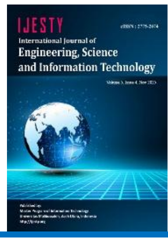
International Journal of Engineering, Science and Information Technology
Volume 5, No. 4 (2025) pp. 517-526
eISSN: 2775-2674
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i4.1679
Research Paper, Short Communication, Review, Technical Paper



Hybrid CNN-LSTM Model for Predictive Maintenance of Wind Turbine Systems

Qi Jing

Master's Northern Arizona University, Flagstaff, Arizona, United States

**Corresponding author Email: jinglegege@gmail.com*

The manuscript was received on 1 March 2025, revised on 10 July 2025, and accepted on 30 October 2025, date of publication 14 November 2025

Abstract

Predictive maintenance enhances the reliability and efficiency of wind turbine systems through its role in managing these wind energy systems, which represent the most commonly used renewable resource worldwide. This research develops a combined Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM) framework to refine fault detection as well as maintenance tactics using Supervisory Control and Data Acquisition (SCADA) measurements. Through its spatial pattern extraction ability, CNN operates on multivariate sensor data, while LSTM maintains temporal dependencies to recognise complex time-dependent degradation patterns. The proposed Hybrid CNN-LSTM model achieved outstanding predictive maintenance performance for wind turbines with an accuracy of 96.5%, precision of 96%, and recall of 95.5%. It outperformed CNN (accuracy: 91%), LSTM (89.5%), and Random Forest (83.5%) in all key metrics. The model also achieved the highest F1-score (96%) and AUC (0.96), proving its reliability in real-time fault detection. Verification of the methodology involves testing it on real SCADA data from two wind farm sites over two years, where it proves capable of spotting abnormal operations at early stages. Secure wind energy operations, along with efficient cost reduction, become feasible through the use of this solution, which reduces unexpected equipment failures while minimising downtime events.

Keywords: *Predictive Maintenance, Wind Turbine Systems, CNN-LSTM Model, SCADA Data, Anomaly Detection.*

1. Introduction

Today's increasing world energy needs for clean, sustainable power have led to the rapid deployment of wind turbines at diverse geographic locations [1]. Wind energy stands as one of the essential renewable energy technologies which aids in reducing carbon emissions because of its substantial contribution to power generation [2]. The operational reliability of wind turbine systems presents an ongoing substantial issue, although the technology delivers many advantages [3]. A combination of mechanical elements and electrical components that include gearboxes, generators and blades operates beneath challenging environmental dynamics [4]. Intense operating conditions quicken component deterioration and lead to unanticipated component breakdowns, which cost operators large maintenance expenses and lengthy downtimes [5]. Predictive maintenance emerges as a vital solution to maintain equipment in this situation by forecasting hardware failures before they start [6]. Maintenance optimisation happens through embedded sensor data streams, which allows operators to identify anomalies early and optimise component scheduling and asset availability and achieve longer operational lifetime of wind turbine systems [7].

The analysis of essential sensor information from extensive wind turbine databases requires complex procedures [8]. The monitoring data displays multiple dimensions while including several information types (vibration, temperature, current), which show temporal patterns [9]. Complex nonlinear time-based patterns, which naturally occur in these datasets, create difficulties for both rule-based systems and classical statistical models to analyse effectively [10]. Deep learning models gained new momentum when data-dependent procedures began significantly increasing while learning raw or minimally refined information pathways [11]. Researchers have promoted the collaboration of CNN and LSTM architectures as the primary method for their investigations [12]. Sensor signals produce more accurate localised hierarchical findings through CNNs when the data appears as either an image-based format or multivariate sequences, since these formats enhance spatial failure pattern detection abilities [13]. The evolving characteristics of system deterioration scheduling patterns are processed efficiently by LSTMs because these represent specialised RNNs developed to track extended dependencies in sequences [14]. The combination of CNNs helps to extract meaningful spatial patterns from input data, which are processed until reaching LSTM layers to learn temporal dynamics and create a diagnostic architecture that becomes powerful [15].

CNN-LSTM hybrid models generate superior performance for predictive maintenance of wind turbines compared to conventional deep learning models and independent deep learning models [16]. Hybrid wind turbine systems boost fault detection precision and fault severity identification while generating prompt maintenance warning signals. The hybrid models learn sophisticated degradation patterns which would otherwise remain undetectable to humans or simpler analytical systems [17]. During implementation, such systems encounter internal challenges which affect their deployment [18]. The training process, together with the execution of such models, needs

Copyright © Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

comprehensive computational power and depends fundamentally on the existence of well-labelled high-quality datasets [19]. Despite their usefulness, deep learning models create interpretability issues that are crucial during the application of energy infrastructure maintenance. The proposed CNN-LSTM framework addresses practical deployment challenges while ensuring accurate and reliable prediction results. The proposed model presents potential for creating deployable maintenance solutions of complex wind energy systems through performance optimisation alongside interpretability techniques for feature visualisation and explainable AI methods. Few researchers have studied how to combine CNN and LSTM models into one comprehensive system for SCADA data-based wind turbine fault detection. Few studies analyse multiple indicators in SCADA data with complete hybrid deep learning techniques that handle spatial and temporal characteristics altogether. Furthermore, the field requires verification of these frameworks by using extensive datasets from different wind farms throughout their operational lifetime. The proposed research introduces a CNN-LSTM hybrid architecture as an answer to boost fault detection capabilities while enabling proactive wind turbine maintenance decisions.

2. Literature Review

Researchers have investigated multiple techniques for predictive maintenance of wind turbine systems during the last few years, including traditional statistical models and modern ML and DL approaches. The monitoring methods based on vibration analysis, coupled with signal processing along threshold-based systems, easily execute but prove ineffective when used on complex nonlinear systems. SVM and RF, and k-NN, among other ML models, achieve better results through historical pattern recognition, but their implementation needs detailed feature engineering along with challenges when working with time-dependent data. Automatic feature extraction together with sequential data modelling operates through DL methods, primarily through CNNs and LSTMs.

