SEWAGE PLANT ODOR MONITORING AND PREDICTION SYSTEM

Presented by: Team Blue Sky

OUR TEAM

Victor Opaleke

Model Performance

An Si-Eun

Project Researcher

Kasimov Azizjon

Project Manager

Min Woo Cho

Project Researcher

Sukhrob Ikromov

Data Scientist

OVERVIEW

01
Background

02
Project details

03
Development scope

04
Development process

05
System overview

06
System Architecture

07
Challenges

08
Development results

09
Completed tasks

BACKGROUND

The odor emitted from a sewage treatment plant is an important environmental problem. Bad odors can lead to public complaints, health issues, and reduce the quality of life in nearby areas. An estimation of odor emission rate is difficult to detect and quantify. The evaluation of odor emission involves huge manpower, time, and cost.

PROJECT DETAILS

Sensor technology

Sensor technology is used to detect concentrations of NH3 (ammonia), H2S (hydrogen sulfide), total volatile organic compounds (TVOCs) and other odor-causing compounds.

Weather synchronization

Predict odor distribution patterns using weather variables such as temperature, humidity, wind direction, and speed.

Predictive Analytics

Use machine learning algorithms to analyze time series data to increase the accuracy of odor intensity prediction.

ORIGINAL SCOPE

DESIGN GOAL 1

Implement a model that enables the prediction of odor levels from some features of a given dataset

DESIGN GOAL 2

Visualize the predictions of models and report their performance.

DEVELOPMENT PROCESS

Data Preprocessing

Normalize data, feature engineer, identify outliers and identify seasonality.

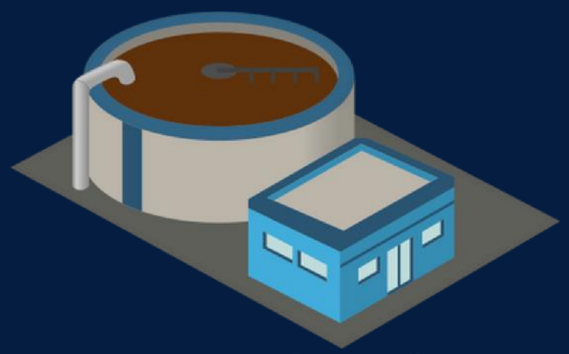
Algorithms

Implementing various prediction models.

