iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 Issue: VI Month of publication: June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44578

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Profiling the Transformer Degradation Using Deep Learning

Dr. P. B. Manoj[1], Monika A[2], Abhay Kumar[3], Abhay Pratap Singh, Ritesh Kumar[5]
[1]Professor Department Of ECE, AMCEC, Bengaluru
[2, 3, 4, 5]ECE Department of AMC Engineering College

Abstract: This paper presents a detailed analysis of Transformer failure reasons and methods in conjunction with a real time data of the electrical transmission and distribution system to find the reasons and its remedies for better capacity utilization and reliability. The analysis is carried out using data collected from one zone of a city based substation chescom. The analysis is carried out in accordance with IEEE standards and summarized to present the reasons and possible remedial measures for prevention of transformer failures.
Keywords: Transformer failure prediction, data mining, systematic literature review, data mining algorithms, classification.

I. INTRODUCTION

Transformer is a static machine with very high efficiency and rugged construction. The rate of failure of distribution transformers in India is higher (12-17%) as compared to developed countries (2-3%). This high failure rate is cause of concern to all the Distribution Companies (Discoms) in the country. Every year, nearly 200 Crore of Indian Rupees (INR) are spent by the Discoms for repair and replacement of distribution transformers [1].

The loss becomes enormous due to the transformer failure, if the revenue loss for supply outage is also taken into consideration. It has become a serious problem

due to increasing failure rate every year [1-8]. The role of transformer begins at generating station as the power is generated at maximum value of 11 KV in generating stations in India, far away from the load centres. This power needs stepping up to extra high voltages for reduction of current thereby the losses during transmission. Thereafter it is stepped down to 66/11KV at substations for primary distribution network and further stepped down to 11KV/400V using distribution transformers for secondary distribution system to feed consumers of different categories e.g. domestic, commercial etc.

Distribution system consists of 11KV feeders; Distribution transformers and low tension (LT) three phase 4 wire systems. So the distribution transformer is a most important component of the distributiion system to provide uninterrupted power supply to the consumers and it should be highly reliable and efficient. The transformer failures result in loss, not only on account of repair or replacement of failed transformer, but also, the revenue loss to the utility on account of power not supplied to the consumers. Other important disadvantage is reduced reliability of the system, because of frequent failure of power supply. The risk of failure is defined as the product of probability of failure and consequences [2]. To improve the reliability of the system and to reduce the risk.

II. TRANSFORMER FAILURE RELATED PARAMETER

A. Aging factor

To assess the state of a transformer, each component would have to be analysed separately. However, the main question remains whether a transformer can perform normally if one of the components is faulty. It is reasonable to assume that this would depend on the state of the individual part. To assess the state of the whole component, it would be beneficial to go with the odds of which part tends to fail most often. Environmental factors used to have a major contribution to a transformer's aging process. Heat and humidity were the main reasons for the degradation of the insulation layer and caused corrosion to the oil tank. With power electronics being introduced into the grid, repetitive transients are becoming an influential player in shortening of the transformer's life span. This paper will discuss the major aging factors affecting the transformer and it will also propose how a predictive health model can be applied to judge the state of this very important component.

B. Acoustic factor

The primary source of acoustic noise generation in a transformer is the periodic mechanical deformation of the transformer core and the winding coils. This is due to the influence of fluctuating electromagnetic flux associated with these parts. The flow line of the acoustic energy generated in the transformer is represented in the figure.1. The propagation of noise from the core and windings through the oil medium is spread to external environment .The spectrum of noise can be captured by the advanced noise level measuring instruments. The physical phenomena associated with this noise generation can be briefed as follows:

1) The material of a transformer core exhibits magnetostrictive properties. The vibration of the core is due to its magnetostrictive strain varying at twice the frequency of the alternating magnetic flux. The frequencies of the magnetic flux are equal to the power system supply frequency and its harmonics.
2) When there are residual gaps between laminations of the core, the periodic magneto-motive force may cause the core laminations to strike against each other and produce noise. Also, the periodic mutual forces between the current-carrying coil windings can induce vibrations if there are any loose turns of the coil.

C. Temperature Factor

Temperature is one of the prime factors that affect a transformer's life. In fact, increased temperature is the major cause of reduced transformer life. Further, the cause of most transformer failures is a breakdown of the insulation system, so anything that adversely affects the insulating properties inside the transformer reduces transformer life. Such things as overloading the transformer, moisture in the transformer, poor quality oil or insulating paper, and extreme temperatures affect the insulating properties of the transformer. Most transformers are designed to operate for a minimum of 20-30 years at the nameplate load, if properly sized, installed and maintained. Transformers loaded above the nameplate rating over an extended period of time may have reduced life expectancy.

