# CAPSTONE PROJECT

# PREDICTIVE MAINTENANCE OF INDUSTRIAL MACHINERY

Presented By:
1. Pragya Srivastava-Delhi Technical Campus-Computer Science and Engineering(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- Provided in the PDF

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.

# PROBLEM STATEMENT

To improve efficiency and reduce unexpected breakdowns, this project focuses on developing a machine learning model that can predict failures in industrial machines before they happen. By analyzing real-time sensor data, the system will identify early signs of issues such as tool wear, overheating, or power failure.

**Key Points:**

- Leverages real-time **sensor data** (temperature, vibration, voltage, etc.)
- Uses **machine learning** to classify types of failures
- Aims to enable **proactive maintenance**
- Reduces **downtime**, **maintenance costs**, and **safety risks**
- Aligns with **Industry 4.0** and **smart manufacturing goals**

# PROPOSED SOLUTION

- **The proposed system aims to prevent unexpected failures in industrial machines by predicting them in advance using real-time sensor data. By applying machine learning techniques, the system will classify different types of failures to enable proactive maintenance and reduce downtime.**
- **Data Collection:**
  - **Use the Kaggle dataset: Predictive Maintenance Classification**
  - **Includes sensor readings like temperature, pressure, rotational speed, torque, and machine status.**
- **Data Preprocessing:**
  - **Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.**
  - **Normalize or scale sensor values for better model performance.**
  - **Apply feature engineering to extract patterns relevant to failure prediction.**
- **Machine Learning Algorithm:**
  - **Implement a machine learning algorithm**, such as a classification model (e.g., Random Forest, XGBoost, or Neural Network), to **predict failure types** (e.g., tool wear, heat dissipation issues, power failure) based on sensor data patterns.**.**
  - **Consider incorporating other factors** such as machine age, operating hours, and workload history to **improve failure prediction accuracy**.

# PROPOSED SOLUTION

- **Deployment:**
  - **Develop a user-friendly interface or application** that provides real-time predictions of machine failure types based on live sensor inputs.
  - **Deploy the solution on a scalable and reliable platform**, such as IBM Cloud Lite, considering factors like response time, system uptime, and accessibility for maintenance teams.
- **Evaluation:**
  - **Assess the model's performance** using appropriate classification metrics such as Accuracy, Precision, Recall, F1-Score, and Confusion Matrix.
  - **Fine-tune the model** based on real-time feedback and continuously monitor prediction performance using updated sensor data.
  - **Result:** A reliable and accurate system that predicts machine failure types in advance, enabling timely maintenance and reducing unplanned downtime.

# SYSTEM APPROACH

**The "System Approach" section outlines the overall strategy and methodology for developing and deploying a machine learning-based predictive maintenance system for industrial machinery. Here's a suggested structure for this section:**

- **System requirements**

  **Hardware Requirements:**

  **Processor:** Intel Core i3 or higher

  **RAM:** Minimum 4 GB (8 GB recommended)

  **Storage:** At least 500 MB of free disk space

  **Operating System:** Windows, macOS, or Linux

  **Software Requirements:**

  1. **IBM Cloud (mandatory)**
  2. **IBM Watson Studio** for data preprocessing, model development, and deployment
  3. **IBM Cloud Object Storage** for storing and managing sensor datasets
  4. **Python 3.8 or above**
  5. **Web browser** (Chrome/Firefox)

# SYSTEM APPROACH

- **Library required to build the model**

  **Pre-installed in IBM Watson Studio Environment:**

  - pandas – for data manipulation
  - numpy – for numerical operations
  - scikit-learn – for model building and evaluation
  - matplotlib & seaborn – for visualizations
  - xgboost – for gradient boosting classifiers
  - tensorflow & keras – for deep learning (optional)

