CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY USING MACHINE LEARNING

Presented By:
1. Gautam Tripathi-Graphic Era Hill University-CSE

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

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.

PROPOSED SOLUTION

- The goal is to develop a predictive maintenance model that can anticipate machine failures before they occur. This model assists organizations in scheduling timely maintenance, reducing downtime and associated costs, and improving operational efficiency. The model should classify whether a machine is likely to fail (binary or multi-class, depending on failure types).
- **Key components:**
  - Data Collection: Use the Kaggle Predictive Maintenance dataset.
  - Preprocessing: Clean, encode, and normalize the dataset.
  - Model Training: Train a classification model (Decision Tree, Random Forest).
  - Evaluation: Validate using accuracy, precision, recall, and F1-score.

SYSTEM APPROACH

- The "System Approach" section outlines the overall strategy and methodology for developing and implementing the power system fault detection and classification. Here's a suggested structure for this section:
- **System requirements :**

  IBM Cloud (mandatory)

  IBM Watson Studio for model development and deployment

  IBM Cloud Object Storage for dataset handling

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**

  Random Forest Classifier (or Decision Tree based on performance).

- **Data Input:**

  UDI, Product ID, Type, Air temperature [K], Process temperature [K], Rotational speed [rpm], Torque [Nm], Tool wear [min], Target.

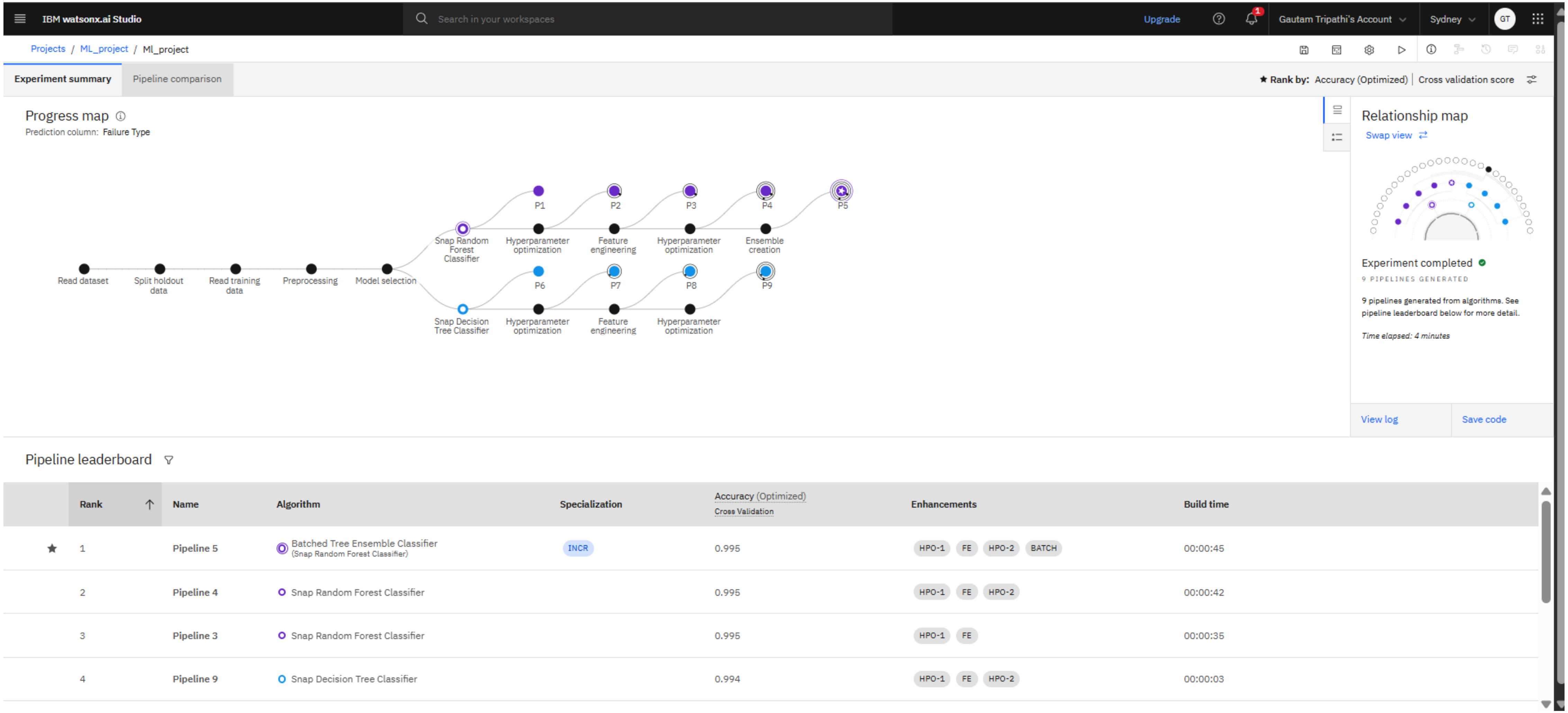
- **Training Process:**

  Supervised learning using labeled failure types.

- **Prediction Process:**

  Model deployed on IBM Watson Studio with API endpoint for real-time predictions

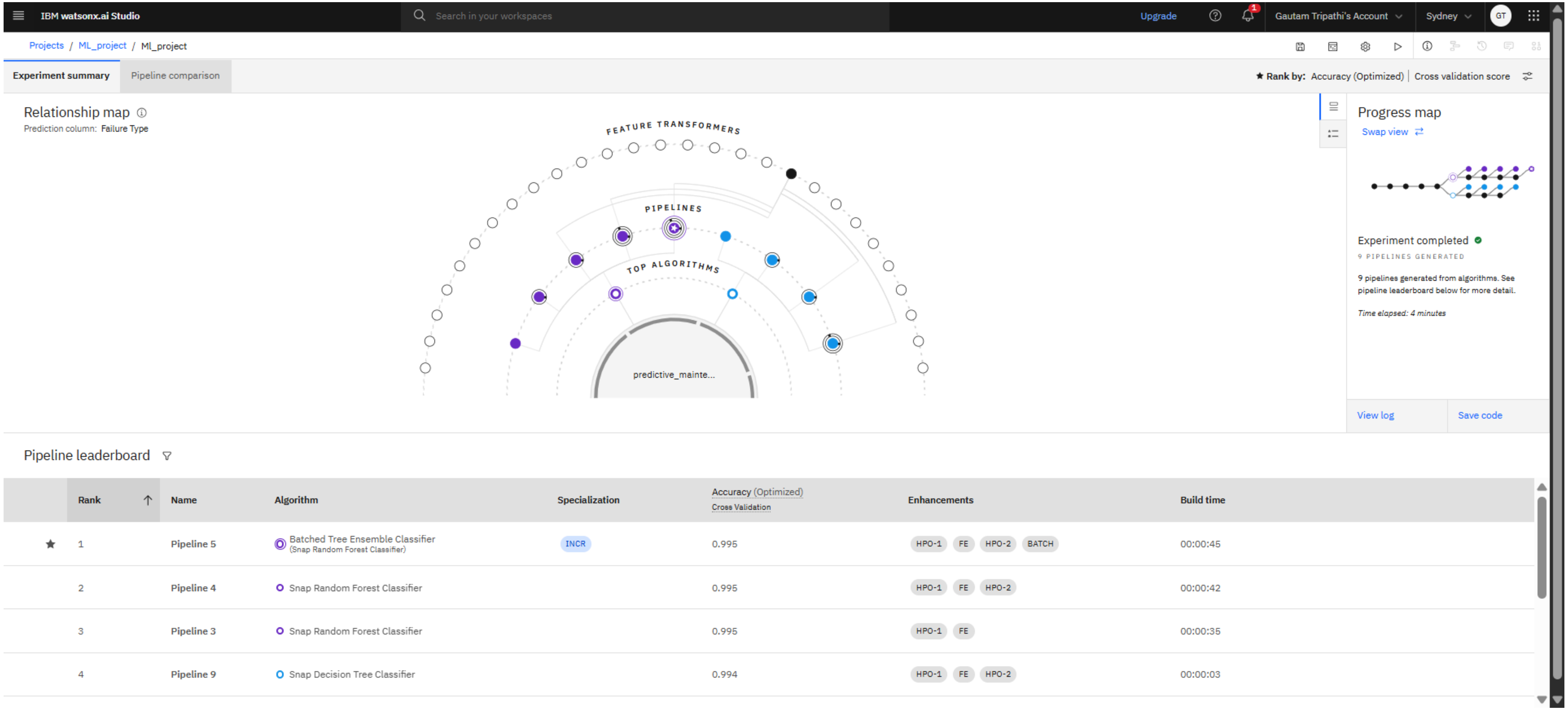
RESULT



	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	00:00:45
	2	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:42
	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	00:00:03

edunet foundation

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Gautam Tripathi's Account
Sydney
Projects / ML_project / Ml_project
Experiment summary
Pipeline comparison
Rank by: Accuracy (Optimized) | Cross validation score
Relationship map
Prediction column: Failure Type
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
predictive_mainte...
Progress map
Swap view
Experiment completed
9 PIPELINES GENERATED
9 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 4 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 00:00:45
2 Pipeline 4 Snap Random Forest Classifier 0.995 HPO-1 FE HPO-2 00:00:42
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 00:00:03

RESULT

IBM watsonx.ai Studio   Search in your workspaces   Upgrade   Gautam Tripathi's Account   Sydney   GT

Deployment spaces / ML_deploy / P5 - Snap Random Forest Classifier: Ml_project /

Ml_Deployment

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	161	L47340	L	298.4	308.2	1282	60.7	216	1
2	1303	L48482	L	298.6	309.8	1505	45.7	144	0
3	169	L47348	L	298.4	308.3	1433	62.3	20	1
4	1501	L48689	L	298	308.5	1429	37.7	220	1
5	3237	M18096	M	300.8	309.4	1342	62.4	113	1
6	1256	L56841	L	254.21	365.7	1896	56.2	214	0
7									
8									
9									
10									

6 rows, 9 columns

Predict

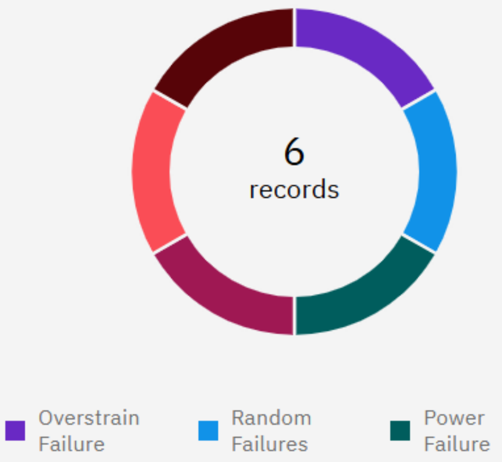
RESULT

Prediction results

Prediction type

Multiclass classification

Prediction percentage



Display format for prediction results

Table view JSON view

Show input data

	Prediction	Confidence
1	Overstrain Failure	96%
2	Random Failures	80%
3	Power Failure	100%
4	Tool Wear Failure	100%
5	Heat Dissipation Failure	100%
6	No Failure	50%
7		
8		
9		

Download JSON file

CONCLUSION

- The project successfully demonstrates a machine learning-based predictive maintenance system capable of forecasting equipment failures using real-time sensor data. By implementing the Random Forest model on IBM Watson Studio, the system achieved reliable classification performance across multiple failure types. Key challenges, such as handling imbalanced classes and maintaining data quality, were effectively addressed through comprehensive preprocessing and model tuning. Overall, the solution enables proactive maintenance scheduling, significantly reducing unplanned downtimes and enhancing operational efficiency.

FUTURE SCOPE

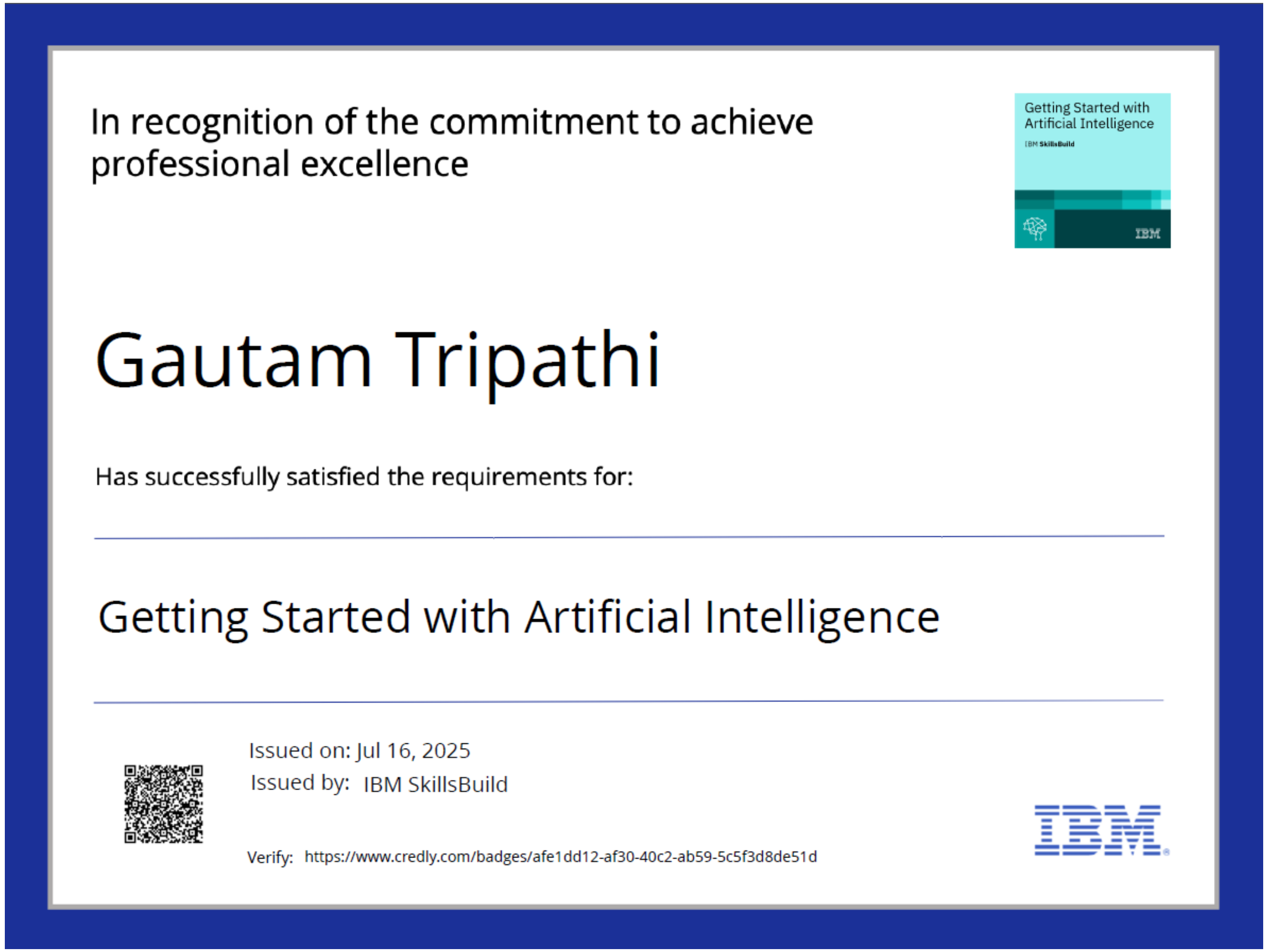
- The current system can be extended by incorporating additional sensor parameters, such as vibration levels or electrical current, to enhance prediction accuracy. The model can be further scaled to support various machine types and failure modes by retraining on more diverse datasets. Integrating real-time data streaming and prediction services through IBM Cloud pipelines can help deploy the solution in live industrial environments. Additionally, building a user-friendly monitoring dashboard for maintenance teams will enable faster decision-making. Exploring lightweight deployment methods, such as containerization with Docker or serverless APIs, can also improve system portability and scalability.

REFERENCES

- IBM Watson Studio Documentation – https://dataplatform.cloud.ibm.com
- Kaggle Predictive Maintenance Dataset – https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification
- Scikit-learn: Machine Learning in Python – https://scikit-learn.org
- Random Forest Algorithm: Breiman, L. (2001). Random Forests. Machine Learning Journal
- IBM Cloud Object Storage Documentation – https://cloud.ibm.com/docs/cloud-object-storage

edunet foundation

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

Gautam Tripathi

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/afe1dd12-af30-40c2-ab59-5c5f3d8de51d

IBM

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild

Gautam Tripathi

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/2328e016-49dc-4f69-a674-fba613f94573

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Gautam Tripathi

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