D. Weather Factor

Temperature, pressure, moisture content, speed and direction in which its moves. One factor is that the sun does not heat all parts of the earth equally; the resulting difference in density and pressure causes the air to move from different places. The ambient temperature plays an important role for the determination of the hot spot temperature of a trans-former determined by its loading profile. The hot spot temperature increases linearly by increasing the ambient temperature. Roughly, for every 1°C increment in ambient temperature, the loading capacity can be decreased by 1% without any loss-of-life and vice versa. On average, the ambient temperature is increasing to some extent every year due to the effect of overall global warming. This has been realized by monitoring ambient temperatures over the last several years. The IEC loading guide is recommended for transformers operating with an ambient temperature of 20°C. However, it does not give the actual hot spot temperature at varying ambient temperatures. Table 2 gives the future ambient temperatures of different locations in Finland which is based on weather forecasts .

E. Load Factor

It is, of course, obvious that a trans- former operating under low-load-factor conditions can carry some overload under peak load conditions with the same degree of aging of its insulation as one carrying rated load continuously. The problem is to determine the amount of overload that will give the same aging, for different load factors. To do this requires: (l) determination of the hottest-spot temperature during one complete 24-hour period, and (2) integration of the hottest-spot temperature curve during the 24-hour period, to obtain the amount of aging of the insulation. When an overload is found (by trial) to cause the same amount of aging as occurs when the transformer is carrying rated load for 24 hours, this represents the permissible overload for that particular kind of load curve.

F. Region factor

This geographical region contains the constant demand to update and upgrade the existing transmission for the transformer monitoring system Market Revenue. The region plays an important role in the constant demand for updating and up-gradation of the existing transmission and distribution of infrastructure for balancing or maintaining electricity demand.

III. PREPARTION FOR MODEL CONSTRUCTION

The dataset consists of information related to 1000 transformers. Dataset have 10 features out of which first 9 Represents the Transformer_id, Region_factor, weather_factor, Temperature_factor, acoustic_factor, load_factor, Aging_factor, 3m_fail_factor,6m_fail_factor 9th represent the actual status of transformer at the time of sample collection. In this article all factors were considered for the analysis based on data availability[5]. There are 8 types of fault and as 9th category for constructing the model.

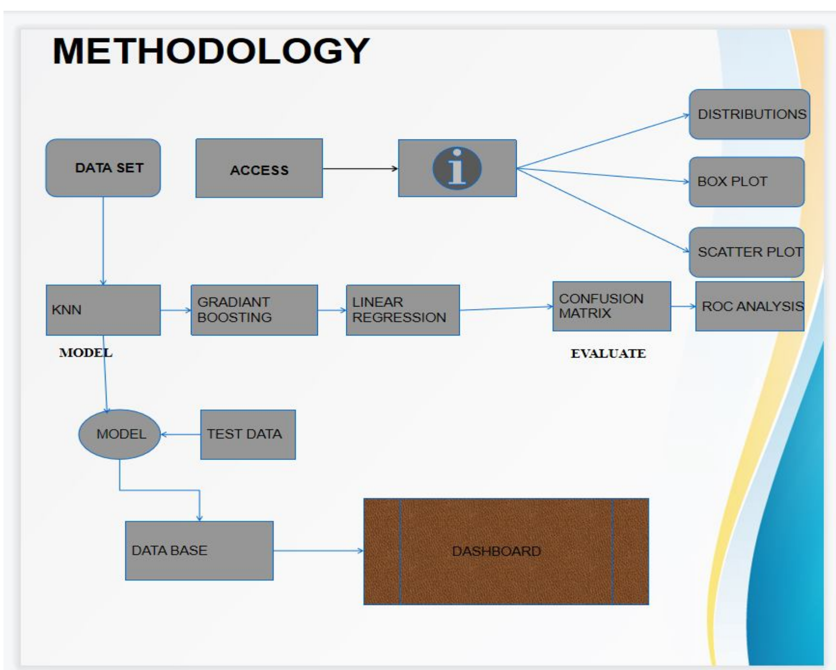
A. Data Pre-Processing

Data preprocessing is a vital part of training a machine learning model as the state of data can directly affect the learning. Mostly data is in drawn from various sources which makes it non uniform and ambiguous. So, it is important to preprocess data before feeding it to a machine leaning model[7].If the data that we are using is having some inadequate or irrelevant information, then the model may present less accurate results, or may fail to discover anything of use at all. Thus, data pre-processing is an important step machine learning. The pre-processing step is used to resolve several types of problems such as noise in data, redundancy data, missing values in data etc. All the machine learning algorithms rely heavily on the product of data-prepossessing, which is the final training set [8].