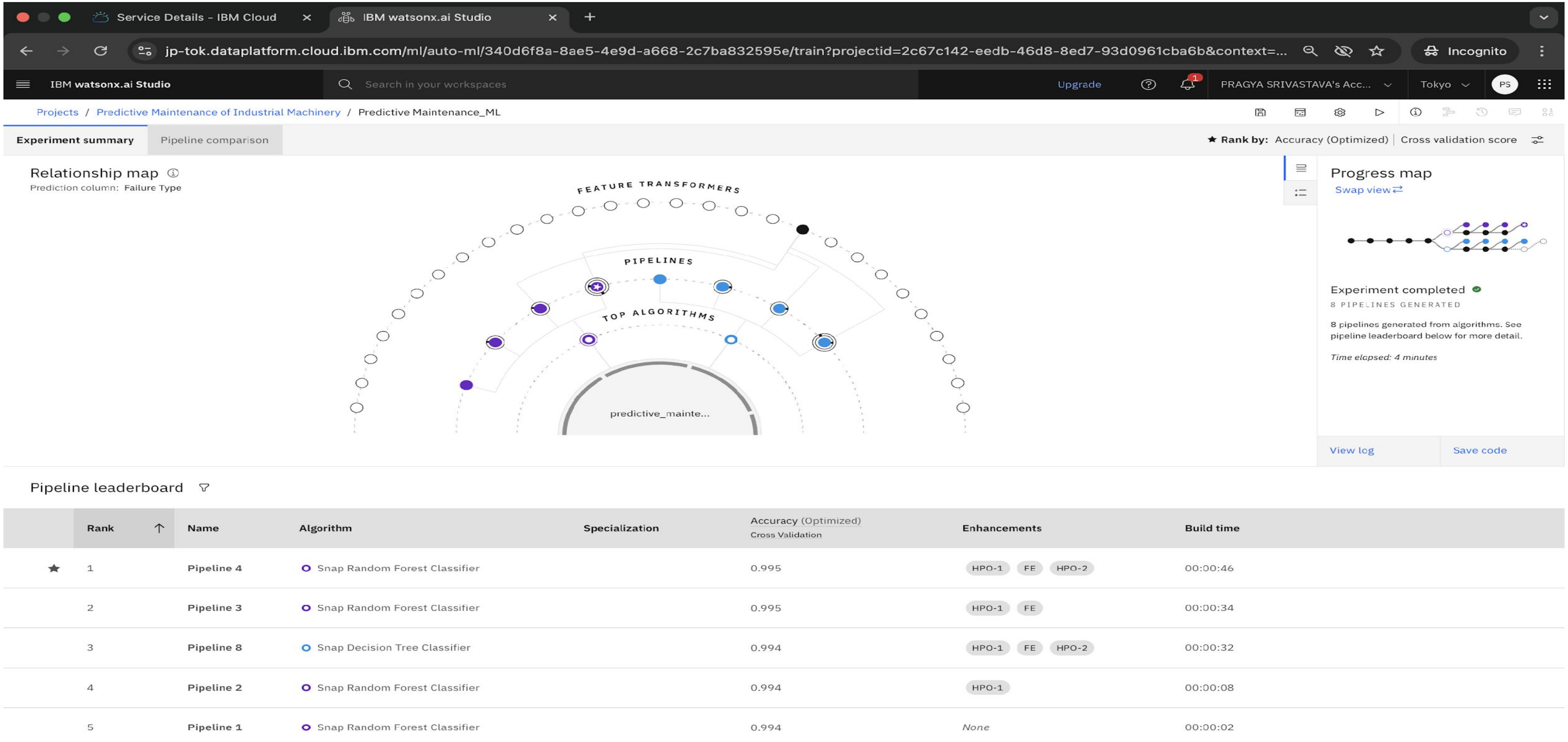
  **For IBM Cloud Services Integration:**

  - ibm-watson-machine-learning – to deploy models and create scoring endpoints
  - ibm-cos-sdk – to access IBM Cloud Object Storage buckets

