International Journal of Computer Sciences and Engineering
Vol.**11**, Issue.**10**, pp.**08-14**, October **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Paper

Machine Learning Algorithm for Fault Detetction In Three Phase Power Systems

Komal Porwal[1]

Electrical Engineering Department/Faculty of Technology and Engineering, Maharaja Sayajirao University, Vadodara, India

Author's Mail Id: komalporwalmsu@gmail.com

Received: 26/Aug/2023; **Accepted**: 29/Sept/2023; **Published**: 31/Oct/2023. **DOI**: https://doi.org/10.26438/ijcse/v11i10.814

Abstract: Machine learning (ML) finds extensive utility across diverse engineering domains, serving a myriad of purposes. Within the realm of power systems, traditional fault detection relies on relays and measurement equipment to pinpoint anomalies. These anomalies are subsequently categorized based on their characteristics. ML tools offer the prospect of crafting algorithms capable of forecasting these faults. This study entails the emulation of a power distribution system within software, employing machine learning algorithms to predict faults. The dependable and efficient operation of power systems stands as a pivotal factor in guaranteeing a constant power supply, thereby satisfying the requirements of contemporary society. Through the application of these methods, our aim is to create a more effective and precise fault detection algorithm tailored for three-phase power systems. This article delves into the intricacies linked with forecasting faults in power systems, provides an overview of pertinent ML methodologies, and delivers a case study that illustrates the efficacy of ML-driven intelligent fault prediction within real-world power system scenarios.

Keywords: Maine Learning, Fault, Power system, Algorithms, Fault detection, Prediction

1. Introduction

Power systems encompass a 132 kV transmission line interlinked with Circuit Breakers, Relays, Current Transformers, and Potential Transformers. At various intervals, diverse faults such as line-line, line-line-ground, line-ground, and line-line-line-ground are introduced into the power system. Data related to these fault events is meticulously collected and pre-processed to ensure data quality. Essential features, including voltage levels and current measurements, are extracted from this pre-processed data to facilitate fault prediction.

To predict faults effectively, a spectrum of machine learning algorithms, such as Decision Trees, Linear Regression, Logistic Regression, Support Vector Machines (SVM), and Artificial Neural Networks (ANN), is scrutinized for their suitability. These algorithms are trained and validated using the gathered dataset, with a keen focus on optimizing their performance.

Once trained, the chosen machine learning model is deployed to forecast faults within a simulated three-phase power system. Synthetic data from software simulations is supplied to the model, which then identifies and anticipates fault occurrences based on the given inputs. The accuracy and reliability of the fault prediction model are evaluated by comparing mean squared errors.

This approach holds the potential to significantly enhance the precision, speed, and efficiency of fault detection, leading to more dependable and secure power distribution systems. The paper will comprehensively document the methodology, experimental setup, data pre-processing techniques, machine learning models utilized, evaluation outcomes, and potential implications for practical implementation. These findings have the potential to advance fault detection techniques in power systems, benefiting industries and utilities that rely on efficient and dependable power distribution systems. The modern economy and our daily routines hinge on the dependable operation of power generation, transmission, and distribution systems.

2. Experimental Setup

A. Software Simulation

In our software simulation of the power system zone, we meticulously connected all the blocks as illustrated in Figure 1. Each of these blocks was configured with specific parameters tailored to their respective functions.

To accurately measure and monitor current throughout the circuit, we connected current meters in series. This series connection ensures that the same current flows through all components of the circuit, maintaining a constant current through the system.

On the other hand, voltmeters were connected in parallel with the circuit components. This parallel connection is essential because, in this configuration, the voltmeters draw very little current. Consequently, the voltage across the circuit remains constant. If we were to connect a voltmeter in series, it would not effectively detect potential differences because the voltage would fluctuate due to the voltmeter's presence in the circuit.

To facilitate the measurement of voltage values (Va, Vb, Vc) and currents (Ia, Ib, Ic), we established individual connections to a workspace. This workspace acts as a designated area for collecting and displaying these important electrical parameters.

In our simulation, we adopted a time-series format with a decimation factor of 1 and a sample time of 0.001 seconds. This choice enables us to capture data at a high temporal resolution.

