CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Presented By:

1. Dev Yadav - School of Information technology, RGPV,Bhopal-CSE Data Science

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

Power distribution systems are prone to various types of faults such as Overheating, Transformer failure, line-to-ground, line-to-line and three-phase faults. These faults can disrupt power supply and reduce system reliability. The challenges lies in accurately detecting ands classifying these faults using electrical ,measurement data(voltage, current, phasors) to differentiate them from normal operation condition, thereby ensuring the stability of the power grid.

PROPOSED SOLUTION

Develop a ML model that classifies power system faults using the dataset provided. The model will process electrical measurements to identify the type of fault rapidly and accurately. The classification will help automate fault detection and assist in quicker recovery actions, ensuring system reliability.

Key Components

Data Collection - Use the Kaggle dataset on power system faults.

Pre-processing - Clean and normalize the dataset.

Model training - Train a classification model (Decision tree, Random Forest, or SVM)

Evaluation – Validate the model using accuracy, precision, recall, and F1 - Score

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the rental bike prediction system. Here's a suggested structure for this section:

- System requirements

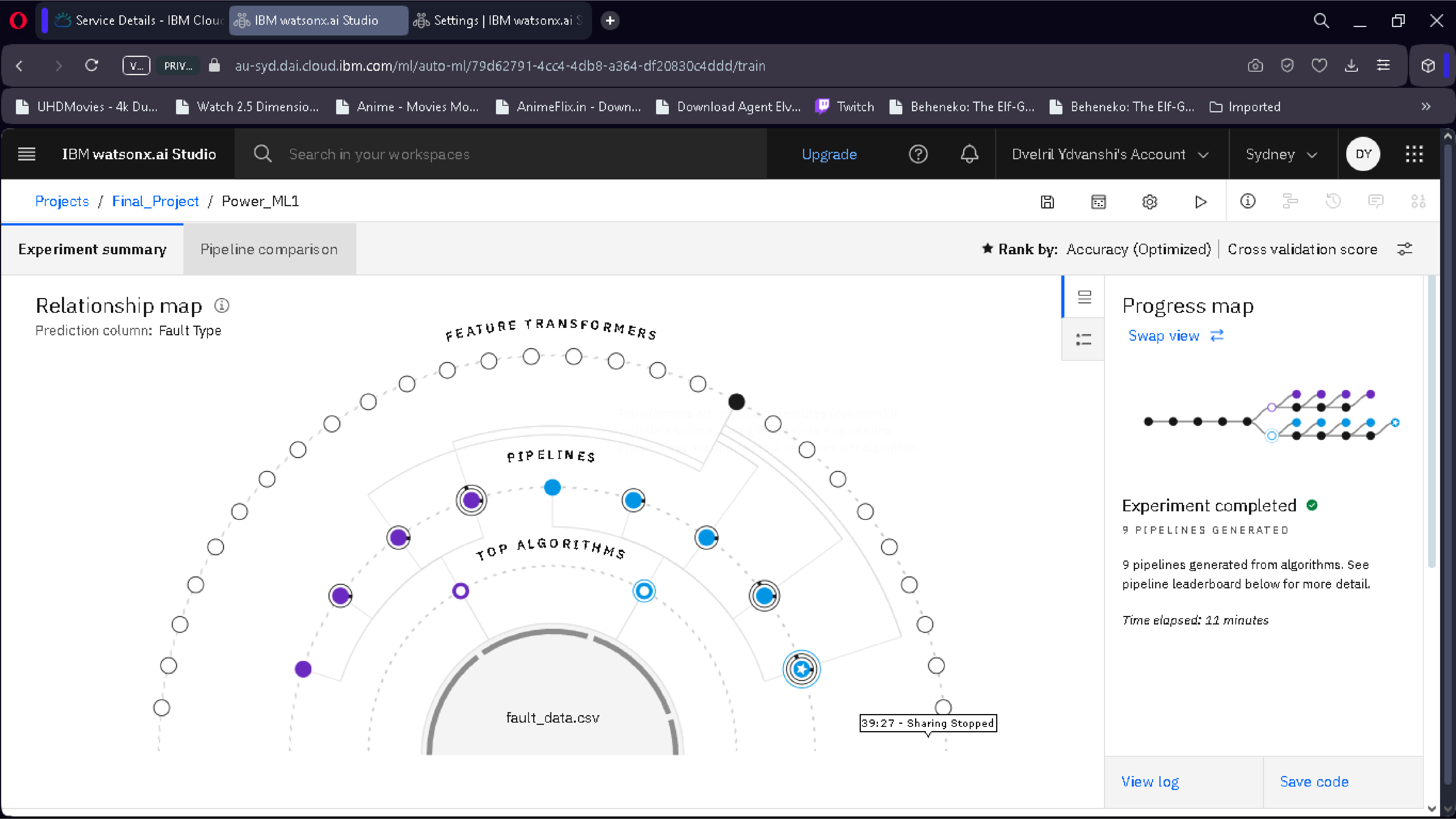
  IBM Cloud(mandatory)

