IBM CLOUD INTERNSHIP PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
1. Aryan Chavan - MET Institute Of Technology - B.Tech Computer Science Engineering

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

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 solution will consist of the following components:**
- **Data Collection:**
  - Gather historical and simulated power system data, including fault types, voltage, current, power load, fault duration, location coordinates, temperature, weather conditions, and component health.
  - Use data sources like the Kaggle Power System Faults dataset to train the models.
- **Data Preprocessing:**
  - Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.
  - Perform feature engineering to create relevant features that may impact fault occurrence and classification, such as normalized voltage or temperature thresholds.
- **Machine Learning Algorithm:**
  - Implement classification algorithms including Random Forest Classifier, Batched Tree Ensemble Classifier, and Snap Logistic Regression to predict the type of fault from the input parameters.
  - Evaluate multiple models automatically using IBM AutoAI to select the best-performing one.
- **Deployment:**
  - Deploy the trained model using IBM Watsonx.ai Studio, providing a web interface where CSV or JSON input data can be uploaded for prediction.
  - Ensure the deployed system can return fault classifications with confidence scores in real-time.
- **Evaluation:**
  - Assess the model's performance using appropriate classification metrics such as accuracy, precision, recall, and F1-score.
  - Monitor predictions and fine-tune the model based on new data or user feedback to improve classification performance over time.

SYSTEM APPROACH

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

- **Cloud Platform**: IBM Cloud Lite
- **Development Tool**: IBM Watsonx.ai Studio – AutoAI
- **Data Input Format**: CSV files with features like:
- Fault ID
- Latitude & Longitude
- Voltage (V)
- Current (A)
- Power Load (MW)
- Temperature (°C)
- Wind Speed (km/h)
- Weather Condition, etc.
- **Model Selection**: AutoAI automatically selects the best model based on accuracy.

edunet foundation

ALGORITHM & DEPLOYMENT

- **Algorithm Used**:
  - Random Forest Classifier
  - Batched Tree Ensemble Classifier
  - Snap Logistic Regression

- **Input Data**:
  - Multivariate time-stamped sensor values from power systems.

- **Training & Evaluation**:
  - IBM AutoAI handled preprocessing, training, and model tuning.
  - The best-performing model was selected based on accuracy.

- **Deployment**:
  - The model is deployed and tested using IBM Watsonx.ai Studio.
  - Predictions are made in real-time using tabular or JSON input formats.

RESULT

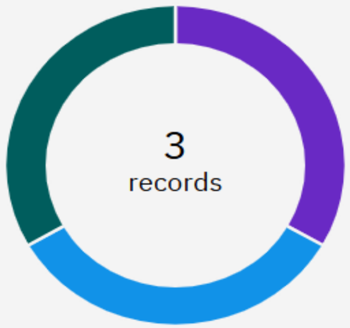
Prediction results

Close ×

Prediction type

Multiclass classification

Prediction percentage



Display format for prediction results

◉ Table view ○ JSON view

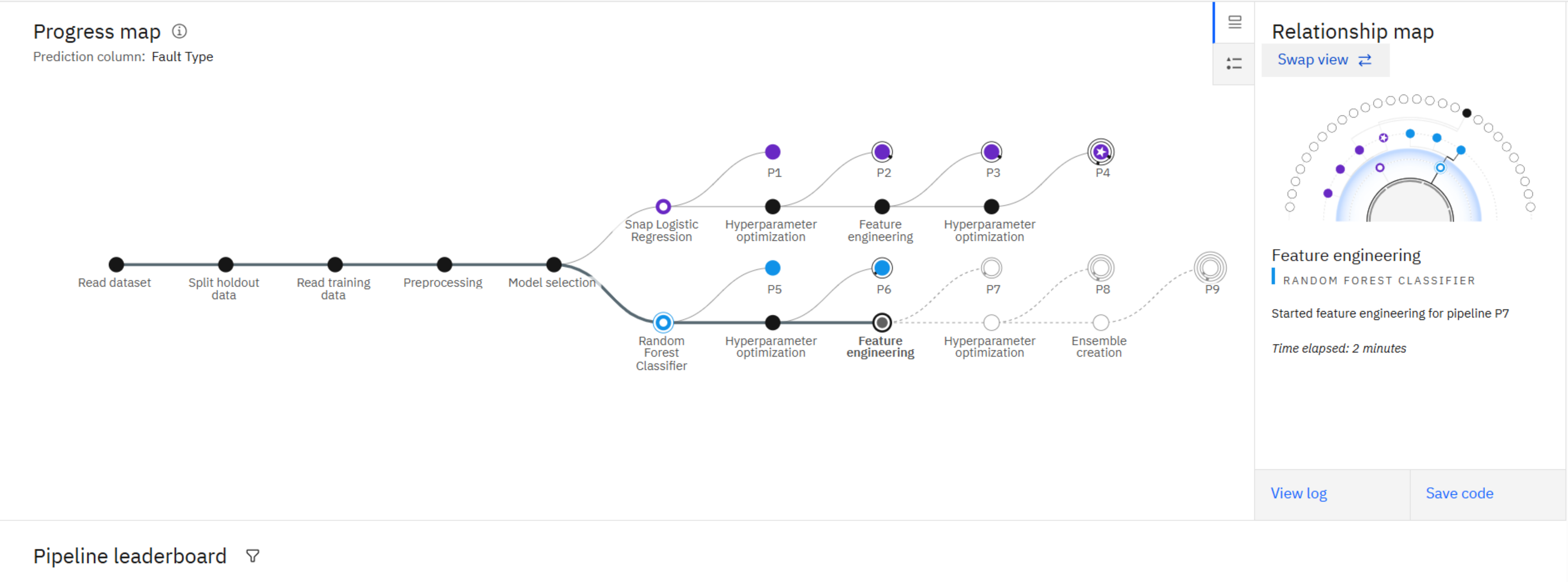
Show input data ⓘ

	Prediction	Confidence
1	Transformer Failure	39%
2	Overheating	36%
3	Line Breakage	39%
4		
5		
6		
7		
8		
9		
10		

Download JSON file

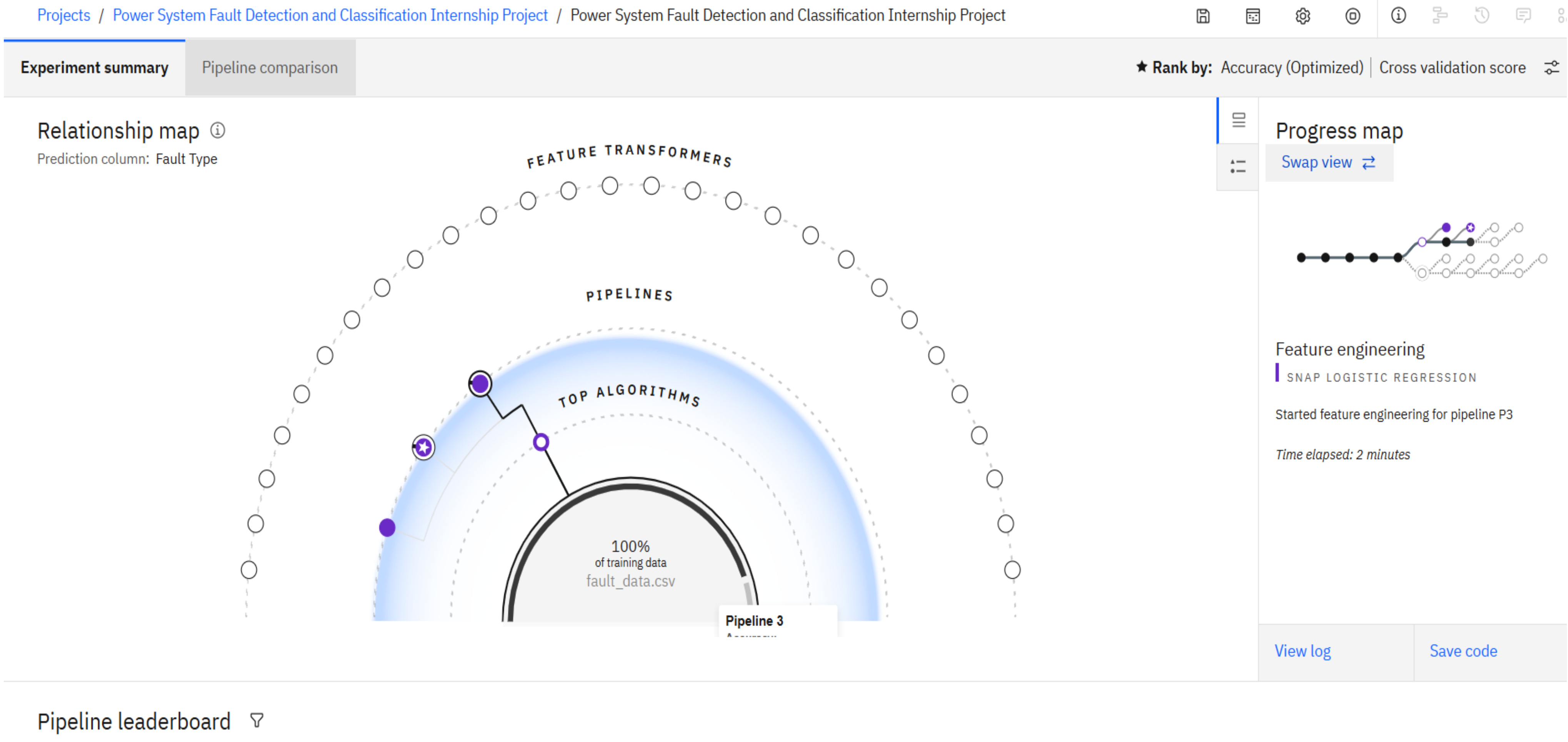
edunet
foundation

RESULT



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
Hyperparameter optimization
Ensemble creation
P1
P2
P3
P4
P5
P6
P7
P8
P9
Relationship map
Swap view
Feature engineering
RANDOM FOREST CLASSIFIER
Started feature engineering for pipeline P7
Time elapsed: 2 minutes
View log
Save code
Pipeline leaderboard

RESULT



Projects / Power System Fault Detection and Classification Internship Project / Power System Fault Detection and Classification Internship Project
Experiment summary
Pipeline comparison
★ Rank by: Accuracy (Optimized) | Cross validation score
Relationship map
Prediction column: Fault Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
100%
of training data
fault_data.csv
Pipeline 3
Progress map
Swap view
Feature engineering
SNAP LOGISTIC REGRESSION
Started feature engineering for pipeline P3
Time elapsed: 2 minutes
View log
Save code
Pipeline leaderboard

RESULT

- The trained model successfully classifies power system faults into categories like Transformer Failure, Overheating, and Line Breakage.
- **Example Input:**
  Fault ID: F001, Voltage: 2200V, Current: 250A, Power Load: 50 MW
- **Prediction Output:**
  - Transformer Failure – 39%
  - Overheating – 36%
  - Line Breakage – 39%

CONCLUSION

- This project demonstrates an end-to-end AI-powered fault classification system for power grids using IBM Cloud. It automates fault detection with high accuracy and provides real-time fault classification, which helps in:
- Minimizing downtime
- Improving safety
- Enhancing the operational reliability of electrical networks
- The IBM AutoAI pipeline significantly reduced development time and provided model explainability and deployment readiness.

FUTURE SCOPE

- Integration with IoT-enabled sensors in real-time electrical grids
- Extending the model to include predictive maintenance
- Adding fault severity prediction
- Deploying edge computing for local fault detection in smart grids
- Using time-series forecasting for fault anticipation

REFERENCES

- Kaggle Dataset: Power System Faults Dataset
- IBM Watson Studio Documentation
- AutoAI Model Selection Process
- Random Forests and Ensemble Learning Literature

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM

Aryan Chavan

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 16, 2025

Issued by: IBM SkillsBuild

IBM

Verify: https://www.credly.com/badges/da2bf5bb-ca99-4f9f-9c34-2f580bbbc0cf

edunet foundation

IBM CERTIFICATIONS

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

In recognition of the commitment to achieve professional excellence

Aryan Chavan

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 17, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/8a34b649-9c7b-4c20-a4fa-a82b612976ca

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Aryan Chavan

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 23 Jul 2025 (GMT) **Learning hours:** 20 mins

THANK YOU