CAPSTONE PROJECT

PROJECT TITLE

Presented By:
1. Raj Kiran Acharyya – MCKV Institute Of Engineering – Computer Science & Engineering

edunet foundation

OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

- **Problem statement No.39 – Predictive Maintenance of Industrial Machinery**
- Example: This project aims to predict equipment failures in industrial machines before they happen using sensor data. By analyzing real-time parameters like temperature, torque, and tool wear, a classification model is built to identify issues like tool wear, overheating, or power failure. This helps enable proactive maintenance, minimizing downtime and operational costs.

PROPOSED SOLUTION

- The proposed system aims to address the challenge of predicting potential machine failures in industrial environments before they occur. This involves leveraging data analytics and machine learning techniques to accurately forecast failure patterns based on operational data. The solution will consist of the following components:
- **Data Collection:**
  - Gather historical sensor data from industrial machines, including parameters like temperature, torque, rotational speed, and tool wear.
  - Focus solely on the available sensor readings from the dataset without relying on external contextual inputs.
- **Data Preprocessing:**
  - Clean and preprocess the raw sensor data to handle missing values, outliers, and inconsistencies.
  - Perform feature engineering to derive relevant indicators (e.g., rate of temperature change, wear thresholds) that may signal upcoming failures.
- **Machine Learning Algorithm:**
  - IBM watsonx.ai's AutoAI generated a Snap Random Forest Classifier (Batched Tree Ensemble) that achieved a cross-validation accuracy of **99%**. This ensemble model aggregates multiple decision trees to classify failure types accurately based on input sensor data.
  - The model pipeline was automatically enhanced using Feature Engineering (FE) and Hyperparameter Optimization (HPO) to improve accuracy and robustness. Real-time features like air temperature, process temperature, torque, and tool wear were key inputs for classification.

- **Deployment:**
  - IBM Watsonx.ai Studio is used to deploy the trained Snap Random Forest Classifier, enabling real-time predictions via an easy-to-use interface that accepts JSON or CSV inputs.
  - IBM's hybrid pipeline, which provides quick, scalable, and dependable access to real-time machine failure predictions, powers the model. It also has a public REST API.
- **Evaluation:**
  - During AutoAI training in IBM Watsonx.ai Studio, accuracy metrics and cross-validation were used to evaluate the model's performance.
  - Accuracy metrics and cross-validation were used to evaluate the model's performance during AutoAI training in IBM Watsonx.ai Studio.
  - Result: The deployed model accurately predicts failure types such as tool wear, power failure, and overheating, enabling proactive maintenance

SYSTEM APPROACH

This section outlines the overall strategy and methodology for developing and implementing the **Predictive Maintenance System using IBM watsonx.ai Studio.** Here's a suggested structure for this section:

- **System requirements:** To build and deploy the predictive maintenance classification model, the following components were required:
  - IBM watsonx.ai Studio for AutoAI pipeline generation and deployment.
  - IBM watsonx.ai Runtime for inferencing and model training.
  - Watson Machine Learning services for deployment and scalability.
  - Online deployment environment with a REST API endpoint for real-time prediction.
  - 10–20 Capacity Unit-Hours (CUH)
  - Compute environments (1–4 vCPUs, 4–16 GB RAM)

edunet foundation

- **Library required to build the model:** Key libraries and tools used in the AutoAI pipeline and backend processing:
  - **AutoAI** – Automated data preprocessing, feature engineering, and model selection
  - **Snap Random Forest Classifier** – Chosen as the top-performing algorithm
  - **Pandas & NumPy** – For initial dataset formatting
  - **Sklearn metrics** (internally used) – For model evaluation (accuracy, precision, recall, F1)
  - **Deployment Space on watsonx.ai** – For real-time online predictions via JSON input
  - **API reference & testing interface** – Provided by IBM watsonx.ai for integration

ALGORITHM & DEPLOYMENT

- The Snap Random Forest Classifier, chosen by IBM watsonx.ai's AutoAI for its high accuracy, it was deployed to deliver real-time failure predictions in the maintenance system. Its robust performance made it ideal for production use.
- **Algorithm Selection:**
  - The model chosen by IBM watsonx.ai’s AutoAI was the **Snap Random Forest Classifier**, which emerged as the best-performing algorithm during experimentation.
  - This ensemble-based tree classifier is well-suited for handling tabular sensor data and performs well in classifying different failure types. Its ability to manage non-linear relationships and categorical variables makes it ideal for predictive maintenance tasks.
- **Data Input:**
  - The model uses key sensor inputs—machine type, air and process temperatures, rotational speed, torque, and tool wear—to assess equipment health and predict failures.
  - These features reflect the machine’s operational state and help the model detect early signs of issues before they escalate.

