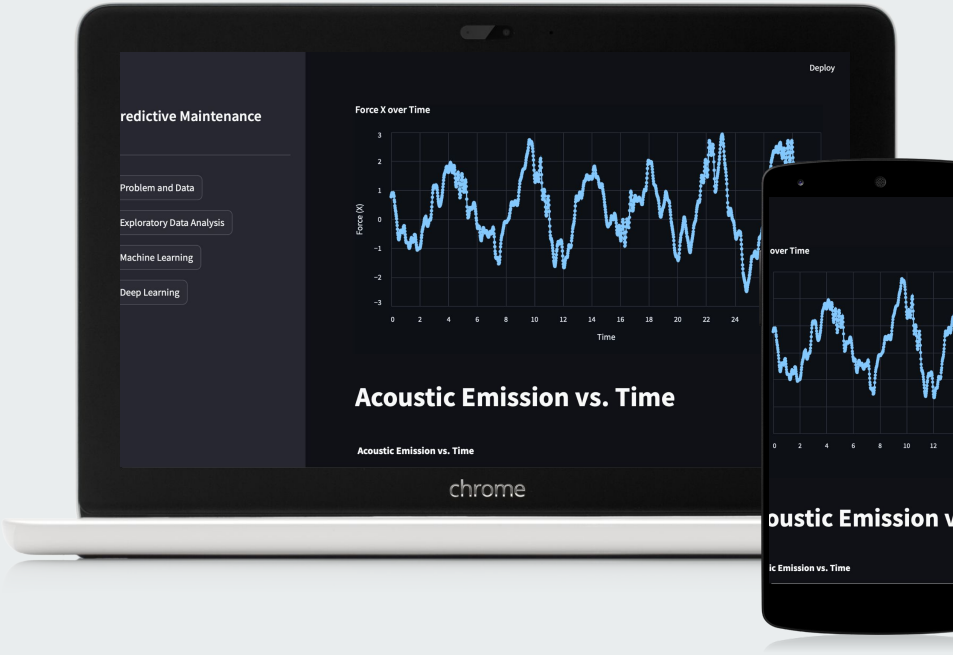
Predictive Maintenance Project

Analyzing an industrial machinery dataset and predicting machine wear



Clement Ogundipe Daniel Martin David Ramos Salamanca Dermot Lyons Edina Petróczki

Project objective

Use machine learning to predict when a machine needs maintenance.

Why?

This kind of data science project has the potential to save manufacturing companies money.

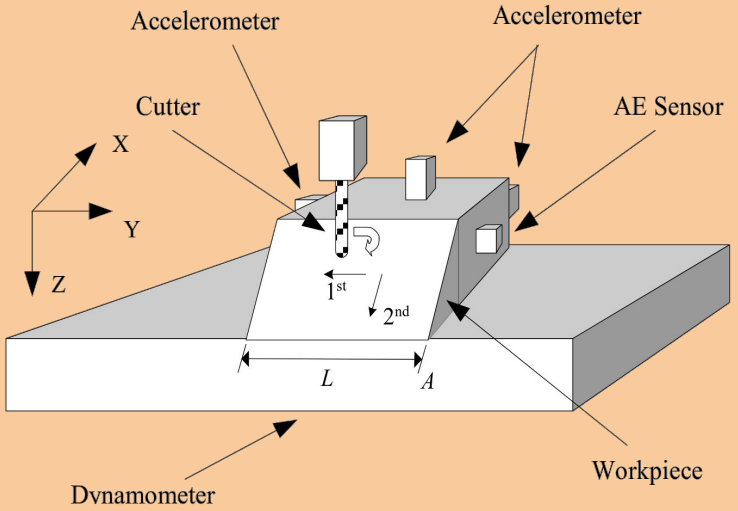
Problems to solve

1. Find a suitable data source
2. Analyse the data
3. Model the data
4. Host the data on the cloud
5. Visualise the data

Overview: Predictive maintenance of a CNC milling machine

In machine cutting processes, in order to ensure surface finish quality, it is imperative to keep the cutting tool used in top operative condition.

One approach to monitor the machine wear is to measure certain process parameters, such as cutting force, tool vibration and acoustic emissions and predict the wear of the cutter.