# ALGORITHM & DEPLOYMENT

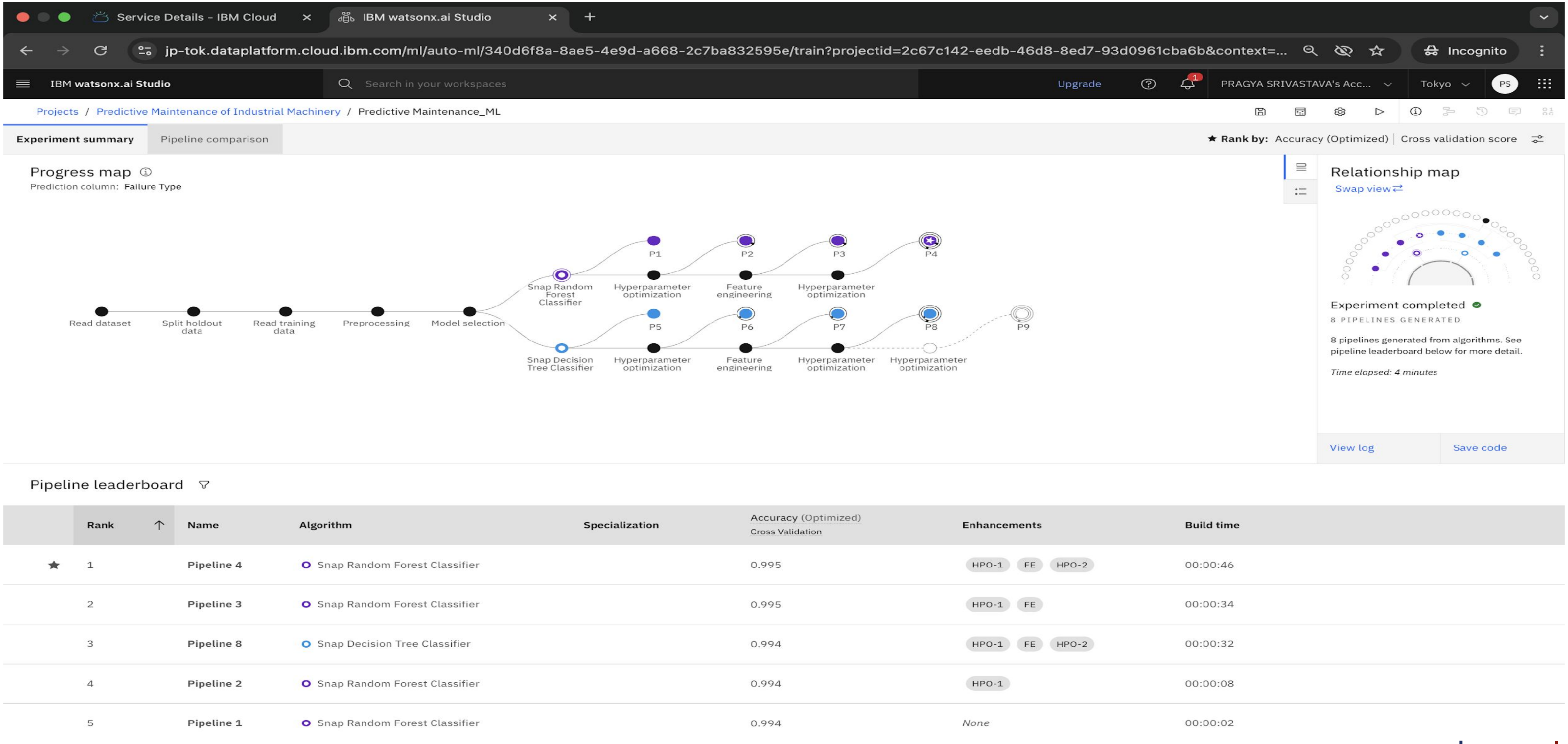
- In the Algorithm section, we outline the machine learning algorithm chosen for predicting machinery failures :
- **Algorithm Selection:**
  - Snap Random Forest Classifier (or SVM based on comparative model performance)
- **Data Input:**
  - Sensor readings such as vibration, temperature, torque, pressure, and rotational speed from industrial machines
- **Training Process:**
  - Supervised learning using labeled maintenance data (e.g., tool wear, power failure, overheating)
- **Prediction Process:**
  - Model deployed on IBM Watson Studio with an API endpoint for real-time failure classification and proactive maintenance alerts

