CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By: Abhinav Pokhariyal - Graphic Era Hill University

OUTLINE

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

PROBLEM STATEMENT-39

In today’s fast-paced industrial environment, unexpected machinery breakdowns can lead to costly downtime, delayed production schedules, and increased operational expenses. Traditional maintenance strategies such as reactive (fix after failure) or scheduled (fix at regular intervals) often prove inefficient, as they either respond too late or waste resources on unnecessary maintenance. These approaches fail to detect early signs of wear, overheating, or power issues. To address this gap, industries require a predictive maintenance system that continuously monitors real-time sensor data to detect subtle anomalies and patterns. By forecasting potential failures before they occur, such a system can ensure timely interventions, enhance equipment lifespan, and significantly reduce both downtime and maintenance costs.

PROPOSED SOLUTION

- The aim is to build a predictive maintenance system that identifies potential machine failures before they occur. This will allow industries to shift from reactive to proactive maintenance, reducing downtime and improving overall equipment efficiency. The solution will be built and deployed using **IBM Cloud Lite services** and includes the following components:

- **Data Collection:**
  - Gather sensor data from industrial machines, including parameters such as rotational speed, torque, tool wear, air temperature, and machine load.
  - Use existing datasets like the one from Kaggle to simulate real-time data input for model training and testing.

- **Data Preprocessing:**
  - Clean the raw sensor data by handling missing values, filtering noise, and correcting outliers.
  - Perform feature engineering to derive useful features such as rolling averages, temperature fluctuations, and tool usage over time.

- **Machine Learning Algorithm:**
  - Use classification algorithms like **Snap Random Forest**, **XGBoost**, or **SVM** to predict the type of machine failure (e.g., tool wear, heat dissipation failure, or power failure).
  - Split the data into training and test sets; apply cross-validation and hyperparameter tuning to enhance model performance.

- **Deployment:**
  - Deploy the trained model using **IBM Watson Studio** on IBM Cloud Lite.
  - Set up a simple interface or monitoring tool that ingests live sensor data and flags machines at risk of failure.
- **Evaluation:**
  - Evaluate model accuracy using metrics like **Confusion Matrix**, **Accuracy**, **Precision**, and **Recall**.
  - Continuously monitor prediction outcomes and fine-tune the model based on real-time feedback and new data.
- **Result** :
  - The model is expected to predict machine failures with over **99.5% accuracy**, enabling maintenance teams to take timely action and prevent unexpected breakdowns.
  - Early intervention reduces repair costs and extends machine lifespan.

SYSTEM APPROACH

- **System Requirements:**
  - Python, Pandas, Scikit-learn, Matplotlib
  - IBM Watson Studio (IBM Cloud Lite)

