The 4th International Seminar of Science and Technology
ISST 2024 Vol 4 (2025) 007
Innovations in Science and Technology to Realize Sustainable Development Goals
Faculty of Science and Technology
Universitas Terbuka

Application of Machine Learning for Predictive Maintenance in Power Transformer Health Assessment: A Comparative Study of Support Vector Machine, Artificial Neural Network, and Random Forest

Endah Septa Sintiya [1]*, Ekojono[1], Azis Rahman Prasojo[2], and Hilda Khoirotul Hidayah[1]

[1]State Polytechnic of Malang, Information Technology Department, Malang, Indonesia, 65141
[2]State Polytechnic of Malang, Electrical Engineering Department, Malang, Indonesia, 65141

Abstract - Electricity plays an irreplaceable role in human daily life. To meet the increasing demand for electrical energy, a reliable electrical system, such as a power transformer, is required. Power transformers hold a crucial role in the electrical power system, where the long-term reliability of the transformer is closely related to the safety and stability of the power system. Therefore, transformer maintenance must be carried out to anticipate sudden failures and ensure the overall reliability of the electrical power system. These assessments can be performed in various ways, including the Health Index and Dissolved Gas Analysis. The Duval Pentagon Method (DPM) and Duval Triangle Method (DTM) are used in Dissolved Gas Analysis to ascertain the condition of transformers. In this development, a comparison of three machine learning models—SVM, ANN, and Random Forest—was made using the DPM and DTM datasets to obtain the model with the highest accuracy. The confusion matrix was applied to each DTM and DPM method with several split ratios for training and testing sets. The splits included 90:10, 80:20, 75:25, and 60:40. The model with the highest accuracy will be implemented in a transformer maintenance information system to determine the transformer's condition. The results of the Health Index and Dissolved Gas Analysis calculations can determine the appropriate recommendations for power transformer actions.

Keywords: predictive maintenance, power transformer health assessment, machine learning, SVM, ANN, random forest

1 Introduction

In modern society, electrical power is a fundamental necessity. To meet the growing energy demands, a dependable electrical infrastructure is crucial. Power transformers are critical components in the electrical distribution system. Consequently, monitoring transformer health is essential to avoid unexpected breakdowns that could disrupt power supply to consumers. To prevent transformer damage and ensure continuous operation, specialists in the field conduct regular maintenance

*Corresponding author: e.septa@polinema.ac.id

procedures. This maintenance is carried out through various methods, including Health Index Assessment and Dissolved Gas Analysis (DGA) [1].

To conduct Health Index Assessment and Dissolved Gas Analysis [1], it is necessary to obtain test data on Oil [2], [3] and Paper parameters [4], which are components of the transformer insulation system. The process of evaluating transformers involves multiple stages to detect faults, requiring considerable time and effort to determine the proper course of action. Additionally, numerous transformers require maintenance by experts. Consequently, there is a need for a system that can expedite the assessment procedure and recommend appropriate actions based on the identification of faults [5].

The transformer insulation system's test outcomes will be utilized for Health Index Assessment and Dissolved Gas Analysis. While the system conducts the evaluation, it does not provide instructions on how to examine the transformer insulation system. It serves as a calculation tool, identifies faults, and stores transformer report data [6]. The system's development will involve prior design and planning to streamline the implementation phase. Each feature will undergo thorough testing to ensure completeness, allowing for immediate correction of any inconsistencies that may arise.

The Health Index Assessment is divided into two types: Oil Index and Paper Index [7]. The test results used for the Oil Index Assessment [2], [3] include Breakdown Voltage (kV), Moisture Content (ppm), Acid Number (MgKOH/mg), Interfacial Tension (Dyne/cm), and Color Scale. For the Paper Index [8], the test results include CO (ppm), CO2 (ppm), CO2/CO ratio, 2FAL, Dpest, CH4 (ppm), C2H6 (ppm), C2H4 (ppm), Interfacial Tension (Dyne/cm), Moisture Content (ppm), Acid Number (MgKOH/mg), and Color Scale. The results of the oil and paper index calculations will determine the transformer's condition and recommended actions.

Various techniques, including the Duval Pentagon Method (DPM) and Duval Triangle Method (DTM), are employed to conduct Dissolved Gas Analysis (DGA) [9]. These approaches involve multiple computational steps, which require considerable time for fault detection. Nevertheless, they demonstrate high precision in identifying faults. This study aims to evaluate three machine learning models trained using DTM and DPM datasets to determine which model yields the highest accuracy.

This research is supported by multiple prior studies, including [6]. The study aims to determine the Health Index [8] by considering three factors: Failure Factor, Oil Quality Factor [10], and Paper Condition Factor. Each of these parameters is evaluated and weighted differently. The findings of this investigation demonstrate a strong correlation between the calculated Health Index and the actual condition of the transformer.

In the research titled [11] purpose of this study is to apply the Health Index method to 143 transformers in order to calculate Health Index values across different transformer groups. Once these values are obtained, a linear regression is conducted to develop a linear function that explains the relationship between the Health Index values and the transformer's age. In this research, the aging condition of transformers is assessed by evaluating their apparent age, which is determined by plotting the Health Index values on the regression line. The apparent age is considered a more accurate reflection of the transformer's aging state compared to its actual age. Transformers whose apparent age exceeds their actual age are deemed to be aging more rapidly. The proposed approach is designed to help experts take preventive measures to ensure the reliability and performance of transformers while minimizing the risk of unexpected failures.

