May 13 Check-In With Siemens Healthineers Meeting Minutes

We prepared some slides to help us go over what we have accomplished in the past week. Here are the comments and feedback we received on each section.

Predictor Wrangling

- Got time series and predictors matched.
- Got rid of all 0 columns.
- On unsuccessful land pin contact:
- Tried k-means on predictor data
    - Got 2 clusters with only 2 records (both outliers)
- Also tried Random Forest
- Only feedback was that they wanted to know if the random forest had been attempted on all the data or just unsuccessful/pin contact.

Time Series Wrangling

- We can exclude the really short readings.
- 0 as sample detection time can mean the sample was never detected or the reading ended early because the operator realized something was wrong.

Filtering

- There's a range of wet-up period lengths – we will want to plot them to make sure we're getting all of it.
- In Finished Goods, samples are always injected around 200 seconds.
- Pin contacts usually occur on the calibration side. They can occur on the sample side but it's rare.
- Probably best to standardize the waveforms and then filter.
- Maybe a stochastic method might work for looking at waveforms if it's true that unsuccessful/pin contact can be separated based on the noise.
- They are not sure about the frequency of the noise. Different materials/enzyme activity – different batches might have different noise. There may not be a set frequency range that is bad.
- User can't actually test the card until wet-up/heating protocol is done. The signal is what the clean result looks like on the successful side – low frequency.

Unbalanced Data

- They are using resampling techniques for other analytes but ours is noisier – resampling is not going to be easy.
    - One way: random over/under sampling – not the best
    - Cluster sampling – could use to undersample. Might not work for our analyte though.
- One option: run neural net, then run again but don't re-initialize weights with each iteration.
- Just because a test returns successful – doesn't mean it's a good test. Critical medical errors occur within successful

- Good waveforms are looked at by operators as context to compare with unsuccessful ones. Eg. Anomaly might only look weird on one side of the waveform.
- Less than 0.5% critical medical errors – clusters within successful are also going to be quite imbalanced.
- Thought with looking at successful – some way of determining different patterns that occur in successful readings.
- If 6 weeks isn't enough time to look at everything – could shift focus and look only at the unsuccessful readings (ignore the successful for now except as reference).

They are trying to synthesize more pin contact errors. Might need to find whether they are representative of 'wild' pin contacts. They're not injecting samples though so they will have 0 sample detect times.

Feedback from the advisory committee on proposal presentation:

- Everyone is excited about the project
- If we find some methods/find clarity on what could potentially work, people will be happy.

New pin contacts – some are a bit older, but someone went in and verified that they are actually pin contacts before giving us the data.