CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
1. Jeya Nivin J-Sathyabama Institute of Science and Technology-B.E. Computer Science and Engineering

edunet
foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development 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

The proposed system uses machine learning to detect and classify faults in a power distribution system based on electrical measurement data (e.g., voltage and current phasors). It consists of the following components

Data Collection:

Acquire voltage and current phasor data under various conditions, including normal operation and different fault types (line-to-ground, line-to-line, three-phase).

Data Preprocessing:

Clean the data, handle missing values, and extract key features such as phasor magnitude, angle changes, and symmetrical components.

Machine Learning Algorithm:

Train a supervised machine learning model (e.g., Random Forest, SVM, CNN, or LSTM) to classify the type of fault based on labeled data.

Deployment:

Deploy the model to monitor incoming data streams and provide real-time fault classification with minimal latency.

Evaluation:

Measure performance using accuracy, precision, recall, and confusion matrix. Continuously update the model with new data for improved accuracy.

edunet
foundation

SYSTEM APPROACH

Hardware:

- Processor: Intel i5 or higher
- RAM: 8–16 GB

Software:

- OS: Windows, Linux, or macOS
- Python 3.7+

Required Libraries

- Data Processing: numpy, pandas.
- Visualization: matplotlib, seabor.
- Machine Learning, scikit-learn, xgboost, tensorflow or keras.
- Utilities: joblib, pickle, scipy.signal.

ALGORITHM & DEPLOYMENT

Algorithm Selection:

An **LSTM network** is selected for its ability to model time-dependent patterns in electrical phasor data, making it effective for detecting and classifying different fault types in power systems.

Data Input:

The model takes voltage and current magnitudes and angles, sequence components, timestamps, and frequency changes as input features to differentiate normal operation from faults like line-to-ground, line-to-line, and three-phase faults.

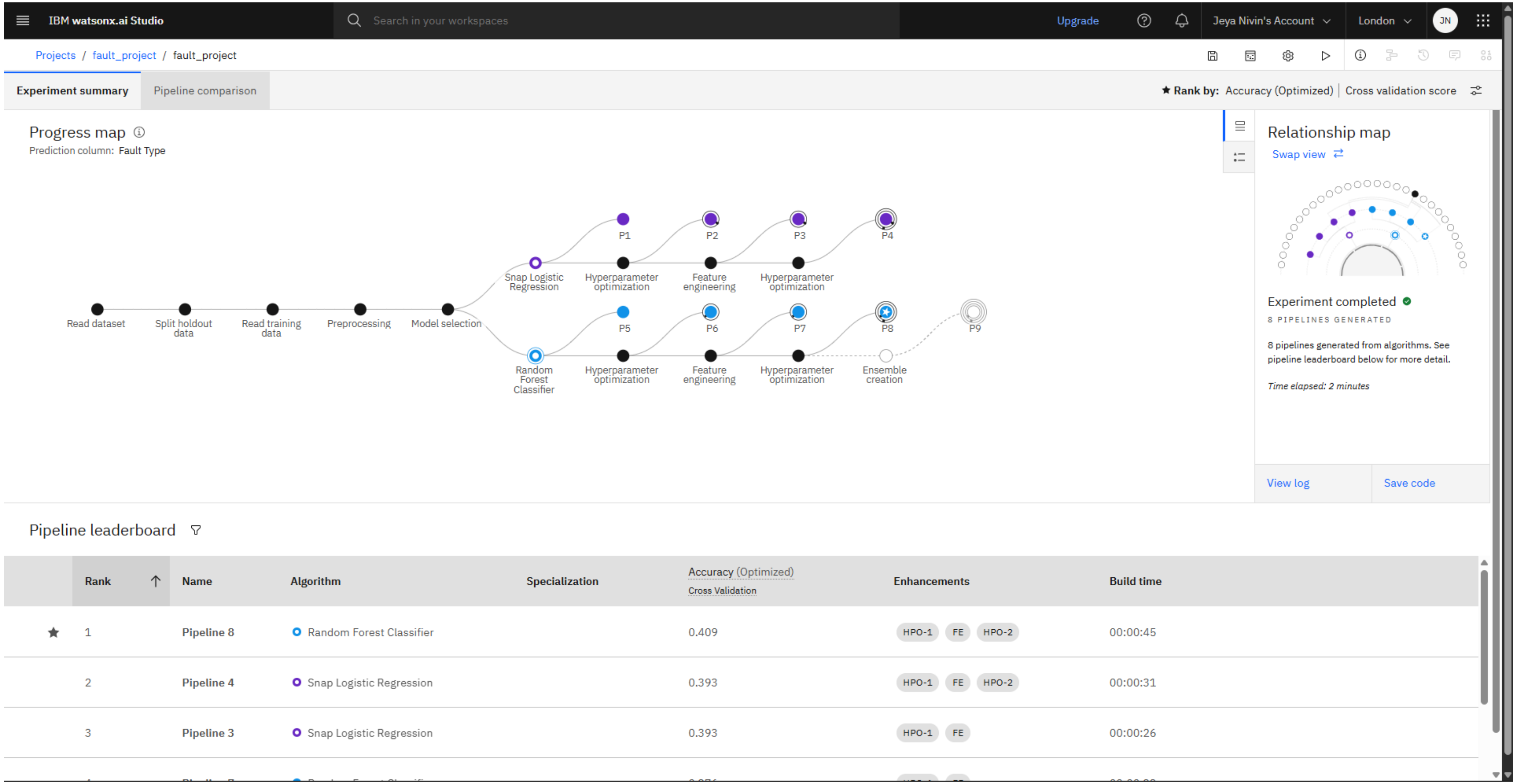
Training Process:

The model is trained on annotated historical or simulated datasets, using techniques like cross-validation and hyperparameter tuning to optimize performance and avoid overfitting.

Prediction Process:

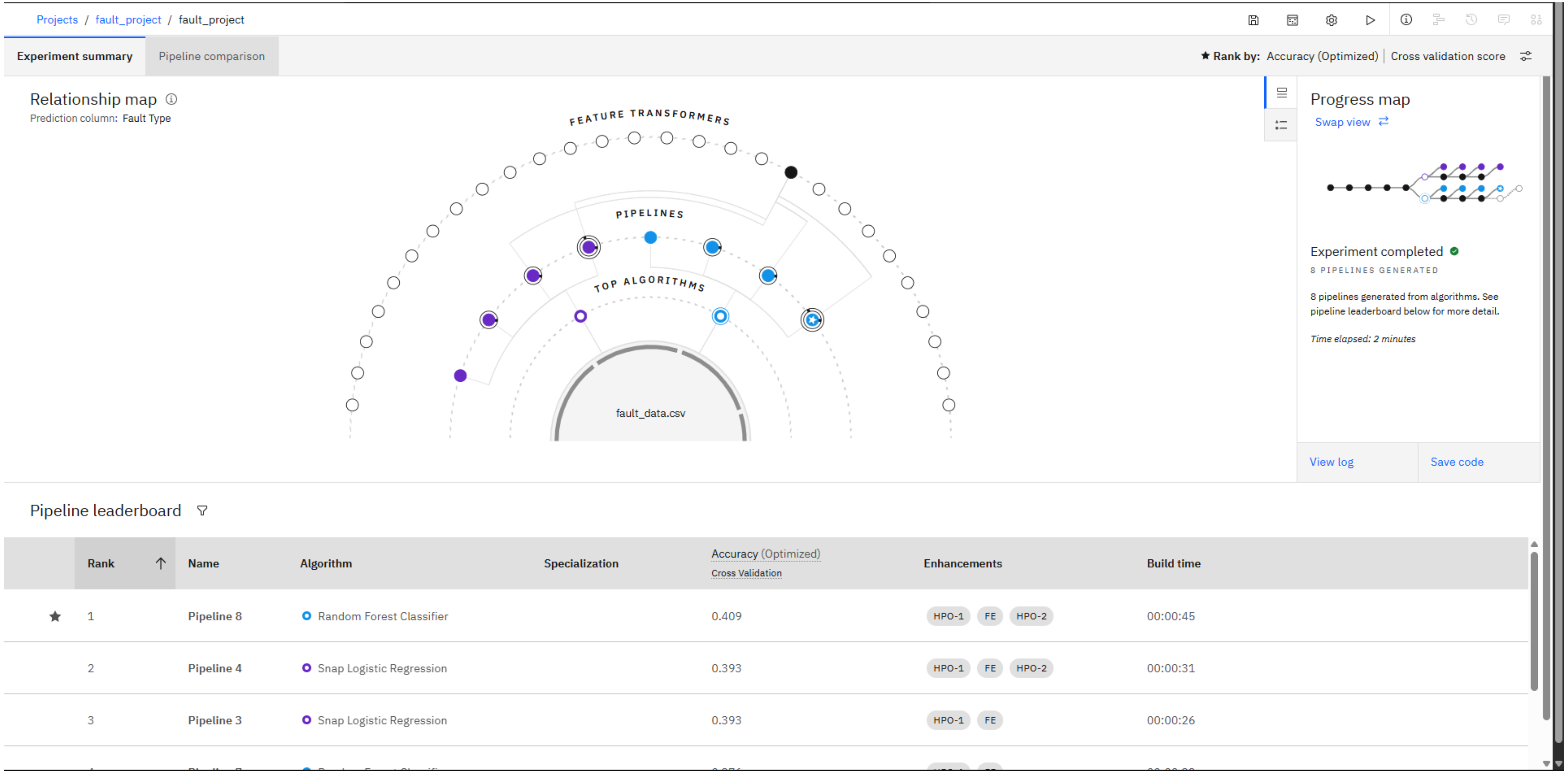
In deployment, the model analyzes streaming sensor data in real time to identify and classify faults quickly, supporting prompt grid protection actions.

RESULT

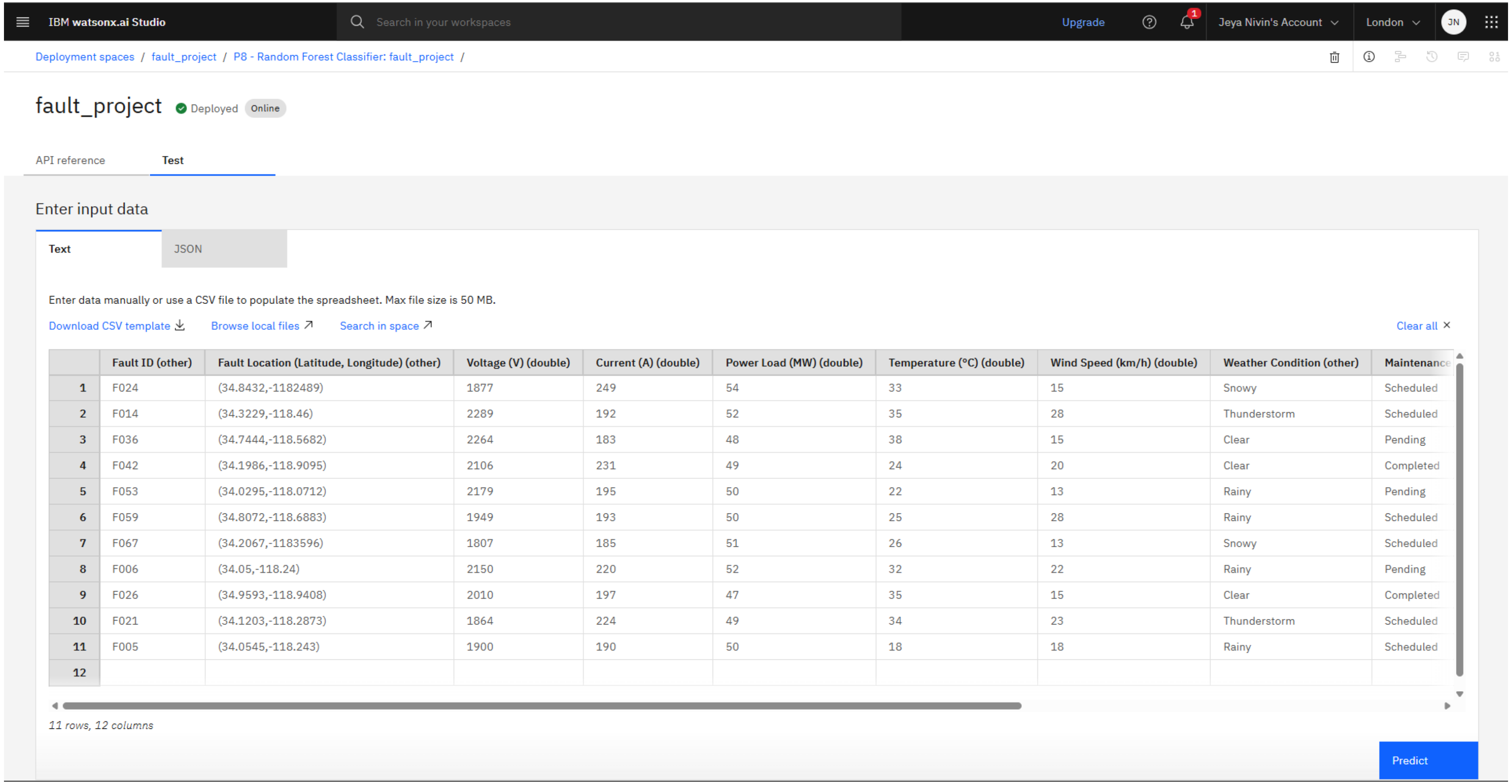


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Jeya Nivin's Account
London
JN
Projects / fault_project / fault_project
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Progress map
Prediction column: Fault Type
Read dataset
Split holdout data
Read 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: 2 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 00:00:45
2 Pipeline 4 Snap Logistic Regression 0.393 HPO-1 FE HPO-2 00:00:31
3 Pipeline 3 Snap Logistic Regression 0.393 HPO-1 FE 00:00:26

RESULT

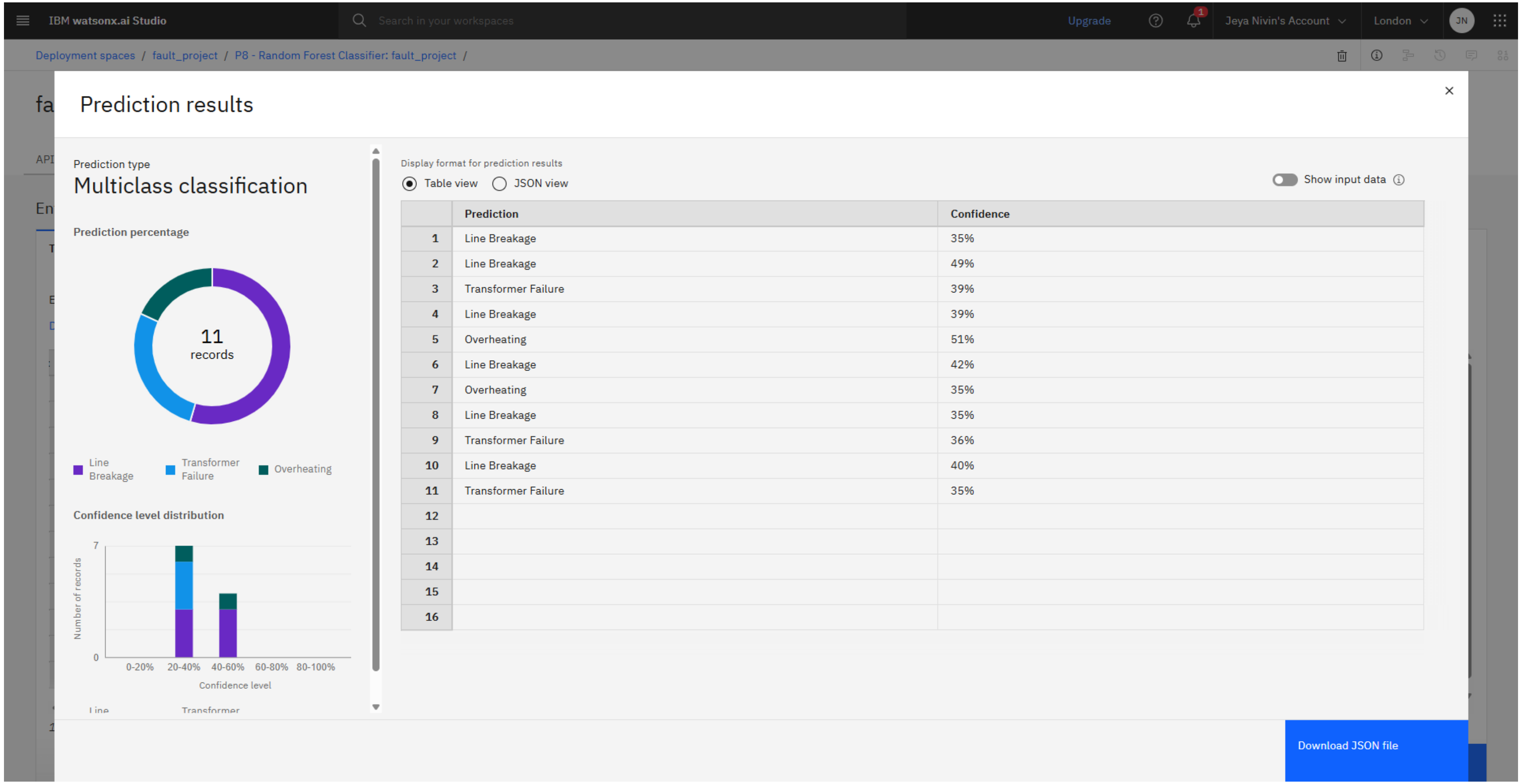

Projects / fault_project / fault_project
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Relationship map
Prediction column: Fault Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
fault_data.csv
Progress map
Swap view
Experiment completed
8 PIPELINES GENERATED
8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 2 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 00:00:45
2 Pipeline 4 Snap Logistic Regression 0.393 HPO-1 FE HPO-2 00:00:31
3 Pipeline 3 Snap Logistic Regression 0.393 HPO-1 FE 00:00:26

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
1	F024	(34.8432,-1182489)	1877	249	54	33	15	Snowy	Scheduled
2	F014	(34.3229,-118.46)	2289	192	52	35	28	Thunderstorm	Scheduled
3	F036	(34.7444,-118.5682)	2264	183	48	38	15	Clear	Pending
4	F042	(34.1986,-118.9095)	2106	231	49	24	20	Clear	Completed
5	F053	(34.0295,-118.0712)	2179	195	50	22	13	Rainy	Pending
6	F059	(34.8072,-118.6883)	1949	193	50	25	28	Rainy	Scheduled
7	F067	(34.2067,-1183596)	1807	185	51	26	13	Snowy	Scheduled
8	F006	(34.05,-118.24)	2150	220	52	32	22	Rainy	Pending
9	F026	(34.9593,-118.9408)	2010	197	47	35	15	Clear	Completed
10	F021	(34.1203,-118.2873)	1864	224	49	34	23	Thunderstorm	Scheduled
11	F005	(34.0545,-118.243)	1900	190	50	18	18	Rainy	Scheduled
12									

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Jeya Nivin's Account
London
Deployment spaces / fault_project / P8 - Random Forest Classifier: fault_project /
Prediction results
Prediction type
Multiclass classification
Prediction percentage
11
records
Line Breakage
Transformer Failure
Overheating
Confidence level distribution
Number of records
Confidence level
0-20%
20-40%
40-60%
60-80%
80-100%
Display format for prediction results
Table view
JSON view
Show input data
Prediction
Confidence
1 Line Breakage 35%
2 Line Breakage 49%
3 Transformer Failure 39%
4 Line Breakage 39%
5 Overheating 51%
6 Line Breakage 42%
7 Overheating 35%
8 Line Breakage 35%
9 Transformer Failure 36%
10 Line Breakage 40%
11 Transformer Failure 35%
12
13
14
15
16
Download JSON file

CONCLUSION

- The machine learning model effectively detects and classifies faults in power distribution systems using voltage and current data. It accurately distinguishes normal conditions from various faults, enabling fast and reliable fault identification to maintain grid stability.
- Challenges included handling noisy data and ensuring real-time detection. Future improvements could involve better data processing and adaptive learning for enhanced accuracy.
- Accurate fault detection is crucial for minimizing outages and protecting equipment, contributing to a more resilient power grid.

