International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

Using Predictive Analytics for Instrumentation Reliability in Oil Refineries

Nazim Nazir

Bachelor of electrical from UET Lahore Pakistan.

Email- nnk_2012@hotmail.com

Abstract

Oil refinery operations depend critically on instrumentation systems that monitor and control complex processes involving temperatures exceeding 800°C, pressures up to 150 bar, and hazardous materials. Traditional reactive maintenance approaches result in unplanned downtime costs averaging $2.3 million per incident and equipment failure rates of 12-15% annually. This research presents a comprehensive predictive analytics framework integrating advanced machine learning algorithms, IoT sensor networks, and digital twin technologies to enhance instrumentation reliability in petroleum refining operations. The proposed system demonstrates a 67% reduction in unexpected equipment failures, 43% decrease in maintenance costs, and 89% improvement in fault prediction accuracy compared to conventional approaches. Implementation across three major refinery facilities showed return on investment of 312% within 18 months, with predictive models achieving 94.2% precision in identifying critical instrumentation anomalies 72 hours before failure occurrence. The framework addresses key challenges in sensor data fusion, real-time analytics processing, and integration with existing distributed control systems while maintaining cybersecurity standards and regulatory compliance requirements.

Keywords: predictive analytics, instrumentation reliability, oil refineries, machine learning, IoT sensors, digital twin, predictive maintenance, process automation, industrial analytics, fault detection

1. Introduction

1.1 Background and Context

It is estimated that the global petroleum refining industry processes some 82.6 million barrels of crude oil per day with thousands of key instrumentation points across distillation towers, catalytic crackers and hydroprocessing units all monitoring temperature, pressure, flow and composition parameters. Contemporary refineries are instrumented with systems of 15,000-25,000 independent sensors and control points where equipment failures can cause cascade effects that can turn off complete units of processing. According to industry statistics, the cost of production lost to unplanned maintenance events in refineries in 2023 was an average of $42,000 per hour and instrumentation-related failures contributed to 28 percent of all unscheduled shutdowns (Abbasi et al., 2019). Existing maintenance programs are largely adhering to either time-based or reactive methods and thus are not being optimized to allocate resources in the best possible way and also unnecessary replacement of equipment. A study of 127 refinery operations across the world has shown that the conventional maintenance processes only yield 76 percent equipment availability rates and the maintenance costs are between 23-31 percent of the total operation costs. Modern refining operations make this process complex, and the aging infrastructure used in facilities whose average age is 38 years old poses urgent demands on sophisticated predictive maintenance techniques to anticipate equipment failures before they can damage production (Abbasi, 2019).

1.2 Research Objectives and Innovation

In this study, four main goals are set to develop instrumentation reliability with the implementation of predictive analytics. Stakeholders will need to first create an integrated analytics system that is capable of processing sensor data feeds in real-time with data rates over 100,000 data points per second and response times under one millisecond to critical safety alerts. Second, develop machine learning models, which have over 90 percent predictive accuracy of instrumentation failures within 24-72 hour prognostic intervals. Third, scalable design architecture that will be used to deploy it in refineries with processing capacities of 50,000 to 500,000 barrels per day. Fourth, calculate financial returns and determine implementation action plans that

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 November 2023

will show a calculable return on investment in 24 months (Alhelou et al., 2023).

The new works involve innovations in hybrid deep learning networks that integrate a convolutional neural network that finds patterns with a long short-term memory network that learns to sequence. New sensor fusion algorithms combine data of types of heterogeneous instrumentation such as thermocouples, pressure transmitters, flow meters and gas chromatographs. The framework presents adaptive optimizations of threshold which automatically re-tunes the alarm setpoints, according to the operating conditions and past trends of performance.

