TRANSFORMERS MAGAZINE'S
INDUSTRY NAVIGATOR

INVESTMENTS, ARTIFICIAL INTELLIGENCE AND SUSTAINABILITY

CONFERENCE 2024

Machine Learning Algorithms in the Transformer Industry Supporting Transformer Monitoring and Condition Assessment

Luiz V. Cheim, PhD

Senior Principal R&D Engineer
Transformers Business
Hitachi Energy

Madrid, 12 June 2024

A Few Examples of ML Application in Transformers

1. Transformer Condition Assessment with ML (Cigre Paris, 2018 Paper)
2. Train ML Algorithms with Duval Pentagons
3. Load Forecast Based on dynamic online monitoring data
4. Use ML to detect bushing insulation failure
5. Probabilistic Bayesian Networks
6. The Superminds Project

PS.: All the above ideas are Patent protected

HITACHI
Inspire the Next

Hitachi Energy

Machine Learning Application 1:

Transformer Condition Assessment with Machine Learning

CIGRE Paper 2018

PS2 - Best Paper Award

21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

A2-206

CIGRE 2018

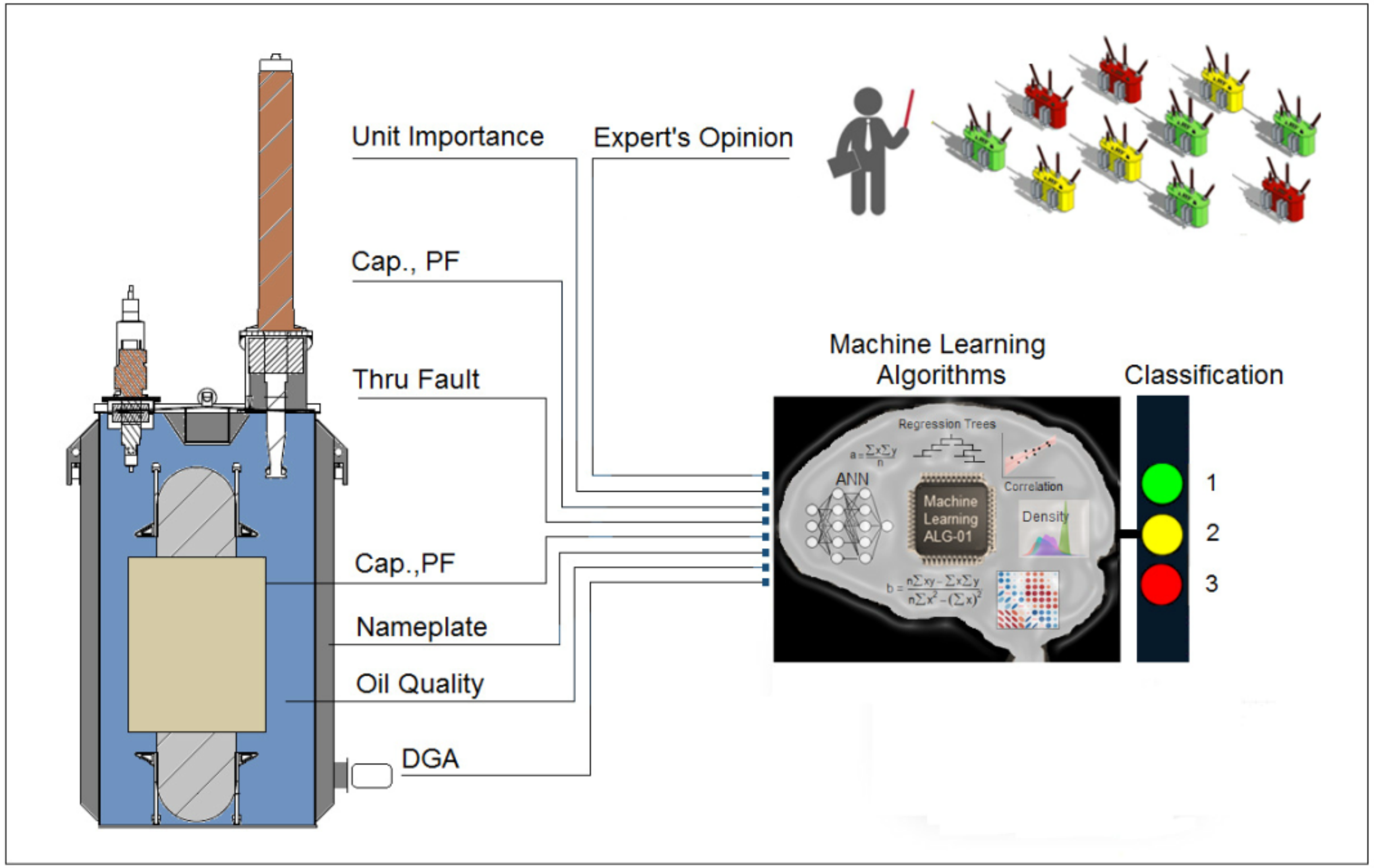
Machine Learning Tools in Support of Transformer Diagnostics

L. CHEIM
ABB
USA

SUMMARY

The paper describes the use of Machine Learning (ML) algorithms as supporting tools for the automatic classification of power transformers operating condition. The successful use of ML tools may find multiple applications in the industry such as providing fast ways of analyzing new data streaming from online sensors, evaluating the importance of individual variables in the context of transformer condition assessment and also the need or adequacy of data imputation in the so widely common problem of missing data. The work consists of training12 ML algorithms with real data from 1,000 (one thousand) transformers that were individually analyzed by human experts. Each transformer in the database was scored with a 'green', 'yellow' or 'red' card depending on the data, the interpretation of human experts, or even after some calculations carried out by the company's internal algorithms frequently utilized by the experts to identify units with technical operational issues. The ML algorithms, however, did not utilize or were given any of the engineering tools employed by the human experts. The algorithms only employed the raw data in a supervised learning process in which a column named 'Class' was added to the transformer information with the classification red, yellow or green provided by the human expert. The paper describes the main steps towards the training of the ML algorithms and the stunning output produced by those algorithms when requested to analyze 200 unseen cases during training. A very important discussion about the common missing data issue and how to handle it is also provided.

The Basic Idea



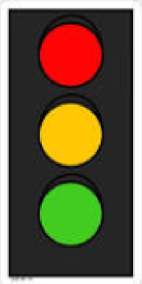
The dataset (800 training cases, 200 test cases)

