CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION UISNG MACHINE LEARNING

Presented By:
1. Raj Kumar Raj –Hemvati Nandan Bahuguna Garhwal University (A Central University) –Computer Science Engineering

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

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.
  - Pre-processing : 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 F!-score.

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 Watson studio for model development and deployment

  IBM cloud object storage for dataset handling

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**

  Random Forest Classifier (or SVM based on performance)

- **Data Input:**

  Voltage , Current , and Phasor measurements from the datasets

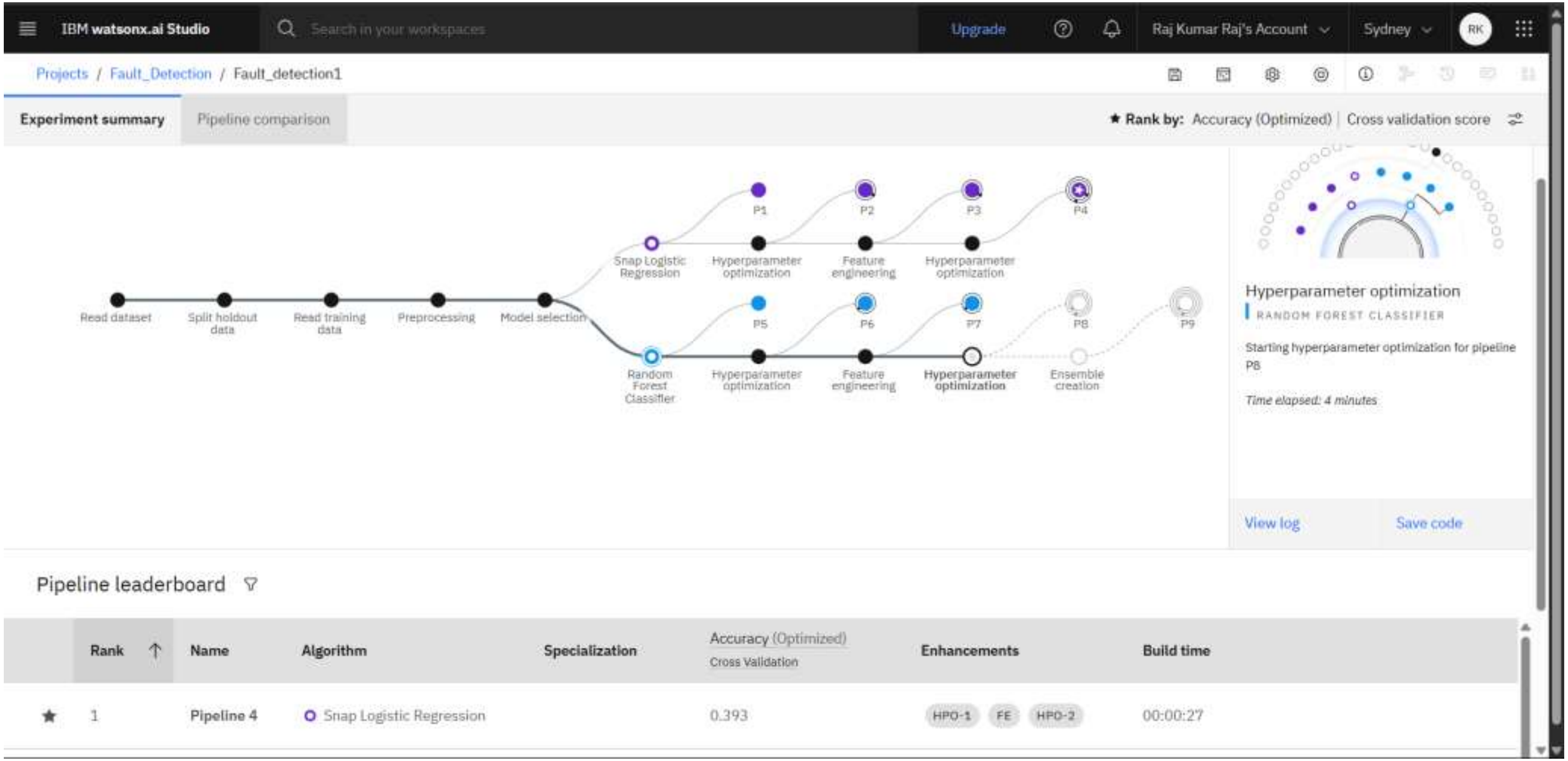
- **Training Process:**

  Supervised learning using labelled fault types

- **Prediction Process:**

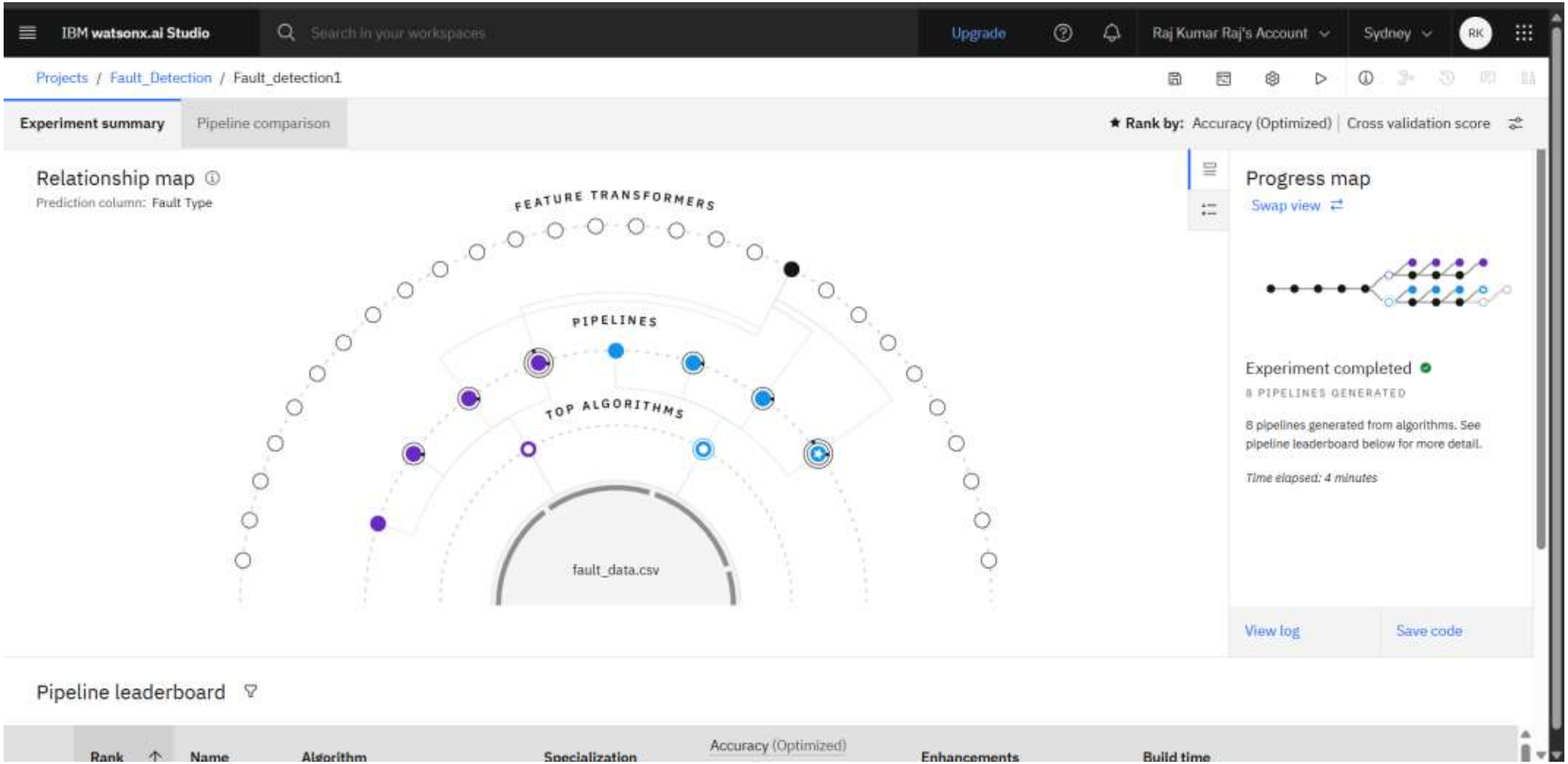
  Model deployed on IBM Watson Studio with API endpoint for real-time predictions

