CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:

1.Student Name- Sujoya Ghosh

2.College Name- Dr. B C Roy Engineering College, Durgapur

3.Department- Electronics and Communication Engineering (ECE)

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

Power distribution systems are vulnerable to various electrical faults due to Overheating, Line Breakage or Transformer faults . These faults can lead to equipment damage, power outages, and safety risks if not detected and addressed promptly. Traditional protection mechanisms rely on predefined rules and may lack adaptability to new or evolving fault patterns.

This project aims to design a machine learning-based model that can accurately detect and classify different power system faults. The ultimate goal is to enable faster and more intelligent fault identification to enhance grid reliability and operational safety.

PROPOSED SOLUTION

The proposed system uses machine learning to detect and classify power system faults based on electrical parameters. The workflow includes data preprocessing, feature extraction, and training classification models (e.g., Random Forest, SVM) to identify fault types such as overheating. The goal is to enable fast, accurate, and automated fault detection to enhance grid reliability and minimize downtime.

- **Data Collection: Use the Kaggle data set on power system faults.**
- **Data Preprocessing : Clean and normalize the data set.**
- **Model Training: Train a classification model (Decision tree, Random Forest, SVM) .**
- **Deployment:**
  - **Develop a user-friendly interface or application that provides real-time predictions and classifications for the type of power system fault.**
- **Evaluation:**
  - **Assess the model's performance using appropriate metrics such as accuracy, precision , recall and F1 score.**
  - **Fine-tune the model based on feedback and continuous monitoring of prediction accuracy.**

SYSTEM APPROACH

- **System Requirements**

  Laptop with 12gb RAM having i5 or Ryzen5 above processor

  Platform: IBM Cloud

  Tool Used: Watsonx.ai Studio

  Dataset Storage: IBM Cloud Object Storage

  Notebook runs entirely in the cloud using Watsonx.ai compute

- **Library required to build the model:**

  Pandas for reading and pre processing dataset (csv)

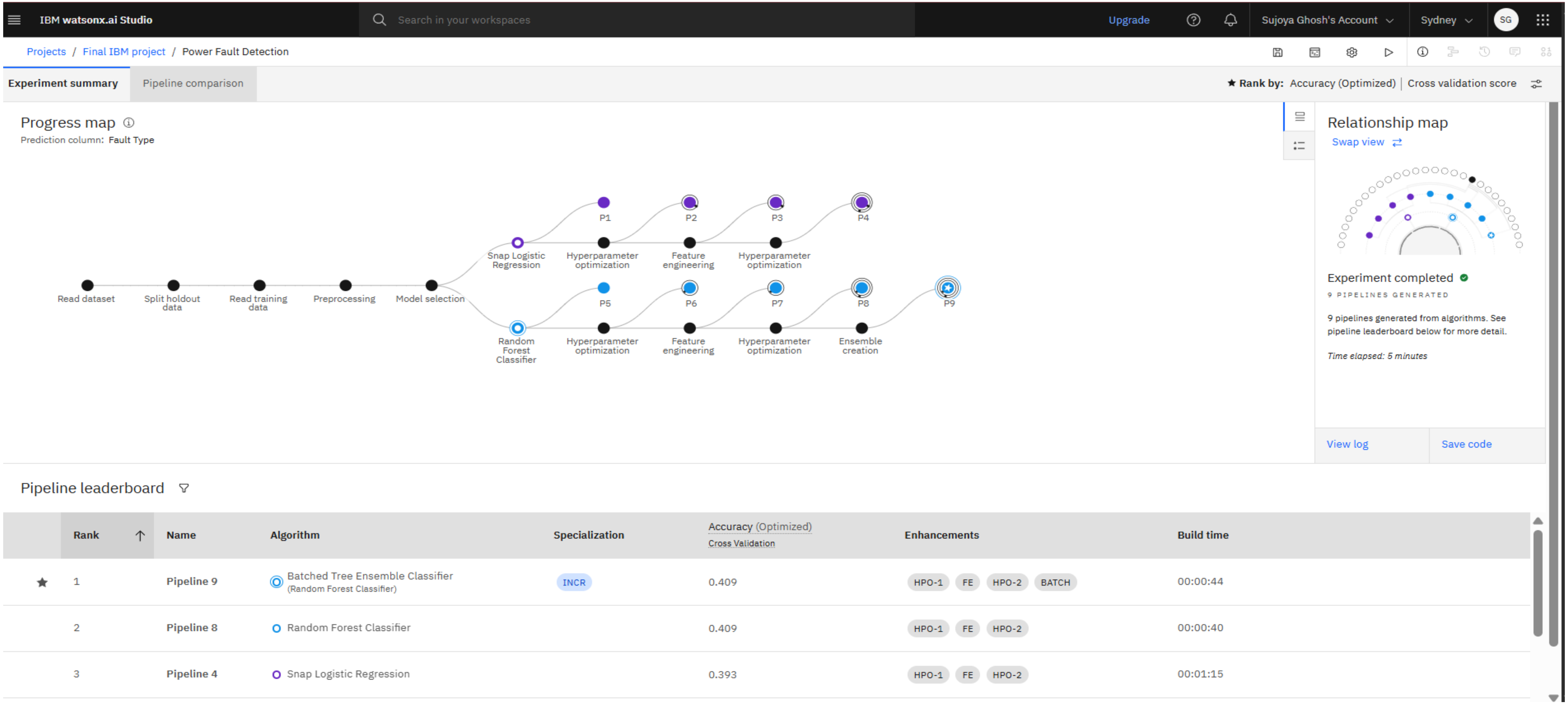
  Language: Python 3 (pre-installed in Watsonx.ai)

