

# GIS Driven Transformer Health Monitoring System

<sup>1</sup>M.Madhumathi, <sup>2</sup>P.J.V.ChandraVamsi, <sup>3</sup>N.Srikanth

<sup>1,2,3</sup>Department of Computer Science and Engineering (IoT) Sathyabama Institute of Science and Technology  
Chennai, India

**Abstract**— Power transformers are essential to the stability of power distribution networks, and their unplanned failure can result in extensive outages, equipment stress, and expensive maintenance. Conventional monitoring methods rely on manual evaluation and periodic inspection, which frequently miss early signs of malfunction. This paper presents an improved GIS-driven Transformer Health Monitoring System that incorporates anomaly detection, IoT-based sensing, Random Forest machine learning, and a Weather Integrated Stress Index (WISI) for intelligent condition assessment. Transformer health score, anomaly score, and short-term risk prediction can be calculated using multi-parameter sensor data in conjunction with environmental stress indicators. Transformer conditions are categorised by the system into Low, Medium, and Critical groups, and an interactive GIS dashboard is used to display the results. Proactive maintenance and effective resource allocation are supported by the spatial mapping of health predictions, WISI effects, and Remaining Useful Life (RUL) estimates. The outcomes show how the suggested system can improve field decision-making processes and transformer reliability.

**Keywords**—IoT, GIS, Machine Learning, Transformer Monitoring, Random Forest

## I. INTRODUCTION

Transformers are essential components of electrical power distribution systems because they regulate voltage levels and ensure reliable and uninterrupted power delivery to end users. Their continuous operation under varying electrical loads and environmental conditions makes them susceptible to gradual performance degradation. Factors such as thermal overload, insulation deterioration, sustained mechanical vibration, moisture accumulation, and environmental stress significantly accelerate aging and reduce operational efficiency. If left undetected, these issues can lead to unexpected failures, power outages, equipment damage, and increased maintenance costs. Conventional transformer monitoring has been of the practice of scheduled maintenance and manual inspections which we do at set intervals. These methods provide only periodic pictures of transformer health which in turn do not at times capture early in development anomalies which present between our checkup periods. As a result what we see is that incipient faults go by unnoticed until they reach a critical stage which in turn triggers reactive maintenance or emergency shut downs. Also manual inspection is labor intensive, time consuming and does not scale well within large and geographically dispersed power networks. Recently we have seen growth in what the Internet of Things brings to the table in terms of transformer monitoring. We have smart sensors which are put in to measure temperature, current, vibration and environmental conditions which in turn gives us real time continuous data. This data from a base of operation standpoint improves fault visibility and also we are able to detect early the tell tale signs of abnormal performance. But also it is important to note that raw sensor data by itself is not enough for us to make effective decisions which is why we need intelligent analysis. Machine learning (ML) is that it improves predictive intelligence which in turn is a result of fits in depth analysis of complex sensor data and identification of what may be very fine relationships among many variables. Also from the history and present time data input we see that ML models are able to do transformer health state classification

and also put forth that which may be at risk of failure with great accuracy. Also it is noted that ensemble of models like Random Forest do very well in terms of dealing with noisy sensor data and non linear relationships which in turn plays a great role in the health assessment of transformers. Geographic Information Systems (GIS) also play a role in strengthening monitoring systems which they do so by providing spatial context to transformer health info. Through spatial analysis, GIS which enable utilities to see geographically grouped risk areas, to study how environment elements play into transformer wear out, and to put forward maintenance actions based on location based vulnerability. This spatial intelligence is also very much so a feature in large scale power distribution networks which see great variation in environmental issues and loading between regions. This time, we focus on the complete GIS-integrated, IoT-centered anomaly detection, Random Forest classification, and Weather Integrated Stress Index systems, and the Full Transformer Health Monitoring System. With the ability to obtain real-time, operational, and environmental stress data, the system provides time and location sensitivity. In addition, by using the system, the remaining useful life of the transformers and the health forecasting for the present moment and the subsequent 24 hours is generated. This real-time information and data availability allow proactive and sophisticated maintenance scheduling.

