CAPSTONE PROJE

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented by :
1. UDAY PATIL - Mit Academy Of Engineering - ENTC

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- 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.

In today's power distribution systems, quickly identifying when something goes wrong — like a short circuit or line fault — is critical to keeping electricity flowing safely and reliably. The challenge is to build a machine learning model that can detect these faults and figure out exactly what type of fault it is, using live data like voltage and current readings. By doing this, we can help grid operators respond faster, prevent bigger failures, and improve the overall stability of the power grid.

PROPOSED SOLUTION

- **The system trains a machine learning model for fault classification using IBM Watson Studio and AutoAI on IBM Cloud. To differentiate between different types of faults, it analyzes environmental data, voltage, and current. A real-time fault event classification for grid operators is the end result.**

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

- **System Requirements:**

  8 cpu & 32GB ram

  IBM Cloud Lite account

  IBM Watson Studio

  Kaggle dataset

- **Technologies Used:**

  IBM Watson AutoAI

  IBM Cloud Object Storage

  Python (optional for preprocessing)

ALGORITHM & DEPLOYMENT

- **Algorithm:** Snap Random Forest Classifier
- **Main Input Features:**

  Voltage

  Current

  Power Load

  Temperature

  Wind Speed
- **Process:**

  Data cleaned and fed into AutoAI

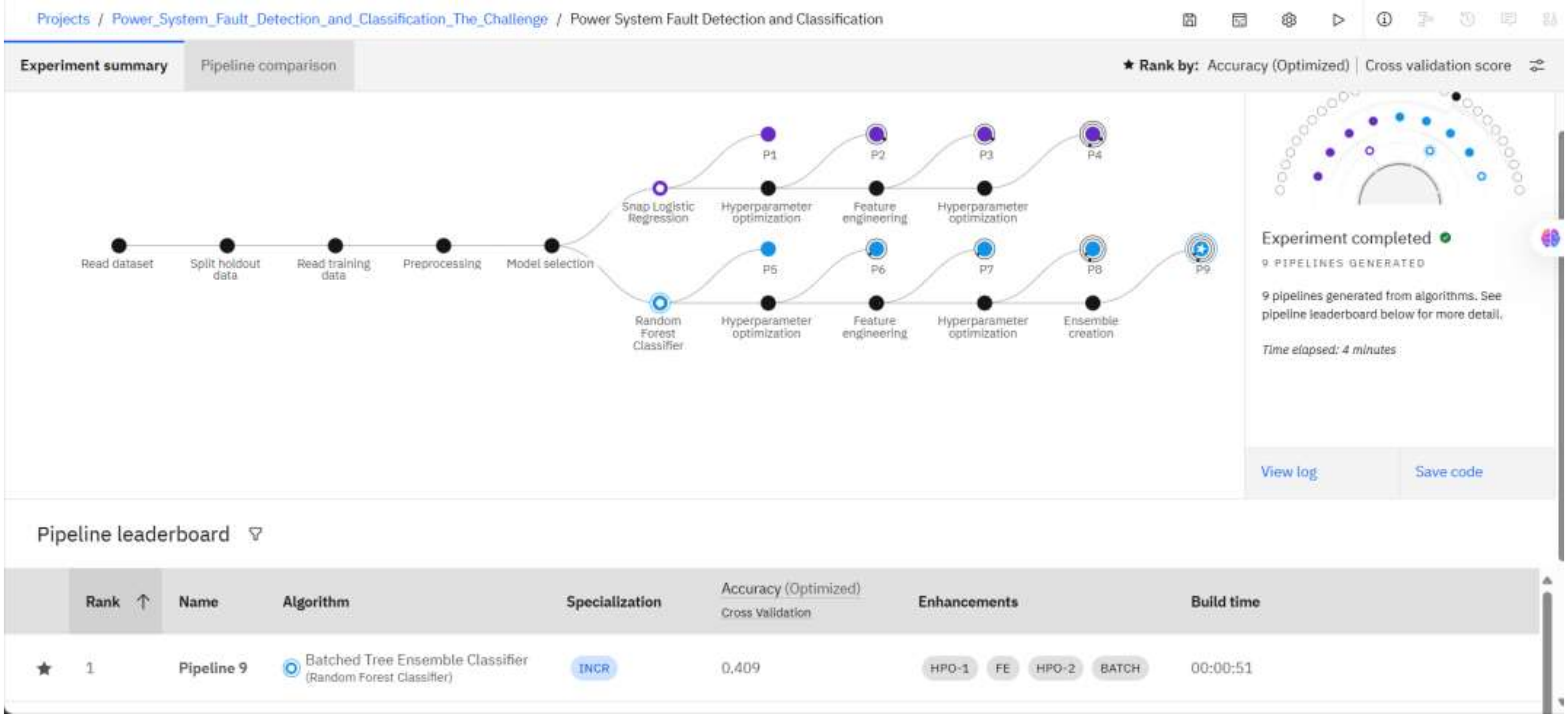
  Best pipeline selected based on accuracy
- **Deployment:**

