CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:

1. NAVANEETHA BASHETTI - ST. PETERS ENGINEERING COLLEGE- CIVIL

edunet foundation

OUTLINE

- **Problem Statement**
- **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

- Develop a machine learning model to detect and classify different types of faults in a power distribution system using electrical measurement data such as voltage and current phasors. The model will identify whether the system is operating normally or if a fault (line-to-ground, line-to-line, or three-phase) has occurred. This automated classification will enable rapid fault detection, support faster decision-making, and enhance the overall reliability of the power grid.
- **Data Collection:** Use a publicly available dataset (e.g., from Kaggle) containing labeled fault types and corresponding electrical measurements.
- **Data Pre-processing:** Clean the data by handling missing values, normalizing features, and preparing it for model training.
- **Model Training:** Build and train classification models such as Decision Tree, Random Forest, or Support Vector Machine (SVM) to learn patterns from the input data.
- **Evaluation:** Assess model performance using metrics like accuracy, precision, recall, and F1-score to ensure reliability and robustness.

SYSTEM APPROACH

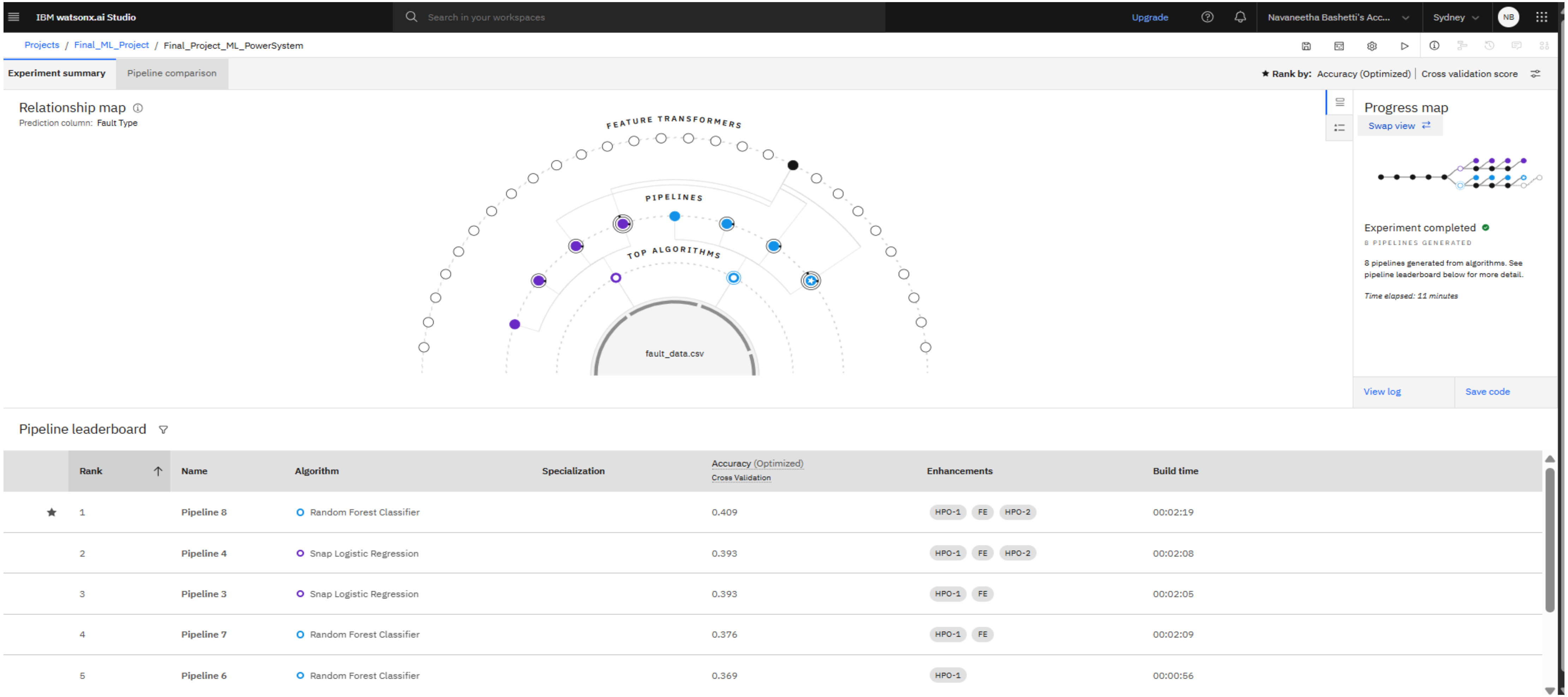
The "System Approach" section outlines the overall strategy and methodology for developing and implementing the rental bike prediction system. Here's a suggested structure for this section:

- **System requirements:**
  - IBM Cloud (Mandatory)
  - IBM Watsonx.ai Studio for model development and deployment
  - IBM Cloud Storage for dataset handling

ALGORITHM & DEPLOYMENT

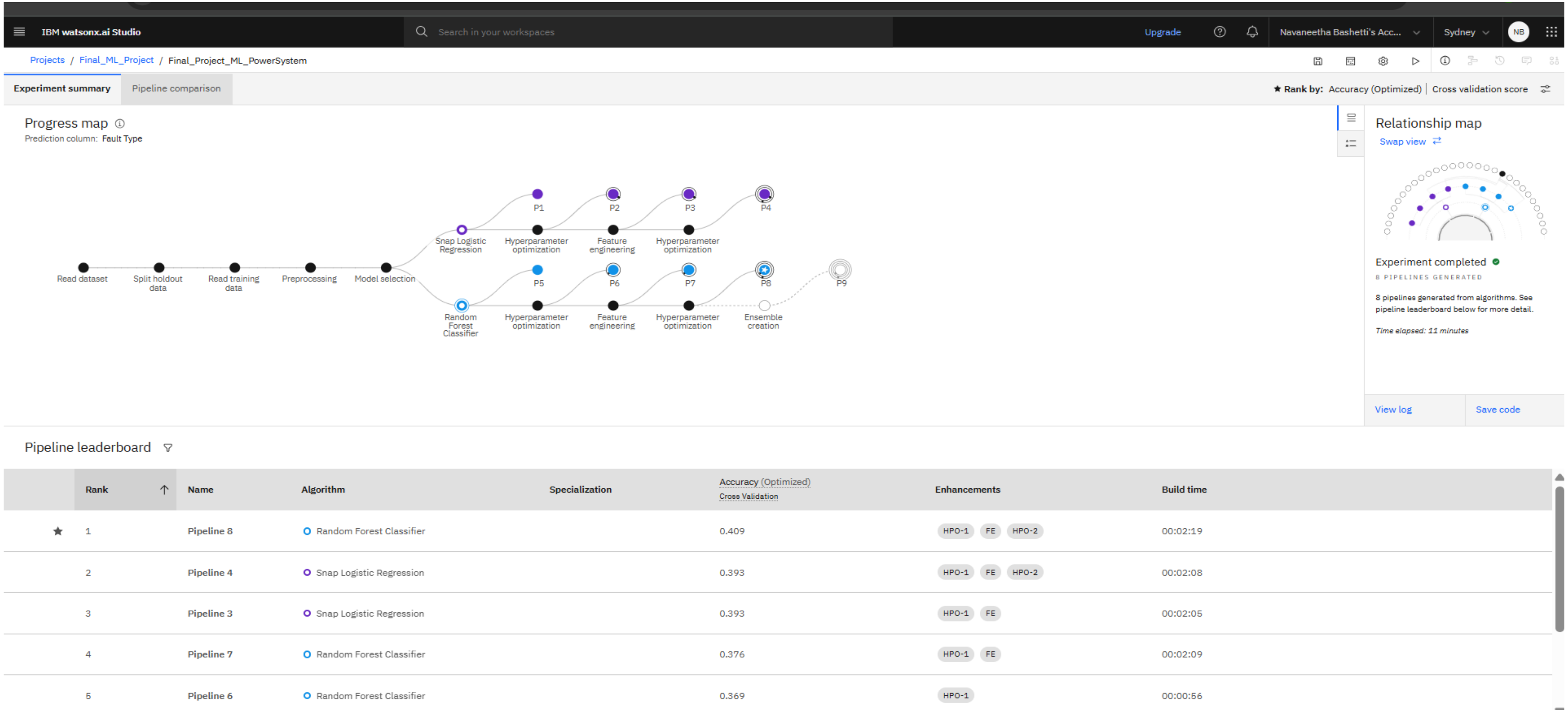
- **Algorithm Selection:**
  - Random Forest Classifier (or Support Vector Machine, based on accuracy and performance)
- **Data Input:**
  - The model uses input features like voltage and current phasors from the dataset, with the fault type as the target label.
- **Training Process:**
  - Supervised learning using labeled fault types for model training
- **Prediction Process:**
  - Model deployed using IBM Watson Studio with an API endpoint for real-time fault classification

RESULT



	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:02:19
	2	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:02:08
	3	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:02:05
	4	Pipeline 7	Random Forest Classifier		0.376	HPO-1 FE	00:02:09
	5	Pipeline 6	Random Forest Classifier		0.369	HPO-1	00:00:56

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Navaneetha Bashetti's Acc...
Sydney
NB
Projects / Final_ML_Project / Final_Project_ML_PowerSystem
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Progress map
Prediction column: Fault Type
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
Relationship map
Swap view
Experiment completed
8 PIPELINES GENERATED
8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 11 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 | 00:02:19
2 | Pipeline 4 | Snap Logistic Regression | | 0.393 | HPO-1 FE HPO-2 | 00:02:08
3 | Pipeline 3 | Snap Logistic Regression | | 0.393 | HPO-1 FE | 00:02:05
4 | Pipeline 7 | Random Forest Classifier | | 0.376 | HPO-1 FE | 00:02:09
5 | Pipeline 6 | Random Forest Classifier | | 0.369 | HPO-1 | 00:00:56

RESULT

IBM watsonx.ai Studio

Search in your workspaces

Upgrade

Navaneetha Bashetti's Acc...

Sydney

NB

Deployment spaces / PowerSystem_DLP1 / P8 - Random Forest Classifier: Final_Project_ML_PowerSystem /

PowerSystem_DLP2

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.

