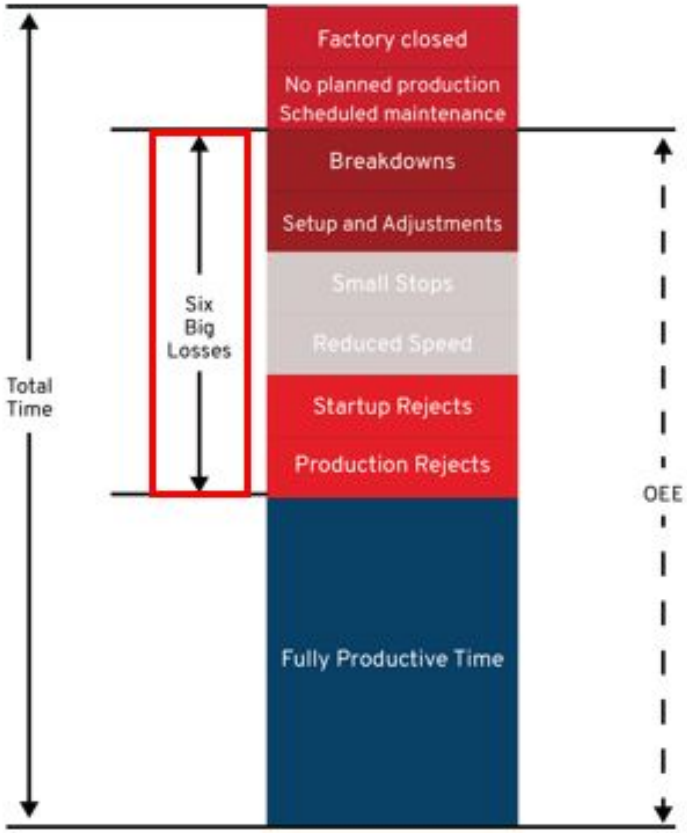
Predictive Maintenance

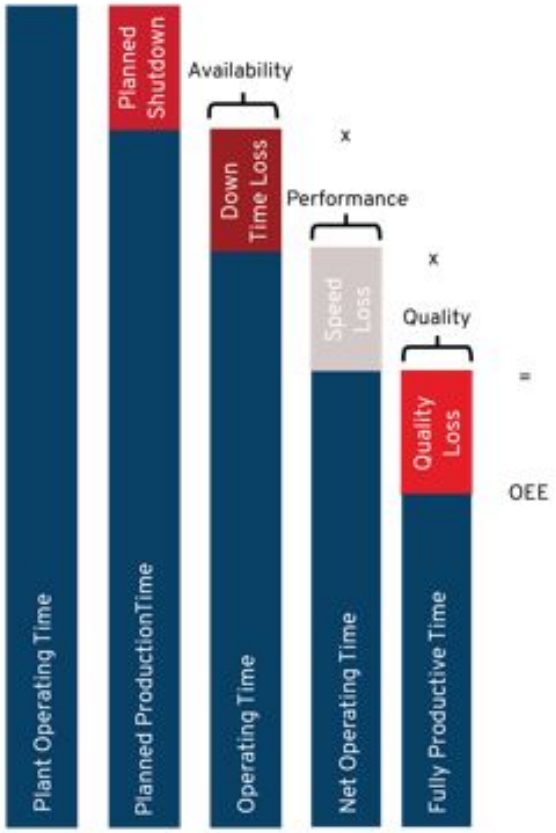
1. ณัฐวุฒิ ทองศรีนุ่น 65056036
2. ทศพล กนกพิพัฒน์วงศ์ 65056040
3. นภัสวรรณ ขันดี 65056049
4. ภูริวัฒน์ แสงระวี 65056071

Business Understanding

Overview

High OEE, Low downtimes, High quality products, High Productivity





Business Understanding

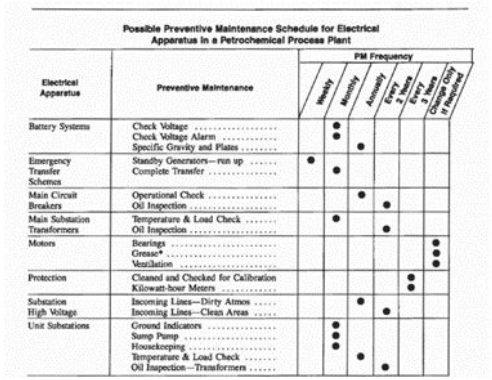
Overview

Level 1 Reactive

- Maintenance based upon running to failure.
- Repair/Replace.
- Little Data.

Level 2 Preventive

- Routine inspection.
- Maintenance upon time.
- Some data collection.

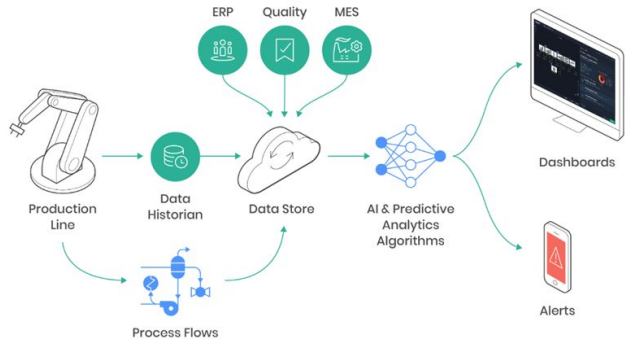


Possible Preventive Maintenance Schedule for Electrical Apparatus in a Petrochemical Process Plant

