NTNU

Norwegian University of Science and Technology

Fault detection, classification, and root cause identification for a manufacturing production line setup

Muhammad Gibran Alfarizi

NTNU

Outline

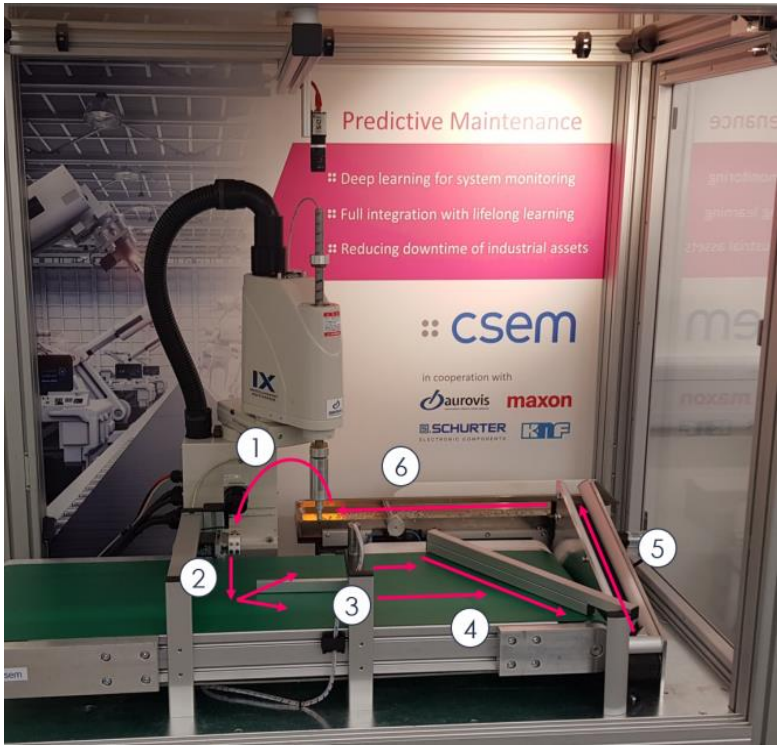
- System Description
- Dataset
- Objectives
- Data Preprocessing
- Methodology
- Results

System Description

On this fuse-test-bench, first it assesses whether the fuse is conducting electricity.

If the first test was successful, the fuse gets heated up by applying a current of 200mA for a time interval of 1.5s.

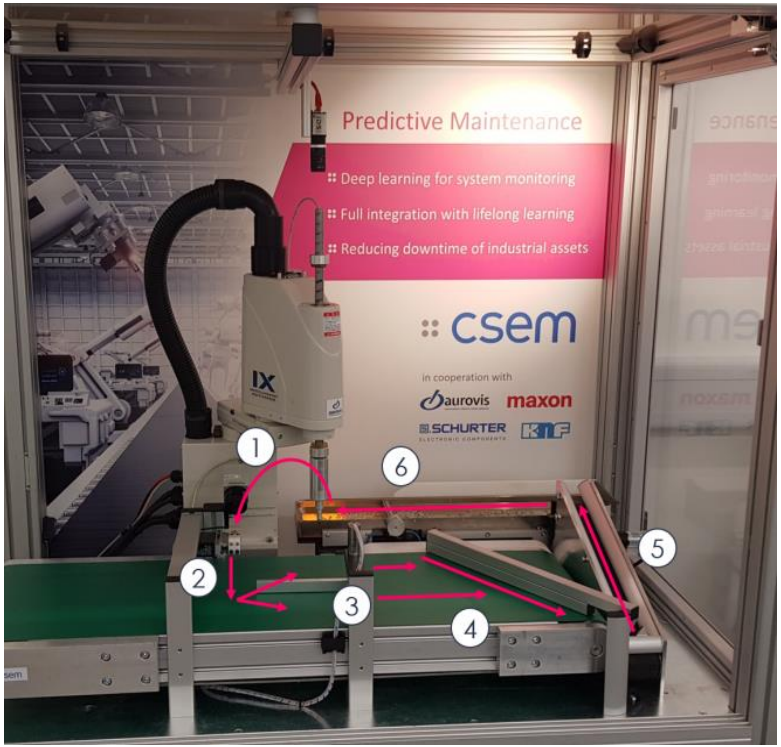
The heating up is measured by a thermal camera. After the tests, the fuse is moved back into the feeder with two conveyor belts.



System Description

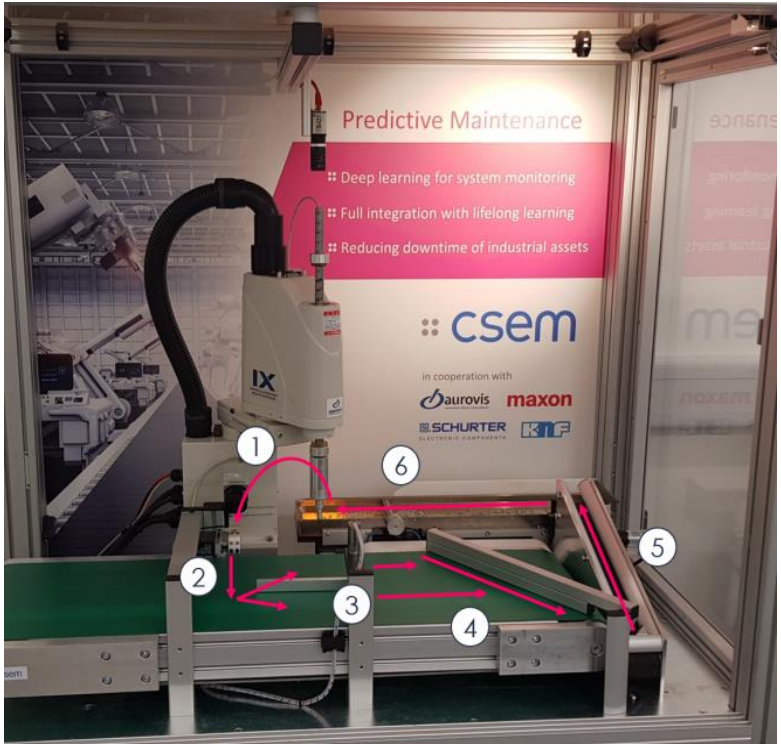
- The machine is monitored by an array of 50 sensors recording the evolution of a number of quantities of interest to establish the health state of the machine in real-time.

- The sensor data is aggregated over a time window of 10s, and one statistical data point is calculated.



System Description

- 4-axis SCARA robot
- fuse feeder,
- thermal camera (382×288 pixel, 0-250 deg C)
- camera to detect fuses on feeder (1280×1024 pixel)
- EC motor for big conveyor belt
- EC motor for fuse selector
- DC motor for small conveyor belt
- Vacuum pump for robot gripper
- Pressure pump for feeder barrier
- several valves for the pneumatics system



System Description

- Under its nominal working regime, there are no throughput defects at any level of the quality-control-pipeline, from picking-up the fuses to their transportation and analysis.

