



PREDICTIVE MAINTENANCE FOR EQUIPMENT FAILURE ANALYSIS USING MACHINE LEARNING APPROACH

Umar, M.S.¹; Bello, A.T.²; Tukur, A.B.³

^{1,3}Department of Computer Science, Faculty of Computing, Sokoto State University, Sokoto, P.M.B. 2134, Sokoto, Nigeria

²Department of Computer Science, Umaru Ali Shinkafi Polytechnic, Sokoto, P.M.B. 2356, Sokoto, Nigeria

*Corresponding author email: mahdeeus@gmail.com

Received: 15-03-2026

Revised: 06-05-2026

Accepted: 16-05-2026

Published: 17-05-2026

Abstract: Predictive maintenance has emerged as an effective strategy for enhancing equipment reliability, minimizing unplanned downtime, and optimizing maintenance resources in modern industrial systems. Conventional maintenance approaches, including reactive and preventive maintenance, remain insufficient for anticipating failures in complex machinery such as hydraulic press equipment. This study proposes a data-driven predictive maintenance framework based on a Long Short-Term Memory (LSTM) neural network for equipment failure prediction using sensor-based operational data. A subset of 1,000 records comprising temperature, vibration, and usage-related features was extracted from a publicly available hydraulic system dataset and pre-processed to ensure suitability for deep learning modelling. The proposed LSTM model was trained and optimized using two optimization algorithms, Adam and RMSprop, to examine their impact on convergence behaviour and predictive performance. Model evaluation was conducted using five widely accepted classification metrics: accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). Experimental results demonstrate that the LSTM model optimized with Adam achieved superior overall performance, recording an accuracy of 97%, precision of 97%, recall of 97%, F1-score of 96%, and ROC-AUC of 99%. The RMSprop-optimized model exhibited comparable performance, achieving a perfect ROC-AUC of 100%. Comparative analysis against benchmark machine learning models reported in prior studies including Random Forest, Gradient Boosting, and Support Vector Machine shows that the proposed LSTM-based approach consistently outperforms conventional techniques. These findings demonstrate strong effectiveness of LSTM networks for predictive maintenance within the evaluated experimental setting in hydraulic press systems and highlight the role of optimizer selection in enhancing model performance.

Key words: Predictive maintenance; Machine learning; LSTM; Hydraulic press; Equipment failure

1 Introduction

The effective maintenance of industrial equipment is essential for ensuring operational continuity, workplace safety, and cost efficiency in manufacturing and production environments. Equipment failures often result in unplanned downtime, production losses, increased maintenance costs, and potential safety risks, all of which negatively affect organizational productivity and profitability (da Costa and Cavalcante, 2022; Zhao et al., 2021). As industrial systems become more complex and operate under demanding conditions, the limitations of conventional maintenance strategies become increasingly evident.

Traditional maintenance approaches, such as reactive maintenance and preventive maintenance, remain widely used across industries. Reactive maintenance addresses faults only after failures occur, frequently leading to unexpected downtime and costly repairs. Preventive maintenance, on the other hand, schedules servicing at predefined intervals regardless of the actual condition of the equipment, often resulting in unnecessary maintenance actions and inefficient use of resources (da Costa and Cavalcante, 2022). These approaches are increasingly inadequate for modern industrial systems that generate large volumes of

operational data and require condition-based decision making.

Predictive maintenance (PdM) has emerged as a more effective alternative by leveraging historical and real time operational data to anticipate equipment failures before they occur. By identifying early signs of degradation, predictive maintenance enables timely interventions that reduce downtime, extend equipment lifespan, and improve overall system reliability (Kane et al., 2022; Othman Alhuqayl et al., 2024). Advances in sensor technologies, Industrial Internet of Things (IIoT) platforms, and data analytics have further accelerated the adoption of PdM across diverse industrial sectors (Keleko et al., 2022). However, the success of predictive maintenance systems depends largely on the robustness of the predictive models and their ability to learn complex temporal patterns from equipment condition data.

Hydraulic press machines represent a critical class of industrial equipment widely used in metal forming, moulding, forging, and heavy manufacturing operations. These machines operate under high pressure and harsh mechanical conditions, making them susceptible to failures such as fluid leakage, excessive vibration, overheating, and pressure instability (Dawande et al., 2023; Prihananto and Sahroni, 2023). Failures in hydraulic press systems can cause severe production disruptions, increased operational costs, and safety hazards, underscoring the need for reliable and data driven predictive maintenance solutions tailored to their operational characteristics.

This study proposes a machine learning based predictive maintenance framework for hydraulic press equipment using a Long Short-Term Memory (LSTM) neural network. The research focuses on evaluating the impact of optimization algorithms, specifically Adam and RMSprop, on model convergence behaviour and predictive performance. The proposed framework is evaluated using five standard classification metrics: accuracy, precision, recall, F1 score, and ROC AUC, to provide a comprehensive assessment of failure prediction capability. In addition, the model is validated using both a hydraulic press condition monitoring dataset and a benchmark predictive maintenance dataset to assess robustness and generalizability. The key contributions of this study include systematic optimizer comparison, transparent experimental design, and comprehensive multi metric evaluation of an LSTM based predictive maintenance model for hydraulic press systems.

2 Literature Review

Predictive maintenance has received increasing attention in recent years as industries transition from reactive and preventive maintenance strategies toward

data driven decision making frameworks (da Costa and Cavalcante, 2022; Tsallis et al., 2025). This transition has been facilitated by advances in sensor technologies, Industrial Internet of Things (IIoT) infrastructures, and machine learning techniques capable of extracting actionable insights from large volumes of operational data (Keleko et al., 2022). Existing predictive maintenance research can broadly be categorized into conventional machine learning based approaches and deep learning-based approaches, each with distinct advantages and limitations (Tsallis et al., 2025).

Early predictive maintenance studies primarily employed traditional machine learning algorithms such as Random Forest (RF), Gradient Boosting Machine (GBM), Support Vector Machine (SVM), and decision tree classifiers. These models have been widely adopted due to their relative interpretability and effectiveness on structured, tabular datasets (Karwa et al., 2024). For instance, Karwa et al. (2024) investigated RF, GBM, and SVM models for equipment failure prediction using a publicly available predictive maintenance dataset, reporting competitive accuracy and precision values. However, their evaluation relied on a limited set of performance metrics and did not explore deep learning architectures capable of modelling temporal dependencies inherent in sequential sensor data.

