CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:

1. Hritik Kumar Vishwakarma

-Lokmata Devi Ahilya bai Holkar Rajkiya Engineering College, Mainpuri

–Computer Science and Engineering Department

edunet foundation

OUTLINE

- Problem Statement
- Proposed Solution
- System Development Approach
- Algorithm & Deployment
- Relation and Progress Map images
- Input image
- Result
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

Modern power distribution systems are prone to various types of electrical faults, such as line-to-ground, line-to-line, and three-phase faults. These faults can cause severe damage to equipment, power outages, and safety hazards. Manual fault detection methods are slow and may lead to prolonged downtime. There is a critical need for a rapid, accurate, and automated way to detect and classify faults to maintain grid stability and reliability.

PROPOSED SOLUTION

- The proposed system aims to address the challenge of detecting and classifying different types of faults in a power distribution system. This involves leveraging electrical measurement data (voltage and current phasors) and applying machine learning techniques to accurately identify fault types. The solution will consist of the following components:
- **Data Collection:**
  - Utilize the Kaggle Power System Faults Dataset containing historical voltage and current phasor readings under both normal and various fault conditions.
- **Machine Learning Algorithm:**
  - Implement a Random Forest classifier using Scikit-learn to train a model that classifies each sample as either normal or one of the fault types (e.g., line-to-ground, line-to-line, three-phase).
  - Optimize hyperparameters (e.g., number of trees, max depth) using GridSearchCV to improve model accuracy.
- **Deployment:**
  - Package the trained Random Forest model and deploy it using IBM Watson Machine Learning services on IBM Cloud Lite.
- **Evaluation:**
  - Evaluate model performance on the test dataset using metrics such as Accuracy, Precision, Recall, and F1-Score.
  - Use a confusion matrix to analyze model performance on individual fault classes and identify misclassification patterns.
- **Result:**
  - The solution will produce real-time predictions of fault types for each measurement input, allowing for rapid identification and classification of faults in power distribution networks.

SYSTEM APPROACH

- **Data Source:** Kaggle's Power System Faults Dataset
- **Visualization:** tree and pipelining
- **IBM Cloud Lite Services:**
  - IBM Watson Studio for model development
  - IBM Cloud Object Storage for dataset storage
  - IBM Watson Machine Learning for model deployment
- **Development Language:** Python

ALGORITHM & DEPLOYMENT

Algorithm Selection:
We selected the **Random Forest classifier**, an ensemble learning algorithm based on decision trees, for fault detection and classification. Random Forest is well-suited for handling tabular data with multiple features, is robust to outliers, and reduces the risk of overfitting by averaging the predictions of many trees. Its ability to handle complex, nonlinear relationships makes it ideal for distinguishing between different power system fault types based on electrical measurements.

