CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By:
Rohit Varshney- Galgotias University-MCA(Cloud Computing)

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

Example: Develop a predictive maintenance model for a fleet of industrial machines to anticipate failures before they occur. This project will involve analyzing sensor data from machinery to identify patterns that precede a failure. The goal is to create a classification model that can predict the type of failure (e.g., tool wear, heat dissipation, power failure) based on real-time operational data. This will enable proactive maintenance, reducing downtime and operational costs.

dataset link – https://www.kaggle.com/datasets/shivamb/machine predictive-maintenance-classification

PROPOSED SOLUTION

- The proposed system aims to address the challenge of predicting potential machine failures in a fleet of industrial machines to enable proactive maintenance and minimize downtime. This involves leveraging data analytics and machine learning techniques to identify patterns in sensor data that precede different types of failures (e.g., tool wear, heat dissipation issues, power failure). The solution will consist of the following components:
- Data Collection
  - Gather historical sensor data from industrial machines, including parameters such as vibration, temperature, pressure, and rotational speed.
  - Use the provided Kaggle dataset (Predictive Maintenance Dataset) to obtain labeled data on machine operational states and failure types.
  - Integrate real-time operational data streams for continuous monitoring and prediction.
- Data Preprocessing
  - Clean and preprocess the collected data to handle missing values, noise, and outliers.
  - Normalize/standardize sensor readings to ensure consistency across different machines.
  - Perform feature engineering to extract relevant patterns and statistical measures (e.g., mean, variance, frequency domain features) that might indicate machine degradation.
- Machine Learning Algorithm
  - Implement a classification model to predict the type of machine failure based on sensor data. Suitable algorithms may include:
  - Random Forest
  - Gradient Boosting (XGBoost/LightGBM)
  - Support Vector Machines (SVM)
  - Neural Networks (for deeper feature extraction)

edunet foundation

PROPOSED SOLUTION

- Compare model performances to select the best-performing classifier.
- Train the model using labeled failure data and test it on unseen data to evaluate predictive accuracy.
- Deployment
  - Develop a real-time monitoring system that processes incoming sensor data and generates failure predictions.
  - Create a dashboard/interface to display machine health status, predicted failure types, and maintenance alerts.
  - Deploy the solution on a scalable infrastructure capable of handling high-frequency sensor data streams.
- Evaluation
  - Assess the model's performance using appropriate classification metrics such as:
  - Accuracy
  - Precision, Recall, F1-Score
  - Confusion Matrix
  - Perform cross-validation and hyperparameter tuning to improve predictive performance.
  - Continuously monitor the system in a production environment and retrain models with new data to maintain accuracy.
- Result

The developed predictive maintenance model is expected to accurately anticipate different types of machine failures before they occur. This will enable maintenance teams to schedule repairs proactively, thereby reducing unplanned downtime, optimizing maintenance costs, and improving overall operational efficiency.

SYSTEM APPROACH

- System Approach
  - The "System Approach" section outlines the overall strategy and methodology for developing and implementing the predictive maintenance system for industrial machines.
- System Requirements
  - To build and deploy the predictive maintenance model, the following requirements are considered:
  - Hardware Requirements
    - Processor: Intel i5 or higher
    - RAM: 8GB or more
    - Storage: 20GB free space
    - Industrial IoT sensors: To collect real-time machine operational data
    - Reliable network connectivity to send data to IBM Cloud
  - Software Requirements
    - IBM Cloud (Lite) account
    - IBM watsonx.ai Studio for model development and deployment
    - IBM Cloud Object Storage (for dataset storage)
- Libraries Required
    - Since watsonx.ai Studio offers built-in machine learning and data handling capabilities, no additional external libraries are required. The platform's AutoAI and built-in ML algorithms are used to
    - Import and clean the dataset
    - Train and evaluate the predictive model
    - Deploy the model as an endpoint for real-time predictions

