International Journal on Advanced Science Engineering Information Technology

Vol.15 (2025) No. 1
ISSN: 2088-5334

Evaluating Machine Learning and Deep Learning Algorithms for Predictive Maintenance of Hydraulic Systems

Ayat Al-Khulaqi [a], Naveen Palanichamy [a,*], Su Cheng Haw [a], S Charles Raja [b]

[a] *Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Selangor, Malaysia*
[b] *Department of Electrical and Electronics Engineering, Thiagarajar College of Engineering, Madurai, India*
*Corresponding author: *p.naveen@mmu.edu.my*

***Abstract*—Hydraulic systems are essential in industries like aerospace and petroleum. However, equipment degradation can lead to failures over time, resulting in costly downtime. Condition monitoring and predictive maintenance can be implemented to predict the equipment's failure before the machine's total failure. Current data-driven methods for predicting faults in hydraulic systems are insufficient due to inaccurate predictions. Our primary objective of this research is to investigate the potential of different classifier models' predictive capabilities in enhancing the reliability of hydraulic systems. This paper implements two machine learning models (ML), random forest (RF) and categorical boost (Catboost), and one deep learning model (DL), long-short-term memory (LSTM), to predict the maintenance needs of hydraulic system equipment using the data from ZeMA gGmbH. The results of the models are evaluated through different metrics, such as Precision, Recall, F1-Score, and Accuracy. The outcomes of the experiments validated the paramount importance of the RF model, which has proven to be the most efficient and successful in accurately predicting the instances of equipment failure before the occurrence of total system failure. Critical hydraulic system condition components revealed their varying performance across different components, with LSTM excelling in predictiveness of the valve, RF dominating pump predictions, and overall reliability observed for Accumulator and Stable Flag. The experimental findings demonstrate that the proposed method for predicting the state of hydraulic systems outperforms alternative approaches.**

***Keywords*— Hydraulic systems; condition monitoring; predictive maintenance; machine learning; deep learning.**

Manuscript received 9 Apr. 2024; revised 28 Sep. 2024; accepted 12 Dec. 2024. Date of publication 28 Feb. 2025.

I. INTRODUCTION

Hydraulic systems are used in numerous industries, including the oil and gas, air transport, construction, mobile vehicles, and factory equipment. Hence, many installations/mechanical components used in these applications need frequent service to continue operating flawlessly. As an innovative approach within industries, predictive maintenance (PdM), through its capability to detect potential equipment breakdown through sensors embedded in them, forms part of the Internet of Things (IoT). The conventional approach to machine maintenance costs a lot and mainly responds to breakdowns. Machine Learning (ML) and Deep Learning (DL) algorithms are used for predicting equipment failures, which have shown promising results in predicting them before they occur [1], [2]. This evolution includes increased operational efficiency, decreased downtime, and better practice in predictive maintenance [3].

Condition monitoring (CM) is the initial step or act that should be done to look for possible conditions and predict maintenance to stop machines from being down for a long time and reduce maintenance expenses [4]. It involves sampling process data, such as temperature, pressure, and so forth, through sampling equipment connected to a computer system [5]. The use of CM is an essential aspect of PdM. PdM entails continuous monitoring of machines for their health status and working conditions, which can lead to quality improvement of products, increased output rates, and overall factory performance [6], [7].

Hydraulic systems consist of five conditioning components: a cooler, which manages the fluid temperature and prevents overheating; the valve controls fluid flow and direction; the pump provides the necessary pressure to circulate fluid; the accumulator stores energy and compensates for fluid leakage; and the stable flag serves as an overall health indicator. Previous studies have adopted diverse methodologies to conduct hydraulic system fault analysis, including ML, DL, or combining both. Some researchers have concentrated on the failure of the entire hydraulic system ([8], [9], [10], [11], [12], [13], [14]), while others have focused on specific hydraulic system conditions such as valves, pumps, coolers, accumulators, and stable flags ([4],[15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]). Certain studies have aimed at diagnosing multiple faults and identifying different degrees of single faults ([26], [27]). A significant portion of the existing literature focuses on identifying the failure of the entire machinery or four specific condition components. However, studies investigating the failure of all five condition components are relatively limited. This paper aims to bridge this gap by incorporating all five condition components that predict the maintenance of hydraulic systems. Also, to explore the predictive maintenance of various ML and DL models to improve the reliability of hydraulic systems.

The paper begins by discussing the literature regarding the predictive maintenance of hydraulic systems. Hydraulic systems necessitate regular maintenance to ensure optimal performance. Implementing Predictive Maintenance (PdM) can significantly enhance the efficiency of these systems. By utilizing ML or DL algorithms, it is possible to predict the degradation state of the conditioning components. Table I shows the ML and DL algorithms discussed in the related works.