22 features →

1,000 transformers ↓

Class

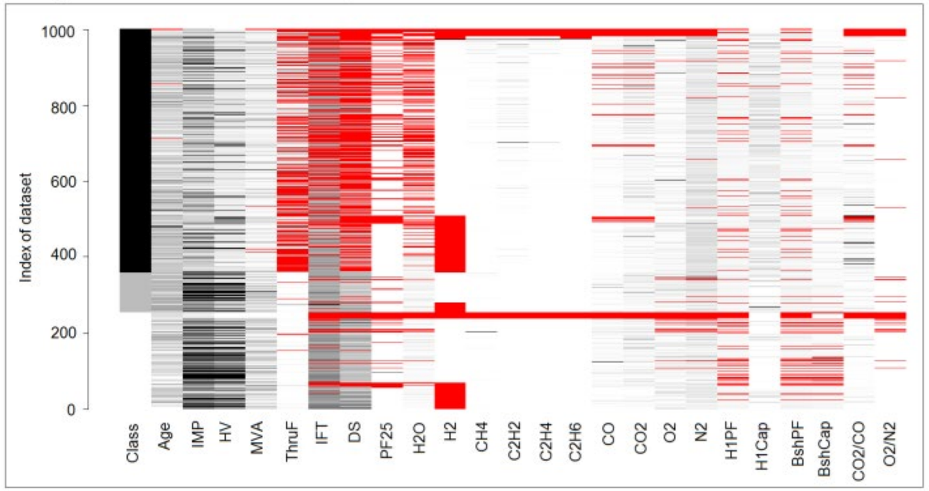
Age	IMP	HV	MVA	TF	IFT	DS	PF25	H2O	H2	CH4	C2H2	C2H4	C2H6	CO	CO2	O2	N2	H1PF	H1Cap	BshPF	BshCap	Class
1	70	345	50	5	48	72	0.002	1	0	0	0	0	0	43	617.6	25427.2	60699.6	0.33	60.481	0.33	604.81	1
2	100	765	750	2.5	49	36	0.003	1	0	0	0	0	0	5	43	9036	33887	0.59	62.11	0.59	621.1	1
33	100	765	500	0.5	39.9	43	0.003	4.6	10	38	0	10	61	59	1145	7387	45707	0.44	58.639	0.44	586.39	3
2	70	345	50	5	48	63	0.003	4	0	0	0	0	0	76.8	655.4	16557.6	40380.6	0.33	57.301	0.33	573.01	1
2	70	345	50	5	47	68	0.003	7	0	0	0	0	0	15.5	162.2	11183	25795.7	0.36	59.85	0.36	598.5	1
2	70	345	50	5	44	69	0.003	7	0	0	0	0	0	16.3	198.3	14992.4	33496.2	0.36	[redacted]	0.36	602.23	1
34	64	230	313.6	[redacted]	32	25	0.003	6	5	2	2	1	0	16	1157	1270	142262	0.68	[redacted]	0.97	145	3
42	56	230	360	[redacted]	31.7	[redacted]	0.003	9	6	9	20	3	0	249	671	16531	37260	0.34	[redacted]	0.288	1127.95	3
35	73.9	138	688	[redacted]	30	[redacted]	0.003	2	78	0	16	26	0	2	100	9151	83255	0.44	[redacted]	0.47	328.8	3
14	93.8	345	590	[redacted]	37	[redacted]	0.003	9	8	11	3	5	0	335	1486	7681	51674	0.32	[redacted]	0.27	5789	3
51	8.1	154	13.33	[redacted]	32	[redacted]	0.003	25	8	0	0	7	0	157	1001	16366	68058	[redacted]	[redacted]	[redacted]	362.05	3
5	15.8	138	10.5	[redacted]	39	[redacted]	0.003	35	7	1	0	0	0	8	314	5800	71050	0.28	[redacted]	0.2	326.9	3
35	100	765	500	0.5	43.7	37	0.004	5.9	66	43	0	2	15	202	2216	19404	55852	0.49	58.268	0.49	582.68	3
24	33	138	56	[redacted]	41	[redacted]	0.004	15	54	5	3	12	0	48	384	2258	66577	0.63	1723.2	0.63	396.95	3
48	28.4	138	180	[redacted]	35	[redacted]	0.004	4	4	0	1	2	0	10	1313	4901	78987	0.5	4990.9	0.68	1545.8	3
39	41.6	500	133	[redacted]	38	[redacted]	0.004	9	33	4	1	4	0	16	1898	4465	82914	0.34	2112	0.28	1008.5	3
29	100	765	50	0.1	40.5	37	0.005	3.2	30	245	0	82	182	[redacted]	[redacted]	614	39099	0.51	43.132	0.51	431.32	1
29	100	765	50	0.1	40.8	40	0.005	6.6	65	283	0	119	166	[redacted]	[redacted]	1037	45507	0.51	42.529	0.51	425.29	1
41	25	230	360	[redacted]	29	[redacted]	0.005	4	24	20	6	4	0	[redacted]	[redacted]	4999	34615	0.47	3531.3	0.39	263.95	3
2	100	765	750	2.5	50.5	33	0.006	1	4	0	0	0	0	5	49	4734	19207	0.36	45.273	0.36	452.73	1
2	100	765	750	1	[redacted]	37	0.006	5	1	0	0	0	0	1	38	1594	3700	0.29	56.401	0.29	564.01	1
29	100	765	50	0.1	39.5	39	0.006	5.5	1	68	0	5	129	136	1897	1337	45997	0.46	42.843	0.46	428.43	1
2	70	345	50	5	48	73	0.006	6	28	0	0	0	2.2	13.8	168.1	14520.1	35966.7	0.37	59.789	0.37	597.89	1
20	20.9	138	44.8	[redacted]	35	[redacted]	0.006	3	0	1	0	0	0	3	793	3657	75581	0.37	1020	0.25	280.5	3
33	15	138	35	[redacted]	41	28	0.007	11	0	1	1	4	0	5	446	2430	93300	0.73	908.2	0.22	156.8	3
27	100	765	50	0.1	35.1	31	0.007	8.1	16	28	0	6	17	326	996	5824	51682	0.28	37.3	0.28	373	1
2	70	345	50	5	47	69	0.007	4	26	0	0	0	0	15	158.4	11516.2	24266.5	0.35	59.318	0.35	593.18	1
33	16	69	35	[redacted]	38	19	0.008	7	8	0	0	0	0	3	132	3900	103000	0.89	941.9	0.26	121.25	3
13	100	765	500	3.7	36.8	34	0.009	6	5	16	0	18	10	543	1554	1395	17057	0.31	58.86	0.31	588.6	1
…																						
9	100	746	250	26.2	35.2	39	0.009	4.7	2	26	0	3	59	300	915	14161	48826	0.305	60.836	0.305	608.36	1
8	100	765	100	0.1	36.7	41	0.009	3	7	8	0	8	5	363	742	2467	12308	0.37	57.149	0.37	571.49	1
39	38.8	345	112	[redacted]	34	[redacted]	0.009	2	8	5	7	3	0	14	2141	4696	74455	0.64	1334	0.58	522.5	3
54	55.6	161	170	0.2	33	45	0.01	10	0	2	3	4	0	5	543	811	76908	0.39	2568.6	0.33	226.15	1
54	4.76	161	25	0.2	39	46	0.01	5	2	1	0	0	0	5	105	13558	44463	1.03	1103.6	0.7	163.15	1



Internal

Data Preparation in Machine Learning (70-80% of effort!)

Missing data



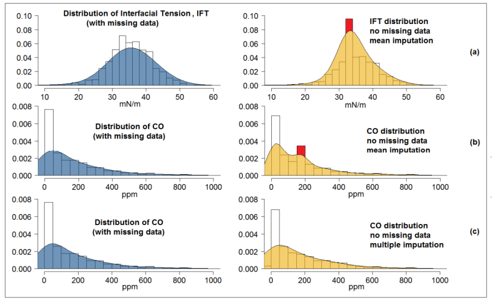
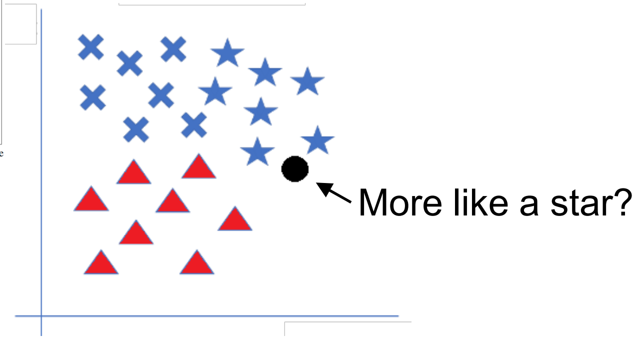


Figure 4 – Illustration of mean imputation (a) and (b) as compared to multiple statistical imputation in (c). Notice the significant effect of mean imputation on the probability density functions of (a) and (b).

Data imputation strategies



Avoid collinearities, Remove low variance features, etc..

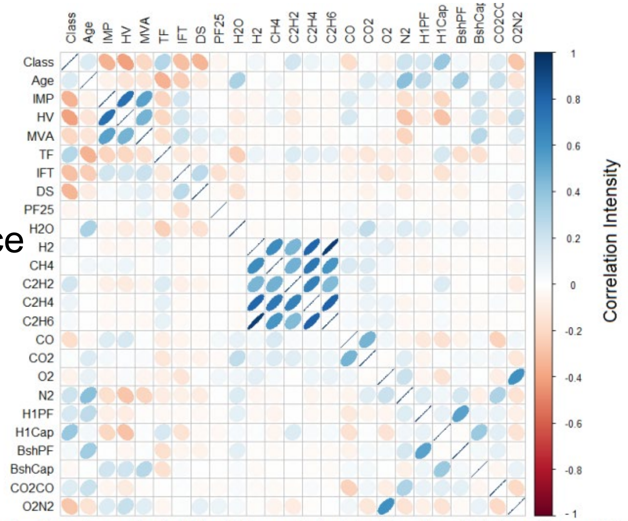


Figure 5 – Correlation matrix of all 24 variables used in the study. Blue ellipses indicate positive correlation; red ellipses show negative correlation. Color intensity is proportional to correlation. Blank squares = no correlation.

How to handle outliers?