In the study titled [12] the objective is to develop AI-based approaches to enrich conventional methods (Weight-sum) in measuring the Health Index of power transformers. Several AI algorithms are compared to find the model with the highest accuracy in predicting the Transformer Health Index. The methods compared include Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Markov Chain, Support Vector Machine (SVM), General Regression Neural Network (GRNN), Principal Component Analysis (PCA), Random Forest [13], [14], and Analytical Hierarchy Process (AHP) [15]. In this study [12], the highest classification accuracy was achieved with the Random Forest method, at 97.3% for AI Full (based on AI for all classification stages) and 98% for AI-SW (based on AI for factor categories and Scoring-Weighting based on HI category). Other algorithms, ranked by accuracy, are decision tree (96%), AdaBoost (94.8%), neural network (91.3%), SVM (89.3%), Naïve Bayes (70.7%), and kNN (70%).

Therefore, this research uses machine learning models such as Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest. These models were chosen because previous studies on the application of machine learning models in Health Index Assessment have shown that these three models have high accuracy. The model with the highest accuracy will be implemented in the transformer maintenance information system to facilitate fault identification. These types of faults will be categorized based on the severity of the condition. Based on these categories, the priority of transformer maintenance can be determined. See the separate guidelines document for further information.

2 Materials and methods

2.1 Methods

The development of a system to monitor transformer health using the Health Index and Dissolved Gas Analysis follows the prototype model of the Software Development Life Cycle (SDLC). The prototype model is a method that allows users to have an initial visualization of the software being developed, and they can perform early testing before the software is released. The stages of the method used are shown in Fig 1. The development process using the prototype model involves several structured and iterative stages.

2.2 Requirement Analysis

At this stage, data and information collection is carried out as a basis for system development. The requirements analysis phase is divided into two parts: functional requirements and non-functional requirements. Functional requirements encompass various processes that will be performed by the system, such as the system’s response to certain situations. The functional requirements can be seen in Table 1.

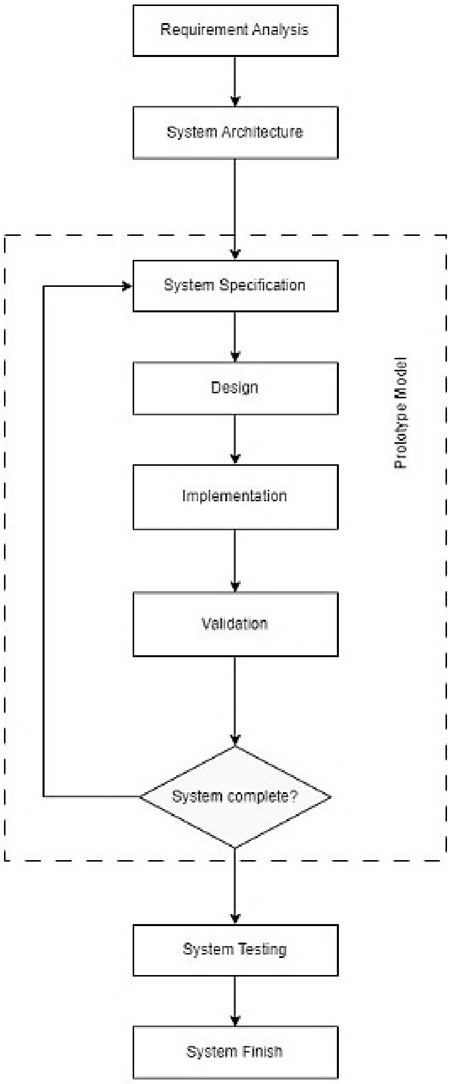

Fig. 1. Prototype Model

Table 1. Functional Requirement

No	User	Description
1	Management Staff	Can input parameter data that affects the transformer's condition during maintenance and view the calculation results of the Health Index and Dissolved Gas Analysis (DGA).
2	Verifier	Can verify the Health Index and Dissolved Gas Analysis (DGA) report data that has been entered.
3	Asset Manager	Can input transformer data containing descriptions of each transformer type and view the recommended actions needed for the transformer.
4	Supervisor	Can customize calculation formulas as needed, view recommended actions for the transformer, and download reports.
5	Admin	Can input parameter and transformer data, verify stored data, customize formulas, view calculation results and recommendations, and download reports.

Table 2. Non-Functional Requirement

No	Parameter	Description
1	Usability	The system is designed to be easily understood by users.
2	Compatibility	All features in the system can be used on all types of browsers.

3	Security	The system can only be accessed by users registered in the system. Each user will have authorization according to their role.

2.2.1 System artchitecture

After requirement analysis, the next stage is system architecture, which is the design or arrangement of a system that may include networks, hardware, and structured software. The transformer asset management system architecture can be seen in Fig 2.