Several survey and review studies have examined the broader landscape of predictive maintenance methodologies. Tsallis et al. (2025) provided an application-oriented review of machine learning based predictive maintenance techniques, highlighting the increasing adoption of hybrid, ensemble, and digital twin supported approaches across industrial domains. Similarly, Serradilla Oscar et al. (2020) and Hector and Panjanathan (2024) reviewed deep learning architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), emphasizing their superior capability in handling time series data compared with traditional machine learning models. While these studies offer valuable conceptual insights, most lack empirical validation on specific industrial equipment or detailed experimental comparisons.

With the growing availability of high-resolution time series sensor data, deep learning models particularly Long Short-Term Memory (LSTM) networks have emerged as promising tools for predictive maintenance applications (Serradilla Oscar et al., 2020). LSTM networks are specifically designed to capture long term temporal dependencies, making them suitable for monitoring equipment degradation patterns over time. Meddaoui et al. (2024) demonstrated the effectiveness of deep learning models for remaining useful life (RUL) prediction using real world aerospace

maintenance datasets. However, such studies are often domain specific and may not generalize effectively to other industrial systems, including hydraulic press equipment.

Despite their widespread industrial use, hydraulic press systems remain relatively underrepresented in predictive maintenance literature. Existing studies related to hydraulic systems tend to focus on fault diagnosis, vibration analysis, thermography, or hardware centric monitoring approaches rather than predictive failure modelling using advanced machine learning techniques (Dawande et al., 2023; Samuel, 2023). Moreover, many studies rely on simulated or synthetic datasets that fail to capture real world operational variability, sensor noise, and complex failure dynamics, limiting the practical applicability of their predictive models (Karwa et al., 2024).

Another notable limitation in existing predictive maintenance research is the lack of comprehensive model evaluation. Many studies emphasize accuracy and precision while neglecting critical metrics such as recall, F1 score, and ROC AUC, which are essential in maintenance decision making scenarios where false negatives can lead to severe equipment damage and operational losses (Zhao et al., 2021). Additionally, the impact of optimizer selection and hyperparameter tuning on deep learning model performance is often underexplored, despite their significant influence on training stability, convergence behaviour, and predictive reliability (Meddaoui et al., 2024).

Based on the reviewed literature, several research gaps are identified: (i) Overheating and Thermal Stress: Temperature induced degradation in hydraulic systems accelerates component wear and necessitates predictive models capable of early overheating detection instability (Prihananto, D. and Sahroni, T.R. 2023). (ii) Mechanical Vibration: Undetected excessive vibration accelerates mechanical degradation and increases the risk of unplanned hydraulic press failure instability (Prihananto, D. and Sahroni, T.R. 2023). (iii) Limited Performance Evaluation: Overreliance on accuracy and precision restricts comprehensive assessment of predictive maintenance models (Zhao et al., 2021). (iv) Synthetic Data Dependency: The use of simulated datasets limits model generalizability to real world industrial conditions (Karwa et al., 2024).

These identified gaps motivate the present study, which proposes an LSTM based predictive maintenance framework for hydraulic press equipment with systematic optimizer comparison and comprehensive multi metric evaluation. By focusing on real sensor-based datasets, transparent experimental design, and optimizer level analysis, this study aims to advance the

reliability and practical applicability of predictive maintenance models for complex industrial equipment. A hydraulic press is a mechanical device that applies compressive force using hydraulic pressure, operating based on Pascal's principle (Dawande, H. et al. 2023). The main components of a typical hydraulic press system are shown in Figure 1.

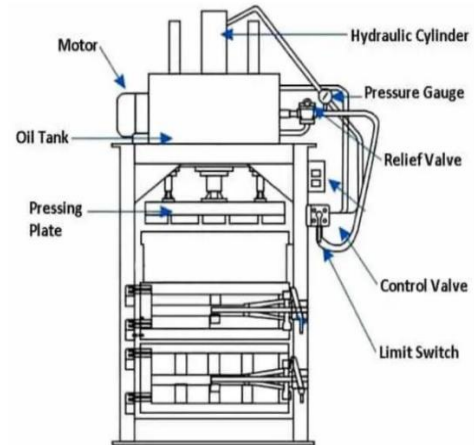


Figure 1: Schematic Diagram Of A Single-Acting Vertical Hydraulic Press

3 Methodology

This section describes the research design, data sources, model architecture, experimental setup, and evaluation strategy employed to develop and assess the proposed predictive maintenance model for hydraulic press equipment. The methodology is designed to ensure reproducibility, transparency, and alignment with best practices in machine learning based predictive maintenance research (Han et al., 2012; Kane et al., 2022).

3.1 Research Design Overview

The study adopts a supervised machine learning approach to predict equipment failure using time dependent sensor data collected from hydraulic systems. A Long Short-Term Memory (LSTM) neural network is employed due to its proven effectiveness in modeling sequential data and capturing temporal dependencies, which are critical in predictive maintenance applications involving time series signals (Serradilla Oscar et al., 2020; Meddaoui et al., 2024).

To evaluate the impact of optimization strategies on model performance, two widely used optimizers, namely Adam and RMSprop, are applied and compared under identical experimental conditions. This optimizer-based comparison is motivated by existing literature that highlights the influence of optimization techniques on training stability and convergence behavior in deep learning models (Meddaoui et al., 2024).

The overall research workflow consists of data acquisition, preprocessing, model development, experimental evaluation, and comparative analysis with benchmark models reported in prior literature. The

system architecture of the proposed framework is illustrated in Figure 2, which outlines the data preprocessing pipeline, LSTM training process, and performance evaluation stage.

