PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
Tanisha Singh – GLA University – B.tech (CSE)

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

- Design a machine learning model to detect and classify different types of faults in a power distribution system. Using electrical measurement data (e.g., voltage and current phasors), the model should be able to distinguish between normal operating conditions and various fault conditions (such as line-to-ground, line-to-line, or three-phase faults). The objective is to enable rapid and accurate fault identification, which is crucial for maintaining power grid stability and reliability.

PROPOSED SOLUTION

- The proposed system addresses the detection and classification of power distribution faults using machine learning to ensure timely fault identification and grid reliability. The solution includes:
- **Data Collection:** Use the Kaggle dataset containing voltage and current phasors under various fault conditions (LG, LL, DLG, 3Φ, Normal).
- **Data Preprocessing:** Handle missing values, normalize features, encode fault types, and split data into train-test sets.
- **ML Algorithm:** Train a **Random Forest Classifier** to distinguish fault types; also evaluate models like Decision Tree and SVM.
- **Deployment:** Deploy the model as a REST API via **IBM Watson Machine Learning** with data stored on **IBM Cloud Object Storage** and developed using **Watson Studio**.
- **Evaluation:** Use metrics like Accuracy, Precision, Recall, Confusion Matrix; apply cross-validation and monitor for retraining.

SYSTEM APPROACH

System requirements

•**IBM Cloud Account** (Lite plan)
•**IBM Watson Studio** – for development and training environment
•**IBM Cloud Object Storage** – to store datasets securely
•**IBM Watson Machine Learning** – for model deployment and serving as API
•Internet browser and stable network connection

Library required to build the model

•**Pandas** – for data manipulation and analysis
•**NumPy** – for numerical operations
•**Matplotlib / Seaborn** – for data visualization (EDA)
•**Scikit-learn** – for preprocessing, model training, and evaluation
•**Joblib or Pickle** – to save the trained model
•**Watson Machine Learning SDK (ibm-watson-machine-learning)** – to deploy the model on IBM Cloud

ALGORITHM & DEPLOYMENT

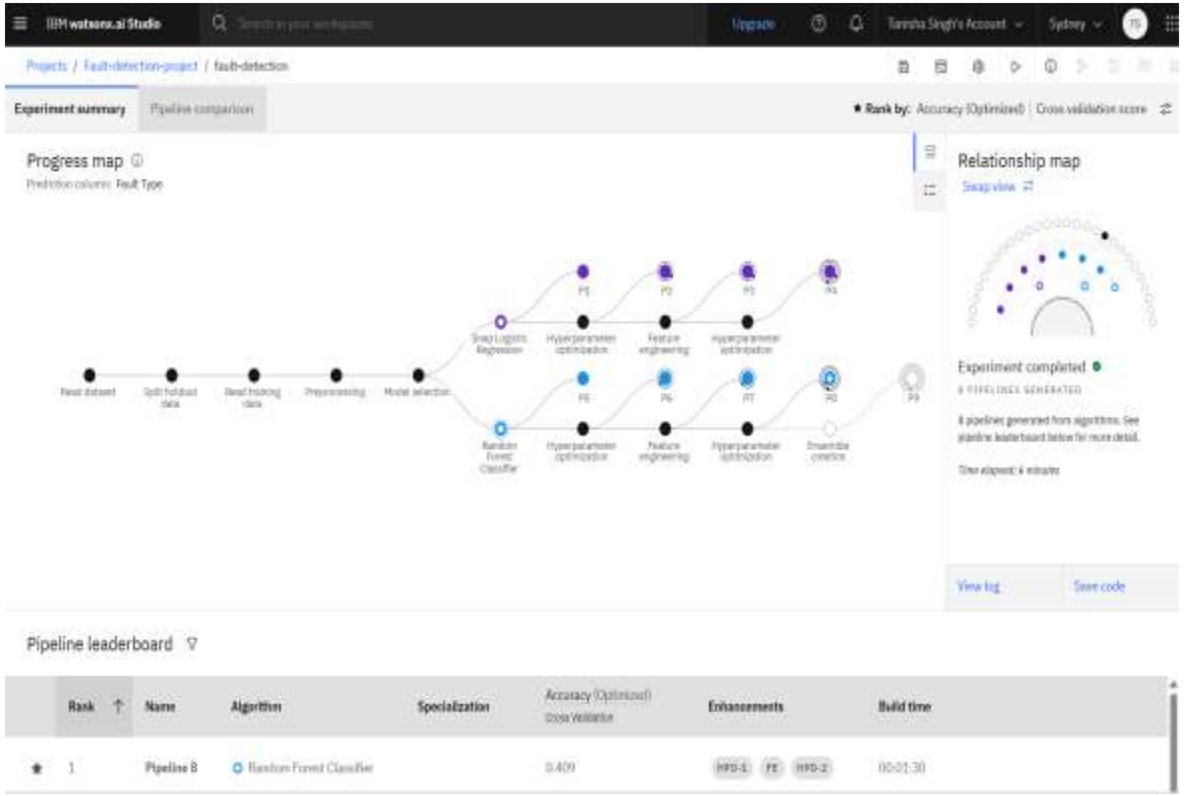
- **Algorithm Used:** Random Forest Classifier — chosen for its accuracy, robustness, and suitability for multi-class classification problems.
- **Input Features:** Voltage & current phasors, real/reactive power, frequency, and fault type labels.
- **Training Process:** Dataset split (80:20), preprocessing done, and model trained using cross-validation for fault classification.
- **Prediction:** Real-time classification of faults as *No Fault*, *LG*, *LL*, or *3-Phase Fault* using phasor input.
- **Deployment:** Model deployed on **IBM Watson Machine Learning** with a REST API for live predictions via IBM Cloud Lite.

