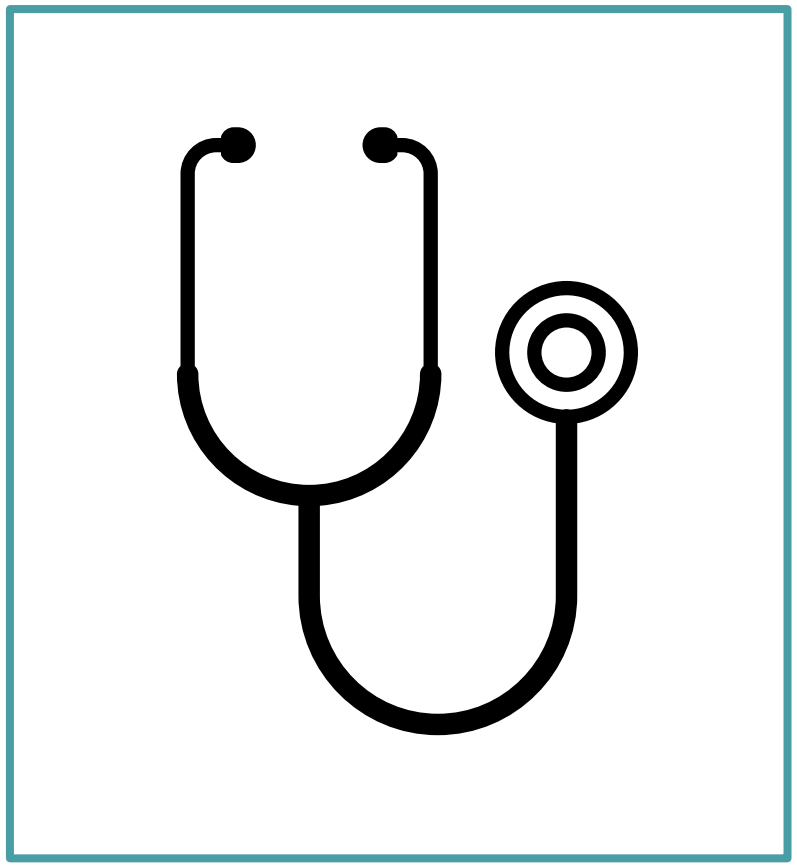
ANOMALY DETECTION IN BIOSENSOR WAVEFORMS

JUSTINE FILION, NEETHU GOPALAKRISHNA, SAISREE GR, SARA HALL

SIEMENS HEALTHINEERS



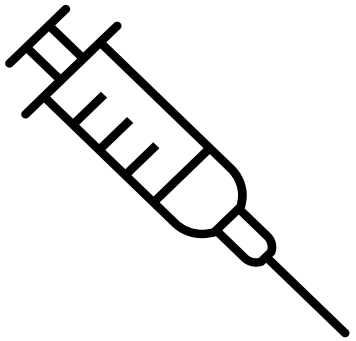
- Company that produces medical devices
- Focused on medical technology innovation
- Based out of Germany

https://www.siemens-healthineers.com/

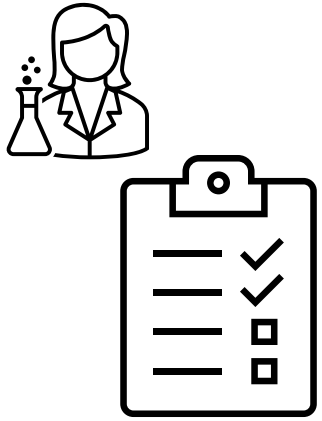
EPOC BLOOD ANALYSIS SYSTEM

https://www.siemens-healthineers.com/blood-gas/blood-gas-systems/epoc-blood-analysis-system

