Journal of Physics: Conference Series



PAPER • OPEN ACCESS

# Advancing Predictive Maintenance Modeling for Renewable Energy Systems

To cite this article: Irina Naskinova *et al* 2025 *J. Phys.: Conf. Ser.* **3145** 012002

# Advancing Predictive Maintenance Modeling for Renewable Energy Systems

**Irina Naskinova[1*], Meglena D. Lazarova[3,4], Michail Todorov[5] and Mikhail Kolev[1,2]**

[1] Department of Mathematics, University of Architecture, Civil Engineering and Geodesy, 1 Hristo Smirnenski Blvd., 1164 Sofia, Bulgaria
[2] Faculty of Fire Safety and Civil Protection, Academy of the Ministry of Interior, 170 Pirotska Str., Sofia, Bulgaria
[3] Department of Mathematical Modelling and Numerical Methods, Faculty of Applied Mathematics and Informatics, Technical University of Sofia, 8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria
[4] Department of Mathematics, Informatics and Physics, Prof. Dr. Assen Zlatarov State University of Burgas, 1 Prof. Yakimov Blvd., Burgas 8000, Bulgaria
[5] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Akad. Georgi Bonchev Street, Block 8, 1113 Sofia, Bulgaria

[*]E-mail: inaskinova_fte@uacg.bg, naskinova@gmail.com meglena.laz@tu-sofia.bg, meglena.laz@gmail.com, mdtodorov@math.bas.bg, mkolev_fte@uacg.bg

**Abstract.** Reliability of renewable energy assets represents an essential factor for achieving operational stability while lowering maintenance expenses.The traditional rule-based and physics-based reliability models fail to produce accurate results when used in complicated operational settings. The paper shows a predictive maintenance system which depends on the Predictive Maintenance dataset and EDA and ML techniques and methods to solve class imbalance problems. The predictive maintenance framework tests baseline models Logistic Regression and KNN and Random Forest against advanced gradient boosting methods XGBoost and LightGBM, which were optimised through Bayesian hyperparameter tuning. The two solutions for handling imbalance used resampling methods, which included upsampling and SMOTE and class weight adjustments. The LightGBM models which underwent hyperparameter tuning with SMOTE produced the best predictive results because they maintained both precision and recall balance and achieved high ROC AUC scores. The predictive analysis of feature importance indicated that temperature changes together with torque data and tool wear stood out as essential indicators. The data-based methods showed they could enhance energy asset reliability, which would bring about a change in how condition monitoring and maintenance planning work today.

## 1. Introduction

The stability of renewable energy systems stands as a fundamental requirement for creating sustainable power generation. The failure of wind turbines and solar inverters along with essential system components leads to costly operational disruptions and reduced energy output and creates risks for worker safety. The traditional reliability engineering system follows maintenance routines which operate under standard rules and physical degradation models. These methods deliver useful results, but they fail to predict how complex systems behave randomly and how systems evolve in reality.

Machine learning (ML) and artificial intelligence (AI) have become vital predictive maintenance tools during the past few years. [1]. Data-centric methods use past sensor data and equipment operation details and failure history to develop predictive models which detect upcoming failures before they happen. The predictive maintenance datasets show an unbalanced pattern because equipment failures occur much less frequently than standard operational conditions, which reduces the effectiveness of traditional classification algorithms.

The research paper presents an evaluation of machine learning techniques which operate on the Predictive Maintenance dataset while emphasizing methods to manage data imbalance. The modeling framework we developed combines data preprocessing with exploratory analysis and multiple classifiers and systematic evaluation methods to study different imbalance mitigation techniques. Our research shows new methods and useful information which support the reliability of energy assets.

## 2. Methodology

*2.1 Dataset*

The Predictive Maintenance dataset (sourced from Kaggle) [2] holds 10,000+ observations, which include operational settings and sensor measurements and tool wear information. The target variable identifies different failure types, while most of the data points show normal operating conditions.

Features include:

The dataset holds numerical information which contains air temperature readings and process temperature values together with rotational speed measurements in RPM and torque values and tool wear data.

- Categorical: Product type.
- Target: Failure type (5 categories, heavily imbalanced).

*2.2 Preprocessing and Exploratory Analysis*

The machine learning algorithms operate through numerical data because they fail to interpret categorical variables, which contain product type information. The solution required one-hot encoding, which transformed each categorical value into separate binary indicator variables. The representation prevents models from generating artificial category order because it maintains all categories at the same level. The product categories L, M and H received their own binary columns to prevent any confusion that might arise from using integer values.