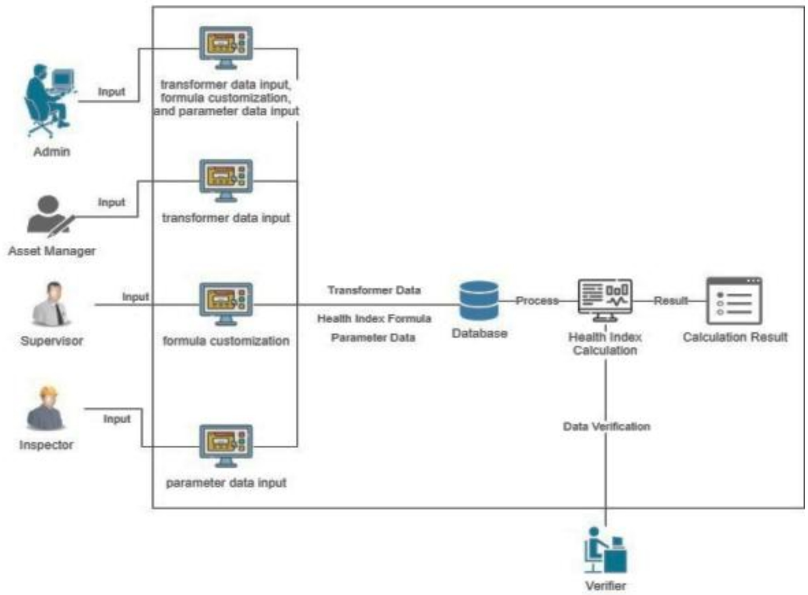

Fig. 2. System Architecture

This system is designed for five types of users, namely admin, maintenance staff, verifier, asset manager, and supervisor. The admin can input transformer data, parameter data, and customize formulas. The asset manager inputs transformer data containing descriptions of transformer types. The supervisor customizes the health index formulas according to calculation needs. The maintenance staff inputs parameter data based on testing conducted on transformers. The entered data will be stored in the database for later processing by the system. Processing is carried out using the health index method integrated into formulas and machine learning. The calculation results from this process will be used to determine the prioritization of transformer maintenance. The overall processing results will be displayed by the system in the form of reports.

2.2.2 System specification

After the system architecture is completed, the next stage is system specification. System specification explains every aspect of the system parts to be built and details the functionality and features required by the system being developed.

The system specifications, tailored to the user requirements in this asset management system, are as follows:

a. The system features authentication and authorization to ensure that only authorized users can access the system.
b. The system can store transformer data, including descriptions such as specifications, brand, code, serial number, and so on.

c. The system can store parameter data entered by users. The parameters data can be added as new input values to the machine learning dataset.
d. The system provides a feature to customize the health index formulas used in calculations.
e. The system can provide options for DPM or DTM method to be used in determining the fault conditions.
f. The system can determine the health condition of transformers using the integrated health index method in customized formulas.
g. The system can visualize a summary of stored data, and users can download it in the form of reports.
h. The system can send notifications to users when there are processes in the system that require specific actions.

2.2.3 Design

The next stage is to design the specified system. In this design stage, functional requirements are designed using the Unified Modeling Language (UML) method, which includes use case diagrams. Use case system diagram as shown in Fig 3.

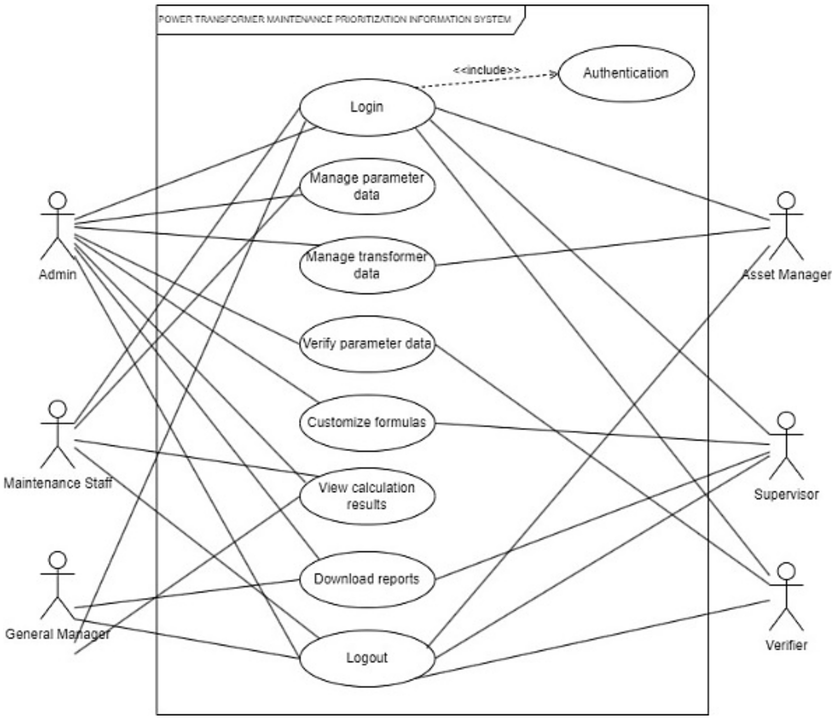

Fig. 3. Use Case Diagram

2.2.4 Implementation

In the implementation stage, a transformer condition assessment system will be built by applying a machine learning model, based on the design results from the previous stage. The system will be web-based. Then, the machine learning model with the highest accuracy will be developed using the Python programming language.