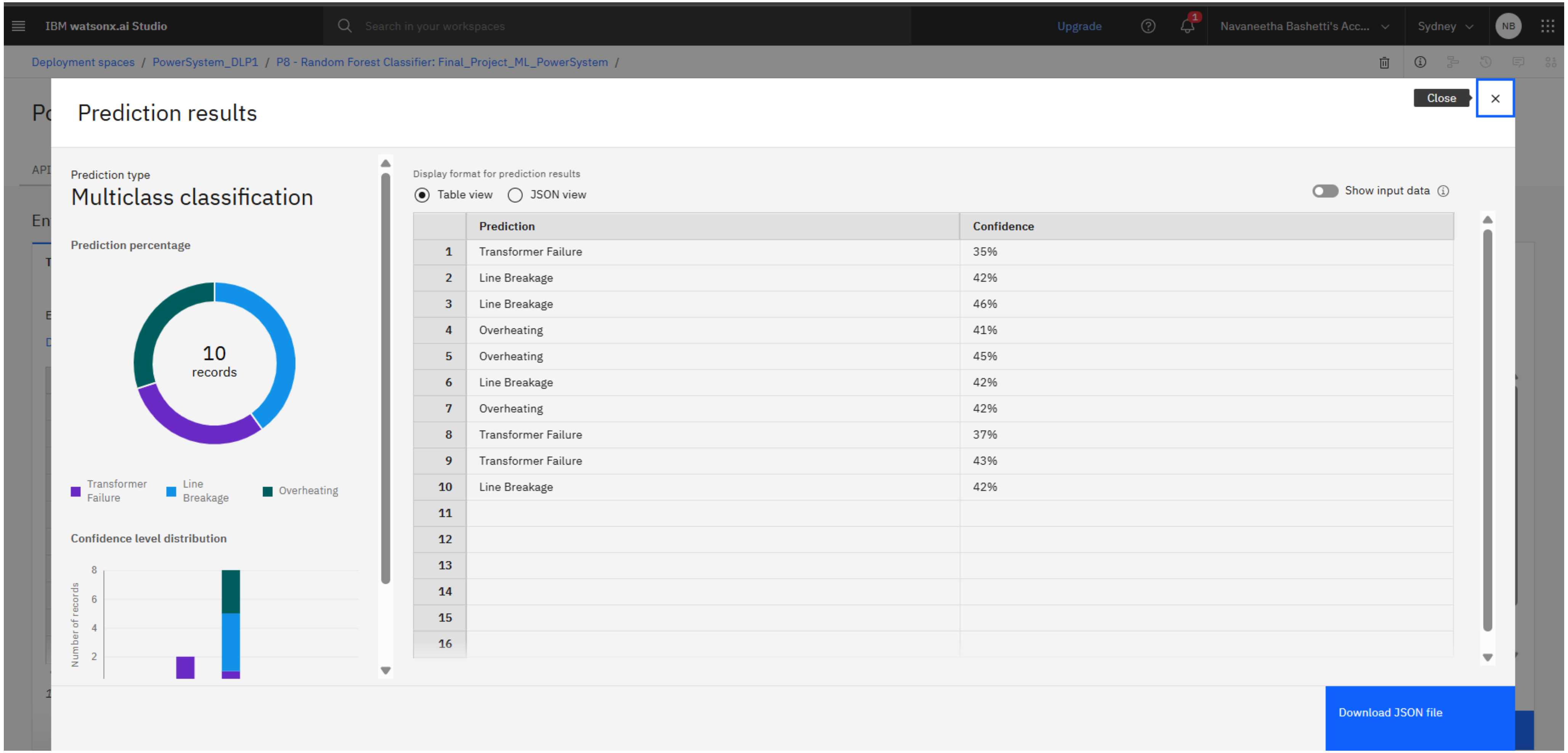
Download CSV template | Browse local files | Search in space | Clear all

	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)	Weather Condition (other)	Maintenanc
1	F002	(34.056, -118.245)	1800	180	45	28	15	Rainy	Completed
2	F165	(34.6646, -118.5539)	2079	227	53	28	25	Clear	Pending
3	F013	(34.9687, -118.5356)	2012	248	52	24	29	Clear	Completed
4	F017	(34.9346, -118.9658)	2263	229	55	21	16	Rainy	Scheduled
5	F033	(34.4338, -118.7342)	2199	242	53	39	28	Windstorm	Pending
6	F038	(34.991, -118.7368)	2206	200	48	37	27	Snowy	Pending
7	F056	(34.0609, -118.4429)	2149	227	46	25	10	Thunderstorm	Scheduled
8	F063	(34.4263, -118.5649)	1842	242	45	35	13	Snowy	Pending
9	F124	(34.6259, -118.4447)	2104	236	50	23	18	Windstorm	Completed
10	F240	(34.9651, -118.8866)	2031	223	46	36	27	Windstorm	Pending

10 rows, 12 columns

Predict

RESULT



edunet foundation

CONCLUSION

In this project, I developed a machine learning model to predict fault types in a power distribution system using voltage and current phasor data from the provided Excel dataset. By training a supervised learning model—such as a Random Forest classifier—on labeled fault types, the system accurately distinguishes between normal and various fault conditions (e.g., line-to-ground, line-to-line, and three-phase faults). This model helps automate fault detection, supports real-time monitoring, and enhances the overall stability and reliability of the power grid.

FUTURE SCOPE

- The model can be integrated with real-time power grid systems for automatic and instant fault detection.
- Advanced deep learning techniques can be explored to further improve fault classification accuracy.
- The system can be scaled to handle more complex faults and support smart grid automation.

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Navaneetha Bashetti

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/9980c4f9-3fa1-4329-b28a-9cd06a6d8753

IBM

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Navaneetha Bashetti

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/2ab2e0fb-9bf8-4814-9ef3-df934138fcf3

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Navaneetha Bashetti

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