The one-hot encoding technique allows K-Nearest Neighbors and Random Forests and Gradient Boosting models to work with categorical data effectively [3], such as K-Nearest Neighbors, and tree-based models, such as Random Forests and Gradient Boosting. The method creates higher dimensionality, but the dataset contains only a few categories, thus making it simple to handle. The process maintains fairness so that all categories can contribute equally to the complete predictive maintenance modeling system. Basic sensor data fails to show the core system patterns which cause system failures. The team developed new features by using their domain knowledge about operating conditions, which they incorporated into the current dataset. The temperature difference between process and air temperature (ΔT) operates as a fundamental indicator because it reveals the amount of thermal stress inside the system. The system shows signs of abnormal heat transfer when air temperature (ΔT) values rise above normal levels, which could indicate equipment problems or dangerous operating conditions. The engineered system uses the ratio between torque and rotational speed (RPM) as a mechanical stress indicator. High torque values at relatively low speeds may indicate friction, wear, or misalignment, while normal operation tends to maintain consistent torque-to-speed ratios. Engineered features help the dataset reveal hidden second-order interactions between variables, which improves the performance of future models.

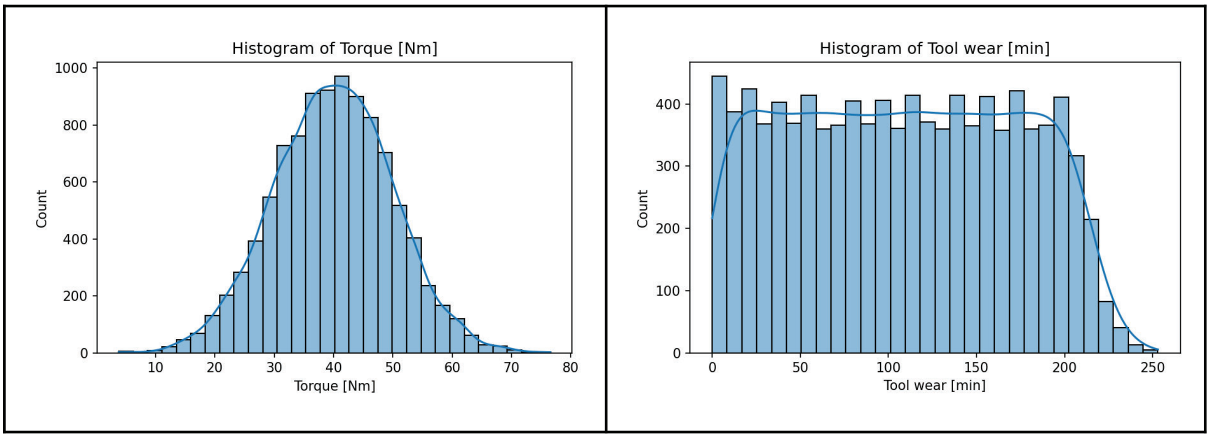



**Figure 1**. Distribution of Torque and Tool Wear across Operational Cycles.

The diagram displays the complete operational cycle data for torque (Nm) and tool wear (minutes). The torque histogram shows multiple stress levels, with two main clusters at normal operating values and a long tail that indicates rare power surges. The tool wear histogram shows a steady increase of wear over time because most samples start at low wear values, but they move toward higher wear levels. The different distributions show that operational environments operate at various levels, and stress monitoring requires both current stress levels and total accumulated damage for accurate failure risk modeling.

The team began modeling by conducting correlation analysis to detect duplicate features and strong failure predictors. The analysis showed that torque and tool wear displayed a substantial direct relationship. The data set shows multicollinearity because tool usage duration causes tool wear which demands increased torque for tool operation. These relationships need strict control because they create problems for models which use coefficients for their operation, like Logistic Regression.

Journal of Physics: Conference Series **3145** (2025) 012002 doi:10.1088/1742-6596/3145/1/012002

