CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By:
Bappaditya Banerjee
College Name :St.Thomas' College of Engineering & Technology
Department : Information Technology

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 proposed system aims to address the challenge of predicting the required bike count at each hour to ensure a stable supply of rental bikes. This involves leveraging data analytics and machine learning techniques to forecast demand patterns accurately. The solution will consist of the following components:

- **Data Collection:**

  Gather historical data on industrial machines along with their failure type from Kaggle dataset

- **Data Preprocessing:**

  Clean and preprocess the collected data from Kaggle dataset to handle missing values, outliers, and inconsistencies.

- **Machine Learning Algorithm:**

  Implement a machine learning algorithm, such as Decision Tree Classifier, Random Forest Classifier or SVM to train the model

- **Deployment:**
  - Develop a user-friendly interface or application that provides real-time predictions for the type of failure in a particular industrial machine
  - 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 metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or other relevant metrics.
  - Fine-tune the model based on feedback and continuous monitoring of prediction accuracy.

Result: This will enable proactive maintenance, reducing downtime and operational costs.

SYSTEM APPROACH

System Requirements :

1)IBM Cloud platform

2)IBM watsonx.ai Studio

3)IBM watsonx.ai Runtime Service

4)IBM cloud object storage

Libraries used :

scikit-learn library of Python

ALGORITHM & DEPLOYMENT

Algorithm Selection:

Snap Random Forest Classifier : The **Snap Random Forest Classifier** is an ensemble learning model that builds multiple decision trees and combines their outputs to improve prediction accuracy and robustness on structured data like CSV files.

It is considered effective for this project because it combines multiple decision trees to handle CSV-based datasets robustly, offering high accuracy, resistance to overfitting, and strong performance on diverse features.

Data Input:

UDI, Product ID, Type, Air Temperature (in Kelvin), Process Temperature(in Kelvin), Rotational Speed(rpm),Torque(Nm), Tool wear(minimum value),

Target

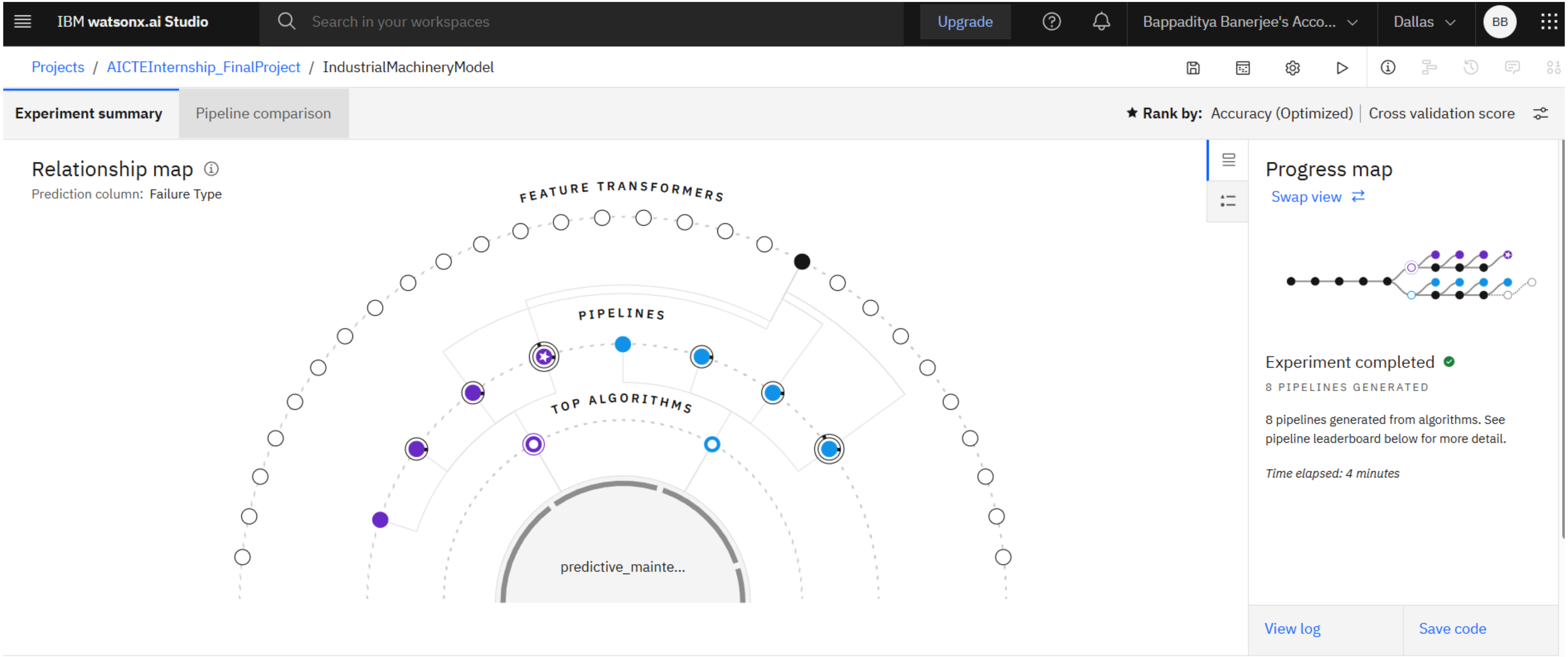
Training Process:

