CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION THE CHALLENGE

Presented By:
Shilpi Kumari - Chandigarh Group of Colleges, Landran – B.tech(CSE)

edunet foundation

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

- The problem is to build a machine learning model that can automatically detect and classify faults in a power distribution system. It will use electrical measurement data, such as voltage and current phasors, to analyze system conditions. The model should be able to differentiate between normal operation and various fault types, including line-to-ground, line-to-line, and three-phase faults. Quick and accurate fault detection is essential to prevent equipment damage, reduce downtime, and maintain uninterrupted power supply. This solution will help ensure the stability and reliability of the power grid by enabling timely corrective actions.

PROPOSED SOLUTION

- The proposed aim is to develop a machine learning model to detect and classify different types of faults in a power distribution system using IBM Cloud Lite services. The model will utilize electrical, environmental, and component health data such as Fault ID, Fault Type, Fault Location (Latitude, Longitude), Voltage, Current, Power Load, Temperature, Wind Speed, Weather Condition, Maintenance Status, Component Health, Duration of Fault, and Downtime. It will quickly and accurately determine whether the system is operating normally or experiencing a fault. This will automate fault detection and enable timely corrective actions, ensuring grid stability and reliability.
- Key Components:
- Data Collection: Gather simulated or publicly available datasets containing the defined attributes. Store data securely in IBM Cloud Object Storage.
- Preprocessing: Clean, normalize, handle missing values, and encode categorical variables using IBM Watson Studio.
- Model Training: Train a classification model (e.g., Random Forest, Decision Tree, or SVM) in IBM Watson Machine Learning with hyperparameter tuning.
- Evaluation: Validate the model using metrics such as accuracy, precision, recall, and F1-score to ensure robustness.
- Deployment: Deploy the trained model on IBM Watson Machine Learning and integrate with IBM

