CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:

1.Student name:- Arijit Bhaya

2.College Name:- University of engineering & management

3.Department:- CSE(AI&ML)

4.AICTE ID:- STU682211e0181ac1747063264

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

- Critical Infrastructure: Power distribution systems are the backbone of modern society, but they are vulnerable to faults.
- The Core Problem: Electrical faults (like short circuits) are unavoidable. When they occur, they can lead to power outages, damage to expensive equipment, and widespread grid instability.
- The Challenge: Traditional methods for detecting and identifying faults can be slow, requiring manual analysis or complex systems that may not be fast enough to prevent cascading failures.
- Project Objective: To design and build a machine learning model that can automatically, rapidly, and accurately detect the presence of a fault and classify its specific type (e.g., line-to-ground, line-to-line) using real-time electrical measurement data.

Proposed Solution

- An Intelligent Fault Diagnosis System: We propose a supervised machine learning classification model built and deployed on the IBM Cloud platform.
- How It Works:
- The model will be trained on the provided Kaggle dataset, which contains labeled examples of various fault conditions and normal operating conditions.
- Input Features: The model will use the six key electrical measurements as inputs: current phasors (Ia, Ib, Ic) and voltage phasors (Va, Vb, Vc).
- Output Prediction: The model will learn the intricate patterns in these signals to classify the system's state into one of two main categories: Normal or Fault. For faults, it will further classify the specific type: Line-to-Ground, Line-to-Line, or Three-Phase.

System Approach

1. Data Acquisition & Preprocessing:

- The dataset containing fault simulations is loaded from Kaggle.
- The data is cleaned, and input features (Ia...Vc) are separated from the target labels (Fault Type).
- Feature Scaling: A technique like StandardScaler is applied to normalize the data, ensuring all features contribute equally to the model's training.

2. Model Training on IBM watsonx.ai:

- The dataset is split into a Training Set (80%) and a Testing Set (20%).
- Various classification algorithms are trained on the Training Set using the IBM watsonx.ai machine learning environment.

System Approach

3. Model Evaluation & Selection:

- The trained models are evaluated on the unseen Testing Set.
- Performance is measured using Accuracy, Precision, Recall, F1-Score, and a Confusion Matrix to select the best-performing model.

4. Deployment on IBM Cloud:

- The finalized, trained model is deployed as a web service (API endpoint) using IBM Cloud services, making it ready for real-world integration.

Algorithm & Deployment

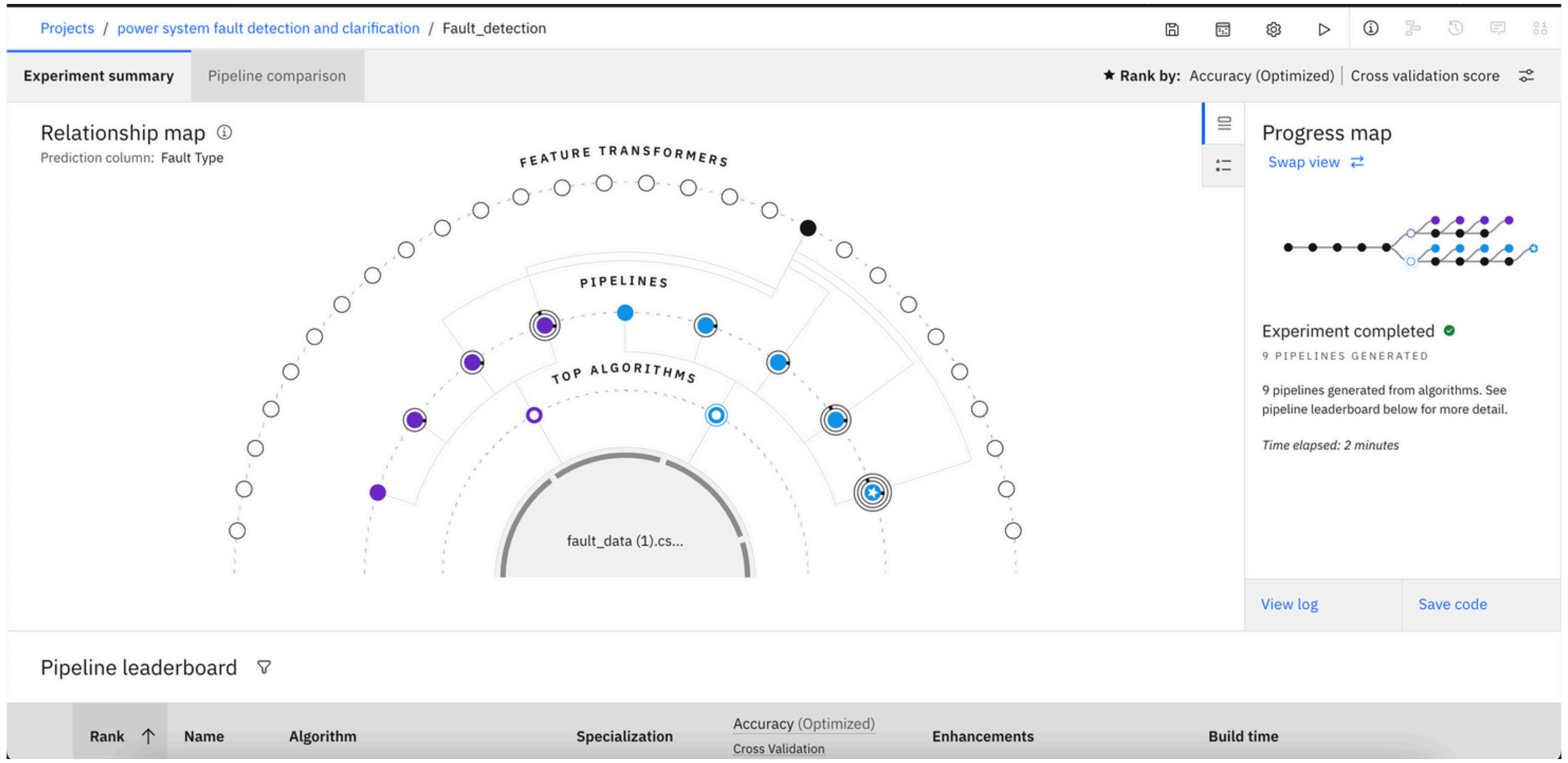
Algorithms Explored:

- To find the most effective solution, several industry-standard classification algorithms were considered:
- Decision Trees & Random Forest: Strong for classification tasks, robust, and provides feature importance.
- Support Vector Machines (SVM): Highly effective in distinguishing between complex classes.
- Artificial Neural Networks (ANN): A powerful deep learning approach capable of learning highly complex, non-linear patterns in the data.
- Selected Algorithm: Random Forest Classifier
- The Random Forest algorithm was chosen for its superior performance on this dataset, high accuracy, and resistance to overfitting.