MOTIVATION, PURPOSE AND PROBLEM

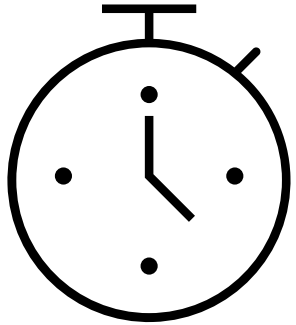


Blood analysis requires accuracy



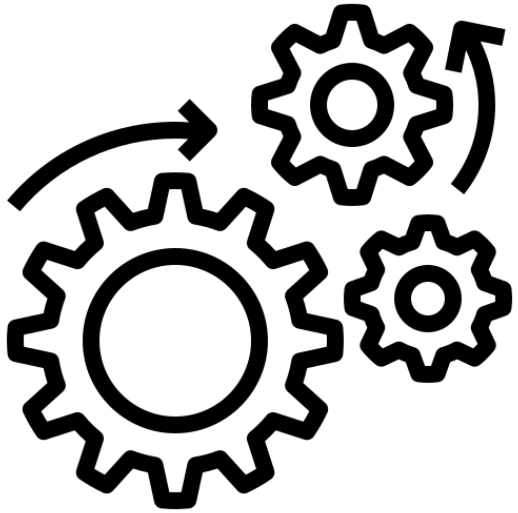
<u>Quality Control:</u>

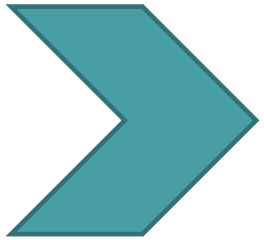
Manually check waveforms for anomalies (like pin contact)