Electrical Apparatus	Preventive Maintenance	PM Frequency: Weekly	Monthly	Annually	Every 2 Years	Every 3 Years	Change Only If Required
Battery Systems	Check Voltage		•				
	Check Voltage Alarm		•				
	Specific Gravity and Plates			•			
Emergency Transfer Schemes	Standby Generators—run up	•					
	Complete Transfer		•				
Main Circuit Breakers	Operational Check			•			
	Oil Inspection				•		
Main Substation Transformers	Temperature & Load Check		•				
	Oil Inspection				•		
Motors	Bearings						•
	Grease*						•
	Ventilation						•
Protection	Cleaned and Checked for Calibration					•	
	Kilowatt-hour Meters					•	
Substation High Voltage	Incoming Lines—Dirty Atmos			•			
	Incoming Lines—Clean Areas				•		
Unit Substations	Ground Indicators		•				
	Sump Pump		•				
	Housekeeping		•				
	Temperature & Load Check			•			
	Oil Inspection—Transformers				•		

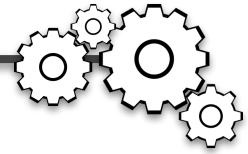
Level 3 Predictive

- Monitor equipment health.
- Online performance intelligence and correction.
- Maintenance based upon condition.
- Data is analyzed.



Business Understanding

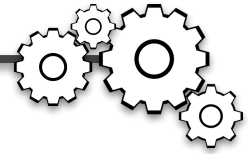
Goal



‘‘Predict Machine health status in order to improve machine effectiveness.’’

Data Understanding

Dataset selection

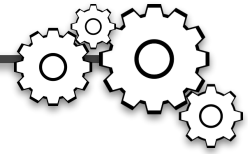


Row No.	Machine fai...	UDI	Product ID	Type	Air temperat...	Process te...	Rotational s...	Torque [Nm]	Tool wear [...	TWF	HDF	PWF	OSF	RNF
1	0	1	M14860	M	298.100	308.600	1551	42.800	0	0	0	0	0	0
2	0	2	L47181	L	298.200	308.700	1408	46.300	3	0	0	0	0	0
3	0	3	L47182	L	298.100	308.500	1498	49.400	5	0	0	0	0	0
4	0	4	L47183	L	298.200	308.600	1433	39.500	7	0	0	0	0	0
5	0	5	L47184	L	298.200	308.700	1408	40	9	0	0	0	0	0
6	0	6	M14865	M	298.100	308.600	1425	41.900	11	0	0	0	0	0
7	0	7	L47186	L	298.100	308.600	1558	42.400	14	0	0	0	0	0
8	0	8	L47187	L	298.100	308.600	1527	40.200	16	0	0	0	0	0
9	0	9	M14868	M	298.300	308.700	1667	28.600	18	0	0	0	0	0
10	0	10	M14869	M	298.500	309	1741	28	21	0	0	0	0	0
11	0	11	H29424	H	298.400	308.900	1782	23.900	24	0	0	0	0	0
12	0	12	H29425	H	298.600	309.100	1423	44.300	29	0	0	0	0	0
13	0	13	M14872	M	298.600	309.100	1339	51.100	34	0	0	0	0	0
14	0	14	M14873	M	298.600	309.200	1742	30	37	0	0	0	0	0
15	0	15	L47194	L	298.600	309.200	2035	19.600	40	0	0	0	0	0
16	0	16	L47195	L	298.600	309.200	1542	48.400	42	0	0	0	0	0
17	0	17	M14876	M	298.600	309.200	1311	46.600	44	0	0	0	0	0

Since real predictive maintenance datasets are generally difficult to obtain and in particular difficult to publish, we present and provide a synthetic dataset that reflects real predictive maintenance encountered in industry to the best of our knowledge.

Data Understanding

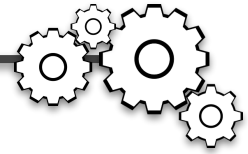
Dataset selection



Feature	Description
UID	unique identifier ranging from 1 to 10000
Product ID	consisting of a letter L, M, or H for low (50% of all products), medium (30%) and high (20%) as product quality variants and a variant-specific serial number
Type	L, M, or H for low (50% of all products), medium (30%) and high (20%)
Air temperature [K]	generated using a random walk process later normalized to a standard deviation of 2 K around 300 K
Process temperature [K]	generated using a random walk process normalized to a standard deviation of 1 K, added to the air temperature plus 10 K.
Rotational speed [rpm]	calculated from a power of 2860 W, overlaid with a normally distributed noise
Torque [Nm]	torque values are normally distributed around 40 Nm with a Ïf = 10 Nm and no negative values.
Tool wear [min]	The quality variants H/M/L add 5/3/2 minutes of tool wear to the used tool in the process.

Data Understanding

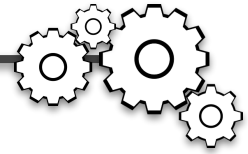
Dataset selection



Feature	Description
Machine Failure	Failure / No Failure Note : 'machine failure' label that indicates, whether the machine has failed in this particular datapoint for any of the following failure modes are true.
Tool wear failure (TWF)	the tool will be replaced of fail at a randomly selected tool wear time between 200 and 240 mins (120 times in our dataset). At this point in time, the tool is replaced 69 times, and fails 51 times (randomly assigned).
Heat dissipation failure (HDF)	heat dissipation causes a process failure, if the difference between air- and process temperature is below 8.6 K and the toolâ€™s rotational speed is below 1380 rpm. This is the case for 115 data points.
Power failure (PWF)	the product of torque and rotational speed (in rad/s) equals the power required for the process. If this power is below 3500 W or above 9000 W, the process fails, which is the case 95 times in our dataset.
Overstrain failure (OSF)	if the product of tool wear and torque exceeds 11,000 minNm for the L product variant (12,000 M, 13,000 H), the process fails due to overstrain. This is true for 98 data points.
Random failures (RNF)	each process has a chance of 0,1 % to fail regardless of its process parameters. This is the case for only 5 data points, less than could be expected for 10,000 data points in our dataset.

