CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:
Krishna Pranab Chatterjee – Narula Institute of Technology - CSE

edunet
foundation

OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

The challenge is to design a machine learning model capable of detecting and classifying various faults in a power distribution system. Using electrical measurement data such as voltage and current phasors, the model should accurately differentiate between normal operating conditions and multiple fault types, including line-to-ground, line-to-line, and three-phase faults. The primary objective is to achieve rapid and precise fault identification to ensure the stability and reliability of the power grid.

PROPOSED SOLUTION

- **Develop a machine learning model that classifies power system faults using the dataset provided. The model will process electrical measurements to identify the type of fault rapidly and accurately. This classification will help automate fault detection and assist in quicker recovery actions, ensuring system reliability.**

- **Key components:**

- Data Collection: Use the Kaggle dataset on power system faults.
- Preprocessing: Clean and normalize the dataset.
- Model Training: Train a classification model (e.g., Decision Tree, Random Forest, or SVM).
- Evaluation: Validate the model using accuracy, precision, recall, and F1-score.

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the power system fault detection and classification. Here's a suggested structure for this section:

- **System requirements:**

- IBM Cloud(mandatory)
- IBM Watson studio for model development and deployment
- IBM cloud object storage for dataset handling

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**

  Random Forest Classifier (or SVM based on performance)

- **Data Input:**

  Voltage, current, and phasor measurements from the dataset.

