# Independent Study Proposal

Fall 2023

Shamsul Haque

Professor Judy Fox

Topic: Big Data Systems Project

## Overview

Help the team with assistance on the new project, which is aimed at tracking the spread of COVID-19 and variables that influence its spread such as age. Also, I helped to create a course site for the course DS 5110: Big Data Systems. I will work on specifically making a website to help display data found through lectures on DS 5110: Big Data Systems in an engaging format. I will work to integrate hands-on learning activities into the project. Thus, the learning capabilities for students taking this course will be enhanced. I also will work to integrate hands-on learning activities into the project.

## Literature Review/Background

1. https://tddg.github.io/ds5110-spring23/lectures/
2. https://nmagee.github.io/ds2002/schedule/
3. https://microsoft.github.io/graspologic/latest/reference/release.html
4. https://jupyterbook.org/en/stable/start/your-first-book.html
5. TFT for Interpretable Multi-Horizon
6. http://rafalab.dfci.harvard.edu/dsbook/getting-started.html
7. https://github.com/tddg/ds5110-spring23/blob/main/assets/ray_API_demo.ipynb
8. https://drive.google.com/file/d/1sdca8xPRLTBumEhoE0OJJB6WvuQ94lWS/view

Located above are all of the reference sites and documents that I have been utilizing throughout the semester. If this is not what is being looked for here, I can provide better documentation.

## Hypotheses to be explored

1. How can we best curate a site that will reimagine ways for students to engage with the material they are learning for DS 5110: Big Data Systems?
    a. What types of ways should this data be structured and displayed?
    b. What activities should these students participate in to best grasp these concepts?
    c. How can the site be made easy to navigate?
2. Which age group has the most significant effect on the spread of COVID-19?

## Expected research results

1. Research with external resources and grasp Big Data Systems through reviewing the course content, which would inform how we design our site
2. To successfully design a web application for DS 5110: Big Data Systems, allowing for others to have a memorable user experience.
3. Submit by end of semester, with integrated ways for hands on learning through activities (AWS Academy, Jupyter Notebook)

# Methodology

## Data

The data mostly came from varying websites and sources on big data systems, primarily coming from courses taught by Professor Fox.

## Primary Goal

A site hosted on github pages that displays all of the necessary information on DS 5110: Big Data Systems, with integrated learning activities (i.e. Jupyter Notebook) and a streamlined design for others to use to increase their knowledge and comprehension of how Big Data Systems function.

## Sub-Goals and Experiments

1. Integrate learning activities through AWS classroom into the book site.
2. Through an interactive process and feedback approach, add elements and adjust the website along the way to appeal to Professor Fox's requirements.
3. Develop a plethora of web development skills along the way.
4. Become better at understanding how to create a seamless UI experience as I strive to pursue a career in that.
5. Construct the site to be simple and enhance the learning experience for students taking DS 5110: Big Data Systems..

# Assignments

## Effort

I will spend 5-10 hours per week for this independent study, as much as I would normally do for a 3 credit CS course.

# Schedule

| Week | Readings: | Code: | Paper Progress | Dr. Fox suggestions / Milestones |
|---|---|---|---|---|
| 1 | | | N/A | |
| 2 | https://pytorch.org/tutorials/beginner/basics/intro.html | Learning PyTorch as assigned by Khairul | N/A | Familiarized with PyTorch |
| 3 | https://pytorch.org/tutorials/beginner/basics/intro.html | Continued to finish PyTorch course | N/A | Made significant progress of PyTorch course |
| | | | | |
| 4 | https://arxiv.org/pdf/2104.00950 | Finished PyTorch course<br>Read paper suggested by Khairul | N/A | Prepared for paper reading on friday and finished PyTorch course |
| 5 | https://github.com/UVA-MLSys/COVID-19-age-groups | Familiarized with codebase and looked through poster submission | N/A | Understand context of poster and codebase |
| 6 | https://github.com/orgs/UVA-MLSys/projects/2 | Watched Luke Tutorial on Rivanna | N/A | Understand Rivanna |