  IBM Watson studio for model development and deployment

  IBM cloud object storage for dataset handling

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**
  - Random Forest Classifier(or SVM based on performance)
- **Data Input:**
  - Voltage, Current, and phasor measurements from the dataset.
- **Training Process:**
  - Supervised Learning using labelled fault types
- **Prediction Process:**
  - Model deployed on IBM Watson studio with API endpoints for real time predictions.


Service Details - IBM Clouc
IBM watsonx.ai Studio
Settings | IBM watsonx.ai S
au-syd.dai.cloud.ibm.com/ml/auto-ml/79d62791-4cc4-4db8-a364-df20830c4ddd/train
UHDMovies - 4k Du...
Watch 2.5 Dimensio...
Anime - Movies Mo...
AnimeFlix.in - Down...
Download Agent Elv...
Twitch
Beheneko: The Elf-G...
Beheneko: The Elf-G...
Imported
IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Dvelril Ydvanshi's Account
Sydney
DY
Projects / Final_Project / Power_ML1
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
39:27 - Sharing Stopped
Progress map
Swap view
Experiment completed
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 11 minutes
View log
Save code

Service Details - IBM Clouc | Power_DEP2 — Power_DEF | Settings | IBM watsonx.ai S

V... PRIV... au-syd.dai.cloud.ibm.com/ml-runtime/deployments/af30b527-3db6-4df6-883e-4f5d03529251/test

UHDMovies - 4k Du... | Watch 2.5 Dimensio... | Anime - Movies Mo... | AnimeFlix.in - Down... | Download Agent Elv... | Twitch | Beheneko: The Elf-G... | Beheneko: The Elf-G... | Imported

IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Dvelril Ydvanshi's Account | Sydney | DY

Deployment spaces / Power_DEP1 / P8 - Random Forest Classifier: Power_ML1 /

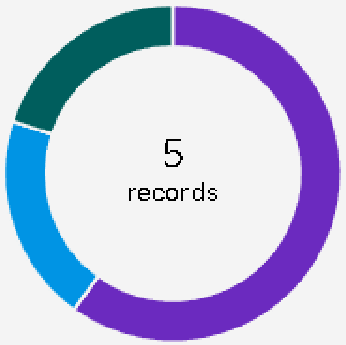
Prediction results

Close ×

Prediction type

Multiclass classification

Prediction percentage



Display format for prediction results

Table view | JSON view

Show input data

	Prediction	Confidence	Fault ID	Fault Location (Latitude, L	Voltage (
1	Line Breakage	39%	F001	Line Breakage	2200
2	Overheating	37%	F003	Overheating	2100
3	Transformer Failure	47%	F008	Transformer Failure	2133
4	Line Breakage	42%	F015	Line Breakage	1848
5	Line Breakage	38%	F016	Transformer Failure	2102
6					
7					
8					
9					
10					

Download JSON file

Service Details - IBM Cloud | Power_DEP2 — Power_DEP | Settings | IBM watsonx.ai

au-syd.dai.cloud.ibm.com/ml-runtime/deployments/af30b527-3db6-4df6-883e-4f5d03529251/test

