
CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:
1. Swarnika-Maharshi Dayanand University-B.TECH CSE

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

Power distribution systems are prone to various types of faults such as line-to-ground, line-to-line, and three-phase faults. These faults can disrupt power supply and reduce system reliability. The challenge lies in accurately detecting and classifying these faults using electrical measurement data (voltage, current, phasors) to differentiate them from normal operating conditions, thereby ensuring the stability of the power grid.

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.
- Preprocessing: Clean and normalize the dataset.
- Model Training: 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:-

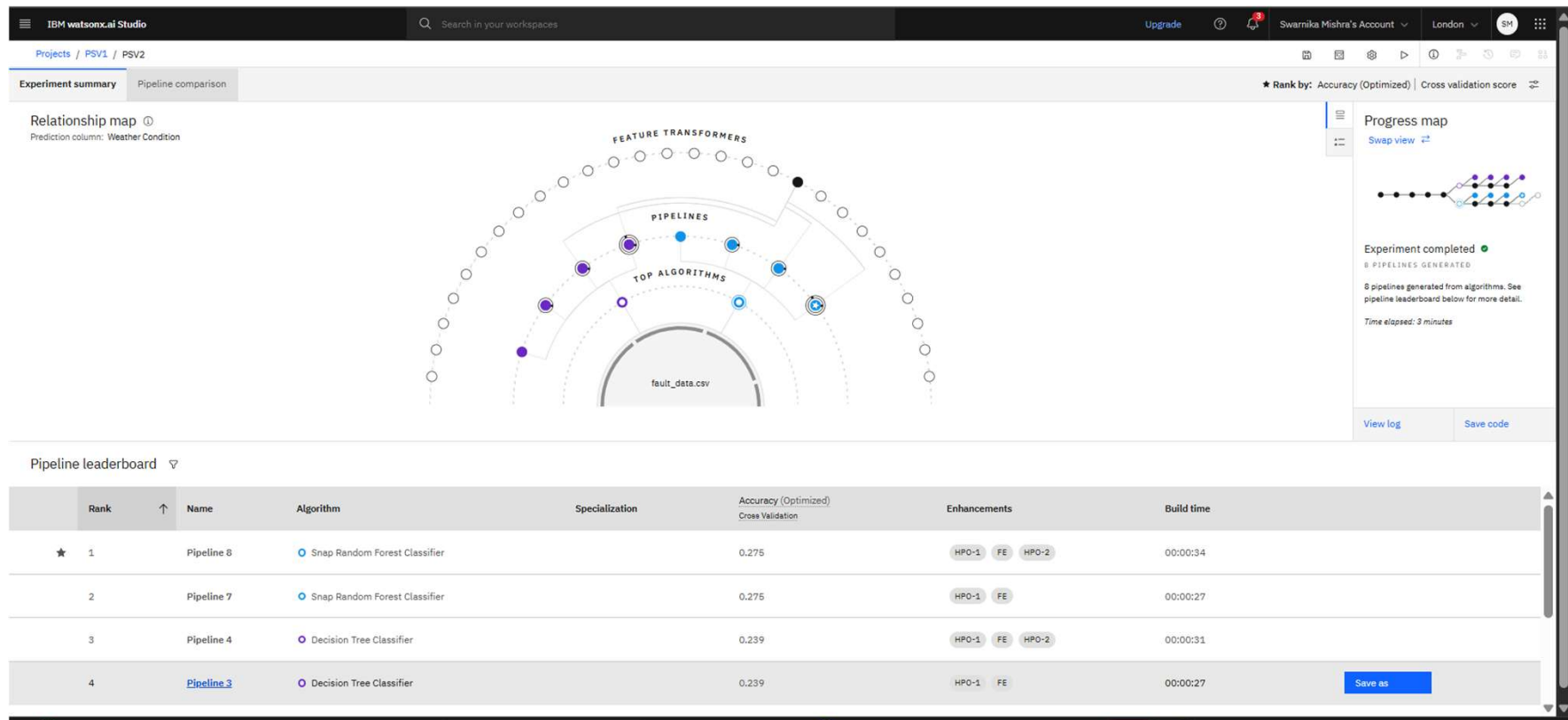
System requirements

- IBM Cloud(mandatory)IBM Watson studio for model development and deployment IBM cloud object storage for dataset handling.