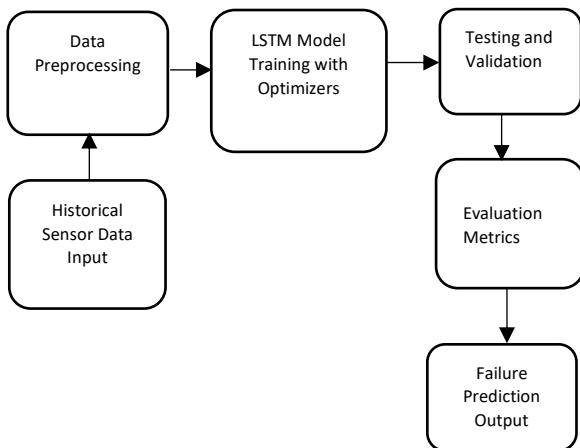


Figure 2: System Architecture of LSTM Based Predictive Maintenance Framework

3.2 Dataset Description and Data Collection

The dataset used in this study was obtained from a publicly available hydraulic system condition monitoring repository, consisting of sensor based operational measurements collected from hydraulic equipment operating under varying working conditions. The dataset includes multiple sensor channels related to temperature variation, vibration behavior, pressure dynamics, and operational usage patterns, which are commonly used in predictive maintenance studies (Karwa et al., 2024; Nsor, 2025).

To ensure computational efficiency while preserving data representativeness, a subset of 1,000 records with 3 feature variables was selected for model training and evaluation. This sampling strategy is consistent with prior machine learning based predictive maintenance studies that balance dataset size with training efficiency and performance stability (Karwa et al., 2024). The selected data were labeled to indicate normal and faulty operating conditions, enabling supervised learning.

Although the original dataset contains a larger number of records, a subset of 1,000 samples was deliberately selected to balance computational efficiency and controlled experimentation. High-performing deep learning systems can be developed using small datasets when appropriate architectures, regularization, and transfer/meta-learning strategies are employed (Rather et al., 2024). The dataset used in this study was obtained from the UCI Machine Learning Repository – Condition Monitoring of Hydraulic Systems, which contains 2,205 operational cycles collected from an

experimental hydraulic test rig operating under diverse load and fault conditions. Each cycle corresponds to a 60-second operational sequence and is annotated with the health state of key hydraulic components, including the cooler, valve, pump, and accumulator, under multiple fault severity levels. The original dataset comprises high-resolution multi-sensor time-series measurements, including temperature (1 Hz), vibration (1 Hz), pressure (100 Hz), and flow signals, recorded over repeated operational cycles.

In this study, a stratified random subset of 1,000 records was derived from the original dataset to preserve the class distribution between normal and faulty operating conditions. The subset selection was performed after temporal aggregation, where average temperature values per hour were computed to reduce high-frequency noise and align the data with maintenance-oriented decision timescales. This aggregation step enables stable LSTM training while retaining the underlying degradation patterns relevant for predictive maintenance. The subset was selected to preserve class distribution and approximate the operating conditions of the original dataset (Schneider et al., 2018).

Similar subset-based evaluations have been adopted in prior predictive maintenance studies to ensure stable training while facilitating detailed optimizer comparison (Karwa et al., 2024). To mitigate overfitting risks associated with small sample sizes, stratified sampling was applied, and model performance was monitored using validation loss and ROC AUC stability across epochs. The consistent convergence observed in the training and validation curves indicates that the selected subset is statistically representative for the evaluated task.

3.3 Data Preprocessing

Data preprocessing was performed to improve data quality and suitability for LSTM based modeling. The preprocessing steps included normalization of feature values to ensure uniform scaling across sensors, and label encoding of categorical attributes where applicable. Normalization is particularly important in neural network training to prevent features with large numeric ranges from dominating the learning process.

3.4 Model Architecture and Development

The predictive maintenance model is based on a Long Short Term Memory neural network, chosen for its ability to retain and learn from long term dependencies in time series data. LSTM networks address the vanishing gradient problem associated with conventional recurrent neural networks, making them well suited for long horizon temporal prediction tasks (Serradilla Oscar et al., 2020).

The LSTM architecture consists of an input layer, one or more hidden LSTM layers, and a dense output layer responsible for binary classification of equipment condition (failure or non-failure). To examine the

influence of optimization strategies on training stability and convergence behaviour, two optimizers were independently evaluated:

- Adam optimizer, known for adaptive learning rates and fast convergence;
- RMSprop optimizer, commonly used for recurrent neural networks due to its ability to handle non stationary objectives.

All other hyperparameters were kept constant to ensure a fair and controlled comparison between optimization techniques (Meddaoui et al., 2024).

3.5 Experimental Setup and Design

A dedicated experimental setup was defined to ensure clarity, reproducibility, and meaningful interpretation of results. The experimental parameters used for training and evaluation are summarized in Table 1.

Table 1: Experimental Setup And Design Table

Parameter	Experimental Setup
Sample Size	1,000 x 6 records
Epochs	50
Batch Size	32
Optimizer	Adam and RMSprop
Learning Rate	1e-4
Train/validation/test splits	80%, 20%
Training Time	200 to 300 sec

The LSTM model was trained for 50 epochs with a batch size of 32 and a learning rate of 1×10^{-4} , values selected to balance convergence speed and model generalization. studies to balance convergence speed and model generalization (Serradilla Oscar et al., 2020). Training time ranged between 200 and 300 seconds, depending on optimizer behavior and convergence speed.

The same dataset, train-test split, and model architecture were maintained across experiments to isolate the effect of optimizer selection. All experiments were implemented using Python and deep learning libraries within a cloud-based environment. Model training and evaluation were performed using Google Collab with GPU acceleration to ensure efficient computation and scalability, consistent with modern machine learning workflows (Kane et al., 2022).

3.6 Evaluation Metric

To provide a comprehensive and reliable assessment of the predictive maintenance model, five standard classification performance metrics were employed:

accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC–AUC). These metrics are widely used in predictive maintenance studies to evaluate overall classification performance, failure detection capability, and class separability (Zhao et al., 2021; Othman Alhuqayl et al., 2024).

