Vol. 08, No. 1 (2026) 181-188, doi: 10.24874/PES08.01.019

**Proceedings on Engineering Sciences**

www.pesjournal.net

# A FAULT DETECTION MODEL FOR PREDICTIVE MAINTENANCE OF A WIND TURBINE BASED ON A HYBRID DEEP LEARNING APPROACH

Noumich Farouk [1]
Abouchabaka Jaafar
Amrani Ayoub

Received 19.06.2025.
Revised 17.09.2025.
Accepted 30.09.2025.

Keywords:

*Industry 4.0, Predictive Maintenance, Wind Turbine, Deep Learning Networks Fault Detection, Genetic Algorithms (GA)*

A B S T R A C T

*In the context of Industry 4.0, predictive maintenance of wind turbines is a crucial issue for optimizing their performance. This article proposes a new hybrid deep learning approach, aligned with the principles of Industry 4.0, to improve the predictive detection of failures in wind turbines. By combining long short-term recurrent neural networks (LSTM), artificial neural networks (ANN), and genetic algorithms (GA), our hybrid model (GA-LSTM-ANN) effectively captures the temporal and nonlinear complexities of sensor data. Genetic algorithms automatically optimize the model configuration. This innovative approach offers a unique solution for predictive fault detection. Applied to a real dataset, our model has significantly outperformed traditional methods, achieving 96.32% precision, 95.91% accuracy, 96.45% F1-score, and 96.41% recall for wind turbine fault detection. These promising results open up new perspectives for optimizing the management of wind farms and contributing to the energy transition.*



## 1. INTRODUCTION

The growing global demand for renewable energies has propelled the wind power sector to the forefront. However, the complexity and cost of wind power systems necessitate the optimization of their operation. Unanticipated failures, particularly in gearboxes, bearings, and blades, result in substantial maintenance expenses and considerable production losses (Guan et al., 2022; Pang, 2023).

Traditional maintenance methodologies, predicated on periodic inspections, exhibit limitations in terms of cost-effectiveness and responsiveness to rapidly evolving operating conditions. To address these constraints, predictive maintenance strategies, founded on data analysis and artificial intelligence, have emerged as prominent solutions. These approaches enable proactive identification of potential faults, thereby reducing operational costs and downtime while maximizing energy production. Furthermore, remote monitoring enhances operator safety and facilitates informed decision-making processes (Al-Aomar, Almazroui, &

[1] Corresponding author: Noumich Farouk
Email: farouk.noumich@uit.ac.ma

Aljeneibi, 2016). The detection of faults in wind turbines represents a major challenge due to the inherent complexities of these systems. The intrinsically nonlinear characteristics of wind turbines, combined with the uncertainty of environmental disturbances and the temporal dependence of the data, make the implementation of effective detection methods particularly complex (Jiang, He, Xie, & Yan, 2018).

Faced with these challenges, various approaches have been explored. Although artificial neural networks and support vector machines have been widely used to diagnose wind turbine failures (Li, Liu, & Shu, 2018), these methods often rely on manual feature engineering, making them prone to trial and error. Moreover, they struggle to manage the complexity and variability of real-world data. Moreover, Fourier analysis techniques, well-suited for periodic signals, show their limitations when faced with non-stationary and noisy data characteristic of wind systems (De Novaes Pires Leite et al., 2021). Univariate analysis, on the other hand, fails to capture the complex interactions between different variables and proves insufficient for a complete defect detection (Arias Velásquez, 2024). Finally, the uncertainties related to system parameters, aerodynamic nonlinearities, and measurement noise exacerbate the difficulty of the task (Tabatabaeipour, Bak, Odgaard, & Stoustrup, 2012).

To address these drawbacks, hybrid approaches, combining supervised and unsupervised learning, have been proposed to improve the robustness and accuracy of defect detection (Wilhelm, Reimann, Gauchel, & Mitschang, 2021). Deep learning, in particular, has demonstrated significant potential in its ability to automatically extract complex and nonlinear features, thereby improving the performance of machine fault detection and diagnosis systems (Saufi, Leong, Ahmad, & Lim, 2019). However, this approach requires large volumes of high-quality data and necessitates careful management of complex temporal sequences, as well as meticulous optimization of model architectures.