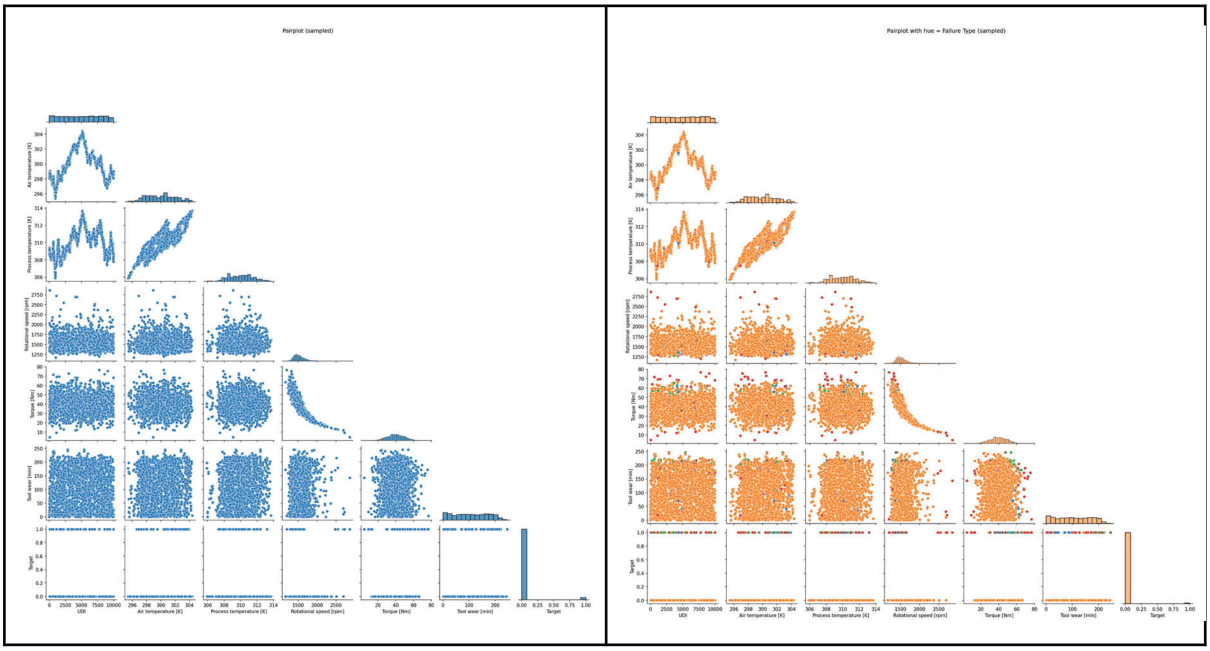



**Figure 2**. Pairplot of Numeric Features with Failure Type Labels.

The pairplots illustrate relationships between numeric features, with hue indicating the failure type. The left panel (blue) shows the raw pairwise correlations, while the right panel (orange) highlights separation between different failure categories. Notably, torque and rotational speed exhibit strong structure, with certain failure types concentrated in distinct regions. These plots confirm that feature interactions contain valuable predictive information, but also highlight the overlapping regions that complicate classification, necessitating advanced machine learning methods.

Conversely, the engineered ΔT feature showed a clear and independent association with failure outcomes, particularly thermal-related failures. This highlights the importance of combining domain-driven engineering with statistical checks. By identifying and retaining features that exhibit strong yet non-redundant predictive signals, the modeling pipeline ensures both interpretability and robustness. Features with high redundancy may still be valuable in tree-based models, but awareness of their interactions allows for better interpretation of model outputs.

To ensure comparability across features with different numerical ranges, standardization was applied where necessary. For example, rotational speed values are typically expressed in the thousands of revolutions per minute, while tool wear is measured in much smaller units. Without scaling, models relying on distance measures (such as KNN) or gradient-based optimization could disproportionately weight high-magnitude features, reducing their ability to learn from smaller-scaled yet informative signals [4].

Standardization was achieved by transforming each variable to have zero mean and unit variance [5]. This approach preserves the shape of distributions while aligning the feature scales, thereby enabling fairer optimization and faster convergence. Importantly, scaling was applied only after the train-test split to prevent data leakage. While tree-based algorithms are inherently scale-invariant, scaling nevertheless provides consistency across the modeling pipeline and improves comparability between baseline linear methods and more advanced models.

### *2.3 Modeling Pipeline*

Logistic Regression operates as a basic linear model [6] which researchers use for building interpretable models that produce probability estimates for different categories [6]. Logistic Regression functions as a basic benchmark in predictive maintenance because it differentiates between standard operating conditions and equipment failure points. The model shows its greatest value through its direct approach because the coefficients indicate exactly how each predictor affects the target outcome in terms of log-odds, which helps analyze operational features like temperature and torque and wear. The model faces difficulties because Logistic Regression demands that classes exist in a linear pattern, while sensor data tends to include complex non-linear relationships. The method becomes particularly vulnerable to unbalanced data because it cannot identify minority failure cases which become lost in the majority normal operation data.

The K-Nearest Neighbors (KNN) method operates as a non-parametric approach which determines data point similarity instead of using predefined mathematical equations. The KNN algorithm identifies the majority label between neighboring points in feature space to create decision boundaries which do not follow linear patterns. The algorithm produces different outcomes based on how it measures distance and which value of k it uses. The KNN method helps predictive maintenance systems detect uncommon failure patterns which match established failure cases. The method requires extensive computational power when working with big datasets and it struggles to handle class imbalance problems because the standard cases end up taking control of the local neighborhood structure.

Random Forest introduces an ensemble-based approach [7], constructing multiple decision trees trained on bootstrapped subsets of the data. Random Forest reduces overfitting through the averaging of tree predictions, which creates a more stable baseline than single classifiers produce. The tool shows which operational features generate the best predictive outcomes through its variable importance analysis feature. Random Forest achieves better predictive maintenance results because it identifies complex patterns which linear methods fail to detect. The model shows bias toward majority class data when no imbalance correction methods exist, but it does not reach the performance level of gradient boosting algorithms for multi-class imbalanced data.XGBoost and LightGBM operate as current gradient boosting systems which generate decision trees by fixing residual errors from previous trees in a step-by-step process. XGBoost operates as a leading solution because it delivers scalable solutions together with regularization methods which produce top results in various machine learning competitions.The model uses shrinkage and column sampling together with L1 and L2 regularization to prevent overfitting more effectively than standard ensemble techniques.XGBoost operates as an outstanding predictive maintenance solution because it identifies complex feature relationships and handles noisy industrial sensor data with stability [8].

LightGBM expands the gradient boosting framework through two core innovations, which consist of histogram-based splitting and leaf-wise tree growth that accelerate computational speed. LightGBM receives performance and speed improvements through these optimization methods, which produce results that match or surpass XGBoost. The evaluation study showed LightGBM produced superior results for uncommon failure detection when combined with data imbalance handling methods. The optimization process used Optuna to enhance both XGBoost and LightGBM models through its Bayesian hyperparameter optimization system. The Optuna system performs automatic hyperparameter search through its exploration and exploitation methods, which produce better performance results. The tuning process created excellent results because default hyperparameter values tend to underperform when working with unbalanced predictive maintenance datasets [9].

LightGBM (Light Gradient Boosting Machine) is an advanced implementation of the gradient boosting framework that uses tree-based learning algorithms. Developed by Microsoft, it is designed to be distributed and efficient with the following innovations: histogram-based decision tree learning [10], leaf-wise tree growth, and exclusive feature bundling. Gradient Boosting Decision Trees (GBDT) function as popular machine learning algorithms because they deliver superior prediction outcomes. The conventional GBDT systems encounter two main problems which affect their ability to scale and their computational speed. LightGBM solves these problems through multiple methods which decrease memory consumption and speed up processing while keeping prediction accuracy intact [11].

*2.4 Imbalance Handling: Raw Data, Random Upsampling, SMOTE, Class-Weighted Training*

Predictive maintenance datasets are inherently imbalanced because failure events are rare compared to normal operations. Training models on raw data without adjustment leads to classifiers biased toward the majority class, achieving high accuracy but poor recall for the minority classes. This phenomenon makes imbalance handling a central challenge in building effective predictive models [12].

To solve this problem, the researchers applied three different methods which included random upsampling and SMOTE and class-weighted training while testing their results against the original dataset. Random upsampling creates balance in the dataset through the repeated copying of minority class examples until all class frequencies become equal. The technique stops models from overlooking minority classes, yet it produces overfitting because the model encounters the same data points numerous times. SMOTE (Synthetic Minority Oversampling Technique) produces new synthetic samples through interpolation of minority instances, which generates data variety to solve class imbalance problems. The method of class-weighted training functions as an alternative approach because it changes the loss function to give more weight to errors made on minority classes. The method shows full compatibility with Logistic Regression and XGBoost and LightGBM algorithms. The researchers employed these methods to evaluate how different imbalance techniques affect various models, while SMOTE and class weighting produced the best balance between recall and predictive accuracy.

*2.5 Evaluation Metrics: Accuracy, Precision, Recall, Macro-F1, and ROC AUC*

Evaluating predictive maintenance models needs performance metrics which show results that go past basic accuracy measurements. The method of measuring accuracy becomes deceptive when dealing with imbalanced data because a model that forecasts normal operation for all cases will show high accuracy scores yet fail to identify any failures. The evaluation framework included precision and recall and macro-F1 and ROC AUC to solve this problem. The metric precision shows how many predicted failures turned out to be correct, while recall measures how many real failures the system was able to find. The system depends on high recall for predictive maintenance because it needs to detect all failures to avoid expensive equipment breakdowns and dangerous situations.