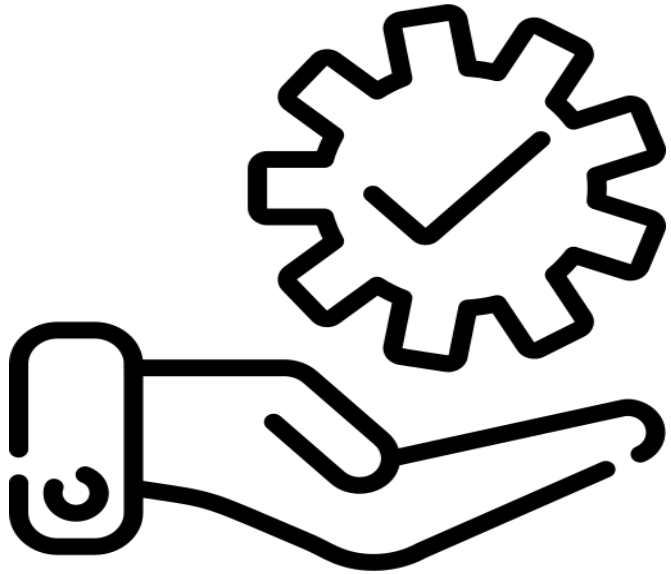


- Time consuming
- Challenging to identify visually

SOLUTION



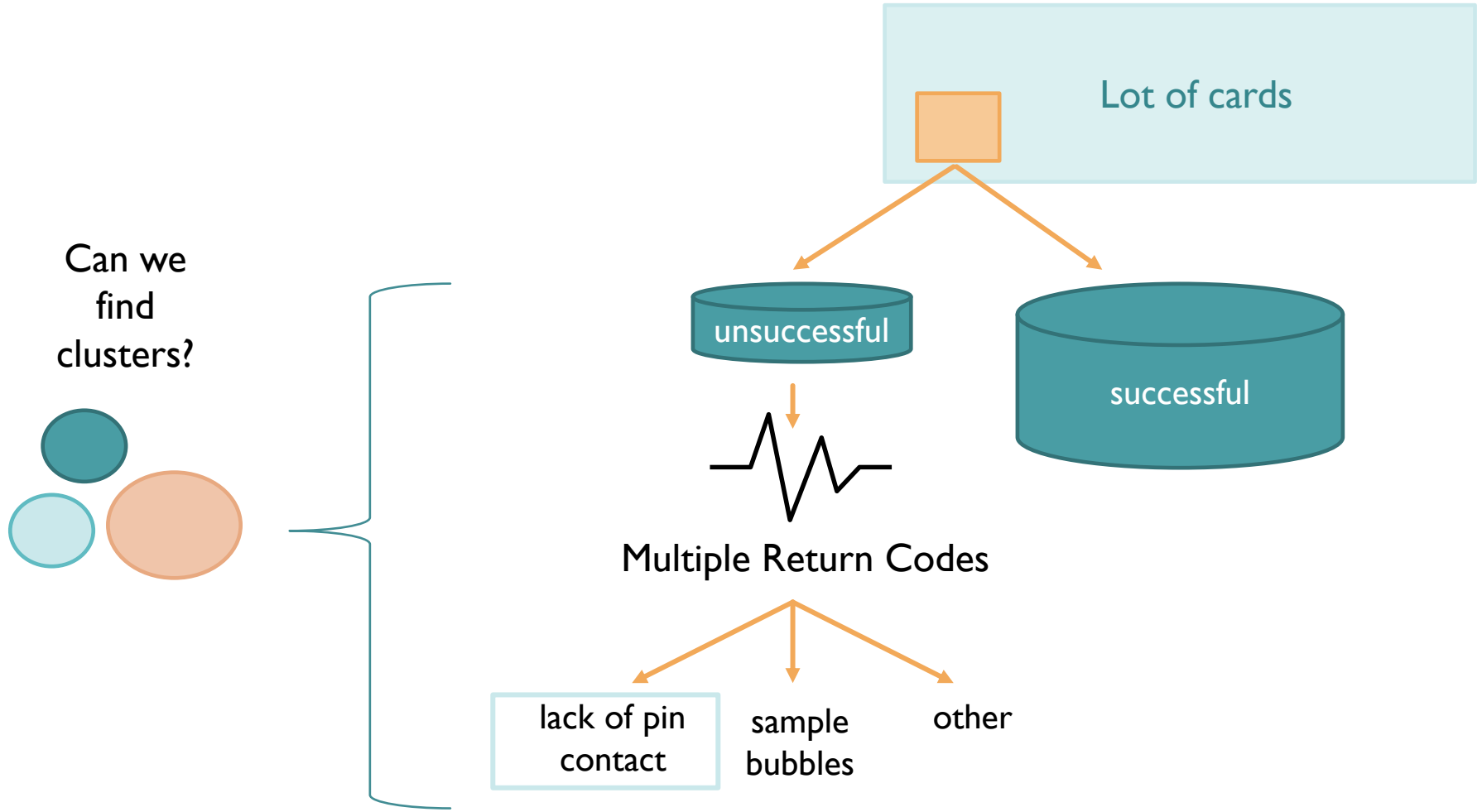




