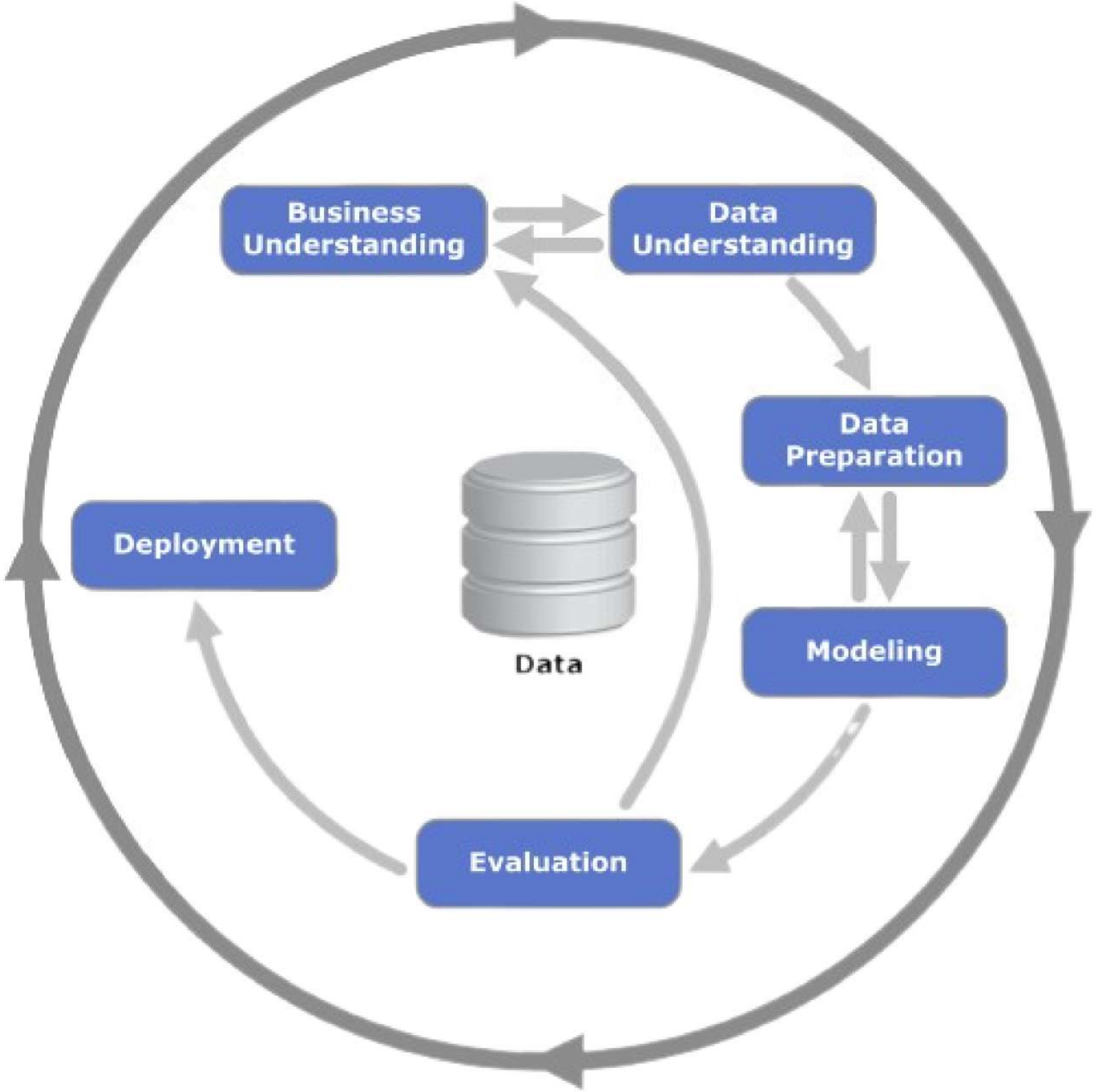
The Champions

CRISP-DM

a process model that serves as the base for a data science process.



The Champions

Business Understanding

Problem	Effect
Sebuah perusahaan X sering terjadi kerusakan mesin (machine failure) tak terduga pada saat berjalannya produksi.	Muncul biaya - biaya tambahan perbaikan, sehingga perusahaan dinilai belum layak untuk melakukan penambahan kapasitas produksi dan melakukan ekspansi bisnis ke pasar mancanegara.