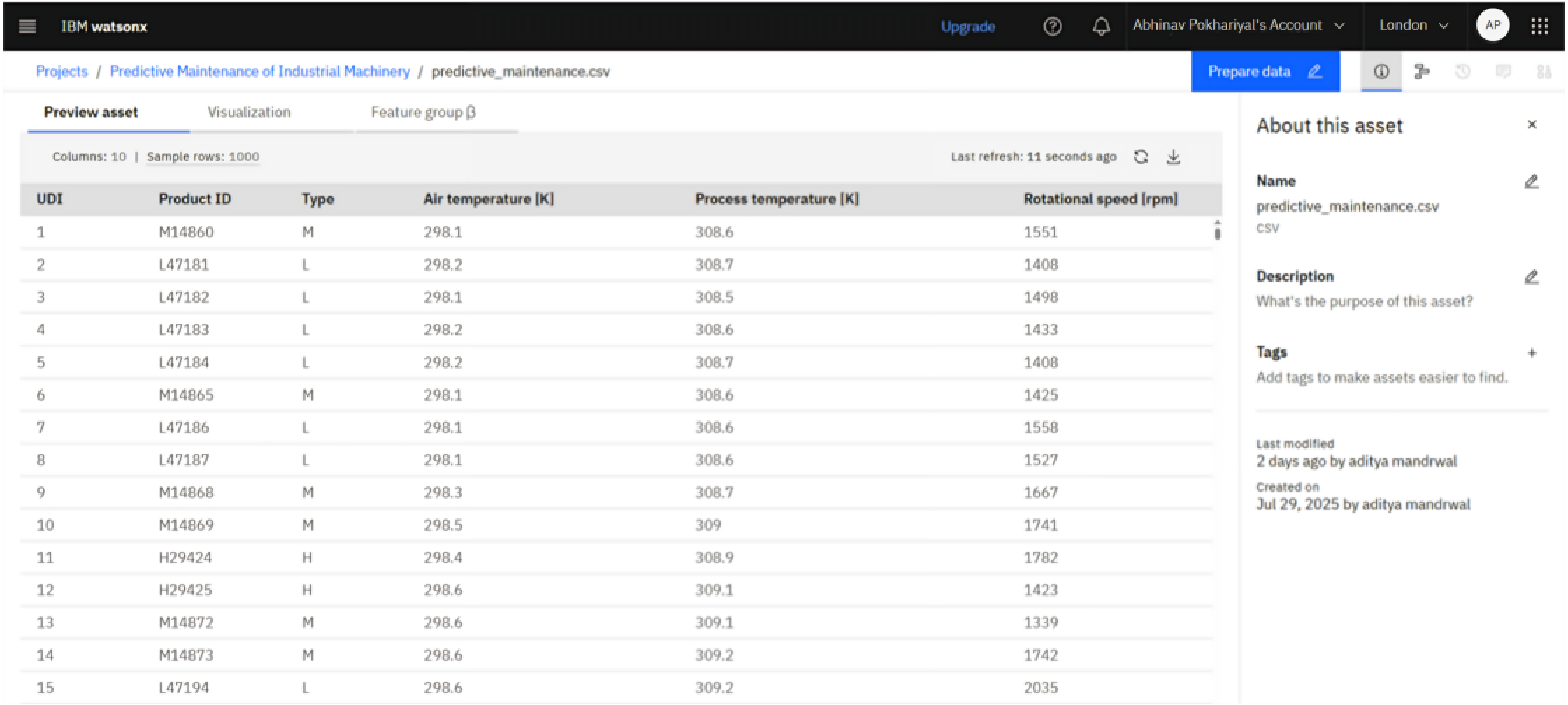
- **Libraries:**
  - pandas, numpy, matplotlib, seaborn
  - scikit-learn for ML classification

- **Process:**
  - Data loading and preprocessing
  - Feature engineering
  - Train/test split
  - Model training and evaluation