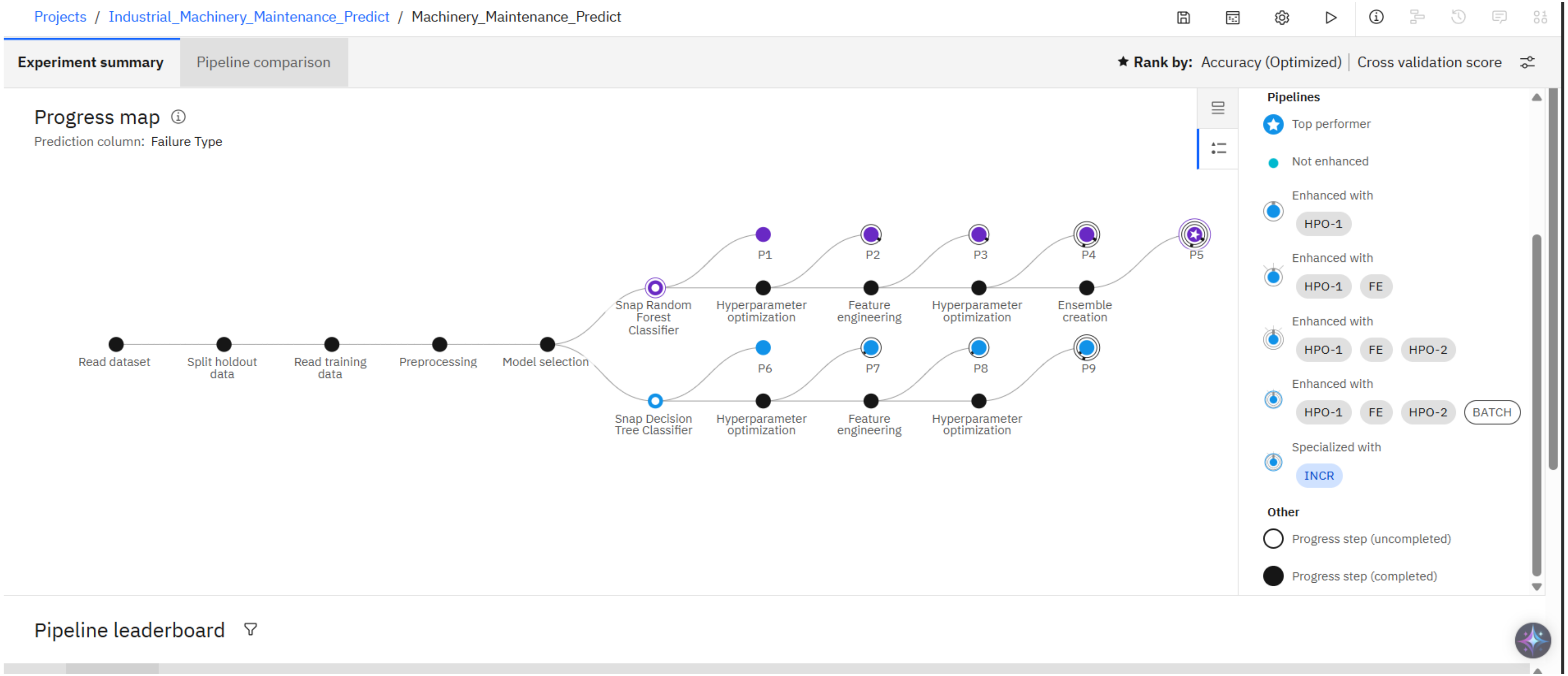
ALGORITHM & DEPLOYMENT

- Algorithm Selection
  - The chosen algorithm for predicting machinery failures is the **Snap Random Forest Classifier.** Random Forest is an ensemble learning method that builds multiple decision trees and combines their predictions to achieve higher accuracy and robustness.
    It was selected because:
  - It handles **multiclass classification** (different types of failures).
  - It is effective with **sensor data**, which may contain noisy or nonlinear patterns.
  - It provides **feature importance**, helping to identify the most critical sensor readings leading to failures.
- Data Input
  - The classifier uses sensor data from the machines, including features such as:
  - Air temperature
  - Process temperature
  - Rotational speed
  - Torque
  - Tool wear
    These features are extracted from the Kaggle dataset and serve as input variables to predict the failure type.

