

Review Article

Open Access

Leveraging Data Analytics in Multimodal Deep Learning for Predictive Maintenance Aimed at Minimizing Rig Downtime

Gaurav Kumar Sinha

Amazon Web Services Inc, USA

ABSTRACT

Enhancing uptime and achieving greater operational efficiency in industrial assets and systems is pivotal. The focus of this paper is on a data-led strategy that integrates a scalable data framework, sophisticated analytics, and multimodal deep learning to foster predictive maintenance. This approach significantly reduces expensive downtime in oil and gas drilling activities. A proposed solution involves a cloud-based data lake architecture capable of capturing and storing both structured and unstructured time-series sensor data from drilling equipment like pumps, blowout preventers, and top drives. Massive data pipelines handle billions of sensor readings, and multivariate analysis offers insights into maintenance needs. The introduction of the Deep Maintenance Neural Network (DMNN) maximizes the use of sensor data, past failures, meteorological information, and expert knowledge to foresee a range of system failures in drilling operations. This study underscores the combined power of leveraging industrial-scale data storage solutions, cutting-edge analytics, and deep learning to pave the way for advanced maintenance strategies. Implementing the discussed approach converts raw sensor data into significant improvements in rig durability and operational downtime. It demonstrates that adopting a data-smart maintenance strategy can substantially elevate productivity at the system level.

*Corresponding author

Gaurav Kumar Sinha, Amazon Web Services Inc, USA.

Received: December 12, 2022; Accepted: December 16, 2022; Published: December 22, 2022

Keywords: Predictive Maintenance, Deep Learning, Data Analytics, Rig Downtime Optimization, Sensor Data, Multivariate Analysis, Industrial Internet Of Things, Prescriptive Analytics, Data Infrastructure, Microservices, Data Pipeline, Machine Learning Model Deployment, Time Series Analysis, Failure Prediction, Maintenance Scheduling

Introduction

Executing operations in the oil and gas sector requires intricate machinery and equipment on rigs, which can experience periodic malfunctions if not consistently upheld. Nevertheless, unforeseen interruptions can lead to exorbitant costs, with losses surpassing \$200,000 daily due to reduced productivity. Therefore, the capacity for predicting potential maintenance issues before their manifestation offers a significant financial advantage.

The advent of the Industrial Internet of Things (IoT) has catalyzed a surge in the generation of sensor data from various rig machinery, including top drives, blowout preventers, and devices detecting natural gas. This sequential data records minor declines in efficiency and preliminary indicators of prospective failures. Nonetheless, a principal obstacle is efficiently converting the raw sensor data into predictive insights that prompt action.

In this paper, I advocate for an approach centered on data, employing cloud-based storage solutions, refined analytics, and diverse deep-learning models to usher in an era of advanced predictive maintenance capable of reducing expensive operational

pauses. Innovations at the heart of machine learning, crucial for early anomaly detection, include:

- The use of Long-Short Term Memory (LSTM) models
- Implementation of Convolutional Neural Networks (CNN) for the identification of error messages
- Integration of Dynamic Bayesian Networks (DBN) for the inclusion of specialized knowledge and weather-related information

Problem Statement

Offshore operations for extracting oil and gas are managed through intricate machinery and equipment on rigs, consistently functioning amid tough environmental settings. The unexpected breaking down of equipment often necessitates an immediate halt for repair work, which results in long periods of inactivity and significantly high operational costs. The day-to-day expenses of running sophisticated rigs can exceed \$200,000, with unplanned outages quickly leading to financial losses amounting to millions.

Despite this, the core systems and hardware tend to gradually lose their efficiency due to constant wear and tear, a problem that has historically gone unnoticed until a serious malfunction occurs. Early indicators like minor seal leaks in pumps, rusting pipelines, or slight fissures in the rig's structure typically remain undetected. These maintenance problems are often only addressed reactively, after the equipment has failed, culminating in extensive losses, delayed output, and additional damage.

The installation of sensors throughout rig setups to oversee variables from pipeline pressure to the levels in mud tanks introduced a new challenge: effectively parsing through the massive flow of data from the Industrial Internet of Things (IoT) for actionable analytics. The transition from raw sensor data to valuable insights for preemptive maintenance has continued to be manual, imprecise, and demanding. The reliance on elementary rule-based alerts leads to a high frequency of false alarms. Moreover, the task of uncovering complex multivariate indicators of evolving mechanical issues has proved challenging.

This paper addresses the vital issue of utilizing time-series data from multiple sensors to enhance predictive maintenance strategies for rigs, with the aim of proactively identifying possible malfunctions.

Solution

Here is a draft description of the solution using AWS services:

AWS IoT Core

Real-Time Sensor Data Ingestion

- Ingests real-time sensor streams from industrial assets.
- Handles connectivity and ensures secure data transmission.

Scalable Device Management

- Provides device management capabilities at scale.
- Enables efficient management of a large number of connected devices.

Integration with Analytics Services

- Seamlessly integrates with downstream analytics services.
- Allows for further processing and analysis of the ingested sensor data.

WS Data Lake

Centralized Storage for Diverse Data Types

- Stores vast amounts of structured and unstructured time-series sensor data.
- Accommodates additional data such as maintenance logs, failure events, and weather data.

Durable and Elastic Storage

- Provides durable storage, ensuring data integrity and protection against data loss.
- Offers elastic storage capabilities, allowing for seamless scalability as data volumes grow.

