MACHINE LEARNING PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
RAHUL KUMAR SINHA-BIT SINDRI, DHANBAD-INFORMATION TECHNOLOGY

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- 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

1 Data Collection

Gather Data: Collect high-resolution voltage and current phasor measurements (e.g., from Phasor Measurement Units - PMUs) under different conditions:

Normal operating conditions

Fault conditions: line-to-ground, line-to-line, three-phase faults

Data Labeling: Label each data segment with the corresponding state for supervised learning.

2. Data Preprocessing

Filtering & Noise Reduction: Use filters (e.g., low-pass filter) to eliminate noise.

Windowing: Segment data into fixed-length windows (e.g., 100-200 ms) for analysis.

Normalization: Scale features for consistent ranges.

3.Model Selection

Use robust classification algorithms suitable for Multiclass-Classification:

Random Forest Classifier : it provides robustness and also haves high accuracy.

4. Validation & Testing :-

Simulate various fault scenarios to validate model robustness.

Conduct field tests under controlled conditions to ensure reliability.

SYSTEM APPROACH

HARDWARE , SOFTWARE AND STORAGE REQUIREMENT :-

1.watsonx.ai Studio:

- It offers a collaborative environment with tools for data preparation, model development, testing, and deployment, all within a cloud-based interface.

2. watsonx.ai Runtime:

- It provides **8 virtual CPU and 32 GB of RAM.**

3. Cloud object Storage:

- In the free lite version, we are provided with 25 GB per month storage

ALGORITHM & DEPLOYMENT

- In this Algorithm section, we would see the machine learning algorithm chosen for predicting Power fault Detection.
- **Algorithm Selection:**
  - Here Random Forest Classifier is chosen because Random Forest tends to deliver excellent predictive performance by combining multiple decision trees to reduce overfitting and increase generalization which is required here for predicting the power fault type.
- **Data Input:**
  - The inputs taken here are fault type, fault location , voltage, current , power load, Temperature, wind speed, weather , component nature, Duration of fault and Down Time these are the parameters taken here as the inputs.

Training Process:

- **Input Data**: The Random Forest Classifier uses the extracted features along with labels indicating the fault type.
- **Model Training**: Train the Random Forest classifier on this labeled dataset so it learns to distinguish patterns associated with different faults.
- **Prediction Process:**-

Prediction: The Random Forest analyzes the incoming features and predicts whether a fault exists and its specific type.