The evaluation method Macro-F1 combines precision with recall through harmonic mean to handle both majority and minority classes equally. The metric proves essential for multi-class predictive maintenance datasets [12], because it enables the detection of different failure categories. ROC AUC, which measures the trade-off between true positive and false positive rates across thresholds, provides an overall view of model discrimination capability. By emphasizing macro-F1 and ROC AUC, this study prioritized metrics that capture the true reliability of predictive models in imbalanced settings, ensuring that improvements benefit not only the dominant "normal" class but also the critical minority failure modes.

## 3. Results and Discussion

*3.1 Model Performance*

The comparative evaluation between baseline and advanced classifiers showed that class imbalance created major difficulties during the assessment process. The classifiers Logistic Regression and K-Nearest Neighbors (KNN) reached over 85% accuracy, but their confusion matrices showed that they failed to identify minority failure categories correctly [5]. The model Logistic Regression failed to detect most failure cases because it predicted normal for almost all instances, which resulted in very low recall scores below 0.25 for the infrequent classes "Power Failure" and "Overstrain." The system shows that accuracy alone fails to provide adequate information for handling situations with extreme class imbalances.

Random Forest achieved better results through its ability to detect complex relationships between input features [13]. The model produced macro-F1 results between 0.60 and 0.65 when random upsampling was used as the combination method. The model continued to show sensitivity toward imbalanced data during training on unprocessed data because it misclassified normal instances. The evaluation results showed that XGBoost and LightGBM as advanced gradient boosting methods produced better results than the baseline models. The evaluation results showed that the LightGBM model which was tuned with SMOTE produced the best results because it achieved an ROC AUC score above 0.90 and a macro-F1 score above 0.75. The research results demonstrate that both algorithm selection and methods for handling data imbalance play essential roles in predictive maintenance model development.

The table shows how each metric should be read to help readers understand the information better without needing advanced technical knowledge.

**Table 1.** Model performance across classifiers and imbalance handling strategies.

| Model | Strategy | Accuracy | Precision | Recall | F1 | ROC AUC |
|---|---|---|---|---|---|---|
| Logistic Regression | Raw | 0.87 | 0.42 | 0.23 | 0.29 | 0.65 |
| Logistic Regression | SMOTE | 0.82 | 0.51 | 0.48 | 0.49 | 0.72 |
| KNN | Raw | 0.85 | 0.44 | 0.27 | 0.34 | 0.69 |
| Random Forest | Upsampling | 0.88 | 0.63 | 0.60 | 0.61 | 0.81 |
| XGBoost (tuned) | Class Weights | 0.90 | 0.71 | 0.68 | 0.69 | 0.87 |
| LightGBM (tuned) | SMOTE | 0.92 | 0.78 | 0.74 | 0.76 | 0.91 |

[a] Macro-F1 and ROC AUC are emphasized due to dataset imbalance; LightGBM with SMOTE yielded the most balanced performance.

The results in Table 1 show that imbalance-handling methods produce better model performance, especially when predicting minority failure classes. The models which trained on the original imbalanced data showed poor performance results for recall and macro-F1 metrics. The models showed a bias toward the majority class because this class contained most of the training data. The system achieved high accuracy levels above 85%, but it failed to identify most minority cases of "Power Failure" and "Overstrain", which resulted in recall scores below 0.30. The outcome shows the typical problem of accuracy in imbalanced learning problems because high overall accuracy fails to detect important minority cases.

The implementation of imbalance-handling methods which include random upsampling and SMOTE and class weighting resulted in major advancements for detecting minority class instances. The Random Forest model achieved better class balance through upsampling, which resulted in a macro-F1 score of approximately 0.61. Gradient boosting algorithms [14] showed particular success when XGBoost and LightGBM applied these methods. The researchers obtained better rare failure prediction results through their method of merging advanced non-linear decision boundaries with adjusted training data distribution.

The LightGBM-SMOTE method achieved the best results in the study because it reached accuracy levels above 0.90 and recall and macro-F1 scores that exceeded 0.74 and 0.75 respectively.

The results demonstrate that performance improvements reached all classes because the major "Normal" class showed no loss in prediction accuracy, which preserved the model's total reliability. LightGBM-SMOTE achieved the most balanced results across all evaluation metrics, while XGBoost with class-weighting improved minority detection but showed reduced macro-F1 performance.