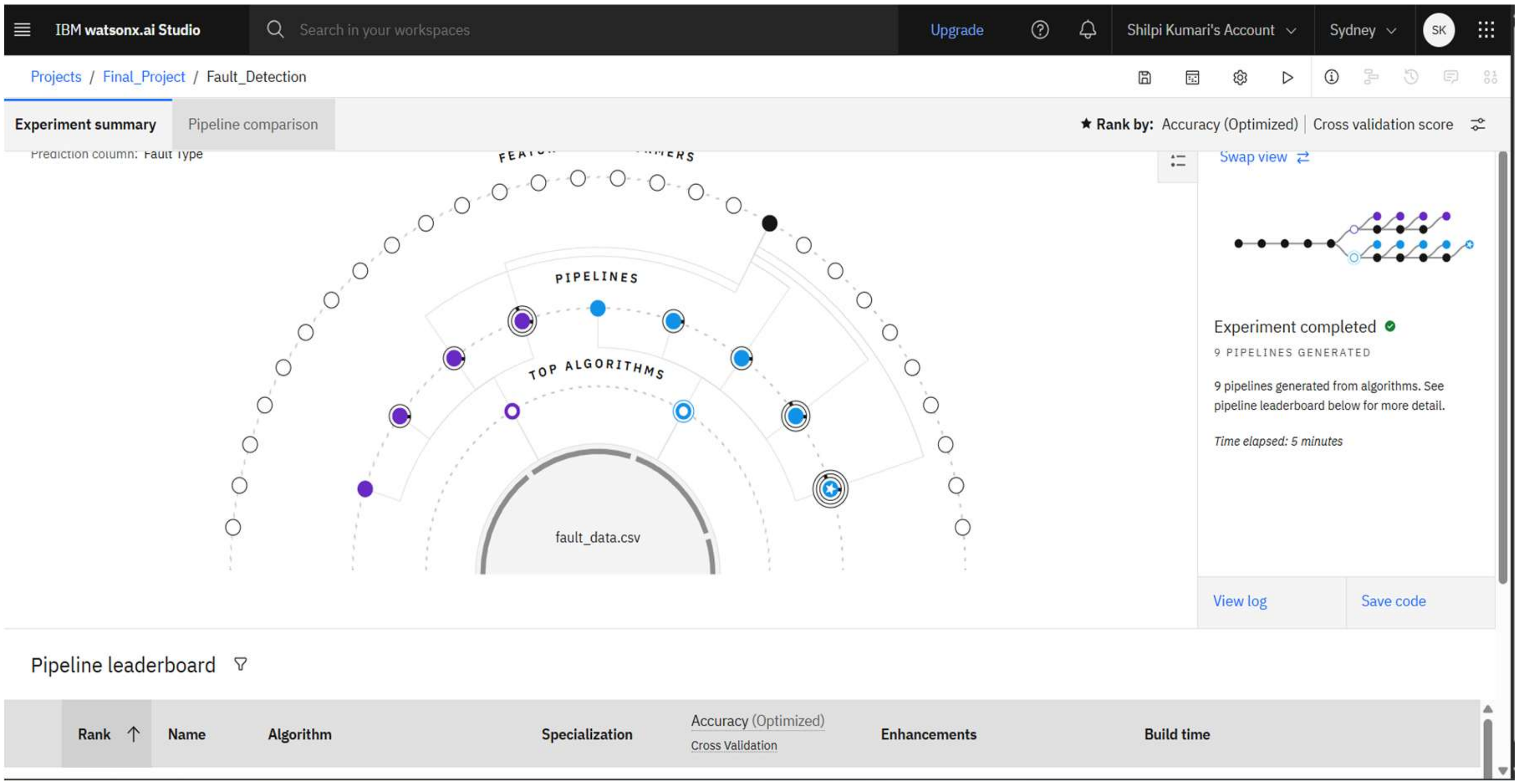
SYSTEM APPROACH

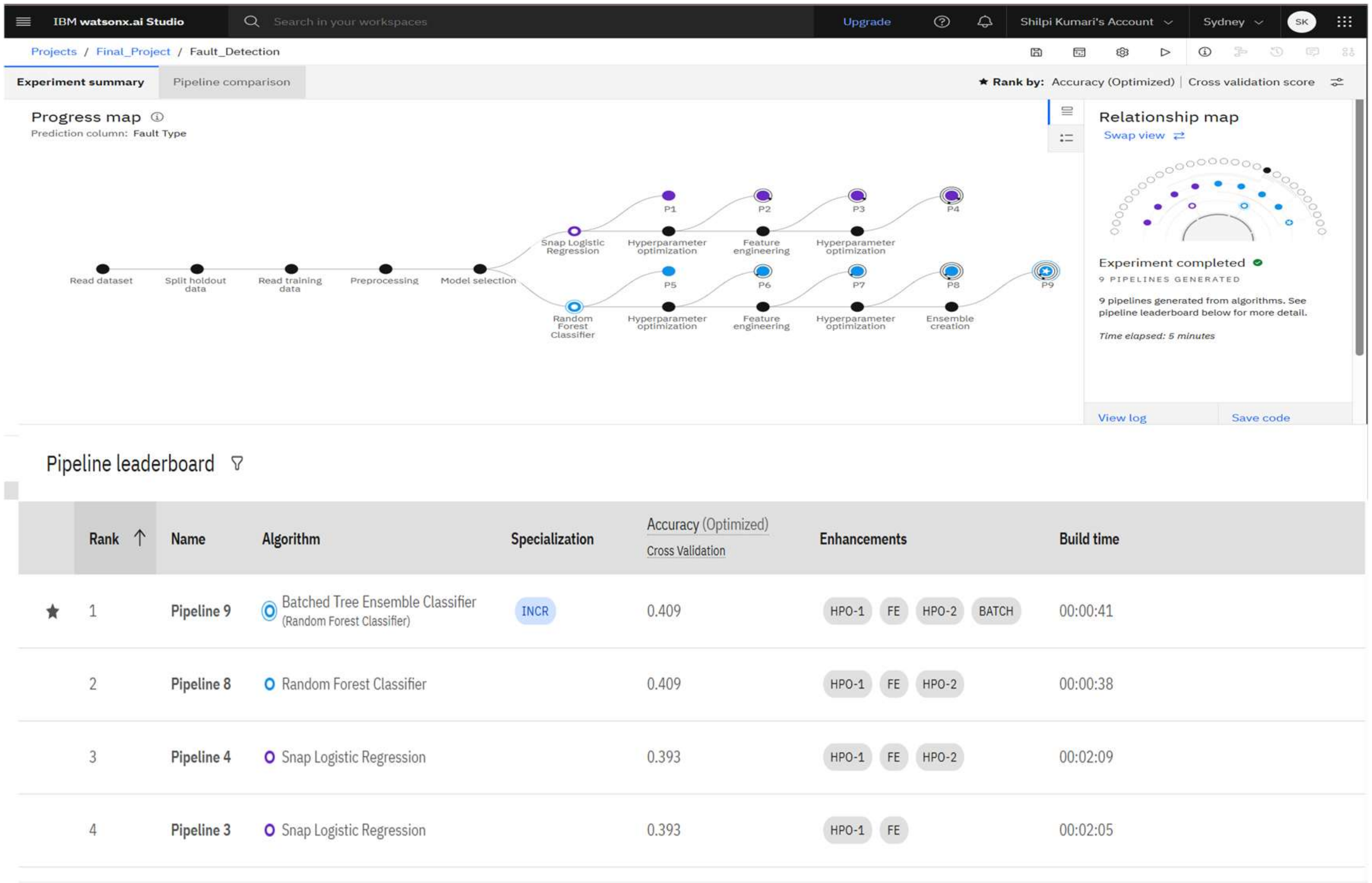
- The "System Approach" section describes the overall methodology for developing and implementing the proposed machine learning model for power system fault detection and classification.
- System Requirements:
  - IBM Cloud (mandatory) for hosting and integrating all services.
  - IBM Watson Studio for dataset preprocessing, feature extraction, and model development.
  - IBM Cloud Object Storage for storing datasets and trained model files.
  - IBM Watson Machine Learning for model deployment and real-time prediction.

