantage of these technologies too. This matters, as we want to be prepared for the future and make the most of these state-of-the-art technologies.

To start off our project, my partner and I had to look into what kind of resources it would take in order to deploy an LLM. The LLM we wanted to deploy was called BLOOM. BLOOM is very similar to ChatGPT, in the way that it is supposed to reply to human prompts and generate responses. What differentiates it is if we can deploy it on our own servers all of the data, we send it stays in the lab and it stays proprietary, solving the privacy issue. We took notes of what we currently have available to us at the LinQuest lab. My partner and I talked to the lab managers and service team about our specs and how much we had to work with and then continued based off of that. We researched the BLOOM model and learned that our lab resources won't be able to handle the deployment of BLOOM due to lack of computing resources. As a result, we had to turn to smaller versions of the BLOOM model, or we had to focus our resources on finding different ways to centralize and automate our data and this led to the focus on developing a different app for data analytics. We learn from this that smaller tech companies struggle with using large language models due to conflicts between data privacy and computing resources.

After realizing that deploying the BLOOM model was not feasible due to resource constraints, my team shifted its focus towards finding alternative solutions. This shift in focus towards centralizing and automating data and developing a different app for data analytics is a pragmatic approach. It highlights the importance of adapting technology to fit the specific needs and limitations of your organization, especially when dealing with emerging technologies like LLMs.

I will contact old companies and reach out to colleagues who work in various companies and ask them about how their integration of LLMs is going. By doing this I hope to gain valuable insights into how different companies are approaching the integration of Large Language Models (LLMs) and how they are managing the challenges related to data privacy and computing resources.

STS Research Problem: How are schools currently using LLMs in teaching and what kind of experiments are they running?

The introduction of Large Language Models (LLMs) is poised to significantly transform the education system, involving extensive experimentation. This shift is like the ongoing integration of robots, where humans continue to maximize benefits while mitigating impacts on quality of life (Baber 2020). In my STS project I am aiming to investigate how professors are experimenting with LLMs in the education scene and what kind of impacts LLM technology is having in classrooms. This topic is important and already shows mutual shaping as we can already see some of these effects taking place around us. For instance, schools at various educational levels are beginning to employ LLMs as personal tutors, a clear indication of their growing role in shaping the educational experience (Extance, 2023). This marks a significant shift in the educational landscape. Exploring how the widespread adoption of LLMs might reshape future learning environments is crucial, particularly in how it may redefine the skills students need for success in a digital world.

LLMs are increasingly being integrated into educational settings, reflecting a significant shift in teaching methodologies. Educators are finding innovative ways to utilize LLMs, such as providing alternative explanations for complex concepts, creating additional examples and practice questions, and assisting in the structuring of courses (Joshi et al). Additionally these models can assist with administrative tasks, including grading and performance evaluation (Manhiça 2023), and their ability to learn from user data makes them adaptable tools capable of catering to the individual needs of students (Gifford 2023).

It's important to remember the other parties involved in this sociotechnical system. Administrative teams must pick a sufficient model, ChatGPT has a lot of hype however there are many different models (Bard, Bing, etc) to pick from with different results in different fields such as math, writing, etc (Amita et al, 2023). This is not only a problem for faculty but also the model designers themselves, they need to pick a model that understands the requests of the education system, making questions that will challenge students, specification for certain fields (STEM, writing, etc) (Baierl 2023). Furthermore, the development of models specifically for education, such as EduChat, demonstrates the potential for customized LLMs to enhance learning experiences (Dan et al, 2023).

However, the integration of LLMs provides many problems. We currently have models with very high writing analyzation accuracy with training data, but poor accuracy with actual human writing, so it's their job to pick a sufficient detector (Michael 2023). As a result, many experiments will have to be conducted to figure out what model works the best in which environment, trial and error seems to be a big theme in my sources so far. Other obstacles

include LLM access, algorithmic bias, and plagiarism (Meyer, 2023). While these are not the focus of my paper, I feel it's important to mention them as they are a big part of the existing sociotechnical system and will arise as schools are conducting experiments with LLM technology.

There will be many actors involved in this socio-technical problem such as students, professors, media, etc. My primary focus will be on smaller groups that will have a big role in facilitating LLM functionality. Model designers will be responsible for creating different models depending on a school's needs. Faculty must create experiments and record feedback of their students to tweak and improve future experiments. IT teams will also be a major player as they must communicate with the professors and the model providers to facilitate connections for the experiments. These are the main actors I want to focus on during my research.