The Snap Random Forest Classifier is trained using historical data by creating multiple decision trees on bootstrapped samples, using historical data by creating multiple decision trees on bootstrapped samples, using techniques like **cross-validation** to assess performance and **hyperparameter tuning** (e.g., grid search) to optimize parameters like tree depth and number of estimators.

Prediction Process:

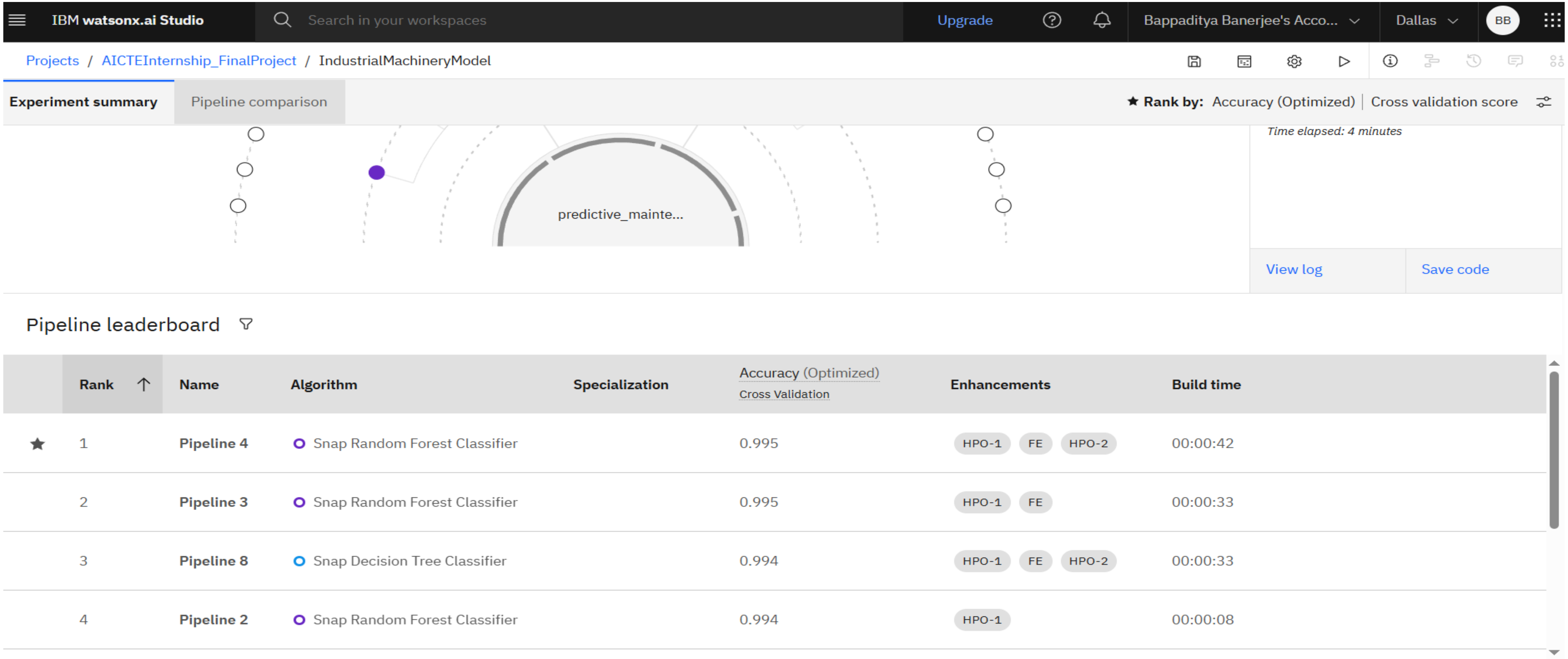
The inputs are given to the ML model and accordingly it predicts the Failure type with a certain accuracy.After training, the model could make correct predictions on the Kaggle dataset values which was given for this project.