# RESULT



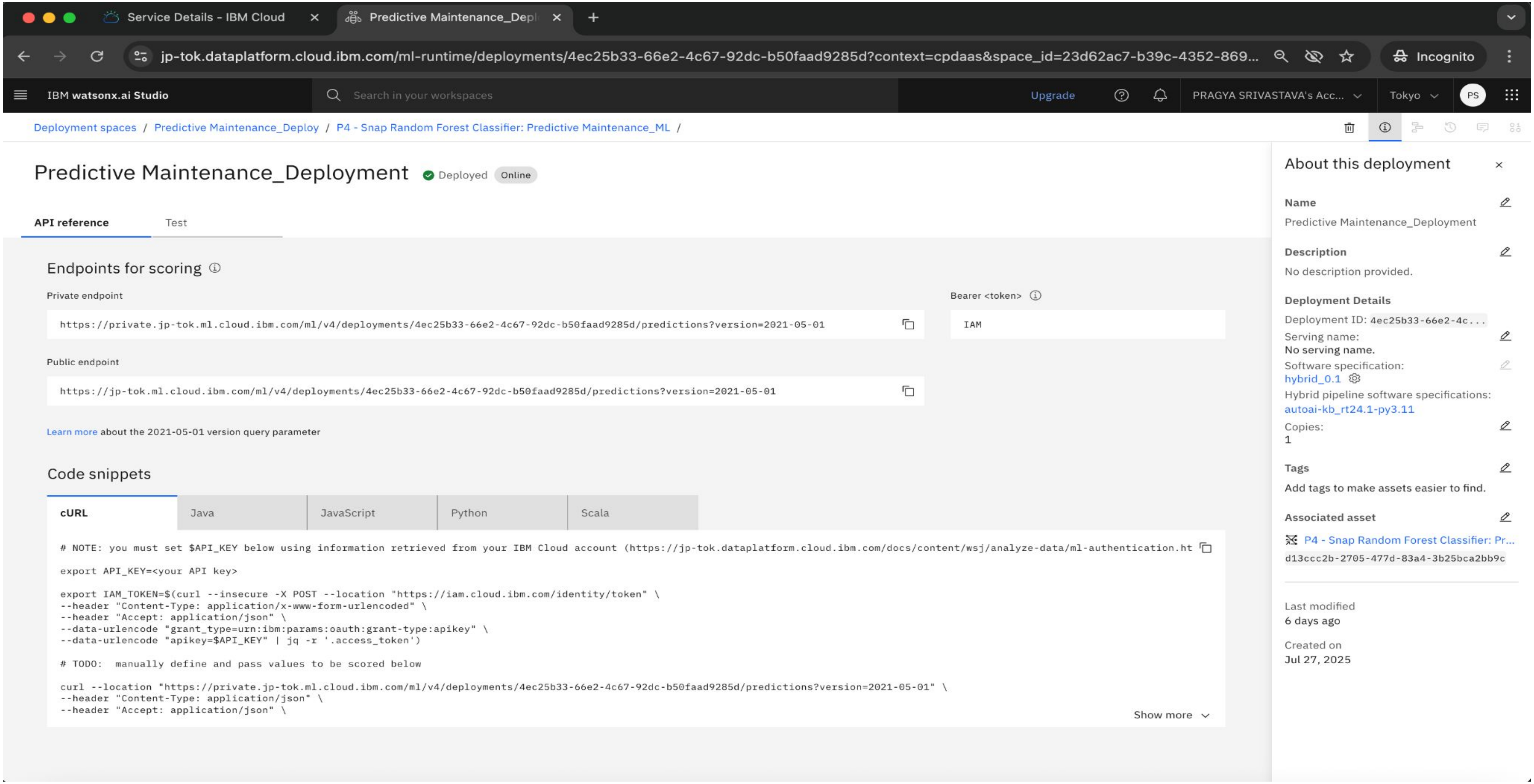

**Relationship Map with pipeline leaderboard**

# RESULT



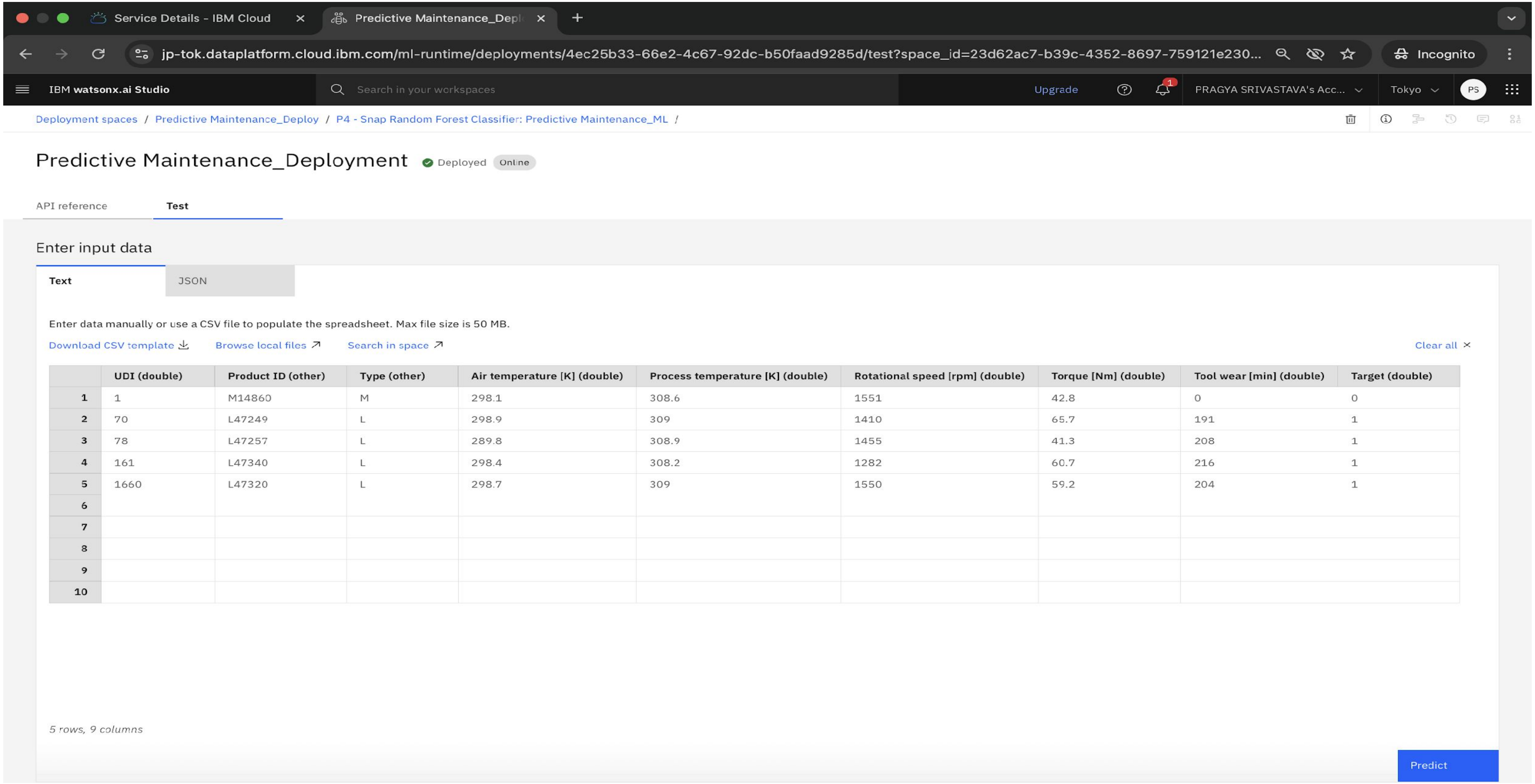

| | Rank | Name | Algorithm | Specialization | Accuracy (Optimized) Cross Validation | Enhancements | Build time |
|---|---|---|---|---|---|---|---|
| ★ | 1 | Pipeline 4 | Snap Random Forest Classifier | | 0.995 | HPO-1 FE HPO-2 | 00:00:46 |
| | 2 | Pipeline 3 | Snap Random Forest Classifier | | 0.995 | HPO-1 FE | 00:00:34 |
| | 3 | Pipeline 8 | Snap Decision Tree Classifier | | 0.994 | HPO-1 FE HPO-2 | 00:00:32 |
| | 4 | Pipeline 2 | Snap Random Forest Classifier | | 0.994 | HPO-1 | 00:00:08 |
| | 5 | Pipeline 1 | Snap Random Forest Classifier | | 0.994 | None | 00:00:02 |