- However, different artificial failure modes can be artificially injected by manually altering the behaviour of one or more components. For instance:
  1. Modification of the operating mode of the robotic arm picking up the fuses
  2. Introduction of a pressure leakage on the pneumatic system
  3. Altering the speed of the conveyor belts

- In total, 8 failure modes are introduced

Dataset

- The experimental dataset is composed of a set of 50 signals. Each experiment can run from ≈ 1 to ≈ 3 hours.
- These signals can be divided into three categories:
  - Machine health monitoring signals: Pressure, Vacuum, FuseHeatSlope, …
  - Environment monitoring signals: Temperature, Humidity, …
  - Others: CPUTemperature, ProcessMemoryConsumption, …
- Statistical measures calculated every 10 seconds:
  - vCnt (number of samples recorded in a fixed time-window)
  - vFreq (Sampling frequency within the same time-window)
  - vMax (Maximum value recorded within the time-window)
  - vMin (Minimum value recorded within the time-window)
  - vTrend (Max-Min/10)
  - value (Mean value recorded within the time-window)

Dataset

- Every signal may not have all the features, for instance, the signal *Temperature* has only the *value* feature.

- Fault-free experiments (having label 0) represent the behaviour of the machine during its normal operating regime. Fault-free experiments have been acquired by using **two different system parameter configurations**. Yet, both system parameter configurations lead to a nominal system behaviour.

- Depending on the fault, unhealthy experiments have been labeled with **8 different labels**. Each fault is characterized by an anomalous behaviour of one or or more signals.

Dataset

Table 1 **Model training and validation** – 70 experiments

Class	Number of Datasets
0	50
2	4
3	4
5	4
7	4
9	4

Table 2 **Model refinement** – 29 experiments

Class	Popularity
0	20
4	3
11	3
12	3

Fault-free data shape:
(36763, 247)

Faulty data shape:
(21208, 247)

Objectives

1. Identify and classify the faults.
2. Rank the input signals to identify the most important ones for the prediction.
3. Predict the correct fault in the shortest time.
4. Identification of system parameter configuration in fault-free experiments.

Data Preprocessing

Signal	CpuTemperature				DurationPickToPick						...	VacuumValveClosed					
feature	0	1	2	3	0	1	2	3	4	5	...	1	2	3	4	5	6
time																	
0	48.50	45.00	1.151086	47.000	0.0	0.000000	NaN	NaN	NaN	NaN	...	0.000000	NaN	NaN	NaN	NaN	NaN
1	55.25	44.25	3.256628	48.975	0.0	0.000000	NaN	NaN	NaN	NaN	...	0.000000	NaN	NaN	NaN	NaN	NaN
2	49.75	44.25	1.901480	46.125	0.0	0.000000	NaN	NaN	NaN	NaN	...	0.000000	NaN	NaN	NaN	NaN	NaN
3	51.25	44.50	2.495120	46.325	0.0	0.000000	NaN	NaN	NaN	NaN	...	0.000000	NaN	NaN	NaN	NaN	NaN
4	54.00	46.00	2.636404	49.325	2.0	0.198834	3.118	2.831	0.143500	0.2870	...	1.690092	-0.376446	-0.743156	0.118686	-0.006051	-0.608128
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1075	55.50	49.25	2.123676	52.050	4.0	0.398570	3.277	2.721	0.208668	0.0976	...	1.893210	-0.308573	-0.735343	0.125341	-0.010031	-0.590577
1076	54.25	49.50	1.508517	51.925	3.0	0.298705	3.294	2.873	0.174146	0.2105	...	1.792229	-0.208472	-0.720206	0.158803	-0.009554	-0.548733
1077	56.50	50.00	1.949359	53.500	3.0	0.297850	3.175	2.774	0.173283	-0.2005	...	2.283519	-0.241188	-0.732902	0.157569	0.001406	-0.539600
1078	55.25	48.75	2.219938	51.125	3.0	0.299037	3.279	3.202	0.031633	-0.0155	...	1.096470	-0.257302	-0.730461	0.150160	0.013326	-0.586236
1079	56.75	48.50	2.516073	51.475	0.0	0.000000	NaN	NaN	NaN	NaN	...	0.000000	NaN	NaN	NaN	NaN	NaN

1080 rows × 247 columns

Data Preprocessing

1. Remove features that mainly contain zero.

ErrorFrame		FeederAction1	FeederAction2	FeederAction3	FeederAction4
0	1	0	0	0	0
0.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0

ErrorFrame: Counter of the error frames from the camera observing the fuses while they're on the feeder

FeederAction: The feeder (which brings the fuses from the conveyor belt to the picking area) can perform 4 possible actions (corresponding to 4 different degrees of freedom)