2.2.5 Validation

The Validation stage aims to ensure that the prototype consistently meets the established requirements, in accordance with the system specifications and requirement analysis detailed earlier. This stage also includes evaluating the consistency of implementation with the designed system. It involves reviews to ensure that each implemented system element aligns with the plan established during the design stage.

2.2.6 System complete

In this stage, a check is conducted to ensure whether the entire system is complete or still requires refinement. This verification involves comparison with the requirements stated in the system requirement analysis stage. If the system meets the requirements well, the process will proceed to the system testing stage. However, if there are discrepancies in terms of features or functions, revisions will be made, starting from the system specification stage. This process repeats until the system is deemed to meet the requirements set in the requirement analysis stage. At this stage, there may be several iterations until the system is declared complete according to the specified standards.

2.2.7 Testing

System testing includes a comprehensive examination of the entire system to identify and rectify any remaining issues. This testing ensures that the entire prototype functions correctly and meets the established criteria.

2.2.8 System finish

The final stage of the SDLC prototype method is the system finish. The system that has been built and passed through the system testing stage can now be used. At this stage, the system has been confirmed to run smoothly.

3 Results and discussion

3.1 Method Testing Result

The dataset contains parameters from the Duval Triangle Method (DTM), Duval Pentagon Method (DPM), and Transformer Fault Types. Before modelling, missing data is handled, and the dataset is split into training and testing sets. Three machine learning models—Artificial Neural Network (ANN), Support Vector Machine (SVM), and Random Forest—are compared. Hyperparameters for each model are optimized using Grid Search and Random Search. Finally, the models are trained with the best hyperparameters. The test results for the machine learning models can be seen in Fig 4, 5, and 6.

Using default hyperparameters, the models showed stable but lower accuracy compared to results from tuning via Grid Search and Random Search. The graph in figure 4 displays the Precision, Recall, F1-score, and Accuracy metrics across different datasets (DTM 1, DTM 4, DTM 5, DPM 1, DPM 2) and their respective hyperparameter tuning methods (No Tuning, Random Search, Grid Search). This helps visualize the performance impact of various tuning methods on the models applied to each dataset. You can observe variations in all four metrics based on the dataset and tuning method.