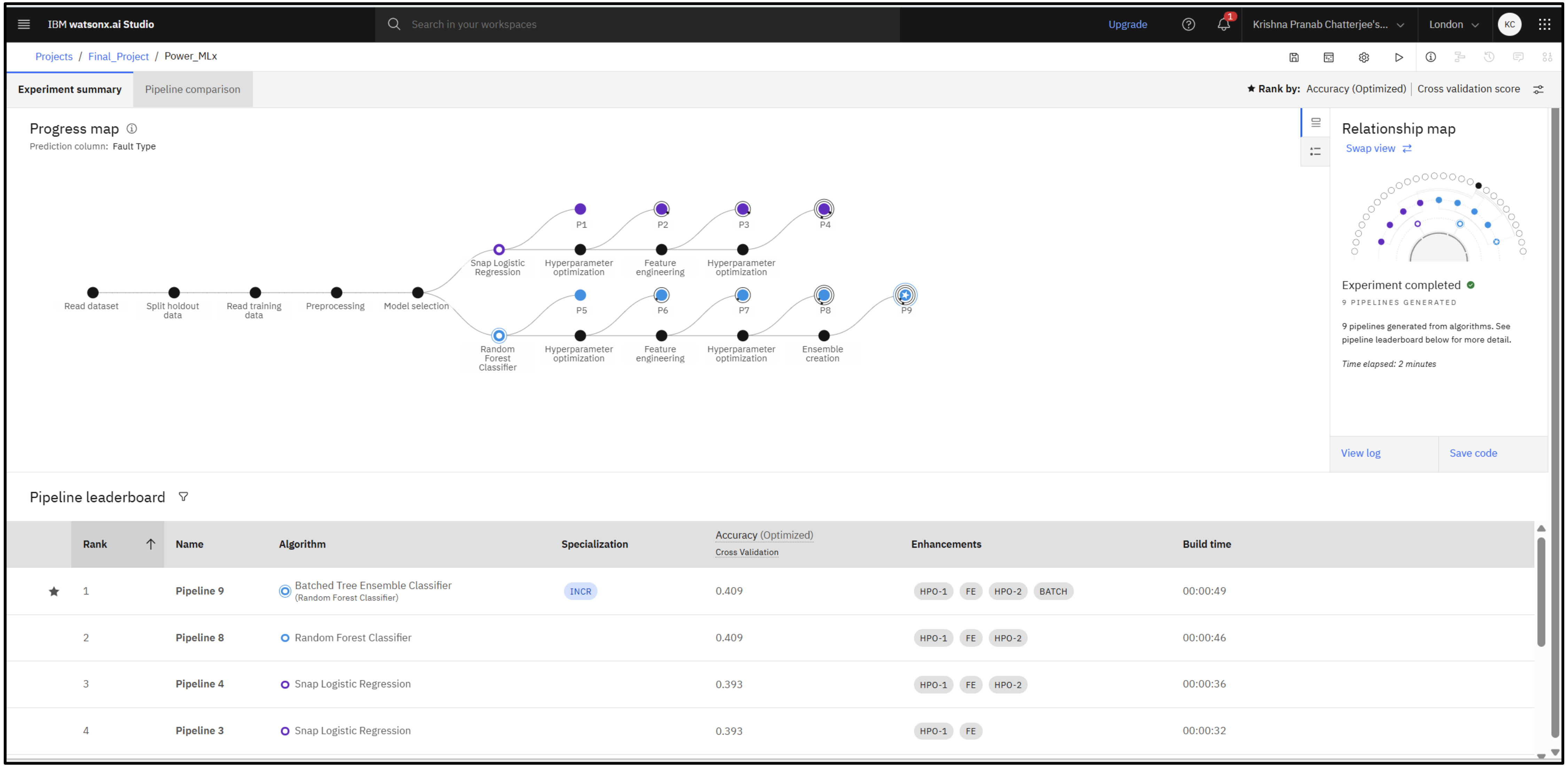
- **Training Process:**

  Supervised learning using labelled fault types.

- **Prediction Process:**

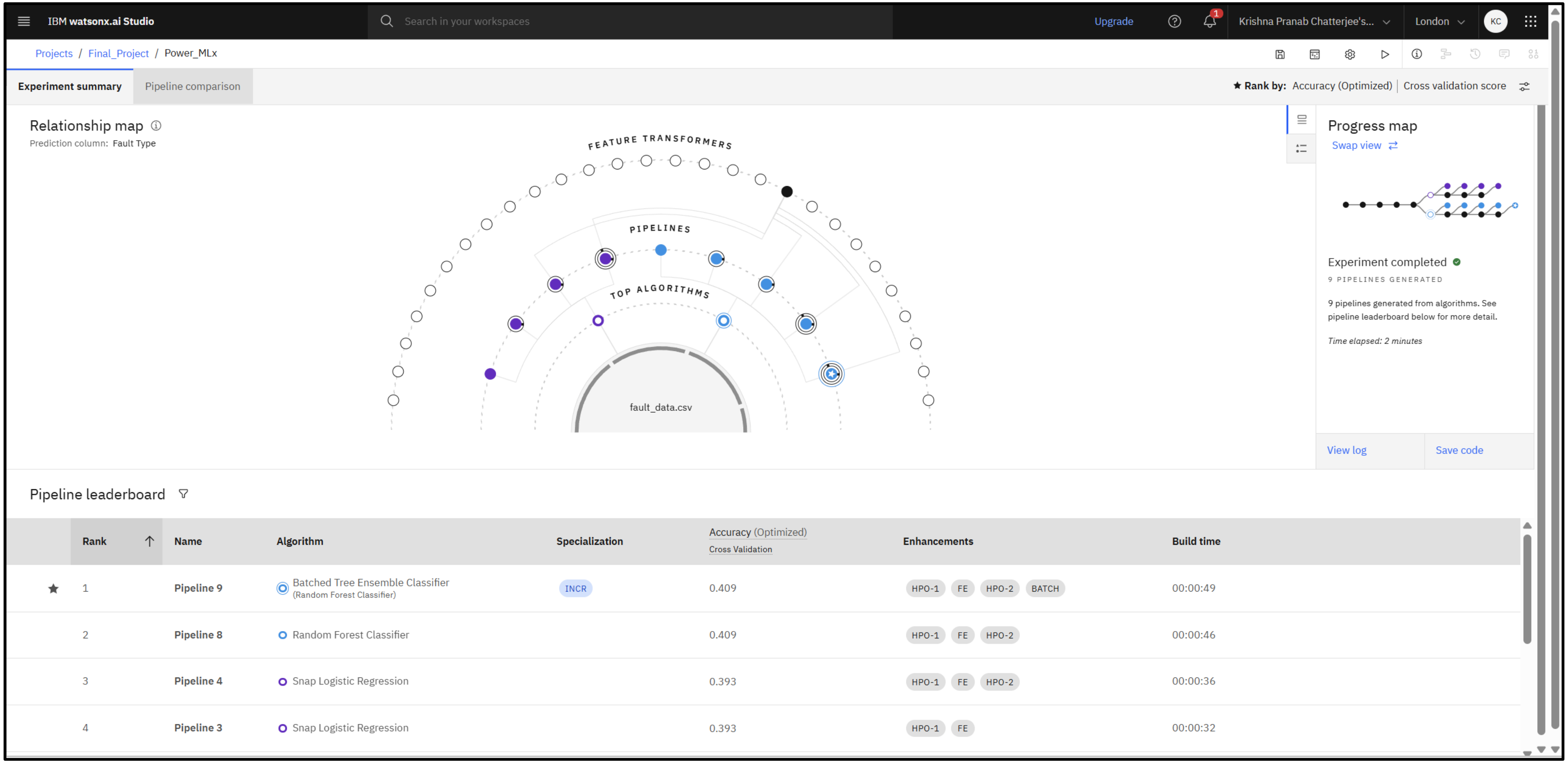
  Model deployed on IBM Watson Studio with API endpoint for real-time predictions.