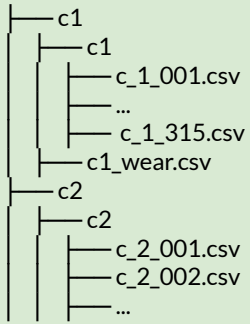
The dataset

We used the dataset from the 2010 PHM Society Conference Data Challenge

It has **~18 GBs** of sensor data from a CNC cutting machine prepared for analysis and modeling.

Data source

The data set is available at Kaggle:
PHM data challenge 2010



The group

Daniel:
GitHub repo owner
Dashboards
Frontend

Clement:
Cloud hosting
Python basics

Dermot:
Scrum master
Dashboards
Python basics

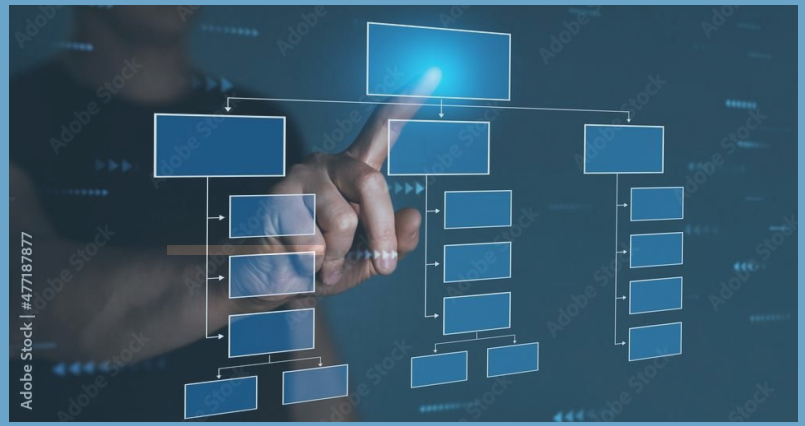
David:
Product owner
EDA
Feature engineering
Deep learning modeling
GitHub repo owner

Edina:
EDA
Time Series analysis
ML modeling

Mascot

Our epics

- Setting up the environment
- Learning about tools
- Learning the domain knowledge
- Data exploration
- Data modeling
- Cloud storage
- Dashboarding



