KnE Social Sciences

The 1st International Conference on Creative Design, Business and Society (1st ICCDBS) 2023

KnE Publishing

Research Article

Detecting Geothermal Operational Asset Anomalies Using the Locality-Sensitive Hashing (LSH) Algorithm

Muhammad Vito Hamza*, Fransisco T.P. Simamora, Efrata Pratenta Meliala, R. Fuad Satrio Ajie, Hanifah Nur Azizah, Fajar Khamim Mustofa, and Adi Suparyanto

PT Geo Dipa Energi (Persero), Indonesia

Abstract.
Geothermal power plants are crucial for sustainable energy generation, necessitating the reliable maintenance of their operating assets. This research proposes an approach for asset maintenance through anomaly detection using the Locality-Sensitive Hashing (LSH) algorithm. The accuracy and coverage of traditional anomaly detection approaches in geothermal power plants may be constrained by sensor monitoring systems. The LSH algorithm is used to improve detection skills and get a full understanding of the state of important assets. The proposed method utilizes historical sensor data collected during geothermal power plant operations. This data is transformed into hash codes using LSH, effectively capturing similarities between various operational states and asset conditions. By comparing the hash codes of the current operational state with a library of precomputed hash codes representing typical operating conditions, the LSH algorithm can identify deviations indicating potential irregularities. This facilitates early detection of anomalies, even in large-scale databases, enabling prompt maintenance interventions. The application of anomaly detection using the LSH algorithm provides benefits such as improved asset maintenance planning, reduced downtime, and increased operational safety. By leveraging data-driven analysis and the effectiveness of LSH, geothermal operators can detect faults early, enabling prompt interventions and optimizing reliability and efficiency. By leveraging historical sensor data and the efficient similarity approximation capabilities of LSH, the proposed approach enables early diagnosis of problems, improving maintenance planning and optimizing geothermal operations.

Keywords: geothermal assets, locality-sensitive hashing, asset condition, fault detection, reliability

Corresponding Author: Muhammad Vito Hamza; email: vito.hamza@geodipa.co.id

Published: 19 November 2024

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the 1st ICCDBS Conference Committee.

1. Introduction

Geothermal is one of the energy sources that are relied upon to achieve Indonesia's 23% renewable energy mix by 2025 [1]. The government plans to develop 1445 MW of Geothermal Power Plant from 2021 - 2025 [2]. Yet Geothermal Projects were constantly facing challenges both in exploration drilling [3] and social acceptance [4]. Still,

How to cite this article: Muhammad Vito Hamza*, Fransisco T.P. Simamora, Efrata Pratenta Meliala, R. Fuad Satrio Ajie, Hanifah Nur Azizah, Fajar Khamim Mustofa, and Adi Suparyanto, (2024), "Detecting Geothermal Operational Asset Anomalies Using the Locality-Sensitive Hashing (LSH) Algorithm" in *The 1st International Conference on Creative Design, Business and Society (1st ICCDBS) 2023*, KnE Social Sciences, pages 239–254. DOI 10.18502/kss.v9i32.17439

Geothermal Power Plants are preferable to PLN due to their base load 24-hour dispatch nature which is expected to generate stable electricity [5]. Therefore, their consistent functioning is key to ensuring a steady and predictable energy supply. Physical Asset Management is one of the frameworks that allow physical assets such as Geothermal Plant to be optimized. One of the key building blocks of Asset Management is Lifecycle Delivery where when an asset is acquired or commercially operated, it will be managed through operation and maintenance [6]. It is known that Power Plant operating capacity is decreased by almost 35% due to downtime and hidden losses [7]. Therefore, anomaly detection is an important part of asset management in power plants because it allows possible problems to be identified early on before they cause substantial losses or downtime.

There are several methods of anomaly detection. Through condition-based maintenance (CBM), the most common technologies are oil analysis, vibration analysis, and infrared thermography [9]. However, CBMs are periodic and may not cover early anomalies that could potentially be identified by using real-time data.

The role of Machine Learning (ML) comes with ingesting data on power plant operating conditions and has been explored and implemented intensively by power plant entities [10]. Several algorithms are used for energy efficiency and forecasting maintenance schedules. For anomaly detection, recognizing the pattern of operating parameters is the baseline to gain insight and decide which deviation is considered an anomaly or not. A conventional machine learning model such as clustering usually relies on data patterns between input and output to obtain pattern recognition. However, when presented with high-dimensional fault data with unclear output and highly interdependent input, it would compensate for interpretation speed and accuracy [11].

Locality-sensitive hashing (LSH) is a potential algorithm optimized for anomaly detection in geothermal power plants, as it can effectively capture similarities between various operational states and asset conditions. LSH is a method that compresses high-dimensional data into one-dimensional data to detect residuals for the whole dataset [12]. This allows the system owner to look at single-value residuals to detect a failure in the system or equipment model while gaining insight into the biggest contributor. Resulting in faster anomaly detection. This paper will present the use case of LSH in anomaly detection of Patuha Geothermal Power Plants which leads to forced outage events.

