**Date:** May 20, 2021

**Previous Week**

**Goal:** Anomaly Detection Model

**Completed:**

* **Understand/Build working framework for anomaly detector with InfluxDB**
  + have two frameworks that should work including use of Telegraf running a Python program
  + the streaming framework is the area we feel we have a handle on at this point (putting work on hold until there is model detail in anomaly detection model), no longer part of the critical path for now
* **Ongoing anomaly detection framework research/short-listing**
* **Ongoing EDA**
* **Ongoing support on data parsing to support streaming**

**Issues:**

* Still don’t have easy access to all data and current access method is time consuming
* Sounds like there won’t be labelled data, we’ll need to come up with our own anomaly labels
* Ongoing time spent on data parsing (hopefully this is now complete)
* Not as far as we’d like to be on the anomaly detection framework selection – we were hoping to build this week but we felt it was more important to continue research and select a method
* Recognized that we need to spend time on selecting performance measures for the system (something that was missed in the proposal)

**Outstanding/Carryover:**

* **Complete ongoing EDA to support identification of anomalies and guide model**
  + Still in progress and will likely somewhat continue early next week
* **Meet with domain experts (Friday)**
* **Completing short-list of anomaly detection methods (Friday)**
  + working on a more thorough selection than what was completed last week (should be complete today/tomorrow)
* **Build the model both in terms of data processing and anomaly detection**
  + this will now be in this coming week as was not completed this week
* **Test various short-listed anomaly detectors and select/have a working model** 
  + this will now be in this coming week as model was not completed previous week
  + includes **selecting performance measures** early in the week
* **Did not have time to look at Azure**
  + Putting this somewhat on hold until the above gets on track

**New Week**

**Goal:** Supposed to Implement Streaming Pipeline but will largely be Building the Anomaly Detection Model

**New Tasks:**

* **Status Presentation**
  + Needs to be complete by end of week
* **Streaming anomaly detection pipeline**
  + original plan for this week and will start being implemented based on model progress once we better understand what the model will look like
* **Model Tuning and Evaluate Performance**
  + original plan for this week and will start being evaluated as week progresses

**Realistically, feel that Week 4 will largely push into Week 5 (Value Week). However, we hope that there will still be time to start at least one of the additional items.**

**Issues/Resources Needed:**

1. Access to all of the data in InfluxDB would be really useful (if possible). Sounds like it might not happen within the timeframe we need to move forward with modelling so we’re committing to downloading say ~15 sensors (15 sensors x 12 months x 2 years = ~360 downloads) + self-labelling to be for supporting the modelling next week.

If we are able to get data in InfluxDB in the next week or so, we can look at additional sensors; otherwise, our study will potentially be more limited to the 15 sensors downloaded.

|  |  |  |
| --- | --- | --- |
| **Person** | **In Progress** | **New** |
| **Nate** | * None | * Support Ryan/Mitch tasks as needed to get them on track * Start implementing streaming anomaly detection model when ready * Status Presentation |
| **Mitch** | * Build/Test Detection Models | * Start implementing streaming anomaly detection model when ready |
| **Ryan** | * EDA (includes anomaly identification) * Select Performance Measures | * Select Performance Measures * Model Tuning and Evaluate Performance |
| **ALL** | * Meet with domain experts |  |
| **Currently Unassigned** | * Review Azure |  |