RESULT



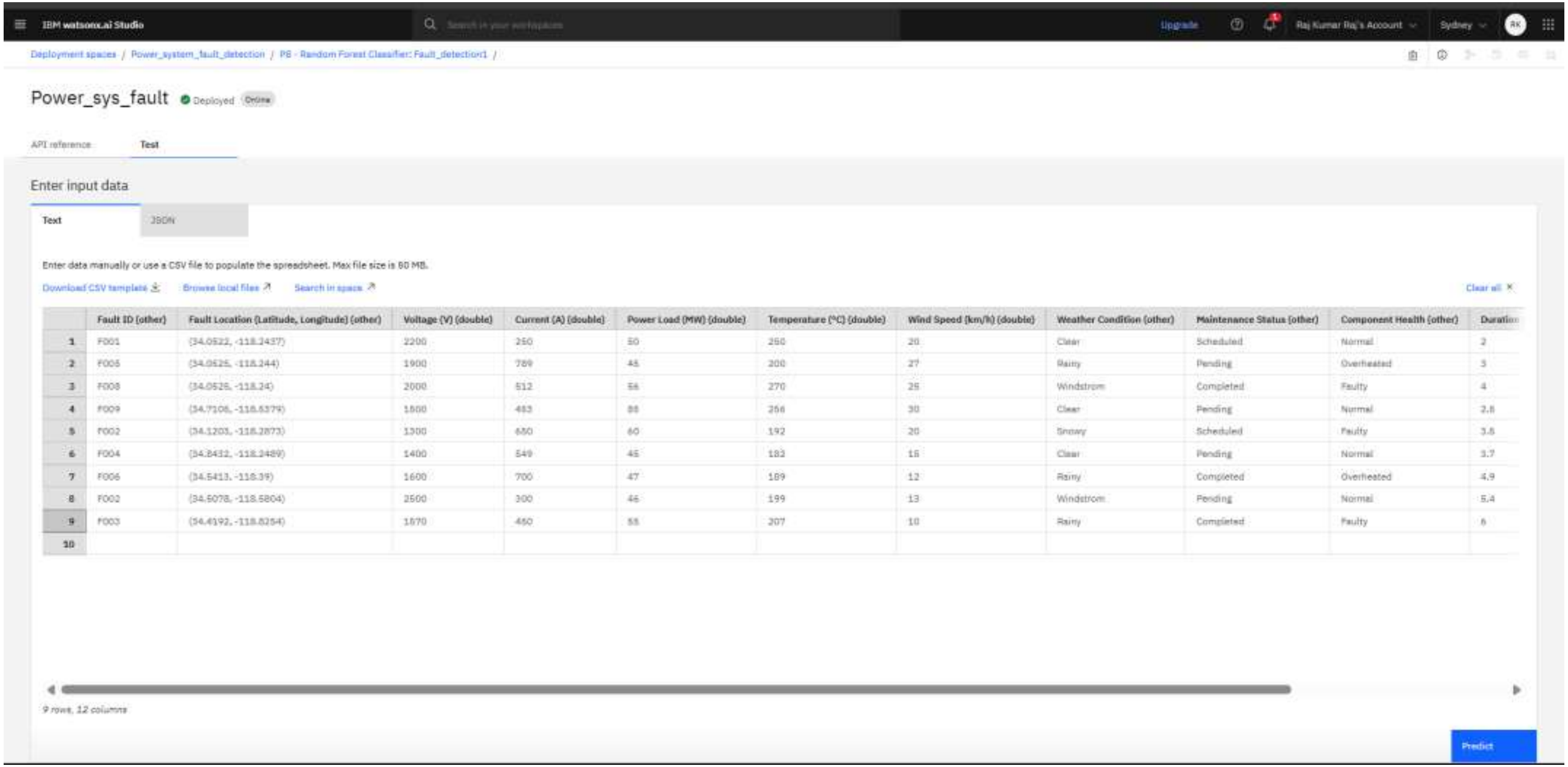
IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Raj Kumar Raj's Account
Sydney
RK
Projects / Fault_Detection / Fault_detection1
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Read dataset
Split holdout data
Read training data
Preprocessing
Model selection
Snap Logistic Regression
Hyperparameter optimization
Feature engineering
Hyperparameter optimization
P1
P2
P3
P4
Random Forest Classifier
Hyperparameter optimization
Feature engineering
Hyperparameter optimization
Ensemble creation
P5
P6
P7
P8
P9
Hyperparameter optimization
RANDOM FOREST CLASSIFIER
Starting hyperparameter optimization for pipeline P8
Time elapsed: 4 minutes
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
Specialization
Accuracy (Optimized)
Cross Validation
Enhancements
Build time
1
Pipeline 4
Snap Logistic Regression
0.393
HPO-1
FE
HPO-2
00:00:27