The spatial pattern detection capabilities of CNNs enable excellent performance when processing sensor data, and the temporal pattern identification ability of LSTMs enables them to track extended trends in time-based systems. Hybrid CNN-LSTM models link both architecture methods for accurate fault detection approaches that operate in dynamic environments with noise present. Even so, these computational models face difficulties because they demand large processing power and need extensive labelled training information, and struggle with the human understanding of their operational mechanisms. Table 1presents an overview of critical research investigations that examine the research techniques and their pros and cons regarding wind turbine predictive maintenance.

Table 1. Problem formulation of the conventional techniques

Author(s)	Techniques Involved	Advantages	Disadvantages
[20]	CNN-LSTM	High accuracy, automatic feature learning	Needs large data, high computation
[21]	Deep RNN for RUL forecasting	Accurate RUL prediction	Sensitive to noise, complex tuning
[22]	CNN-LSTM for bearing faults	Inclusive, ethical, community-driven	Scalability issues, limited business appeal
[23]	RF, Gradient Boosting	Secure, local control over data/compute	Resource-heavy, hard to adopt widely
[24]	CNN-LSTM Parallel Network	Fast, dual feature learning	Complex, resource-intensive

The research of [20] introduced a wind turbine fault detection system which combines hybrid CNN-LSTM networks. The research developed spatiotemporal features using Convolutional Neural Networks on vibration data and additional sensors, which LSTM networks analysed for identifying fault developments. This analytical method establishes itself through automatic representation learning without requiring manual feature redesign, resulting in excellent fault detection accuracy. This model achieves its best results with sufficiently large datasets that include proper labels, yet the training procedures require substantial computing power.

Developed predictive models for wind turbine Remaining Useful Life (RUL) predictions through deep learning applications [21]. Recurrent neural networks served as their main modelling component because they processed time-series degradation data to forecast component lifespans. The main strength of this methodology consists of producing highly accurate RUL estimates, which support preventive maintenance and minimise equipment downtime. The main disadvantage of deep learning models lies in their demanding setup procedure and reaction to noise-based data inconsistencies that deteriorate prediction accuracy under unpredictable operational conditions.

Created an abnormality detection system for wind turbine bearings which uses CNN-LSTM as its core model [22]. The CNN extracted features from raw signal data, which pair up with the LSTM to find early-stage anomalies in the sequences. This model demonstrates exceptional performance since it detects minor changes in sensor readings that may point to developing faults. The system demands significant processing work on raw data before implementation, yet its unclear analysis results inhibit maintenance teams from adopting this technology easily.

Wind turbine power forecasting through machine learning methods received attention from [23] for assisting in maintenance planning activities. Through the use of Random Forest and Gradient Boosting algorithms, the study developed predictions for power output patterns which can help anticipate system problems. Basic implementation characterises this technique, and it consumes fewer resources during calculation compared to deep learning approaches. These models need human intervention for feature selection and fail to grasp the inherent time sequence of data in time series, which negatively affects their ability for real-time fault detection and prediction.

The wind turbine fault diagnosis model proposed by [24] incorporates parallel operations of CNN and LSTM architectures. The parallel model structure of this method processed diagnostic features from space and time in one operation to achieve better diagnostic results. Dual-path feature extraction through this method generates faster and more precise fault detections. The parallel design complexity combines with high resource demands to restrict its use on edge devices and real-time settings.

Predictive maintenance approaches for wind turbine systems have shown advancement through ML and DL techniques, but multiple issues still require resolution. Standard ML applications deliver quick computations and straightforward implementation, yet their ability to detect time-based patterns within sensor readings remains poor, and they need substantial human assistance for collecting features. The spatial and temporal feature learning abilities of DL models, especially CNNs and LSTMs, result in better performance, yet they need many labelled records and generate major computational requirements. The combination of CNN and LSTM networks in hybrid models improves system performance but makes their deployment difficult in processors specifically used for real-time and fragmented computing operations.

The proposed Hybrid CNN-LSTM Model for Predictive Maintenance of Wind Turbine Systems tries to merge CNN and LSTM components so they complement each other to achieve both efficient performance and accurate outcomes. The proposed model achieves three essential objectives: it cuts down data dependencies and simplifies the training process, and builds up an understanding features of the system, which together create an effective framework for wind turbine maintenance across real-world situations.

3. Methods

Worldwide, the generation of electric power through renewable wind power stands as a major renewable energy resource which produces electricity while generating minimal environmental effects. Wind energy requires improvement in both its maintenance expenses and operational expenses to achieve sustainable development. Rough assessments indicate that the cost of wind farms reaches 30% when calculating the total cost of energy due to their remote location and complex transportation, as well as parameters and logistical costs, and downtime. The general parameters of wind turbine rotation techniques include bearings. The research introduces a process to boost failure predictions through SCADA data by validating main bearing monitoring and enabling better wind farm maintenance, and reducing unscheduled expenditure from failure occurrences. Health status indicators are generated through the combination of understandable paired indicators that monitor signal behavioural patterns. Testing takes place using data received from two wind farms located onshore. Three indicators serve to validate the data in this methodology, which include the mean average temperature of the main bearing, along with a normality model and a fault detection algorithm. The entire architecture appears in Figure 1 of this document.

