Electrical Faults Analysis and Classification

DIVESHWARAN P
ELECTRICAL AND COMPUTER ENGINEERING
AMRITA SCHOOL OF ENGINEERING
AMRITA VISHWA VIDHYAPEETHAM, COIMBATORE
cb.en.u4elc22010@cb.students.amrita.edu

ABHIRM REDDY K
ELECTRICAL AND COMPUTER ENGINEERING
AMRITA SCHOOL OF ENGINEERING
AMRITA VISHWA VIDYAPEETHAM, COIMBATORE
cb.en.u4elc22018@cb.students.amrita.edu

S KARTHIK
ELECTRICAL AND COMPUTER ENGINEERING
AMRITA SCHOOL OF ENGINEERING
AMRITA VISHWA VIDYAPEETHAM, COIMBATORE
cb.en.u4elc22049@cb.students.amrita.edu

ANIL K
ELECTRICAL AND COMPUTER ENGINEERING
AMRITA SCHOOL OF ENGINEERING
AMRITA VISHWA VIDYAPEETHAM, COIMBATORE
cb.en.u4elc22065@cb.students.amrita.edu

ABSTRACT: **Electrical faults pose significant risks in power systems, leading to equipment damage, operational downtime, and safety concerns. Traditional fault classification techniques are time-consuming and require domain expertise. This project explores the application of Machine Learning (ML) models—Support Vector Machine (SVM), XGBoost, Random Forest, and Multi-Layer Perceptron (MLP)—to classify electrical faults based on sensor data. The dataset consists of class data and fault detection data, where dimensionality reduction using Principal Component Analysis (PCA) is applied to improve efficiency. Performance is evaluated using F1-score and confusion matrix to ensure accurate classification.**

Keywords: **Electrical Fault Classification, Machine Learning, Support Vector Machine (SVM), XGBoost, Random Forest, Multi-Layer Perceptron (MLP), Principal Component Analysis (PCA), Fault Detection, Power System Reliability.**

I. INTRODUCTION:

Electrical faults in power systems can lead to severe disruptions, including equipment failure, economic losses, and safety hazards. Identifying and classifying electrical faults accurately and promptly is crucial to ensuring the reliability and efficiency of electrical power networks. Traditional fault classification techniques, such as rule-based systems and manual expert analysis, often fail to meet the needs of modern power grids due to their complexity and scale.

Machine Learning (ML) has emerged as a powerful alternative for fault analysis, offering automated classification and real-time monitoring capabilities. This project employs ML models, including Support Vector Machine (SVM), XGBoost, Random Forest, and Multi-Layer Perceptron (MLP), to analyze and classify electrical faults. By leveraging historical fault data, these models can identify fault types such as Line-to-Ground (LG), Line-to-Line (LL), Double Line-to-Ground (LLG), Triple Line-to-Ground (LLLG), and Three-Phase (LLL) faults.

To improve computational efficiency and classification accuracy, Principal Component Analysis (PCA) is used for dimensionality reduction. The effectiveness of the models is evaluated using performance metrics such as the F1-score and confusion matrix, ensuring a robust and accurate fault classification system.

II. LITERATURE SURVEY:

Traditional Fault Detection Methods

Conventional techniques such as rule-based expert systems and phasor measurement units (PMUs) rely on predefined thresholds and domain expertise. While effective for basic fault classification, these methods struggle with large-scale power grids and complex fault scenarios.

Digital fault recorders (DFRs) have improved real-time monitoring, but their accuracy depends on empirical tuning and may not generalize well to unseen fault conditions.

Machine Learning in Fault Classification

Several studies have demonstrated the effectiveness of ML models in classifying electrical faults. Support Vector Machine (SVM) is widely used for its robustness in high-dimensional spaces and strong generalization capability.

XGBoost, a gradient boosting algorithm, has shown superior performance in classification tasks by efficiently handling complex, nonlinear relationships within the data.

Random Forest, as an ensemble learning method, enhances fault detection accuracy by reducing overfitting and improving the stability of predictions.

Multi-Layer Perceptron (MLP), a deep learning approach, has proven beneficial in capturing nonlinear patterns and dependencies between electrical parameters.

