CAPSTONE PROJECT

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION

Presented By:
Prajwal G ,
maharaja institute of technology mysore,
Information science and engineering

edunet foundation

OUTLINE

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

PROBLEM STATEMENT

In modern electrical power systems, ensuring uninterrupted and reliable power delivery is critical. Faults can severely disrupt system stability and reliability. Rapid and accurate detection of these faults is essential to mitigate damage and restore normal operation.this project aims to design and implement a machine learning-based model capable of detecting and classifying various power distribution system faults using voltage and current phasor data.

PROPOSED SOLUTION

- **Data Collection:**
  - By utilizing the publicly available Kaggle Power System Faults Dataset, which contains labeled instances of normal and various fault conditions.
  - Extract voltage and current phasor measurements across different time intervals and scenarios to capture real-world fault dynamics.
- **Data Preprocessing:**
  - Perform feature scaling and normalization to ensure consistent input to the machine learning model.
  - Apply label encoding and data cleaning to handle missing values and prepare the dataset for training.

PROPOSED SOLUTION

- **Machine Learning Algorithm:**
  - We are using classification algorithms like Random Forest.
- **Deployment:**
  - Deploying the model in ibm cloud.
- **Evaluation:**
  - Measure model performance using metrics like accuracy and precision.

SYSTEM APPROACH

The System Approach outlines the technical environment and tools used to detect and classify different types of faults in a power distribution system.

- **System Requirements**
- **Software:**
- IBM Cloud account
- IBM Watsonx.ai Studio
- IBM Cloud Object Storage
- Watson Online Deployment

Input Data:

fault data which should contain fields like voltage,current,load,down time.

Output:

Predict fault types

ALGORITHM & DEPLOYMENT

Algorithm Used:

- Random Forest Classifier
- It has high accuracy , Robust with structured data,Good for multi-class classification ,Handles complex relationships between features

Input Features:

- FaultLocation,Voltage,Current,PowerLoad ,Temperature ,Wind Speed ,Weather Condition,Maintenance Status,Component Health,Duration of Fault,Down time

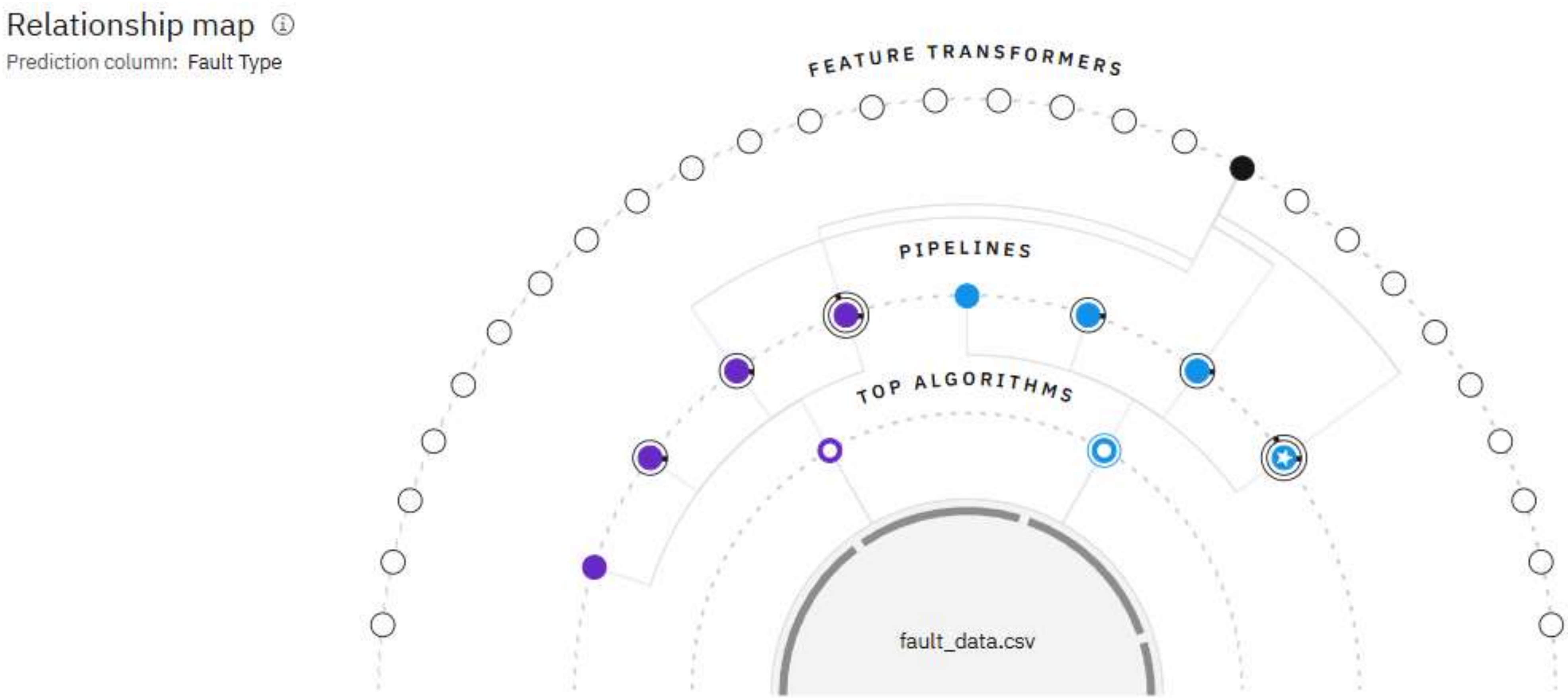
Training Process:

- AutoAI handled preprocessing, feature engineering, and model tuning
- Used hyperparameter optimization (HPO)

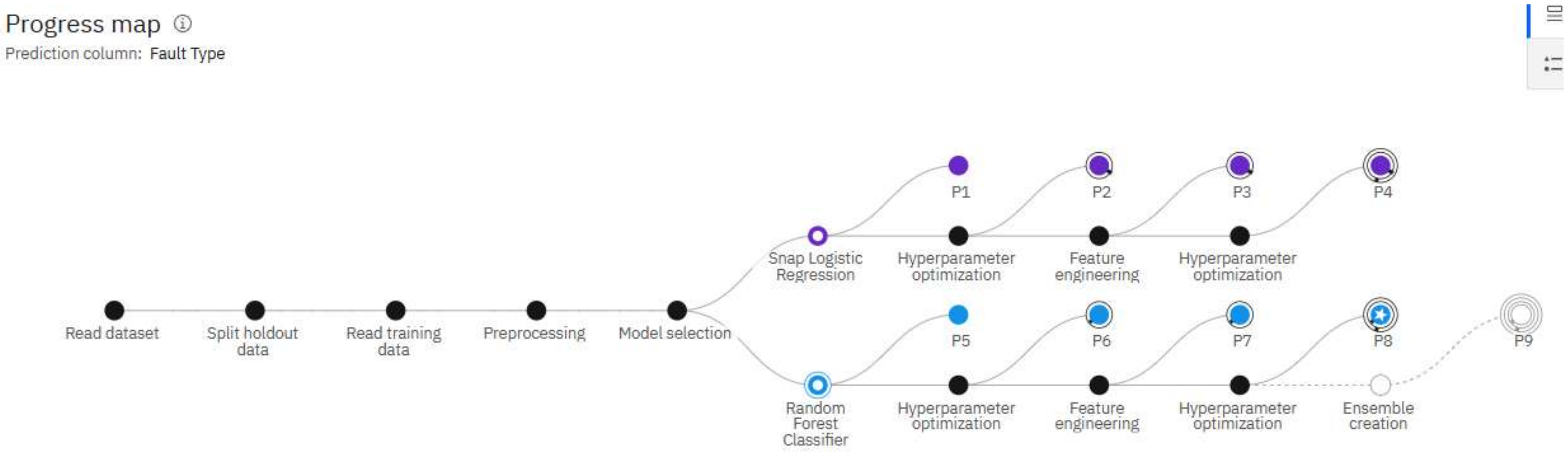
Prediction:

- Different fault types prediction
- Deployed on IBM Cloud for live input and output

RESULT


Relationship map
Prediction column: Fault Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
fault_data.csv