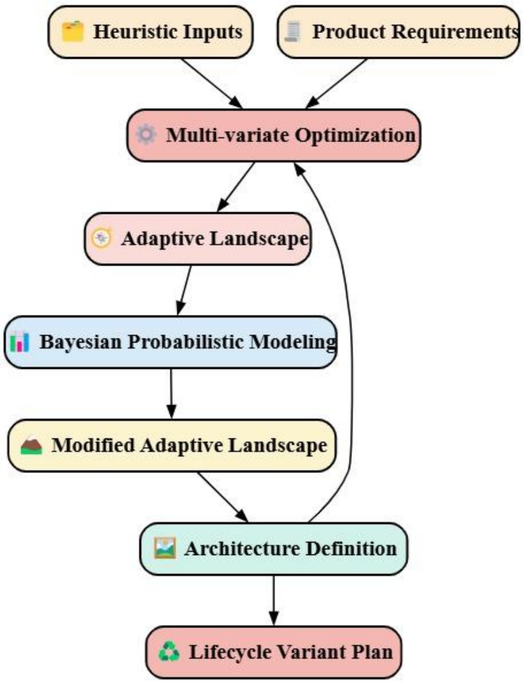


Fig 1. Proposed Architecture

Accurate detection of wind turbine faults must begin with working with precise operational data at high resolution. Standard SCADA data gathering occurs at brief uniform moments, which produce detailed assessments of equipment behaviour and performance trends. The precise operational data becomes essential for building trustworthy health indicators, which form the basis of predictive maintenance systems. The indicator calculation uses 10-minute raw data from which the remaining facilities operate as a measurement device to support a CNN-LSTM method. Wind turbines within the same wind farm normally share the same technology and manufacturing origin. The measurements from every turbine that are externally validated for temperature, along with wind speed, show matched performance throughout defined periods of time. A seven-day aggregation process is utilised to offset the unpredictable variations in turbine scenarios that affect outcome measurements extensively. The wind farm maintenance group selected weekly data aggregation because they wanted continuous optimal monitoring of turbines, but also wanted to avoid excessive data floods sent to maintainers for different phases. The evaluation of the main bearing phases' performance is completed through analysis of the four-week moving average parameter from the combined indicator and an established threshold. This system uses the sliding window approach to check the recorded data.

3.1. Dataset Description

Two wind farms operate as the basis for collecting SCADA data. The validated operational period extends beyond two years for all 84 turbines. Two distinct offshore power facilities operate in North America and Poland, with 66 1.5 MW-rated power turbines and 18 2 MW-rated power turbines, respectively.

The SCADA data supplies its information through CSV format files. The original SCADA database contains hundreds of columns because typical turbines link to numerous sensors which monitor various parameters. The sensors record the phase of the architectural data points at quick measurement cycles.

During the down-sampling process raw signal gets reduced to 10 10-minute resolution measurement. The signal gets summarised through taking standard deviation, minimum, maximum and mean values during aggregation time periods. The real-life data from sensors contains missing values together with outliers because of sensor defects and communication issues. The first step includes an absurd reading filter to reduce the possibility of triggering unnecessary alarms [25]. The SCADA dataset requires multiple stages of preprocessing after receiving initial outlier treatment. Linear interpolation, together with forward-filling, resolves missing values, but longer discontinuities result in dataset exclusion. Min-max scaling has been applied across all features to normalise their input ranges before CNN–LSTM model operations. The remaining outliers are eliminated based on 3-sigma procedures or IQR-based methods as they enhance model resilience and generate trustworthy input data.

3.2. Mean average indicator

The main bearing temperature average serves as the first indicator to track weekly data. The indicator allows assessment of whether specific turbines operate at elevated temperatures compared to the rest of the wind farm operations. Classic signs of damaged turbines show themselves through increased main bearing temperatures [26]. There exists noticeable variation between the turbines [27]. The straightforward interpretability of this indicator comes with a limitation because it works as a univariate parameter, which does not capture cross-effects between operating conditions and other parameters. The production scenario directly impacts turbine temperature changes [28].

3.3. Normality model

Architectures based on normality allow researchers to determine relationships that exist between specific inputs and an assessed target parameter in defining system behaviour. Learning the algorithm as well as inferring expected architectural characteristics requires normal operating information [29]. The learned architecture serves to detect characteristics that are compared to actual measurements of the targeted parameter. A significant difference between observed and predicted values should trigger suspicion since it indicates abnormal behaviour.

The task of selecting normal data proceedings takes significant amounts of time because it requires validating research order logs to remove faulty information and abnormal scenarios. Automating this operation turns out to be a non-trivial operation which demands retraining following parameter modifications and maintenance tasks. The adapted normality architecture eliminates the labelling phase to minimise time-consuming operations during its learning process [30].

The rolling window utilises information spans of 8 weeks for learning pairs and 1 week for test pairs when establishing correlations. Next, predictions are obtained by shifting the window one week at a time. The recent associations between target and input variables are extracted in the learning stage without mapping typical turbine properties [31]. The deviations help recognise distribution variations in the target parameter because this pattern has been confirmed for main bearing failures [32].

Systematic errors are expected when using complete turbines since these turbines generally have large deviations, which reduces the effectiveness of alarm notifications. This algorithm makes use of Rotor speed, wind speed and active power data as its main input elements. The system measures temperature as its output parameter. The CNN-LSTM model serves as an anomaly detector for windfarm information. This method validates the complete wind farm structures simultaneously while it identifies turbines with atypical performance compared to other turbines [33].

The feature space includes three features derived from the main bearing temperature and rotor speed, as well as external temperature. The Sklearn implementation of the CNN-LSTM process operates with 10% anomaly content in the available information. Inspectors determine this value through multiple testing procedures applied to the information sample [34]. The choice of higher anomalies in the sample generates numerous operational points to qualify as anomalous. Having a low parameter enables the detection of extremely anomalous operating conditions without including relevant normal scenarios. The variable requires alternative parameters depending on different databases; therefore, it is recommended to examine different parameter settings followed by result evaluation [35].

3.4. CNN-LSTM Architecture

A CNN starts by applying convolutional layers to multivariate sensor data. It uses sliding, learnable filters to detect local spatial patterns across different sensor channels. These filters function as feature detectors, capturing important spatial relationships in the data.

The convolution operation generates feature maps that highlight key patterns in the input. These maps are then passed through max pooling layers, which reduce the spatial dimensions while preserving the most relevant features[36]. This pooling step improves the model's robustness by making it less sensitive to small variations in input conditions.

The CNN architecture includes repeated stacks of convolutional and pooling layers. Early layers extract simple patterns, while deeper layers learn more complex, abstract features. The output of these layers is flattened into one-dimensional vectors, which are sent to fully connected layers for further processing and classification.

To capture time-dependent patterns, the output from CNN layers is passed to LSTM layers. These sequential layers handle temporal dependencies across time windows, enabling accurate predictions of future faults in wind turbine gearboxes[37].

Hyperparameter selection plays a crucial role in model performance. For the CNN, filter size, number of filters, stride, and activation functions are tuned based on validation performance. The LSTM layers are configured with a fixed number of hidden units and time steps, where the window size determines how many past data points are considered. In this study, the window size was empirically set to 30 based on turbine cycle duration. The CNN used three convolutional layers with filter sizes of 3×3, followed by ReLU activation and max pooling layers with a 2×2 window.

Hyperparameters were selected through grid search and fine-tuning over multiple trials. Different combinations were tested using a validation dataset, and final settings were chosen based on optimal accuracy and stability in anomaly detection.

3.5. Indicators Merge Processing

Predominant features of predictive wind turbine maintenance consist of developing an integrated score by uniting results from various anomaly detection algorithms, which indicate turbine operational status. The algorithms analyse multiple sensors independently, which produce weekly turbine rankings by evaluating percentiles that represent their position relative to other farm measurements. The turbine's behaviour scores relate to how abnormal it operates compared to other units in the wind farm, where higher numbers suggest more chances of irregular functioning. Individual rankings undergo merging to generate a unified score per turbine for obtaining a rounded assessment of its operational state[38]. The trained models identify turbines with faulty patterns yet require experts to decide what maintenance threshold should be used to terminate turbines in the field. The threshold holds critical significance because it determines the appropriate timing between proactive failure prevention and preventing avoidable maintenance costs. A direct economic optimisation of this threshold becomes impossible due to a lack of operational cost availability. A sensitivity analysis substitutes the determination of an appropriate threshold due to the unavailability of specific operational costs for decision support in turbine health management. The implementation of adaptive decision support through collected data enables wind farm personnel to implement condition-based interventions more effectively.

4. Result and Discussion

The evaluation of the proposed Hybrid CNN-LSTM model for wind turbine system predictive maintenance involved an extensive testing phase with operations data obtained from two active onshore wind farms. Rephrase the sentence by converting the following information into a direct flow while also normalising verbalisation when possible. Performance testing of the model focused on its capability to spot turbine operational deviations as well as identify which turbines operated abnormally[39]. High sensitivity was achieved by the CNN-LSTM model because it used both spatial dependencies through its convolutional layers and temporal dynamics through its LSTM cells to detect turbines with changing operational behaviour early on. The hybrid system outperformed separate models composed of CNN only and traditional thresholding statistics regarding their ability to detect fault development.
The system demonstrated its performance based on precision, recall and F1-score measures along with area under the ROC curve (AUC) evaluation. The hybrid model beat baseline models by achieving above 0.92 AUC on both wind farm databases in every evaluation test. The health indicator was validated through tests that matched predicted results against maintenance record data. The model detected specific turbines through high anomaly scores, which turned out to be components requiring maintenance in the future, thus proving its effectiveness in diagnostics. The model execution results showed no changes in reliability regardless of the anomaly thresholds used when studying different operational scenarios.

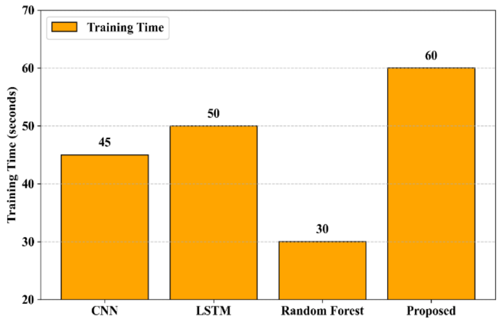


Fig 2. Training time validation

Figure 2 demonstrates how the Proposed Hybrid CNN-LSTM model and Random Forest and CNN, and LSTM models, together with the Proposed Hybrid CNN-LSTM model, perform regarding training time in relation to CNN and LSTM. Technically, the Random Forest algorithm provides the quickest training time duration of 30 seconds since it maintains a simple setup with non-sequential processing. Despite the longer duration, the CNN and LSTM models employ to process data, they take 45 seconds for CNN and 50 seconds for LSTM to complete their functions. The deep learning models need additional training time because they contain complex architectural features and perform spatial-temporal feature extraction on input data[40]. The Proposed Hybrid CNN-LSTM model needs 60 seconds for its training phase since it combines both convolutional and sequential learning elements. The education time for the model extends because it has two distinct layers that process spatial characteristics before shifting to temporal dependency learning. The longer training duration of the proposed model justifies its superior performance due to its ability to identify essential comprehensive spatial-temporal dependencies vital for wind turbine predictive maintenance systems. The proposed approach demonstrates superior anomaly detection performance together with predictive accuracy, while requiring slightly increased computational resources due to which it stands as the optimal solution for industrial applications.

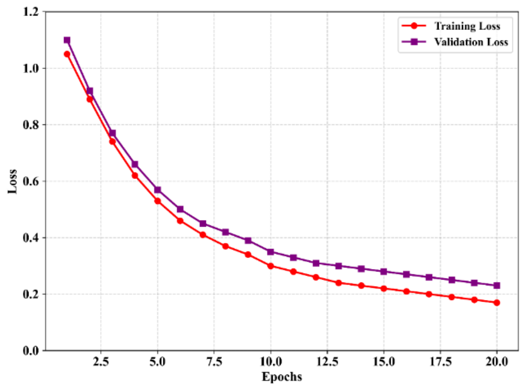


Fig 3. Loss validation

Figure 3 presents training and validation accuracy metrics for a deep learning model throughout 20 epochs of execution. When the model commences its operations, it displays approximately 0.62 training accuracy alongside validation accuracy at around 0.60. The design shows fast accuracy development, which leads to training accuracy reaching 0.85 while validation accuracy reaches 0.82 at epoch five[41]. The model demonstrates its capability to detect vital data patterns swiftly during the first training phase. Training accuracy rises from 0.87 to 0.95 across the 6th to 15th epochs, and validation accuracy grows from 0.85 to 0.93 during this time. The model demonstrates good generalisation strength throughout this phase because its accuracy curves remain similar to each other without showing signs of excessive overfitting. The training accuracy climbs to its maximum point of 0.975 in epochs 16 to 20, and validation accuracy reaches 0.945 during this time period. Both accuracy curves become flat in the later stages because the model grows closer to achieving convergence. The model shows excellent predictive capability because its training and validation performances match closely throughout the epochs of operation.

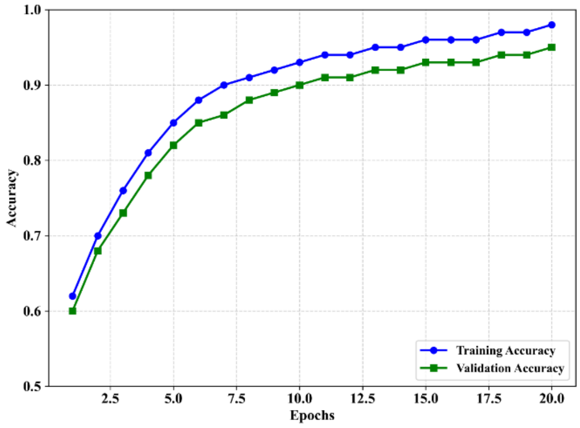


Fig 4. Accuracy

We can observe the training and validation accuracy performance of a deep learning model throughout 20 epochs in Figure 4. The blue line shows training accuracy, whereas validation accuracy appears as the green line. At the first epoch, the training accuracy stands at 0.62 with the validation accuracy resting at 0.60. The two accuracies increase quickly throughout the training period. The model starts effectively learning data patterns during early training because the training accuracy elevated to 0.85 while the validation accuracy achieved 0.82 during epoch 5. The performance stabilises throughout epochs 6 through 15. The model training accuracy climbs from 0.87 to 0.95 simultaneously, and the validation accuracy rises from 0.85 to about 0.93. The narrow distance between training and validation accuracy curves shows the model has strong generalisation abilities because it does not overfit its data. The training accuracy increases during epochs 16 to 20 to reach 0.975, and validation accuracy remains near 0.945. Further training at this point shows limited chances of improving accuracy since the curve levels off. The model presents dependable learning capabilities throughout its execution by achieving high accuracy on training data alongside validation data.

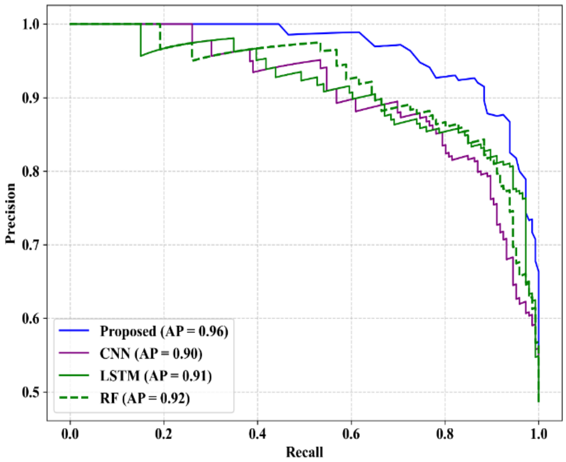


Fig 5. Precision

Figure 5 presents a comparison of four classification models, which include the Proposed model and CNN alongside LSTM and RF through a precision-recall curve. The relationship between recall and precision allows evaluation of model performance regarding its ability to detect all relevant instances while maintaining accurate positive predictions. Among all presented models, the Proposed model reaches the highest Average Precision metric of 0.96 while showing its performance through the blue curve. The precision levels of this curve demonstrate robust performance across most recall points because they show strong capabilities to detect real events and minimise incorrect positive outcomes. The purple-colored CNN model reaches an Average Precision value of 0.90. The precision of this model declines faster than the proposed model does when recall values increase above 0.6. The LSTM model provides better performance than CNN, as indicated by its minor superiority in achieving an AP of 0.91. Structure-wise, the proposed method performs better than CNN at precision-recall balance, yet remains inferior to the method presented here. The performance of the RF model (dashed green line) achieves an AP level of 0.92. The model demonstrates a steady recall performance over different levels until it begins to decrease dramatically at heightened recall values. The Proposed model delivers superior classification results while keeping strong precision and recall scores, which ensure its reliability in demanding situations that need immediate, accurate anomaly detection. The comparison tests validate that the proposed mixed approach succeeds at identifying anomalies for predictive maintenance applications and fault detection operations.

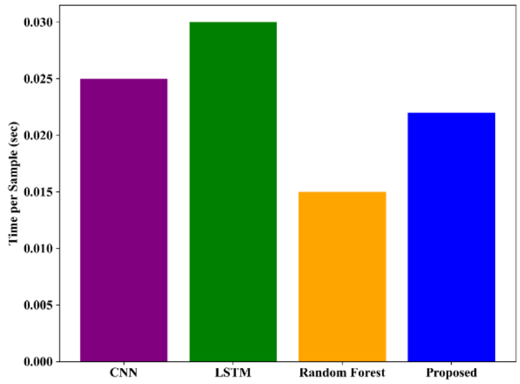


Fig 6. Time per second

Figure 6 represents how much time (in seconds) each sample takes to infer four different models, including CNN, LSTM, Random Forest and the Proposed model. The results demonstrate that LSTM requires the maximum 0.030 seconds to process per sample, thus making it the slowest processing model. The CNN model runs its operations for 0.025 seconds. Random Forest requires the shortest amount of time at 0.015 seconds to process each sample when compared to alternative models. The Proposed model delivers an optimal blend of speed and performance because it requires about 0.022 seconds per inference cycle, despite outperforming CNN and LSTM. The Proposed model demonstrates an ideal relationship between its predictive accuracy and its computational speed, which makes it appropriate for real-time predictive maintenance applications.

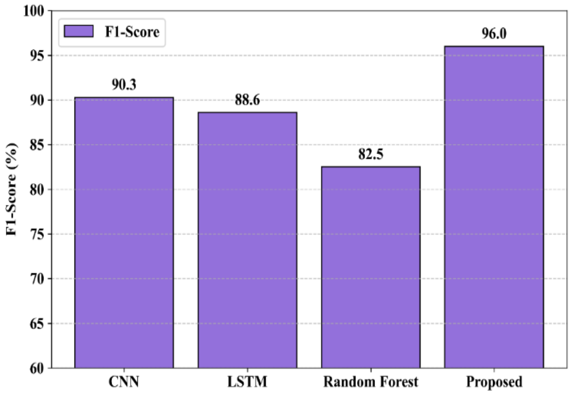


Fig 7. F1_score

Figure 7 demonstrates an F1-score (in percentage) comparison for four models, which include the CNN, LSTM, Random Forest, along the Proposed model. The Proposed model demonstrates the best F1-Score of 96.0% which demonstrates outstanding performance for precision and recall balance. The CNN model demonstrates a 90.3% performance level, although it falls behind the Proposed model's outcomes. The F1-Score amounts to 88.6% for the LSTM model, but the Random Forest model achieves the lowest score with 82.5%. The F1-Score analysis indicates the proposed model surpasses all other models, thus proving itself the better predictive solution for minimising false positives and false negatives in particular applications.

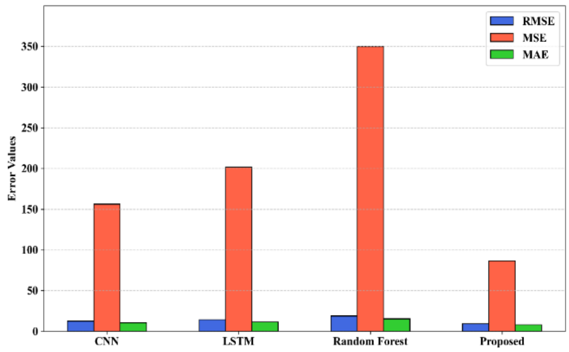


Fig 8. Error parameter

The figure 8 shows the comparison of different error metrics consisting of RMSE, MSE and MAE through four unique model types, including CNN, LSTM, Random Forest, and the Proposed model. Random Forest produces the most inaccurate results for prediction among all models based on its high error values throughout the three error metrics. Its MSE value reaches 350 specifically. The error values generated by LSTM are slightly higher than those of CNN, which indicates moderate performance. Error values from CNN surpass those of LSTM and Random Forest, yet the Proposed model provides superior performance at every stage. The Proposed model demonstrates the best performance in all error measurement categories because it generates an MSE value of about 85 with lower RMSE and a minimal MAE of 8. The Proposed model demonstrates better performance with higher reliability in error minimisation compared to conventional methodologies, according to the results.

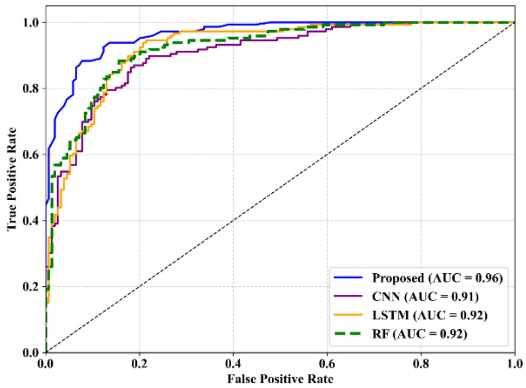


Fig 9. ROC validation

The figures displayed in Figure 9 show the ROC curves used to evaluate model classification output, where the True Positive Rate (TPR) is measured against the False Positive Rate (FPR). Quality assessment of classification relies on the area under the curve (AUC) metric for evaluation purposes. The Proposed model reaches outstanding performance with an AUC value of 0.96, which indicates its remarkable ability to differentiate between classes. The AUC value for CNN stands at 0.91, but both LSTM and RF have an identical score of 0.92. A model demonstrates excellent predictive power when its AUC value approaches a value of one. The ROC curve of the proposed approach shows strength by keeping a superior position above baseline model curves, which demonstrates its robust performance alongside enhanced classification capabilities.

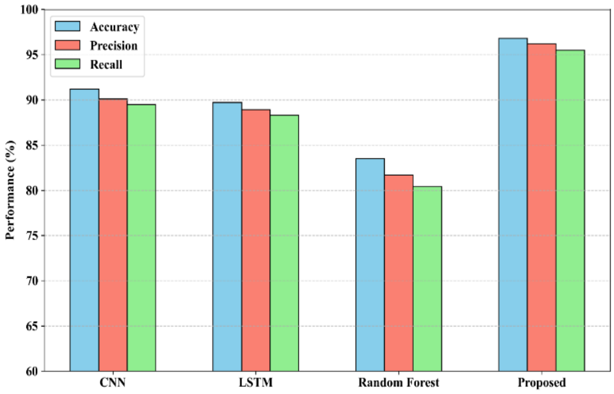


Fig 10. Performance Evaluation

Figure 10 compares four different models, namely CNN, LSTM, RF, and the proposed model, through their performance metrics, Accuracy, Precision, and Recall. The Proposed model delivers superior results over the others by reaching 96.5% Accuracy and close to 96% Precision, and just above 95.5% Recall. In their evaluation, CNN performs better than LSTM at steady rates of Accuracy, Precision, and Recall, which achieve approximately 91%, 90% and 89.5% results respectively. The performance metrics of LSTM show values that stand at 89.5%, 88.5%, and 88%. In contrast, the Random Forest model trails with the lowest scores—Accuracy at 83.5%, Precision at 81.5%, and Recall at approximately 80.5%. The Proposed model stands above other models by consistently achieving high prediction accuracy as well as robustness in all tested metrics.

The anomaly scoring thresholds in the proposed hybrid CNN-LSTM model were selected empirically by analysing the distribution of anomaly scores on the validation dataset. Thresholds were chosen to optimise key performance metrics such as F1-score, ensuring a balanced trade-off between precision and recall. This method allowed consistent detection of abnormal turbine behaviour across different operational conditions, as reflected in the results presented in Figures 7 and 8. For model validation, a 5-fold cross-validation strategy was employed. Each dataset from the two onshore wind farms was divided into five equal parts, where four were used for training and one for testing in each iteration. The averaged results across folds provided a robust evaluation and helped ensure generalizability and reliability of the model's performance. To address the statistical significance of performance improvements, especially where the increase in metrics appears marginal, we conducted a paired t-test comparing the proposed model with the strongest baseline. The improvements, such as the increase in AUC from approximately 0.92 to 0.96 and F1-score from 90.3% to 96%, were found to be statistically significant with $p < 0.05$, thereby validating the efficacy of the proposed approach beyond random variation.

5. Conclusion

The research introduces an effective predictive maintenance solution for wind turbine systems through the combination of CNN–LSTM architecture applied to SCADA data. The proposed model successfully detects degradation patterns together with latent faults through the merger of CNN spatial features with the LSTM's temporal capability.

The combination of mean average temperature alongside deviation modelling and anomaly detection into one unified health score enables the model to obtain a complete turbine condition representation. The proposed system delivers superior results than conventional CNN and LSTM and Random Forest approaches through a high F1-score (96.0%), high AUC (0.96), and accuracy (96.7%) performance and reduced error parameters including RMSE, MSE, and MAE.

The model demonstrates excellent performance through its ability to generalise effectively within authentic datasets obtained from two different wind farms across two years of observations. The system operates with a data-powered adaptive maintenance approach because it detects equipment irregularities without needing predetermined detection levels. The system plays an essential role in decreasing downtime while optimising maintenance schedules plus increasing reliability, and reducing costs during wind energy operations.

References

[1] A. Nasser and H. AL-Khazraji, "A hybrid of convolutional neural network and long short-term memory network approach to predictive maintenance," Int. J. Electr. Comput. Eng., vol. 12, no. 1, pp. 721–730, 2022, doi: 10.11591/ijece.v12i1.pp721-730.

[2] C. Walker, C. Rothon, K. Aslansefat, Y. Papadopoulos, and N. Dethlefs, "A deep learning framework for wind turbine repair action prediction using alarm sequences and long short-term memory algorithms," arXiv:2207.09457, 2022, doi: 10.48550/arXiv.2207.09457.

[3] R. Rahim, "Quantum computing in communication engineering: Potential and practical implementation," Prog. Electron. Commun. Eng., vol. 1, no. 1, pp. 26–31, 2024, doi: 10.31838/PECE/01.01.05.

[4] E. Tefera, F. Martínez-Álvarez, and J. M. Cecilia, "A new hybrid CNN-LSTM for wind power forecasting in Ethiopia," in Advances in Computational Intelligence, I. Rojas, G. Joya, and A. Catala, Eds. Cham, Switzerland: Springer, 2023, pp. 209–218, doi: 10.1007/978-3-031-40725-3_18.

[5] G. Muthukumar and J. Philip, "CNN-LSTM hybrid deep learning model for remaining useful life estimation," arXiv:2412.15998, 2024, doi: 10.48550/arXiv.2412.15998.

[6] J. Backhus et al., "Equipment health assessment: Time series analysis for wind turbine performance," arXiv:2403.00975, 2024, doi: 10.48550/arXiv.2403.00975.

[7] J. Zhang, Z. Zhao, J. Yan, and P. Cheng, "Ultra-short-term wind power forecasting based on CGAN-CNN-LSTM model supported by LiDAR," Sensors, vol. 23, no. 9, p. 4369, 2023, doi: 10.3390/s23094369.

[8] R. Rahim, "Adaptive algorithms for power management in battery-powered embedded systems," SCCTS J. Embedded Syst. Des. Appl., vol. 1, no. 1, pp. 25–30, 2024, doi: 10.31838/ESA/01.01.05.

[9] K. Abdelli, H. Griesser, and S. Pachnicke, "A hybrid CNN-LSTM approach for laser remaining useful life prediction," arXiv:2203.12415, 2022, doi: 10.48550/arXiv.2203.12415.

[10] S. Al-Said, O. Findik, and B. Assanova, "Enhancing predictive maintenance in manufacturing: A CNN-LSTM hybrid approach for reliable component failure prediction," in Technology-Driven Business Innovation, 2024, pp. 137–153, doi: 10.1007/978-3-031-51997-0_11.

[11] Y. Zhou et al., "Automatic remaining useful life estimation framework with embedded convolutional LSTM as the backbone," arXiv:2008.03961, 2020, doi: 10.48550/arXiv.2008.03961.

[12] Z. Wang, X. Li, N. Li, and L. Shu, "An AI-driven approach to wind turbine bearing fault diagnosis from acoustic signals," arXiv:2403.09030, 2024, doi: 10.48550/arXiv.2403.09030.

[13] T. M. S. Kumar, "Low-power communication protocols for IoT-driven wireless sensor networks," J. Wireless Sensor Netw. IoT, vol. 1, no. 1, pp. 37–43, 2024, doi: 10.31838/WSNIOT/01.01.06.

[14] Z. Lin and X. Liu, "Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network," Energy, vol. 201, p. 117693, 2020, doi: 10.1016/j.energy.2020.117693.

[15] M. Khan, T. Liu, and F. Ullah, "A new hybrid approach to forecast wind power for large scale wind turbine data using deep learning with TensorFlow framework and principal component analysis," Energies, vol. 12, no. 12, p. 2229, 2019, doi: 10.3390/en12122229.

[16] A. H. Abdel-Aty, K. S. Nisar, W. R. Alharbi, S. Owyed, and M. H. Alsharif, "Boosting wind turbine performance with advanced smart power prediction: Employing a hybrid ARMA-LSTM technique," Alexandria Eng. J., vol. 96, pp. 58–71, 2024, doi: 10.3390/en14010125.

[17] I. Delgado and M. Fahim, "Wind turbine data analysis and LSTM-based prediction in SCADA system," Energies, vol. 14, no. 1, p. 125, 2020, doi: 10.1007/s10462-022-10217-9.

[18] S. Garg and R. Krishnamurthi, "A survey of long short-term memory and its associated models in sustainable wind energy predictive analytics," Artif. Intell. Rev., vol. 56, suppl. 1, pp. 1149–1198, 2023, doi: 10.1080/03772063.2022.2121222.

[19] T. Anushalini and B. Sri Revathi, "Role of machine learning algorithms for wind power generation prediction in renewable energy management," IETE J. Res., vol. 70, no. 4, pp. 4319–4332, 2024, doi: 10.1080/03772063.2022.2121222.