2. Literature Review

There are several studies and implementations regarding anomaly detection with machine learning (ML). Both are using supervised and unsupervised learning [13]. Supervised methods such as random forest [14] are highlighted would be more difficult without proper data ingestion. Conventionally, unsupervised learning relied upon to detect patterns in data is clustering. Yet clustering faces predictive accuracy challenges in high-dimensional data [15]. While other clustering methods such as OPTICS are facing accuracy challenges during anomaly detection [16].

Using a supervised neural network, a thermal power plant successfully detected anomaly readings with a slight increase in feedwater flow and pressure decrease [17]. Although the result is satisfying in terms of accurately detecting slight changes in data, the paper didn't provide enough information on the result interpretation. This is important since in power plant system operation, there are a lot of parameters related to each other. For example, Table 1 shows that a turbine-generator efficiency formula involves six dynamic transmitter data (parameter). Figure 1 explains the relationship between the efficiency losses and its building parameter.

TABLE 1: Turbine-Generator Efficiency Parameter (DiPippo, 2016) [18].

No	Value	Parameter	Data Source
A	Steam Turbine Power (kW)	Mass Flowrate x (h4 – h5)	Calculated
		Inlet Turbine Mass Flowrate (T/h)	Flow Transmitter
		Inlet Enthalpy (h4):	**Calculated from:**
		Inlet Pressure (Barg)	Pressure Transmitter
		Inlet Temperature (°C)	Temperature Transmitter
		Outlet Enthalpy (h5):	**Calculated from:**
		Outlet Pressure (Barg)	Pressure Transmitter
		Outlet Temperature (°C)	Temperature Transmitter
B	Gross Generator Power (kW)	Active Power (kW)	kWh Meter
D	Generator Efficiency	$\frac{B}{A}$	Calculated

Neural network, random forest, and clustering are black-box model which does not use any particular prior knowledge of the data relationship to provide a prediction [19]. Hence, a black-box model of turbine-generator efficiency would be able to find anomalies such as efficiency losses. However, the model won't be able to explain

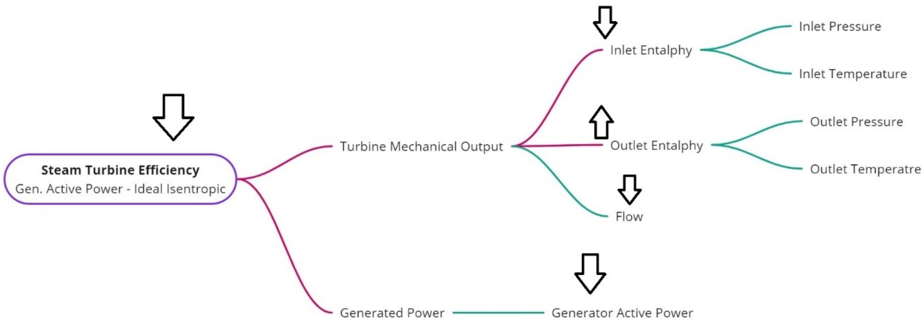

Figure 1: Potential parameters to address during efficiency anomaly.

where the losses came from to the domain expert due to the mathematical complexities involved [20].

Due to this challenge, the machine learning models no longer emphasized just based on its predictive accuracy, but also descriptive accuracy. Hence as described in Figure 2, an interpretable model or white-box model will include a post hoc analysis such as histograms and scatter plots to accurately describe the data relationship [21].

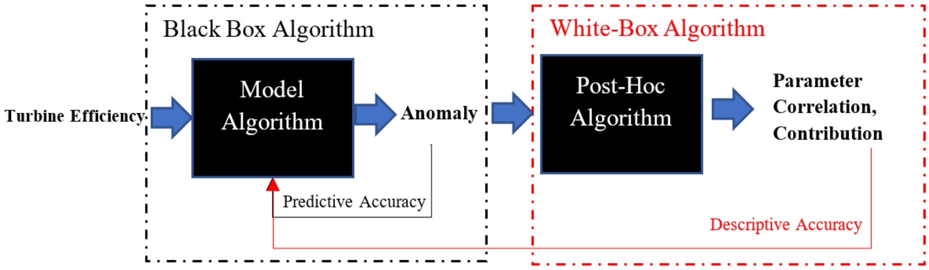

Figure 2: Interpretable (White-Box) Machine Learning includes Post-Hoc Algorithm [21].

There are various advantages to using white-box algorithms. First, they are more accountable since white-box algorithms are transparent in how they process data input by giving in weights and relationships. Second, white box algorithms are more justifiable because the process can be adjusted to match the demands of a specific domain challenge. Third, because white-box algorithms also provide correlation and contribution between their input, it is easier to interpret why the algorithm generates the predictions [22]. Interpretability is the main component of explainable AI which embody five different perspective: the right to explanation, trust, model improvement, and discovering novel concepts [23].