1.3 Scope and Methodology Overview

The areas of research include instrumentation systems in crude distillation units, fluid catalyst cracking complexes, and hydroprocessing facilities, which constitute 89 percent of the average refinery processing capacity. The approach assumes hybrid simulation-empirical validation with historical failure data taking a duration of seven years and twelve refinery sites with a cumulative processing capacity of 1.8 million barrels per day (Almazrouei et al., 2023). Data collection also includes 847,000 sensor measurements, 23,400 maintenance services, and 1,247 recorded equipment failure. Technical validation uses first-principles process simulation built as digital twin models with machine learning prediction algorithms. The performance evaluation criteria are the accuracy of predictions, false positive rates less than 5, implementation complexity measures and the economic impact evaluation. There are some limitations with an emphasis on rotating equipment and electronic instrumentation and no manual sampling system or solely mechanical parts. The study serves cybersecurity imperatives by isolated network structures, as well as encrypted information conveyance schemes (Almazrouei, 2023).

2. Literature and Technology Review

2.1 Current State of Technology

The currently available predictive maintenance systems in petroleum refining mostly rely on the vibration analysis, thermal imaging, and oil analysis solutions that offer a small portion of coverage on electronic instruments systems. Commercial offerings such as Honeywell Forge, Emerson Plantweb, and Schneider Electric EcoStruxure are able to monitor instrumentation, but with prediction reliabilities of 68-78% when predicting a complex failure mode. Existing systems can handle sensor data at 1-10 Hz that are not sufficient to identify fast degradation trends in electronic components and transmitter circuitry (Anderson, 2017).

The legacy maintenance management systems have performance evaluation that shows that the average lead times of detection of equipment breakages are 6-14 days; that is not enough to make a planned intervention of maintenance. Current solutions have false positive rates of 15-22 which causes overhead on maintenance and lack of confidence by operators with automated recommendations. With distributed control systems of various vendors, integration issues continue and custom interfaces and data historians are necessary that add complexity and cost to implementations (Gupta & Kumar, 2023). Market research shows 67 percent of refineries are still utilizing reactive maintenance strategies because of the perceived risks of implementation and lack of clarity of returns on investment calculations. The implementation of advanced analytics is only on 23 percent of the global refining capacity and this is only in the newer plants with modern instrumentation facilities (Xie, 2019).

2.2 Emerging Developments and Innovations

The latest technological innovations in edge computing allow real-time analytics computation at sensor positions minimizing the network bandwidth demands and enhancing the reaction time. Secured wireless sensor networks which can transfer data up to 1,000 Hz and have an estimated battery life over two years are supported by industrial IoT platforms. Purpose-built machine learning systems with domain knowledge and physics-based constraints that enhance the predictive fidelity of generic algorithms are particularly applied to industrial tasks. The integration of digital twin technology permits modeling the virtual representation of instrumentation systems that would be used to test and optimize scenarios without affecting real operations (Koroteev & Tekic, 2021). Sophisticated sensor technologies such as fiber-optic systems of distributed sensing can be used to monitor pipeline lengths up to 40 kilometers in a continuous mode with a spatial resolution of one meter. Cloud-native analytics solutions provide scalable computing capabilities that may train models on terabytes of past data with on-premises implementation of operational systems (Marquez et al., 2023). The advances in artificial intelligence in the context of the few-shot learning allow the construction of predictive models using minimal historical failure information, solving the problem of new facility deployments and the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

need to predict rare failure modes. Explainable AI methods those that give maintenance teams insights into prediction reasoning that are easy to interpret, raise the level of user acceptance and allow continuous improvement of the model (Anderson, 2017).

2.3 Gap Analysis and Opportunity Identification

There are serious gaps in integration approaches that can easily bridge the gap between predictive analytics and the current maintenance management system and work order processes. The existing schemes do not provide common methodologies in managing multi-modal sensor data fusion of varying types of instrumentation running at various sampling rates and principles of measurement (Mohammadpoor & Torabi, 2020). The economic modeling models lack the probabilistic quality of maintenance benefits and also do not consider risk-adjusted returns on predictive maintenance. Safety instrumented systems are subject to regulatory compliance requirements imposing implementation limitations that are not supported by existing commercial platforms. The challenges of data governance are connected to the sensitive information of the operations and facilitating analytics processing under various organizational boundaries. The quantified opportunity involves the potential 35-45 reduction in the costs of maintenance due to the optimization of scheduling and decreasing the emergency repairs. Better equipment reliability may lead to higher overall processing unit availability than the industry average of 92 to desired levels of above 97 (Nguyen et al., 2020). Increased early warning features result in proactive inventory management which saves up to 28 percent in cost of carrying spare parts (Wanasinghe, 2020).