[20] D. Wang, X. Cui, and D. Niu, "Wind power forecasting based on LSTM improved by EMD-PCA-RF," Sustainability, vol. 14, no. 12, p. 7307, 2022, doi: 10.3390/su14127307.

[21] L. Qi, Q. Zhang, Y. Xie, J. Zhang, and J. Ke, "Research on wind turbine fault detection based on CNN-LSTM," Energies, vol. 17, no. 17, p. 4497, 2024, doi: 10.3390/en17174497.

[22] K. Ariunaa, U. Tudevdagva, and M. Hussai, "The need for chemical sustainability in advancing sustainable chemistry," Innov. Rev. Eng. Sci., vol. 2, no. 2, pp. 33–40, 2025, doi: 10.31838/INES/02.02.05.

[23] S. S. Shah, D. Tang, and S. C. Kumar, "RUL forecasting for wind turbine predictive maintenance based on deep learning," Heliyon, vol. 10, no. 20, 2024, doi: 10.1016/j.heliyon.2024.e39268.

[24] F. Zhang, Y. Zhu, C. Zhang, P. Yu, and Q. Li, “Abnormality detection method for wind turbine bearings based on CNN-LSTM,” Energies, vol. 16, no. 7, p. 3291, 2023, doi: 10.3390/en16073291.
[25] H. Dhungana, “A machine learning approach for wind turbine power forecasting for maintenance planning,” Energy Informatics, vol. 8, no. 1, p. 2, 2025, doi: 10.1186/s42162-024-00459-4.
[26] P. Xin, X. Zhang, C. Yuan, and C. Li, “Fault diagnosis of wind turbine based on CNN-LSTM parallel network model,” Acad. J. Eng. Technol. Sci., vol. 6, no. 12, pp. 115–121, 2023, doi: 10.25236/AJETS.2023.061218.
[27] P. A. Buestán-Andrade, M. Santos, J. E. Sierra-García, and J. P. Pazmiño-Piedra, “Comparison of LSTM, GRU and transformer neural network architecture for prediction of wind turbine variables,” in Int. Conf. Soft Comput. Models Ind. Environ. Appl., Cham, Switzerland: Springer, 2023, pp. 334–343, doi: 10.1007/978-3-031-44064-8_31.
[28] M. Kavitha, “Enhancing security and privacy in reconfigurable computing: Challenges and methods,” SCCTS Trans. Reconfigurable Comput., vol. 1, no. 1, pp. 16–20, 2024, doi: 10.31838/RCC/01.01.04.
[29] Ö. A. Karaman, “Prediction of wind power with machine learning models,” Appl. Sci., vol. 13, no. 20, p. 11455, 2023, doi: 10.3390/app132011455.
[30] A. Lawal, S. Rehman, L. M. Alhems, and M. M. Alam, “Wind speed prediction using hybrid 1D CNN and BLSTM network,” IEEE Access, vol. 9, pp. 156672–156679, 2021, doi: 10.1109/ACCESS.2021.3129812.
[31] F. Shahid, A. Zameer, A. Mehmood, and M. A. Z. Raja, “A novel WaveNets long short term memory paradigm for wind power prediction,” Appl. Energy, vol. 269, p. 115098, 2020, doi: 10.1016/j.apenergy.2020.115098.
[32] J. Wang, Y. Qian, L. Zhang, K. Wang, and H. Zhang, “A novel wind power forecasting system integrating time series refining, nonlinear multi-objective optimized deep learning and linear error correction,” Energy Convers. Manag., vol. 299, p. 117818, 2024, doi: 10.1016/j.enconman.2024.117818.
[33] F. Shahid, A. Zameer, and M. Muneeb, “A novel genetic LSTM model for wind power forecast,” Energy, vol. 223, p. 120069, 2021, doi: 10.1016/j.energy.2021.120069.
[34] Z. Zhao et al., “Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features,” Eng. Appl. Artif. Intell., vol. 121, p. 105982, 2023, doi: 10.1016/j.engappai.2023.105982.
[35] L. Xiang, J. Liu, X. Yang, A. Hu, and H. Su, “Ultra-short term wind power prediction applying a novel model named SATCN-LSTM,” Energy Convers. Manag., vol. 252, p. 115036, 2022, doi: 10.1016/j.enconman.2021.115036.
[36] M. Kavitha, “Energy-efficient algorithms for machine learning on embedded systems,” J. Integr. VLSI, Embedded Comput. Technol., vol. 1, no. 1, pp. 16–20, 2024, doi: 10.31838/JIVCT/01.01.04.
[37] A. Kisvari, Z. Lin, and X. Liu, “Wind power forecasting – A data-driven method along with gated recurrent neural network,” Renew. Energy, vol. 163, pp. 1895–1909, 2021, doi: 10.1016/j.renene.2020.10.099.
[38] E. T. Habtemariam, K. Kekeba, M. Martínez-Ballesteros, and F. Martínez-Álvarez, “A Bayesian optimization-based LSTM model for wind power forecasting in the Adama district, Ethiopia,” Energies, vol. 16, no. 5, p. 2317, 2023, doi: 10.3390/en16052317.
[39] O. Abedinia et al., “Wind power forecasting enhancement utilizing adaptive quantile function and CNN-LSTM: A probabilistic approach,” IEEE Trans. Ind. Appl., vol. 60, no. 3, pp. 4446–4457, 2024, doi: 10.1109/TIA.2024.3350945.
[40] S. M. Malakouti et al., “Advanced techniques for wind energy production forecasting: Leveraging multi-layer perceptron + Bayesian optimization, ensemble learning, and CNN-LSTM models,” Case Stud. Chem. Environ. Eng., vol. 10, p. 100881, 2024, doi: 10.1016/j.cscee.2024.100881.
[41] C. Xie, X. Yang, T. Chen, Q. Fang, J. Wang, and Y. Shen, “Short-term wind power prediction framework using numerical weather predictions and residual convolutional long short-term memory attention network,” Eng. Appl. Artif. Intell., vol. 133, p. 108543, 2024, doi: 10.1016/j.engappai.2023.108543.