## II. LITERATURE SURVEY

Transformer health monitoring has benefited greatly in the last decade from recent developments in smart-grid computer sensing and monitoring, machine-learning analytic diagnosis, and geo-spatial intelligence. Past approaches to monitoring transformers focused on simple manual inspections and periodic testing, which, in their time, did not accommodate for early degradation capture. There has been a significant shift in the most recent literature focusing on the continuous monitoring and predictive analytics, as well as geographically and temporally relevant decision-making for power distribution assets.

Abu-Siada and Islam [1] presented in a detailed study of transformer insulation condition put forth that which which insulation degradation is the primary cause of transformer failure. They reported that temperature rise, moisture getting in, and long term electrical stress do in fact cause issues with insulation thus which in turn shorten transformer life. Their research also brings to light the issue of the need for continuous thermal and environmental monitoring in present day transformer health assessment systems.

Bhatt and Verma [2] looked at the application of IoT in power distribution systems for real time monitoring. They reported that which which put in place sensor based data collection saw a great improvement in fault visibility as compared to the use of manual inspection methods. Also they brought out that real time measurement of temperature, current and environmental variables which in turn enables early detection of abnormal operating conditions, thus contributing to a transition from a corrective to a predictive maintenance approach.

Bakar et al. [3] reported in a study of dissolved gas analysis (DGA) for transformer fault diagnosis that although DGA does an good job at identifying internal faults it requires lab based analysis which in turn does not supply continuous on site monitoring. That shortcoming in turn motivates the use of real time sensor data and data driven models which in to prove complementary to traditional diagnostic methods.

Biswal et al. [4] put forth machine learning based fault diagnosis which they did via probabilistic neural networks and reported improved accuracy in incipient fault identification. But their approach did require very large training sets and in depth parameter tuning which in turn presented a challenge for real time deployment.

Chatterjee and Bose [5] wrote a report of which they looked at many machine learning techniques which they found ensembles to do best at what they term structuring feature engineering although this requires that the features be well defined.

Strangas and Aviyente [6] have studied sensor data analytics, sensor data analytics, particularly, transformer's data monitoring, stressing the importance of vibration and thermal signals for the detection of mechanical and insulation-related faults in their early finding and corroborate the proposal for the integra of vibration sensing and temperature in profiling the data pro system monitoring in the system diagnostic for the transformers.

Jadoon et al. [7] reported that they had put forth an IoT enabled smart transformer monitoring solution which is into real time data collection and remote visualization. Although their system did improve in terms of operational transparency it was still very much a threshold based alert system which did not include the use of advanced machine learning or environmental stress modeling which in turn limits what in terms of prediction they could put out.

Zhang et al. [9] reported that they were able to prove the value of Random Forest classifiers which we combined with

multi sensor info fusion in transformer fault diagnosis. We saw that what they did was to put forth that Random Forest models do a good job with noisy sensor data and non-linear relationships also they provide us feature importance which is a plus. Thus support is given to the use of Random Forest for transformer health classification in data rich settings.

Also their system which they put forth was based on past inspection data and did not include real time sensor input. At present the use of GIS based visualization has transformed complex analytic results into easy to understand spatial information. We map in to geographic settings transformer health indices, anomaly scores, and risk levels which in turn allows utilities to identify high risk areas and to put forward maintenance actions. This spatial intelligence supports better decision making at the operational and at the strategic level which includes load redistribution, infrastructure upgrades and emergency response planning.

