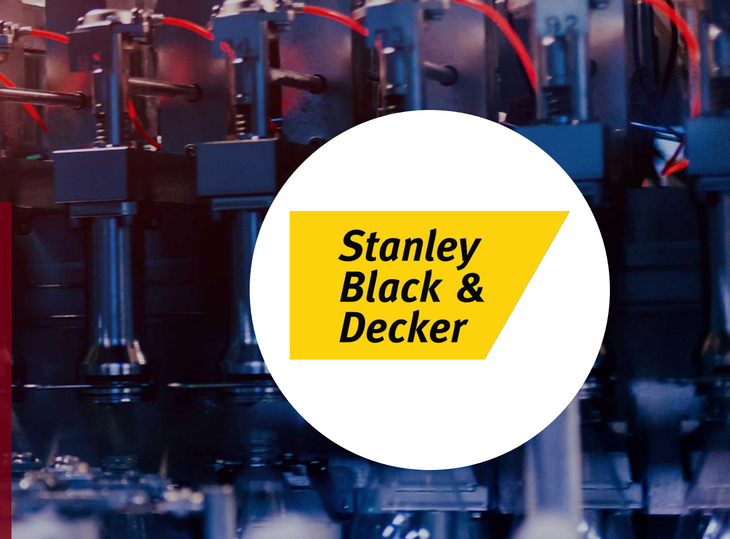


Adaptive Predictive Maintenance via Integrated Data Architecture

BUSINESS PROBLEM

Unplanned machine breakdowns continued to disrupt production across the plant despite recent investments in production monitoring and maintenance software. While large volumes of operational and maintenance data were being collected, the systems operated independently and were rarely used to support proactive decision making. The organization wanted to understand whether these existing enterprise data sources could be integrated and used to anticipate breakdowns earlier and support predictive maintenance planning.

DATA SOURCES

Production Monitoring Productivity - hourly production output information

Production Monitoring Downtime Events - event log of machine interruptions, downtime classifications, and process-level disruptions.

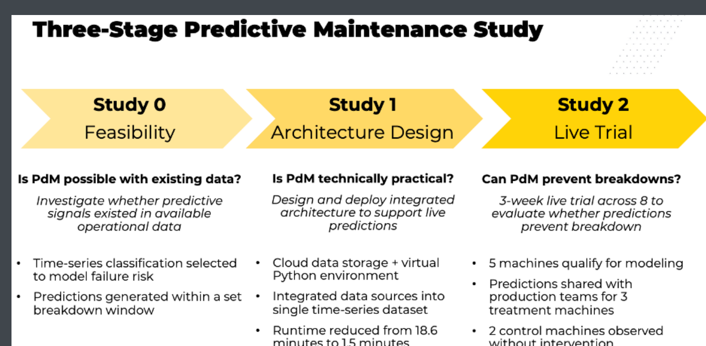
CMMS Maintenance Work Orders - event logs of maintenance repairs on each machine

Data Types and Format

Boolean, Categorical (nominal), Count (integer), Numeric, Ordered (ordinal), Relational Data, Structured, Text, Time series (Dynamic)

APPROACH

This project developed an integrated data architecture linking production monitoring and CMMS records to enable predictive maintenance modeling. Event-based machine histories were constructed from operational logs and used to train classification models to predict breakdowns within a defined future window. Gradient boosted tree models were selected due to their performance on structured industrial datasets. The system was deployed in a live pilot environment where predictions were delivered through prototype visualizations.



Maci Prescott
MIT Leaders for Global Operations, Class of 2026

Leaders for Global Operations

IMPACT

This project demonstrated that predictive maintenance (PdM) can be developed using existing enterprise operational data without deploying additional sensors. During the first week of the live pilot, the model achieved 88% balanced accuracy in predicting machine breakdowns, with offline validation performance above 0.75 AUC across several assets. These results showed that production and maintenance event logs can contain meaningful predictive signals. A major contribution of the work was the design of an integrated data architecture linking production monitoring and CMMS systems that previously operated independently. Improvements to data storage, encoding, and decoding methods reduced overall system runtime by 92%, including a 72% reduction in feature encoding time. These changes significantly improved the feasibility of running predictive models on large enterprise datasets and created a scalable foundation for future PdM applications. The project also contributed to organizational alignment around predictive analytics. Findings from the study were presented to 300+ data and analytics power users across the organization, outlining the architectural and data governance changes required to support PdM and other machine learning initiatives. Finally, the live deployment clarified how PdM systems must operate in practice. While models can signal increased breakdown risk, maintenance teams still require accessible historical diagnostics to interpret and act on those predictions. The study reinforced that effective PdM requires both advanced machine learning and historical failure patterns to support operational decision making.

DRIVERS	Unplanned machine breakdowns continued to disrupt production despite recent investments in production monitoring and CMMS systems. The systems operated independently and rarely supported proactive maintenance decisions. The organization wanted to understand whether these existing data sources could be integrated to anticipate failures earlier.
BARRIERS	Production and maintenance data were stored in separate systems with inconsistent machine identifiers and different data structures. Operational logs were often aggregated or updated in batches, limiting temporal resolution. Data continuity disruptions and sparse failure events also created class imbalance challenges for predictive modeling.
ENABLERS	The organization had already deployed production monitoring and CMMS systems across the plant, creating a large base of operational and maintenance records. Support from the data analytics and operations technology teams enabled access to these systems, collaboration on architecture design, and participation in the live pilot and organizational review sessions.
ACTIONS	Production monitoring and maintenance records were integrated into a unified machine history using a new data architecture linking the two enterprise systems. Event-driven datasets were constructed from operational logs and used to train classification models predicting breakdowns within a future time window. The data was tested for predictive signals, then deployed into a live pilot to evaluate real-world performance.
INNOVATION	Most predictive maintenance work assumes dense sensor data and tightly integrated digital systems. This project instead demonstrated how predictive models can be developed using existing enterprise operational logs. The approach combined event-driven machine histories, enterprise data architecture integration, and machine learning to enable predictive maintenance in a traditional manufacturing environment.
IMPROVEMENT	Architecture design made predictive maintenance analysis practical on enterprise data. Separating model training and prediction reduced runtime by 92%, while encoding optimizations reduced feature processing time by 72%. Supported by this architecture, the first week of the live pilot achieved 88% balanced accuracy in predicting breakdowns. The work also built shared understanding of predictive maintenance across the analytics community.
BEST PRACTICES	Predictive maintenance implementation depends as much on system architecture and operational workflows as on the model itself. Organizations should prioritize reliable data pipelines, integrated machine histories, and accessible diagnostics so maintenance teams can interpret and act on predictive signals.
OTHER APPLICATIONS	The architecture and event-driven modeling approach can support other predictive analytics applications using enterprise operational data. Examples include predicting production interruptions, identifying recurring downtime drivers, forecasting maintenance workload, and supporting reliability analysis across assets and plants.