- **Training Process:**
  - Training was done via IBM watsonx.ai AutoAI, which automatically handled feature engineering, cross-validation, and hyperparameter tuning.
  - It generated multiple pipelines, with the best one (Snap Random Forest) achieving ~99.5% accuracy.
- **Prediction Process:**
  - It accepts input as a structured JSON or CSV format through the online deployment interface in watsonx.ai.
  - The model then returns the predicted failure type, enabling preventive action before actual breakdowns occur.
  - Once deployed, the trained Snap Random Forest model classifies real-time operational data into potential failure categories.

RESULT

IBM watsonx.ai Studio

Search in your workspaces

Upgrade

Raj Kiran Acharyya's Accou...

London

RA

Deployment spaces / DeployingPredictiveMaintenance / P5 - Snap Random Forest Classifier: predictive_maintenance /

Deploying_Machine_Predictive_Maintenance_Classification

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 ×

	UDI (double)	Product ID (other)	Type (other)	Air temperature [K] (double)	Process temperature [K] (double)	Rotational speed [rpm] (double)	Torque [Nm] (double)	Tool wear [min] (double)	Target (double)
1	20	M14879	M	298.6	309.3	1632	32.4	55	0
2	364	L47543	L	297.6	308.3	1438	47.8	84	0
3	904	L48083	L	295.7	306.2	2270	14.6	149	1
4	1212	H30625	H	296.9	308	1416	46.9	105	0
5	4436	L51615	L	302.3	310.1	1321	52.7	134	1
6	4817	L51996	L	303.4	312	1521	35.9	215	1
7	2951	M17810	M	300.7	309.5	1445	46.4	24	0
8	3866	H33279	H	302.6	311.5	1629	34.4	228	1
9									
10									

8 rows, 9 columns

Predict

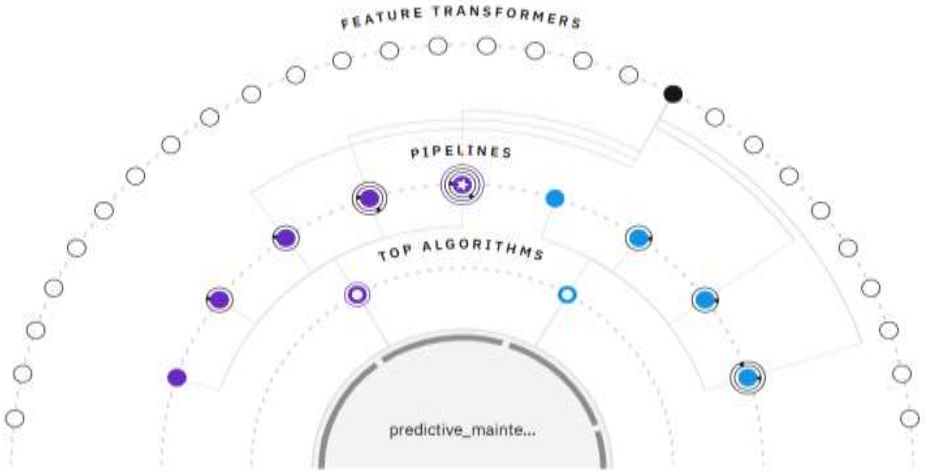
IBM watsonx.ai Studio   Search in your workspaces   Upgrade   Raj Kiran Acharyya's Account   London   RA

Projects / Smart_Predictive_Maintenance_for_Industrial_Machines / predictive_maintenance

Experiment summary   Pipeline comparison   ★ Rank by: Accuracy (Optimized) | Cross validation score

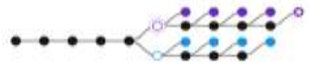
Relationship map

Prediction column: Failure Type



Progress 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 5	Batched Tree Ensemble Classifier (Snap Random Forest Classifier)	INCR	0.995	HPO-1 FE HPO-2 BATCH	00:00:47
	2	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:44
	3	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:35
	4	Pipeline 9	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:04
	5	Pipeline 2	Snap Random Forest Classifier		0.994	HPO-1	00:00:09
	6	Pipeline 1	Snap Random Forest Classifier		0.994	*None*	00:00:05

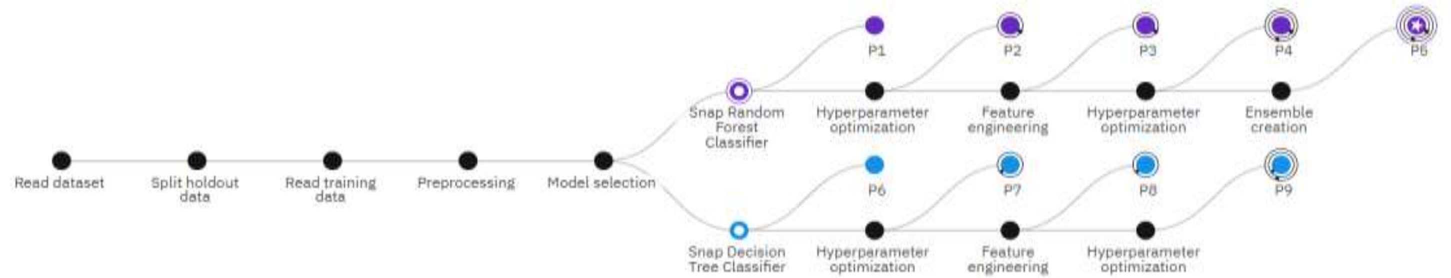
IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Raj Kiran Acharyya's Account | London | RA