In this study, we present a new hybrid approach for fault detection in wind turbines, combining robustness and precision. By integrating the strengths of long short-term memory (LSTM) networks, artificial neural networks (ANNs), and genetic algorithms (GAs), we propose an innovative model capable of capturing the complexities of wind turbine data and automatically optimizing its parameters. LSTMs, thanks to their ability to retain information over long temporal sequences, are particularly well-suited for modeling the complex data of wind turbines. Artificial neural networks, on the other hand, excel at automatically extracting relevant features, even from noisy and nonlinear data. Finally, genetic algorithms allow for precise optimization of the model's hyper parameters, ensuring optimal performance. This synergistic combination overcomes the limitations of traditional methods, particularly in terms of accuracy and generalization. Our model thus contributes to improving the reliability of wind turbines and optimizing maintenance operations, supporting the energy transition towards more sustainable and efficient solutions.

The article is organized as follows: Section 2 provides a critical review of existing defect detection methods, highlighting their strengths and weaknesses. Section 3 presents in detail our innovative hybrid model. Section 4 presents the experimental results, comparing the performance of our model to that of traditional methods. Finally, section 5 concludes the article.

## 2. RELATED WORKS

The maintenance of wind turbines represents a major challenge for the renewable energy industry. Inadequate maintenance can lead to considerable financial losses, unexpected production shutdowns, and compromised safety of the installations. To overcome these problems, predictive maintenance has emerged as an innovative solution. By leveraging data analysis, this approach allows for anticipating failures and optimizing interventions, thereby contributing to maximizing the availability and profitability of wind farms (Garan, Kovalenko, & Tidriri, 2022). This section of the article examines previous work dedicated to fault detection in the context of predictive maintenance of wind turbines. We review the significant advancements, particularly the improvements made to deep learning algorithms for wind turbine maintenance. Moreover, we discuss the current challenges and future prospects in this constantly evolving field.

In recent years, numerous studies have explored the application of machine learning and deep learning methods to fault detection in wind turbines. These approaches have been widely applied to various subsystems of wind turbines, such as gearboxes, generators, and power transmission systems (Snitchler, King, Winn, & Gamble, 2011).

Several studies have focused on the use of long short-term memory (LSTM) neural networks for early fault detection, due to their ability to process time series data. For example, a study demonstrated the ability of these networks to predict main bearing failures up to 4 months in advance (Vidal, Puruncajas, Tutivén, & Castellani, 2023). In the same direction, (Yang, Teng, Li, & Feng, 2020) used LSTM networks to accurately predict time series data from operating wind turbines and effectively detect failures, potentially reducing downtime and repair costs.

Furthermore, research has been conducted on the optimization of hyperparameters in LSTM networks. (Chen, Rao, Feng, & Zuo, 2022) propose a new approach for hyperparameter selection, based on physical principles, to improve the detection of gearbox faults. Although promising, this approach, which minimizes the average validation error, does not optimally differentiate between defective and healthy states, thus revealing potential for improving detectability. Moreover, since the results were obtained in a laboratory environment, their generalization to real-world conditions remains to be confirmed.

Other approaches have also been explored for fault detection in wind turbines. De Novaes Pires Leite et al., (2021) propose an innovative methodology to assess the health of wind turbines, using quantifiers derived from information theory. This approach, based on the analysis of temporal vibration waveforms, addresses the limitations of Fourier analysis when applied to the non-stationary and noisy signals characteristic of wind turbines. Although promising, this methodology also presents challenges related to the complexity of wind turbine signals.

In parallel, Selvaraj and Selvaraj (2021) examine the application of Optimized Artificial Neural Networks (OANN) for small wind turbines and highlight the potential of neural networks to identify complex patterns in sensor data. Amini, Kanfoud and Gan (2022) conducted an in-depth comparative study of different artificial neural network (ANN) architectures, highlighting the superior performance of Model 2 in predicting generator failures. Similarly, Laguna et al. (2020) developed a predictive maintenance tool based on artificial neural networks for gearboxes, a key component in wind turbines where significant energy losses are concentrated.

Furthermore, Riquelme, Fuentes and Chavez (2020) review the various synthetic inertia techniques for wind turbines and highlight their limitations. In particular, the authors highlight the modeling difficulties, such as managing non-linearities and the absence of a clear objective function, which makes it difficult to optimize the model parameters. These limitations make it impossible to achieve optimal performance and adequate representativeness of the model.

Nunes, Morais and Sardinha, (2021) provide a comprehensive review of the most promising machine learning methods for predictive maintenance of wind turbines. The authors highlight the central role of artificial neural networks, including CNN, LSTM, and SVM, in fault detection. They also emphasize the crucial importance of the quality of SCADA data and the implementation of rigorous preprocessing to optimize model performance.