Data Understanding

Quality issue

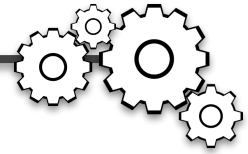


Discrete data

Row No.	Machine fail...	UDI	Product ID	Type	Air temperat...	Process te...	Rotational s...	Torque [Nm]	Tool wear [...	TWF	HDF	PWF	OSF	RNF
1	0	1	M14860	M	298.100	308.600	1551	42.800	0	0	0	0	0	0
2	0	2	L47181	L	298.200	308.700	1408	46.300	3	0	0	0	0	0
3	0	3	L47182	L	298.100	308.500	1498	49.400	5	0	0	0	0	0
4	0	4	L47183	L	298.200	308.600	1433	39.500	7	0	0	0	0	0
5	0	5	L47184	L	298.200	308.700	1408	40	9	0	0	0	0	0
6	0	6	M14865	M	298.100	308.600	1425	41.900	11	0	0	0	0	0
7	0	7	L47186	L	298.100	308.600	1558	42.400	14	0	0	0	0	0
8	0	8	L47187	L	298.100	308.600	1527	40.200	16	0	0	0	0	0
9	0	9	M14868	M	298.300	308.700	1667	28.600	18	0	0	0	0	0
10	0	10	M14869	M	298.500	309	1741	28	21	0	0	0	0	0
11	0	11	H29424	H	298.400	308.900	1782	23.900	24	0	0	0	0	0

Data Understanding

Quality issue



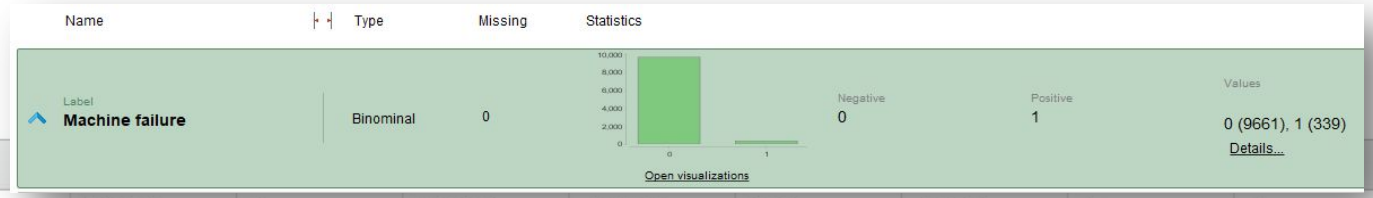
Different scale

Row No.	Machine fail...	UDI	Product ID	Type	Air temperat...	Process te...	Rotational s...	Torque [Nm]	Tool wear [...	TWF	HDF	PWF	OSF	RNF
1	0	1	M14860	M	298.100	308.600	1551	42.800	0	0	0	0	0	0
2	0	2	L47181	L	298.200	308.700	1408	46.300	3	0	0	0	0	0
3	0	3	L47182	L	298.100	308.500	1498	49.400	5	0	0	0	0	0
4	0	4	L47183	L	298.200	308.600	1433	39.500	7	0	0	0	0	0
5	0	5	L47184	L	298.200	308.700	1408	40	9	0	0	0	0	0
6	0	6	M14865	M	298.100	308.600	1425	41.900	11	0	0	0	0	0
7	0	7	L47186	L	298.100	308.600	1558	42.400	14	0	0	0	0	0
8	0	8	L47187	L	298.100	308.600	1527	40.200	16	0	0	0	0	0
9	0	9	M14868	M	298.300	308.700	1667	28.600	18	0	0	0	0	0
10	0	10	M14869	M	298.500	309	1741	28	21	0	0	0	0	0
11	0	11	H29424	H	298.400	308.900	1782	23.900	24	0	0	0	0	0

Data Understanding

Quality issue

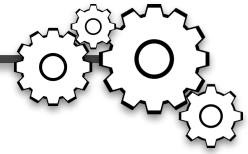
Imbalanced data



Row No.	Machine fail...	UDI								HDF	PWF	OSF	RNF
1	0	1	M14860	M	298.100	308.600	1551	42.800	0	0	0	0	0	0
2	0	2	L47181	L	298.200	308.700	1408	46.300	3	0	0	0	0	0
3	0	3	L47182	L	298.100	308.500	1498	49.400	5	0	0	0	0	0
4	0	4	L47183	L	298.200	308.600	1433	39.500	7	0	0	0	0	0
5	0	5	L47184	L	298.200	308.700	1408	40	9	0	0	0	0	0
6	0	6	M14865	M	298.100	308.600	1425	41.900	11	0	0	0	0	0
7	0	7	L47186	L	298.100	308.600	1558	42.400	14	0	0	0	0	0
8	0	8	L47187	L	298.100	308.600	1527	40.200	16	0	0	0	0	0
9	0	9	M14868	M	298.300	308.700	1667	28.600	18	0	0	0	0	0
10	0	10	M14869	M	298.500	309	1741	28	21	0	0	0	0	0
11	0	11	H29424	H	298.400	308.900	1782	23.900	24	0	0	0	0	0