Dimensionality Reduction with PCA

Feature selection and dimensionality reduction techniques such as PCA help improve model efficiency by removing redundant and correlated features.

Studies indicate that applying PCA before classification enhances model generalization, reduces computational complexity, and mitigates the risk of overfitting.

III. METHODOLOGY:

This study follows a structured approach to predict power consumption using AI techniques. The methodology includes data preprocessing, training machine learning models, evaluating their performance using key metrics, and implementing a GUI for real-time predictions. The models used include XGBoost, Random Forest, MLP and SVM, each optimized for handling weather-based power consumption forecasting.

- **Dataset Preparation**

  The dataset includes electrical fault detection and classification data with sensor readings representing system parameters.

  The target variable corresponds to different types of electrical faults. The dataset is split into training and testing sets using stratified sampling to maintain class balance.

- **Data Cleaning and Transformation**

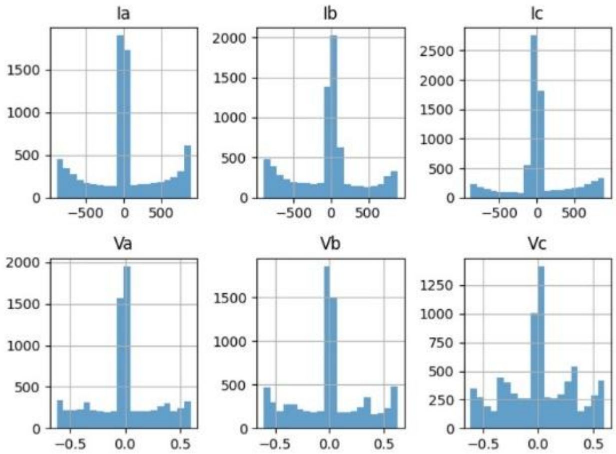
  Handling missing values through imputation or removal.

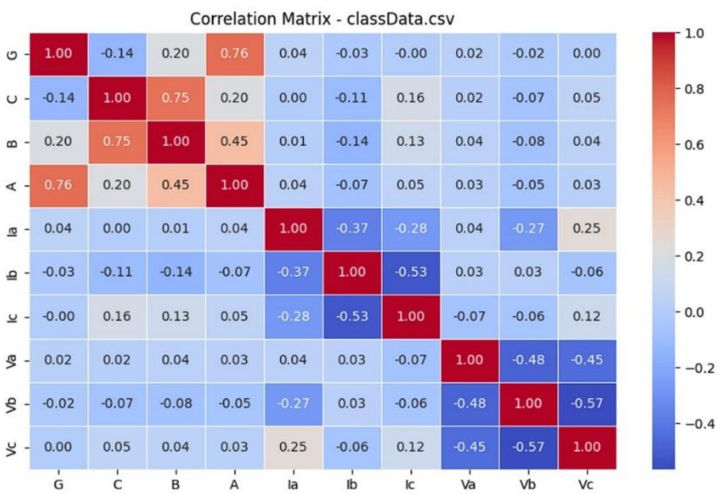
  Dropping irrelevant features, such as unnamed or non-informative columns.

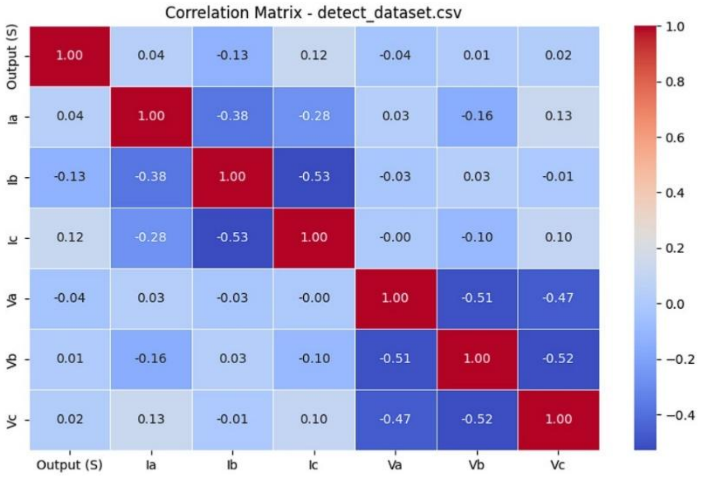
  Applying normalization techniques while avoiding StandardScaler, as per project constraints.