Comparison

Comparing and choosing the best performing prediction algorithm

MODEL OPERATION



Sensory data from sewage plant

Prediction model

Visualization module

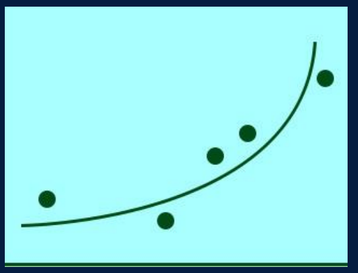
Anomaly notice

MODELS TESTED

XGBoost

Cone-shaped neural network

Linear regression

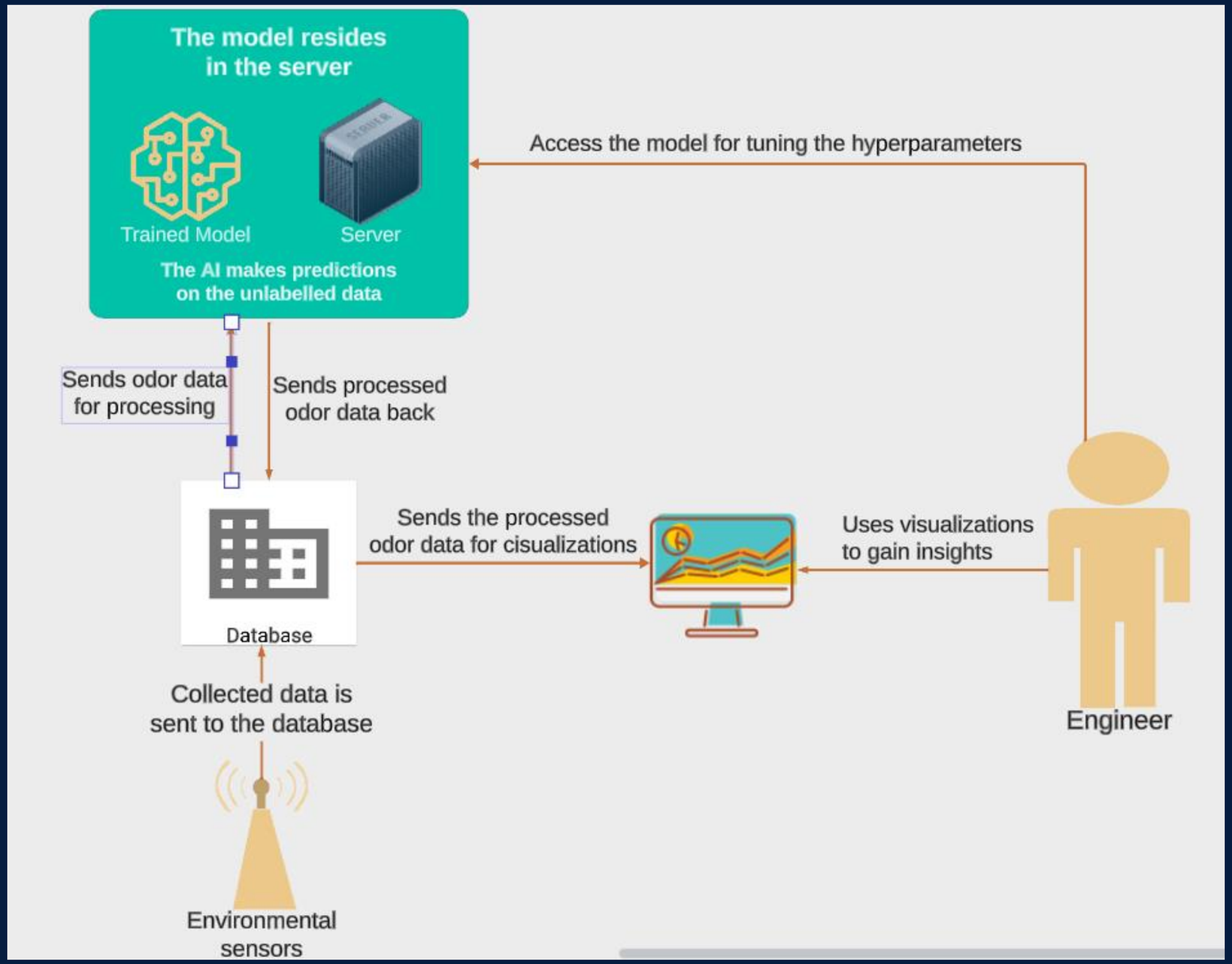


Polynomial regression

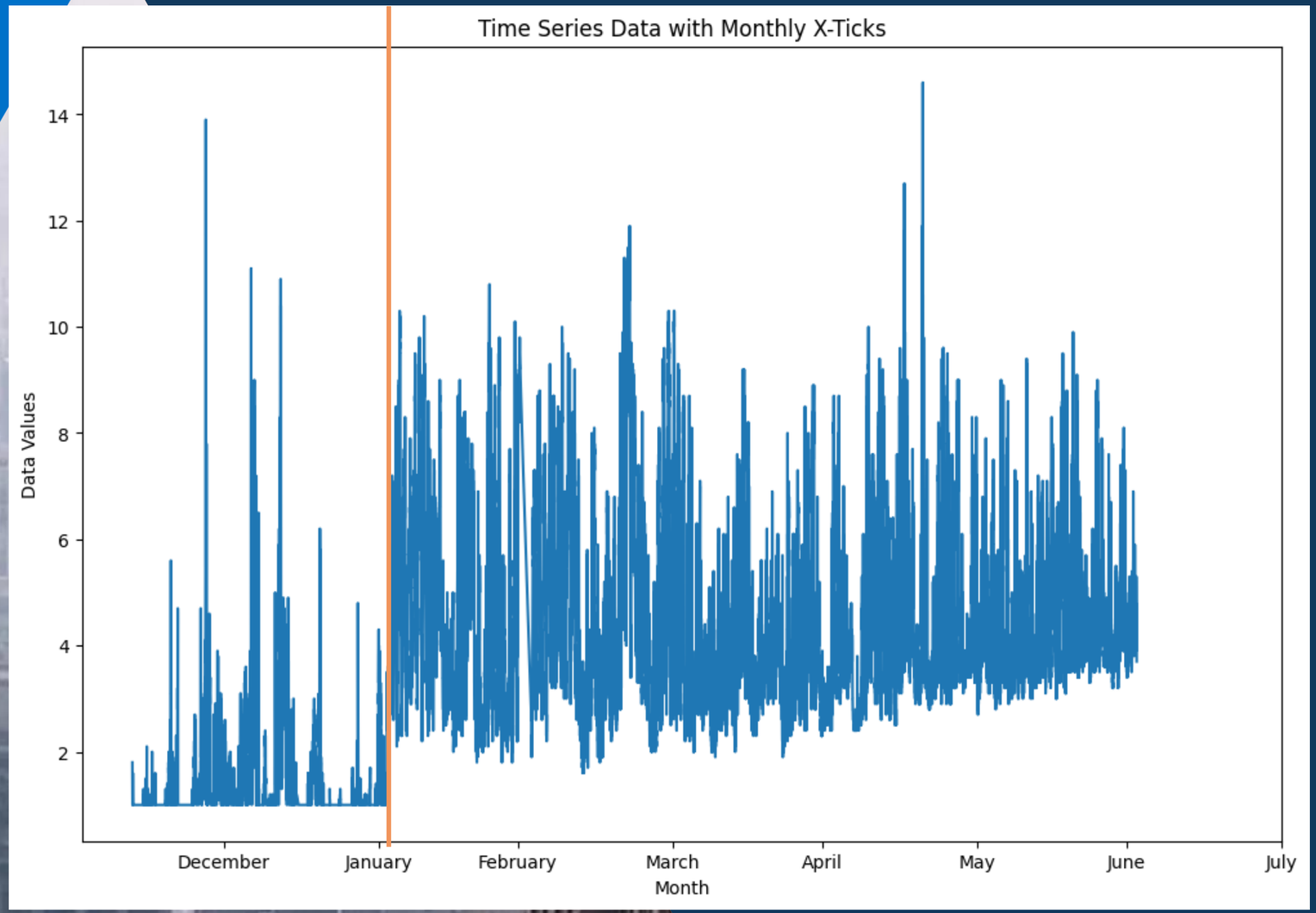
LST
M

Extreme randomized trees regressor

SYSTEM ARCHITECTURE DIAGRAM



Implementation problem

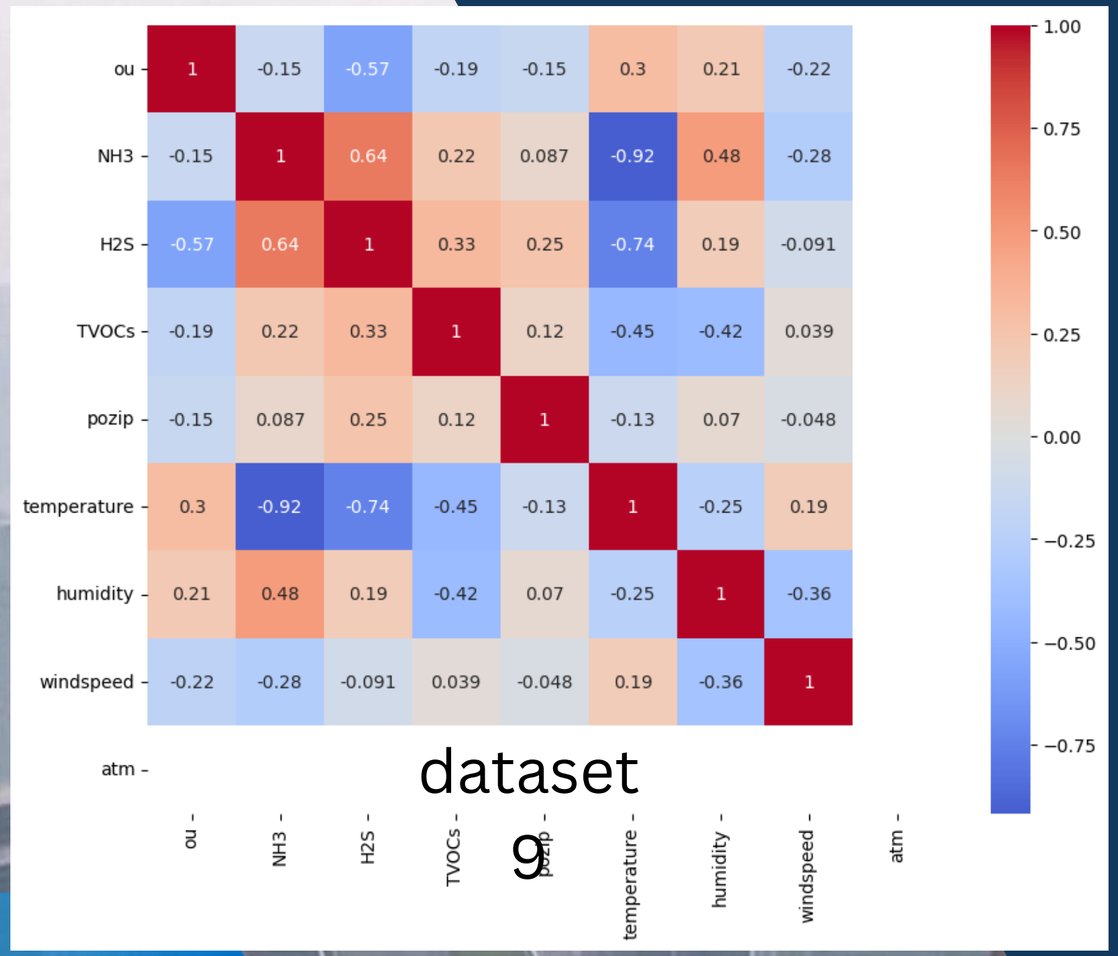


Implementation problem No. 2