Locality Sensitive Hashing (LSH) is a model that works well to interpret high-dimensional data [24]. The approach is based on the notion that comparable objects are likely to be assigned to the same bucket by splitting the high-dimensional space into several smaller buckets. The technique accomplishes this by using a hash function to convert each object in the high-dimensional space into a bucket. Since comparable things are likely to have the same hash value, the hash function is created in this way.

The algorithm can run a similarity search after the items have been placed in buckets by simply comparing the things in the same bucket. Compared to comparing every object in the high-dimensional space to every other object, this is substantially faster. The authors of the research demonstrate that even with limited data, their LSH algorithm can perform well for similarity search in high-dimensional domains [24]. The capability of LSH for anomaly detection both in synthetic and real datasets is comparable to other available methods [25].

2.1. Hypotheses

In this paper, the LSH algorithm will be tested to detect both prediction accuracy and descriptive accuracy. The process of detecting the anomaly of an equipment or system is shown by overall model residuals (OMR) that of clustering. Then it would create a feature selection that ranks sensor contribution and fault diagnosis of the OMR. In short, the LSH algorithm must be able to: predict changes in OMR and predict contributors to the changes.

3. Methodology Research

The model development method used is based on anomaly detection in an integrated parameter system [25] which involves historical data collection, data pre-processing (configuration), statistical learning (training), model evaluation (testing), and model fine-tuning (see Figure 3).

Historical data collection: The first step is to gather a dataset that is relevant to the problem that the model is trying to solve. The dataset should be large enough and representative of the real-world data that the model will be used. The datasets also be used as a baseline of normal operating conditions. This means there must be incorporation of experts that justify the data.

Data pre-processing (configuration): Once the dataset has been gathered, the next step is to configure the ML model by doing data pre-processing. This involves selecting the appropriate model architecture and hyperparameters. The model architecture is the overall structure of the model, while the hyperparameters are the tuning parameters that control the learning process. This would also involve certain data cleansing such as filtering, trimming, and clearing outliers.

Statistical learning (training): Once the model has been configured, it needs to be trained on the dataset. This involves feeding the dataset to the model and allowing it to learn the patterns in the data.

Model evaluation (testing): Once the model has been trained, it needs to be evaluated on a held-out test set. This helps to assess the performance of the model on unseen data and to identify any potential areas for improvement.

Model fine tuning (re-training): By evaluation, the model should be re-cleansed if it fails in anomaly detection. If the model failed to describe failure or sensor contribution accordingly, the model data inputs would need to be re-weighted.

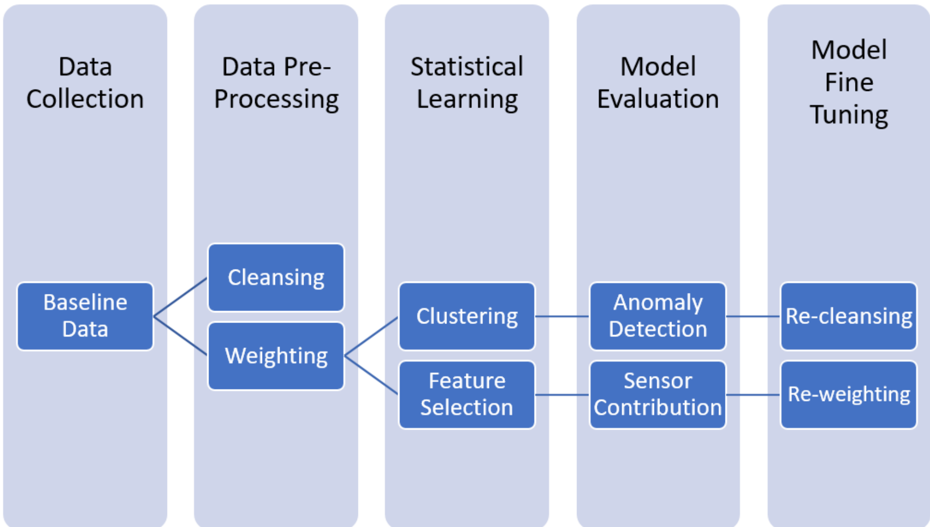

Figure 3: The process of Geothermal Power Plant Model Development.

3.1. Data Gathering

The quality and quantity of the data used to train an ML model have a significant impact on its performance. Therefore, it is important to gather a dataset that is relevant, large enough, and representative of the real-world data that the model will be used. When gathering data, it is important to consider the following factors:

1. Relevance: The data should be relevant to the problem that the model is trying to solve. For example, if the model is trying to predict the anomaly of an electric motor, then the data present should be the parameters of the motor such as winding temperatures, electrical current, and terminal voltage.

