HOME | Timeline | Next Week

Tuesday 05/24

This was my first day of the internship. I had to work on the causality project. As I was instructed I started setting up the environment on my local windows machine.

- I downloaded & installed the HPCC Platform
- Installed ECL version community_8.6.28-1
- Followed the instructions from https://hpccsystems.com/download/free-modules/machine-learning-library to install the machine learning library
- Installed ML_core bundle 3.2.2, HPCC_Causality bundle 1.0

In the team meeting with Roger and Zheyu from 1130-1300, we had a discussion on causality focusing on d-separation, Causal models. I am introduced to Causal effects (average, direct and indirect). I was told about my focusing part in this project will be on causal inferences and causal metrics.

In a 1:1 meeting with Roger, I was instructed to set up the environment in a virtual machine, as I had faced some issues with the HPCC Platform in windows. For a better understanding of Causality concepts, I started reading Causality Toolkit Blog.

Wednesday 05/25

This is the day of setting up the environment, so I was trying to install the virtual box and ubuntu 18.04. The installation failed twice for Ubuntu 18.04. And then I tried installation with ubuntu 20.04 which also failed. So re-installed ubuntu 18.04. This time it worked with slight modifications in the VM. Unfortunately, it failed to install the HPCC Platform in that VM. Exhausted with this I decided to discuss with Roger setting up the environment. I continued reading causal blogs.

Along with this, I was searching for a dataset to use in this project. I found one health dataset from Kaggle.

Thursday 05/26

I found an HPCC platform image (older version). Installed HPCC platform image in the Oracle Virtual Box and ran ECL watch, ECL IDLE. For understanding ECL, I ran the Anagram program in the ECL playground. I tried to install HPCC Client tools, but it failed, so I reached out to Roger. I was advised to install the HPCC Platform in Ubuntu using the Hyper-V machine, instead of an image file.

In the ream meeting with Roger and Zheyu, I am introduced to a few more causality concepts like causal notations and causal metrics. We discussed rewriting the expectation calculus, interventions and counterfactuals.

I cloned because repository and started testing the models. When I faced some difficulty in testing the models, I reached out to Zheyu. And set up the meeting with Zheyu and cleared it out.

Friday 05/27

I followed the instructions sent by roger in the pdf, originally sent by Lili Xu. I installed HPCC_Platform in ubuntu in HyperV and installed HPCC_clienttools, Because module, ML_Core bundle, and Causality bundle. Finally, it looked like the environment has been set up for the work.

I found some bugs in Testing the models in because module and reported it to Roger.

I have decided to use covid19.hpccsystems.com for applying the causality toolkit. So I have started analyzing the dataset looking for a possible hypothesis. I have found two hypotheses.

- What was the effect of full vaccination on testing positive for the second time?
- What is the impact of weather on the spreading of COVID-19?

HOME | Timeline | Next Week