Internal

Best ML Algorithms (10-fold CV training, with 3 repeats)

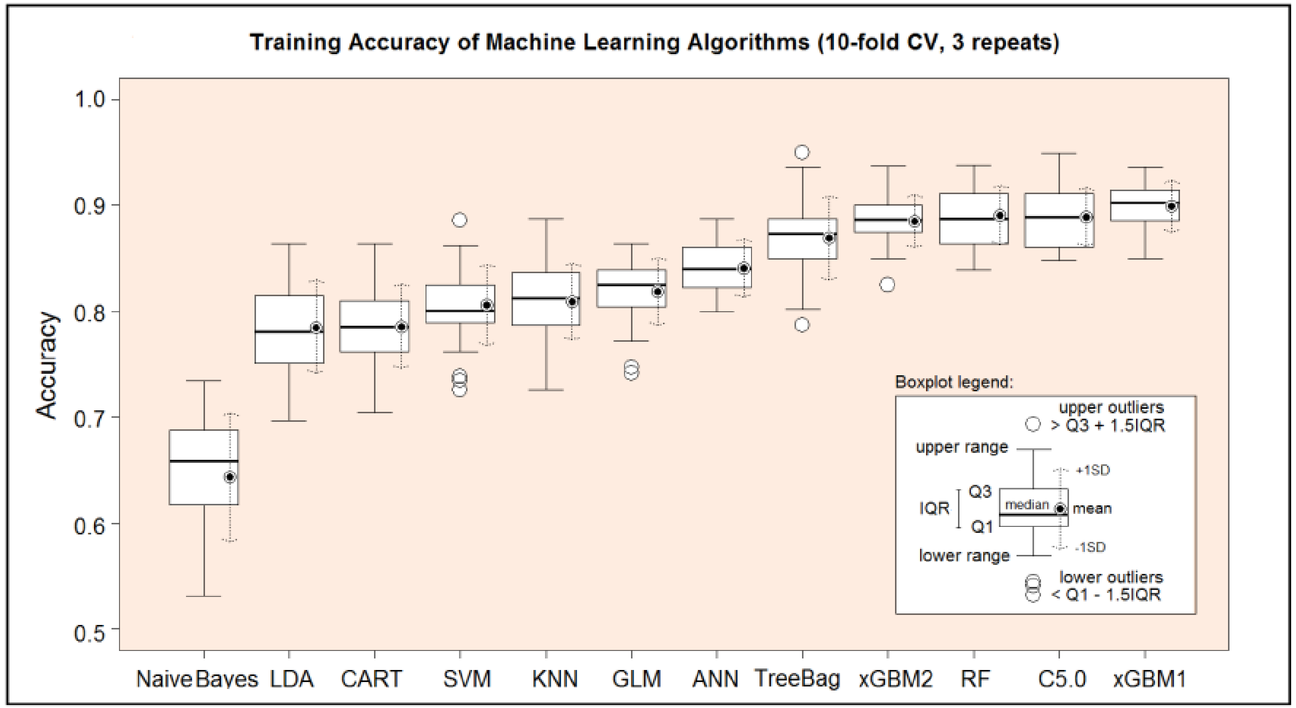


Figure 9 – Comparative accuracy of Machine Learning algorithms after training 12 models with 80% of the available data, by 10-fold cross validation (CV) and 3 repeats. The ML algorithms were Naïve Bayes, Linear Discriminant Analysis (LDA), Classification and Regression Trees (CART), General Linear Model (GLM), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Artificial Neural Networks (ANN), Tree Bagging, Extreme Gradient Boosting Machine (xGBM1 and xGBM2), Random Forest (RF) and C5.0.

ML Prediction ↓ / Human expert classification →	Green	Yellow	Red
Green	61	3	0
Yellow	0	14	0
Red	1	3	118
Totals	62	20	118

Algorithm Accuracy = (61 + 14 + 118) / 200 ≈ 97%

Machine Learning Application 2:

Train ML Algorithms with the Combined Duval Pentagons

Combined Duval Pentagons – Energies 2020 Joint Paper with M. Duval

energies

MDPI

Article

Combined Duval Pentagons: A Simplified Approach

Luiz Cheim [1,*], Michel Duval [2] and Saad Haider [1]

1 Transformer Technology Center, ABB Power Grids, St. Louis, MO 63127, USA; saad.haider@us.abb.com
2 Institut de Recherche d'Hydro-Québec (IREQ), Varennes, QC J3X 1S1, Canada; Duval.Michel@ireq.ca
* Correspondence: luiz.v.cheim@us.abb.com; Tel.: +1-314-452-1445

Received: 10 April 2020; Accepted: 2 June 2020; Published: 3 June 2020

check for updates

Abstract: The paper describes a newly proposed combination of the two existing Duval Pentagons method utilized for the identification of mineral oil-insulated transformers. The aim of the combination is to facilitate automatic fault identification through computer programs, and at the same time, apply the full capability of both original Pentagons, now reduced to a single geometry. The thorough classification of a given fault (say, of the electrical or thermal kind), employing individual Pentagons 1 and 2, as originally defined, involves a complex geometrical problem that requires the build-up of a convoluted geometry (a regular Pentagon whose axes represent each of five possible combustible gases) to be constructed using computer language code and programming, followed by the logical localization of the geometrical centroid of an irregular pentagon, formed by the partial contribution of individual combustibles, inside two similar structures (Pentagons 1 and 2) that, nonetheless, have different classification zones and boundaries, as more thoroughly explained and exemplified in the main body of this article. The proposed combined approach results in a lower number of total fault zones (10 in the combined Pentagons against 14 when considering Pentagons 1 and 2 separately, although zones PD, S, D1 and D2 are common to both Pentagons 1 and 2), and therefore eliminates the need to solve for two separate Pentagons.

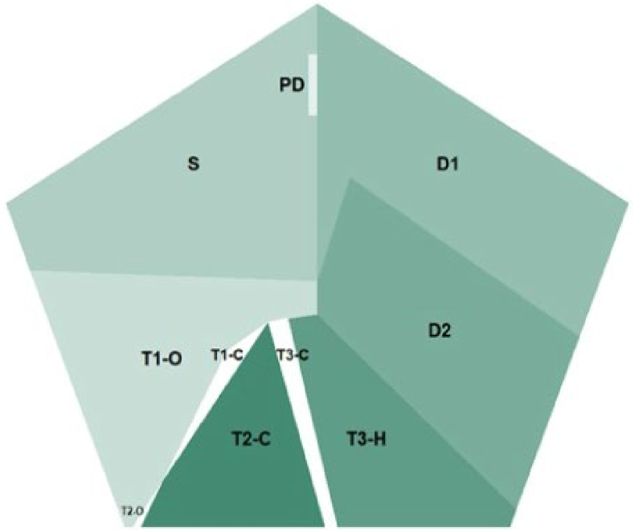


Figure 9. Final geometriy of the newly Combined Pentagons. The 10 zones are PD, S, D1, D2, T1-O, T1-C, T2-O, T2-C, T3-C, T3-H.

HITACHI
Inspire the Next

Train ML Algorithms with the Combined Pentagons – Artificial Data Generated

cigre
For power system expertise

Here your Paper ID – 3 or 4 numbers **Session 2022**
Study Committee D1
PS1 – Testing, Monitoring and Diagnostics

Machine Learning Algorithm Trained by the Duval Pentagons
A Simplified Duval Pentagons Approach

Luiz CHEIM
HITACHI ENERGY
USA
luiz.v.cheim@hitachienergy.com

SUMMARY

Artificial intelligence (AI) and Machine Learning (ML) algorithms have been successfully used in many applications in support of human-expert activities, notably in situations that require routine work and, sometimes, excessive number of repetitive processes. The application presented in this paper considers the need for continuous condition assessment of transformers based on Dissolved Gas Analysis (DGA) and applying the so-called Duval Pentagons criteria. The approach taken in the present work utilizes the so-called Combined Pentagons 1 and 2 to train a Machine Learning Algorithm to simultaneously classify any combination of the five combustible gases used in the Duval Pentagons method, without the need for the actual Pentagons, after the ML algorithm is properly trained. To achieve a high accuracy in DGA classification, and to cover a significant area of the combined pentagons during the training of the ML algorithm, a very large number of artificially generated centroids inside the pentagons was utilized.