RESULT



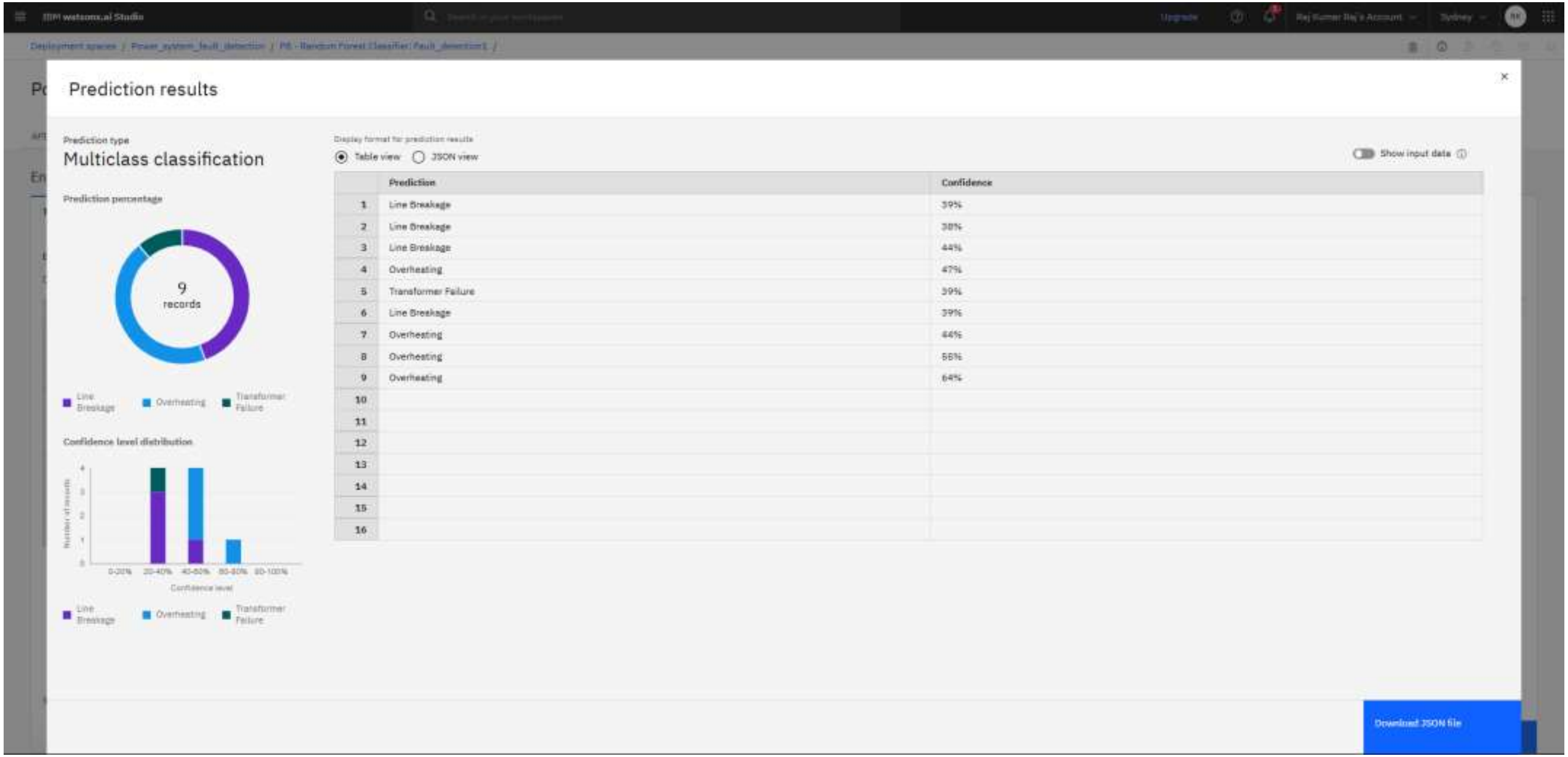
IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Raj Kumar Raj's Account
Sydney
RK
Projects / Fault_Detection / Fault_detection1
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Relationship map
Prediction column: Fault Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
fault_data.csv
Progress map
Swap view
Experiment completed
8 PIPELINES GENERATED
8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 4 minutes
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
Specialization
Accuracy (Optimized)
Enhancements
Build time