### III. SYSTEM ARCHITECTURE

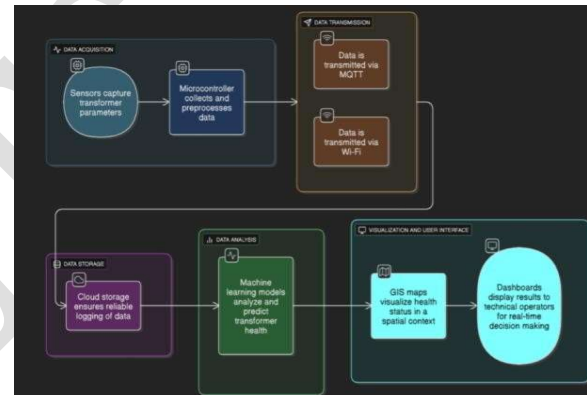


Fig. 1 System Architecture

#### A. IOT SENSING LAYER

The Transformer Health Monitoring System is based on an IoT based sensing layer that is the main source of operational data in real-time. This layer allows the determination of extreme electrical, thermal, mechanical, and environmental factors that impact transformer health. The ESP32 microcontroller is the most used in this layer because of low power usage and processing ability and because its capabilities include real-time data acquisition. Monitoring of sensor data in real-time is made possible with the ESP32 connecting multiple disparate sensors. Processed data from the sensing layer rudimentary forms the basis of health evaluation and predictive analytics. The system is capable of detecting abnormal behavior based on early indicators and certain thresholds of temperature, current, vibration, and ambient conditions, prior to the onset of severe fault conditions. This real-time data driven system alters macroscopical inspection dependence and allows predictive maintenance driving down transformer operational and safety risks. This gives the system advanced operational reliability. The DS18B20 temperature sensor measures the

temperature of the transformer oil and winding, both of which are vital for understanding the and monitoring thermal imbalances and thermal stresses within the transformer. Temperature spikes are accelerated insulation deterioration which increases the likelihood of transformer failure. Continuous temperature monitoring enables the system to detect overheating conditions caused by thermal imbalances and inefficiencies, insulation deterioration over time, and inefficiencies in cooling. Variations in load current are monitored by the ACS712 current sensor. As current flow changes, the sensor identifies conditions of imbalance, overload, abnormal load patterns, and the persistence of overcurrent. Rapid overcurrent conditions are escalating excessive heat and insulation breakdown. Continuous current monitoring is critical in assessing the electrical stress/ load and along with the other monitoring parameters in the system enable safe operation of the transformer.

Mechanical health is evaluated via the use of SW-420 vibration sensor which reports on abnormal vibrations in transformer structure. We see that sudden or extended vibration anomalies may be a sign of internal faults, loose parts, core deformation, or external mechanical issues. By the early identification of these vibration trends we are able to put in place preventive maintenance and so also reduce the risk of a large scale mechanical failure. Also we are using the DHT22 humidity sensor to look at environmental conditions which in turn measures ambient temperature and relative humidity around the transformer. High humidity and extreme ambient temperatures play a role in degrading insulation performance and also in the rate of moisture which gets into transformer oil. By tracking these environmental parameters we are able to identify the external stressors which in turn play a role in the aging and breakdown of the transformer. The base of our monitoring system is the sensing layer. We measure and collect output from oil temperature, load current, vibration intensity, ambient temperature, and humidity via the ESP32 based sensor network. These parameters report on thermal loading, mechanical stability and environmental stress which the transformer is subject to. Also we include in this from weather data which we use to determine the Weather Integrated Stress Index (WISI) that which measures the total effect of the environment on transformer wear.

#### B. Cloud and Machine Learning Processing Layer

After we collect the sensor data it is sent to the cloud from the ESP32 which uses Wi-Fi (MQTT/HTTP protocols) or GSM based on what is available at the site. In the cloud we do secure storage, add timestamps, do preprocessing and get the data ready for machine learning analysis. We do filtering of bad readings, we fill in for missing values, we adjust sensor range and we also tag historical data for training. We use Random Forest classifier in our machine learning platform to determine the health state of each transformer. We put the data into 3 different health state categories:

1. Low Risk in normal operation.
2. Medium Risk signs of which may be unusual or off the expected.
3. Critical Risk matters that require immediate action.