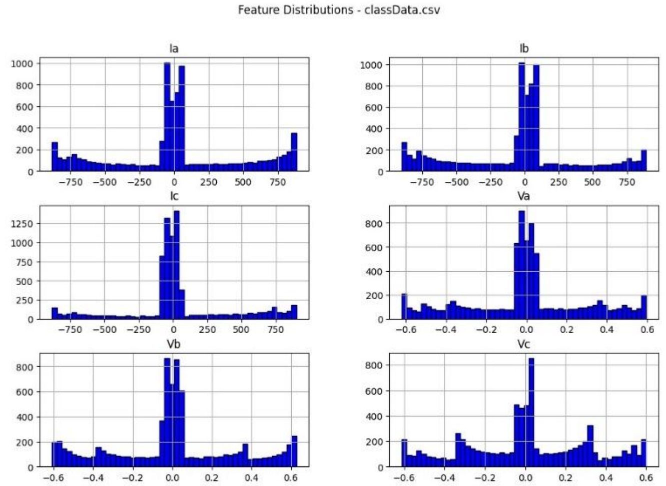
  Using PCA to retain 95% variance while reducing dimensionality.

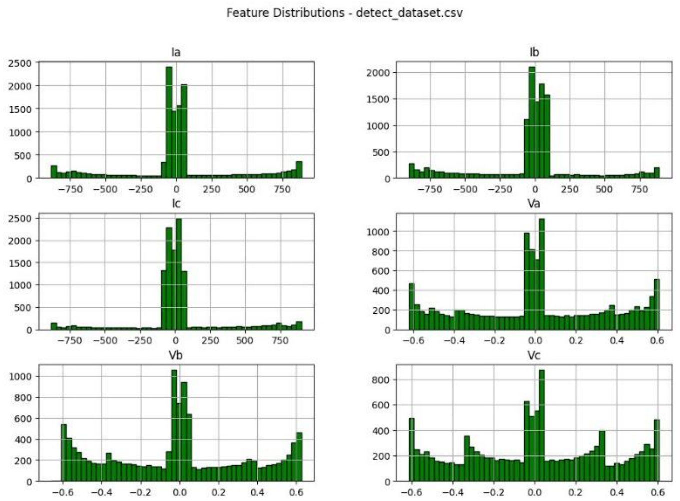
  Implementing K-Fold Cross-Validation (K=5) to enhance model robustness and reduce bias.











Model Implementation

Support Vector Machine (SVM)

- SVM is used for classification by constructing hyperplanes that best separate fault categories.
- The radial basis function (RBF) kernel is employed to handle non-linearity in the dataset.
- Hyperparameter tuning is performed using grid search to optimize `C` and `gamma` values.

XGBoost

- XGBoost applies gradient boosting to iteratively refine fault classification performance.
- The model uses decision trees as weak learners and enhances them with boosting mechanisms.
- Cross-validation is used to optimize hyperparameters such as learning rate and tree depth.

Random Forest

- This ensemble learning model builds multiple decision trees to improve fault classification accuracy.
- Bootstrap aggregation (bagging) is used to minimize overfitting and enhance model stability.
- Feature importance analysis helps in understanding the impact of different parameters on classification.

Multi-Layer Perceptron (MLP)

- MLP is implemented with multiple hidden layers using backpropagation for training.
- Dropout layers are incorporated for regularization to prevent overfitting.
- The Adam optimizer and categorical cross-entropy loss function are used for efficient training.
- Learning rate tuning is performed using a scheduler to accelerate convergence.

By leveraging these ML techniques and preprocessing steps, this project aims to develop a robust, scalable, and accurate fault classification system for electrical power networks.

