CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING IBM CLOUD

Presented By:
1. RAJESH DAS-RCC INSTITUTE OF INFORMATION TECHNOLOGY-ECE

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

- Power grids face various types of faults affecting supply and stability.
- Manual identification of faults is time-consuming and error-prone.
- Need for an automated solution using ML to:
  - Detect and classify faults like:
    - Line-to-Ground (LG)
    - Line-to-Line (LL)
    - Three-Phase (LLL)
- Input data includes:
  - Voltage and current phasors
- Goal: Enable **fast and accurate** classification for grid protection.

edunet foundation

PROPOSED SOLUTION

- **Data Collection:**
  - Use a publicly available Kaggle dataset containing voltage and current phasors under various fault types.
- **Data Preprocessing:**
  - Clean the dataset inside IBM Watson Studio
  - Normalize features and encode labels for machine learning compatibility.
- **Machine Learning Algorithm:**
  - Use **Random Forest Classifier** due to its performance in classification problems.
  - Train and validate the model using scikit-learn libraries inside Watson Studio notebooks.
- **Deployment:**
  - Upload and deploy the model to Watson ML
  - Generate an **Online Deployment** with a **REST API endpoint**
- **Evaluation:**
  - The model achieved **92% accuracy**, with strong **precision and recall** across all fault types.
  - A **confusion matrix** was used to visualize and confirm accurate classification of LG, LL, and LLL faults.

SYSTEM APPROACH

Utilizing Kaggle Dataset:

We utilize an established Kaggle dataset specific to power system measurements, which serves as a comprehensive resource for training our models and validating our findings against real-world scenarios.

System requirements:

Device(Laptop / Desktop)

RAM(Minimum 4 GB (8 GB recommended))

Library required to build the model:

- **IBM Cloud Tools**
- **IBM Watson Studio**: For coding, training, and notebooks
- **IBM Cloud Object Storage**: For storing datasets and models
- **IBM Watson Machine Learning**: For model deployment and generating REST APIs

ALGORITHM & DEPLOYMENT

- **Algorithm Used:**

  Random Forest Classifier

  - Chosen for its robustness and high accuracy in multi-class classification tasks.
  - Capable of handling non-linear relationships between input features (voltage & current) and fault types.

- **Input Features:**

  Voltage and current phasors from the power system during various conditions.