**Pipeline leaderboard with Progress Map**

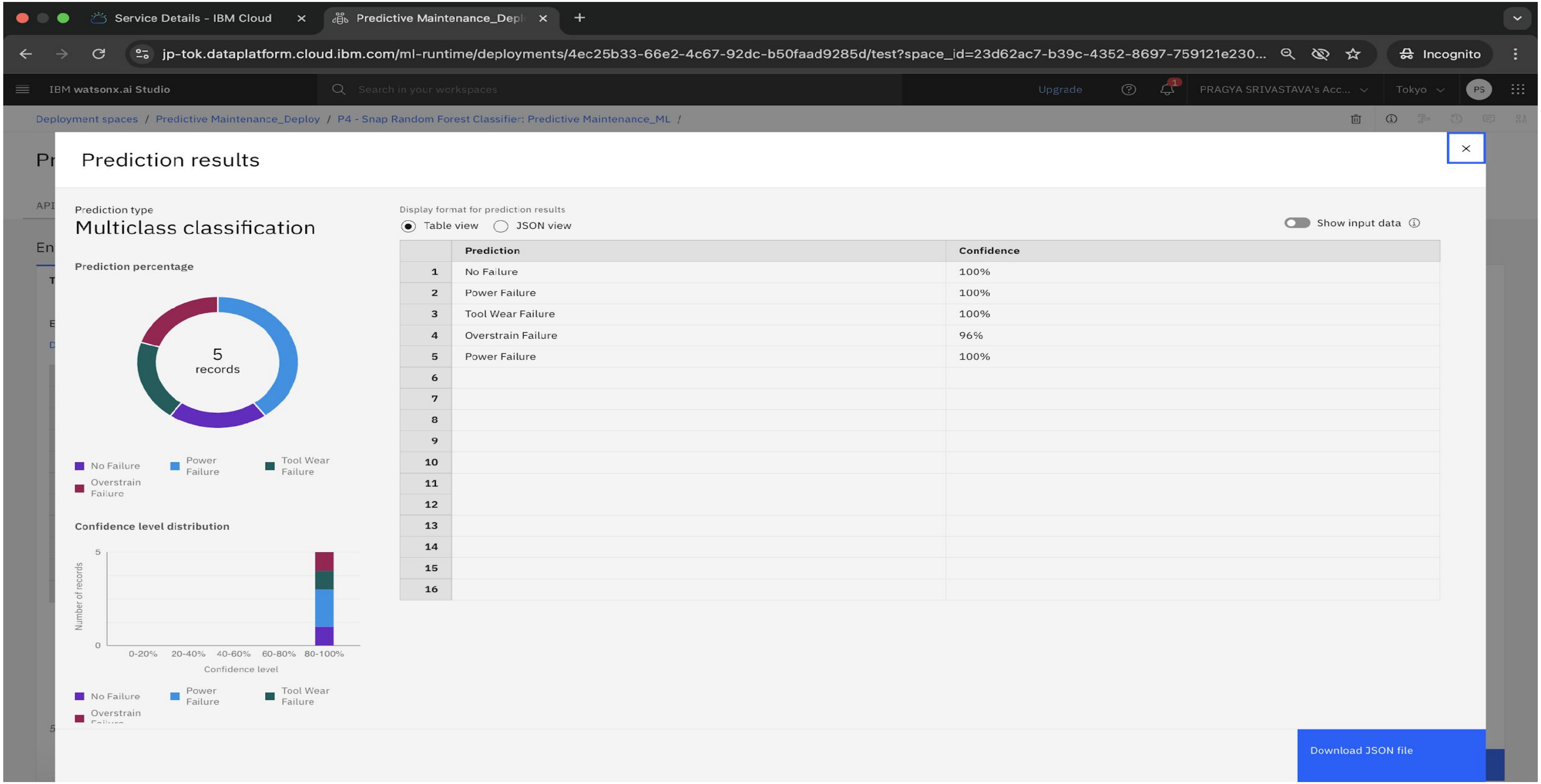



**Completed deployment of the best pipeline and now the project is online**

edunet foundation




# Deployment of best pipeline ML model and input values for testing the ML model




# Results of the Deployed ML model

edunet foundation

# RESULT EXPLANATION

Model Building & Deployment Workflow (4 Key Steps)

**1.Relationship Map**

- Visualizes how different ML pipelines are constructed
- Shows interaction of algorithms (e.g., Random Forest) with feature engineering and tuning

**2.Pipeline Leaderboard**

- Ranks pipelines based on accuracy
- Best model: Random Forest with 99.5% accuracy (Pipeline P4)
- Includes HPO (Hyperparameter Optimization) and FE (Feature Engineering)

**3.Pipeline Progress Map**

- Step-by-step AutoAI model building process:
  Data Ingestion → Preprocessing → Model Selection → HPO → Final Pipelines

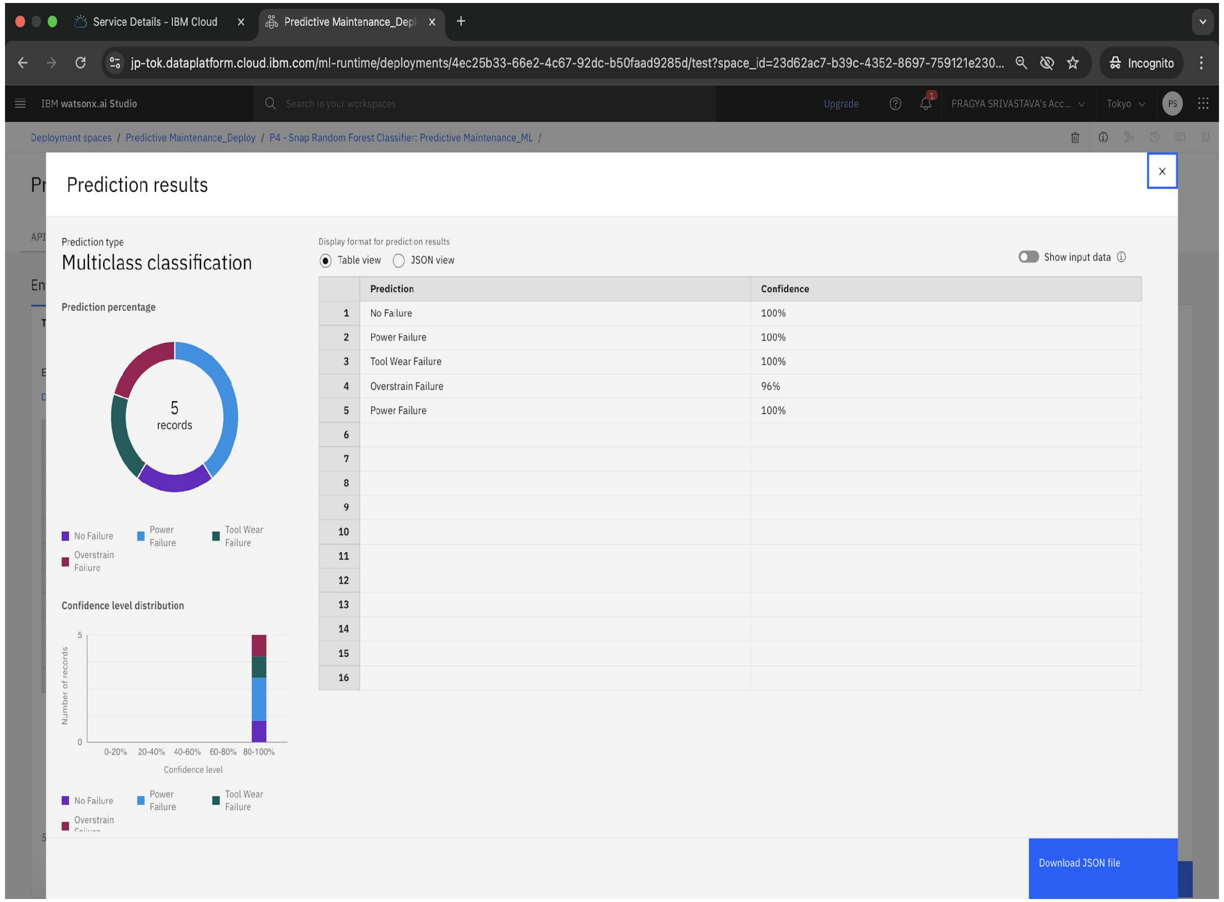
**4.Deployment & Testing**

- Deployed model tested with new machine data
- Predicts failure types in real time based on sensor inputs

# Explanation of Prediction Results – Multiclass Classification (IBM Watsonx.ai)

**Overview:**

- **Prediction Goal**: Identify type of machinery failure using trained ML model
- **Input**: 5 test records
- **Model Type**: Multiclass classification
- **Output**: Failure type with confidence score




**Pie Chart - Prediction Distribution:**

- Visual breakdown of how many times each class was predicted
- Shows 1 record each for No Failure, Tool Wear Failure, Overstrain Failure
- 2 records predicted as Power Failure

**Bar Chart - Confidence Level Distribution:**

- All predictions are **high-confidence** (96%–100%)
- Indicates that the model is very certain about its classifications

**Interpretation:**

- The deployed model is correctly identifying **multiple types of failures**, showing its **multiclass capabilities**.
- High confidence across all predictions implies the model is **well-trained and reliable** for industrial use.
- **Useful for maintenance teams** to proactively respond to specific failure types before they occur.

# CONCLUSION

- The predictive maintenance model developed in this project successfully demonstrated the potential of machine learning in identifying and classifying industrial machinery failures before they occur.
- By leveraging sensor data such as temperature, vibration, and torque, the Snap Random Forest classifier was trained to predict failure types including tool wear, power failure, and overheating.
- Deployment on IBM Cloud Lite enabled real-time monitoring through API endpoints, making the system both scalable and accessible. Challenges such as noisy data and model tuning were addressed through preprocessing and performance evaluation.
- Overall, the project underscores the value of data-driven maintenance strategies in minimizing machine downtime, reducing costs, and improving operational efficiency in industrial settings.

# FUTURE SCOPE

- The predictive maintenance system can be enhanced by integrating additional data sources such as acoustic emissions, oil quality, and ambient environmental conditions to improve failure prediction accuracy.
- Algorithmic improvements, including hyperparameter tuning and ensemble learning techniques, can further optimize model performance and reduce false positives.
- The system can be scaled to monitor machinery across multiple plants, cities, or even countries, enabling centralized and consistent maintenance strategies across large industrial networks.
- Incorporating **edge computing** would allow real-time processing of sensor data directly on-site, reducing latency and dependency on cloud infrastructure.
- Future upgrades could also explore **advanced machine learning models** such as deep neural networks or reinforcement learning to handle complex failure patterns and adapt to evolving machine behavior.

# REFERENCES


- **Carvalho, T. P., Soares, F. A. A. M., Vita, R., et al.** (2019). *A systematic literature review of machine learning methods applied to predictive maintenance*. Computers & Industrial Engineering, 137, 106024.
 https://doi.org/10.1016/j.cie.2019.106024
- **Zhang, W., Yang, D., Wang, H.** (2019). *Data-driven methods for predictive maintenance of industrial equipment: A survey*. IEEE Systems Journal, 13(3), 2213–2227.
 https://doi.org/10.1109/JSYST.2018.2813800
- **Kaggle Dataset - Predictive Maintenance Classification**
 Shivam Bansal (2018). *Machine Predictive Maintenance Classification Dataset*.
 https://www.kaggle.com/datasets/shivamb/machine-predictive-maintenance-classification
- **Breiman, L.** (2001). *Random Forests*. Machine Learning, 45(1), 5–32.
 https://doi.org/10.1023/A:1010933404324
- **IBM Watson Studio Documentation**
 IBM Cloud Docs. *Deploying Machine Learning Models Using IBM Watson Studio*.
 https://www.ibm.com/cloud/watson-studio
- **Scikit-learn Documentation**
 Pedregosa, F., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, 2825–2830.
 https://scikit-learn.org/stable/documentation.html
- **García, S., Luengo, J., Herrera, F.** (2015). *Data Preprocessing in Data Mining*. Springer.
 https://doi.org/10.1007/978-3-319-10247-4

# IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild
IBM

# Pragya Srivastava

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/e8bcc579-c13d-4799-85ba-dd623684c7c5

IBM

# IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence


## Pragya Srivastava

Has successfully satisfied the requirements for:

---

### Journey to Cloud: Envisioning Your Solution

---

Issued on: Jul 16, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/0f51f731-0513-48cc-8b34-cf2f8f815566

IBM

# IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Pragya Srivastava

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

# IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Pragya Srivastava

for the completion of

**Introduction to Retrieval Augmented Generation**

(ALM-COURSE_3825461)

According to the Adobe Learning Manager system of record

**Completion date:** 30 Jul 2025 (GMT)

**Learning hours:** 1 hr 30 mins

# IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Pragya Srivastava

for the completion of

**Ethical Considerations for Generative AI**

(MDL-571)

According to the Moodle system of record

**Completion date:** 31 Jul 2025 (GMT)

**Learning hours:** 1 hr 30 mins

# THANK YOU