Current trends in predictive maintenance research for wind turbines include the development of hybrid deep learning models combining different neural network architectures. Farsoni, Castaldi and Simani (2021) propose a hybrid approach to detect failures in wind turbines, with the aim of improving the reliability and availability of these energy systems. The researchers developed a piecewise affine model to capture the nonlinear characteristics and mitigate measurement uncertainties. Moreover, they implemented a noise reduction scheme to address the effects of inertial measurements associated with wind speed and turbulence. In a related study, Arias Velásquez (2024) proposes a multi-step machine learning approach, combining models such as Random Forest, XGBoost, and logistic regression, to improve the early detection of bearing failures in wind turbine generators. This method, surpassing traditional single-variable approaches, achieves high classification accuracy. However, the authors acknowledge that other modes of failure, such as lubrication or contamination issues, are not fully addressed by this approach and require further investigation.

Furthermore, Wang and Yin (2023) propose an innovative approach to detect complex failures in wind turbines using a hybrid 3DSE-CNN-2DLSTM model. This method, leveraging SCADA data, surpasses traditional techniques by enabling the detection of fault combinations. By integrating spatial and temporal correlations, this model offers improved diagnostic accuracy. However, since the evaluation is based on a single dataset, additional validation on different wind farms would be necessary to generalize the results.

The application of deep learning techniques to the predictive maintenance of wind turbines has revealed considerable potential in previous studies. Nevertheless, these models face major challenges, including inherent temporal variability and the pervasive noise in wind data, as well as the complexities related to hyper parameter optimization. To overcome these limitations, this study proposes a new hybrid GA-LSTM-ANN model. Long Short-Term Memory (LSTM) networks excel at capturing complex long-term temporal correlations, making them particularly well suited for modeling sequential data from wind turbines. Artificial neural networks (ANN) complement this approach by offering exceptional flexibility in representing nonlinear relationships between variables and demonstrating resilience to noise, thereby improving the model's generalization capabilities. Moreover, the integration of genetic algorithms ensures a global optimization of hyper parameters, allowing for an optimal model configuration. This innovative hybrid methodology significantly improves the robustness and accuracy of anomaly detection and fault prediction, even when applied to complex and dynamic wind systems.

## 3. METHODOLOGY

In order to improve the reliability of wind farms, a new hybrid deep learning approach for failure prediction has been developed, combining the strengths of long short-term memory (LSTM) recurrent neural networks and artificial neural networks (ANN). Genetic algorithms (GA) have been implemented to optimize the hyperparameters of this hybrid model, allowing it to achieve optimal performance. The methodology is based on real data from a Portuguese wind farm. After a rigorous data reprocessing, including data normalization, handling class imbalances, and selecting discriminative features, the data were split into training and test sets. The GA-LSTM-ANN model was then trained on the training set before being evaluated on the test set using classification metrics such as accuracy, precision, recall, and F1 score. Figure 1 presents a schematic representation of this methodology.

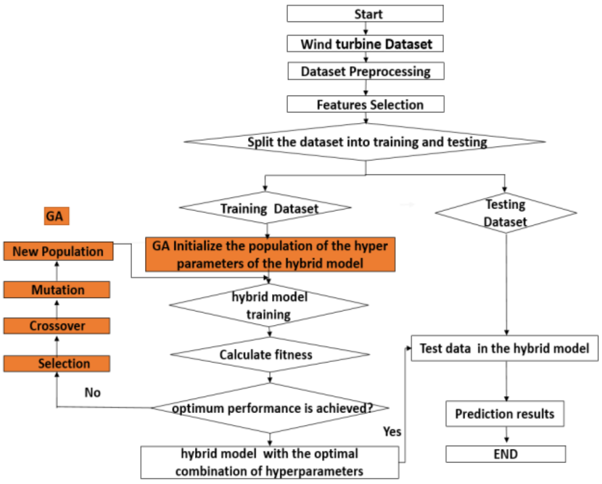


**Figure 1.** Overall process diagram of the suggested GA-LSTM-ANN method

### 3.1. Dataset

To evaluate the effectiveness of our methodology, we used a subset of data from the onshore wind farm "Park A" located in Portugal (Gück, Roelofs, & Faulstich, 2024). This dataset includes 22 complete time series, each representing an operational period of one year for one of the five wind turbines in the park. The data were collected at 10-minute intervals, capturing a range of operational parameters. The anomaly departures were identified using the EDP failure (EDP Inovação. 2018.), which provides the start timestamps for each fault, classified as unplanned maintenance incidents. For each turbine, the corresponding SCADA dataset includes a full year of data for model training, as well as an additional 4 to 98 days of data for forecasting.