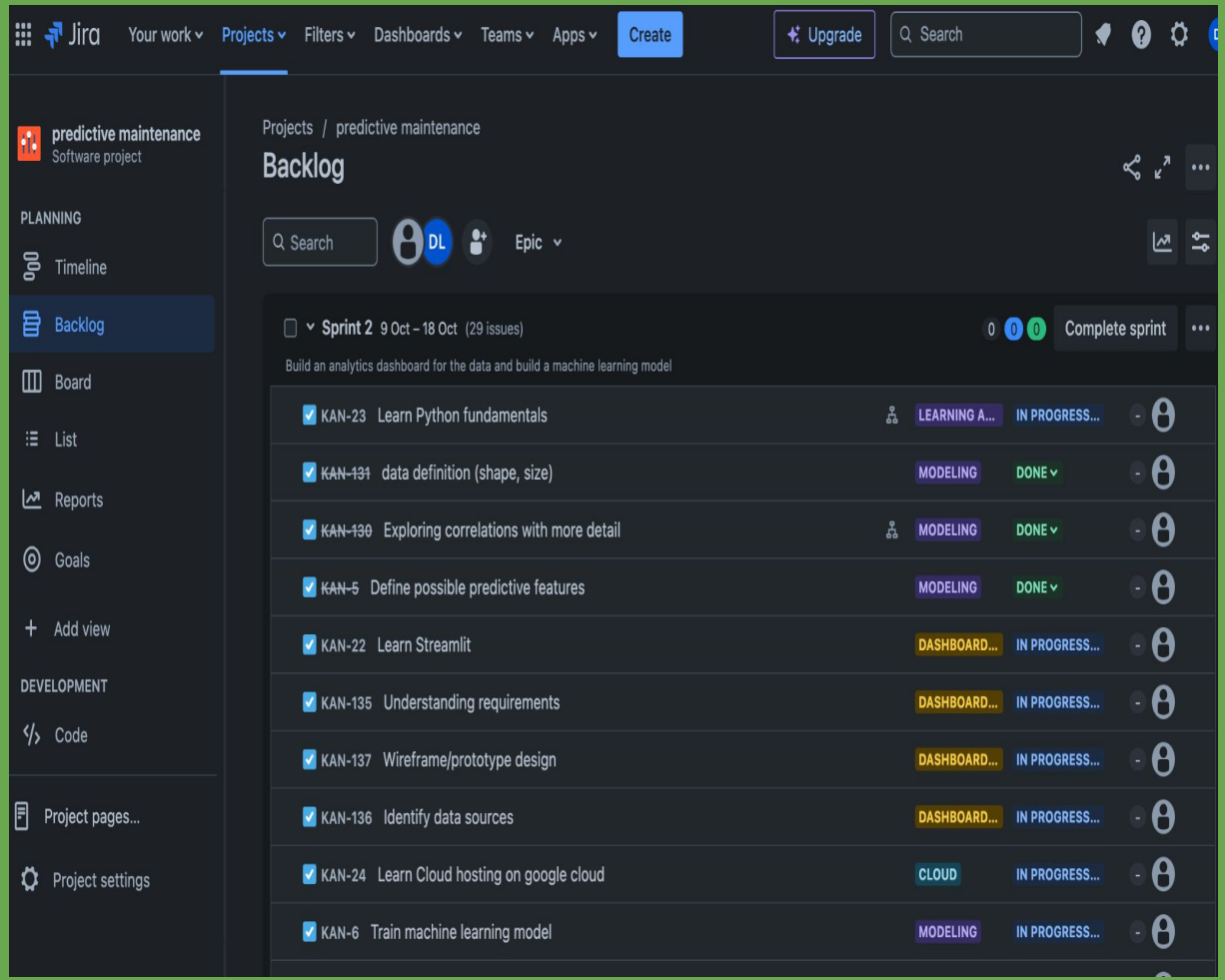
Workflow

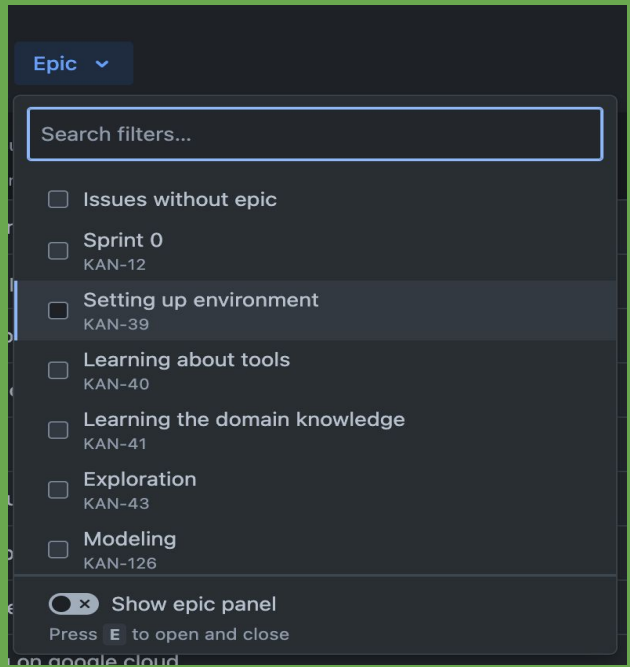
Jira

Slack

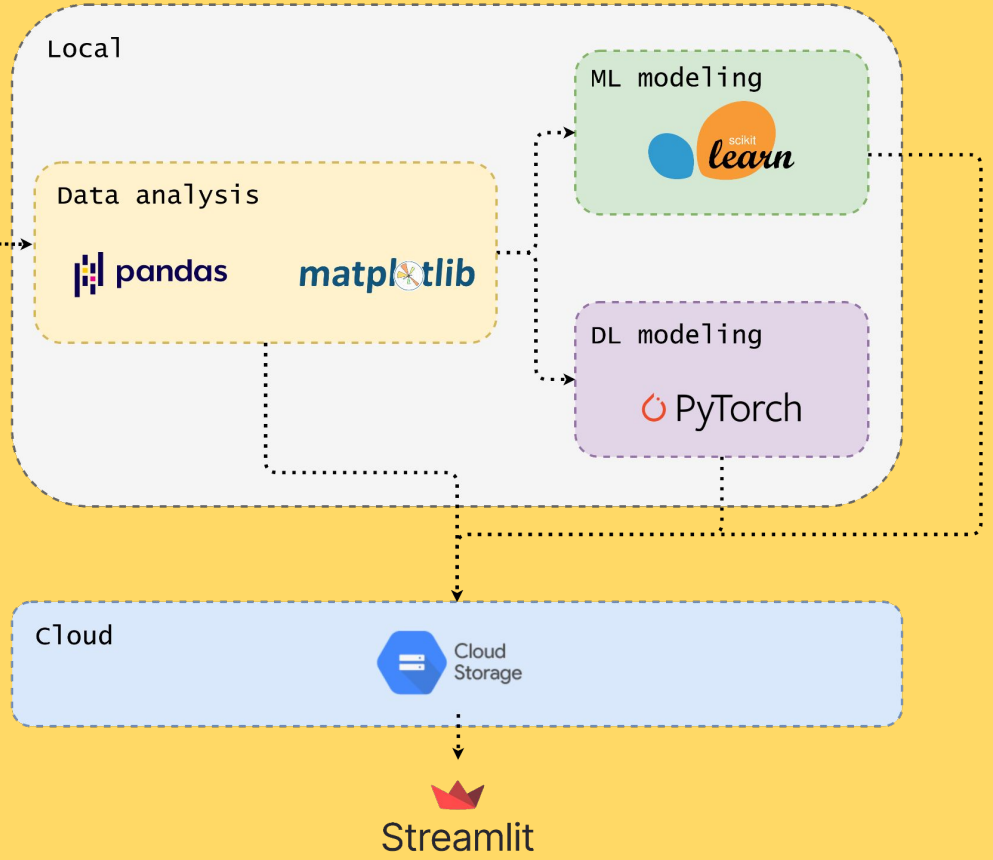
Google docs

Google meet


Jira
Your work
Projects
Filters
Dashboards
Teams
Apps
Create
Upgrade
Search
predictive maintenance
Software project
PLANNING
Timeline
Backlog
Board
List
Reports
Goals
Add view
DEVELOPMENT
Code
Project pages...
Project settings
Projects / predictive maintenance
Backlog
Search
DL
Epic
Sprint 2 9 Oct – 18 Oct (29 issues)
0 0 0
Complete sprint
Build an analytics dashboard for the data and build a machine learning model
KAN-23 Learn Python fundamentals LEARNING A... IN PROGRESS...
KAN-131 data definition (shape, size) MODELING DONE
KAN-130 Exploring correlations with more detail MODELING DONE
KAN-5 Define possible predictive features MODELING DONE
KAN-22 Learn Streamlit DASHBOARD... IN PROGRESS...
KAN-135 Understanding requirements DASHBOARD... IN PROGRESS...
KAN-137 Wireframe/prototype design DASHBOARD... IN PROGRESS...
KAN-136 Identify data sources DASHBOARD... IN PROGRESS...
KAN-24 Learn Cloud hosting on google cloud CLOUD IN PROGRESS...
KAN-6 Train machine learning model MODELING IN PROGRESS...