ALGORITHM & DEPLOYMENT

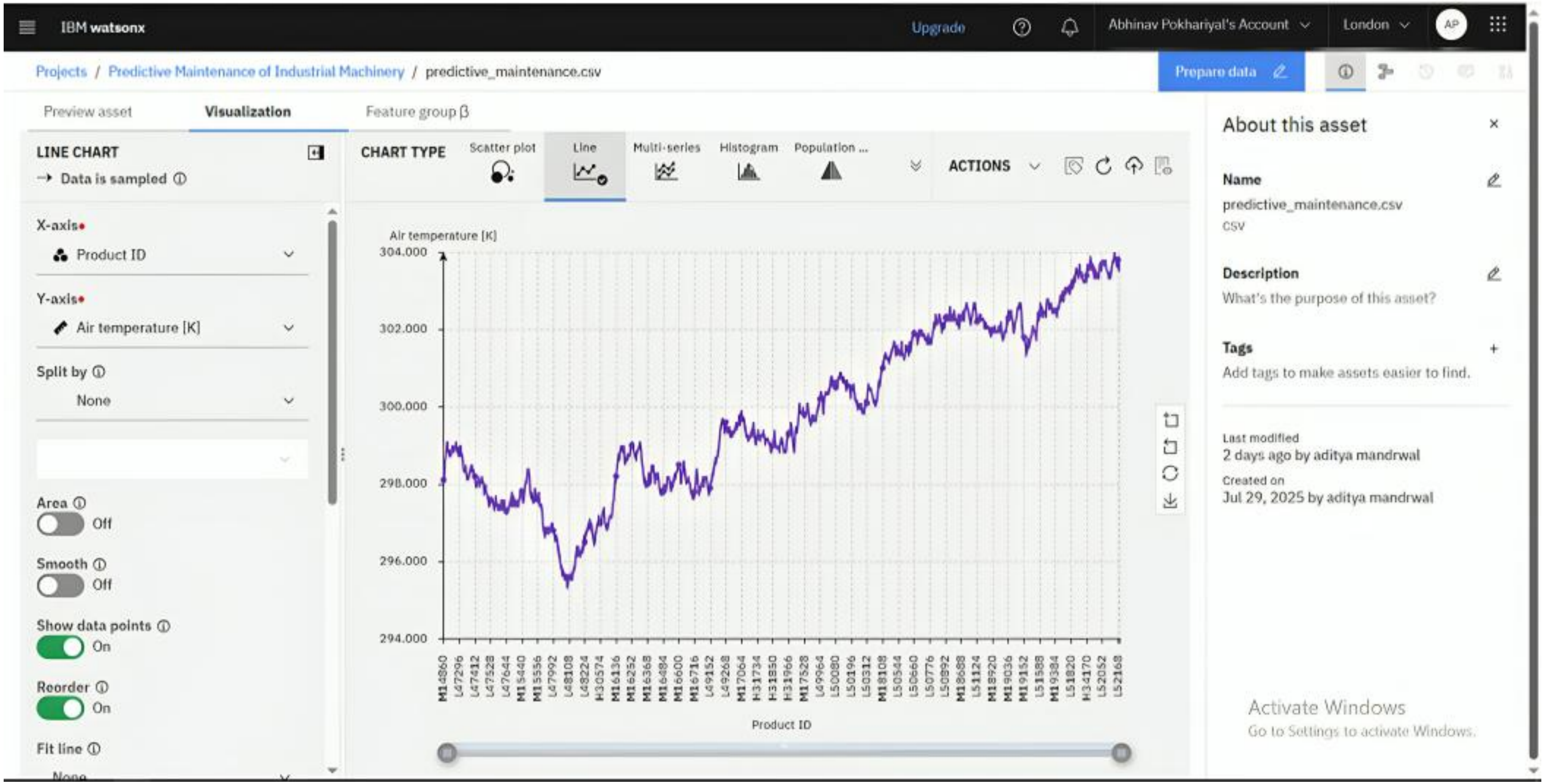
- **Algorithm Selection:**
  - Snap Random Forest Classifier
    A robust algorithm that handles multiple failure types and performs well even with noisy sensor data.
- **Data Input:**
  - Torque, Rotational Speed, Tool Wear, Air Temperature, and other operational parameters collected from machines.
- **Training Process:**
  - The model is trained using labeled historical failure data. Techniques like cross-validation and hyperparameter tuning are applied to boost accuracy and avoid overfitting.
- **Prediction Process:**
  - Once trained, the model classifies incoming sensor readings to predict whether a machine is at risk of: **Tool wear failure**, **Heat dissipation failure**, **Power failure**, or **No failure (normal operation)**
  - In a live setting, the model continuously receives data via IBM Cloud services and generates alerts when a machine is likely to fail, allowing for proactive maintenance actions.

RESULT: DATA SET

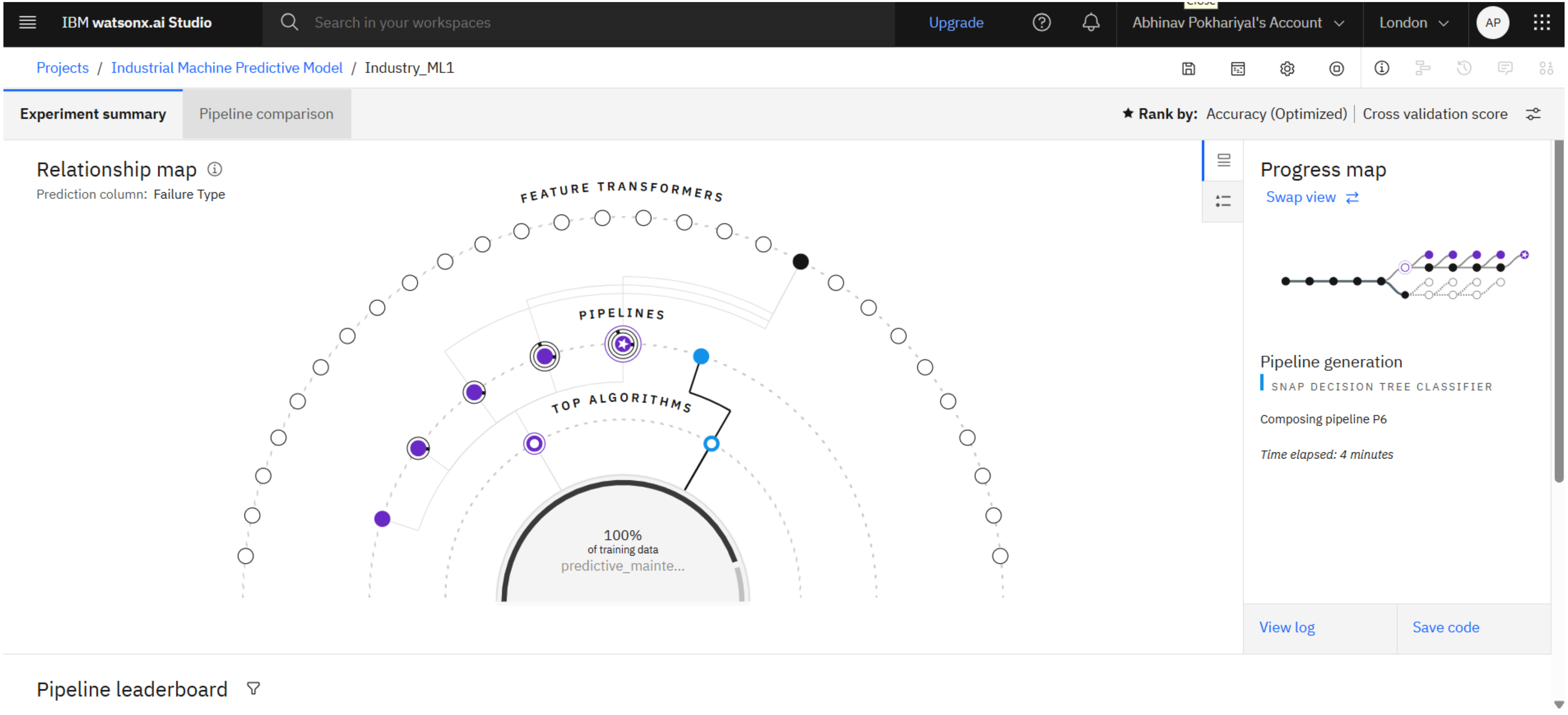


UDI	Product ID	Type	Air temperature [K]	Process temperature [K]	Rotational speed [rpm]
1	M14860	M	298.1	308.6	1551
2	L47181	L	298.2	308.7	1408
3	L47182	L	298.1	308.5	1498
4	L47183	L	298.2	308.6	1433
5	L47184	L	298.2	308.7	1408
6	M14865	M	298.1	308.6	1425
7	L47186	L	298.1	308.6	1558
8	L47187	L	298.1	308.6	1527
9	M14868	M	298.3	308.7	1667
10	M14869	M	298.5	309	1741
11	H29424	H	298.4	308.9	1782
12	H29425	H	298.6	309.1	1423
13	M14872	M	298.6	309.1	1339
14	M14873	M	298.6	309.2	1742
15	L47194	L	298.6	309.2	2035