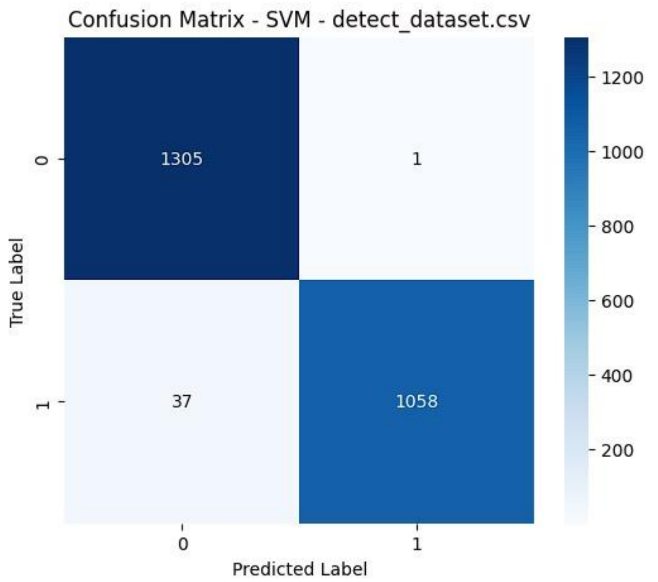
IV. RESULT:

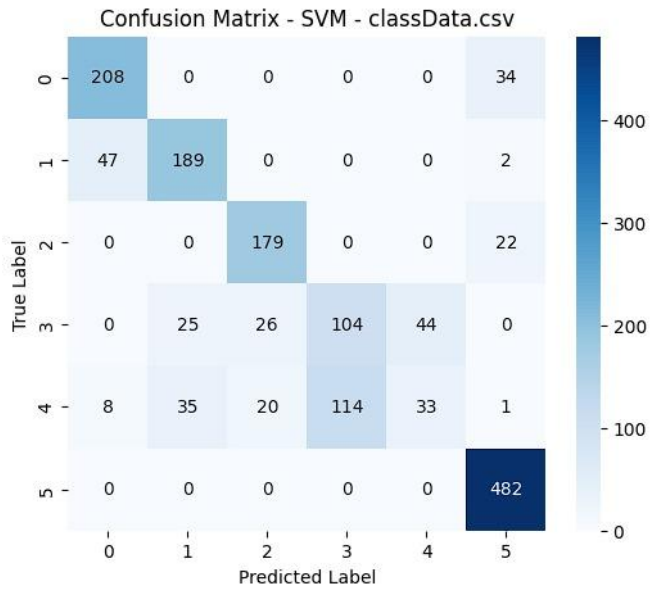
The models were evaluated using MAE, Accuracy, Precision, Recall, F1 Score, and Confusion Matrix.

(i) SVM:

```
--- SVM - detect_dataset.csv Results ---
Accuracy: 0.9841732611411912
Classification Report:
              precision    recall  f1-score   support

           0       0.97      1.00      0.99      1306
           1       1.00      0.97      0.98      1095

    accuracy                           0.98      2401
   macro avg       0.99      0.98      0.98      2401
weighted avg       0.98      0.98      0.98      2401
```

```
--- SVM - classData.csv Results ---
Accuracy: 0.7596948506039415
Classification Report:
              precision    recall  f1-score   support

           0       0.79      0.86      0.82       242
           1       0.76      0.79      0.78       238
           2       0.80      0.89      0.84       201
           3       0.48      0.52      0.50       199
           4       0.43      0.16      0.23       211
           5       0.89      1.00      0.94       482

    accuracy                           0.76      1573
   macro avg       0.69      0.70      0.69      1573
weighted avg       0.73      0.76      0.73      1573
```

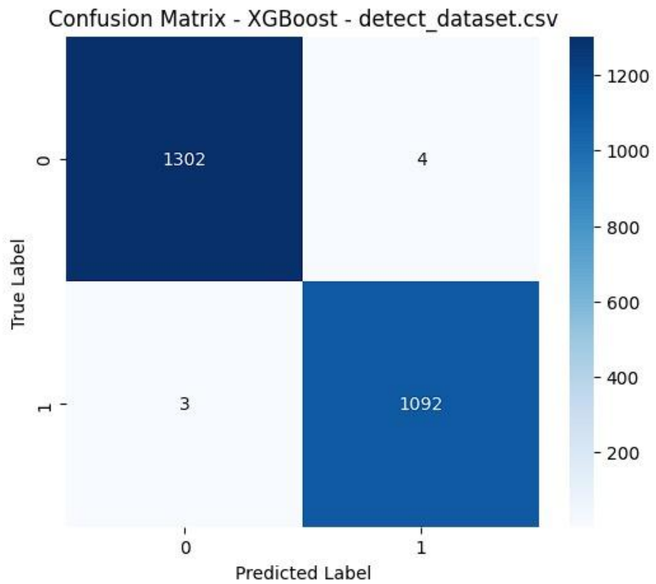


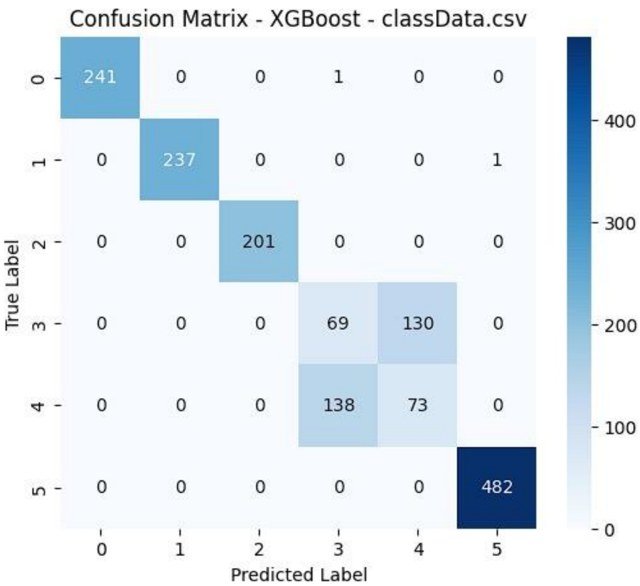


(ii) XGBOOST:

```
--- XGBoost - detect_dataset.csv Results ---
Accuracy: 0.9970845481049563
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1306
           1       1.00      1.00      1.00      1095

    accuracy                           1.00      2401
   macro avg       1.00      1.00      1.00      2401
weighted avg       1.00      1.00      1.00      2401
```

```
--- XGBoost - classData.csv Results ---
Accuracy: 0.828353464717101
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       242
           1       1.00      1.00      1.00       238
           2       1.00      1.00      1.00       201
           3       0.33      0.35      0.34       199
           4       0.36      0.35      0.35       211
           5       1.00      1.00      1.00       482

    accuracy                           0.83      1573
   macro avg       0.78      0.78      0.78      1573
weighted avg       0.83      0.83      0.83      1573
```

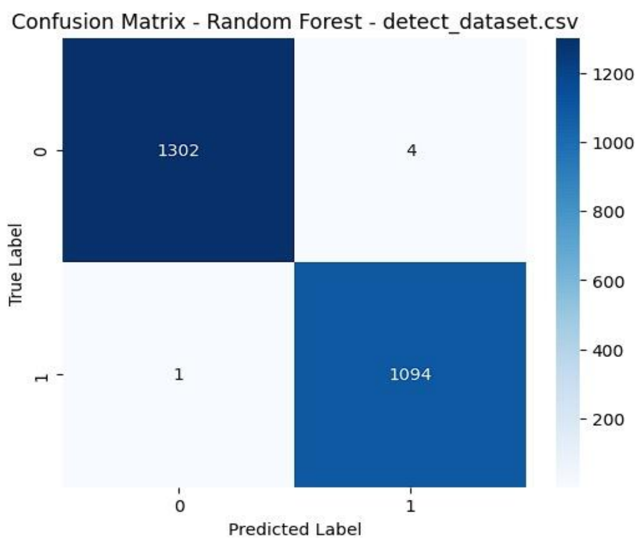


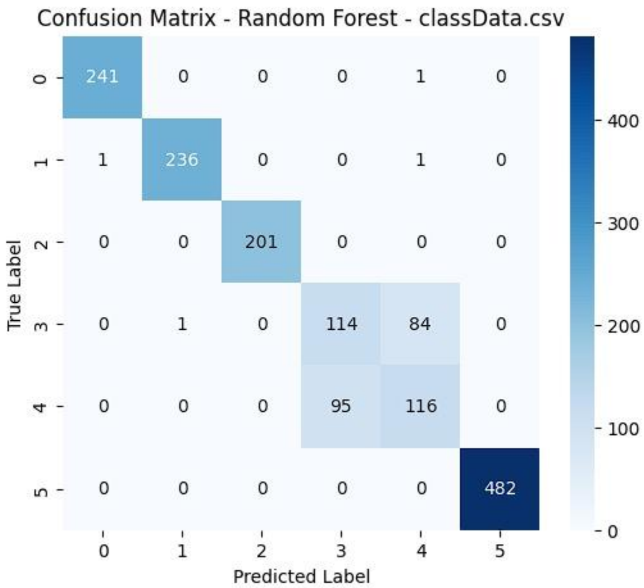


(iii) RANDOM FOREST:

```
--- Random Forest - detect_dataset.csv Results ---
Accuracy: 0.9979175343606831
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1306
           1       1.00      1.00      1.00      1095

    accuracy                           1.00      2401
   macro avg       1.00      1.00      1.00      2401
weighted avg       1.00      1.00      1.00      2401
```

```
--- Random Forest - classData.csv Results ---
Accuracy: 0.8836617927527018
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       242
           1       1.00      0.99      0.99       238
           2       1.00      1.00      1.00       201
           3       0.55      0.57      0.56       199
           4       0.57      0.55      0.56       211
           5       1.00      1.00      1.00       482

    accuracy                           0.88      1573
   macro avg       0.85      0.85      0.85      1573
weighted avg       0.88      0.88      0.88      1573
```

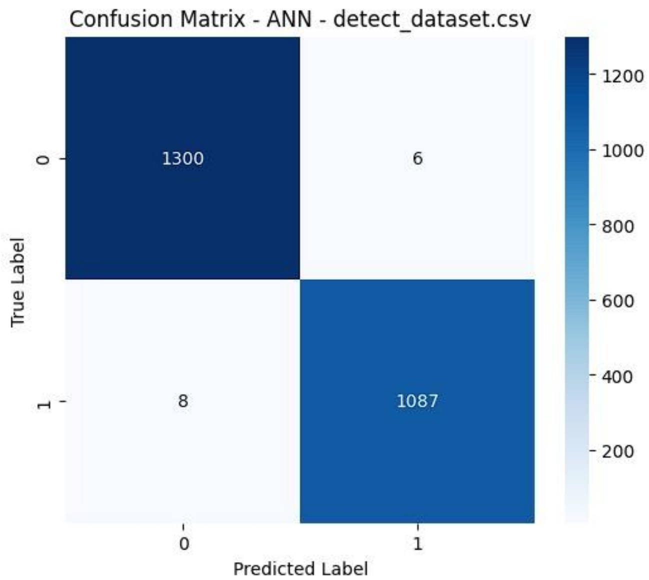


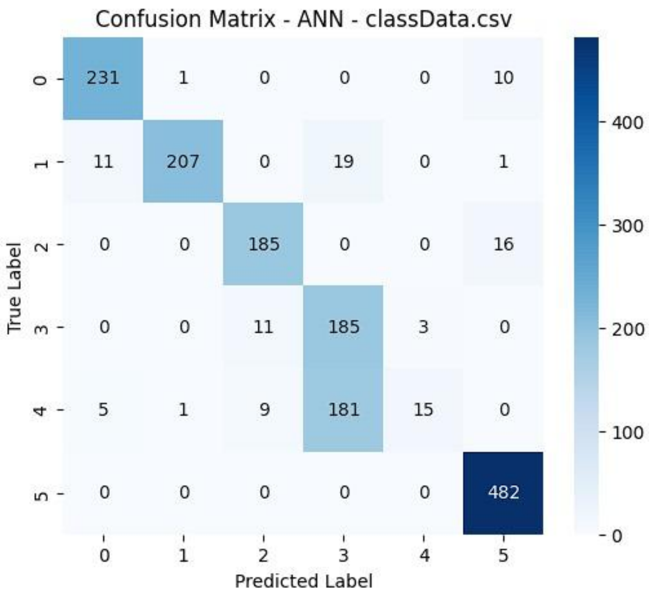


(iv)MLP:

```
--- ANN - detect_dataset.csv Results ---
Accuracy: 0.9941690962099126
Classification Report:
              precision    recall  f1-score   support

           0       0.99      1.00      0.99      1306
           1       0.99      0.99      0.99      1095

    accuracy                           0.99      2401
   macro avg       0.99      0.99      0.99      2401
weighted avg       0.99      0.99      0.99      2401
```

```
--- ANN - classData.csv Results ---
Accuracy: 0.8296249205340115
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.95      0.94       242
           1       0.99      0.87      0.93       238
           2       0.90      0.92      0.91       201
           3       0.48      0.93      0.63       199
           4       0.83      0.07      0.13       211
           5       0.95      1.00      0.97       482

    accuracy                           0.83      1573
   macro avg       0.85      0.79      0.75      1573
weighted avg       0.87      0.83      0.80      1573
```





V. CONCLUSION:

In this study, we explored the application of machine learning techniques—Support Vector Machine (SVM), XGBoost, Random Forest, and Multi-Layer Perceptron (MLP)—for the classification of electrical faults in power systems. Traditional fault detection methods are often time-consuming and reliant on expert knowledge, whereas ML models offer an automated and scalable approach for fault analysis.

Our methodology included data preprocessing techniques such as missing value handling, feature selection, and dimensionality reduction using Principal Component Analysis (PCA) to improve model efficiency. The models were trained and evaluated using performance metrics such as accuracy, precision, recall, F1-score, and confusion matrices.

The results indicate that ML models effectively classify different types of electrical faults, with each model showing varying levels of performance. The ensemble-based methods, such as XGBoost and Random Forest, provided robust and stable predictions, while MLP leveraged deep learning techniques to capture complex patterns in fault data. SVM demonstrated strong classification capabilities, particularly in high-dimensional spaces.

This study demonstrates that ML-based fault classification can significantly enhance the reliability of power systems by enabling real-time fault detection and reducing manual intervention. Future work can focus on integrating real-time data from IoT-enabled sensors, optimizing model performance with advanced feature engineering, and deploying the trained models in edge computing environments for on-site fault analysis.

VI. REFERNCES:

[1] Baskar. D, Dr. Selvam. P,“ Machine Learning Framework For Power System Fault Detection And Classification” ,20 issn 2277-8616- 2032,2020.

[2]AlokMukherjee, Palash Kumar Kundu,Arabinda Das, “ Transmission Line Faults in Power System and the Different Algorithms for Identification, Classification and Localization:” J. Inst. Eng. India Ser. B ,2018, in press.

[3] S.A. Aleem, N. Shahid, I.H. Naqvi, “Methodologies in power systems fault detection and diagnosis”. Energy Syst. 6(1), 85–108 ,2017, in press.

[4] K. Chen, C. Huang, J. He, “Fault detection, classification and location for transmission lines and distribution systems: a review on the methods”,2016,in press.

[5]Jamil, M., Sharma, S.K. & Singh, R. “Fault detection and classification in electrical power transmission system using artificial neural network.” SpringerPlus 4, 334(2015).

[6] AvagaddiPrasad , J. Belwin Edward, K. Ravi, “ A review on fault classification methodologies in power transmission systems”, School of Electrical Engineering, VIT University,2018, in press.

[7] Wang, Xinan, "Machine learning applications in power systems". Electrical Engineering Theses and Dissertations,2018,in press.

[8] R. S. Meena & M. K. Lodha, “Unsymmetrical fault analysis & protection of the existing power system”, Department of Electrical Engineering, Sri Balaji College of Engineering & Technology, Jaipur, Rajasthan Rajasthan Technical University Kota,International Journal of Multidisciplinary Research and Modern Education (IJMRME) ISSN : 2454 - 6119 ,Volume I, Issue I, 2015.

[9] K. Moloi and A. O. Akumu, "Power distribution fault diagnostic method based on machine learning technique," 2019 IEEE PES/IAS PowerAfrica, 2019, pp.238242,doi:10.1109/PowerAfrica.2019.8928633