ALGORITHM & DEPLOYMENT

- ## Training Process
  - The dataset is uploaded to **IBM watsonx.ai Studio**.
  - Data preprocessing (cleaning, normalization, and feature selection) is performed using the platform's built-in tools.
  - The Snap Random Forest Classifier is trained on labeled historical data, where the target variable is the **failure type**.
  - Cross-validation is applied to avoid overfitting, and the algorithm automatically tunes parameters (e.g., number of trees, max depth) to improve accuracy.
- ## Prediction Process
  - Once trained, the model can process **real-time sensor data** to predict whether a machine is at risk of failure and classify the **type of failure**.
  - The model runs within IBM watsonx.ai and, when deployed, uses an API endpoint to receive live data from IoT sensors.
  - The prediction output helps maintenance teams take proactive measures, reducing downtime and costs.

RESULT



Projects / Industrial_Machinery_Maintenance_Predict / Machinery_Maintenance_Predict
Experiment summary
Pipeline comparison
★ Rank by: Accuracy (Optimized) | Cross validation score
Progress map
Prediction column: Failure Type
Read dataset
Split holdout data
Read training data
Preprocessing
Model selection
Snap Random Forest Classifier
Snap Decision Tree Classifier
Hyperparameter optimization
Feature engineering
Hyperparameter optimization
Ensemble creation
P1
P2
P3
P4
P5
P6
P7
P8
P9
Pipelines
Top performer
Not enhanced
Enhanced with
HPO-1
Enhanced with
HPO-1 FE
Enhanced with
HPO-1 FE HPO-2
Enhanced with
HPO-1 FE HPO-2 BATCH
Specialized with
INCR
Other
Progress step (uncompleted)
Progress step (completed)
Pipeline leaderboard

RESULT

Deployment spaces / Machinery_Deploy / P5 - Snap Random Forest Classifier: Machinery_Maintenance_Predict /

Machinery_Deploy

Deployed Online

API reference | Test

Enter input data

Text | JSON

Enter data manually or use a CSV file to populate the spreadsheet. Max file size is 50 MB.