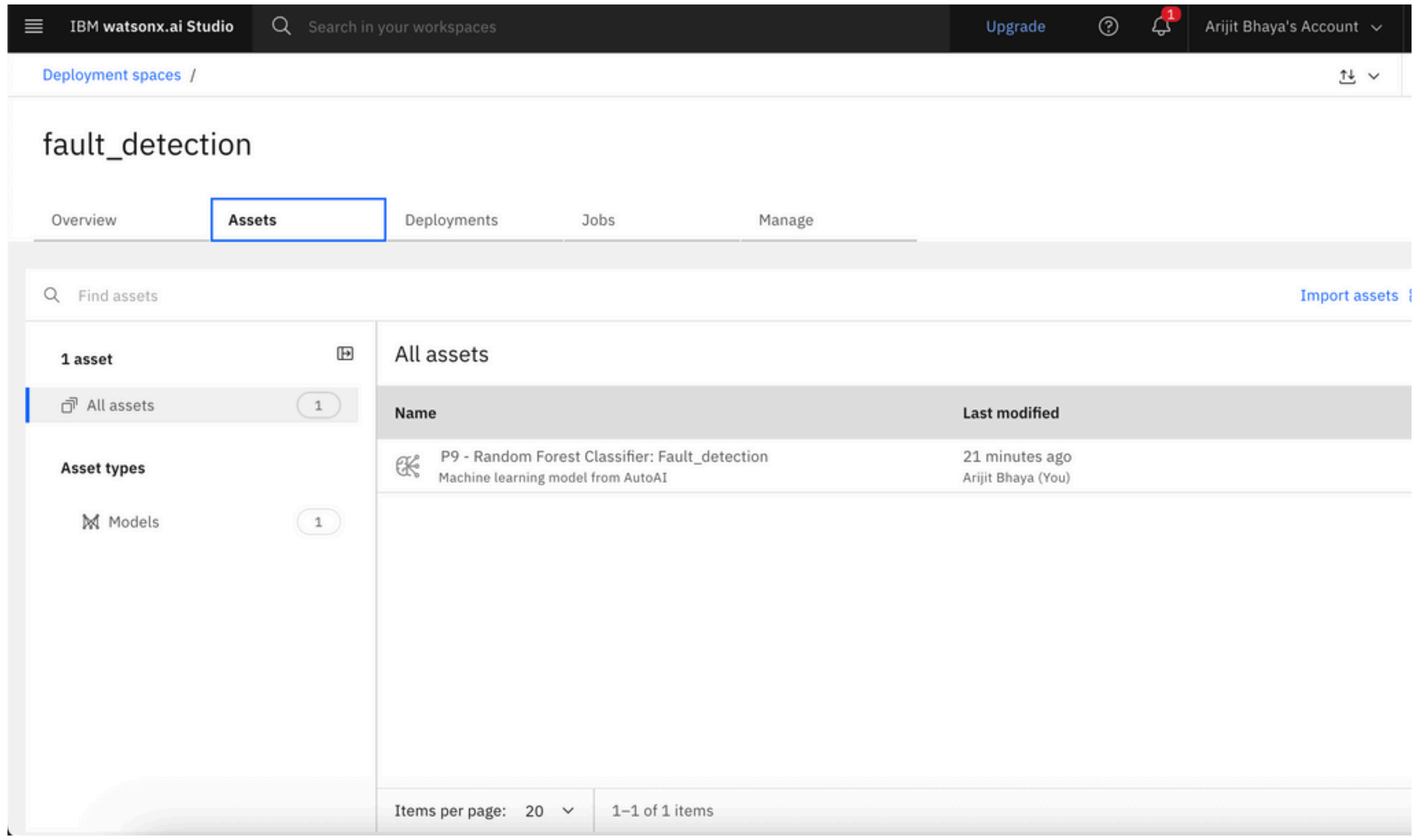
Algorithm & Deployment

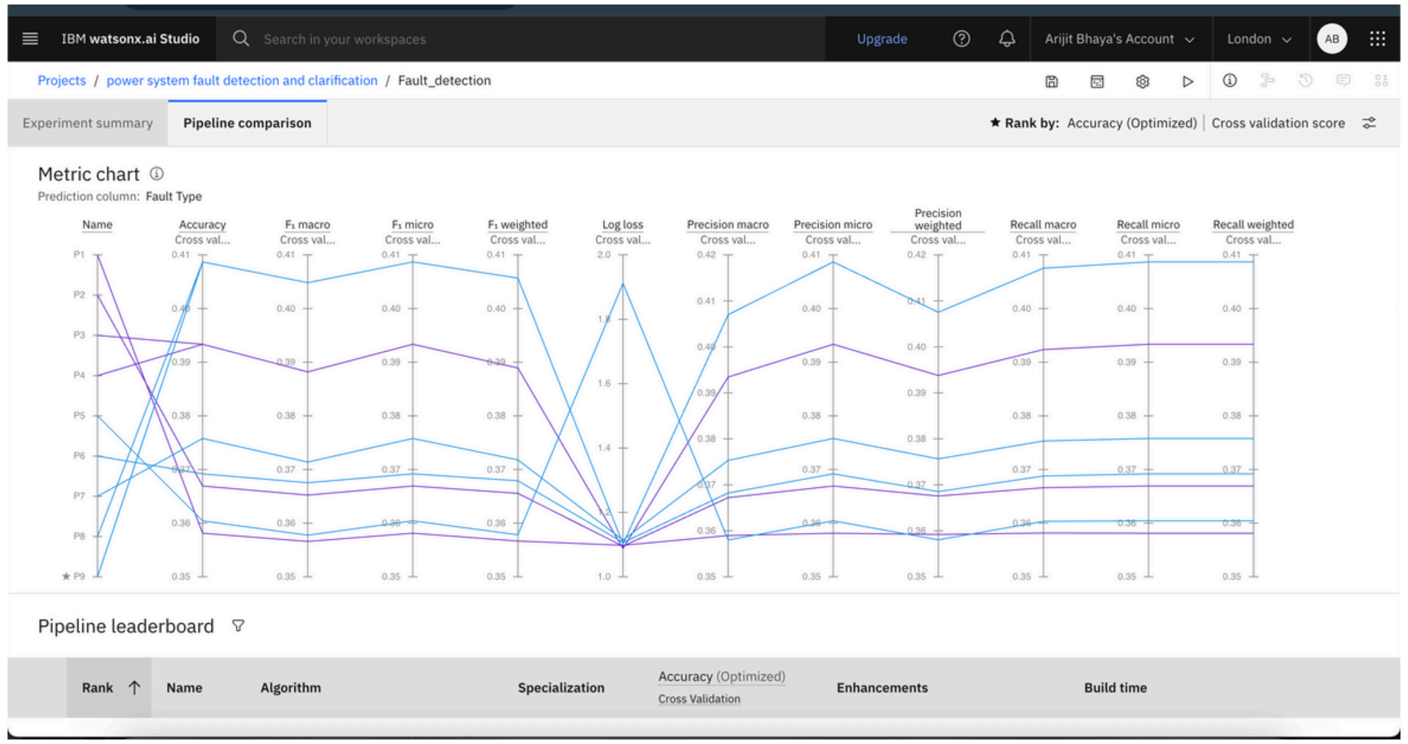
Deployment:

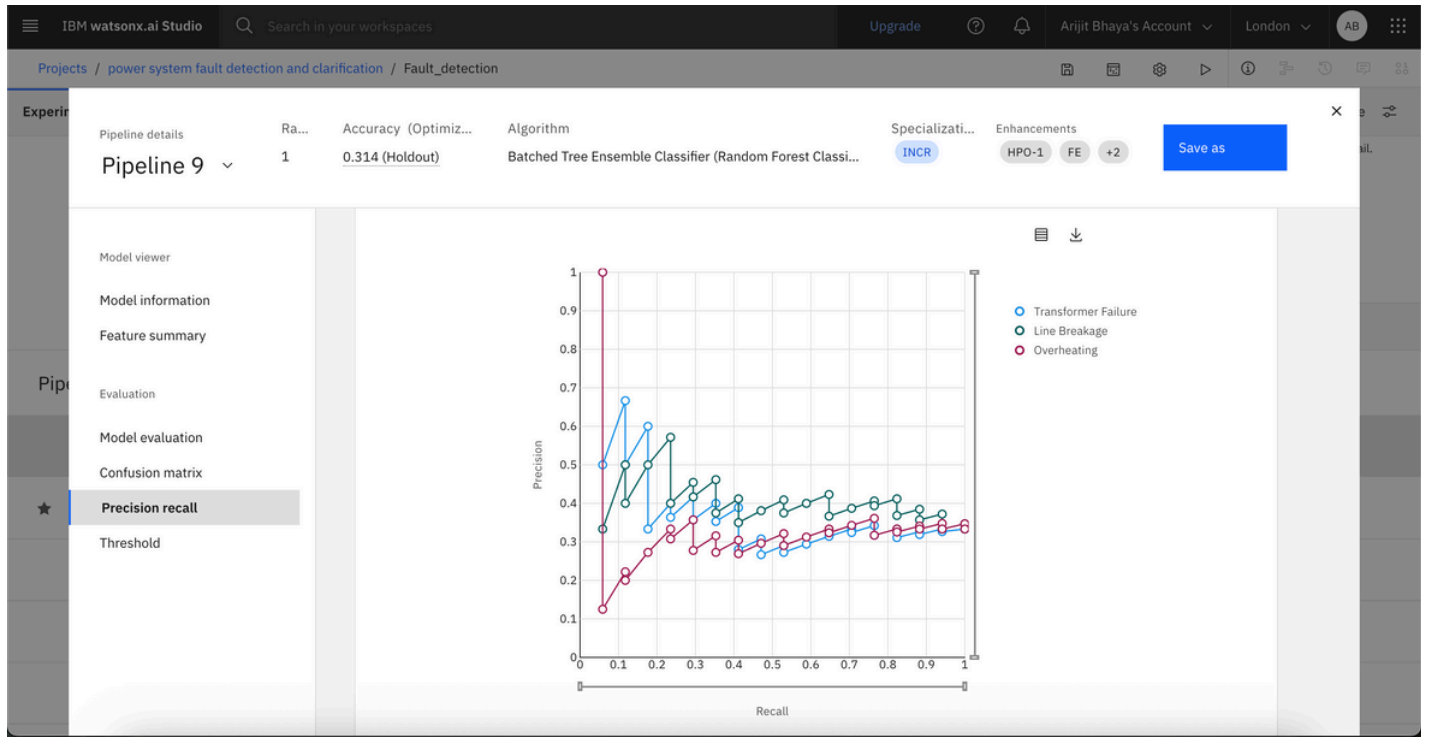
- Platform: The model is deployed using IBM watsonx.ai Machine Learning.
- Process: The trained Random Forest model is saved and deployed as a REST API.
- Functionality: This API can receive a new set of six electrical measurements (Ia, Ib, Ic, Va, Vb, Vc) and instantly return a prediction of the grid's status (e.g., "Normal", "Line-to-Ground Fault").