edunet
foundation

ALGORITHM & DEPLOYMENT

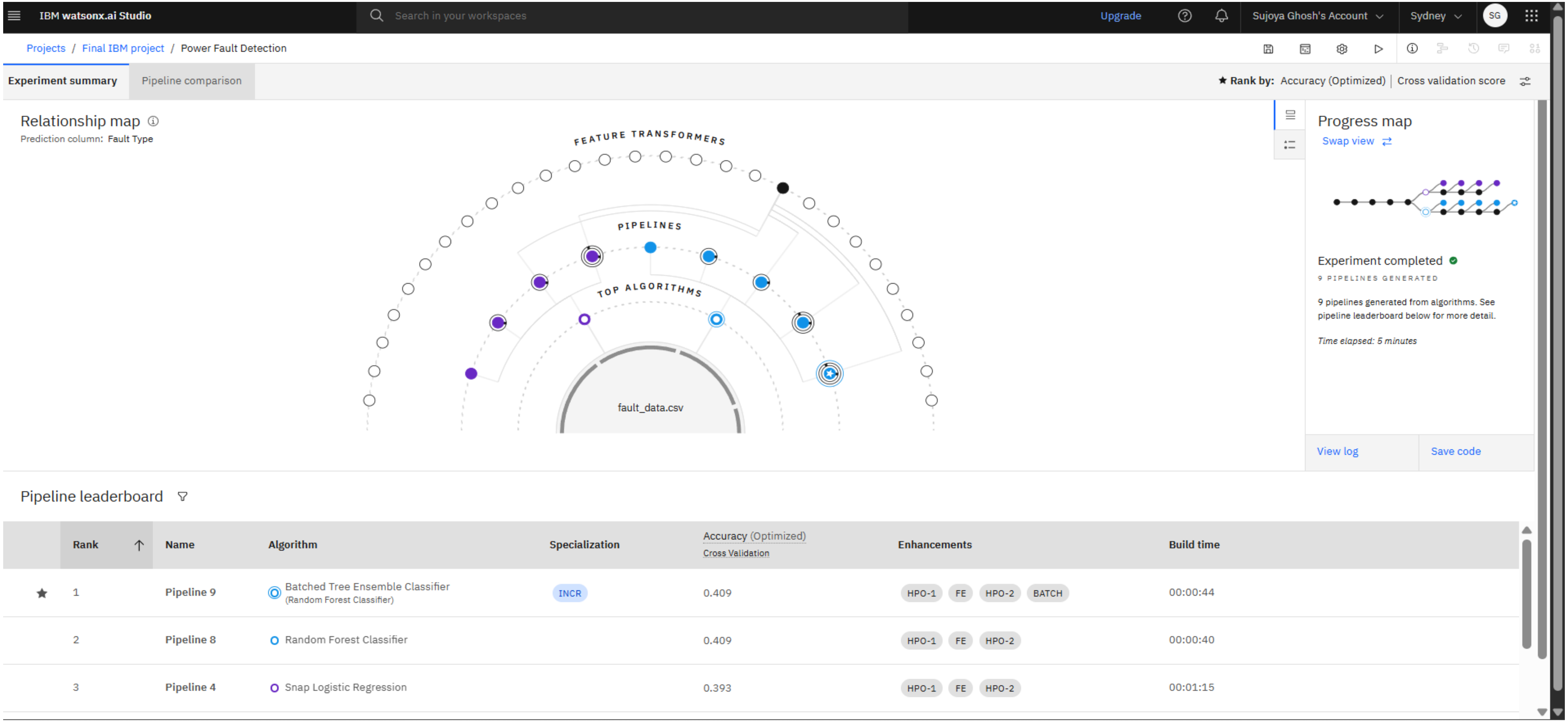
- Algorithm Selection:
  - Random Forest Classifier is being used (or SVM based on performance).
- Data Input:
  - Volate, Current, Power Load, Temperature, Wind Speed etc from the Kaggle data sheet.
- Training Process:
  - Supervised learning using labelled fault data types.
- Prediction Process:
  - Model is deployed on IBM Watson Studio with API endpoint for real time predictions.