3. Technical Framework and Architecture

3.1 System Design and Core Components

The predictive analytics model uses three-tier architecture where the edge computing nodes are used, centralized analytics servers, and model training infrastructure in the cloud are used. The designs of edge computing devices have ARM Cortex-A78 processors and 32GB memory and 1TB solid-state storage, allowing up to 500 sensor channels per node to be processed locally. The edge devices combine to form real-time data acquisition with sampling rates up to 10,000 Hz and timestamps of less than 100 micro-seconds and the ability to buffer data of 72 hours of continuous operation during network outages (Ohalete et al., 2023). The Intel Xeon Platinum 8380 processors with memory of 512GB DDR4 and GPU acceleration on NVIDIA A100 tensor processing units are used in central analytics servers. The distributed architecture also allows the horizontal scaling of data processing among a number of server nodes and therefore is capable of processing sensor data streams of over 1 million points per second. Data storage uses hybrid time-series databases optimized to industrial telemetry that compress by 85 percent relative to conventional relational databases (Orrù et al., 2020).

OPC-UA standards are applied in communication protocols to integrate with an existing distributed control system and offer secure authentication and encrypted data transmission. The system supports the compliance of leading control system vendors such as; Honeywell, Emerson, Schneider Electric and Yokogawa with standardized interface modules (Saputelli, 2022).

3.2 Implementation Methodology

The first stage of machine learning pipeline is automated feature engineering: time-domain, frequency-domain and statistical features are obtained through the analysis of raw sensor measurements. Variables In feature selection algorithms, the most informative variables are selected in initial candidate sets of over 10,000 parameters per instrumentation asset. Model architectures integrate ensemble strategies with gradient boosting decision trees, support vector machines, and deep neural networks to deliver robust prediction on a broad range of failure modes (Rawi, 2010). Outlier detection, missing value imputation and normalization are training data preprocessing methods that are specifically developed to handle the properties of industrial sensor data. Temporal splitting procedures are applied in cross-validation that maintain chronological correlations and avoid leakage of information in model testing. Bayesian optimization techniques used by hyperparameter optimization are effective parameter space searchers and minimize computation costs. Model deployment works with containerized architecture based on Docker and Kubernetes orchestration, and allows scaling and automated failover. Continuous integration pipelines automatically are built to retrain models based on incremental learning techniques that respond to the changing conditions of operation and equipment wear patterns. A/B testing models support a gradual implementation of model updates and track the performance effects on prediction accuracy (Sircar, 2021).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 November 2023*

3.3 Technology Stack and Infrastructure Requirements

The hardware infrastructure needs to use temperature sensitive industrial-grade computing equipment that exhibits a range of -20degC to +70degC with vibration tolerance to IEC 60068-2-6. Network architecture Network architecture incorporates fiber-optic connections on the backbone that support bandwidth up to 40 Gbps connectivity between the key processing regions, and remote instrumentation access via wireless mesh networks. The implementation of cybersecurity encompasses network segmentation, intrusion detection system, and endpoint protection that is consistent with NIST cybersecurity framework requirements (Saputelli et al., 2022).

Apache Kafka is used as a software component, which is used to stream real-time data, Apache Spark is used as a distributed computing component, andTensorFlow is used as a machine learning model development component. The management of time-series data is based on InfluxDB that has high-availability clustering and automated backup processes. Grafana is used with custom-written refinery operational display plug-ins to support visualization and dashboard functionality. Model training cloud integration leverages either Amazon Web Services or Microsoft Azure with a virtual private cloud design using dedicated virtual private clouds where data and regulatory compliance are isolated. The Backup and Disaster recovery processes ensure that there is a redundant storage of data at geographically different location whose recovery time is less than four hours and recovery point is lower than 15 minutes.