| | | | | |
|---|---|---|---|---|
| | | Attempt at Rivanna Replication | | |
| 7 | https://github.com/orgs/UVA-MLSys/teams/bigdata | Learned CSS and HTML | N/A | Joined Big Data Systems Team |
| 8 | https://github.com/UVADS/ds5110<br>https://tddg.github.io/ds5110-spring23/lectures/ | Familiarized with codebase and establish plan for contributions to codebase | N/A | Learn structure of course and go through lectures |
| 9 | https://microsoft.github.io/graspologic/latest/reference/release.html<br>https://jupyterbook.org/en/stable/start/your-first-book.html | Helped with video submission<br>Helped redesign layout and hero page of website | N/A | Turn in video submission and start contributing to booksite |
| 10 | https://arxiv.org/abs/1912.09363 | Helped embed elements such as presentations into the course site<br>This is direct and is more convenient for users as they have to click less to get to the lectures and information they need<br>Now, we know how to embed more interactive elements | N/A | Continue to add content to website |

| | | | | |
|---|---|---|---|---|
| | | going into the future. | | |
| 11 | https://tddg.github.io/ds5110-spring23/lectures/ | Adjusted Book site drastically, more readable and appealing to users by adding better spacing between headers and paragraphs, picture backgrounds, etc. Left and right buttons implemented in addition to navbar which helps user to easily go from section to section on website | N/A | Continue to add content to website |
| 12 | http://rafalab.dfci.harvard.edu/dsbook/getting-started.html | Implemented dropdown system for selecting lectures which makes it more efficient for users to get to the lecture they need to view rather than have to scroll through the entire page | N/A | Continue to add content to website |
| 13 | https://github.com/tddg/ds5110-spring23/blob/main/assets/ray_API_demo.ipynb | Added in animations to announcements section which increase UI, included icons to enhance user comfortability and engagement with the website, helped with structure of schedule page, lecture page is | N/A | Final touches to website and think about next steps with direction for improvements with site |

| | | | | |
|---|---|---|---|---|
| | | more dynamic and easier to navigate. | | |
| 14 | | Reviewed the codebase and made final stylistic adjustments to the site before displaying at presentation<br><br>Compare the first repository to final as many changes and improvements have been made through this iterative process<br><br>Site structure is complete, the only thing left is to integrate Preetham's interactive activities into the website | N/A | Prepare 5-10 Presentation detailing results of project |

# Reflection

As I conclude my independent study on the "Big Data Systems Project," I am drawn to reflect on the journey and the myriad of skills and insights I have gained. This project, centered around creating an effective educational tool for DS 5110: Big Data Systems, provided a platform for me to delve deeply into the realms of web development, UI/UX design, and data analysis, all while contributing to a meaningful cause at the beginning of the semester—the understanding of COVID-19's spread and its influencing factors.

One of the primary aspects of this project was the design and development of a website to facilitate learning. The challenge was to make complex data and concepts accessible and engaging to students. This endeavor required me to develop technical skills with an invested understanding of user experience. Crafting a user-friendly interface, integrating interactive learning activities like Jupyter Notebook, and ensuring the site's navigability were paramount.

The iterative process of receiving and implementing feedback was crucial in this regard. It simulated what a software engineer would have to deal with in a typical agile software development lifecycle. It wasn't just about building a website; it was about creating an educational experience.

In the process, my web development skills flourished. I primarily focused on front-end development. I learned to appreciate the nuances of creating a seamless UI experience, a skill I aspire to carry forward in my career. The use of technologies like AWS Classroom and Jupyter Notebook not only enhanced the site but also expanded my technical capabilities. I became better at using these tools, understanding their potential to transform educational content into interactive experiences.

Throughout this project, I also honed my research skills. Delving into various resources, from academic papers to online courses, I learned the importance of continuous learning and staying updated with the latest developments in the field. If you want to be valuable in this field, you should never stop learning. This aspect of the project was not just about gathering information; it was about synthesizing it to create something valuable and practical.

The independent study was also a journey of personal growth. Managing my time effectively between this project and other academic responsibilities taught me the importance of discipline and time management. I wanted to ensure I had the necessary time to devote my full effort to this project. It also fostered a sense of responsibility, as the outcome of this project had implications beyond my academic grade—it impacted the learning experience of my peers.

Above all, this project was not just an academic endeavor but a profound learning experience that spanned technical skills, research acumen, and personal growth. It has equipped me with tools and insights that I am eager to apply in my future endeavors in the field of computer science. As I move forward, I carry with me not only the knowledge I have gained but also the understanding of its potential impact on the world around me. I look forward to continuing my growth as a part of this research team going into the future.