Accuracy represents the proportion of correctly classified instances over the total number of predictions and is defined as:

$$Accuracy = \frac{TP+TN}{TP + TN + FP + FN} \quad (1)$$

Equation 3.1: Equation of Accuracy

Precision measures the reliability of failure predictions by quantifying the proportion of correctly predicted failure cases among all predicted failures:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Equation 3.2: Equation of Precision

Recall, also referred to as sensitivity, evaluates the model's ability to correctly detect actual failure cases and is expressed as:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Equation 3.3: Equation of Recall

The F1 score provides a balanced evaluation by computing the harmonic mean of precision and recall, particularly useful when class imbalance exists:

$$F1 - Score = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (4)$$

Equation 3.4: Equation of F1 Score

The ROC–AUC metric assesses the model's discriminative capability across varying classification thresholds. It is derived from the relationship between the true positive rate (TPR) and false positive rate (FPR), defined respectively as:

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

Equation 3.5a: Equation of ROC-AUC

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

Equation 3.5b: Equation of ROC-AUC

A higher ROC–AUC value indicates stronger class separability and improved robustness of the predictive model in distinguishing between normal and faulty operating conditions. These metrics collectively provide a comprehensive evaluation framework suitable for maintenance decision-making scenarios where failure misclassification can result in significant operational and safety consequences.

These metrics were selected to address limitations observed in prior predictive maintenance studies that

relied primarily on accuracy and precision, which may mask poor failure detection performance (Othman Alhuqayl et al., 2024).

3.7 Comparative Benchmarking Strategy

To evaluate both the optimizer behaviour and the generalizability of the proposed LSTM-based predictive maintenance framework, experiments were conducted in two complementary phases using different datasets.

In the first phase, the LSTM model optimized with Adam and RMSprop was trained and evaluated on the Condition Monitoring of Hydraulic Systems dataset obtained from the UCI Machine Learning Repository. This dataset represents real hydraulic press operating conditions and serves as the primary dataset for analysing optimizer performance and model effectiveness in a hydraulic press context.

In the second phase, the LSTM model optimized with Adam was trained and evaluated on a benchmark predictive maintenance dataset (Bansal, 2020), namely the Machine Predictive Maintenance Classification dataset hosted on Kaggle by Bansal, S. (2020). *Machine predictive maintenance classification dataset*. This benchmark dataset has been widely used for evaluating classical machine learning models such as Random Forest, Gradient Boosting Machine, and Support Vector Machine. The performance of the LSTM model obtained in this study was compared with the reported RF, GBM, and SVM results on the same benchmark dataset, enabling a fair comparison across learning paradigms.

This two-stage evaluation strategy allows the study to assess optimizer effects on a domain-specific hydraulic dataset while also demonstrating the robustness and comparative advantage of the proposed LSTM approach on a widely used benchmark dataset.

4 Results

This section presents the experimental results obtained from training and evaluating the proposed LSTM based predictive maintenance models. The results are organized to first report quantitative performance outcomes, followed by visual analysis of benchmark comparisons, optimizer performance, and training behaviour. Detailed interpretation of these results is provided in the subsequent Discussion section.

4.1 Performance of Benchmark Machine Learning Models

To establish a baseline for comparison, the reported performance of conventional machine learning models Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Machine (SVM) from the benchmark study is first presented. These models were evaluated on an equipment failure prediction task using accuracy and precision metrics, as reported by Karwa et al. (2024).

Figure 3 shows the visualization of the benchmark models' accuracy and precision scores. Random Forest

achieved the highest accuracy among the benchmark models, while Gradient Boosting slightly outperformed RF in terms of precision. The SVM model recorded the lowest scores across both metrics. These results are consistent with previously reported findings where ensemble-based models demonstrate improved performance over kernel-based methods in equipment failure classification tasks (Karwa et al., 2024).

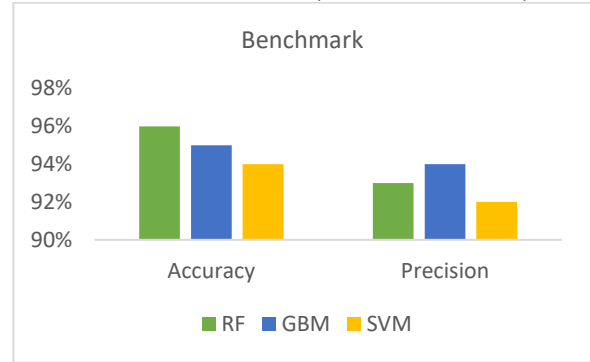


Figure 3: Performance Comparison Of Benchmark Machine Learning Models (RF, GBM, SVM) For Equipment Failure Prediction Using Accuracy And Precision Metrics.

4.2 Performance of Proposal LSTM Models

The predictive performance of the proposed LSTM model was evaluated using two optimization algorithms: Adam and RMSprop. Both models were assessed across five performance metrics accuracy, precision, recall, F1-score, and ROC-AUC to provide a comprehensive evaluation, addressing limitations of prior studies that relied on fewer metrics (Othman Alhuqayl et al., 2024).

Figure 4 presents a comparison of the LSTM models optimized with Adam and RMSprop across all evaluation metrics. The results indicate that both optimizers achieved very similar overall performance. The LSTM+Adam configuration demonstrated slightly higher precision, while the LSTM+RMSprop achieved a perfect ROC-AUC score, indicating excellent class separability between faulty and non-faulty states (Zhao et al., 2021).

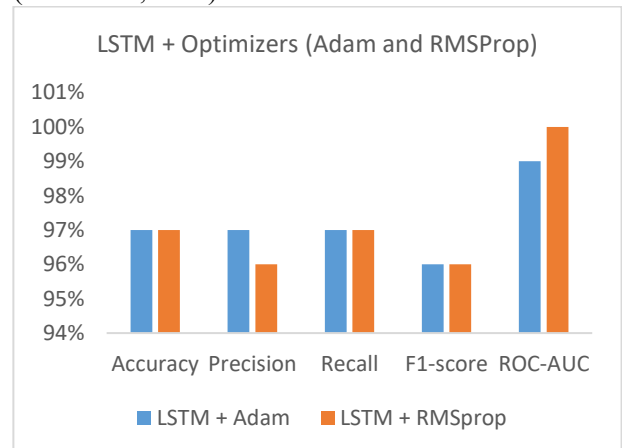


Figure 4: Comparative Performance Of LSTM Models Optimized With Adam And RMSprop Across Accuracy, Precision, Recall, F1-Score, And ROC-AUC Metrics For Hydraulic Press Failure Prediction.

4.3 Training Accuracy Analysis

The convergence behaviour of the LSTM models during training was analysed through accuracy curves generated over 50 training epochs. This analysis provides insight into the learning stability and convergence speed associated with each optimization algorithm.