Data pre-processing includes:

1) Loading Dataset.
2) Handling missing data.
3) Handling text labels.
4) Separating dependent and independent variables.
5) Handling data with categories.
6) Normalizing the data.
7) Splitting training and testing data

Dataset is loaded and is divided into the input features(X) and output features(y). Spitting is very important in order to validate the performance of the machine leaning model. The consider data is normalized and it will be split into training and testing data [10]. The training data is selected from the splitted dataset to feed the classifier model



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

IV. RELATED WORK

A. Artificial Neural Networks (ANN)

The concept of ANN is basically introduced from the subject of biology where neural network plays an important role in human brain [23]. neural network is just a web of inter connected neurons. in the neural network, the most basic information processing unit is the neuron. they are organized in three or more layers, such as the input layer, one or several hidden layers and single output layer [37]. ANN method can be used to recognize the hidden relationships between the dissolved gases and the fault types through training process.

Ann was introduced by [26] in the power systems fault detection in 1996. reference [22] used ANN to predict the lightning surge and it was found that ANN can be used directly to assess a certain network's risk of failure or indirectly to determine which type of lightning arr-ester, maximum cable length, or cable insulation level should be used to meet a certain risk of failure. based on the study conducted by , ANN’s successful development was able to predict the incipient fault in the power transformer using dissolved gas concentration.

B. Decision Tree (DT)

The decision tree is a tree structure, which is mainly composed of nodes and branches, and the nodes contain leaf nodes and intermediate nodes. The intermediate nodes are used to represent a feature, and leaf nodes are used to represent a class label.The attributes that appear in the decision tree nodes provide important information to promote classification. Decision tree (dt) based artificial intelligence techniques were already used in transformer fault prediction [29–31].

These studies demonstrate the effectiveness of the dt based algorithms moreover, dt based techniques were also used for condition monitoring, assessment, fault diagnosis and repair actions [32, 33]. these studies prove the efficiency of the dt technique in similar areas. reference [27] applied dt for feature selection and then carried on the bearing fault diagnosis with the kernel neighborhood fractional multiple support vector machine (msvm).

In na study pertaining to fault detection in three phase transformer, [28] explored decision tree using differential protection scheme. based on the test results, it was found that the decision tree method is the best in classifying fault prediction with higher sensitivity and accuracy as compared with linear model.

C. Naïve Bayes (NB)

Bayesian classifier is a statistical classifier and supervised learning technique. It predicts class membership probabilities. reference [41] mentioned that nb classifiers show high accuracy and speed when applied to large databases. according to references [42, 43], nb uses not only a small amount of training data, but also works at simple structure, fast calculation speed and high accuracy. reference [36] used nb classifier for the transformer fault analysis the different testing scenarios showed that the nb diagnosis model constructed has a good performance given complete testing data in a study carried out by [34], few algorithms were used namely naïve bayes, random forest, j48, bagging, ibk (knn in weka tool) and logistic regression. It was found that the random forest algorithm performs better than naïve bayes and naïve bayes has an advantage over ibk. in another study, [35] employed knn and naïve bayes classifiers for the diagnosis of insulating oil used in power transformers. based on the evaluation using duval triangle reports, it was found that the knn algorithm provides the highest accuracy rate than the naïve bayes algorithm.

D. Gradient Boosting

Gradient boosting is a method standing out for its prediction speed and accuracy, particularly with large and complex datasets. From kaggle competitions to machine learning solutions for business, this algorithm has produced the best results. We already know that errors play a major role in any machine learning algorithm there are mainly two types of error, bias error and variance error. Gradient boost algorithm helps us minimize bias error of the model before getting into the details of this algorithm we must have some knowledge about adaboost algorithm which is again a boosting method. this algorithm starts by building a decision stump and then assigning equal weights to all the data points, then it increases the weights for all the points which are misclassified and lowers the weight for those that are easy to classify or are correctly classified. A new decision stump is made for these weighted data points. the idea behind this is to improve the predictions made by the first stump.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

E. Random Forest

Random forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both classification and regression problems in ml. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.As the name suggests, "Random forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.