Hitachi Energy

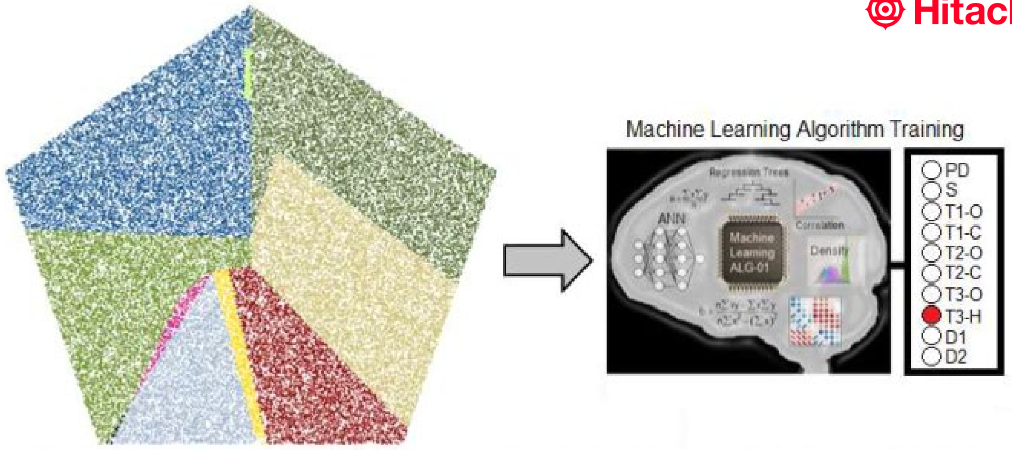
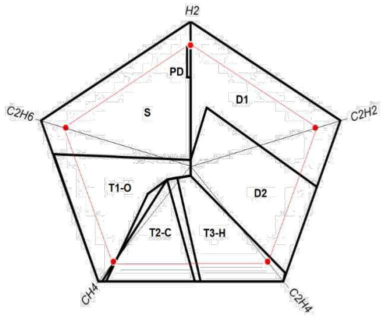


Figure 3 – Illustration of Machine Learning algorithm trained with artificially generated centroids, covering the whole area of the Combined Pentagons shown in Figures 1 and 2b, in all 10 fault regions. The illustration contains 50,000 artificially generated centroids.



The Catch

Artificially generated points inside the Pentagon may not correspond to actual transformer data BUT actual data will necessarily fall within the limits of the Pentagon → covering the whole Pentagon with artificially generated points will necessary represent all possible cases of real transformers!!!!

Machine Learning Application 3:

Short term load forecast – Optimizing Loadability

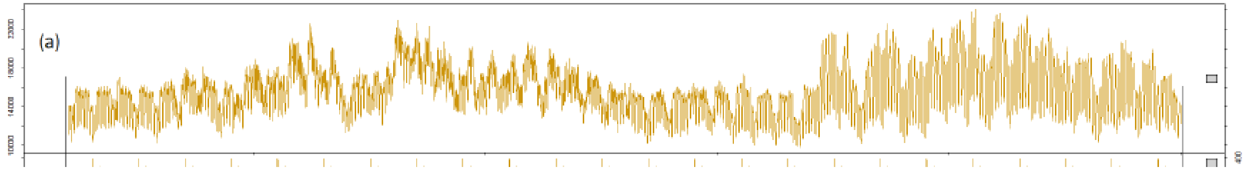
Short-time Load Forecast

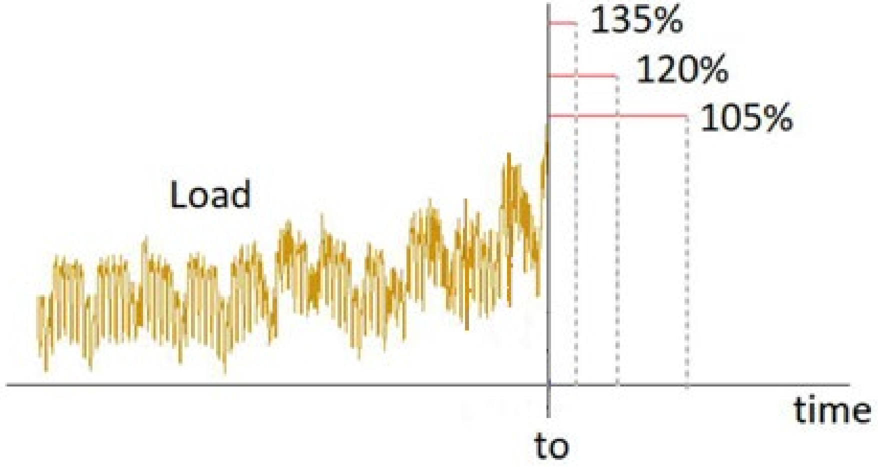
Short term load forecast with Machine Learning

Hitachi ABB Power Grids
Transformer Technology Center
By Luiz Cheim, PhD
2019

1. Introduction

Load forecast is a field of investigation on its own, having many applications in the electricity market, and it has been the subject of a large number of studies and applied statistical tools, such as for example ARIMA (auto-regressive integrated moving average) and others, both for long term as well as short term forecast. In the process of searching for the ideal tool, it is a common practice to break the time serie that represents the historical load into a number of components, such as the trend component, th seasonal and the random variation as illustrated in Figure 1 below. This allows researchers to fin adequate models to each component and search for the ideal aggregation of all components for th final forecast. In trying to minimize the prediction error researchers look into the addition of othe correlated parameters, such as for example daily average temperature, seasonal festivals and sport events, etc. This carries additional complication to the already complex problem, considering that th identification of such events in the training and validation datasets requires highly specialized tools tha can achieve such data tagging in generic applications (for example, when testing the algorithm in tw completely different states of a given country or in different countries with their specifics cultures an holidays).

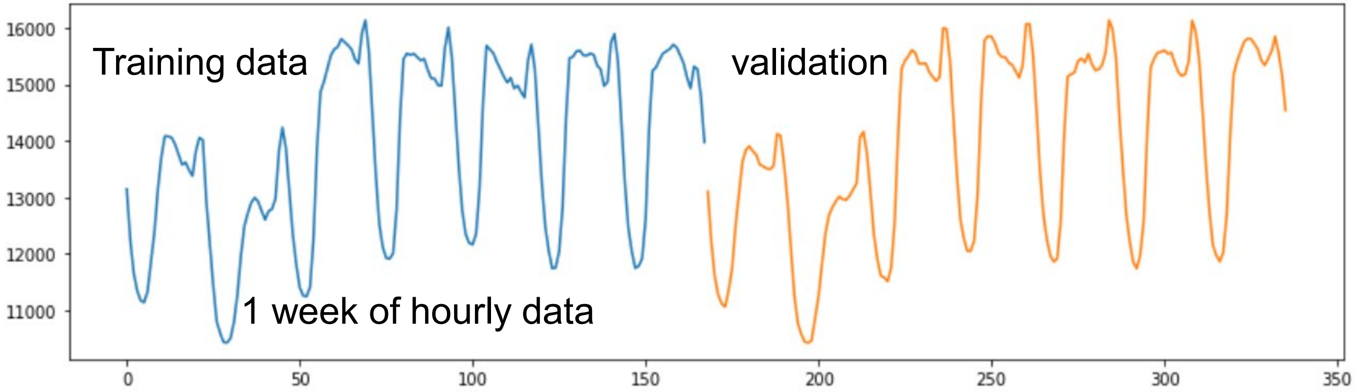




HITACHI
Inspire the Next

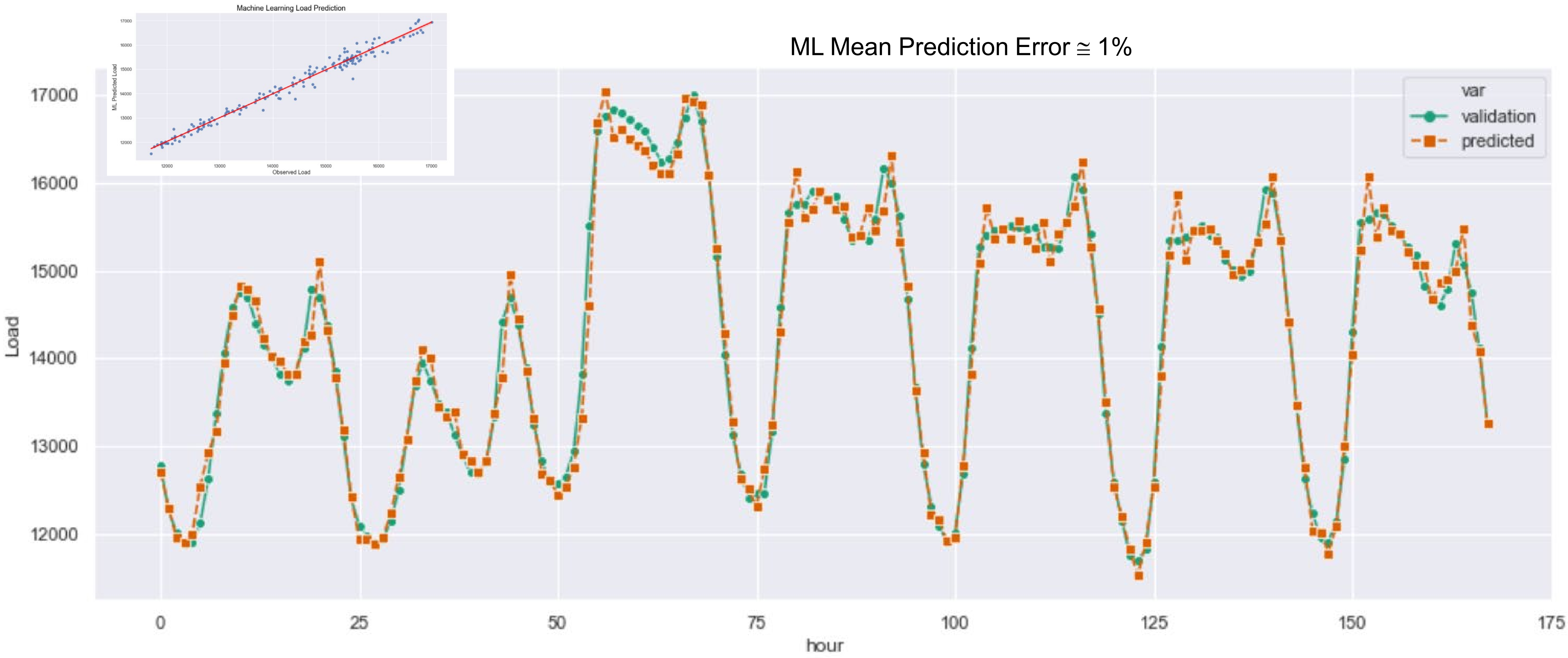
Short-time Load Forecast

Hitachi Energy



Short-time Load Forecast (no retraining)

ML Mean Prediction Error ≅ 1%



Predicting 2h, 4h and 8h ahead

7. Prediction of 2h Ahead

For the prediction of 2h ahead by the ML a dataset training equivalent to 1 year of hourly load was utilized. Figure 23 below shows the still impressive result of prediction for 1 week ahead, after the 1-year training set, where every hour inside the validation week was predicted 2h earlier by the ML algorithm.

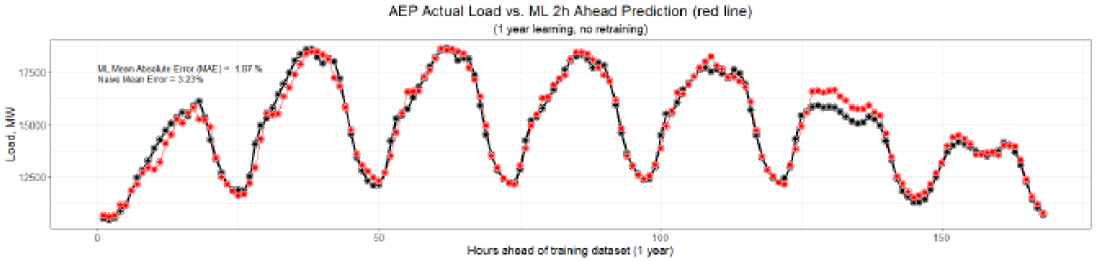


Figure 23 – Comparison between actual load (black dots) and the Machine Learning prediction, 2h ahead, for a whole week (168h) ahead of the 1-year training dataset. Average error still below 2%.

8. Prediction of 4h Ahead

For the prediction of 4h ahead by the ML a dataset training equivalent to 1 year of hourly load was utilized. Figure 24 below shows the result of prediction for 1 week ahead, after the 1-year training set, where every hour inside the validation week was predicted 4h earlier by the ML algorithm.

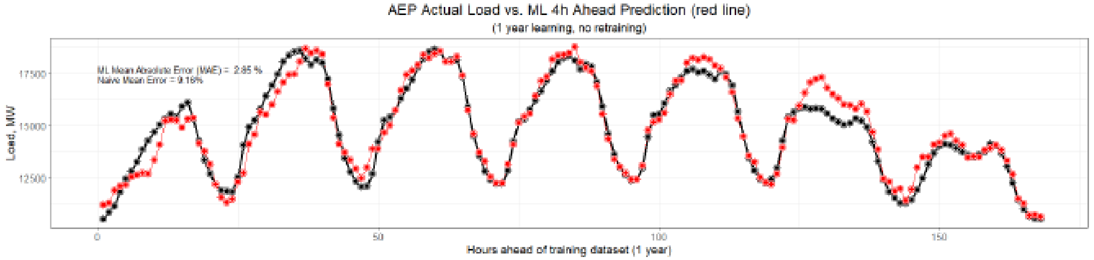


Figure 24 - Comparison between actual load (black dots) and the Machine Learning prediction, 4h ahead, for a whole week (168h) ahead of the 1-year training dataset. Average prediction error still below 3%. Naïve predictor average prediction error near 10%.

10. Prediction of 8h Ahead

For the prediction of 6h ahead by the ML a dataset training equivalent to 1 year of hourly load was utilized. Figure 26 below shows the result of prediction for 1 week ahead, after the 1-year training set, where every hour inside the validation week was predicted 8h earlier by the ML algorithm.

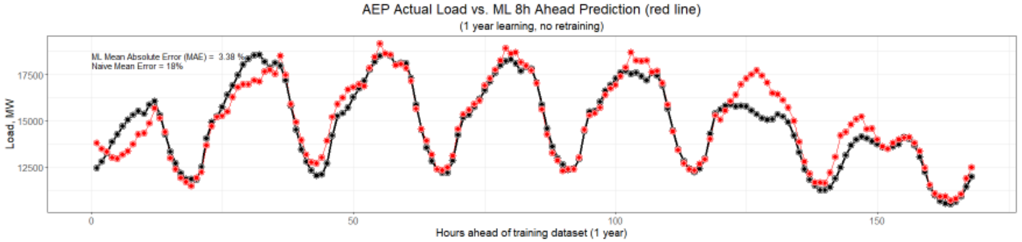


Figure 25 - Comparison between actual load (black dots) and the Machine Learning prediction, 6h ahead, for a whole week (168h) ahead of the 1-year training dataset. Average prediction error still just above 3%. Naïve predictor average prediction error 18%.

Machine Learning Application 4:

Dynamically detect insulation failure (cap./pf)
using ML forecasting capability

Detect Bushing Capacitance and Power Factor failure with ML

Machine Learning Forecasting Tool in the Detection of Insulation Degradation (Insulation Capacitance and Power Factor)

Hitachi ABB Power Grids
By Luiz Cheim, Roberto Zannol and Nilanga Abeywickrama

1. Introduction

Any type of electrical insulation can be represented by an electrical equivalent circuit as illustrated in Figure 1 below. A perfect insulation submitted to an applied voltage, a shown, should not allow any loss current (I_R) but only the natural capacitive current (I_C), given the nature of the physical structure. In practice, however, no insulation is ideal and there is always some leakage current, as illustrated in the electrical equivalent circuit. This leakage current tends to be very small but its ratio to the capacitive current or the total current is an important indication of the actual condition or quality of the insulation.

This is illustrated by the vector diagram in Figure 1. The angle δ between the capacitive and total current defines the so-called dissipation factor, $\tan\delta = I_R / I_C$. For small δ (expected in any good insulation material) the complementary angle ϕ defines the so called "power factor" given by $\cos\phi = I_R / I$, which numerically is very close to tan δ, since $I \cong I_C$, as it can be seen in the vector diagram.

The power factor as sensitive to defects in the metallic plates, voids in the insulation, short-circuited plates and water/particle contamination, as illustrated in Figure 2.

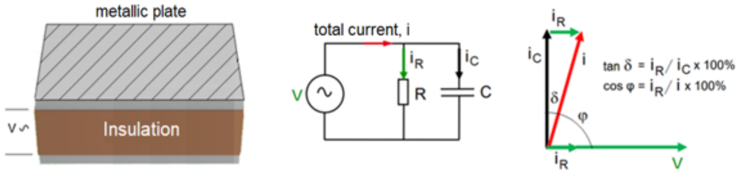
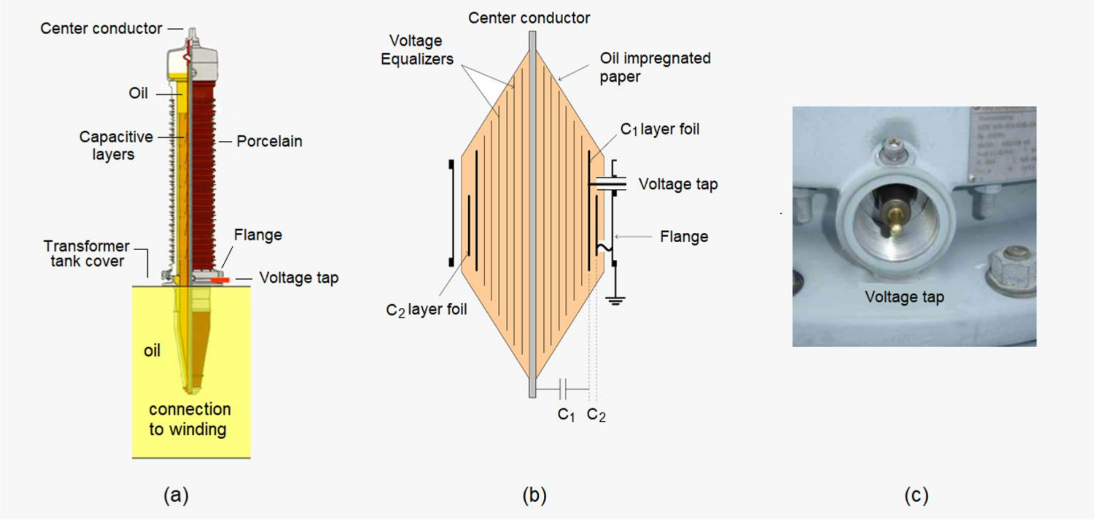


Figure 1 – Illustrative example of a generic insulation, showing the applied voltage (V), the leakage current i_R, the capacitive current, i_C and vector diagram with the relative position of each current vector. An ideal insulation would have i_R = 0. For small



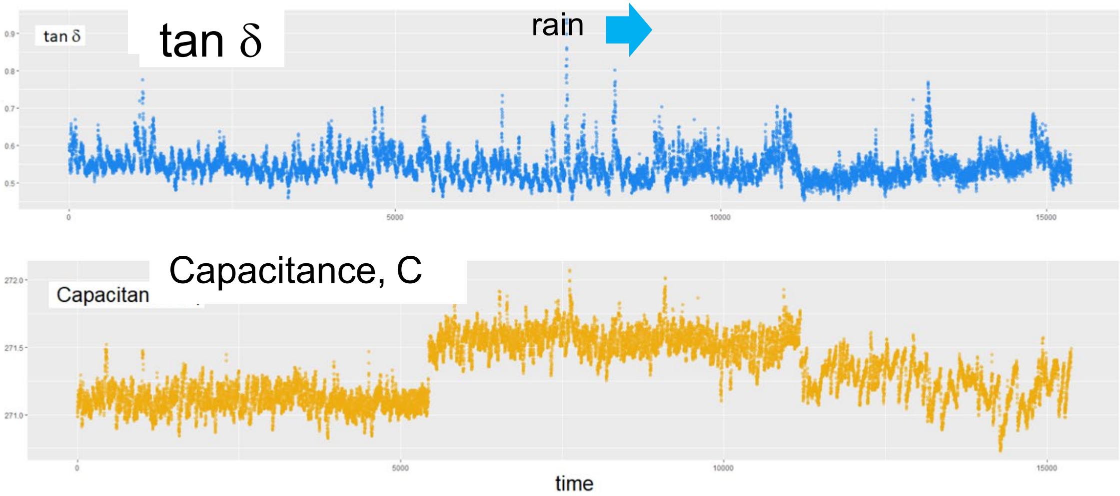


Figure 7 – Bushing tan δ and capacitance C_1 used to train the Machine Learning algorithms

Dramatic Deterioration in ML Prediction with Change in Cap.

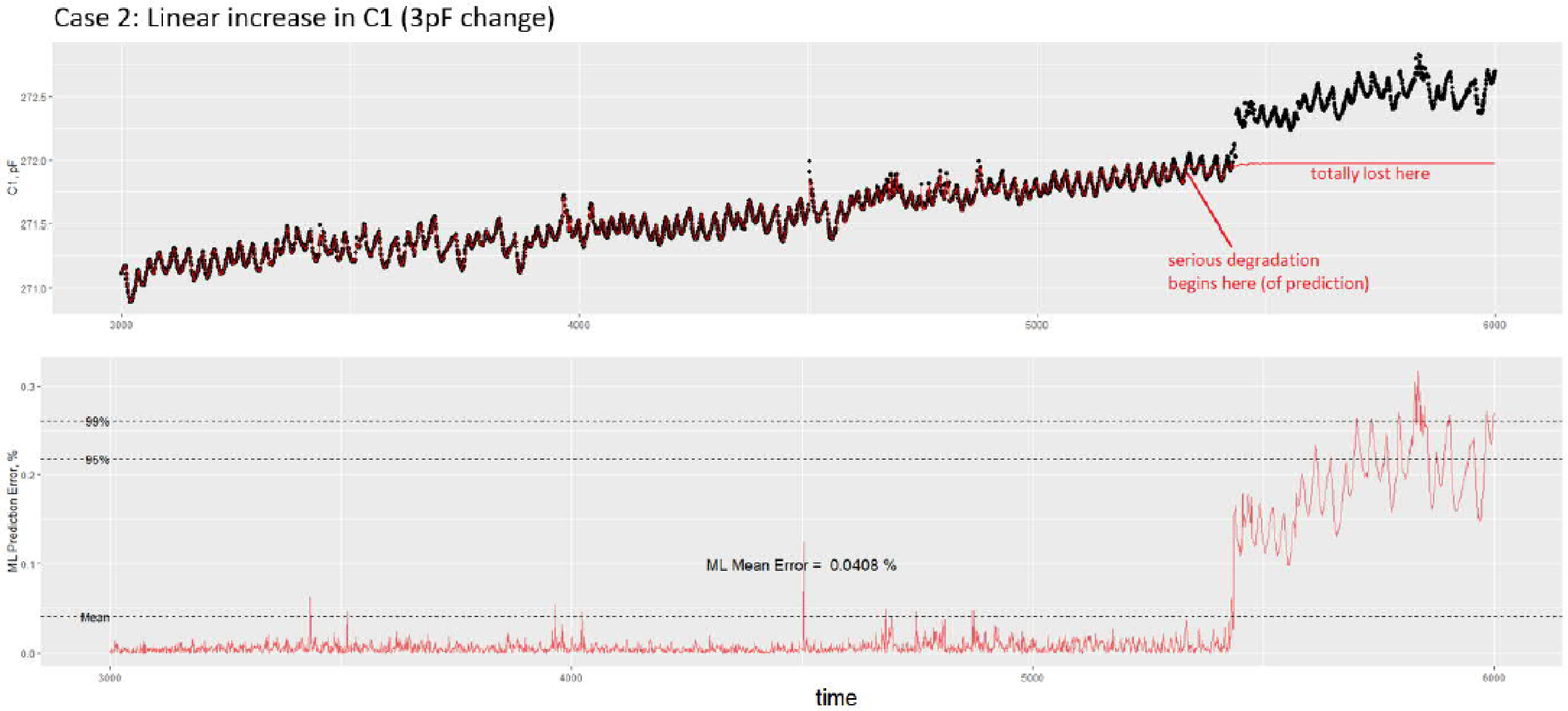
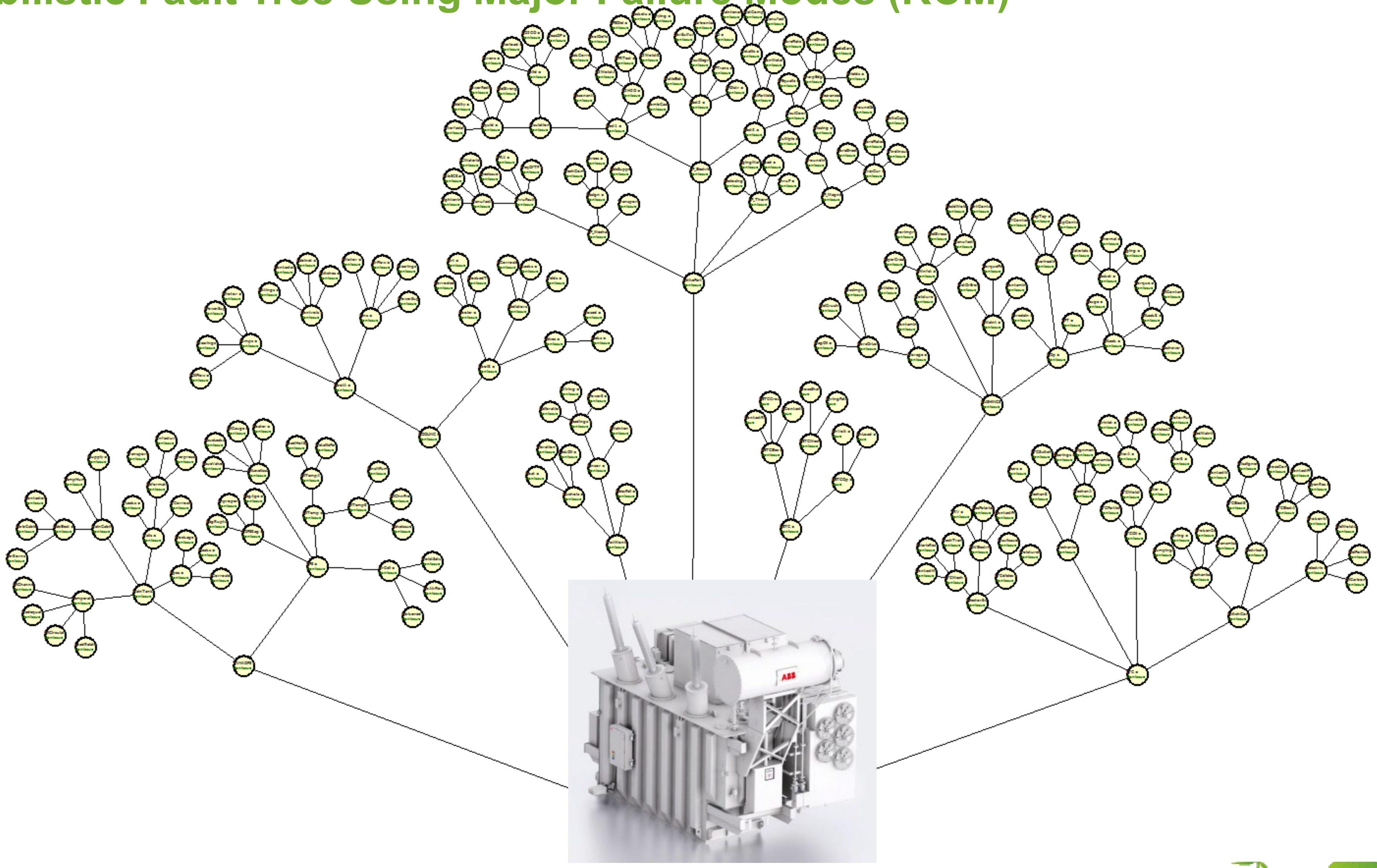