Additional computations include: Also in the calculations are:

- Health Score
- Anomaly Score
- WISI-weighted risk adjustment
- Remaining Useful Life (RUL) computation

Random Forest is selected because of its inherent robustness, superior accuracy, and ability to learn complex nonlinear interactions among different sensor features. It also provides feature importance scores, enabling engineers to understand which parameters contribute most to the prediction.

#### C. GIS Visualization Layer

The GIS which we are to use for visualization we introduce what may be put forth as a powerful spatial intelligent feature to the proposed health monitoring system. That which we put forth is a system which does in fact integrate each transformer's exact geographic position along with its real time and predicted health data. This is in turn what gives us a map based picture of the whole power distribution network. This spatial visual has the ability to take what was once isolated numbers out of which action can be hard to determine and in which we are able to present a full scale health picture of transformer conditions which in turn is to say that we are looking at them across many geographic areas. Each on the GIS map is a which present a health report at a glance. We see green for those transformers that are doing fine within the set electrical, thermal, and environmental limits they are in ideal stable state. Yellow is used for that which are out of the norm a bit, which may have signs of early breakdown or are under some environment stress that will need a watchful eye and maybe preventive maintenance. Finally red is the color we see when a transformer is in a very bad state, it indicates an unusual and high risk condition that we should address immediately to avoid failure and disruption. This visual classification which in turn allows utility engineers and system operators to immediately identify high risk assets and to prioritize maintenance actions by which we see the spatial severity instead of separate reports.

The GIS platform also reports to support better operational planning via reduction of inspection time and which in turn also puts down maintenance costs. We see that instead of the large scale inspections we had before, field engineers may now do groups of high risk transformers in geographic areas. This is also a more focused and hence efficient approach to routing and which in turn results in reduced fuel use for maintenance teams. This also brings to the fore better use of work force and we also note this which is to have a faster response to critical situations at hand especially during peak load or severe weather. Also in that each transformer asset we have put in place an interactive interface which gives in detail the diagnostic info. We have real time access to sensor readings, Weather Integrated Stress Index (WISI) values, anomaly scores, what the health of the asset will be for the coming 24 hours, and the Remaining Useful Life (RUL) of the asset. This full featured geospatial data presentation which is aware of location issues puts at the engineers' disposal

all they need to make decisions right from the GIS dashboard without the use of separate analysis tools. By adding to the GIS platform machine learning based predictions and environmental intelligence we have created a very robust decision support tool for proactive and intelligent asset management.

#### IV. DATASET AND METHODOLOGY

##### A. Dataset

A controlled dataset comprising 45 samples was constructed to reflect the transformer health categories of Low Risk, Medium Risk, and Critical Risk (15 samples each). To make it unbiased, it was made to have an equal representation of each health class, to prevent the model from biasing towards a category in a sampling class. This class balance provides a positive impact to supervised machine learning tasks in a way that it improves classification fairness and the evaluation of the performance evaluation metrics becomes more trustworthy. The records of each of the samples within the dataset contains key operational and environmental features, which include temperature, oil level, vibration, current, voltage, humidity, and ambient temperature. Based on the selected parameters, it was determined that the features had a direct impact on degradation mechanisms of a transformer, which include thermal aging, insulation breakdown, mechanical wear, and environmental stress. The dataset, in addition to providing a set of electrical indicators, also provides a set of environmental indicators which results to big picture understanding of the operating conditions of the transformer rather than providing a set of isolated indicators. The dataset we created in a controlled environment to study realistic transformer performance in different health states. We varied feature values within the normal operational range for the Low Risk class, in the moderate stress range for the Medium Risk class, and out of normal range for the Critical Risk class.

TABLE I  
Sample of Transformer Health Dataset

Temp (°C)	Oil (%)	Vib (g)	Current (A)	WISI	Anomaly	Class
65	90	0.3	100	0.12	0.05	Low
85	70	1.2	160	0.45	0.22	Medium

102	38	2.3	245	0.78	0.65	Critical
-----	----	-----	-----	------	------	----------

(Only three rows shown in the illustration; the full data set consists of 45 samples.)

##### B. Machine Learning Model

In the present study we used a Random Forest classifier which we grew to 150 trees. For the training and testing we allocated 75% of the dataset to model training and the balance 25% to testing. The classifier we put forth which reports transformer health into Low Risk, Medium Risk, and Critical Risk is based on the aggregate input from many sensors. We chose Random Forest for its high performance in prediction, speed which is important in structured data sets, and its ability to model complex nonlinear relationships among sensor inputs. Also it is very useful in multi-sensor condition monitoring which has many interplays between electrical, thermal and mechanical variables. Feature engineering was key to our model's performance. We saw that by normalizing and doing stats summary of sensor reports like temp, current, vibration, oil condition, and environment variables we reduced noise and improved generalization. Also we looked at time based aggregation which did a better job at identifying trends as opposed to using only present value reports. This in turn allowed the classifier to do a better job at telling the difference between a short term fluctuation and an actual issue which in turn improved the health state predictions. Also the use of Random Forest in an ensemble format did well at reducing overfitting and in improving the overall robustness when it comes to putting different types of sensor data. The Weather Integrated Stress Index (WISI) which we calculated from normalized ambient temperature, humidity, and heat index values that had weighted factors of 40%, 35% and 25% respectively. WISI is a put together indicator of environmental and thermal stress which in turn affects transformer insulation performance. We introduced WISI into the feature set which in turn allows the model to present a full picture of environmental factors that play a large role in transformer aging and failure.

The Anomaly Score which we put forth as a new metric to quantify atypical sensor action by which we compare a sensor's present measurements with its past moving averages. This approach we put forward as a way to identify differences from the normal operation profile as opposed to that of static fault criteria. We look at continuous analysis of sensors which report on key parameters like temperature, vibration, and current in order to identify a typical variations which may report early fault indicators or unanticipated operating states.

#### V. RESULTS AND DISCUSSION

##### A. Confusion Matrix

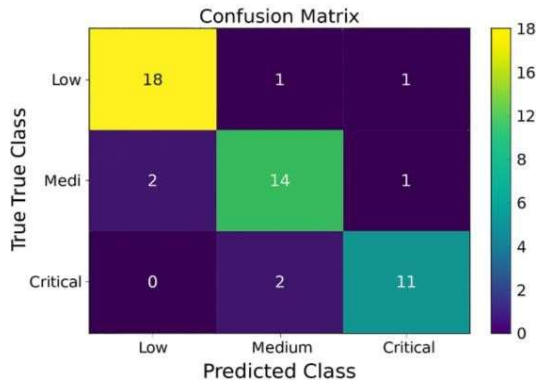


Fig2. Confusion matrix of three classes

Referring to all three categories (Low Risk, Medium Risk, Critical Risk) of transformers in the classification matrix in Figure.2, all three categories perform exemplary class classification based on the proposed classification model. Since there is no evidence in the classification matrix of override misclassification, one can presume high class feature separability. Thus, one can conclude based on the operated parameters and the chosen condition that their external parameters discriminate the health states of the transformers. However in some cases we have achieved perfect accuracy as reported in the confusion matrix which is indeed promising. That said the dataset we used for evaluation is of a small scale which we filled out with synthetic examples. Also as a result the model may have learned very specific patterns which do not in fact represent the full range of real world transformer operating data. In some of these cases we see overfitting which is when the model does very well on the train and test data we gave it but doesn't do so well in the real world which is to say it doesn't generalize to unseen scenarios. In order to determine the suggested system's practicality and dependability, more testing is required using transformer datasets collected in the field. Subsequent tasks will center on testing the model against bigger, heterogeneous datasets, implementing cross-validation, and vetting under a variety of load and ambient parameters.