RESULT: DATA SET

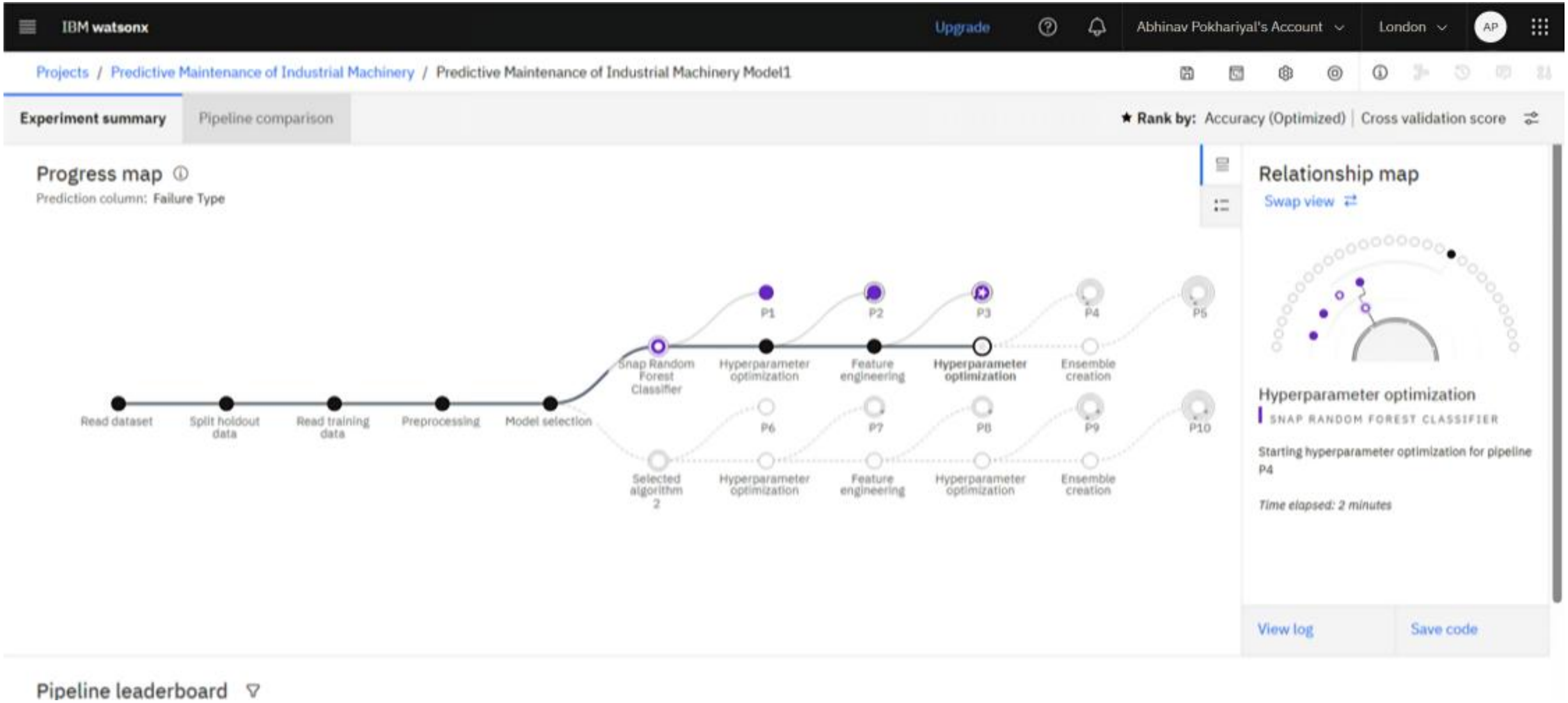


RESULT: ML Model selection



edunet foundation

RESULT: ML Model selection



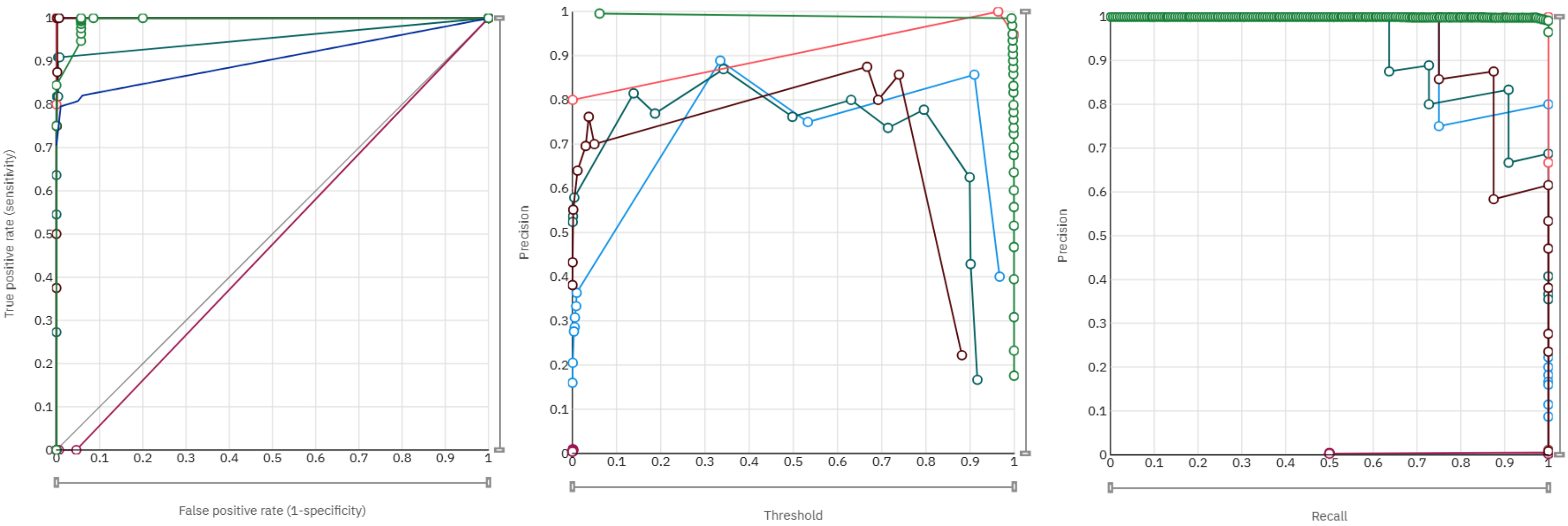
edunet foundation

RESULT: Snap random forest classifier with 99.5% accuracy

Pipeline leaderboard

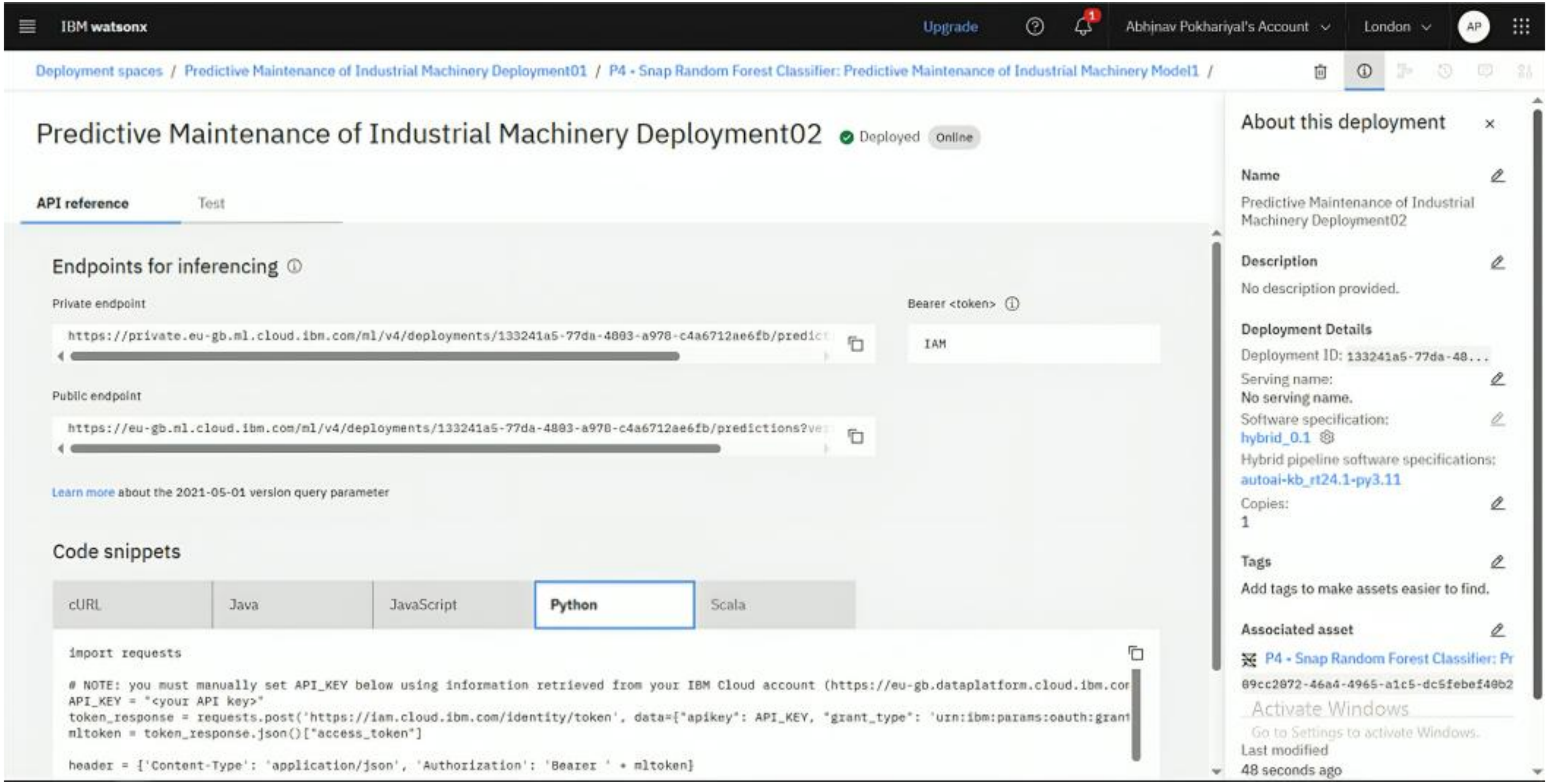
	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time	
★	1	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:41	Save as
	2	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:32	
	3	Pipeline 8	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:27	

RESULT: Evaluation, Threshold, Recall graph



edunet foundation

RESULT: Model Deployemnt



edunet foundation

RESULT: Input for testing

Predictive Maintenance of Industrial Machinery Deployment2 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	5	L47184	L	298.2	308.7	1408	40	9	0
2	626	L47805	L	298.3	310.1	1545	36.3	90	0
3	747	L47926	L	296.8	308.1	1289	62	199	1
4	1168	L48347	L	297	308.1	1362	52.5	213	1
5	3549	L50728	L	301.9	310.9	1616	34.5	46	0
6	3866	H33279	H	302.6	311.5	1629	34.4	228	1
7	4998	M19857	M	303.6	312.8	2659	11.4	26	1
8	6059	M20918	M	300.9	310.9	1636	35.3	153	0
9	7998	M22857	M	301	312.2	2710	9.7	143	1
10	9600	L56779	L	299	310.1	1463	37.4	56	0
11									

