Machine Learning-Based Model to Enhance Smart Grid Reliability

[1]Bhushan Prataprao Patil, [2]Salunkhe Umakant Manohar, [3]Marathe Sushant Sanjay, [4]Patil Jayesh Navnit, [5]Bhavsar Rushikesh Rajendra, [6]Patil Ashwin Sanjay

R. C. Patel Institute of Technology
Shirpur, Maharashtra, India.

Abstract- **The electrical power system is a complex network of interconnected components that can experience disturbances or electrical faults. One challenging task is to detect these faults and determine their underlying causes. In this paper, a new approach using semi-supervised machine learning techniques is presented for fault detection and classification on transmission lines. The approach employs Decision Tree (DT), Random Forest (RF), Logistic Regression and SVM classifiers, which are designed to analyze and classify different types of faults. To extract relevant information from the fault current and voltage signals, a technique called Discrete Wavelet Transform (DWT) is applied for feature extraction. The Decision Tree classifier makes decisions based on constructing an optimal classification tree. It uses a set of rules to determine the most appropriate fault classification. The Random Forest and Logistic Regression also contribute to the fault analysis process. To evaluate the performance of the proposed system, various datasets are used. The results obtained from the two classifiers are compared based on their accuracy. This helps identify the most suitable technique for fault analysis in the given context. This paper presents an innovative approach to classify faults in transmission lines using machine learning techniques. By applying the Decision Tree, Random Forest, and Logistic Regression, along with the feature extraction capabilities of the Discrete Wavelet Transform. The Random Forest and Decision Tree improve the accuracy and effectiveness of fault analysis in the electrical power system**

1. INTRODUCTION

Smart Grid is an Electrical Grid with Automation, Communication and IT systems that can monitor power flows from points of generation to points of consumption (even down to appliances level) and control the power flow or curtail the load to match generation in real time or near real time. A smart grid uses sensing, embedded processing and digital communications to enable the electricity grid to be observable (able to be measured and visualized), controllable (able to manipulated and optimized), automated (able to adapt and self-heal), fully integrated (fully interoperable with existing systems and with the capacity to incorporate a diverse set energy sources). A smart grid uses digital technology to improve reliability, security, and efficiency (both economic and energy) of the electric system from large generation, through the delivery systems to electricity consumers and a growing number of distributed-generation and storage resources.” From the above mentioned definitions the Smart Grid can be described as the transparent, seamless and instantaneous two-way delivery of energy information, enabling the electricity industry to better manage energy delivery and transmission and empowering consumers to have more control over energy decisions. A Smart Grid incorporates the benefits of advanced communications and information technologies to deliver real-time information and enable the near-instantaneous balance of supply and demand on the electrical grid. Fault detection of smart grid is an important research problem that has attracted increasing attention from both academia and industry. It is essential to improve the performance and reduce disruptions of smart power systems. The presence of faults or incipient faults in smart grid can be determined by measuring and analyzing changes in electrical power (e.g., current and voltage), environmental and equipment parameters, a process known as smart grid fault detection. Achieving autonomous smart grid fault detection is crucial to system status awareness, maintenance and operation. With the continuous innovation of power systems and equipment, and the carbon neutral goals, more devices and technologies are used for fault detection.

2. LITERATURE OVERVIEW

Power generation, transmission, and distribution systems are essential for our modern economy and daily lives. It's crucial for these electrical power networks to provide continuous and reliable service. However, a lack of awareness about the network's condition can lead to abnormal operation or even a blackout. That's why power system protection has been a focal point of research since the advent of electricity. Detecting and classifying faults in the power system is a crucial task, and advancements in technology have made significant progress in analyzing these events over the years. In today's interconnected power system, it's important to have a reliable, sensitive, and fast-acting protection scheme to minimize damage to transmission lines caused by disturbances. Additionally, identifying the location of the disturbance is vital to reduce system outage time. Numerous efforts have been made to develop various methods for protecting power systems. These efforts aim to ensure the reliability and stability of the electrical network and enhance its resilience against faults and disruptions. The study [1] relies on analyzing the zero sequence and negative sequence values of currents and voltages to determine the type of fault, such as single phase to ground or two phases to ground. In [2], the Multiresolution analysis (MRA) technique is utilized to extract important features from the current signal. These features are then used for fault classification. The location of fault and type of fault were determined in [3] using a combination of the principle and sequence component analysis. Since expert knowledge-based techniques can be limited without real data, Machine Learning (ML) algorithms are being used to overcome this challenge. ML algorithms can automatically extract knowledge from available data,