Projects / Smart_Predictive_Maintenance_for_Industrial_Machines / predictive_maintenance

Experiment summary | Pipeline comparison | ★ Rank by: Accuracy (Optimized) | Cross validation score

Progress map

Prediction column: **Failure Type**



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 5	Batched Tree Ensemble Classifier (Snap Random Forest Classifier)	INCR	0.995	HPO-1 FE HPO-2 BATCH	00:00:47
	2	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:44
	3	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:35
	4	Pipeline 9	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:04
	5	Pipeline 2	Snap Random Forest Classifier		0.994	HPO-1	00:00:09
	6	Pipeline 1	Snap Random Forest Classifier		0.994	*None*	00:00:05

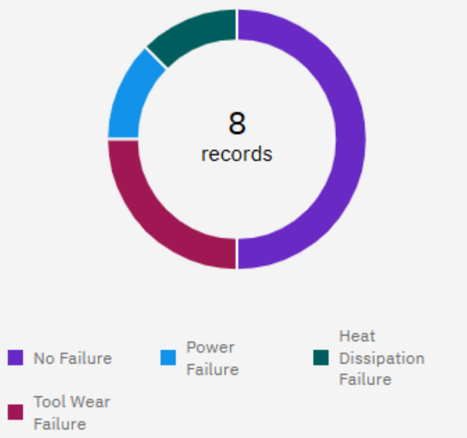
Prediction results

×

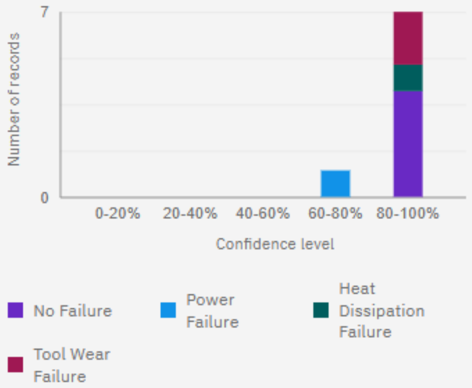
Prediction type

Multiclass classification

Prediction percentage



Confidence level distribution



Display format for prediction results

(•) Table view ( ) JSON view

Show input data ⓘ

	Prediction	Confidence
1	No Failure	100%
2	No Failure	100%
3	Power Failure	70%
4	No Failure	100%
5	Heat Dissipation Failure	100%
6	Tool Wear Failure	90%
7	No Failure	100%
8	Tool Wear Failure	100%
9		
10		
11		
12		
13		
14		
15		
16		

CONCLUSION

- In this project, I developed a predictive maintenance model for industrial machinery using IBM watsonx.ai Studio and AutoAI. By working with real-time sensor data and applying the Snap Random Forest Classifier, I was able to accurately classify failure types such as tool wear, overheating, and power failure. The model showed strong predictive performance, making it suitable for proactive maintenance that can reduce downtime and operational costs.
- Through this work, I've seen firsthand how AI can play a powerful role in improving machine reliability and efficiency. This project helped me understand the value of predictive analytics in industrial settings and how intelligent systems can support smarter decision-making.

FUTURE SCOPE

- **Integration with IoT Platforms:** Real-time sensor data can be streamed using IoT devices to enhance live monitoring and quicker failure detection.
- **Model Enhancement with Deep Learning:** Future versions can use LSTM or CNN models for improved accuracy, especially on sequential or time-series data.
- **Expansion to Other Industries:** The model can be adapted for predictive maintenance in sectors like aviation, automotive, and manufacturing.
- **Automated Alert System:** Integration of the model with maintenance scheduling systems can trigger alerts and maintenance tickets automatically.
- **Edge Computing Deployment:** For faster, on-site processing without relying on cloud latency, models can be deployed on edge devices.
- **Continuous Learning Loop:** With more data, the model can be retrained periodically to improve performance and adapt to new failure patterns.
- **Incorporating Cost Analysis:** Adding a cost-benefit layer to prioritize maintenance actions based on economic impact.

REFERENCES

- **Kaggle dataset link – https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification**
- IBM Cloud Documentation -- https://www.ibm.com/cloud/watsonx-ai

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

IBM

Raj Kiran Acharyya

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 20, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/1b8f5e3e-12b2-4c7f-997f-807f1a752027

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Raj Kiran Acharyya

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/be7e656b-c6f7-456b-93b4-29edc46026de

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Raj Kiran Acharyya

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 26 Jul 2025 (GMT)

Learning hours: 20 mins

edunet foundation

THANK YOU