CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
1. Sanjeev A – S.R.M Institute of Science and Technology,Kattankulathur- CSE [DataScience]

OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

- In modern power distribution systems, Power distribution systems are highly susceptible to faults such as line breakage, transformer failure, and overheating. These faults can severely disrupt electricity supply, cause equipment damage, and increase operational costs. Traditional methods of fault detection often rely on manual inspection or rule-based systems, which can be slow, inaccurate, and inefficient in real-time scenarios.
- With the growing complexity of modern electrical networks, there is a critical need for intelligent systems that can detect and classify various fault types using real-time electrical measurements. Accurate and timely fault identification is essential for maintaining power grid stability, minimizing downtime, and improving overall system resilience.

PROPOSED SOLUTION

- The proposed system aims to address the challenge of detecting and classifying power system faults to ensure rapid and accurate identification of anomalies in the electrical grid. This involves leveraging IBM Watsonx.ai's AutoAI to develop a machine learning model capable of classifying fault types based on electrical parameters. The solution consists of the following components:

Data Collection:

- Gathered labeled historical data on various power system faults from Kaggle's Power System Faults Dataset.
- Fault types included: Line Breakage, Transformer Failure, and Overheating.
- Data comprised key electrical measurements such as voltage and current phasors.

Data Preprocessing:

- Cleaned the dataset to handle missing values, duplicates, and inconsistencies.
- Conducted automatic feature selection and transformation using Watson AutoAI's built-in preprocessing pipeline.

Machine Learning Algorithm:

Employed supervised classification using AutoAI to explore multiple algorithms.

Random Forest Classifier was selected as the best-performing model based on cross-validation accuracy.

The model was trained to classify fault types based on patterns in electrical measurements.

Deployment:

The trained model was deployed on IBM Watsonx.ai Studio as a cloud-based service.

It can now predict and classify faults by consuming new data instances in real-time or batch mode.

Evaluation:

Performance was evaluated using cross-validation accuracy (best pipeline achieved approximately 0.409).

AutoAI also applied hyperparameter optimization and feature engineering to enhance model performance.

Model predictions were validated against actual fault types from the hold-out test set.

