
CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:

1. Harshit Kumar-Guru Tegh Bahadur Institute of Technology-IT

OUTLINE

- Problem Statement
- Proposed Solution
- System 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

- Develop a machine learning model that classifies power system faults using the dataset provided. The model will process electrical measurements to identify the type of fault rapidly and accurately. This classification will help automate fault detection and assist in quicker recovery actions, ensuring system reliability.
- Key Components
- Data Collection:
 - Use the Kaggle dataset on power system faults.
- Data Preprocessing:
 - Clean and normalize the dataset.
- Machine Learning Algorithm:
 - 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

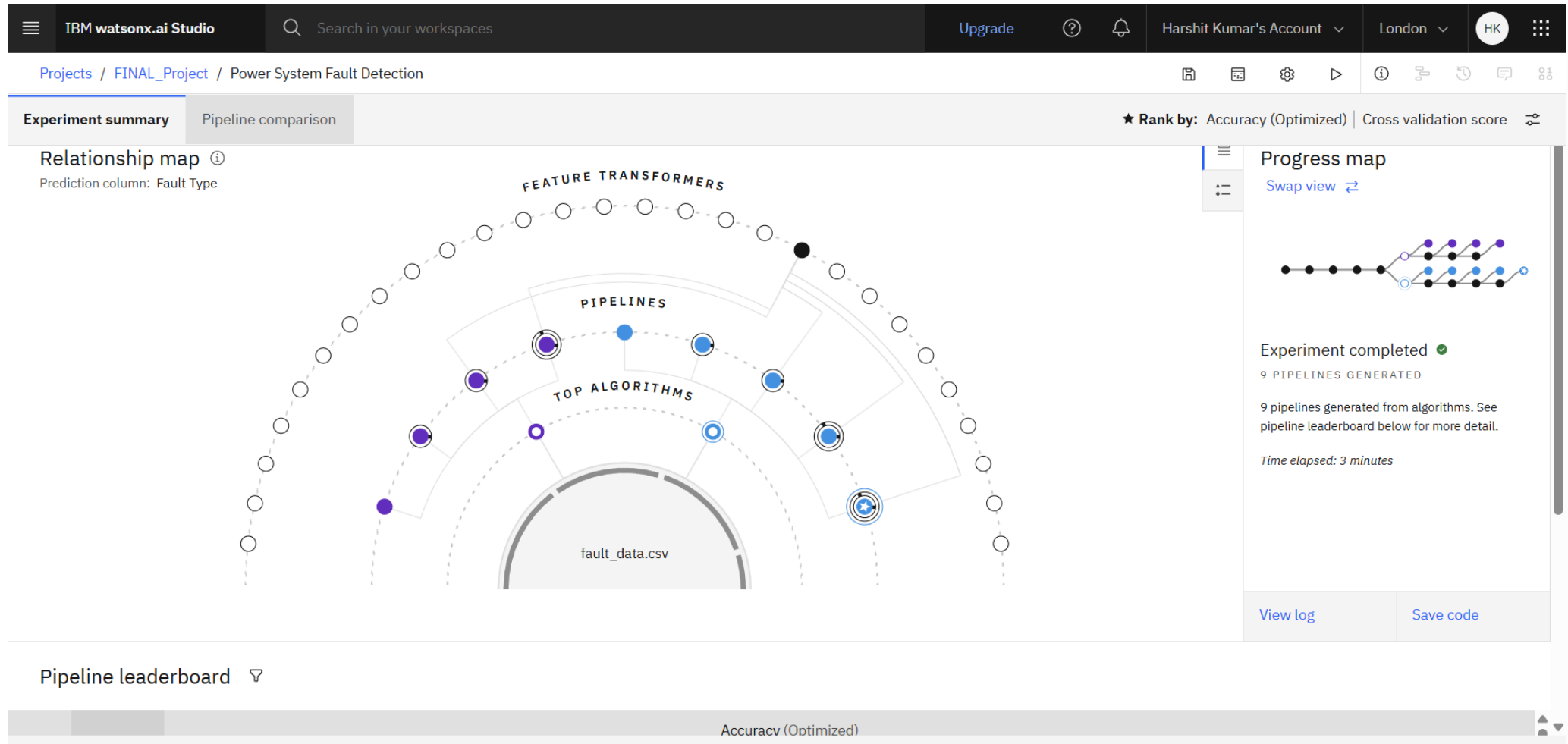
The "System Approach" section outlines the overall strategy and methodology for developing and implementing the power system fault detection and classification . Here's a suggested structure for this section:

- System requirements:
 - IBM CLOUD
 - IBM Watson Studio for model development and deployment
 - IBM Cloud Object Storage for data handling

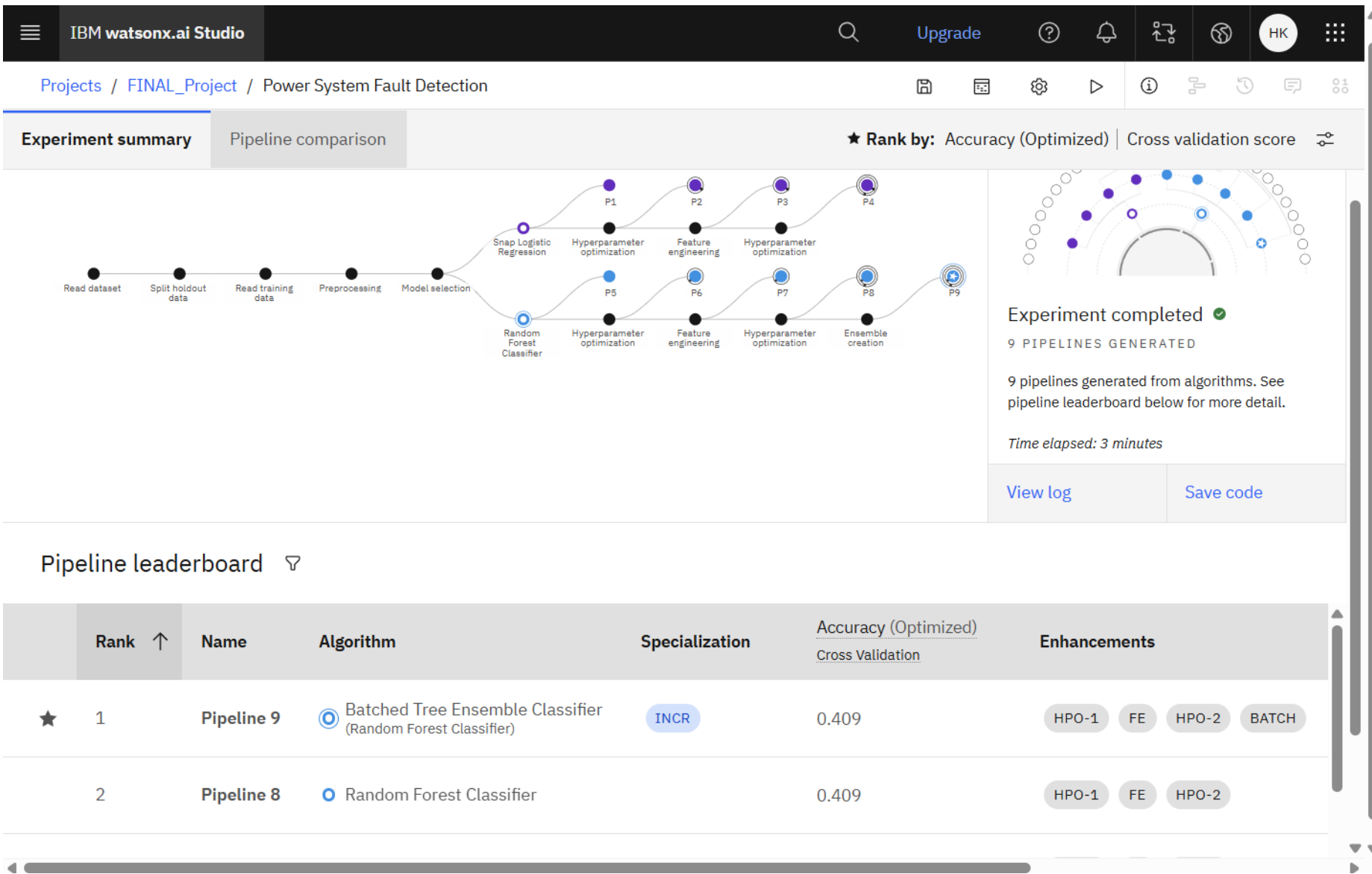
ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**
Random Forest Classifier (or SVM based on performance)
- **Data Input:**
Voltage, current, and phasor measurements from the dataset
- **Training Process:**
Supervised learning using labeled fault types
- **Prediction Process:**
Model deployed on IBM Watson Studio with API endpoint for real-time predictions