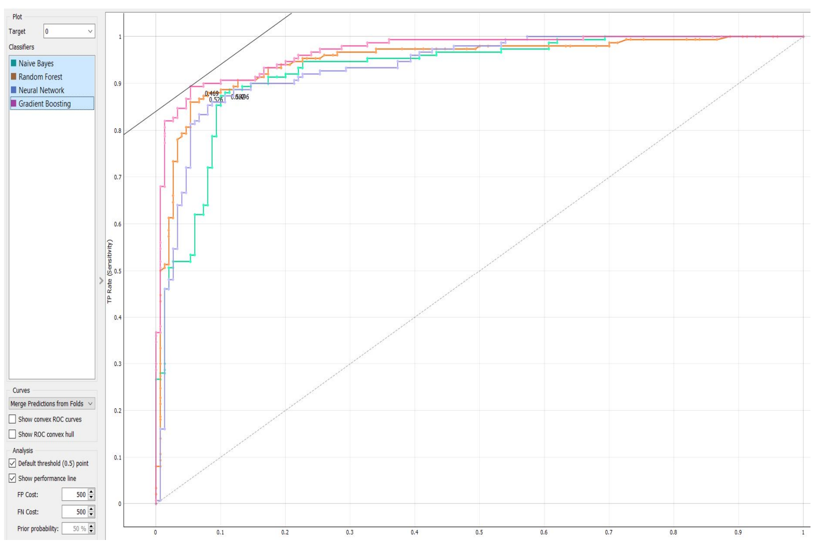
F. Confusion Matrix

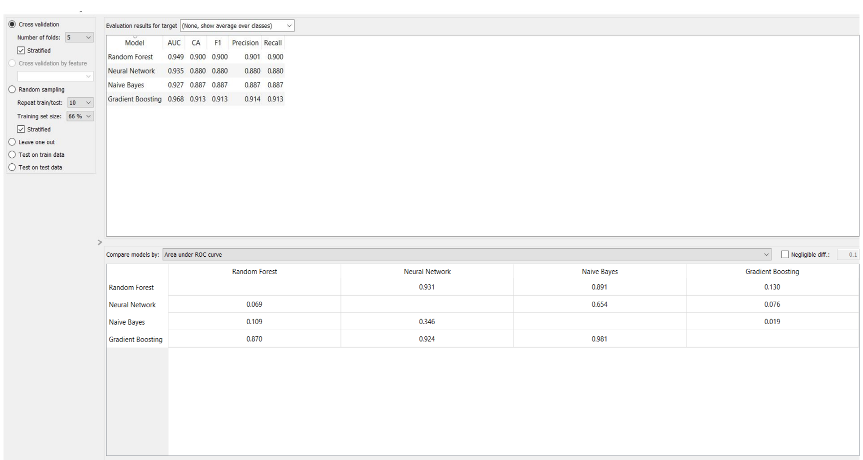
The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known. The matrix itself can be easily understood, but the related terminologies may be confusing. Since it shows the errors in the model performance in the form of a matrix, hence also known as an error matrix.

G. The Receiver Operator Characteristic (ROC)

Curve is an evaluation metric for binary classification problems. It is a probability curve that plots the tpr against fpr at various threshold values and essentially separates the 'signal' from the 'noise'. The area under the curve (auc) is the measure of the ability of a classifier to distinguish between classes and is used as a summary of the roc curve.

V. EXPECTED RESULTS





Evaluation results for target (None, show average over classes)

Model	AUC	CA	F1	Precision	Recall
Random Forest	0.949	0.900	0.900	0.901	0.900
Neural Network	0.935	0.880	0.880	0.880	0.880
Naive Bayes	0.927	0.887	0.887	0.887	0.887
Gradient Boosting	0.968	0.913	0.913	0.914	0.913

Compare models by: Area under ROC curve

	Random Forest	Neural Network	Naive Bayes	Gradient Boosting
Random Forest		0.931	0.891	0.130
Neural Network	0.069		0.654	0.076
Naive Bayes	0.109	0.346		0.019
Gradient Boosting	0.870	0.924	0.981	