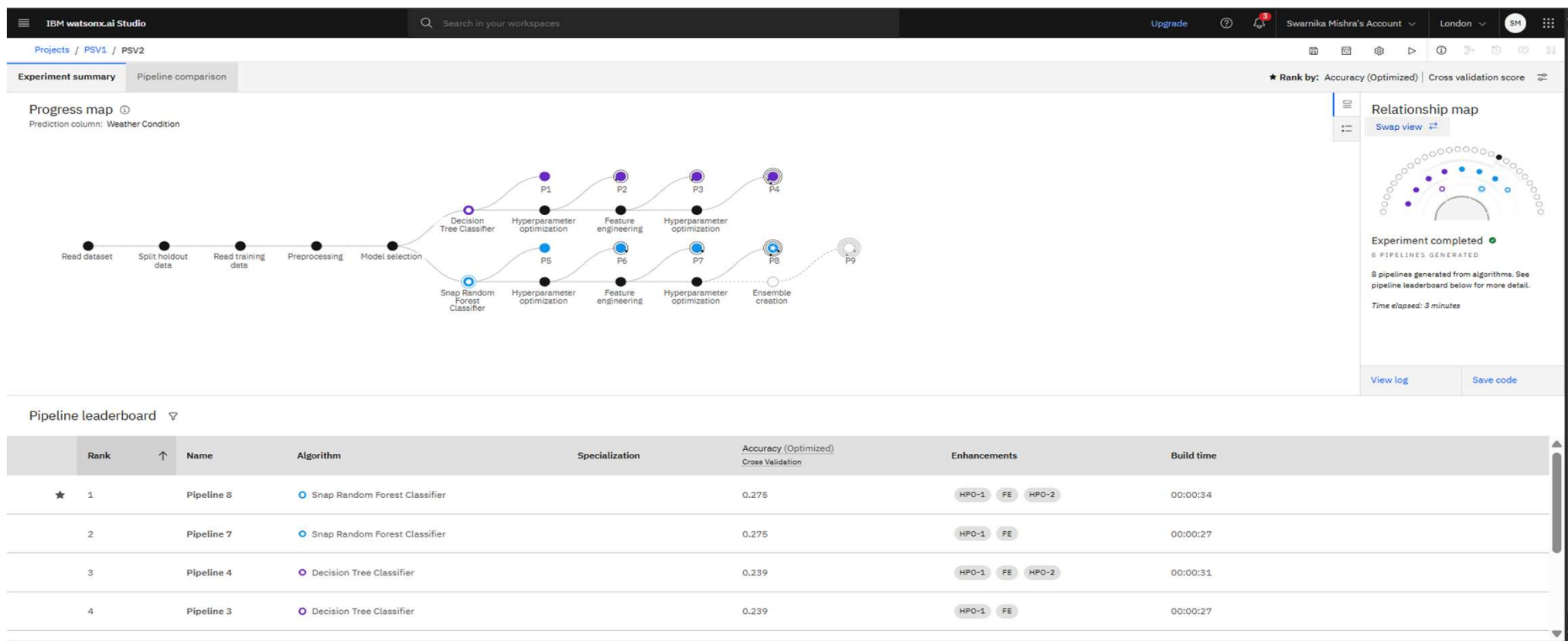
ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting bike counts. Here's an example structure for this section:
- **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



fault_detect 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 Type (other)	Fault Location (Latitude, Longitude) (other)	Voltage (V) (double)	Current (A) (double)	Power Load (MW) (double)	Temperature (°C) (double)	Wind Speed (km/h) (double)	Maintenance Status (other)
1	F010	overheating	(34.9687, -118.5356)	2200	250	70	33	15	pending
2	F011	Line Breakage	(34.5459, -118.8822)	2006	199	50	34	30	completed
3	F001	Transformer Failure	(34.3229, -118.46)	2093	100	60	29	19	scheduled
4	F013	overheating	(34.7708, -118.3011)	1873	200	69	45	27	completed
5	F019	line breakage	(34.2486, -118.6496)	1977	238	77	39	30	pending
6	F002	line breakage	(34.3413, -118.6266)	2005	150	80	42	29	scheduled
7	F077	Transformer Failure	(34.3032, -118.1977)	1890	129	50	24	35	pending
8	F050	over heating	(34.0761, -118.5825)	2006	239	74	38	28	completed
9									
10									

8 rows, 12 columns

Predict

X

Display format for prediction results

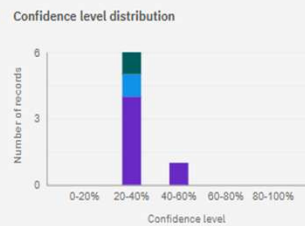
☒ Table view ☐ JSON view

☐ Show input data ⓘ

	Prediction	Confidence
1	Clear	26%
2	Clear	26%
3	Clear	40%
4	Clear	29%
5	Rainy	24%
6	Clear	26%
7	Thunderstorm	23%
8		
9		
10		
11		
12		
13		
14		
15		
16		



■ Clear
 ■ Rainy
 ■ Thundersto...



■ Clear
■ Rainy
■ Thundersto...

[Download JSON file](#)

CONCLUSION

- Design a machine learning model to detect and classify different types of faults in a power distribution system.

FUTURE SCOPE

- 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.

REFERENCES

- <https://www.kaggle.com/datasets/ziya07/power-systemfaults-dataset>
- **Technology** – Use of IBM Cloud Lite services is mandatory. ML Project in Watsonx.ai.studio in IBM cloud

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Swarnika Mishra

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence



Issued on: Jul 18, 2025
Issued by: IBM SkillsBuild

Verify: <https://www.credly.com/badges/d023367e-8fb0-4f8b-b869-afa82231ad4e>



IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Swarnika 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/9d5e6edb-1120-4db3-9be9-c1d1eb7c8778>



IBM CERTIFICATIONS



https://skills.yourlearning.ibm.com/certificate/ALM-COURSE_3824998

1/1

THANK YOU