

Weekly Check-In with Levannia, Aditya, and Mike Meeting Minutes – June 1

Levannia asked for us to do a code review. We showed her some of our most important notebooks.

- It would be helpful for them if we make a glossary describing the notebooks and where everything is.
- If we have time, it would be useful for them if we pull our most important functions (especially the noise reduction ones) into a python script and then demonstrate importing and using them in a notebook.
- Overall they are happy with our process and they don't need us to change our workflow or practices.

Levannia said that the ability to cluster into shapes based on the difference in window means was promising. Also she said an F1 score of 0.6 on a random forest is not nothing and that we have to walk before we can run. She also mentioned that we are free to try supervised methods as we see fit.

We discussed our next steps:

- Finalize our clustering models by this Friday
- Spend next week characterizing and comparing the clusters
- Reserve the final week for writing the report. Both Aditya and Levannia thought that one week for report writing and final presentation creation was a good amount of time.
 - o We can send them the report in pieces for feedback.
- Levannia mentioned that they are very happy with where we are at, and we can freeze the project and start focussing on the report whenever at this point.

We also asked what would be most useful for them for us to accomplish over the next two weeks:

- Levannia said that for them one of the major roadblock is that they don't know the different ways the waveform can look so characterizing those would be very useful for them moving forward. Especially if we find a standard way of describing different categories of waveforms would be useful for them.
- The process of removing the noise and what is most useful will also be beneficial for them.
- Mike agreed – having different clusters is good – it's an existing catch all category of 'sensor issue' so being able to see how the anomalous shapes appear or disappear over time will be useful for informing the manufacturing team in the future.

We asked if they think there's anything particular that would be useful in characterizing the waveforms

- They said this is really their first foray into examining them so they aren't sure but start simple looking at things like mean / first derivative / second derivative / min / max / for different windows
- Mike said sharper looking shapes will be electrical issues, smoother ones potentially mechanical failures

If we have unsuccessful tests that look like pin contacts we can send the ids and they will re-classify.

- They don't know what they have as a benchmark for accuracy in finding pin contacts since they don't look at lots that pass, but just having a start to a modelling approach is helpful.

Aditya will share the code of how he trained a kernel for convolution last week.

Levannia stressed a few times that they are very happy with our work and how much effort we have put into looking at the waveforms.

We learned that the three biggest pieces of feedback from the data review were:

- I think my biggest question was the same as what John Walker asked - mainly how do we know that some of the "unsuccessful" tests weren't pin contacts? I think it's important to look at the two classifications as separate with a grain of salt, since humans have done the labelling and it's fairly easy to mislabel.
- [Monday 1:52 PM] Schoepp, Katherine
I guess another question I just thought of - generally pin contacts on the amps are when the current goes to zero (since there is no electrical connection). I didn't check this, but are the students looking at the data as normalized or raw? Pin contacts are usually quite easy to spot by looking at the raw data for a flatline at 0 nA.
- [Monday 1:53 PM] Schoepp, Katherine
Another complication to consider is that pin contacts can be intermittent, i.e. they can read 0 nA for part of the waveform (e.g. only cal, or only sample, or some subset) and then go away/reappear. So one test may not look the same as another depending on when the pin contact appeared.

One thing we'll try moving forward is seeing if any unsuccessful are consistently clustered in with pin contacts in clusters that have a high proportion of pin contacts across different methods.

We discussed our midterm presentation and how everyone understood it was complicated problem. We also told them that one issue was that our problem statement was not adequately concise, possibly because of our dual focus on describing waveforms and separating out pin contacts. We mentioned how we need to find a better way of formulating it for the final presentation.

- Levannia said that ultimately, they would want the pin contacts separated out, but it might not actually be possible.
 - o They don't know much about the waveforms so the work that we've done with them has been very helpful.
- Aditya said to have a clear statement of work in our final report.