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	00:00:44
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:40
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:01:15

RESULT



edunet foundation

RESULT

IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Sujoya Ghosh's Account | Sydney | SG

Deployment spaces / Power_DEP1 / P9 - Random Forest Classifier: Power Fault Detection /

Power_D2

Deployed Online

API reference | Test

Enter input data

Text | JSON

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 ×

	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	48	23	10	Clear	Completed	Normal	2.5	3
2	46	27	29	Thunderstorm	Completed	Overheated	2.6	4.3
3	51	30	14	Snowy	Pending	Faulty	4.2	3.9
4	52	21	30	Rainy	Pending	Faulty	2.9	1.1
5	46	24	16	Clear	Completed	Overheated	3.5	1.2
6	54	35	10	Clear	Scheduled	Faulty	2.5	2.7
7	56	32	12	Rainy	Pending	Normal	2.3	1.6
8	54	37	28	Rainy	Completed	Faulty	2.5	2.5
9								
10								

8 rows, 12 columns

Predict

RESULT

IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Sujoya Ghosh's Account | Sydney | SG

Deployment spaces / Power_DEP1 / P9 - Random Forest Classifier: Power Fault Detection /

Prediction results

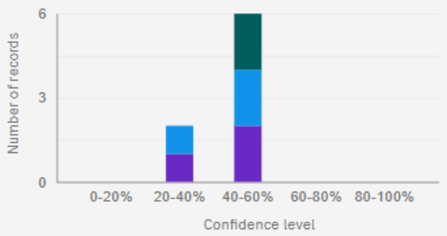
Close ×

Prediction type

Multiclass classification

Prediction percentage

Confidence level distribution



Display format for prediction results

◉ Table view ○ JSON view

Show input data

	Prediction	Confidence
1	Line Breakage	54%
2	Transformer Failure	38%
3	Transformer Failure	43%
4	Transformer Failure	43%
5	Overheating	45%
6	Line Breakage	46%
7	Line Breakage	39%
8	Overheating	58%
9		
10		
11		
12		
13		
14		
15		
16		

Download JSON file

edunet foundation

CONCLUSION

The developed machine learning-based fault detection system effectively identifies and classifies common power system faults such as overheating, transformer breakage, and line breakage. By leveraging voltage and current phasor data, the model enhances the speed and accuracy of fault diagnosis compared to traditional protection mechanisms. While the system shows promising performance in controlled conditions, real-time deployment and broader dataset integration remain areas for improvement. This project demonstrates the potential of AI in creating intelligent, data-driven protection systems, paving the way for smarter and more resilient power grids.

FUTURE SCOPE

The Power System Fault Detection can be significantly enhanced to handle real-world challenges:

- **Real-Time Grid Integration:**
  Deploy the model within SCADA systems for live fault monitoring.
- **Edge Deployment:**
  Implement lightweight ML models on embedded devices at substations.
- **Advanced Learning Models:**
  Use LSTM/CNN to improve detection of complex and evolving fault patterns.
- **Fault Localization:**
  Extend functionality to identify the exact location of faults.
- **Hybrid Protection Systems:**
  Combine rule-based and AI-based methods for enhanced grid reliability.

REFERENCES

- IEEE Xplore – Articles on machine learning for power system fault detection

- Kaggle – Power system fault detection dataset link – https://www.kaggle.com/datasets/ziya07/power-system faults-dataset

- NSL-KDD and other synthetic datasets for classification tasks

- ResearchGate – Papers on voltage/current-based fault classification

IBM CERTIFICATIONS
(getting started with AI)

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

IBM

Sujoya Ghosh

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/b36a008e-a5e7-4ed4-9f8e-89693185ce72

IBM

edunet
foundation

IBM CERTIFICATIONS

Journey to Cloud

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild

In recognition of the commitment to achieve professional excellence

Sujoya Ghosh

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 19, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/561ca082-6513-4119-b383-e825029af702

IBM

edunet foundation

IBM CERTIFICATIONS (RAG Lab)

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Sujoya Ghosh

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

edunet foundation

THANK YOU