CAPSTONE PROJECT

PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By:
Rahul Paul
-Department of Computer Science and Engineering
- Tezpur University.

edunet foundation

OUTLINE

edunet
foundation

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 will utilize IBM's AutoAI on the watsonx.ai platform to build a high-accuracy predictive maintenance model. AutoAI automates the end-to-end process of machine learning model development, allowing for the rapid creation and deployment of a solution.

- The solution will consist of the following components:
- **Data Collection:**
  - The project will use the "Machine Predictive Maintenance Classification" dataset from Kaggle. This dataset contains historical operational data and failure records for industrial machinery.
- **Data Preprocessing (Automated by AutoAI):**
  - The AutoAI tool will automatically handle essential data preparation steps, including cleaning the data, handling any missing values, and encoding features for optimal model performance
- **Automated Model Building (AutoAI):**
  - AutoAI will streamline the model creation process through:
    - **Algorithm Selection:** It will test and evaluate multiple classification algorithms to find the most suitable one for predicting machine failures.
    - **Feature Engineering:** It will automatically create new, insightful features from the raw sensor data to improve the model's predictive power.
    - **Hyperparameter Optimization:** It will fine-tune the selected algorithms to achieve the highest possible accuracy.

PROPOSED SOLUTION

- **Deployment:**
  - The best-performing model pipeline discovered by AutoAI will be deployed as a web service on IBM Cloud. This creates a scalable and reliable API endpoint for making real-time failure predictions.
- **Evaluation:**
  - The model's ability to predict failures will be rigorously evaluated using key classification metrics:
    - **Accuracy:** The overall percentage of correct predictions.
    - **Precision and Recall:** To measure the model's performance for specific failure types.
    - **F1-Score:** The harmonic mean of precision and recall, providing a single score for model quality.
    - **Confusion Matrix:** To visualize the model's accuracy across different failure classes.

SYSTEM APPROACH

- SYSTEM APPROACH
- The development strategy for the predictive maintenance system is centered around the integrated and automated environment of IBM's cloud services.
- **System Requirements:**
  - An active IBM Cloud Lite account.
  - Provisioned instances of the following core services:
    - watsonx.ai Studio
    - Watson Machine Learning
- **Libraries and Frameworks Utilized by AutoAI:**
  - While AutoAI provides a no-code interface, it leverages a powerful stack of open-source machine learning libraries in the background, including:
    - **Scikit-learn:** For a comprehensive suite of data preprocessing tools and classification models.
    - **XGBoost & LightGBM:** For state-of-the-art gradient boosting algorithms known for their high performance.

ALGORITHM & DEPLOYMENT

- **Algorithm Selection (Automated by AutoAI):**
  - AutoAI will run a competition among various classification algorithms such as Logistic Regression, Decision Trees, Random Forests, and Gradient Boosting Machines. It will automatically select and present the top-performing algorithms based on their predictive accuracy on the dataset.
- **Data Input:**
  - The model will be trained on the Kaggle "Machine Predictive Maintenance Classification" dataset. The input features include sensor readings like air temperature, process temperature, rotational speed, torque, and tool wear, along with the corresponding failure type.
- **Training Process (Automated by AutoAI):**
  - The training process is fully automated. AutoAI ingests the dataset, splits it for training and validation, and then systematically applies a wide range of data transformations, feature engineering techniques, and model training routines. It performs hyperparameter optimization to ensure the final model is highly tuned for this specific task.
- **Prediction Process:**
  - Once deployed, the model will function via an API. A system can send a set of current sensor readings from a machine to this API. The model will then return a prediction, indicating whether the machine is operating normally or is likely to experience a specific type of failure soon.
- **Deployment:**
  - The final, optimized model pipeline from AutoAI will be promoted to a deployment space within the Watson Machine Learning service. An online (real-time) deployment will be created, which provides a REST API endpoint for integration with other applications.

RESULT

DataSet:

Projects / Predictive Maintenance of Industrial Machinery / predictive_maintenance.csv