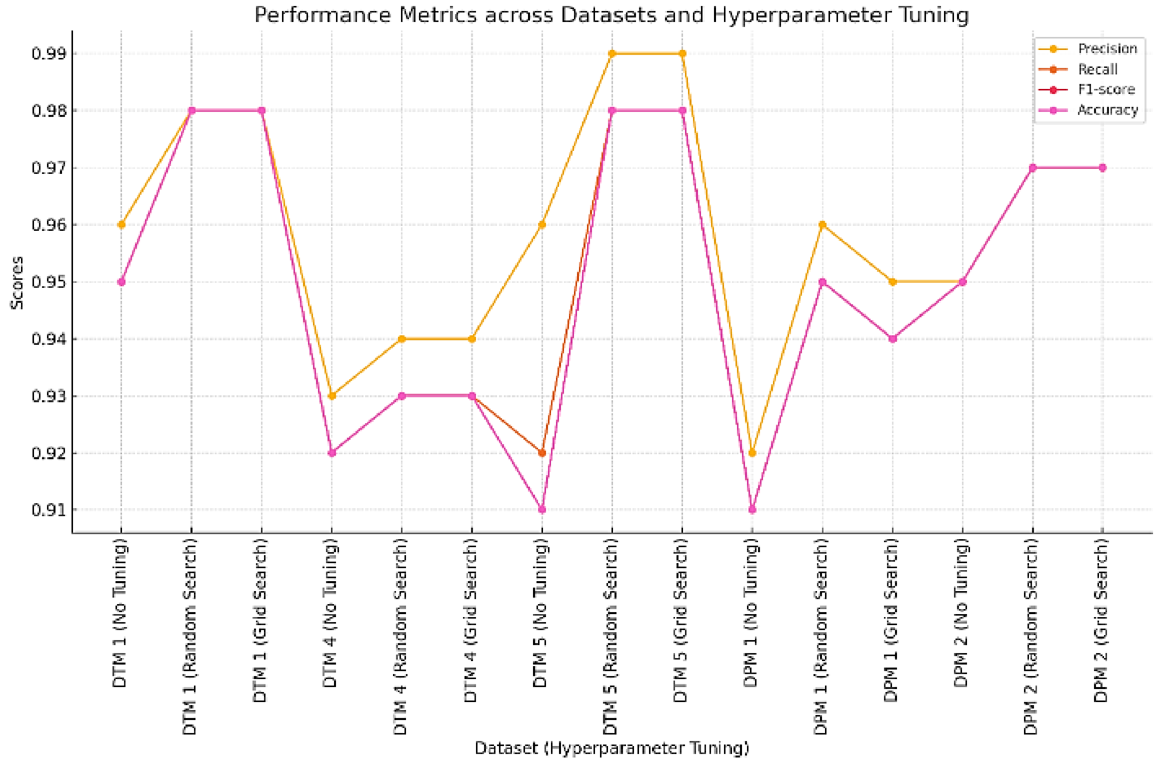

Fig. 4. Confusion Matrix Result for Artificial Neural Network

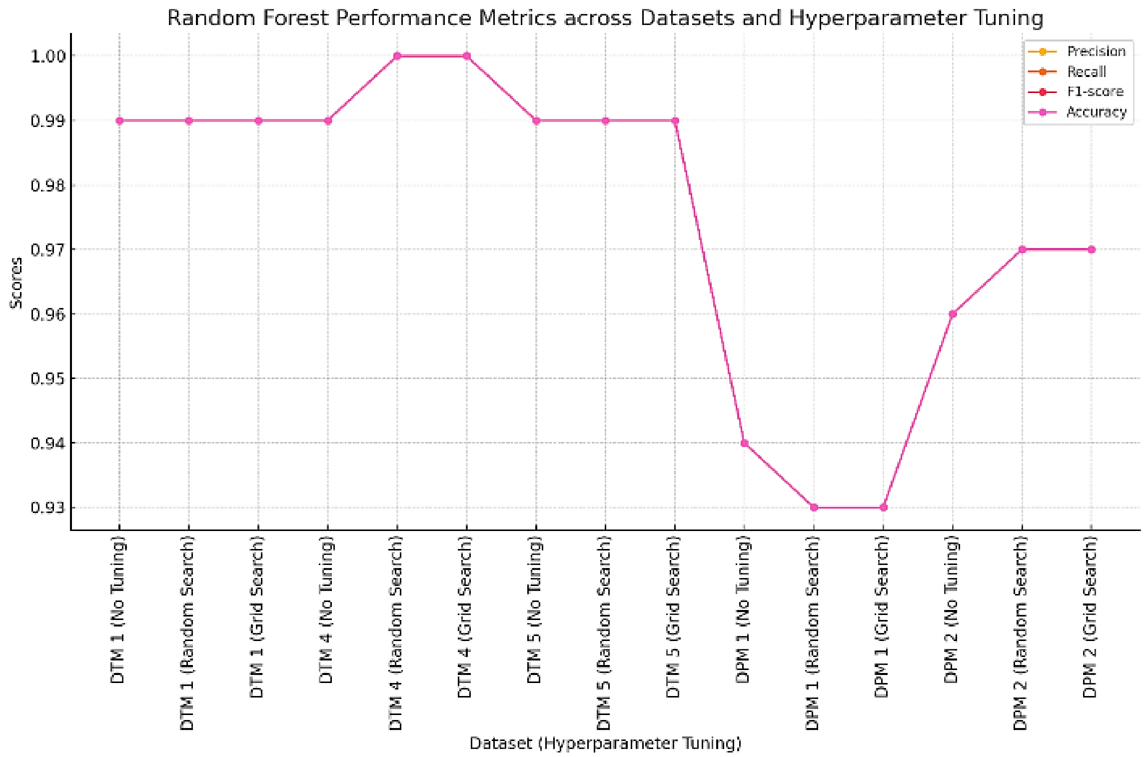

Fig. 5. Confusion Matrix Result for Random Forest

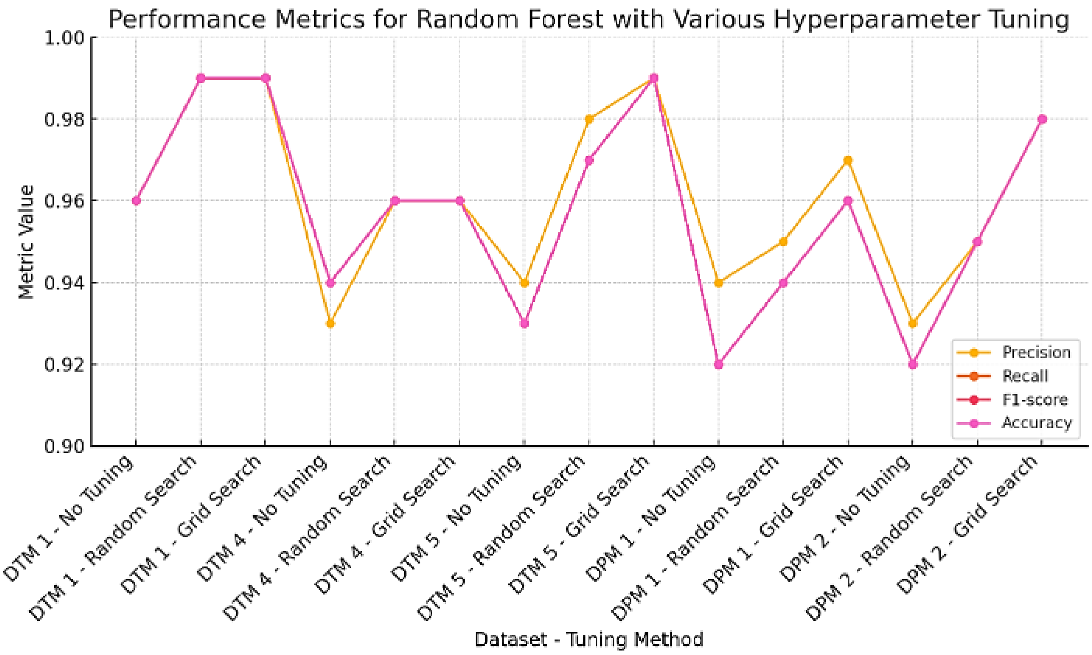

Fig. 6. Confusion Matrix Result for SVM

3.2 Functional Test Result

System testing using blackbox testing is a method to discover functions or features that may not be working correctly and to test whether the system operates according to the designed features. The testing aims to ensure that the system aligns with the previously established design. One of the test cases, detailed in Table 3, examines the functionality before the user enters the system.

Table 3. Functional testing before the user logs into the system

Page	Feature	Testing Scenario	Expected Outcome	Result
Login	Login form	User logs in with correct email and password registered in the system	User successfully logs into the system	Successful
		User logs in with incorrect email and password not registered in the system	User cannot log into the system and remains on the login page	Successful
	Registration button	User clicks on the "Register" button	User is directed to the registration page	Successful
Register	Registration form	User enters name, email, password, confirm password, and checks the checkbox agreeing to terms and conditions to register	User successfully registers and is redirected to the dashboard page	Successful
		User enters name, email, password, confirm password, but does not check the checkbox agreeing to terms and conditions	Registration fails and the system displays an error message	Successful
	Login button	User clicks on the "Login" button	User is directed to the login page	Successful

3.3 System Test Result

The power transformer maintenance system, using Health Index and DGA methods, was tested using User Acceptance Testing (UAT) with seven participants from the Diploma-IV Electrical Systems program. The results showed that the system successfully met its intended functions, with users finding the interface easy to understand, the features easy to operate, and the information clear. Most participants agreed that the system effectively determines the health index and identifies transformer faults, though some minor bugs were noted. Users also provided suggestions for further development to enhance the system and expand its application to other areas like power distribution. Overall, the system received a high satisfaction score of 92.85%, reflecting its strong usability and functionality.

3.4 Discussion

Based on the research data testing conducted by the examiner, the discussion outcomes indicate that the power transformer maintenance prioritization information system, using Health Index and Dissolved Gas Analysis (DGA) methods, has been well-received by users. The application interface meets user needs and is easy to understand, with all features being user-friendly and functioning as intended. Users found the information provided by the system clear and helpful, particularly in implementing Health Index and DGA methods to determine transformer conditions and recommend appropriate maintenance actions. The custom formula features also proved beneficial for users, allowing them to calculate the Health Index according to specific research needs. Additionally, the use of Machine Learning in identifying transformer disturbances within DGA has been advantageous for users. Feedback from users suggests further development of the application, including the addition of more methods for comprehensive analysis or the creation of similar systems with different objectives.

3.5 System Implementation

The Random Forest and SVM models are implemented on the Dissolved Gas Analysis (DGA) calculation page under the Transformer Fault Type determination section. Users must input transformer data and gas parameter test results, from which the system calculates deltas and rates to determine the transformer's status and provide recommendations. Users can choose from methods DTM 1, DTM 4, DTM 5, DPM 1, or DPM 2. After selecting a method, the system displays the transformer's Fault Type. Figure 4 shows the flowchart for determining Transformer Fault Type using the Random Forest and SVM models.

The Random Forest and SVM models are implemented on the Dissolved Gas Analysis page. Users must input transformer data and gas parameter test results to determine the fault type, as shown in Fig 7. The delta and rate calculations, along with the transformer's status and recommended actions, are displayed in Fig 8. Fig 9 shows the implementation of Random Forest and SVM models in selecting fault type determination methods, with the fault type result appearing as shown in Fig 8.