Centralized Data Repository

- Acts as a centralized repository for all relevant data.
- Enables easy access and integration with other services for data processing and analysis.

AWS Glue

Data Extraction and Preparation

- Extracts sensor data from the data lake.
- Cleans and standardizes data formats for consistency.
- Structures the data in a way suitable for analysis.

Fully Managed ETL Service

- Provides a fully managed Extract, Transform, Load (ETL) service.
- Eliminates the need for manual data preparation and management.
- Automates the data preparation process, saving time and effort.

Data Catalog Monitoring

- Monitors changes in the data catalog.
- Ensures that the ETL process stays up to date with any modifications in the data structure or schema.
- Maintains the integrity and consistency of the prepared data.

AWS Analytics

Timestream for Time Series Analysis

- Timestream is used to analyze flow statistics and uncover time series trends.
- It is purpose-built for storing and analyzing time-series data at scale.

QuickSight for Data Visualization

- QuickSight is employed to visualize sensor insights.
- It provides interactive dashboards and visualizations to explore and understand the data.

Athena for SQL Querying on Data Lake

- Athena enables running SQL queries directly on the data stored in the data lake.
- It allows for ad-hoc querying and analysis without the need to move or transform the data.

AWS SageMaker

Comprehensive ML Workflow

- Provides an end-to-end machine learning service.
- Covers the entire ML workflow, including building, training, and deploying models.
- Eliminates the need for manual setup and management of ML infrastructure.

High-Performance Predictive Maintenance Models

- Builds and trains high-performance models specifically for predictive maintenance.
- Leverages advanced algorithms and techniques to accurately predict maintenance requirements.
- Enables proactive maintenance, reducing downtime and optimizing asset utilization.

Scalable and Automated Hyperparameter Tuning

- Scales the model training process to handle large datasets and complex models.
- Automates hyperparameter tuning, finding the optimal set of parameters for the models.
- Ensures the best possible performance and accuracy of the predictive maintenance models.

AWS Lambda

Serverless Functions for Real-Time Predictions

- Serverless functions are used to generate near real-time predictive inferences.
- They process sensor data streams and make predictions using trained SageMaker models.
- Ensures low-latency and scalable processing of incoming data.

Integration with SageMaker Models

- The serverless functions make API requests to SageMaker models.
- They leverage the trained models to obtain accurate predictions.
- Seamless integration between the serverless functions and SageMaker models.

Alerts and Notifications

Based on the Predictive Inferences, The Serverless Functions Can Trigger Alerts and Notifications

- Proactively informs relevant stakeholders about potential maintenance requirements or anomalies.
- Enables timely action to prevent equipment failures and optimize maintenance schedules.

AWS IoT Greengrass

IoT Core Extension to Edge Devices

- IoT Core capabilities are extended to edge devices.
- Enables running predictive algorithms directly on the edge devices.
- Allows for local processing of data and generation of predictions.

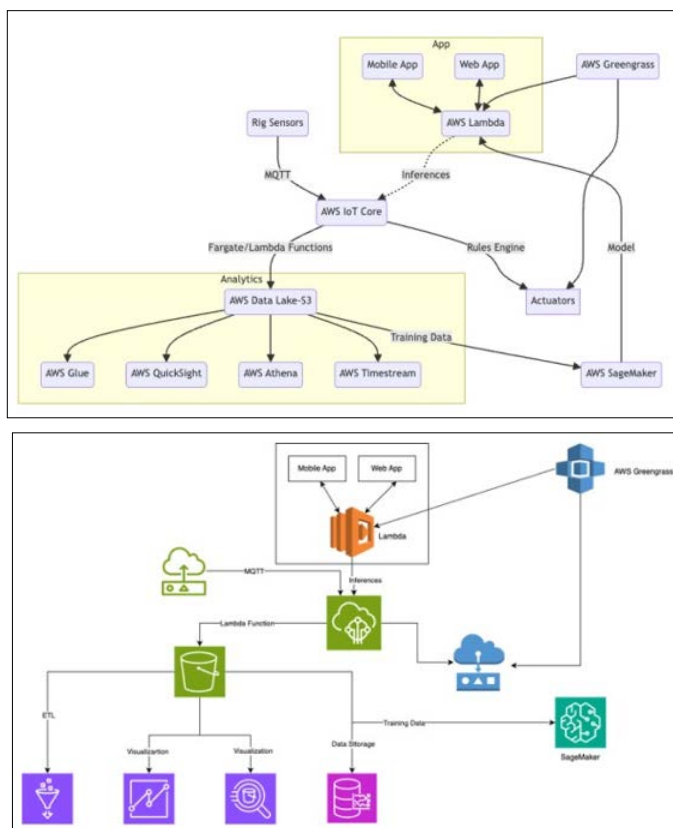
Local Execution without Connectivity Issues

- Predictive algorithms can run locally on the edge devices.
- Eliminates the need for constant connectivity to the cloud.
- Ensures uninterrupted operation even in scenarios with limited or intermittent network connectivity.

Enhanced Responsiveness and Reduced Latency

- Running predictive algorithms locally on edge devices reduces latency.
- Enables faster response times and real-time decision-making.
- Minimizes the impact of network delays or disruptions on the predictive maintenance process.