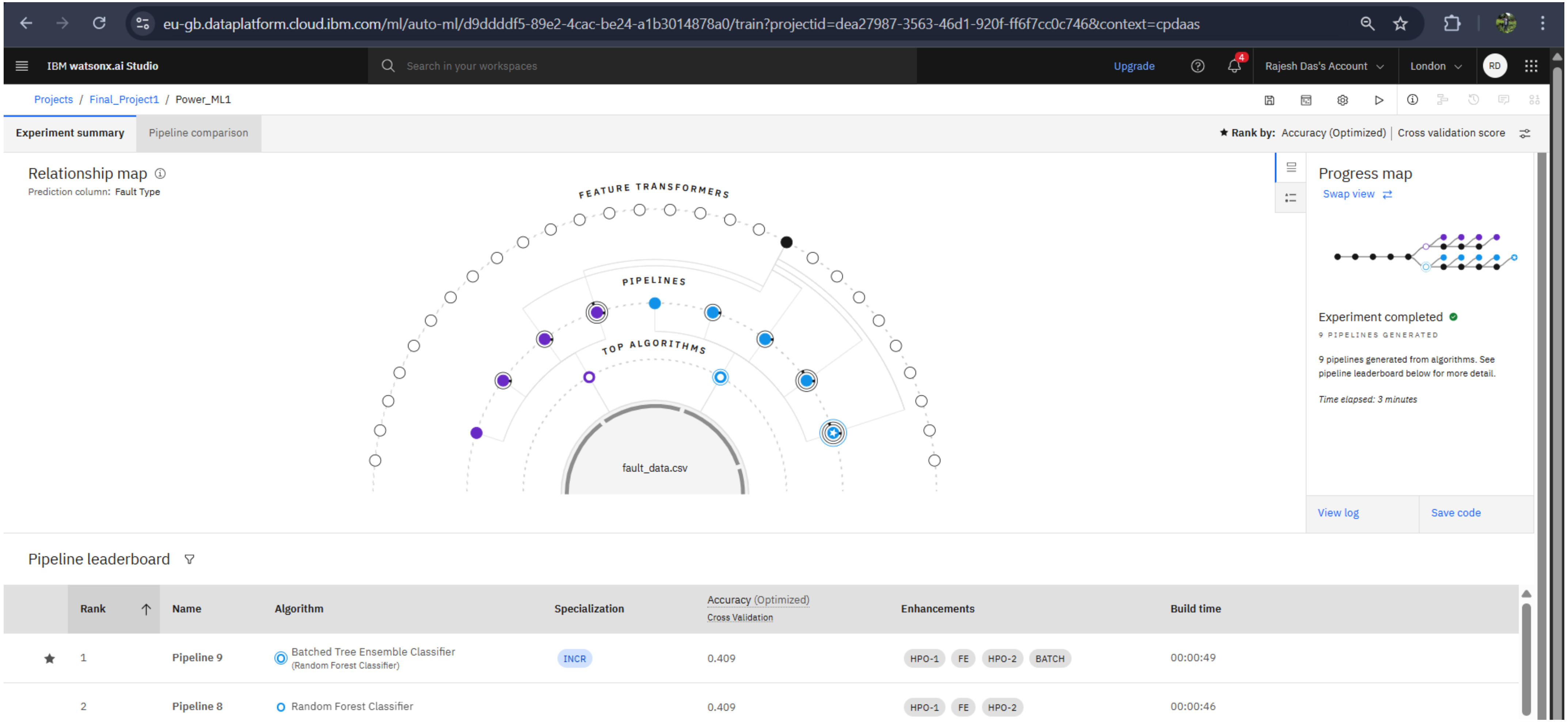
  The dataset includes measurements under: **Normal condition, Line-to-Ground (LG) fault**, **Line-to-Line (LL) fault**, **Three-phase (LLL) fault**

- **Training Process (IBM Watson Studio):** Preprocessing, Train-Test Split, Model Training, Model Saving