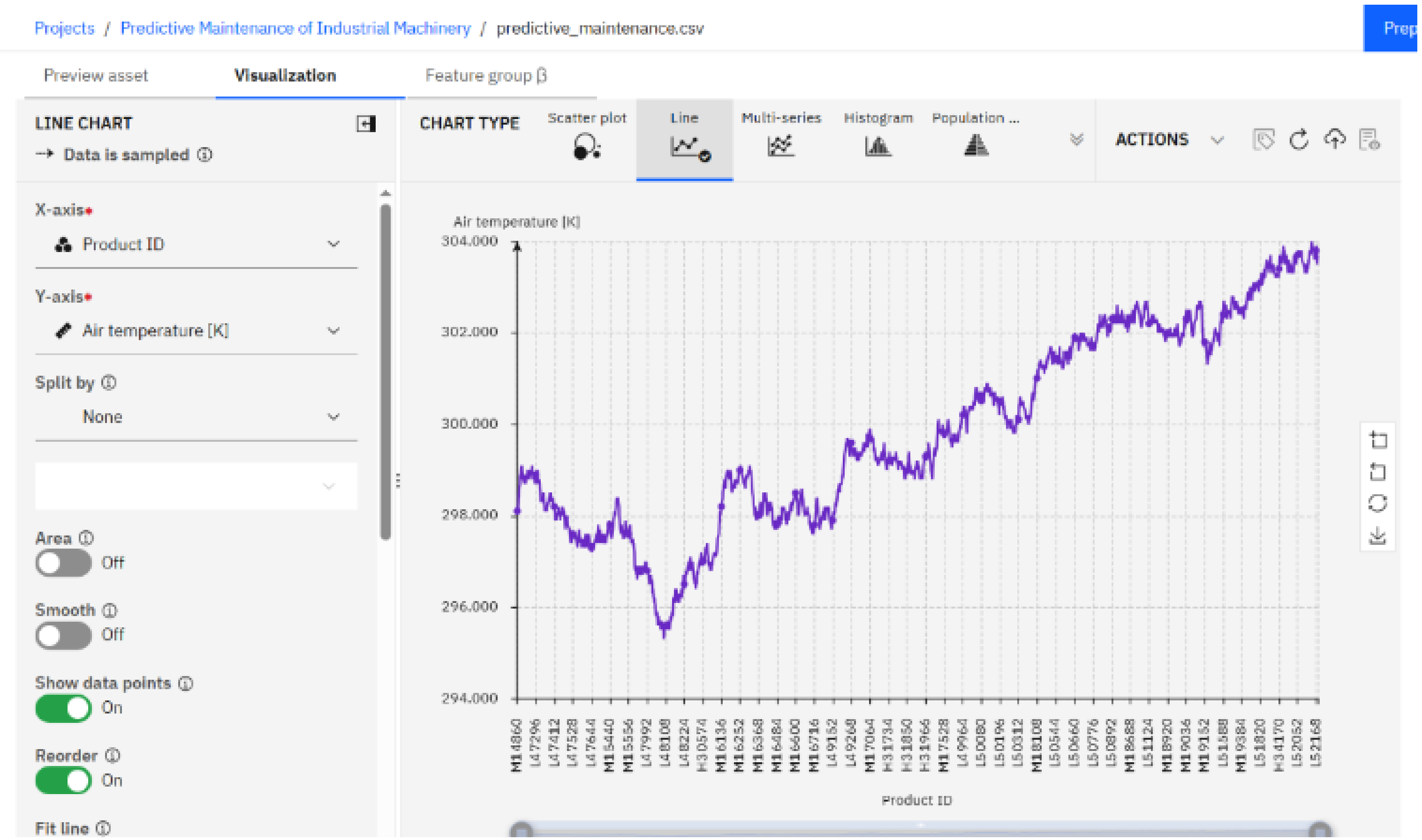
Preview asset | Visualization | Feature group β

Columns: 10 | Sample rows: 1000 — Last refresh: 11 seconds ago

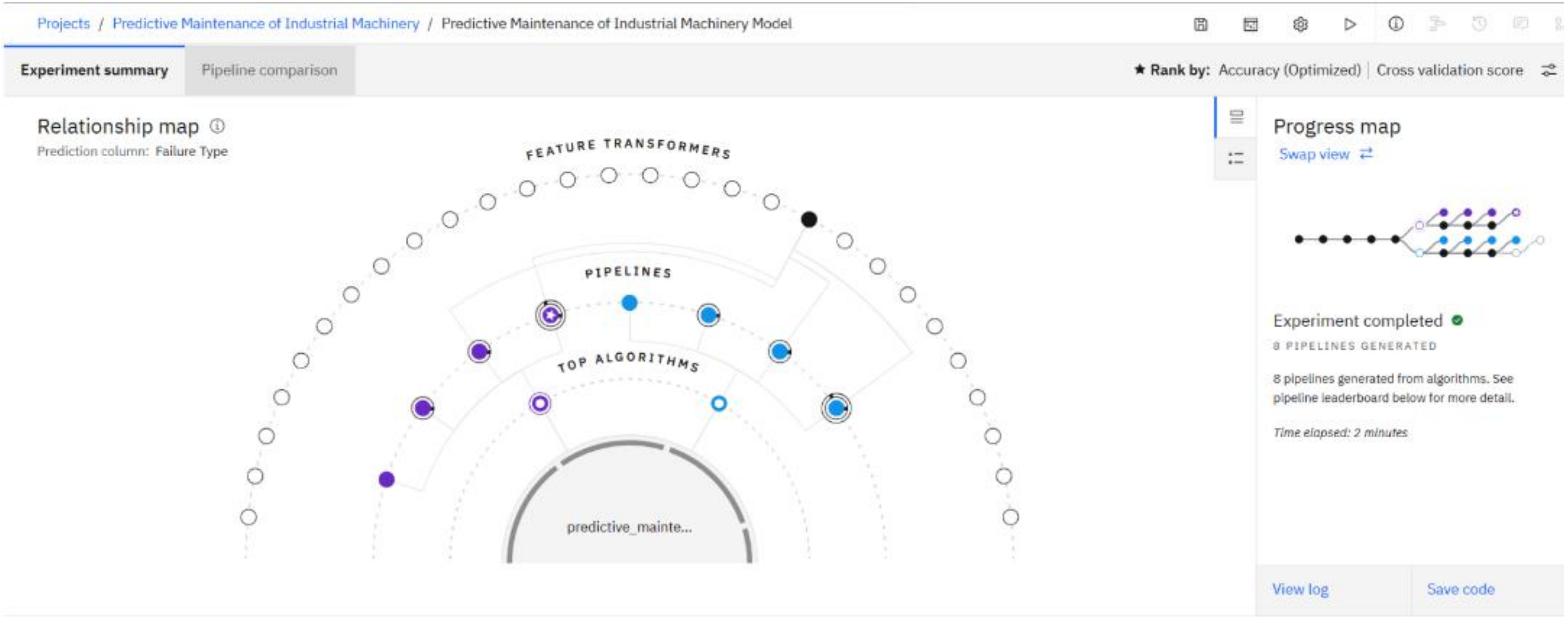
UDI	Product ID	Type	Air temperature [K]	Process temperature [K]	Rotational speed [rpm]
1	M14860	M	298.1	308.6	1551
2	L47181	L	298.2	308.7	1408
3	L47182	L	298.1	308.5	1498
4	L47183	L	298.2	308.6	1433
5	L47184	L	298.2	308.7	1408
6	M14865	M	298.1	308.6	1425
7	L47186	L	298.1	308.6	1558
8	L47187	L	298.1	308.6	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

DATASET ANALYSIS:

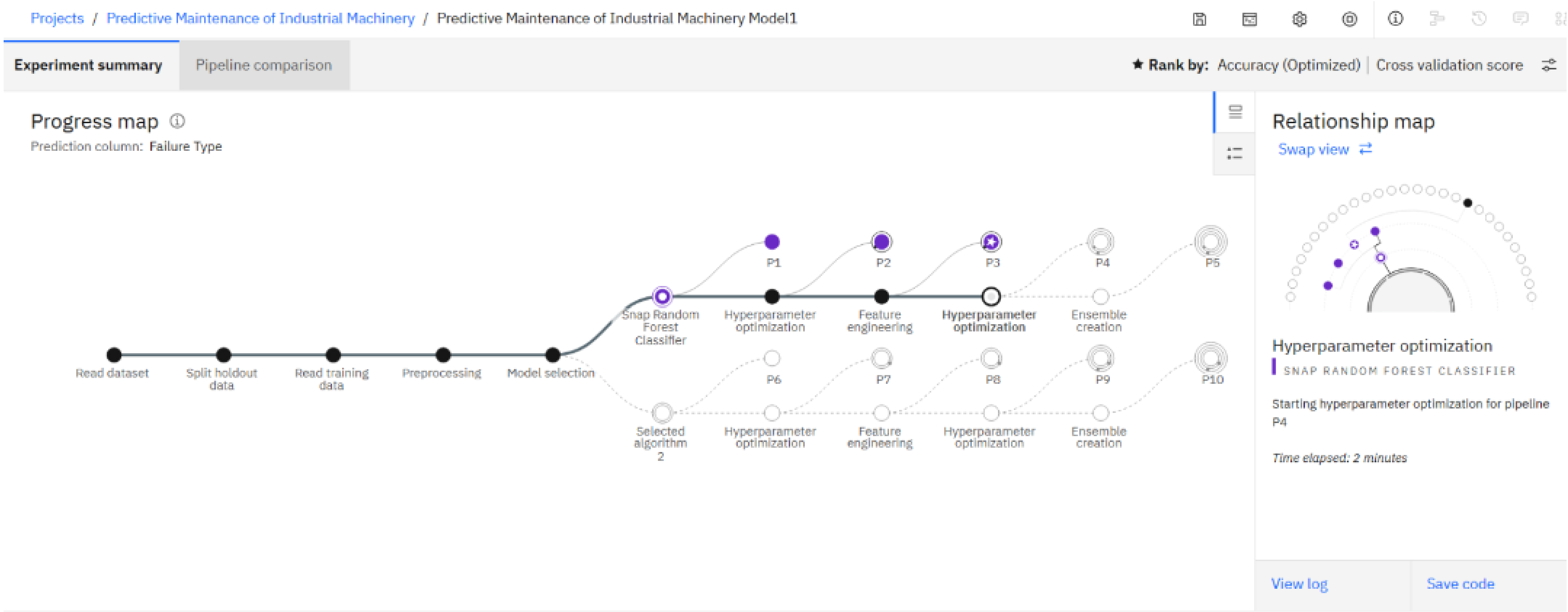


RESULT



Projects / Predictive Maintenance of Industrial Machinery / Predictive Maintenance of Industrial Machinery Model
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
8 PIPELINES GENERATED
8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 2 minutes
View log
Save code