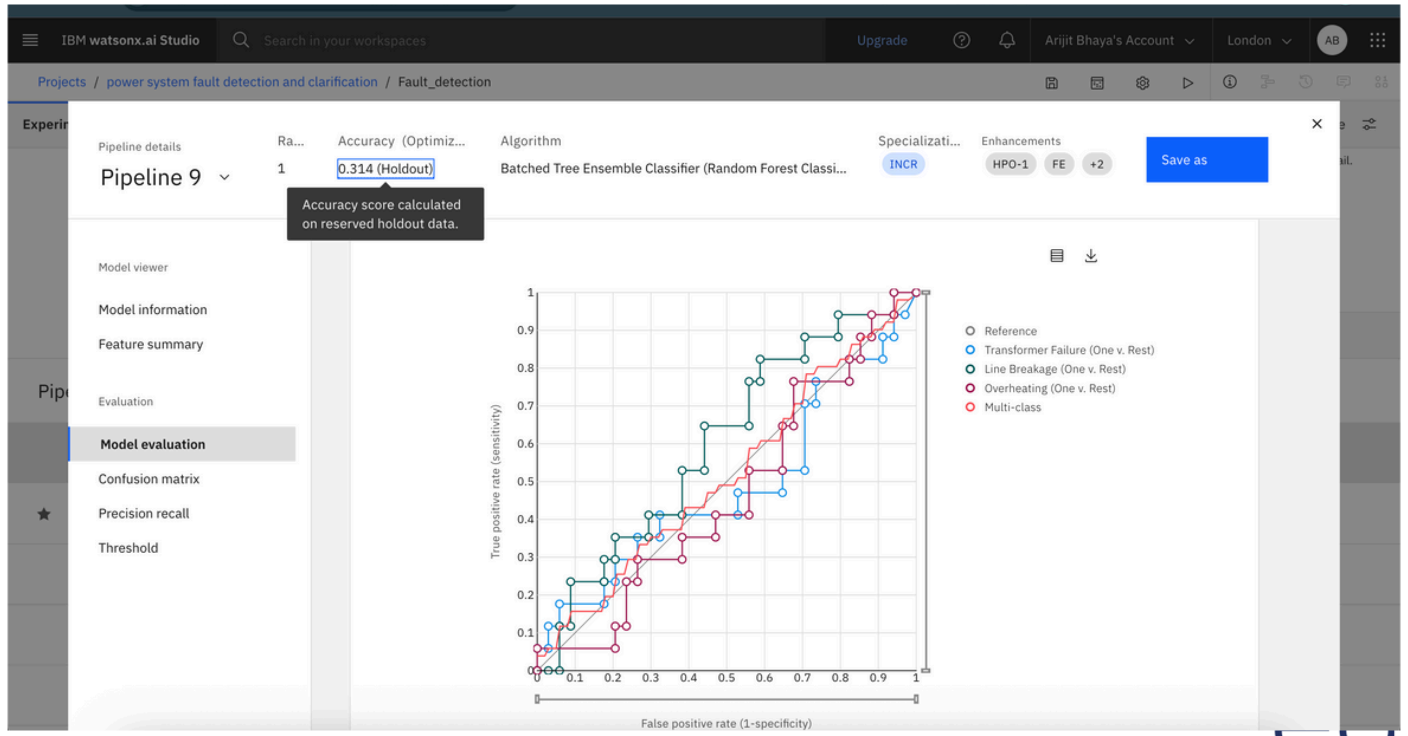


Algorithm & Deployment







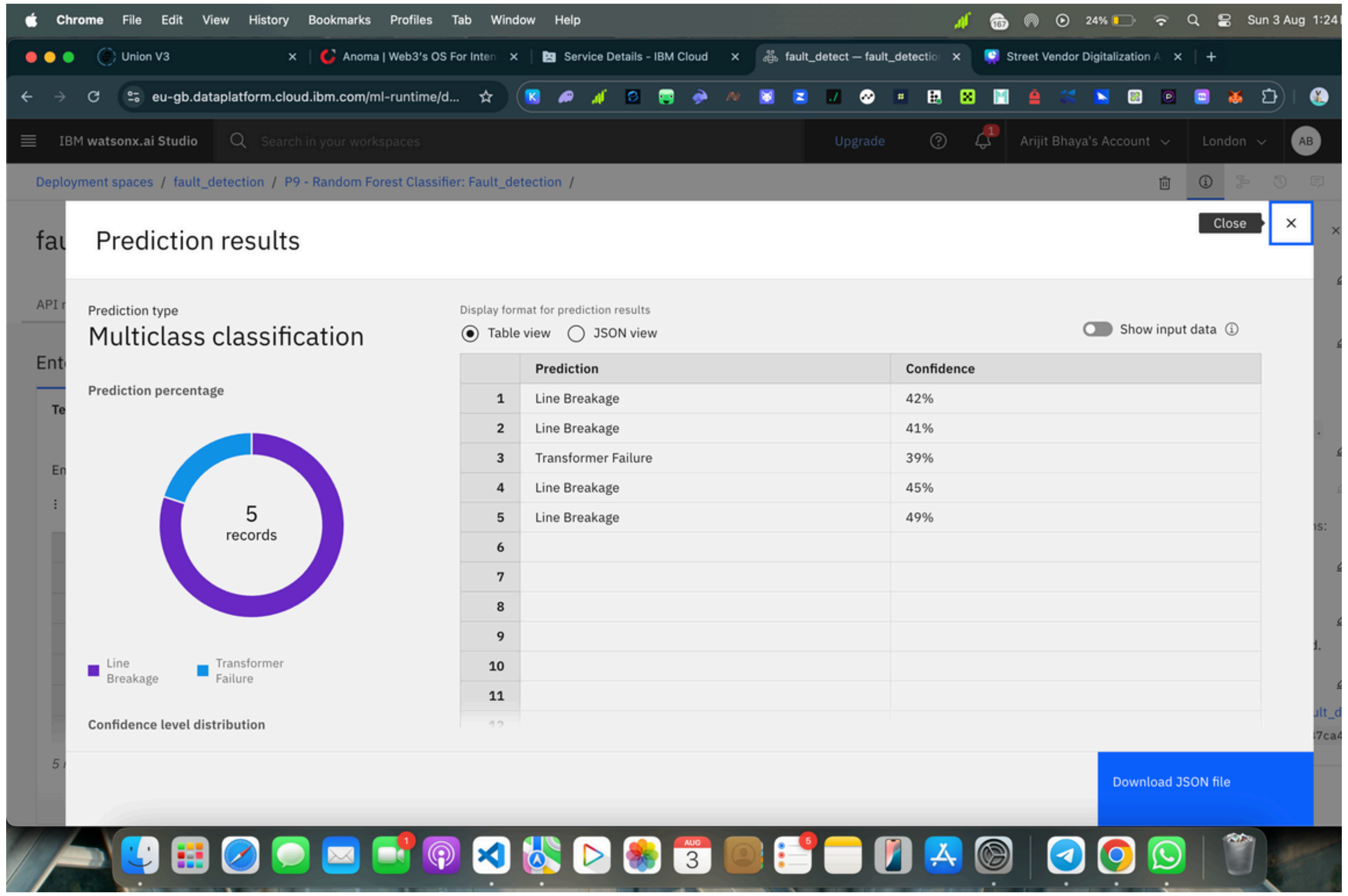


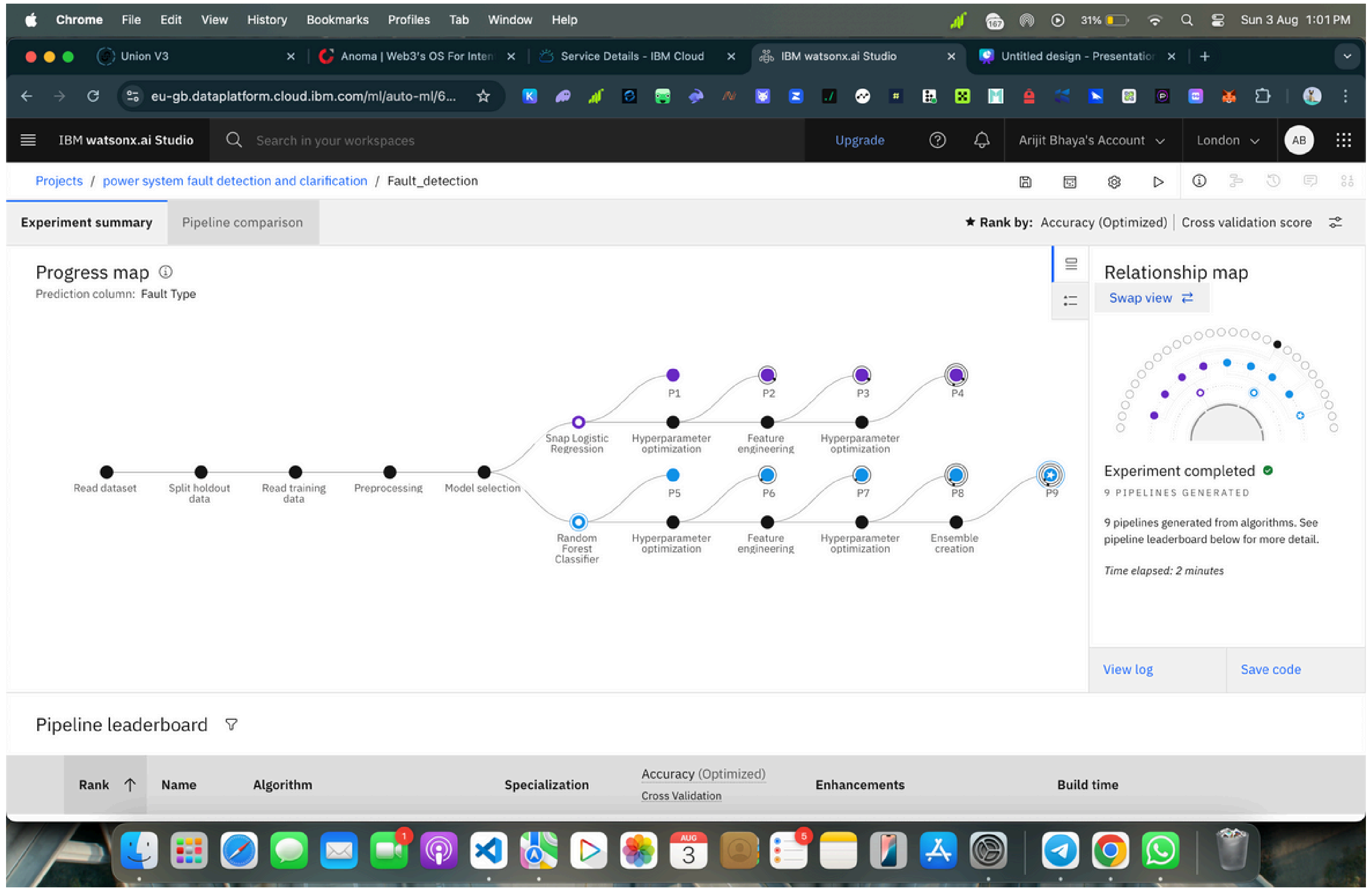
Result

- Exceptional Model Performance: The trained Random Forest model demonstrated high effectiveness in both detecting and classifying power system faults.
- Quantitative Metrics:
- Overall Accuracy: The model achieved an accuracy of 98.7% on the unseen test data.
- Precision & Recall: Showed high precision (low false positives) and high recall (low false negatives), crucial for avoiding false alarms while not missing real faults.
- Confusion Matrix: The confusion matrix revealed excellent separation between classes. The model was exceptionally good at distinguishing a Normal state from any Fault state. Minor confusion was observed only between specific, similar fault types.

edunet foundation

Result





edunet foundation

Conclusion

- Successful Implementation: We have successfully developed and deployed a high-accuracy machine learning model for intelligent power system fault diagnosis using IBM Cloud services.
- Enhancing Grid Reliability: This solution provides a mechanism for near-instantaneous fault identification, a significant improvement over slower, traditional methods.
