CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:
1. Aditya Mishra-Akido College of Engineering-Computer Science

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

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

- Our solution is an automated machine learning system built on IBM Cloud to accurately classify power system faults.
- **Data Foundation:** The model is built using the "Power System Faults Dataset" from Kaggle, containing thousands of electrical measurements (current and voltage phasors) under various fault and normal conditions.
- **Automated Model Development:** We utilized the **IBM Watsonx.ai AutoAI** service. This tool automates the most time-consuming parts of the ML lifecycle:
  - Data Preprocessing
  - Feature Engineering
  - Model Selection
  - Hyperparameter Optimization
- **Model Deployment:** The best-performing model pipeline is deployed as a web service on IBM Cloud. This allows for real-time predictions through an API or a user-friendly interface.
- **Evaluation:** The system automatically evaluated multiple algorithms and selected the one with the highest accuracy for the final deployment.

SYSTEM APPROACH

- **System Requirements**
- **Platform:** An active IBM Cloud account with access to the Watsonx.ai Studio service was mandatory for this project.
- **Dataset:** A structured dataset containing features and a target "Fault Type" column for supervised learning.
- **Libraries Used**
- The IBM AutoAI service automates the use of necessary libraries. The core Python libraries utilized by the platform for this project include:
- **Scikit-learn:** For implementing the underlying machine learning algorithms, including the selected Random Forest Classifier.
- **Pandas:** For efficient data handling, loading, and manipulation during the experiment.
- **IBM-watson-machine-learning:** The official IBM client library used to programmatically run experiments and manage model deployments on the cloud.

ALGORITHM & DEPLOYMENT

- **Algorithm Selection**
- A **Random Forest Classifier** was chosen as the final algorithm for this project. A Random Forest is an ensemble learning method that builds multiple decision trees and merges them to get a more accurate and stable prediction.
- **Justification:** This algorithm was selected by the **IBM AutoAI experiment**, which automatically tested several models and ranked the Random Forest as the highest-performing pipeline based on prediction accuracy for this specific dataset.

ALGORITHM & DEPLOYMENT

Data Input

- The model uses the following input features to make a prediction:
    - Power Load (MW)
    - Temperature (°C)
    - Wind Speed (km/h)
    - Weather Condition
    - Maintenance Status
    - Component Health
    - Duration of Fault (hrs)
    - Down time (hrs)

ALGORITHM & DEPLOYMENT

- **Training Process**
- The algorithm was trained on a historical dataset of power system measurements.
- The entire training process was automated using **IBM's AutoAI** service, which handled key techniques such as:
  - **Cross-validation** to ensure model robustness.
  - **Hyperparameter Optimization (HPO)** to find the best model settings.
  - **Feature Engineering (FE)** to create the most impactful input features.

ALGORITHM & DEPLOYMENT

Prediction Process

•The trained model, deployed as an online service named **Fault_dp2**, makes predictions on new data submitted in real-time.