Figure 11– Detection of capacitance deviation by significant variation in the ML forecast error, for a linear change in capacitance C_1 to a total of 3pF. Top figure shows linear increase in capacitance (black dots) as compared to the ML prediction (red line). The bottom chart shows the actual deviation in error of prediction in the ML algorithm, increasing more than 10x when compared to the reference case shown in Figure 8.

Machine Learning Application 5:

Dynamic transformer condition assessment
using Probabilistic Bayesian Networks (digital twin)

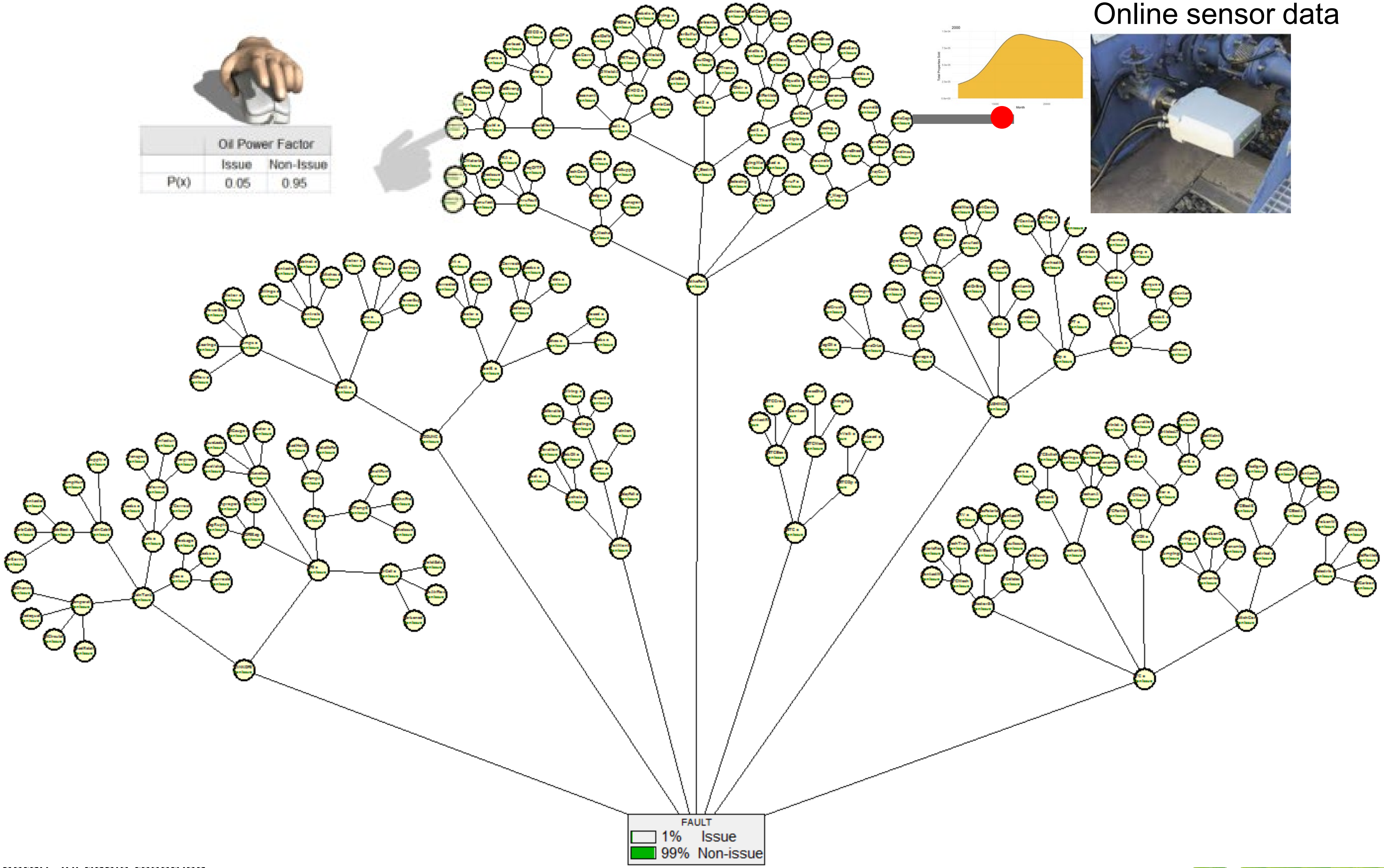
Build a Probabilistic Fault Tree Using Major Failure Modes (RCM)



Typical customer failure rate

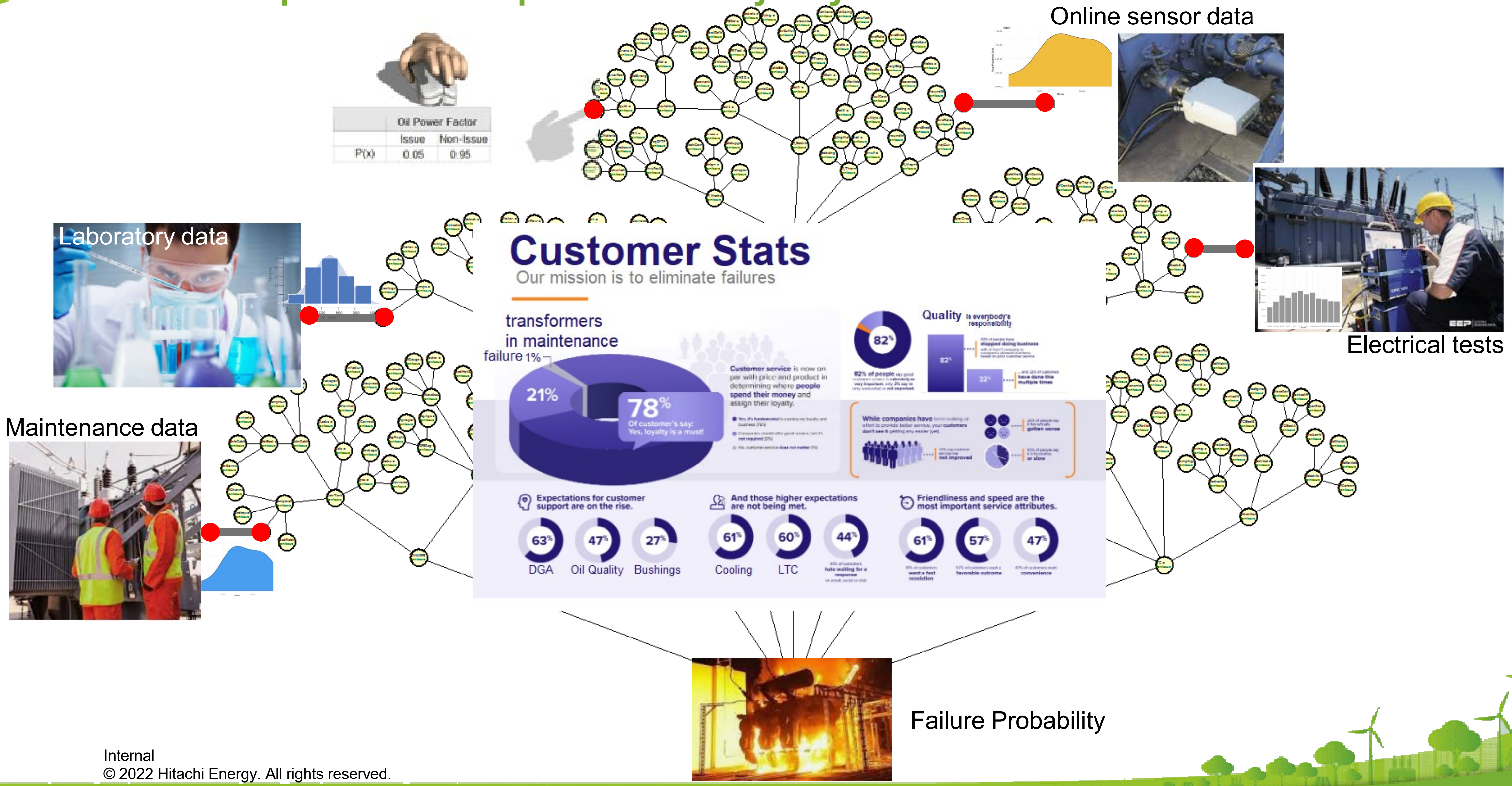
Nodes probabilities dynamically updated in many ways