RESULT



edunet foundation

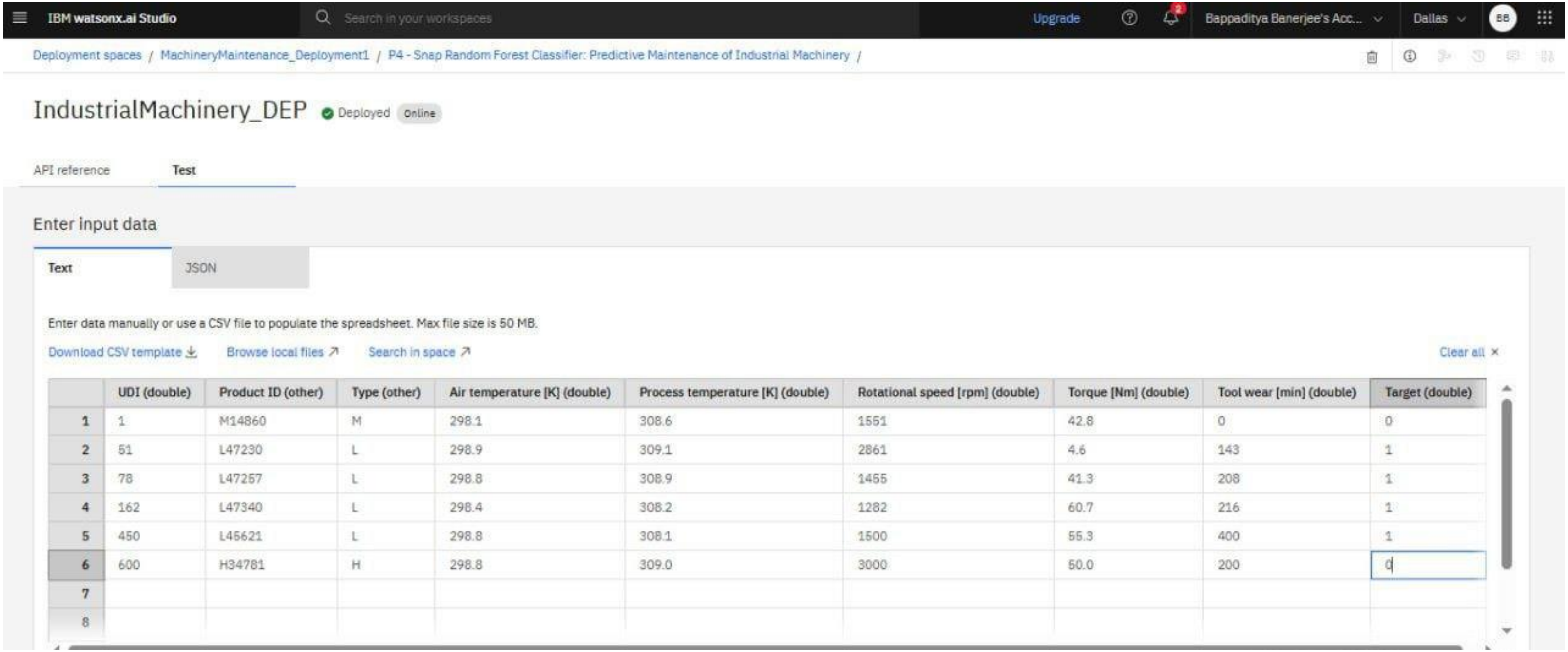
RESULT



	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	00:00:42
	2	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:33
	3	Pipeline 8	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:33
	4	Pipeline 2	Snap Random Forest Classifier		0.994	HPO-1	00:00:08

edunet foundation

RESULT


IBM watsonx.ai Studio
Search in your workspaces
Upgrade
Bappaditya Banerjee's Acc...
Dallas
BB
Deployment spaces / MachineryMaintenance_Deployment1 / P4 - Snap Random Forest Classifier: Predictive Maintenance of Industrial Machinery /
IndustrialMachinery_DEP
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	M14860	M	298.1	308.6	1551	42.8	0	0
2	51	L47230	L	298.9	309.1	2861	4.6	143	1
3	78	L47257	L	298.8	308.9	1455	41.3	208	1
4	162	L47340	L	298.4	308.2	1282	60.7	216	1
5	450	L45621	L	298.8	308.1	1500	55.3	400	1
6	600	H34781	H	298.8	309.0	3000	50.0	200	0
7									
8									

RESULT

Prediction results

×

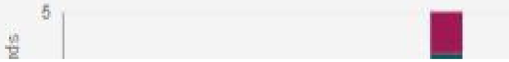
Prediction type

Multiclass classification

Prediction percentage

No Failure
Power Failure
Tool Wear Failure
Overstrain Failure

Confidence level distribution



Display format for prediction results

Table view JSON view

Show input data

	Prediction	Confidence
1	No Failure	100%
2	Power Failure	100%
3	Tool Wear Failure	100%
4	Overstrain Failure	96%
5	No Failure	30%
6	No Failure	90%
7		
8		
9		
10		
11		
12		
13		
14		

CONCLUSION

Findings:

We could predict the type of the failure in the industry machinery with this Machine Learning project. It is very effective in identifying the patterns that precede a failure so we can take precautionary measures from beforehand

Challenges encountered:

i)IBM cloud platform is sometimes slow due to lot of traffic

ii)The ML model can predict the type of failure but not with 100% accuracy all the time.

FUTURE SCOPE

- Additional data sources could be incorporated to ensure better prediction outcomes
- Algorithm could be optimized for better performance
- The type of failures that this model can predict could be increased
- Other ML techniques can also be integrated for better outcomes

REFERENCES

Research Paper:

https://www.researchgate.net/publication/354386593_Adoption_of_machine_learning_technology_for_failure_prediction_in_industrial_maintenance_a_systematic_review

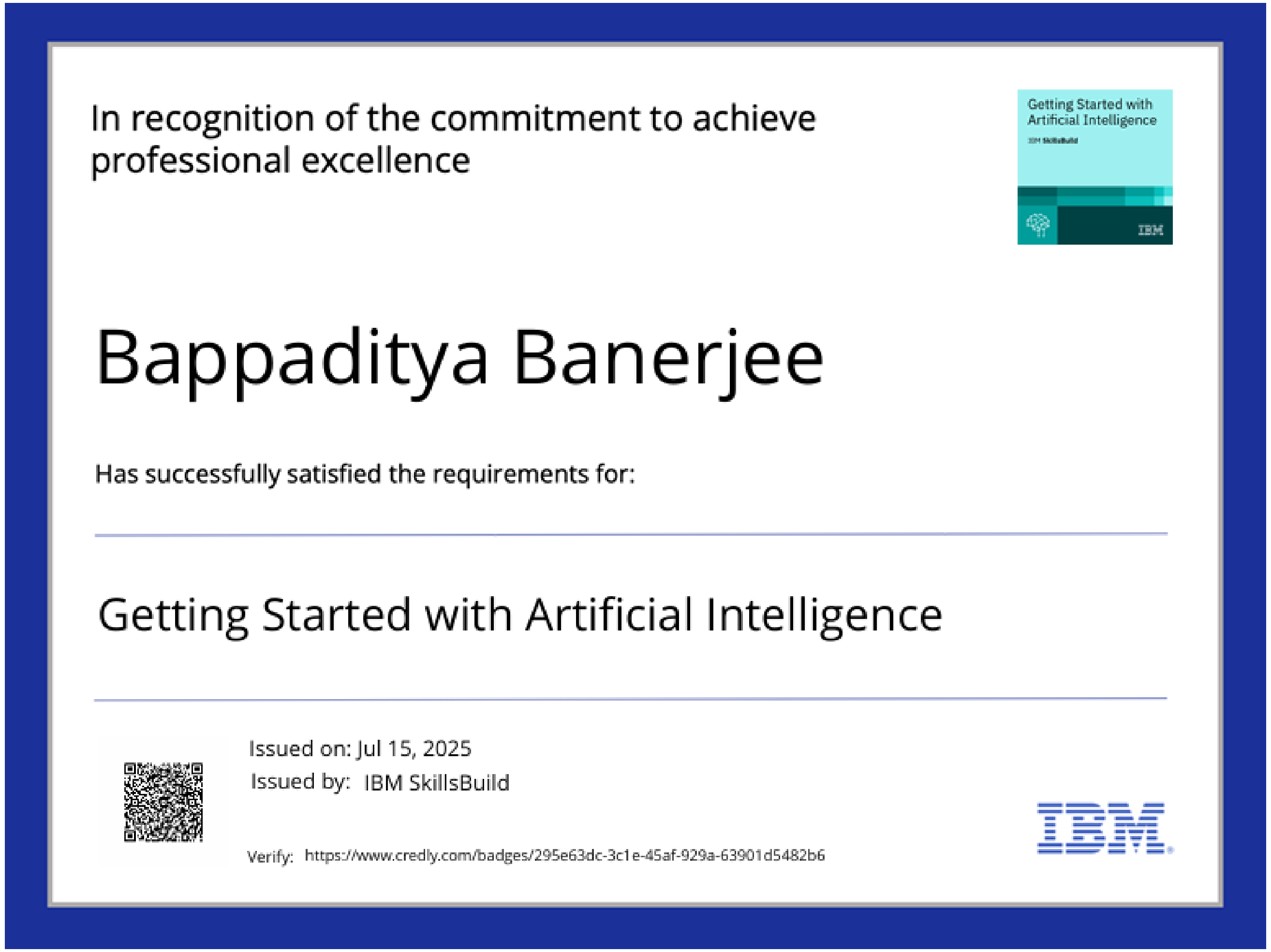
Machine Learning Model :

https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

Evaluation of ML model:

https://www.geeksforgeeks.org/machine-learning/machine-learning-model-evaluation/

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Bappaditya Banerjee

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 15, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/295e63dc-3c1e-45af-929a-63901d5482b6

IBM CERTIFICATIONS

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Bappaditya Banerjee

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