Data Understanding

Quality issue



Can Eliminate

Can Eliminate

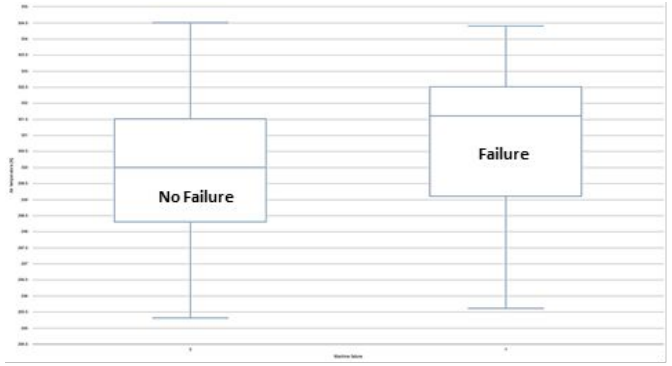
Row No.	Machine fail...	UDI	Product ID	Type	Air temperat...	Process te...	Rotational s...	Torque [Nm]	Tool wear [...	TWF	HDF	PWF	OSF	RNF
1	0	1	M14860	M	298.100	308.600	1551	42.800	0	0	0	0	0	0
2	0	2	L47181	L	298.200	308.700	1408	46.300	3	0	0	0	0	0
3	0	3	L47182	L	298.100	308.500	1498	49.400	5	0	0	0	0	0
4	0	4	L47183	L	298.200	308.600	1433	39.500	7	0	0	0	0	0
5	0	5	L47184	L	298.200	308.700	1408	40	9	0	0	0	0	0
6	0	6	M14865	M	298.100	308.600	1425	41.900	11	0	0	0	0	0
7	0	7	L47186	L	298.100	308.600	1558	42.400	14	0	0	0	0	0
8	0	8	L47187	L	298.100	308.600	1527	40.200	16	0	0	0	0	0
9	0	9	M14868	M	298.300	308.700	1667	28.600	18	0	0	0	0	0
10	0	10	M14869	M	298.500	309	1741	28	21	0	0	0	0	0
11	0	11	H29424	H	298.400	308.900	1782	23.900	24	0	0	0	0	0

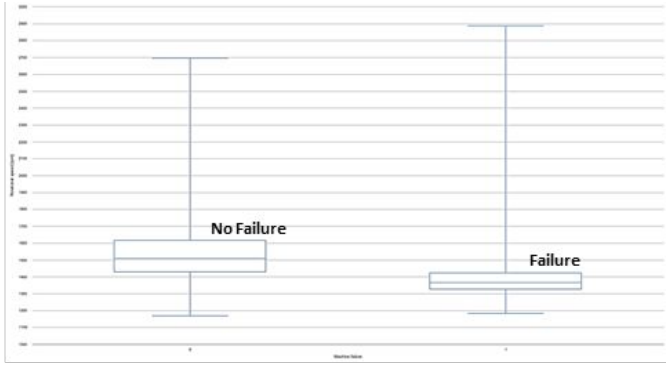
Data Understanding

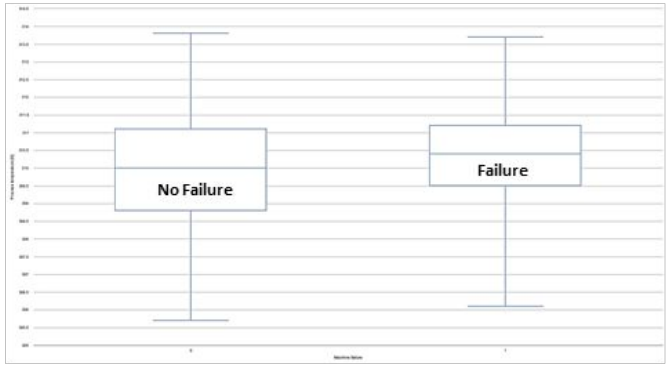
Preliminary analysis

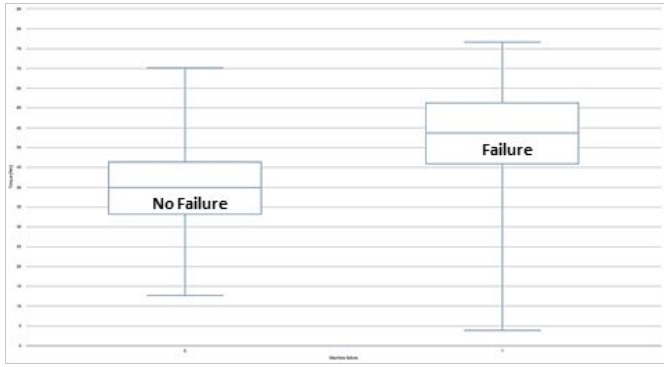
No.	Attribute	Description
3	Type	L, M or H for low, medium, high
4	Air temperature [K]	Environment temperature
5	Process temperature [K]	Temperature during operation
6	Rotational speed [rpm]	spindle speed for cutting
7	Torque [Nm]	how much force can be applied to the cutting tool and the material being machined
8	Tool wear [min]	the minimum amount of time that a cutting tool can be used
9	Machine failure	Failure / No Failure