Our analysis focused on the data from the turbine "asset_id" = 0. These data include 86 features and cover the period from August 3, 2021, to August 23, 2022, totaling 54,986 records. Each record is identified by a row ID and an accurate timestamp. In addition to the characteristics of the sensor data, each time series includes a unique asset identifier identifying the turbine, a "train-test" column indicating whether the record belongs to the training or test data, and a status identifier specifying the state of the turbine.

In order to categorize the different states of the turbine, we have defined two main statuses: "anomaly" and "normal." The "anomaly" status includes "status-type-id" 3 and 5, corresponding respectively to a maintenance (service) state and other unusual operational events. The "normal" status ("status-type-id" = 0) indicates optimal turbine operation, meaning energy production without incident. These status labels allow determining whether a specific data point corresponds to normal or abnormal behavior of the turbine. Table 1 provides a detailed summary of all the "status-type-ids," including their description and classification (normal or abnormal).

**Table 1.** Status descriptions (Gück, Roelofs, & Faulstich, 2024)

| Status ID | Description | Approved as Normal |
|---|---|---|
| 0 | Normal operation without limitations | True |
| 3 | Asset is in service mode/service team is at the site | False |
| 5 | Other operational states | False |

The raw data underwent a multi-step reprocessing procedure. First, the outliers were identified and removed. Next, min-max normalization was applied to ensure that all features have a comparable scale. Feature selection was performed using Pearson correlation to identify the most relevant features. We calculated the correlation between each numerical feature and the "status-type-id". The 18 features with an absolute correlation greater than 0.5 were selected. Figure 2 illustrates the results of this selection, presenting a heat map that highlights the strong correlations between the input features and the target variable. To address the class imbalance, SMOTE was applied to generate synthetic samples for the minority class ("anomaly"). Finally, the data was split into training and test sets to evaluate the performance of the hybrid model

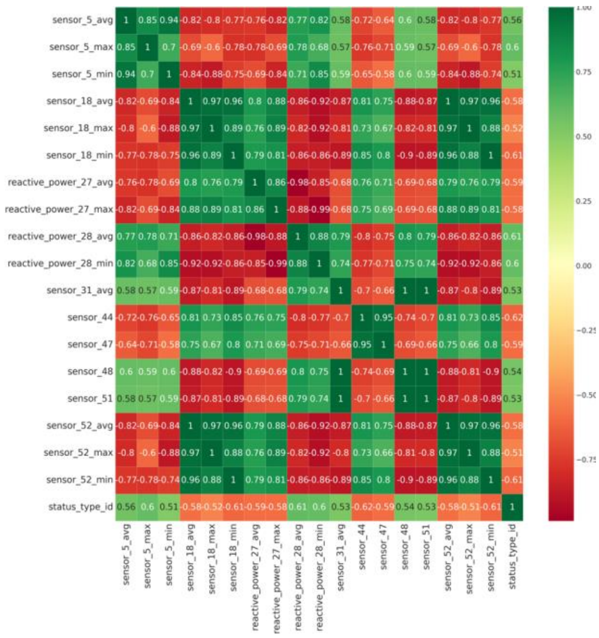


**Figure 2.** Correlation Matrix

### 3.2. Proposed Model

In this study, we propose an innovative hybrid model to predict wind turbine failures, combining the strengths of recurrent neural networks (LSTM), artificial neural networks (ANN), and genetic algorithms (GA). LSTMs excel at capturing the temporal dependencies inherent in wind data, while ANNs model the complex relationships between variables. GA optimize the model architecture, ensuring maximum performance.

Our hybrid model (Figure 3) is structured in four layers: an input layer, followed by a first LSTM hidden layer for temporal feature extraction. The extracted features are then transmitted to a second hidden layer of the artificial neural network (ANN) where they are combined and transformed to reveal more complex and nonlinear relationships. Finally, an output layer of the ANN generates the final prediction. Genetic algorithms optimize the entire model by selecting the most relevant features and adjusting the hyperparameters for optimal accuracy.