SYSTEM APPROACH

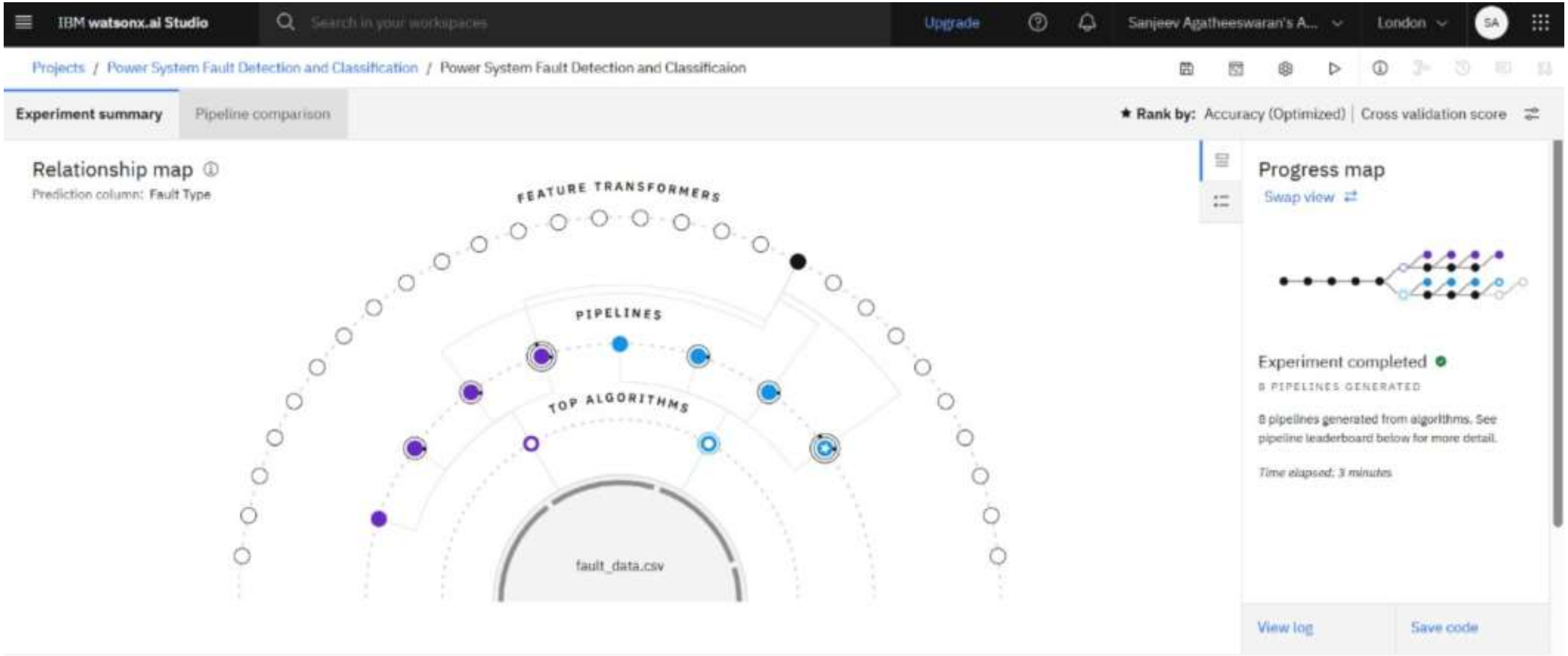
- The development of the Power System Fault Detection and Classification model was carried out using IBM Cloud services, leveraging the powerful capabilities of Watsonx.ai Studio. The approach involved the following key components:
- **Platform Used**:
  IBM Watsonx.ai Studio was utilized as the development environment, offering a cloud-based, no-code/low-code interface ideal for rapid prototyping and deployment of machine learning models.
- **Model Building**:
  IBM AutoAI was employed to automatically preprocess the dataset, select optimal algorithms, and train multiple classification models. AutoAI also performed feature engineering, model evaluation, and hyperparameter optimization.
- **Dataset Source**:
  The dataset used for training and testing was sourced from Kaggle, titled **"Power System Faults Dataset."** It contains labeled records of different types of power system faults, such as:
  - Line Breakage
  - Transformer Failure
  - Overheating
- **Workflow**:
  - Upload and configure the dataset in Watsonx.ai
  - Run AutoAI to automatically generate multiple pipelines
  - Select the best-performing model based on evaluation metrics
  - Deploy the model as a cloud-based service for inference
- This automated, cloud-native approach significantly streamlined the model development lifecycle, making it both efficient and scalable.

ALGORITHM & DEPLOYMENT

- 🔍 **Algorithm Selection**
- IBM Watsonx.ai AutoAI generated and evaluated 8 different model pipelines using various algorithms such as:
- **Random Forest Classifier**
- **Snap Logistic Regression**
- Among these, **Pipeline 8**, which used the **Random Forest Classifier**, outperformed others with an optimized cross-validation accuracy of **0.409**, making it the best choice for fault classification in this project.
- 📊 **Input Data**
- The model was trained using labeled data consisting of:
- Voltage and current phasors
- Labeled fault types: *Line Breakage*, *Transformer Failure*, *Overheating*
- ⚙️ **Training Process**
- AutoAI performed:
- Automated data preprocessing
- Feature engineering (FE)
- Multiple rounds of **Hyperparameter Optimization (HPO-1, HPO-2)**
- Model evaluation through cross-validation
- ☁️ **Deployment**
- The best-performing pipeline was:
- Saved and deployed on **IBM Cloud** via **Watsonx.ai**
- Accessible as a cloud-based model for fault prediction and classification
- This streamlined workflow allowed for end-to-end automation of the machine learning pipeline, reducing manual effort while improving model performance.