	Oil Power Factor	
	Issue	Non-Issue
P(x)	0.05	0.95

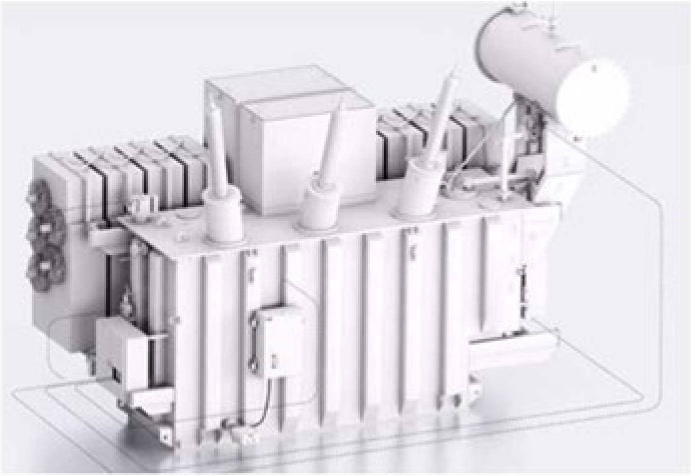


Online sensor data

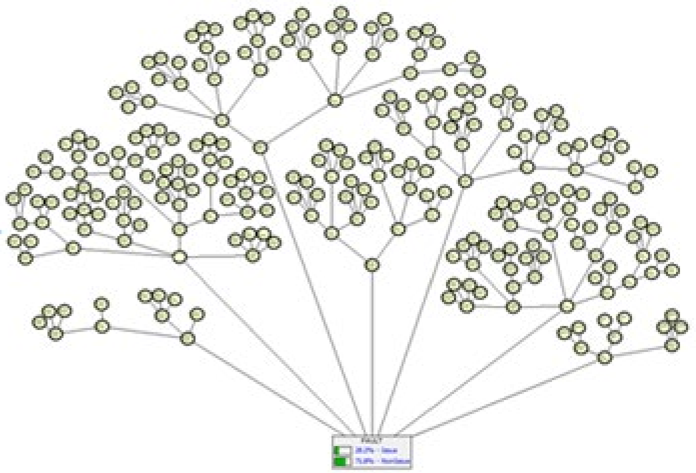
Nodes probabilities updated in many ways



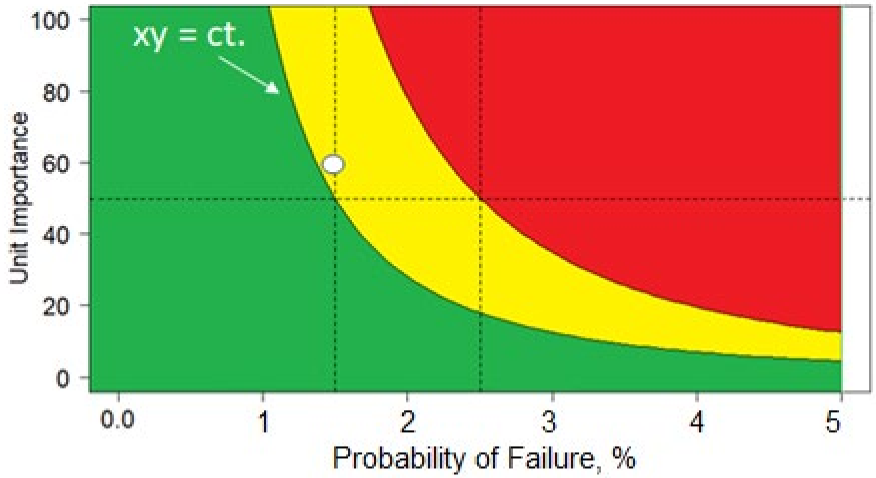
Integrate Solution into APM (the digital twin)



Probabilistic Fault Tree



Customizable boundaries between green, yellow, red
xy = constant



The Superminds Project (Patent granted in 2022): **6**

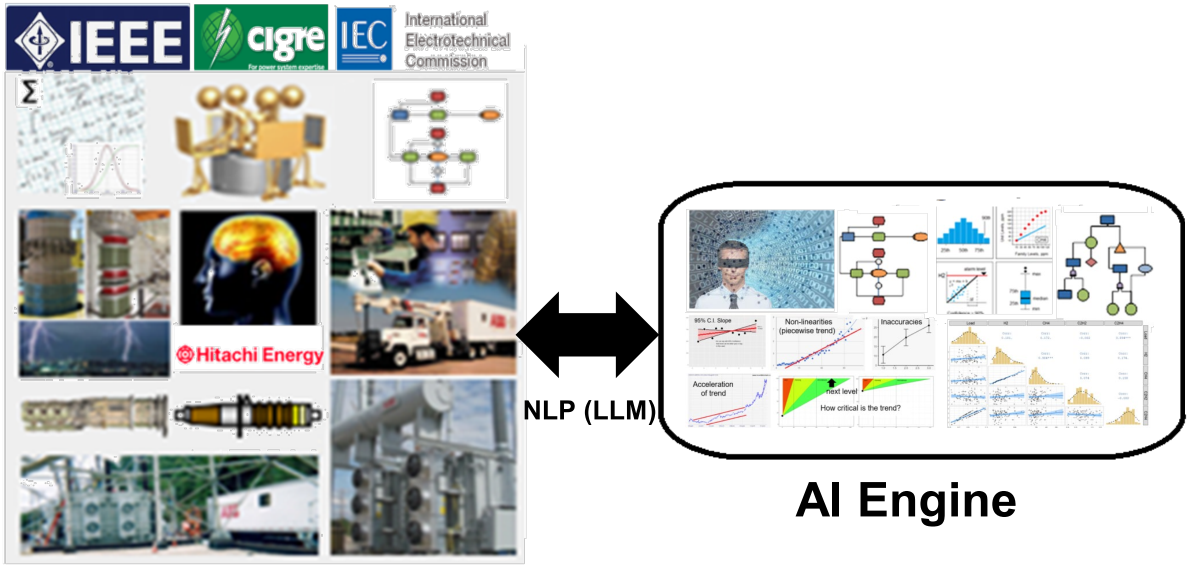
Create the Infrastructure to offer Transformer Users an ***AI Powered Engine***
for Statistical Analysis/Condition Assessment of their Assets

Superminds Idea

Build a SME guided knowledge pool and users' anonymized database from which AI/ML Applications may extract information, learn about statistical patterns of "**similar assets**", associate user's experience to knowledge and HE experience, build possible scenarios and correlations and provide the end user with recommended actions!

AI Engine Respond to user's queries, providing
most likely scenarios, stats, recommendations for action, etc.

SME Knowledge Pool



AI Engine

HITACHI
Inspire the Next

Hitachi Energy

Thank You!

Hitachi Energy

HITACHI
Inspire the Next