UHDMovies - 4k Du... | Watch 2.5 Dimensio... | Anime - Movies Mo... | AnimeFlix.in - Down... | Download Agent Elv... | Twitch | Beheneko: The Elf-G... | Beheneko: The Elf-G... | Imported

IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Dvelril Ydvanshi's Account | Sydney | DY

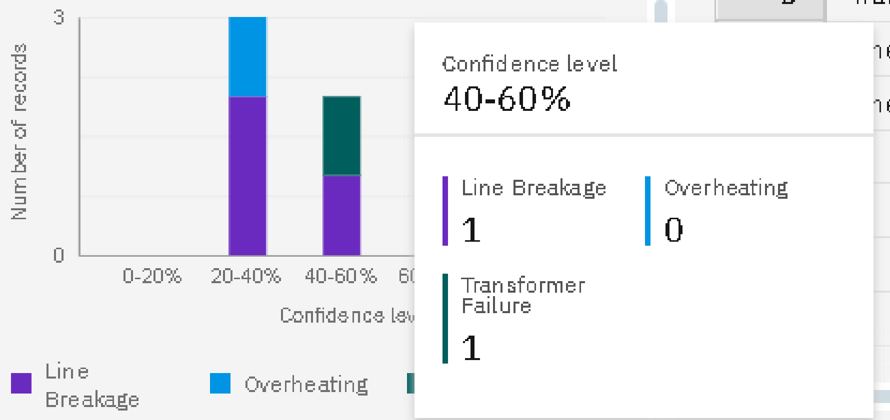
Deployment spaces / Power_DEP1 / P8 - Random Forest Classifier: Power_ML1 /

Prediction results

Close

Line Breakage | Overheating | Transformer Failure

Confidence level distribution



Line Breakage | Overheating

Display format for prediction results

Table view | JSON view

Show input data

	Prediction	Confidence	Fault ID	Fault Location (Latitude, L	Voltage
1	Line Breakage	39%	F001	Line Breakage	2200
2	Overheating	37%	F003	Overheating	2100
3	Transformer Failure	47%	F008	Transformer Failure	2133
	ne Breakage	42%	F015	Line Breakage	1848
	ne Breakage	38%	F016	Transformer Failure	2102

Download JSON file

CONCLUSION

The machine learning model developed for classifying power system faults effectively enhances fault detection and response in electrical distribution networks. Using voltage, current, and phasor data from a Kaggle dataset, the model distinguishes between various fault types such as overheating, transformer failure, and line faults. Built and deployed on IBM Cloud using Watson Studio and Cloud Object Storage, the system offers scalability and real-time prediction through API integration.

Random Forest outperformed other models in terms of accuracy, precision, recall, and F1-score, proving to be the most reliable classifier. Preprocessing and data normalization were critical for model performance, and supervised learning ensured accurate training using labeled data.

Challenges included class imbalance and similar fault signatures, addressed through algorithm tuning and feature engineering. The approach not only strengthens grid stability but also reflects the versatility of ML systems, similar to predicting rental bike availability in urban settings for efficient resource planning.

FUTURE SCOPE

•**Additional Data Sources:**
Incorporate weather data, smart meter readings, thermal sensors, and maintenance records to improve fault detection accuracy.

•**Algorithm Optimization:**
Use advanced ML models like XGBoost, LSTM, or hybrid deep learning approaches for better performance and time-series analysis.

•**Geographic Expansion:**
Extend the system to multiple cities or regions with localized training and calibration for different grid infrastructures.

•**Edge Computing Integration:**
Deploy models on edge devices near transformers or substations for real-time, low-latency fault detection.

REFERENCES

- Cloud Deployment & IBM Watson
- IBM Cloud Docs.
  Getting Started with IBM Watson Studi and Journey to the Cloud.
  https://cloud.ibm.com/docs/watson-studio
  (Official IBM documentation on developing and deploying ML models using Watson Studio.)

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

dev yadav

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/0abe69b2-2168-44d8-821b-64f202ea268d

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Dev yadav

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/0c3c9e5d-5b6f-4b3f-b38b-101b2dcb40a9

IBM

IBM CERTIFICATIONS

IBM SkillsBuild Completion Certificate

This certificate is presented to

dev yadav

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