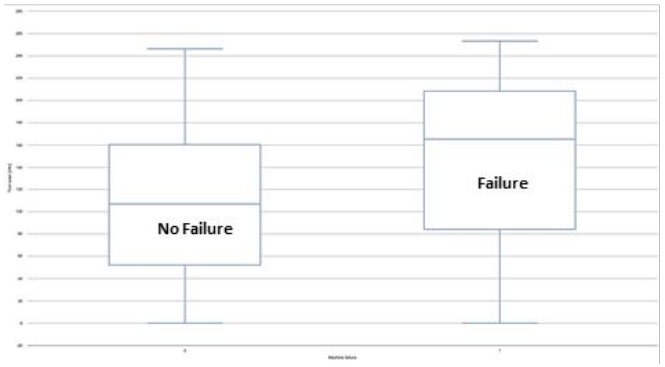
Attribute vs Machine failure









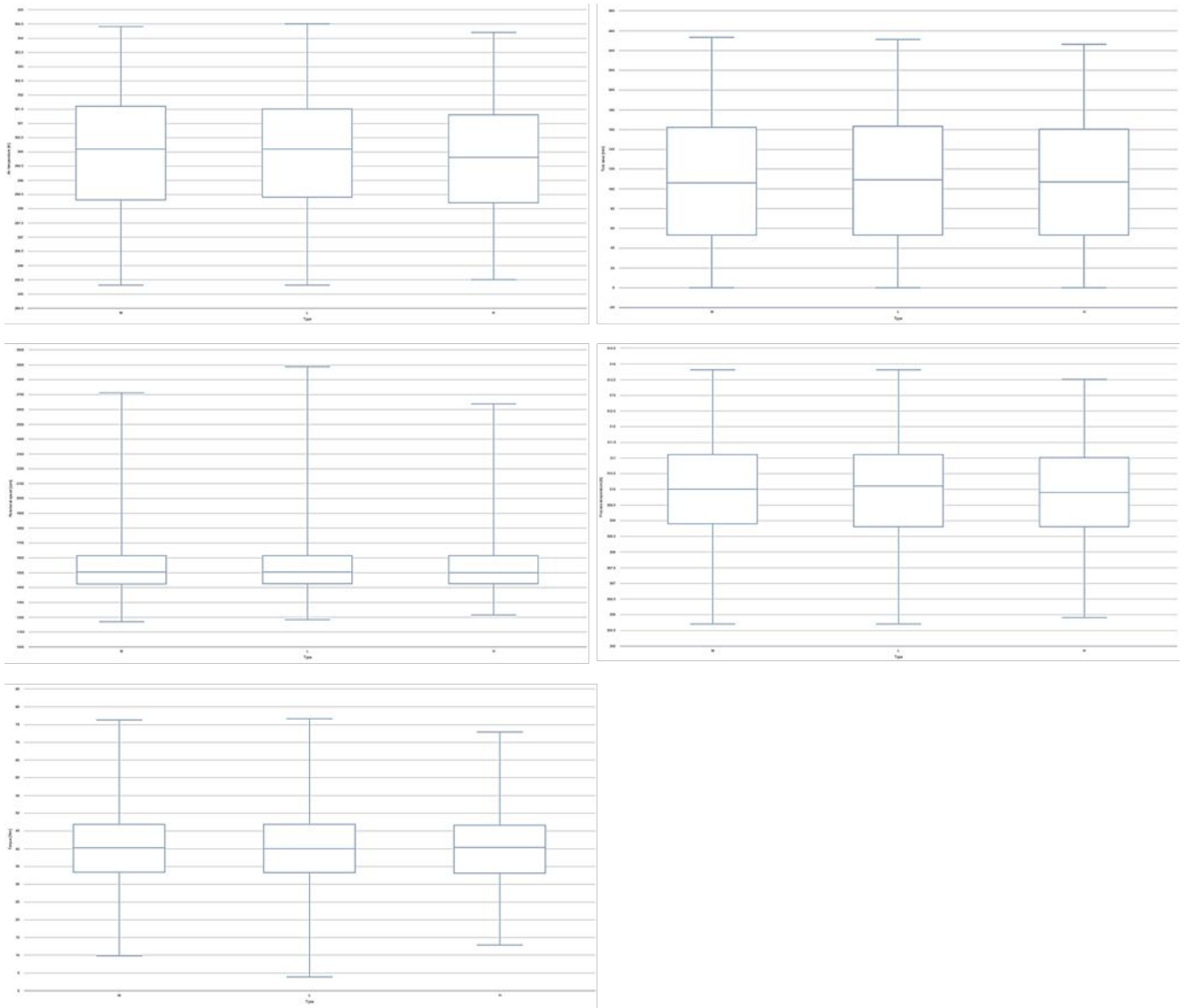


Data Understanding

Preliminary analysis

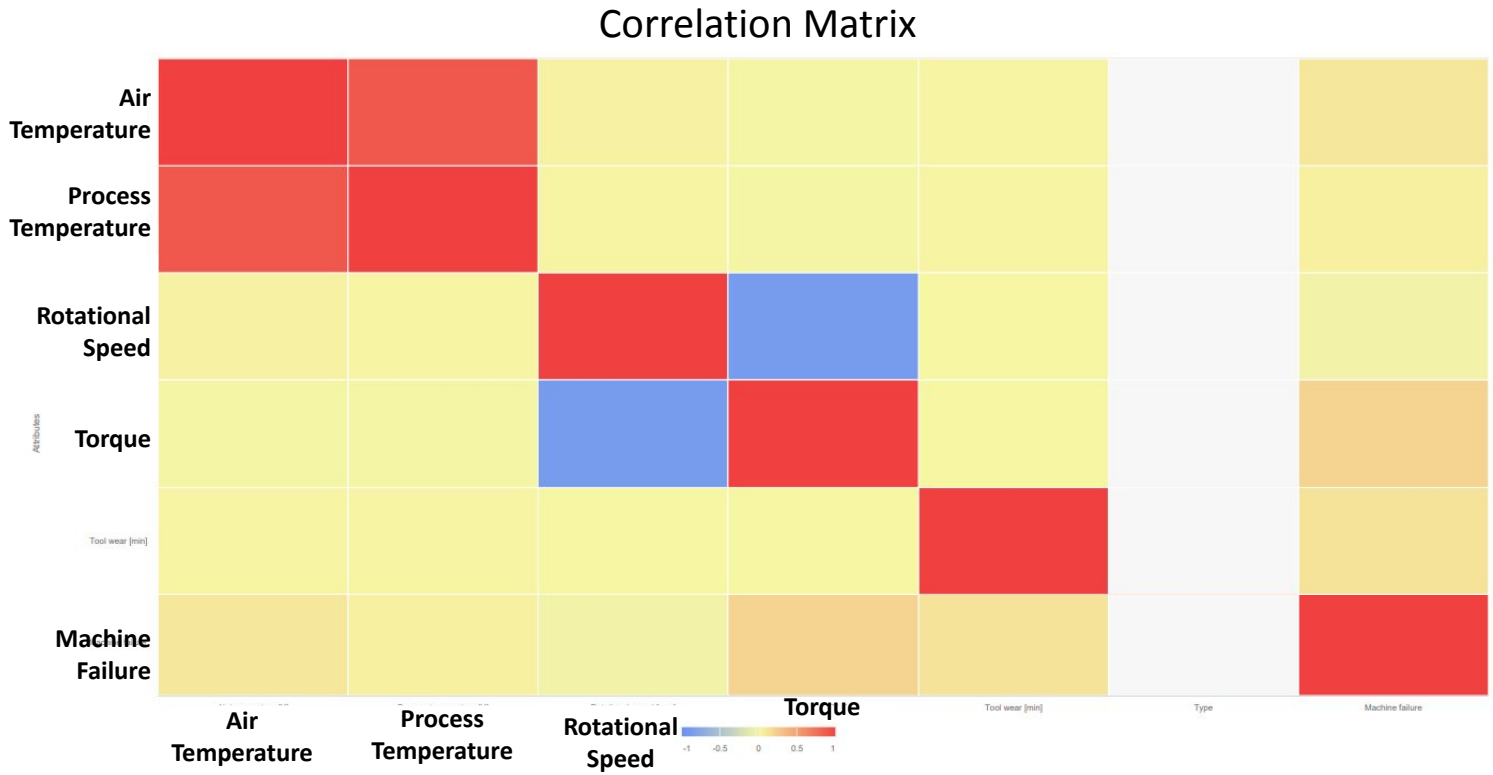
No.	Attribute	Description
3	Type	L, M or H for low, medium, high
4	Air temperature [K]	Environment temperature
5	Process temperature [K]	Temperature during operation
6	Rotational speed [rpm]	spindle speed for cutting
7	Torque [Nm]	how much force can be applied to the cutting tool and the material being machined
8	Tool wear [min]	the minimum amount of time that a cutting tool can be used
9	Machine failure	Failure / No Failure

