CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By: Yuvraj Vilas Shete (E&TC)
College Name & Department : MIT Academy Of Engineering

edunet foundation

OUTLINE

- Problem Statement
- Technology used
- Wow factor
- End users
- Result
- Conclusion
- Git-hub Link
- Future scope
- IBM Certifications

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 system trains a machine learning model for fault classification using IBM Watson Studio and AutoAI on IBM Cloud. To differentiate between different types of faults, it analyzes environmental data, voltage, and current. A real-time fault event classification for grid operators is the end result. 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.

TECHNOLOGY USED

System Requirements: 8 cpu & 32GB ram

IBM Cloud Lite account IBM Watson Studio Kaggle dataset

Technologies Used: IBM Watson AutoAI

IBM Cloud Object Storage

IBM CLOUD SERVICES USED

- IBM Cloud Watsonx AI Studio
- IBM Cloud Watsonx AI runtime
- IBM Cloud Agent Lab

ALGORITHM & DEVELOPMENT

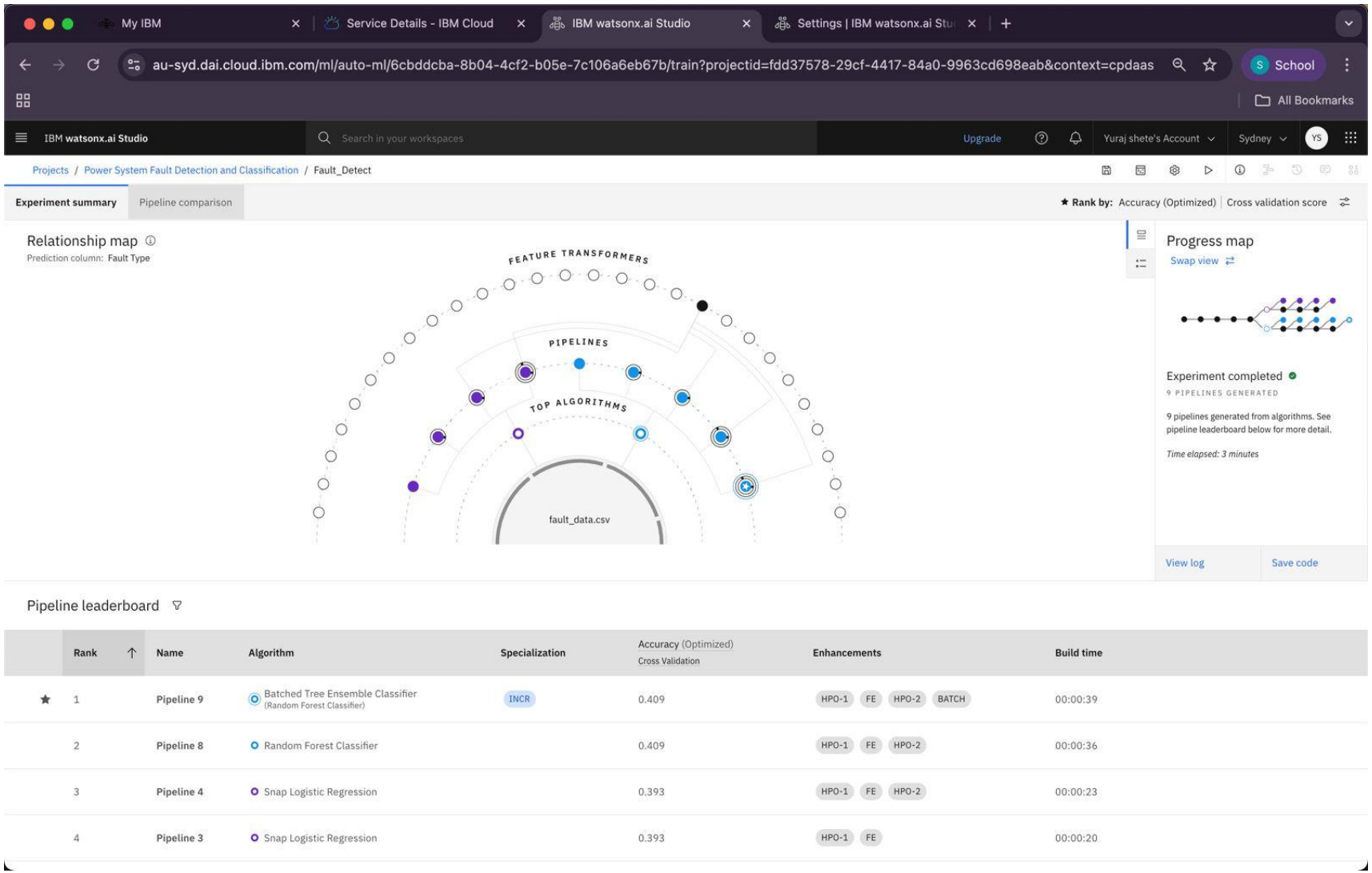
Algorithm: Snap Random Forest Classifier

#Main Input Features: Voltage Current Power Load Temperature Wind Speed

#Process: Data cleaned and fed into AutoAI Best pipeline selected based on accuracy

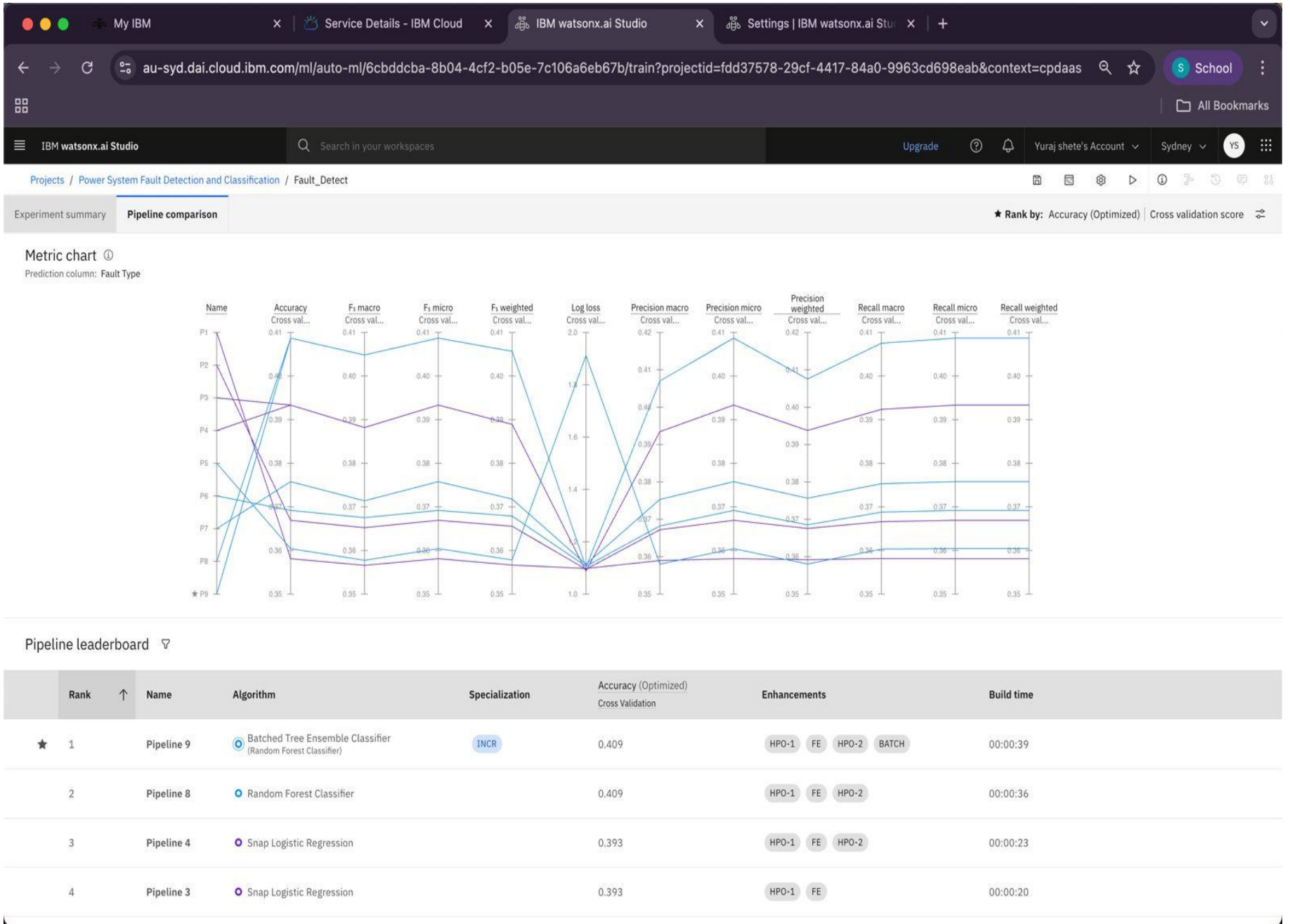
#Deployment: Watson ML deployed endpoint used for real-time predictions

RESULTS



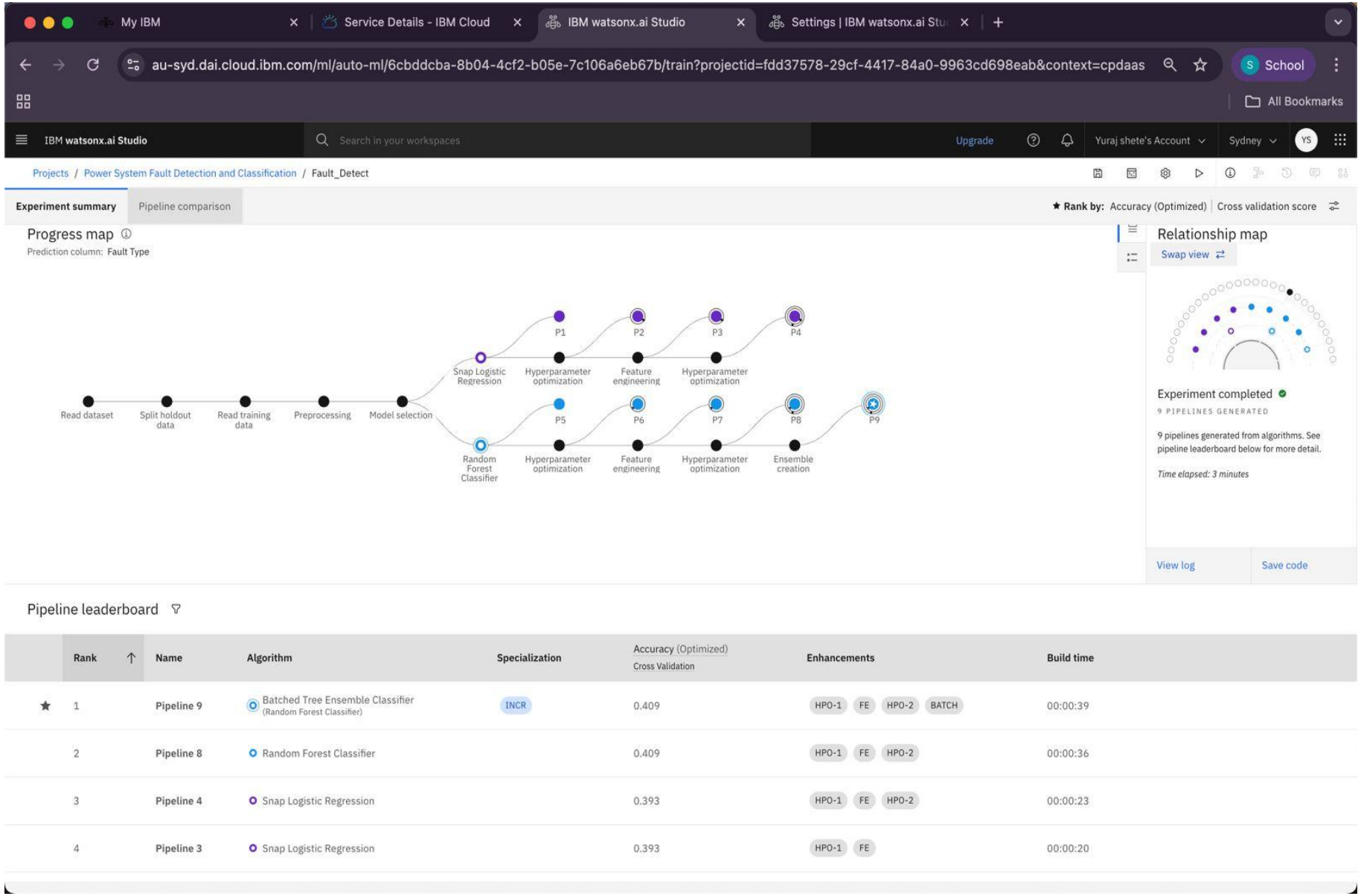
edunet foundation

RESULTS



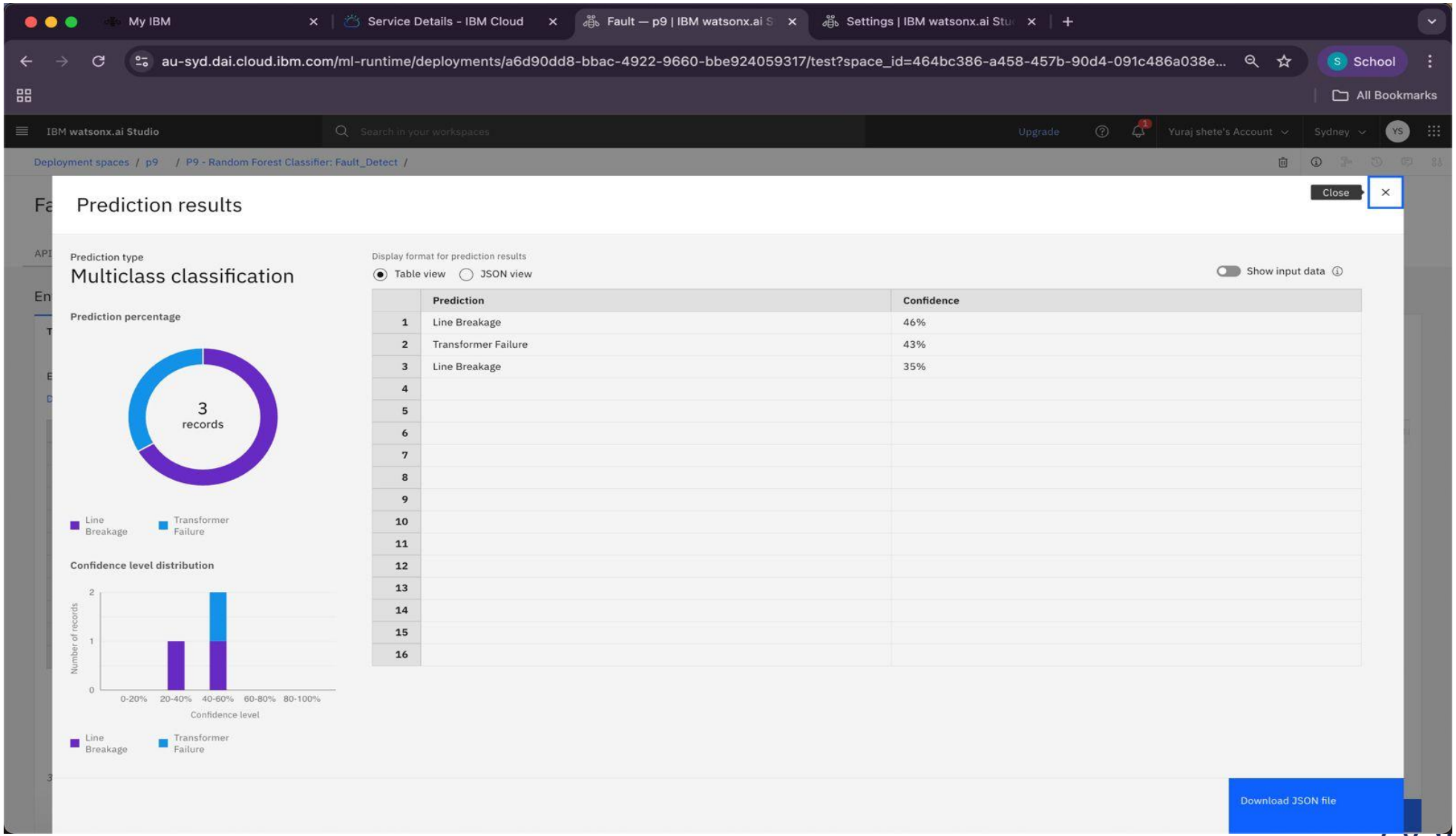
	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 9	Batched Tree Ensemble Classifier (Random Forest Classifier)	INCR	0.409	HPO-1 FE HPO-2 BATCH	00:00:39
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:36
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:23
	4	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:00:20

RESULTS



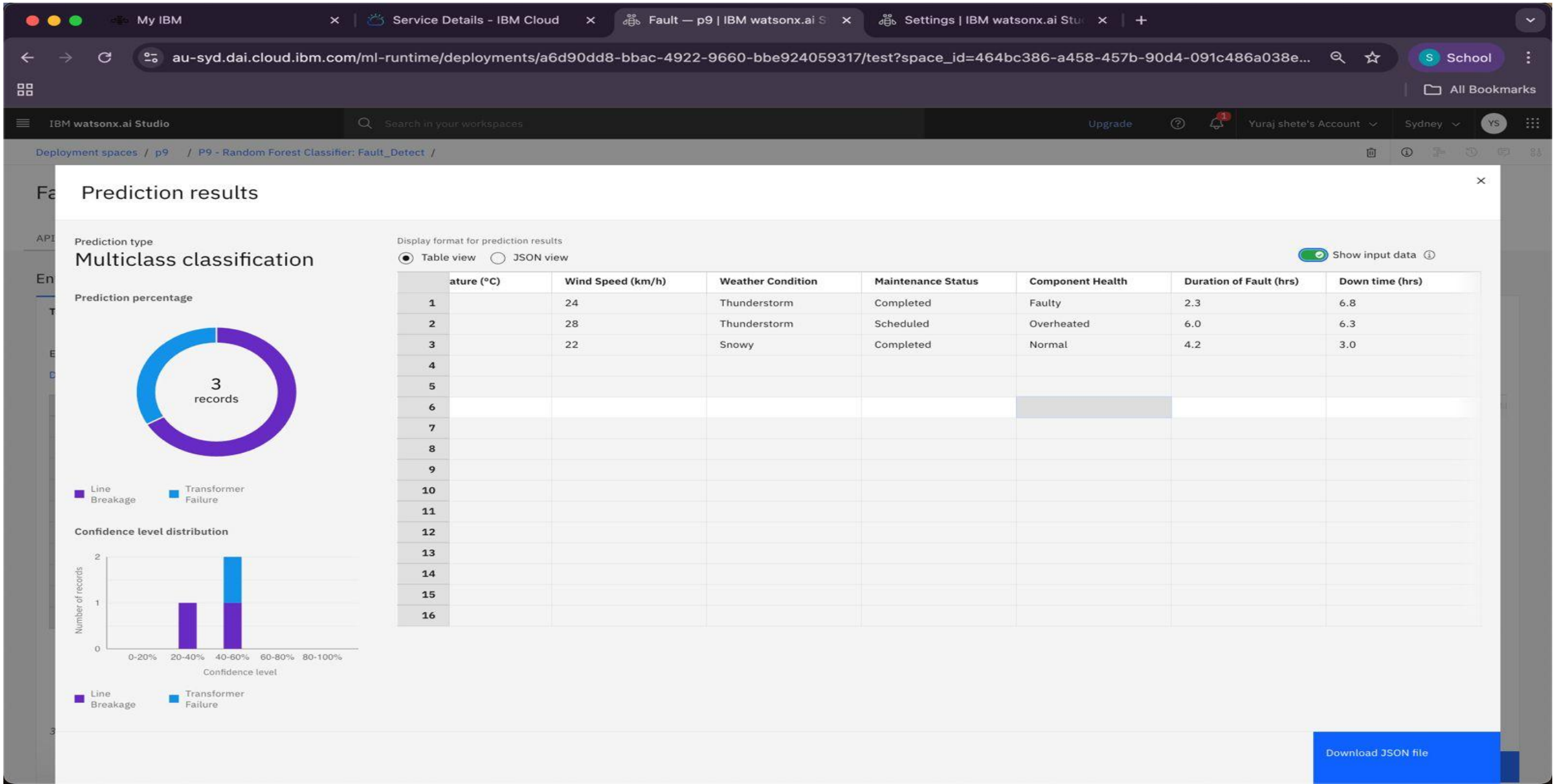
	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 9	Batched Tree Ensemble Classifier (Random Forest Classifier)	INCR	0.409	HPO-1 FE HPO-2 BATCH	00:00:39
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:36
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:23
	4	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:00:20

RESULTS



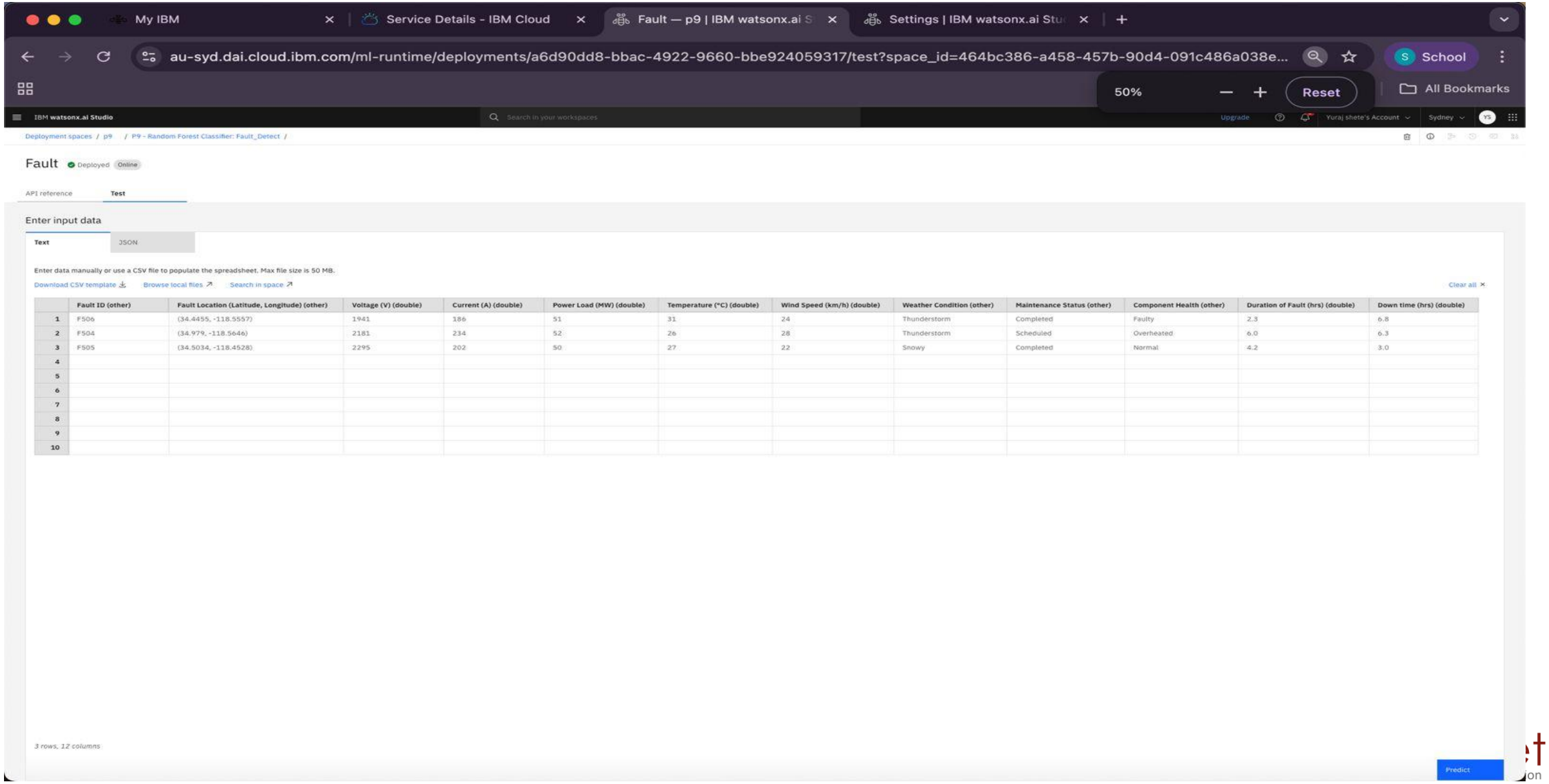
	Prediction	Confidence
1	Line Breakage	46%
2	Transformer Failure	43%
3	Line Breakage	35%
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

RESULTS



	ature (°C)	Wind Speed (km/h)	Weather Condition	Maintenance Status	Component Health	Duration of Fault (hrs)	Down time (hrs)
1		24	Thunderstorm	Completed	Faulty	2.3	6.8
2		28	Thunderstorm	Scheduled	Overheated	6.0	6.3
3		22	Snowy	Completed	Normal	4.2	3.0
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							

RESULTS



	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)	Duration of Fault (hrs) (double)	Down time (hrs) (double)
1	F506	(34.4455, -118.5557)	1941	186	51	31	24	Thunderstorm	Completed	Faulty	2.3	6.8
2	F504	(34.979, -118.5646)	2181	234	52	26	28	Thunderstorm	Scheduled	Overheated	6.0	6.3
3	F505	(34.5034, -118.4528)	2295	202	50	27	22	Snowy	Completed	Normal	4.2	3.0
4												
5												
6												
7												
8												
9												
10												

CONCLUSION

- In this project, a machine learning model was developed to detect and classify power system faults using sensor and environmental data. After testing multiple pipelines, the best-performing model was the **Batched Tree Ensemble Classifier (Random Forest)** with an optimized accuracy of **40.9%.** The model successfully classified faults into categories like **Line Breakage** and **Transformer Failure** using input features such as voltage, current, weather condition, and component health.

FUTURE SCOPE

- Add GPS-based fault location analysis
- Improve model accuracy with deep learning
- Integrate real-time sensor streaming
- Expand system to classify subtypes of each fault
- Add dashboard for visualization

IBM CERTIFICATIONS

edunet foundation

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Yuvraj shete

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 01 Aug 2025 (GMT)

Learning hours: 20 mins

edunet foundation

In recognition of the commitment to achieve professional excellence

YUVRAJ SHETE

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/98363115-71d5-42ef-90fe-ed9114d57591

IBM

edunet foundation

GITHUB LINK

- https://github.com/Yuvraj9600/-Machine-learning-project-

THANK YOU

edunet foundation