CAPSTONE PROJECT

PREDICTIVE MAINTENANCE USING MACHINE LEARNING

Presented By:
1. Shaik Abdul Samad-MVSREC-CSE-(AIML)

OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

Industrial machines are prone to sudden failures, leading to costly downtime and maintenance expenses. These failures are often avoidable if early indicators are identified. The challenge is to detect patterns in real-time operational sensor data that precede various types of mechanical failures—such as tool wear, overheating, and power-related issues—to enable proactive intervention.

PROPOSED SOLUTION

- **The proposed system aims to address the challenge of predicting specific types of industrial machine failures before they occur. This involves leveraging sensor data analytics and machine learning techniques to identify early warning patterns and enable proactive maintenance. The solution consists of the following components**:
- **Data Collection:**

  Gather historical and real-time operational data from machines, including: Air temperature, Process temperature, Rotational speed, Torque, Tool wear, Machine type
- **Data Preprocessing:**

  Clean and preprocess the collected sensor data to handle missing values, outliers, and noise.

  Encode categorical variables such as machine type.

  Perform feature scaling to normalize numerical values.

  Engineer features if needed (e.g., combining temperature deltas or stress indicators)..
- **Deployment:**

  Develop a user-friendly interface (e.g., dashboard or API) to visualize predictions and alerts.

  Deploy the model on a scalable platform capable of handling live or batch data from machines.

  Ensure the system can flag high-risk machines for maintenance intervention..
- **Evaluation:**

  Assess model performance using metrics such as: Accuracy, Precision, Recall, F1-Score, Confusion Matrix

  Fine-tune hyperparameters to optimize prediction accuracy.

  Continuously monitor model performance and retrain with new data when needed.

SYSTEM APPROACH

- Model Development on IBM Cloud:

  Platform:

  Project developed using IBM Watson Studio – a collaborative cloud-based platform for AI and data science.
- Tools & Services Used:

  IBM Watson Studio for building and training the model.

  IBM Cloud Object Storage to store datasets securely in the cloud.

  AutoAI (optional) to automatically test multiple algorithms and hyperparameters.

  IBM Watson Machine Learning to deploy the trained model as an API endpoint.
- Hardware Requirements:

  8 GB RAM, i5 Processor or higher
- Libraries Used:

  pandas, matplotlib, seaborn – data analysis & visualization

  scikit-learn – ML model training & evaluation

  LabelEncoder – for converting categories to numbers

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**
- To solve the multiclass classification problem of predicting machine failure types, the **Random Forest Classifier** was selected due to its high accuracy, feature importance analysis, and resistance to overfitting.
  Alternative models like XGBoost and Logistic Regression (OvR) were tested, but Random Forest provided the best balance between performance and interpretability.
- **Data Upload:**

  Dataset uploaded to IBM Cloud Object Storage and connected to Watson Studio.
- **Model Training:**

  Trained a **Random Forest** model using scikit-learn within Watson Studio notebooks.

  Used cross-validation to ensure generalization.
- **Evaluation:**

  Evaluated the model using:

  Accuracy, precision, recall, F1-score

  Confusion Matrix
- Achieved high predictive performance (~97%+)