Attribute vs Product Type



Data Understanding

Preliminary analysis

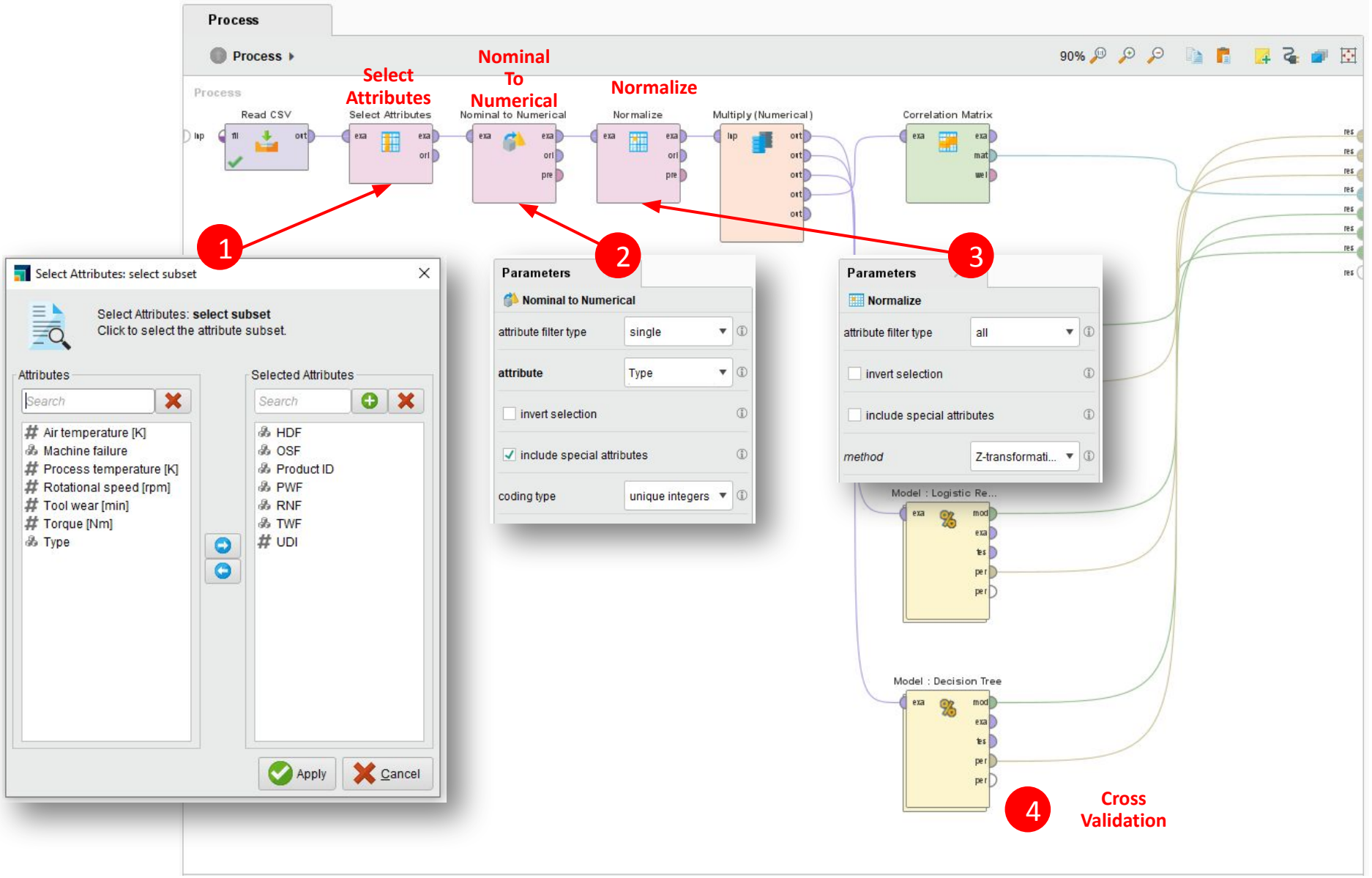


- Process temperature(k) & Air temperature(k)

- Torque(Nm) & Rotational speed(rpm)
the torque of a CNC machine is inversely proportional to the speed. That is, as the speed increases, the torque decreases, and vice versa.