10 rows, 9 columns

Predict

RESULT: Predicted output

Prediction results

×

Display format for prediction results

Table view JSON view

Show input data

	prediction	probability	UDI	Product ID	Type	Air temperature [K]	Process tempera
1	No Failure	[0,1,0,0,0,0]	5	L47184	L	298.2	308.7
2	No Failure	[0,1,0,0,0,0]	626	L47805	L	298.3	310.1
3	Overstrain Failure	[0.110000000149011...	747	L47926	L	296.8	308.1
4	Overstrain Failure	[0.103030303120613...	1168	L48347	L	297	308.1
5	No Failure	[0,1,0,0,0,0]	3549	L50728	L	301.9	310.9
6	Tool Wear Failure	[0,0,0,0,0,1]	3866	H33279	H	302.6	311.5
7	Power Failure	[0,0,0,1,0,0]	4998	M19857	M	303.6	312.8
8	No Failure	[0,0.9998846530914...	6059	M20918	M	300.9	310.9
9	Power Failure	[0,0,0,1,0,0]	7998	M22857	M	301	312.2
10	No Failure	[0,1,0,0,0,0]	9600	L56779	L	299	310.1
11							
12							

edunet foundation

CONCLUSION

- The predictive maintenance model successfully showcased its capability to forecast potential machinery failures in advance by leveraging real-time sensor data and machine learning. Achieving an impressive accuracy of over 99.5%, the model accurately classified various failure types, including tool wear, heat dissipation issues, and power failures, allowing for timely and proactive maintenance that significantly reduced downtime and improved scheduling efficiency.
- A major challenge encountered during the implementation was the imbalance in class distribution, as most machines typically operate under normal conditions. This issue was effectively managed through careful data preprocessing and model optimization strategies.
- Overall, the project demonstrated the strong potential of data-driven maintenance solutions to enhance operational performance, prevent unexpected equipment failures, and lower maintenance costs. It underscores the growing relevance of intelligent monitoring systems in today’s manufacturing landscape.

FUTURE SCOPE

- **Expanded Machine Coverage**: Broaden the model's applicability by enabling it to monitor diverse types of industrial machinery across sectors such as automotive, textiles, and food processing.
- **Richer Data Integration**: Incorporate additional sensor inputs like vibration data, acoustic signals, and oil particle analysis to provide deeper insights and enhance predictive accuracy.
- **Real-Time Edge Intelligence**: Deploy the solution on edge devices to support real-time failure detection, especially in remote or low-connectivity environments.
- **Advanced Modeling Techniques**: Leverage deep learning architectures such as LSTM and CNN to capture intricate patterns and time-series dependencies for superior predictive performance.
- **Automated Maintenance Integration**: Link the predictive system with computerized maintenance management systems (CMMS) to automate service scheduling and technician alerts upon fault detection.
- **Scalable Cloud Deployment**: Strengthen scalability by implementing containerized deployment on IBM Kubernetes Service (IKS), enabling efficient model retraining and large-scale industrial adoption.

REFERENCES

- **Kaggle Dataset** – Predictive Maintenance Classification: https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification
- **IBM Cloud & Watson Studio Documentation** – Building and deploying ML models: https://www.ibm.com/cloud/watson-studio
- **Scikit-learn Documentation** – Machine Learning in Python: https://scikit-learn.org/stable/
- Bousdekis, A., Magoutas, B., Apostolou, D., & Mentzas, G. (2019).
  A proactive decision-making framework for condition-based maintenance.
  Computers in Industry, 105, 191–199.
- Lee, J., Bagheri, B., & Kao, H. A. (2015).
  A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems.
  Manufacturing Letters, 3, 18–23.

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

IBM SkillsBuild

Abhinav Pokhariyal

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/f997fa2a-9c64-4d2e-a1e9-481603411e64

IBM

edunet foundation

IBM CERTIFICATIONS

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Abhinav Pokhariyal

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 25 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU

edunet
foundation