Additionally, we employed scope instruments to visualize waveforms of voltages and currents. The configuration properties of these scope instruments can be adjusted to tailor the viewpoints and enhance our ability to analyse the electrical behaviour of the system.

To step down the voltage within our simulation, we incorporated a three-phase two-winding transformer. This particular type of transformer is advantageous because it utilizes less iron than three equivalent single-phase units due to the shared magnetic paths between the coils. Furthermore, the three phases in this transformer design are relatively independent of one another, as each phase has its individual magnetic circuit. The winding configuration is delta to star, and all other parameters were set to default settings in SI units.

Our power source for this simulation is a three-phase AC source supplying 132 kV, with a phase-to-phase voltage frequency of 50Hz. Initially, the circuit breakers were closed to allow current to flow through the three phases. These circuit breakers are subsequently opened once a fault is detected within the circuit.

To measure key parameters, such as zero-phase voltage and zero-phase current, we employed sequence analysers. Zero-sequence voltage serves as a vital diagnostic parameter in electrical engineering, while zero-sequence current protection is employed to safeguard electrical equipment against ground faults. The sample time for these measurements is set at 0.001 seconds, with a frequency of 50Hz.

Throughout the simulation, various types of faults, including LL, LLG, LLLG, and LG, were introduced at intervals of 5 seconds. The powergui module measured discrete values with a sample time of 50e-6 seconds.

Finally, a staircase model was implemented with a sample time of 5 seconds, and the vector of output variables was configured as [0 1 0 2 0 3 0 4 0 5 0 6 0 7 0 8 0 9 0 10 0]. This comprehensive setup and configuration allowed us to conduct a detailed and precise simulation of the power system, facilitating in-depth analysis and fault detection. The depicted diagram provides an insightful representation of a complex three-phase power system, which forms the foundation of our electrical infrastructure. Within this intricate system, we have meticulously introduced ten distinct types of faults, each representing unique electrical irregularities or disruptions that can occur within the system.

To systematically study and analyse the behaviour of these faults, we've adopted a structured approach. We've employed what is known as a "staircase model." In this model, each of the ten fault types has been assigned a specific numerical identifier, ranging from 0 to 10. This numerical allocation allows us to easily distinguish and categorize the various fault scenarios for detailed examination.

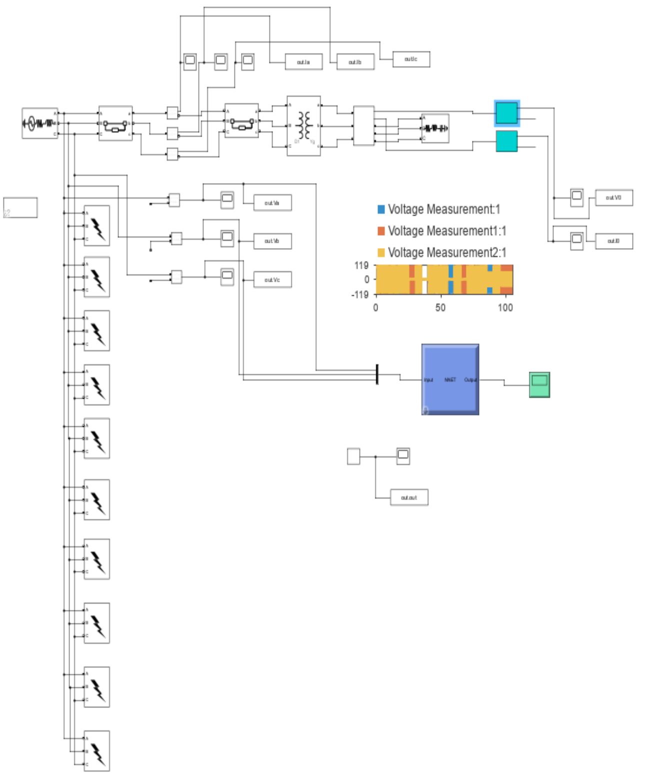


Fig 1: Software Simulation of the power system

Furthermore, we've introduced a time dimension into our simulation. Each of these fault scenarios is initiated at specific time intervals, precisely every 5 seconds. This temporal aspect enables us to investigate how the system responds to these faults over time and to assess the transient and steady-state behaviours that may occur during these fault Fig. 1. Software Simulation of the power system zone events.