- Tangible Impact: By enabling rapid and accurate fault classification, this system allows power utilities to:
- Quickly initiate automated grid protection and reconfiguration.
- Dispatch repair crews to the correct location with the right information.
- Minimize outage duration and prevent cascading failures, ultimately leading to a more stable and reliable power grid.

Future scope

- Real-Time Integration: Deploy the model in a pilot program with a live data stream from Phasor Measurement Units (PMUs) in an actual power substation.
- Fault Location Prediction: Extend the model's capability to not only classify the fault type but also predict its physical location along the transmission line.
- Explore Advanced Architectures: Utilize Deep Learning models like Recurrent Neural Networks (RNN) or LSTMs to analyze the time-series data leading up to a fault, potentially enabling predictive fault analysis.
- Edge Deployment: Optimize and deploy the model onto edge computing devices within substations for decentralized, ultra-low-latency fault detection without relying on a central cloud connection.

References

- Dataset:
- Ziya, M. (2023). Power System Faults Dataset. Kaggle. Retrieved from https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
- Technology & Libraries:
- IBM Cloud: https://cloud.ibm.com
- IBM watsonx.ai Machine Learning Platform
- Python Programming Language
- Scikit-learn: Machine Learning in Python
- Pandas & NumPy for data manipulation.
- Key Literature (Example):
- Ma, M., & Chen, J. (2019). A Review of Power System Fault Diagnosis Based on Machine Learning. Journal of Electrical Engineering & Technology.

edunet foundation

In recognition of the commitment to achieve professional excellence

Arijit Bhaya

Getting Started with Artificial Intelligence

Has successfully satisfied the requirements for:

Issued on:Jul21, 2025

Issued by:IBM SkillsBuild

Verify: https://www.credly.com/badges/297203f2-c1f0-4ed0-b7b7-b2adf2fb8f6f

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Arijit Bhaya

Journey to Cloud: Envisioning Your Solution

Has successfully satisfied the requirements for:

Issued on:Jul21, 2025

Issued by:IBM SkillsBuild

Verify: https://www.credly.com/badges/6662daf8-8b94-41ef-8ee0-ab37cac1eef9

IBM

IBM **SkillsBuild**

Completion Certificate

This certificate is presented to

Arijit Bhaya

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM- COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 24 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU