DIAGNOSTICS

Artificial intelligence and machine learning in the era of digital transformer monitoring: Exciting developments at Hitachi Energy

ABSTRACT

The era of digitalization brings new challenges and new paradigms since transformer users and manufacturers alike are moving towards digital solutions. This transition requires new approaches, new architectures, and new ways of looking at data collection, storage, and assessment. Speed and reliability of actionable information become essential at a time when data is ubiquitous, loads are more complex, and energy production moves from traditional plants to distributed generation. This article intends to show some of the ongoing efforts at Hitachi Energy to address these and other demanding technical and economic issues.

Our wind power forecast approach deals with the problem of uncertainty in upcoming power demand. We propose a machine learning model to predict power demand to improve the calculation of loadability and cooling / hotspot calculations. Similarly, our Bushing Tan δ and Capacitance Fault Detection solution uses the error of a model to detect problems with Tan δ and capacitance. Our Probabilistic Fault Tree describes an open-source approach that uses Bayesian networks to find the probability of failure of a specific transformer. Finally, we describe two publications made by our team regarding the use of synthetic data created using the Duval Pentagons to generate a model that diagnoses transformer faults; and a patent regarding the creation of an infrastructure that uses blockchain to anonymize users and provide them with information about their transformer fleet using artificial intelligence.

KEYWORDS:

transformers, artificial intelligence, machine learning, power forecast, fault diagnosis

Dr. Luiz CHEIM, Dr. Mauricio SOTO, Tucker REED

The analysis of this data allows us to better understand, diagnose, maintain, and support transformers throughout their life cycles

1. Introduction

It's all about the Data.

Data gives deeper insight that allows us to answer questions that otherwise could not be answered. With the rapid pace of the adoption of digitalization, generation and access to data have never been easier. Today, transformer manufacturers are making "digitally enabled" transformers the standard. Transformers are being outfitted with the sensors and monitors needed to generate, collect, analyze, and store critical transformer asset details in near real-time. The analysis of this data allows us to better understand, diagnose, maintain, and support transformers throughout their life cycles. We can now even capture data surrounding the transformer, including ambient temperature, weather patterns, wind speed, humidity, geomagnetically induced current (GIC) activity, etc. and analyze how this surrounding data influences the performance of the transformers. We are finding ways to interpret these ever-growing datasets to help system operators solve day-to-day problems and answer questions that would require vast resources and human-intensive expertise to be resolved. This is the next big hurtle! We have extensive amounts of data; how can we use it to the users' advantage while respecting the privacy and requirements of the data owner? Machine learning and artificial intelligence (AI) are exciting area of exploration for Hitachi Energy as we are starting to see the massive potential of what these types of analytics can provide to our asset owners. The techniques we are looking at, whether it be multivariate regressions, neural networks, or decision trees, allow us to leverage these massive and growing datasets to catch intricacies hidden within the data, and flag or expose areas to prevent failures, predict behavior, and increase performance.

What if we could improve our ability to forecast demand? What if we could accurately guess the life expectancy of a transformer? What if we could answer questions regarding the health of a transformer that were previously impossible to answer? What new insights would that bring? How would that change the way we monitor transformers in the field? What other questions could it answer? This is what Hitachi Energy's research is focused on: innovation, digitalization, and creating approaches to shift the way transformers are monitored. This article discusses some of the use-cases based on the initial results and the huge potential of digitalization applications.

To improve power demand prediction, we have explored different artificial intelligence approaches using data from conventional generation and wind power sources

2. Wind power forecast

The upcoming power demand is unknown. This is a common problem in the energy sector that utilities have to deal with, thus having to constantly balance electricity supply and demand. Previous approaches have been proposed [1,2] given that an improper balance between power generation and consumption may lead to technical and economic difficulties associated with, for example, a lack of power supply when the demand is at its peak or system overload that can diminish the transformers' expected life, among other problems.

To improve power demand prediction, we have explored different artificial intelligence approaches using data from conventional generation and wind power sources. Based on our experimentation, we train a deep neural network, as shown in Fig. 1, with the goal of predicting and informing our users of upcoming power demand. Even though our original datasheet is composed of thousands of records of wind power generation sampled every 15 minutes, we were able to train a neural network that requires only a small amount of data and still maintains high accuracy in its prediction. The training takes a few seconds, and the model can predict future values instantaneously once provided with the most recent data. This enables our customers to obtain a more accurate power demand prediction and take decisions based on the information provided, such as load shedding, if necessary.