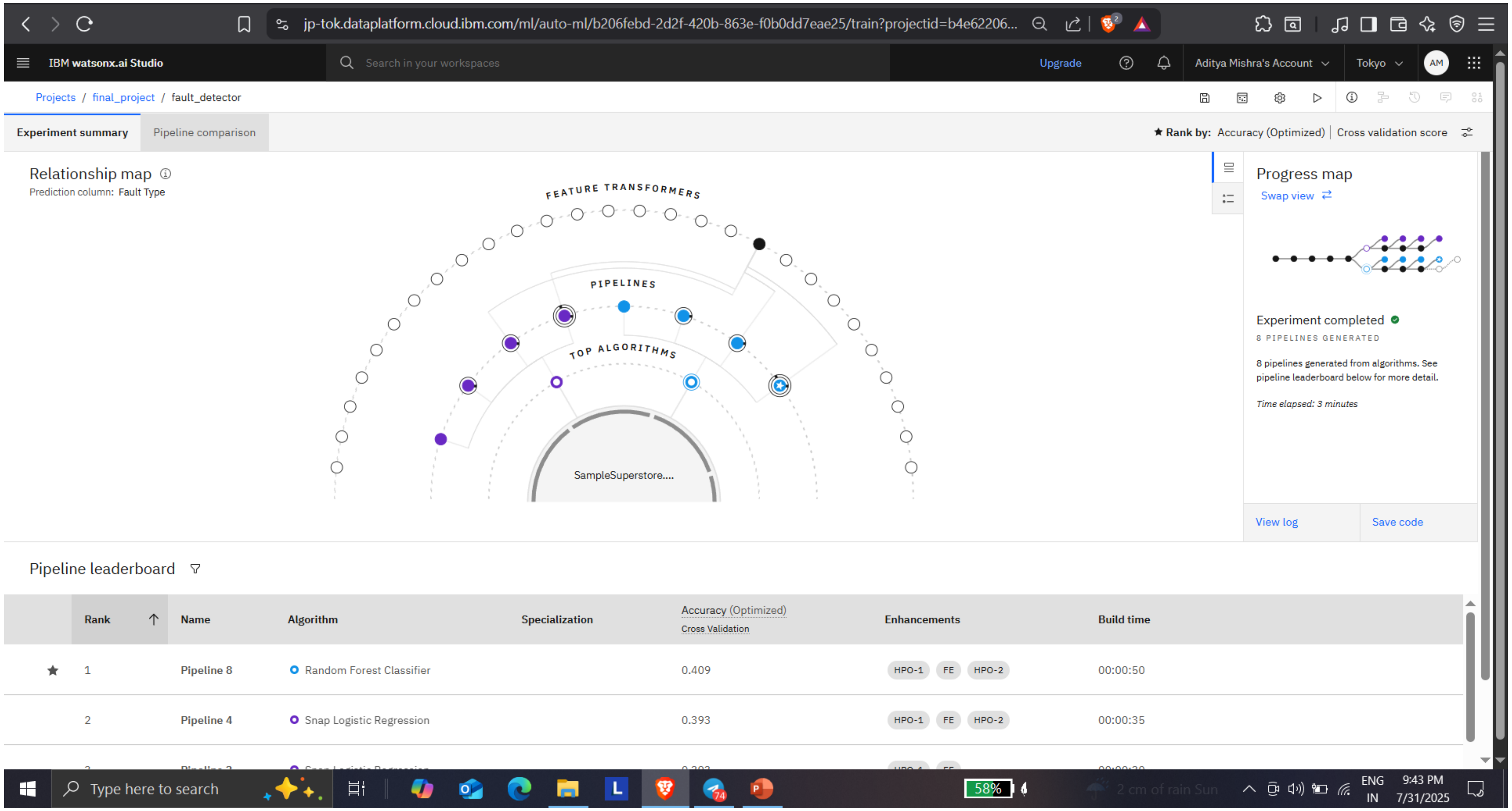
•When a user provides new input values, the model analyzes them and classifies the event into a specific **Fault Type**, such as "Transformer Failure" or "Overheating".

•The system also returns a **confidence score** for each prediction,

• indicating the model's certainty in its answer.

edunet foundation

RESULT


jp-tok.dataplatform.cloud.ibm.com/ml/auto-ml/b206febd-2d2f-420b-863e-f0b0dd7eae25/train?projectid=b4e62206...

