International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(16s) 2026 pp. 1547-1560
https://doi.org/10.70917/ijcisim-2026-4683
Article

A Hybrid Random Forest–XGBoost Framework for Power Transformer Health Index Prediction

Pushpinder Kumar[1], Ved Parkash[2], Naveen Kumar Sharma[3]

[1]Electrical Engineering Department, MRSPTU, Bathinda, Punjab, India. pushpinder.numvas@gmail.com
[2]Electrical Engineering Department, MRSPTU, Bathinda, Punjab, India. dr.vparora@mrsptu.ac.in
[3]Electrical Engineering Department, I. K. Gujaral Punjab Technical University Main Campus Kapurthala, Punjab, India. naveen31.sharma@gmail.com

Abstract: In this study, a health index prediction framework for the power transformer is developed based on machine learning (ML) for evaluating the reliability and operational state of it. The feature engineering and Gaussian noise-based data augmentation are performed on the dataset to select the appropriate features and generate a large amount of data for ML algorithms. Several ML algorithms, including Linear Regression (LR), Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Lightweight Gradient Boosting Method (LightGBM), are employed. Among these algorithms, the two most superior are chosen based on their individual performance. The Optuna optimizer and residual learning strategy are applied to improve the performance of the proposed method and combine the predictions of superior ML algorithms for the final prediction. The simulation evaluation is done by collecting the dataset from the previous studies. The proposed method achieves an RMSE value of 2.49, an MAE value of 1.45, a MedAE value of 0.72, and an R^2 value of 0.98. Finally, the feature interpretation is done using the SHAP analysis, which shows that Dibenzyl Disulfide (DBDS), Interfacial V, Methane, and DBDS-Water exhibit the largest impact on transformer life expectancy.
Keywords: Health Index, Machine Learning, Power Transformer, Prediction, Regression

1. Introduction

Power transformers play a vital role in electrical grids by ensuring reliable, efficient, and high-quality power delivery to consumers [1]. Given the high capital cost of transformers, condition monitoring (CM) techniques are widely deployed to detect and mitigate premature failures caused by thermal, chemical, electrical, and environmental stress [2]. Various online and offline diagnostic techniques are utilized to assess transformer health, including Dissolved Gas Analysis (DGA), diagnostic oil testing, Recovery Voltage Measurement (RVM), Frequency Domain Spectroscopy (FDS), and Partial Discharge (PD) monitoring [3-4].

Traditionally, Health Index (HI) evaluation relied on weighted sum models, expert judgment, and core diagnostic parameters [5-6]. However, these approaches are prone to human subjectivity and struggle to handle complex, non-linear, or incomplete data [6]. Machine learning (ML) algorithms have recently emerged as an effective alternative to overcome these limitations and accurately predict the health index of power transformers [7-10].

Generally, an ML-based condition monitoring framework has three phases, including data collection, pre-processing, and health index prediction, as shown in Figure 1 [11]. In the first phase, the historical data of the power transformer are collected. Following that, preprocessing is performed on the dataset to prepare it for the ML algorithm by handling the missing data, feature engineering, and data splitting into training and testing sets. In the final phase, the selected features of the data are utilized to train and evaluate the ML algorithm for predicting the health index of the transformer.

Figure 1. ML-based Condition Monitoring Framework

In the previous studies, several ML algorithms employed for predict the health index of the power transformer. For instance, Arsya et al. [11] collected the 656 samples of oil-immersed power transformers for health index prediction using the ML, scoring, and weighted methods. Among these methods, the GB regression model demonstrates superior performance compared to the other methods. Initially, Soliman et al. [12] performed the data balancing and feature engineering to prepare the dataset for ML algorithms. Furthermore, the hyper-parameter tuning of the ML algorithms was performed using the grid search method. The result indicates that the LightGBM algorithm achieves the accuracy value of 97.9%. Zahra et al. [13] utilized the deep learning autoencoder and ML gradient boosting regressor for feature extraction from the dataset and estimating the health index of the transformer. The dataset was collected from Mendeley and consists of 470 samples of transformers and 14 features. The result indicates that they have reduced the dimension of the features by 35.7% and accomplished the minimum RMSE value of 1.466. Furthermore, they developed an additional approach to address the dataset's imbalance issue by using the SMOTE-ENN method; LightGBM was employed for HI classification and achieved an accuracy of 99% [14]. In a similar study, Chandramohan et al. [15] used the four ensemble learning algorithms, including RF, CatBoost, LightGBM, and XGBoost, for health index prediction of power transformers. The 470 records of the transformer were taken for evaluation purposes, and the RF algorithm achieves the lowest MAPE value of 1.24%. Mogos et al. [16] used LightGBM for health index prediction, whereas the robust expectation-maximization algorithm was used for handling the missing data. They have achieved the RMSE value of 6.23.

The previous studies indicate that the sample size (656, 470) used to evaluate the health index of the power transformer is limited [11, 13,16]. Furthermore, the hyperparameter tuning of the parameter is performed using the grid search method, which searches for the parameter value at predefined values, and each solution is independent of the others. Furthermore, the computational complexity of the grid search algorithm is high if the predefined values are large. To address these limitations, this study focused on developing a machine learning framework for predicting the health index of the power transformer. This framework employs data augmentation and feature engineering to enlarge the limited sample size of the dataset and enriches the feature representation used for training the ML algorithms. Furthermore, we aim to achieve superior predictive accuracy by benchmarking and optimizing the hyperparameters of two ML algorithms against others. The results indicate that the proposed method achieves a lower error metric value and a higher coefficient of determination compared to previous approaches.

The remaining paper is subdivided into five sections. Section 2 defines the proposed methodology. Section 3 shows the results and discussion. Finally, this study is concluded, and future work is defined in Section 4.

2. Proposed Methodology

Figure 2 shows the framework of the proposed method is developed for health index prediction of the power transformer.