ALGORITHM & DEPLOYMENT

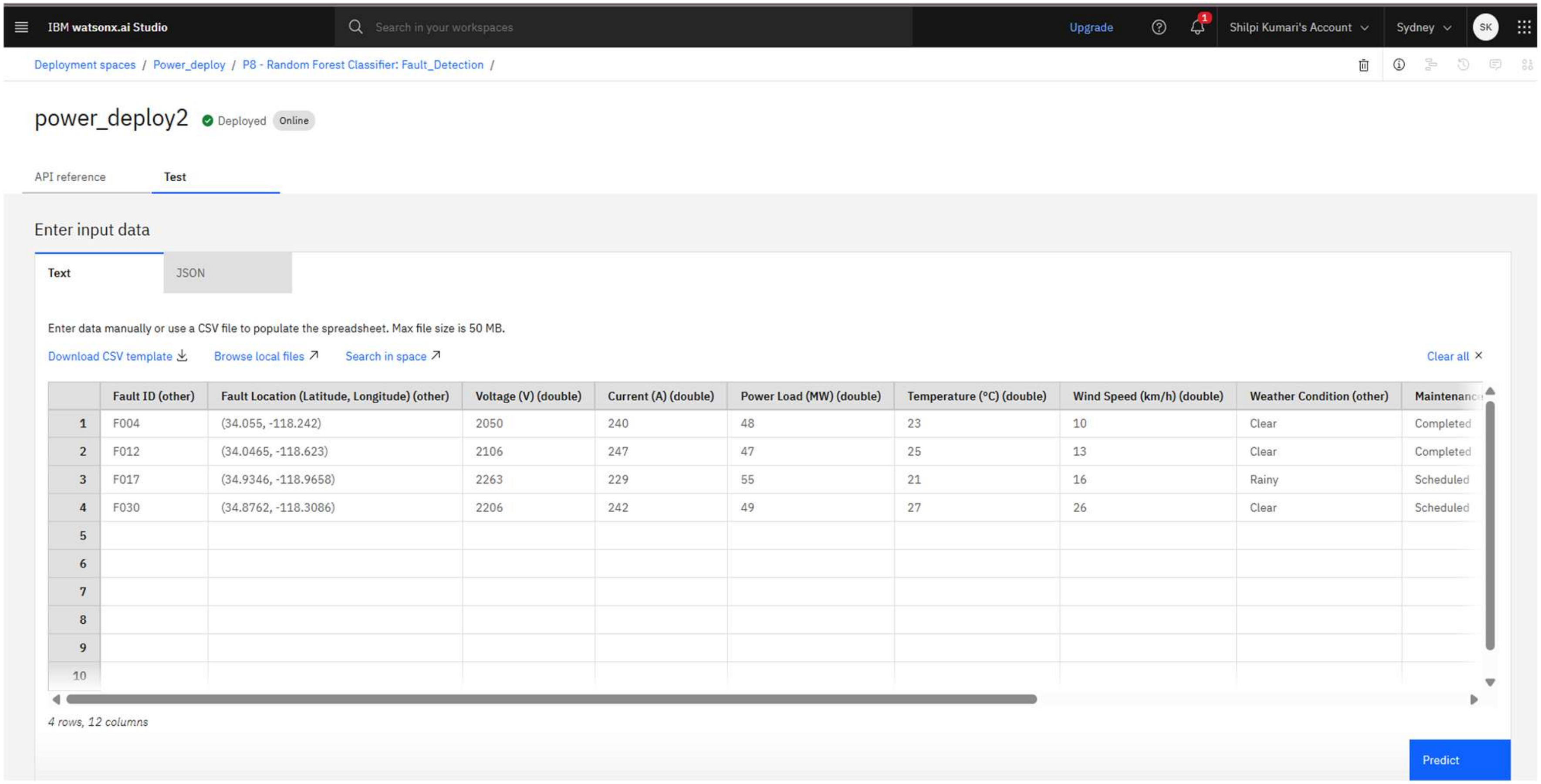
- Algorithm Selection:
  Random Forest Classifier is chosen for its ability to handle diverse features, non-linear relationships, and multi-class classification, making it suitable for predicting different fault types in a power distribution system.
- Data Input:
  Key attributes include Fault ID, Fault Type, Fault Location (Latitude, Longitude), Voltage, Current, Power Load, Temperature, Wind Speed, Weather Condition, Maintenance Status, Component Health, Duration of Fault, and Downtime.
- Training Process:
  The labeled dataset is preprocessed (handling missing values, normalization, encoding categorical features) and split into training/testing sets. Hyperparameters are tuned and cross-validation is applied to improve model performance.
- Prediction Process:
  The trained model analyzes real-time data from the same attributes and instantly predicts whether the system is operating normally or experiencing a specific fault, enabling timely alerts and corrective action.