  Watson ML deployed endpoint used for real-time predictions
- **Accuracy Achieved:** 40.9%

RESULT

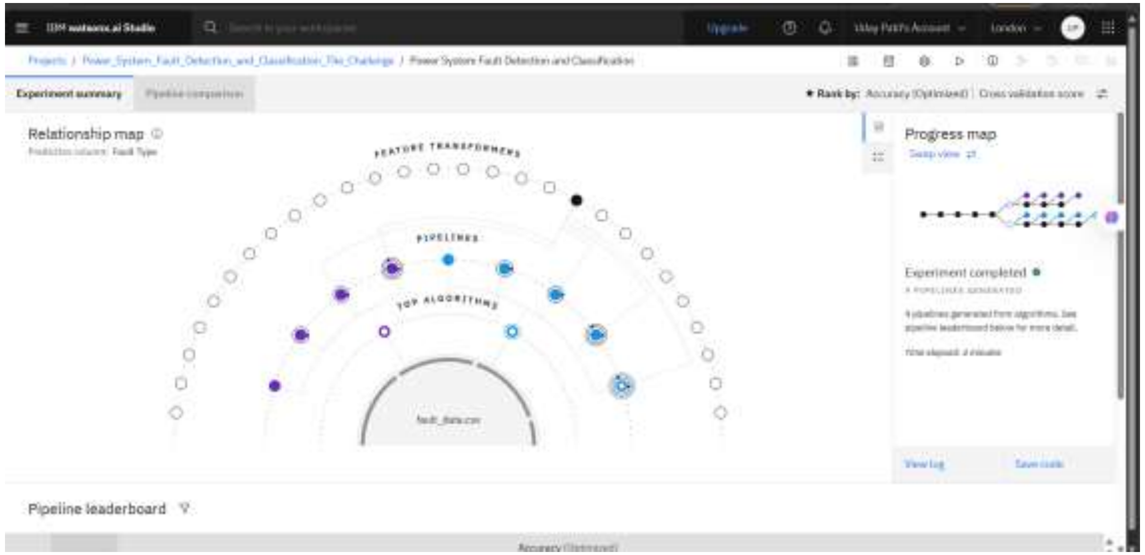
- The model successfully classifies different fault types using input data. Output is presented in tabular and JSON format through the deployed service.
  Accuracy: 40.9% (Snap Random Forest pipeline)