To collect evidence to learn more about my research topic, I have many different data collection methods. I will be interviewing people that I have connections with at different universities and schools. I hope to be able to include a diverse range of interviewees (colleagues, educators, administrators) when conducting my interviews. I want to learn about a few main things, how teachers will be using LLMs to enhance teaching, what kind of results have they garnered, and what they have planned for future semesters. I plan to review papers and reports related to my STS topic, aiming to diversify my perspective and compare findings with insights from my interviews. These papers and reports will hopefully be able to provide me with more data and variety than I can obtain with just my interviews. Media sources such as press releases

and news media are currently heavily covering AI topics, so I feel they will be worth at least looking into.

Conclusion

In conclusion, the introduction of Large Language Models (LLMs) has led to an era filled with both excitement and uncertainty. This prospectus delves into two significant research problems related to the integration of LLMs: one focusing on their deployment and usage in small tech companies and the other exploring their impact on schools and universities. For my technical report, I will look into how companies are struggling with the integration of LLMs and its drawbacks. For my STS research problem, I will try to see how schools are trying to implement and experiment with these new technologies. As we look ahead, the integration of LLMs into our daily lives, whether in the workplace or the classroom, presents numerous opportunities and challenges.

References:

Amita, K., Kumari, A., Singh, A., Singh, S. K., Ayesha, J., Dhanvijay Anup Kumar, D., Pinjar, M. J., & Himel, M. (2023). Large Language Models in Hematology Case Solving: A Comparative Study of ChatGPT-3.5, Google Bard, and Microsoft Bing. *Cureus*, 15(8). https://doi.org/10.7759/cureus.43861

Baber, T. (2020). 2020 ASME Human-Powered Vehicle Competition University of Virginia Orange Team; Robots and Society: Robots Influence on Manufacturing. [Bachelor of Science, University of Virginia, School of Engineering and Applied Science]. *University of Virginia*. https://doi.org/10.18130/v3-6300-2n19

Baierl, J. D. (2023). Applications of Large Language Models in Education: Literature Review and Case Study (Order No. 30528143). *ProQuest Dissertations & Theses Global*. https://escholarship.org/uc/item/6kf0r28s

Dan, Y., Lei, Z., Gu, Y., Li, Y., Yin, J., Lin, J., Ye, L., Tie, Z., Zhou, Y., Wang, Y., Zhou, A., Zhou, Z., Chen, Q., Zhou, J., He, L., & Qiu, X. (2023). EduChat: A Large-Scale Language Model-based Chatbot System for Intelligent Education. Cornell University Library, arXiv.org. https://arxiv.org/abs/2308.02773

Extance, A. (2023, November 15). ChatGPT has entered the classroom: how LLMs could transform education. *Nature*.
Retrived from https://www.nature.com/articles/d41586-023-03507-3

Gifford, A. (2023, October 11). Generative AI hits education, ushering in a sea change for Schools. *GovTech*.
Retrived from https://www.govtech.com/cde/generative-ai-hits-education-ushering-in-a-sea-change-for-schools

Joshi, I., Budhiraja, R., Pranav, D. T., Jain, L., Deshpande, M., Srivastava, A., Rallapalli, S., Akolekar, H. D., Challa, J. S., & Kumar, D. (2023). "With Great Power Comes Great Responsibility!": Student and Instructor Perspectives on the influence of LLMs on Undergraduate Engineering Education. Cornell University Library, arXiv.org. https://arxiv.org/abs/2309.10694

Manhiça, R., Santos, A., & Cravino, J. (2022). The use of artificial intelligence in learning management systems in the context of higher education: Systematic literature review. *CISTI (Iberian Conference on Information Systems & Technologies / Conferência Ibérica de Sistemas e Tecnologias de Informação) Proceedings*, 17. 10.23919/CISTI54924.2022.9820205

Mearian, L. (2023, June 28). Ai will kill these jobs (but create new ones, too). *Computerworld*. Retrived from https://www.computerworld.com/article/3700857/ai-will-kill-these-jobs-but-create-new-ones-too.html

Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., Karen O'Connor, Li, R., Pei-Chen, P., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: opportunities and challenges. *Biodata Mining*, 16. https://libkey.io/libraries/1691/10.1186/s13040-023-00339-9

Michael, S. O., Karnalim, O., Carlos, A. S., & Liut, M. (2023). Detecting LLM-Generated Text

in Computing Education: A Comparative Study for ChatGPT Cases. Cornell University Library, arXiv.org. https://arxiv.org/abs/2307.07411

Piñeiro-Martín, A., García-Mateo, C., Docío-Fernández, L., & María del Carmen López-Pérez. (2023). Ethical Challenges in the Development of Virtual Assistants Powered by Large Language Models. *Electronics*, 12(14). https://doi.org/10.3390/electronics12143170