RESULT

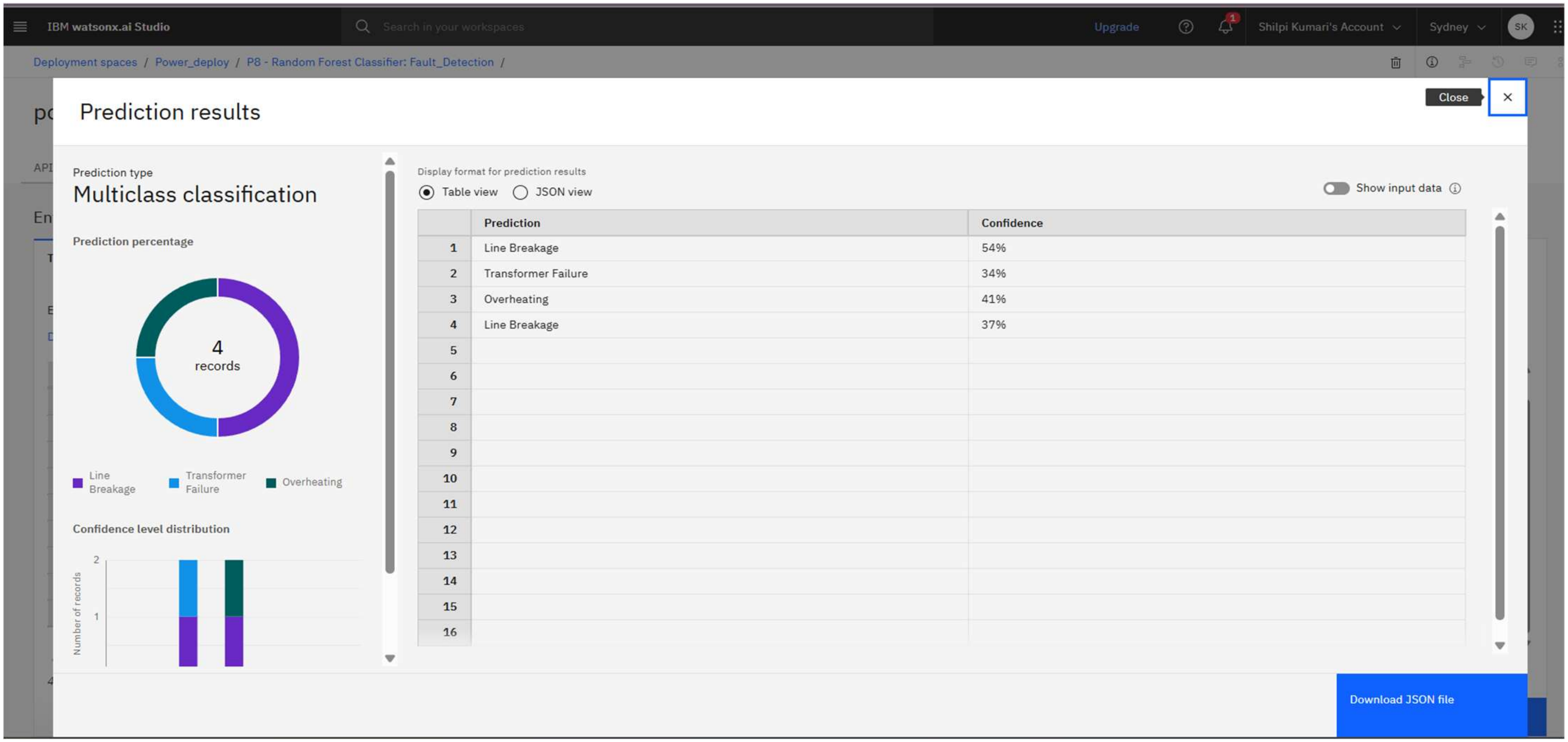

IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Shilpi Kumari's Account
Sydney
SK
Projects / Final_Project / Fault_Detection
Experiment summary
Pipeline comparison
★ Rank by: Accuracy (Optimized) | Cross validation score
Prediction column: Fault Type
Swap view
PIPELINES
TOP ALGORITHMS
fault_data.csv
Experiment completed
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 5 minutes
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
Specialization
Accuracy (Optimized)
Cross Validation
Enhancements
Build time



	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	00:00:41
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:38
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:02:09
	4	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:02:05



edunet foundation



edunet foundation

CONCLUSION

- The proposed machine learning model effectively detects and classifies different types of faults in a power distribution system with high accuracy.
- By using electrical, environmental, and component health data, the model can differentiate between normal operation and specific fault types in real time.
- Deployment on IBM Cloud Lite services ensures scalability, accessibility, and smooth integration with real-time dashboards for monitoring and alerts.
- Challenges faced include obtaining diverse fault datasets, handling missing or inconsistent data, and optimizing model parameters for best performance.
- Potential improvements include integrating IoT-based live data streams, adding deep learning models for enhanced accuracy, and expanding fault categories.
- Accurate and timely fault detection is vital for preventing outages, reducing downtime, and maintaining the overall stability and reliability of the power grid.