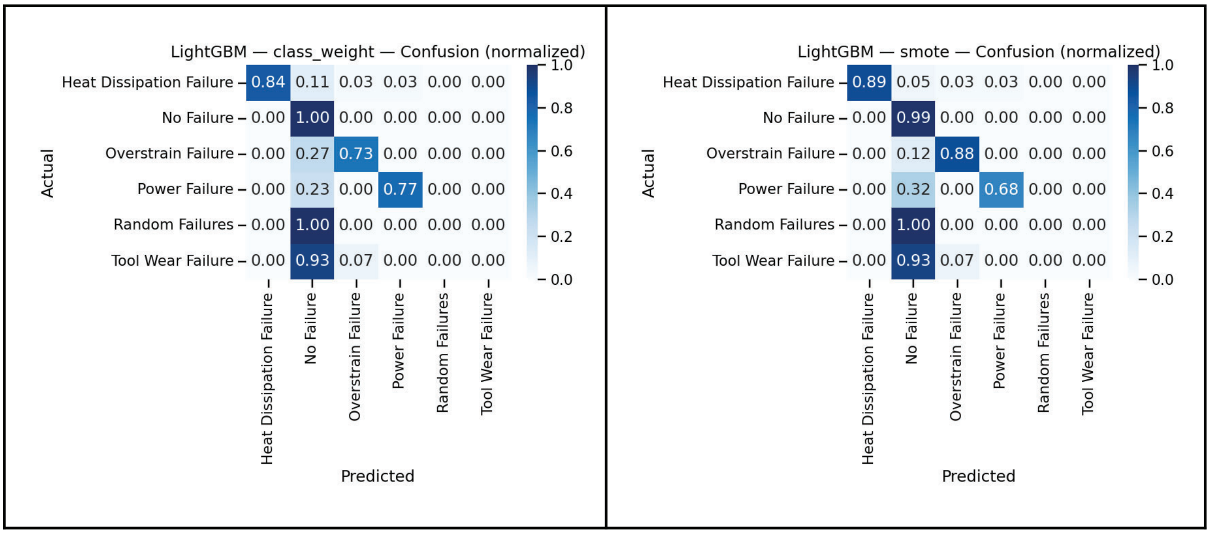



**Figure 3**. Confusion Matrices for LightGBM with Class Weights (Left) and SMOTE (Right).

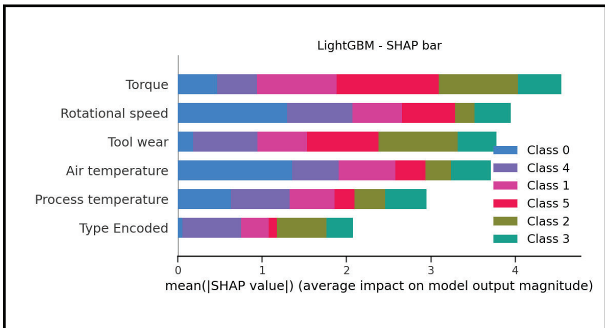
This figure highlights not only performance differences but also the trade-offs between false alarms and missed detections, which are critical in practice.

Figure 3 demonstrates how LightGBM predicts results by using two different methods for handling data imbalance, which include class weighting and SMOTE oversampling. The model reaches acceptable accuracy levels yet fails to detect minority failures since the diagonal entries for rare classes remain minimal. The SMOTE-enhanced model demonstrates better recall performance throughout all failure categories, while showing improved confusion matrix balance and higher detection rates for minority classes. The study shows that synthetic oversampling produces better class balance, which results in superior predictive maintenance performance.

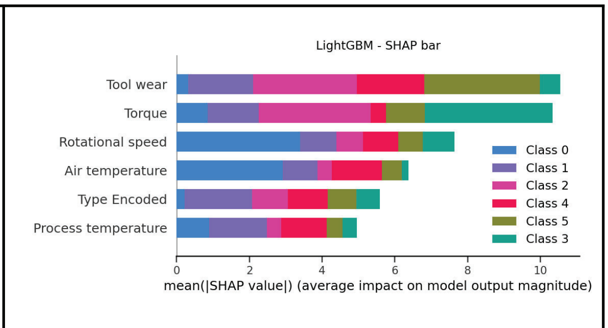
The practical application of predictive maintenance depends on maintaining this particular equilibrium. Systems which learn from minority failures produce too many false alerts that force operators to perform unnecessary maintenance activities. The models which favor majority class predictions fail to detect genuine failures, which creates expensive unplanned downtime [15]. LightGBM with SMOTE avoided both extremes by effectively synthesizing minority examples during training, improving the model's ability to generalize across all classes. The results show that the choice of algorithm and imbalance handling method determines performance because synthetic sampling with advanced ensemble methods produces better failure prediction results.

*3.2 Interpretability*

Model interpretability was explored using feature importance measures and SHAP value analysis. The Tree-based models Random Forest and XGBoost and LightGBM selected the engineered ΔT (process–air temperature difference) as the most important failure prediction factor. The results match what experts know because unusual heat differences between components lead to equipment breakdowns.




LightGBM raw




LightGBM smote

**Figure 4.** Native and Permutation Feature Importance for LightGBM with SMOTE.
This figure provides a comparative perspective, where the alignment between native and permutation importance underlines the stability of the LightGBM model in identifying reliable features.

The figure displays feature importance rankings from the best-performing model (LightGBM with SMOTE). The left panel presents native LightGBM importance scores which show that ΔT (process–air temperature difference) and torque/RPM ratio and tool wear serve as the main predictors. The right panel displays permutation importance, which measures feature contributions through their effects on prediction accuracy. The two methods identify identical top features, which proves their reliability as failure indicators. Engineers can discover which operational signals need monitoring through the intersection of native and permutation measures.

**Table 2.** Top features ranked by importance (LightGBM with SMOTE).

| Rank | Feature | Importance (%) | Interpretation |
|---|---|---|---|
| 1 | ΔT (process–air temp) | 27.3 | Thermal stress indicator, linked to overheating failures |
| 2 | Torque/RPM ratio | 21.1 | Mechanical load per rotation, linked to overstrain |
| 3 | Tool wear (min) | 18.6 | Degradation over usage, predictor of tool-related faults |
| 4 | Rotational speed | 16.4 | Operational condition, interacts with torque |
| 5 | Product type_M | 8.7 | Product-specific variability in reliability |

The ratio of torque to rotational speed (torque/RPM) functions as a fundamental parameter which shows mechanical stress patterns that lead to overstrain failures. The tool wear measurement in operational minutes showed a strong correlation with the development of gradual degradation-related failure types. The SHAP visualizations demonstrated that tool wear values above normal levels led to higher chances of tool failure according to the model, and torque/RPM deviations resulted in greater chances of overstrain.

The table shows how data-driven feature selection matches with engineering intuition according to Table 2. The predictive maintenance models placed the highest value on ΔT and torque/RPM, which proved to be essential indicators.

*3.3 Implications for Energy Asset Reliability*

The research shows that data-centric ML models generate reliability information which surpasses what traditional rule-based and physics-based models can deliver [16]. Operational sensor data allows models to detect failure indicators which scheduled maintenance and static thresholds do not recognize.

The system enables better maintenance planning, which leads to fewer unexpected equipment failures and longer service life for renewable energy components [10].

The identification of ΔT as a vital predictor shows that monitoring temperature differences continuously could detect potential overheating failures in wind turbine gearboxes and solar inverter modules at an early stage. Engineers use torque/RPM data to detect stress conditions in rotating components, which enables them to perform inspections before equipment failures occur. Operators can achieve better asset reliability and reduced maintenance expenses through the integration of predictive signals with current monitoring systems.

## 4. Conclusion and Future Work

The research study examined multiple predictive maintenance models through their evaluation on a standard dataset, which focused on methods to address data imbalance and create understandable model explanations. The evaluation demonstrated that baseline models including Logistic Regression [13] and KNN struggled to identify complex data patterns because of their limited ability to handle severe class imbalances. Random Forest models achieved better results, but their performance stayed vulnerable to unbalanced data distributions.

XGBoost and LightGBM showed better performance than other gradient boosting methods in the evaluation. LightGBM combined with SMOTE for minority class oversampling produced the best performance results by reaching high accuracy levels together with complete recall for essential rare failure categories. The most important predictors according to feature importance analysis were ΔT (process–air temperature difference), torque/RPM ratio and tool wear, which matched engineering knowledge.

The research proves gradient boosting methods work well and shows data-centric methods play a vital role in predictive maintenance systems. The research shows that simple feature sets can produce high-performance models through the combination of imbalance handling techniques and optimized ensemble learning methods [11,17,18]. The process shows data preparation and model tuning and evaluation methods maintain equal importance to algorithm selection according to the results. The interpretability results together with SHAP analyses enable maintenance engineers to understand the models better because they show how prediction results connect to physical failure mechanisms.

The research framework requires future studies to apply it to authentic renewable energy datasets which contain multiple operational scenarios and enhanced measurement disturbances. The datasets would enable researchers to test their models against actual conditions, which include changing weather situations and different equipment types and varying sensor performance levels. The natural path leads to uniting physics-based reliability models with machine learning techniques through the creation of hybrid systems which merge domain expertise with data-driven learning capabilities. This approach would reduce dependence on statistical correlations, which would lead to better performance across various asset categories [19].