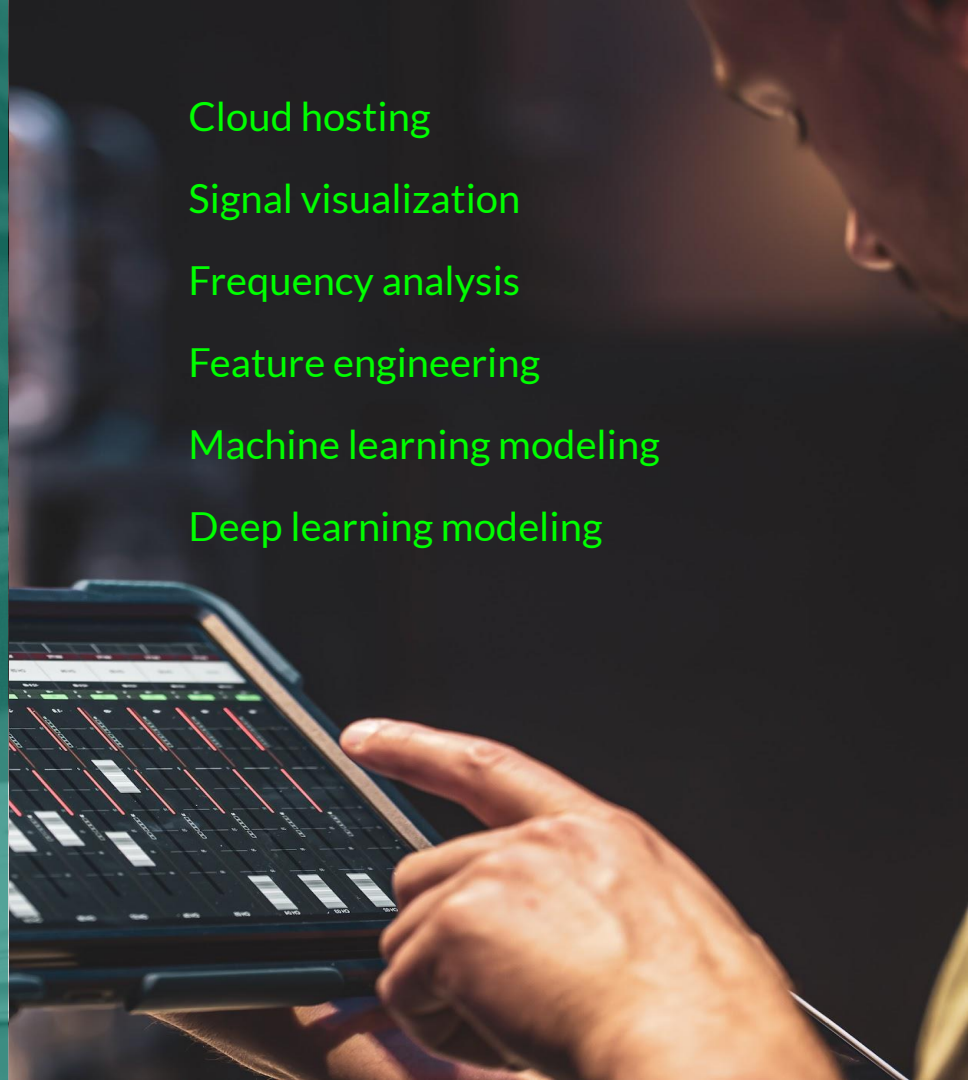

Epic
Search filters...
Issues without epic
Sprint 0
KAN-12
Setting up environment
KAN-39
Learning about tools
KAN-40
Learning the domain knowledge
KAN-41
Exploration
KAN-43
Modeling
KAN-126
Show epic panel
Press E to open and close