making them valuable tools in various applications within the power system field [4]. The paper [5] discusses the use of Machine Learning (ML) for fault analysis in power systems. Specifically, Supervised Machine Learning, a common type of ML, has been widely employed for fault classification. However, one challenge is that a significant amount of power system data is recorded without labels, making it difficult for Supervised ML to accurately classify faults. To address this, the paper proposes using a Semi-supervised ML approach, which can effectively handle the large amount of unlabeled power system data. In the previous work [7], a semi-supervised Machine Learning (ML) approach called KNN for fault identification and classification in a small transmission system. They used the amplitude of the fault current as the feature to detect and classify faults. The results showed that the classifier performed well for small systems with fewer attributes.
In the present work, the focus is on a new technology where only a small amount of data is labeled, which is then used to label the unlabeled data using suitable classifiers. To extract the features from the three-phase current and voltage signals, they implemented a technique called Discrete Wavelet Transform (DWT). The DWT helps uncover hidden information in the fault current and voltage signals, and the derived frequency components carry the fault-related information. In [8] the decomposed current and voltage signals are used as feature vectors to predict the fault class in transmission lines. The K Nearest Neighbor (KNN) and Decision Tree (DT) classifiers are employed for fault detection and classification. The performance of both classifiers is compared to determine which one is most suitable for accurately detecting and classifying transmission line faults.

3. METHODOLOGY

Machine learning is a growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information. Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more. With the help of sample historical data, which is known as training data, machine learning algorithms build a mathematical model that helps in making predictions or decisions without being explicitly programmed. Machine learning brings computer science and statistics together for creating predictive models. Machine learning constructs or uses the algorithms that learn from historical data. The more we will provide the information, the higher will be the performance. A Machine Learning system learns from historical data, builds the prediction models, and whenever it receives new data, predicts the output for it. The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately. Suppose we have a complex problem, where we need to perform some predictions, so instead of writing a code for it, we just need to feed the data to generic algorithms, and with the help of these algorithms, machine builds the logic as per the data and predict the output. Machine learning has changed our way of thinking about the problem. The below block diagram explains the working of Machine Learning algorithm.

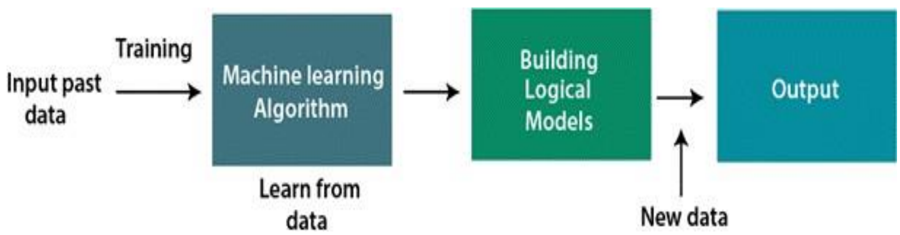

Figure: 1ML Model

4. MACHINE LEARNING ALGORITHMS

Machine Learning algorithms are the programs that can learn the hidden patterns from the data, predict the output, and improve the performance from experiences on their own. Different algorithms can be used in machine learning for different tasks are Logistic Regression, SVM (Support Vector Machine), Decision Tree and Random Forest. Classification result and based on the majority votes, the algorithm predicts the final output. Random forest is a fast algorithm, and can efficiently deal with the missing incorrect data. The below Figure shows the Dataset of the Transmission Line faults For working on the dataset we use Jupiter Notebook and for reading the dataset we import Pandas library, Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data.

Figure2: Dataset of Transmission Line

G	B	Y	R	Ia	Ib	Ic	Va	Vb	Vc
1	0	0	1	-151.292	-9.67745	85.80016	0.40075	-0.13293	-0.26781
1	0	0	1	-336.186	-76.2833	18.3289	0.312732	-0.12363	-0.1891
1	0	0	1	-502.892	-174.648	-80.9247	0.265728	-0.1143	-0.15143
1	0	1	1	-622.611	-220.052	-45.9412	-0.04311	0.377573	-0.33446
1	0	1	1	-614.49	-232.13	-46.5326	-0.0426	0.377181	-0.33458
1	0	1	1	-603.458	-246.876	-47.0167	-0.038	0.3746	-0.3366
0	1	1	0	-3.65168	-696.03	702.1821	-0.36555	-0.01723	0.382776
0	1	1	0	-4.6445	-690.086	697.2292	-0.37106	-0.01778	0.388841
0	1	1	0	-5.63786	-683.979	692.1135	-0.37648	-0.01834	0.394813
0	1	1	1	5.148249	-781.9	778.915	-0.03607	0.001009	0.035061
0	1	1	1	19.03908	-788.628	771.7516	-0.03534	0.000255	0.035081
0	1	1	1	32.67772	-794.986	764.4705	-0.03495	-0.0004	0.035349
1	1	1	1	-391.2	-507.078	898.3879	-0.03862	0.019274	0.019345
1	1	1	1	-381.394	-516.573	898.0726	-0.04097	0.019332	0.021633
1	1	1	1	-370.802	-526.476	897.3817	-0.04234	0.019111	0.023226
0	0	0	0	-71.3533	56.51298	11.36905	-0.02182	-0.49953	0.521349
0	0	0	0	-71.1438	55.56138	12.11123	-0.0123	-0.50432	0.516623
0	0	0	0	-70.919	54.60509	12.84284	-0.00278	-0.50899	0.511768

5. RESULTS AND DISCUSSION

A. Logistic Regression

The Logistic Regression Machine gives the Training Accuracy is 33.59 % and the Model Accuracy is 31.98

LogisticRegression

In [28]:
```
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LogisticRegression

lreg = LogisticRegression()
lreg.fit(X_train, y_train)
y_pred_lr = lreg.predict(X_test)

l_train = round(lreg.score(X_train, y_train) * 100, 2)
l_accuracy = round(accuracy_score(y_pred_lr, y_test) * 100, 2)

print("Training Accuracy    :",l_train ,"%")
print("Model Accuracy Score :",l_accuracy ,"%")
```
```
Training Accuracy    : 33.59 %
Model Accuracy Score : 31.98 %
```

B. Random Forest

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. The Random Forest gives the Training Accuracy is 100.0% and the Model Accuracy is 89.07

RandomForestClassifier

In [29]:
```
from sklearn.ensemble import RandomForestClassifier

random_forest = RandomForestClassifier(n_estimators=100)
random_forest.fit(X_train, y_train)
y_pred_rf = random_forest.predict(X_test)
random_forest.score(X_train, y_train)

random_forest_train = round(random_forest.score(X_train, y_train) * 100, 2)
random_forest_accuracy = round(accuracy_score(y_pred_rf, y_test) * 100, 2)

print("Training Accuracy    :",random_forest_train ,"%")
print("Model Accuracy Score :",random_forest_accuracy ,"%")
```
```
Training Accuracy    : 100.0 %
Model Accuracy Score : 89.07 %
```

C. Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. The objective of the SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The Support Vector Machine gives the Training Accuracy is 76.46% and the Model Accuracy is 74.13

SVM

In [30]:
```
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score

svc = SVC()
svc.fit(X_train, y_train)
y_pred_svc = svc.predict(X_test)

svc_train = round(svc.score(X_train, y_train) * 100, 2)
svc_accuracy = round(accuracy_score(y_pred_svc, y_test) * 100, 2)

print("Training Accuracy    :",svc_train ,"%")
print("Model Accuracy Score :",svc_accuracy ,"%")
```
```
Training Accuracy    : 76.46 %
Model Accuracy Score : 74.13 %
```

D. Decision Tree

Decision Tree is the most powerful and popular tool for classification and pre-diction. A Decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a classlabel. The Support Vector Machine gives the Training Accuracy is 100.0% and the Model Accuracy is 89.64

DecisionTreeClassifier

In [31]:
```
from sklearn.tree import DecisionTreeClassifier

decision = DecisionTreeClassifier()
decision.fit(X_train, y_train)

y_pred_dec = decision.predict(X_test)

decision_train = round(decision.score(X_train, y_train) * 100, 2)
decision_accuracy = round(accuracy_score(y_pred_dec, y_test) * 100, 2)

print("Training Accuracy    :",decision_train ,"%")
print("Model Accuracy Score :",decision_accuracy ,"%")
```
```
Training Accuracy    : 100.0 %
Model Accuracy Score : 89.64 %
```

CONCLUSION

This paper focuses on the fault detection in smart grid using Machine Learning techniques. The machine learning is efficiently used for estimating Instruments loss of life, detecting power quality events and faults, making optimal energy dis-patch decisions to reduce cost of energy, efficient electricity market operations, and securing data and preventing attacks on smart grid. The application Machine Learning of in smart grids provides powerful technical support to the digital power systems.These include limitation of finding past labelled data, rapid changes in failure patterns, issues with generating high resolution synthetic data for training, finding efficient feature selection techniques. By using the ML models we can predict which type of fault is occurred in the System. We used different Machine Learning Algorithms, Which has different accuracy. The Random Forest and Decision Tree improve the accuracy and effectiveness of fault analysis in the electrical power system

REFERENCES:

1. A. Rahmati and R. Adhami, quot;A fault detection and classification techniquebased on sequential components,quot; IEEE Trans. Ind. Appl., vol. 50, DOI10.1109/TIA.2014.2313652,no.6,pp.4202-4209,Dec.2014.
2. M.J.B. Reddy, D.V. Rajesh and D.K. Mohanta, "Robust transmission line faultclassification using wavelet multi-resolution analysis," Computers amp; Elec-trical Engineering, vol. 39, DOI 10.1016/j.compeleceng.2013.02.013, no. 4,pp.1219-1247,May2013.
3. Q. Alsafasfeh, I. Abdel-Qader and A. Harb, quot;Fault classification and Loc-alizationinPowerSystemsUsingFaultSignaturesandPrincipalCompon-entsAnalysisquot;,EnergyandPowerEngineering,vol.4,DOI10.4236/epe.2012.46064,no.6,pp.506-522,Nov.2012.
4. SeyedMahdiMiraftabzadeh,FedericaFoiadelli,MichelaLongo,MarcoPa-setti, "A Survey of Machine Learning Applications for Power System Analyt-ics", IEEE International Conference on Environment and Electrical Engineer-ingand2019IEEEIndustrialandCommercialPowerSystemsEurope(EEEIC /Iamp;CPSEurope),DOI:10.1109/EEEIC.2019.8783340,2019.
5. Halil Alper Tokel, Rana Al Halaseh, Gholamreza Alirezaei, and Rudolf Mathar,"A New Approach for Machine Learning-Based Fault Detection and ClassificationinPowerSystems",IEEEPoweramp;EnergySocietyInnovative Smart Grid

Technologies Conference (ISGT), DOI: 10.1109/ ISGT.2018. 8403343,2018.
6. Maitreyee Dey, Soumya Prakash Rana, Sandra Dudley, "Semi SupervisedLearning Techniques for Automated Fault Detection and Diagnosis of HVACSystems", 30th International Conference on Tools with Artificial Intelligence(ICTAI),DOI:10.1109/ICTAI.2018.00136,August2018.
7. PreranaP.Wasnik,Dr.N.J.Phadkule,Dr.K.D.Thakur,"FaultDetectionAndClassification Based On Semisupervised Machine Learning Using KNN", inProc, International Conference on Innovative Trends and Advances in Engin-eeringandTechnology(ICITAET),2019.
8. SumanDevi,NagendraK.Swarnkar,SheeshRamOlaandOmPrakashMahela,"Detection of Transmission Line Faults Using Discrete Wavelet Transform",Conference on Advances in Signal Processing (CASP),DOI: 10.1109/ CASP.2016.7746152,2016.
9. M. Gowrishankar, P. Nagaveni and P. Balakrishnan, "Transmission Line FaultDetection and Classification Using Discrete Wavelet Transform and ArtificialNeural Network", Middle-East Journal of Scientific Research, DOI: 10.5829/idosi.mejsr.2016.24.04.23063,2016.
10. Papia Ray, B. K. Panigrahi and N. Senroy, "Extreme Learning Machine basedFault Classification in a Series Compensated Transmission Line", IEEE Inter-national Conference on Power Electronics,Drives and Energy Systems,DOI:10.1109/PEDES.2012.6484297,2012.
11. Aida Asadi Majd,Haidar Samet and Teymoor Ghanbari,"k-NN based faultdetection and classification methods for power transmission systems", Protec-tionandControlofModernPowerSystems,DOI10.1186/s41601-017-0063-z,2