Automate the process of quality control

Improving efficiency of workforce productivity

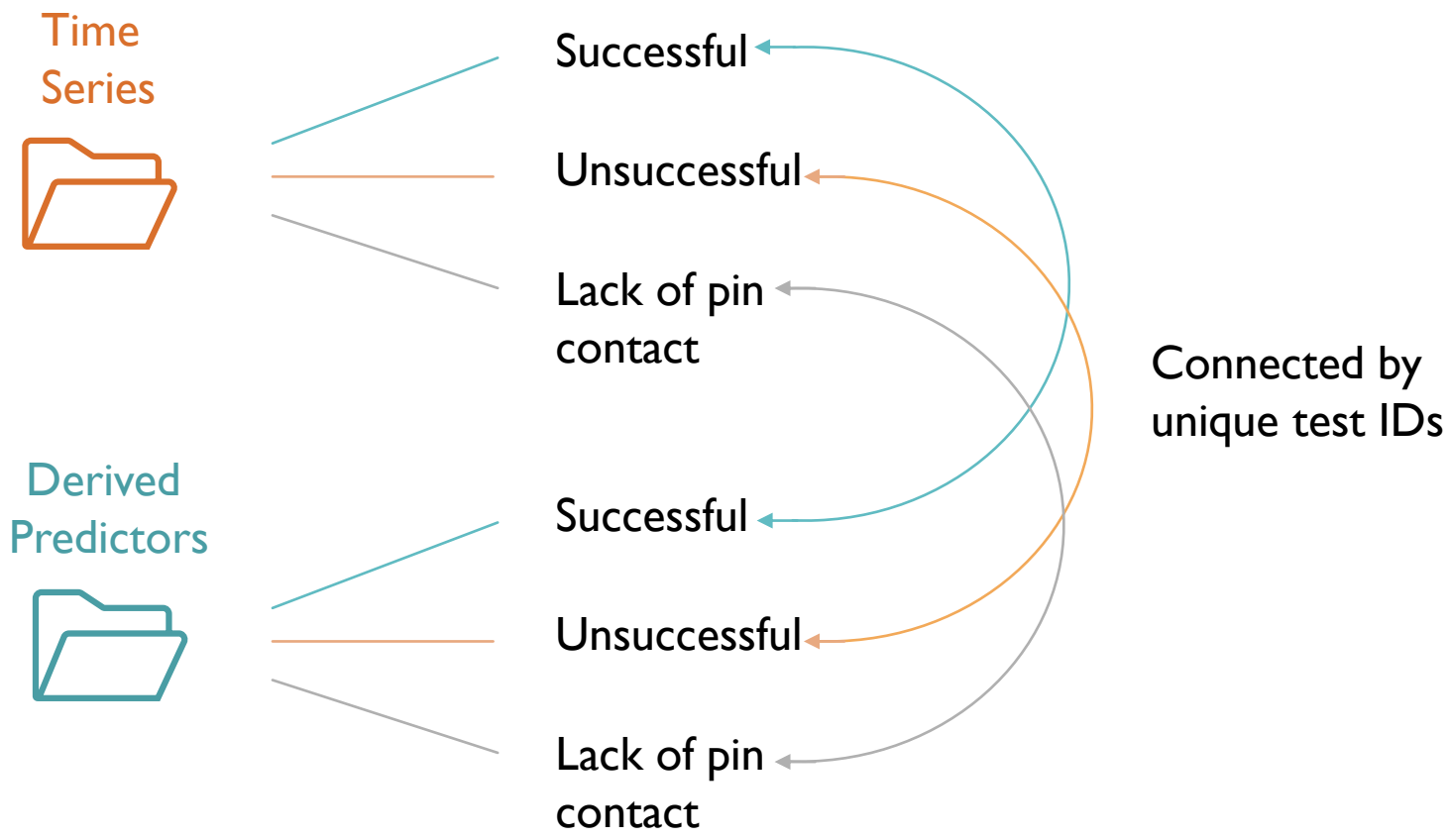
AUTOMATION OF THE FINISHED GOODS PROCESS



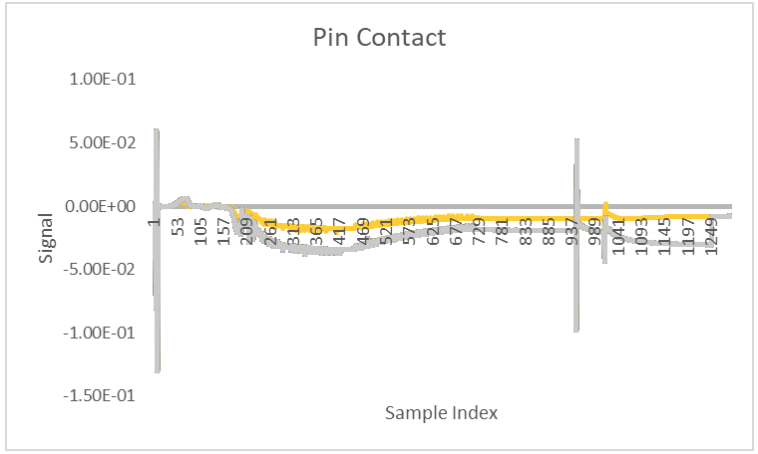
AIMS AND OBJECTIVES

1. **Develop machine learning pipelines to cluster readings**
2. **Determine which methods are effective and which are not for identifying anomalies in biosensor readings**

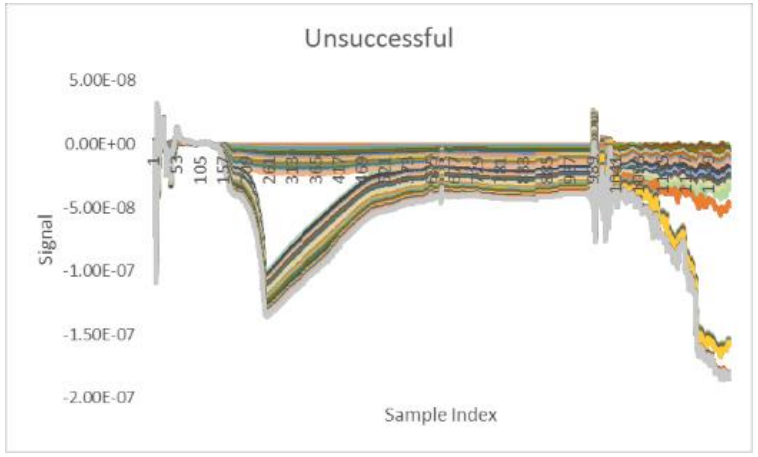
DATA


Time Series
Derived Predictors
Successful
Unsuccessful
Lack of pin contact
Successful
Unsuccessful
Lack of pin contact
Connected by unique test IDs

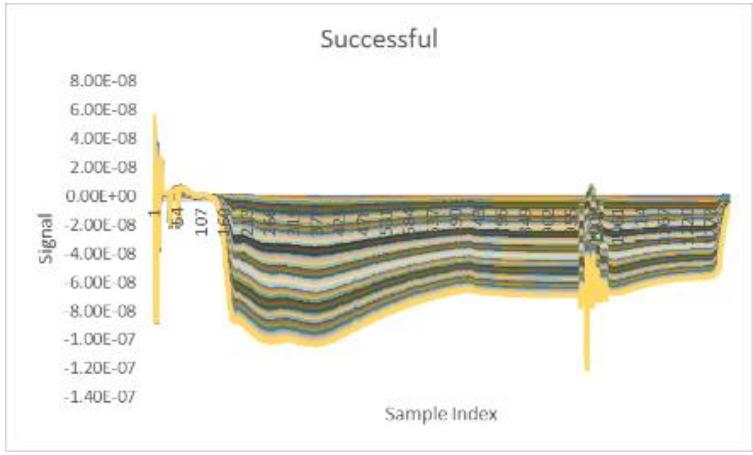
DATA



~ 80 Readings



~ 10 000 Readings

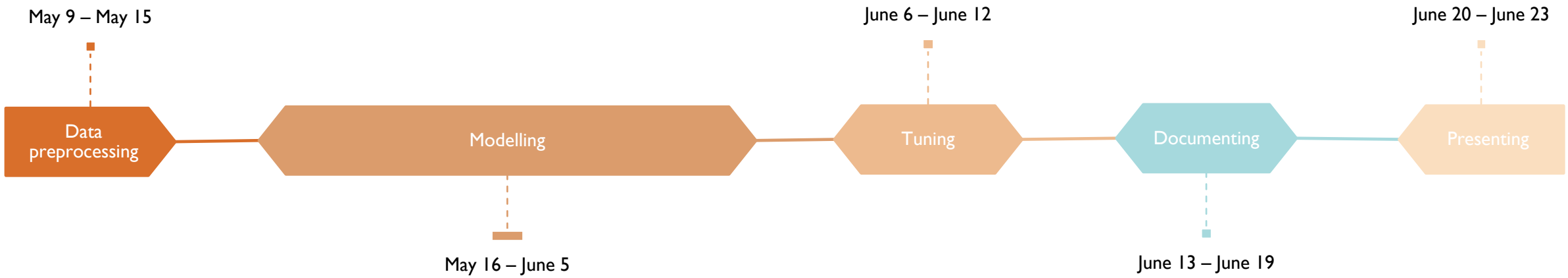


~ 400 000 Readings

DATA

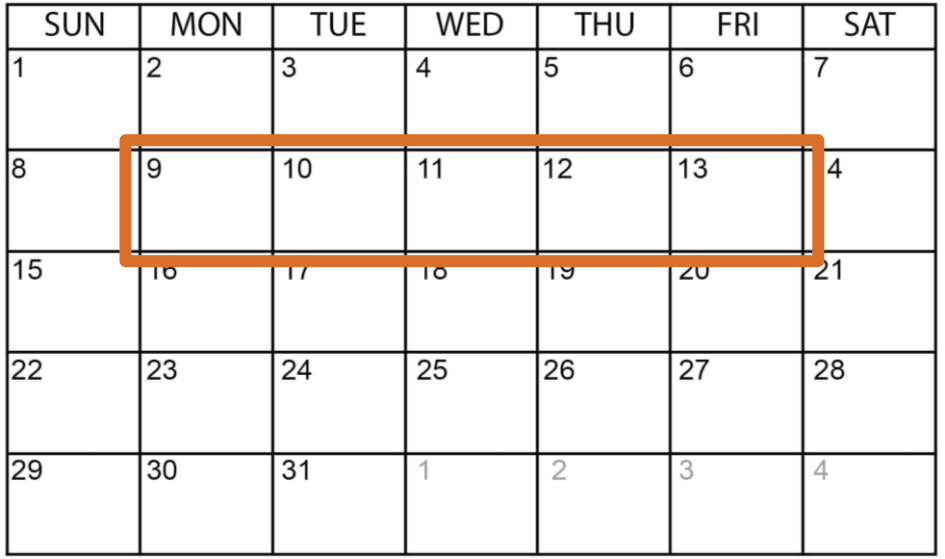
	TestID	ReturnCode	CExtrapolation	CMean	CDrift	CNoise	CSecond	CWindowMovedBack	SExtrapolation	SMean	...	TransDrift	TransNoise	TransSecond	AMean	ADrift	ANoise
0	8040472	SampleDipLate	-0.363469	-0.126184	-0.005467	0.006113	-0.001081	0	-0.133205	-0.133205	...	72	-0.000747	0	0.146503	-0.006869	0.020550
1	8040613	CannotCalculate	-0.723230	-0.732675	0.000218	0.003244	0.000032	0	-0.587647	-0.587647	...	0	0.000047	0	-0.740345	-0.000411	0.003191
2	8042716	AdditionalDriftHigh	-0.683783	-0.686323	0.000058	0.002718	0.000005	0	-1.541267	-1.541267	...	37	0.001581	0	-0.679148	0.008325	0.011853
3	8042793	CalMeanQCLow	-48.398226	-48.397023	-0.000027	0.001028	0.000006	0	-48.396699	-48.396699	...	3	-0.000026	0	-48.396549	0.000024	0.001141
4	8043617	CannotCalculate	0.000000	0.000000	0.000000	0.000000	0.000000	0	0.000000	0.000000	...	43	0.000000	0	0.000000	0.000000	0.000000

