
CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By:

- 1. Aman Dhiman**
- 2. S.I.E.T Panchkula**
- 3. B.Tech CSE (Artificial Intelligence and Machine Learning)**

OUTLINE

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

PROBLEM STATEMENT

- The aim is to use sensor data collected from machines—such as temperature, pressure, vibration, and rotational speed—to predict the type of failure that might occur. These failures could include tool wear, power failure, or issues with heat dissipation. By analyzing this data, the model should be able to detect early warning signs of mechanical problems and classify the likely type of failure before it happens. This proactive approach will help reduce unexpected machine breakdowns, minimize downtime, and lower maintenance costs. The project involves building a classification model that can learn from historical machine data and support real-time decision-making for industrial operation

PROPOSED SOLUTION

- The proposed system aims to address the challenge of predicting machine failure using sensor data such as air temperature, process temperature, rotational speed, and torque, was uploaded and explored.
- Data Collection:
 - Gather historical data based on parts wearout phase.
 - Utilize real-time data sources, air temperature, process temperature, rotational speed, and torque.
- Data Preprocessing:
 - Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.
 - Feature engineering to extract relevant features from the data.
- Machine Learning Algorithm:
 - Random Forest Algorithm is used to accurately classify different types of machine failures based on sensor data patterns.
 - Consider incorporating additional factors like machine age, maintenance history, and operating environment to improve the accuracy of failure predictions.
- Deployment:
 - The predictive maintenance model was deployed on IBM Cloud .
 - Deploy the solution on a scalable and reliable platform, considering factors like server infrastructure, response time, and user accessibility.
- Evaluation:
 - Assess the model's performance using appropriate metric(Accuracy).
 - Result: Accuracy = 0.995

SYSTEM APPROACH

- **System Approach**

- The "System Approach" outlines the overall methodology and technical framework used to develop and deploy the predictive maintenance model for industrial machinery.

- **System Requirements**

- **IBM Cloud Lite Account** (for Watson Studio and Machine Learning services)
- **Internet Browser** (Chrome/Firefox for accessing IBM Watson Studio)
- **Cloud Storage** (IBM Cloud Object Storage for data storage and access)
- **Optional:** Local Python environment (for exploratory data analysis and offline experimentation)

ALGORITHM & DEPLOYMENT

- **Algorithm Selection**

- Random Forest was chosen for its high accuracy, robustness to noise, and ability to handle complex, non-linear data.

- **Data Input**

- Input features included air temperature, process temperature, torque, rotational speed, tool wear, and machine type

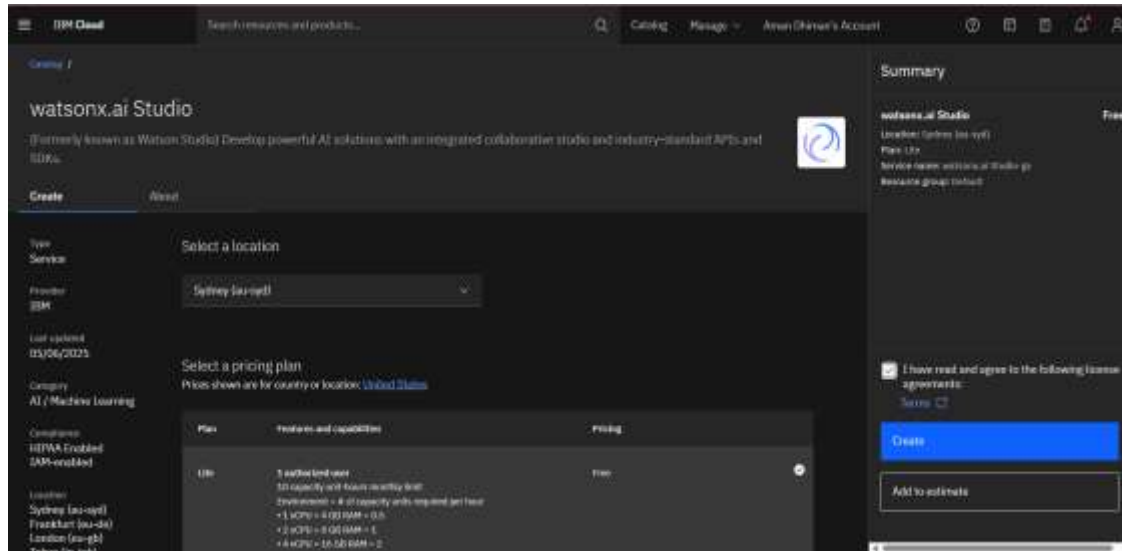
- **Training Process**

- The model was trained using labeled failure data with cross-validation and hyperparameter tuning to prevent overfitting and improve generalization.