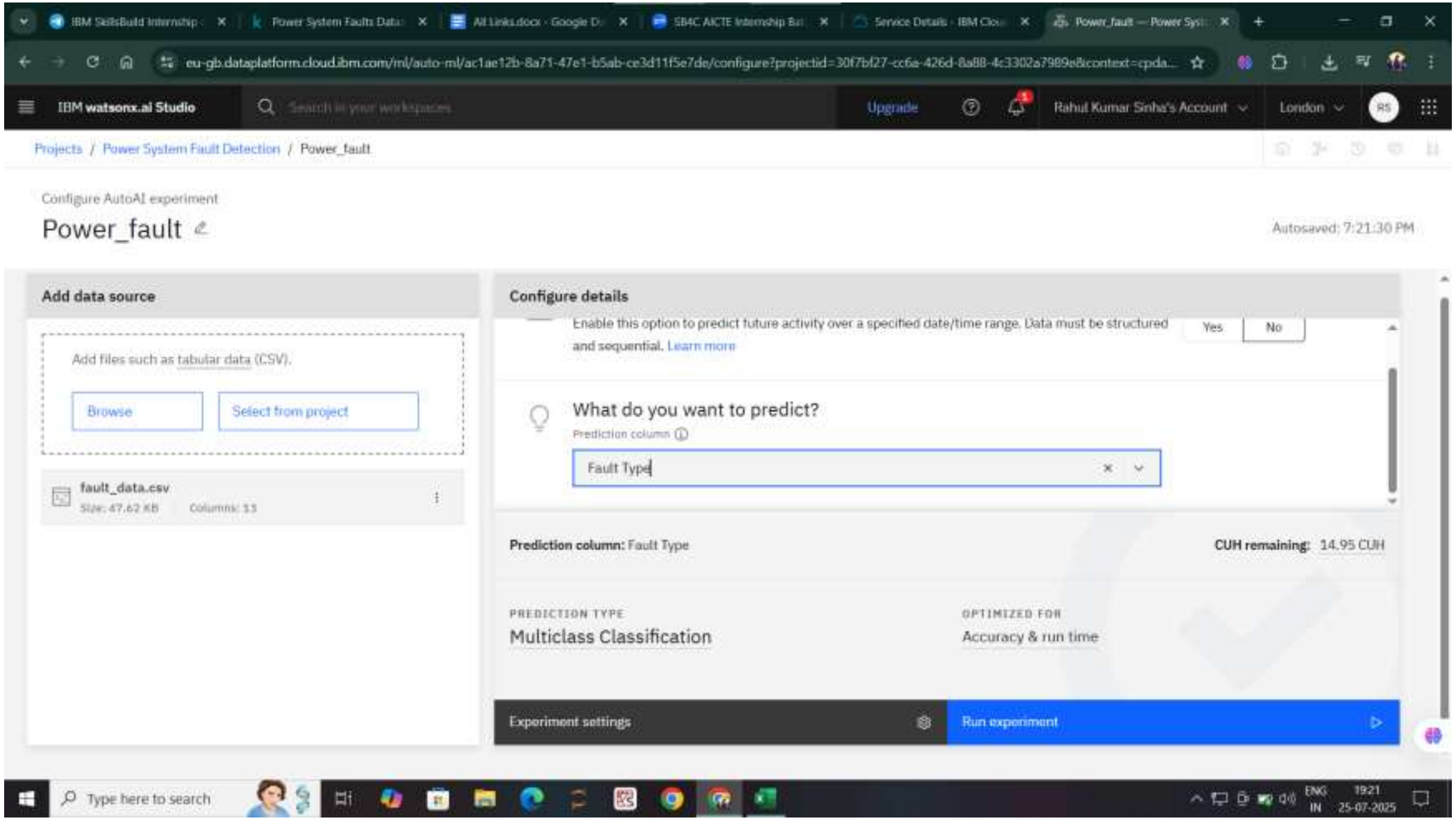
Decision-Making: The system can trigger alerts, isolate faulty sections, or initiate corrective actions based on the prediction.

.