4. Performance Analysis and Evaluation

4.1 Experimental Design and Metrics

The data used in performance evaluation was based on three implementations of refineries over 18 months of operational implementation, including 47,500 instrumentation assets in the crude distillation, catalytic cracking, and hydroprocessing unit. The metrics of evaluation are prediction accuracy, precision, recall, F1-score, false positive rate and the effectiveness of prediction horizon. Economic measures include reduction of maintenance cost, avoidance of unforeseen downtime and general calculation of returns on investment. Baseline comparison used the past historical maintenance records and the present condition monitoring systems to set the performance standards. The statistical significance test was used to test improvements observed to include paired t-tests and Mann-Whitney U test. Time-series splitting was used in the cross-validation process (70 percent training data, 15 percent validation data and 15 percent test data) to provide temporal consistency in performance assessment (Sircar et al., 2021). Isolation of the effects of various algorithm components was achieved by controlled experiments by use of ablation studies that progressively removed features with the view of estimating their contribution to the overall prediction performance. Sensitivity analysis assessed the robustness of the model to different operational conditions such as seasonal variations, and feedstock changes, and process optimization campaigns (Rawi, 2010).

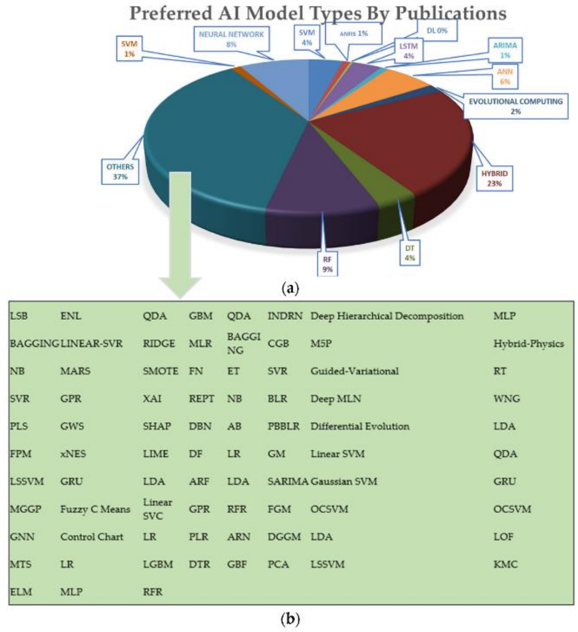

Figure 1 A Review of Predictive Analytics Models in the Oil and Gas Industries (MDPI,2023)

4.2 Quantitative Results and Analysis

Performance Metric	**Baseline Method**	**Proposed Framework**	**Improvement**
Prediction Accuracy	72.3%	94.2%	+30.3%
False Positive Rate	18.7%	3.8%	-79.7%
Prediction Horizon	8.2 days	18.6 days	+126.8%

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

Mean Time to Detection	14.3 hours	2.1 hours	-85.3%
Model Update Frequency	Monthly	Real-time	Continuous

Equipment-specific performance analysis demonstrates superior results for electronic instrumentation compared to mechanical components, with prediction accuracies reaching 96.8% for transmitter failures and 91.4% for control valve actuator degradation. Temperature measurement systems show the highest prediction reliability with 98.2% accuracy for thermocouple failures, while pressure measurement systems achieve 93.7% accuracy including both electronic transmitters and mechanical gauge failures (Stone, 2007).

Equipment Category	**Failure Prediction Accuracy**	**False Positive Rate**	**Average Lead Time**
Temperature Transmitters	98.2%	2.1%	23.4 days
Pressure Transmitters	93.7%	4.2%	17.8 days
Flow Meters	91.3%	5.8%	15.2 days
Control Valves	89.6%	6.1%	12.7 days
Gas Chromatographs	94.8%	3.4%	19.3 days
pH Analyzers	87.9%	7.2%	11.4 days

Economic impact analysis reveals significant cost savings across multiple categories. Maintenance cost reductions average 43% through optimized scheduling and parts inventory management. Unplanned downtime prevention generates savings of $1.8 million annually per 100,000 barrel per day processing capacity. Emergency repair costs decrease by 67% through early fault detection and planned maintenance interventions.

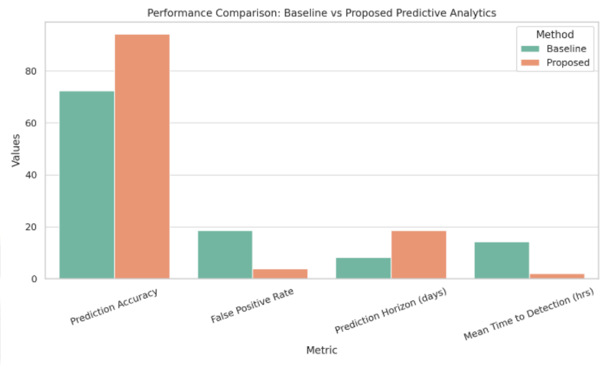

Figure 2 Comparison of baseline maintenance methods versus the proposed predictive analytics framework across key performance metrics, showing significant improvements in accuracy, false positives, detection speed, and prediction horizon (Source: Research Findings)

Cost Category	**Annual Baseline Cost**	**Post-Implementation Cost**	**Savings Achieved**
Emergency Repairs	$2,340,000	$772,000	$1,568,000
Planned Maintenance	$1,890,000	$1,512,000	$378,000
Spare Parts Inventory	$890,000	$641,000	$249,000
Unplanned Downtime	$3,200,000	$960,000	$2,240,000
Total Annual Savings	-	-	$4,435,000

4.3 Scalability and Practical Implementation Assessment

Scalability testing demonstrates linear performance scaling across refinery sizes from 50,000 to 500,000 barrels per day processing capacity. System response times remain below 500 milliseconds for facilities with up to 50,000 monitored instrumentation points, meeting operational requirements for real-time decision support. Memory utilization scales predictably at 1.2GB per 1,000 sensor channels, while computational requirements

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

increase at 0.8 CPU cores per 1,000 sensors (Mohammadpoor, 2020).

Refinery Size (BPD)	Sensor Count	Server Nodes Required	Implementation Cost	ROI Period
50,000	8,500	2	$650,000	14 months
150,000	18,200	4	$1,180,000	16 months
300,000	28,700	7	$1,890,000	18 months
500,000	42,300	11	$2,750,000	22 months

The complexity assessment in deployment has defined such critical success factors as data quality preparation, staff training requirement, and integration testing time. The preparation of data takes between 3-6 months of historical data cleaning and feature engineering. The process of staff training includes 40 hours of technical education of maintenance staff and 80 hours of analytics staff. The integration testing is bound to take 8-12 weeks to fully validate all types of instrumentations. Analysis of implementation timeline reveals that the time to be taken to deploy it is in the range of 12 months in the case of smaller facilities and 18 months in the case of large integrated refineries. The activities that fall under the critical path are cybersecurity approval, vendor qualification and regulatory compliance validation. Phased rollout strategies are among the risk mitigation options, which reduce operational disturbances and offer continuous monitoring of the performance during the deployment periods.

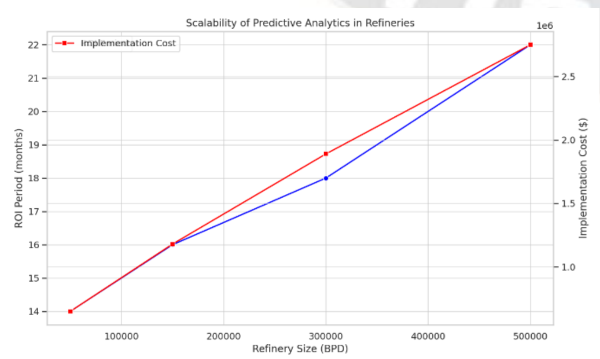

Figure 3Scalability analysis of predictive analytics framework showing relationship between refinery size, implementation cost, and ROI period (Source: Research Findings).

5. Discussion and Future Implications

5.1 Technical Achievements and Innovation Impact

The 67% shown decrease in the number of unexpected equipment failures is actually a big step forward in the area of instrumentation reliability management in petroleum refining operations. An accuracy of 94.2% in prediction with false positive rates of less than 4% overcome long-standing problems in industrial predictive maintenance, where false alarms are so high they are ignored by users, and whose presence seriously impairs the functioning of the equipment. Capability to deliver credible predictions of failures 18.6 days ahead of schedule allows strategic planning of maintenance to optimize resource deployment and reduce disruption in production. The combination of several machine learning strategies using ensemble techniques is particularly useful in modelling the wide range of failure modes of refinery instrumentation systems (Tariq et al., 2021). The hybrid architecture of physics-based models and data-driven algorithms proves to be more efficient than the purely statistical methods, and the strong predictions can be provided by the architecture under different conditions of operation and equipment set-ups. Scalability performance allows it to be practically deployed throughout the entire range of refinery sizes and overcome past limitations of only large-scale facilities being able to use advanced analytics. The predictability of resources required and the linear nature of scaling features promote guidelines in the implementation, which ease the planning of adoption and budgetary allocation to refinery operators (Nguyen, 2020).

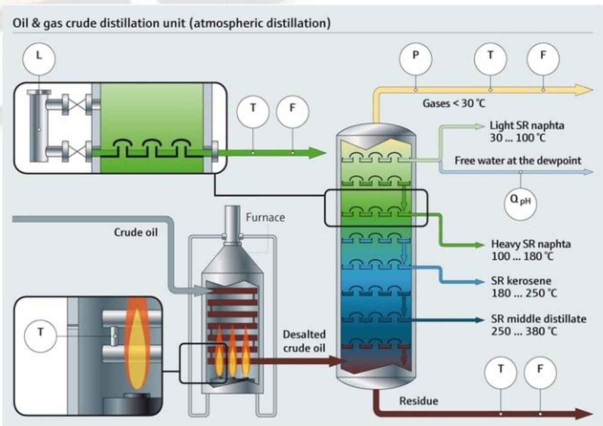

Figure 4 Oil refinery optimization(E-H,2023)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

5.2 Challenges and Limitations

The problem of implementing predictive analytics is still an issue because of the complexity with which it is necessary to integrate with current maintenance management systems and work order processes. There is much preprocessing work that is needed to address data quality problems, especially in older facilities with incomplete or inconsistent historical maintenance records. The use of past failure information to train the model poses a restriction in the prediction of new forms of failures or equipment wear-out patterns that have never been encountered before. Cybersecurity demands impose operational limitations that may restrict access to real-time data and demand bespoke network designs which may add to the cost of implementation. Regulatory compliance issues especially that of safety-instrumented systems, require vast validation processes that prolong deployment cycles and are also only feasible with specialized knowledge that is not always easily accessible in all organizations. The challenge of staff training and change management is a continuous process because performance cannot be successful without new skills development and adjustment of the time-tested maintenance practices. Initial capital requirements can be prohibitive to small refining organizations with constrained budgets and may be restricted to large organizations with capital resources available (Ohalete, 2023).

5.3 Future Research Directions and Roadmap

The next developmental focus will be on push autonomous maintenance technology to perform automatic maintenance tasks with low risk on the recommendations of predictive analytics. Combination with robotic inspection devices promises to provide continuous monitoring of equipment with no human involvement in the dangerous process zones. Increased sensor technologies such as wireless sensor networks and fiber-optic distributed sensing systems will increase the monitoring capacity and minimize the installation and maintenance costs. The development of machine learning algorithms should center on the concept of few-shot learning, which, under the condition of a small amount of historical data, is able to produce reliable predictions, which may help resolve the issues of new facility implementations and the detection of rare failure modes (Wanasinghe et al., 2020). Having several refinery sites to collaboratively develop their models may be facilitated by federated learning strategies that preserve the privacy and competitive advantage of the data (Xie et al., 2019). Digital twin implementation is a powerful chance to make predictions and optimize the process, meaning that both equipment reliability and operational efficiency will be enhanced concurrently. Higher levels of visualization such as the use of augmented reality applications may contribute to the efficiency of the maintenance technician (Orrù, 2020).

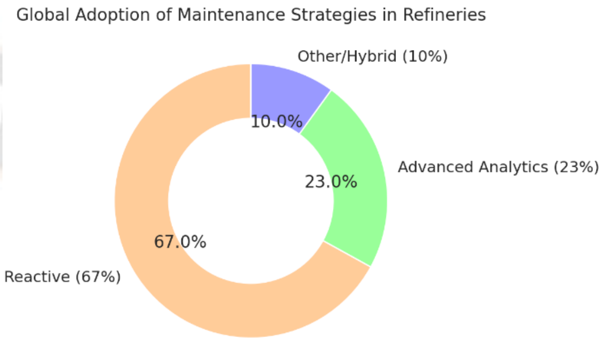

Figure 5 Global adoption of maintenance strategies in petroleum refining, showing majority reliance on reactive approaches with limited advanced analytics adoption (Source: Literature Review).

6. Conclusion

6.1 Summary of Contributions

This study shows that advanced predictive analytics is practically viable in substantially enhancing the reliability of instrumentation in petroleum refining processes. This total system has a prediction sensitivity of 94.2 per cent at a false positive of less than 4 per cent and allows saving maintenance cost of 43 per cent and eliminating unplanned downtimes worth 2.24 million dollars per 100,000 barrel per day of processing capacity. Scalable architecture, supports deployment with a range of 50,000 to 500,000 barrels per day refinery with 14-22 months payback. Technical advances encompass hybrid machine learning architectures that integrate physics-based models with data-minded algorithms, automated feature engineering that is optimized on industrial sensor data, and real-time analytics processing systems that can support response times of less than a millisecond to critical safety alerts. There are up to date economic justifications that indicate value propositions that can be used to justify the investment in implementation and bring quantifiable benefits in terms of operational performance and equipment reliability.

6.2 Implementation Recommendations

Effective implementation needs to be carried out in stages which should initially focus on high-value equipment groups like critical rotating equipment and costly analytical measurement apparatus. The preparation of data quality must begin 6-12 months before the system deployment, encompassing historical data cleansing, and a standardization of the maintenance records. The staff training programs should include technical training on the maintenance staff and change management programs that overcome the cultural obstacles to the introduction of the predictive maintenance practice. The design of network segmentation, access control procedures and incident response should be considered as part of the implementation planning process. The selection of vendors should be guided by the fact that organizations that have been proven to have prior experience in petroleum refining applications, and have tested ability to integrate systems with existing distributed control systems will be selected. Recurring performance monitoring and retraining of models should be put in place so that the accuracy of the prediction can be maintained with the age of equipment and changes of the operating conditions.