RESULT



	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)	Maintenance Status (other)	Component Health (other)	Duratio
1	F001	(34.0522, -118.2437)	2200	250	50	250	20	Clear	Scheduled	Normal	2
2	F005	(34.0525, -118.244)	1900	789	45	200	27	Rainy	Pending	Overheated	3
3	F008	(34.0526, -118.24)	2000	512	56	270	25	Windstrom	Completed	Faulty	4
4	F009	(34.7106, -118.6379)	1800	483	88	256	30	Clear	Pending	Normal	2.8
5	F002	(34.1203, -118.2873)	1300	650	60	192	20	Snowy	Scheduled	Faulty	3.6
6	F004	(34.8432, -118.2489)	1400	549	45	182	15	Clear	Pending	Normal	3.7
7	F006	(34.5413, -118.39)	1600	700	47	189	12	Rainy	Completed	Overheated	4.9
8	F002	(34.5078, -118.5804)	2500	300	46	199	13	Windstrom	Pending	Normal	5.4
9	F003	(34.4192, -118.8254)	1670	450	55	207	10	Rainy	Completed	Faulty	6
10											

RESULT



Prediction results
Prediction type
Multiclass classification
Prediction percentage
9
records
Line Breakage
Overheating
Transformer Failure
Confidence level distribution
Display format for prediction results
Table view
JSON view
Show input data
Prediction
Confidence
1 Line Breakage 39%
2 Line Breakage 38%
3 Line Breakage 44%
4 Overheating 47%
5 Transformer Failure 39%
6 Line Breakage 39%
7 Overheating 44%
8 Overheating 56%
9 Overheating 64%
Download JSON file

CONCLUSION

- The developed machine learning model successfully classifies different types of power system faults using electrical measurements such as voltage, current, and phasor data. By utilizing supervised learning techniques and training the model on a Kaggle dataset, the system can accurately distinguish between normal and various fault conditions, including line-to-ground, line-to-line, and three-phase faults. The integration of the model with IBM Watson Studio and IBM Cloud Object Storage ensures seamless model deployment and accessibility for real-time fault detection. This approach can significantly reduce the time required to identify and respond to faults, thereby improving the stability, reliability, and efficiency of power distribution systems.

FUTURE SCOPE

- **Enhanced Dataset** – Future work can focus on using real-time data from smart grids and IoT-enabled sensors to improve model generalization and robustness.
- **Deep Learning Models** – Advanced deep learning techniques such as LSTM and CNN can be explored for handling time-series electrical data to enhance accuracy.
- **Fault Severity Analysis** – The model can be extended to not only detect the type of fault but also estimate fault severity to prioritize maintenance actions.
- **Integration with SCADA Systems** – Real-time integration with SCADA and energy management systems can automate fault detection and isolation in practical power grids.
- **Edge Deployment** – Deploying lightweight models on edge devices for real-time decision-making in remote substations can further enhance fault response speed.

REFERENCES

- Kaggle Dataset: Power System Fault Detection Dataset (
- J. Zhao, Z. Zhang, et al., *“Machine Learning-Based Fault Detection in Power Distribution Systems,”* IEEE Transactions on Smart Grid, 2021.
- A. Sahoo, R. Mohanty, *“Fault Detection and Classification in Power Systems using Machine Learning Techniques,”* International Journal of Electrical Power & Energy Systems, 2020.
- IBM Cloud Documentation – https://cloud.ibm.com/docs
- IBM Watson Studio Documentation – https://www.ibm.com/cloud/watson-studio

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Raj Kumar Raj

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/92cedd3f-7456-42f0-84bf-0ea98d038f2b

edunet foundation

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

Raj Kumar Raj

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/0992e7fa-3b22-4d02-bb71-9ddbd0d49478

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Raj Kumar Raj

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