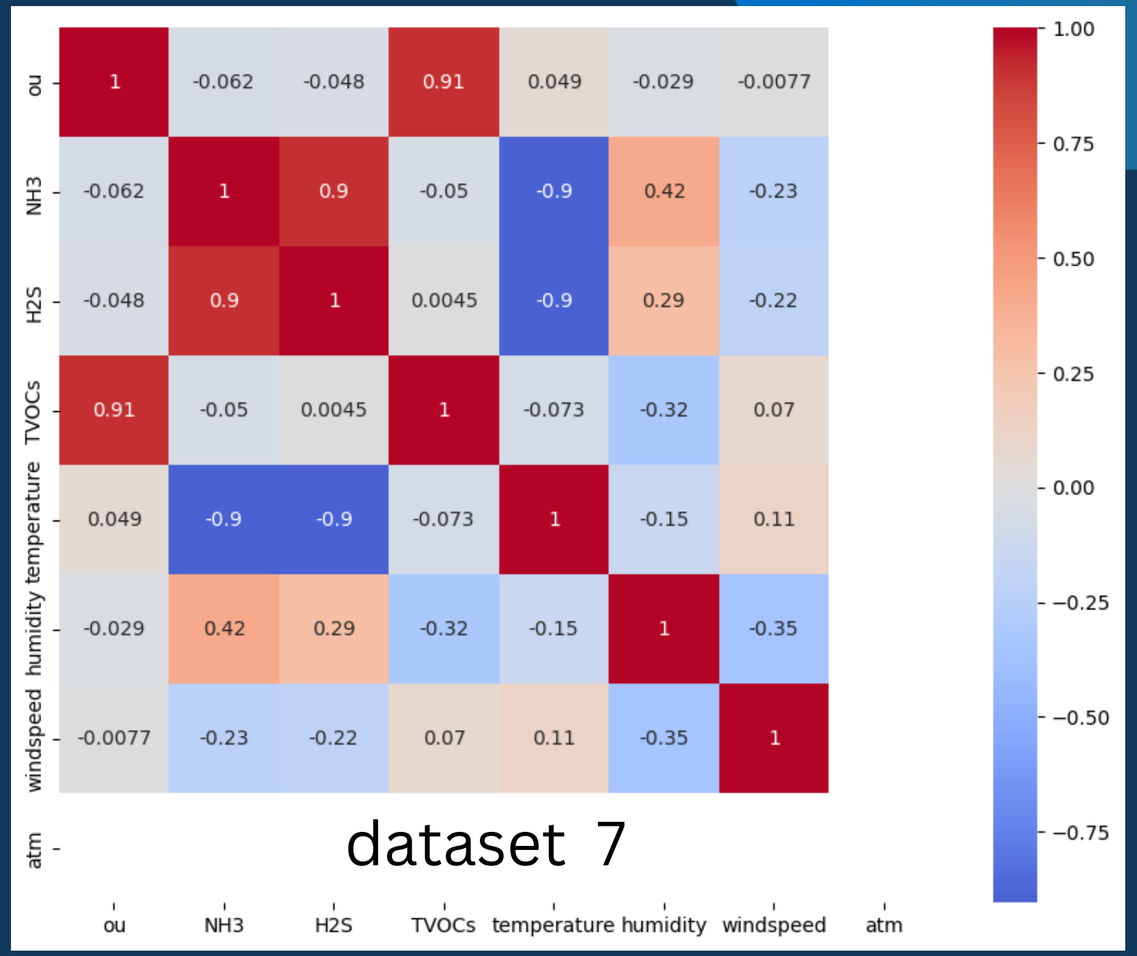
```
data[data["ou"] == 1.0].describe()
```

	ou	NH3	H2S	TVOCs	temperature	humidity
count	8854.0	8854.000000	8854.000000	8854.000000	8854.000000	8854.000000
mean	1.0	0.846675	0.272148	0.085093	-2.434448	6.210594
std	0.0	0.071850	0.031618	0.009110	5.593167	1.865886
min	1.0	0.638000	0.156000	0.018000	-16.200000	0.000000
25%	1.0	0.796000	0.250000	0.085000	-6.600000	4.900000
50%	1.0	0.838000	0.270000	0.086000	-2.100000	6.100000
75%	1.0	0.904000	0.296000	0.087000	1.900000	7.700000
max	1.0	1.014000	0.367000	0.110000	12.700000	10.000000