The GA-LSTM-ANN model is trained using the fit () method. The SCADA data from the wind turbine, previously selected based on a correlation analysis, is provided as input. The LSTM layer, designed to handle temporal sequences, exploits the sequential nature of the SCADA data. Each LSTM cell captures long-term temporal dependencies by sequentially updating its internal state based on the current input and its previous state. This mechanism allows the LSTM to memorize relevant information and forget those that have become obsolete. At the output of the LSTM layer, a feature vector is obtained. This vector represents a compact representation of the input sequence, encapsulating the learned temporal information. This vector is then transmitted to a fully connected artificial neural network (ANN). The artificial neural network (ANN) learns higher-level representations by combining the temporal features extracted by the LSTM with other potentially present information in the data. Finally, an output layer of the artificial neural network (ANN) performs a binary classification task, associating the learned representations with the two predefined classes: "normal" and "anomaly."

Inspired by nature, John Holland opened new perspectives in the field of optimization with genetic algorithms, an approach inspired by the mechanisms of natural evolution. These algorithms aim to find optimal or near-optimal solutions to complex problems by simulating biological processes such as crossover, mutation, and selection to identify the most effective combinations. A population of individuals, each representing a potential solution, evolves over generations thanks to these operations. The quality of an individual is evaluated using an objective function, which measures its ability to solve the problem. This evaluation allows for the selection of the best individuals to generate the population of the next generation, thereby fostering the emergence of increasingly effective solutions.

In our study, a genetic algorithm (GA) is used to adaptively determine the optimal size of the window, the number of neurons in the LSTM and ANN layers, as well as the batch size. Inspired by evolutionary mechanisms, this approach optimizes the model's ability to capture the complexities of temporal data and maximize prediction accuracy. The LSTM layer, on the other hand, is designed to leverage past information through its long-term memory cells. The choice of the sliding window size is crucial for capturing long-term temporal dependencies without introducing excessive noise. The GA optimally determines this size, finding a balance between the amount of past information to consider and the model's complexity. LSTM cells are responsible for managing the information in this window, determining which information should be remembered or forgotten, thus allowing the model to adapt to the different dynamics of the data. This approach enables the implementation of effective predictive maintenance for wind turbines

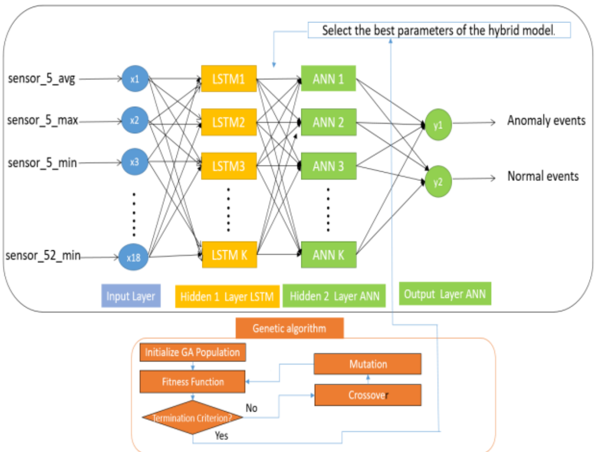


**Figure 3.** The organization of the proposed model GA-LSTM-ANN

### 3.3. Evaluation Metrics

To evaluate the performance of our models, we will use the classic classification metrics: precision, recall, F1 score, and accuracy. The equations for all the evaluation metrics were obtained from (Grandini et al., (2020)) and are briefly presented below:

- Precision: Proportion of correct positive predictions relative to all positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

- Accuracy: Proportion of correct predictions (positive and negative) relative to the total predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (2)$$

- Recall: Proportion of true positives correctly identified compared to all actual positives.

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

- F1-score: Harmonic mean of precision and recall, providing a balance between these two Measures.

$$\text{F1 score} = 2 * \frac{(Precision * Recall)}{Precision + Recall} \qquad (4)$$

## 4. RESULT AND DISCUSSION

In this section, we present the results obtained with our hybrid GA-LSTM-ANN model, designed for fault detection in wind turbines for predictive maintenance. This innovative model combines long and short-term recurrent neural networks (LSTM), artificial neural networks (ANN), and genetic algorithms (GA). Genetic algorithms (GA) are particularly well-suited for optimizing complex models with numerous parameters, as they allow for the efficient exploration of large search spaces and the discovery of optimal combinations.

Using a genetic optimization algorithm, we have determined the optimal configuration of our hybrid model. This includes a time window of 50 minutes (5), 154 neurons in the first LSTM hidden layer, and 68 neurons in the second ANN layer, as well as a batch size of 78 samples. These parameters were selected with the aim of maximizing the model's performance while minimizing its computational complexity. The detailed parameters of the genetic algorithm and the neural network are presented in Table 2.