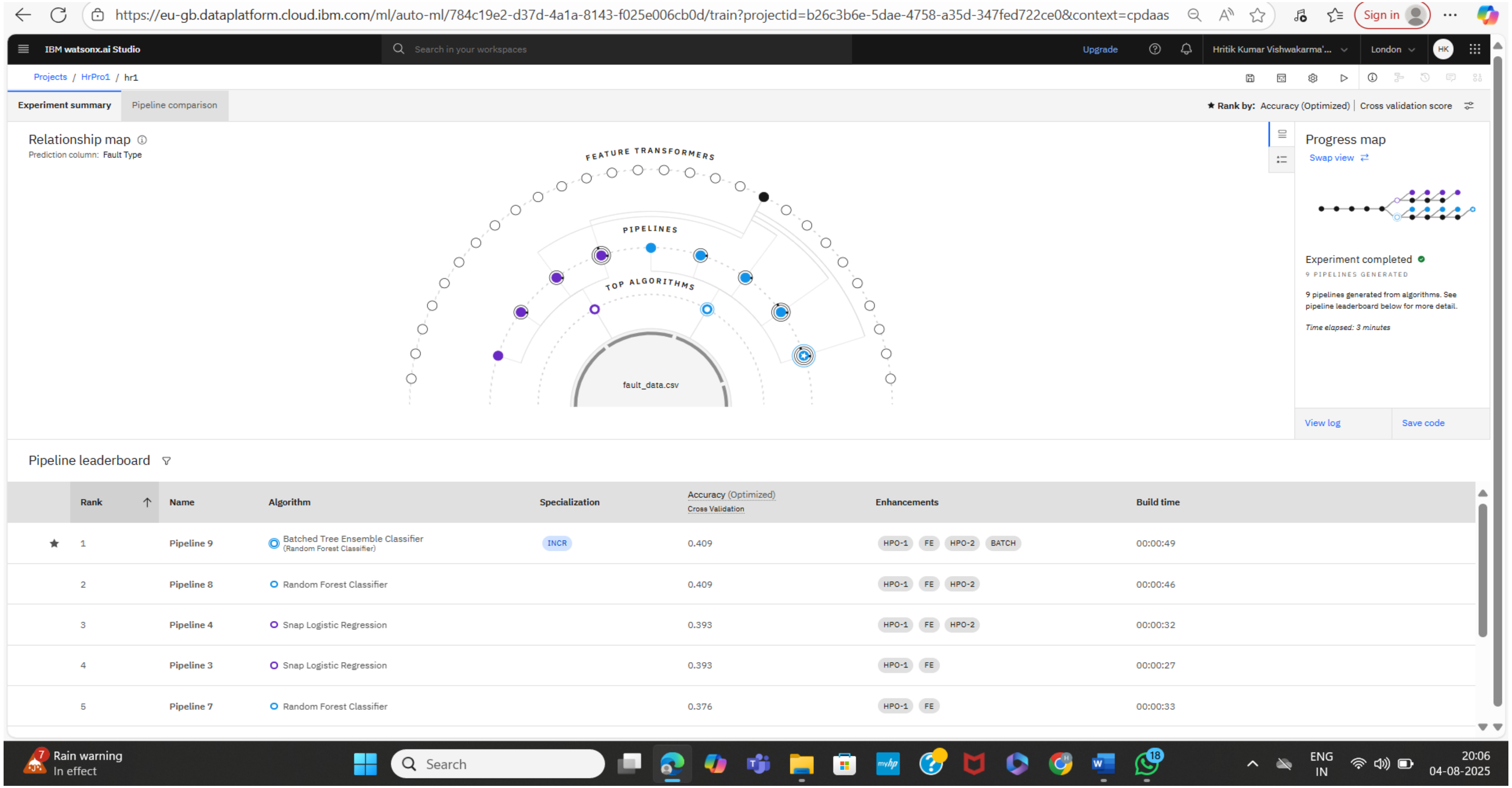
- **Data Input:**

- Voltage and current phasor measurements from all three phases (magnitudes and angles),
- Derived features such as differences between phases, phase imbalances, and combined highlighting abnormal conditions,
- Categorical labels for each sample indicating the fault type (e.g., normal, line-to-ground, line-to-line, three-phase fault).