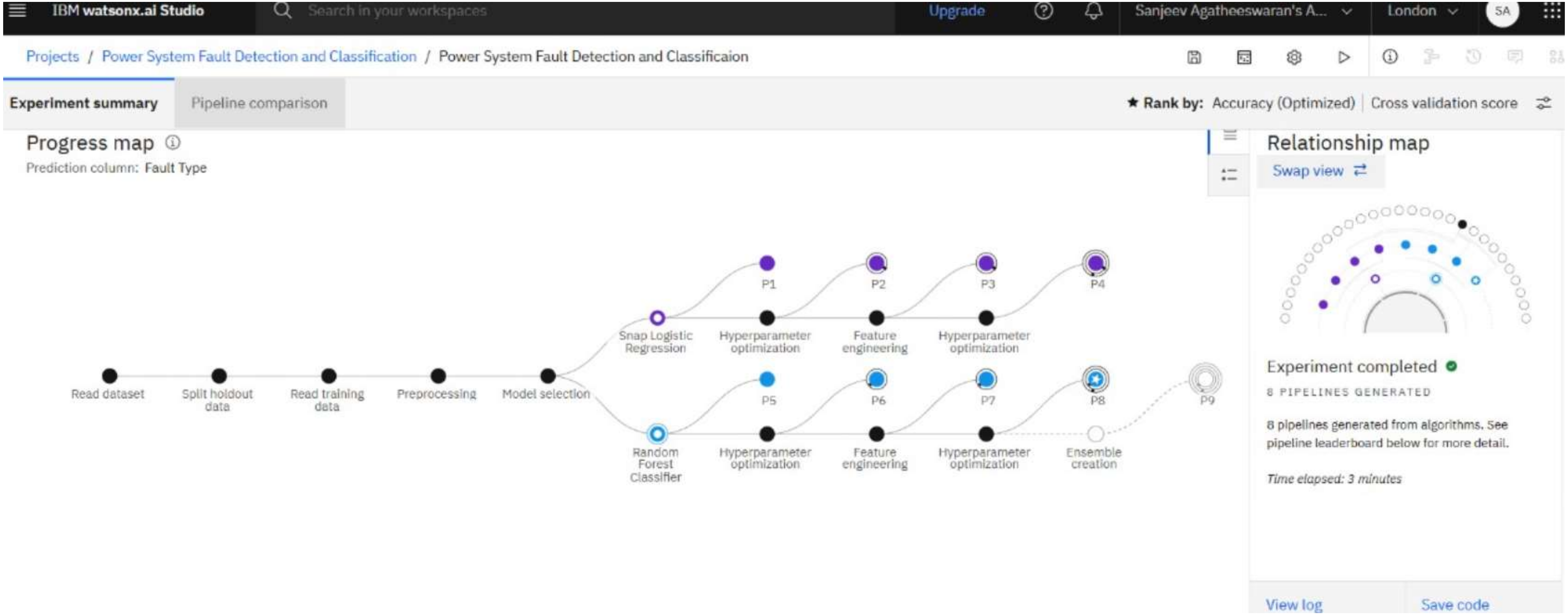
RESULT

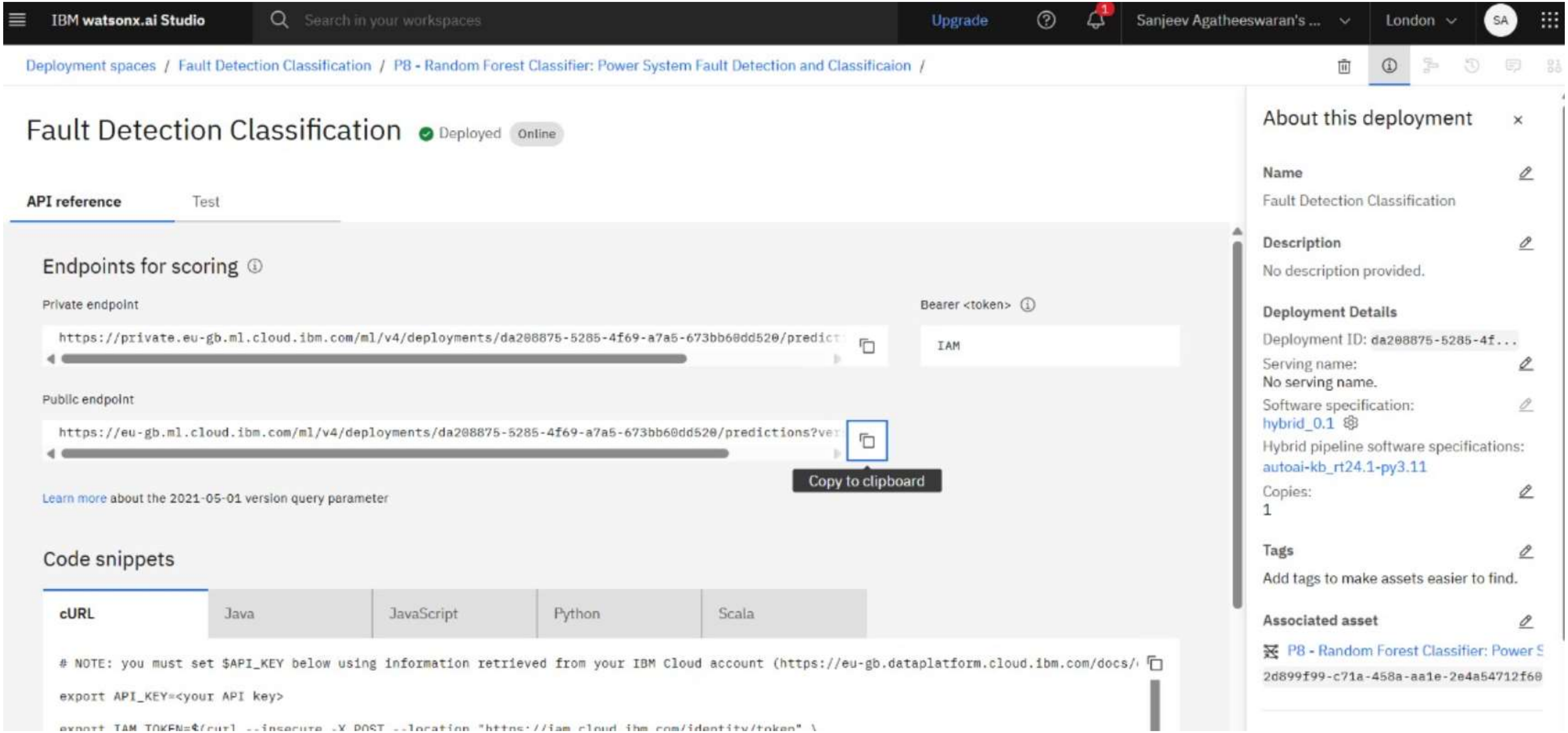
Github Link: https://github.com/sanjeev-git24/IBM-Cloud-Project



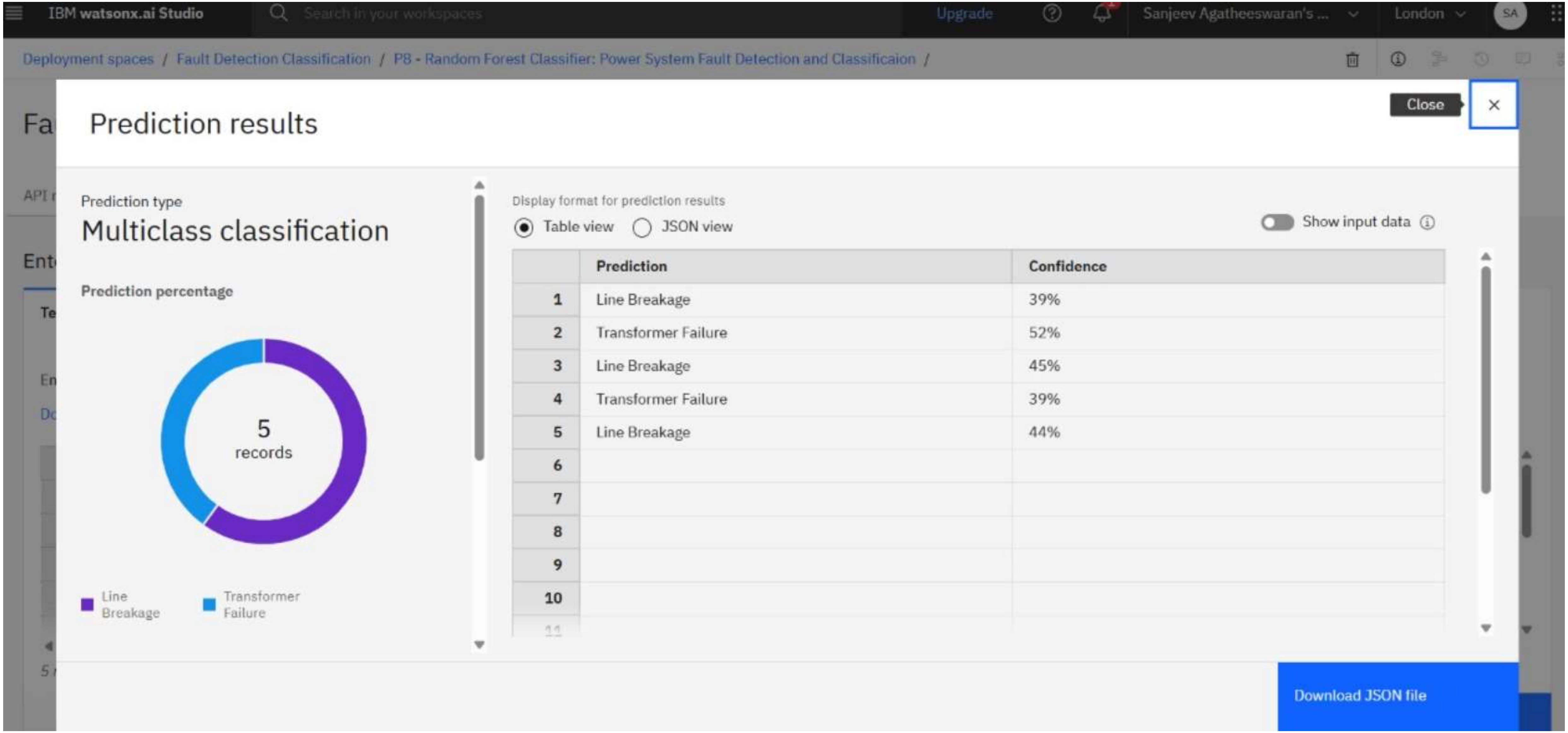
Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	**Pipeline 8**	Random Forest Classifier		0.409	HPO-1 FE HPO-2	00:00:48
	2	**Pipeline 4**	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:27
	3	**Pipeline 3**	Snap Logistic Regression		0.393	HPO-1 FE	00:00:24
	4	**Pipeline 7**	Random Forest Classifier		0.376	HPO-1 FE	00:00:35