RESULT



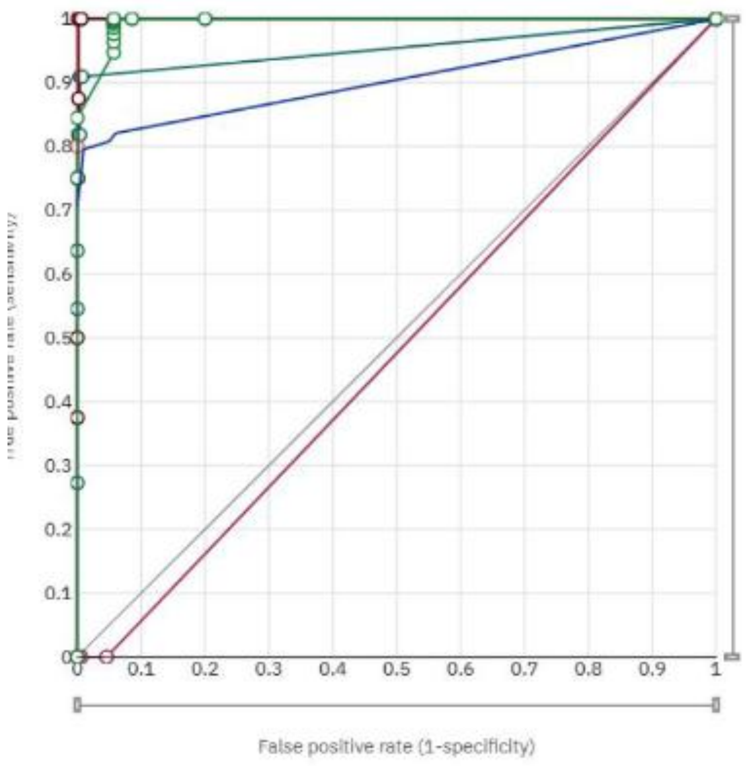
edunet foundation

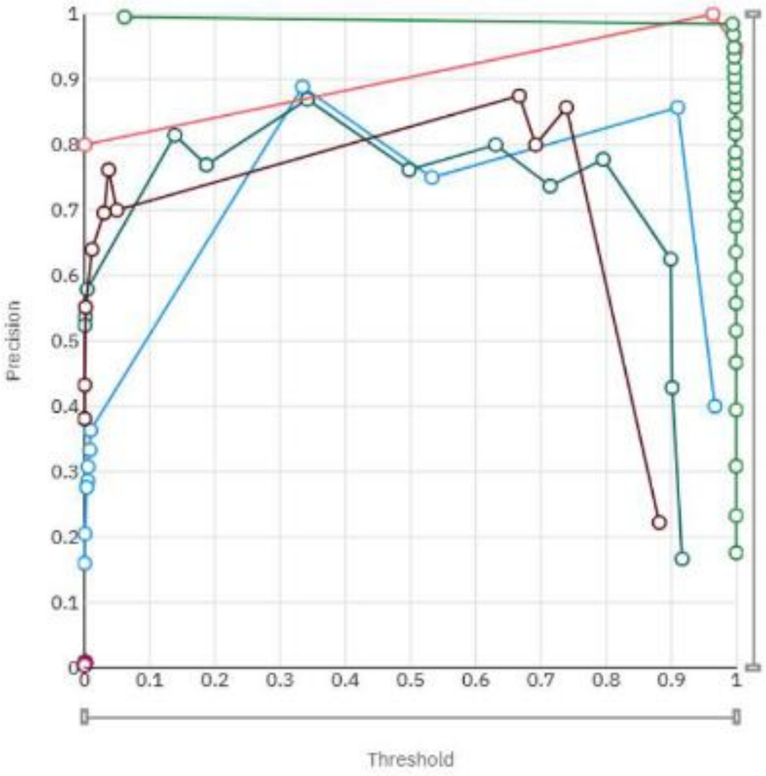
RESULT

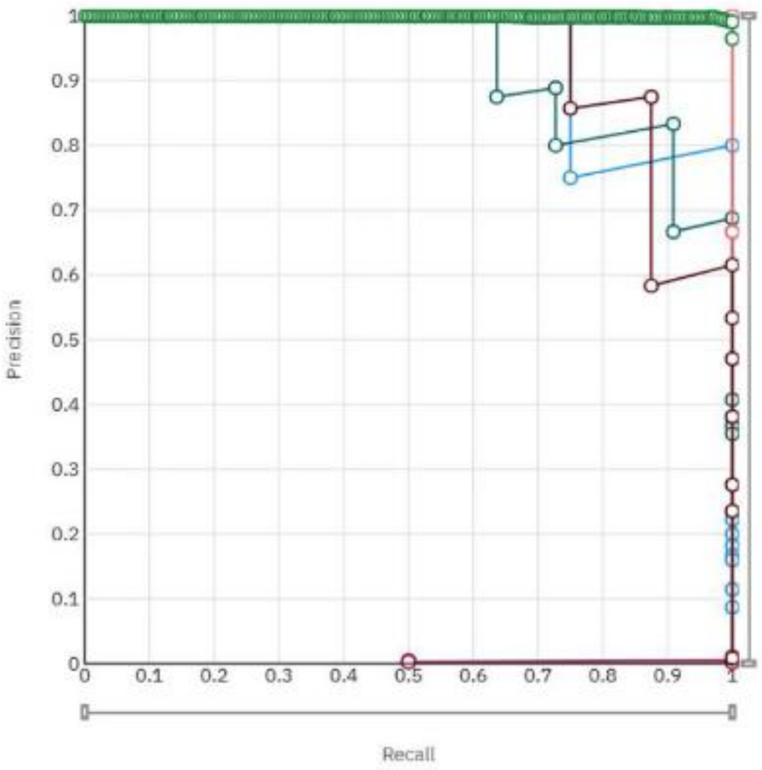
Pipeline leaderboard

	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


False positive rate (1-specificity)


Precision
Threshold


Precision
Recall

RESULT

TESTING:

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

RESULT

- Final 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

edunet foundation

CONCLUSION

This project successfully demonstrates the development of a predictive maintenance system using IBM's AutoAI on watsonx.ai. The use of an automated platform like AutoAI proves to be highly effective in tackling complex industrial challenges, drastically reducing the time and expertise required to build and deploy a high-performance machine learning model. The resulting solution provides a clear pathway for businesses to minimize operational downtime, reduce maintenance costs, and improve overall industrial efficiency.

FUTURE SCOPE

- The current system provides a strong foundation that can be expanded upon in several ways:
- **Real-time Data Streaming:** Integrate the deployed model with live data streams from IoT sensors on actual factory floor machinery for immediate, real-time failure prediction.
- **Remaining Useful Life (RUL) Prediction:** Develop a regression model to predict the remaining time before a component is expected to fail, allowing for even more precise maintenance scheduling.
- **Root Cause Analysis Integration:** Enhance the system to not only predict a failure but also to suggest the most likely contributing factors based on sensor data patterns.
- **Edge Deployment:** Deploy the trained model onto edge computing devices located near the machinery. This would reduce latency and allow for predictions even without a constant internet connection.

REFERENCES

- REFERENCES
- B, S. (2020). *Predictive Maintenance Dataset*. Kaggle. Retrieved from https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification
- IBM. (n.d.). *AutoAI*. IBM. Retrieved from https://www.ibm.com/watson/ai-and-data/autoai
- IBM. (n.d.). *Watson Machine Learning*. IBM. Retrieved from https://www.ibm.com/cloud/watson-machine-learning
- GeeksforGeeks. (2024). *Predictive Maintenance using Machine Learning*.
- Towards Data Science. (2020). *A practical guide for Predictive Maintenance*.

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

Rahul Paul

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 21, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/e5f1f705-1577-4e4f-a0b8-eaca021c9bb4

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Rahul Paul

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 21, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/76755d2a-dc5b-44b2-86e5-699c4727a750

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rahul Paul

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