VI. RESULTS AND DISCUSSIONS

Based on the review performed, it is found that several algorithms are explored in each study and the models' accuracy is studied. The most frequently used algorithms for predicting failures are artificial neural network(ann), decision tree (dt), random forest,gradient boosting and naïve bayes (nb) where gradient boosting(gb) produced the most Accurate outputs for most of the studies.Using real test cases, demonstrated that the trained gradient boosting is able to detect Irregularities that could potentially lead to positive results With an accuracy of 96.8 %. However, it is hard to choose between high performing algorithms like random forest,naïve bayes (nb) and gradient boosting (gb) as their performance is comparable, and after Conducting few repeated experiments, gradient boosting (gb) is found to be the best classifier. The second most popular algorithm Adopted by researchers is random forest with 94.9% of accuracy.These ML systems are trained using massive datasets of existing transformer testing and monitoring data. A comprehensive dataset will typically use information from new, in-use and failed transformers, exposing the ML algorithm to various operational states and failure conditions. Information from the field can also help manufacturers build better transformers. If one fails or begins to behave unusually, businesses may have access to months of sensor data leading up to the failure.While still experimental, there's evidence that predictive maintenance can be much more effective than preventive maintenance. Preventive maintenance can be costly, but it's almost always less expensive than a major failure. Maintaining transformers isn't always a straightforward process, however. Technicians can miss minor issues that may eventually develop into serious problems, even with a theoretically effective maintenance schedule.

VII. CONCLUSION

This paper focused on reviewing data mining techniques and algorithms for developing the transformer failure prediction model. The model's performance is evaluated based on its accuracy. the most prominent technique used is Gradient boosting(GB).Whereas for other failure types, there is a requirement for further experiment using different types of input and algorithms in order to have a more reliable prediction. Further research also needs to be carried on to discover more transformer failure types and the best algorithms to predict those failures.

REFERENCES

[1] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," IEEE Electrical Insulation Magazine. 2002, doi: 10.1109/MEI.2002.1161455.

[2] R. R. Rogers, "Ieee And IEc Codes To Interpret Incipient Faults In Transformers/ Using Gas In Oil Analysis," IEEE Trans. Electr. Insul., 1978, doi: 10.1109/TEI.1978.298141.

[3] H. Ma, T. K. Saha, C. Ekanayake, and D. Martin, "Smart transformer for smart grid - Intelligent framework and techniques for power transformer asset management," IEEE Trans. Smart Grid, 2015, doi: 10.1109/TSG.2014.2384501.

[4] M. Duval and J. J. Dukarm, "Improving the reliability of transformer gas-in-oil diagnosis," IEEE Electr. Insul. Mag., 2005, doi: 10.1109/MEI.2005.1489986.

[5] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," IEEE Electr. Insul. Mag., 2001, doi: 10.1109/57.917529.

[6] R. Hierons, "Machine learning. Tom M. Mitchell. Published by McGraw-Hill, Maidenhead, U.K., International Student Edition, 1997. ISBN: 0-07-115467-1, 414 pages. Price: U.K. £22.99, soft cover.," Softw. Testing, Verif. Reliab., 1999, doi: 10.1002/(sici)1099-1689(199909)9:3<191::aid-stvr184>3.0.co;2-e.

[7] 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON) Galgotias University, Greater Noida, UP, India. Oct 2-4, 2020

[8] Singh, R. and Singh, A, "Causes of failure of distribution transformers in India," in Proc. 9th International Conference on Environment and Electrical Engineering (EEEIC), Prague, Czech Republic ,2010.

[9] Mohsen Akbari, P. Khazaee, I. Sabetghadam and P. Karimifard "Failure Modes and Effects Analysis (FMEA) for Power Transformers," in Proc. 28th International Conference on Power System, Tehran, Iran, 4-6, November 2013.

[10] D. Linhjell T. J. Painter L. E. Lundgaard, W. Hansen, "Aging of oil impregnated paper in power transformers," IEEE transactions on power delivery, 19(1), January 2004.