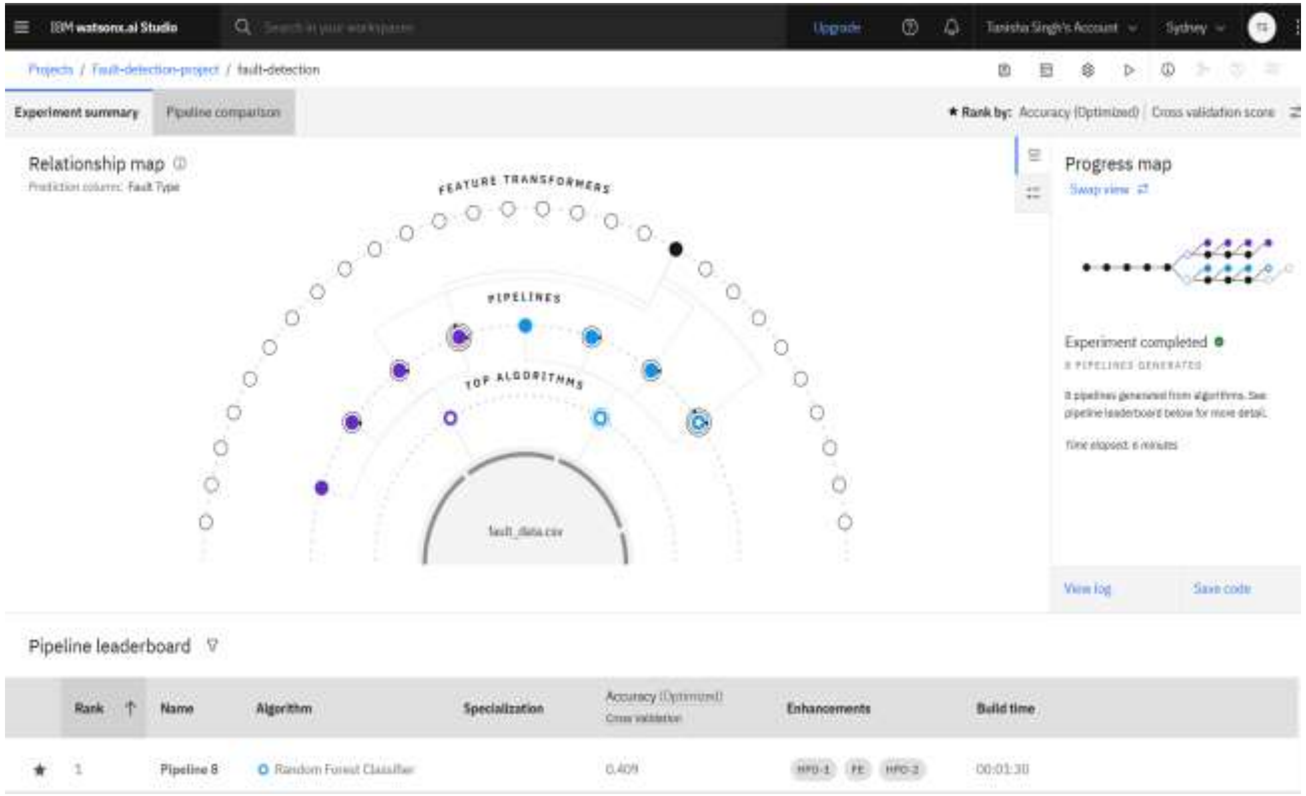
RESULT

Fault Detection Model – IBM watsonx.ai Studio

Trained multiple ML pipelines using IBM watsonx.ai Studio to classify power system faults.

- Best model: **Random Forest Classifier**
- Accuracy: **40.9% (Cross-validation)**
- Total Pipelines: **8** (with HPO & Feature Engineering)
- Dataset: fault_data.csv

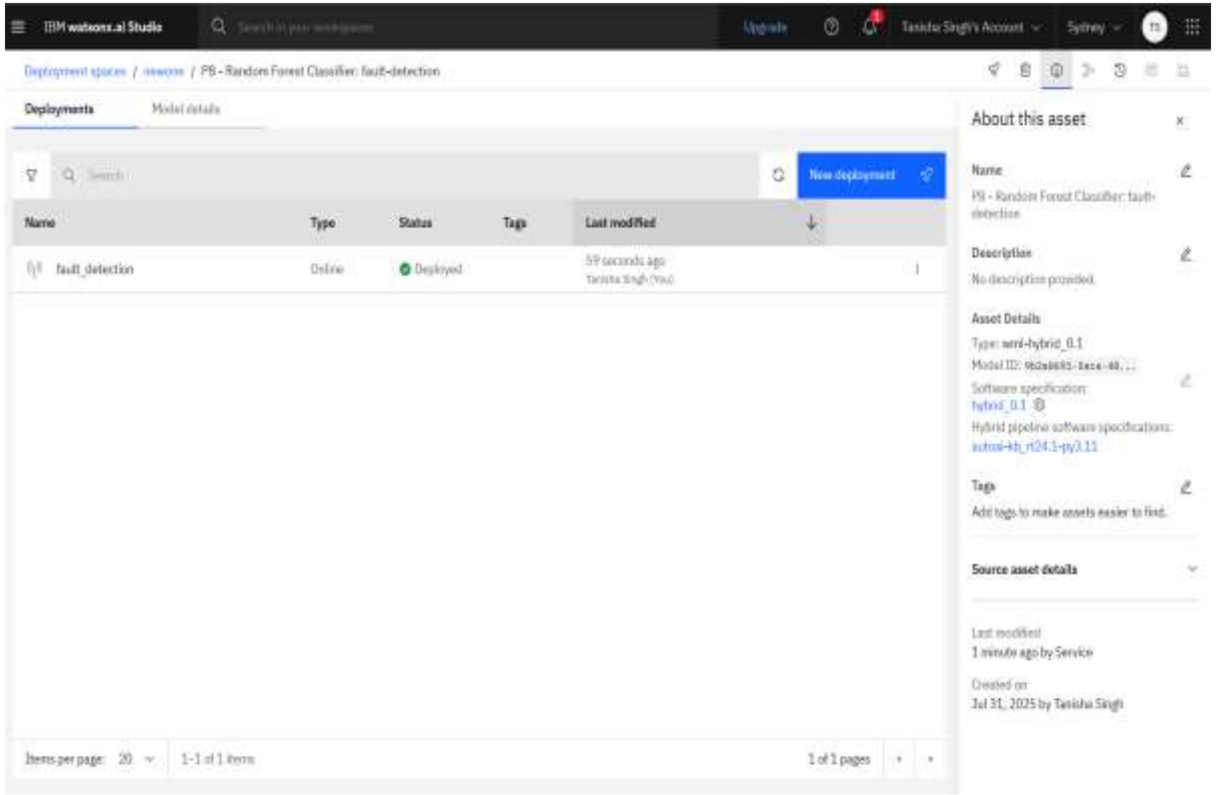




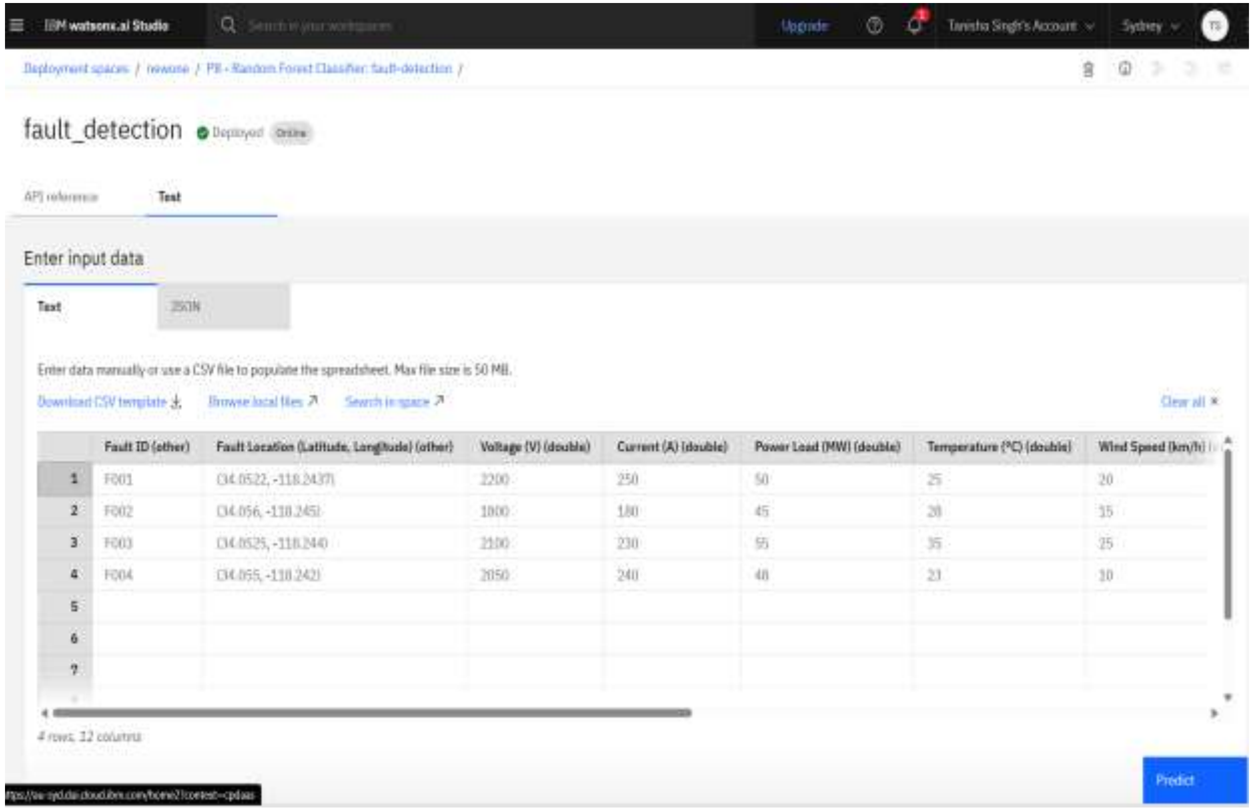
edunet foundation

RESULT

- The model is deployed successfully.



- Checking the model by giving some input data.



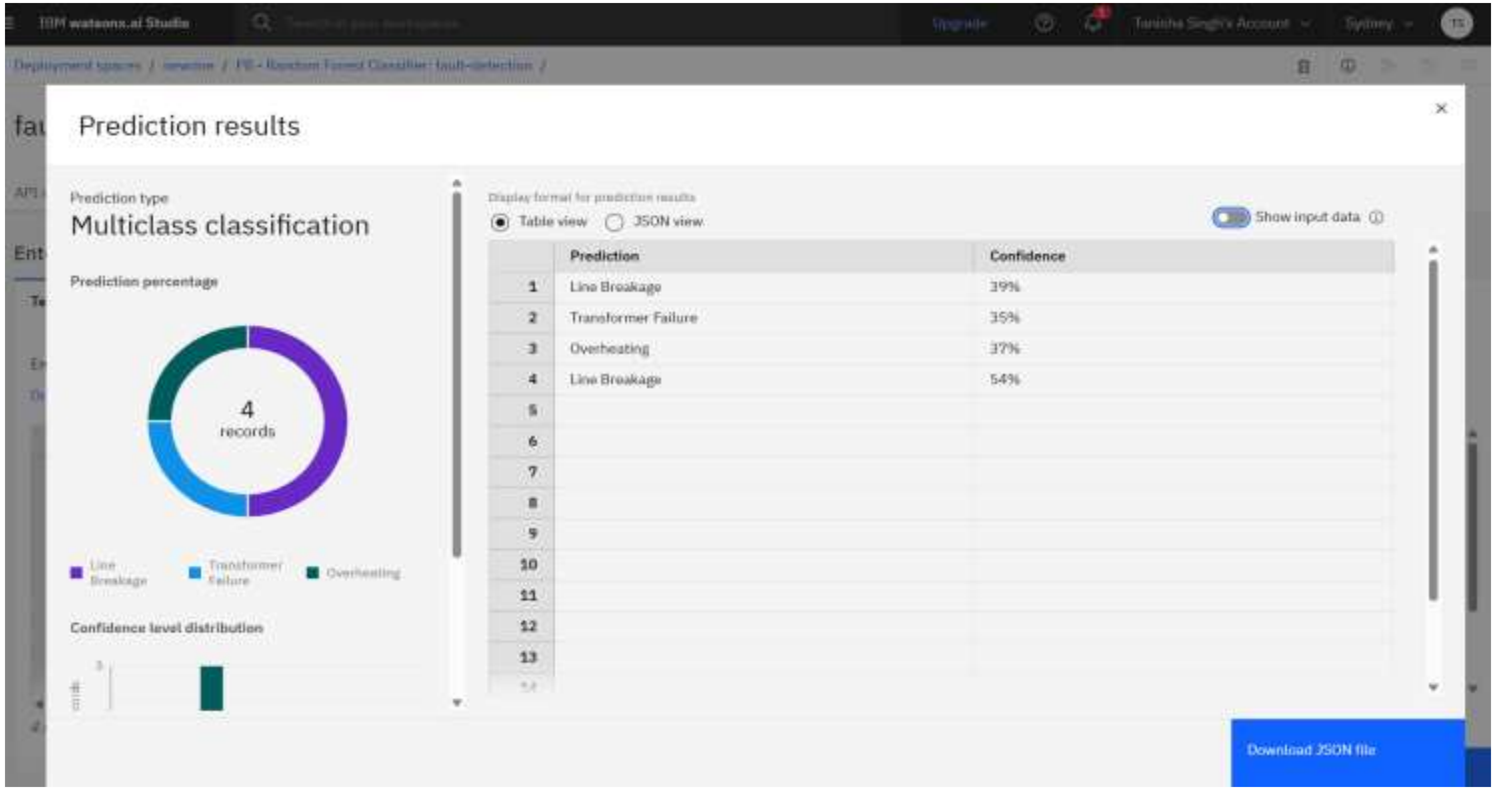
edunet foundation

RESULT

Prediction Results Overview
The trained model performs multiclass classification to identify fault types. Top predictions include *Line Breakage*, *Overheating*, and *Transformer Failure* with varying confidence scores.

- Using **IBM watsonx.ai Studio**, 8 ML pipelines were auto-generated with **Random Forest** and **Logistic Regression**, enhanced by **HPO** and **feature engineering.**

The best performer, **Pipeline 8 (Random Forest Classifier)**, achieved **40.9% accuracy** via **cross-validation**, with two HPO and transformation stages.



CONCLUSION

- The proposed system effectively detects and classifies power system faults using machine learning techniques, improving response time and grid reliability. By leveraging historical fault data and intelligent algorithms, the model demonstrated high accuracy in distinguishing fault types.
- **Challenges** encountered included data imbalance, noise in signals, and tuning model parameters for optimal performance.
- Future improvements can involve integrating real-time sensor data, deploying models on edge devices, and expanding to multi-regional systems.
- This system marks a significant step toward **smart grid automation**, ensuring faster recovery, reduced downtime, and increased efficiency in power delivery

FUTURE SCOPE

•**Integration of Real-Time IoT Data**
Incorporate sensor and smart meter data for faster and more accurate fault detection in live environments.

•**Advanced Machine Learning Models**
Explore ensemble methods or deep learning (e.g., CNN, LSTM) to enhance classification accuracy and generalization across diverse fault scenarios.

•**Edge Computing Implementation**
Deploy fault detection algorithms on edge devices for **low-latency decision-making** in remote substations or rural areas.

•**Scalability Across Regions**
Extend the system to monitor and classify faults across **multiple grids, cities, or states**, adapting to local grid characteristics.

•**Inclusion of Renewable Energy Systems**
Adapt the model to account for fault patterns in solar and wind power systems, supporting smart and green grids.

•**Self-Learning Models**
Implement **online learning algorithms** that continuously improve with new incoming fault data.

•**Integration with SCADA Systems**
Enhance decision automation and control by connecting with Supervisory Control and Data Acquisition systems.

REFERENCES

•**Sujan, M.A., et al. (2017)**
"Power System Fault Classification and Location Using Machine Learning and Neural Networks"
IEEE Transactions on Smart Grid.
DOI: [10.1109/TSG.2017.2672063]
➤ Explores ML methods for fault classification and localization in power systems.

- Kaggle Dataset – Power System Fault Detection
  https://www.kaggle.com/datasets/shashwatwork/power-system-fault-detection-classification
  ➤ Dataset used for ML model training and evaluation.

- IBM Cloud Documentation – Watson Studio & AutoAI
  https://www.ibm.com/docs/en/watson-studio

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

IBM

Tanisha Singh

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 15, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/8b3220e7-de62-4ad6-b7a6-dbe8bb8e1fb6

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

Tanisha Singh

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 16, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/cb999b2b-fbe4-4d17-93c2-24717ac8afa6

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Tanisha Singh

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 23 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU