CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY USING MACHINE LEARNING

Presented By:
1. BIKKUMALLA UDAYASREE – MALLA REDDY COLLEGE OF ENGINEERING AND TECHNOLOGY – COMPUTER SCIENCE AND ENGINEERING.

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

Predictive maintenance is critical in industrial settings to avoid unexpected machine failures, reduce downtime, and minimize operational costs. This project focuses on developing a predictive maintenance model for a fleet of industrial machines to anticipate failures before they occur. The challenge lies in analyzing historical and real-time sensor data from machinery to detect patterns and signals that typically precede different types of failures. The goal is to build a robust classification model capable of accurately predicting the specific type of failure—such as tool wear, heat dissipation issues, or power failures—before it happens. By enabling proactive intervention through intelligent insights, the system will help maintenance teams take timely action, improve operational efficiency, enhance machine longevity, and significantly reduce unplanned interruptions.

PROPOSED SOLUTION

- To address the challenge of anticipating machinery failures in industrial environments, we propose building a **predictive maintenance system** powered by machine learning. The solution will focus on collecting and analyzing real-time and historical **sensor data** from a fleet of machines to detect patterns and anomalies that precede failures. The project will involve several key stages:
- **Data Collection and Preprocessing**:
  Sensor data such as temperature, pressure, vibration, power consumption, and rotational speed will be gathered. The raw data will be cleaned, normalized, and synchronized. Time-based features, statistical metrics, and trend indicators will be engineered to enhance predictive performance.
- **Exploratory Data Analysis (EDA)**:
  We will perform detailed EDA to understand sensor behavior under normal and failing conditions. Visualization techniques and correlation analysis will help uncover the relationships between features and failure types.
- **Model Development**:
  A **supervised machine learning classification model** will be trained to predict specific failure types (e.g., tool wear, heat dissipation, power failure). Models such as Random Forest, XG Boost, and LSTM (for time-series patterns) will be evaluated. Class imbalance, if present, will be addressed using techniques like SMOTE or class weighting.
- **Model Evaluation**:
  The model's accuracy, precision, recall, F1-score, and confusion matrix will be used to evaluate its performance. Cross-validation and hyperparameter tuning will be applied to improve robustness and generalizability.

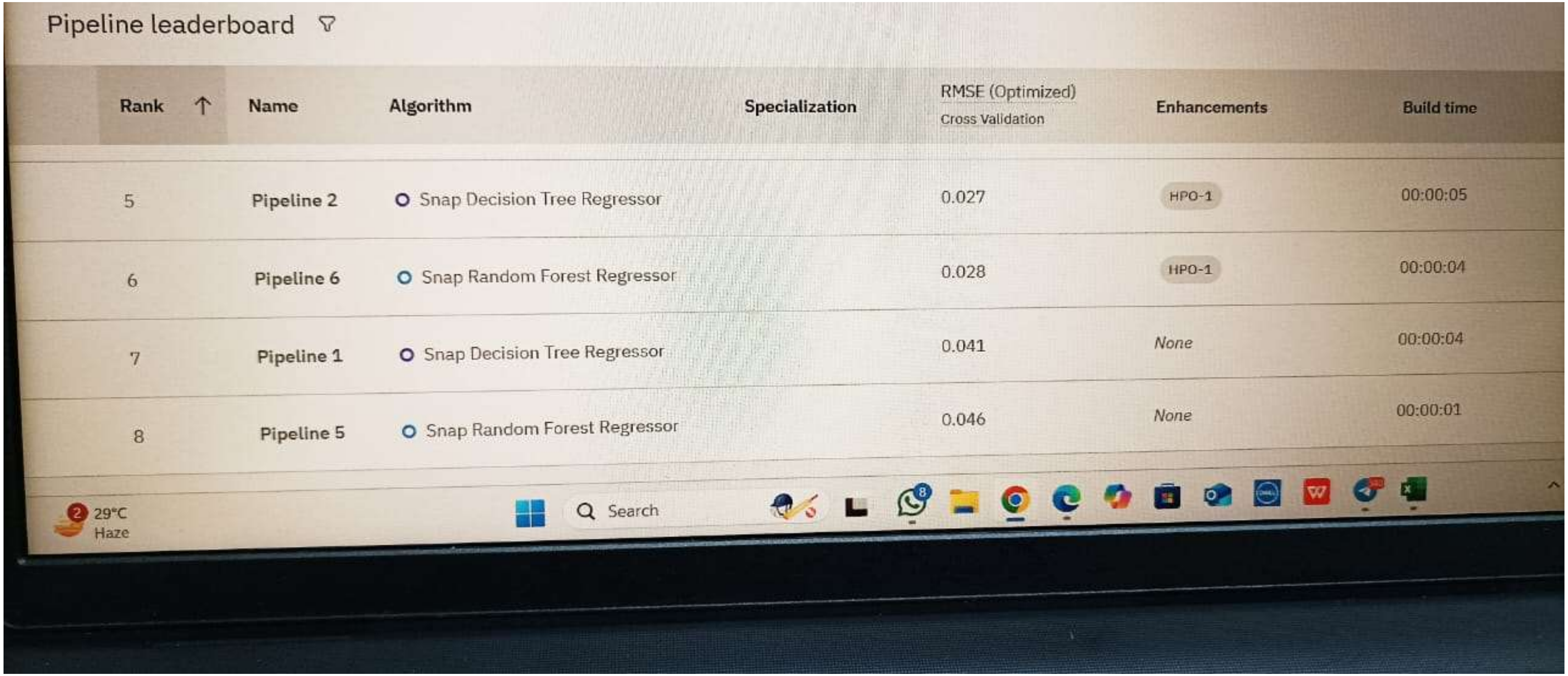
SYSTEM APPROACH

- **1. Sensor Data Collection:**
  Gather real-time and historical data from machines (e.g., temperature, pressure, vibration).
- **2. Data Preprocessing:**
  Clean the data by handling missing values, outliers, and normalize it for consistency.
- **3. Feature Engineering:**
  Extract useful features like statistical metrics and time-based trends to improve prediction.
- **4. Labeling Failures:**
  Assign labels to data based on failure types such as tool wear, power failure, etc.
- **5. Model Building:**
  Train classification models (e.g., Random Forest, XGBoost, LSTM) to predict failure types.
- **6. Model Evaluation:**
  Use metrics like accuracy, precision, recall, and F1-score to evaluate model performance.

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting bike counts. Here's an example structure for this section:
- **Algorithm Selection:**
  - Provide a brief overview of the chosen algorithm (e.g., time-series forecasting model, like ARIMA or LSTM) and justify its selection based on the problem statement and data characteristics.
- **Data Input:**
  - Specify the input features used by the algorithm, such as historical bike rental data, weather conditions, day of the week, and any other relevant factors.
- **Training Process:**
  - Explain how the algorithm is trained using historical data. Highlight any specific considerations or techniques employed, such as cross-validation or hyperparameter tuning.
- **Prediction Process:**
  - Detail how the trained algorithm makes predictions for future bike counts. Discuss any real-time data inputs considered during the prediction phase.

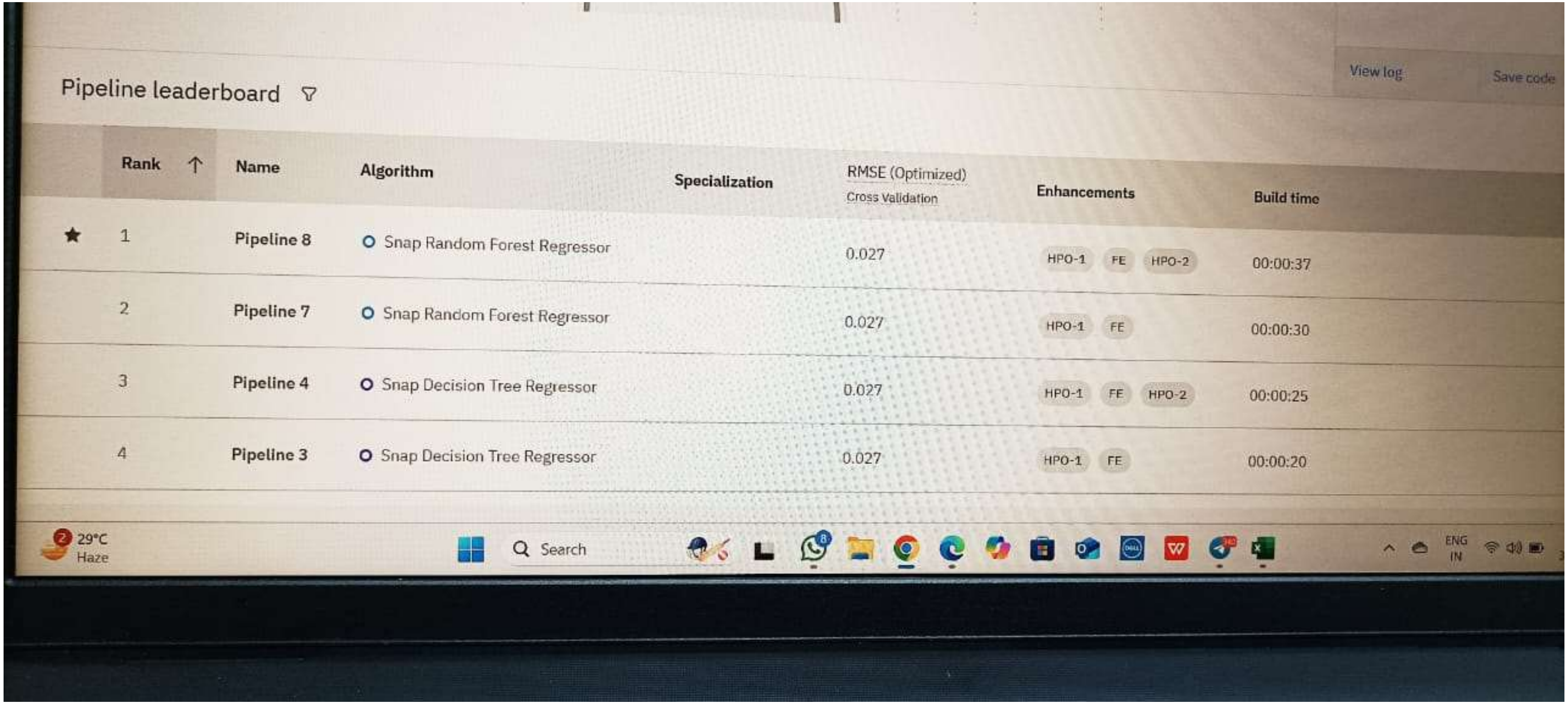
RESULT



Pipeline leaderboard

Rank	Name	Algorithm	Specialization	RMSE (Optimized) Cross Validation	Enhancements	Build time
5	Pipeline 2	Snap Decision Tree Regressor		0.027	HPO-1	00:00:05
6	Pipeline 6	Snap Random Forest Regressor		0.028	HPO-1	00:00:04
7	Pipeline 1	Snap Decision Tree Regressor		0.041	None	00:00:04
8	Pipeline 5	Snap Random Forest Regressor		0.046	None	00:00:01

RESULT

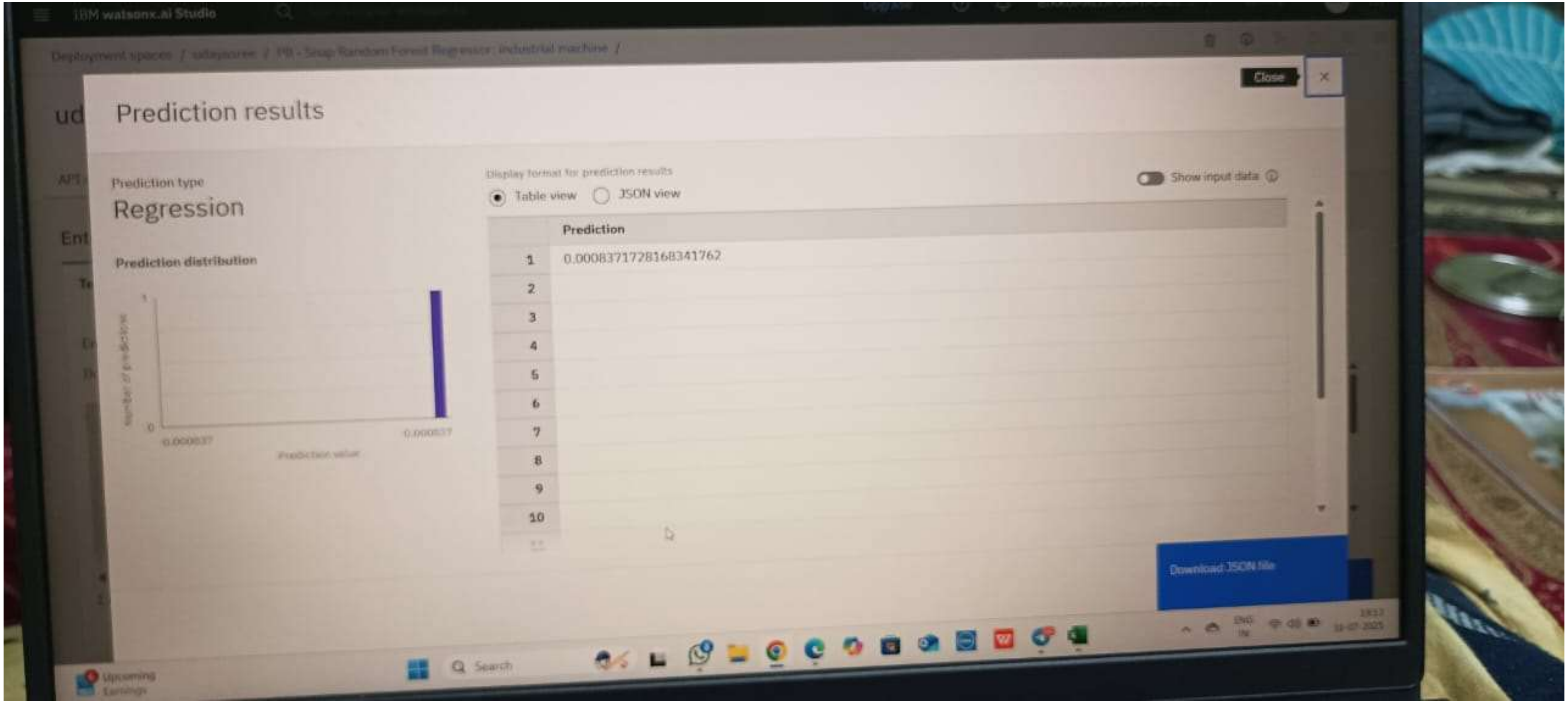


Pipeline leaderboard

View log Save code

	Rank	Name	Algorithm	Specialization	RMSE (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 8	Snap Random Forest Regressor		0.027	HPO-1 FE HPO-2	00:00:37
	2	Pipeline 7	Snap Random Forest Regressor		0.027	HPO-1 FE	00:00:30
	3	Pipeline 4	Snap Decision Tree Regressor		0.027	HPO-1 FE HPO-2	00:00:25
	4	Pipeline 3	Snap Decision Tree Regressor		0.027	HPO-1 FE	00:00:20

RESULT


IBM watsonx.ai Studio
Prediction results
Close
Prediction type
Regression
Prediction distribution
Display format for prediction results
Table view
JSON view
Show input data
Prediction
1 0.0008371728168341762
2
3
4
5
6
7
8
9
10
Download JSON file

CONCLUSION

This project demonstrates the effectiveness of using machine learning techniques to predict types of failures in industrial machinery based on sensor data. By analyzing real-time operational parameters and identifying patterns preceding failures like tool wear, heat dissipation, and power issues, the developed classification model enables proactive maintenance. Implementing such a system helps reduce unexpected downtime, optimize maintenance schedules, and lower operational costs, ultimately improving the reliability and efficiency of industrial operations. Continuous monitoring and model updates ensure adaptability to evolving machine conditions, making predictive maintenance a valuable tool for modern industries

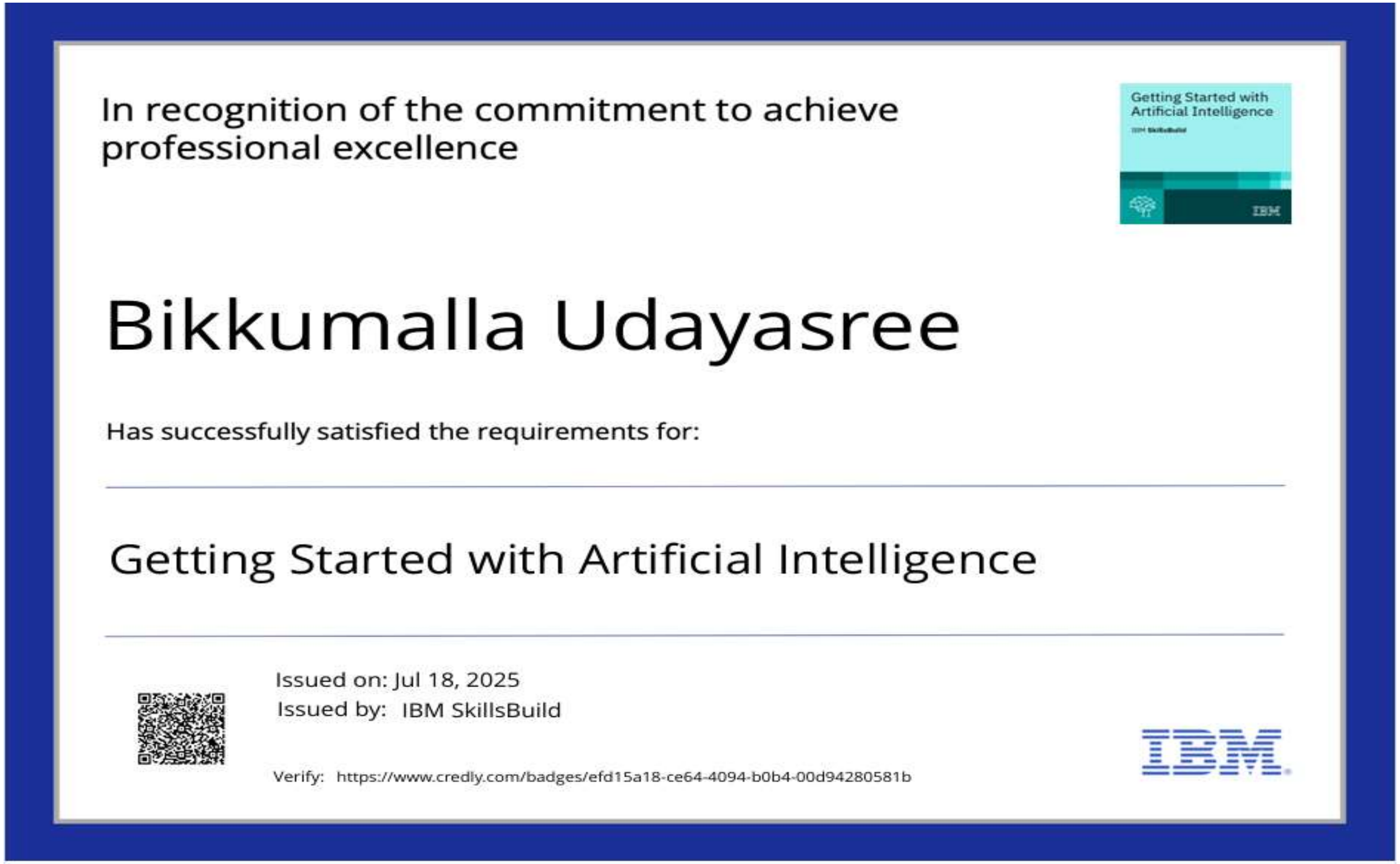
FUTURE SCOPE

- Integrate with IoT systems for real-time monitoring.
- Use advanced deep learning models for better accuracy.
- Add anomaly detection for unknown failure types.
- Deploy models on edge devices for faster predictions.
- Automate maintenance scheduling based on predictions.
- Improve model explainability to build user trust.
- Combine sensor data with other data sources for richer insights.

REFERENCES

- IBM Developer. Build and Deploy Machine Learning Models on IBM Watson Studio.
- https://developer.ibm.com
- IBM Cloud. IBM Watson Machine Learning Service – Model Deployment and Scoring.
- https://cloud.ibm.com/catalog/services/watson-machine-learning

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Bikkumalla Udayasree

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

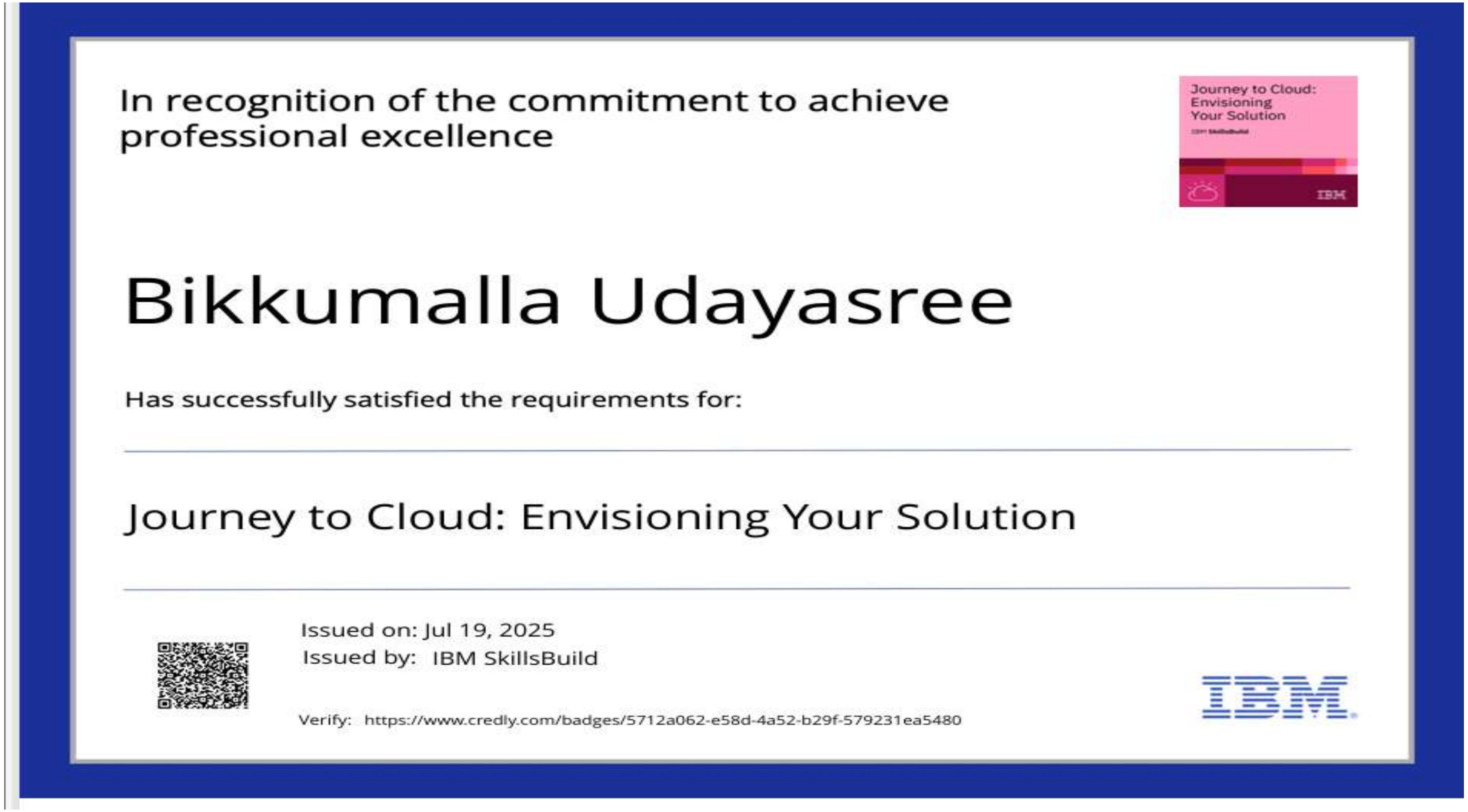
Issued on: Jul 18, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/efd15a18-ce64-4094-b0b4-00d94280581b

IBM

edunet foundation

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM

Bikkumalla Udayasree

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/5712a062-e58d-4a52-b29f-579231ea5480

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

BIKKUMALLA UDAYASREE

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