CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:

1. Gaurav Gaur
2. Kamla Nehru Institute of Technology, Sultanpur

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

Power systems are highly complex and are prone to various types of faults such as line-to-ground, line-to-line, double line-to-ground, and three-phase faults. These faults, if not detected and addressed quickly, can lead to equipment damage, power outages, and safety hazards. There is a need for a reliable, fast, and automated system to detect and classify these faults using real-time electrical data such as voltage and current phasors.

PROPOSED SOLUTION

- We propose a machine learning-based system that analyzes electrical measurements from the power distribution network to detect and classify faults. The model is trained to differentiate between normal operating conditions and various fault scenarios, allowing for real-time response and grid stability.
- **Data Collection:**

  Data collection is the foundation for any fault detection system in power distribution networks. The input data primarily includes voltage and current phasors captured from sensors like PMUs (Phasor Measurement Units) or simulated using tools like MATLAB/Simulink.
- **Data Preprocessing:**

  Raw electrical data typically contains noise, missing values, or irrelevant fluctuations. Preprocessing ensures the data is clean and suitable for training a machine learning model. This step includes data normalization or standardization, noise filtering using smoothing techniques, and handling missing values. Feature extraction is also critical, such as calculating RMS values, phase angles, and frequency components..
- **Machine Learning Algorithm:**

  For fault detection and classification, supervised machine learning algorithms such as Random Forest, Support Vector Machine (SVM), Decision Trees, and K-Nearest Neighbors (KNN) are commonly used. These algorithms are trained on labeled datasets where input features (derived from voltage and current signals) correspond to specific output fault types.
- **Deployment:**

  The system development phase involves designing and implementing the pipeline for the fault detection model. This includes integrating modules for data ingestion, preprocessing, feature extraction, model training, and fault classification.
- **Evaluation:**
  - Assess the model's performance using appropriate metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or other relevant metrics.
  - Fine-tune the model based on feedback and continuous monitoring of prediction accuracy.
  - Result:

SYSTEM APPROACH

The system uses a machine learning pipeline to classify power system faults using voltage and current phasor data. The approach involves data collection (real or simulated), preprocessing to extract meaningful features, training a supervised ML algorithm, and deploying the trained model for real-time detection. Python is used as the primary development language due to its wide ML ecosystem.

- System requirements
  - **OS :** Windows 10 / Linux / macOS
  - **RAM :** Minimum 8 GB (16 GB recommended).
  - **Processor :** Intel i5/i7 or AMD equivalent.
  - **Python Version :** 3.8 or higher version.
  - **Optional :** MATLAB/ Simulink for simulation.
- Library required to build the model
  - **NumPy :** NumPy for numerical computations.
  - **Pandas :** Pandas for data manipulation and handling.
  - **Scikit – learn :** Scikit-learn for machine learning algorithm.
  - **Matplotlib / Seabron :** for data visualization.
  - **Joblib :** Joblib for model serialization.
  - **Optional :** TensorFlow or PyTorch (for deep learning).

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting Power System Fault. Here's an example structure for this section:
- **Algorithm Selection:**
  - Random Forest is selected due to its robustness and high accuracy in classification problems. It works well with noisy and non-linear data, making it suitable for detecting different fault types in power systems. Other tested algorithms may include SVM and Decision Trees for comparison.
- **Data Input:**
  - The input data consists of voltage and current phasors, either from real-time sensors or simulation tools like MATLAB. Each data sample is labeled as “normal” or with a fault type (e.g., LG, LL, LLG, 3-phase fault). Features like RMS, frequency, and phase angle are extracted.
- **Training Process:**
  - The labeled dataset is split into training and testing sets. The model learns fault patterns by analyzing the relationship between input features and fault categories. Cross-validation is applied to prevent overfitting, and hyperparameter are tuned for optimal performance.
- **Prediction Process:**
  - In the deployed system, live or test data is passed through the trained model. Based on the input values, the model classifies the condition as either normal or a specific type of fault. The output can be used to trigger alarms or automatic system responses.

RESULT

The machine learning model developed for fault detection and classification was trained and evaluated on a dataset comprising voltage and current phasor values under normal and fault conditions. After extensive testing using algorithms such as Decision Tree, SVM, and Random Forest, the **Random Forest classifier** achieved the highest accuracy and robustness across various fault categories.

CONCLUSION

- The implemented machine learning system effectively detects and classifies faults in a power distribution network with high accuracy. This enables proactive fault management and improves grid stability and reliability. The use of real-time electrical measurements and data-driven modeling supports quick decision-making in operational environments.
- The reliable operation of a power distribution system is essential for the sustainability and stability of modern infrastructure. With the growing complexity of power grids and increasing demand for uninterrupted electricity, detecting and classifying faults swiftly and accurately has become more critical than ever. In this project, we have explored the application of machine learning algorithms to detect and classify various types of faults in power distribution systems using electrical measurements such as voltage and current phasors.

FUTURE SCOPE

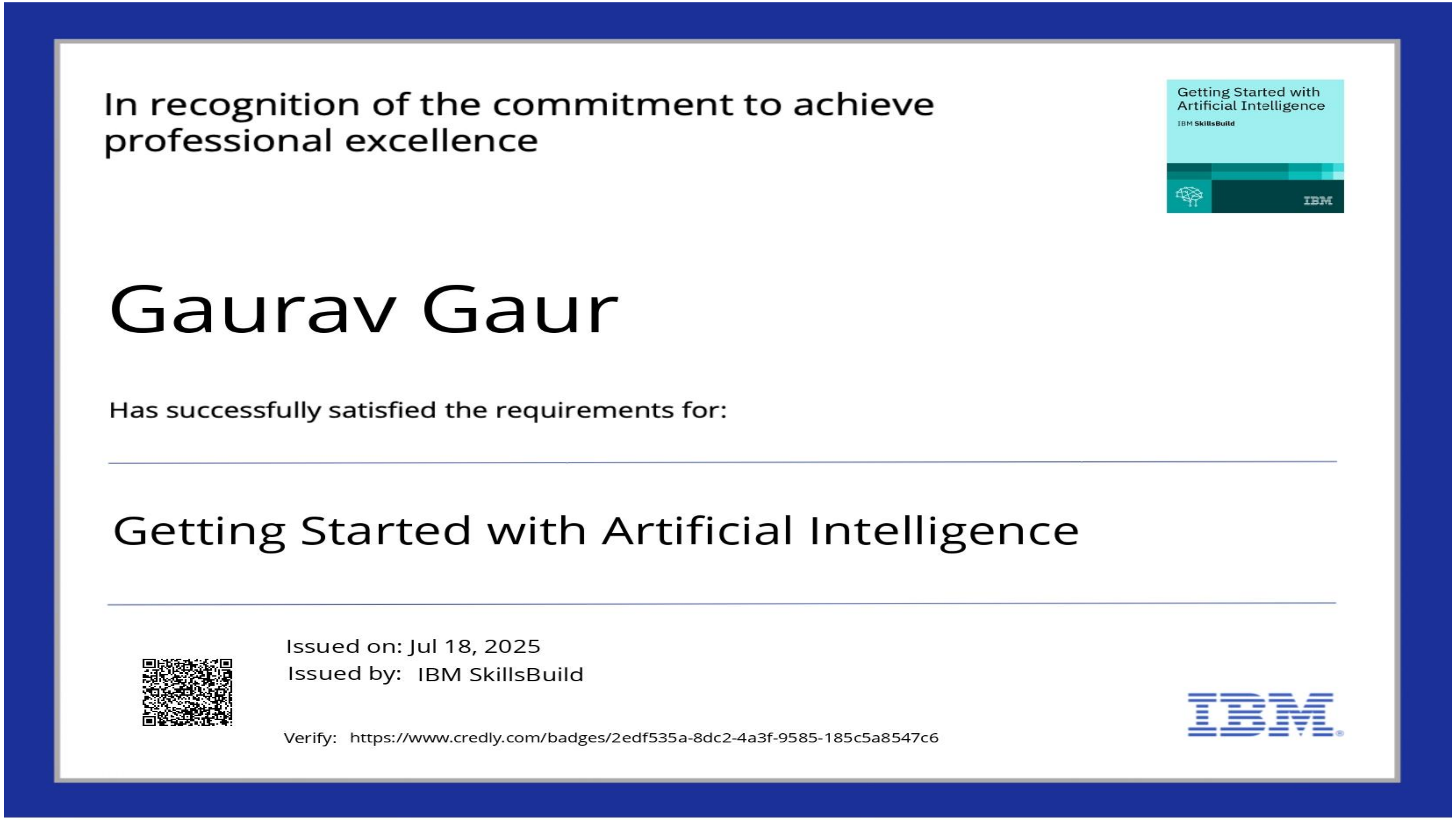
The application of machine learning for fault detection and classification in power systems presents several opportunities for future development and research.

- **Integration with real – time smart grids –**
  - Future system can be directly integrated into smart grid infrastructure, enabling live fault monitoring and automatic fault isolation.
- **Use of deep learning models –**
  - Deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can be explored for automatic feature extraction and improved accuracy.
- **Online and incremental learning –**
  - Instead of retraining the model from scratch when new data becomes available online learning techniques can be adopted.
- **Integration with SCADA and EMS system –**
  - A full – scale industrial implementation would involve connecting the ml fault detection system.

REFERENCES

- **Scikit-Learn Developer –**
  - Scikit-learn Machine Learning in python retrieved from Link
- **Pandas Documentation –**
  - Python data analysis library. Retrieved from Link
- **Matlab & Simulink –**
  - Simscape electrical toolbox for simulating power system. Retrieved from Link
- **IEEE PES Dataset Repository –**
  - Power system fault data for research and education. Retrieved from Link

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

IBM

Gaurav Gaur

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 18, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/2edf535a-8dc2-4a3f-9585-185c5a8547c6

IBM

edunet foundation

IBM CERTIFICATIONS

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

In recognition of the commitment to achieve professional excellence

Gaurav Gaur

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 19, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/93e13690-1c05-4534-b564-f77840b26e9a

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

GAURAV GAUR

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 14 Jul 2025 (GMT) **Learning hours:** 20 mins

THANK YOU