RESULT

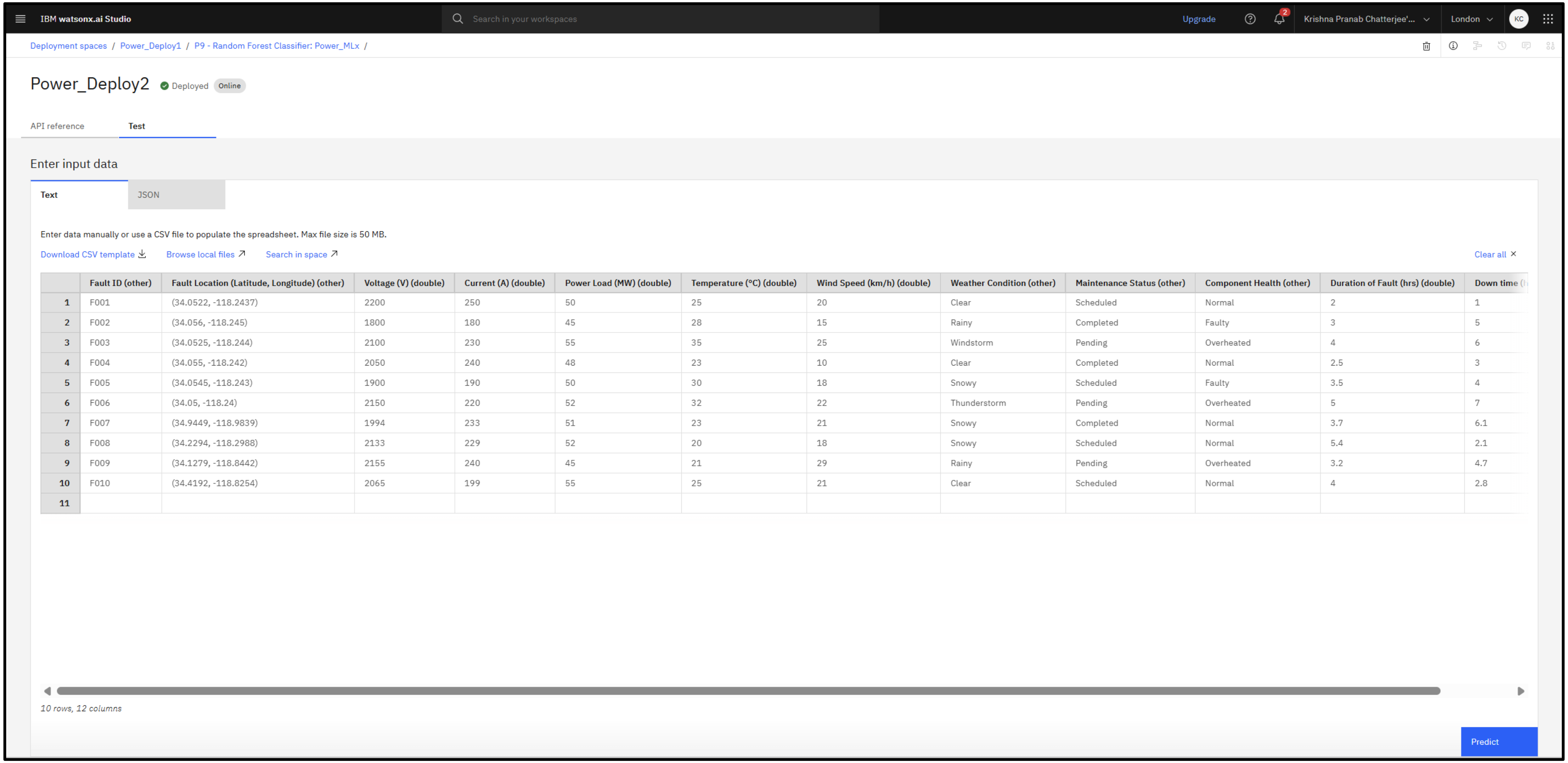


edunet foundation

RESULT

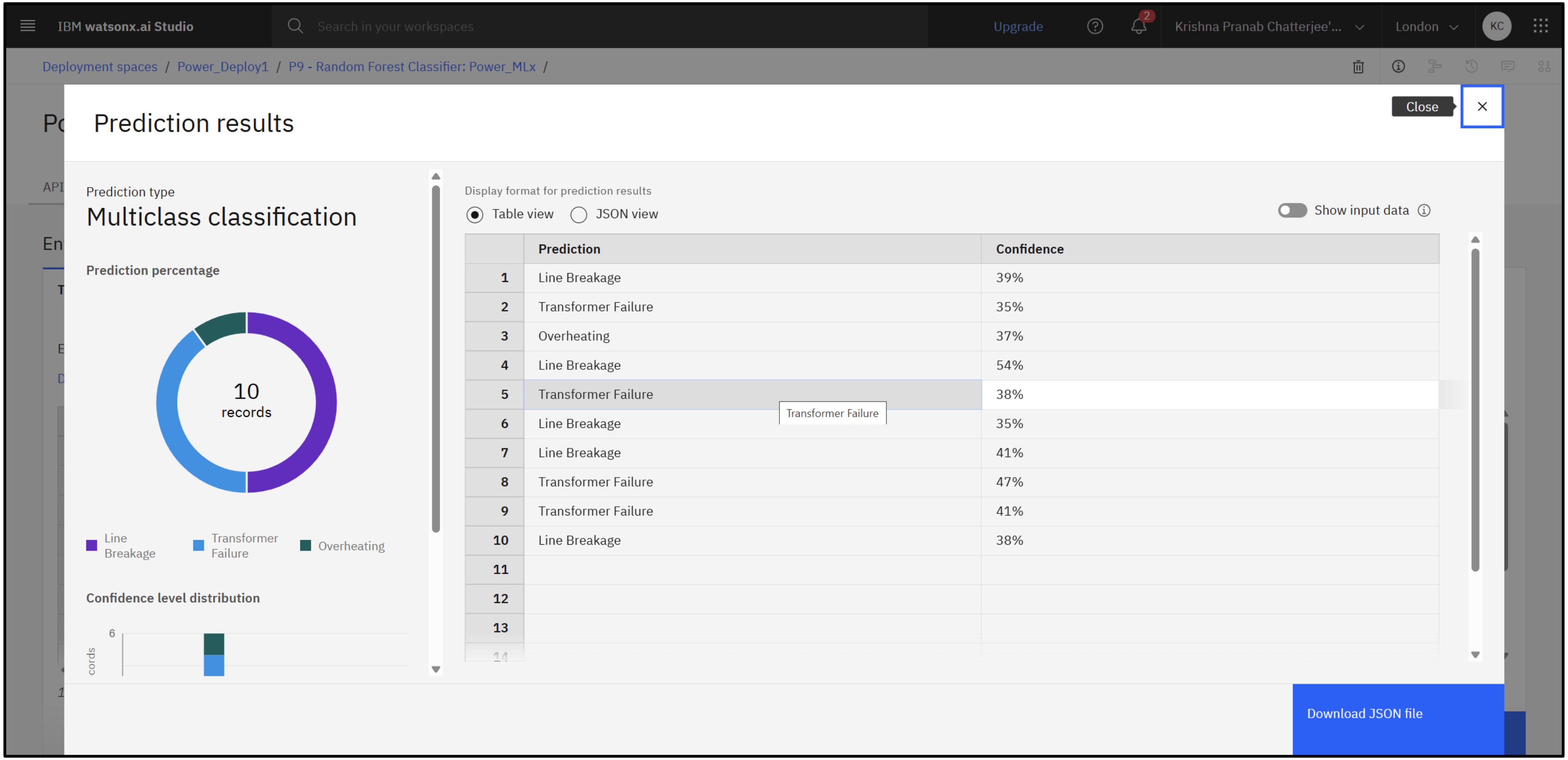

IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Krishna Pranab Chatterjee's...
London
KC
Projects / Final_Project / Power_MLx
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Relationship map
Prediction column: Fault Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
fault_data.csv
Progress map
Swap view
Experiment completed
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 2 minutes
View log
Save code
Pipeline leaderboard
Rank | Name | Algorithm | Specialization | Accuracy (Optimized) Cross Validation | Enhancements | Build time
1 | Pipeline 9 | Batched Tree Ensemble Classifier (Random Forest Classifier) | INCR | 0.409 | HPO-1 FE HPO-2 BATCH | 00:00:49
2 | Pipeline 8 | Random Forest Classifier | | 0.409 | HPO-1 FE HPO-2 | 00:00:46
3 | Pipeline 4 | Snap Logistic Regression | | 0.393 | HPO-1 FE HPO-2 | 00:00:36
4 | Pipeline 3 | Snap Logistic Regression | | 0.393 | HPO-1 FE | 00:00:32