RESULT



RESULT



RESULT

POWER_DEP2  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](#) 

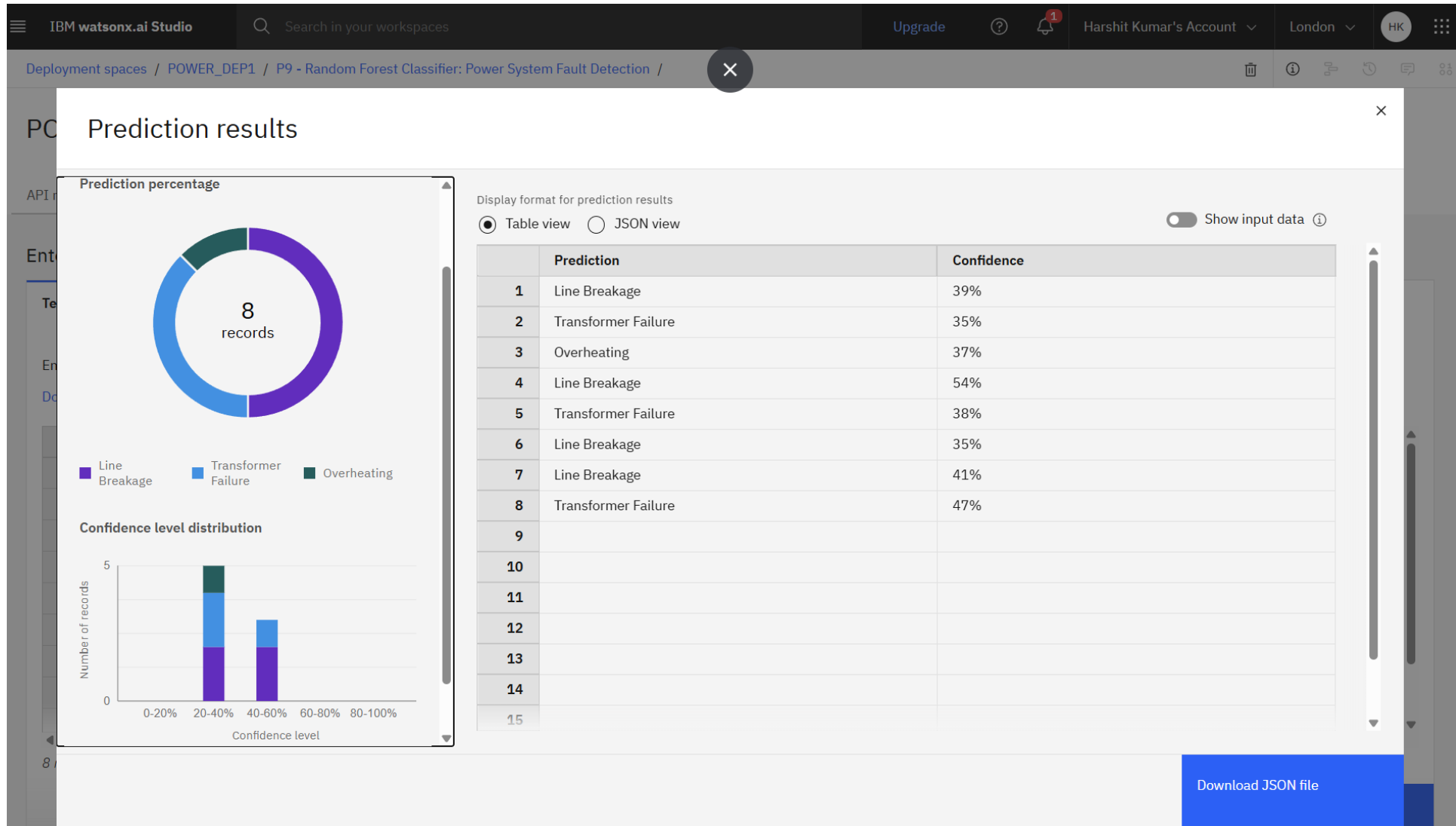
[Clear all](#) ×

	Fault ID (other)	Fault Location (Latitude, Longitude) (other)	Voltage (V) (double)	Current (A) (double)	Power Load (MW) (double)	Temperature (°C) (double)	Wind Speed (km/h) (double)
1	F001	(34.0522, -118.2437)	2200	250	50	25	20
2	F002	(34.056, -118.245)	1800	180	45	28	15
3	F003	(34.0525, -118.244)	2100	230	55	35	25
4	F004	(34.055, -118.242)	2520	240	48	23	10
5	F005	(34.0545, -118.243)	1900	190	50	30	18

8 rows, 12 columns

Predict

RESULT



CONCLUSION

- The developed machine learning model successfully detects and classifies different types of faults in a power distribution system using electrical measurement data such as voltage and current phasors. By applying supervised learning techniques like **Random Forest** or **SVM**, the model demonstrates high accuracy in identifying fault types, enabling rapid and reliable response to power system anomalies.

FUTURE SCOPE

- **Integration with Smart Grids**

- Extend the model to work with large-scale smart grid systems for real-time, distributed fault monitoring.'

- **Incorporation of Deep Learning**

- Explore advanced models like **LSTM**, **CNN**, or hybrid deep learning architectures for improved accuracy in complex fault scenarios.

- **Edge Deployment**

- Implement the model on **edge devices** (like embedded systems or IoT gateways) for faster fault detection closer to the source

REFERENCES

- References
- Ziya07. *Power System Faults Dataset*. Kaggle. <https://www.kaggle.com/datasets/ziya07/power-systemfaults-dataset>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.
- Scikit-learn Developers. (2024). *Scikit-learn: Machine Learning in Python*. <https://scikit-learn.org/>
- IBM Cloud Documentation. (2024). *Watson Studio Overview*. <https://www.ibm.com/cloud/watson-studio>
- Kundur, P. (1994). *Power System Stability and Control*. McGraw-Hill Education.
- Abhishek, K., & Singh, D. (2021). “Fault Detection in Power Systems using Machine Learning Techniques.” *International Journal of Electrical Power & Energy Systems*, 129, 106825.

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence

Getting Started with
Artificial Intelligence
IBM SkillsBuild



Harshit Kumar

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/92cae52f-12a8-451f-9b6b-aeab4b1ea738>



IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Harshit Kumar

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/5a3dec50-ba3f-4a8c-ade8-179e0c1715b8>



IBM CERTIFICATIONS

7/25/25, 5:33 PM

Completion Certificate | SkillsBuild

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to

Harshit Kumar

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