Although an increase in the population size and the number of generations could potentially improve the convergence of the genetic algorithm, it also results in a significant increase in computation time. Therefore, we opted for a compromise between the accuracy of the results and computational complexity. The use of the genetic algorithm in this context allowed us to achieve optimal performance for our model.

The choice of these specific parameters is crucial to avoid both under fitting and overfitting. A number of neurons that is too low can prevent the model from capturing the complexity of the data, while a number that is too high can Lead to overfitting, where the model memorizes the noise present in the training data rather than learning the general characteristics. The size of the time window influences the amount of historical information used to make predictions. A window that is too short may not contain enough relevant information, while a window that is too long can make the model too sensitive to short-term variations. The batch size, on the other hand, affects the convergence speed of the optimization algorithm and the stability of the training.

**Table 2.** Main parameters to adjust in a GA-LSTM-ANN model

| Parameter | Component | Value |
|---|---|---|
| Population size | GA | 10 |
| Generation | GA | 5 |
| Crossover rate | GA | 0.4 |
| Mutation rate | GA | 0.1 |
| Number of neurons | LSTM | 154 |
| Number of neurons | ANN | 68 |
| Window size | LSTM-ANN | 5 |
| batch size | LSTM-ANN | 78 |

The results, presented in Table 3, demonstrate that our hybrid model significantly outperforms the classical LSTM and ANN models for fault detection in wind turbines. It achieves an accuracy of 95.91%, a recall of 96.41%, an F1 score of 96.45%, and a precision of 96.32%. These results are significantly superior to those obtained by the LSTM model (accuracy: 92.98%, precision: 93.45%, recall 92.91%, F1 score: 93.51%) and the ANN model (accuracy: 89.25%, precision: 92.04%, recall: 85.91%, F1 score: 88.88%). Figure 4 clearly illustrates this superiority by presenting the classification performance of each algorithm in a comparative bar chart. These strong performance indicators demonstrate the ability of the GA-LSTM-ANN model to accurately detect failure events while minimizing false positives and false negatives

**Table 3.** Comparison of performance metrics for the GA-LSTM-ANN, ANN, and LSTM models

| Metrics | ANN | LSTM | GA-LSTM-ANN |
|---|---|---|---|
| Accuracy | 89.25% | 92.98% | 95.91% |
| Precision | 92.04% | 93.45% | 96.32% |
| Recall | 85.91% | 92.91% | 96.41% |
| F1-score | 88.8% | 93.51% | 96.45% |

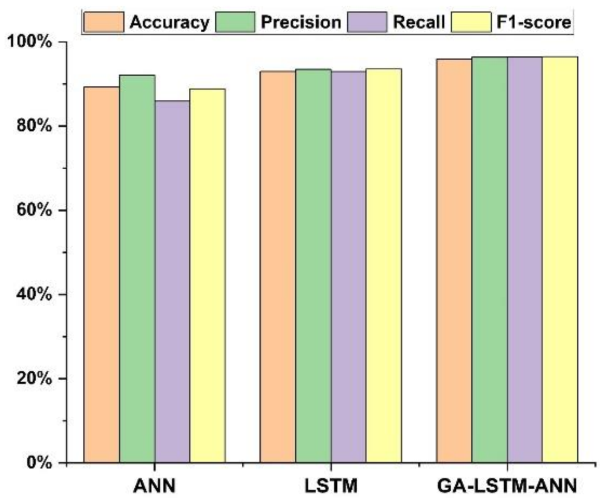


**Figure 4.** The total classification metrics of each model

The improvement in the performance of the GA-LSTM-ANN model is due to the optimization of hyperparameters by the genetic algorithm, which allows for better capture of data complexities and an improvement in the model's generalization. The use of shorter time windows, also optimized by the genetic algorithm, has contributed to increasing the accuracy of predictions, highlighting the crucial importance of data quality in this context.

The results of our study demonstrate that the GA-LSTM-ANN hybrid model constitutes a very effective solution for the predictive maintenance of wind turbines. This model has surpassed several state-of-the-art traditional methods in deep learning on a complex and nonlinear dataset. It thus reveals its ability to capture the temporal dependencies present in the multivariate sensor data. This increased accuracy allows for early detection of failures, which promotes proactive and optimized maintenance.