Data Preprocessing

2. Replace NaN (missing value) with 0

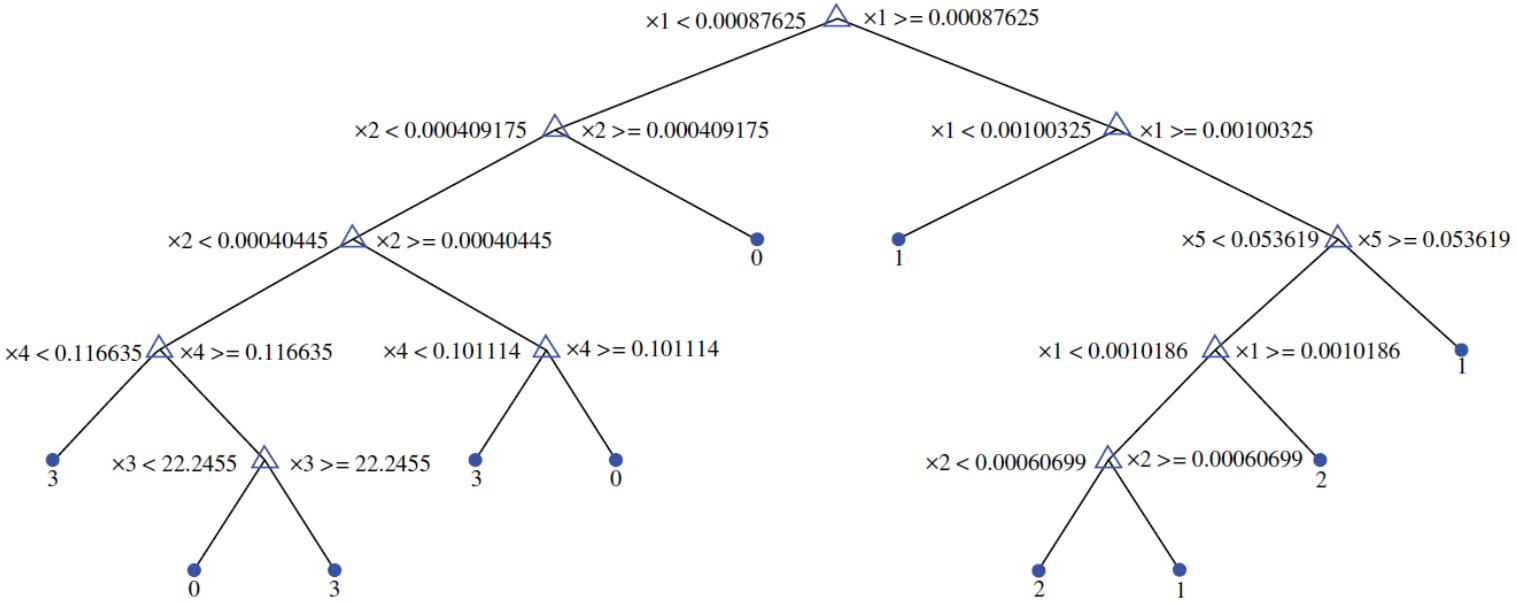
Signal	CpuTemperature				DurationPickToPick						...	VacuumValveClosed					
feature	0	1	2	3	0	1	2	3	4	5	...	1	2	3	4	5	6
time																	
0	48.50	45.00	1.151086	47.000	0.0	0.000000	0.000	0.000	0.000000	0.0000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	55.25	44.25	3.256628	48.975	0.0	0.000000	0.000	0.000	0.000000	0.0000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	49.75	44.25	1.901480	46.125	0.0	0.000000	0.000	0.000	0.000000	0.0000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	51.25	44.50	2.495120	46.325	0.0	0.000000	0.000	0.000	0.000000	0.0000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	54.00	46.00	2.636404	49.325	2.0	0.198834	3.118	2.831	0.143500	0.2870	...	1.690092	-0.376446	-0.743156	0.118686	-0.006051	-0.608128
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1075	55.50	49.25	2.123676	52.050	4.0	0.398570	3.277	2.721	0.208668	0.0976	...	1.893210	-0.308573	-0.735343	0.125341	-0.010031	-0.590577
1076	54.25	49.50	1.508517	51.925	3.0	0.298705	3.294	2.873	0.174146	0.2105	...	1.792229	-0.208472	-0.720206	0.158803	-0.009554	-0.548733
1077	56.50	50.00	1.949359	53.500	3.0	0.297850	3.175	2.774	0.173283	-0.2005	...	2.283519	-0.241188	-0.732902	0.157569	0.001406	-0.539600
1078	55.25	48.75	2.219938	51.125	3.0	0.299037	3.279	3.202	0.031633	-0.0155	...	1.096470	-0.257302	-0.730461	0.150160	0.013326	-0.586236
1079	56.75	48.50	2.516073	51.475	0.0	0.000000	0.000	0.000	0.000000	0.0000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

1080 rows × 247 columns

3. Combine all the experiment data and divide them into 80% training 20% test dataset randomly.

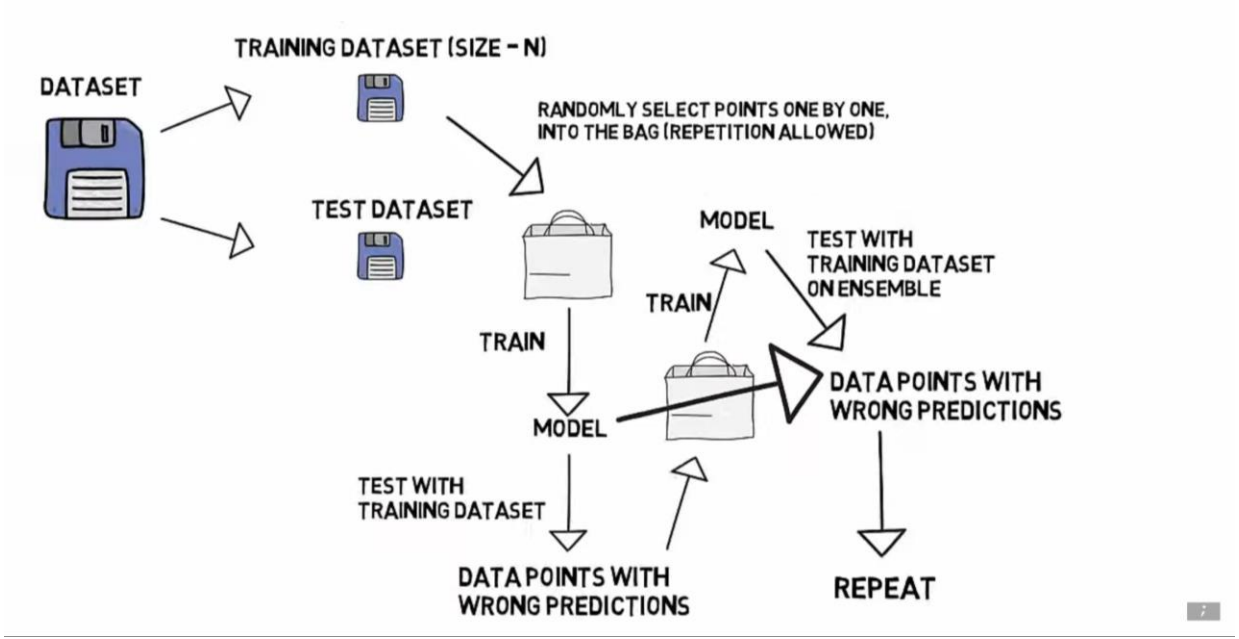
Methodology: Gradient Boosting

Decision Tree



Methodology: Gradient Boosting

Boosting: Boosting is a method of converting weak learners into strong learners.