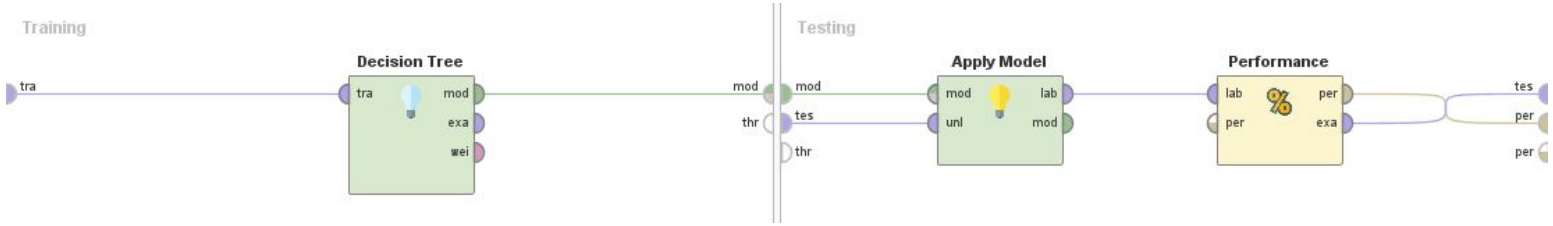
- Torque(Nm) & Machine Failure
excessive torque can cause mechanical stress and wear on the components of the machine. Over time, this can lead to component failure.

Data Preparation

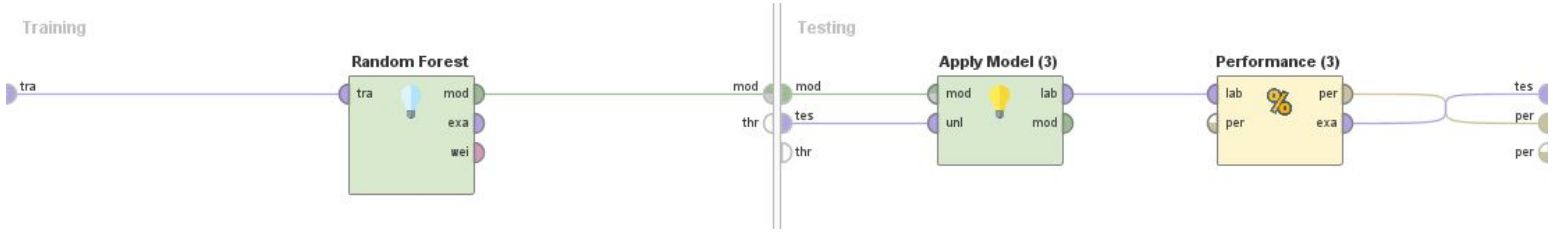


Modeling

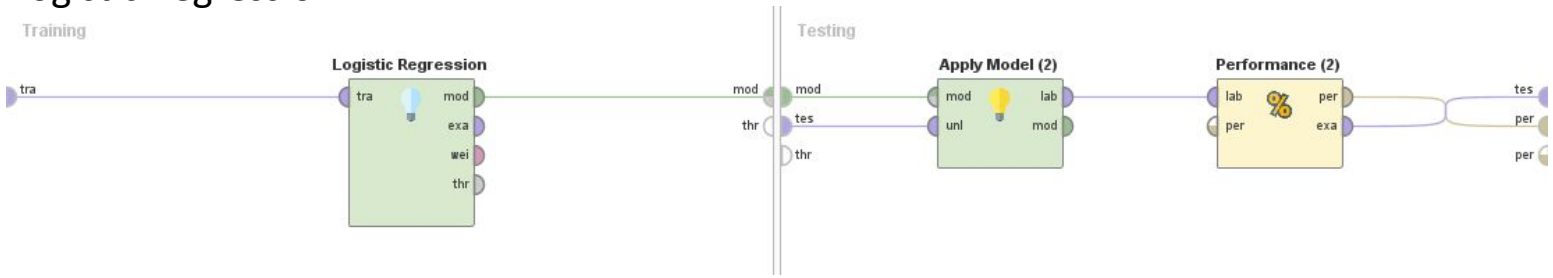
Decision Tree



Random Forest

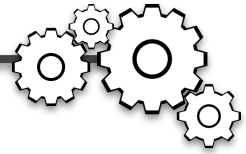


Logistic Regression



Modeling

Logistic Regression



Performance

accuracy: 96.97% +/- 0.26% (micro average: 96.97%)

	true 0	true 1	class precision
pred. 0	9632	274	97.23%
pred. 1	29	65	69.15%
class recall	99.70%	19.17%	

Model

Attribute	Coefficient	Std. Coefficient	Std. Error	z-Value	p-Value
Type	0.026	0.026	0.066	0.398	0.691
Air temperature [K]	1.544	1.544	0.145	10.675	0
Process temperature [K]	-1.102	-1.102	0.143	-7.705	0.000
Rotational speed [rpm]	2.068	2.068	0.095	21.758	0
Torque [Nm]	2.800	2.800	0.114	24.485	0
Tool wear [min]	0.849	0.849	0.072	11.748	0
Intercept	-4.892	-4.892	0.124	-39.405	0

```
MSE: 0.02396534
RMSE: 0.15480743
R^2: 0.26825157
AUC: 0.89692694
pr_auc: 0.45385846
logloss: 0.09613681
mean_per_class_error: 0.24473485
default threshold: 0.1935194432735443
```