Although this study focused on the wind energy sector, the flexible architecture of the GA-LSTM-ANN model and its training methodology can be adapted to other industrial sectors. For example, the model could be applied to predictive maintenance of machine tools or transportation systems. Future work could explore new optimization strategies to further improve the model's performance, as well as assess its applicability to other complex systems.

## 4. CONCLUSION

This study presents a new hybrid approach, GA-LSTM-ANN, for predictive maintenance in Industry 4.0. By integrating long short-term memory networks (LSTM), artificial neural networks (ANN), and a genetic algorithm (GA) to optimize the window size, the number of neurons, and the batch size, we have developed an advanced methodology for detecting failures in wind turbines. The genetic algorithm, inspired by the principles of natural evolution, facilitated the identification of the optimal model configuration.

The results obtained on a real dataset from a Portuguese wind farm demonstrated the superior performance of our approach compared to traditional LSTM and ANN models. In fact, the GA-LSTM-ANN model had rates of 96.32% for precision, 96.41% for recall, 96.45% for accuracy, and 95.91% for F1 score. This shows a big improvement. The tests also confirmed the significant superiority of GA-LSTM-ANN, thus validating the genetic algorithm's contribution to optimizing the model architecture.

This research highlights the effectiveness of the GA-LSTM-ANN approach in capturing both the complexities and non-linearities within the multivariate data from wind turbine sensors. This innovative methodology, which leverages the combined strengths of LSTM, ANN, and genetic algorithm optimization, paves the way for improved predictive maintenance not only in the wind energy sector but also in other industrial fields.

## References:


Al-Aomar, R., Almazroui, S., & Aljeneibi, S. (2016). Reducing Operational Downtime in Service Processes: A Six Sigma Case Study. Institute of Electrical and Electronics Engineers. doi: 10.1109/icimsa.2016.7504029

Amini, A., Kanfoud, J., & Gan, T.-H. (2022). An Artificial Intelligence Neural Network Predictive Model for Anomaly Detection and Monitoring of Wind Turbines Using SCADA Data. *Applied Artificial Intelligence, 36*(1). doi: 10.1080/08839514.2022.2034718

Chen, Y., Rao, M., Feng, K., & Zuo, M. J. (2022). Physics-Informed LSTM hyperparameters selection for gearbox fault detection. *Mechanical Systems and Signal Processing, 171*, 108907. doi: 10.1016/j.ymssp.2022.108907

De Novaes Pires Leite, G., Da Cunha, G. T. M., Dos Santos Junior, J. G., Araújo, A. M., Rosas, P. A. C., Stosic, T., Stosic, B., & Rosso, O. A. (2021). Alternative fault detection and diagnostic using information theory quantifiers based on vibration time-waveforms from condition monitoring systems: Application to operational wind turbines. *Renewable Energy, 164*, 1183–1194. doi: 10.1016/j.renene.2020.10.129

EDP Inovação. (2018). EDPR Wind Farm Open Data: Wind Turbine SCADA Signals and Historical Failure Logbook from 2016 and 2017. EDP Inovação: Lisbon, Portugal.

Farsoni, S., Castaldi, P., & Simani, S. (2021). Fuzzy and Neural Network Approaches to Wind Turbine Fault Diagnosis. *Applied Sciences, 11*(11), 5035. doi: 10.3390/app11115035

Garan, M., Kovalenko, I., & Tidriri, K. (2022). A Data-Centric Machine Learning Methodology: Application on Predictive Maintenance of Wind Turbines. *Energies, 15*(3), 826. doi: 10.3390/en15030826

Grandini, M., Bagli, E., & Visani, G. (2020, August 13). Metrics for Multi-Class Classification: An Overview. [Cornell University]. doi: 10.48550/arxiv.2008.05756

Guan, B., Su, Z., Yu, Q., Li, Z., Feng, W., Yang, D., & Zhang, D. (2022). Monitoring the blades of a wind turbine by using videogrammetry. *Optics and Lasers in Engineering, 152*, 106901. doi:10.1016/j.optlaseng.2021.106901

Gück, C., Roelofs, C. M. A., & Faulstich, S. (2024). CARE to Compare: A Real-World Benchmark Dataset for Early Fault Detection in Wind Turbine Data. *Data, 9*(12), 138. doi: 10.3390/data9120138

Jiang, G., Xie, P., He, H., & Yan, J. (2018). Wind Turbine Fault Detection Using a Denoising Autoencoder With Temporal Information. *IEEE/ASME Transactions on Mechatronics, 23*(1), 89-100. doi:10.1109/TMECH.2017.2759301