IBM watsonx.ai Studio

Projects / final_project / fault_detector

Experiment summary | Pipeline comparison

★ Rank by: Accuracy (Optimized) | Cross validation score

Relationship map
Prediction column: Fault Type

Progress map
Swap view

Experiment completed
8 PIPELINES GENERATED

8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 3 minutes

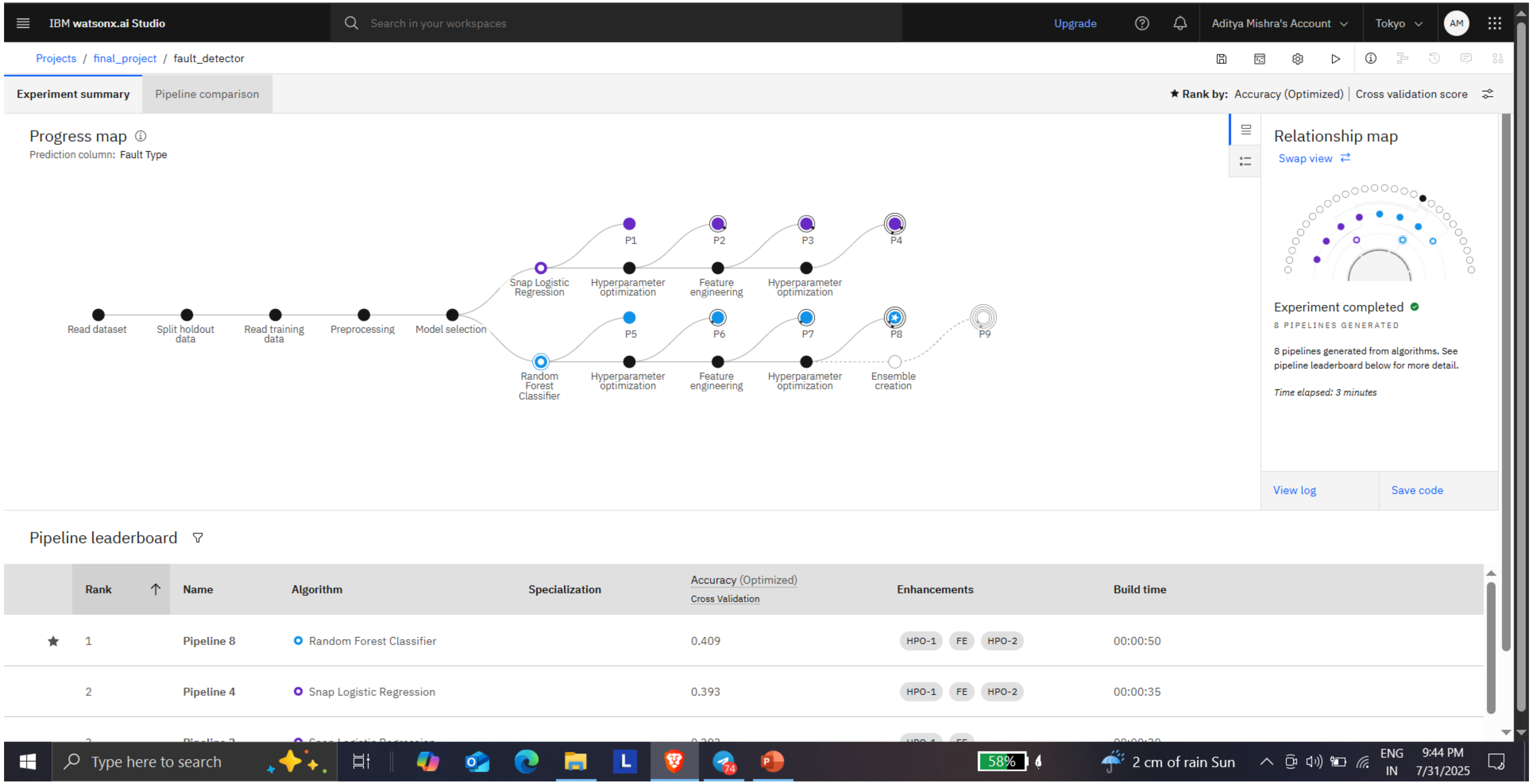
View log | Save code

Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:50
	2	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:35