Similarly, this prediction helps the monitoring and maintenance of transformers, given that we can provide a better calculation of transformer loadability and better usage / control of cooling and hotspot temperature.

In our approach, we use a suitable volume and quality of data collected anonymously to train a deep neural network and predict the following power readings, as shown in Fig. 2.

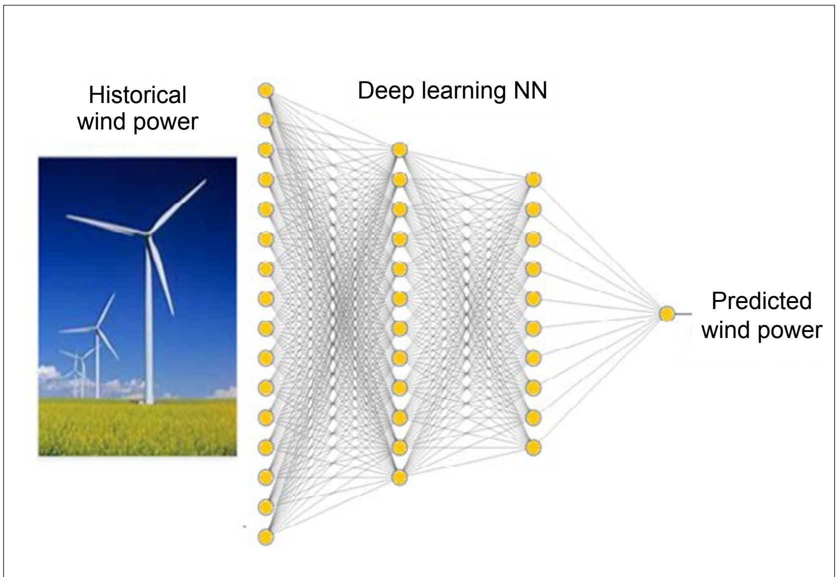


Figure 1. Deep learning artificial neural network utilized to predict wind power

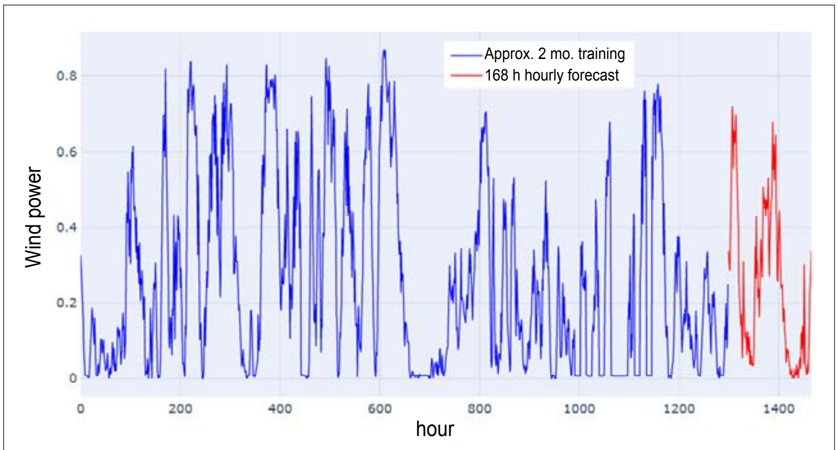


Figure 2. Wind power training data (blue) and one week of validation dataset (red) utilized to verify the machine learning algorithm (wind power is measured in per unit based on the maximum kW generated).

Fig. 3 displays the normalized fit between the predicted power readings and the actual power readings. In an optimal prediction, the red dots would be perfectly aligned with the blue line, meaning that the predicted power and actual power match perfectly. As seen in the graph, the alignment between predicted and actual power is very close to optimal with a coefficient of determination (r-squared) of 0.99 and root mean squared error (RMSE) of 52.17 kW. R-squared is a value between [0,1] that represents to what extent the variance of one variable explains the variance of the second variable. Similarly, the RMSE indicates that, on average, our predictions were just 52.17 kW away from the actual values. Finally, the mean absolute percentage error (MAPE) describes the ratio between predicted and actual values as a percentage of the actual values. We used 10-fold cross-validation to avoid overfitting concerns, where data is evaluated 10 times in a held-out portion of the data. Additionally, our dataset is composed of wind power from a variety of sources, which also implies our model is more likely to properly generalize to future instances.

Overall, our approach can very accurately predict power demand, becoming a very powerful tool for our customers and for the improvement of monitoring of transformers. Furthermore, the use of wind turbines and their potential as an energy source can greatly vary depending on different factors, such as geographical location and capacity. Thus, our model can also be retrained to optimize for such factors.

The historical data was used to train machine learning algorithms, and artificially generated defects were added to the data to assess the prediction error of the algorithms

3. Bushing tan δ and capacitance fault detection

The same methodology applied to the load forecast above was applied to historical data of bushing capacitance, C1 and tan δ, illustrated in Fig. 4a–4e. Figures 4d and 4e show real online data acquired over a six-month period from an oil-impregnated paper (OIP) bushing on a 3-phase, 130 kV transformer. The historical data was used to train machine learning algorithms, and artificially generated defects were added to the data to assess the prediction error of the algorithms.

Fig. 5 illustrates the significant increase in the prediction error starting when the artificial defect was introduced. Fig. 5a illustrates a linear increase in tan δ and Fig. 5b a step change in bushing capacitance, C1. Notice the abrupt change in machine learning prediction error. In this case, the error in prediction is used to indicate significant changes in both tan δ and capacitance.

4. Probabilistic transformer fault tree

The implementation of our proprietary monitoring algorithms is confidential; thus, we are not able to share the details of how these algorithms operate, even though some of our customers would like to understand how they are implemented. Hitachi Energy's goal is to provide our customers with the best insight and experience possible when it comes to feedback from our monitoring algorithms. To provide the maximum understanding

Our approach can very accurately predict power demand, becoming a very powerful tool for our customers and for the improvement of monitoring of transformers

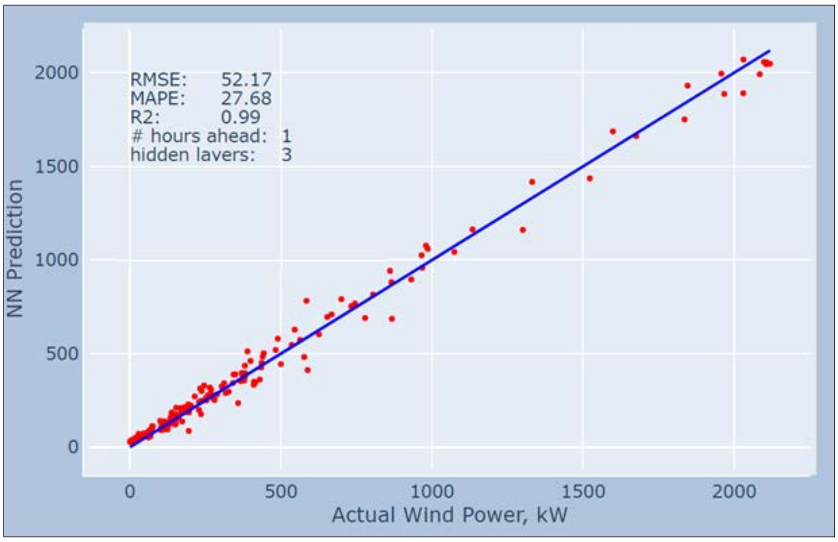


Figure 3. Correlation between actual wind power (horizontal axis) and predicted wind power by the machine learning algorithm. Predictions were made for 1h ahead, short-term forecast.