Data correlation

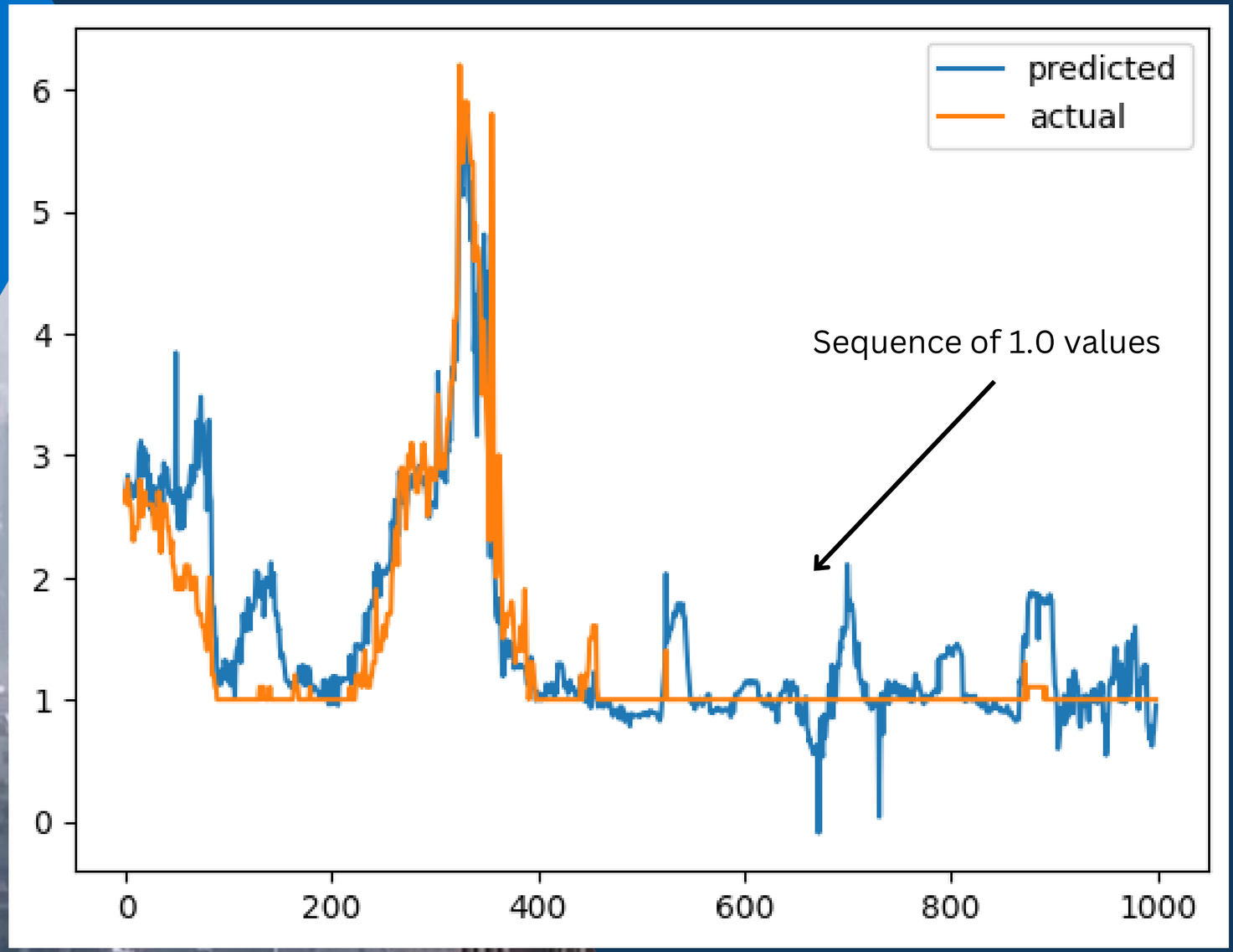


dataset 9



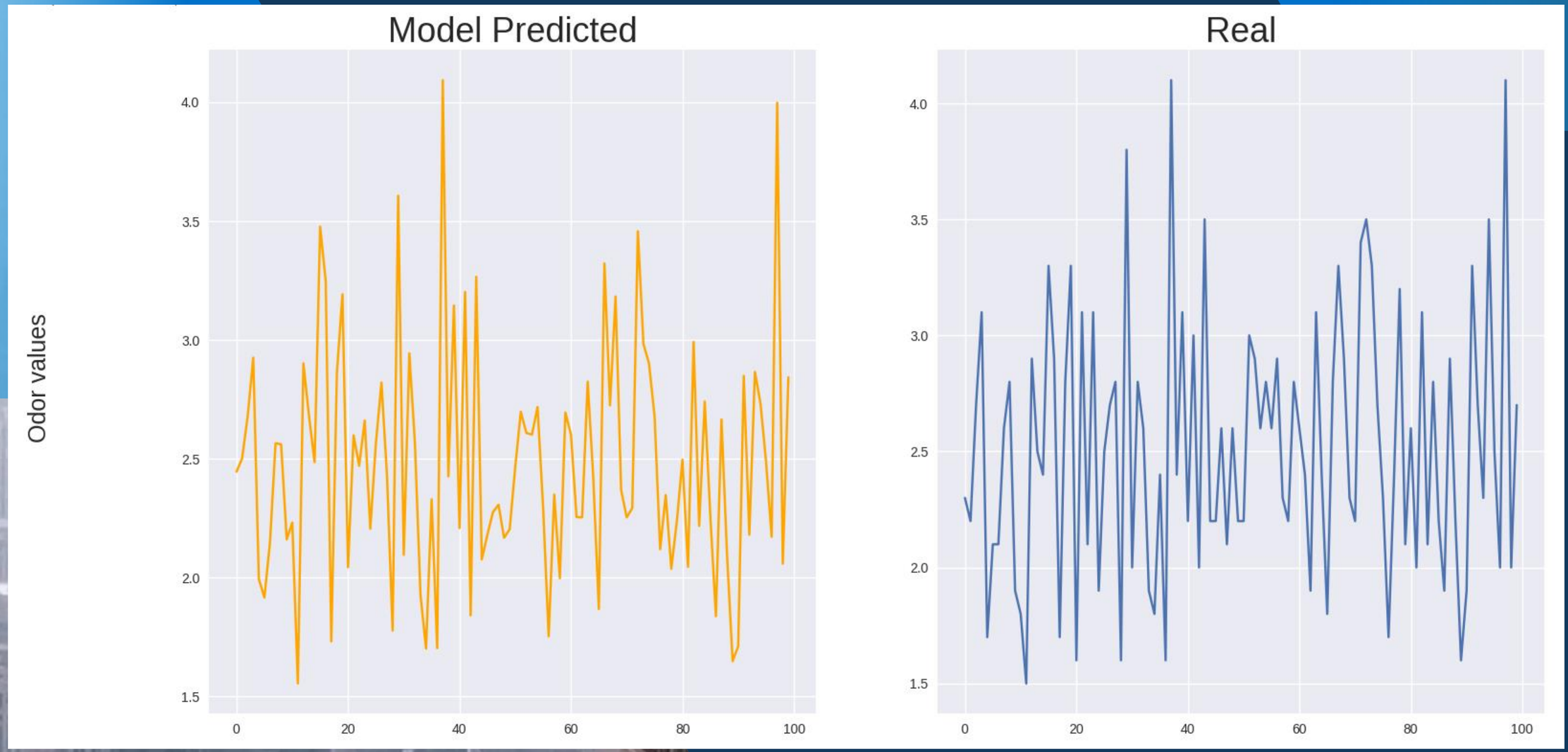
dataset 7

Implementation status



Polynomial regression.
R2 value: 0.8
MSE: 0.15

Development Result



Extra trees regressor model.

Prediction accuracy: 80%. Error value: 0.07

FACED CHALLENGES

STRUGGLING TO UNDERSTAND THE DATA

The provided data was hard to interpret and understand given features that were not strongly correlated with the target variable

LACK OF DATA

The labelled data had a period of less than a year, that created limitations to identify the seasonality

NO KNOWLEDGE ABOUT EQUIPMENT

Did not know the location and models of the sensors

COMPLETED TASKS

1. EXPLORED THE DATASET

We implemented statistical methods to find data distribution and feature correlation

2. STUDIED SIMILAR PROJECTS

We looked at studies done on the same problem worldwide. After studying them, we gained insights and proceeded our development.

3. PREPROCESING

Based on the papers, recommendations from mentor and general conventions we preprocessed the dataset

4. IMPLEMENTED PREDICTION ALGORITHMS

Based on the research papers and recomendations from mentor, we trained certain selection of algorithms

5. COMPARED PERFORMANCES

After training the models, we compared their performance and reported to our mentor.

Thank For Your Attention