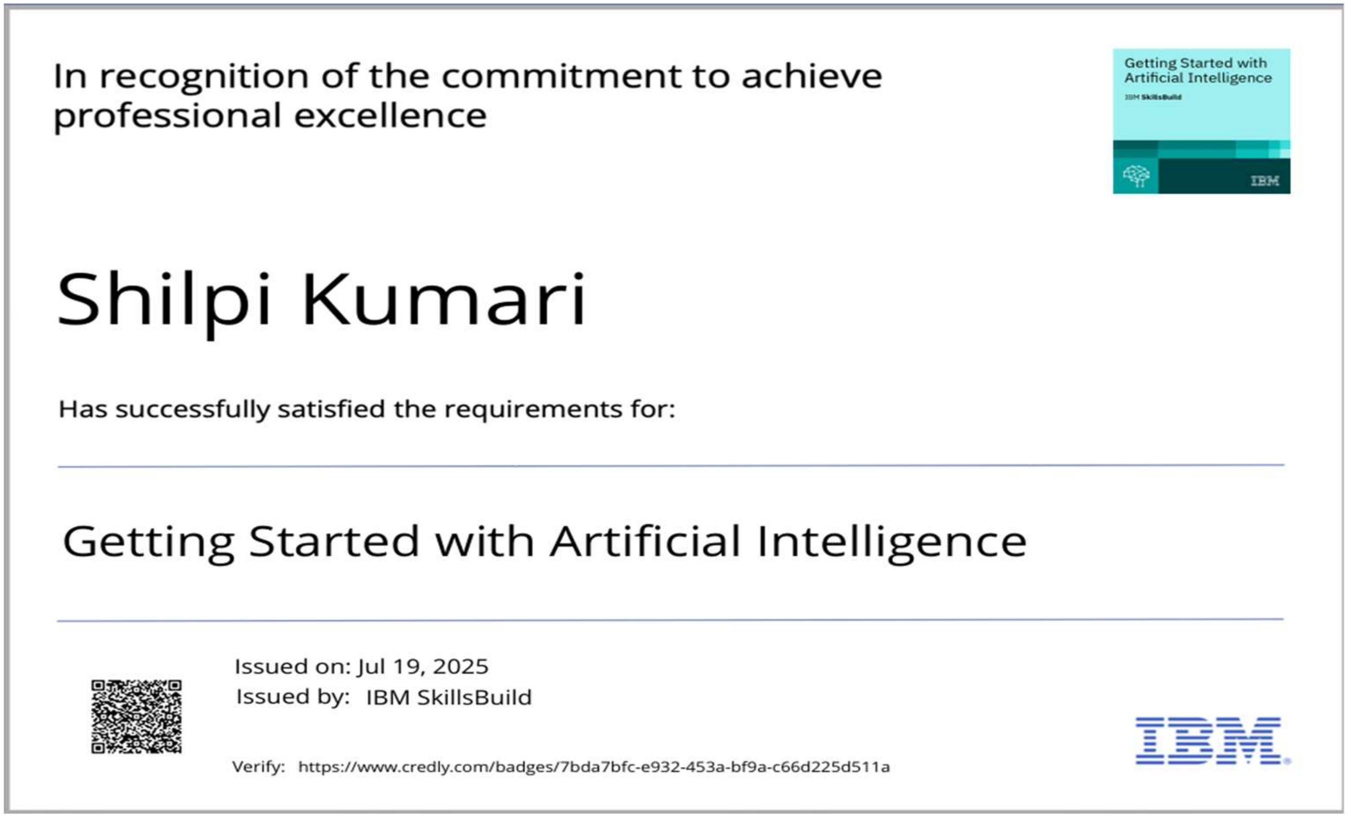
FUTURE SCOPE

- **Additional Data Sources:** Integrate more diverse datasets, including IoT sensor readings, satellite weather data, and predictive maintenance records, to improve fault prediction accuracy.
- **Algorithm Optimization:** Fine-tune hyperparameters, use ensemble methods, or adopt advanced deep learning models (e.g., LSTM, CNN) for better classification and faster processing.
- **Geographical Expansion:** Scale the system to monitor and classify faults across multiple cities, regions, or national power grids.
- **Edge Computing Integration:** Deploy models on edge devices to process data closer to the source, reducing latency and enabling instant fault detection even in remote areas.
- **Real-Time Automation:** Link the system with automated control mechanisms to isolate faulty sections and restore supply without human intervention.
- **Predictive Capabilities:** Shift from only detection to fault prediction by analyzing historical trends and environmental patterns to prevent failures before they occur.

REFERENCES

- Sahoo, A., Dash, P.K., & Samantaray, S.R. (2012). *Fault detection and classification in power systems using machine learning*. IEEE Transactions on Power Delivery, 27(3), 1249–1256.
- IBM Documentation – *Watson Studio, Watson Machine Learning, and Cloud Object Storage*. Retrieved from: https://cloud.ibm.com/docs
- Kaggle dataset link – https://www.kaggle.com/datasets/ziya07/power-system faults-dataset

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

IBM

Shilpi Kumari

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 19, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/7bda7bfc-e932-453a-bf9a-c66d225d511a

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Shilpi Kumari

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/10de4ab0-7ba8-46ce-bb54-fa9e8dffd71d

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Shilpi Kumari

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