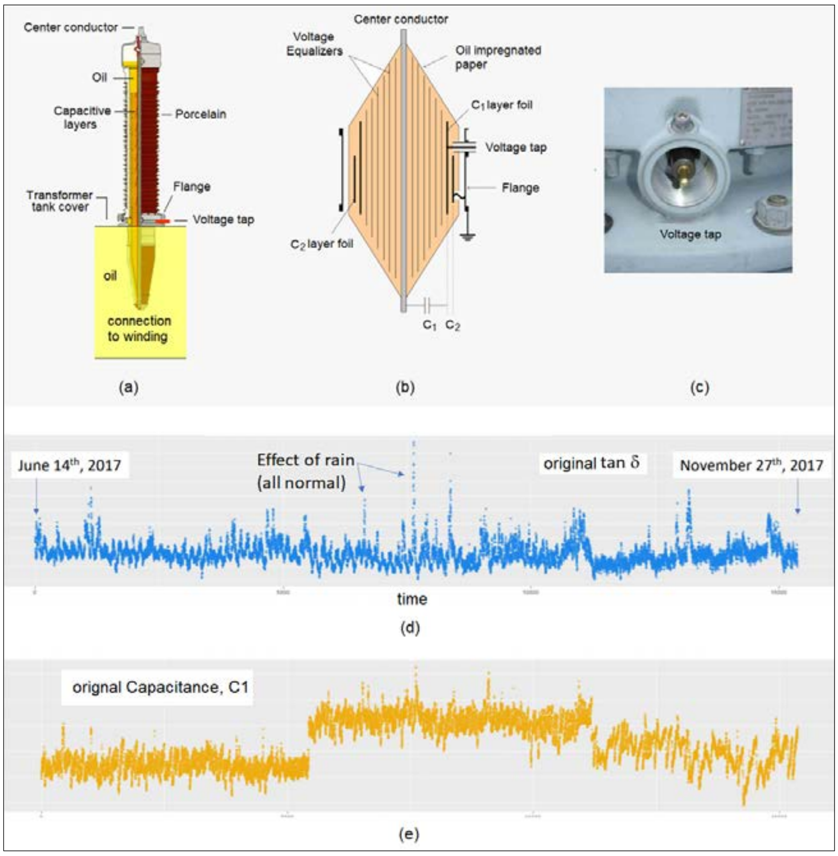


Figure 4. Illustration of (a) transformer bushing and its (b) capacitive layers, also showing the (c) voltage tap utilized to install online monitoring systems; (d) 6 months of online data collection for the tan δ and (e) capacitance, C1.

We are proposing an approach that combines our vast transformer domain knowledge with an open-source alternative based on a Bayesian network to help our customers determine the probability of failure of a transformer

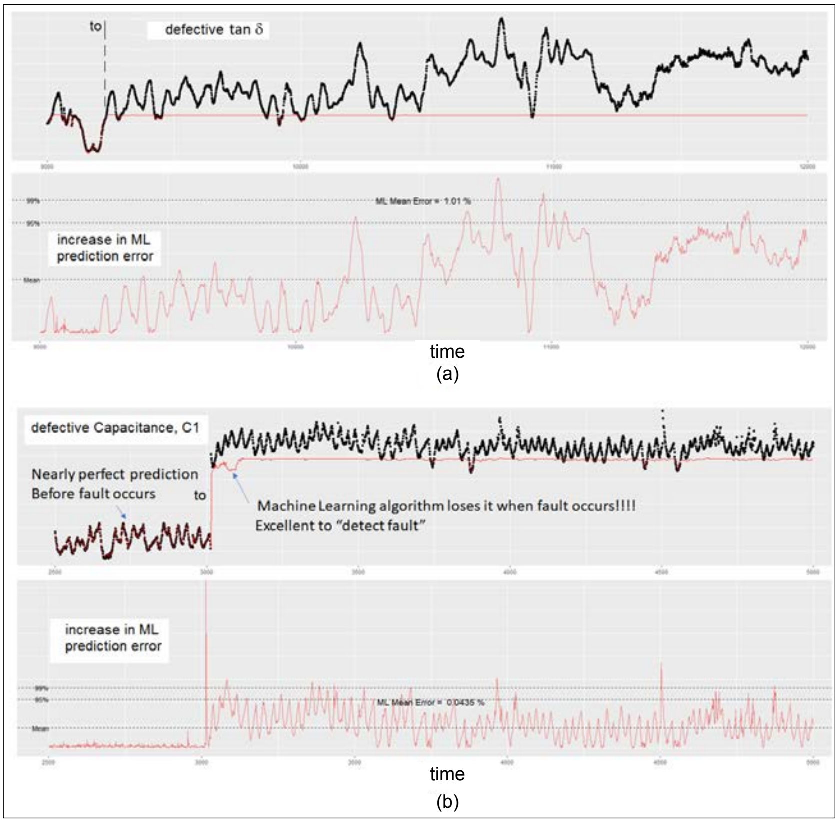


Figure 5. Illustration of the technique employed to detect (a) significant change in bushing tan δ and (b) significant change in bushing capacitance C1. The sudden error increase in the prediction of the model is used to detect the fault.