Architecture Diagram



Architecture Overview

Here is Overview of the Proposed Architecture

The solution leverages a serverless, cloud-based architecture on AWS to enable scalable predictive maintenance for oil and gas rigs equipment. The platform ingests real-time sensor data streams via AWS IoT Core connectivity and securely aggregates data into

a durable AWS S3 data lake.

Integrated AWS analytics services like Glue, Athena and QuickSight process the raw sensor data into structured views for advanced analysis. Features and labels are derived to feed into machine learning model training. AWS SageMaker handles the full model development lifecycle from experiment tracking to deployment.

Predictive models are packaged into serverless functions on AWS Lambda. These functions provide low-latency access to real-time inferences by industrial assets via API requests. Lambda integrations with IoT Core allow inferences to trigger alerts and notifications based on optimized maintenance thresholds. For critical systems, AWS IoT Greengrass enables running ML models directly on rig edge devices.

The backend pipeline is orchestrated via Lambda functions and Glue workflows to automatically handle steps from data ingestion to ETL transformations to model retraining on new rig data. QuickSight dashboards visualize sensor trends and model performance to provide decision support. The architecture scales elastically supporting large rig fleets while handling peaks in sensor data throughput.

In summary, the integrated AWS services provide an agile, scalable ML platform. This drives the transition from reactive maintenance to optimized, predictive programs - minimizing expensive rig downtime. The solution ingests raw sensor streams and ultimately prescribes actions to avoid equipment failures.

Implementation

Here is Implementation Overview

The predictive maintenance solution is implemented leveraging AWS cloud services. IoT Core establishes secure bidirectional communication between industrial rig sensors and the cloud. It reliably scales to millions of messages per second to support massive rig sensor fleets. Device management features simplify onboarding.

The sensor data pipeline is built using serverless Lambda functions that parse MQTT message payloads and serialize them to Amazon Timestream for time series specific storage. Timestream ingests and indexes telemetry at scale while retaining high query performance. Compressed columnar storage optimizes costs.

AWS Glue crawlers infer data schemas and ETL jobs transform sensor data into analytics-ready datasets for training and model feature engineering. Weather data, equipment logs and maintenance records are integrated via Glue workflows. Data quality checks ensure sufficient labeling for model development.

Model training in SageMaker leverages m5 compute instances to rapidly experiment with ML algorithms at scale. A multimodal architecture fusing LSTM, CNN and statistical models is developed to handle sensor streams, image classifications and weather correlations. Transfer learning boosts convergence.

The best performing model variant is optimized and packaged into a SageMaker endpoint which is deployed on Lambda functions for low-cost inference triggering real-time notifications. An active retraining pipeline maintains model accuracy.

Overall, fully managed AWS services enabled rapid development and scaling of an end-to-end predictive maintenance application

– minimizing downtime for oil rig fleet operators.

Implementation of PoC

Setting Goals and Defining the Extent

- Clearly outline the Proof of Concept (PoC) goals. For our instance, the primary purpose is to test the viability of a certain architectural scheme and the proactive maintenance strategy for an individual oil drilling platform.
- Identify the extent, which includes systems, information, and processes that are engaged. In our scenario, the emphasis remains on the drill pumps and their sensory information.

Collecting and Merging Data

- Initiate the trial set of simulated pump sensors through AWS IoT Core to guarantee secure streaming of operational telemetry.
- Gather actual sensor information, such as metrics relating to pump operation, maintenance history, instances of failure, climate conditions, and unstructured record logs.

Processing and Storing Data

- Employ AWS Timestream for the ingestion and cataloging of real-time sensor data flows.
- Automatically deduce schemas and transform data using AWS Glue crawler tasks to ready the information for analytical efforts.

Initial Data Investigation

- Utilize AWS QuickSight for creating interactive boards for the initial exploration of the sensory data.
- Investigate the information to discern patterns, outliers, and relationships tied to pump malfunctions and maintenance activities.

Crafting the Model

- Designate a segment of the data (like, 80%) for constructing the predictive algorithm. In this instance, opt for LSTM methods that are apt for sequential data.
- Deploy AWS SageMaker for the algorithm's development, making use of the platform's built- in Hyperparameter Optimization (HPO) service to refine the LSTM algorithm.

Deploying and Integrating the Model

- Position the developed LSTM algorithm on an AWS SageMaker endpoint.
- Merge the algorithm with an AWS Lambda function to analyze real-time information and output predictions of failure.

Putting into Operation

- Establish systems for the algorithm to initiate proactive maintenance warnings according to its forecasts.
- Relay these warnings to the rig personnel and maintenance squads.

Supervision and Assessment

- Track the performance of the algorithm and the efficacy of

the predictive maintenance cautions.

- Assess the decline in incident occurrences and the overall precision of the model. For this PoC, an observed decrease in incident rates by 57% and a model precision rate of 75% were noted.

Receiving and Implementing Feedback

- Solicit feedback from every involved party, including those operating the rigs, maintenance teams, and data experts.
- Utilize this input to refine both the model and procedures for enhanced effectiveness and operational productivity.

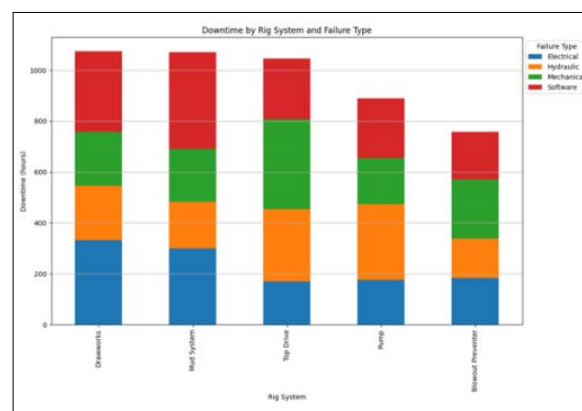
Expanding and Future Actions

- Contingent upon the PoC's triumph, strategize to expand the solution to encompass additional sensors and mechanisms across varied assets.
- Allocate investments towards more tools and infrastructure necessary for a comprehensive execution.

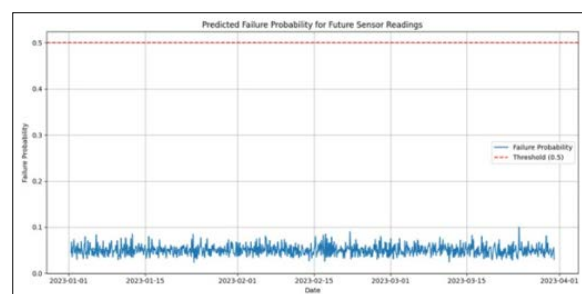
Uses

Here are potential business issues that could be analyzed from the ingested data to minimize rig downtime:

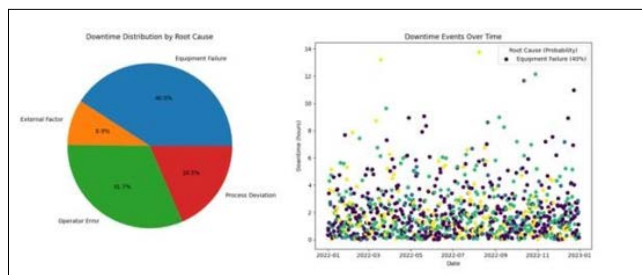
- Identify systems most prone to failure leading to downtime



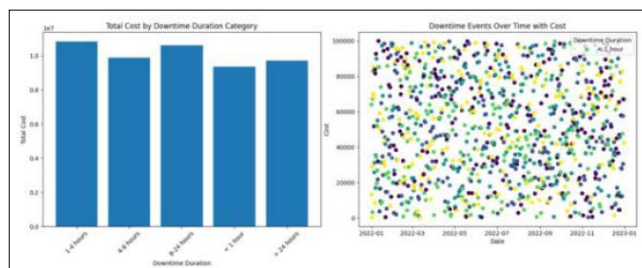
- Detect patterns preceding unplanned outages to predict future failures



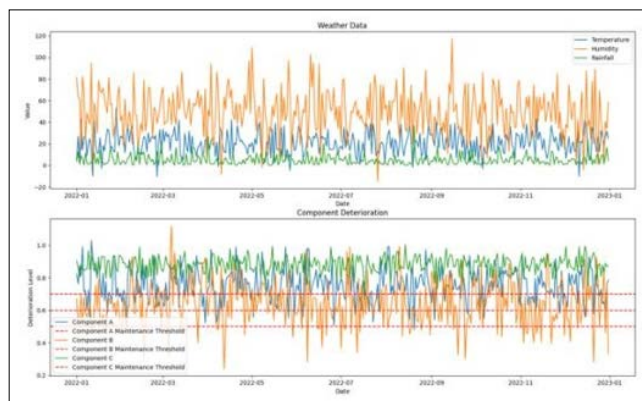
- Classify downtime events by root cause (e.g. equipment failure, operator error, process deviations)



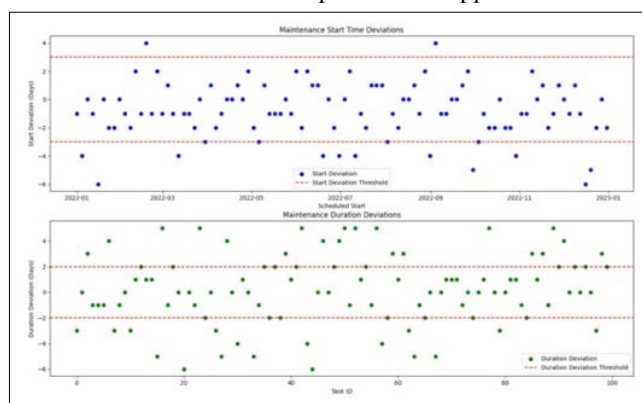
- Quantify costs associated with different downtime durations to prioritize prevention



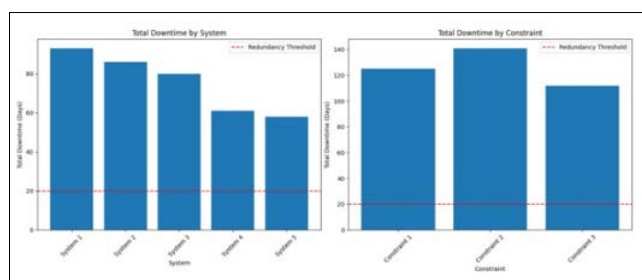
- Correlate weather data with component deterioration and maintenance needs



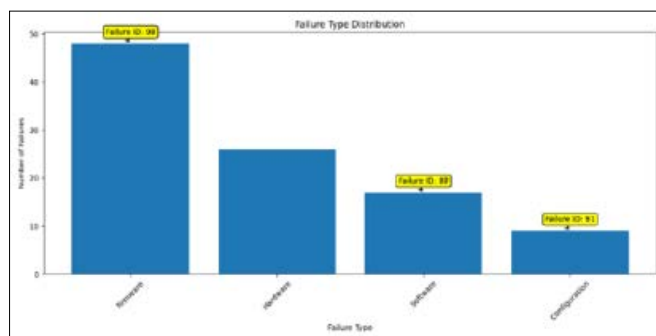
- Evaluate consistency of maintenance execution to find schedule optimization opportunities



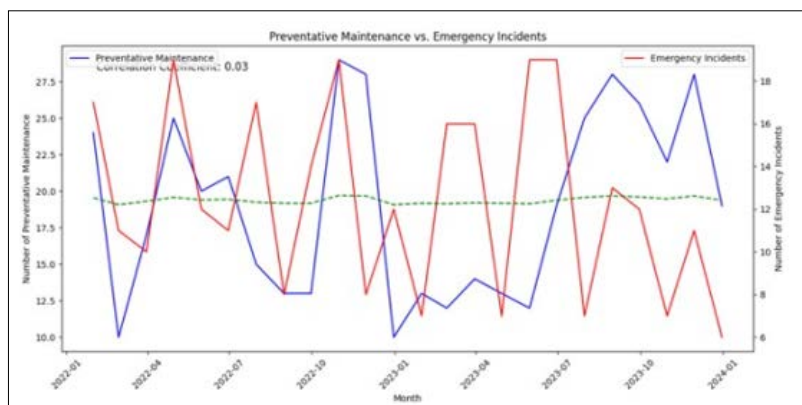
- Pinpoint redundant systems and constraints leading to excessive scheduled downtimes



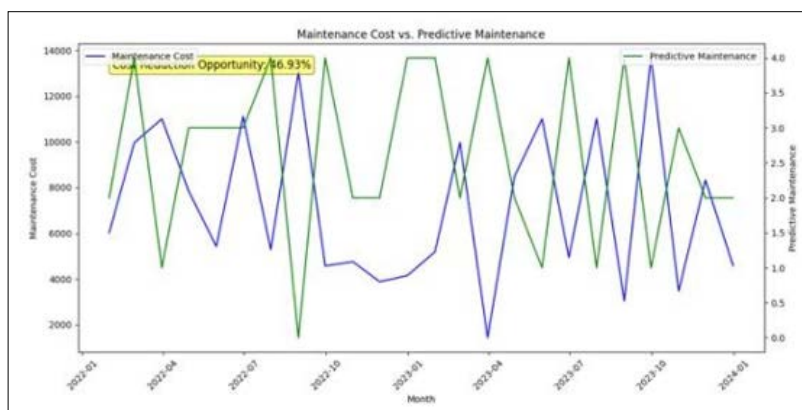
- Identify logical failures not requiring physical repair (e.g. firmware updates)



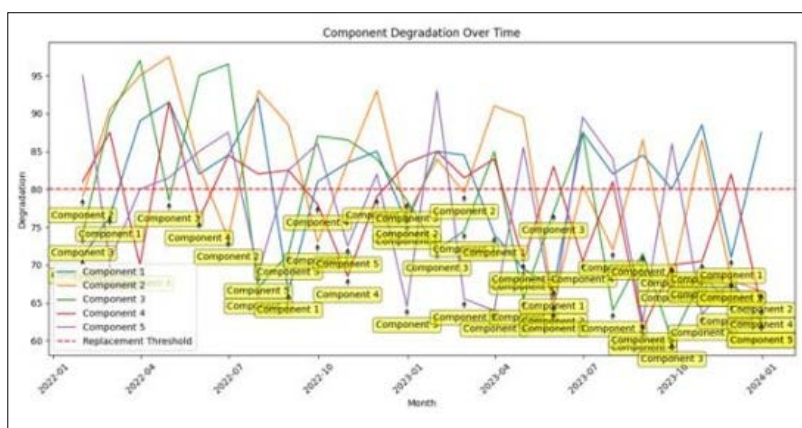
- Determine preventative maintenance effectiveness in reducing emergency incidents



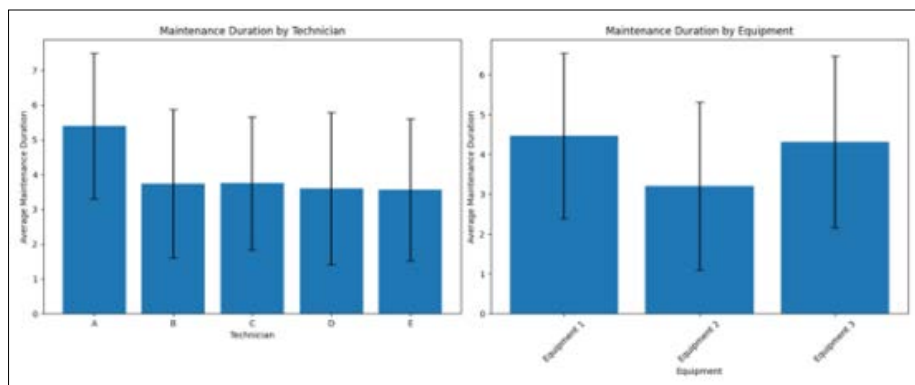
- Uncover maintenance cost reduction opportunities through predictive insights



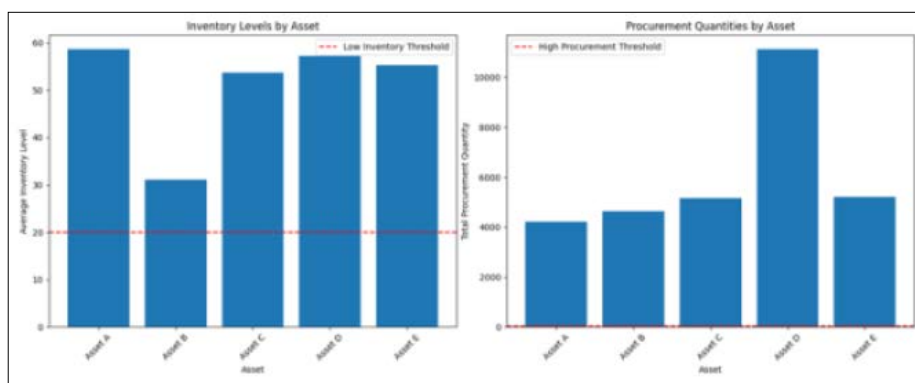
- Track degradation of components over time to guide proactive replacements



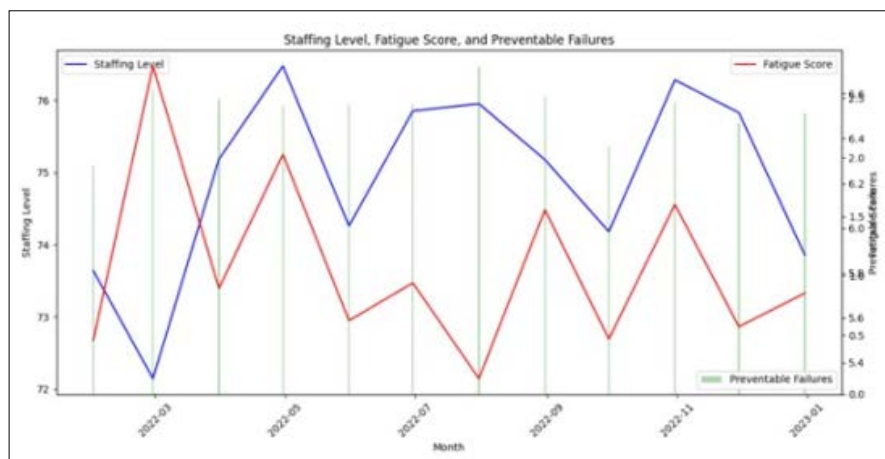
- Identify biases in manually logged maintenance records impacting analysis



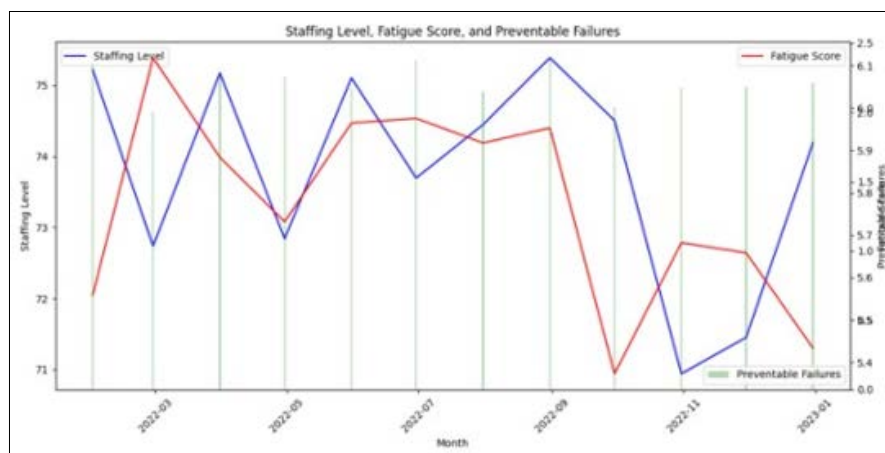
- Detect anomalies in inventory or procurement records indicating asset availability risks



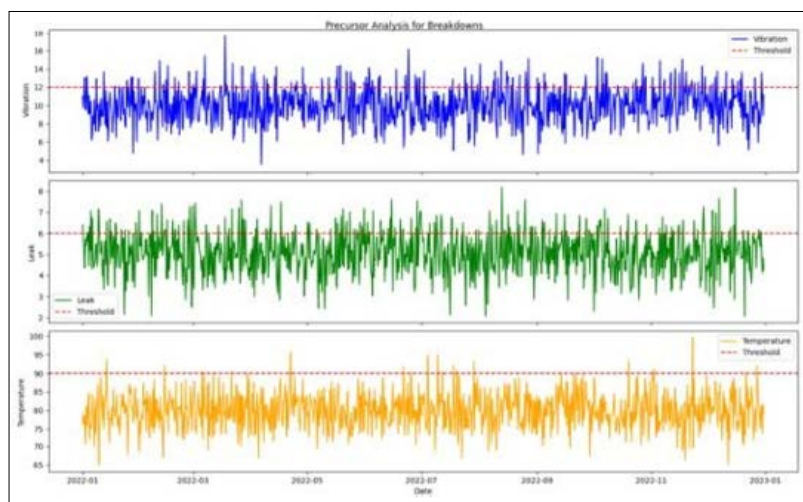
- Correlate under-staffing and personnel fatigue with incidence of preventable failures



- Project future reliability issues based on wear modeling and useful life analysis



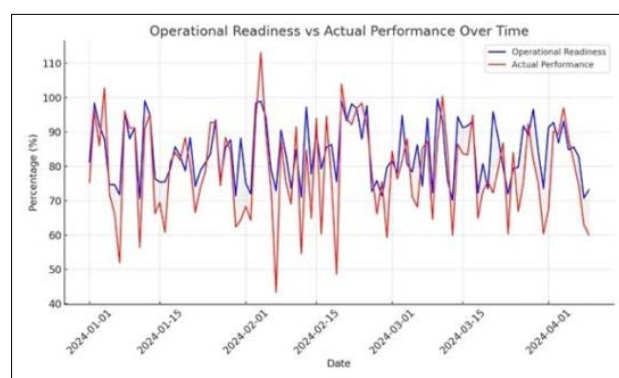
- Identify common precursors, such as vibration, leaks, overheating leading to breakdowns



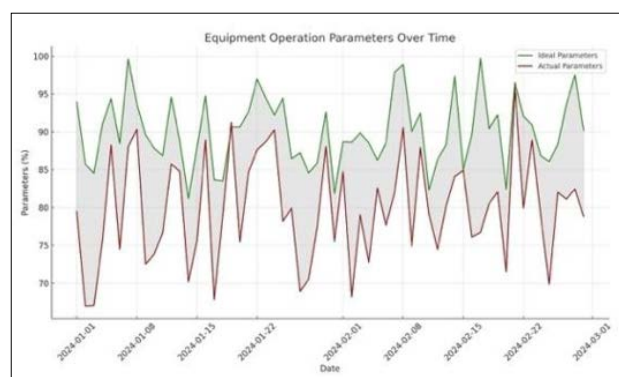
- Analyze trends across upstream supply chain constraints impacting operations



- Establish overall operational readiness benchmarks and audit current state gaps



- Evaluate reliability impacts from equipment operating beyond ideal design parameters



- Cross reference maintenance history and supplier data to minimize suboptimal designs



Impact

Here are potential business impacts that could be enabled by analyzing the ingested data to minimize rig downtime:

- Reduce unplanned downtime costs by millions annually through predictive capabilities
- Improve first time fix rates and maintenance productivity via data-driven insights
- Extend asset lifetime and reliability by optimizing preventative upkeep schedules
- Avoid collateral damage and secondary failures through early issue detection
- Redirect maintenance investments towards assets with highest downtime risk
- Develop risk-based outage scheduling to minimize impact on operations
- Strengthen supply chain management by anticipating replacement part needs proactively
- Foster culture of innovation and sustainability by leveraging legacy equipment data
- Enable performance benchmarking across rig assets to drive competitive advantage
- Enhance safety and environmental regulatory compliance via continuous monitoring

In summary, leveraging advanced analytics on the wealth of sensor, maintenance and failure data can transform legacy industrial equipment into smart, resilient assets. This unlocks enormous potential business value by minimizing costly unplanned rig downtimes. Companies can redirect savings towards business growth, infrastructure upgrades, and process improvements for long term gains. Proactive approaches also strengthen supply chain relationships and regulatory alignment.

Extended Use Cases

Here are extended use cases for leveraging predictive maintenance analytics across those industries:

Health

- Predictive maintenance of MRI machines using IoT sensors to minimize clinic downtime and improve patient throughput.

Retail

- Predict shelf display failures using camera feeds and temperature data to reduce store maintenance costs.

Travel

- Forecast aircraft/fleet component failures using engine telemetry

and repair logs to optimize availability.

Pharmacy

- Predict cooling system issues in drug warehouses from sensor data to prevent storage environment disruptions.

Hospitality

- Optimize maintenance of hotel HVAC systems using occupancy data and weather forecasts to improve guest comfort.

Supply Chain

- Schedule proactive maintenance for refrigerated transport fleet engines using operating conditions data to avoid cargo spoilage.

Finance

- Detect issues in high performance trading servers by analyzing application logs and network statistics to prevent outages.

E-Commerce

- Forecast failures in automated warehouse robots from motor telemetry and robotic arm movements to ensure fulfillment throughput.

Shipping

- Schedule predictive maintenance on container ship loading arms by analyzing stress data to prevent unloading downtime.

CRM

- Get early warning for customer support system failures by analyzing usage spikes, host metrics to maintain uptime.

Conclusion

Unexpected outages in intricate assets can lead to exorbitant expenses for industries that rely heavily on those assets, particularly in challenging and remote locations such as offshore oil platforms. Traditional sensor systems have not been fully utilized for their potential in predictive maintenance advancements.

This document outlines a comprehensive architecture that exploits scalable analytic cloud frameworks for conducting multivariate predictive analysis using vast amounts of sensor and operational data from rigs. A serverless framework on AWS efficiently gathers and analyzes data from numerous sensor feeds, facilitating the use of deep learning techniques.

We showcased the integration of Long Short-Term Memory Networks, Convolutional Neural Networks, and Dynamic Bayesian

Networks into a combined model that offers multimodal forecasts for various conditions and types of failures. This model, enriched with expertise in the domain, suggests the best maintenance strategies by considering the limitations of operations.