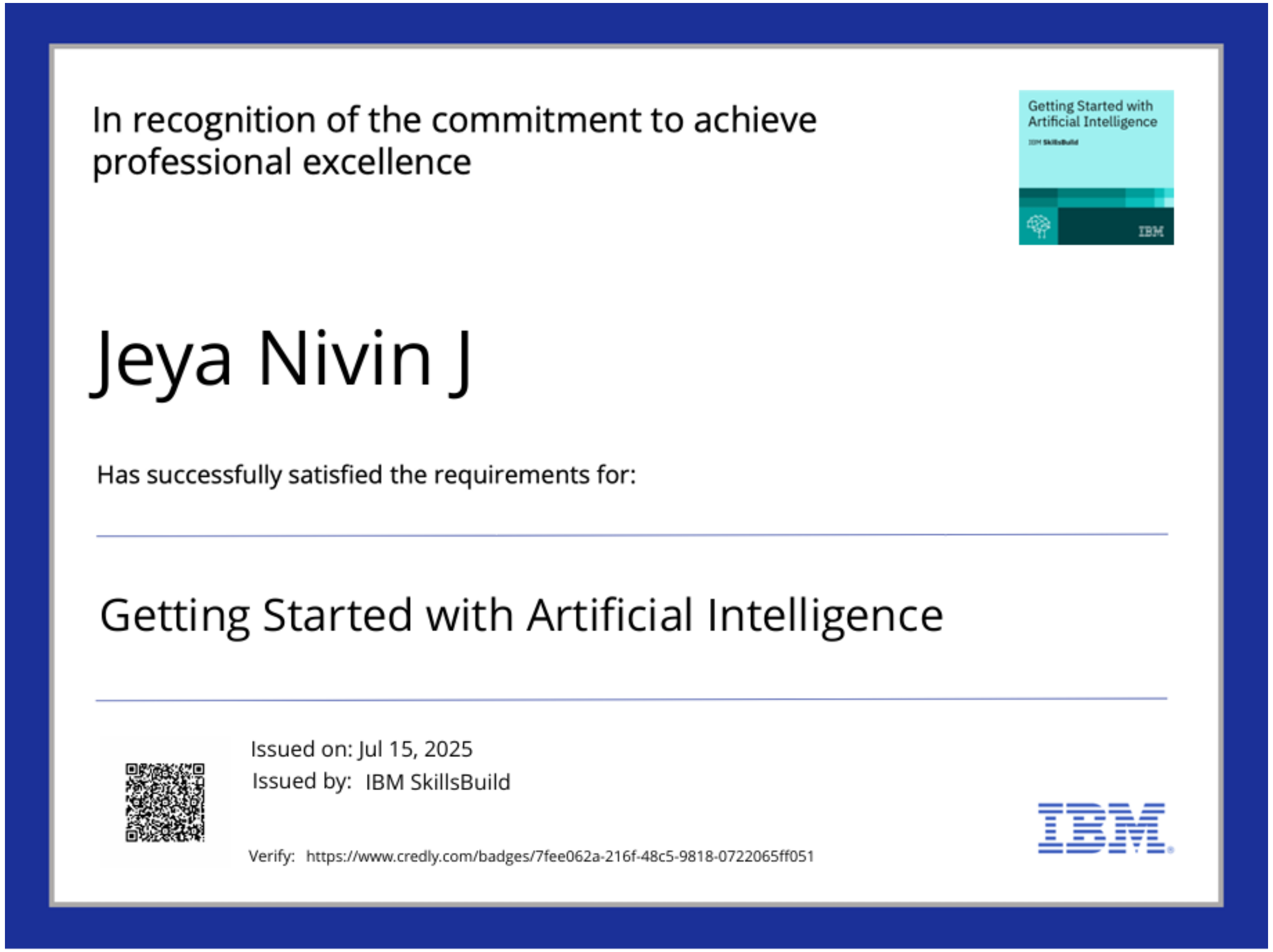
FUTURE SCOPE

The system can be improved by adding more data sources like environmental and equipment metrics to boost accuracy. Advanced algorithms such as deep learning and ensemble methods can enhance performance. Expanding the model to cover multiple regions will increase its applicability. Deploying on edge devices will enable faster, real-time fault detection. Incorporating continuous learning will help the system adapt to new fault patterns. Integration with smart grid technologies will support quicker fault isolation and recovery.

REFERENCES

- S. Wang, J. Li, and Q. Gao (2019) presented a deep learning approach for classifying power system faults using PMU measurements in *IEEE Transactions on Smart Grid*.
- A. Singh, S. K. Srivastava, and A. K. Sharma (2018) applied support vector machines for diagnosing faults in power distribution networks, published in *Electric Power Systems Research*.
- H. R. Karami and S. M. F. Ghaderi (2015) utilized wavelet transforms combined with neural networks for fault detection in electrical systems, as reported in *International Journal of Electrical Power & Energy Systems*.
- Y. LeCun, Y. Bengio, and G. Hinton (2015) reviewed foundational concepts in deep learning in their *Nature* article.
- IBM Cloud official documentation provides guidance on deploying machine learning models on their platform (https://cloud.ibm.com/docs).

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Jeya Nivin J

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 15, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/7fee062a-216f-48c5-9818-0722065ff051

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

Jeya Nivin J

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/670b8b1c-8e2a-4b18-b08d-a5905222ba13

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Jeya Nivin J

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