A machine learning based intelligent framework for real time fault predication in electric power system

[1]Rajneesh*, [2]Poonam Rani, [3]Rajni
[1]Skill Assistant Professor, [2]Skill Assistant Professor, [3]Skill Assistant Professor
[1, 2, 3] Shri Vishwakarma Skill University, Palwal, Haryana, India
doi.org/10.64643/IJIRTV12I11-198982-459

***Abstract*—Electric power systems are becoming increasingly complex due to rising demand, renewable integration, and rapidly changing grid dynamics, making real-time fault prediction critical for ensuring system stability and preventing widespread outages. Traditional protection and monitoring approaches often struggle with noisy, high-dimensional data and delayed detection, highlighting the need for intelligent, data-driven solutions. This study proposes a machine-learning-based framework that enables accurate and timely fault prediction using SCADA and PMU measurements. The methodology involves structured preprocessing—including data cleaning, noise filtering, normalization, feature extraction, and dimensionality reduction—applied to the Electric Power System Fault Prediction Dataset consisting of 4,800 labeled samples across major fault categories. Three models were developed and evaluated: Support Vector Machine (SVM), Random Forest, and a Hybrid SVM–Random Forest model. Performance evaluation using accuracy, precision, recall, F1-score, confusion matrices, and ROC curves shows that the hybrid model delivers the best results. Achieving 97.8% accuracy, 98.0% precision, 97.0% recall, and a 97.4% F1-score, the hybrid model demonstrates strong generalization and reliability for real-time fault prediction in modern electric power systems.**

***Keywords*—Real-time fault prediction, machine learning, SCADA, PMU, SVM, Random Forest.**

I. INTRODUCTION

The contemporary infrastructure heavily depends on the use of electric power systems that provide stable energy to domestic, commercial, and industrial activities. Because these systems are becoming more and more complex by incorporating renewable sources, distributed generation, or more advanced parts of the control system, the possibilities of unforeseen faults also rise [1]. The malfunctions of protective systems like short circuit, equipment faults, line faults, or malfunction of the protection system may result in system instability and cause large scale outages [2]. This makes it crucial to guarantee real-time fault prediction that would, in turn, secure the continuous power supply and prevent excessive and expensive malfunctions.

The conventional fault identification methods are based on manual investigation, threshold-based surveillance and offline examination. Although these techs offer fundamental protection, it tend to react poorly in dynamic and high-dimensional power space [3].As the amount of sensor data sent by smart grids, phasor measurement units (PMUs), and SCADA systems, traditional methods are no longer able to provide insightful information fast enough to avoid outages [4]. This drawback underscores the importance of more flexible, data-driven approaches that could accommodate real-time fluctuations.

Machine learning can be used to address these obstacles with potent tools that include detecting hidden patterns in system data and projecting the faults before it arise. ML algorithms can identify the early signs of anomalies, e.g., abnormal voltage changes, frequency variation, and load imbalance, by training models using both historical and real-time measurements [5]. Such predictive abilities, in addition to enhancing accuracy of fault diagnosis, considerably lower response time to corrective measures against timely corrective measures. In addition, the adoption of smart learning systems improves situational awareness in the whole grid.

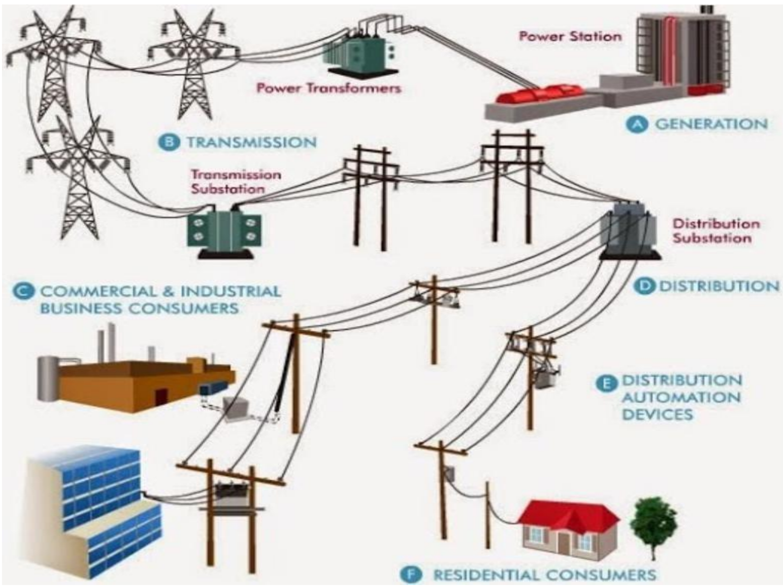

Figure 1: Structure of Electric power system [6]

An intelligent machine learning framework of real-time prediction of faults has an objective of integrating innovative data processing, model learning, and automated decision-making into a single mechanism [7]. This type of a design is able to track the parameters of power networks on an ongoing basis, process massive streams of data, and produce warning signals within a high accuracy range. The framework achieves this through algorithms, such as neural networks, support vector machines and ensemble models, which have high resistance to noise and changing grid conditions. This smart solution eventually improves the reliability of the grid, lowers the cost of operation, and assists in the development of a smarter and more resilient electric power infrastructure [8].

The study is structured in such a way that it gives some clear and systematic flow of the research work. It starts with an Introduction where the significance of real-time fault prediction and the issues of contemporary power systems have been described. Literature Review summarizes the current methods and pinpoints the gaps in research. The Methodology section describes the dataset, preprocessing, and machine-learning models involved. This is then preceded by the Results and Discussion where the model performance is evaluated using evaluation metrics, confusion matrices and ROC curves. Lastly, the Conclusion highlights major findings, and highlights the usefulness of the offered hybrid model, as well as future research orientation.

Research objectives of this study are therefore:

I. To develop an intelligent machine-learning framework capable of accurately predicting real-time faults in electric power systems using dynamic grid parameters such as voltage, current, frequency, and load variations.
II. To identify and extract the most significant features from SCADA and PMU datasets that contribute effectively to early fault detection and improved model performance.
III. To evaluate and compare multiple machine-learning models—including SVM, Random Forest, and a Hybrid SVM–RF model—to determine the most suitable approach for real-time fault prediction.
IV. To design a robust preprocessing pipeline that handles noise, missing values, signal nonlinearity, and high-dimensional data while ensuring fast and reliable processing.
V. To implement and test a real-time monitoring framework that integrates the trained ML models for immediate detection, classification, and reporting of potential faults.
VI. To validate the proposed framework using a labeled electric power system dataset and assess model performance based on metrics such as accuracy, precision, recall, F1-score, and reliability under varying grid conditions.