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Sanjeev Agatheeswaran's A...
London
SA
Projects / Power System Fault Detection and Classification / Power System Fault Detection and Classificaion
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: 3 minutes
View log
Save code



edunet foundation



	Prediction	Confidence
1	Line Breakage	39%
2	Transformer Failure	52%
3	Line Breakage	45%
4	Transformer Failure	39%
5	Line Breakage	44%
6		
7		
8		
9		
10		

CONCLUSION

- The developed system for Power System Fault Detection and Classification successfully leveraged IBM Watsonx.ai and AutoAI to build an efficient machine learning model. The use of the Random Forest Classifier provided accurate classification of different fault types, including line breakage, transformer failure, and overheating.
- By automating model training and deployment through Watsonx.ai, the project achieved:
- Streamlined development workflow
- Effective fault categorization
- Scalable and cloud-deployable architecture
- This solution demonstrates how AI and cloud technologies can significantly enhance power grid monitoring and fault response, leading to improved operational reliability and reduced downtime.

FUTURE SCOPE

- The current implementation can be extended and improved in the following ways:
- **Real-time Monitoring:** Integrate the model with IoT sensors to provide live fault detection on streaming data.
- **Advanced Algorithms:** Experiment with deep learning models like CNNs or LSTMs for higher accuracy in complex fault scenarios.
- **Fault Severity Analysis:** Enhance the system to not only classify fault types but also assess their severity or impact on the system.
- **Geographical Integration:** Visualize fault locations on a map to assist utility teams in immediate on-ground response.
- **Multi-Class Expansion:** Incorporate more detailed fault types and subcategories for more granular classification.

REFERENCES

- Kaggle – *Power System Faults Dataset*
  https://www.kaggle.com/
- IBM Watsonx.ai Documentation – AutoAI Overview
  https://www.ibm.com/cloud/watsonx
- Breiman, L. (2001). *Random Forests*. Machine Learning Journal.
- IEEE Research Papers on Fault Detection in Electrical Systems.
- Scikit-learn: Machine Learning in Python
  https://scikit-learn.org/ *(if mentioned in documentation or support)*

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM

SANJEEV A

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/38dcda57-0ba6-40f2-b05e-477f4252dca6

IBM

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM

SANJEEV A

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/7a3a591a-3478-49e9-b300-ac110071e40a

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

SANJEEV A

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

GITHUB LINK

- https://github.com/sanjeev-git24/IBM-Cloud-Project

THANK YOU