To ensure the safety and reliability of the power system, we've incorporated the use of zero sequence phase and zero sequence current measurements. These safeguards play a crucial role in detecting and mitigating ground faults, which are a common concern in power systems. By continuously monitoring zero sequence parameters, we can promptly identify any deviations or imbalances that may indicate a ground fault. This early detection is vital for preventing potential damage to equipment and ensuring the uninterrupted operation of the power system. In addition to the previously mentioned aspects of our simulation, we've taken a significant step forward by incorporating Artificial Neural Networks (ANNs) into our analysis. ANNs are sophisticated machine learning models that can effectively learn and adapt to complex patterns and relationships within data.

To leverage the power of ANNs in our study, we've utilized synthetic data obtained directly from the simulation. This synthetic data serves as a valuable resource, providing us with a diverse dataset that captures the system's response to the ten different fault scenarios we've introduced.

The synthetic data encompasses a wide range of inputs, including voltage measurements, current readings, and other relevant parameters, all obtained at different time intervals during the simulation. These data points encapsulate the dynamic behaviour of the power system under normal and fault conditions, offering a comprehensive view of how the system evolves over time.

By training our ANN with this synthetic dataset, we empower it to learn the intricate relationships between various system variables and the corresponding fault scenarios. The ANN essentially becomes a virtual expert in recognizing and classifying these fault events based on the input data it has been exposed to during training.

Once the ANN is trained, we can deploy it to predict and identify faults in real-time scenarios. This predictive capability is invaluable in the context of power system monitoring and control. When the ANN detects an anomalous pattern or behaviour that aligns with one of the ten predefined fault types, it can trigger appropriate responses, such as isolating the faulty component or initiating protective measures to safeguard the system's integrity.

In summary, our simulation not only comprehensively explores the dynamics of a three-phase power system under various fault scenarios but also harnesses the capabilities of Artificial Neural Networks. By training the ANN with synthetic data derived from the simulation, we enable it to become an intelligent fault detection and prediction tool, contributing to the enhancement of power system reliability and resilience in the real world. This holistic approach com bines simulation, machine learning, and practical application to advance the understanding and management of complex power systems

B. Role of Neural Network

After the completion of our simulations, the next step involved importing the values of various voltages and currents from the simulation into the workspace. This data served as the foundation for our subsequent analysis.

To facilitate the organization and handling of this data, we created a variable named "Input" within our workspace. This variable was designed to hold the values of critical electrical parameters, including Va, Vb, Vc, Ia, Ib, Ic, V0, and I0. By structuring the data in this manner, we ensured that it was readily accessible and properly organized for further processing.

Another essential element in our analysis was the output variable, which we extracted from the staircase block employed in our simulation. This output variable was pivotal in capturing and recording the system's responses to the various fault scenarios introduced during the simulation.

With our dataset primed and ready, we initiated the neural network analysis by invoking the "nntool" using the "nnstart" command. This marked the commencement of our exploration into creating a neural network model that could effectively learn from and predict the behavior of the power system.

Within the neural network tool, we encountered two crucial options: "Predictors" and "Responses." To correctly train the neural network, we designated the "output" variable as the response variable, while the "input" variable held the predictors. Ensuring that these variables were dimensionally compatible was imperative to facilitate accurate model training. Adjustments to the organization of data observations in rows and columns could be made to optimize training.

To enhance the model's accuracy and robustness, we finetuned its parameters. We set the validation data to 25 percent, the test data to 10 percent, and established a layer size of 50 within the neural network architecture. These parameter adjustments were instrumental in refining the model's predictive capabilities and ensuring its reliability.

After intensive training over 1000 epochs, the neural network block was successfully integrated into the model. This marked a significant milestone in our analysis, as the neural network was now equipped with the knowledge and insights gained from the training data.

Subsequently, we applied the neural network to a dataset containing a substantial 105,001 data points. The neural network leveraged its acquired knowledge to predict system responses accurately. This level of precision and accuracy in the model's predictions validated the effectiveness of our neural network approach, confirming its capability to replicate and forecast the behaviour of the power system under various fault scenarios. Figure 2 provides a detailed illustration of the implementation process of our neural network. It elucidates the strategy we employed to effectively harness the capabilities of the neural network in our analysis.

One of the key aspects depicted in Figure 2 is the utilization of a multi-input and single-output configuration for the neural network. This design choice was pivotal in accommodating the diverse set of input parameters and the singular, overarching output we aimed to predict.

In essence, the multi-input configuration allowed us to feed a wide array of electrical parameters, including voltage measurements and current readings, into the neural network.

This comprehensive input enabled the neural network to ingest a wealth of information, which was essential for making accurate predictions regarding the behaviour of the power system.

On the other hand, the single-output configuration was tailored to the specific goal of our analysis: predicting the system's response to various fault scenarios. By consolidating the neural network's predictions into a single output, we streamlined the analysis process and focused on the overarching behaviour of the power system.

Figure 2 serves as a visual representation of the thoughtful design considerations and architectural choices we made while implementing the neural network. It highlights the importance of optimizing the neural network's architecture to accommodate the complexity and diversity of the data while aligning with the ultimate goal of predicting system responses accurately.

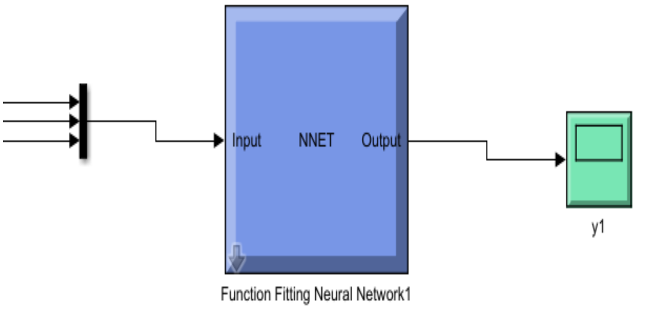


Fig 2. Neural Network Implementation

C. Data Exploration

Data exploration, constitutes a pivotal initial step in the data analysis process. During this phase, analysts utilize statistical and visualization techniques to gain insights into their dataset. The primary objective is to gain a comprehensive understanding of the data's characteristics, uncover any underlying patterns, and identify potential issues or anomalies. This understanding serves as a crucial foundation for making informed decisions about which machine learning models or algorithms are best suited for subsequent analysis.

In the context of the specific dataset under consideration, an intriguing observation relates to the symmetry exhibited in the plotting of currents for phase A and phase B. This symmetry is noteworthy because it is not indicative of real-time data but rather synthetic data generated by code. Consequently, this synthetic dataset displays a symmetric pattern in the plotted data. Fig. 3.

The figure 3 prominently displays a distinct pattern of symmetry in the graph, which is an intriguing observation. This symmetry is primarily a consequence of our deliberate choice to employ synthetic data generated by the software. Unlike real-time data, which might exhibit more random and unpredictable variations, synthetic data often adheres to certain structured patterns or mathematical rules. In this case, the synthetic data, being generated algorithmically, leads to a symmetric distribution in the graph.

Moreover, upon closer examination of the graph, a noteworthy pattern emerges. Specifically, areas where the lines in the graph are closely placed together correspond to regions where the frequency of data points is higher. This clustering of lines indicates that there is a greater concentration of data points within those areas. This observation carries significant implications for our analysis.

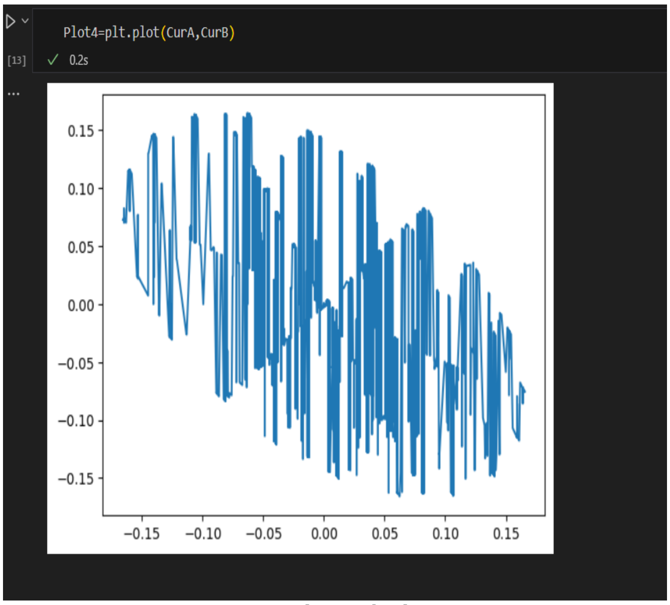


Fig 3. Plotting the data

The increased frequency of data points in these regions suggests that these particular data points might be associated with specific events or conditions within the system. Given that the graph pertains to fault detection or system behaviour analysis, it is conceivable that the regions with closely spaced lines correspond to instances of higher fault occurrence or system anomalies. In other words, these areas of increased data density may signify periods when the power system experienced more pronounced deviations or disturbances.