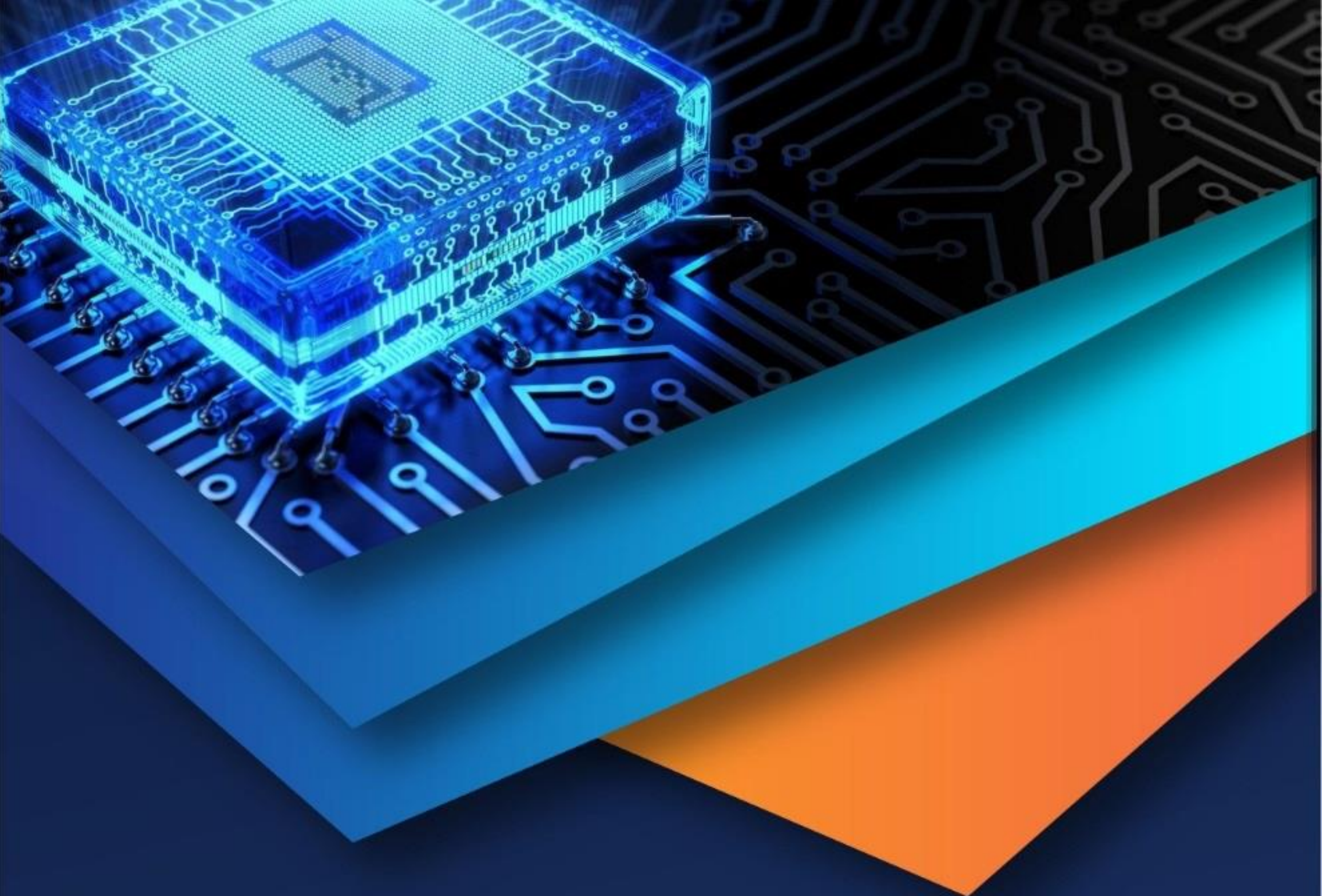
[11] William H. Bartley, HSB, "Analysis of Transformer Failures," Proceedings of the Thirty Six Annual Conference, Stockholm, 2003..

[12] S.S.Rajurkar and Amit R. Kulkarni, "Analysis of Power Transformer failure in Transmission utilities" in Proc. 16th National Power Systems Conference, 15th-17th December, 2010.

[13] A.K. Lokhanin G.Y. Shneider V.V. Sokolov V.M. Chornogostsky, "Internal insulation failure mechanisms of HV equipment under service conditions", 15-201 CIGRE Session 2002.

[14] George Eduful1 and Godfred Mensah, "An Investigation into Protection Integrity of Distribution Transformers - A Case Study," in Proc. of the World Congress on En

[15] Jiang Long1 , Li Shiyong1 , Yang Chao1 , Wang Dejun1 , Yao Yang2 , Wang Kai2 , Zhang Hongru2 , Li Qingquan2 (1. Guizhou Power Grid Co., LTD. Guiyang Power Supply Bureau, Guiyang 550081, China 2. Shandong Provincial Key Laboratory of UHV Transmission Technology and Equipment, School of Electrical Engineering, Shandong University, Jinan 250061, China)



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY