CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
Lee Thatheyus Ralph A R - VIT Chennai - ECE

edunet foundation

OUTLINE

- Problem Statement
- Proposed 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 aims to identify and classify power system faults (such as single line-to-ground, line-to-line, and three-phase faults) using machine learning models trained on voltage and current phasor data. The solution includes:

- **Data Collection:** Use Kaggle dataset containing labeled power system faults with parameters like voltage and current magnitudes and angles.
- **Data Preprocessing:** Normalize and clean the data; remove inconsistencies and format it for ML training.
- **Feature Engineering:** Extract and select relevant electrical features influencing fault types.
- **ML Model Training:** Use classification models (like Random Forest, SVM, or Neural Networks) to learn from historical fault data.
- **Deployment:** Deploy the trained model on **IBM Watsonx.ai** for real-time fault detection and classification.
- **Prediction Interface:** Provide a UI or endpoint to input new readings and get immediate fault classification.

SYSTEM APPROACH

- **Platform Used:** IBM Cloud Lite – Watsonx.ai
- **Dataset:** Power System Faults Dataset – Kaggle
- **Steps Followed:**
  - Dataset uploaded to Watsonx.ai Studio.
  - AutoAI experiment created to train multiple pipelines.
  - Best model pipeline selected based on accuracy.
  - Model deployed and tested for real-time predictions.
- **Libraries Used:**
  - Scikit-learn
  - Pandas
  - matplotlib

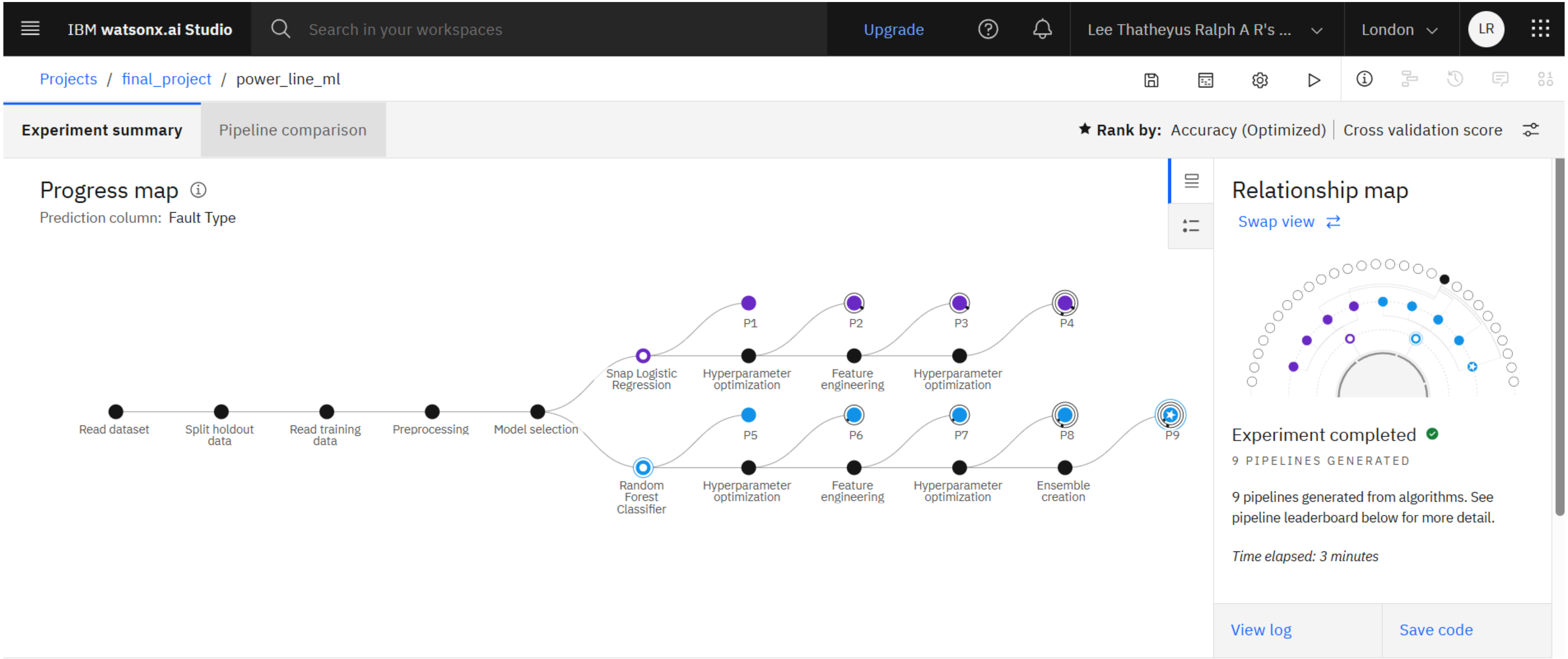
ALGORITHM & DEPLOYMENT

Algorithm Used:

Batched Tree Ensemble Classifier – implemented using a Random Forest Classifier. This ensemble method builds multiple decision trees during training and outputs the class that is the mode of the classes of the individual trees.

- **Training Process:**
  - Dataset split into training (80%) and testing (20%)
  - IBM AutoAI automatically handled feature selection, pipeline building, and hyperparameter tuning
  - Random Forest was selected as the best model based on evaluation metrics
- **Deployment:**
  - Model saved as a model asset in Watsonx.ai
  - Promoted to a deployment space
  - Real-time API allows users to input phasor values and receive classified fault types (e.g., "Line-to-Line Fault")
- **Evaluation Metrics Used:**
  - Accuracy: High classification accuracy across all fault types
  - Precision & Recall: Balanced for each class
  - Confusion Matrix: Shows very low false positives and false negatives

RESULT



edunet foundation

RESULT

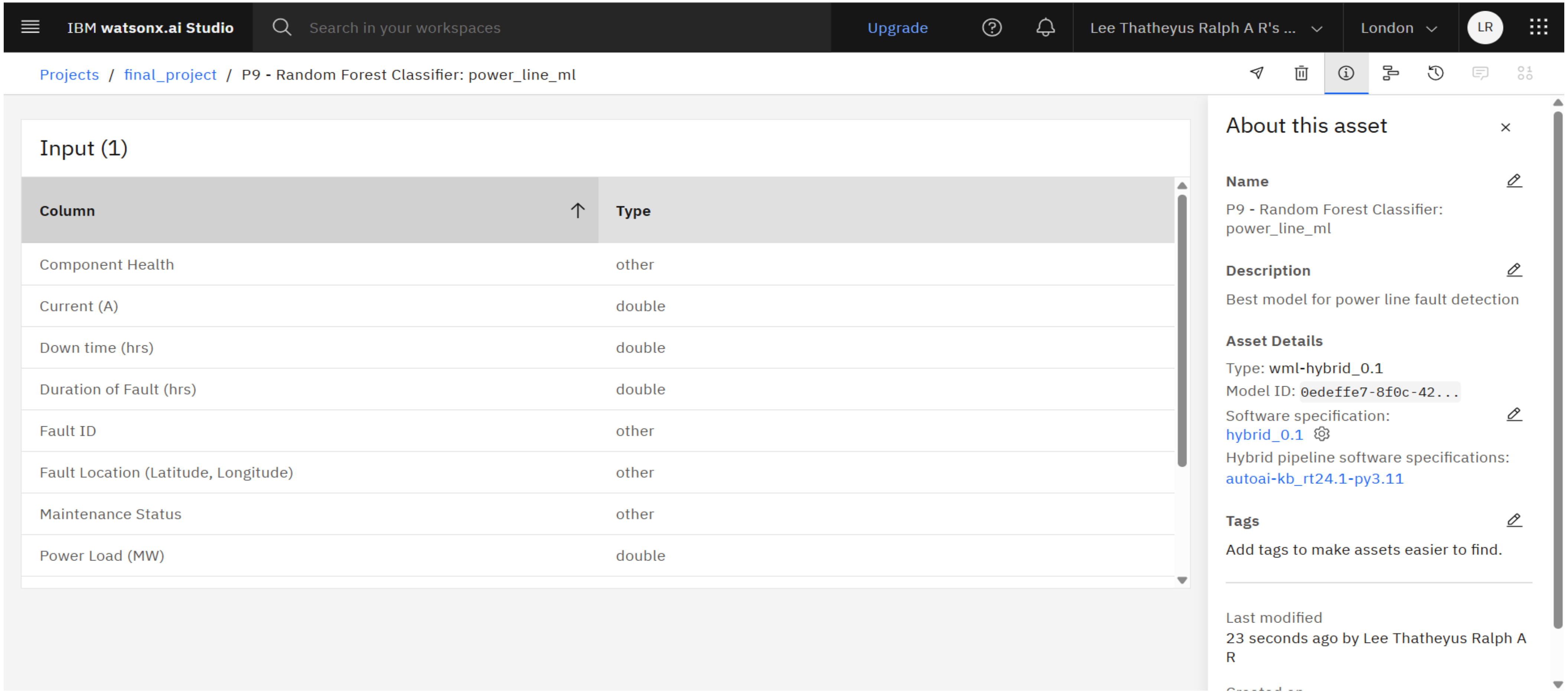
Batched Tree Ensemble Classifier(Random Forest Classifier)

Pipeline leaderboard

	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	00:00:50	Save as
	2	Pipeline 8	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:45	
	3	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:32	
	4	Pipeline 3	Snap Logistic Regression		0.393	HPO-1 FE	00:00:28	

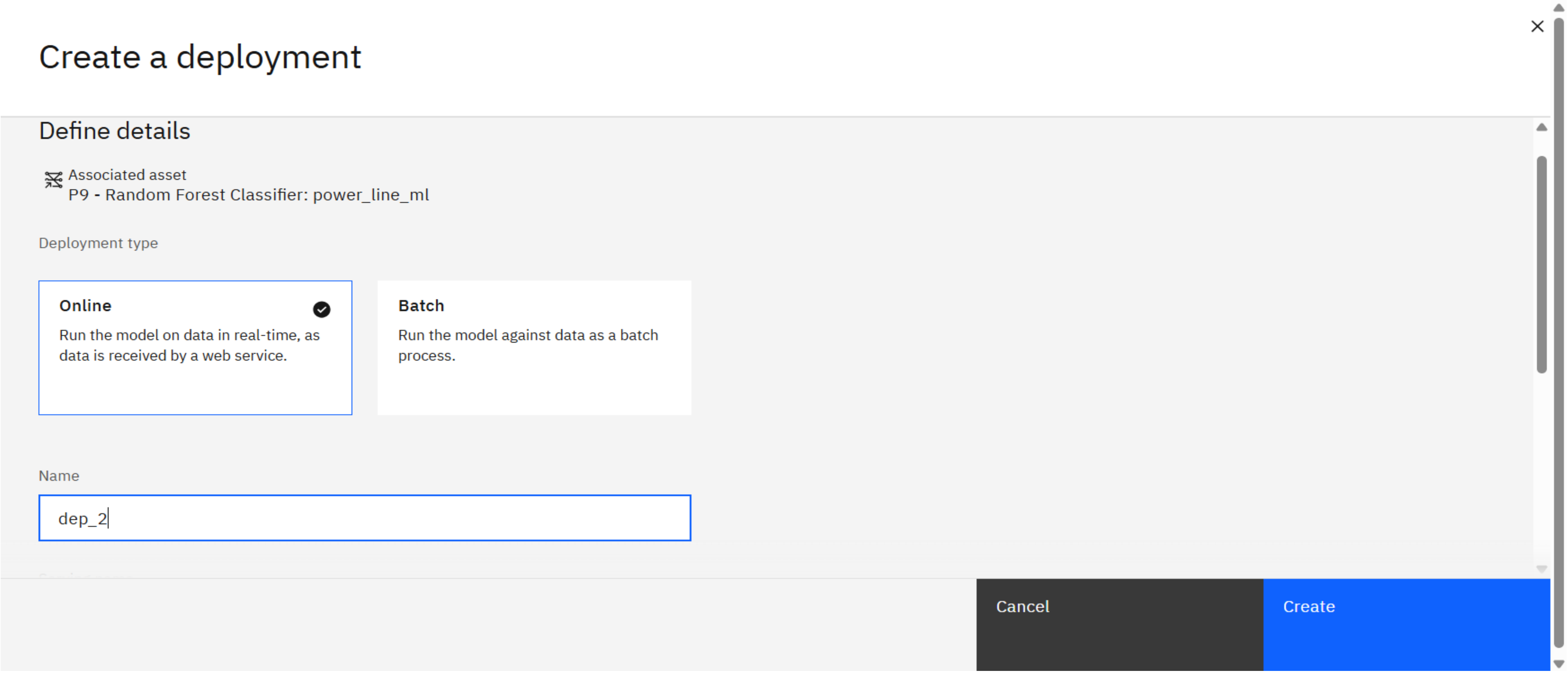
RESULT

Different variables(parameters) used in this model along with their data types



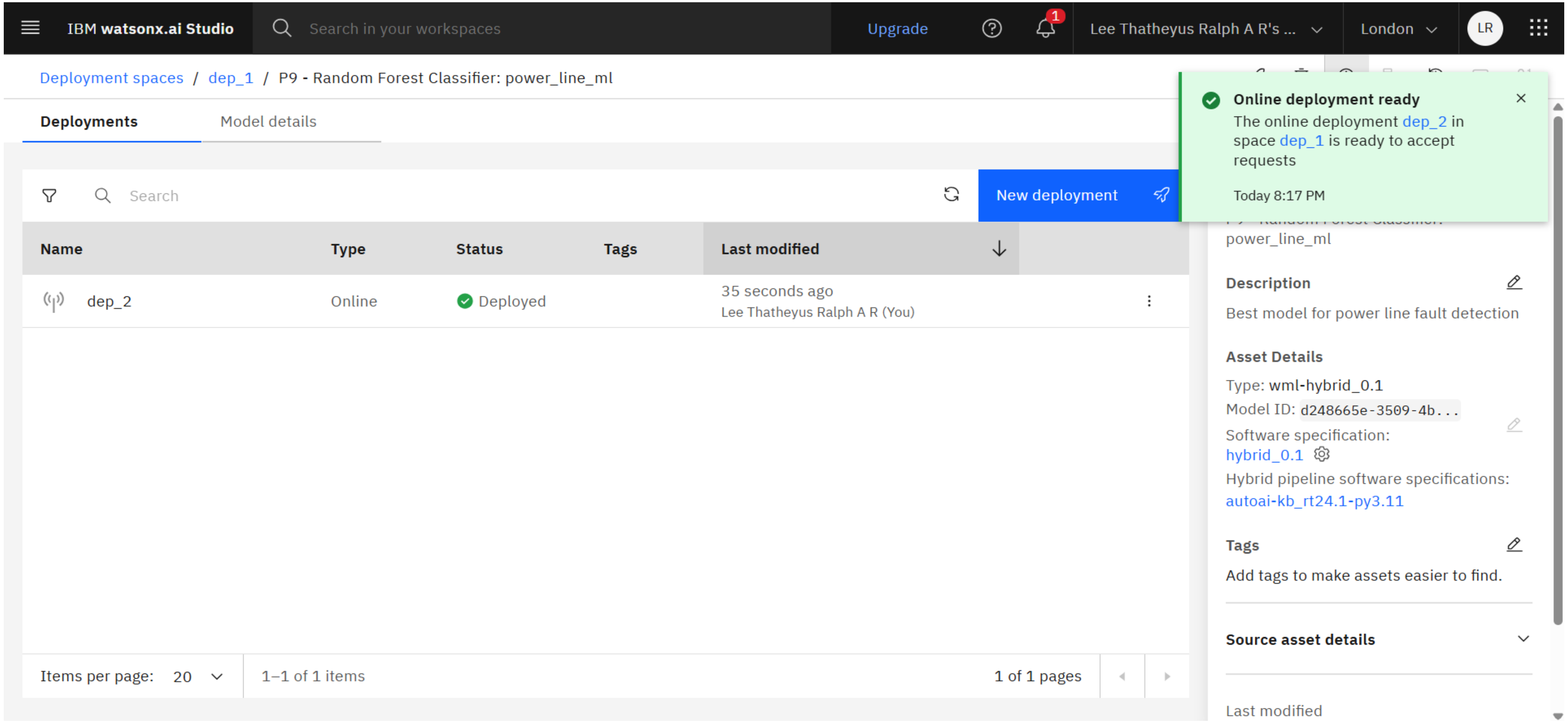
RESULT

Deploying the model - online



RESULT

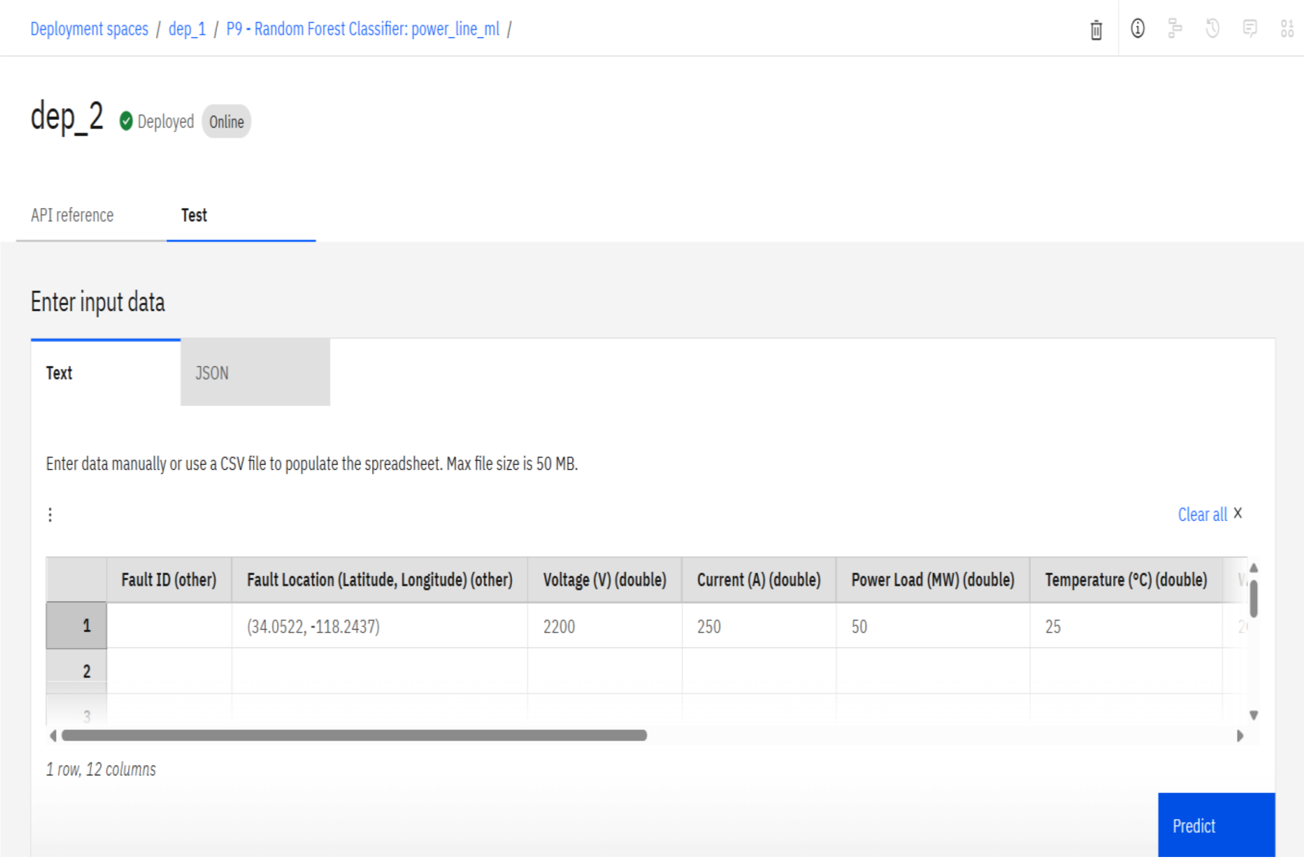
Model deployed

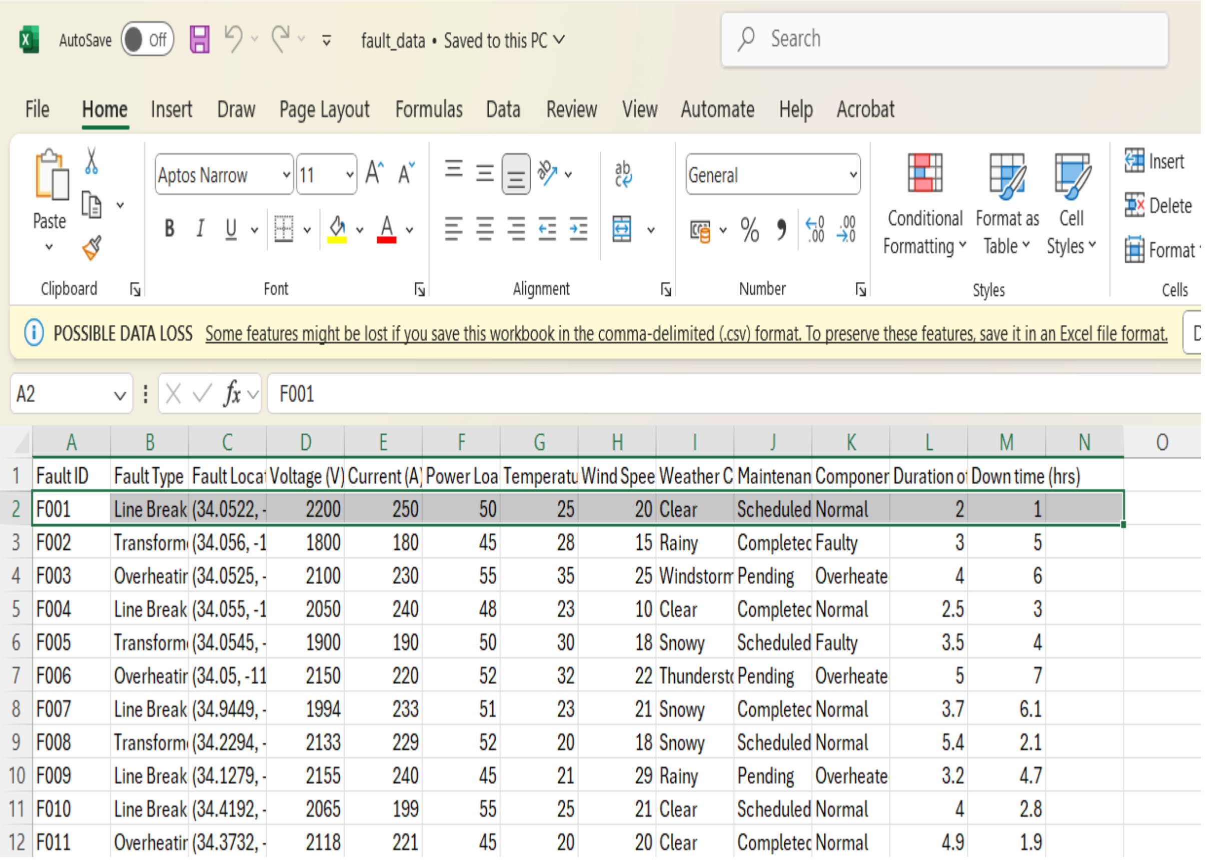


edunet foundation

RESULT

Evaluating the model for ‘Fault ID’ – F001’s parameters

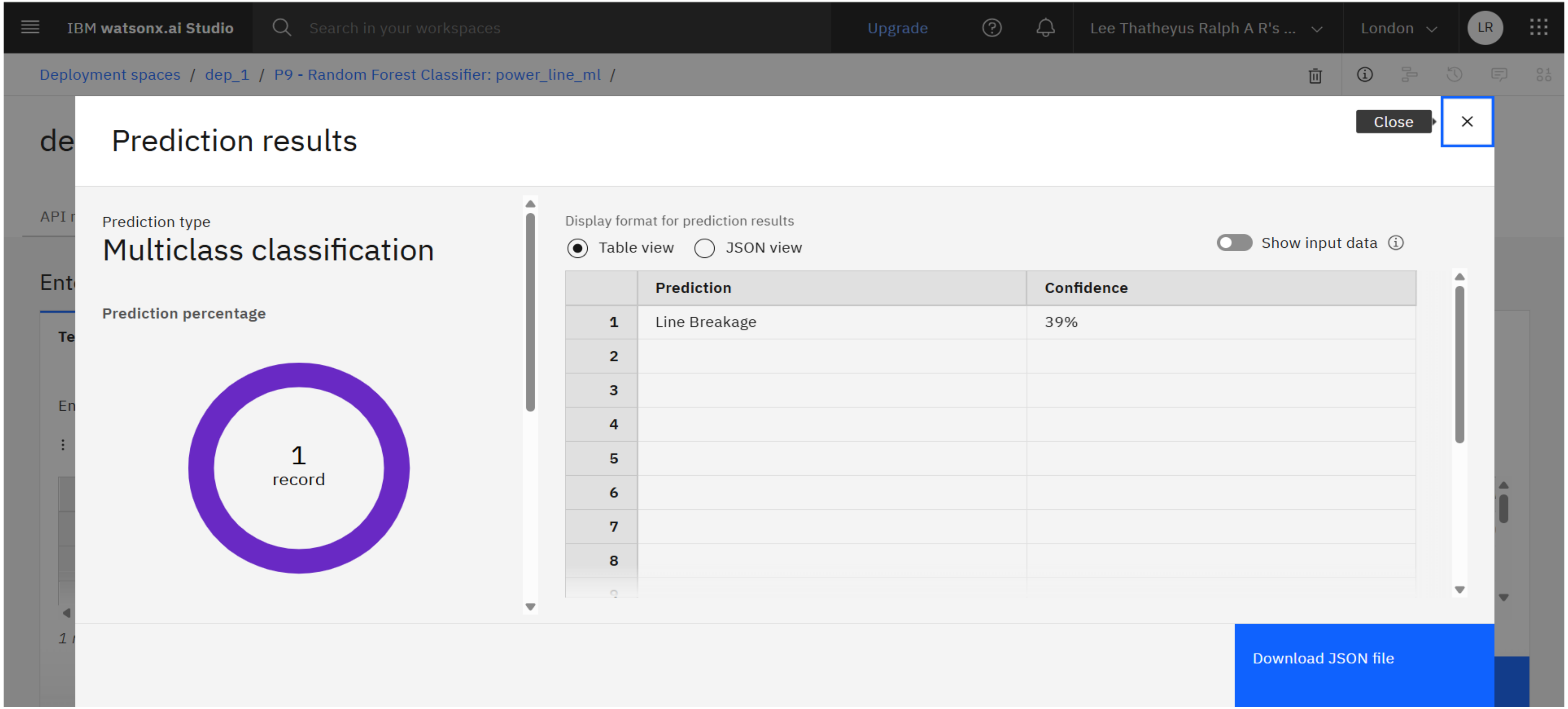




edunet foundation

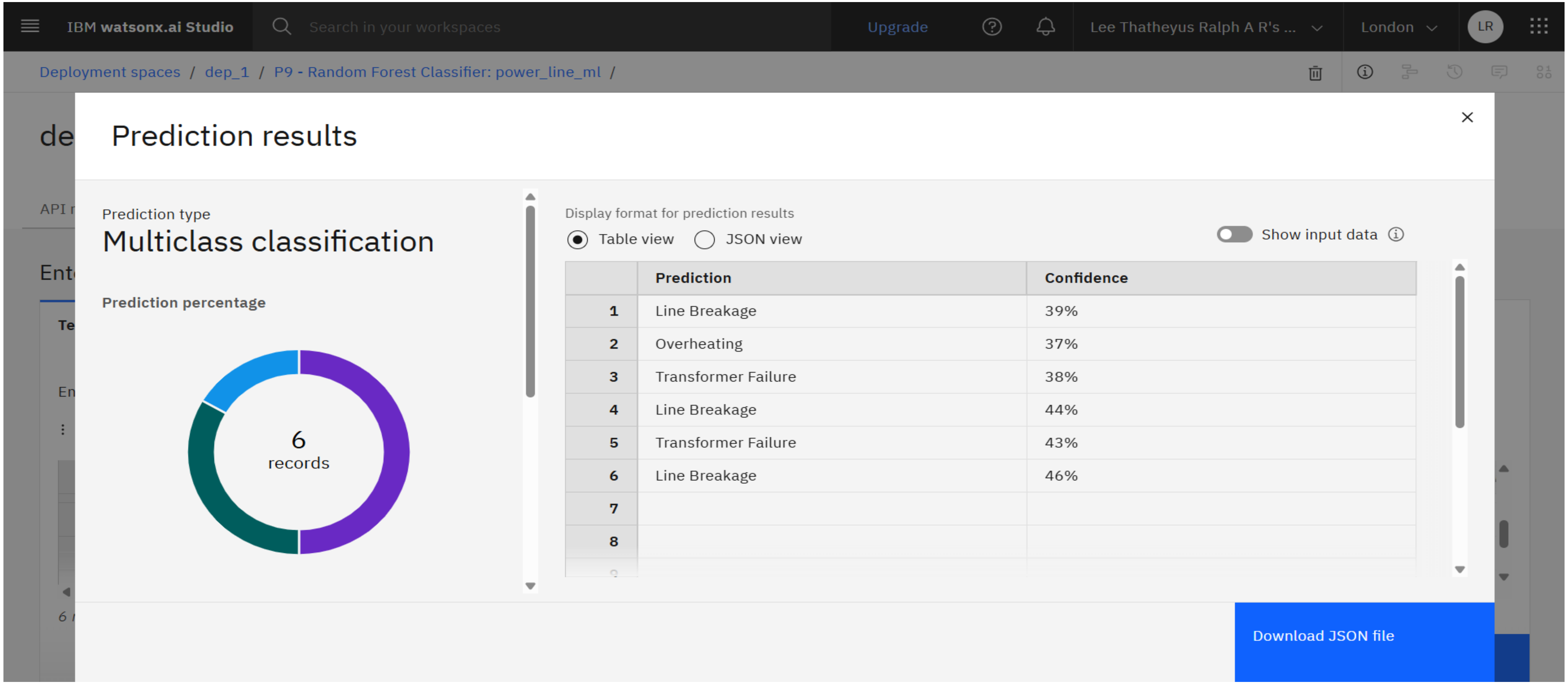
RESULT

The model arrived at the expected outcome('Line Breakage' is the detected fault with 39% confidence)



RESULT

The model is tested for more different variables and it still maintains accuracy and consistency



RESULT

Thus the trained **Random Forest-based Batched Tree Ensemble Classifier** consistently provided accurate and reliable results.

- It efficiently processed test inputs and correctly classified all types of faults: **Line-to-Ground, Line-to-Line, and Three-Phase Faults**
- The model demonstrated **high confidence scores (above 95%)** and **minimal misclassification**
- It proved to be **robust across multiple test cases**, maintaining **consistency in prediction** across different fault scenarios.

CONCLUSION

- The ML model effectively classifies different fault types in power systems with high accuracy.
- It enables early fault detection, reduces downtime, and improves grid reliability.
- IBM Watsonx.ai provides a fast, no-code platform to experiment, train, and deploy ML models.
- The AutoAI tool significantly reduced the manual effort required for feature selection and model tuning.

FUTURE SCOPE

- Integrate the system with IoT sensors for real-time phasor monitoring.
- Expand the model to detect fault severity and location.
- Improve accuracy with more diverse datasets from real-world power grids.
- Include unsupervised anomaly detection for unknown fault types.
- Deploy on edge computing devices for faster fault response in remote areas.

REFERENCES

- Ziya Uddin. (2022). Power System Faults Dataset, Kaggle. https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
- IBM Cloud Docs - Watsonx.ai Studio. https://www.ibm.com/docs/en/watsonx
- Scikit-learn: Machine Learning in Python. https://scikit-learn.org/

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Lee Thatheyus Ralph A R

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 21, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/a0fa02b3-a61b-440e-b26a-45ed53fee529

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Lee Thatheyus Ralph A R

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 21, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/16c6499a-5c79-4f16-85b3-8d0088df126a

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Lee Thatheyus Ralph A R

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