As illustrated in Figure 5 and Figure 6, both LSTM models exhibited steady improvements in training accuracy as the number of epochs increased. Significant gains were observed within the first 40 epochs, after which the accuracy stabilized. By the 50th epoch, both configurations achieved a final accuracy of 97%, indicating effective learning of temporal patterns present in the hydraulic press sensor data. Such convergence behaviour aligns with previous deep learning-based predictive maintenance studies employing LSTM networks.

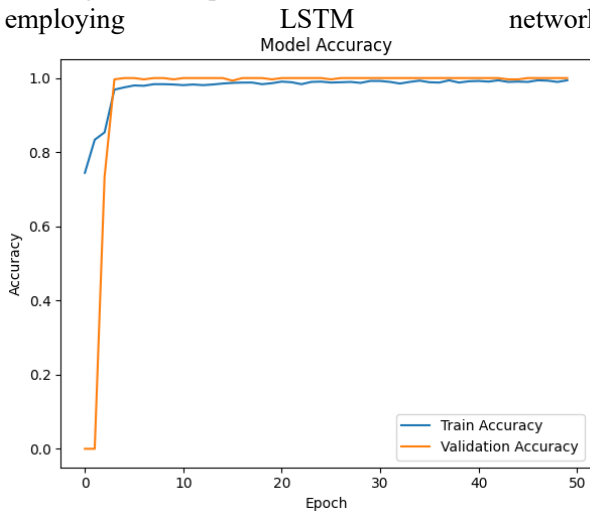


Figure 5: Training Accuracy Curve Of The LSTM Model Optimized With Adam Over 50 Epochs.

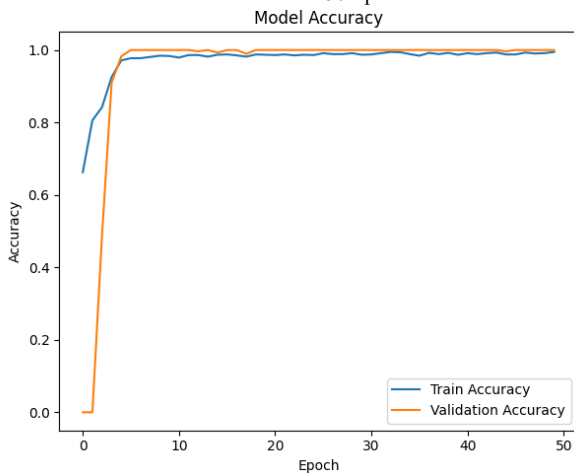


Figure 6: Training Accuracy Curve Of The LSTM Model Optimized With RMSProp Over 50 Epochs.

4.4 Training Loss Analysis

To further assess model learning behaviour, training and validation loss curves were examined. Loss metrics provide insight into how effectively the models minimize prediction errors and whether overfitting occurs during training.

Figure 7 and Figure 8 show the loss curves for the Adam- and RMSprop-optimized LSTM models, respectively. In both cases, loss values decreased consistently as training progressed, with sharp reductions observed during the early epochs. The gradual flattening of the loss curves in later epochs indicates convergence and stable learning. The close alignment between training and validation loss suggests that neither model suffered from significant overfitting, consistent with best practices reported in deep learning maintenance literature.

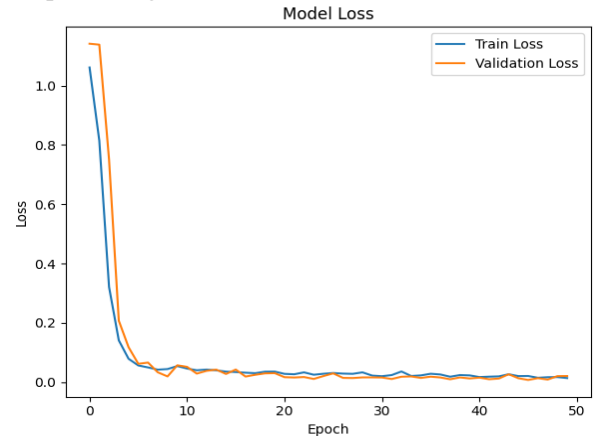


Figure 7: Training And Validation Loss Curves For The LSTM Model Using The Adam Optimizer.

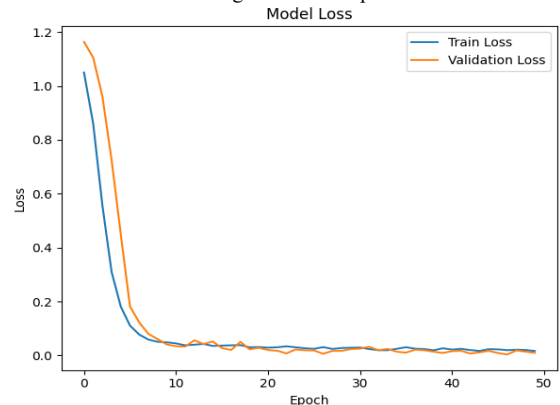


Figure 8: Training And Validation Loss Curves For The LSTM Model Using The RMSProp Optimizer.

4.5 Model Performance on Benchmark Dataset

To further validate model generalizability, the LSTM model optimized with Adam was evaluated using benchmark data reported in prior literature. The training accuracy and loss behaviour under these conditions were examined to assess robustness across datasets.

As shown in Figure 9, the training and validation accuracy curves exhibit consistent improvement and convergence, indicating good generalization capability. Similarly, Figure 10 demonstrates a steady reduction in loss values with minimal divergence between training and validation loss. This behaviour confirms stable model performance and reflects the robustness of LSTM networks in handling time-series data across varying datasets.

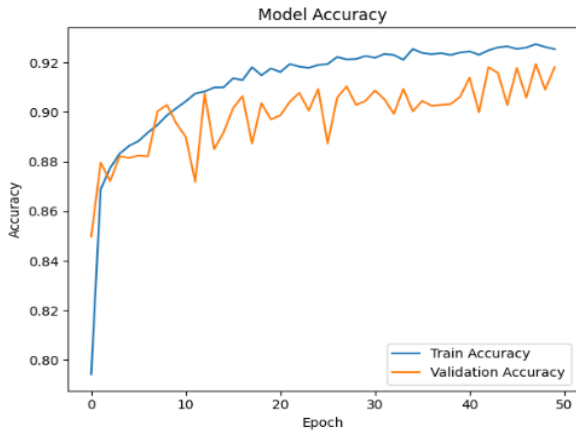


Figure 9: Training Accuracy Curve Of The LSTM + Adam Model Evaluated On Benchmark Data.

Metric	LSTM		Benchmark			LSTM
Technique and Optimizer	LST M + Ada m	LST M + RMS prop	RF	GB M	SV M	LSTM + Adam Optimize r + Benchmark data
Accuracy	97 %	97%	96 %	95 %	94 %	94%
Precision	97 %	96%	93 %	94 %	92 %	97%
Recall	97 %	97%	-	-	-	94%
F1-score	96 %	96%	-	-	-	95%
ROC-AUC	99 %	100 %	-	-	-	-

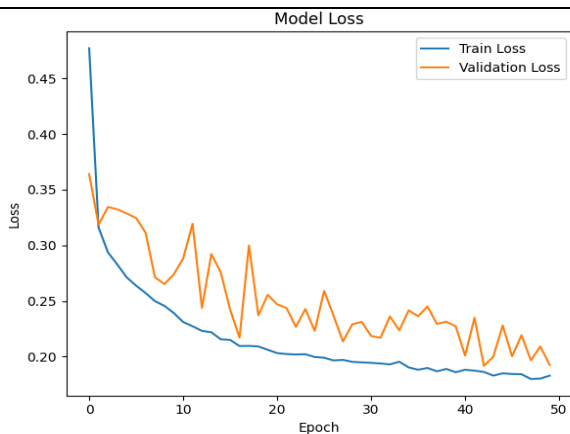


Figure 10: Training Loss Curve Of The LSTM + Adam Model Evaluated On Benchmark Data.

4.6 Overall Performance Summary

A summary of the quantitative performance metrics obtained from the proposed LSTM models is presented in Table 3. The table consolidates results across all evaluation metrics for both optimization strategies.

Table 3: Performance Comparison Of LSTM Models On The Hydraulic Press Dataset And Comparison With Conventional Machine Learning Classifiers (RF, GBM, SVM) Evaluated On The Same Benchmark Predictive Maintenance Dataset.

The first set of results focuses on the performance of the proposed LSTM model trained using Adam and RMSprop optimizers. Both configurations demonstrated strong and consistent predictive capability across all evaluated metrics. The LSTM with Adam achieved an accuracy of 97%, matching the accuracy obtained using RMSprop. Precision was slightly higher for Adam at 97%, compared to 96% for RMSprop, indicating better control of false positive predictions. Recall values were identical for both optimizers at 97%, showing that both models were equally effective in identifying true fault conditions. The F1-score further confirmed this balance, with both optimizers achieving 96%. In terms of separability, RMSprop recorded a ROC–AUC of 100%, marginally outperforming Adam, which achieved 99%, indicating superior discrimination capability across classification thresholds.

The second part of the analysis compares the LSTM models with benchmark machine learning techniques, namely Random Forest (RF), Gradient Boosting Machine (GBM), and Support Vector Machine (SVM). While these benchmark models showed reasonably strong performance, their results were consistently below those achieved by the LSTM-based approaches. RF reached an accuracy of 96%, followed by GBM at 95%, and SVM at 94%. Precision values for the benchmark models ranged from 92% to 94%, which are notably lower than those obtained by the LSTM models. Recall, F1-score, and ROC–AUC values were not available for the benchmark techniques, limiting their overall comparative evaluation. These results indicate that although traditional machine learning models are effective, they do not fully capture the temporal dependencies inherent in the data when compared to LSTM-based models.

The final part of the evaluation examines the LSTM model trained with the Adam optimizer using benchmark data. In this configuration, the model achieved an accuracy of 94%, a precision of 97%, recall of 94%, and an F1-score of 95%. While the accuracy is slightly lower than the proposed LSTM models trained on the primary dataset, the high precision demonstrates reliable fault prediction capability, suggesting the potential generalizability of the LSTM–Adam framework across datasets with similar characteristics when applied to external benchmark data.

Overall, the results confirm that LSTM models optimized with Adam and RMSprop offer superior

predictive performance compared to conventional machine learning benchmarks, with optimizer selection influencing specific performance aspects such as precision and class separability. Additionally, the consistent performance observed when using benchmark data further validates the robustness and applicability of the proposed LSTM-based predictive maintenance approach.

5 Discussion

This section interprets the experimental results presented in Section 4 and discusses their implications for predictive maintenance of hydraulic press equipment. The findings are analysed in relation to model performance, optimizer behaviour, benchmarking outcomes, and their practical relevance to industrial maintenance applications.

The experimental results demonstrate that the proposed LSTM based predictive maintenance model achieves consistently high predictive performance across all evaluated metrics. Both LSTM configurations optimized using Adam and RMSprop recorded an accuracy of 97%, confirming the model's ability to correctly classify equipment failure conditions using sensor-based time series data. These results indicate that deep learning architectures are well suited to capturing the complex temporal patterns characteristic of hydraulic press operations, which are often difficult to represent using conventional machine learning methods (Meddaoui et al., 2024).

A key observation from the results is the superiority of the LSTM based models over conventional machine learning benchmarks reported in prior studies. As shown in Section 4, the Random Forest, Gradient Boosting, and Support Vector Machine models reported in the benchmark study achieved lower accuracy and precision, and did not provide performance values for recall, F1 score, or ROC AUC (Karwa et al., 2024). This limitation reduces their effectiveness in predictive maintenance scenarios, where missed failure detections can result in severe operational consequences. The higher recall values achieved by the LSTM models in this study indicate a stronger ability to detect impending failures, thereby reducing the likelihood of undetected faults that could lead to costly downtime and equipment damage (Zhao et al., 2021).

The inclusion of multiple evaluation metrics provides a more balanced and practical assessment of predictive performance than accuracy alone. Predictive maintenance systems must not only classify equipment states correctly but also minimize false negatives, which can undermine operational safety and reliability. The strong F1 score values obtained in this study indicate that the proposed model maintains an effective trade-off between precision and recall. Furthermore, the high ROC AUC values, particularly the perfect score achieved by the RMSprop optimized LSTM demonstrate excellent class separability across varying

decision thresholds. This characteristic is particularly valuable in real world maintenance applications where alert thresholds may need to be adjusted dynamically based on operational priorities and risk tolerance (Othman Alhuqayl et al., 2024; Zhao et al., 2021).

The optimizer comparison reveals meaningful insights into model training behaviour and convergence characteristics. While both Adam and RMSprop produced comparable predictive performance, subtle differences were observed. Adam demonstrated slightly better precision and more stable convergence behaviour, suggesting smoother optimization of weight updates. RMSprop, on the other hand, achieved marginally faster convergence and superior ROC AUC performance, indicating improved sensitivity to class discrimination. These findings suggest that optimizer selection can influence not only training efficiency but also specific aspects of predictive performance, supporting observations reported in prior deep learning based predictive maintenance studies (Meddaoui et al., 2024).

The training accuracy and loss analyses further confirm the stability and generalization capability of the proposed model. The smooth convergence patterns observed in both accuracy and loss curves indicate effective learning without excessive overfitting. The close alignment between training and validation performance supports the reliability of the model for deployment in operational environments. Additionally, the strong performance observed when the model was evaluated on benchmark data highlights its potential to generalize beyond a single dataset, an important requirement for industrial predictive maintenance systems (Serradilla Oscar et al., 2020).

From a practical perspective, the results reinforce the suitability of LSTM based models for predictive maintenance of hydraulic press systems. The ability to anticipate failures using sequential sensor data enables maintenance teams to transition from schedule-based servicing to condition-based decision making. This shift can reduce unplanned downtime, improve maintenance planning, and extend equipment lifespan, which are key objectives in modern industrial maintenance strategies (da Costa and Cavalcante, 2022; Othman Alhuqayl et al., 2024). By explicitly reporting experimental parameters, training behaviour, and comprehensive evaluation metrics, this study also addresses reproducibility concerns commonly highlighted in predictive maintenance literature (Tsallis et al., 2025).

Overall, the findings confirm that deep learning driven predictive maintenance, when supported by appropriate optimization strategies and comprehensive evaluation, offers significant advantages over traditional machine learning approaches. These advantages are particularly relevant for complex industrial equipment such as hydraulic presses, where temporal dependencies, operational variability, and failure dynamics play a

critical role in maintenance decision making (Prihananto, D. and Sahroni, T.R. 2023).

6 Contribution

This study makes several important contributions to the field of predictive maintenance and industrial equipment monitoring. First, it presents a structured and reproducible LSTM based predictive maintenance framework tailored specifically for hydraulic press equipment, a class of industrial machinery that has received limited attention in existing deep learning driven predictive maintenance studies. While hydraulic systems are widely used in manufacturing and heavy industries, most prior research has focused on fault diagnosis or hardware centric monitoring rather than data driven failure prediction using advanced learning models (Prihananto, D. and Sahroni, T.R. 2023; Samuel, 2023). By addressing hydraulic press equipment explicitly, this research extends the applicability of sequential deep learning models to an under explored but industrially critical domain.

Second, the study provides a systematic comparison of optimization strategies by evaluating the impact of Adam and RMSprop optimizers on LSTM model performance under identical experimental conditions. This comparative analysis offers practical insight into optimizer selection, highlighting differences in convergence behaviour, precision stability, and classification capability. Existing predictive maintenance studies often omit optimizer level analysis despite its strong influence on training dynamics and performance stability in deep learning models (Meddaoui et al., 2024). By isolating optimizer effects, this work contributes actionable guidance for practitioners deploying LSTM based predictive maintenance systems.

Third, this research advances predictive maintenance evaluation practices by employing a comprehensive set of performance metrics, including accuracy, precision, recall, F1 score, and ROC AUC. Unlike many existing studies that rely primarily on accuracy and precision, the inclusion of recall and ROC AUC improves the relevance of the findings for real world maintenance decision making, where missed failure detections can have severe operational consequences (Zhao et al., 2021; Othman Alhuqayl et al., 2024). This multi metric evaluation provides a more balanced assessment of model reliability and failure sensitivity. In addition, the study strengthens reproducibility and transparency by explicitly documenting experimental parameters such as training epochs, batch size, learning rate, optimizer configuration, training duration, and evaluation procedures. The lack of detailed experimental reporting is a commonly identified limitation in predictive maintenance literature, hindering replication and comparative validation across studies (Tsallis et al., 2025). By addressing this gap, the proposed framework supports future replication and methodological consistency.

Finally, through benchmarking against conventional machine learning models reported in prior work, this study demonstrates the performance advantages of LSTM based predictive maintenance models in capturing temporal dependencies within sensor data. The results confirm that deep learning-based approaches, when carefully configured and comprehensively evaluated, offer superior predictive capability and practical value for industrial maintenance applications compared with traditional machine learning techniques (Karwa et al., 2024; Serradilla Oscar et al., 2020.).

7 Conclusion

This study investigated the application of a machine learning-based predictive maintenance framework for hydraulic press equipment using a Long Short-Term Memory (LSTM) neural network. By leveraging sensor-based operational data, the proposed model demonstrated strong capability in predicting equipment failure conditions, thereby addressing the limitations associated with traditional reactive and preventive maintenance strategies that often fail to anticipate equipment degradation effectively (da Costa and Cavalcante, 2022; Zhao et al., 2021). The findings confirm that deep learning approaches are well suited for modelling the complex temporal dynamics inherent in industrial maintenance data, particularly for systems operating under variable and high-stress conditions. Experimental results showed that the proposed LSTM model achieved consistently high performance across multiple evaluation metrics. The Adam-optimized configuration produced superior convergence stability and slightly higher precision, while the RMSprop-optimized model exhibited excellent class discrimination, achieving a perfect ROC-AUC score. Both configurations outperformed benchmark machine learning models reported in prior studies, including Random Forest, Gradient Boosting, and Support Vector Machines, highlighting the effectiveness of LSTM architectures in capturing time-dependent patterns that conventional models often fail to represent adequately (Karwa et al., 2024; Serradilla Oscar et al., 2020).