DATA UNDERSTANDING

Dataset in General

The Champions

10.000 data

14 feature

5 independent failures

- UID
- product ID
- type
- air temperature [K]
- process temperature [K]
- Machine failure
- rotational speed [rpm]
- torque [Nm]
- tool wear [min]

Power Label

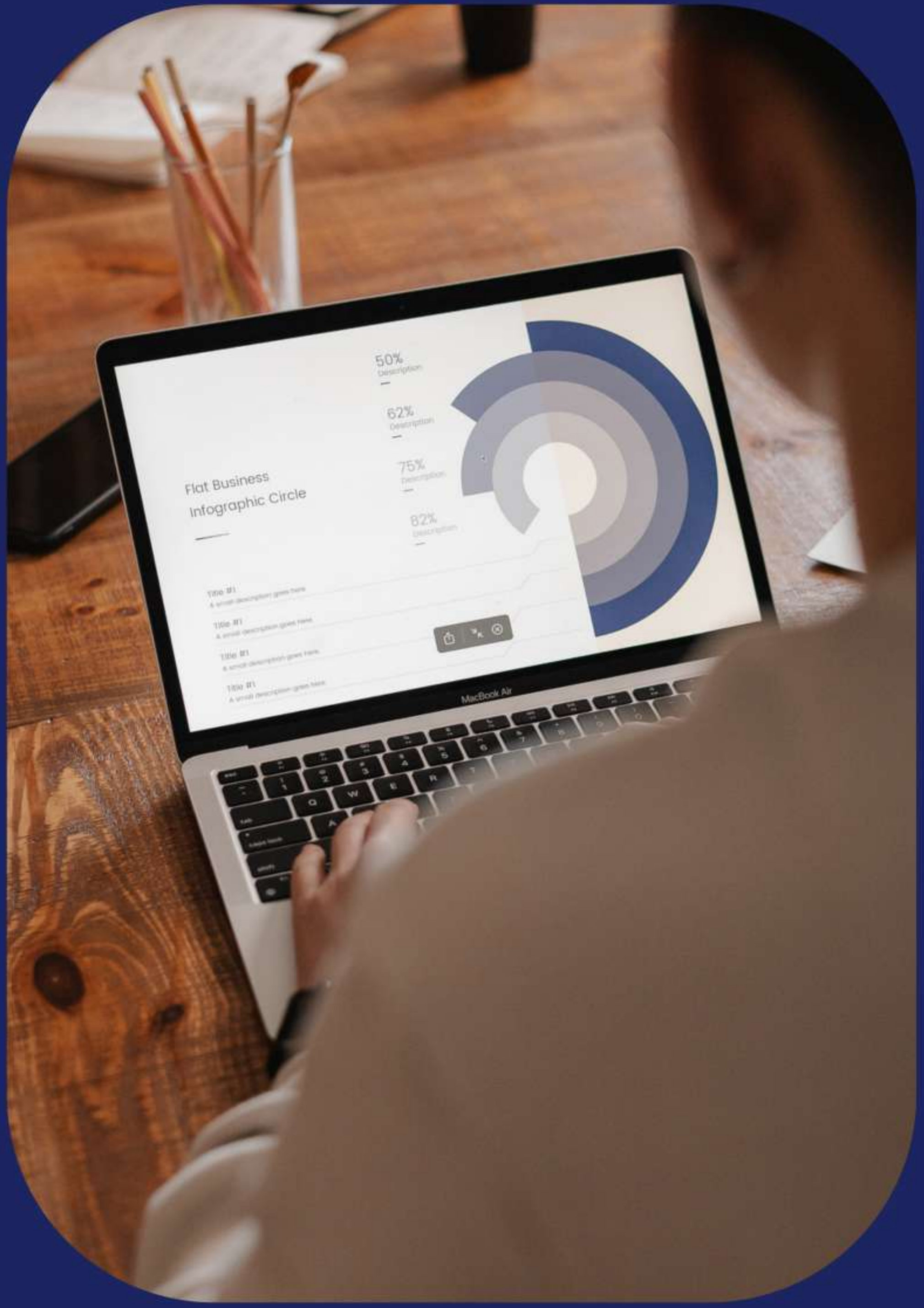
Label Power Set — Multilabel to Multiclass