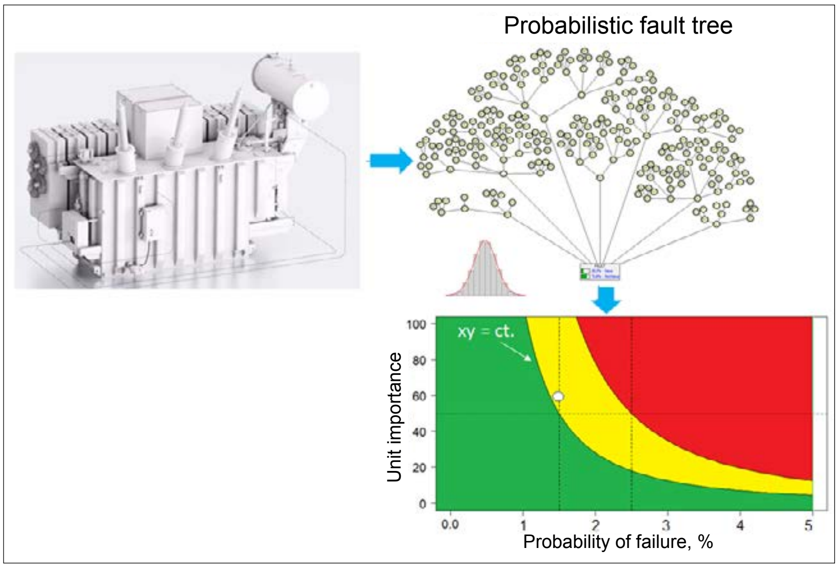


Figure 6. A probabilistic fault tree is used to "mimic" transformer condition based on the dynamic assessment of evidence coming from online sensors, laboratory analysis, site tests, electrical tests, etc. The probabilistic fault tree propagates belief in all directions based on the mathematical formulation of Bayes conditional probabilities and also based on the prior fault statistics provided by individual users.

of our algorithms logic, we are proposing an approach that combines our vast transformer domain knowledge with an open-source alternative based on a Bayesian network to help our customers determine the probability of failure of a transformer, as seen in Fig. 6. This solution is based on information from public standards and statistical analysis that can be discussed freely with customers and users alike.

Our Bayesian network is generated from a set of transformer components. An initial reference was taken from the fault tree described in IEEE PES Standard C57.14-2017 [3] – Guide for Evaluation and Reconditioning of Liquid Immersed Power Transformers. Each node gets assigned a customer-defined weight, and using the monitoring information from each specific transformer (online sensors, historical data, lab tests, etc.) we can generate a Bayesian network to calculate the transformer's probability of failure in real-time. Weights are tailored to each user given that the failure rate of each electrical component depends on many user-specific factors such as frequency of data sampling and tests performed, geographical location, capacity and usage, etc.

Some of the main advantages of this approach include:

- Comprehensive: includes a large number of failure modes.
- Dynamic: the probability of failure is updated based on new information.
- Flexible: customers can vary the weight and number of nodes.
- Generic: Applies to any power grid asset.

Finally, one of the main advantages of using Bayesian networks is that their belief propagation works in all directions. For our approach, this implies that we can use the network in two ways: To calculate the probability of failure of the transformer based on information in-coming from its components (belief propagating top down), and to identify the cause of failure of a component based on the weights of its nodes (belief propagating bottom-up).

5. Machine learning algorithm trained by the combined Duval pentagons

In order to train a machine learning model to classify DGA data (five combustible gases H_2, CH_4, C_2H_2, C_2H_4, C_2H_6) using the Combined Duval Pentagons [4] one would need a large corpus of dissolved gas analysis (DGA) readings from real transformers with labels for the supervised training process. The label in this training process represents the issue the transformer suffered associated to each training DGA reading. Using synthetic transformer data for five combustible gases, we trained a model as seen in Fig. 7. The generation of many thousands of centroids inside the Pentagon implies that any real DGA reading would lead to a centroid already generated through the artificial process.