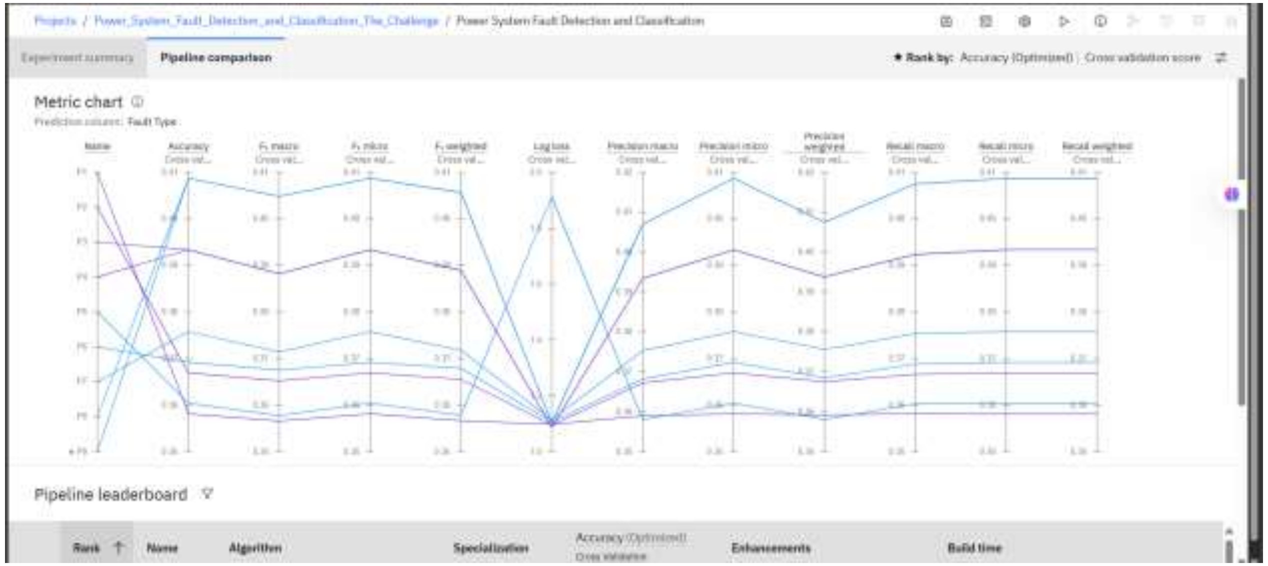
RESULT



Projects / Power_System_Fault_Detection_and_Classification_The_Challenge / Power System Fault Detection and Classification
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Read dataset
Split holdout data
Read training data
Preprocessing
Model selection
Snap Logistic Regression
Random Forest Classifier
Hyperparameter optimization
Feature engineering
Ensemble creation
P1
P2
P3
P4
P5
P6
P7
P8
P9
Experiment completed
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 4 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
00:00:51