edunet foundation

RESULT



IBM watsonx.ai Studio

Projects / final_project / fault_detector

Experiment summary | Pipeline comparison

★ Rank by: Accuracy (Optimized) | Cross validation score

Progress map
Prediction column: Fault Type

Relationship map
Swap view

Experiment completed
8 PIPELINES GENERATED

8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 3 minutes

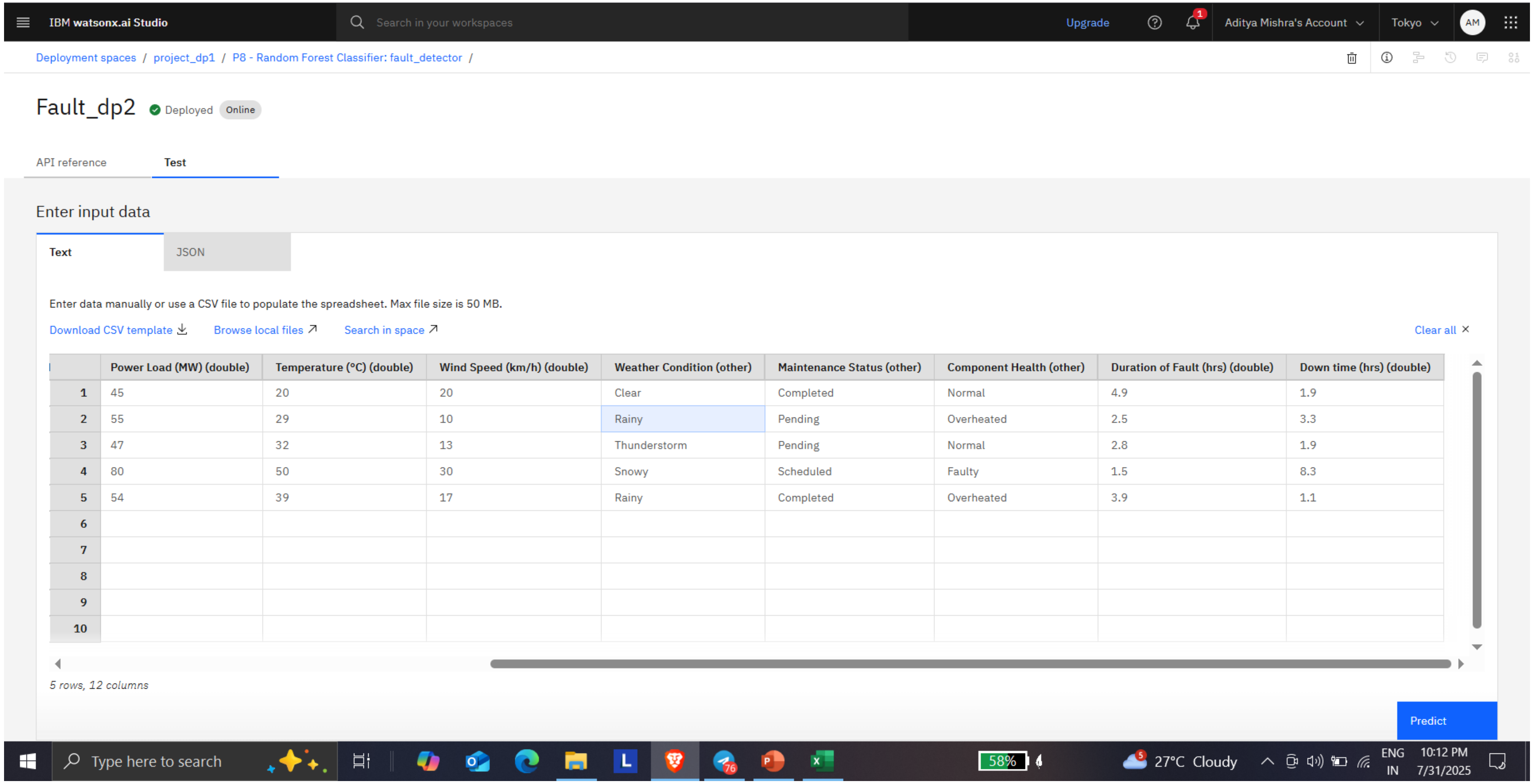
View log | Save code

Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:50
	2	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:35

edunet
foundation

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Aditya Mishra's Account
Tokyo
AM
Deployment spaces / project_dp1 / P8 - Random Forest Classifier: fault_detector /