The machine learning algorithm was trained using the artificial centroids and then validated using 552 real DGA cases from real transformers. The classification of these real cases is described in Fig. 8b where the instances in black boxes were correctly classified (547 out of 552), and the ones in red were incorrectly classified (5 out of 552).

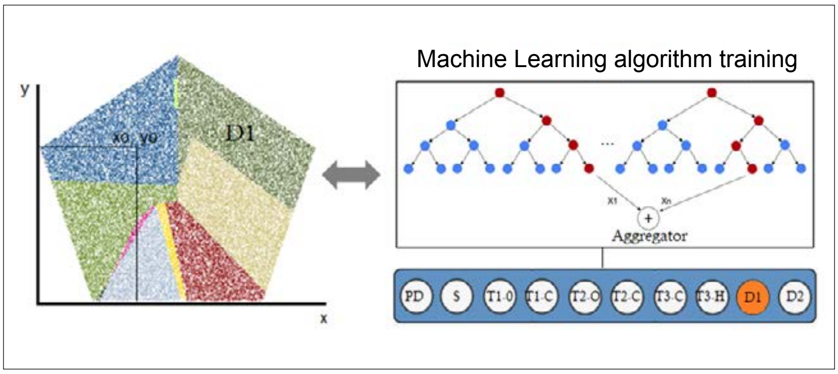


Figure 7. Illustration of machine learning algorithm trained with artificially generated DGA centroids, covering the whole area of the Combined Pentagon (all 10 fault regions). The illustration contains 50,000 artificially generated centroids.

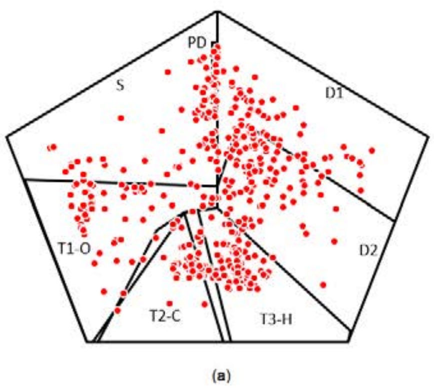
To train a machine learning model to classify DGA data using the Combined Duval Pentagons, one would need a large corpus of DGA readings from real transformers with labels for the supervised training process

6. AI-driven super-minds for industrial electrical equipment communities [5]

Transformer users typically prefer not to share their transformer data due to the risks associated with it. The lack of data sharing then leads to a lack of validation of comprehensive methodologies to monitor and diagnose transformers and a lack of visibility of how their transformers perform in comparison to other users.

In this project, we propose a smart infrastructure that combines a subject-matter-expert-guided knowledge pool, an anonymized transformer user database (similarly applicable to any other asset) aggregated by a set of machine learning algorithms and statistical tools to learn from both ends (knowledge pool and users' data). This infrastructure identifies the most likely scenario involving a given asset type (e.g., large transmission transformers or generator step-up transformers) and returns information to the anonymized user, including statistical support information, reasoning, and recommendations for action. A simplified illustration of the whole process is presented in Fig. 9. Users would join by subscription, and their data is only visible to the data wrangling tool. Users of the same type of asset would have access to statistical analysis and machine learning algorithms output and recommendations for similar issues or related matters.

Note: Hitachi Energy was granted a patent on this technology [5].



Confusion Matrix

Machine Learning Classification →; Duval Pentagons Diagnostics ↓

	D1	D2	PD	S	T1-C	T1-O	T2-C	T2-O	T3-C	T3-H
D1	56	1	0	0	0	0	0	0	0	0
D2	1	100	0	0	0	0	0	0	0	0
PD	0	0	15	0	0	0	0	0	0	0
S	0	0	0	76	0	0	0	0	0	0
T1-C	0	0	0	0	5	0	0	0	0	0
T1-O	0	0	0	0	0	64	0	0	0	0
T2-C	0	0	0	0	0	0	78	0	1	0
T2-O	0	0	0	0	0	0	0	0	0	0
T3-C	0	0	0	0	0	0	1	0	29	1
T3-H	0	0	0	0	0	0	0	0	0	124

(b)

Figure 8. Machine learning algorithm test phase (a) test dataset plotted on the Combined Pentagon with 552 DGA samples from real transformers (b) Confusion matrix providing the output of the machine learning classification for the 552 test cases (not seen during training of the ML algorithm); the red rectangles show misclassification (total of 5 cases) whereas the black rectangles in the diagonal show the correct classifications per fault region, with a resulting prediction accuracy of 99% plus. The horizontal lines of the confusion matrix show the classification in the original Combined Pentagon, whereas the columns show the output of the ML algorithm.