A. Machine Learning

This paper proposes a technique to improve the fault tolerance and accuracy of a hydraulic system by using logistic regression (LR), K-nearest Neighbor (KNN), decision tree (DT), RF, and naïve Bayes (NB) to predict faults. The proposed technique is implemented using the Spyder IDE software tool on a Raspberry Pi 3 Model B+ controller [8]. The authors proposed a new design of multi-layer stacking ensemble models to enhance the fault detection of manufacturing plants through the use of data from hydraulic systems by combining five ML algorithms and Linear Discriminant Analysis (LDA) together for better classification performance than the traditional stacking ensemble methods [28]. Proposed for hydraulic system failure prediction enhancement was a Time-based Imbalanced Data Synthesis Technique (TIDS) process and an XGboost classifier used to generate time domain features from data and to synthesize minority samples to address data imbalance [10].

The authors found faults in the pneumatic system. They predicted them with a hybrid semi-supervised learning model combined with traditional classification methods such as support vector machine (SVM), LR, DT, NB, and RF [12]. On the other hand, different research proposes a probability-based algorithm for analyzing the time-series data of hydraulic systems and evaluating multiple conditions using the Gaussian Mixture Model (GMM) for high accuracy [20]. The authors presented a method to diagnose various faults in the condition components of a hydraulic system based on principal component analysis and a multi-output, multi-class SVM for effective fault identification [22].

B. Deep Learning

One proposed data-driven approach focuses on deep neural networks (DNN) for multi-class classification degradation levels of each state of the hydraulic system [26]. The authors have described a strategy designed to diagnose many faults in hydraulic systems using time-series representations with FCN to acquire instantaneous features throughout multi-rate data [27]. The authors presented an advanced neural network model, Auto-NAHL, with automated hyperparameter tuning through Particle Swarm Optimization (PSO) for predicting maintenance in hydraulic systems [19]. Aa study focused on AI technique based on a Residual Network (ResNet-18) is presented for the high detection classification of faulted cooling circuitry in hydraulic systems [29]. Researchers proposed a Multirate Sensor Information Fusion Strategy (MRSIFS) for fault diagnosis under multitask conditions, which would address the condition of the hydraulic system. The method proposed in this study uses multidimensional convolutional blocks and integrates multisource information fusion into the architecture of CNN [25].

C. Machine Learning and Deep Learning

A study focusing on predictive maintenance was conducted to diagnose faults and predict the condition of the components in hydraulic systems using LR, RF, ANN, LightGBM, and Catboost [15]. A web application was also developed to demonstrate the exploratory data analysis of the system's condition. Researchers suggested a method for monitoring the health condition of hydraulic systems. Ensemble learning was used to improve the predictive precision of SVM classifiers, forming this method's essence. Besides, LDA and ANN were incorporated to compare the results against ensemble SVM [16]. A proposed method using machine and DL algorithms for fault detection and tolerance in each condition of the hydraulic systems. These algorithms include LR, KNN, DT, RF, and NB [4].

This includes the predictive study of industrial machine mechanical part conditions by using machine and deep learning algorithms, particularly LSTM and RF algorithms [17], [30]. In this research [31], the authors employed LSTM-based machine learning algorithms for business energy management aimed at the efficient use of energy in the case of electric vehicle (EV) charging, enhancing the planning and reducing the computation requirements compared with conventional methods. In another study, the authors presented an approach to monitoring the conditions of hydraulic systems using LDA, ANN, Linear SVM, and RBF SVM algorithms. The proposed technique applies multivariate statistics in sensor data analysis and fault detection. The data extracted from this is used in the training [18]. A proposed method for fault classification in hydraulic systems using a combination of Nearest Centroid (NC) with Dynamic time warping Barycenter Averaging (DBA) and RF algorithm to enhance the accuracy and speed of diagnosis [11]. A study on the predictive maintenance system on the innovative health assessment framework for hydraulic systems is undertaken using ensemble general multiclass support vector machines (EGMSVM) for stacking several GMSVMs as sub-models and one RF as a metamodel [21].

TABLE I
MACHINE AND DEEP LEARNING ALGORITHMS

	[4]	[8]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[21]	[23]	[24]	[25]	[26]	[27]	[28]
LR	/	/		/		/	/							/				/
KNN		/				/												
DT		/	/	/										/				
RF		/	/	/		/	/		/			/		/				
NB		/		/		/												
ANN					/		/	/		/								
XGBOOST														/				/
LIGHTGBM							/							/				
CATBOOST							/											
LDA					/	/		/		/		/		/				
Ensemble								/										
SVM																		
DNN																/		
NN	/																	
SVM	/			/	/	/												
DF	/																	
LSTM						/			/		/	/						
SVM (Linear)			/							/								
SVM (RBF)										/								
SoftMax			/															
FCN																	/	
Auto-NAHL											/							
1D - CNN											/	/						
DBN											/							
GMM																		
GMSVM												/						
E-SVM												/						
EGMSVMs												/						
Heterogeneous Stacking												/						
MO-RPELM													/					
MO-DSAE													/					
MO-RF													/					
MO-SVM													/					
MO-KNN													/					
CNN						/									/			
LSVM					/													
SVC														/				
Multi-layer Perceptron														/				
CART						/												

In this work, the authors presented algorithms such as LDA, RF, GMSVM, E-SVMs, LSTM, 1D-CNN, and heterogeneous stacking. This paper proposed a technique in hydraulic system fault diagnosis based on multiple output classification. They combined LDA and Hybrid Kernel Extreme Learning Machine to get high classification accuracy using MO-RPELM kernel. On the other hand, its authors compared performance in different models with MO-DSAE, MO-RF, MO-SVM, and MO-KNN, as shown in [21]. The present proposed hybrid artificial intelligence technique for predictive maintenance of the hydraulic system is composed of some algorithms like LSSVM, LDA SVM, and ANN, along with the combination of two additional feature selection techniques with ICEEMDAN-PCA or PCA without ICEEMDAN in combination with these algorithms [11]. A study was carried out on the anomaly detection system for the condition of hydraulic machinery signals, using eight algorithms: Multi-layer Perceptron, LDA, DT, RF, XGBoost, SVC, LightGBM and LR, along with three feature selection methods in the form of Pearson Correlation Coefficient, Spearman's Rank Correlation Coefficient, and the Boruta Algorithm [22]. It studies the fault detection and diagnosis framework for hydraulic machinery with nine algorithms: CNN, LR, CART, LDA, SVM, KNN, LSTM, RF and NB. The algorithms are implemented for four feature selections—Feature Importance (FI), RkSE, Time Domain Features, and Principal Component Analysis—without feature selection to compare different results [12].

To conclude, most authors implemented fault analysis on the degradation states of the hydraulic system condition components since monitoring the health is crucial. However, a comprehensive analysis of all five condition components is often overlooked. This study aims to fill this void, thereby maintaining energy efficiency and material savings and enhancing quality. In previous research papers, most implemented both ML and DL algorithms; in our research paper, the implementation of the ML model will be based on the most and least frequent ones: RF and Catboost. Catboost achieved high accuracy for all the hydraulic system conditions compared to other ML algorithms [12]. For DL, the most frequent will be implemented, where LSTM is chosen over ANN due to LSTM performing better in previous papers for predicting the degradation states of the hydraulic system conditions. For example, LSTM achieved high accuracy in predicting the conditions of the hydraulic system with results of cooler (100%), valve (95%), pump (99%), and accumulator (97%) [14]. In another research paper, LSTM achieved results for coolers (100%), valves (100%), pumps (100%), and accumulators (73%) [18]. Compared to ANN results of cooler (100%), valve (100%),15 pump (80%), and accumulator (50.4%) [15]. In another paper, ANN had better results due to the feature selection technique, which resulted in a cooler (100%), valve (97%), pump (94%), and accumulator (92%) [12].

This study presents a predictive maintenance framework that significantly advances from current methods. Most of the earlier studies concentrated on either one or a few condition components of the hydraulic systems, which may affect their predictive abilities. In contrast, our proposed approach consists of all five condition components of hydraulic systems, offering a comprehensive and integrated solution. Therefore, this paper attempts to develop and validate a model which predicts the degradation state of certain vital parts of a hydraulic system using ML and DL techniques. This approach bridges the gap in current predictive maintenance strategies and sets a new standard for hydraulic systems reliability improvement.

II. MATERIALS AND METHOD

The methodology proposed includes the ZeMA gGmbH dataset, data pre-processing, modelling, and evaluation of the hydraulic system's predictive maintenance; the methodology flow chart is shown in Fig. 1.

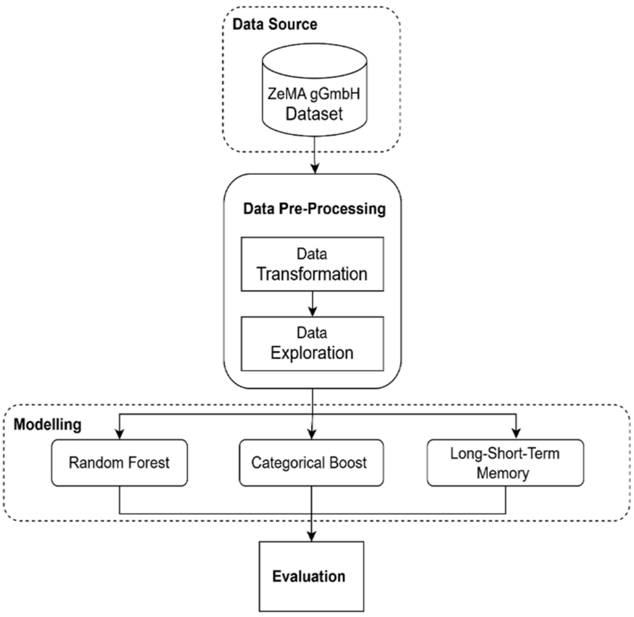

Fig. 1 Hydraulic System Methodology Chart

A. Data Source

The hydraulic system health condition data was sourced from a hydraulic test rig by the Centre for Mechatronics and Automation Technology (ZeMA), Saarbruecken, Germany [32]. The dataset encapsulates sensor data from the test rig, designed to simulate four distinct types of faults and system stability, each with varying severity levels. The test rig includes a primary fluid working circuit and an auxiliary cooling-filtration circuit connected by an oil tank. The working circuit, powered by main pump MP1 (with an electrical motor power of 3.3kw), undergoes cyclic repetition of different load levels regulated by the proportional pressure relief valve.

The collection of process measures like pressure, volume flow and temperature are carried out by the system every 60 seconds, while all the sensors are programmed to log process values in a cyclic manner. In the same experiment, five hydraulic components of the system's cooling circuit were systematically changed: cooler, valve, pump, accumulator, and stable flag. The dataset consists of 2205 instances which constitute a useful basis for analysis. It includes multiple feature types, such as multivariate time-series data with numerical and categorical attributes. It consists of raw sensor data in matrix form, where every row is a cycle while columns are the data points within a cycle.

The test rig was designed to emulate many fault scenarios, the details of which are shown in Table II. The test rig is equipped with sensors that record a variety of process values, including pressure (PS1-6), motor power (EPS1), volume flow (FS1-2), temperature (TS1-5), vibration (VS1), cooling efficiency (CE), cooling power (CP), and efficiency factor (SE). These sensors operate independently, meaning their readings remain unaffected by the system's state. Instead, they serve the crucial function of monitoring the system's condition.

TABLE II
SENSOR DATA

Physical Dimension	Sensor	Measuring Value	Units	SF
Pressure	PS1-PS6	Pressure	bar	100 Hz
MotorPower	ESP1	Motor Power	W	100 Hz
Flow Rate	FS1-FS2	Volume Flow	1/min	10 Hz
Temperature	TS1-TS5	Temperature	°C	1 Hz
Vibration	VS1	Vibration	mm/s	1 Hz
Cooling Efficiency	CE	Cooling Efficiency(virtual)	%	1 Hz
Cooling Power	CP	Cooling Power (virtual)	kW	1 Hz
System Efficiency	SE	Efficiency Factor	%	1 Hz

The control parameters in this hydraulic system are deemed dependent variables, as their conditions are directly influenced by the independent variables within the system. These dependent variables include the cooler, valve, pump, and hydraulic accumulator, which are crucial to the system's operation and are detailed in Table III. The cooler regulates the temperature, the valve controls the fluid flow, the pump maintains pressure, and the accumulator compensates for pressure fluctuations.

B. Data Pre-Processing

Data preprocessing is essential since it aligns the collected data and lays a base for any data-driven technique, allowing an accurate and detailed analysis [15]. For this research paper, the data preprocessing implement was first extracting data from text files into arrays. Calculating the mean for all sensors per cycle. Setting the data frame for the processed data includes data from sensors and condition components. Lastly, explore the data to identify any missing or null values. Data exploration is visualizing data by identifying patterns and outliers, data distribution, and its importance. During the exploration, it was found that no data points were missing or outliers. Since the dataset does not include any outliers or missing data, this helps improve the predictions' accuracy.

TABLE III
CONTROL PARAMETERS

Component	Condition
Cooler	Cooling power decrease
Valve	Switching degradation
Pump	Internal leakage
Accumulator	Gas leakage
Stable Flag	Condition stability

C. Modelling

Upon completing data pre-processing, training was assigned 80% of the data, while 20% was reserved for testing. This allocation is consistent with research suggesting that the best results are obtained using 20-30% of all available test data sets and 70-80% on training [33]. Predictive maintenance data modelling contains the prediction model development through ML and DL algorithms- RF, Catboost, and LSTM. Complex pattern recognition and relationship identification effectiveness guide the choice of these algorithms. This process entails choosing appropriate algorithms as well as training them using information that will help bring out patterns in this regard.

RF is a versatile approach that outperforms regression- and classification-related tasks. It is created by combining several decision trees built on random samples and features. It is fast-fitting, low parameter-sensitive, has built-in error estimation, and is effective even when dimensions are extremely high [34]. Catboost, a well-known algorithm for gradient-boosting trees, utilises a random permutation process and unbiased boosting to mitigate information loss and variance. It considers any possible combination involving categorial to facilitate its performance and generalisation [35]. LSTM architecture contains memory blocks equipped with self-connected memory cells and three gates for ruling the information flow. It is suitable for long term dependencies clustering in sequential data, including forecasting time series, natural language processing and speech recognition [36].

D. Evaluation

Evaluation metrics are among the most critical tools for estimating model effectiveness and providing thrilling insights into a number of aspects of the functionality of making predictions by the model. Precision can also be referred to as positive predictive value. It provides a measure of correctness for those positive predictions made by a model in measuring the ratio of accurate positive outcomes to all positive forecasts. Recall, sometimes called sensitivity, is a measure of the model's capability to capture all relevant instances from a dataset. It finds the ratio of correct positive predictions divided by the number of actual positive predictions. The F1-Score encapsulates both precision and recall in one. The score is calculated as the harmonic mean of both, giving a good indicator across classes in an imbalanced data set. Finally, accuracy evaluates the overall correctness of the model based on the number of proper outcomes, including both true positives and true negatives, out of the whole number of cases considered. Table IV shows the evaluation metrics and their formulas.