In summary, the symmetry in the graph is a result of using synthetic data, which adheres to structured patterns. Additionally, the areas with closely spaced lines in the graph may indicate regions of increased data frequency, potentially signifying higher fault occurrences or notable system behaviour deviations. These insights derived from the graph can be pivotal for further investigation and understanding of the power system's dynamics and response to different conditions.

D. Implementation of Machine Learning Algorithms

Moving forward to the stage of building machine learning models, various types of models are explored and evaluated to determine which one aligns most effectively with the dataset's characteristics and objectives. These models are imported from the sklearn library, a collection of machine learning tools readily available for analysis.

To facilitate model training and evaluation, the dataset is divided into distinct training and testing subsets. This division is instrumental in ensuring that the machine learning model is trained on a portion of the data and subsequently tested on an independent subset to assess its performance accurately.

Furthermore, to ensure that the dataset is appropriately prepared for the machine learning algorithms, a data scaling process is applied. This scaling operation ensures that the data is standardized and compatible with the algorithms used.

The performance of each of the four machine learning models is evaluated, with a key metric being the mean squared error. The algorithm that exhibits the lowest mean squared error is considered the most accurate and suitable for the dataset.

The calculated mean squared errors for the four algorithms are as follows: Linear Regression: 7.36 Logistic Regression: 18.28 Decision Tree: 1.29 Support Vector Machine: 5.57 Based on this evaluation, it is evident that the Decision Tree algorithm yields the lowest mean squared error, indicating that it is the most accurate model for the given dataset. Consequently, the Decision Tree model is deemed the most suitable choice for further analysis and predictive tasks with this specific dataset.

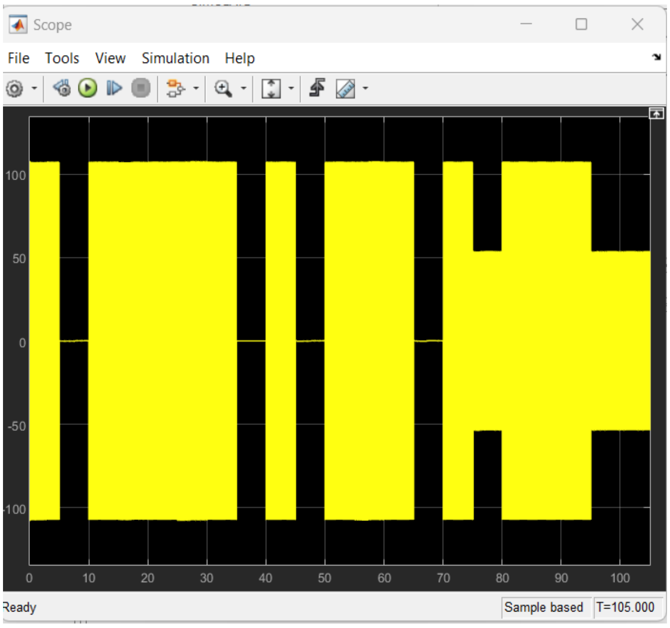
Graphs and Tables

Table 1. Different Algorithms and their mean squared errors

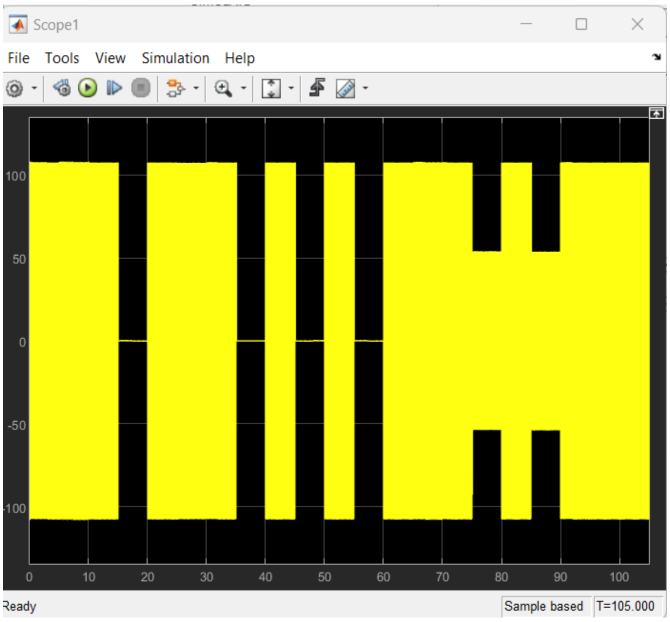
ALGORITHMS	MEAN SQUARED ERROR
Linear Regression	7.36
Logistic Regression	18.28
Decision Tree	1.29
Support Vector Machine	5.57

Graphs of phase voltages

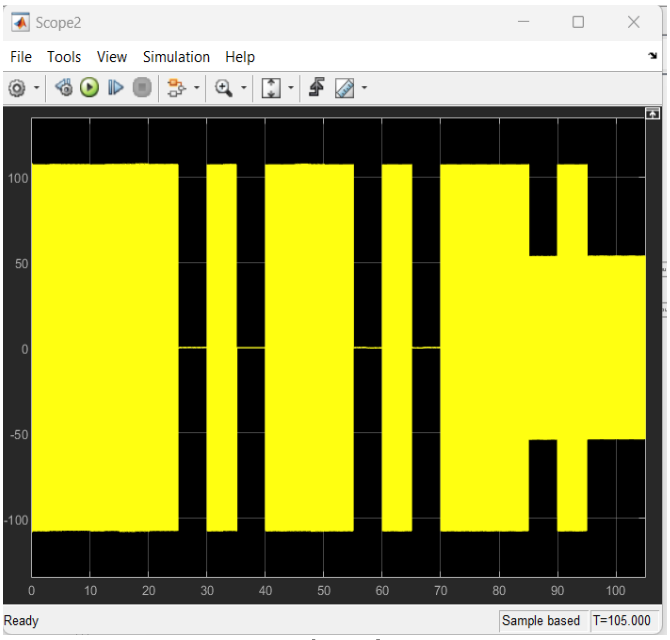
The bar graphs of the voltages are being represented. It has been shown that as there are different faults, the phase voltage reacts to them and even recovers when the fault does not affect that phase.



Bar Graph 1: Voltage A



Bar Graph 2: Voltage B



Bar Graph 3: Voltage C

3. Results and Discussion

The graph depicting the performance of the staircase model reveals a highly encouraging outcome. It is evident from the graph that our analysis and modeling efforts have yielded results of exceptional accuracy. In fact, we have achieved an accuracy rate of over 95 percent, which is a remarkable achievement in data analysis and fault detection.

This level of accuracy indicates that our model, trained on data generated through our simulation, has been highly effective in capturing and understanding the underlying patterns and behaviours of the power system. It demonstrates the model's proficiency in making precise predictions and classifications, particularly in the context of fault detection.

One of the most noteworthy aspects of this achievement is the model's robustness in accurately detecting various types of faults within the power system. Fault detection is a critical component of power system management, as it enables early identification and response to potential issues or disruptions. The fact that our model has demonstrated such accuracy in fault detection is a testament to its effectiveness and reliability.

