

Course Title: Research Methodology in Computer Science

Course Code: AI800

Lecturer: Dr. Mohammed Anbar

Assignment Report: Formulating a clear problem statement and research objectives

Student ID: 26200088

Student Name: Salem Ahmed Buhluli

Generalized ML for Power Plant Prediction and Fault Detection

Problem Statement

Machine learning is now widely used in power plants, both to predict plant output and emissions from sensor data (Tüfekci, 2014; Tan et al., 2019) and to detect equipment faults before they cause shutdowns (Khalid et al., 2021; Fahmi et al., 2024). The problem is that most published models are trained and tested on one plant only, or on a short period of data: Tüfekci (2014) covered one combined cycle plant at full load only, Tan et al. (2019) used seven days of data from one boiler, and Fahmi et al. (2024) one month of data from a single gas turbine. Labelled fault data are also rare and imbalanced (Stetco et al., 2019), and most studies use private datasets and different evaluation metrics, so the methods cannot be compared fairly (Carvalho et al., 2019). As a result, there is no way to be sure that a published model will work well on a new plant.

This problem is important because the faults these models target, such as boiler tube leakage and turbine failures, are among the main causes of power plant shutdowns, and every shutdown means lost generation and costly emergency repairs (Khalid et al., 2021), while accurate NO_x prediction is needed to control emissions when plants change load frequently (Tan et al., 2019). The existing solutions each have clear limits: the Bagging model works only at full load (Tüfekci, 2014), the sensor selection method was tested on one plant and two fault types (Khalid et al., 2021), the LSTM stops at prediction (Tan et al., 2019), and the TCN–Autoencoder reports no precision, recall, or false alarm rate (Fahmi et al., 2024).

Therefore, further research is needed to build a model that is tested on more than one unit and operating condition, that can work with few labelled faults, and that is evaluated with metrics that reflect real operation. This research aims to fill this gap.

Research Objectives

The main aim of this research is to develop and evaluate a machine learning framework for performance prediction and fault detection in power plants that stays reliable across different units and operating conditions. To achieve this aim, the following objectives are defined:

1. To identify the main limitations of existing machine learning models used in power plant operations and asset management, in terms of validation scope, data availability, and evaluation practice.
2. To collect and prepare operational data from more than one unit, covering steady and transient conditions, and to handle noise, missing values, and class imbalance in the data.
3. To propose and design a hybrid machine learning model that combines sensor selection with a temporal deep learning network and can work with few labelled fault examples.

4. To evaluate the proposed model under steady and transient conditions using MAE and RMSE for prediction quality, and precision, recall, and false alarm rate for fault detection quality.
5. To compare the proposed model with existing methods, such as bagged tree ensembles, LSTM, and TCN–Autoencoder, and to test it on a unit not used in training to verify its generalizability.

Summary of the Research Papers

Paper 1: Prediction of Full Load Electrical Power Output of a Combined Cycle Power Plant Using Machine Learning Methods

Tüfekci (2014) predicted the full-load output of a combined cycle power plant from four ambient variables, using 9,568 hourly records collected over six years. Among fifteen regression methods, Bagging with REP Tree gave the best result (MAE 2.818 MW, RMSE 3.787 MW), showing that ambient data alone can predict output accurately. The main limitation is that the model covers full-load operation of a single plant only, so nothing is known about part-load behaviour or other plants.

Paper 2: Machine Learning Methods for Wind Turbine Condition Monitoring: A Review

Stetco et al. (2019) reviewed machine learning studies on wind turbine condition monitoring. They found that most studies use SCADA or simulated data, that neural networks, SVMs, and decision trees are the most common models, and that scarce labelled fault data and class imbalance are the main obstacles in the field. As a review it adds no new empirical evidence, and its scope is limited to wind turbines.

Paper 3: A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance

Carvalho et al. (2019) presented a systematic review of machine learning in predictive maintenance of industrial equipment. Random Forest, neural networks, SVM, and k-means are the most used methods, and hybrid combinations of models give the most promising results. The review also shows that most studies rely on private datasets, which blocks replication and fair comparison between methods.