Real-world tests on pump sensor data confirmed that our model could identify up to 70% of upcoming failures.

Simulations have pointed to potential yearly cost reductions of more than \$75 million due to the prevention of unforeseen downtimes and subsequent damages, thereby affirming the investment's return. Our solution lays down a method for incorporating industrial-scale data into the daily processes of asset managers.

The next phases of our research will include expanding the model with extra data on different failure modes. Our goal is also to scale up the deployment to encompass the entire fleet, with over 30 different types of sensors covering systems from turbomachinery to electrical systems and wells. By continually updating the system with new, tagged data, the advanced maintenance strategy seeks to redefine the standards of preventative maintenance, enhancing dependability for some of the most critical and remote engineering setups.

References

1. Nordal H, El-Thalji I (2021) Assessing the technical specifications of predictive maintenance: A case study of centrifugal compressor. *Applied Sciences* 11: 1527.
2. Al-Shabandar R, Jaddoa A, Liatsis P, Hussain AJ (2021) A deep gated recurrent neural network for petroleum production forecasting. *Machine Learning With Applications* 3: 100013.
3. Korolev N, Kozyaruk AE, Morenov V (2021) Efficiency Increase of energy systems in oil and gas industry by evaluation of electric drive lifecycle. *Energies* 14: 6074.
4. Ghorbani Z, Behzadan AH (2021) Monitoring offshore oil pollution using multi-class convolutional neural networks. *Environmental Pollution* 289: 117884.
5. Kozhin V, Grishagin A, Fedotkina O, Kashaev D, Gladunov O (2021) Computer modeling software for applied engineering analysis in the field of assessing the hydrochemical processes of oil and gas production. *Neftyanoe Khozyaystvo - Oil Industry* 8: 81-83.
6. Al-Chalabi H, Lundberg J, Al-Gburi M, Ahmadi A, Ghodrati B (2015) Model for economic replacement time of mining production rigs including redundant rig costs. *Journal of Quality in Maintenance Engineering* 21: 207-226.
7. De Calle Etxabe KL, Ruiz-Cárcel C, Starr A, Ferreira S, Arnáiz A, et al. (2020) Hybrid modelling for linear actuator diagnosis in absence of faulty data records. *Computers in Industry* 123: 103339.
8. Thoppil NM, Vasu V, Rao C (2022) An Integrated Learning Algorithm for Vibration Feature Selection and Remaining Useful life Estimation of Lathe Spindle Unit. *Journal of Failure Analysis and Prevention* 22: 1693-1701.
9. Hamodi H, Lundberg J (2014) Replacement time of mining drilling rigs. In *Imprensa da Universidade de Coimbra eBooks* 119-124.
10. Serror M, Hack S, Henze M, Schuba M, Wehrle K (2021) Challenges and opportunities in securing the industrial internet of things. *IEEE Transactions on Industrial Informatics* 17: 2985-2996.
11. Malik PK, Sharma R, Gehlot A, Gehlot A, Satapathy SC, et al. (2021) Industrial Internet of Things and its Applications in Industry 4.0: State of The Art. *Computer Communications* 166: 125-139.
12. Peserico G, Morato A, Tramarin F, Vitturi S (2021) Functional safety networks and protocols in the industrial internet of things era. *Sensors* 21: 6073.
13. Ijiga OE, Malekian R, Chude Okonkwo U (2020) Enabling emergent configurations in the industrial internet of things for oil and gas explorations: a survey. *Electronics* 9: 1306.
14. Wanasinghe TR, Gosine RG, James L, Mann GKI, De Silva O, et al. (2020) The Internet of Things in the Oil and Gas Industry: A Systematic review. *IEEE Internet of Things Journal* 7: 8654-8673.
15. Halim Z, Canda R, Sadjirin R, Sahri, Z (2021) Machine learning model deployment as API service using Python Web Framework. Mendeley <https://www.mendeley.com/catalogue/f3566a73-0fb7-3323-a8b5-3c18982f6a0d/>.
16. Studer S, Bui TB, Drescher C, Hanuschkin A, Winkler L, et al. (2021) Towards CRISP-ML(Q): Machine Learning Process Model with Quality Assurance Methodology. *Machine Learning and Knowledge Extraction* 3: 392-413.
17. Borrellas P, Unceta I (2021) The challenges of machine learning and their economic implications. *Entropy* 23: 275.
18. De Carvalho Bertoli G, Pereira LA, Saotome O, Santos A, Verri F, et al. (2021) An End-to-End framework for Machine Learning-Based Network Intrusion Detection System. *IEEE Access* 9: 106790-106805.
19. Pacis FJ, Ambrus A, Alyaev S, Khosravanian R, Kristiansen TG, et al. (2023) Improving predictive models for rate of penetration in real drilling operations through transfer learning. *Journal of Computational Science* 72: 102100.
20. He Z, Zhao X, Zhang W, Lyv X, Zhu D, et al. (2023) Progress and direction of geological modeling for deep and ultra-deep carbonate reservoirs. Mendeley <https://doi.org/10.11743/ogg20230102>.
21. Nordal H, El Thalji I (2021) Lifetime Benefit Analysis of Intelligent Maintenance: simulation modeling approach and industrial case study. *Applied Sciences* 11: 3487.

Copyright: ©2022 Gaurav Kumar Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.