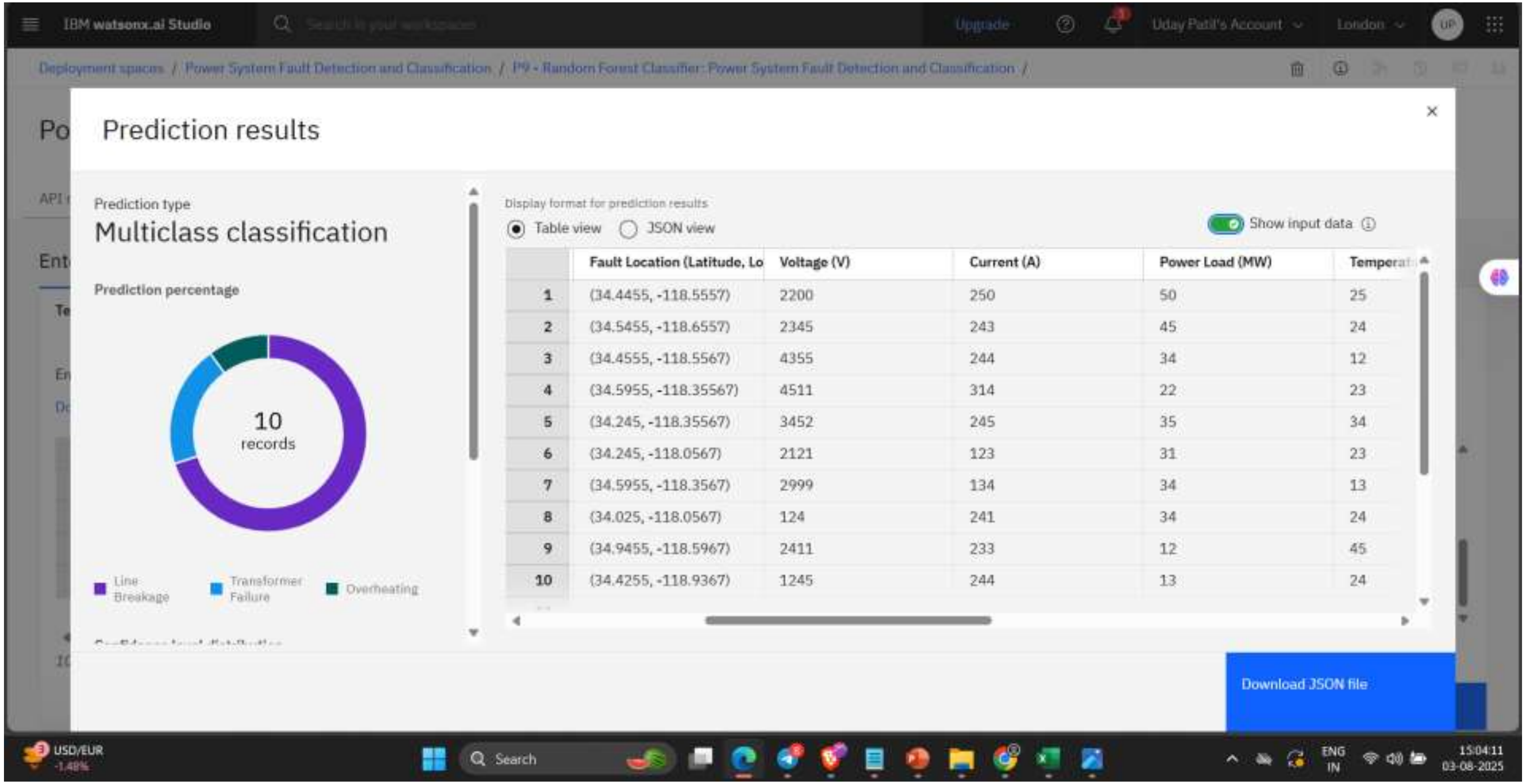
RESULT



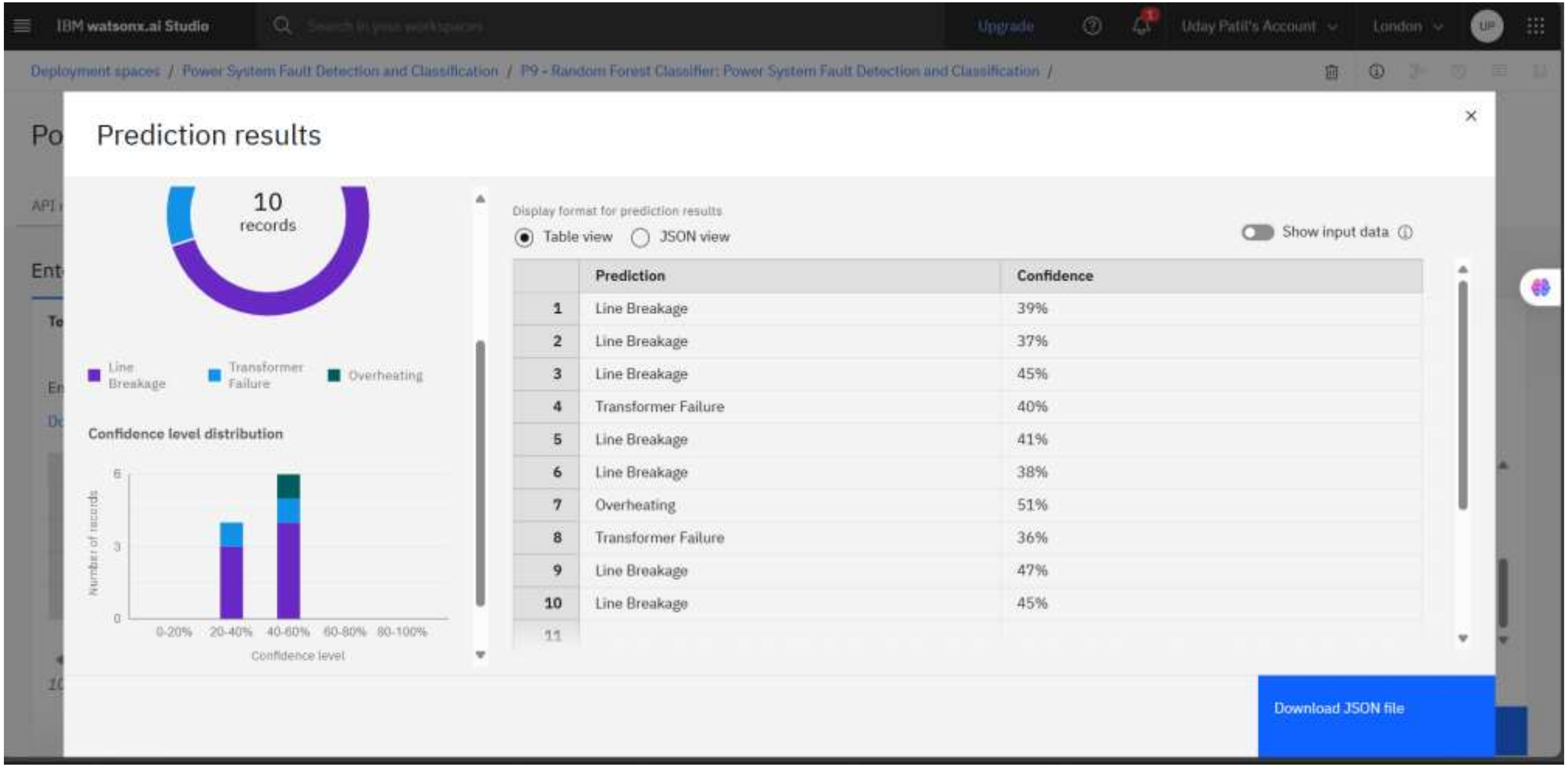


	Rank ↑	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 9	Batched Tree Ensemble Classifier (Random Forest Classifier)	INCR	0.409	HPO-1 FE HPO-2 BATCH	00:00:51
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:48
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:31
	4	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:00:28

RESULT


IBM watsonx.ai Studio
Upgrade
Uday Patil's Account
London
Deployment spaces / Power System Fault Detection and Classification / P9 - Random Forest Classifier: Power System Fault Detection and Classification /
Prediction results
Prediction type
Multiclass classification
Prediction percentage
10
records
Line Breakage
Transformer Failure
Overheating
Display format for prediction results
Table view
JSON view
Show input data
Fault Location (Latitude, Lo | Voltage (V) | Current (A) | Power Load (MW) | Temperat
1 | (34.4455, -118.5557) | 2200 | 250 | 50 | 25
2 | (34.5455, -118.6557) | 2345 | 243 | 45 | 24
3 | (34.4555, -118.5567) | 4355 | 244 | 34 | 12
4 | (34.5955, -118.35567) | 4511 | 314 | 22 | 23
5 | (34.245, -118.35567) | 3452 | 245 | 35 | 34