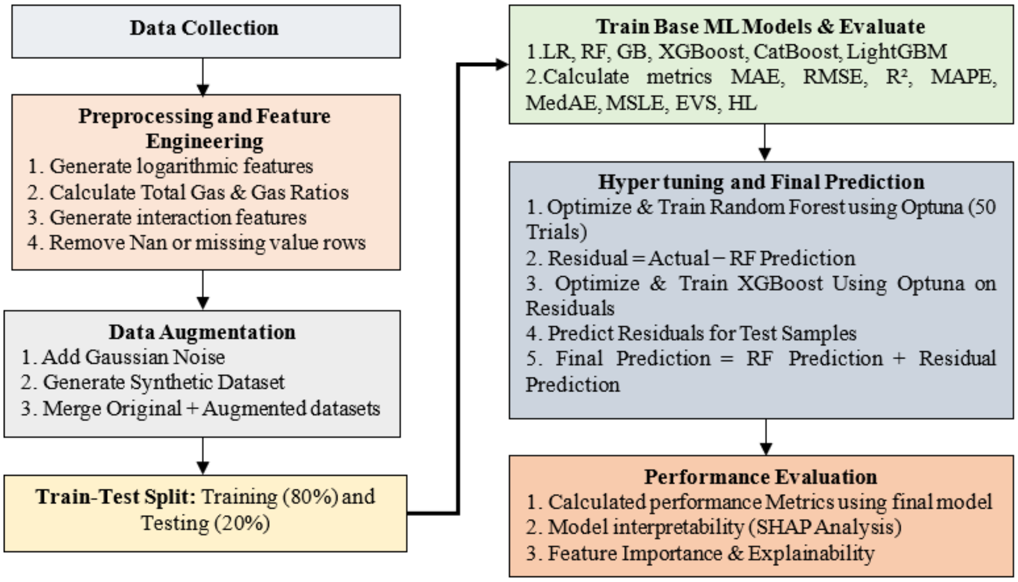


Figure 2. Framework of the Proposed Methodology

2.1 Data Collection

The dataset was collected from the previous studies [13, 17]. This dataset contains the 470 samples of the transformers and 16 features. Figure 3 shows the heat map of correlation between features.

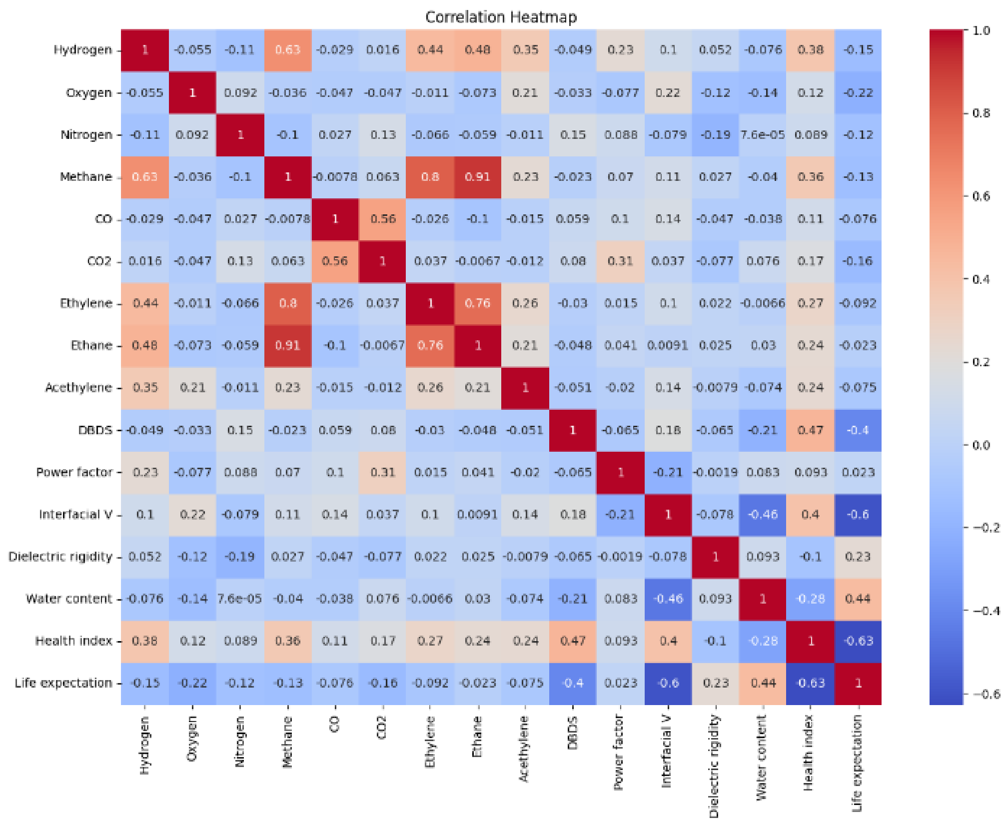
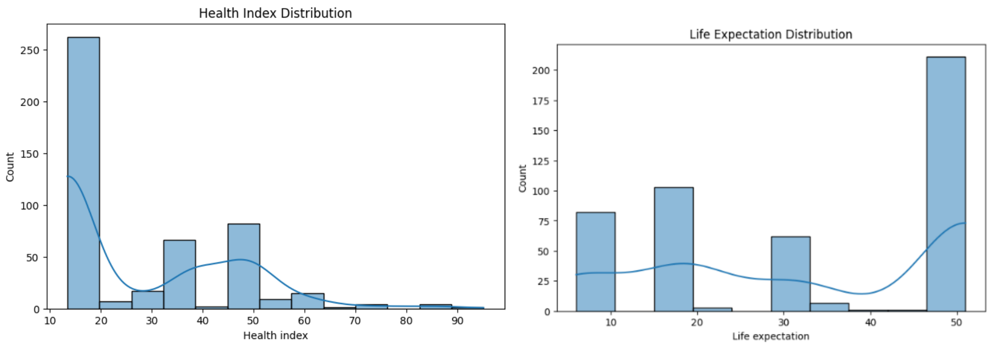


Figure 3. Correlation Heatmap

Figure 3 shows that there is a strong negative correlation between 'Health index' and 'Life expectancy', which makes intuitive sense (as health index increases, life expectancy decreases). This indicates that transformers in poorer health are expected to have shorter remaining lifetimes. Acetylene and Hydrogen showed positive correlations with the Health Index. On the other hand, Nitrogen and Oxygen showed negative correlations. Furthermore, a few gases show positive correlations with one another. This shows that under some conditions, they might increase or decrease together. These gases also show changing correlations with ‘Health index.’