Tech stack



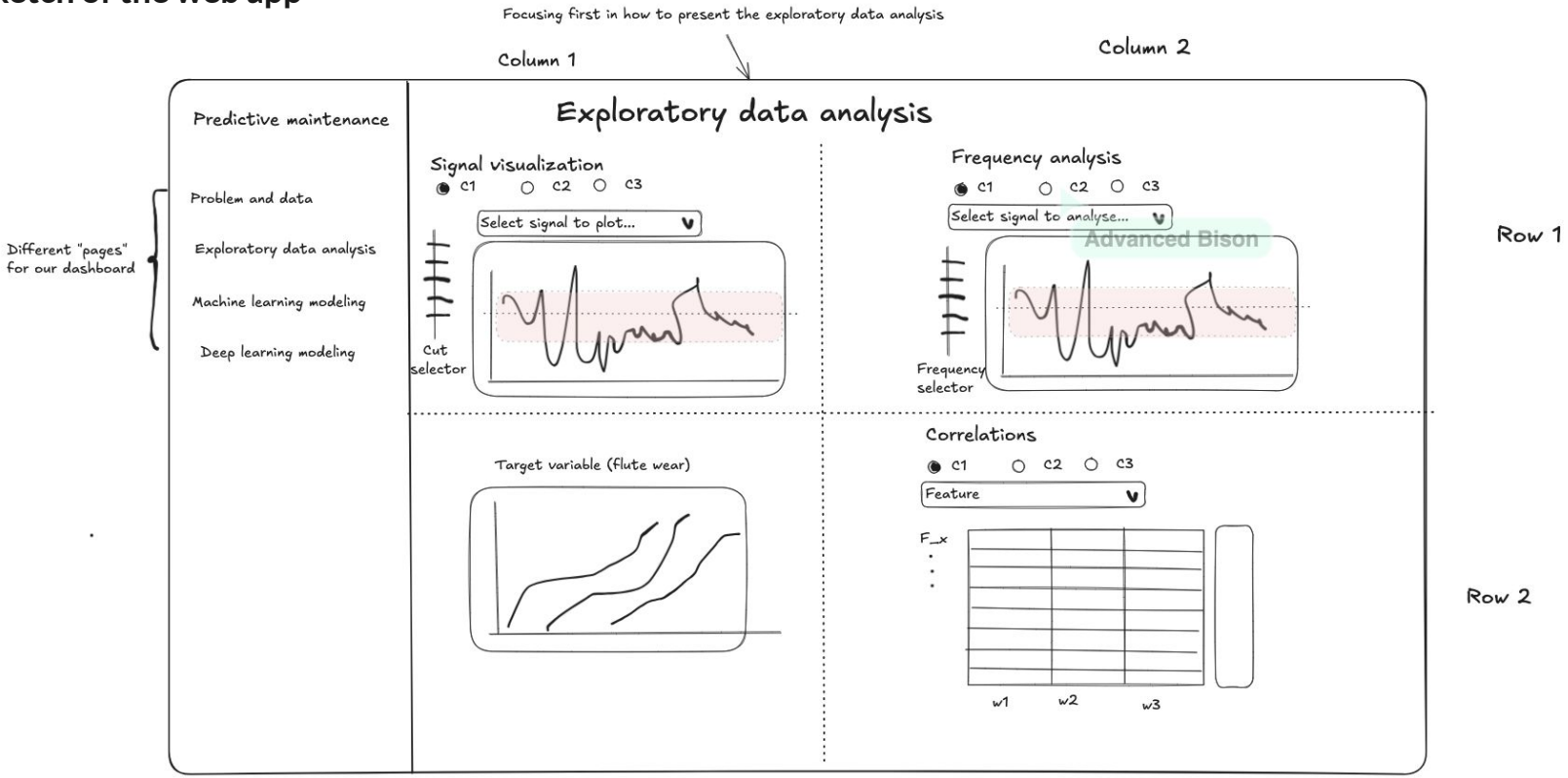
Other tools

MVP features



- Cloud hosting
- Signal visualization
- Frequency analysis
- Feature engineering
- Machine learning modeling
- Deep learning modeling

Sketch of the web app



Predictive maintenance

Problem and data

Exploratory data analysis

Machine learning modeling

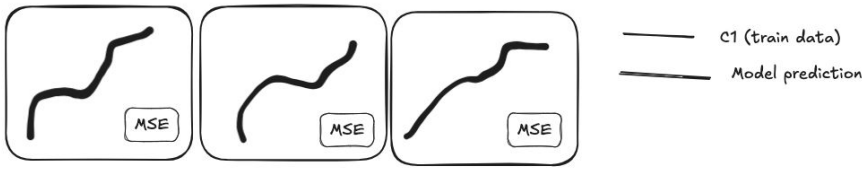
Deep learning modeling

Machine learning modeling

About the model:
* Features used
* Multi smothing lasso something

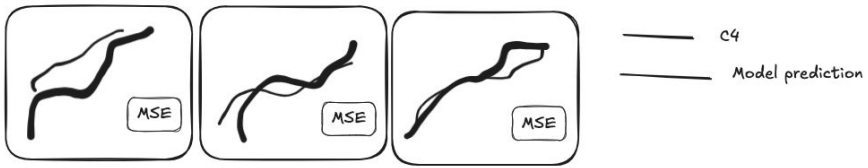
$y = Ax + B$

Train with C1

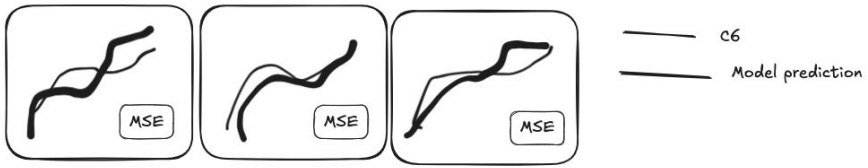


Predicting the other datasets

Predicting c4



Predicting c6



Summary of the project experience

- Lots of learning
- Difficult to predict tasks' time-frame
- Hard to keep Jira up-to-date
- Planning could have been more flexible and inclusive
- Communication and workflow improved as project developed
- One-on-one meetings were crucial
- Project was challenging and ambitious

Possible next steps

- Fine-tune the ML and DL models
  - use more of data to train
  - explore the parameter space further
- Deploy models to the cloud
- Simulate using the model for real-time monitoring
  - generate synthetic data in real-time
  - upload it to google drive
  - display it in the dashboard in real-time
  - run the model in real-time to add an estimate of the wear