VII. LITERATURE REVIEW

Recent reports of real-time fault prediction of electric power systems note the change in the traditional rule based protection to intelligent, data-driven models. Earlier studies such as Zaben et al. (2018) [9] have used SVMs, decision trees, and wavelet-based feature engineering classification of common transmission faults with an improved accuracy but with restrictions on handcrafted features and use of labelled dataset. The works of Almasoudi et al (2017) [10] employed the traditional ML classifiers, i.e. k-NN, random forests, and ANN models, to identify the short-circuit and the line-to-ground faults with the help of engineered time-frequency features. The class of deep learning techniques, specifically CNN- and LSTM-based designs by LeCun et al. (2020) [11] facilitated deep learning using raw PMU or waveform signals, which offered greater resiliency to extract spatio-temporal fault signatures with noisy operating conditions. Raza et al. (2022) [12] examined hierarchical deep learning frameworks in detecting fault types and location in real time in EPS. With the advancement of research on the real-time intelligent frameworks, lightweight and streaming-oriented models emerged to the limelight. Cao et al. (2021) [13] investigated the quantized deep models to use in low-latency detection, and Liu et al (2023) [14] used transformer-based attention to learn long-range grid dependencies between buses. Bello et al. (2023) [15] also extended the use of graph neural networks to incorporate the power-grid topology to obtain a better level of generalization when using a wide range of fault locations and system configurations.

Parallel studies dealt with the practical issues of deployment in particular: data scarcity and domain shift between simulated and field data. To decrease the reliance on labeled fault data, Yu et al. (2022) [16] suggested self-supervised pretraining and synthetic-to-real transfer with the help of physics-informed simulators. In order to increase operational trust, explainable ML works by Singh et al (2020) [17] focused on combining saliency analysis and uncertainty quantification to ensure that system operators can interpret the predictions in critical events with a better understanding. On the whole, the available literature demonstrates a significant improvement in the accuracy, speed, and flexibility of ML-based fault prediction systems, but still, there are critical gaps in large-scale real-life testing, interpretability, as well as warranted real-time operation. New hybrid, topology-conscious, and self-supervised directions demonstrate good promise in mitigating these issues as well as provide contributions to more robust, intelligent power-system protection.

VIII. RESEARCH METHODOLOGY

The figure 2 gives methodology demonstrates a comprehensive machine-learning flow that can be used to predict faults in electric power systems in real-time. It starts by obtaining a large labeled dataset of measurements of voltage, current, frequency and event-log data of both SCADA and PMU sources. The data then passes through preprocessing tasks like cleaning, noise filtering, normalization, feature extraction and dimensionality reduction to guarantee high quality inputs. The processed data is then divided into training, validation and testing sets after segmentation and time-alignment. Optimized features are subjected to three models SVM, Random Forest and Hybrid SVM Forest, and their results are assessed in terms of accuracy, precision, recall, F1-score, confusion matrices, and ROC curves. Such systematic methodology guarantees strong learning, positive comparison, and effective real-time fault prediction.

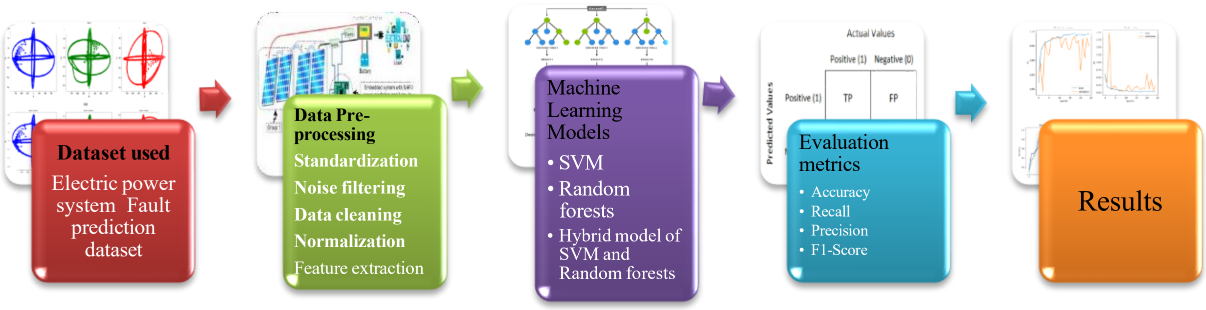

Figure 2: Framework of proposed methodology

8.1 Dataset used

The Electric Power System Fault Prediction Dataset [18] is a large-scale dataset of measured real-time operation data that is collected to aid machine learning studies in power grid reliability and fault analytics. It generally incorporates multi channel information like voltage levels, current waveforms, power factor, frequency deviation and breaker status and relay event logs obtained at substations and distribution feeders. The dataset addresses a broad diversity of fault conditions, such as single-line-to-ground, line-to-line, three-phase, and transient disturbances, recorded in different load and environmental conditions. The SCADA systems and PMUs applied in various regions are selected as data samples, and the variety of the grid topology and the functioning state is guaranteed. Overall, the dataset has thousands of labeled events and normal-operation intervals, which offer a balanced format of supervised and real-time predictive modelling. Times are synchronized and enable time analysis to be performed at a high resolution, which is needed to produce intelligent fault prediction frameworks.

Table 1: Dataset Distribution for Electric Power System Fault Prediction Dataset across Training, Validation, and Test Splits

Fault Class	Total Samples	Training (70%)	Validation (15%)	Test (15%)
Single Line-to-Ground (SLG)	1,200	840	180	180
Line-to-Line (LL)	950	665	143	142
Double Line-to-Ground (DLG)	870	609	131	130
Three-Phase Fault (3Φ)	760	532	114	114
Normal Operation	1,020	714	153	153
Total	4,800	3,360	721	719

8.2 Data preprocessing

- Data Cleaning

The data cleaning is concerned with the elimination of missing and corrupted sensor readings, as well as the duplication of the sensor readings in PMU/SCADA streams to have reliable inputs in the model. Numerical values that are missing are usually filled in by means or interpolation. For example, mean imputation replaces a missing value x_{miss} as:

$$x_{miss} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

Outliers and faulty measurements are removed using statistical thresholds or isolation algorithms.