6 | (34.245, -118.0567) | 2121 | 123 | 31 | 23
7 | (34.5955, -118.3567) | 2999 | 134 | 34 | 13
8 | (34.025, -118.0567) | 124 | 241 | 34 | 24
9 | (34.9455, -118.5967) | 2411 | 233 | 12 | 45
10 | (34.4255, -118.9367) | 1245 | 244 | 13 | 24
Download JSON file
USD/EUR
-1.48%
Search
ENG
IN
15:04:11
03-08-2025

RESULT



IBM watsonx.ai Studio
Upgrade
Uday Patil's Account
London
Deployment spaces / Power System Fault Detection and Classification / P9 - Random Forest Classifier: Power System Fault Detection and Classification /
Prediction results
10
records
Line Breakage
Transformer Failure
Overheating
Confidence level distribution
Number of records
6
3
0
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 Line Breakage 37%
3 Line Breakage 45%
4 Transformer Failure 40%
5 Line Breakage 41%
6 Line Breakage 38%
7 Overheating 51%
8 Transformer Failure 36%
9 Line Breakage 47%
10 Line Breakage 45%
11
Download JSON file

CONCLUSION

- The machine learning model shows how artificial intelligence (AI) can support fault classification in real time. Even though the accuracy is 40.9% now, results can be improved with more data and fine-tuning. Decision-making in power grid operations is supported by the system.

FUTURE SCOPE

•Add GPS-based fault location analysis

•Improve model accuracy with deep learning

•Integrate real-time sensor streaming

•Expand system to classify subtypes of each fault

•Add dashboard for visualization

REFERENCES

- https://www.kaggle.com/datasets/ziya07/

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

UDAY PATIL

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

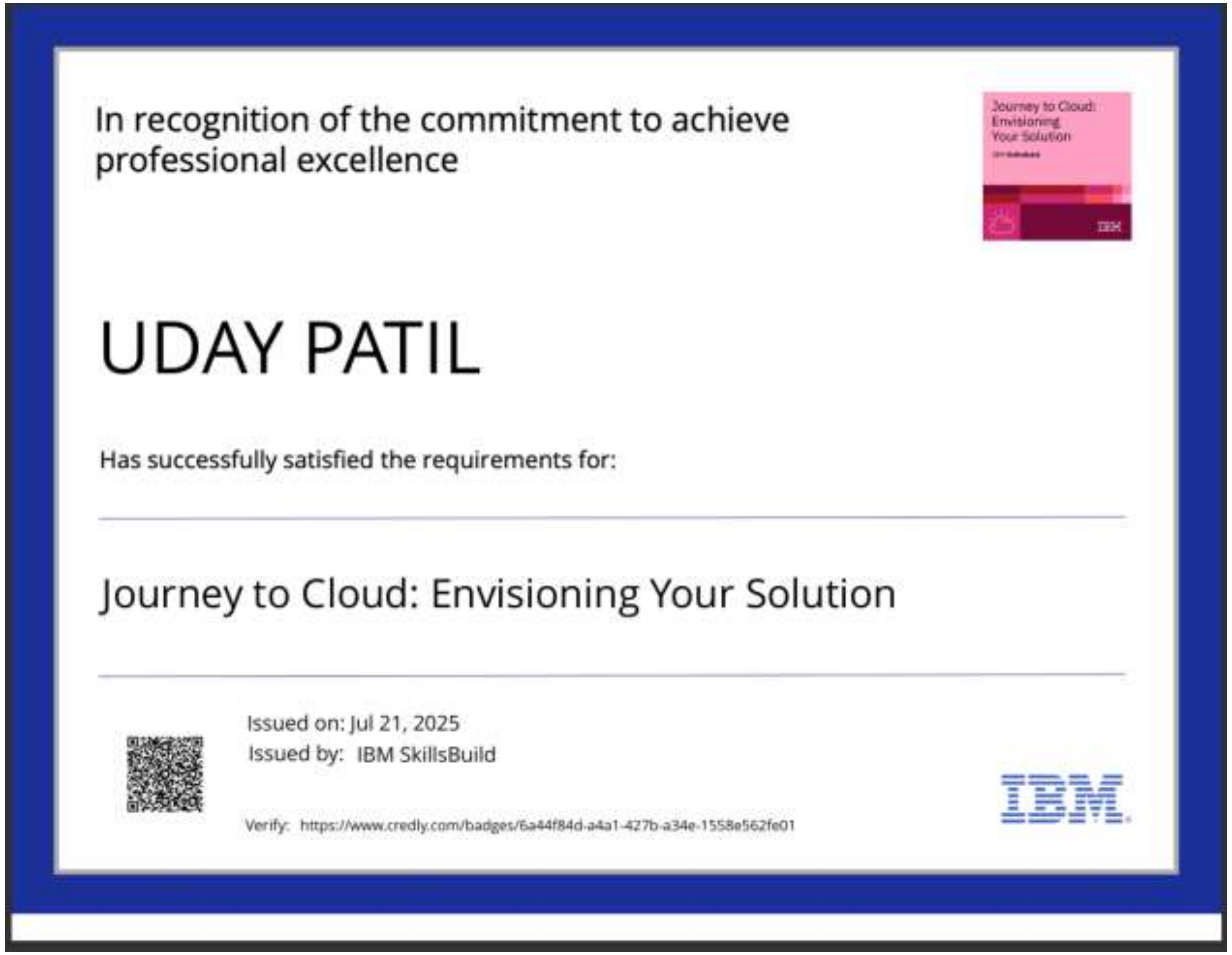
Issued on: Jul 20, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/a3d0f9e6-9f0e-4169-8697-77628633ccd2

IBM

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

UDAY PATIL

Has successfully satisfied the requirements for:

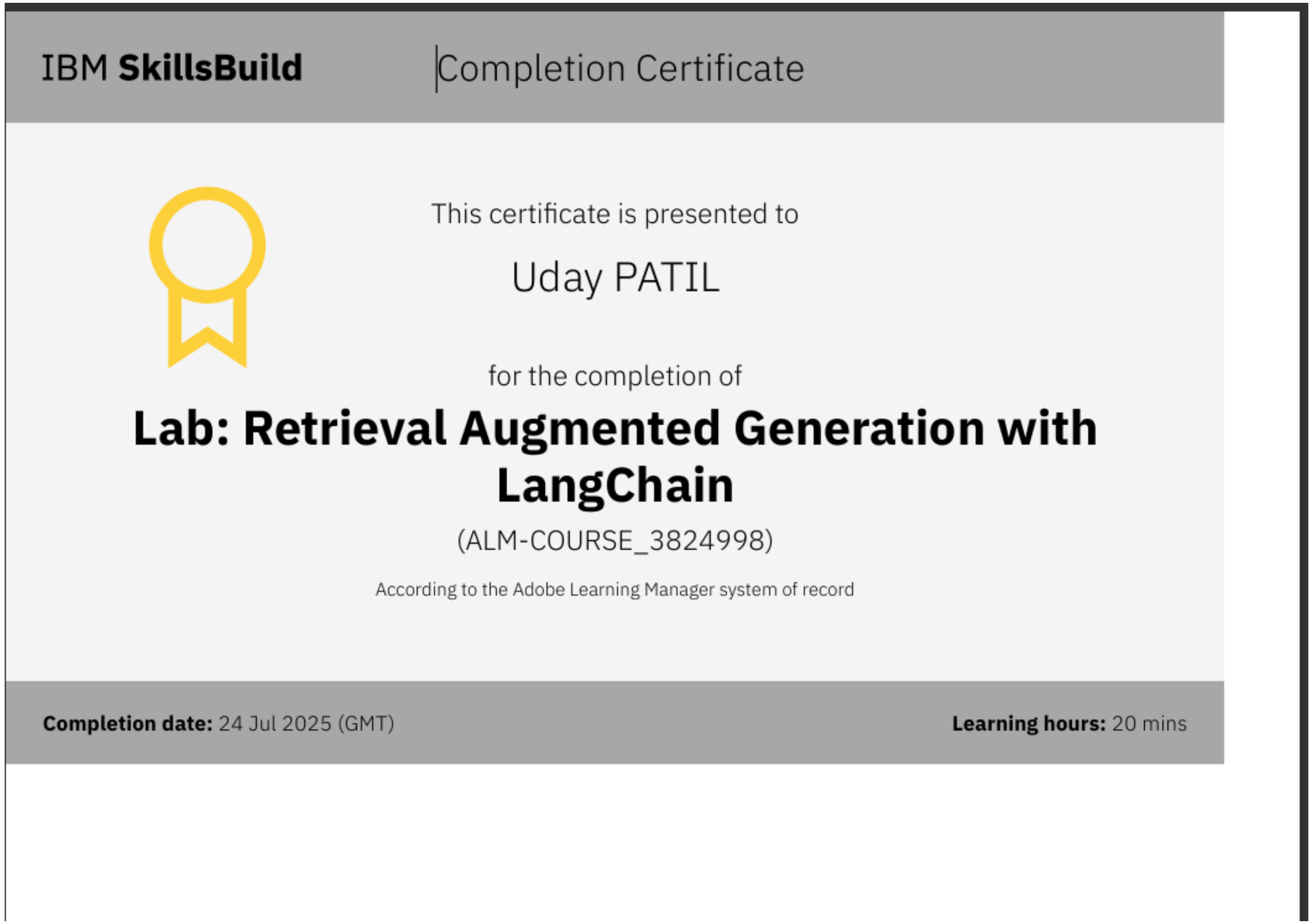
Journey to Cloud: Envisioning Your Solution

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

Verify: https://www.credly.com/badges/6a44f84d-a4a1-427b-a34e-1558e562fe01

IBM

IBM CERTIFICATIONS



IBM **SkillsBuild** | Completion Certificate

This certificate is presented to

Uday PATIL

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