(a) (b)

Figure 4. (a) Distribution of health index (b) Distribution of Life Expectancy

Figures 4 and 5 show the frequency distribution of the input and target features. In the distribution of ‘Health index’, the values are concentrated around the lower end, showing the cases of low health (Figure 4(a)). On the other hand, a binomial distribution is seen in the ‘Life Expectancy’ distribution. Its distribution shows that values of life expectancy are concentrated around 19 and 51 (Figure 4(b)).

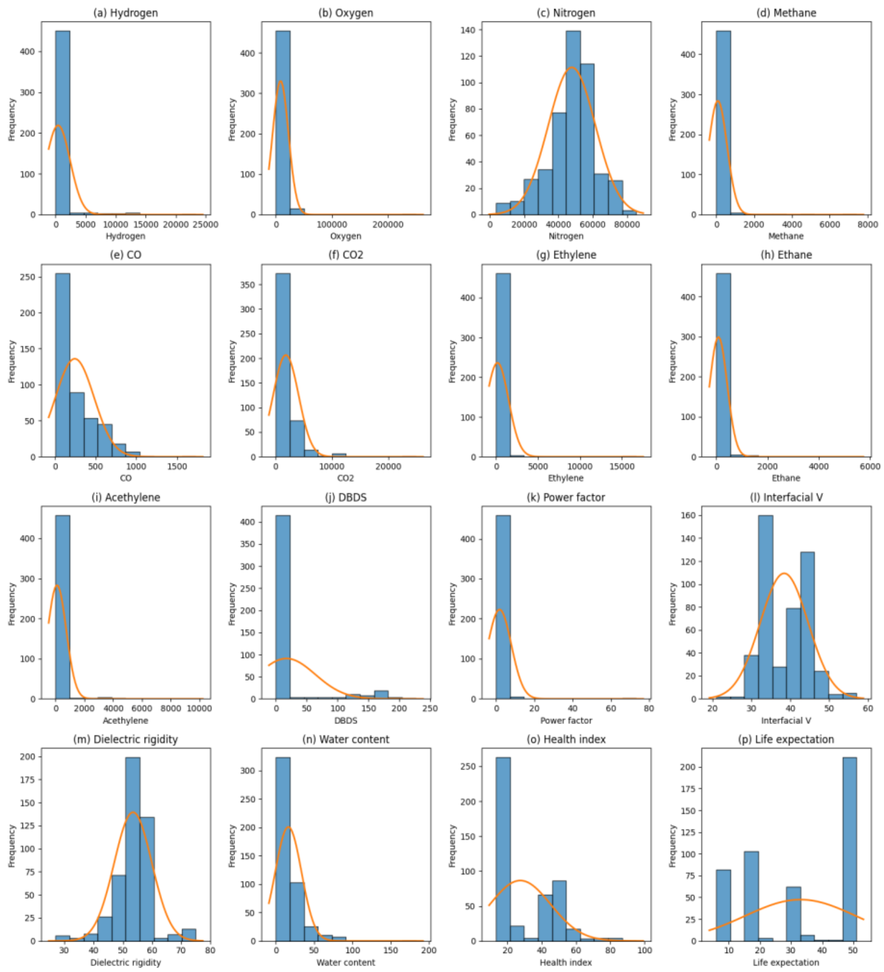


Figure 5. Input features' frequency distribution

Furthermore, the distribution of transformer parameters is shown in Figure 5 using the Kernel Density Estimate (KDE) curves in a histogram. From the graphs, it is evident that most dissolved gas features like Methane, Ethylene, Hydrogen, Oxygen, and CO2 show right-skewed distributions with some extreme values. On the other hand, Interfacial Voltage, Nitrogen, and Dielectric Rigidity show normal distributions.

2.2 Pre-Processing and Feature Engineering

In the pre-processing stage, the dataset is loaded and processed. After that, it is enhanced via feature engineering to extract some complex patterns. Some newly developed features include gas concentration ratios, logarithmic functions, interaction features, and total concentration of gases.

2.3 Data Augmentation

In the dataset, the observations were limited. So, to increase the number of observations, a Gaussian noise-based augmentation technique is used to expand the dataset. This helped to create synthetic samples and increase the training data. In this, small random noise is added to feature values using a normal distribution. The original dataset is copied multiple times before adding noise. Feature values are slightly modified to create synthetic samples. Also, negative feature values are replaced with 0 using clipping, as shown in Table 1. Initially, in the dataset, there are 470 samples, and after augmentation, the sample size increases to 1880. This indicates a 300% increment in the dataset for evaluating the generalization and robustness of the proposed method.

Table 1. Pseudocode for the Data Augmentation

Input: Dataset D Output: Augmented Dataset
1. Select feature columns and target columns 2. Set augmentation factor = 3 3. Create an empty augmented dataset 4. Repeat 3 times: a. Copy original dataset b. Add small Gaussian noise to feature columns c. Store modified dataset 5. Combine original and augmented datasets 6. Replace negative feature values with 0 7. Save the augmented dataset as a CSV file

2.4 Train-Test Split

The train and test splitting of the dataset is done to evaluate the performance of the ML algorithm. In this study, the split was done in an 80:20 ratio. This reflects that 80% of the random data was used to train the ML algorithm to learn the dataset's features, while 20% of the data was used for testing.