Modeling

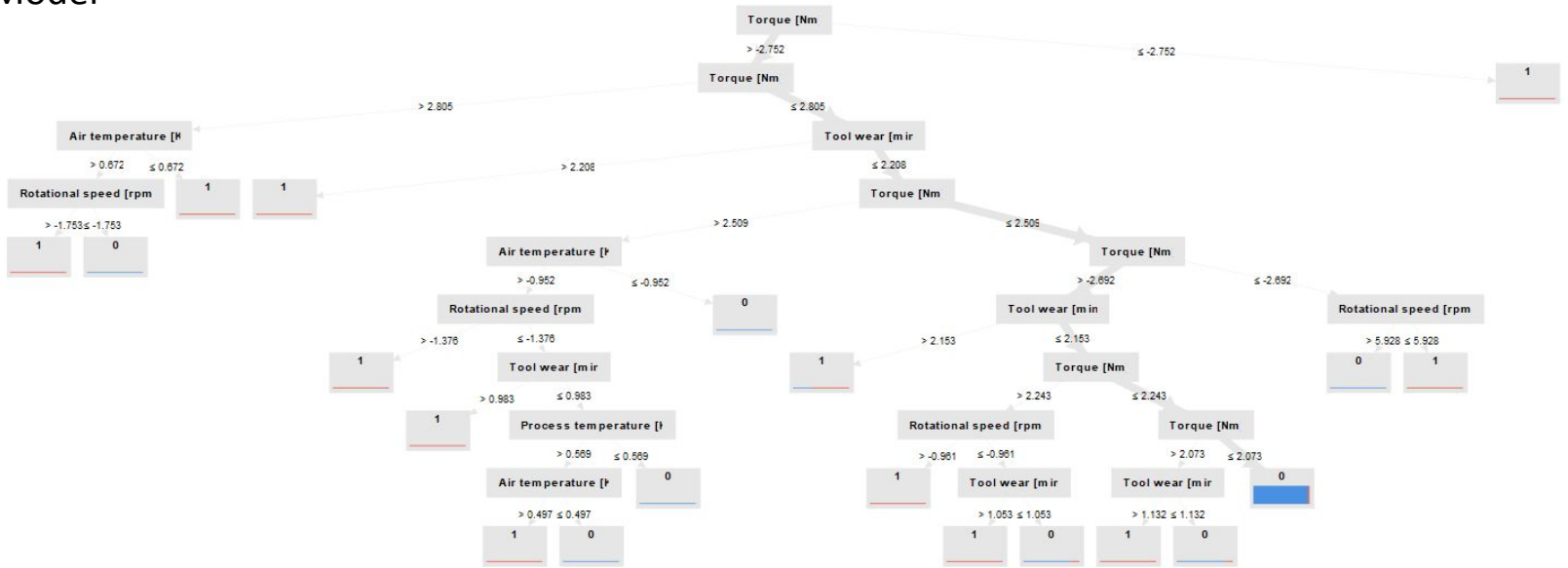
Decision Tree

Performance

accuracy: 97.39% +/- 0.29% (micro average: 97.39%)

	true 0	true 1	class precision
pred. 0	9649	249	97.48%
pred. 1	12	90	88.24%
class recall	99.88%	26.55%	

Model



Modeling

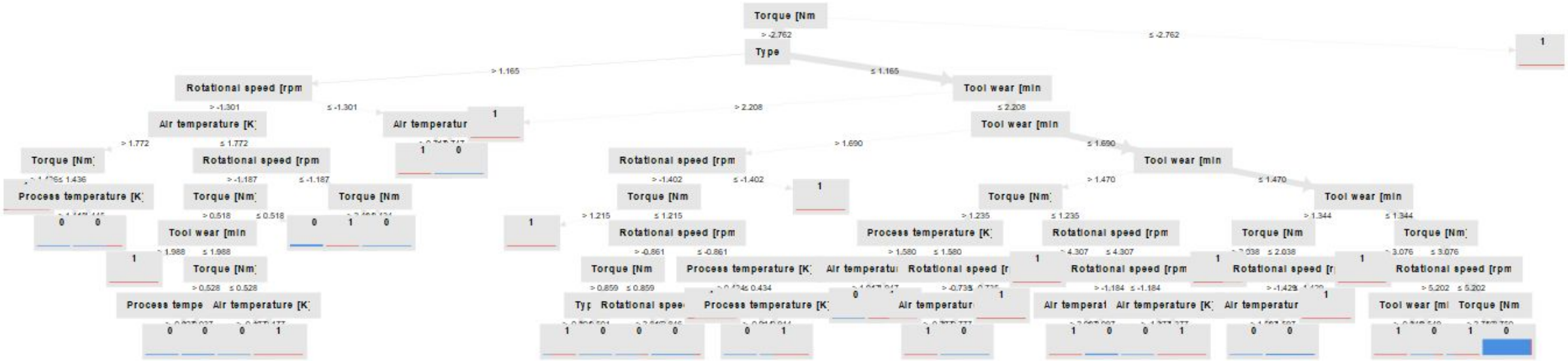
Random Forest

Performance

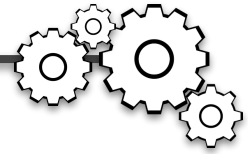
accuracy: 97.24% +/- 0.20% (micro average: 97.24%)

	true 0	true 1	class precision
pred. 0	9647	262	97.36%
pred. 1	14	77	84.62%
class recall	99.86%	22.71%	

Model (1 of 100 sample dataset from model)



Conclusion



CRISP-DM approach :

Business understanding	Data understanding	Data preparation	Modeling
• **Predict Machine health status in order to improve machine effectiveness.**	• **Predictive maintenance datasets, 14 columns (CNC machine)** • **Quality issue** 1.) Discrete data 2.) Different scale 3.) Imbalanced data 4.) Useless attribute • **Insight from preliminary analysis** 1.) Data range between failure and no failure is different. 2.) Negative correlation between - Torque(Nm) & Rotational speed(rpm) - Torque(Nm) & Machine failure. 3.) Positive correlation between Process & Air temperature.	• **Solve the quality issue** 1.) Nominal to Numerical attribute. 2.) Select attribute 3.) Normalize 4.) Cross validation	• **Classification machine health status** 1.) Decision Tree (97.39%) 2.) Logistic Regression (96.97%) 3.) Random forest (97.24%)

Future work

How does predictive maintenance work?