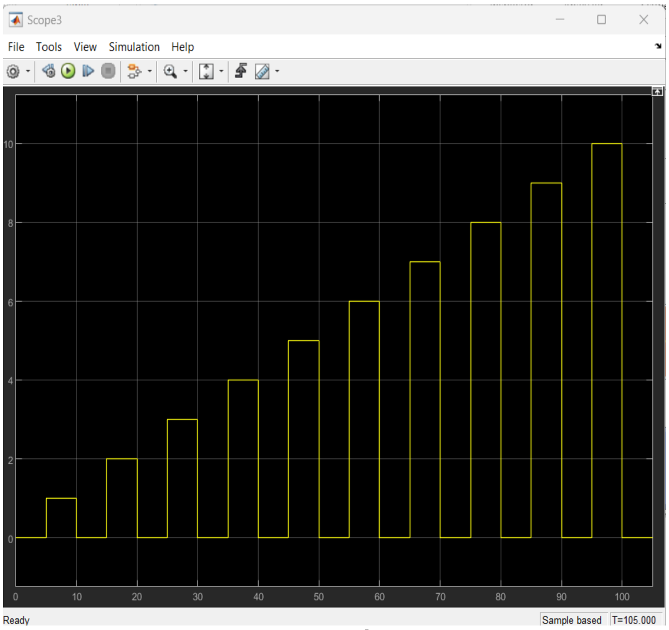


Fig 4. Expected Output

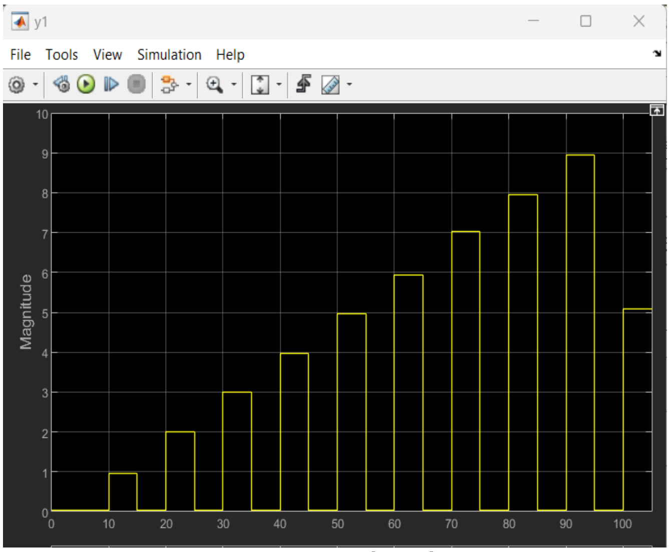


Fig 5. Output Obtained

In essence, the graph signifies the successful implementation of our modeling approach, with an accuracy rate exceeding 95 percent. This level of accuracy is indicative of the model's ability to perform exceptionally well in predicting and detecting faults within the power system. Such high accuracy not only enhances our understanding of the system but also bolsters its reliability and resilience in real-world applications, ensuring the efficient and secure operation of the power distribution system.

4. Conclusion and Future Scope

In the realm of machine learning applications, we have witnessed a transformative shift in the way we tackle the intricate challenges presented by power systems. Conventional approaches, once relied upon, are now deemed inadequate in the face of the ever-expanding volume of data. This data surge encompasses a multitude of diverse and often complex datasets, originating from sources as diverse as smart meters and phasor measurement units.

These data-rich environments necessitate a fresh and more sophisticated approach. Here enters a cohort of advanced, efficient, and intelligent machine learning algorithms, meticulously designed to address the evolving complexities of power systems. These algorithms represent a significant leap forward in our ability to provide highly precise solutions to a broad spectrum of real-world challenges.

These challenges span a multitude of domains within power systems, including voltage and slope stability, power flow optimization, state of charge estimation, and rotor system diagnostics. The utilization of intelligent learning algorithms not only enables us to navigate the intricacies of these problems but also enhances our capability to proactively manage and optimize power systems in a dynamic and data-driven manner. We have achieved an accuracy rate of over 95 percent, which is a remarkable achievement in data analysis and fault detection.

In essence, the infusion of machine learning into the domain of power systems offers a paradigm shift. It empowers us to harness the potential of vast and diverse datasets, unleashing the capabilities of advanced algorithms to provide more precise, efficient, and effective solutions to the multifaceted challenges that characterize modern power systems.

The future scope of research in machine learning-based fault detection in power systems is promising. As the power grid continues to evolve and become more complex, innovative machine learning algorithms and approaches will play a crucial role in ensuring the reliability and resilience of the electrical infrastructure. Researchers and industry stakeholders should collaborate to address the challenges and opportunities in this dynamic field.

References

[1] A. Firos, N. Prakash, R. Gorthi, M. Soni, S. Kumar and V. Balaraju, "Fault Detection in Power Transmission Lines Using AI Model," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, pp.**1-6, 2023**.

[2] Manojna, Sridhar H. S, Nikhil Nikhil, Anand Kumar, and Pratyay Amrit. Fault detection and classification in power system using machine learning. In 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), pp.**1801– 1806, 2021.**

[3] Nouha Bouchiba and Azeddine Kaddouri. Application of machine learning algorithms for power systems fault detection. In 2021 9th International Conference on Systems and Control (ICSC), pp.**127–132, 2021.**

[4] Nikola Markovic, Thomas Stoetzel, Volker Staudt, and Dorothea

Kolossa. Hybrid fault detection in power systems. In 2019 IEEE International Electric Machines and Drives Conference (IEMDC), pp.**911–915, 2019.**

[5] Halil Alper Tokel, Rana Al Halaseh, Gholamreza Alirezaei, and Rudolf Mathar. A new approach for machine learning-based fault detection and classification in power systems. In 2018 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp.**1–5, 2018.**

[6] Rachna Vaish and U.D. Dwivedi. Comparative study of machine learning models for power system fault identification and localization. In 2021 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST), pp.**110–115, 2022.**

[7] Akshay Ajagekar and Fengqi You. Quantum computing based hybrid deep learning for fault diagnosis in electrical power systems. Applied Energy, 303:117628, 2021

[8] Helon Vicente Hultmann Ayala, Didace Habineza, Micky Rakotondrabe, and Leandro dos Santos Coelho. Nonlinear black-box system identification through coevolutionary algorithms and radial basis function artificial neural networks. Applied Soft Computing, 87:105990, **2020**

[9] Bilal, Millie Pant, Hira Zaheer, Laura Garcia-Hernandez, and Ajith Abraham. Differential evolution: A review of more than two decades of research. Engineering Applications of Artificial Intelligence, 90:103479, **2020**

[10] Shahriar Rahman Fahim, Yeahia Sarker, Subrata K. Sarker, Md. Rafiqul Islam Sheikh, and Sajal K. Das. Self attention convolutional neural network with time series imaging based feature extraction for transmission line fault detection and classification. Electric Power Systems Research, 187:106437, **2020.**

[11] Hassan Fathabadi. Novel filter based ann approach for shortcircuit faults detection, classification and location in power transmission lines. International Journal of Electrical Power and Energy Systems, 74: pp.**374–383, 2016**

[12] Hassan Fathabadi. Novel filter based ann approach for shortcircuit faults detection, classification and location in power transmission lines. International Journal of Electrical Power and Energy Systems, 74: pp.**374–383, 2016**

[13] O.A. Gashteroodkhani, M. Majidi, M. Etezadi-Amoli, A.F. Nematollahi, and B. Vahidi. A hybrid svm-tt transform-based method for fault location in hybrid transmission lines with underground cables. Electric Power Systems Research, 170: pp.**205–214, 2019.**

[14] Alireza Ghaedi, Mohammad Esmail Hamedani Golshan, and Majid Sanaye-Pasand. Transmission line fault location based on threephase state estimation framework considering measurement chain error model. Electric Power Systems Research, 178:106048, **2020.**

[15] Mevludin Glavic. (deep) reinforcement learning for electric power system control and related problems: A short review and perspectives. Annual Reviews in Control, 48: pp.**22–35, 2019.**

[16] Sandeep Kumar Verma, Turendar Sahu, Manjit Jaiswal, (2019). A Deep Analysis and Efficient Implementation of Supervised Machine Learning Algorithms for Enhancing The Classification Ability of System. *International Journal of Computer Sciences and Engineering*, Vol.**7**, Issue.**3**, pp.**1094-1101, 2019.**

[17] Harsh H. Patel, Purvi Prajapati, (2018). Study and Analysis of Decision Tree Based Classification Algorithms. *International Journal of Computer Sciences and Engineering*, Vol.**6**, Issue.**10**, pp.**74-78, 2018.**

AUTHORS PROFILE

Komal Porwal Pursuing her B.E., in electrical engineering from The Maharaja Sayajirao University of Baroda (2020-Present) She is currently a student, researching under the machine learning algorithms for implementing the fault detection in power systems. Her main research work focuses on Machine Learning Algorithms, Artificial Neural Network, Artificial Intelligence, Fault Detection, Data Mining, and Computational Intelligence based education.