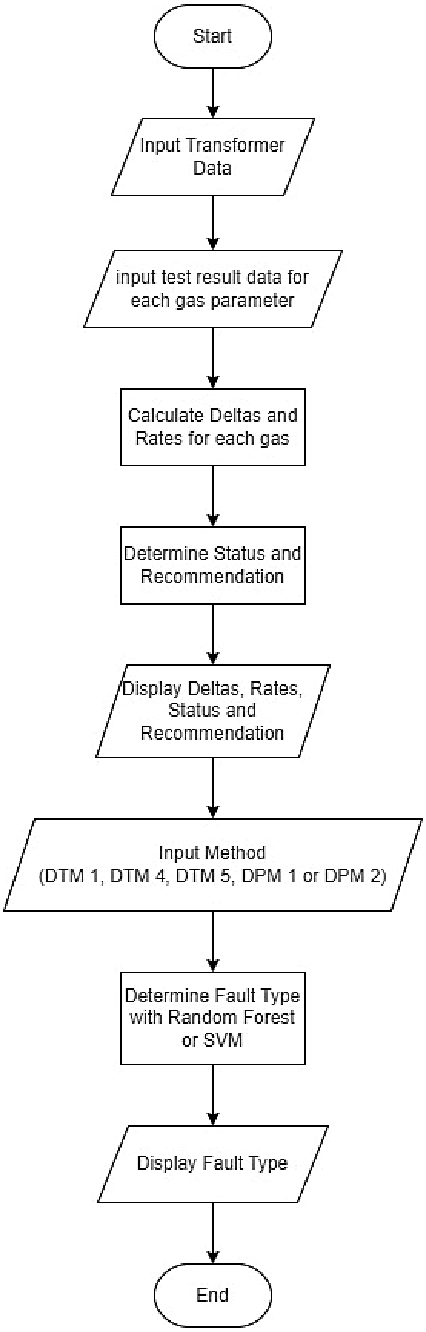

Fig. 7. Flowchart to Determine Fault Type

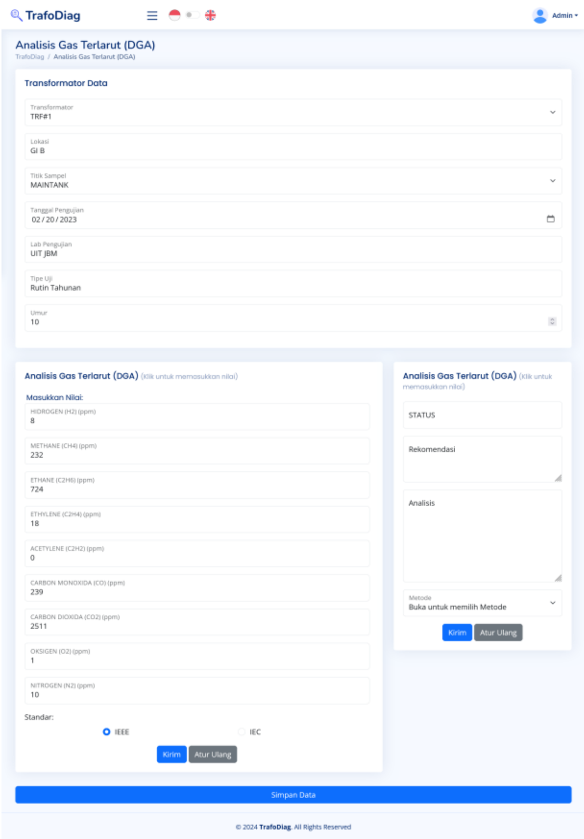

Fig 8 Input Transformer Data and Testing Result

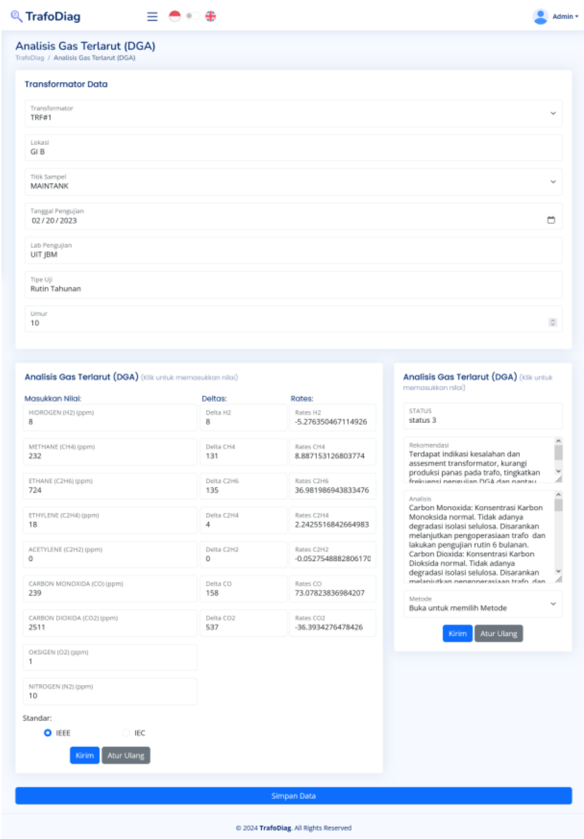

Fig. 9. Calculation Result

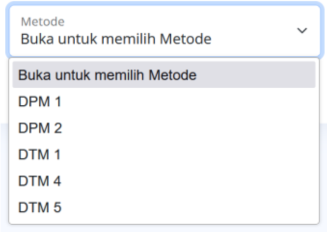

Fig. 10. Selection of Fault Type Determination Method

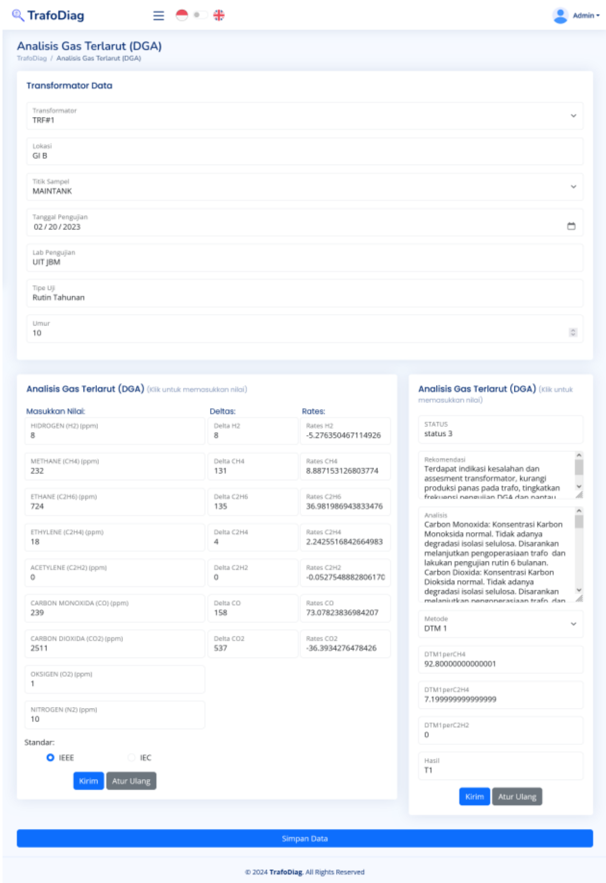

Figure 11 Fault Type Result

4 Conclusion

The research on implementing machine learning in a power transformer maintenance prioritization system concluded that the system effectively supports maintenance by providing action recommendations based on transformer conditions, using calculations from Health Index and Dissolved Gas Analysis (DGA). The study compared Artificial Neural Network (ANN), Random Forest, and Support Vector Machine (SVM) for classifying transformer fault types using data from DTM 1, DTM 4, DTM 5, DPM 1, and DPM 2. Random Forest showed the highest accuracy for DTM datasets, while SVM excelled in DPM datasets. Random Forest accuracy on DTM: 99% and SVM accuracy on DPM: 98%. Significant improvement in accuracy after tuning using Grid Search and Random Search for all models. The system successfully integrates these models, allowing users to select DGA methods for fault type determination, thus offering flexibility to meet various calculation needs. Future improvements can include expanding the system to analyze other electrical equipment and adding more advanced analysis methods.

References

[1] S. Bustamante, M. Manana, A. Arroyo, P. Castro, A. Laso, and R. Martinez, "Dissolved Gas Analysis Equipment for Online Monitoring of Transformer Oil: A Review," Sensors, vol. 19, no. 19, p. 4057, Sep. 2019, doi: 10.3390/s19194057.
[2] E. Ekojono, H. V. Banjarnahor, and E. S. Sintiya, "Multi-expert GMM decision support system with AHP method for determining weight of transformer oil insulation quality index," Proceeding Int. Semin. Sci. Technol., vol. 3, pp. 273–285, Feb. 2024, doi: 10.33830/isst.v3i1.2314.
[3] IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers. doi: 10.1109/IEEESTD.2019.8890040.
[4] E. Ekojono, M. R. T. Rizkia, and E. S. Sintiya, "Multi-criteria decision support system using AHP method to determine the weight of transformer paper insulation index," Proceeding Int. Semin. Sci. Technol., vol. 3, pp. 257–271, Feb. 2024, doi: 10.33830/isst.v3i1.2299.
[5] A. Abdullah, R. Ali, S. Yaacob, T. Mansur, and N. Baharudin, "Prediction of Transformer Health Index Using Condition Situation Monitoring (CSM) Diagnostic Techniques," J. Phys. Conf. Ser., vol. 1878, no. 1, p. 012007, May 2021, doi: 10.1088/1742-6596/1878/1/012007.
[6] W. R. Tamma, R. A. Prasojo, and Suwarno, "High voltage power transformer condition assessment considering the health index value and its decreasing rate," High Volt., vol. 6, no. 2, pp. 314–327, Apr. 2021, doi: 10.1049/hve2.12074.
[7] P. Cole, T. L. Macarthur, and A. Sbravati, "Condition assessment of power transformers".
[8] School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung 40132, Indonesia, R. A. Prasojo, S. Suwarno, and School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung 40132, Indonesia, "Power Transformer Paper Insulation Assessment based on Oil Measurement Data using SVM-Classifier," Int. J. Electr. Eng. Inform., vol. 10, no. 4, pp. 661–673, Dec. 2018, doi: 10.15676/ijeei.2018.10.4.4.
[9] O. E. Gouda, S. H. El-Hoshy, and H. H. E.L.-Tamaly, "Condition assessment of power transformers based on dissolved gas analysis," IET Gener. Transm. Distrib., vol. 13, no. 12, pp. 2299–2310, Jun. 2019, doi: 10.1049/iet-gtd.2018.6168.
[10] R. Furqaranda, A. Bintoro, A. Asri, W. K. A. Al-Ani, and A. Shrestha, "Analysis Oil Condition of Transformer PT-8801-A by Using the Method TDCG, Rogers Ratio, Key Gas, and Duval Triangle: A Case Study at PT. Perta Arun Gas," J. Renew. Energy Electr. Comput. Eng., vol. 2, no. 2, p. 47, Nov. 2022, doi: 10.29103/jreece.v2i2.8567.
[11] W. R. Tamma, R. Azis Prasojo, and S. Suwarno, "Assessment of High Voltage Power Transformer Aging Condition Based on Health Index Value Considering Its Apparent and Actual Age," in 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia: IEEE, Oct. 2020, pp. 292–296. doi: 10.1109/ICITEE49829.2020.9271778.
[12] A. Alqudsi and A. El-Hag, "Application of Machine Learning in Transformer Health Index Prediction," Energies, vol. 12, no. 14, p. 2694, Jul. 2019, doi: 10.3390/en12142694.
[13] B. R. Khalil, "HEART FAILURE PATIENTS ARE CLASSIFIED USING THE RANDOM FOREST AND NAÏVE BAYES ALGORITHMS," J. Northeast. Univ., vol. 26, no. 03, 2023.
[14] P. R. Togatorop, M. Sianturi, D. Simamora, and D. Silaen, "Optimizing Random Forest using Genetic Algorithm for Heart Disease Classification," Lontar Komput. J. Ilm. Teknol. Inf., vol. 13, no. 1, p. 60, Aug. 2022, doi: 10.24843/LKJITI.2022.v13.i01.p06.
[15] M. Mohri, A. Rostamizadeh, and A. Talwalkar, Foundations of machine learning, Second edition. in Adaptive computation and machine learning. Cambridge, Massachusetts London, England: The MIT Press, 2018.