2. Quantity: The dataset should be large enough to allow the model to learn the patterns in the data. The minimum dataset size depends on the complexity of the problem and the type of ML model being used.

3. Representativeness: The dataset should be representative of the real-world data that the model will be used on. This means that the dataset should include a variety of different examples, and it should not be biased towards any subgroup.

In this particular paper, data from the power plant condensing system will be used with its dependent parameters. The data will be limited by the number of sensors gathered for the unit. Outliers will also be filtered out so the model can fit with the normal operation of the system. The model involves several pieces of equipment each of them, as seen as Table 2:

Table 2: Equipment and Parameter of Circulating Water System.

No	Equipment	Parameter	Data Source	Data Type
1	Main Condenser	Main Condenser Level (mm)	Level Transmitter	Dynamic
		Main Condenser Pressure (Barg)	Pressure Transmitter	Dynamic
		Main Condenser Temperature (°C)	Temperature Transmitter	Dynamic
2	Condenser Spray Valve	Valve Position (%)	Position Indicator	Dynamic
		Valve Command (%)	Logic Diagram	Dynamic
3	Hot Well Pump	Discharge Pressure (Barg)	Pressure Transmitter	Dynamic
		Motor Current (A)	Amperemeter	Dynamic
		LCV Position (%)	Position Indicator	Dynamic

3.2. Model Configuration

Once the dataset has been gathered, the next step is to configure the ML model. This involves selecting the appropriate model architecture and hyperparameters. The model architecture is the overall structure of the model. It determines how the model will learn from the data and how it will make predictions. Since LSH leveraging the white box model, the model would be configured with pre-knowledge of hot well pump failure mode which will be mapped with its sensor causal relationship such as increase, decrease, and constant such as Table 3. The model would also involve weighting sensor contribution to certain faults.

TABLE 3: Fault Diagnostic relationship input.

Name	COOLING WATER FLOW TO CONDENSER	CONDENSER VAC	CONDENSER TEMPERATURE	CONDENSER SPRAY VALVE POSITION	CONDENSER LEVEL	HOTWELL PUMP B DISCHARGE CV POS	HOTWELL PUMP A DISCHARGE CV POS
Spray Valve Fail to Open	↓	↑	↑	↓	↓		
Temperature Transmitter Fail to Function			↑↓				
Vessel High Temperature (PROD)			↑				
Condenser High Level	↓			↓	↑	↓	↓

The data that will be used for training are from 13 - 31 August 2023 with intervals of 5 minutes. Since training data are the baseline or normal reference of operating conditions, any anomaly present during training is cleansed so it can be verified during model evaluation. As shown in Figure 4, the anomaly present on the condenser spray valve is cleansed. During training, known operation intervention will create a bias toward the model sensor contribution. Due to this, it is important to choose a date interval without any intervention.

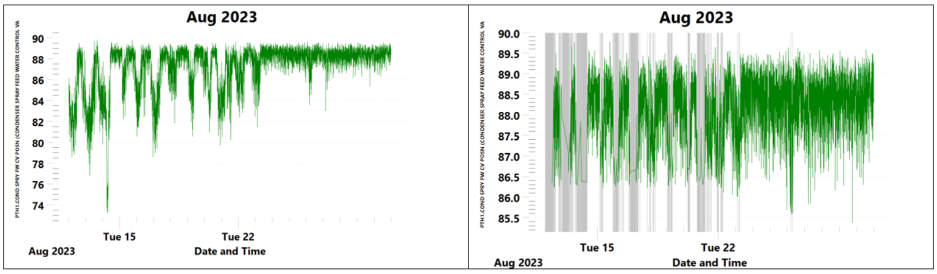

Figure 4: Data without cleansing (left), data used for model training (right).

3.3. Model Training

Once the model has been configured, it needs to be trained on the dataset. This involves feeding the dataset to the model and allowing it to learn the patterns in the data. The pattern that the model learns is the “normal” operating pattern so the model will fit prediction and residuals with normal reference as the baseline. The data is then clustered to get a working model that will be used to test data.

3.4. Model Evaluation

Once the model has been trained, it needs to be evaluated on a held-out test set. This helps to assess the performance of the model on unseen data and to identify any potential areas for improvement. In this test, the model will be presented with actual data from 13 August – 01 September 2023 (without cleansing) to detect anomalies in the condenser spray valve (see Figures 5 and 6). This component was known to be the cause of plant forced outage on September 1, 2023.

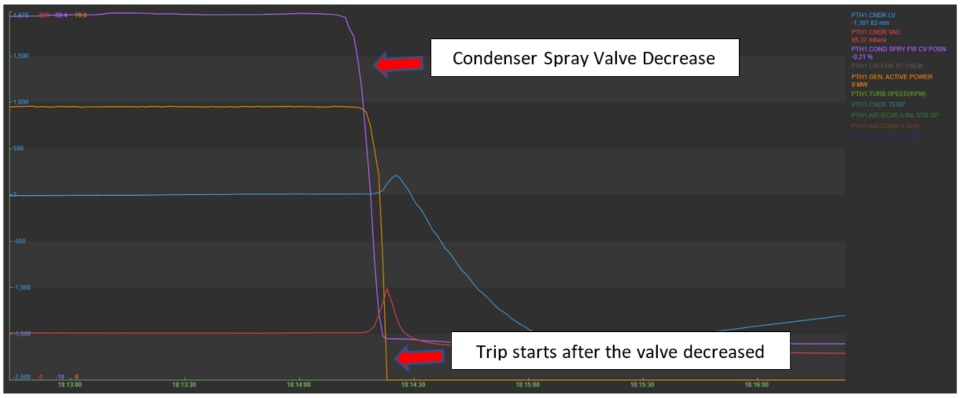

Figure 5: Actual trend of spray valve decrease.

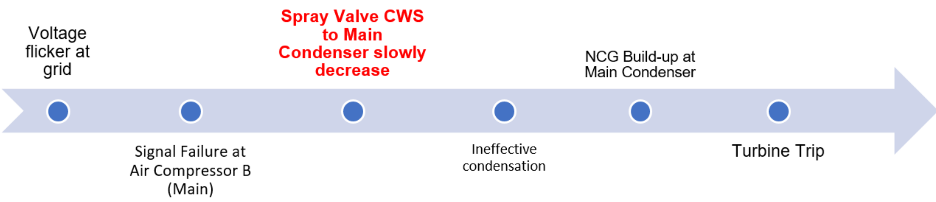

Figure 6: Simple 5 Whys analysis of Trip event during September 1, 2023.

First, the model needs to detect earlier by using overall model residuals (OMR) which are based on the regular RMSE formula:

$$OMR\ (\%) = \sqrt{\sum_{i=1}^{n} \frac{(y'_i - y_i)}{n}} (1)$$

y'_i = Predicted value

y_i = Actual value

n = Number of observations (parameters)

Second, the model needs to be able to predict the sensor contributions of the OMR. This would involve using a statistical histogram or Pareto analysis. To be descriptive accurate, the model would have to predict the highest sensor contributor as the Condenser Spray Valve. And fault diagnostic would predict the highest fault to be the condenser spray valve fail to open.

3.5. Model Fine Tuning

If the model is not performing well on the evaluation set, it may need to be re-tuned. This involves adjusting the model architecture or hyperparameters and then training the model again. The process of re-tuning a model can be iterative, and it may be necessary to retune the model several times before it achieves the desired performance. However, if the model is showing the desired result, the model is not necessarily fine-tuned. The practice of fine tuning is usually done after maintenance of equipment which makes the equipment run in different baselines.

4. Results and Discussion

The model of the condensing system resulted in the overall residual model (OMR) which summarizes all the observations in each sensor into a single feature. The feature then generated an area of predicted value in which tolerable around 5% deviation to the OMR (see Figure 7).

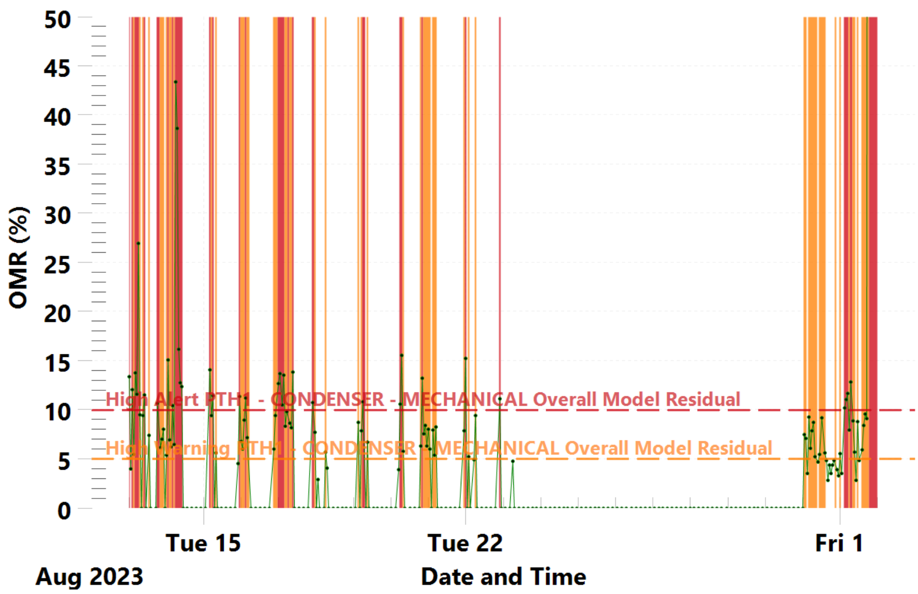

Figure 7: OMR profile of test data (13 August – 01 September 2023).

Figure 8 shows OMR above 5% during 13 – 23 August 2023. This means the model has an accurate prediction of anomaly detection.

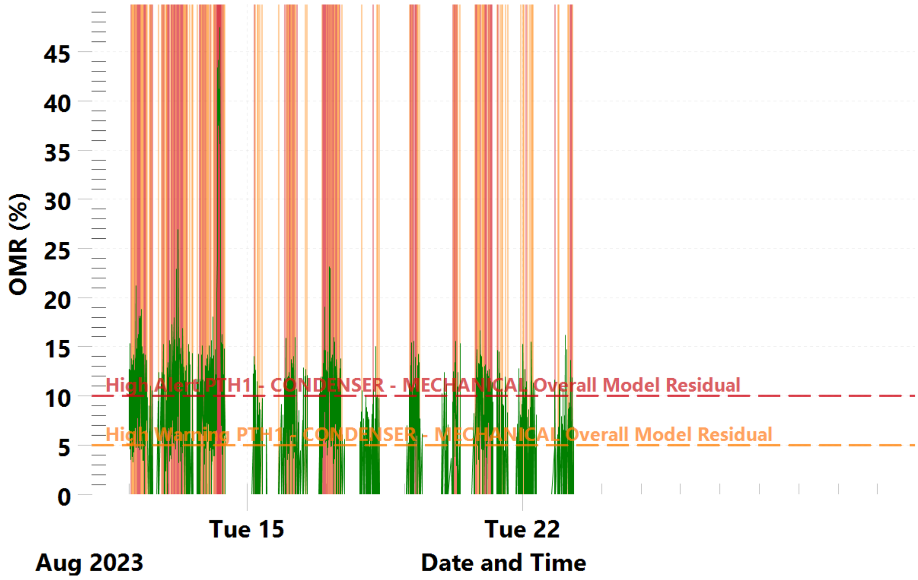

Figure 8: Shows OMR above 5% at 13- 23 August 2023.

Figure 9 shows that during the period of 13 – 23 August 2023, the highest sensor contributor is the condenser spray valve at 49.6%. This would validate the model's descriptive accuracy. Following with Figure 10 detects 47.8% of fault potentially from Spray Valve Fail to Open.

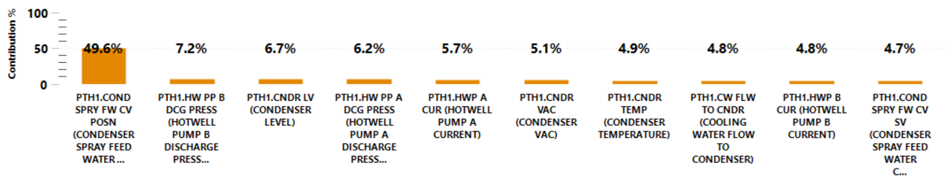

Figure 9: Sensor Contribution.

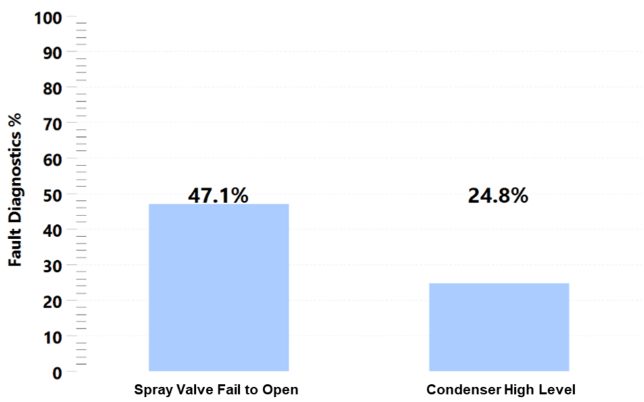

Figure 10: Predicted fault diagnosis at 13 – 23 August 2023.

Figure 11 shows the OMR increase before the forced outage event on 01 September. In this anomaly finding, the signal contribution is dispersed evenly with the highest contributor still coming from the condenser spray valve. While the predicted fault leaned towards the condenser's high level. The close gap between the contributions means the analysis is still descriptive accurate because it shows accountable results that the sensor deviation and fault diagnosis are inconclusive. Further analysis by the domain expert is indeed needed but fine-tuning is not necessary. Figure 12 shows signal contribution near the forced outage event, and Figure 13 depicts predicted fault diagnosis near trip event.

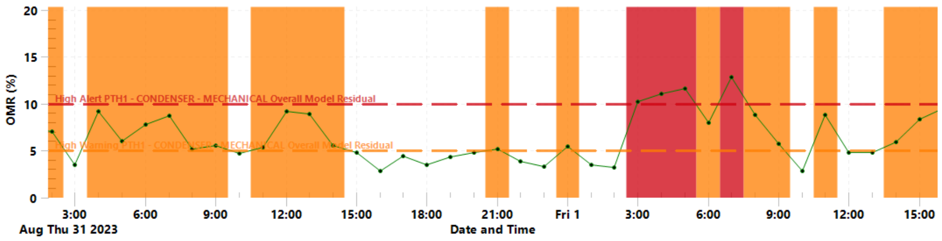

Figure 11: Shows OMR above 5% at 31 August – 01 September (before outage).

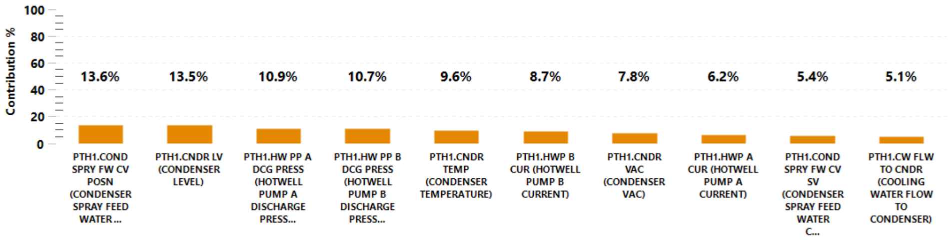

Figure 12: Signal contribution near the forced outage event.

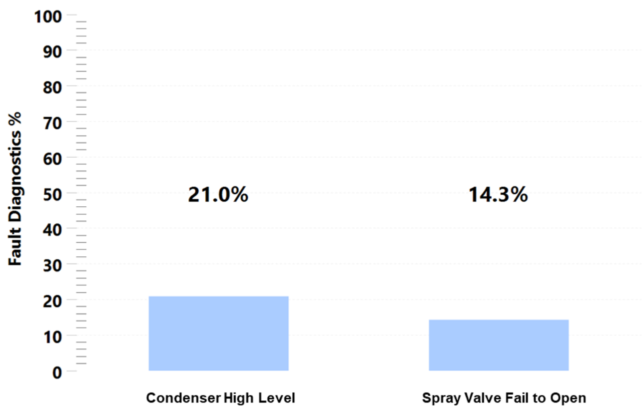

Figure 13: Predicted fault diagnosis near trip event.

5. Conclusion

The development of a white-box model aims to show an accountable, justifiable, and interpretable result. That is why a white-box model has to be directed by subject matter experts (SMEs) and domain information to be deployed. One of the white-box methods that are implemented in this paper is locality-sensitive hashing (LSH). The LSH model in Patuha Geothermal Power Plant can successfully detect anomalies earlier before the plant shutdown. Not only predicting the anomalies, but the model also could describe the parameter that most contributes to the anomaly and diagnose anomaly to a certain fault condition. This contribution and diagnosis are developed from the data pre-processing process which involves weighting and creating relationships between data before deploying the model. The key to accurate prediction is actually in the weighting and relationship of the data which is done by subject matter experts (SMEs) from the domain.

As a white-box model, the model should have to be evaluated and fine-tuned by the SMEs frequently. The main evaluation parameter monitored in the LSH model is overall model residuals (OMR) and its sensor contribution (Pareto analysis). The usage of OMR is helpful to give SMEs a preliminary insight into what's happening in the plant/equipment. The Pareto analysis then would help to find the sensor contribution of the anomaly. However, SMEs also need to verify what is happening in the field. As the interpreter of machine learning, the responsibility lies on its reader to evaluate the bias.

For further research, evaluation and comparison analysis are needed to benchmark model results with algorithms such as independent component analysis (ICA), clustering, and other black-box methods. Furthermore, this model would be potentially expanded into the sub-surface of geothermal fields. Like other renewable energies, geothermal would depends heavily on its resources which in practice, is usually done in engineering technologies in the form of well simulations that result in well intervention. As the capacity assurance of well, the early detection would help detect faults in the subsurface such as casing leaks, brine carry-over, and equipment fishing.

Acknowledgments

Acknowledgment is given to Patuha Geothermal Field for their openness in discussion and data gathering to analyze the model. Acknowledgement also be given to the digital transformation team in the head office from the Operation, Maintenance & Technical

Services Division and Information Technology (IT) Department. Also dedicated this paper to the hard work that Tiara Vibrasindo Pratama (TVP) did during the model development phase.

References

[1] IESR. Indonesia energy transition outlook 2023: Tracking progress of energy transition in indonesia: Pursuing energy security in the time of transition. Institute for Essential Services Reform (IESR), Jakarta. 2022

[2] ESDM. *National Electricity Supply Business Plan* (RUPTL) PT PLN Persero. Ministry of Energy and Mineral Resources (ESDM), Jakarta; 2021.

[3] Purba D, Chandra VR, Fadhillah FR, et al. Drilling infrastructure construction challenges in geothermal exploration project in Eastern Indonesia. In: Proceedings World Geothermal Congress 2020+1; International Geothermal Association (IGA), Iceland. 2021.