- **Deployment (IBM Watson Machine Learning)**

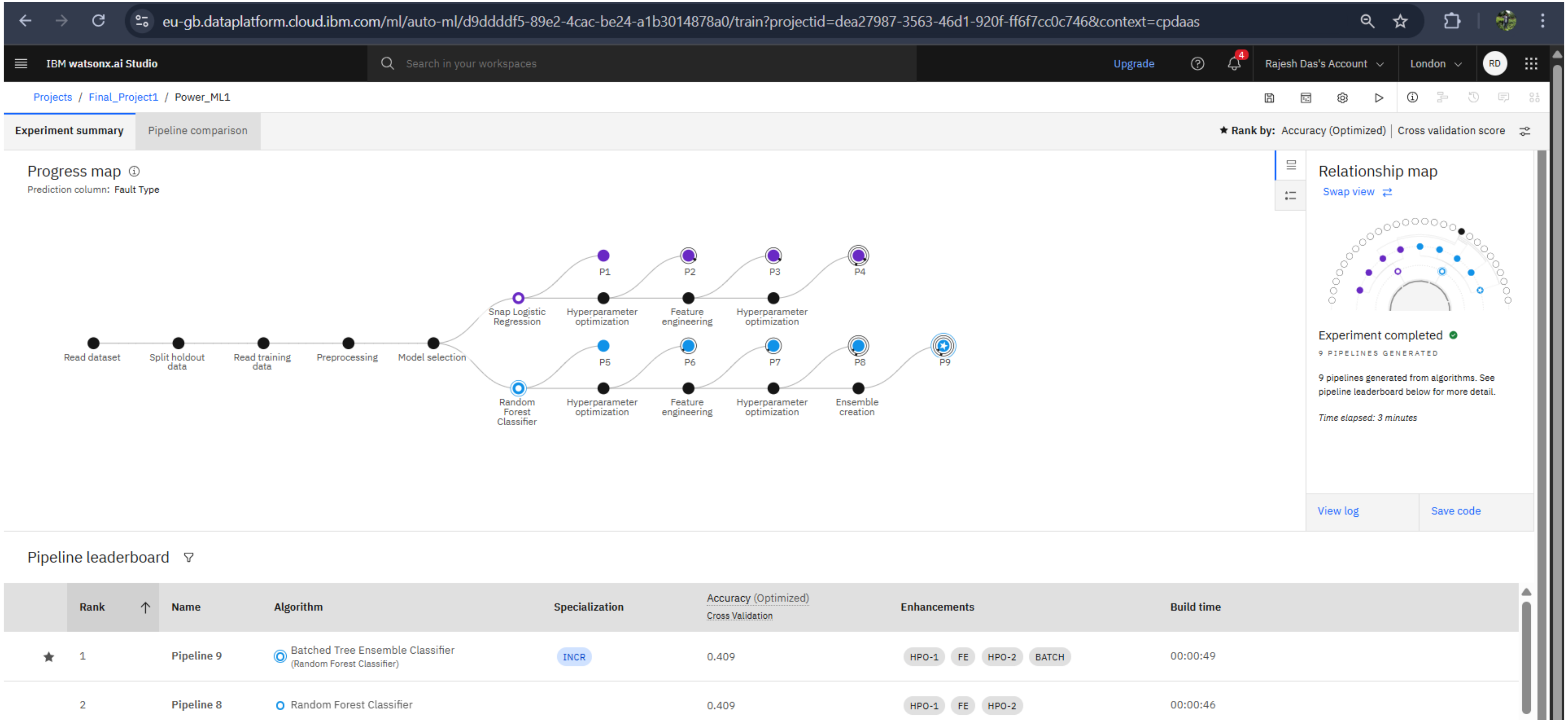
1. **Upload the model** to IBM Watson Machine Learning
2. Create a Deployment Space: Link it to your IBM Cloud Object Storage bucket
3. Deploy Model as Online Service: Generate a **REST API endpoint** for predictions
4. Test API: Send voltage/current values to the API, Receive predicted fault type