SCHEDULE


May 9 – May 15
Data preprocessing
Modelling
May 16 – June 5
June 6 – June 12
Tuning
Documenting
June 13 – June 19
June 20 – June 23
Presenting

I – DATA PREPROCESSING

- Perform noise reduction on the waveforms
- Split the waveforms into different windows
- Resample the data to obtain multiple training sets with more balanced classes
- Clean and wrangle the aggregate predictors data



2 - MODELLING

- Build various unsupervised machine learning pipelines to cluster different types of readings and identify pin contact errors.
  - Example:
    - Mixture Model?

3 – TUNING

- Improve the promising model(s)
- If we have time, look at data augmentation methods

4 - DOCUMENTATION

- Write the final report
- Prepare the final presentation

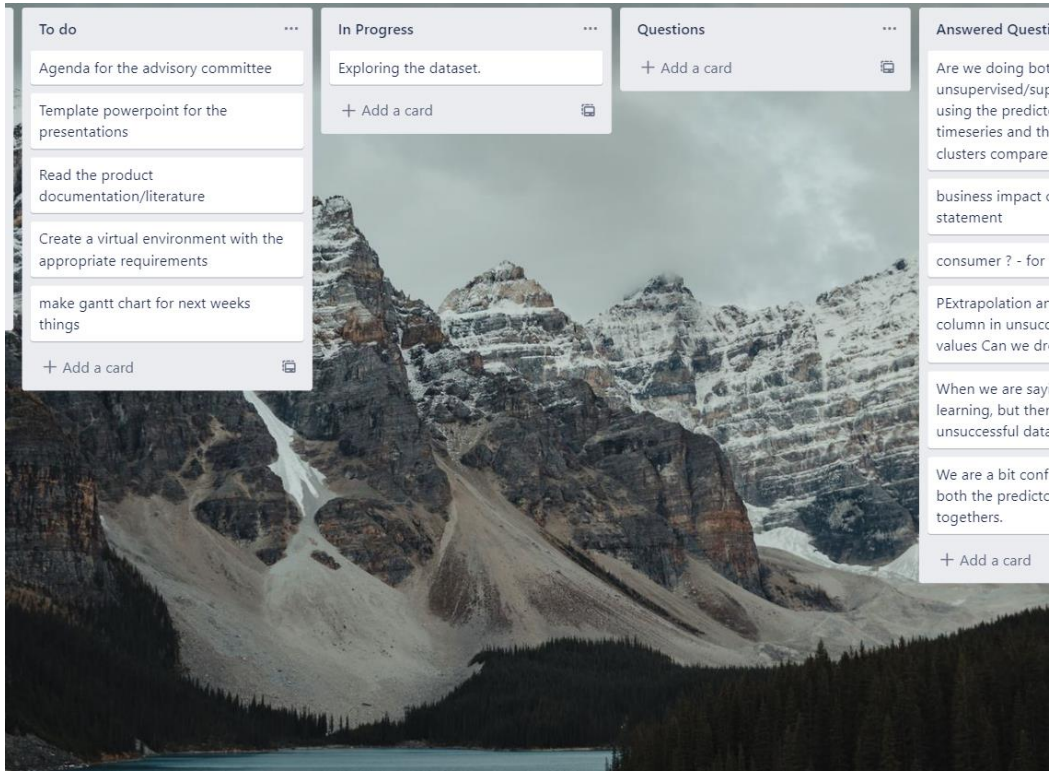
5 - PRESENTING

- Presenting the final project

JUNE 2022

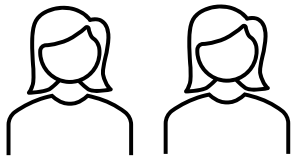
SUN	MON	TUE	WED	THU	FRI	SAT
29	30	31	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	1	2

TEAM ORGANIZATION

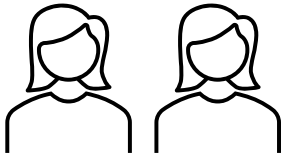


- Trello board
- Daily team meetings
- Monday/Friday meetings with client
- Bi-monthly meetings with advisory committee
- Gantt chart for weekly progress
- Private slack channel for daily communication

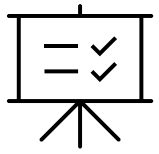
ROLES AND RESPONSIBILITIES

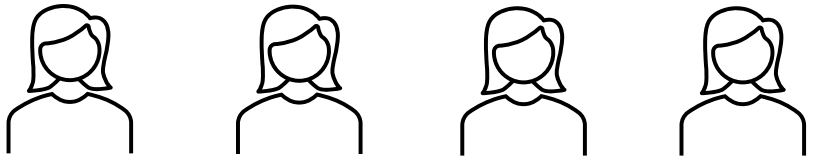


Team working with aggregate predictors

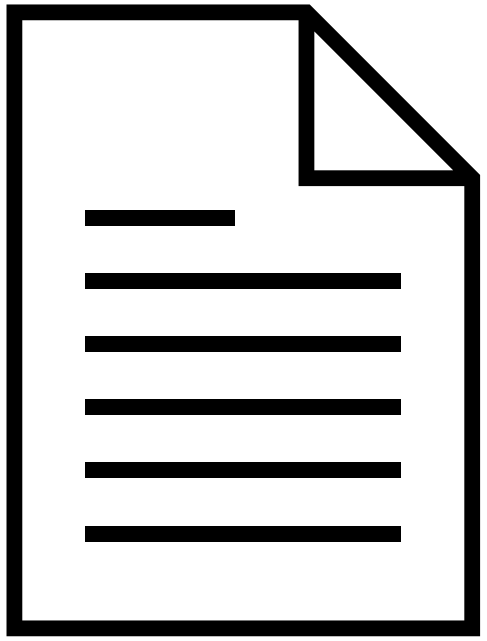


Team working with biosensor waveforms





DELIVERABLES



- Well commented Python code for everything we have tried
- A final report detailing our attempts

Weekly log

Week 1 : May 2-7th

Date	Task	Description
May02	Setup github repo	• Sara configured the github repo, all team members joined.
May03	First Client Meeting	• Introduction to the Siemens Healthnieers and ice breaker bingo game to build team dynamics. • Presented with the problem at hand, expectations and data to explore. • Exploration of data.
May04	Q&A session with client	• Clarification on questions arised during exploration of data from previous day • Contiuned research on understanding data and expectations.
May05	Q&A session with client & Proposal writeup	• Further questions to understand the problem better. • Each team member worked on apart of proposal and proof read each others work, mailed it over to clients for approval
May06	Feedback on proposal & Presentation slide	• Client approved to move forward with the proposal • Prepared tbe presentation slides for course instructors meeting and Siemens Advisory committe meeting.

THIS WEEK'S UPDATE

BREAKDOWN OF THIS WEEK

TASK	ASSIGNED TO	PROGRESS	START	END
LITERATURE REVIEW/DATA PRE-PROCESSING				
Statistical test to use subset of successful	Justine	0%	5-9-22	5-11-22
Building training set with bootstrap	Justine	0%	5-12-22	5-13-22
Filtering (noise reduction)	Sara, Saisree	0%	5-9-22	5-10-22
Windowing of time series	Sara, Saisree	0%	5-9-22	5-13-22
Cleaning and wrangling the predictor file	Neethu	0%	5-9-22	5-14-22