We propose a smart infrastructure that combines a subject-matter-expert-guided knowledge pool, an anonymized transformer user database aggregated by a set of machine learning algorithms and statistical tools

Conclusion

In this article, we have described a series of artificial intelligence approaches currently being developed in Hitachi Energy. Our wind power forecast approach deals with the problem of uncertainty in upcoming power demand by proposing a machine learning model to predict power demand to improve the calculation of loadability and cooling / hotspot calculations. Our Bushing Tan δ and Capacitance Fault Detection solution uses the error of a model to detect problems with Tan δ and capacitance. Our Probabilistic Fault Tree describes an open-source approach that uses Bayesian networks to find the probability of failure of a specific transformer. Finally, we describe two publications made by our team regarding the use of synthetic data created using the Duval Pentagons to generate a model that diagnoses transformer faults; and a patent regarding the creation of an infrastructure that uses blockchain to anonymize users and provide them with information about their transformer fleet using artificial intelligence. These exciting developments show the commitment of Hitachi Energy to disruptive innovation and our assurance to provide the highest quality products to the needs of our customers.

Bibliography

[1] X. Wang, P. Guo, X. Huang, *A review of wind power forecasting models*, Energy Procedia, Volume 12, 2011, p. 770–778, https://doi.org/10.1016/j.egypro.2011.10.103

[2] Sh. Hanifi et al., *A critical review of wind power forecasting methods—past, present and future*, Energies 2020, 13(15), 3764; https://doi.org/10.3390/en13153764

[3] C57.140-2017 – *IEEE Guide for Evaluation and Reconditioning of Liquid Immersed Power Transformers*, IEEE PES Transformers C57.14-2017

[4] L. Cheim, M. Duval, S. Haider, *Combined duval pentagons – a simplified approach*, MDPI Energies, 2020, 13, 2859; doi:10.3390/en13112859

[5] L. Cheim, L. Tehini, S. Pearce, *Artificial intelligence / machine learning driven assessment system for a community of electrical equipment users*, US Patent #1,262,739B2, March 2022

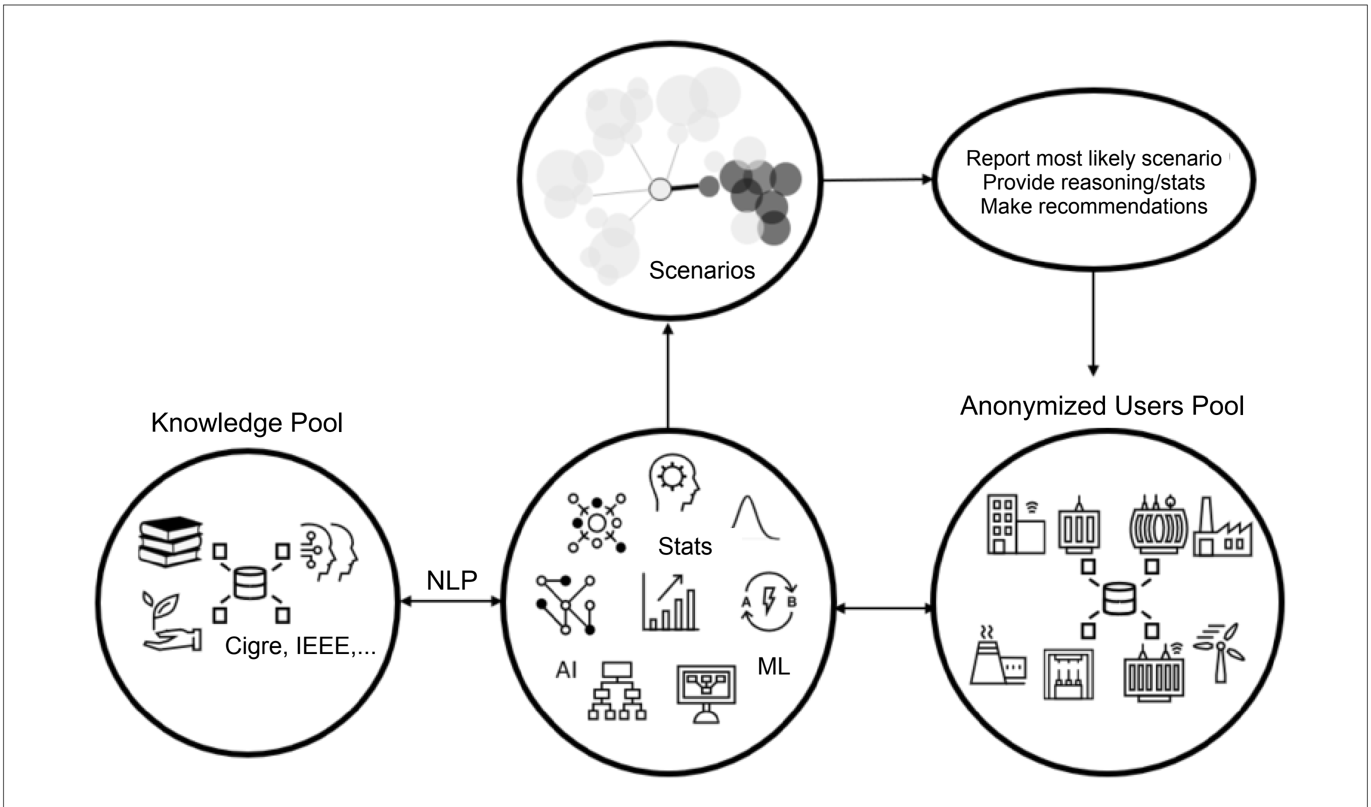


Figure 9. The "Super Mind" infrastructure. The blocks on the left illustrate the build-up of a knowledge pool guided by subject matter experts. The block on the right shows multiple transformer users that provide anonymized data to the data pool and who are subscribers to the service. The lower block in the middle illustrates the "data wrangling" phase information from the knowledge pool extracted through Natural Language Processing using multiple statistical tools and Machine Learning Algorithms that will ultimately build multiple scenarios of possible events/occurrences and inform the specific user of the most likely scenario, given the data and evidence. The solution will also contain a machine-built expert system that will make recommendations for actions.

Recent exciting developments show the commitment of Hitachi Energy to disruptive innovation and our assurance to provide the highest quality products to the needs of our customers

PEER REVIEWED

Authors

Dr. Luiz Cheim has been with Hitachi Energy as a Sr. Principal R&D Engineer for a number of years, having over 30 years' of experience in the power transformers industry. His major activities as part of a global R&D team are in the development of transformers condition assessment and performance models and algorithms, as well as the development of new sensors and state-of-the-art monitoring technologies. Dr. Cheim is the proponent of the Hitachi Energy Transformer Inspection Robot (TXploreTM). In August 2018 Luiz was granted the Best Paper Award by the Cigre organization in Paris, Study Committee A2/PS2 on the use of AI/Machine Learning techniques in support of transformer diagnostics. Luiz is in the editorial board of the new Cigre Green Book on Transformer Life Management, a member of the Cigre WG D2.52 AI Application and Technology in the Power Industry, and responsible for Chapter 5 – Applicability and Maturity of AI Technologies. Dr. Cheim has filed over 20 patents in the last 10 years alone, including the most recently granted by the US Patent Office on AI Superminds (reference 5). Luiz is also the Guest Editor of the Transformer Magazine Special Edition on AI and Machine Learning, November 2022 issue.

Dr. Mauricio Soto is the Artificial Intelligence Program Manager at Hitachi Energy's transformers business unit. He obtained his Ph.D. and M.Sc. from the School of Computer Science at Carnegie Mellon University. His work involves creating AI models for the monitoring, diagnosis, and manufacturing of transformers and driving artificial intelligence innovation within Hitachi Energy. Dr. Soto was recognized as "Individual with Exceptional Ability in the National Interest" by the U.S. Government, has served as a program committee member for four international conferences, published 10+ research papers, and currently serves as author and leader of IEEE/PES Transformers.

Tucker Reed is the Global Product Manager for Hitachi Energy's transformer diagnostics software. He is also Hitachi Energy's Digital Services Manager within the US market. His work is to provide value by driving the transition of transformer digitalization through the implementation of asset management solutions.