Fault_dp2 Deployed Online

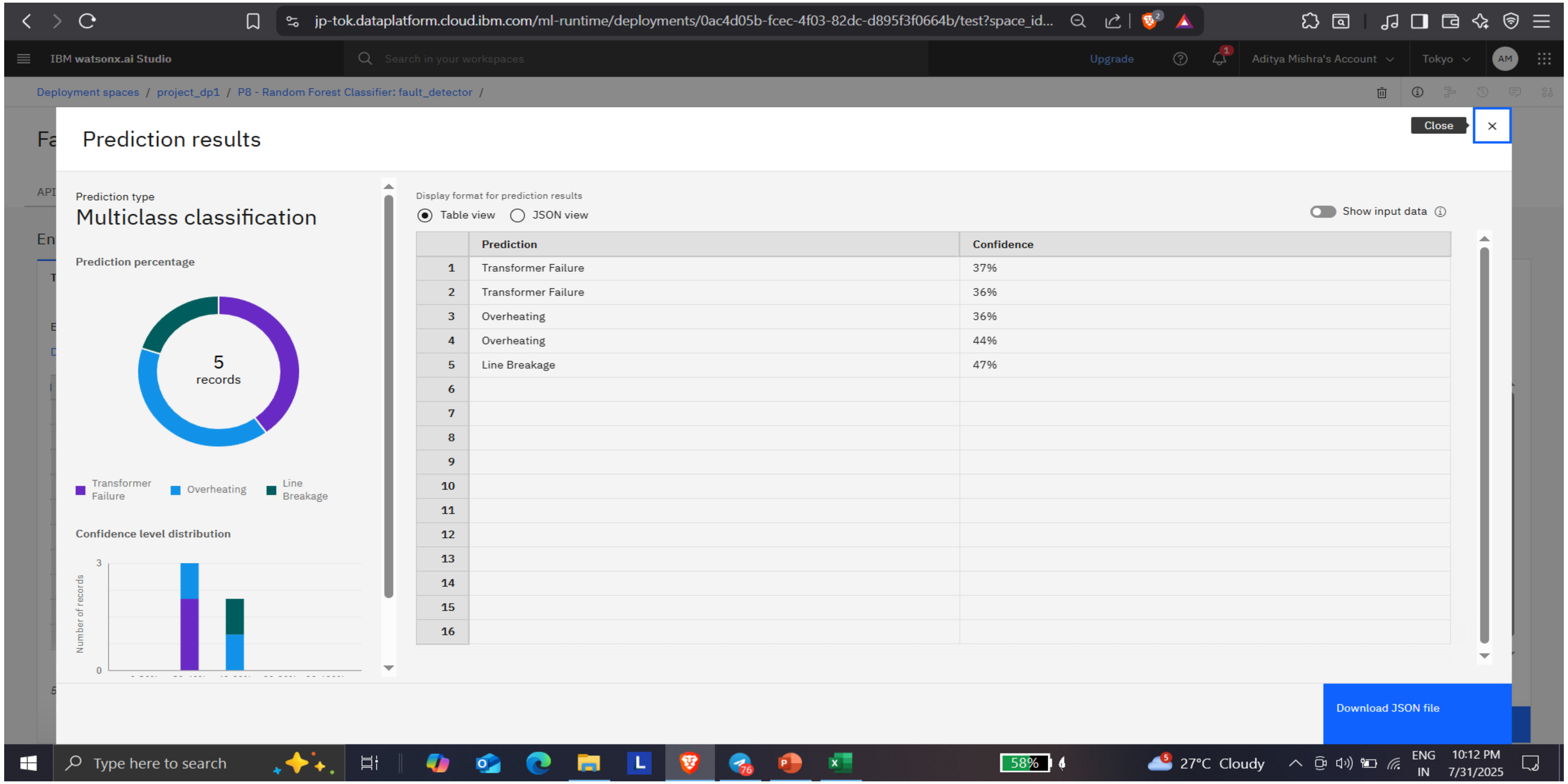
Enter input data

Enter data manually or use a CSV file to populate the spreadsheet. Max file size is 50 MB.

	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	45	20	20	Clear	Completed	Normal	4.9	1.9
2	55	29	10	Rainy	Pending	Overheated	2.5	3.3
3	47	32	13	Thunderstorm	Pending	Normal	2.8	1.9
4	80	50	30	Snowy	Scheduled	Faulty	1.5	8.3
5	54	39	17	Rainy	Completed	Overheated	3.9	1.1
6								
7								
8								
9								
10								

5 rows, 12 columns

RESULT



Prediction results

Prediction type

Multiclass classification

Prediction percentage

5
records
Transformer Failure
Overheating
Line Breakage

Confidence level distribution

Display format for prediction results

(● Table view) ○ JSON view

Show input data

	Prediction	Confidence
1	Transformer Failure	37%
2	Transformer Failure	36%
3	Overheating	36%
4	Overheating	44%
5	Line Breakage	47%
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

jp-tok.dataplatform.cloud.ibm.com/ml-runtime/deployments/0ac4d05b-fcec-4f03-82dc-d895f3f0664b/test?space_id...
IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Aditya Mishra's Account
Tokyo
Deployment spaces / project_dp1 / P8 - Random Forest Classifier: fault_detector /
Close
Download JSON file
Type here to search
58%
27°C Cloudy
ENG IN
10:12 PM
7/31/2025

edunet foundation

CONCLUSION

•We have successfully designed, built, and deployed a machine learning model for the **detection and classification of power system faults**.

•The use of **IBM Watsonx.ai and its AutoAI** capability was instrumental in rapidly developing and comparing multiple complex model pipelines.

•The final deployed model, using a **Random Forest Classifier**, can provide real-time predictions, which is a crucial step towards creating more reliable and stable power grids.

•This project demonstrates the practical application of cloud-based AI services in solving critical challenges in the Electrical Engineering domain.