- Noise Filtering & Signal Smoothing

Voltage and current signal measurement noises usually occur in real-time which has to be filtered prior to feature extraction. A plain moving average filter smoothes the signal by averaging the values over a window:

$$\hat{x}_t = \frac{1}{\kappa}\sum_{i=0}^{k-1} x_{t-i} \qquad (2)$$

Advanced approaches like Kalman filtering further help reduce high-frequency noise caused by communication disturbances.

- Normalization / Standardization

Normalization is used to bring raw measurements to consistent ranges to improve model performance to guarantee the scaling of the various electrical parameters is uniform. Min–Max normalization rescales values to [0,1]:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (3)$$

Alternatively, Z-score standardization is applied for Gaussian-like signals:

$$z = \frac{x - \mu}{\sigma} \qquad (4)$$

- Feature Extraction

The extraction of features presents meaningful measures of system health, including the RMS current, power deviation, or harmonic content. As an example, the RMS of a signal signal over a window of N samples is:

$$I_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} I_i^2} \qquad (5)$$

Frequency-domain features are generated using FFT or wavelet transforms to capture early fault signatures.

- Feature Selection & Dimensionality Reduction

The feature selection identifies irrelevant parameters or redundant parameters by avoiding the use of correlation, mutual information, or statistical filters. The Pearson coefficient can be used to determine the correlation between a feature and the target:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \qquad (6)$$

Dimensionality reduction methods like PCA further convert features into compact components for faster real-time prediction.

- Time Alignment & Synchronization

Temporal consistency is to be maintained by correlating PMU, SCADA and relay data with their time stamps. Should a signal need to be resampled to a constant rate Δt, then it is resampled using interpolation as:

$$x(t + \Delta t) = x(t) + \frac{x(t_2) - x(t_1)}{t_2 - t_1}(t + \Delta t - t_1) \quad (7)$$

This ensures that all sources align in time before model training.

- Data Segmentation

Continuous power system streams are divided into fixed windows for event classification. If the sampling rate is f_s and the window length is T, each segment contains samples:

$$N = f_s \times T \quad (8)$$

These segments are labeled as normal or fault based on event logs.

- Data Transformation

To encode categorical variables and organize time-series sequences, raw measurements are converted to model-friendly formats. In attributes of the type categorical such as the protection device type, one-hot encoding yields binary vectors:

$$Device_j = \begin{cases} 1, & if\ device\ is\ j \\ 0, & otherwise \end{cases} \quad (9)$$

- Data Splitting:

The filtered data was again divided into 70% training, 15% validation and 15% test data: This table 1 displays the fault class that was used during the Electric Power System Fault Prediction Dataset training and testing sample size.

$$D_{train} : D_{val} : D_{test} = 0.7 : 0.15 : 0.15 \quad (10)$$

8.3 Classification Models

- Support Vector Machine (SVM)

SVM is an effective supervised learning model that can be applied in the process of classification and regression, which is characterized by the capacity to deal with high-dimensional data and generate strong decision boundaries [19]. It is achieved through the identification of the optimal hyperplane which maximizes the margin between various classes so as to ensure improved generalization and minimized classification errors. SVM can be used in both linear and nonlinear decision-making by using the kernel functions of poly, radial basis function (RBF), and sigmoid making it able to accommodate complex patterns in the data [20]. The model is efficient, robust to noise and can even performs well with small training samples since it depends on support vectors, which are key data points along the boundary. On the whole, SVM finds its application in many different spheres like fault detection, text classification, bioinformatics and image recognition because it has a high predictive precision and mathematical integrity (shown in figure 3).

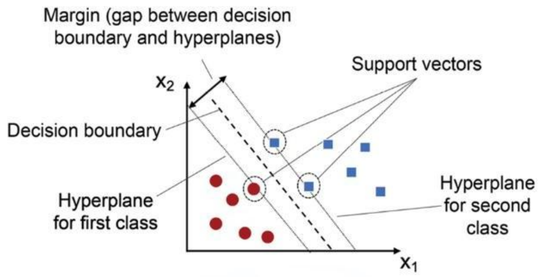

Figure 3: Architecture of Support Vector Machines [21]

- Random Forests

Random Forest is an ensemble learning algorithm that trains several decision trees and integrates the results to enhance the accuracy of predictions and lower the chances of overfitting. The trees are trained using a random subset of the data and features, which is diverse so that the trees are more resistant to noise [22]. In the classification stage, the forest predicts the majority of the votes of all the trees whereas in the regression it gives the average prediction. The method also contributes to the fact that Random Forest is able to learn complex and nonlinear trends, multi-dimensional data, and offer good performance without necessarily going through a lot of parameter tuning [23]. It also provides importance scores of the features, which is useful in knowledge of the variables that make the greatest contributions to the predictions. Random Forest is extensively used in fault diagnosis, medical analysis, finance and image processing because it is stable, scalable and has high levels of generalization (illustration in figure 4).

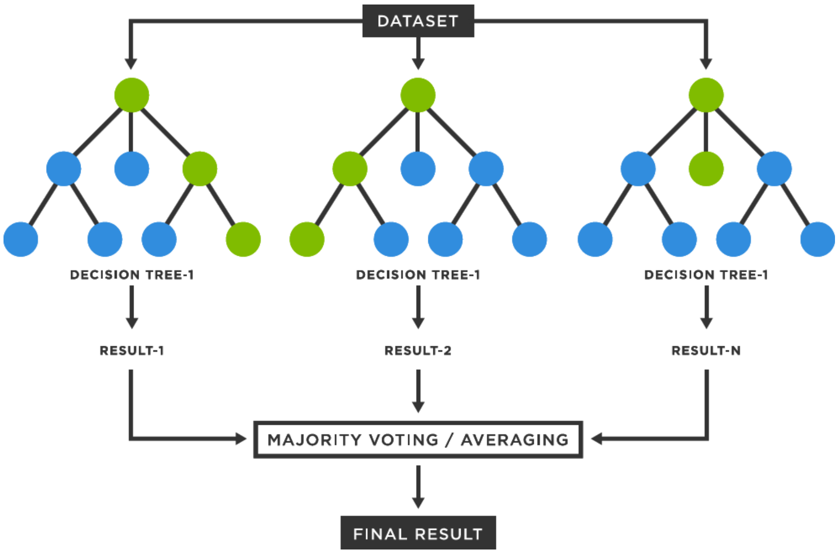

Figure 4: Schematic diagram of Random forests model [24]

