

John Doros

## **4900 Weekly Accomplishments**

### **Week 1**

My accomplishment for week 1 was to understand the CnDetect research paper by reading and researching concepts that I did not understand or concepts that I struggled with, so I focused on taking proper notes and making sure that I understood the ideas that were presented.

### **Week 2**

My accomplishments for week 2 was to again understand and analyze concepts in the CnDetect paper, I discussed with my Supervisor understanding concepts from the paper, such as what type of censorship events can occur, and what needed to be researched for the project such as researching aggregate measurements on censorship events.

### **Week 3**

My accomplishments for week 3 was to complete diagrams and slides that were necessary for the project such as implementing concepts on how data processing is going to be done for the software.

### **Week 4**

My accomplishments for week 4 was to start working on the software project first by organizing a proper workspace such as setting up github, dashboard and setting up slack connections to my github. Then I designed a UI interface for my software that is simple and intuitive for the user. I then connected this interface with the backend of the software and had a full stack software structure that I could begin to add services that are going to be useful for the data processing portion of this project.

### **Week 5**

My accomplishments for week 5 was to start creating and implementing data loading into the software itself as I cannot just load in all the datasets at once, I came up with an efficient “minibatch” approach concept utilizing threads as a background process in order to efficiently load in small batches of datasets that we can analyze further. Then I worked on organizing my software files properly, fixing bugs and making the code more organized as well.

### **Week 6**

My accomplishments for week 6 was coming up with an effective and high probabilistic approach in order to classify censored dataset events with a new “measurement point” approach in which I have created a diagram for and implemented this concept into the software itself. This was the biggest accomplishment as my approach in fact worked and is aligning with my hypothesis part of the project

which is that if censorship data that is preprocessed with aggregation techniques and with metadata integration then censorship events can become more accurately identified while reducing false positives.

### **Week 7**

My accomplishments for week 7 was to come up with new features that will refine our user interface and the improve interactions for the software, I came up with different features to implement into the interface such as a data quality dashboard to see the progress of our data processing, a graphic visual feature in which we can see censored events vs non censored events, an optional decision tree feature in which makes our measurement point approach more refined, and a database storage feature in which can save processed datasets. My accomplishment was as well as reorganizing methods within the classes to make the codebase cleaner.

### **Week 8**

My accomplishment for week 8 was to incorporate the suggested features that I came up with into the software itself such as implementing the data quality dashboard and connecting this component to the backend with a manager class that handles this dashboard. I as well worked on required project deliverables that were needed to be completed such as timelogs.

### **Week 9**

My accomplishment for week 9 was to make sure my slides deliverable was aligned with the deliverable instructions. I was able to meet this challenge by looking and reading the General Expectations of Presentation Slides and the PLOs pdf as well I made sure I have included my learning outcomes in the slides such as I learned how to better implement algorithms, how components can work together in the application, and how to handle data processing more efficiently through the use of threads. So at the end of this time period I was happy with the results as I planned to work and make sure everything is ready for submission.

### **Week 10**

My accomplishment for week 10 was to implement these summary classes in a way that we can utilize them in a main database feature which will help us combine all summaries of datasets into one viewable component which can make analysis much easier, and data extraction much easier as well. The focus is to maintain data extraction in a more easier fashion so the user does not have to worry about how to extract the data in which the database component will handle this making data extraction a smoother process.

### **Week 11**

My accomplishment for week 11 was to figure out a way to solidify if our measurement score approach is reliable since we need to compare to censored events that do not come from Censored Planet as this would be biased in the data itself. So I needed to find an outside source that has confirmed reports and we

can then match with our measurement score and this is what I am having trouble with since no datasets are matching with the dataset that was given to me, but I will continue finding a solution for it.

## **Week 12**

My accomplishment for week 12 was to study and find a solution to properly checking if the measurement approach is reliable I tried many methods and the only reliable approach I found was by utilizing a well known censored database called OONI where there are confirmed reports of censorship events happening at a specific time period, they have a measurement tool in which I can utilize via api requests in order to find the same dataset and exact time of our predicted censored events to theirs and if the report summary matches then we can say our approach is reliable with some score percentage in other words how close did our approach predict if those censored events.

## **Week 13**

My accomplishment for week 13 was to implement the solution that we found by utilizing the OONI measurement toolkit on week 12 and its api requests in Python I needed to format the api requests in a way that gives us the exact datasets that the software processes so that our datasets match in terms of dates/country and domain. After that there was formatting that needed to be done with the dataset from the api requests in order to accurately find the censored datasets that the measurement score produced, the problem was that I needed confirmed reports from the OONI so I needed to check if a specific value in each dataset was marked which was the confirmed report, this made analysis much easier since those datasets are confirmed to be censored or not.

## **Week 14**

My accomplishment for week 14 was to implement the QT PieChart into the score report. I was able to meet this challenge by reading the QT documentation on QT PieCharts, and was able to easily implement this into the software after I learned what was needed.

## **Week 15**

My accomplishment for week 15 was to complete the final deliverable project and making sure everything was organized. I was able to meet this challenge because I followed the rubric that was provided in the deliverable, and I made sure my slides were well organized, structured and had as many details as possible. I also made sure my video recording was well put together.