Predictive Maintenance for Automatic Production Systems

CSE 495
2nd Presentation

Abdurrahman BULUT

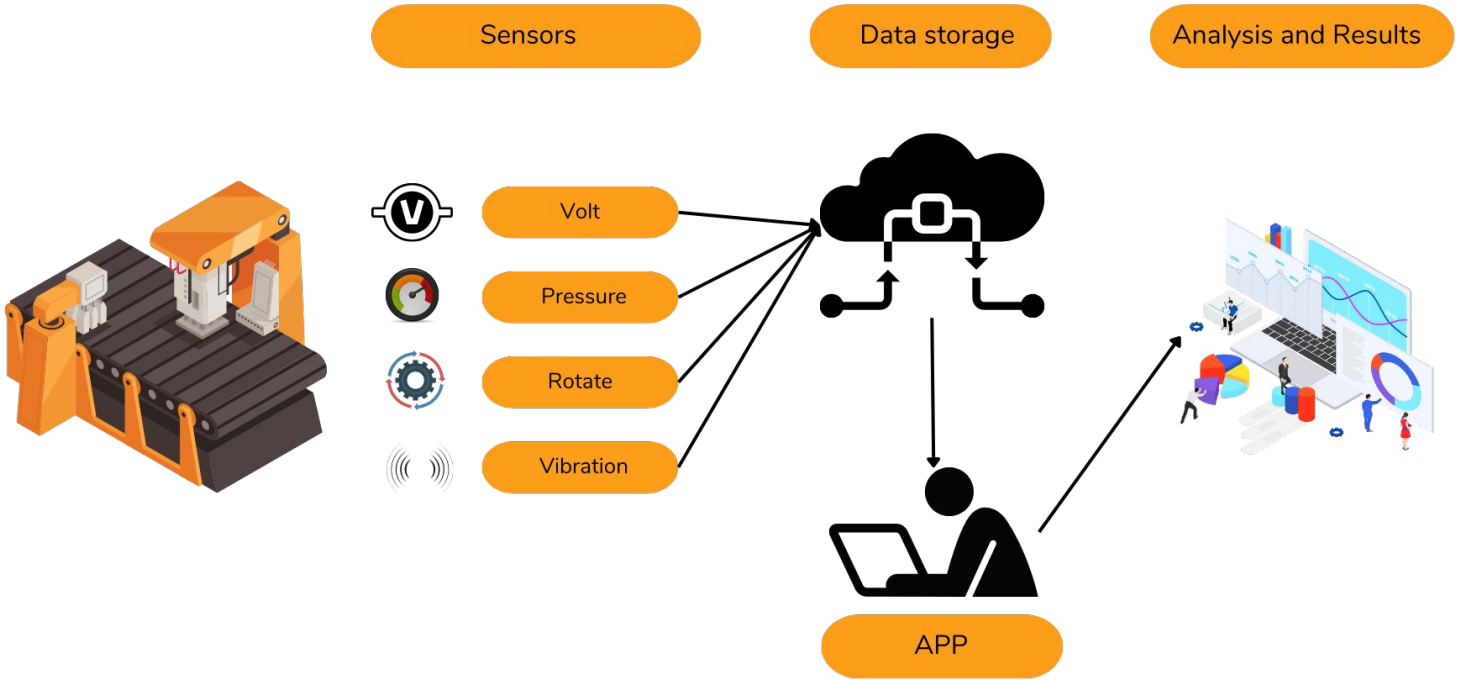
Supervisor: Pr. Dr. İbrahim SOĞUKPINAR

June 2023

Contents

- Project Definition
- The Dataset
- EDA
- Data Pre-processing
- Modelling
- Project Timeline
- Success Criterias
- Resources

Project Definition



The Dataset

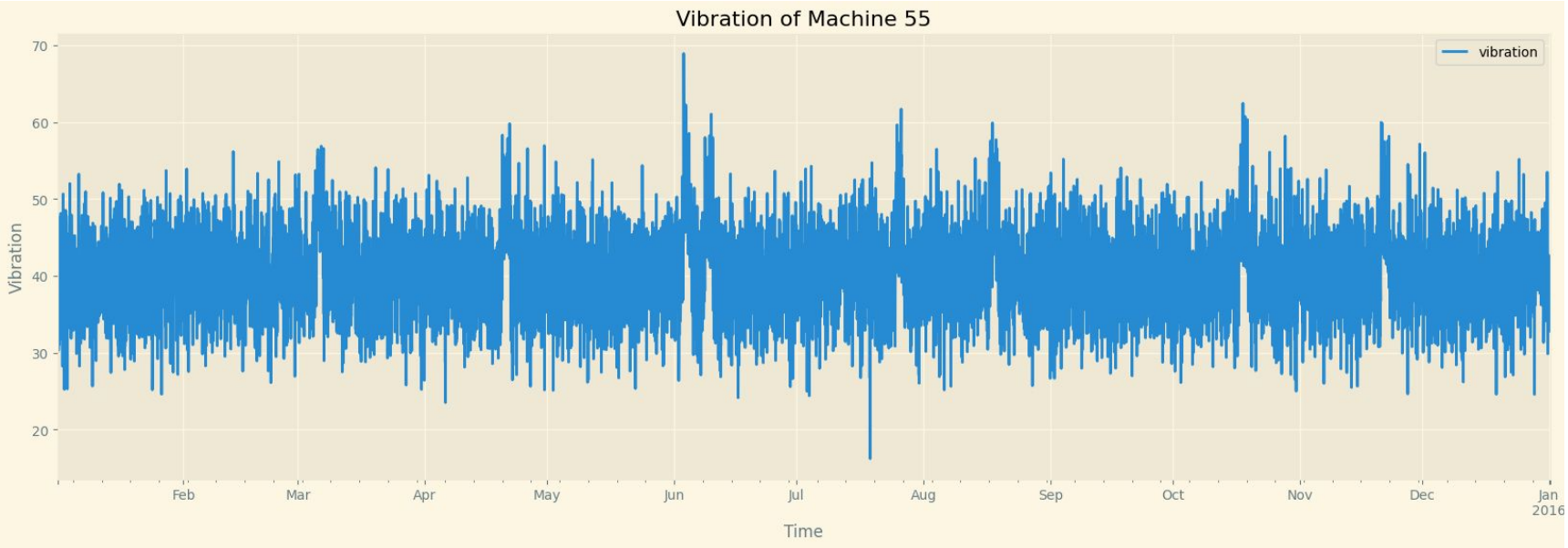
There are 5 CSV Files

The dataset is available in Kaggle. This dataset was available as a part of Azure AI Notebooks for Predictive Maintenance.

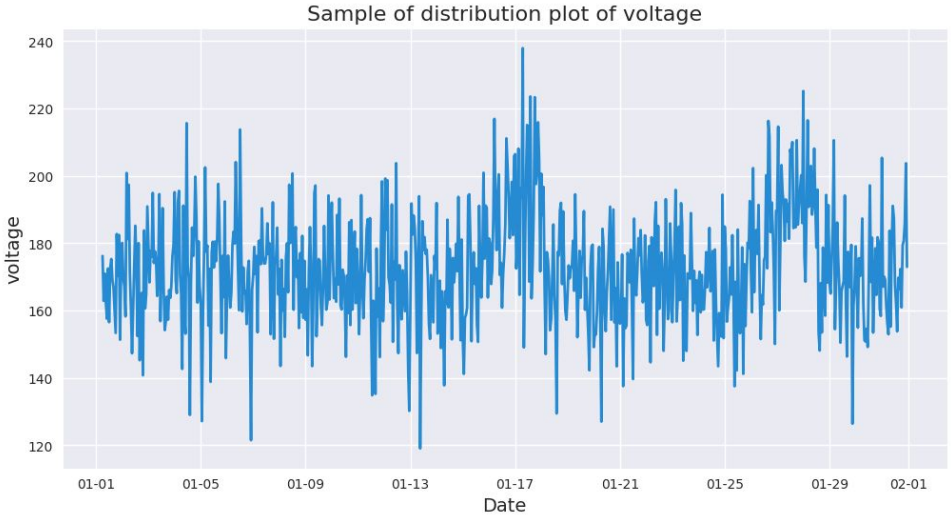
01	**Telemetry Time Series Data (PdM_telemetry.csv)**	It consists of hourly average of voltage, rotation, pressure, vibration collected from 100 machines for the year 2015.
02	**Error (PdM_errors.csv)**	These are errors encountered by the machines while in operating condition.
03	**Maintenance (PdM_maint.csv)**	If a component of a machine is replaced, that is captured as a record in this table
04	**Failures (PdM_failures.csv)**	Each record represents replacement of a component due to failure.
05	**Metadata of Machines (PdM_Machines.csv)**	Model type & age of the Machines.