Beyond predictive accuracy, this research contributes to practical maintenance decision-making by emphasizing comprehensive model evaluation using accuracy, precision, recall, F1-score, and ROC-AUC. This multi-metric assessment provides a more realistic understanding of model performance in real-world scenarios, where failure detection sensitivity and risk mitigation are critical to reducing unplanned downtime and preventing catastrophic equipment failure (Zhao et al., 2021; Othman Alhuqayl et al., 2024). The explicit reporting of experimental parameters, training behaviour, and evaluation procedures further

strengthens the transparency and reproducibility of the proposed approach, addressing limitations frequently identified in predictive maintenance literature (Tsallis et al., 2025).

From an industrial perspective, the results demonstrate that LSTM-based predictive maintenance systems can support earlier fault detection, reduce unplanned downtime, and optimize maintenance scheduling for hydraulic press equipment. The insights gained from optimizer comparison also provide practical guidance for deploying deep learning models in maintenance operations where computational efficiency, training stability, and predictive reliability are important considerations (Meddaoui et al., 2024).

Future research can extend this work in several directions. Incorporating additional operational and environmental variables, such as ambient temperature, load variations, and human operational inputs, may further improve prediction accuracy and robustness (Othman Alhuqayl et al., 2024).

Cost-based optimization strategies could be integrated to balance maintenance decisions against financial impact, enabling more economically informed maintenance planning (da Costa and Cavalcante, 2022). In addition, evaluating the proposed framework on larger real-time datasets and deploying it within live industrial environments would further strengthen its practical applicability and support the advancement of intelligent predictive maintenance systems for complex industrial equipment.

8 References

- Abbas, A. A. (2024). AI for predictive maintenance in industrial systems. *International Journal of Advanced Engineering Technologies and Innovations* Vol. 1 No. 1: pp. 1–10.
- Alshahrani, N. (2024). Machine learning approaches for predictive maintenance in industrial operations, In: *Proceedings of the International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 1–6.
- Coandă, P., Avram, M. and Constantin, V. (2020). A state-of-the-art review of predictive maintenance techniques. *IOP Conference Series: Materials Science and Engineering* Vol. 997 No. 1: Article 012039.
- da Costa, L.Q.M. and Cavalcante, C.A.V. (2022). A review on the study of maintenance effectiveness. *Pesquisa Operacional* Vol. 42 Special Issue 1: Article e263613.
- Dawande, H., et al. (2023). A review paper on hydraulic press machine. *International Journal of Research and Technology Innovations (IJRTI)*.
- Han, J., Kamber, M. and Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann (Elsevier), Waltham, MA, USA.
- Hector, I. and Panjanathan, R. (2024). Predictive maintenance in Industry 4.0: A survey of planning models and machine learning techniques. *PeerJ Computer Science* Vol. 10: Article e2016.
- Kane, A., Kore, A., Khandale, A., Nigade, S. and Joshi, P.P. (2022). Predictive maintenance using machine learning. *International Journal of Engineering Research and Technology* Vol. 11 No. 6: pp. 120–125.
- Karwa, R.R., Bamnote, G.R., Dhumale, Y.A., Deshmukh, P.P., Meshram, R.A. and Iqbal, S.M. (2024). Predictive maintenance: Machine learning approaches for enhanced equipment reliability, In: *Proceedings of the 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI 2024)*.
- Keleko, A.T., Kamsu Foguem, B., Ngouna, R.H. and Tongne, A. (2022). Artificial intelligence and real time predictive maintenance in Industry 4.0: A bibliometric analysis. *AI and Ethics* Vol. 2 No. 4: pp. 553–577.
- Meddaoui, A., Hachmoud, A. and Hain, M. (2024). Advanced machine learning for predictive maintenance: A case study on remaining useful life prediction and reliability enhancement. *The International Journal of Advanced Manufacturing Technology* Vol. 124: pp. 1–15.
- Nsor, M. (2025). Predictive maintenance using machine learning for engineering systems through real time sensor data and anomaly detection models. *International Journal of Research Publication and Reviews* Vol. 6 No. 7: pp. 5167–5183.
- Othman Alhuqayl, S., Alenazi, T., Alabduljabbar, H.A. and Haq, A. (2024). Improving predictive maintenance in industrial environments via IIoT and machine learning. *International Journal of Advanced Computer Science and Applications* Vol. 15 No. 4: pp. 423–432.
- Prihananto, D. and Sahroni, T.R. (2023). Effectiveness analysis of hydraulic torque wrench machine using failure mode and effect analysis. *Journal of Applied Engineering and Technological Science* Vol. 5 No. 1: pp. 474–494.
- Rathi, B., Thapaswi, S., Kambhampati, M., Jain, V., Akshay, P., Pandey, T.N. and Pradhan, S.K. (2025). Realizing the potential of Internet of

- Things in industrial applications. Discover Internet of Things Vol. 5 No. 1: Article 14.
- Rather, I.H., Kumar, S. and Gandomi, A.H. (2024). Breaking the data barrier: A review of deep learning techniques for democratizing AI with small datasets. *Artificial Intelligence Review*.
- Samuel, S.A. (2023). Holistic predictive maintenance framework leveraging AI, vibration analysis, and thermography for proactive asset reliability. *International Journal of Engineering Technology Research and Management* Vol. 7 No. 3: pp. 18–27.
- Serradilla Oscar, Zugasti, E. and Zurutuza, U. (2020). Deep learning models for predictive maintenance: A survey, comparison, challenges and prospects. *CEUR Workshop Proceedings* Vol. 2657: pp. 1–9.
- Tanveer Publications (2010). *Introduction to Hydraulic Presses and Press Body*. Tanveer Publications, New Delhi, India.
- Tsallis, C., Papageorgas, P., Piromalis, D. and Munteanu, R.A. (2025). Application wise review of machine learning based predictive maintenance: Trends, challenges and future directions. *Applied Sciences* Vol. 15 No. 9: Article 4898.
- Zhao, Z., Wu, J., Li, T., Sun, C., Yan, R. and Chen, X. (2021). Challenges and opportunities of AI enabled monitoring, diagnosis and prognosis: A review. *Chinese Journal of Mechanical Engineering* Vol. 34 No. 1: Article 43.
- Schneider, O., Helwig, N. and Schütze, A., 2018. Condition monitoring of hydraulic systems: A multi-sensor hydraulic test rig dataset. *UCI Machine Learning Repository*, Dataset 447.
- Bansal, S. (2020). *Machine predictive maintenance classification dataset*. Kaggle.