RESULT

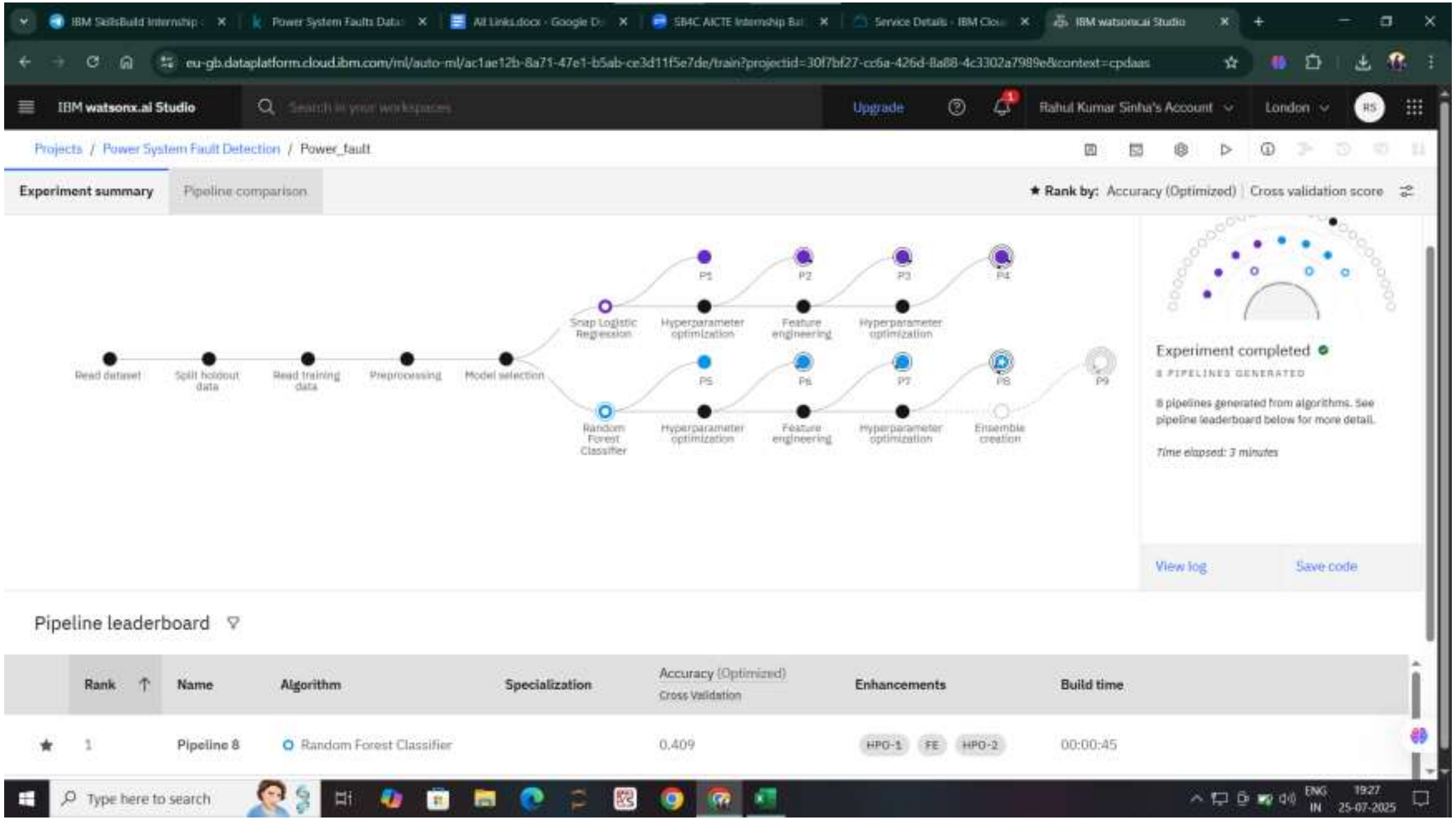
THE RESULTS OF THE POWER SYSTEM FAULT DETECTION AND CLASSIFICATION BY THE USAGE OF WATSONX.AI STUDIO AS A MACHINE LEARNING MODEL ON IBM CLOUD PLATFORM ARE GIVEN AS THE FOLLOWING:-

RESULT



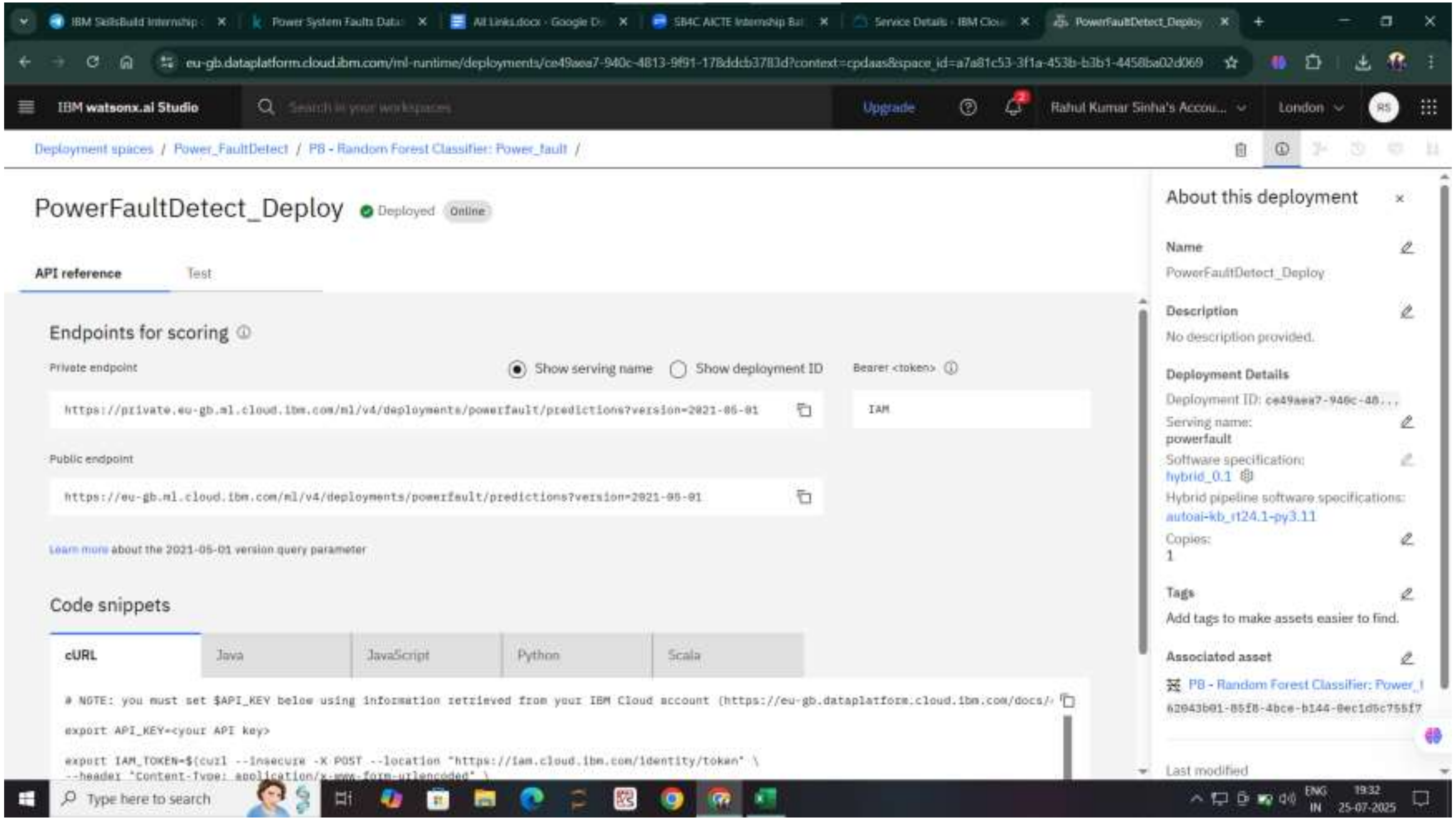
edunet foundation

RESULT



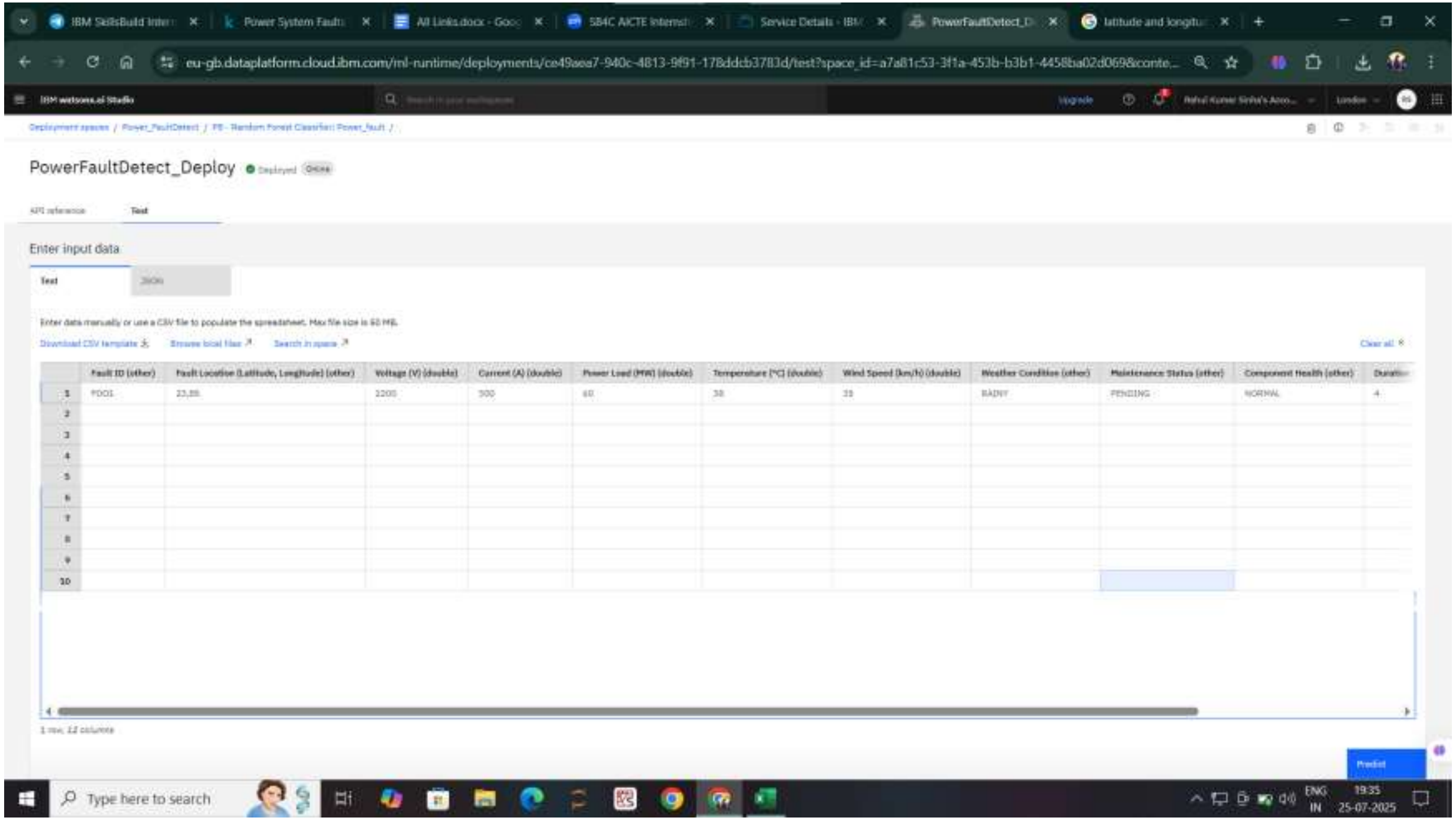
edunet foundation

RESULT



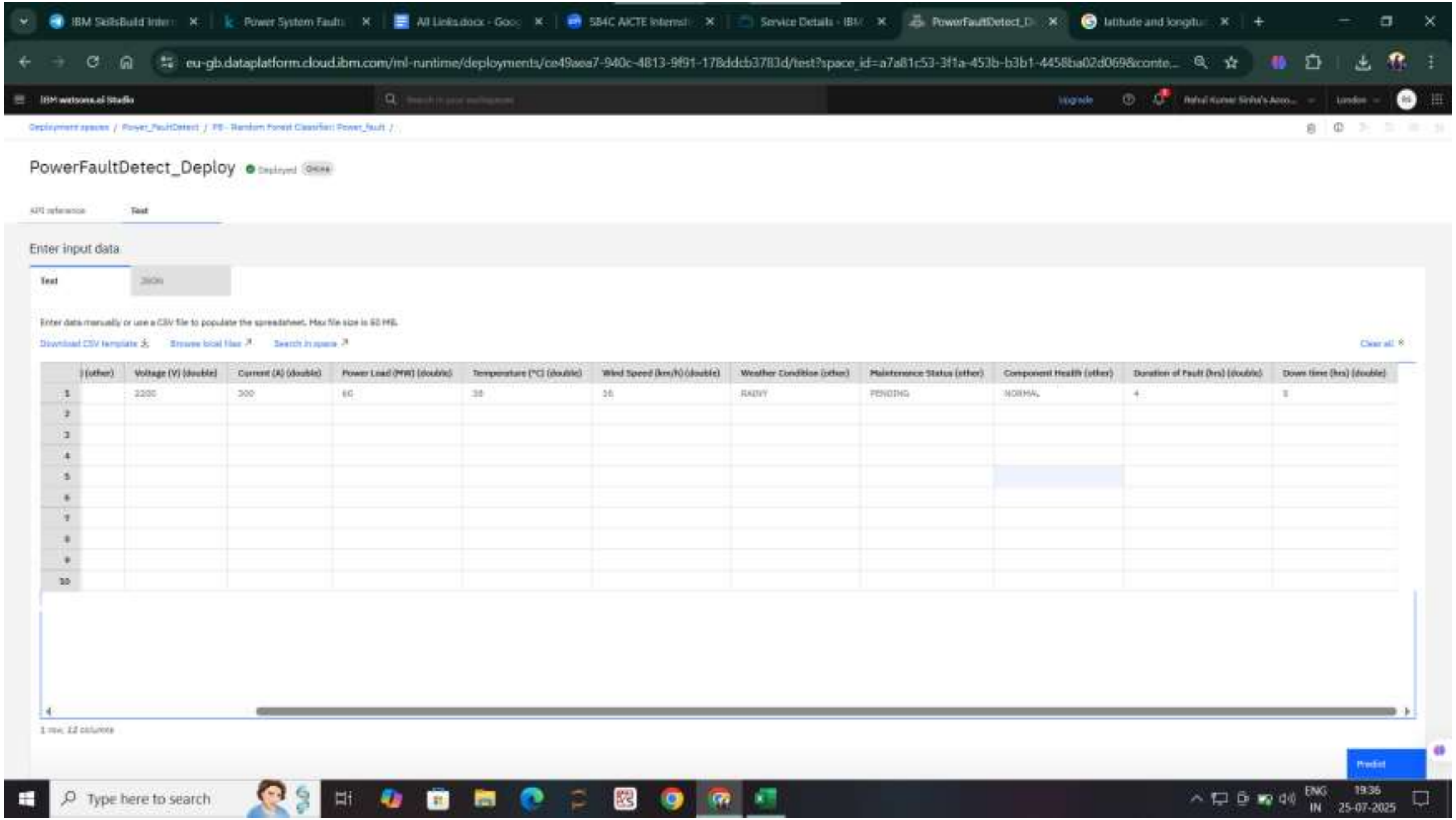
edunet foundation

RESULT



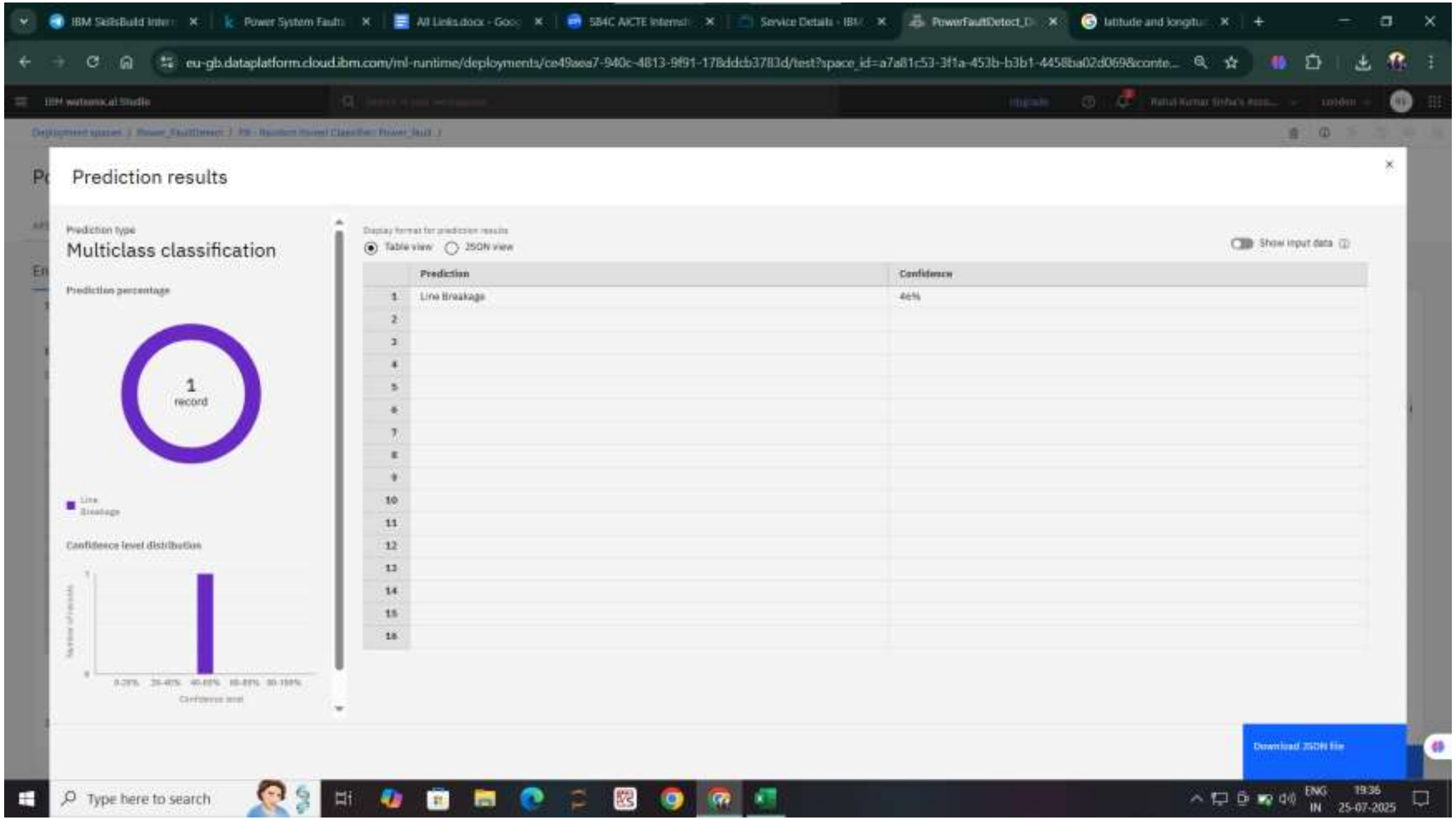
edunet foundation

RESULT



eu-gb.dataplatform.cloud.ibm.com/ml-runtime/deployments/ce49aea7-940c-4813-9f91-178ddcb3783d/test?space_id=a7a81c53-3f1a-453b-b3b1-4458ba02d0698conte...
PowerFaultDetect_Deploy
Enter input data
Enter data manually or use a CSV file to populate the spreadsheet. Max file size is 50 MB.
Download CSV template
Browse local files
Search in space
Clear all
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)
2200
300
60
38
35
RAINY
PENDING
NORMAL
4
5
Type here to search
19:36
25-07-2025

RESULT


eu-gb.dataplatform.cloud.ibm.com/ml-runtime/deployments/ce49aea7-940c-4813-9f91-178ddcb3783d/test?space_id=a7a81c53-3f1a-453b-b3b1-4458ba02d069&conte...
Prediction results
Prediction type
Multiclass classification
Prediction percentage
1
record
Table view
JSON view
Show input data
Prediction
Confidence
Line Breakage
46%
Download JSON file
Type here to search
19:36
25-07-2025

CONCLUSION

- In practical scenarios, utilities might implement this by training a Random Forest with labeled fault data collected during system tests or historical event logs. Once trained, sensors detect anomalies, and the model classifies the faults, enabling rapid and accurate response.
- Implementing a Random Forest classifier for power system faults faces challenges like limited and imbalanced fault data, noise in sensor measurements, selecting relevant features, ensuring real-time processing, adapting to system changes, integration with existing systems, maintaining model accuracy over time, and explaining model decisions. Overcoming these requires good data management, feature engineering, computational resources, and ongoing model updates.

FUTURE SCOPE

This System can be further modified in future by the implementing the following standards:-

1. Integration with IoT and Smart Grids

- Combining fault detection with IoT sensors for real-time, high-resolution monitoring.
- Enabling more accurate and faster fault diagnosis across extensive smart grid networks.

2. Hybrid Models

- Integrating Random Forest with other machine learning algorithms (like deep learning) to improve accuracy and handle complex fault scenarios.
- Using ensemble techniques for more robust fault classification.

3. Automated Data Collection and Self-Training

- Developing systems that automatically collect and label data, allowing continuous model learning and adaptation to evolving system conditions.

4. Enhanced Interpretability

- Improving explainability of model decisions for better trust, regulatory compliance, and system diagnostics.

5. Predictive Maintenance

- Transitioning from fault detection to fault prediction, enabling preventive maintenance and reducing downtime.

REFERENCES

1.Websites :-

- www.google.com
- www.Wikipedia.com
- www.Kaggle.com

2.Books:-

- “Mathematics for Machine Learning" by Deisenroth, Faisal, and Ong

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

RAHUL KUMAR SINHA

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 16, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/c42d9857-ada1-4867-8fde-b37d376163fc

IBM

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

RAHUL KUMAR SINHA

Has successfully satisfied the requirements for:

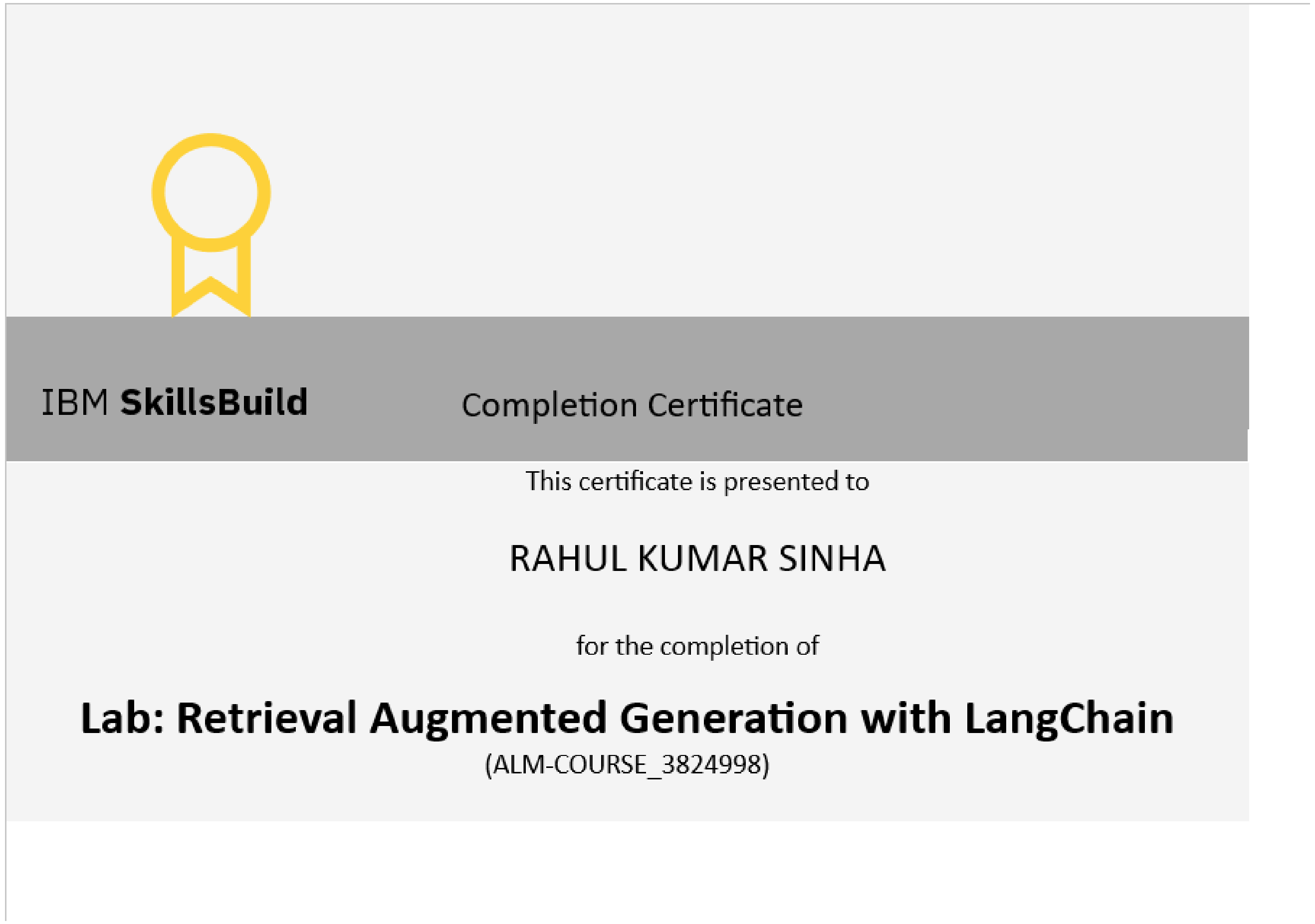
Journey to Cloud: Envisioning Your Solution

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/8f028284-f0e5-427b-85d4-96bbb3d2a167

IBM

IBM CERTIFICATIONS



IBM **SkillsBuild** Completion Certificate

This certificate is presented to

RAHUL KUMAR SINHA

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

THANK YOU !!

edunet
foundation