[4] Fadhillah FR, Asyari MR, Bagaskara A, et al. Challenges in Getting Public Acceptance on Geothermal Project in Indonesia. PROCEEDINGS, 48^{th} Workshop on Geothermal Reservoir Engineering. Stanford University, California. 2023.

[5] Institute for Essential Services Reform (IESR). A roadmap for Indonesia's power sector: How renewable energy can power Java-Bali and Sumatra. Jakarta: IESR; 2019.

[6] ISO Technical Committee 251. ISO 55000:2014 - Asset management. International Organization for Standardization (ISO) 2014. Link:https://www.iso.org/obp/ui/#iso:std:iso:55000:ed-1:en

[7] Mitchell JS. Physical asset management Handbook Fourth Edition. Jakarta: MTS Indonesia; 2009.

[8] Gulati R. Maintenance and reliability best practices. Third edition. Industrial Press, Inc.; 2020.

[9] Bisset C, Venter PV, Coetzer R. A systematic literature review on machine learning applications at coal-fired thermal power plants for improved energy efficiency. Int J Sustain Energy. 2023;42(1):845–72.

[10] Fung PL, Zaidan MA, Timonen H, et al. Evaluation of white-box versus black-box machine learning models in estimating ambient black carbon concentration. J. Aerosol Sci. 2021;152:105694. DOI: 10.1016/j.jaerosci.2020.105694

[11] Shen H, Li T, Schweiger T, et al. Locality sensitive hashing. In [eds]. Encyclopedia of Computer Science and Technology. CRC Press; 2016, pp. 562-574. Available from: https://www.routledgehandbooks.com/doi/10.1081/E-ECST2-120052698

[12] Gangwani P, Joshi S, Upadhyay H, Lagos L. AI-Based anomaly detection on IoT data-driven thermal power plants for condition monitoring and preventive maintenance. In: Bhardwaj T, Upadhyay H, Sharma TK, Fernandes SL, editors. Artificial intelligence in cyber security: Theories and applications. Intelligent systems reference library. Volume 240. Cham: Springer; 2023. DOI: 10.1007/978-3-031-28581-3_8

[13] S. Joshi, H. Upadhyay, L. Lagos, N. S. Akkipeddi, and V. Guerra, "Machine learning approach for malware detection using random forest classifier on process list data structure," in ACM International Conference Proceeding Series, Association for Computing Machinery, Apr. 2018, pp. 98–102. doi: 10.1145/3206098.3206113.

[14] Xu D, Tian Y. A comprehensive survey of clustering algorithms. Ann. Data. Sci. 2015;2(2):165–93.

[15] Y. F. Wang, Y. Jiong, G. P. Su, and Y. R. Qian, "A new outlier detection method based on OPTICS," Sustainable Cities and Society, vol. 45, pp. 197–212, Feb. 2019, doi: 10.1016/j.scs.2018.11.031.

[16] Hajdarevic A, Dzananovic I, Banjanovic-Mehmedovic L, et al. Anomaly detection in thermal power plant using probabilistic neural network. In: 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO); Croatia. 2015. pp. 1118-1123.

[17] Ronald. DiPippo, Geothermal power plants: principles, applications, case studies, and environmental impact. Butterworth-Heinemann, imprint of Elsevier, 2016.

[18] L. Ljung, "Black-box models from input-output measurements," IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188), Budapest, Hungary, 2001, pp. 138-146 vol.1, doi: 10.1109/IMTC.2001.928802.

[19] Loyola-Gonzales O. Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view. IEEE Access. 2019;7:154096–113.

[20] Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Definitions, methods, and applications in interpretable machine learning. Proc Natl Acad Sci USA. 2019 Oct;116(44):22071–80.

[21] V. Hassija et al., "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," Cognitive Computation, vol. 16, no. 1. Springer, pp. 45–74, Jan. 01, 2024. doi: 10.1007/s12559-023-10179-8.

[22] Saeed W, Omlin C. Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. Knowl.-Based Syst. 2023;263:110273. DOI: 10.1016/j.knosys.2023.110273.

[23] Aristides Gionis, Piotr Indyk, and Rajeev Motwani. 1999. Similarity Search in High Dimensions via Hashing. In Proceedings of the 25th International Conference on Very Large Data Bases (VLDB '99). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 518–529. DOI: https://dl.acm.org/doi/10.5555/645925.671516

[24] Meira J, Eiras-Franco C, Bolón-Canedo V, et al. Fast anomaly detection with locality-sensitive hashing and hyperparameter autotuning. Inf. Sci. 2022;607:1245-1264. DOI: 10.1016/j.ins.2022.06.035.

[25] Mihnev A. Method and apparatus for detection of anomalies in integrated parameter systems. US PATENT. US 2015/0199224 AI. 2015. Available from: https://patents.google.com/patent/US20150199224A1/en