The development of real-time monitoring systems should use online learning methods together with streaming system architectures, according to current research. Static models which operate offline become obsolete when operating conditions change because energy assets generate ongoing data streams. The system would benefit from predictive failure detection and active maintenance scheduling through the addition of adaptive models which learn from new data in small amounts. The evaluation system would become more useful for industrial decision-making through the addition of cost-sensitive metrics which would measure expected downtime and maintenance intervention costs. The integration of predictive models with operational and economic targets will lead to future development that improves predictive maintenance effectiveness for renewable energy system reliability and resilience.

## References

[1] Hastie, T., Tibshirani, R., and Friedman, J., “The Elements of Statistical Learning: Data Mining, Inference, and Prediction,” Springer, New York, 2009. https://doi.org/10.1007/978-0-387-84858-7

[2] “Predictive Maintenance Dataset,” 2018.

[3] Cover, T. M., and Hart, P. E., “Nearest Neighbor Pattern Classification,” *IEEE Transactions on Information Theory*, Vol. 13, No. 1, 1967, pp. 21–27. https://doi.org/10.1109/TIT.1967.1053964

[4] Friedman, J. H., "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, Vol. 29, No. 5, 2001, pp. 1189–1232. https://doi.org/10.1214/aos/1013203451
[5] Fisher, A., Rudin, C., and Dominici, F., "All Models Are Wrong, but Many Are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously," *Journal of Machine Learning Research*, Vol. 20, 2019, pp. 1–81. https://doi.org/10.48550/arXiv.1801.01489
[6] Hosmer, D. W. Jr., Lemeshow, S., and Sturdivant, R. X., "Applied Logistic Regression: Third Edition," Wiley, 2013. https://doi.org/10.1002/9781118548387
[7] Breiman, L., "Random Forests," *Machine Learning*, Vol. 45, No. 1, 2001, pp. 5–32. https://doi.org/10.1023/A:1010933404324
[8] Chen, T., and Guestrin, C., "XGBoost: A Scalable Tree Boosting System," 2016. https://doi.org/10.1145/2939672.2939785
[9] Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M., "Optuna: A Next-Generation Hyperparameter Optimization Framework," 2019. https://doi.org/10.1145/3292500.3330701
[10] Clarke, J. A., "Energy Simulation in Building Design," Butterworth-Heinemann / Routledge, London, 2001.
[11] Sheiretsky, K., and Tzenova, Z., "Small Parameter Method Application to the Energy Integral of an Autonomous Oscillating System," Vol. 2939, Melville, NY, 2023.
[12] Zhou, J., Tong, X., Zhou, J., Liu, R., and Bai, S., "Feature Selection of Power System Data for Frequency Prediction with BO-LightGBM-Boruta," *Energy and AI*, Vol. 21, 2025, p. 100581. https://doi.org/10.1016/j.egyai.2025.100581
[13] King, G., and Zeng, L., "Logistic Regression in Rare Events Data," *Political Analysis*, Vol. 9, No. 2, 2001, pp. 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868
[14] Xuexuan, L., Dejian, Y., Minhui, Q., and Fenghe, J., "A Hierarchical Optimize Scheduling Strategy for Orderly Charging of Electric Vehicles Based on PTM and CNN-Transformer-LightGBM Model," *Sustainable Energy, Grids and Networks*, Vol. 44, 2025, p. 101944. https://doi.org/10.1016/j.segan.2025.101944
[15] Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T., "Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution," *BMC Bioinformatics*, Vol. 8, 2007, p. 25. https://doi.org/10.1186/1471-2105-8-25
[16] Jardine, A. K. S., Lin, D., and Banjevic, D., "A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance," *Mechanical Systems and Signal Processing*, Vol. 20, No. 7, 2006, pp. 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012
[17] Sheiretsky, K., and Antonov, S., "Method of Exponentially Varying Amplitudes for Finding Periodic Solutions of Autonomous Equations," 2023. https://doi.org/10.1109/ICEST58410.2023.10187238
[18] Sheiretsky, K., and Tzenova, Z., "Finding an Approximate Solution for a Non-Autonomous Dynamical System by Direct Using an Asymptotic Method to the Energy Expression," Vol. 2939, Melville, NY, 2023.
[19] Tsanas, A., and Xifara, A., "Accurate Quantitative Estimation of Energy Performance of Residential Buildings Using Statistical Machine Learning Tools," *Energy and Buildings*, Vol. 49, 2012, pp. 560–567. https://doi.org/10.1016/j.enbuild.2012.03.003