Paper 4: Optimal Sensor Selection for Equipment Fault Detection in a Thermal Power Plant

Khalid et al. (2021) proposed a correlation-based sensor selection method for fault detection in a real thermal power plant, targeting boiler waterwall tube leakage and turbine faults, which are leading causes of plant shutdowns. The reduced sensor set kept detection performance

while clearly improving computational efficiency. The evaluation covers one plant and two fault classes only, using classical models without deep temporal architectures.

Paper 5: Dynamic Modeling of NO_x Emission in a 660 MW Coal-Fired Boiler with Long Short-Term Memory

Tan et al. (2019) built an LSTM model to predict NO_x emissions of a real 660 MW coal-fired boiler, using 10,000 samples from seven days of operation. The model reached an RMSE of 12.2 mg/Nm³ on test data and clearly beat an SVM baseline under both steady and transient conditions. The data cover only one week of one boiler, and the work stops at prediction without closing the loop into combustion control.

Paper 6: Gas Turbine Fault Detection with a Temporal Convolutional Network–Autoencoder Model

Fahmi et al. (2024) proposed a TCN–Autoencoder with multi-head attention for anomaly detection in gas turbine vibration data, using about one million real samples from the Kirkuk power plant. The model outperformed GRU–Autoencoder, LSTM–Autoencoder, and VAE baselines and flagged expert-confirmed faults such as blade misalignment. The data cover one month of one turbine, and only reconstruction errors are reported, without precision, recall, or false alarm rate.

Critical Analysis

The main strength of these studies is their grounding in real operational data, from six years of plant records to high-frequency vibration measurements, together with a wide methodological range that spans classical ensembles, sensor selection, and modern deep temporal models. Their shared weaknesses, however, are serious: almost every model is validated on a single plant or a short time window, labelled faults are scarce and imbalanced, evaluation metrics differ from paper to paper, and there are no public benchmarks. Both reviews confirm that these weaknesses run across the whole field (Stetco et al., 2019; Carvalho et al., 2019), which makes it impossible to rank the methods objectively or to trust that any of them transfers to a new site.

Connections

- All six papers apply machine learning to real operational plant data instead of purely physics-based models.
- Reliable operation and early fault detection are the shared goal, approached from the prediction side (Tüfekçi, 2014; Tan et al., 2019) and the asset management side (Khalid et al., 2021; Fahmi et al., 2024).
- The same model families recur across the papers — neural networks, SVMs, and tree ensembles — and careful data preparation is treated as decisive in all of them.

Conflicts between the Findings

- Classical ensembles won for static regression (Tüfekci, 2014), while deep temporal models won for time-series tasks (Tan et al., 2019; Fahmi et al., 2024), so the papers do not agree on a single best model family.
- Khalid et al. (2021) used supervised learning on a reduced sensor set, while Fahmi et al. (2024) used unsupervised detection on high-dimensional raw data — two opposite philosophies on labels and feature reduction.
- No two empirical papers share a metric and a dataset, so their reported results cannot be compared directly, a gap both reviews attribute to the absence of public benchmarks (Stetco et al., 2019; Carvalho et al., 2019).

References

- Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- Fahmi, A.-T. W. K., Reza Kashyzadeh, K., & Ghorbani, S. (2024). Advancements in gas turbine fault detection: A machine learning approach based on the temporal convolutional network–autoencoder model. *Applied Sciences*, 14(11), 4551. <https://doi.org/10.3390/app14114551>
- Khalid, S., Hwang, H., & Kim, H. S. (2021). Real-world data-driven machine-learning-based optimal sensor selection approach for equipment fault detection in a thermal power plant. *Mathematics*, 9(21), 2814. <https://doi.org/10.3390/math9212814>
- Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D., Barnes, M., Keane, J., & Nenadic, G. (2019). Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy*, 133, 620–635. <https://doi.org/10.1016/j.renene.2018.10.047>
- Tan, P., Xia, J., Zhang, C., Fang, Q., & Chen, G. (2019). Dynamic modeling of NO_x emission in a 660 MW coal-fired boiler with long short-term memory. *Energy*, 176, 429–436. <https://doi.org/10.1016/j.energy.2019.04.020>
- Tüfekci, P. (2014). Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *International Journal of Electrical Power & Energy Systems*, 60, 126–140. <https://doi.org/10.1016/j.ijepes.2014.02.027>