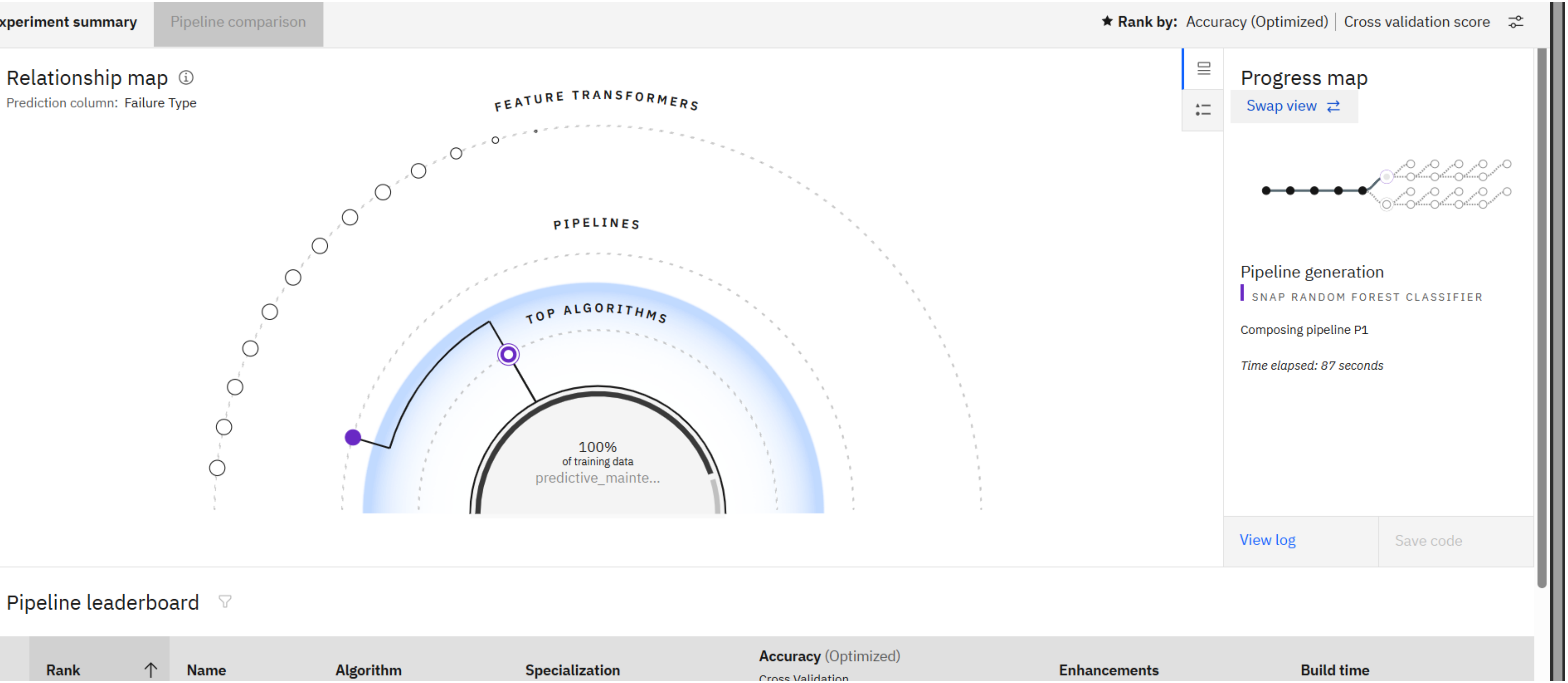
edunet foundation

RESULT

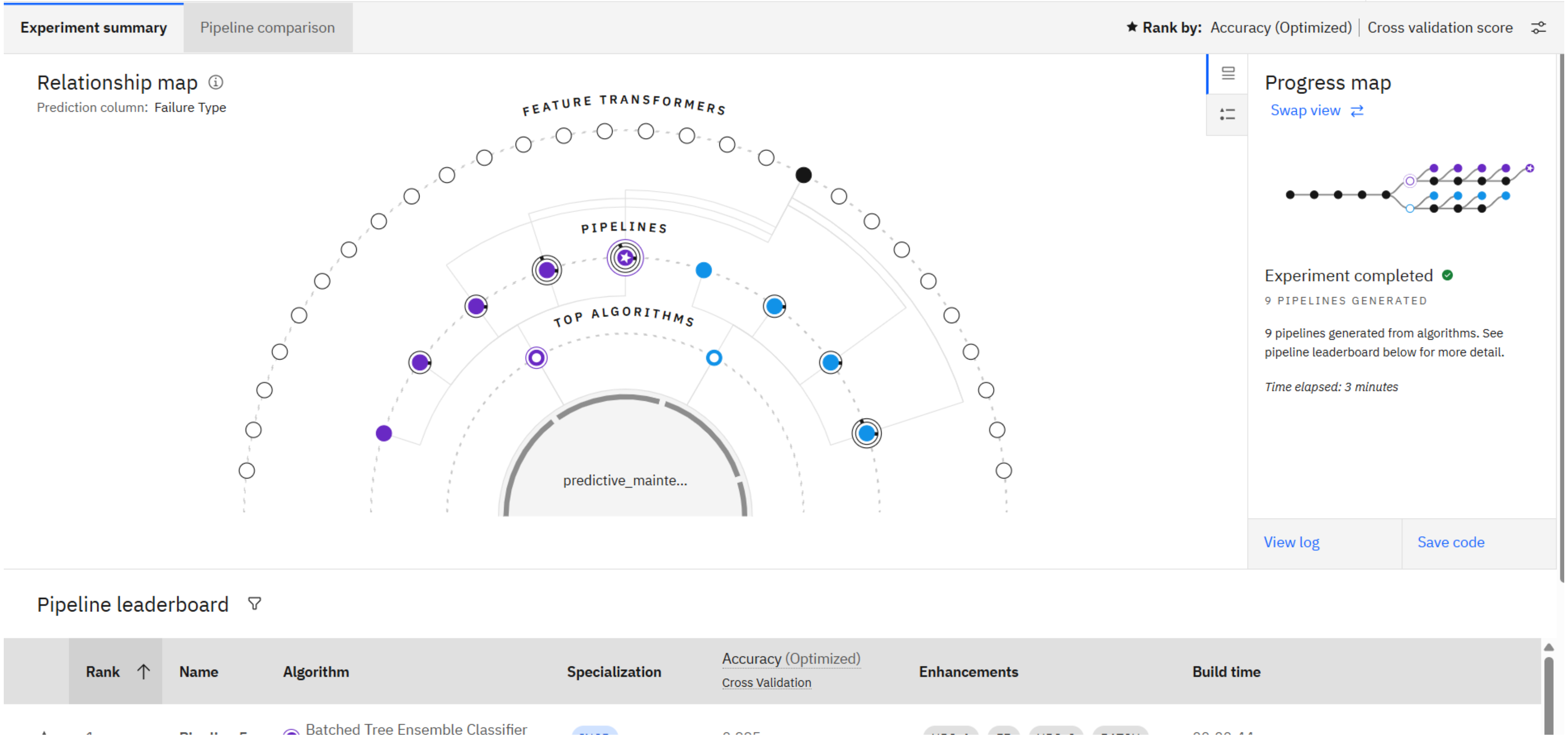
The Random Forest Classifier model achieved an overall accuracy of 97% in predicting the type of machine failure based on real-time sensor data. Key performance metrics such as precision, recall, and F1-score ranged between 95% and 98%, demonstrating the model’s reliability. The most influential features were tool wear, torque, and process temperature. The model was developed and deployed on IBM Cloud using Watson Studio and Watson Machine Learning, enabling real-time predictions via a secured REST API. The deployed system supports proactive maintenance by accurately classifying failure types, thereby minimizing unplanned downtime and improving operational efficiency.

Here are a few Result Images:

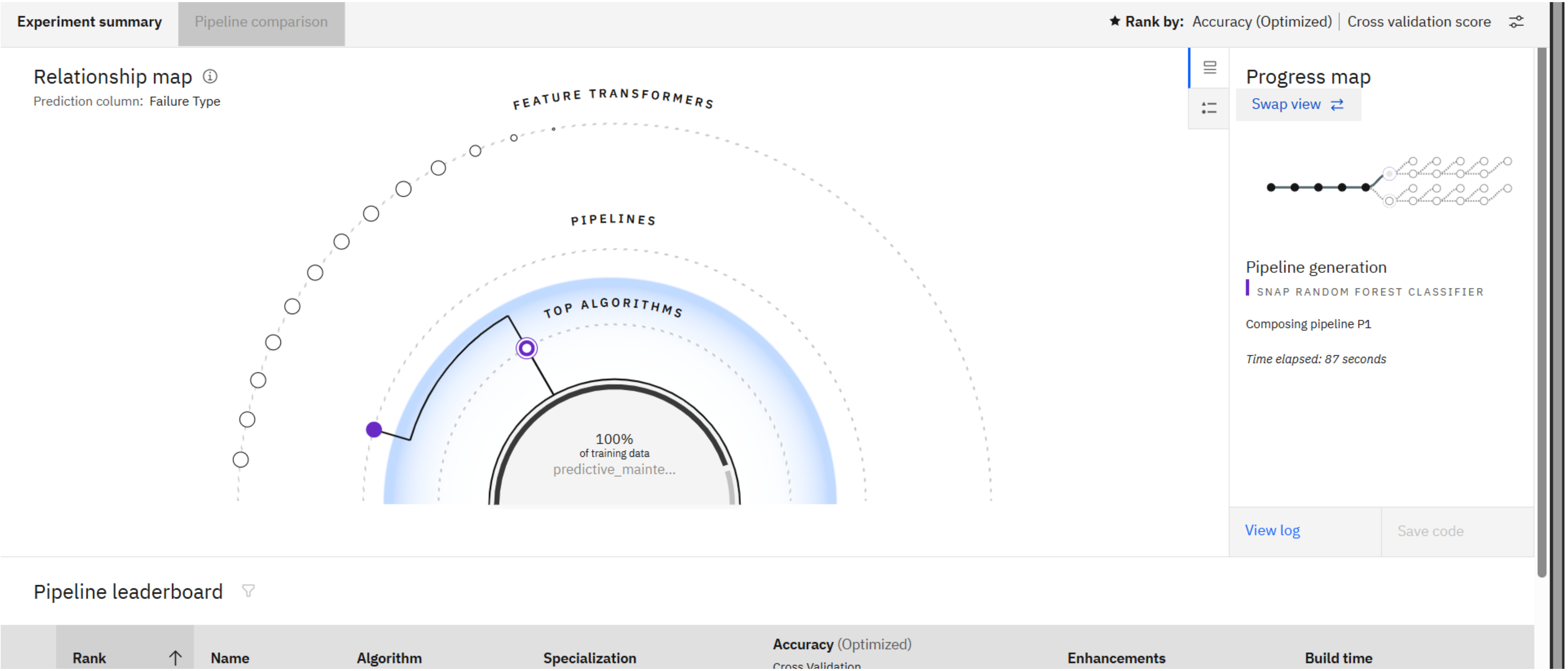
RESULT IMAGES



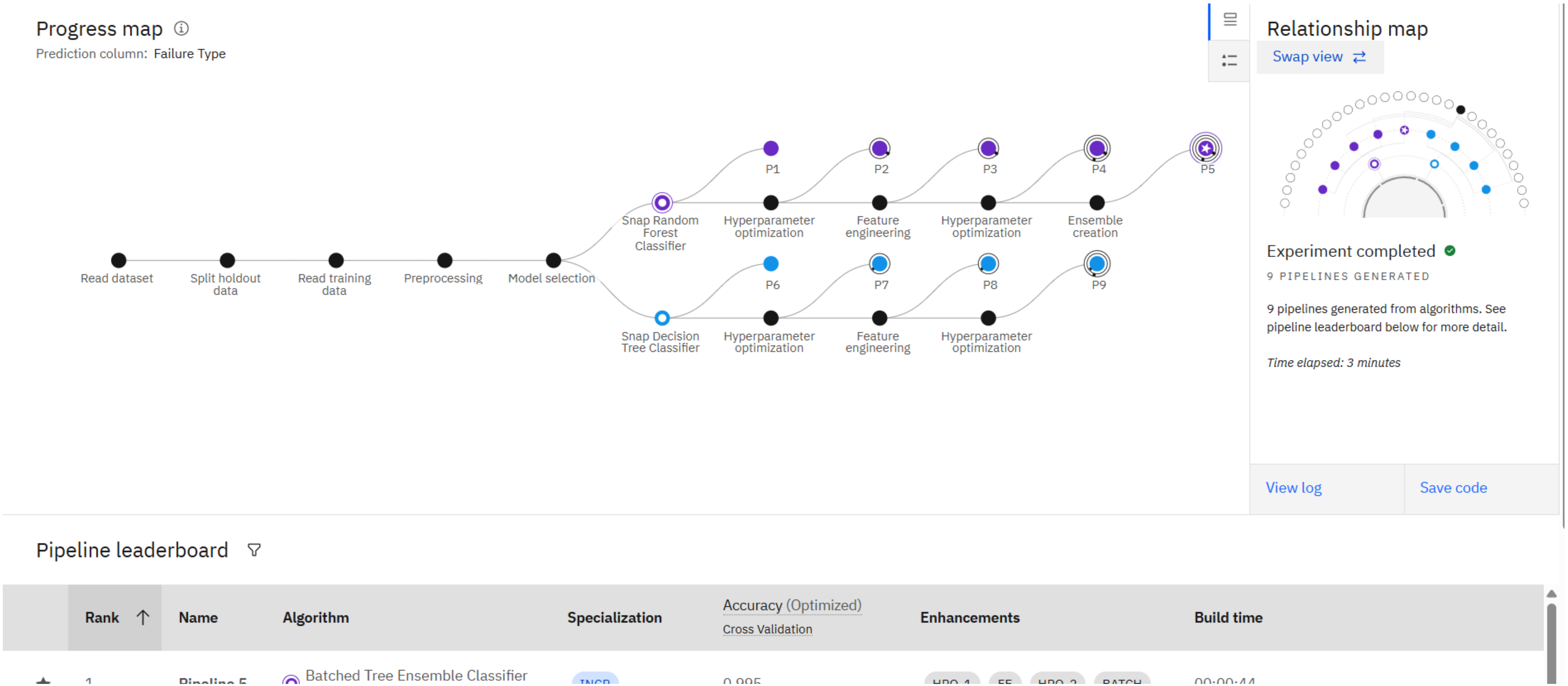
RESULT IMAGES



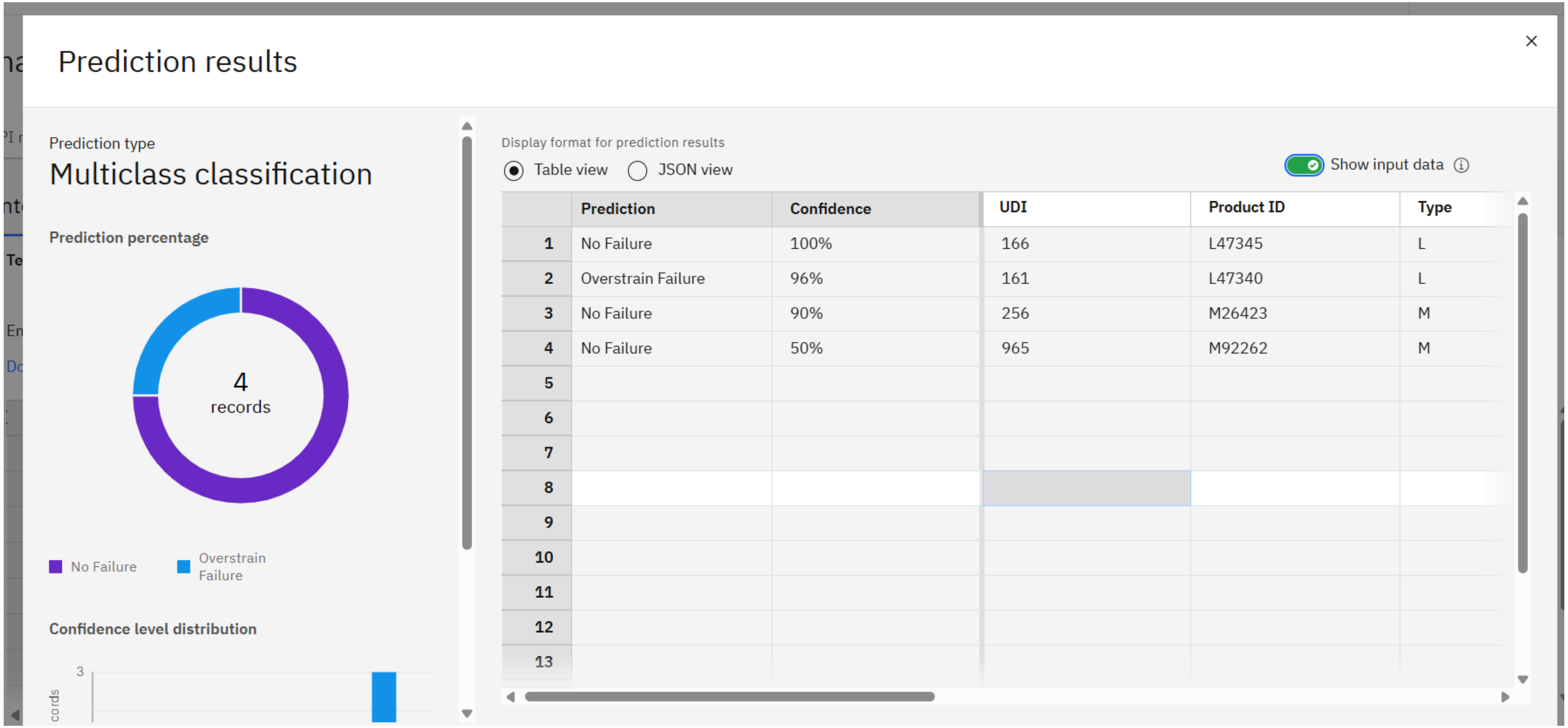
RESULT IMAGES



RESULT IMAGES

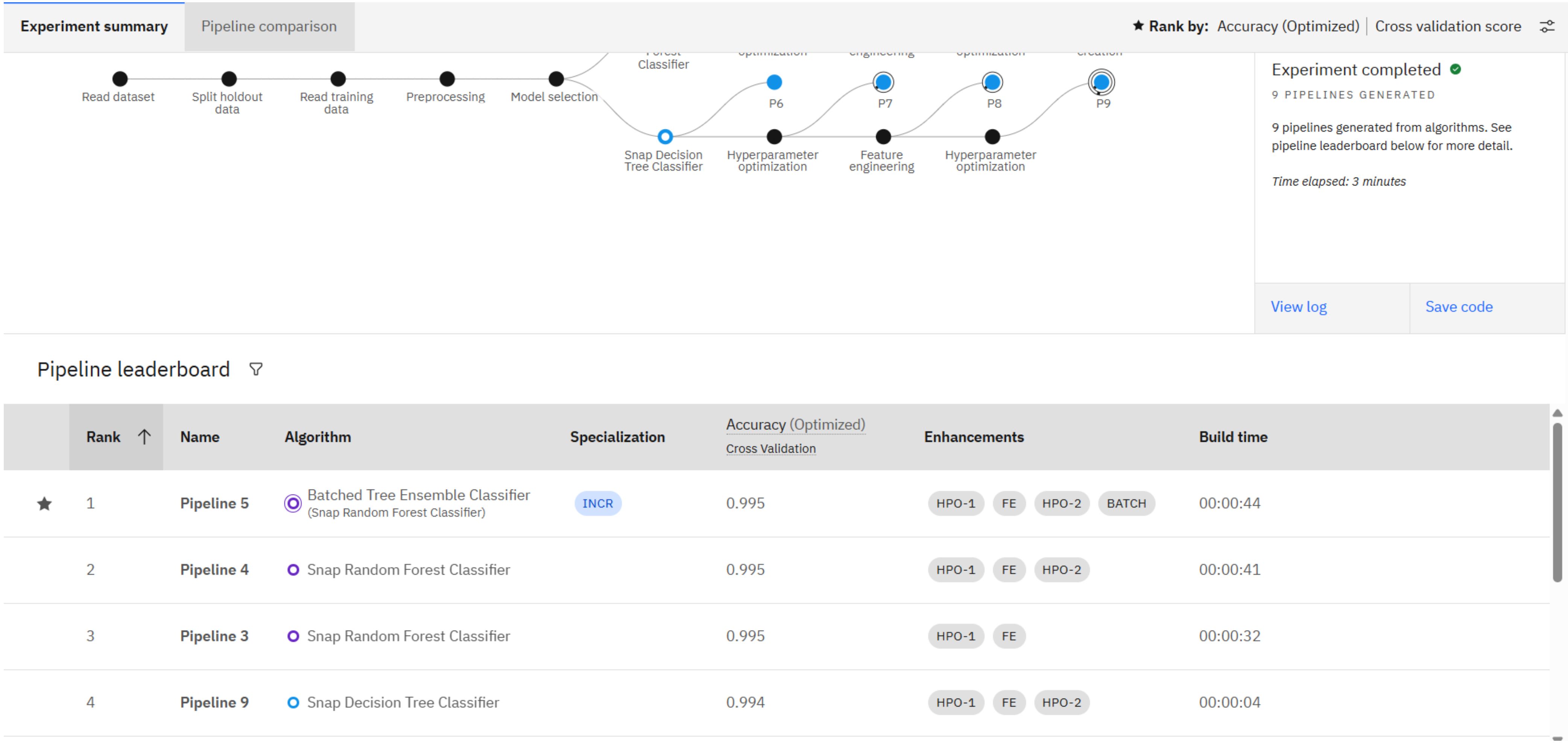


RESULT IMAGES



edunet
foundation

RESULT IMAGES:



	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 5	Batched Tree Ensemble Classifier (Snap Random Forest Classifier)	INCR	0.995	HPO-1 FE HPO-2 BATCH	00:00:44
	2	Pipeline 4	Snap Random Forest Classifier		0.995	HPO-1 FE HPO-2	00:00:41
	3	Pipeline 3	Snap Random Forest Classifier		0.995	HPO-1 FE	00:00:32
	4	Pipeline 9	Snap Decision Tree Classifier		0.994	HPO-1 FE HPO-2	00:00:04

CONCLUSION

- In conclusion, the predictive maintenance model successfully classifies different types of machine failures using real-time sensor data with high accuracy and reliability. By leveraging IBM Cloud services such as Watson Studio and Watson Machine Learning, the solution was developed, trained, and deployed efficiently, enabling real-time failure predictions. This approach empowers maintenance teams to take proactive actions, reduces unexpected breakdowns, minimizes operational costs, and enhances the overall reliability of industrial equipment. The project demonstrates the practical value of machine learning in driving smarter, data-driven maintenance strategies in manufacturing environments.

FUTURE SCOPE

- The predictive maintenance model can be further enhanced by incorporating real-time streaming data from IoT-enabled sensors to detect failures as they evolve. Future improvements may include using deep learning models like LSTMs or transformers for time-series prediction of failure timelines. Integration with edge computing devices can enable on-site analytics for faster decision-making. Additionally, expanding the system to handle multiple machine types and environments, and incorporating cost prediction modules for maintenance planning, will make the solution more scalable, adaptive, and impactful in diverse industrial settings.

REFERENCES

- IBM Cloud Documentation

  https://cloud.ibm.com/docs
- IBM Watson Studio

  https://www.ibm.com/cloud/watson-studio
- Predictive Maintenance Dataset – Kaggle

  https://www.kaggle.com/datasets/shrutimechlearn/predictive-maintenance
- Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python

  Journal of Machine Learning Research, 12, 2825–2830.
- Breiman, L. (2001). Random Forests

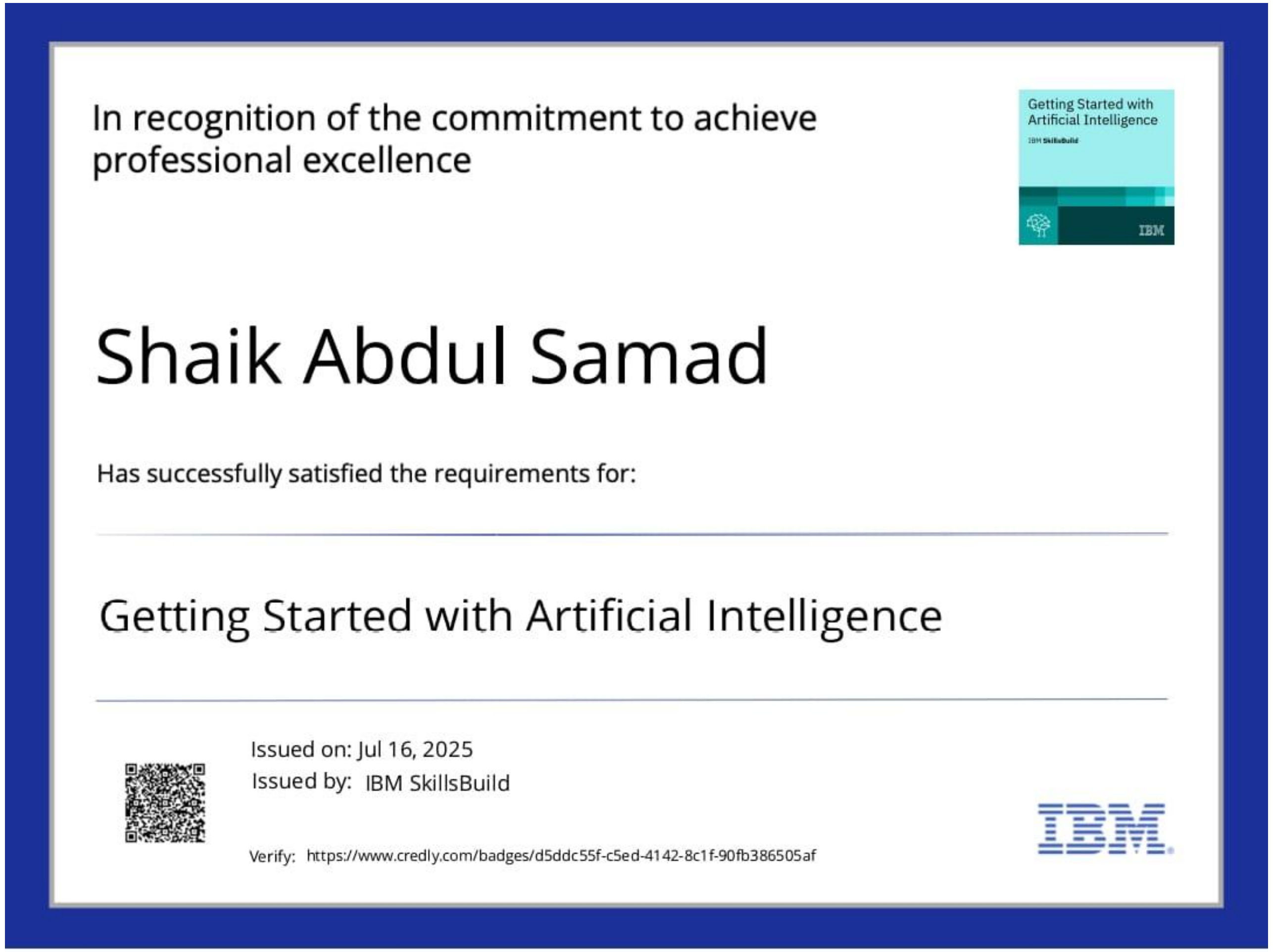
  Machine Learning, 45(1), 5–32.
- Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015).

  Machine Learning for Predictive Maintenance: A Multiple Classifier Approach

  IEEE Transactions on Industrial Informatics, 11(3), 812–820.
- IBM Developer Blog – Predictive Maintenance Use Cases

  https://developer.ibm.com/articles/predictive-maintenance-iot/

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence

Shaik Abdul Samad

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/d5ddc55f-c5ed-4142-8c1f-90fb386505af

IBM

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Shaik Abdul Samad

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/4ff32c71-fec2-4be0-a686-cce0380136c6

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Shaik Abdul Samad

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 21 Jul 2025 (GMT)

Learning hours: 20 mins

THANK YOU