2.5 Train the ML Algorithms and Evaluate

In this study, LR, RF, GB, XGBoost, CatBoost, and LightGBM ML algorithms were investigated, and two superior algorithms were chosen to predict the health index of the power transformer. This section covers the description of these algorithms and how it utilized in the proposed methodology.

- LR: It is a supervised machine learning algorithm that works by finding the relation between input features and target feature using a linear function. This relation is discovered by finding a straight-line relationship between the two types of variables. In this study, LR is used as baseline model to evaluate the model performance.

- RF: This is an ensemble-based algorithm that works by including several DTs using sampling and random selection of features. The final prediction in the RF algorithm is generated using majority vote, which increases the algorithm’s robustness and reduces the prediction variance. In this study, the RF model was trained using n_estimators and max_depth parameters [17]. After analyzing the individual performance, the RF model’s parameters were optimized using OPTUNA, which includes several parameters like n_estimators, max_depth, min_samples_split, min_samples_leaf, and max_features. These optimized parameters of the model are then used as the final model.

- GB: this is an ensemble based algorithm which works by combining multiple weak DTs to develop a strong prediction model. In this, the algorithm develops and train new DT to reduce the error made by previous DT. For this study, GB can helpful to discover nonlinear and complex relationships among health index and features of transformer.

- XGBoost: In this algorithm, the trees are created in a sequential manner. Every new tree learns from the mistakes of the previous model and correct it to develop a new tree. This process is repeated until all the mistakes are corrected. Moreover, in this algorithm, regularization techniques are used to optimize the process of tree learning. With this, the overall model performance is increased via robust scalability [17]. In this study, the XGBoost model is used to train on residual errors of the RF model so that the performance of the proposed model can be increased. The parameters of the XGBoost model are also optimized using OPTUNA. Several parameters of the XGBoost model were optimized, including n_estimators, learning_rate, max_depth, min_child_weight, subsample, colsample_bytree, gamma, reg_alpha, reg_lambda, and objective. The optimized parameters are then used to prepare the final model to learn from residual errors and make predictions.

- CatBoost: This is an advanced algorithm based on gradient boosting. This algorithm helps to improve prediction performance and decrease the possibility of overfitting. This algorithm implements symmetric trees and uses ordered boosting strategies [17]. Moreover, this algorithm requires very little data processing and is resistant to noise. In this analysis, the CatBoost model was used for model training with a fixed number of iterations and random seed. The CatBoost model was used as the base model, and from all base models, RF and XGBoost were selected.

- LightGBM: This algorithm is useful for faster and more efficient training by using a histogram-based strategy and leaf-wise tree growth strategy [17]. This model is designed for large-scale and high-dimensional datasets. In this study, LightGBM was used as the base model, and due to the lower performance of the data compared to the RF and XGBoost models, it was not included in the final model.

2.6 Hyper-Parameter and Final Prediction

Among these algorithms, Random Forest was selected as the primary predictor due to its superior performance. The Optuna optimizer was used to optimize the RF hyperparameters through cross-validation, resulting in an optimized model for health index prediction. To further improve accuracy, a residual learning strategy was adopted. The prediction errors (residuals) from the optimized Random Forest were calculated and used as targets for an optimized XGBoost model. The XGBoost model learned the remaining information that the Random Forest did not capture. During prediction, the Random Forest generates the initial output, while XGBoost predicts the residual correction. The final Health Index is obtained by combining both predictions:

Final Prediction=RF Prediction + Residual Prediction (1)

2.7 Performance Evaluation

MAE, RMSE, R^2, MAPE, MedAE, MSLE, EVS, and HL parameters are determined to evaluate the performance of the proposed method and perform a comparative analysis with the previous approaches. These parameters are determined using the following equations [18-20].

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{4}$$

$$MAPE(\%) = \frac{100}{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \tag{5}$$

$$MedAE = median(|y_1 - \hat{y}_1|, |y_2 - \hat{y}_2|, \ldots |y_n - \hat{y}_n|) \tag{6}$$

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}[\log(1 + y_i) - \log(1 + \hat{y}_i)]^2 \tag{7}$$

$$EVS = 1 - \frac{\mathrm{Var}(y-\hat{y})}{\mathrm{Var}(y)} \tag{8}$$

Let $\alpha = y_i - \hat{y}_i$

Then, $$l(\alpha) = \begin{cases} \frac{1}{2}\alpha^2, & | \alpha | \leq \xi \\ \xi\left(| \alpha | - \frac{\xi}{2}\right), & | \alpha | > \xi \end{cases} \tag{9}$$

- y_i = actual value
- $\hat{y}_i$= predicted value
- $\bar{y}$= mean of actual values
- $\mathrm{Var}(y - \hat{y})$= variance of prediction errors
- $\mathrm{Var}(y)$= variance of actual values
- α is the prediction error,
- ξ is the threshold parameter controlling the transition between quadratic and linear loss.

The feature interpretation of the dataset is done using the SHAP analysis to find out the features that play an important role in the health index of the power transformer.

3. Results and Discussion

This section presents the simulation results for the developed method and comparative analysis with the previous approaches. Python was used for coding purposes and Google Colab for simulation purposes. Initially, we have imported the libraries, including pandas, numpy, matplotlib.pyplot, seaborn, and ML packages. The proposed method was evaluated on a local system with configurations, an Intel(R) Core (TM) i7-7500 CPU, a 64-bit Windows 10 operating system, and 16 GB of RAM. Table 2 shows the hyper-parameter configuration of the proposed method.

Table 2. Hyper-Parameter Configuration

Model	**Parameter**	**Value**
Linear Regression	Default	
Random Forest	n_estimators	200
	max_depth	10
Gradient Boosting	n_estimators	200
	learning_rate	0.05
	max_depth	4
XGBoost	n_estimators	200
	learning_rate	0.05
	max_depth	4
	objective	reg:squarederror
CatBoost	iterations	200
	learning_rate	0.05
	depth	4
	verbose	0
LightGBM	n_estimator	200
	learning_rate	0.05
	max_depth	4
	verbose	-1

Table 3 shows the performance of the ML models using various metrics like MAE, RMSE, R2, MAPE (%), MedAE, MSLE, EVS, and HL. Among the baseline models, XGBoost achieved the best overall performance with the lowest MAE and RMSE and the highest R^2, closely followed by Gradient Boosting and Random Forest. Linear Regression produced the weakest results because of its inability to model nonlinear relationships.

Table 3. Performance Comparison of Base Machine Learning Models

Model	**MAE**	**RMSE**	**R^2**	**MAPE (%)**	**MedAE**	**MSLE**	**EVS**	**HL**
LR	8.4918	11.1023	0.5663	56.8800	6.4153	0.3853	0.5685	8.0058
RF	4.3944	6.8650	0.8342	27.8865	3.3873	0.1099	0.8351	4.0254
GB	4.0432	6.4028	0.8558	22.7116	1.9397	0.1058	0.8608	3.6383
XGBoost	**3.9601**	**6.1688**	**0.8661**	**25.3994**	**1.9779**	**0.1117**	**0.8679**	**3.5544**
CatBoost	4.6337	6.5171	0.8506	31.4771	3.6850	0.1322	0.8507	4.1732
LightGBM	5.2806	7.5657	0.7986	34.4651	3.6485	0.1472	0.7991	4.8040

Random Forest also demonstrated competitive performance, whereas LightGBM and CatBoost achieved moderate results. In contrast, Linear Regression exhibited the highest prediction errors and the lowest goodness-of-fit. This confirms that ensemble methods are more effective for modelling the nonlinear relationships present in transformer health assessment.

3.1 Feature Importance Analysis

Feature importance (Figure 6) demonstrates that Interfacial Voltage is the most influential predictor, followed by Health Index, Water Content, Power Factor, and Hydrogen. These features mostly contribute to the prediction of the life expectancy of the transformer. This confirms the importance of merging electrical characteristics with dissolved gas measurements.

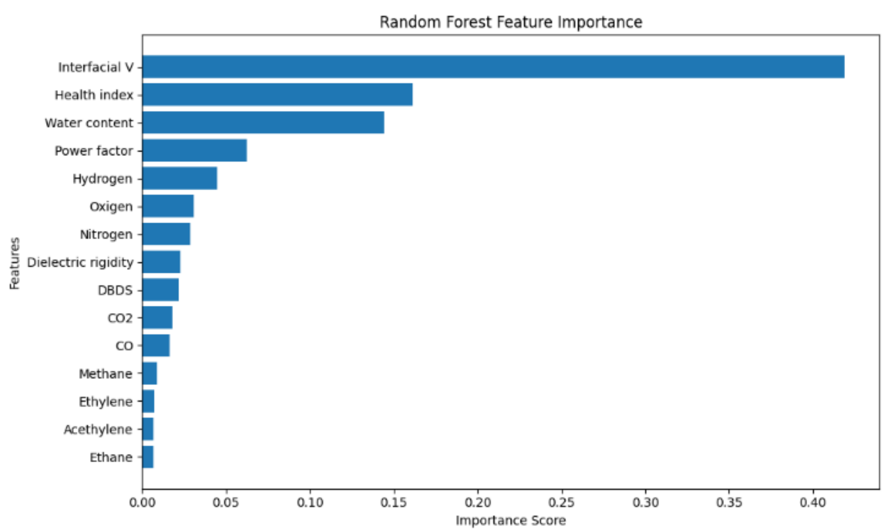


Figure 6. Feature Importance Analysis of Input Features

3.2 *Effect of Data Augmentation*

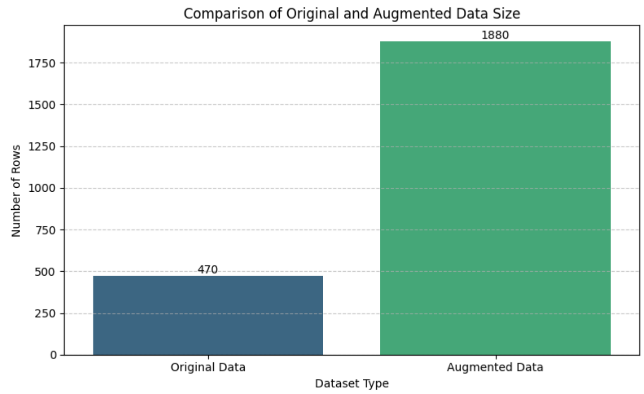


Figure 7. Size of dataset before and after augmentation

Figure 7 shows the dataset size before and after augmentation. It can be seen that the original dataset contained 470 samples. After the data augmentation, the size increased to 1880 samples. This increased-sized dataset provides more diverse samples for training. Thus, helping the ML models reduce the risk of overfitting and increase their generalization capability.

3.3 *Performance of Augmented Models*

The ML algorithms performed more efficiently after augmentation. Among all ML algorithms, Random Forest performed superiorly with $R^2 \approx 0.9465$ and MAE≈2.23 (refer to Table 4). Other ML algorithms like XGBoost, GB, and LightGBM also achieved significant improvements. This confirms the proposed augmentation strategy's effectiveness. When compared to results of without augmentation, RF performance is significantly enhanced by achieving nearly 49% MAE reduction and over 42% RMSE improvement.

Table 4. Performance Comparison of the model with Augmented + Original data

Model	**MAE**	**RMSE**	**R^2**	**MAPE (%)**	**MedAE**	**MSLE**	**EVS**	**HL**
LR	8.1186	10.4435	0.6269	37.9677	6.0072	0.1567	0.6270	7.6424
RF	2.2278	3.9540	0.9465	10.0152	0.8999	0.0233	0.9467	1.8862
GB	2.7086	4.3382	0.9356	12.8954	1.2538	0.0308	0.9357	2.2978
XGBoost	2.7567	4.5652	0.9287	13.0485	1.3220	0.0331	0.9287	2.3391
CatBoost	4.1423	6.5014	0.8554	19.3033	2.2826	0.0632	0.8554	3.7138
LightGBM	3.2563	5.1050	0.9108	15.3718	1.7799	0.0409	0.9110	2.8220

3.4 *Optuna Hyperparameter Optimization*

From the results of individual models, the top two models, Random Forest and XGBoost, are selected to develop a final model to reduce errors and improve the overall accuracy of the model. For this, Optuna was used for optimizing the model's parameters. Several optimization trials were run to find the best combination of parameters that minimizes the error (Figure 8). Then, the residual learning framework using XGBoost was developed to learn from the residual errors of Random Forest. Finally, the predictions of both optimized models are combined to develop the final model.

```python
# ====================================
# STAGE 1 → RANDOM FOREST
# ====================================

rf = RandomForestRegressor(
    **study_rf.best_params,
    random_state=42,
    n_jobs=-1
)

rf.fit(X_train, y_train)

# RF predictions
rf_train_pred = rf.predict(X_train)

rf_test_pred = rf.predict(X_test)

# ====================================
# STAGE 2 → RESIDUALS
# ====================================

residuals = y_train - rf_train_pred
```

```python
# ========================================
# TRAIN FINAL MODEL
# ========================================

best_xgb = XGBRegressor(
    **study_xgb.best_params,
    random_state=42,
    n_jobs=-1
)

best_xgb.fit(X_train, residuals)

# ========================================
# PREDICTION
# ========================================

y_pred_xgb = best_xgb.predict(X_test)
final_pred = rf_test_pred + y_pred_xgb

# ========================================
# METRICS
# ========================================
```

Figure 8. Code of Hypertuning and Optimized Final Model using RF and XGBoost

Thus, the final model combines the advantages of both RF and XGBoost using a residual learning framework. Table 5 shows the proposed model's performance using several performance metrics. The model achieved a significantly higher R2 value of 0.9788. The model also achieved lower error rates as compared to individual baseline models. Figure 9 shows the visualization of the performance metrics of the final proposed model.

Table 5. Performance Metrics of the final proposed model using RF and XGBoost

Metric	**Value**
MAE	1.4479
RMSE	2.4898
R^2	0.9788
MAPE (%)	7.0549
MedAE	0.7187
MSLE	0.0116
EVS	0.9788
HL	1.1077

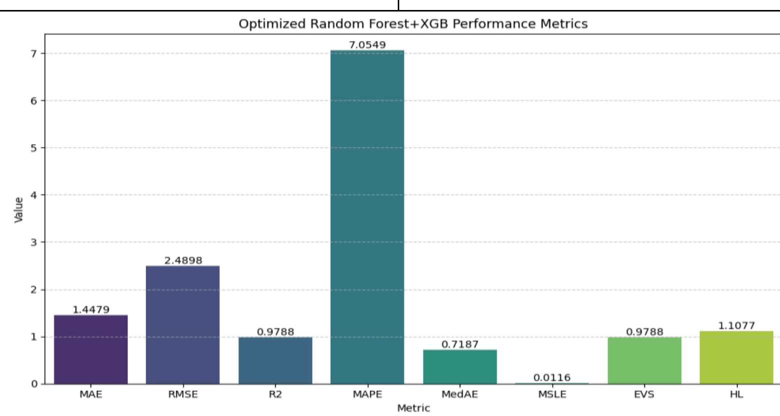


Figure 9. Performance metrics comparison of the final model (RF + XGBoost)

Figure 10 illustrates the distribution of prediction errors for the proposed Optimized RF–XGBoost hybrid model. Most prediction errors are concentrated around zero, indicating that the model produces highly accurate predictions with minimal bias.

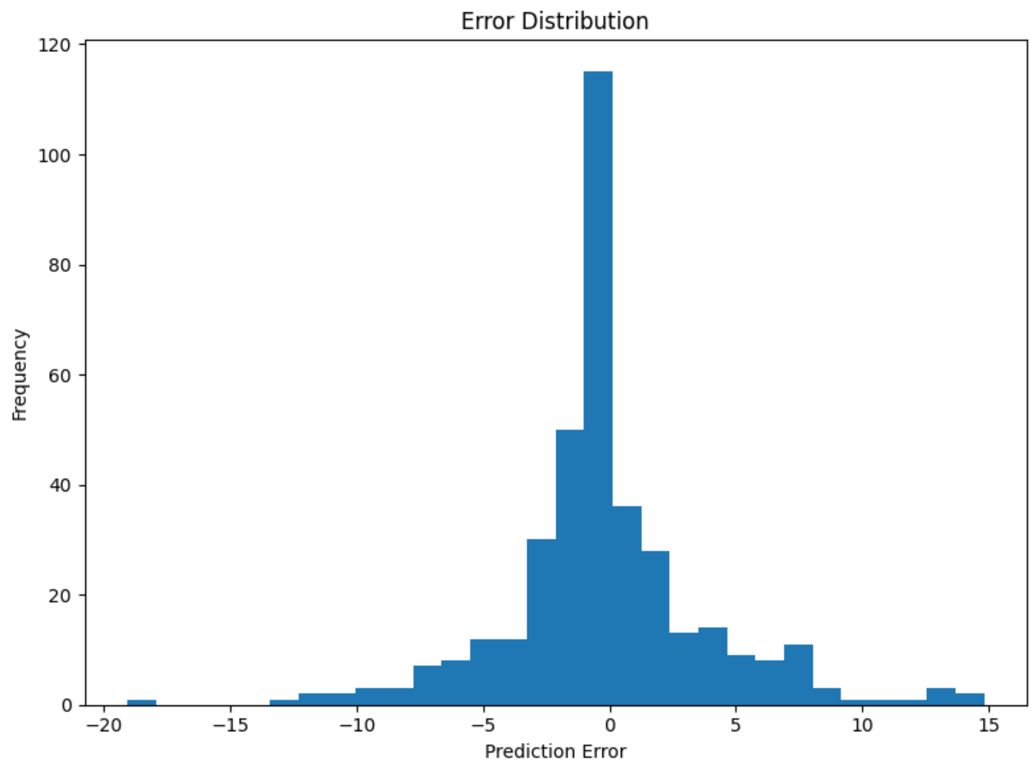


Figure 10. Error Distribution Analysis

Table 6 shows the comparison of the final hybrid model (RF + XGBoost) with baseline models [16]. It can be seen that the proposed hybrid model outperforms the baseline model in terms of all the performance metrics.

Table 6. Comparative Analysis with Baseline Models [16]

Study	**Method**	**RMSE**	**MAE**	**R2**	**MedAE**
Mogos et al. [16]	LightGBM + robust EM	6.23	3.59	0.97	1.19
Proposed method	Optimized RF + XGBoost residual learning	2.49	1.45	0.98	0.72

3.5 SHAP Explainability Analysis

SHAP analysis explains the contribution of each feature toward the final prediction, as shown in Figure 11. Higher values of the SHAP show the higher influence on model's output. The positive and negative values of the SHAP show the direction of influence on the model. Figure 11 shows the SHAP analysis of the proposed method. It can be seen that DBDS_PF, DBDS, Interfacial V, Methane, and DBDS_Water are the most influential features in the model.

4. Conclusion and Future Work

In this study, a health index prediction method is developed for power transformers. In this approach, the sample size of the dataset is increased by performing the data augmentation, and performs the feature engineering on the dataset to select the most appropriate features from it to train the ML algorithms. After that, several ML algorithms are employed, and the two best-performing algorithms (RF and XGBoost) are selected based on their individual performance. Finally, the hyperparameter tuning is performed using the Optuna optimizer to enhance the error value between the actual and predicted values. The simulation evaluation shows that the proposed method accomplishes lower values of error metrics and a high value of R^2 over the baseline models. The feature interpretation using the SHAP analysis shows that DBDS_PF, DBDS, Interfacial V, Methane, and DBDS_Water exhibit the largest impact on transformer life expectancy. In the future, deep learning and other optimization approaches will be explored to enhance the proposed method.

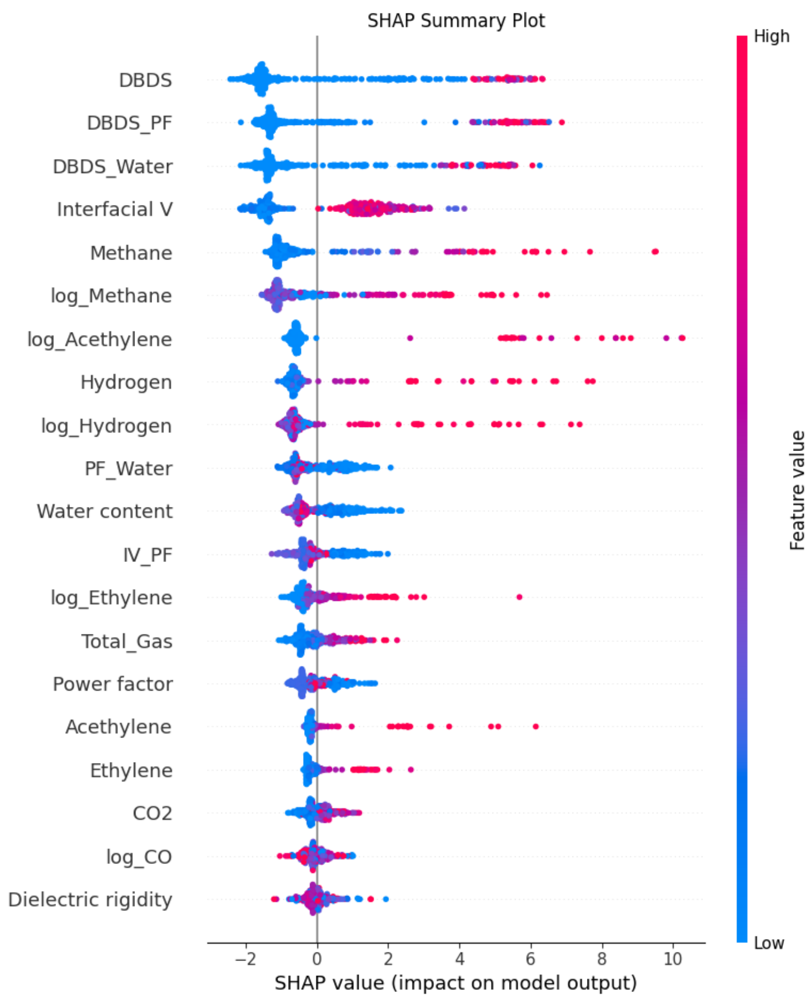


Figure 11. SHAP Analysis

REFERENCES

1. H. Guo and L. Guo, “Health index for power transformer condition assessment based on operation history and test data,” Energy Reports, vol. 8, pp. 9038–9045, Jul. 2022, doi: 10.1016/j.egyr.2022.07.041. Available: https://doi.org/10.1016/j.egyr.2022.07.041
2. R. Soni and B. Mehta, “Evaluation of power transformer health analysis by internal fault criticalities to prevent premature failure using statistical data analytics approach,” Engineering Failure Analysis, vol. 136, p. 106213, Mar. 2022, doi: 10.1016/j.engfailanal.2022.106213. Available: https://doi.org/10.1016/j.engfailanal.2022.106213
3. S. Li, X. Li, Y. Cui, and H. Li, “Review of Transformer Health Index from the Perspective of Survivability and Condition Assessment,” Electronics, vol. 12, no. 11, p. 2407, May 2023, doi: 10.3390/electronics12112407. Available: https://doi.org/10.3390/electronics12112407

4. Z. Xing, Y. He, J. Chen, X. Wang, and B. Du, “Health evaluation of power transformer using deep learning neural network,” Electric Power Systems Research, vol. 215, p. 109016, Nov. 2022, doi: 10.1016/j.epsr.2022.109016. Available: https://doi.org/10.1016/j.epsr.2022.109016
5. A. A. Adekunle, I. Fofana, P. Picher, E. M. Rodriguez-Celis, and O. H. Arroyo-Fernandez, “Analyzing Transformer Insulation Paper Prognostics and Health Management: A Modeling Framework Perspective,” IEEE Access, vol. 12, pp. 58349–58377, Jan. 2024, doi: 10.1109/access.2024.3391823. Available: https://doi.org/10.1109/access.2024.3391823
6. N. R. Arsya, Suwarno, R. A. Prasojo, G. A. Sudrajad, and A. Abu-Siada, “A holistic integration of conventional and machine learning techniques to enhance the analysis of Power Transformer Health Index considering data unavailability,” IEEE Access, vol. 12, pp. 124549–124561, Jan. 2024, doi: 10.1109/access.2024.3450110. Available: https://doi.org/10.1109/access.2024.3450110
7. S. T. Zahra, S. K. Imdad, S. Khan, S. Khalid, and N. A. Baig, “Power transformer health index and life span assessment: A comprehensive review of conventional and machine learning based approaches,” Engineering Applications of Artificial Intelligence, vol. 139, p. 109474, Oct. 2024, doi: 10.1016/j.engappai.2024.109474. Available: https://doi.org/10.1016/j.engappai.2024.109474
8. R. Zemouri, “Power Transformer Prognostics and Health Management Using Machine Learning: A review and future Directions,” Machines, vol. 13, no. 2, p. 125, Feb. 2025, doi: 10.3390/machines13020125. Available: https://doi.org/10.3390/machines13020125
9. Q. T. Tran et al., “A review of health assessment techniques for distribution transformers in smart distribution grids,” Applied Sciences, vol. 10, no. 22, p. 8115, Nov. 2020, doi: 10.3390/app10228115. Available: https://doi.org/10.3390/app10228115[10] D. Rediansyah, R. A. Prasojo, Suwarno, and A. Abu-Siada, “Artificial Intelligence-Based Power Transformer Health Index for handling data uncertainty,” IEEE Access, vol. 9, pp. 150637–150648, Jan. 2021, doi: 10.1109/access.2021.3125379. Available: https://doi.org/10.1109/access.2021.3125379
10. N. R. Arsya, Suwarno, R. A. Prasojo, G. A. Sudrajad, and A. Abu-Siada, “A holistic integration of conventional and machine learning techniques to enhance the analysis of Power Transformer Health Index considering data unavailability,” IEEE Access, vol. 12, pp. 124549–124561, Jan. 2024, doi: 10.1109/access.2024.3450110. Available: https://doi.org/10.1109/access.2024.3450110
11. A. Soliman, M. Elhemaly, and S. S. M. Ghoneim, “Integrating machine learning models for enhancing power transformer health Assessment based on the insulating Paper State,” Measurement, vol. 256, p. 118222, Jun. 2025, doi: 10.1016/j.measurement.2025.118222. Available: https://doi.org/10.1016/j.measurement.2025.118222
12. S. T. Zahra, S. K. Imdad, S. Arif, M. F. Qureshi, H. Altammar, and A. Ahmad, “Power Transformer Health Index and Life span Prediction using Supervised Machine Learning: a hybrid Autoencoder enhanced gradient boosting approach,” IEEE Access, vol. 14, pp. 7185–7203, Jan. 2026, doi: 10.1109/access.2026.3651906. Available: https://doi.org/10.1109/access.2026.3651906
13. S.T. Zahra, S.K Imdad, M.F. Qureshi, and N.A. Baig, “Power Transformer Health Index Assessment: An Explainable LightGBM based minority aware learning using SMOTE-ENN hybrid resampling,” Available at SSRN 7162474.
14. J. Chandramohan, K. Karthick, A. S. K, and G. Ponkumar, “A robust explainable machine learning pipeline for transformer health index prediction addressing data pathologies and redundancy,” Electric Power Systems Research, vol. 259, p. 113275, May 2026, doi: 10.1016/j.epsr.2026.113275. Available: https://doi.org/10.1016/j.epsr.2026.113275
15. A. S. Mogos, X. Liang, and C. Y. Chung, “Enhancing transformer health index prediction using dissolved gas analysis data through integration of LightGBM and Robust EM algorithms,” IEEE Access, vol. 12, pp. 1–12, 2024. Available https://doi.org/10.1109/ACCESS.2024.3439248
16. H. Thamtam, S. Rathee, S. K. Joshi, K. Nagaraj, S. K. Bhat, and A. Kumar, “Explainable AI–driven sustainable transformer health monitoring for resilient power distribution systems,” Energy Conversion and Management X, vol. 31, p. 102133, Jul. 2026, doi: 10.1016/j.ecmx.2026.102133. Available: https://doi.org/10.1016/j.ecmx.2026.102133
17. F. Tian, X. Fan, R. Wang, H. Qin, and Y. Fan, “A power forecasting method for Ultra-Short-Term photovoltaic power generation using transformer model,” Mathematical Problems in Engineering, vol. 2022, pp. 1–15, Oct. 2022, doi: 10.1155/2022/9421400. Available: https://doi.org/10.1155/2022/9421400
18. Z. Arslan, M. E. Sahin, and H. Ulutas, “Comparative evaluation of machine learning and Transformer-Based deep learning for solar energy forecasting on a Real-World dataset,” Arabian Journal for Science and Engineering, vol. 51, no. 5, pp. 6833–6863, Nov. 2025, doi: 10.1007/s13369-025-10769-8. Available: https://doi.org/10.1007/s13369-025-10769-8
19. M. A. A. Putra, Suwarno, and R. A. Prasojo, “Improving Transformer Health Index Prediction Performance Using Machine Learning Algorithms with a Synthetic Minority Oversampling Technique,” Energies, vol. 18, no. 9, p. 2364, May 2025, doi: 10.3390/en18092364. Available: https://doi.org/10.3390/en18092364