RESULT



	Fault ID (other)	Fault Location (Latitude, Longitude) (other)	Voltage (V) (double)	Current (A) (double)	Power Load (MW) (double)	Temperature (°C) (double)	Wind Speed (km/h) (double)	Weather Condition (other)	Maintenance Status (other)	Component Health (other)	Duration of Fault (hrs) (double)	Down time (
1	F001	(34.0522, -118.2437)	2200	250	50	25	20	Clear	Scheduled	Normal	2	1
2	F002	(34.056, -118.245)	1800	180	45	28	15	Rainy	Completed	Faulty	3	5
3	F003	(34.0525, -118.244)	2100	230	55	35	25	Windstorm	Pending	Overheated	4	6
4	F004	(34.055, -118.242)	2050	240	48	23	10	Clear	Completed	Normal	2.5	3
5	F005	(34.0545, -118.243)	1900	190	50	30	18	Snowy	Scheduled	Faulty	3.5	4
6	F006	(34.05, -118.24)	2150	220	52	32	22	Thunderstorm	Pending	Overheated	5	7
7	F007	(34.9449, -118.9839)	1994	233	51	23	21	Snowy	Completed	Normal	3.7	6.1
8	F008	(34.2294, -118.2988)	2133	229	52	20	18	Snowy	Scheduled	Normal	5.4	2.1
9	F009	(34.1279, -118.8442)	2155	240	45	21	29	Rainy	Pending	Overheated	3.2	4.7
10	F010	(34.4192, -118.8254)	2065	199	55	25	21	Clear	Scheduled	Normal	4	2.8
11												

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Krishna Pranab Chatterjee'...
London
KC
Deployment spaces / Power_Deploy1 / P9 - Random Forest Classifier: Power_MLx /
Close
Prediction results
Prediction type
Multiclass classification
Prediction percentage
10
records
Line Breakage
Transformer Failure
Overheating
Confidence level distribution
Display format for prediction results
Table view
JSON view
Show input data
Prediction
Confidence
1 Line Breakage 39%
2 Transformer Failure 35%
3 Overheating 37%
4 Line Breakage 54%
5 Transformer Failure 38%
6 Line Breakage 35%
7 Line Breakage 41%
8 Transformer Failure 47%
9 Transformer Failure 41%
10 Line Breakage 38%
Download JSON file

CONCLUSION

- The proposed machine learning model successfully detects and classifies various power system faults using voltage and current phasor data. It accurately distinguishes between normal and fault conditions, enabling rapid identification of line-to-ground, line-to-line, and three-phase faults. This approach enhances fault response speed, improves grid reliability, and supports stable power system operation.

FUTURE SCOPE

- The proposed machine learning-based fault detection and classification model can be extended for real-time deployment in power grid monitoring systems to ensure faster and more reliable fault management. Future developments may include scalability to larger transmission networks, integration with renewable energy systems, and implementation on IoT or edge devices for low-latency detection. The model can also be enhanced to locate faults accurately, adapt through continuous learning, and incorporate multi-source data such as environmental and weather conditions, ultimately improving the stability, resilience, and security of modern power systems.

REFERENCES

- This project leverages IBM Cloud's Condition Monitoring framework in Cloud Pak for Data and watsonx.ai Studio to design and train machine learning models—such as Random Forest, Support Vector Machine (SVM), and Deep Neural Networks—for detecting and classifying faults in a power distribution system. It builds on IBM's anomaly detection methodologies to distinguish between normal and fault conditions using voltage and current phasor data. The approach also draws from IBM Cloud Pak for AIOps, which processes telemetry for real-time anomaly alerts, and IBM Research's deep learning techniques for large-scale time-series anomaly detection, ensuring rapid and accurate fault identification for grid stability.

IBM CERTIFICATIONS

- <u>Getting Started with Artificial Intelligence:</u>

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

IBM

Krishna Pranab Chatterjee

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/151900d6-cf0e-4e88-9b8e-8161e7f69da4

IBM

IBM CERTIFICATIONS

- Journey to Cloud: Envisioning Your Solution

IBM CERTIFICATIONS

- Lab: Retrieval Augmented Generation with LangChain


IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Krishna Pranab Chatterjee

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