- TWF (tool wear failure)
- HDF (heat dissipation failure)
- PWF (power failure)
- OSF (overstrain failure)
- RNF (random failures)
- Another Failure

The Champions

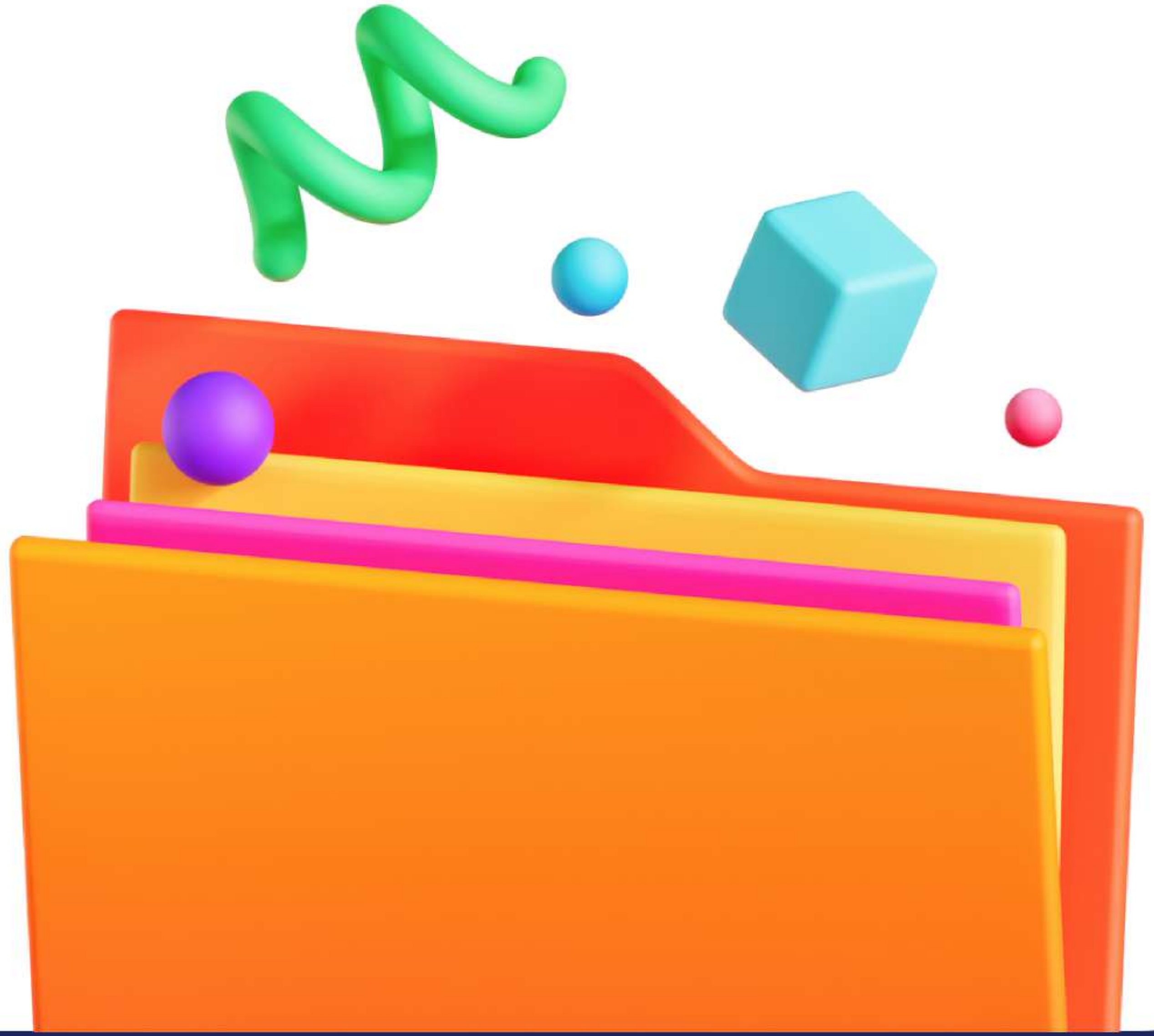
DATA UNDERSTANDING

EDA



- No missing value.
- No duplicate data.
- Many outlier values (but in machine failure category 1). Therefore, this outlier value can be left for analysis purposes, being transformed, or separated into new columns.
- There are several machine failures that have more than 1 independent failure category
- There are 9 data that fall into the category of machine damage based on "Machine Failure", but do not fall into the five independent damage categories.
- There is an imbalanced class.

DATA PREPARATION



1. Dummy Variable
2. New Class
3. Problem Transformation
4. Feature Selection
5. Data Transformation
6. Handling Imbalanced Dataset
7. Hyperparameter Tuning

The Champions

MODELING

The selected model is the model with the highest Recall value

Model	MLPClassifier	RandomForestClassifier	LGBMClassifier	DecisionTreeClassifier	XGBClassifier	LogisticRegression	SVC	GaussianNB	RidgeClassifier	AdaBoost Classifier
Accuracy	0.960333	0.977	0.979667	0.948333	0.976667	0.380667	0.641667	0.030333	0.078	0.002
Recall	0.362839	0.328871	0.28157	0.267313	0.257733	0.189284	0.170234	0.12939	0.090134	0.001237
Precision	0.389869	0.255542	0.2557	0.26318	0.25517	0.356303	0.364393	0.126899	0.195592	0.164167
F1-Score	0.354589	0.280579	0.267215	0.235179	0.255874	0.184003	0.189412	0.047569	0.034081	0.002359

Evaluation

Train vs Test

	Before Hyperparameter Tuning	After Hyperparameter Tuning
Train	0.998477	0.999840
Test	0.958000	0.970400

Evaluation

Classification Report

Before Hyperparameter Tuning

Class	Precision	Recall	F1-Score	Support
0	0.99	0.97	0.98	2422
1	0	0	0	2
2	0	0	0	6
3	0.68	1	0.81	15
4	0.67	0.82	0.74	17
5	1	0.25	0.4	4
6	0.47	0.8	0.59	20
8	0	0	0	2
9	0.12	0.45	0.19	11
11	0	0	0	1
12	0	0	0	0
Accuracy			0.96	2500
Macro Average	0.36	0.39	0.34	2500

After Hyperparameter Tuning

Class	Precision	Recall	F1-Score	Support
0	0.99	0.98	0.98	2422
1	0	0	0	2
2	0	0	0	6
3	0.67	0.8	0.73	15
4	0.58	0.65	0.61	17
5	0.75	0.75	0.75	4
6	0.54	0.7	0.61	20
8	0	0	0	2
9	0.21	0.27	0.24	11
11	0	0	0	1
Accuracy			0.97	2500
Macro Average	0.37	0.42	0.39	2500

Evaluation

Value

Class	Data Test	Before Hyperparameter Tuning	After Hyperparameter Tuning
0	2422	2362	2410
1	2	8	4
2	6	7	5
3	15	22	18
4	17	21	19
5	4	1	4
6	20	34	26
7	0	0	0
8	2	1	0
9	11	41	14
10	0	0	0
11	1	1	0
12	0	0	0

Conclusion

From the value of hyperparameter tuning and metrics evaluation, the results given by the model between before HT and after HT did not have major differences, but when viewed the values in each model class, model with HT were less affected by the majority class. So, we will use the model from Hyperparameter Tuning and this model gives the best recall too.

Evaluation

AUC Best Model

Class	0	1	2	3	4	5	6	7	8	9	10	11	12
AUC	0.81887	0.4992	0.499	0.89879	0.82192	0.8748	0.84758	0	0.5	0.63415	0	0.5	0
Support	2422	2	6	15	17	4	20	0	2	11	0	1	0

Judging from the AUC value of each class, it can be said that the model can classify well (there are 5 classes with AUC values above 80%). However, for some classes due to certain conditions, such as a small number of class members, the AUC value in the class is also low, in this case the model can be said to still work well because it can still classify even in conditions with a small amount of data.

The Champions

Feature Importance

We can't say specifically which feature gives the most effect because neural network is more like a black box model. In this case we only show feature importance based on spesific row.

Feature	Feature Value	Importance
Tool Wear	-0.32	0.01
Torque	-1.24	0.17
Process Temperature	-1.72	0.1
Rotational Speed	0.81	0.14
Air Temperature	1.82	0.02
Type_L	-5.2	0.19
Type_H	5.2	0.03

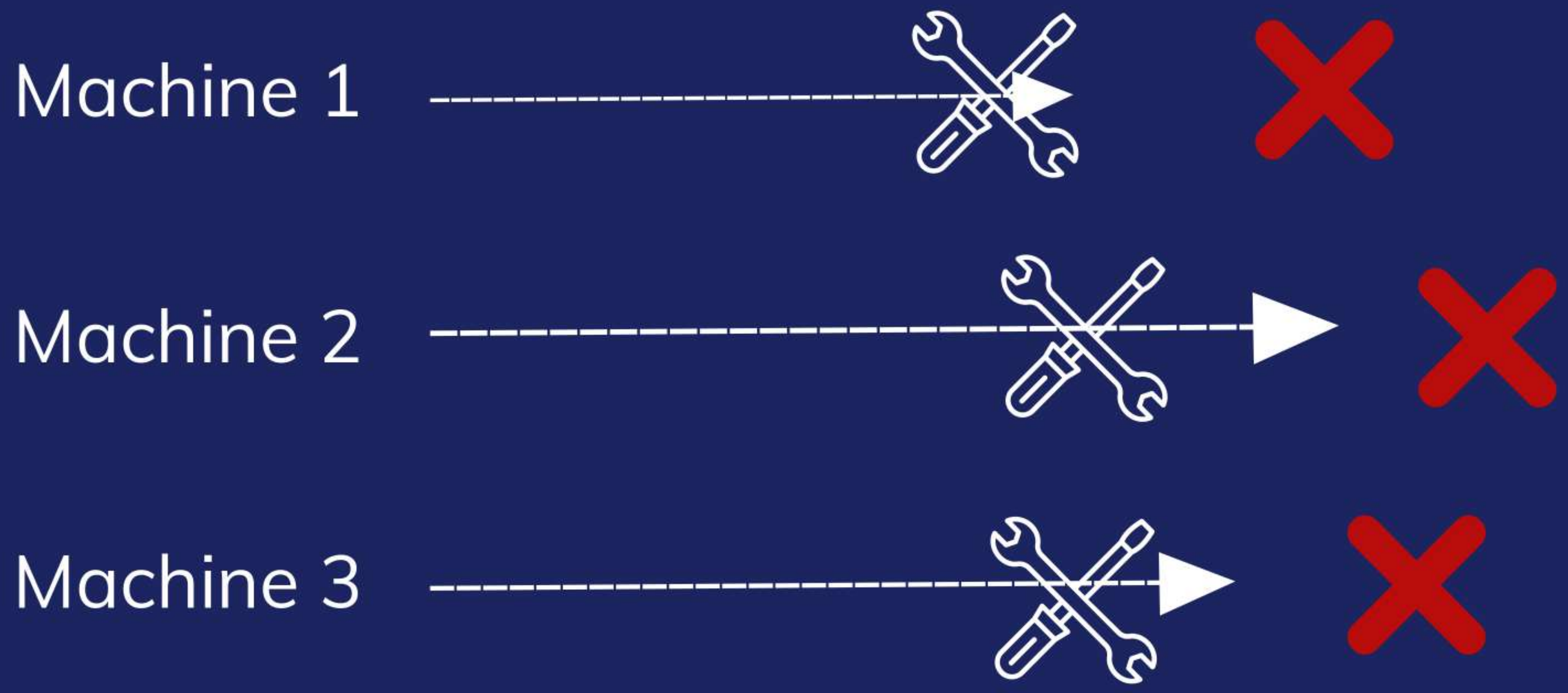
Row 450

Prediction Results

Maintenance Strategy

Predictive Maintenance

Fixes issues before they occur



Back to Agenda

Maintenance Strategy

Predictive maintenance for fixes issues before they occur.

Advantages

- Lowering the operating risk
- Control cost of maintenance
- Find trends related to various maintenance issues.

Open New Business Strategies

- By providing Predictive Maintenance services for its products, a company can proposes extra value for its consumers.

Common Ways To Reach Predictive Maintenance

- Internet of Things
- Cloud Computing
- Artificial intelligence (AI)

Back to Agenda

THANK YOU