TABLE IV
EVALUATION METRICS

Metrics	Formula	
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$	(1)
Precision	$\frac{TP}{TP+FP}$	(2)
Recall	$\frac{TP}{TP+FN}$	(3)
F1-Score	$\frac{2 * Precision * Recall}{Precision + Recall}$	(4)

Using Precision, Recall, F1-Score, and Accuracy metrics will be implemented to evaluate the performance of ML and DL models.

III. RESULTS AND DISCUSSION

A comprehensive prediction of the maintenance of hydraulic systems is carried out using ZeMA gGmbH data [32]. After the pre-processing for accuracy and alignment. ML and DL models are applied in predicting the failures in hydraulic systems., including Random Forest, Catboost, and Long Short-Term Memory. In this case, performance-based assessment metrics would be presented together with the model performances, such as Precision, Recall, F1-Score, and Accuracy. The results obtained for each of the five critical components of hydraulic systems, cooler, valve, pump, accumulator, and stable flag, are clearly compared by the algorithms used as shown in Tables V, VI and VII.

A. Random Forest

The RF model predicts the component 'Cooler' with extraordinary accuracy on all metrics as shown in Table V and Fig. 2. The valve component had a consistent performance of 95.92% except for the F1-Score, which is slightly lower at 95.90%. The pump component, however, displayed sorted scores which are near perfect: 99.55% for accuracy and recall, 99.56% for precision, and F1 score. The Accumulator components and Stable Flag performance was also high, ranging from 97.01% to 97.60%. RF model was excellent at predicting several conditions, especially Pump and Cooler. Those high marks under all these standards signify that this model can be a useful resource when it comes to predictive maintenance. The slightly lower scores for the Valve, Accumulator and Stable Flag components, while still high, suggest areas where the model's performance could potentially be improved.

TABLE V
RF COMPARISON OF EVALUATION METRICS

	Accuracy	Precision	Recall	F1 Score
Cooler	100%	100%	100%	100%
Valve	95.92%	95.92%	95.92%	95.90%
Pump	99.55%	99.56%	99.55%	99.55%
Accumulator	97.51%	97.60%	97.51%	97.52%
Stable Flag	97.01%	97.13%	97.05%	97.02%

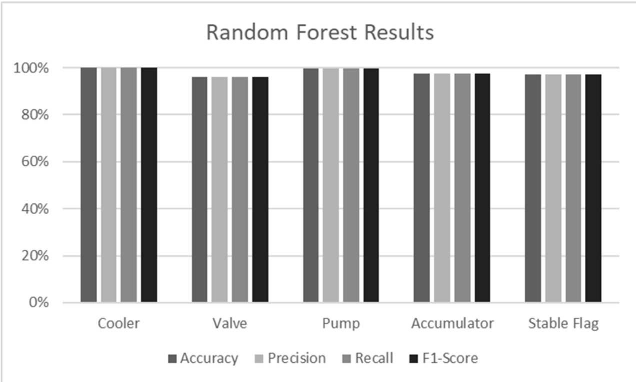

Fig. 2 RF Evaluation Metrics Results

B. *Catboost*

The Catboost model predicts with extraordinary accuracy for component ‘Cooler’, reaching 100% predictor rank in all dimensions as displayed in Table VI and Fig. 3. The Valve achieved scores between 89.02% and 89.12%, suggesting a good estimation rate, though not as perfect. The scores for the Pump component ranged from 98.87% to 98.92%, showing prediction accuracy that is close to perfect. High prediction accuracy is indicated as the Accumulator and Stable Flag component scores range from 95.01% to 95.63%. These findings prove how good our model is at predicting the status of different hydraulic system components: it can do so with assurance due to results obtained through its use, indicating the robustness and reliability of the Catboost model. The model may assist in the predictive maintenance and monitoring of hydraulic systems by preventing potential system failures and enhancing operational efficiency. There are corresponding areas where the valve component could be used to enhance the overall model’s performance. Therefore, further modifications will target and meet this specific design need.

TABLE VI
CATBOOST COMPARISON OF EVALUATION METRICS

	Accuracy	Precision	Recall	F1-Score
Cooler	100%	100%	100%	100%
Valve	89.12%	89.02%	89.12%	88.77%
Pump	98.87%	98.92%	98.87%	98.87%
Accumulator	95.01%	95.06%	95.01%	95.02%
Stable Flag	95.69%	95.83%	95.69%	95.63%

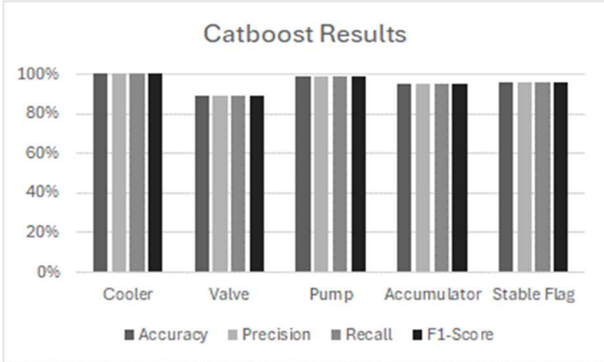

Fig. 3 Catboost Evaluation Metrics Results

C. *LSTM*

The Cooler, Valve and Accumulator components turned to be the best predicted by an LSTM model with this observation pointing out that there exists an outstanding forecasting ability for such components as seen in Table VII and Fig. 4. In terms of performance, however, it is evident that the Cooler, Valve and Accumulator yield better scores as compared to the Pump, with 73.92% for accuracy rate, 82.00% for Precision rate and F1 Scores suggesting potential improvements. The lowest performing component is the Stable Flag component with an accuracy rate of about 66% for recall and 67% precision. The corresponding 53% F1 Score, while the related 44% precision. From these findings, it can be deduced that the LSTM model has some strongholds that should be strengthened to increase its ability to predict outcomes of the hydraulic system more accurately. Future research can be focused on optimizing LSTM architecture, for instance, by introducing more methods of feature scaling and selection in order to enhance the prediction of these constituents.

Fig. 5 illustrates the three ML algorithms—RF, Catboost, and LSTM—across five critical components of hydraulic systems: Pump, Cooler, Hydraulic Accumulator, Valve and Stable Flag. Notably, the Cooler predictions achieved perfect accuracy (100%) across all mod LSTM outperformed the other algorithms for Valve maintenance with 100% accuracy, while Pump predictions favoured the RF model (99.55%). RF and Catboost performed well for Accumulator maintenance (97.51% and 95.01%, respectively). However, in the case of Stable Flag predictions, reliability was observed with RF and Catboost (97.05% and 95.69%), while LSTM lagged (66.67%).

TABLE VII
LSTM COMPARISON OF EVALUATION METRICS

	Accuracy	Precision	Recall	F1-Score
Cooler	100%	100%	100%	100%
Valve	100%	100%	100%	100%
Pump	73.92%	82.23%	73.92%	71.81%
Accumulator	100%	100%	100%	100%
Stable Flag	66.67%	44.44%	66.67%	53.33%

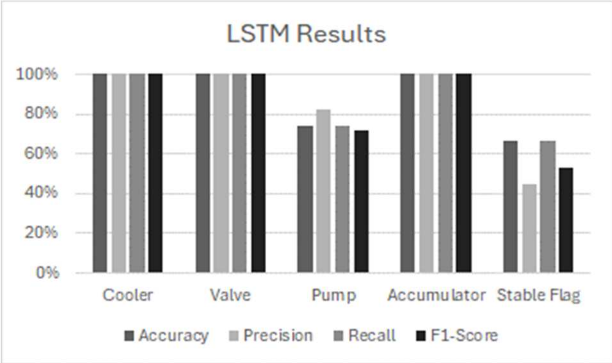

Fig. 4 LSTM Evaluation Metrics Results

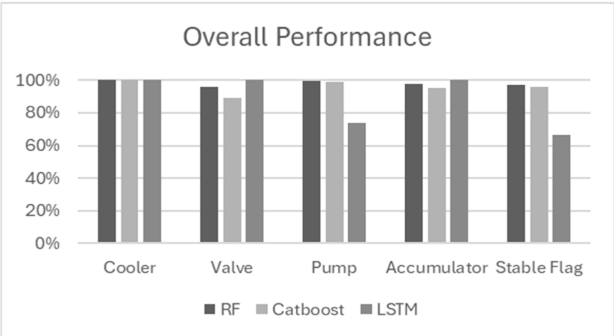

Fig. 5 Overall Accuracy of Classifier Models

RF performed well for the cooler, but the valve, accumulator, and stable flag had lower accuracy. It should be addressed through better hyperparameter tuning and feature selection. Catboost showed high accuracy in the cooler component but lower performance on the valve, indicating the need for feature selection analysis. LSTM had excellent accuracy for valve, cooler, and accumulator prediction results; however, lower accuracy for the pump and stable flag components suggests a need for more hyperparameter tuning or incorporating additional context into the model.

IV. Conclusion

In conclusion, the aim was to explore the potential of various types of classifier models in predicting hydraulic equipment failure. This was done using ML and DL models. RF, Catboost and LSTM algorithms were used in the research. The results show that valve predictions are made by LSTM with excellent performance levels whereas pumps have the best outcome in terms of RF. Besides, accumulators, as well as stable flags, could be predicted accurately by both RF and Catboost. The recommendation for future research is to continue implementing feature scaling, selection, and extraction methods to improve the accuracy of fault prediction models for hydraulic systems.

References

[1] W.-W. Tay, S.-C. Chong, and L.-Y. Chong, "DDoS Attack Detection with Machine Learning," *Journal of Informatics and Web Engineering*, vol. 3, no. 3, pp. 190–207, Oct. 2024, doi:10.33093/jiwe.2024.3.3.12.

[2] S. A. Lashari, M. M. Khan, A. Khan, S. Salahuddin, and M. N. Ata, "Comparative Evaluation of Machine Learning Models for Mobile Phone Price Prediction: Assessing Accuracy, Robustness, and Generalization Performance," *Journal of Informatics and Web Engineering*, vol. 3, no. 3, pp. 147–163, Oct. 2024, doi:10.33093/jiwe.2024.3.3.9.

[3] F. Raza, "AI for Predictive Maintenance in Industrial Systems," *Cosmic Bulletin of Business Management*, vol. 2, pp. 167–183, 2023, doi: 10.13140/RG.2.2.27313.35688.

[4] E. Quatrini, F. Costantino, C. Pocci, and M. Tronci, "Predictive model for the degradation state of a hydraulic system with dimensionality reduction," *Procedia Manuf*, vol. 42, pp. 516–523, 2020, doi:10.1016/j.promfg.2020.02.039.

[5] R. S. Beebe, *Predictive maintenance of pumps using condition monitoring.* Elsevier, 2004. Accessed: Jun. 22, 2024. [Online]. Available: https://openlibrary.org/books/OL3301501M/Predictive_maintenance_of_pumps_using_condition_monitoring

[6] R. Keith Mobley, *An Introduction to Predictive Maintenance*, Second. Elsevier, 2002. doi: 10.1016/B978-0-7506-7531-4.X5000-3.

[7] C. Palanisamy and T. Gangadharan, "Review on Development of Digital Twins for Predicting, Mitigating Faults and Defects in Solar Plants," *International Journal on Robotics, Automation and Sciences*, vol. 6, no. 2, pp. 1–5, Sep. 2024, doi: 10.33093/ijoras.2024.6.2.1.

[8] M. Yugapriya, A. K. J. Judeson, and S. Jayanthy, "Predictive Maintenance of Hydraulic System using Machine Learning Algorithms," in *2022 International Conference on Electronics and Renewable Systems (ICEARS)*, IEEE, Mar. 2022, pp. 1208–1214. doi:10.1109/ICEARS53579.2022.9751840.

[9] J. Wang, T. Zhang, and W. Wang, "Health classification of hydraulic system based on k-means," in *2020 International Conference on Urban Engineering and Management Science (ICUEMS)*, IEEE, Apr. 2020, pp. 550–553. doi: 10.1109/ICUEMS50872.2020.00121.

[10] C.-H. Chen, Y.-K. Chan, and S.-S. Yu, "Hydraulic System Failure Prediction Method with Limited Failure Data," in *2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)*, IEEE, Jun. 2023, pp. 171–173. doi:10.1109/IS3C57901.2023.00053.

[11] Z. Peng, K. Zhang, and Y. Chai, "Multiple fault diagnosis for hydraulic systems using Nearest-centroid-with-DBA and Random-Forest-based-time-series-classification," in *2020 39th Chinese Control Conference (CCC)*, IEEE, Jul. 2020, pp. 29–86. doi:10.23919/CCC50068.2020.9189401.

[12] B. Askari, G. Cavone, R. Carli, A. Grall, and M. Dotoli, "A Semi-Supervised Learning Approach for Fault Detection and Diagnosis in Complex Mechanical Systems," in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, IEEE, Aug. 2023, pp. 1–6. doi: 10.1109/CASE56687.2023.10260469.

[13] A. Buabeng, A. Simons, N. K. Frempong, and Y. Y. Ziggah, "Hybrid intelligent predictive maintenance model for multiclass fault classification," *Soft comput*, vol. 28, no. 15–16, pp. 8749–8770, Aug. 2024, doi: 10.1007/s00500-023-08993-1.

[14] A. Mallak and M. Fathi, "Sensor and Component Fault Detection and Diagnosis for Hydraulic Machinery Integrating LSTM Autoencoder Detector and Diagnostic Classifiers," *Sensors*, vol. 21, no. 2, p. 433, Jan. 2021, doi: 10.3390/s21020433.

[15] V. L. Chekkala and V. Milosavljevic, "Predictive Maintenance for Fault Diagnosis and Failure Prognosis in Hydraulic System MSc Research Project Data Analytics," 2020. Accessed: Dec. 22, 2024.

[16] P. Guo, J. Wu, X. Xu, Y. Cheng, and Y. Wang, "Health condition monitoring of hydraulic system based on ensemble support vector machine," in *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, IEEE, Oct. 2019, pp. 1–5. doi:10.1109/PHM-Qingdao46334.2019.8942981.

[17] S. Gaurkar, A. Kotalwar, and S. Gabale, "Predictive Maintenance of Industrial Machines using Machine Learning," *International Research Journal of Engineering and Technology*, 2021.

[18] N. Helwig, E. Pignanelli, and A. Schutze, "Condition monitoring of a complex hydraulic system using multivariate statistics," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, IEEE, May 2015, pp. 210–215. doi: 10.1109/I2MTC.2015.7151267.