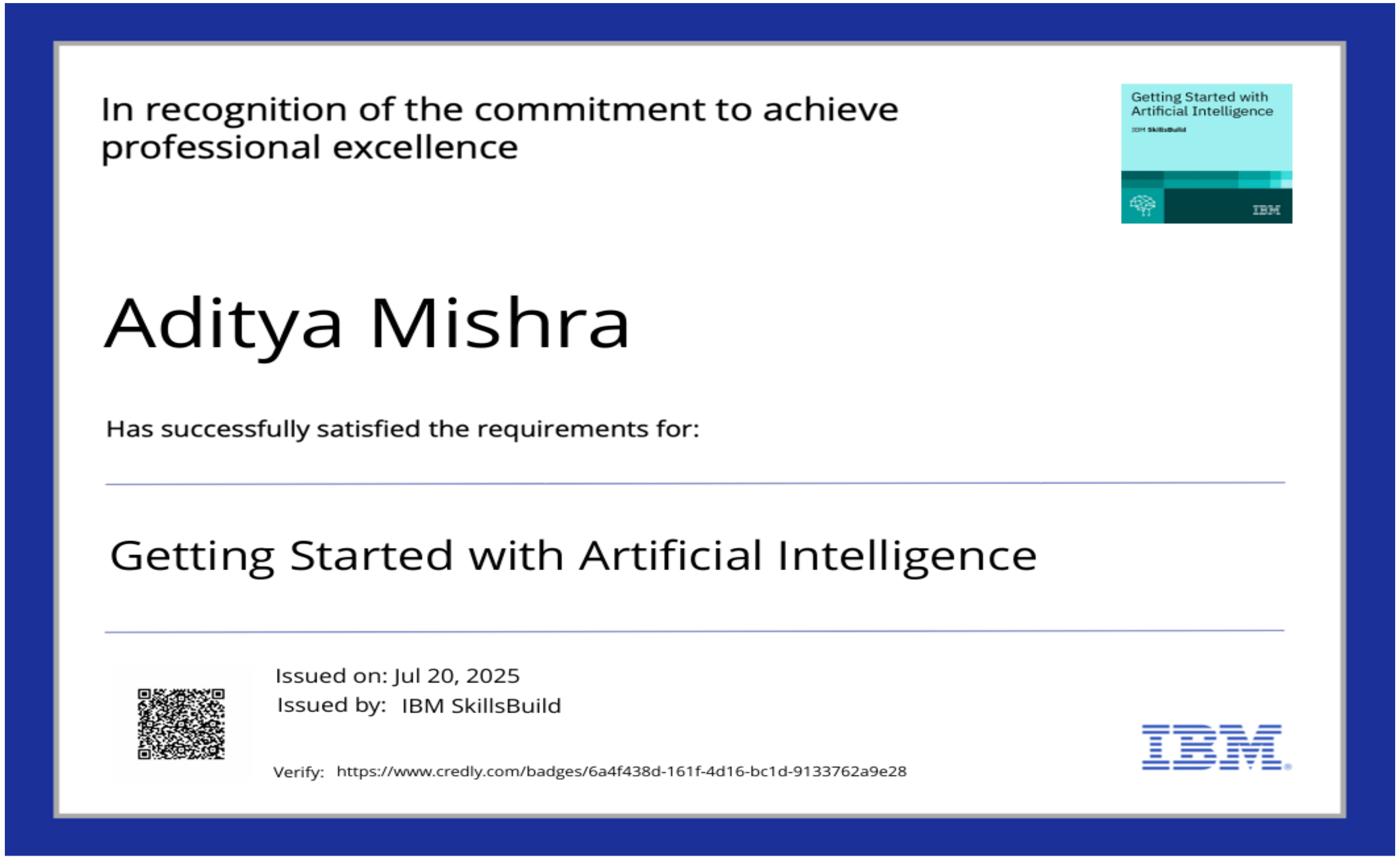
FUTURE SCOPE

- **Real-Time Data Integration:**
- Connect the model directly to live data streams from power grid sensors (SCADA systems) to enable fully autonomous, real-time fault detection instead of manual data entry.
- **Integration and Alerting:**
- Integrate the deployed API with an automated alerting system (via email, SMS, or a live dashboard) to instantly notify grid operators and engineers of detected faults.
- **Develop a Prescriptive System:**
- Move beyond predictive analytics (what will happen) to prescriptive analytics (what to do). The system could suggest immediate response actions or automatically generate maintenance tickets based on the detected fault type.