Laguna, R., Mauiri, D., Geremias, N., Ezzatti, P., Gutierrez, A., Alfonso, J., & Chiruzzo, L. (2020). Early Detection of Gearbox Damage, Development of ANN Forecast Maintenance Tool. In 2020 IEEE 5th International Conference on Technologies for Distributed Learning (TDLA) (Vol. 135, pp. 1–6). doi: 10.1109/tdla47668.2020.9326244

Li, Y., Liu, S., & Shu, L. (2018). Wind turbine fault diagnosis based on Gaussian process classifiers applied to operational data. *Renewable Energy, 134,* 357–366. doi: 10.1016/j.renene.2018.10.088

Nunes, A. R., Morais, H., & Sardinha, A. (2021). Use of Learning Mechanisms to Improve the Condition Monitoring of Wind Turbine Generators: A Review. *Energies, 14*(21), 7129. doi: 10.3390/en14217129

Pang, J. L. (2023). Adaptive fault prediction and maintenance in production lines using deep learning. *International Journal of Simulation Modelling*, *22*(4), 734–745. https://doi.org/10.2507/IJSIMM22-4-CO20

Riquelme, E., Fuentes, C., & Chavez, H. (2020). A Review of Limitations of Wind Synthetic Inertia Methods. In 2020 IEEE 5th International Conference on Technologies for Distributed Learning (TDLA) (Vol. 9, pp. 1–6). doi: 10.1109/tdla47668.2020.9326180

Saufi, S. R., Leong, M. S., Ahmad, Z. A. B., & Lim, M. H. (2019). Challenges and Opportunities of Deep Learning Models for Machinery Fault Detection and Diagnosis: A Review. *IEEE Access, 7,* 122644–122662. doi: 10.1109/access.2019.2938227

Selvaraj, Y., & Selvaraj, C. (2021). Proactive maintenance of small wind turbines using IoT and machine learning models. International Journal of Green Energy, ahead-of-print(ahead-of-print), 463–475. doi:10.1080/15435075.2021.1930004.

Snitchler, G., King, C., Winn, P., & Gamble, B. (2011). 10 MW Class Superconductor Wind Turbine Generators. *IEEE Transactions on Applied Superconductivity, 21*(3), 1089–1092. doi: 10.1109/tasc.2010.2100341

Tabatabaeipour, S. M., Bak, T., Odgaard, P. F., & Stoustrup, J. (2012). Fault Detection of Wind Turbines with Uncertain Parameters: A Set-Membership Approach. *Energies, 5*(7), 2424–2448. doi: 10.3390/en5072424

Velásquez, R. M. A. (2024). Bearings faults and limits in wind turbine generators. *Results in Engineering, 21*, 101891. doi: 10.1016/j.rineng.2024.101891

Vidal, Y., Puruncajas, B., Tutivén, C., & Castellani, F. (2023). Predictive maintenance of wind turbine's main bearing using wind farm SCADA data and LSTM neural networks. *Journal of Physics: Conference Series, 2507*(1), 012024. doi: 10.1088/1742-6596/2507/1/012024

Wang, T., & Yin, L. (2023). A hybrid 3DSE-CNN-2DLSTM model for compound fault detection of wind turbines. *Expert Systems with Applications, 242*, 122776. doi: 10.1016/j.eswa.2023.122776

Wilhelm, Y., Reimann, P., Gauchel, W., & Mitschang, B. (2021). Overview on hybrid approaches to fault detection and diagnosis: Combining data-driven, physics-based and knowledge-based models. *Procedia CIRP, 99,* 278–283. doi: 10.1016/j.procir.2021.03.041

Yang, T., Teng, J., Li, C., & Feng, Y. (2020). Wind turbine fault detection and diagnosis using LSTM neural network (pp. 4042–4047). In 2020 China Conference on Communication and Control (CCC) (IEEE Cat. No. 9188709). doi: 10.23919/ccc50068.2020.9188709

**Noumich Farouk**
Faculty of Sciences, IbnTofail University,
Kenitra,
Morroco
farouk.noumich@uit.ac.ma
ORCID 0009-0009-9729-3447

**Abouchabaka Jaafar**
Faculty of Sciences, IbnTofail University,
Kenitra,
Morroco
jaafar.abouchabaka@uit.ac.ma
ORCID 0000-0003-3193-8416

**Amrani Ayoub**
Faculty of Sciences, IbnTofail University,
Kenitra,
Morroco
ayoub. amrani @uit.ac.ma
ORCID 0000-0001-6301-0976

.