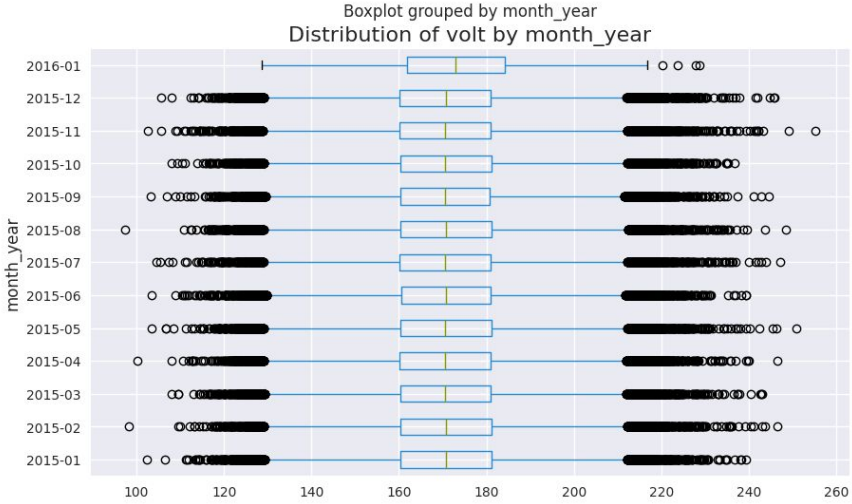
EDA

	datetime	machineID	volt	rotate	pressure	vibration
count	876100	876100.000000	876100.000000	876100.000000	876100.000000	876100.000000
mean	2015-07-02 18:00:00	50.500000	170.777736	446.605119	100.858668	40.385007
min	2015-01-01 06:00:00	1.000000	97.333604	138.432075	51.237106	14.877054
25%	2015-04-02 12:00:00	25.750000	160.304927	412.305714	93.498181	36.777299
50%	2015-07-02 18:00:00	50.500000	170.607338	447.558150	100.425559	40.237247
75%	2015-10-02 00:00:00	75.250000	181.004493	482.176600	107.555231	43.784938
max	2016-01-01 06:00:00	100.000000	255.124717	695.020984	185.951998	76.791072
std	NaN	28.866087	15.509114	52.673886	11.048679	5.370361