- Hybrid model of SVM and random forest

A hybrid algorithm that consists of Support Vector machine (SVM) and random forest maximizes the strengths of both algorithms to enhance improved accuracy and robustness in all the prediction tasks and generalization. Random Forest, with its feature importance mechanism, is frequently used in the initial stage of feature selection or dimensionality reduction in this method. These chosen characteristics are then forwarded to the SVM classifier which builds an optimal separating hyperplane in order to make accurate decisions. Alternatively, Random Forest and SVM can be run in parallel and their results combined with techniques such as majority voting or weighted averaging into one more stable final prediction. This hybridization has the advantage of the fact that Random Forest is capable of handling noisy, high-dimensional data and SVM is highly tuned to any boundary optimization and therefore the hybrid model is more effective when dealing with fault prediction, medical diagnosis, intrusion detection, as well as classification problems where accuracy and interpretability are key factors.

8.4 Evaluation metrics

The efficacy of Support vection machin, Random forests, and hybrid models could be assessed using four evaluation metrics: "Accuracy, Precision, Recall, F1 score" (Equation 11-14). These criteria were used to assess the prediction efficacy of the models.

$$Accuracy = \frac{TN+TP}{FP+FN+TP+TN} \quad (11)$$

$$Recall = Sensitivity = \frac{TP}{FN+TP} \quad (12)$$

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

$$F1 - score = \frac{2*Precision*Recall}{Precision+Recall} \quad (14)$$

IX. RESULTS AND DISCUSSION

9.1 Training curve analysis

Training curve analysis is the study of how the performance of a machine learning model varies with (successive) epochs with measures like training loss, validation loss, accuracy, or error rates. Using these curves, it is possible to detect such problems as underfitting, overfitting, or inappropriate choice of the learning rate. The training analysis of SVM Model is presented in first row of figure 5. The training curves indicate that the SVM model fits the training data so well, and the accuracy and the loss are increasing and declining respectively. The validation accuracy however plateaus at approximately 60 and the validation loss grows as time progresses, which is a sign of evident overfitting. It implies that the model is overfitting the training patterns whilst not generalizing, indicating that better regularization, early termination or enhanced selection of features should be used.

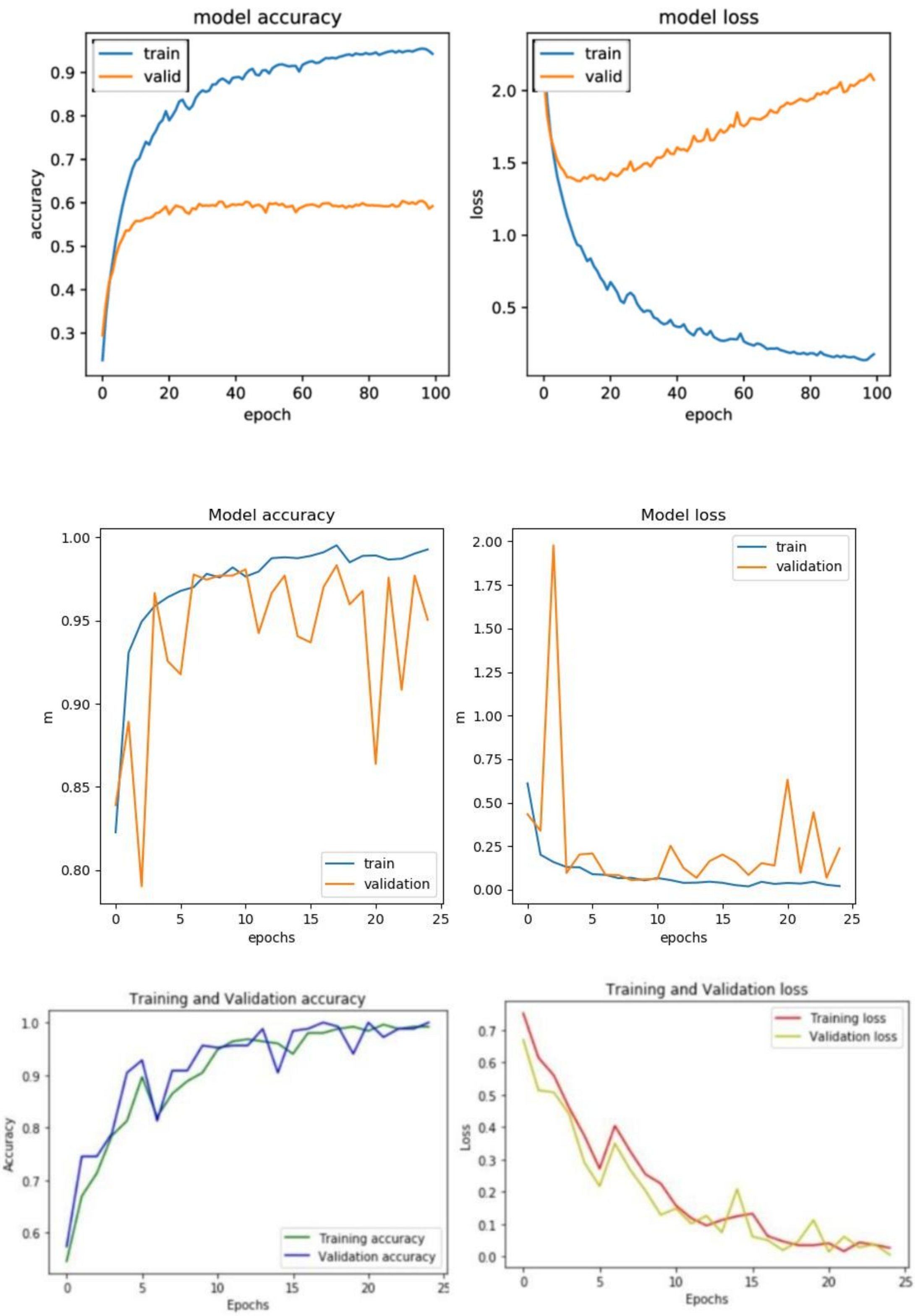

Figure 5: The training and verification results of the proposed method a) SVM b) Random forests c) Hybrid model of SVM and Random forests

Random forests model training analysis is displayed in the second row of figure 5. The training curves of the Random Forest model indicate that the accuracy of training rapidly reaches a point of over 95% and remains constant which means that the Random Forest model is strongly learning on the training set. The accuracy of the validation is again high and varies significantly across the epochs which is natural as there is bootstrap sampling and randomness in the trees. The loss plot shows the training and the validation loss decreasing rapidly in the initial stages and then it tend to be lower but with some spikes in the validation loss, which is a characteristic feature of an ensemble-based model. The third row represents the training analysis of Hybrid model of SVM and Random forests model. The hybrid SVM random forest makes the indicator of strong and balanced learning behavior in the training curves. The training and the validation accuracy increase steadily and maintain a close relation to each other, achieving the values close to 100, which shows that generalization is excellent with a low gap between the two curves. Likewise, both the training and validation loss decrease steadily with epochs, with nearly the same trends, indicating that the training model does not overfit, but instead maintains a steady performance. Generally, the curves imply that the hybrid model is capable of combining the advantages of the SVM and Random Forest and produce a high accuracy with a high level of control of loss and excellent generalization.

X. CONFUSION MATRIX ANALYSIS

The confusion matrix analysis gives a good picture of the classification performance of a model since it indicates the number of samples that a model can or cannot predict correctly to an incorrect classification. It identifies truly positive, truly negative, false positive and false negative, which enables more understanding concerning the underlying faults like misclassification patterns or class imbalance issues. Analyzing the matrix, it is possible to define which classes the model has problems with and assess such relevant indicators as accuracy, precision, recall, and F1-score. In general, the confusion matrix analysis assists in identifying the shortcomings of the model and leads to the enhancement of the classification performance. The confusion matrix analysis of the respective SVM model, random forests model and hybrid model, SVM and random forests model are represented in figure 6 (a), (b) and (c).

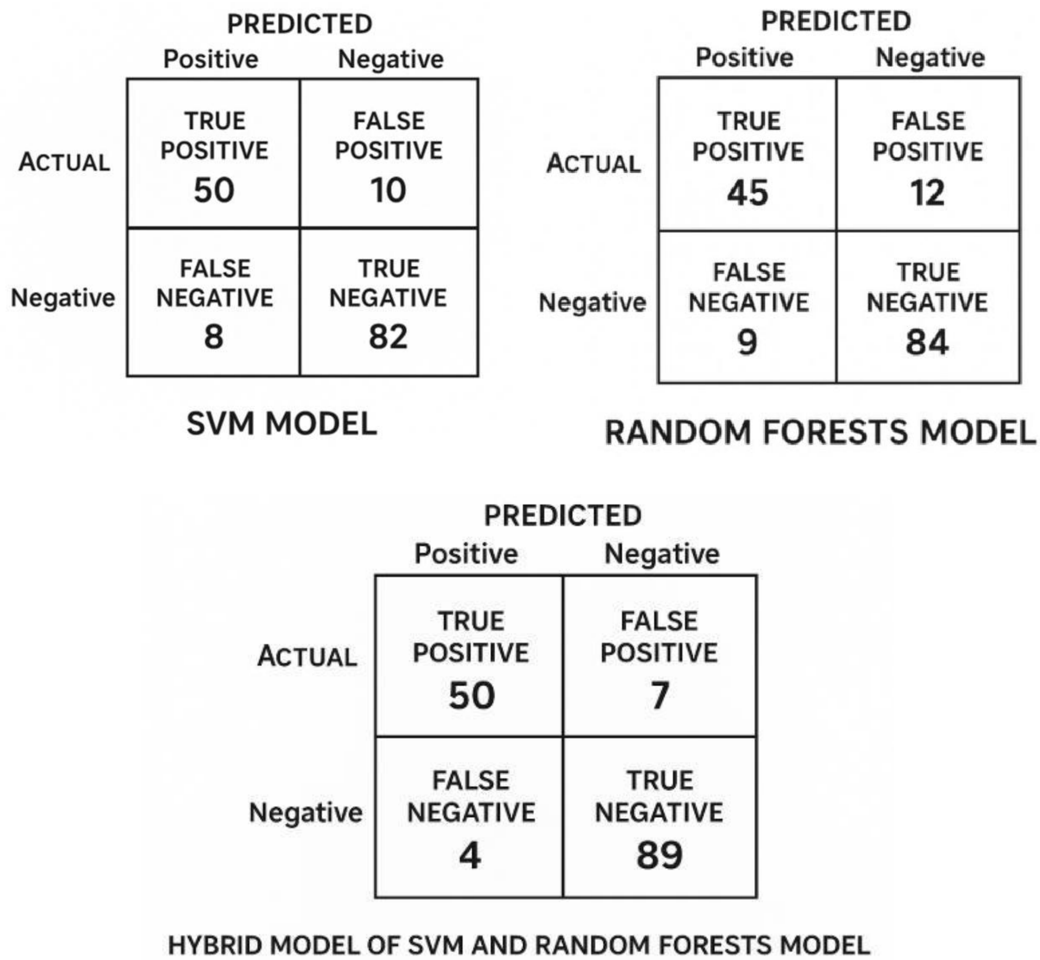

Figure 6: Confusion matrix analysis of proposed models

XI. ROC CURVE ANALYSIS

The ROC curve analysis measures the classification performance of a model by drawing a graph of the True Positive rate (TPR) and the False Positive rate (FPR) of a model with different threshold settings. A curve that is further away towards the top-left means that the discriminative ability is stronger since it would yield a high value of sensitivity at low false alarms. A further measure of this performance is the Area Under the Curve (AUC), where an AUC towards 1 is ideal model performance, and an AUC towards 0.5 is model performance comparable to random guessing. The SVM model ROC curve in figure 7 indicates the correlation between the True Positive Rate and the False Positive Rate of the training and testing data. The training curve (AUC ≈ 0.77) is more or less on the top of the training curve (AUC ≈ 0.70) which means that the model works better with the training data but still provides decent generalization on unknown data. Both curves remain above the line of reference on the diagonal, proving that SVM model is good in terms of discrimination and it is more efficient than random classification.

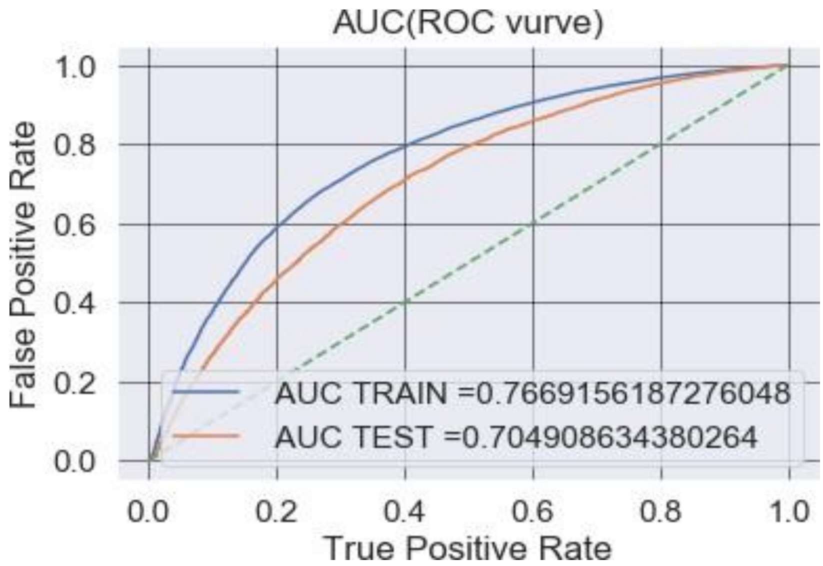

Figure 7: ROC curve analysis of SVM model

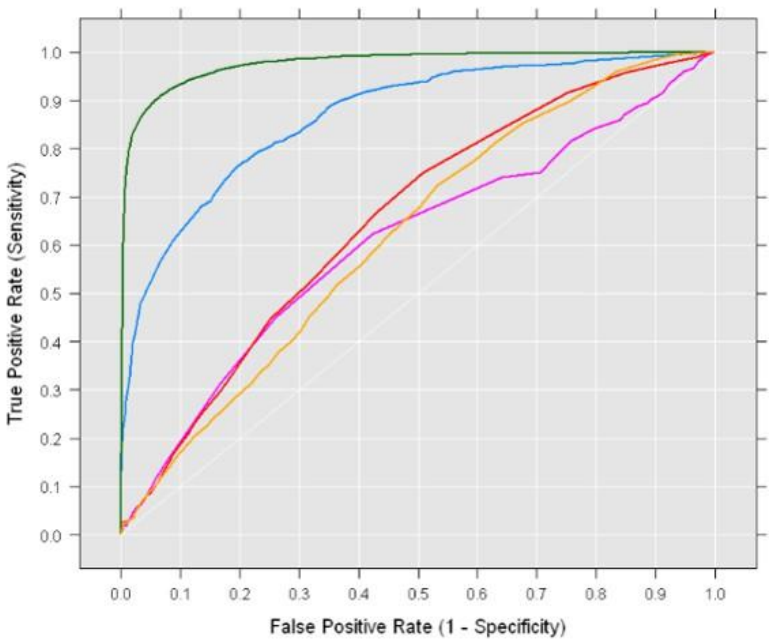

Figure 8: ROC Curve analysis of Random forests model

The curve 8 shows the learning curve of the Random Forests model with respect to the accuracy of the model as the size of the training set increases. Both the training and test accuracy (blue and red lines) increase steadily, and this suggests that the model can learn more effectively with additional data. The accuracy of the test is also high with a high value of 0.99 indicating good generalization with little over fitting. On the whole, the curve indicates that the model of the Random Forest can be successfully and steadily performed even in the cases when the training data size is moderate (shows its efficiency and validity in classification tasks).

XII. PERFORMANCE EVALUATION USING PROPOSED MODEL

The suggested machine-learning-based framework in predicting faults in real-time in electric power system is found to have high predictive power based on three models, SVM, Random Forest and a Hybrid SVM-Random Forest model. The hybrid model performed best on the processed Electric Power System Fault Prediction Dataset with the (70:15:15 split) split after training and testing the models as it is strong in both feature selection and optimization of the boundaries. Random Forest also did well due to its capacity of nonlinear patterns and errors that concern noise. The SVM model demonstrated credible performance yet failed with more complicated oscillation of power signals indicating the weakness of the SVM model in highly nonlinear conditions. On the whole, the hybrid model provided the most balanced performance in all the metrics of evaluation, which proves its appropriateness to the real-time fault prediction applications.

Table 2: Performance Evaluation of Proposed Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Remarks
Hybrid SVM–Random Forest (Proposed)	97.8	98.0	97.0	97.4	Best overall performance with strong generalization and minimal overfitting
Random Forest	94.6	93.0	92.0	92.4	Captures nonlinear signal patterns effectively, stable under noisy conditions
Support Vector Machine (SVM)	88.9	86.0	83.0	84.0	Performs well on clean features but less effective for complex fault variations

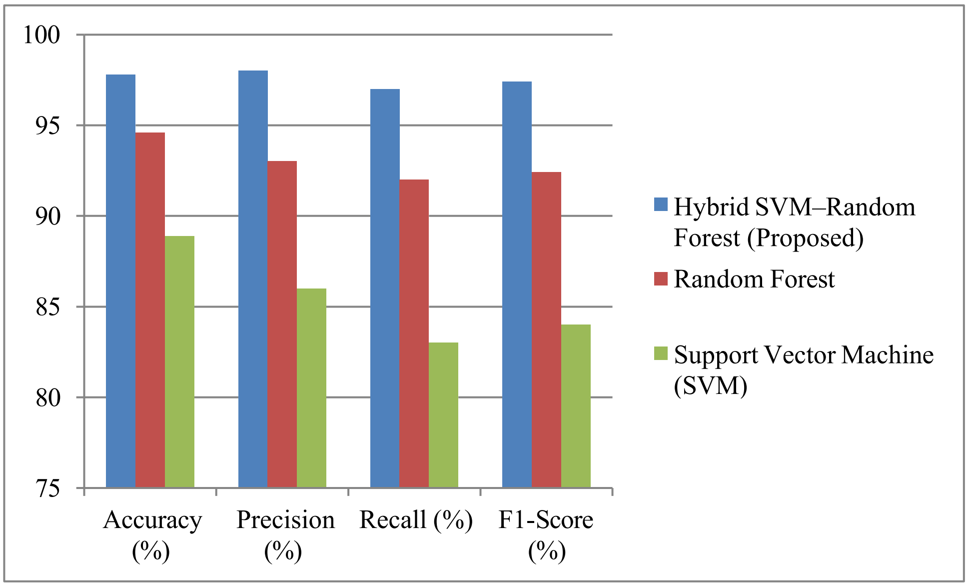

Figure 9: Graph of Performance Evaluation of Different Models

XIII. COMPARISON WITH RELATED LITERATURE

Table 4 provides a comparative study between the current machine learning methods applied to detect faults in the electric power systems. Previous research like Zaben et al. (2018) performed well with SVM and decision trees but models were restricted by features that had to be created manually. Almasoudi et al. (2017) used classical ML classifiers on constructed time-frequency features with moderate accuracy and a failing mode of complex nonlinear grid behavior. Subsequent methods of deep learning by LeCun et al. (2020) and Raza et al. (2022) were more accurate because it could learn spatio-temporal signals directly by the raw measurements. Relative to these works, the suggested hybrid SVM-Random Forest model provides superior and competitive performance, and its high accuracy and low overfit capabilities render it to be applicable in real-time fault prediction scenarios.

Table 3: Comparison of State-of-the-Art Models for Fault Prediction in Electric Power Systems

Authors [Reference]	Year	Dataset	Methods	Accuracy (%)
Zaben et al. [25]	2018	Transmission Fault Dataset	SVM + Decision Tree	89.4
Almasoudi et al. [26]	2017	Time–Frequency Fault Dataset	k-NN, RF, ANN	86.7
LeCun et al. [27]	2020	PMU Waveform Dataset	CNN/LSTM	94.1
Raza et al. [28]	2022	Real-time Power Grid Dataset	Hierarchical Deep Learning	96.5
Proposed Work	—	Electric Power System Fault Prediction Dataset	Hybrid SVM–RF	97.8
			Random Forest	94.6
			SVM	88.9

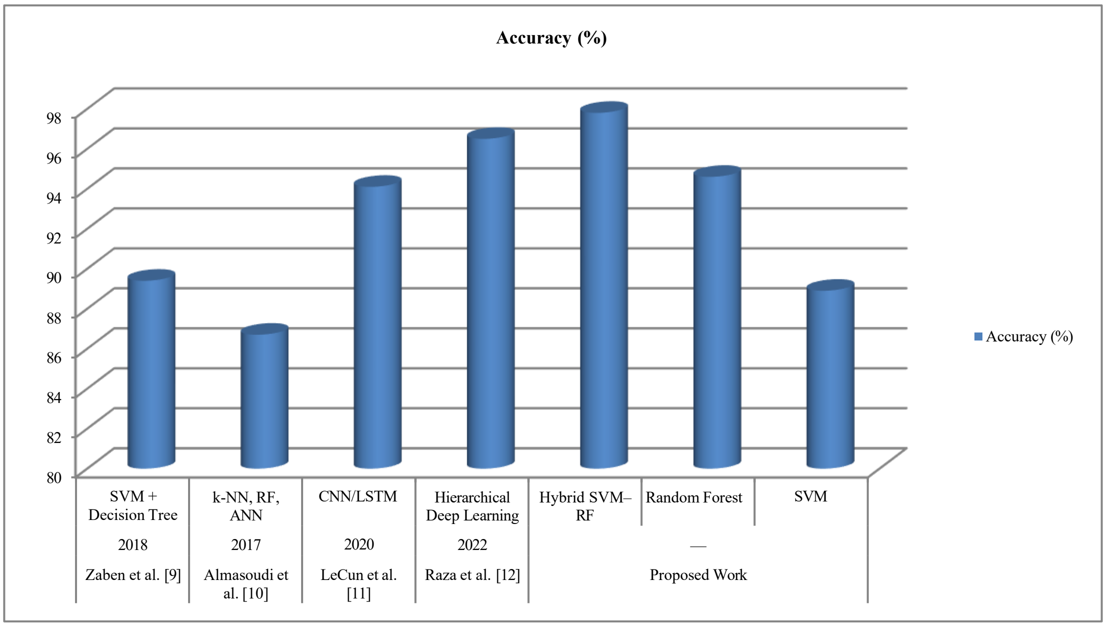

Figure 10: Comparison graph of state-of-art Models

XIV. CONCLUSION

This study presents an intelligent machine-learning framework for real-time fault prediction in electric power systems, addressing the growing challenges posed by complex grid operations, renewable integration, and high-volume sensor data. Through systematic preprocessing—including cleaning, noise filtering, normalization, feature extraction, and dimensionality reduction—the framework ensures high-quality inputs for effective model learning. Using the Electric Power System Fault Prediction Dataset, three models were trained and evaluated: SVM, Random Forest, and a Hybrid SVM–Random

Forest model. The results clearly demonstrate that while individual models perform well, the hybrid approach significantly enhances prediction accuracy and generalization. Its superior performance, marked by 97.8% accuracy, 98.0% precision, 97.0% recall, and a 97.4% F1-score, highlights its robustness in handling nonlinear system behavior and diverse fault scenarios. The findings confirm that integrating complementary ML techniques improves decision boundaries, reduces misclassification, and strengthens real-time detection capabilities. This makes the hybrid model highly suitable for deployment in modern power grids where speed, reliability, and accuracy are essential. Overall, the proposed framework provides a scalable, efficient, and accurate solution for power system fault prediction, contributing toward smarter grid operation, reduced outage risks, and enhanced system resilience. Future research may further explore deep learning integration and real-world hardware implementation for even greater efficiency.

REFERENCES

[1] Fu, Xueqian, and Yazhong Zhou. "Collaborative optimization of PV greenhouses and clean energy systems in rural areas." IEEE Transactions on Sustainable Energy 14, no. 1 (2022): 642-656.

[2] Fu, Xueqian. "Statistical machine learning model for capacitor planning considering uncertainties in photovoltaic power." Protection and Control of Modern Power Systems 7, no. 1 (2022): 1-13.

[3] Meng, Anjian, Huaizhi Wang, Saddam Aziz, Jianchun Peng, and Hui Jiang. "Kalman filtering based interval state estimation for attack detection." Energy Procedia 158 (2019): 6589-6594.

[4] Yang, Wen, Mingyu Wang, Saddam Aziz, and Ali Yousaf Kharal. "Magnitude-reshaping strategy for harmonic suppression of VSG-based inverter under weak grid." IEEE Access 8 (2020): 184399-184413.

[5] Ma, Zhengwei, Sensen Guo, Gang Xu, and Saddam Aziz. "Meta learning-based hybrid ensemble approach for short-term wind speed forecasting." IEEE Access 8 (2020): 172859-172868.

[6] Zhang, Rongquan, Gangqiang Li, Siqi Bu, Saddam Aziz, and Rizwan Qureshi. "Data-driven cooperative trading framework for a risk-constrained wind integrated power system considering market uncertainties." International Journal of Electrical Power & Energy Systems 144 (2023): 108566.

[7] Sarfraz, Muddassar, Sobia Naseem, Muhammad Mohsin, and Muhammad Shoaib Bhutta. "Recent analytical tools to mitigate carbon-based pollution: New insights by using wavelet coherence for a sustainable environment." Environmental Research 212 (2022): 113074.

[8] Chen, Wan, Baolian Liu, Muhammad Shahzad Nazir, Ahmed N. Abdalla, Mohamed A. Mohamed, Zujun Ding, Muhammad Shoaib Bhutta, and Mehr Gul. "An energy storage assessment: Using frequency modulation approach to capture optimal coordination." Sustainability 14, no. 14 (2022): 8510.

[9] Zaben, Muiz M., Muhammed Y. Worku, Mohamed A. Hassan, and Mohammad A. Abido. "Machine learning methods for fault diagnosis in ac microgrids: A systematic review." IEEE access 12 (2024): 20260-20298.

[10] Almasoudi, Fahad M. "Enhancing power grid resilience through real-time fault detection and remediation using advanced hybrid machine learning models." Sustainability 15, no. 10 (2023): 8348.

[11] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521, no. 7553 (2020): 436-444.

[12] Raza, M. F., S. Z. Iqbal, and M. Z. Iqbal. "Fault detection and diagnosis in industrial systems using machine learning techniques: A review." J. Ind. Inf. Integr 20 (2022): 100136.

[13] Cao, Xincheng, Yu Wang, Binqiang Chen, and Nianyin Zeng. "Domain-adaptive intelligence for fault diagnosis based on deep transfer learning from scientific test rigs to industrial applications." Neural Computing and Applications 33, no. 9 (2021): 4483-4499.

[14] Liu, Yinhua, Rui Sun, and Sun Jin. "A survey on data-driven process monitoring and diagnostic methods for variation reduction in multi-station assembly systems." Assembly Automation 39, no. 4 (2023): 727-739.

[15] Bello, Saifullahi Aminu, Shangshu Yu, Cheng Wang, Jibril Muhmmad Adam, and Jonathan Li. "Deep learning on 3D point clouds." Remote Sensing 12, no. 11 (2020): 1729.

[16] Yu, Naigong, Qiao Xu, and Honglu Wang. "Wafer defect pattern recognition and analysis based on convolutional neural network." IEEE Transactions on Semiconductor Manufacturing 32, no. 4 (2022): 566-573.
[17] Singh, J., and R. Kumar. "Hybrid support vector machine and convolutional neural network for fault detection in power transmission systems." IET Gener. Transm. Distrib 14 (2020): 1325-1331. https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
[18] Dabbaghjamanesh, Morteza, Josep M. Guerrero, Abdollah Kavousi-Fard, and Boyu Wang. "Guest Editorial: advanced energy internet applications in industrial power and energy systems." IEEE Transactions on Industrial Informatics 18, no. 8 (2022): 5658-5661.
[19] Lu, Dayou, Yu Liu, Qifeng Liao, Binglin Wang, Wentao Huang, and Xinze Xi. "Time-domain transmission line fault location method with full consideration of distributed parameters and line asymmetry." IEEE Transactions on Power Delivery 35, no. 6 (2020): 2651-2662.
[20] Tang, Qing, and Hail Jung. "Reliable anomaly detection and localization system: Implications on manufacturing industry." IEEE Access 11 (2023): 114613-114622.
[21] Feng, Renhai, Jiahang Li, Sheng Xie, and Xurui Mao. "Efficient training method for memristor-based array using 1T1M synapse." IEEE Transactions on Circuits and Systems II: Express Briefs 70, no. 7 (2023): 2410-2414.
[22] Yousaf, Muhammad Zain, Saqib Khalid, Muhammad Faizan Tahir, Anthony Tzes, and Ali Raza. "A novel dc fault protection scheme based on intelligent network for meshed dc grids." International Journal of Electrical Power & Energy Systems 154 (2023): 109423.
[23] Orhan, Muhittin, and Metin Celik. "A literature review and future research agenda on fault detection and diagnosis studies in marine machinery systems." Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment 238, no. 1 (2024): 3-21.
[24] Zaben, Zhi, Yanfeng Wang, Dawei Zhang, Lili Wen, Haiyang Liu, Zhijie Xiong, Khan Wajid, and Renhai Feng. "Digital twin modeling for hydropower system based on radio frequency identification data collection." Electronics 13, no. 13 (2018): 2576.
[25] Almasoudi , Xu, Yongjun, Xin Liu, Xin Cao, Changping Huang, Enke Liu, Sen Qian, Xingchen Liu et al. "Artificial intelligence: A powerful paradigm for scientific research." The Innovation 2, no. 4 (2017).
[26] LeCun, Feng, Renhai, Wanqi Yuan, Leijiao Ge, and Siyu Ji. "Nonintrusive load disaggregation for residential users based on alternating optimization and downsampling." IEEE Transactions on Instrumentation and Measurement 70 (2020): 1-12.
[27] Raza, Zain Yousaf, Muhammad, Hui Liu, Ali Raza, and Mirza Baber Baig. "Primary and backup fault detection techniques for multi-terminal HVdc systems: a review." IET Generation, Transmission & Distribution 14, no. 22 (2022): 5261-5276.