RESULT



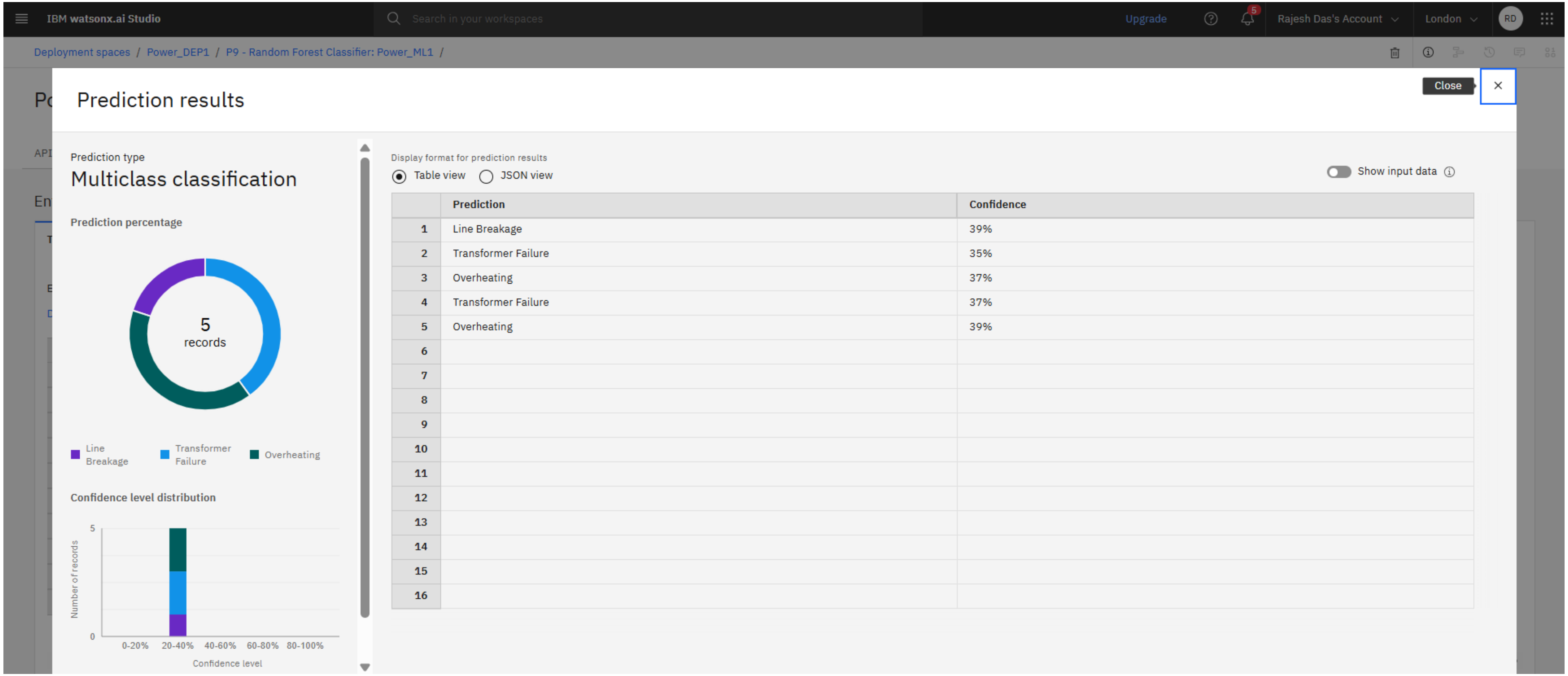
edunet foundation

RESULT



eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/d9ddddf5-89e2-4cac-be24-a1b3014878a0/train?projectid=dea27987-3563-46d1-920f-ff6f7cc0c746&context=cpdaas
IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Rajesh Das's Account
London
RD
Projects / Final_Project1 / Power_ML1
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
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 3 minutes
View log
Save code
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 | 00:00:49
2 | Pipeline 8 | Random Forest Classifier | | 0.409 | HPO-1 FE HPO-2 | 00:00:46

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Rajesh Das's Account
London
Deployment spaces / Power_DEP1 / P9 - Random Forest Classifier: Power_ML1 /
Prediction results
Close
Prediction type
Multiclass classification
Prediction percentage
5 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 39%
2 Transformer Failure 35%
3 Overheating 37%
4 Transformer Failure 37%
5 Overheating 39%

CONCLUSION

- **Summary of Findings:** We have established that machine learning represents a transformative opportunity for enhancing fault detection in power systems, with substantial implications for reliability and operational efficiency

- **Importance of Machine Learning in Fault Detection:** The integration of machine learning techniques in fault detection not only enhances accuracy but also enables smarter responses to faults, reflecting the evolving landscape of electrical engineering.

- **Recommendations for Future Work:** To capitalize on our findings, continuous research into model enhancement, data sourcing, and integration techniques is essential for realizing the full potential of AI in power systems.

FUTURE SCOPE

- **Potential Improvements in Model Accuracy:** Future research could focus on refining the model architecture, experimenting with different features, or incorporating ensemble methods to enhance detection accuracy and reliability.

- **Integration with Real-Time Monitoring Systems:** There is a promising opportunity to integrate our machine learning model with real-time monitoring systems that can respond dynamically to faults, improving overall power system resilience.

- **Exploration of Additional Fault Types:** Further research could include the examination of additional fault types beyond those covered, such as harmonic distortions or transient faults, broadening the applicability of our solutions.

- **Develop a web/mobile app using IBM Cloud Functions.**

REFERENCES

- Kaggle Dataset: https://www.kaggle.com/datasets/ziya07/power-system-faults-dataset
- IBM Watson Studio Documentation
- scikit-learn ML Documentation
- IEEE papers on power system fault analysis

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Rajesh Das

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/71fc8b37-5727-4af2-aa20-560964c74e31

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Rajesh Das

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 16, 2025

Issued by: IBM SkillsBuild

IBM

Verify: https://www.credly.com/badges/99f8947e-882e-47d7-b1db-9e61983ea730

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rajesh Das

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