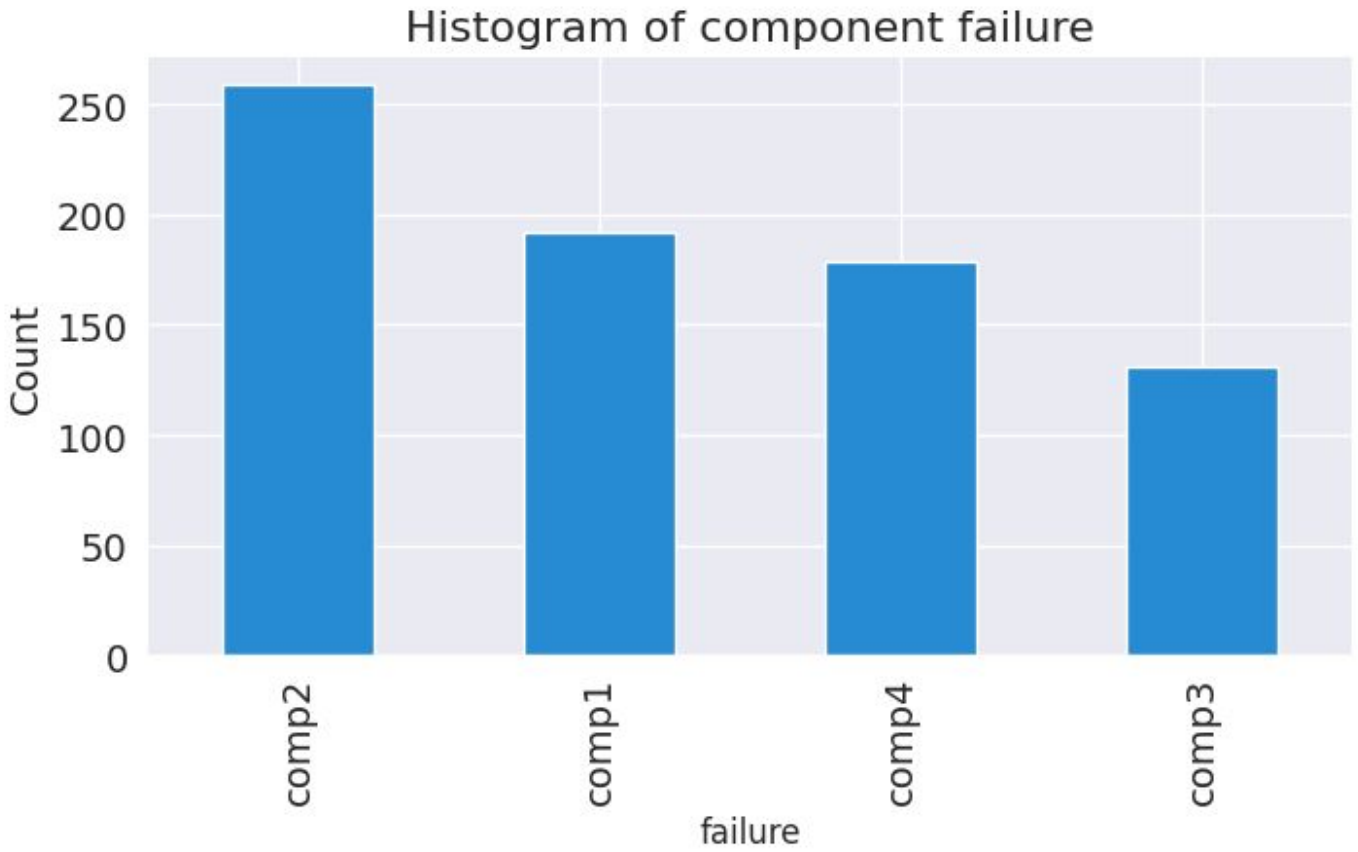


EDA


Sample of distribution plot of voltage
voltage
Date
240
220
200
180
160
140
120
01-01
01-05
01-09
01-13
01-17
01-21
01-25
01-29
02-01


Boxplot grouped by month_year
Distribution of volt by month_year
month_year
2016-01
2015-12
2015-11
2015-10
2015-09
2015-08
2015-07
2015-06
2015-05
2015-04
2015-03
2015-02
2015-01
100
120
140
160
180
200
220
240
260

EDA


Histogram of component failure
Count
250
200
150
100
50
0
comp2
comp1
comp4
comp3
failure

EDA

- There are 5 numbers of errors (total errors: 3919) are encountered by machines
- Type 1 & 2 errors are most frequent which is more than double the numbers of error-5
- The age of Machines is distributed between 0 to 20 years. The median age is 12.5 years. Around 75% of machines age is less than 16 years and around 25% of machines age is less than 7 years.
- There are total 761 numbers of failures are stated in the Failures data set. The most failures happened due to component-2.

Data Pre-processing

- Failure columns added to the telemetry data
- Model and age added to the telemetry data
- Nan values, duplicated values are checked.
- datetime is converted from string to datetime.
- Comp1,Comp2,Comp3,Comp4 means there is a problem with the related component.
- Comp0 means there is no failure.