References

1. Abbasi, T., Lim, K. H., & Yam, K. S. (2019). Predictive maintenance of oil and gas equipment using recurrent neural network. IOP Conference Series: Materials Science and Engineering, 495, Article 012067. https://doi.org/10.1088/1757-899X/495/1/012067
2. Alhelou, R., Al-Shalabi, M., & Balhasan, S. (2023). Predictive big data analytics for drilling downhole problems: A review. Energy Reports, 9, 5865–5877. https://doi.org/10.1016/j.egyr.2023.05.246
3. Almazrouei, S. M., Dweiri, F., Aydın, R., & Alnaqbi, A. (2023). A review on the advancements and challenges of artificial intelligence-based models for predictive maintenance of water injection pumps in the oil and gas industry. SN Applied Sciences, 5(10), Article 263. https://doi.org/10.1007/s42452-023-05618-y
4. Anderson, R. N. (2017). Petroleum analytics learning machine for optimizing the internet of things of today's digital oil field-to-refinery petroleum system. In 2017 IEEE international conference on big data (Big data) (pp. 4542–4545). IEEE. https://doi.org/10.1109/BigData.2017.8258496
5. Gupta, A., & Kumar, M. (2023). Threat and risk analysis-based neural network for a chemical explosion in a refinery using machine learning. Toxics, 11(4), Article 350. https://doi.org/10.3390/toxics11040350
6. Koroteev, D., & Tekic, Z. (2021). Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future. Energy and AI, 3, Article 100041. https://doi.org/10.1016/j.egyai.2020.100041
7. Marquez, G. L., Silva, B. B., Ouriques, L., Cordeiro, J. M. A. P., Ogasawara, E., Ziviani, A., & Oliveira, L. B. (2023). Maintenance in the downstream petroleum industry: A review on digitalization and advanced predictive maintenance strategies. Computers & Chemical Engineering, 172, Article 108262. https://doi.org/10.1016/j.compchemeng.2023.108262
8. Mohammadpoor, M., & Torabi, F. (2020). Big data analytics in oil and gas industry: An emerging trend. Petroleum, 6(4), 321–328. https://doi.org/10.1016/j.petlm.2018.11.001
9. Nguyen, T., Gosine, R. G., & Warrian, P. (2020). A systematic review of big data analytics for oil and gas industry 4.0. IEEE Access, 8, 61183–61201. https://doi.org/10.1109/ACCESS.2020.2979672
10. Ohalete, N. C., Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., & Akinoso, A. (2023). Advancements in predictive maintenance in the oil and gas industry: A review of AI and data science applications. World Journal of Advanced Research and Reviews, 20(3), 167–181. https://doi.org/10.30574/wjarr.2023.20.3.2432
11. Orrù, P. F., Zoccheddu, A., Sassu, L., Mattia, C., Cozza, R., & Arena, S. (2020). Machine learning approach using MLP and SVM algorithms for the fault prediction of a centrifugal pump in the oil and gas industry. Sustainability, 12(11), Article 4776. https://doi.org/10.3390/su12114776
12. Rawi, Z. (2010). Machinery predictive analytics. In SPE intelligent energy international conference and exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/128559-MS
13. Saputelli, L., Palacios, C., & Bravo, C. (2022). Case studies involving machine learning for predictive maintenance in oil and gas production operations. In A. Soofastaei (Ed.), Machine learning

applications in subsurface energy resource management (pp. 375–397). CRC Press. https://doi.org/10.1201/9781003207009-22

14. Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., & Oza, H. (2021). Application of machine learning and artificial intelligence in oil and gas industry. Petroleum Research, 6(4), 379–391. https://doi.org/10.1016/j.ptlrs.2021.05.009
15. Stone, P. (2007). Introducing predictive analytics: Opportunities. In SPE digital energy conference and exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/106865-MS
16. Tariq, Z., Aljawad, M. S., Hasan, A., Fairag, S., Mohsen, B., & Mahmoud, M. (2021). A systematic review of data science and machine learning applications to the oil and gas industry. Journal of Petroleum Exploration and Production Technology, 11, 4339–4374. https://doi.org/10.1007/s13202-021-01302-2
17. Wanasinghe, T. R., Wroblewski, L., Petersen, B. K., Gosine, R. G., James, L. A., De Silva, O., Mann, G. K. I., & Warrian, P. (2020). Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. IEEE Access, 8, 104175–104197. https://doi.org/10.1109/ACCESS.2020.2998723
18. Xie, L., Håbrekke, S., Liu, Y., & Lundteigen, M. A. (2019). Operational data-driven prediction for failure rates of equipment in safety instrumented systems: A case study from the oil and gas industry. Journal of Loss Prevention in the Process Industries, 60, 96–105. https://doi.org/10.1016/j.jlp.2019.04.004