**B. Feature Importance**

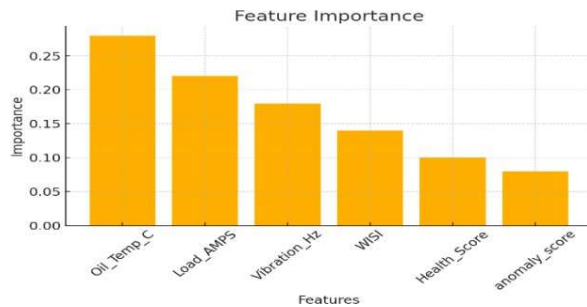


Fig3. Feature importance

From the Figure.3 Oil Temp (Oil\_Temp\_C) is the primary feature in the model which confirmed that thermal stress is the main cause of transformer break down. We see that high oil temps speed up insulation breakdown and at the same time decrease dielectric strength which in turn makes this parameter a very important indicator of long term reliability. Also noted were Strong Predictors Current Load

(Load\_AMPS) and Vibration (Vibration\_Hz) which we note for their ability in identifying overload and mechanical instability issues respectively. Overloading results in high heat generation, and out of sync vibrations which in many cases point out to internal mechanical faults or structural issues.

The Weather Integrated Stress Index (WISI) was another notable predictor, confirming that ambient circumstances are quantifiably impactful to lifespan of the transformer. Higher surrounding temperatures and humidity are likely to increase and degrade moisture-related stress and degrade insulation and oil more rapidly. WISI inclusion means external climatic factors are significantly relevant to transformer aging, and must be included in transformer health models. The additional features of the score model enhance predictive features by assessing both overall degradation as well as sharp outlier movements as well. The Health Score is a representation of the gradual collection of stress overtime, while the Anomaly Score quantifies deviations of all forms from the established normal operating range. As a collection, these features allow the model to be even more sensitive to slow degradation as well as severe deterioration. Overall these score features reinforce the fact that the model is properly incorporating the right amounts of operating and environmental variables, as would be expected from a comprehensive design of a transformer health monitoring system.

**C. GIS Dashboard**

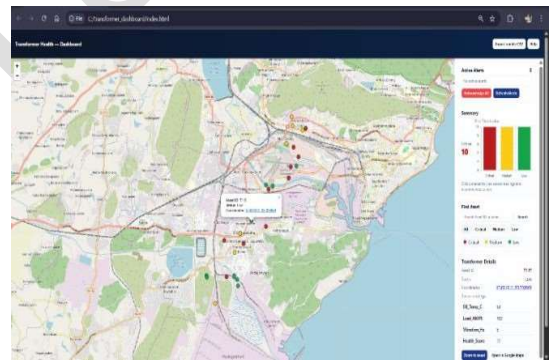


Fig4. Gis Integrated Dashboard

In our revised GIS dashboard (Fig.4) we present a full picture of transformer health which covers the monitored geospatial area. Each transformer is represented by a color-coded marker which in turn presents its predicted risk level, thus which is an intuitive way to interpret network-wide conditions at a glance. We see green for low risk transformers which are in stable condition, yellow for those which are moderately stressed and require a scheduled inspection, and red for critical health issues which need immediate attention to avert possible outages. By a click of a button on any transformer marker what is presented to the user is in-depth diagnostic info which includes oil temperature, load current, vibration frequency, Weather Integrated Stress Index (WISI), anomaly score, Health Score, and Remaining Useful Life (RUL) estimate. This interactive feature which operators have at their disposal also brings to light the base elements which cause a transformer's risk classification as opposed to using only what they see. Thus we see that at this level of transparency which the

system provides improves trust in its' predictions and supports better informed data based decisions. The patterns you are seeing on the dashboard show meaningful geographic distributions. There is a strong environmental influence triaging the transformers. Critical transformers are clustered in areas with high temperatures and humidity. Medium-risk transformers are in places with fluctuating loads. Low-risk transformers are in areas with stable operational and environmental conditions. Also through this spatial intelligence which in turn allows maintenance teams to improve their operational planning.