Results: Classification

Accuracy: 99.89%

Precision: 99.94%

Recall: 98.76%

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Actual \ Prediction	Class 0	Class 2	Class 3	Class 5	Class 7	Class 9
Class 0	[[5051,	0,	0,	0,	0,	1],
Class 2	[1,	552,	0,	0,	0,	0],
Class 3	[5,	0,	576,	0,	0,	0],
Class 5	[0,	0,	0,	598,	0,	0],
Class 7	[0,	0,	0,	0,	573,	0],
Class 9	[2,	0,	0,	0,	0,	564]],

Results: Rank Input Signal

feature	importance
(NumberFuseEstimated, 0)	0.485619
(IntensityTotalImage, 1)	0.215992
(SharpnessImage, 6)	0.067085
(SmartMotorSpeed, 4)	0.023941
(SmartMotorPositionError, 3)	0.018032
(FeederBackgroundIlluminationIntensity, 1)	0.017635
(NumberFuseEstimated, 1)	0.016401
(VacuumValveClosed, 1)	0.014316
(FeederBackgroundIlluminationIntensity, 0)	0.007262
(IntensityTotalThermoImage, 6)	0.007224

feature	sum_importance
NumberFuseEstimated	0.506673
IntensityTotalImage	0.217565
SharpnessImage	0.075362
SmartMotorSpeed	0.029001
FeederBackgroundIlluminationIntensity	0.028714
VacuumValveClosed	0.022246
SmartMotorPositionError	0.021947
LightBarrierPassiveTaskDuration1	0.012694
IntensityTotalThermoImage	0.010411
TotalCpuLoadNormalized	0.009712

Results: Rank Input Signal

- To achieve specific feature ranking for each class, we train the model specifically to predict each class. For instance:
  - Class 0 feature ranking: all fault exp. labeled as 1, so now the model will do binary classification, 0 or 1,
  - Class 2 feature ranking: the training data only consist class 0 and 2,
  - Class 3 feature ranking: the training data only consist class 0 and 3,
  - etc.

Results: Rank Input Signal

	Class_0	Class_2	Class_3	Class_4	Class_5
0	SharpnessImage	FeederBackgroundIlluminationIntensity	SharpnessImage	SmartMotorPositionError	TotalCpuLoadNormalized
1	FeederBackgroundIlluminationIntensity	NumberFuseEstimated	TotalMemoryConsumption	TotalCpuLoadNormalized	VacuumValveClosed
2	SmartMotorSpeed	EPOSPosition	NumberFuseEstimated	Temperature	IntensityTotalThermoImage

Class_7	Class_9	Class_11	Class_12
EPOSPosition	SmartMotorSpeed	DurationRobotFromFeederToTestBench	SmartMotorSpeed
Vacuum	Temperature	DurationTestBenchClosed	Temperature
TemperatureThermoCam	SmartMotorPositionError	Temperature	TotalCpuLoadNormalized

Results

SmartMotorSpeed

0	1	2	3	4	5	6
0.0	0.000000	NaN	NaN	NaN	NaN	NaN
5.0	0.498758	0.0	0.0	0.0	0.0	0.0
5.0	0.498957	0.0	0.0	0.0	0.0	0.0
5.0	0.499137	0.0	-700.0	280.0	-140.0	-140.0
5.0	0.497086	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.496719	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.494143	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.498009	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.496704	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.497169	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.497616	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.497616	-700.0	-700.0	0.0	0.0	-700.0
5.0	0.497209	-700.0	-700.0	0.0	0.0	-700.0

Class 0

SmartMotorSpeed

0	1	2	3	4	5	6
3.0	0.509038	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.398562	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399732	0.0	0.0	0.000000	0.000000e+00	0.000000
3.0	0.299817	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399612	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399739	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399787	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399668	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.399640	0.0	0.0	0.000000	0.000000e+00	0.000000
4.0	0.398943	0.0	-706.0	351.506401	-2.818000e+02	-351.500000
4.0	0.398141	-710.0	-719.0	3.391165	-3.000000e+00	-715.000000
4.0	0.398280	-713.0	-719.0	2.291288	2.000000e+00	-716.500000
4.0	0.396669	-694.0	-708.0	5.147815	4.600000e+00	-701.000000
4.0	0.398120	-681.0	-689.0	3.112475	2.700000e+00	-684.250000
4.0	0.398531	-681.0	-691.0	3.905125	-3.200000e+00	-684.500000
5.0	0.496473	-694.0	-709.0	5.238320	-3.700000e+00	-701.400000
5.0	0.497378	-713.0	-719.0	2.280351	-1.500000e+00	-717.000000
5.0	0.497883	-709.0	-719.0	3.633180	2.500000e+00	-715.000000

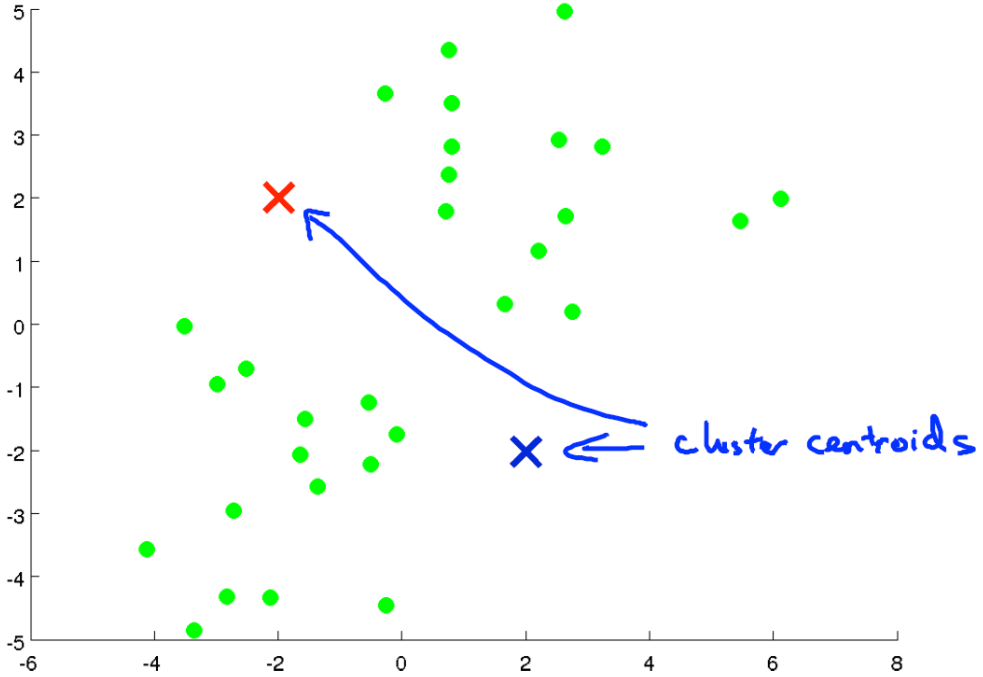
Class 9

Results: Time to classification

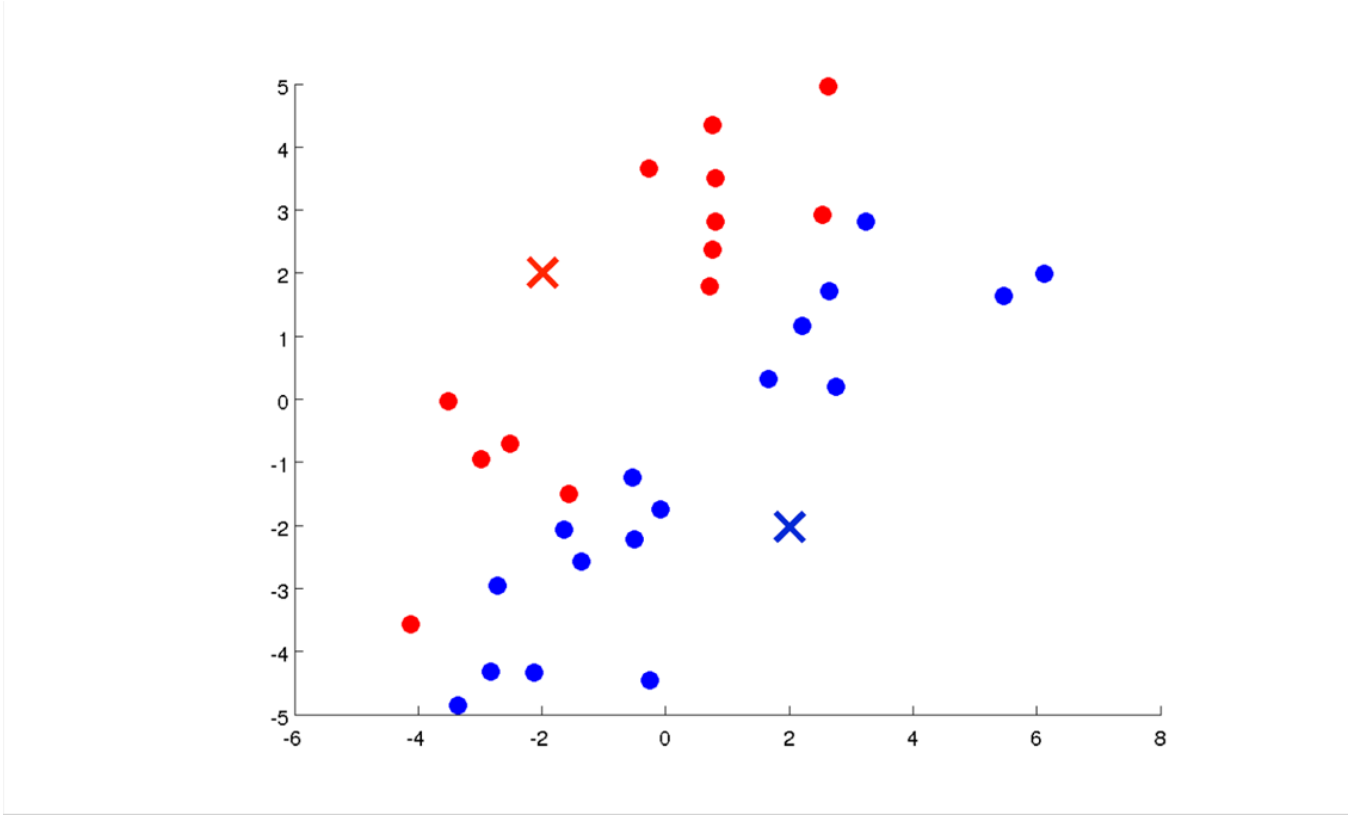
- Time to classification: 1

	Experiment	Label	Time	Ranking
0	class_12_0_data	12	1	['SmartMotorSpeed', 'Temperature', 'TotalCpuLo...

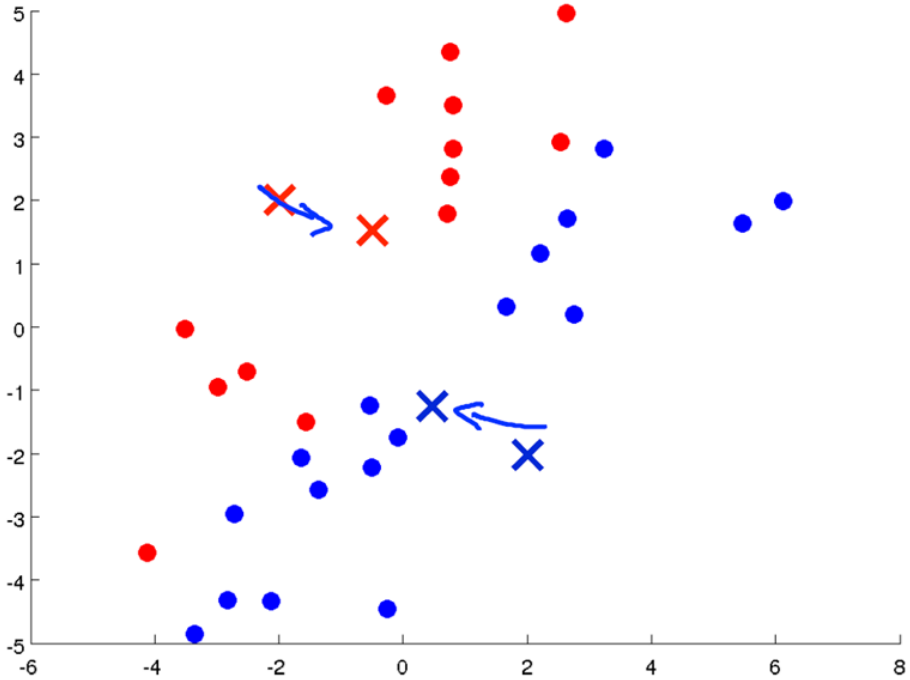
Methodology: k-Means Clustering



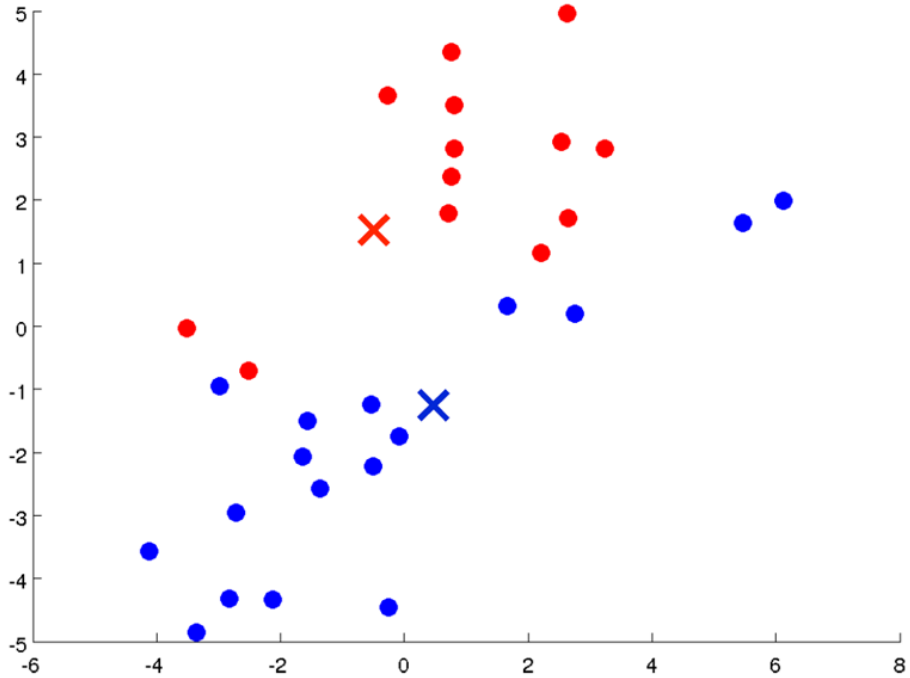
Methodology: k-Means Clustering



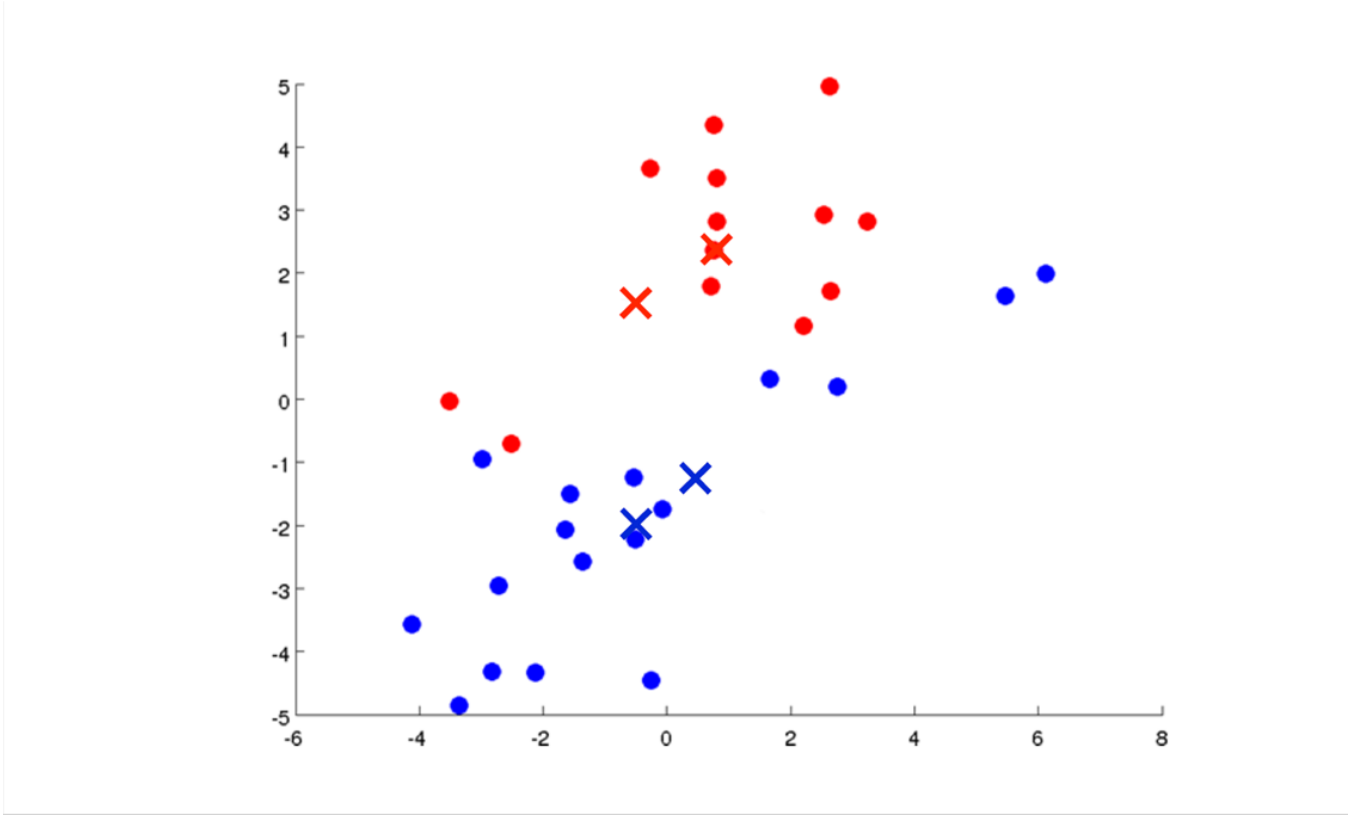
Methodology: k-Means Clustering



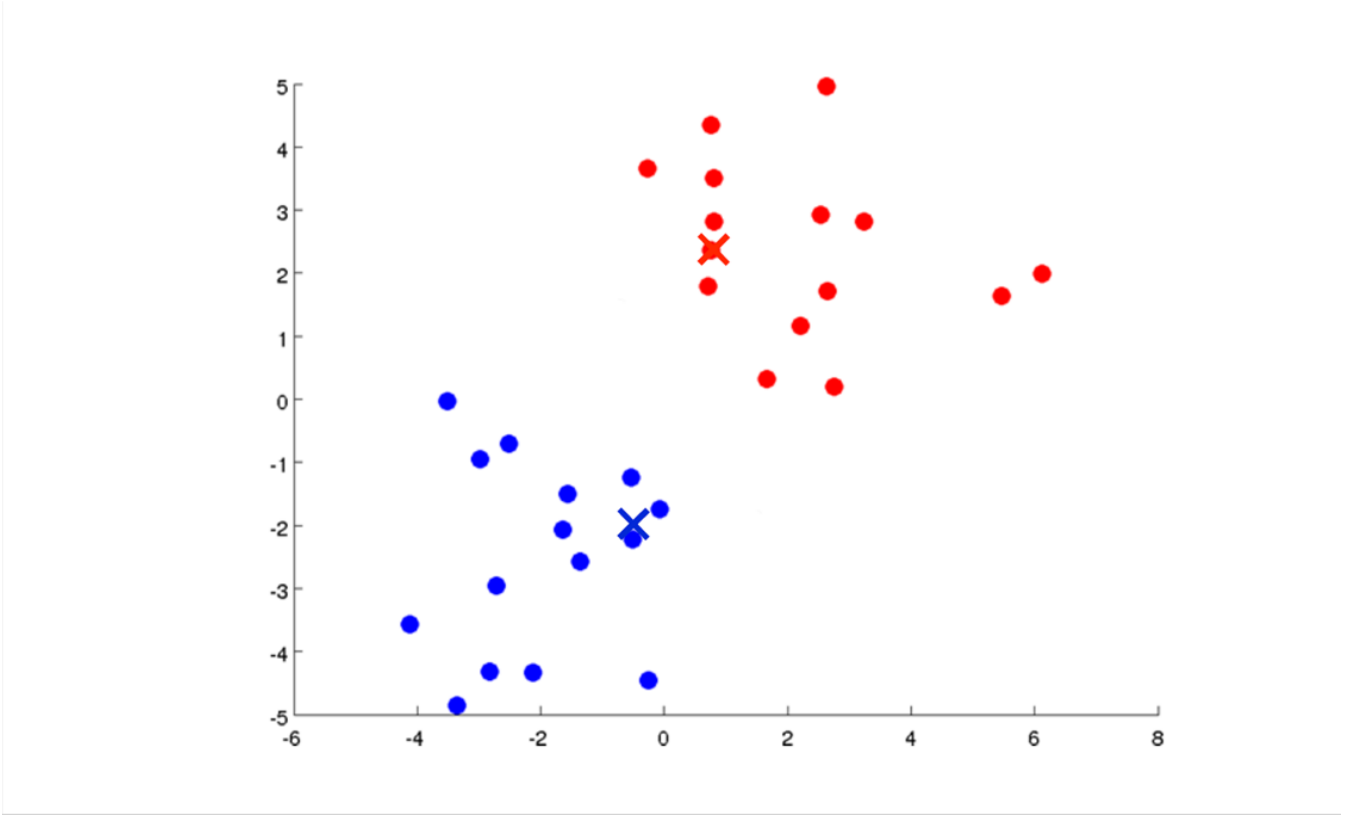
Methodology: k-Means Clustering



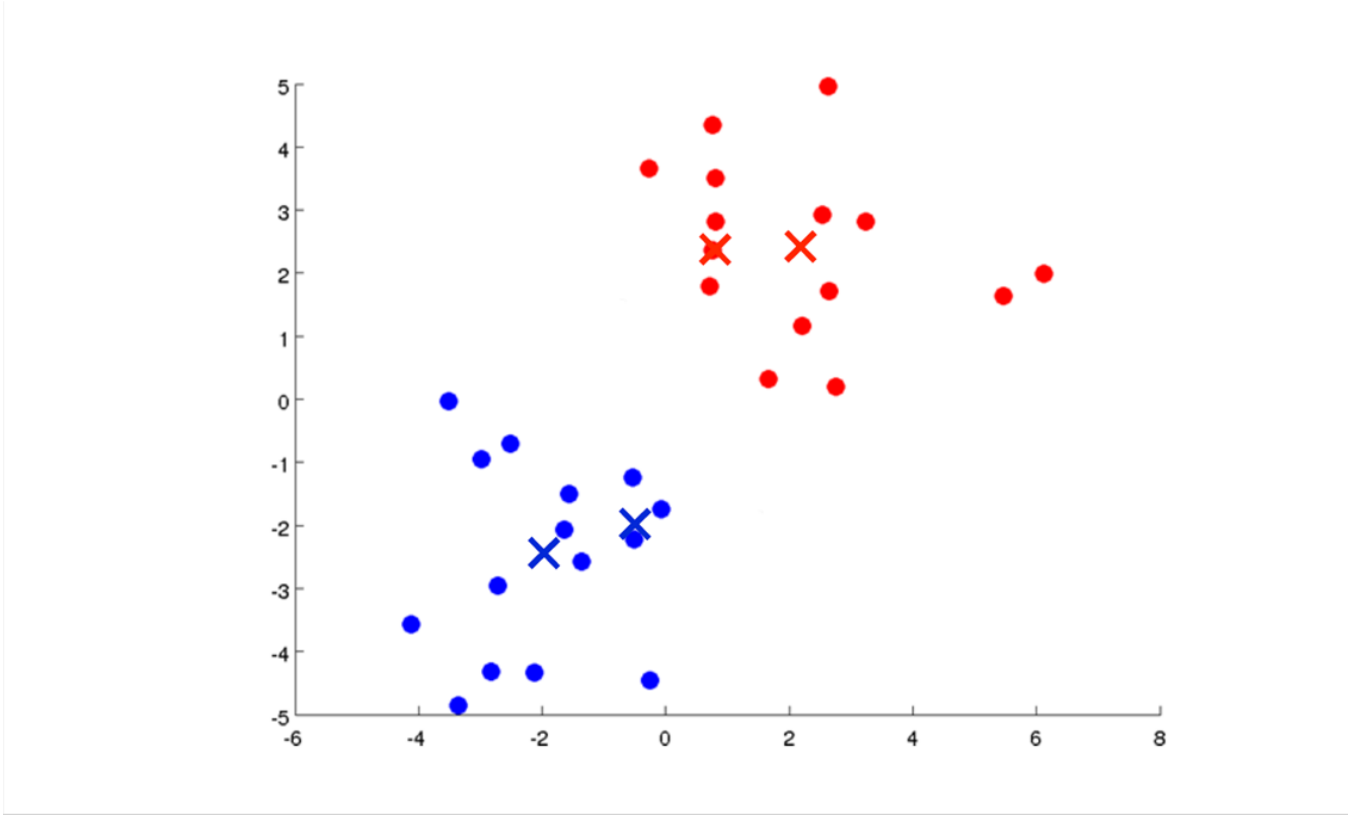
Methodology: k-Means Clustering



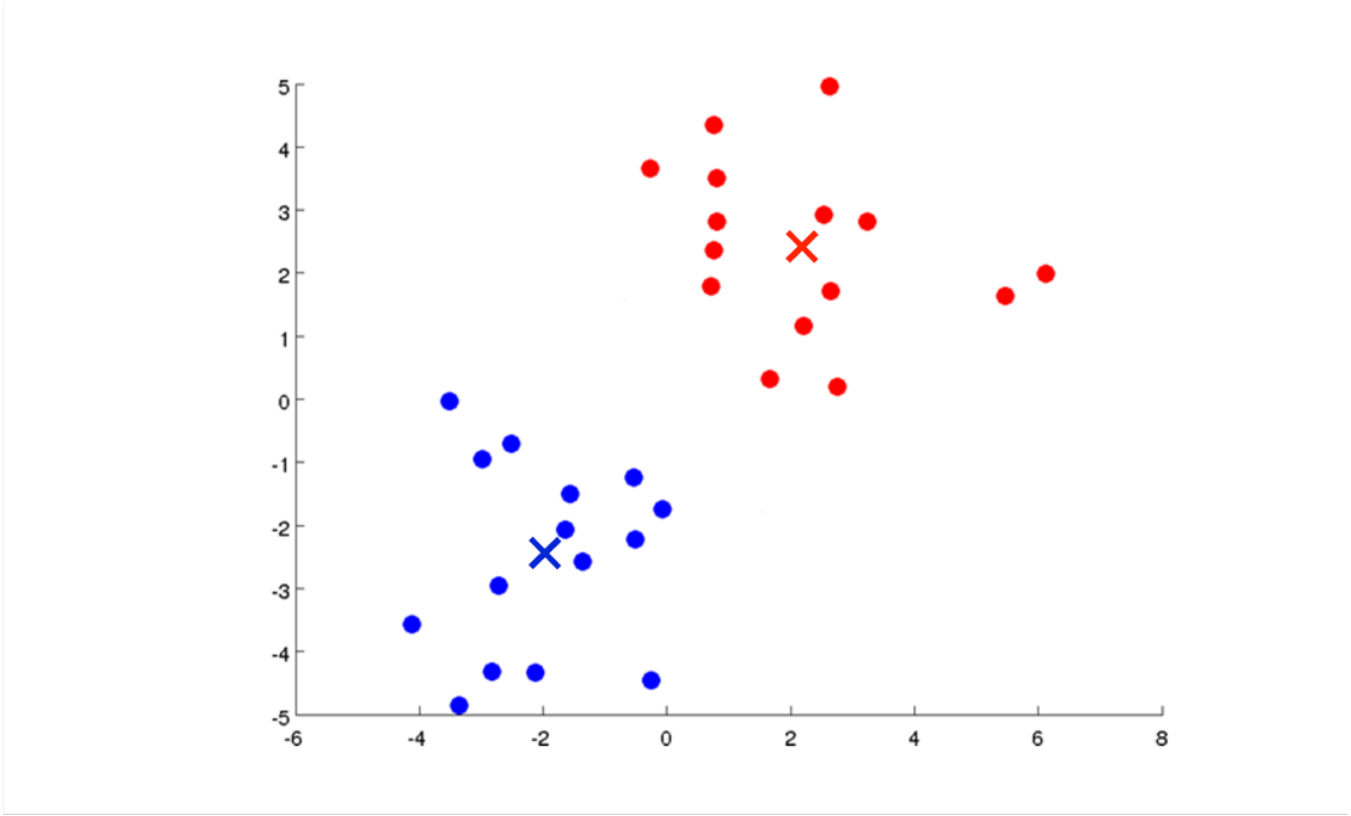
Methodology: k-Means Clustering



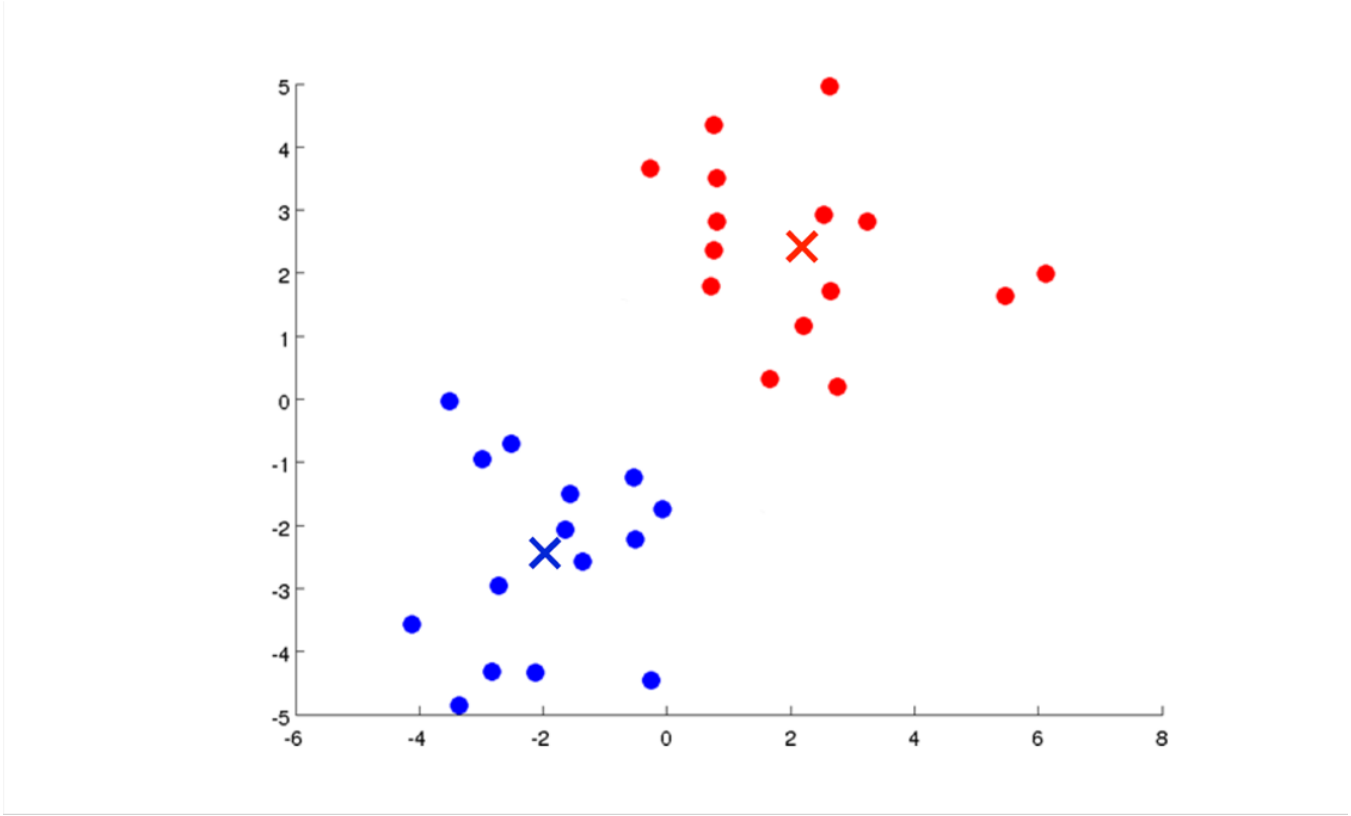
Methodology: k-Means Clustering



Methodology: k-Means Clustering



Methodology: k-Means Clustering



Results

- Identification of system parameter configuration in fault-free experiments.

	Experiment	ClusterLabel
0	class_ 0_0_data.csv	1.0
1	class_ 0_100_data.csv	0.0
2	class_ 0_101_data.csv	0.0
3	class_ 0_103_data.csv	0.0
4	class_ 0_10_data.csv	1.0
5	class_ 0_11_data.csv	1.0
6	class_ 0_13_data.csv	1.0
7	class_ 0_17_data.csv	1.0
8	class_ 0_18_data.csv	1.0
9	class_ 0_19_data.csv	1.0
10	class_ 0_1_data.csv	1.0
11	class_ 0_20_data.csv	1.0
12	class_ 0_23_data.csv	1.0
13	class_ 0_25_data.csv	1.0

```
Running time for kmeans: 10.491950511932373 Seconds
 silhouette_score: 0.2556469133950757
```

$$s = \frac{b - a}{\max(a, b)}$$

a: The mean distance between a sample and all other points in the same class.

b: The mean distance between a sample and all other points in the next nearest cluster.