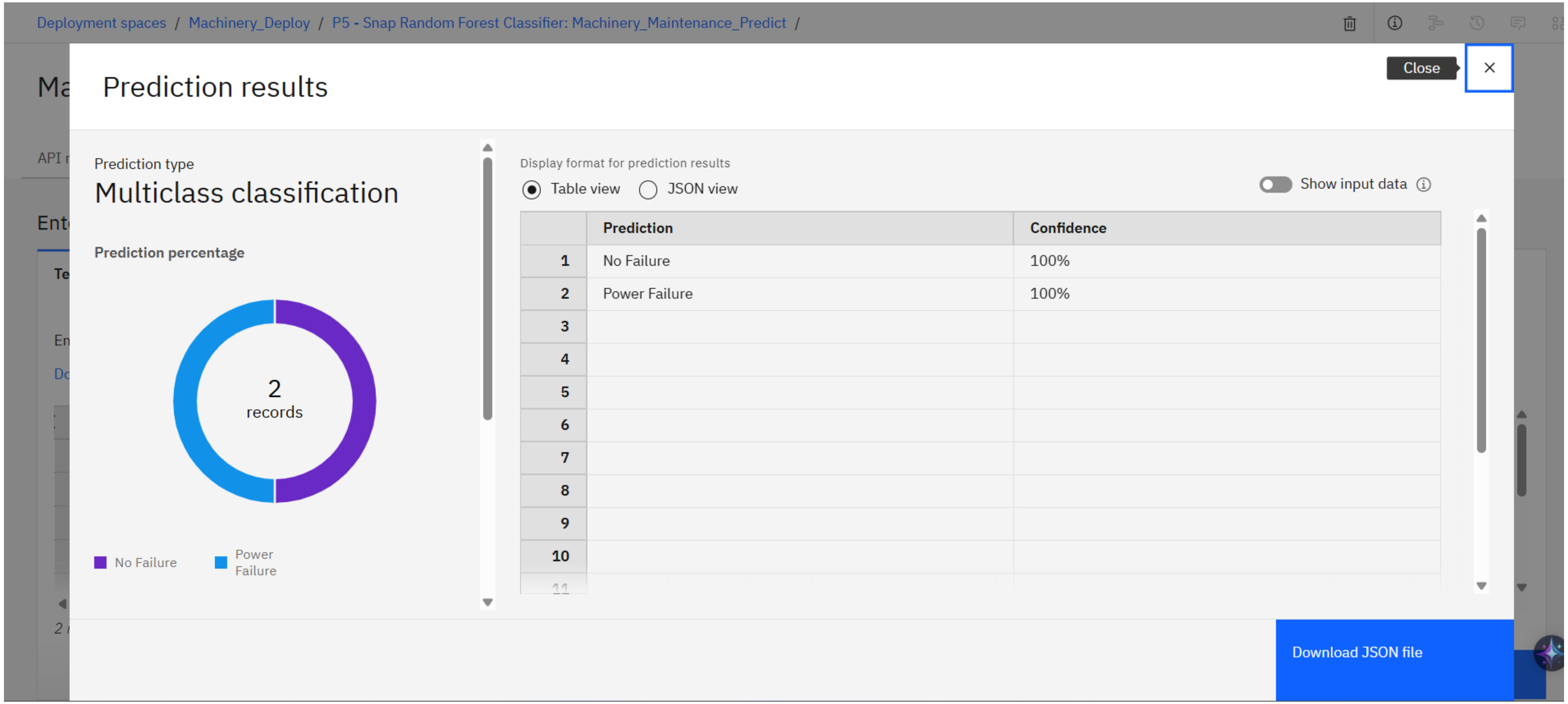
Download CSV template | Browse local files | Search in space | Clear all ×

	Type (other)	Air temperature [K] (double)	Process temperature [K] (double)	Rotational speed [rpm] (double)	Torque [Nm] (double)	Tool wear [min] (double)	Target (double)
1	J	298.3	302.5	1524	25.3	2	0
2	L	298.9	309.1	2861	4.6	143	1
3							
4							
5							