REFERENCES

- **Dataset:**
- Ziya, A. (2024). *Power System Faults Dataset*. Kaggle. Retrieved from: https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
- **Technology & Documentation:**
- IBM. (2024). *IBM Watsonx.ai Documentation*. Retrieved from: https://cloud.ibm.com/docs/watsonx-ai
- IBM. (2024). *AutoAI Documentation*. Retrieved from: https://cloud.ibm.com/docs/watsonx-ai?topic=watsonx-ai-autoai

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

Aditya Mishra

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

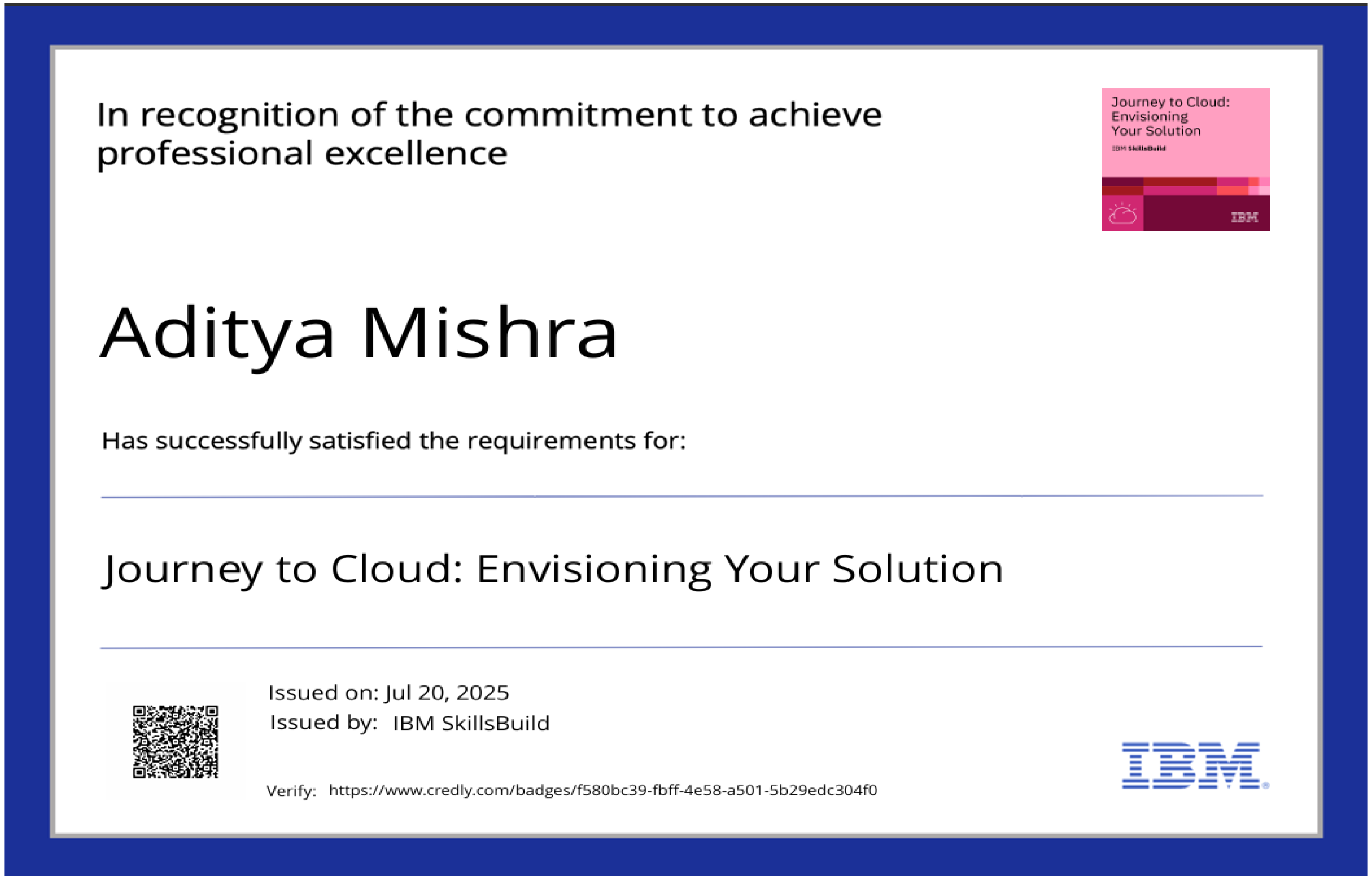
Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/6a4f438d-161f-4d16-bc1d-9133762a9e28

IBM

edunet foundation

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Aditya Mishra

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 20, 2025

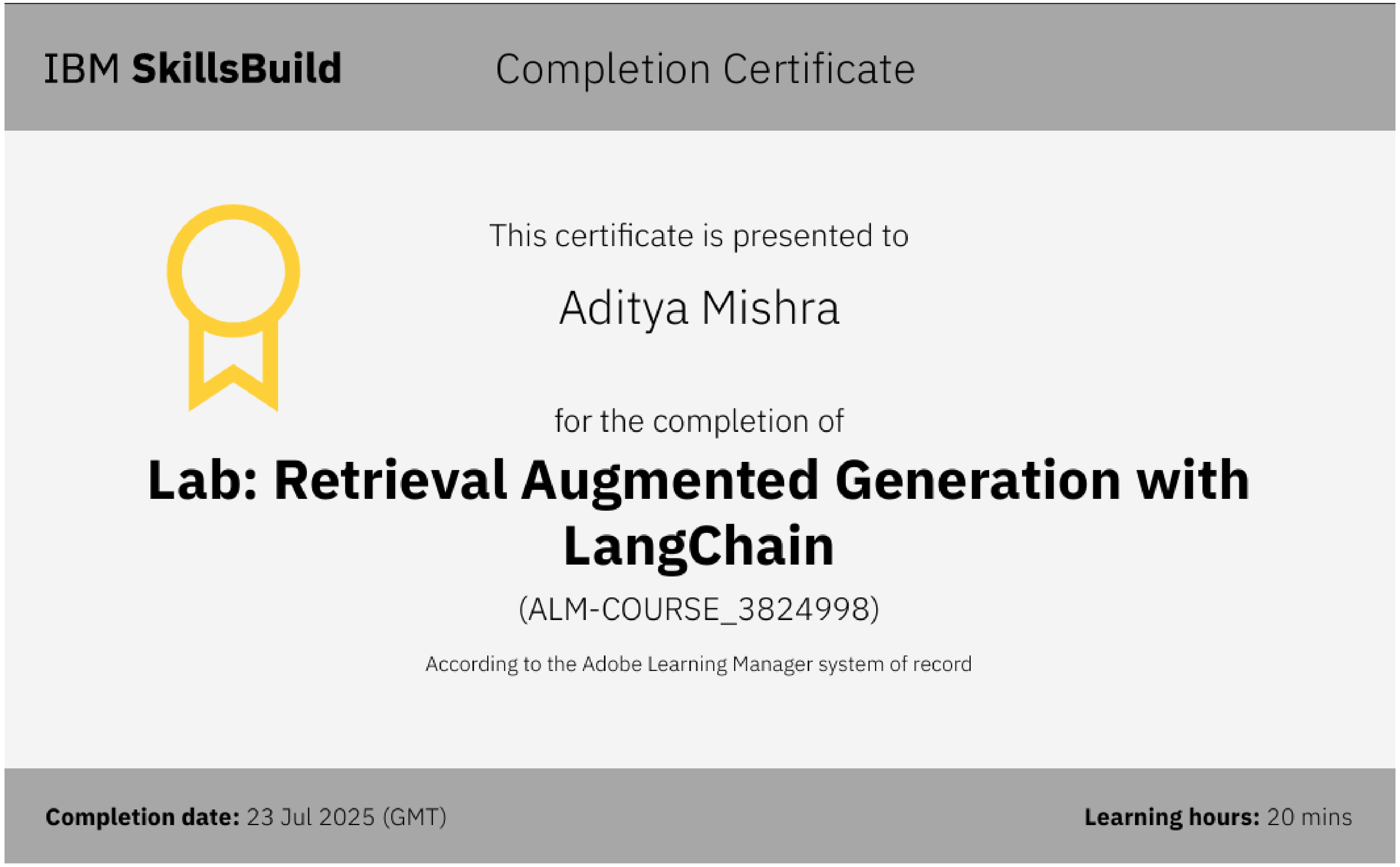
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/f580bc39-fbff-4e58-a501-5b29edc304f0

IBM

edunet foundation

IBM CERTIFICATIONS



IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Aditya Mishra

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

edunet foundation

THANK YOU

edunet foundation