[19] T. Berghout, M. Benbouzid, S. M. Muyeen, T. Bentrcia, and L.-H. Mouss, "Auto-NAHL: A Neural Network Approach for Condition-Based Maintenance of Complex Industrial Systems," *IEEE Access*, vol. 9, pp. 152829–152840, 2021, doi: 10.1109/access.2021.3127084.

[20] J. Soh and D. Kim, "Condition Monitoring with Time Series Data Based on Probabilistic Model," in *2021 24th International Conference on Electrical Machines and Systems (ICEMS)*, IEEE, Oct. 2021, pp. 2630–2634. doi: 10.23919/ICEMS52562.2021.9634481.

[21] J. Wu, P. Guo, Y. Cheng, H. Zhu, X.-B. Wang, and X. Shao, "Ensemble Generalized Multiclass Support-Vector-Machine-Based Health Evaluation of Complex Degradation Systems," *IEEE/ASME*

Transactions on Mechatronics, vol. 25, no. 5, pp. 2230–2240, Oct. 2020, doi: 10.1109/TMECH.2020.3009449.

[22] Z. Xu, H. Yu, J. Tao, and C. Liu, "Compound fault diagnosis in hydraulic system with multi-output SVM," *IET Conference Proceedings*, vol. 2020, no. 3, pp. 84–89, May 2021, doi:10.1049/icp.2021.0470.

[23] J. Liu, H. Xu, X. Peng, J. Wang, and C. He, "Reliable composite fault diagnosis of hydraulic systems based on linear discriminant analysis and multi-output hybrid kernel extreme learning machine," *Reliab Eng Syst Saf*, vol. 234, p. 109178, Jun. 2023, doi:10.1016/j.ress.2023.109178.

[24] D. Kim and T.-Y. Heo, "Anomaly Detection with Feature Extraction Based on Machine Learning Using Hydraulic System IoT Sensor Data," *Sensors*, vol. 22, no. 7, p. 2479, Mar. 2022, doi:10.3390/s22072479.

[25] X. Ma, P. Wang, B. Zhang, and M. Sun, "A Multirate Sensor Information Fusion Strategy for Multitask Fault Diagnosis Based on Convolutional Neural Network," *J Sens*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/9952450.

[26] A. T. Keleko, B. Kamsu-Foguem, R. H. Ngouna, and A. Tongne, "Health condition monitoring of a complex hydraulic system using Deep Neural Network and DeepSHAP explainable XAI," *Advances in Engineering Software*, vol. 175, p. 103339, Jan. 2023, doi:10.1016/j.advengsoft.2022.103339.

[27] P. Zhang, W. Hu, W. Cao, L. Chen, and M. Wu, "Multi-Fault Diagnosis of Hydraulic Systems Based on Fully Convolutional Networks," in *2022 13th Asian Control Conference (ASCC)*, IEEE, May 2022, pp. 631–636. doi: 10.23919/ASCC56756.2022.9828182.

[28] K. Kim and J. Jeong, "Multi-layer Stacking Ensemble for Fault Detection Classification in Hydraulic System," in *2022 26th International Conference on Circuits, Systems, Communications and Computers (CSCC)*, IEEE, Jul. 2022, pp. 341–346. doi:10.1109/CSCC55931.2022.00066.

[29] H. A. Khan, U. Bhatti, K. Kamal, M. Alkahtani, M. H. Abidi, and S. Mathavan, "Fault Classification for Cooling System of Hydraulic Machinery Using AI," *Sensors*, vol. 23, no. 16, p. 7152, Aug. 2023, doi: 10.3390/s23167152.

[30] M. R. Abdurrahman, H. Al-Aziz, F. A. Zayn, M. A. Purnomo, and H. A. Santoso, "Development of Robot Feature for Stunting Analysis Using Long-Short Term Memory (LSTM) Algorithm," *Journal of Informatics and Web Engineering*, vol. 3, no. 3, pp. 164–175, Oct. 2024, doi: 10.33093/jiwe.2024.3.3.10.

[31] J. Jayaram, J. Chetan, and B. Nayak, "Electric Vehicle Health Monitoring with Electric Vehicle Range Prediction and Route Planning," *Journal of Informatics and Web Engineering*, vol. 3, no. 1, pp. 265–276, Feb. 2024, doi: 10.33093/jiwe.2024.3.1.18.

[32] T. Schneider, S. Klein, and M. Bastuck, "Condition monitoring of hydraulic systems Data Set at ZeMA." Accessed: Jun. 22, 2024.

[33] A. Gholamy, V. Kreinovich, and O. Kosheleva, "A Pedagogical Explanation A Pedagogical Explanation Part of the Computer Sciences Commons," 2018. [Online]. Available: https://scholarworks.utep.edu/cs_techrephttps://scholarworks.utep.edu/cs_techrep/1209

[34] C. Zhang and Y. Ma, Eds., *Ensemble Machine Learning*. New York, NY: Springer New York, 2012. doi: 10.1007/978-1-4419-9326-7.

[35] S. Das, *Artificial Intelligence in Highway Safety*. Boca Raton: CRC Press, 2022. doi: 10.1201/9781003005599.

[36] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, vol. 385. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-24797-2.