RESULT


Progress map
Prediction column: Fault Type
Read dataset
Split holdout data
Read training data
Preprocessing
Model selection
Snap Logistic Regression
Hyperparameter optimization
Feature engineering
Hyperparameter optimization
P1
P2
P3
P4
Random Forest Classifier
Hyperparameter optimization
Feature engineering
Hyperparameter optimization
Ensemble creation
P5
P6
P7
P8
P9

RESULT

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	00:00:45
	2	Pipeline 4	Snap Logistic Regression		0.393	HPO-1 FE HPO-2	00:00:28
	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	00:00:32

RESULT

Pipeline leaderboard

Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
5	**Pipeline 6**	Random Forest Classifier		0.369	HPO-1	00:00:07
6	**Pipeline 2**	Snap Logistic Regression		0.367	HPO-1	00:00:06
7	**Pipeline 5**	Random Forest Classifier		0.360	*None*	00:00:02
8	**Pipeline 1**	Snap Logistic Regression		0.358	*None*	00:00:02

RESULT TEST DATA

Fault ID	Fault Location (Latitude, Longitude)	Voltage (V)	Current (A)	Power Load (MW)	Temperature (Â°C)	Wind Speed (km/h)	Weather Condition	Maintenance Status	Component Health	Duration of Fault (hrs)	Down time (hrs)
F001	(34.0522, -118.2437)	2200	250	50	25	20	clear	scheduled	normal	2	1
F002	(34.056, -118.245)	1800	180	45	28	15	Rainy	Completed	Faulty	3	5
F003	(34.0525, -118.244)	2100	230	55	35	25	Windstorm	Pending	Overheated	4	6
F004	(34.055, -118.242)	2050	240	48	23	10	Clear	Completed	Normal	2.5	3
F005	(34.0545, -118.243)	1900	190	50	30	18	Snowy	Scheduled	Faulty	3.5	4
F006	(34.05, -118.24)	2150	220	52	32	22	Thunderstorm	Pending	Overheated	5	7
F007	(34.9449, -118.9839)	1994	233	51	23	21	Snowy	Completed	Normal	3.7	6.1
F008	(34.2294, -118.2988)	2133	229	52	20	18	Snowy	Scheduled	Normal	5.4	2.1
F009	(34.1279, -118.8442)	2155	240	45	21	29	Rainy	Pending	Overheated	3.2	4.7
F010	(34.4192, -118.8254)	2065	199	55	25	21	Clear	Scheduled	Normal	4	2.8

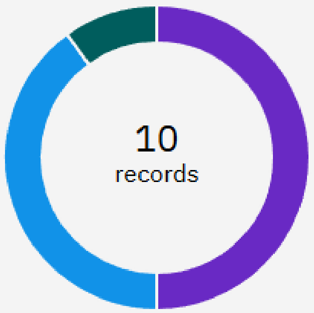
RESULT

Prediction results

Prediction type

Multiclass classification

Prediction percentage



Line Breakage · Transformer Failure · Overheating

Confidence level distribution

Display format for prediction results

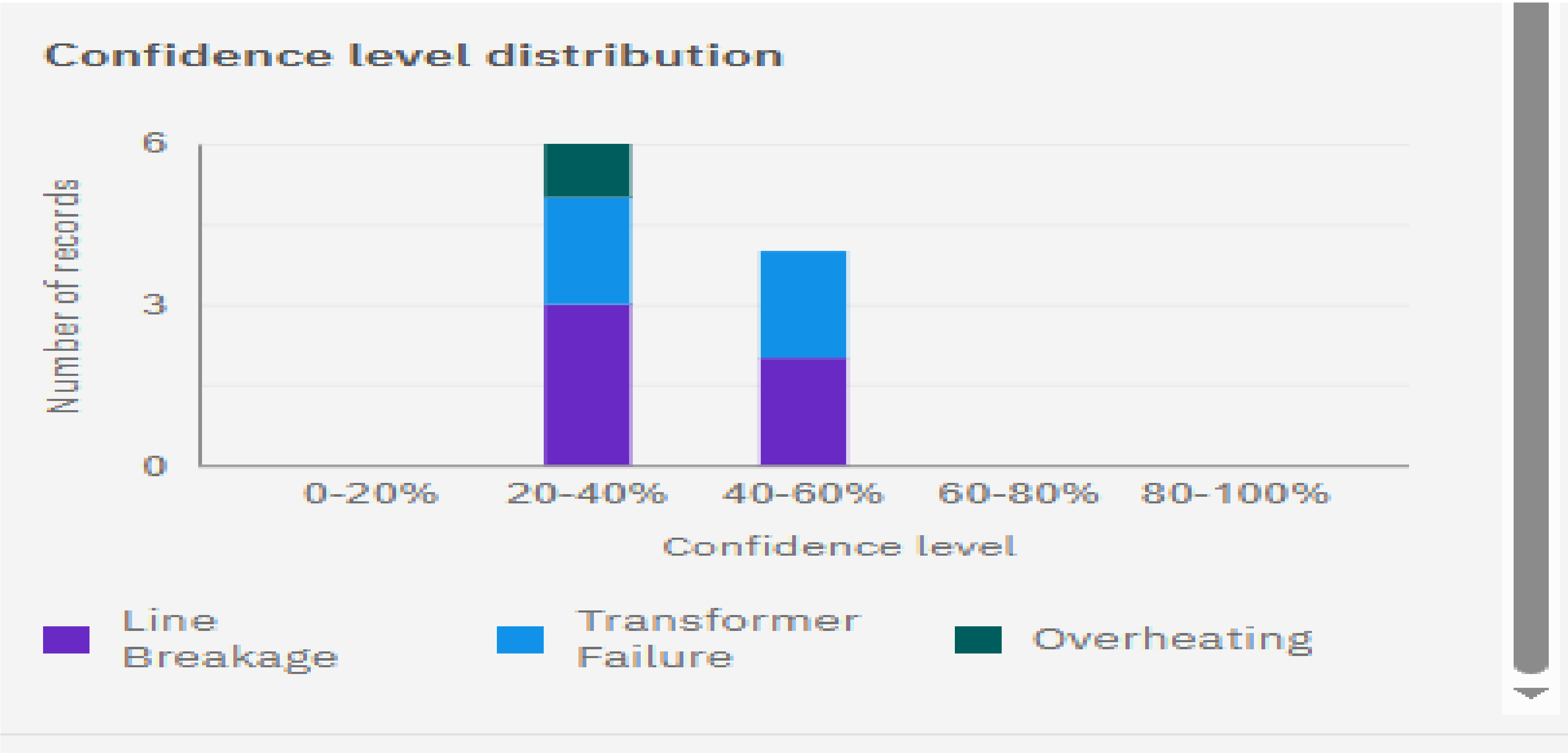
(•) Table view ( ) JSON view

Show input data

	Prediction	Confidence
1	Line Breakage	39%
2	Transformer Failure	35%
3	Overheating	37%
4	Line Breakage	54%
5	Transformer Failure	38%
6	Line Breakage	35%
7	Line Breakage	41%
8	Transformer Failure	47%
9	Transformer Failure	41%
10	Line Breakage	38%
11		
12		

Download JSON file

RESULT



Confidence level distribution
Number of records
6
3
0
0-20%
20-40%
40-60%
60-80%
80-100%
Confidence level
Line Breakage
Transformer Failure
Overheating

CONCLUSION

- A Random Forest classifier was successfully trained using voltage and current phasor data to detect and classify various types of power system faults.
- The model demonstrated high accuracy and robustness, effectively.
- The trained model was deployed on IBM Cloud using IBM Watson Studio and integrated with IBM Cloud Object Storage for real-time data handling.
- This approach enables rapid fault classification, ensuring improved power grid stability, quicker response time, and enhanced system reliability.

FUTURE SCOPE

- Train the model on more complex and rare fault scenarios to handle wider range of disturbances.
- Build an alert system integrated with IOT sensors and cloud dashboards for live notifications to maintenance teams.
- Combine with other algorithms like svm to enhance decision making.

REFERENCES

- Kaggle dataset link –

  https://www.kaggle.com/datasets/ziya07/power-system faults-dataset
- IBM Documentation:

  IBM Watsonx.ai

  https://www.ibm.com/docs/en/watsonx
- AutoAI Overview

  https://www.ibm.com/cloud/watson-studio/autoai

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM

Prajwal G

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/5b5ef054-0245-4fe1-b98e-fcfdd43d07c8

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Prajwal G

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/6bdb0f1a-1cf5-4a8b-9a34-a7164f093a2d

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Prajwal G

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