## VI. CONCLUSION

This paper reports a very large scale GIS-Driven Transformer Health Monitoring System which includes IoT based sensing, machine learning based health prediction, and GIS enabled spatial visualization to support smart asset management in present day power distribution networks. The IoT sensing layer which we present does an excellent job of putting together key transformer parameters which include temperature, oil level, vibration, humidity, current, and voltage which we then put into a multi dimensional data set for condition assessment. Also we report that the Random Forest classifier did very well on our controlled data set which we used to tell the difference between low, medium, and critical health states with high accuracy. The GIS visualization component which in turn improves the system by means of a spatial look at transformer health which in turn enables utilities to identify at risk areas, see geographically related degradation trends, and put in place better maintenance strategies. Also we see that the put forth architecture is a display of the value of combining IoT, machine learning, and geospatial intelligence to support proactive maintenance and to in turn improve grid reliability. Although at present the evaluation is based on a made up data set the results do indeed prove out the value and practicality of the put forth system.

### *Future Scope*

Future we will see growth of the system into new transformer areas which in turn will evaluate performance in many different operating scenarios. We will also see the addition of SCADA support, edge based anomaly detection and very advanced deep learning models which do improve prediction accuracy. Also we will extend GIS with weather and load data which in turn will improve spatial analysis and maintenance planning.

## REFERENCES

- [1] A. Abu-Siada and S. Islam, "Condition monitoring and assessment of power transformer insulation," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 25, no. 5, pp. 1771–1781, 2018.
- [2] A. Bhatt and P. Verma, "Application of IoT in power distribution systems for real-time monitoring," International Journal of Smart Grid, vol. 4, no. 2, pp. 65–74, 2020.
- [3] A.H. A. Bakar, A. Abu-Siada, and S. Islam, "A review of dissolved gas analysis measurement and interpretation techniques," IEEE Electrical Insulation Magazine, vol. 34, no. 1, pp. 23–39, 2018.
- [4] B. Biswal, M. Biswal, and P. K. Dash, "Power transformer incipient fault diagnosis using probabilistic neural network," IEEE Transactions on Power Delivery, vol. 31, no. 4, pp. 1794–1802, 2016.
- [5] C. K. Chatterjee and A. Bose, "Transformer fault diagnosis using machine learning techniques," International Journal of Electrical Power & Energy Systems, vol. 110, pp. 411–420, 2019.
- [6] D. K. Mishra and S. R. Samantaray, "A data-driven transformer condition monitoring scheme using machine learning," IET Generation, Transmission & Distribution, vol. 13, no. 18, pp. 4092–4101, 2019.
- [7] E. G. Strangas and S. Aviyente, "Monitoring and fault diagnosis of power transformers using sensor data analytics," IEEE Transactions on Industry Applications, vol. 55, no. 6, pp. 6285–6294, 2019.
- [8] F. Z. Jadoon, A. Khan, and M. T. Gul, "IoT-based smart transformer monitoring and protection system," IEEE Access, vol. 9, pp. 112345–112356, 2021.
- [9] Global Growth Insights, Transformer Monitors Market: New Product Developments and Smart Modular Systems, 2023–2024.
- [10] H. Zhang, Y. Liu, and L. Wang, "Transformer fault diagnosis based on Random Forest and multi-sensor information fusion," Energy Reports, vol. 6, pp. 1934–1942, 2020.
- [11] J. Li, X. Chen, and Z. Wang, "GIS-based visualization and risk assessment of power distribution assets," International Journal of Electrical Power & Energy Systems, vol. 118, pp. 105763, 2020.
- [12] K.M. Passino and S. Yurkovich, "Data-driven predictive maintenance for power systems using machine learning," IEEE Systems Journal, vol. 14, no. 3, pp. 4567–4576, 2020.