2 rows, 9 columns

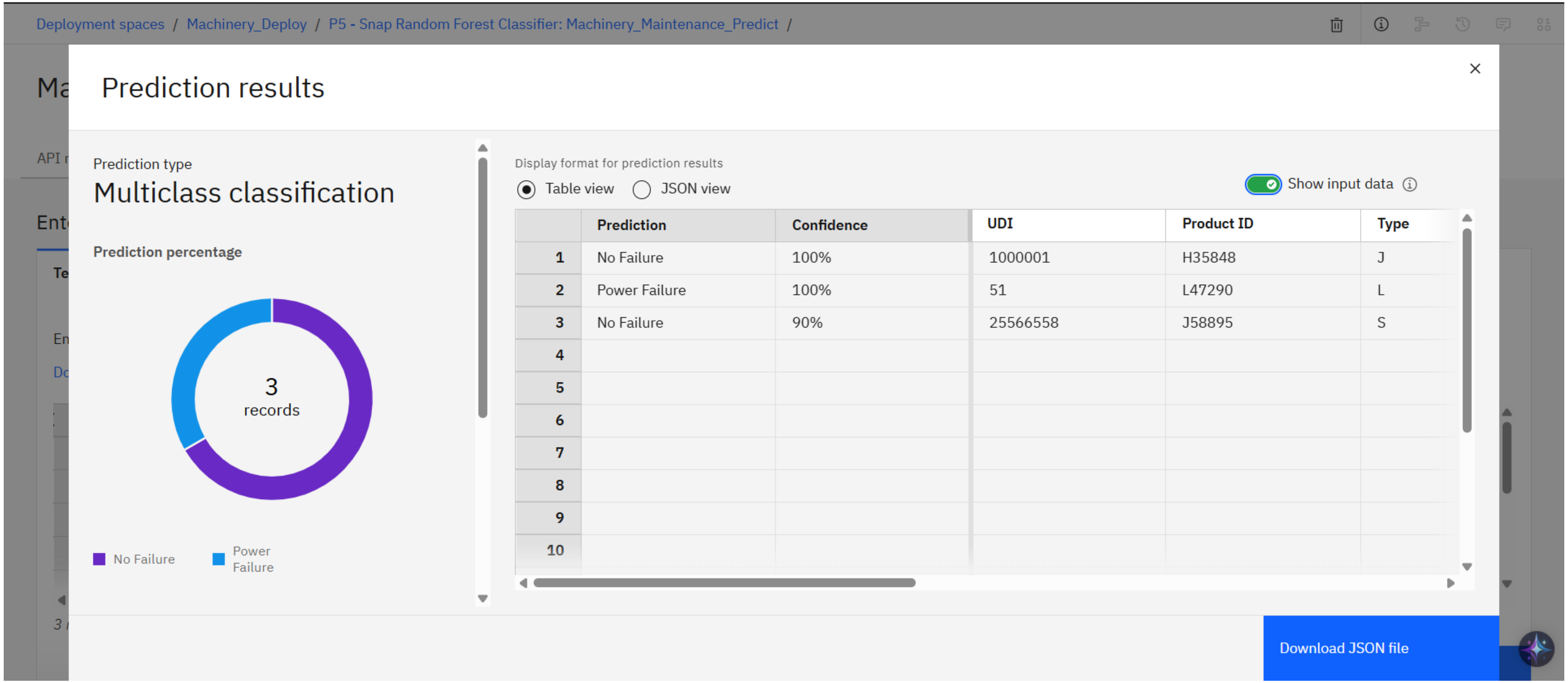
Predict

RESULT



edunet foundation

RESULT



edunet foundation

CONCLUSION

- The predictive maintenance system developed using **IBM Cloud watsonx.ai Studio (Lite)** successfully demonstrates the ability to anticipate machinery failures before they occur. By leveraging the **Snap Random Forest Classifier**, the model effectively analyzes sensor data to classify potential failure types such as tool wear, heat dissipation issues, and power failures.
- The deployment on IBM Cloud enables **real-time monitoring** through an API endpoint, allowing maintenance teams to take proactive actions. This approach minimizes unplanned downtime, optimizes maintenance schedules, and reduces operational costs.
- Overall, the solution proves that integrating **machine learning with cloud-based tools** provides an efficient and scalable way to enhance industrial machine reliability.

FUTURE SCOPE

- The predictive maintenance system can be enhanced by:
- **Integrating IoT sensors** for real-time data streaming and instant failure alerts.
- **Using advanced algorithms** (e.g., deep learning) to improve prediction accuracy.
- **Scaling on full IBM Cloud** to support larger datasets and multiple machine types.
- **Adding a visualization dashboard** for monitoring machine health and predictions.
- **Automating maintenance scheduling** through integration with CMMS.
- **Deploying on edge devices** to enable faster, offline predictions.

REFERENCES

- IBM Cloud watsonx.ai Documentation – https://www.ibm.com/products/watsonx-ai
- Kaggle Dataset – Predictive Maintenance Classification – https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification
- Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32.
- IBM Developer. Building and Deploying Machine Learning Models on IBM Cloud – https://developer.ibm.com

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Rohit Varshney

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/48d14f95-1715-4f7c-8d98-5d4264a6e544

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

Rohit Varshney

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/9a776de3-75f3-4bfc-8db9-e4892b3b321b

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rohit Varshney

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 25 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU

edunet
foundation