**DDate:** May 13, 2021

**Previous Week**

**Goal:** Data and Systems Understanding

**Completed:**

* Understand data systems – still also in progress, understanding how streaming data will go into InfluxDB
* Learn InfluxDB, also able to get telgraf/InfluxDB running locally via Docker
* Create simple offline simulator for anomaly detection testing
* Anomaly detection research
* Partial EDA/Data selection
* Look at various potential methods for running a Python anomaly detector with InfluxDB, included running simple code examples

**Issues:**

* Data access, working on telegraf parsing of JSON input data (which is still ongoing) put where the EDA needs to be behind schedule
* There doesn’t appear to be any straightforward methods for running an online Python program, several methods were looked at all of which aren’t perfect methods (see ppt ‘streaming\_options.ppt’).

**Outstanding/Carryover:**

* Understanding how the data will stream into InfluxDB (should be complete today/tomorrow)
* Completing short-list of anomaly detection methods (should be complete today/tomorrow)
* Get data and complete more comprehensive EDA (data access should be complete today potentially with use of csv file instead, EDA will likely continue into early next week)

**New Week**

**Goal:** Anomaly Detection Model

**New Tasks:**

* Data cleaning/features steps for anomaly detection pipeline
* Test various short-listed anomaly detectors and select/have a working model
* Build working framework for anomaly detector with InfluxDB
* Meet with domain experts – end of week
* Start reviewing Azure

**Issues/Resources Needed:**

1. Need the data access issue resolved, either needs to be resolved ASAP or we need to move into using csv files. Ryan has started working with csv’s to get the EDA going and Jiachen is discussing the file formatting with EWS.
2. Need a good understanding of how streaming data will report to InfluxDB. Coming in via a port (8080), each file will be for a different sensor. Exact format is being discussed.
3. Meet and discuss anomaly detection streaming architecture – this should be done ASAP and may be useful to have Ibrahim available for this. Invite Ibrahim to the Monday meeting.

|  |  |  |
| --- | --- | --- |
| **Person** | **In Progress** | **New** |
| **Nate** | * Determine streaming method to be worked on for anomaly detector (should be complete today/tomorrow) | * Build working framework for streaming anomaly detection with InfluxDB |
| **Mitch** | * Create short-list for work next week (should be complete today/tomorrow) | * Work on data cleaning/features pipeline * Build anomaly detection model |
| **Ryan** | * EDA and select data (will go into early next week) | * Work on data cleaning/features pipeline * Build anomaly detection model |
| **ALL** |  | * Meet with domain experts |
| **Currently Unassigned** |  | * Review Azure |