Data Pre-processing

Lag Features

	machineID	datetime	volt_min_3h	rotate_min_3h	pressure_min_3h	vibration_min_3h
0	1	2015-01-01 09:00:00	162.879223	402.747490	75.237905	34.178847
1	1	2015-01-01 12:00:00	157.610021	346.149335	95.927042	25.990511
2	1	2015-01-01 15:00:00	156.556031	398.648781	101.001083	35.482009
3	1	2015-01-01 18:00:00	160.263954	382.483543	96.480976	38.543681
4	1	2015-01-01 21:00:00	153.353492	402.461187	86.012440	39.739883

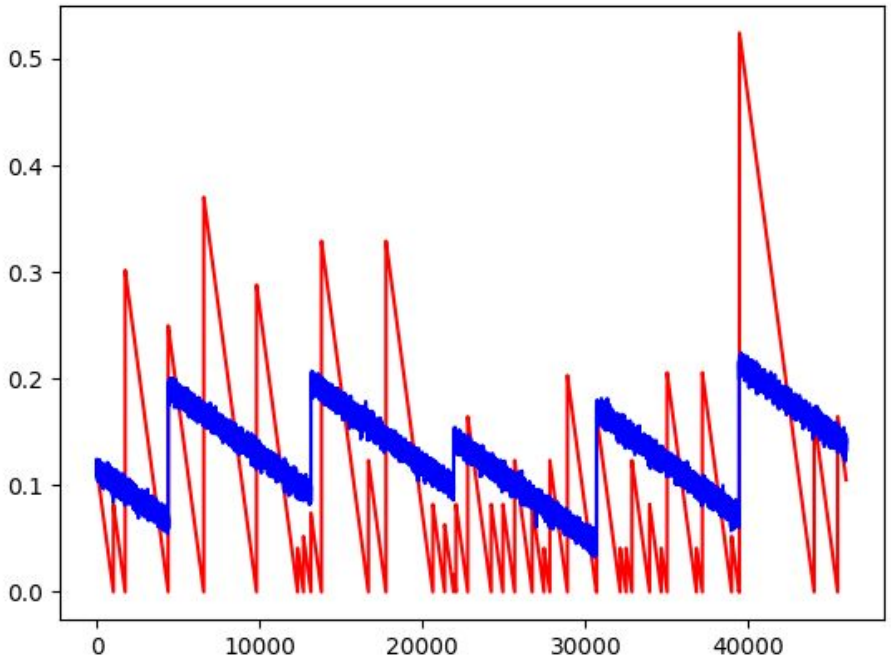
```
telemetry_feat.describe()
```

✓ 0.1s

	machineID	datetime	volt_min_3h	rotate_min_3h	pressure_min_3h	vibration_min_3h	volt_max_3h	rotate_max_3h	pressure_max_3h	vibration_max_3h
count	292100.000000	292100	292100.000000	292100.000000	292100.000000	292100.000000	292100.000000	292100.000000	292100.000000	292100.000000
mean	50.500000	2015-07-02 20:59:59.999999744	158.083149	404.146613	92.384876	36.149107	183.473398	489.026431	109.348395	44.627644
min	1.000000	2015-01-01 09:00:00	97.333604	138.432075	51.237106	14.877054	134.008631	237.641009	75.559302	28.581727
25%	25.750000	2015-04-02 15:00:00	150.370836	379.505260	86.862827	33.457643	175.270539	463.319490	103.502795	41.787201
50%	50.500000	2015-07-02 21:00:00	158.172629	406.891158	92.118675	36.088473	182.769923	489.161104	108.505756	44.291338
75%	75.250000	2015-10-02 03:00:00	165.846384	432.062323	97.281361	38.667734	190.895306	515.375351	114.007202	47.039826
max	100.000000	2016-01-01 09:00:00	235.726785	565.962115	160.026994	68.001841	255.124717	695.020984	185.951998	76.791072
std	28.866119	NaN	11.878952	40.828783	8.783319	4.210114	11.909021	40.802190	8.863881	4.220441

Modelling

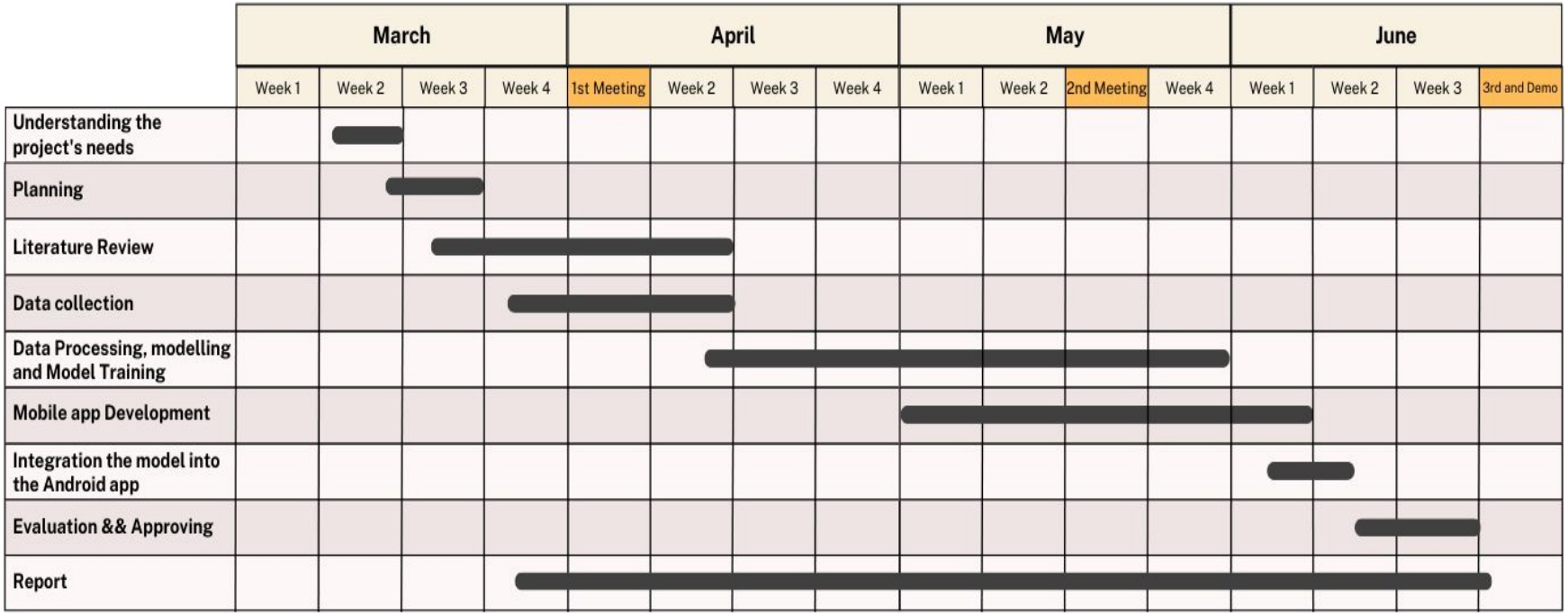
- Linear Regression
  - 70% train, 30% test
  - Data Normalized using MinMaxScaler
  - y test size: 45997
  - second_to_fail



Modelling

- Linear SVC
  - Independent variables: age, volt values, rotate values, pressure values, vibration values
  - Dependent Variables: Failure informations
  - 80% -> training
  - 20% -> test data
  - Accuracy is 89%

Roadmap


March
April
May
June
Week 1
Week 2
Week 3
Week 4
1st Meeting
Week 2
Week 3
Week 4
Week 1
Week 2
2nd Meeting
Week 4
Week 1
Week 2
Week 3
3rd and Demo
Understanding the project's needs
Planning
Literature Review
Data collection
Data Processing, modelling and Model Training
Mobile app Development
Integration the model into the Android app
Evaluation && Approving
Report

Success Criterias

- The success rate of the created model is at least %80
- Calculation of analysis results in less than 20 miliseconds
- 2 file types should be supported to upload the data. (.xlsx - .csv)

Resources

- “Microsoft Azure Predictive Maintenance”, ARNAB, “https://www.kaggle.com/datasets/arnabbiswas1/microsoft-azure-predictive-maintenance”

- “Predictive Maintenance using Machine Learning”, Medini Kumar Bora, “https://medium.com/@Medini_2020/predictive-maintenance-using-machine-learning-3d8b62d5df8e”

- “Artificial intelligence for fault diagnosis of rotating machinery: A review, “https://www.researchgate.net/publication/326742898_Artificial_intelligence_for_fault_diagnosis_of_rotating_machinery_A_review”

- “Fault Handling in Industry 4.0: Definition, Process and Applications”, “https://www.mdpi.com/1424-8220/22/6/2205”

- “An Industry 4.0 Dataset of Contextual Faults in a Smart Factory”, “https://www.sciencedirect.com/science/article/pii/S187705092100314 8?via%3Dihub”