Prediction Process

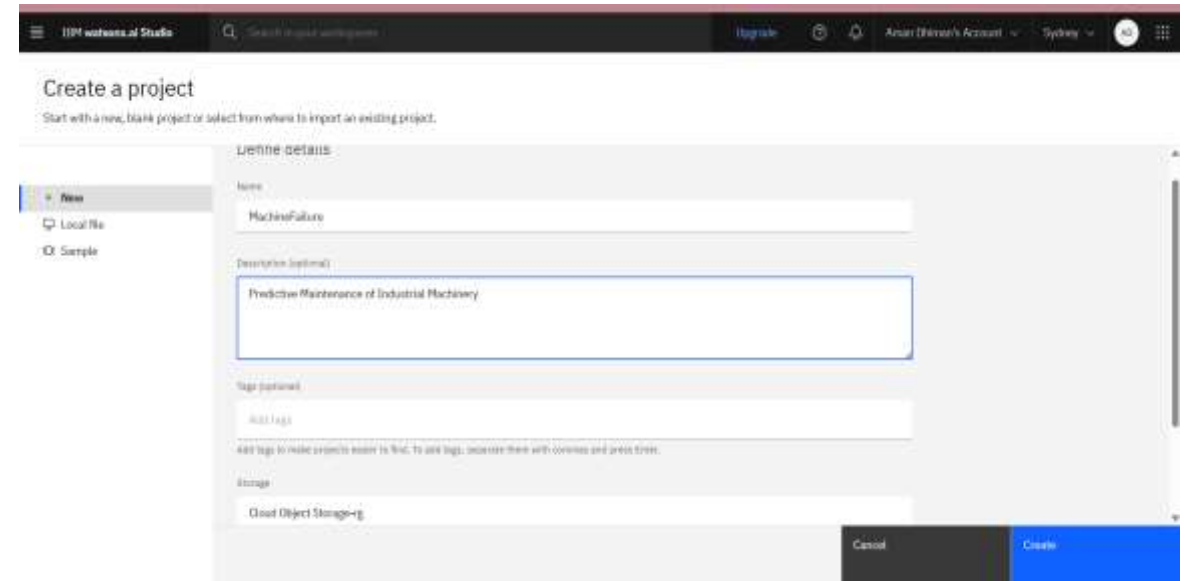
- The trained Random Forest model predicts the type of machine failure based on live sensor data, enabling real-time, proactive maintenance decisions.

RESULT



Creating a project in IBM Cloud helps organize and manage resources, services, and team collaboration under a single workspace. It allows you to track usage, assign roles, and structure workflows efficiently across AI, data, and app development environments.

IBM's **watsonx.ai Studio** is an end-to-end generative AI and ML development environment, enabling AI builders to build, train, fine-tune and deploy foundation models—including IBM's Granite and open-source LLMs—via Prompt Lab, Tuning Studio, AutoAI, SDKs and REST APIs.



Build machine learning models automatically

Define the details to create an AutoAI experiment and open it in the AutoAI tool.

+ New

Sample

Define details

Name

Description (optional)

What's the purpose of this AutoAI experiment?

Tags (optional)

Add tags to make assets easier to find.

Define configuration

watsonx.ai Runtime service instance

Environment definition

This environment definition consumes 20 capacity units per hour for training. For details, see [watsonx.ai Runtime plans](#).

Cancel

Back

Create

Build machine learning model automatically help to build machine learning model on ibm cloud without any prior knowledge of machine learning and coding.

predictFailure is the name of the machine learning model used to predict Failure Type In industrial machines to predict the faults in machine before they occurs.

IBM watsonx.ai Studio

Search for your workspaces

Upgrade

Amas Ofiman's Account

Sydney

AD

Projects / MachineFailure / predictFailure

Configure AutoAI experiment

predictFailure

Autosaved: 11:20:49 AM

Add files such as tabular data (CSV).

predictive_maintenance (1).csv

Size: 525.57 KB Columns: 10

Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential. [Learn more](#)

Yes

What do you want to predict?

Prediction column

Prediction column: Failure Type

CUH remaining: 20 CUH

PREDICTION TYPE

Multiclass Classification

OPTIMIZED FOR

Accuracy & run time

Experiment settings

Rank	Name	Algorithm	Specialization	Accuracy (Optimized Cross Validation)	Enhancements	Build time
1	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:30
2	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:20
3	Pipeline 3	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:40
4	Pipeline 2	Snap Random Forest Classifier		0.994	HPO-1	00:00:06

It ranks different machine learning pipelines by their optimized cross-validation accuracy. The top-performing models use the **Snap Random Forest Classifier**, with accuracies up to **0.995**, and include enhancements like **hyperparameter optimization (HPO)** and **feature engineering (FE)**.

Save as interface in IBM watsonx.ai, where a trained model pipeline is being saved as a **Model asset**. The selected pipeline, **P4 - Snap Random Forest Classifier: predictFailure**, will be saved for future use—enabling deployment for predictions, testing with new data, and lineage tracking.

Select asset type

Model

Create a watsonx.ai Runtime model asset that you can test with new data, deploy to generate predictions, and trace lineage activity.

Notebook

Create a notebook if you want to view the code that created this model pipeline or interact with the model programmatically.

Define details

Name

P4 - Snap Random Forest Classifier: predictFailure

Description (optional)

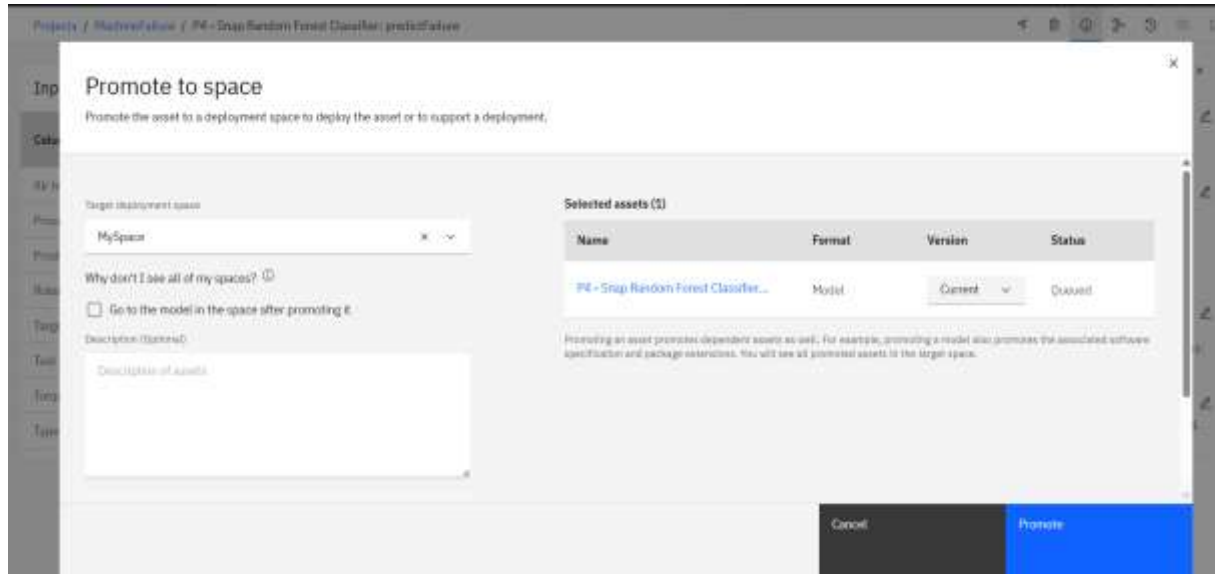
Model description

Tags

Add tags to make assets easier to find.

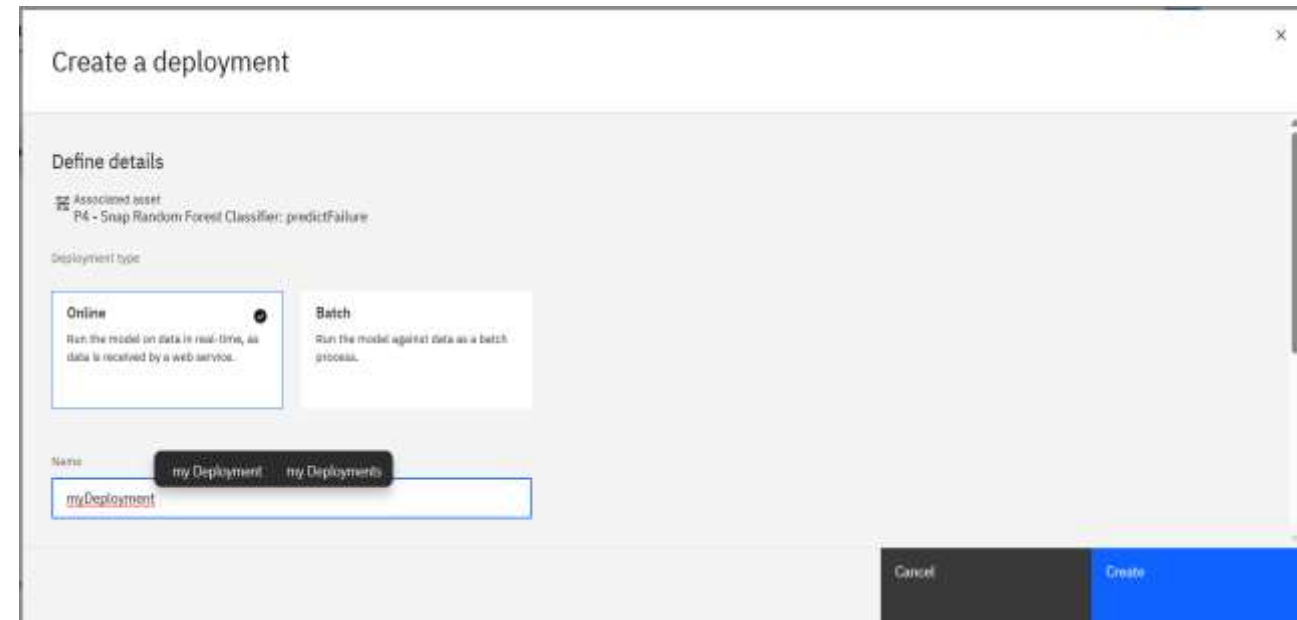
Add a tag

Cancel
Create



"Promote to space" step in IBM watsonx.ai, where the saved model asset (**P4 - Snap Random Forest Classifier**) is being promoted to a **deployment space** (in this case, *MySpace*). This step is essential to enable the model's deployment, allowing it to serve predictions and be managed in a production-like environment.

"Create a deployment" interface in IBM watsonx.ai. The selected model asset (**P4 - Snap Random Forest Classifier**) is being deployed using the **Online** option, which allows real-time predictions via a web service endpoint.



IBM watsonx.ai Studio

Search in your workspaces

Upgrade

Aman Dhiman's Account

Sydney

AD

Deployment spaces / MySpace / P4 - Snap Random Forest Classifier: predictFailure /

myDeployment
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.

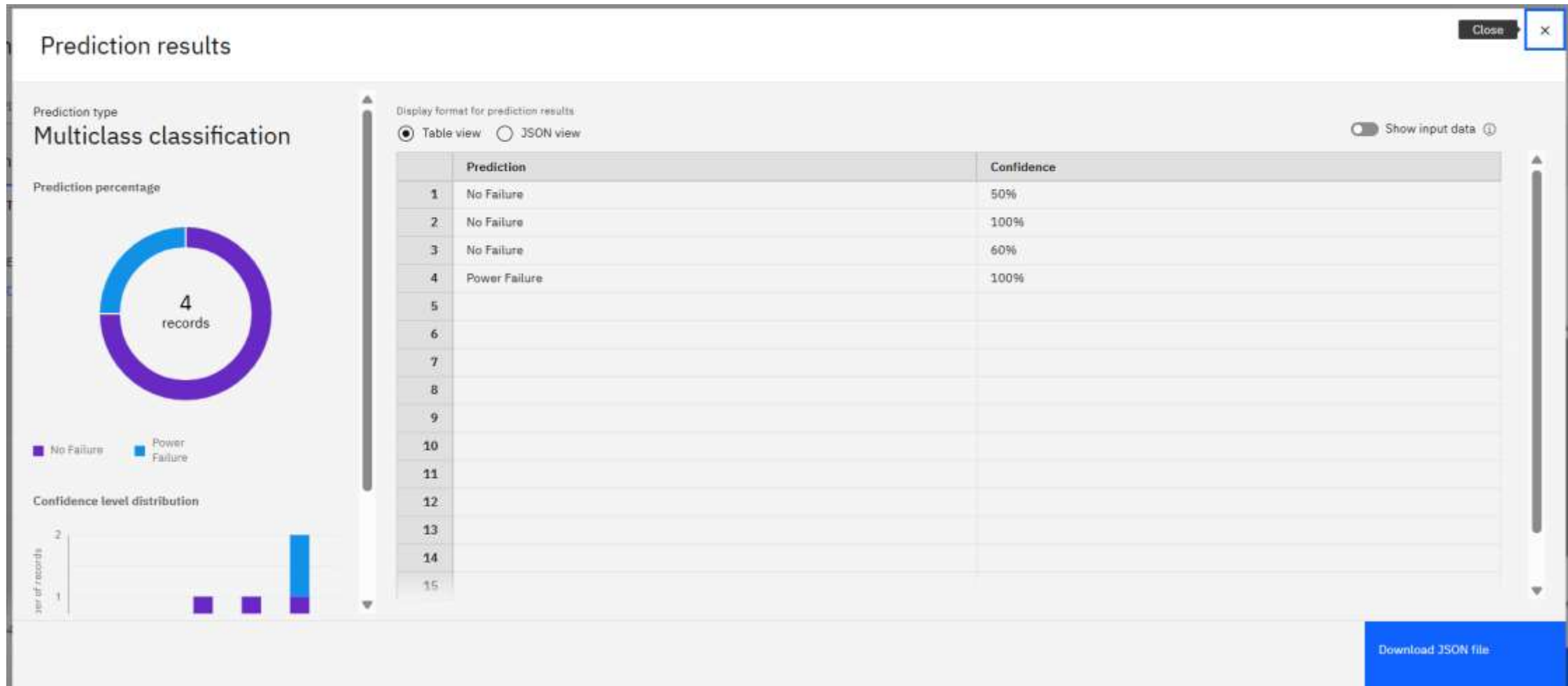
Download CSV template
Browse local files
Search in space
Clear all

	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	1	L69482	L	392	400	3500	46	165	1
2	2	M67496	M	326	296	2500	3500	200	0
3	3	L57398	L	298	290	2100	4	140	1
4	4	L47230	L	299	309	2861	3.5	140	1
5									
6									
7									
8									
9									

4 rows, 9 columns

Predict

The **testing phase of a deployed model** in IBM watsonx.ai Studio. The deployed asset, **myDeployment**, is set to **Online**, allowing real-time predictions. Users can manually input or upload tabular data (like a CSV) to test the model. In this example, input variables such as **air temperature**, **process temperature**, **rotational speed**, **torque**, and **tool wear** are provided. After populating the data, clicking **Predict** will generate failure predictions using the trained **Snap Random Forest Classifier** model.



The **prediction results** from an IBM watsonx.ai model deployed for **multiclass classification**. Out of four test records, three were predicted as "**No Failure**", and one as "**Power Failure**", with corresponding confidence levels shown alongside. The donut chart and confidence distribution graph on the left provide a quick visual summary of outcomes, while the **table view** on the right details predictions per input. These insights help evaluate how well the model is performing on real or simulated data.

CONCLUSION

- The predictive maintenance model developed using the Random Forest algorithm successfully identifies potential machine failures before they occur. By leveraging real-time sensor data and deploying the model on IBM Cloud using watsonx.ai Studio services, the solution enables timely maintenance actions, minimizes unexpected downtimes, and enhances operational efficiency. The automated approach ensures scalability and can be integrated seamlessly into industrial systems for continuous monitoring and decision-making.

FUTURE SCOPE

- **Integration with IoT platforms** to automate data collection and real-time model inference.
- **Addition of new features** such as machine age, maintenance history, and environmental conditions to improve prediction accuracy.
- **Addition of new features** such as machine age, maintenance history, and environmental conditions to improve prediction accuracy.
- **Model retraining pipelines** for continuous learning as new data is generated.
- **Expansion to multiple machine types** and failure modes across diverse industrial environments

REFERENCES

- **Kaggle Dataset – Machine Predictive Maintenance Classification**

<https://www.kaggle.com/datasets/shivamb/machinepredictive-maintenance-classification>

- The dataset used to train and evaluate the failure prediction model.

- **IBM AutoAI Documentation**

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

- Official documentation for IBM AutoAI, used to automate model training and selection.

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Aman Dhiman

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/e73c1cf4-0c30-4c91-9402-7aab066eaf08>



IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



Aman Dhiman

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/df6ac2ef-4857-4ca1-845f-2fbd6ebd4f1b>



IBM CERTIFICATIONS

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to

Aman Dhiman

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