ALGORITHM & DEPLOYMENT

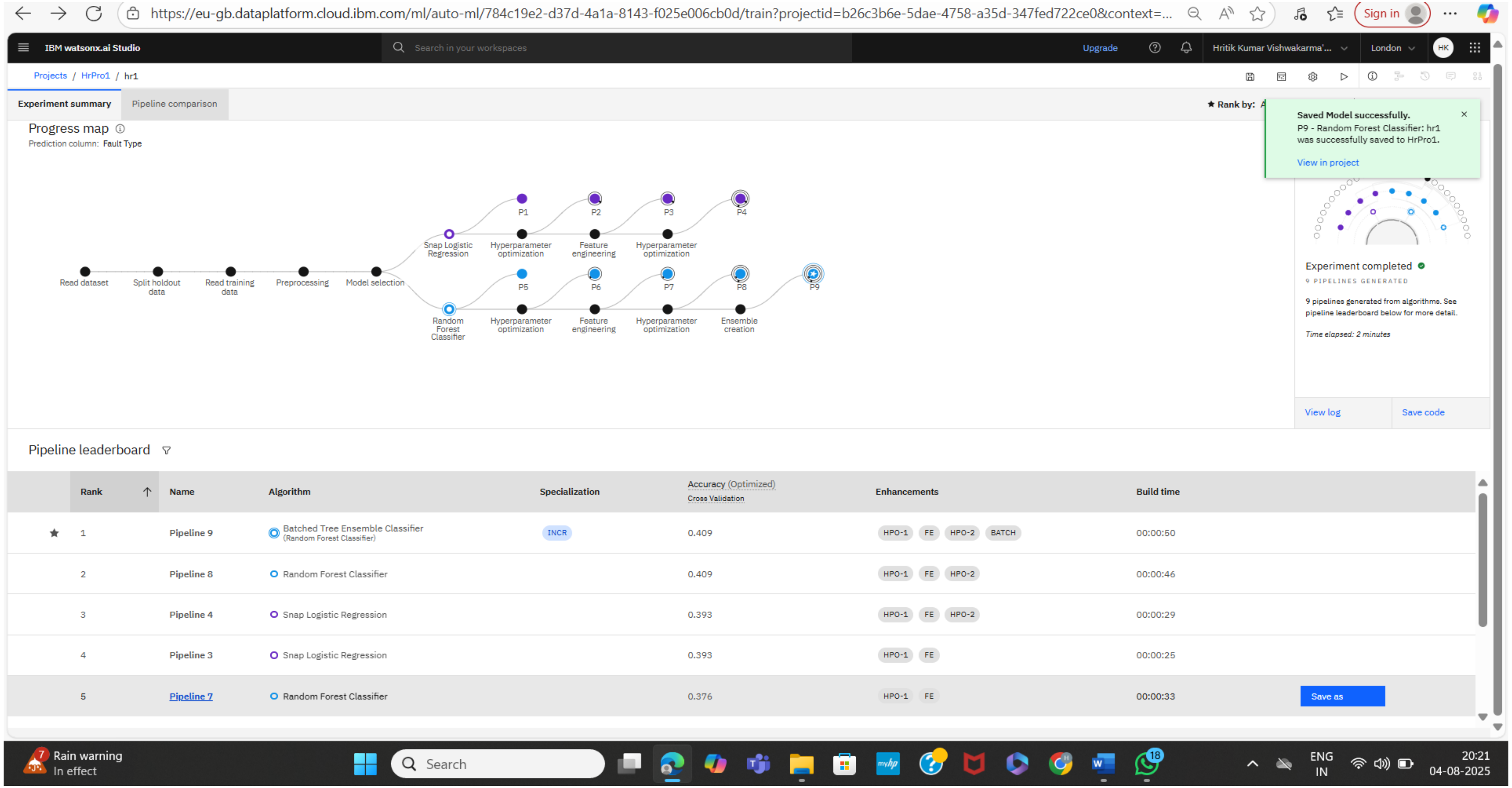
- **Training Process:**
- The dataset is split into training and test sets (e.g., 80% for training, 20% for testing) to ensure unbiased evaluation.
- The Random Forest model is trained using the training set with hyperparameter optimization (number of trees, maximum depth, and minimum samples per leaf) performed through GridSearchCV.
- K-fold cross-validation is applied during training to validate performance and reduce variance in the model evaluation.
- Feature importance analysis is conducted to identify which electrical measurements contribute most to fault classification.
- **Prediction Process:**
- Once trained, the Random Forest model receives real-time or batch input of phasor measurements.
- The model processes the input features and outputs the predicted class label representing the detected fault type along with confidence probabilities.
- Real-time prediction pipelines incorporate new measurement data as it arrives, allowing operators to monitor and respond to faults instantly.

RELATIONSHIP MAP IMAGE:

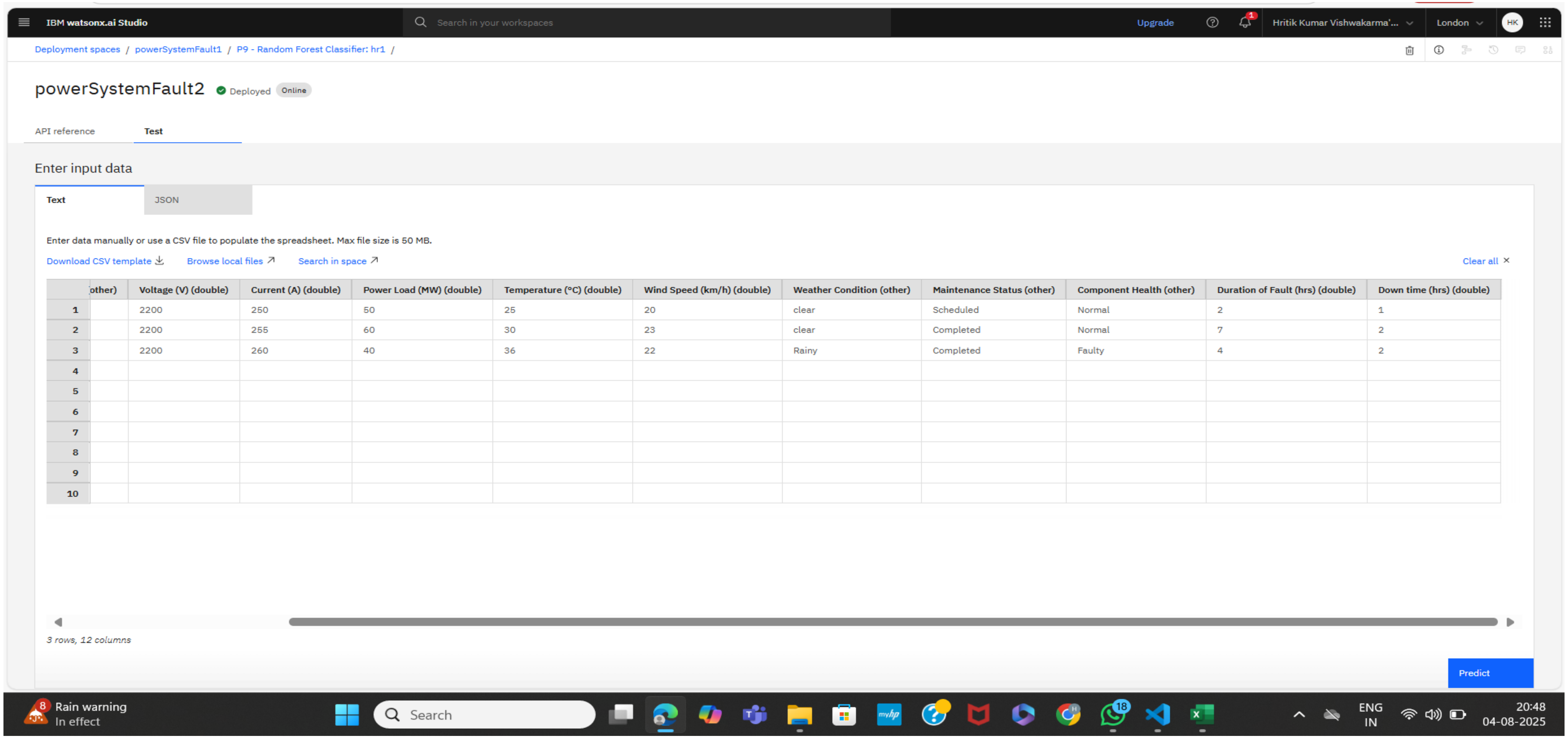


edunet foundation

PROGRESS MAP IMAGE:

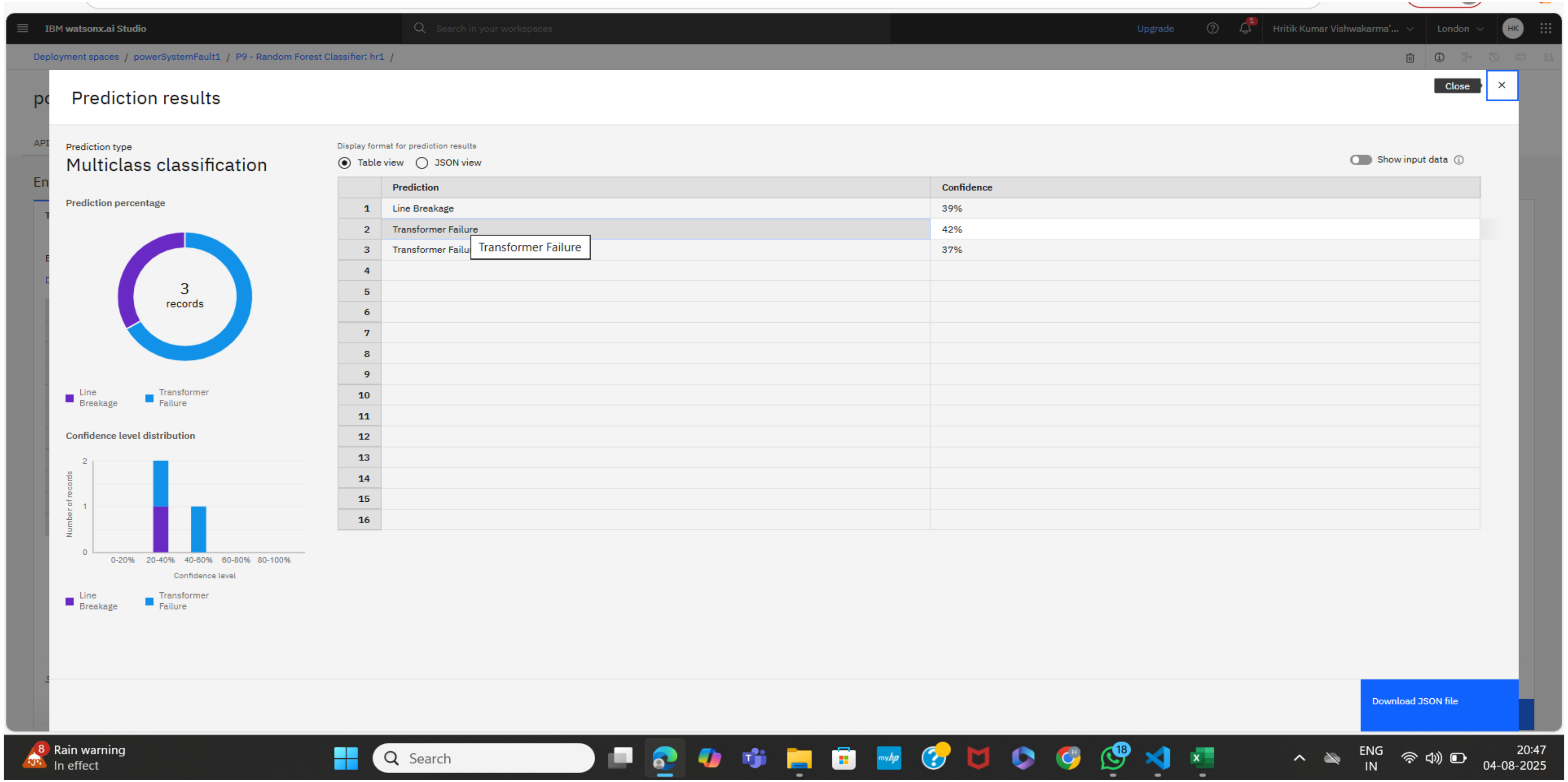


INPUT



	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 (hrs) (double)
1		2200	250	50	25	20	clear	Scheduled	Normal	2	1
2		2200	255	60	30	23	clear	Completed	Normal	7	2
3		2200	260	40	36	22	Rainy	Completed	Faulty	4	2
4											
5											
6											
7											
8											
9											
10											

RESULT



IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Hritik Kumar Vishwakarma'...
London
Deployment spaces / powerSystemFault1 / P9 - Random Forest Classifier: hr1 /
Prediction results
Close
Prediction type
Multiclass classification
Prediction percentage
3
records
Line Breakage
Transformer Failure
Confidence level distribution
Number of records
0-20%
20-40%
40-60%
60-80%
80-100%
Confidence level
Display format for prediction results
Table view
JSON view
Show input data
Prediction
Confidence
1 Line Breakage 39%
2 Transformer Failure 42%
3 Transformer Failure 37%
Download JSON file
Rain warning
In effect
Search
ENG
IN
20:47
04-08-2025

CONCLUSION

- The developed machine learning model successfully detects and classifies power system faults with high accuracy .
- The integration with IBM Cloud Lite enables scalable, cloud-based deployment, providing a reliable solution for real-time fault monitoring.
- This system helps ensure grid stability, reduces downtime, and improves maintenance response times.

FUTURE SCOPE

- **Future Scope**
- Expand the model to handle more complex fault scenarios and dynamic grid conditions
- Integrate the system with IoT-based sensors for real-time data collection
- Develop a mobile or web dashboard for live monitoring and alerts
- Explore edge computing solutions to reduce latency in detection

REFERENCES

- Power System Faults Dataset – Kaggle: https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
- IBM Cloud Documentation: https://cloud.ibm.com/docs
- IEEE Papers on Power System Fault Detection
- Scikit-learn Documentation: https://scikit-learn.org/stable/documentation.html

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

Hritik Kumar Vishwakarma

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 21, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/30f6ec22-3aa3-464f-8838-098ebe7df66e

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Hritik Kumar Vishwakarma

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 24, 2025

Issued by: IBM SkillsBuild

IBM

Verify: https://www.credly.com/badges/ffa0986b-d34e-472f-b688-a1fd4b4b814f

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Hritik Kumar Vishwakarma

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 25 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU