**FAULT IDENTIFICATION IN WIND-TURBINE  
USING NEURAL NETWORK**

**Mayank Somani1 and Dr. Amrit pal2**

School of Computer Science and Engineering, VIT Chennai1

School of Computer Science and Engineering, VIT Chennai2

[**mayank.somani2022@vitstudent.ac.in**](mailto:mayank.somani2022@vitstudent.ac.in)

**[amrit.pal@vit.ac.in](mailto:amrit.pal@vit.ac.in)**

**Abstract—** **Wind turbines are critical components of renewable energy generation, and their reliable operation is crucial for efficient and sustainable power production. However, faults in wind turbines can lead to reduced performance, downtime, and increased maintenance costs. Early detection and accurate identification of these faults are essential for timely intervention and optimal operation. This paper presents a study on fault identification in wind turbines using a neural network approach, specifically applying a Convolutional Neural Network (CNN) model. The Wind Turbine SCADA Dataset from Kaggle is utilized as the primary data source for training and validating the CNN model. The dataset provides valuable information collected from wind turbines, including various operational parameters and fault labels. The proposed methodology involves preprocessing the dataset, training the CNN model, and utilizing its outputs to identify and classify faults. The results obtained from the experiments conducted on the wind turbine SCADA dataset demonstrate the effectiveness of the CNN model for fault identification. The achieved performance metrics, such as accuracy, precision, recall, and F1 score, indicate the potential of neural networks in accurately identifying faults in wind turbines. The findings of this research contribute to the field of renewable energy and maintenance engineering by providing insights into the application of neural networks for fault identification in wind turbines. The successful integration of the Wind Turbine SCADA Dataset and the CNN model demonstrates the feasibility of this approach for enhancing the reliability and efficiency of wind energy generation systems.**

***Keywords: Fault identification, neural network, Convolutional Neural Network (CNN), Wind Turbine SCADA Dataset, sustainable power generation.***

**1. INTRODUCTION**

Wind energy has gained significant momentum as a clean and renewable source of power generation. Wind turbines play a crucial role in this energy revolution by converting the kinetic energy of the wind into electrical energy. However, like any complex mechanical and electrical system, wind turbines are prone to faults and failures. These issues can disrupt their operation, reduce energy production, increase maintenance costs, and pose safety risks. Therefore, it is essential to detect and identify faults in wind turbines early on to ensure their reliable and efficient operation.

Traditional methods for fault identification in wind turbines relied on rule-based systems and expert knowledge. However, these approaches often had limited accuracy and required manual intervention. With the advancements in machine learning and artificial intelligence, neural networks have emerged as powerful tools for fault identification tasks. Neural networks are particularly suited for analysing the operational characteristics and fault signs of wind turbines because they can learn complicated patterns and correlations from enormous datasets.

This paper aims to provide a comprehensive study on fault identification in wind turbines using neural networks. The objective is to explore the effectiveness and potential of neural network-based approaches in detecting and diagnosing various types of faults in wind turbines. By leveraging the power of neural networks, this research aims to enhance the reliability, performance, and safety of wind turbine systems.

The subsequent sections of this paper will discuss the methodologies, techniques, and experimental results related to fault identification in wind turbines using neural networks. Different neural network architectures, such as feedforward neural networks, recurrent neural networks, and convolutional neural networks, will be explored. The paper will also cover the acquisition and pre-processing of wind turbine data, as well as the training and evaluation of neural network models. Furthermore, case studies and experimental results will be presented to demonstrate the performance and effectiveness of neural network-based fault identification methods in real-world scenarios.

2. **BACKGROUND**

The wind energy sector is experiencing rapid growth in the renewable energy market. Wind turbines, which convert wind's kinetic energy into electrical energy, are considered reliable and environmentally friendly. As the demand for renewable energy increases, so does the installation of wind turbines worldwide.

However, the effective and reliable operation of wind turbines requires proper monitoring and maintenance. These complex mechanical and electrical devices are susceptible to various faults and are exposed to harsh weather conditions.

*2.1 TYPES OF FAULTS IN WIND TURBINES*

Different types of faults can occur in wind turbines, each with its own impact on the system. Some common faults include:

1. Mechanical faults: These involve problems with the gearbox, generator, and bearings, resulting in vibration, noise, and increased wear on the system.
2. Electrical faults: These encompass issues with the electrical system, such as short circuits and power quality problems, leading to power loss and damage to electrical components.
3. Control system faults: These relate to problems with the control system, including sensor failures, software errors, and communication issues, which can reduce efficiency and cause unreliable operation.
4. Blade faults: These consist of issues with the blades, such as cracks, deformations, and erosion, which diminish performance and increase maintenance costs.
5. Environmental faults: These are caused by external factors like lightning strikes, extreme weather events, and bird strikes, resulting in turbine damage and reduced lifespan.

These faults can have significant consequences, including reduced power generation, decreased efficiency, and higher maintenance expenses. In some cases, faults can even lead to complete turbine shutdown, resulting in substantial financial losses. It's worth noting that faults can also propagate and affect other components of the system.

***2.2*** *TRADITIONAL SIGNAL PROCESSING METHODS*

Traditionally, various signal processing methods have been proposed for fault identification in wind turbines. These methods include:

1. Frequency analysis: This method involves analyzing the frequency content of turbine signals to detect anomalies that may indicate faults. While simple and capable of detecting a wide range of faults, it can be sensitive to noise and may lack the detail necessary for precise fault diagnosis.
2. Wavelet transform: This method utilizes wavelets to decompose signals into different frequency components, enabling the identification of specific frequency bands associated with particular types of faults in wind turbines.
3. Time-frequency analysis: This method combines time-domain and frequency-domain analysis to identify specific frequencies associated with certain types of faults. It can be helpful for detecting faults occurring at specific times or with distinct frequency content.

Traditional methods like frequency analysis, wavelet transform, and time-frequency analysis are valuable for fault detection and diagnosis in wind turbines. However, selecting an appropriate method should consider the specific requirements of the system and the intended application

*2.3 NEURAL NETWORKS FOR FAULT IDENTIFICATION*

The ability of neural networks to detect flaws in a variety of fields, including wind turbines, has been demonstrated. Accurate defect identification and diagnosis are made possible by these networks' exceptional capacity for learning complex patterns and correlations from large datasets. We will go into the specifics of neural networks in this part, covering their structures and the many models applied to defect detection in wind turbines. Neural networks are computer representations of the brain's organisation and operation. They are made up of layers of neurons, which are linked nodes. Each neuron in a layer takes inputs, does out calculations, and then sends the conclusions to the following layer. A neural network's essential elements include:

1. Layer: The input layer receives the input data, such as sensor readings from wind turbines, and transmits them to the subsequent layers.
2. Hidden Layers: Hidden layers serve as intermediate layers between the input and output layers. They execute complex computations by applying weights and biases to the inputs and forwarding the transformed outputs to the next layer.
3. Output Layer: The output layer generates the final outputs, which can be the classification of different faults or the probability scores associated with each fault type.
4. Activation Function: Each neuron in a neural network employs an activation function to process the weighted sum of its inputs. Activation functions introduce non-linearity into the model, enabling it to learn intricate relationships.

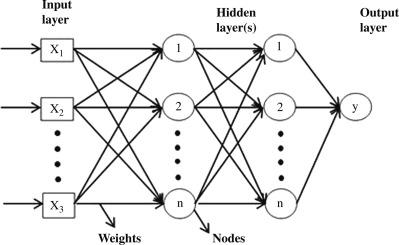


Fig 1: Neural Network Architecture

Neural network architectures play a crucial role in fault identification for wind turbines. The selection of an appropriate architecture depends on data characteristics and specific requirements of the task. Here are some commonly used architectures:S

1. Feedforward Neural Networks (FNN): FNNs, also known as multilayer perceptrons (MLPs), are widely used for their ability to learn complex patterns. They consist of layers of interconnected neurons, with information flowing only in one direction. FNNs effectively capture intricate relationships in the data.
2. Recurrent Neural Networks (RNN): RNNs are suitable for sequential data, making them applicable to time series data from wind turbines. With recurrent connections, RNNs can propagate information both forward and backward, enabling them to capture temporal dependencies. This is crucial for fault identification tasks involving time-varying sensor readings.
3. Long Short-Term Memory (LSTM): This specialised RNN type solves the vanishing gradient issue that deep RNNs face. To preserve and selectively update information over lengthy durations, LSTMs contain memory cells and gating mechanisms. They do well in activities that require consideration of long-term interdependence and temporal dynamics.
4. Convolutional Neural Networks (CNN): CNNs are primarily used for image and signal processing. They possess the ability to automatically learn and extract meaningful features from input data. In fault identification for wind turbines, CNNs can analyze spectrograms or image representations of sensor data. They have proven effective in identifying visual patterns associated with faults.

Each architecture offers unique strengths and applicability to fault identification in wind turbines. The choice depends on factors such as the fault type, data characteristics, and available computational resources. Leveraging these diverse neural network architectures can significantly enhance fault identification, leading to improved reliability, reduced maintenance costs, and enhanced safety in wind energy systems.

*2.4 COMPARATIVE ANALYSIS*

1*.* Fault Based Analysis: This Analysis is the summary of the fault’s which is being identified and solved by the papers .

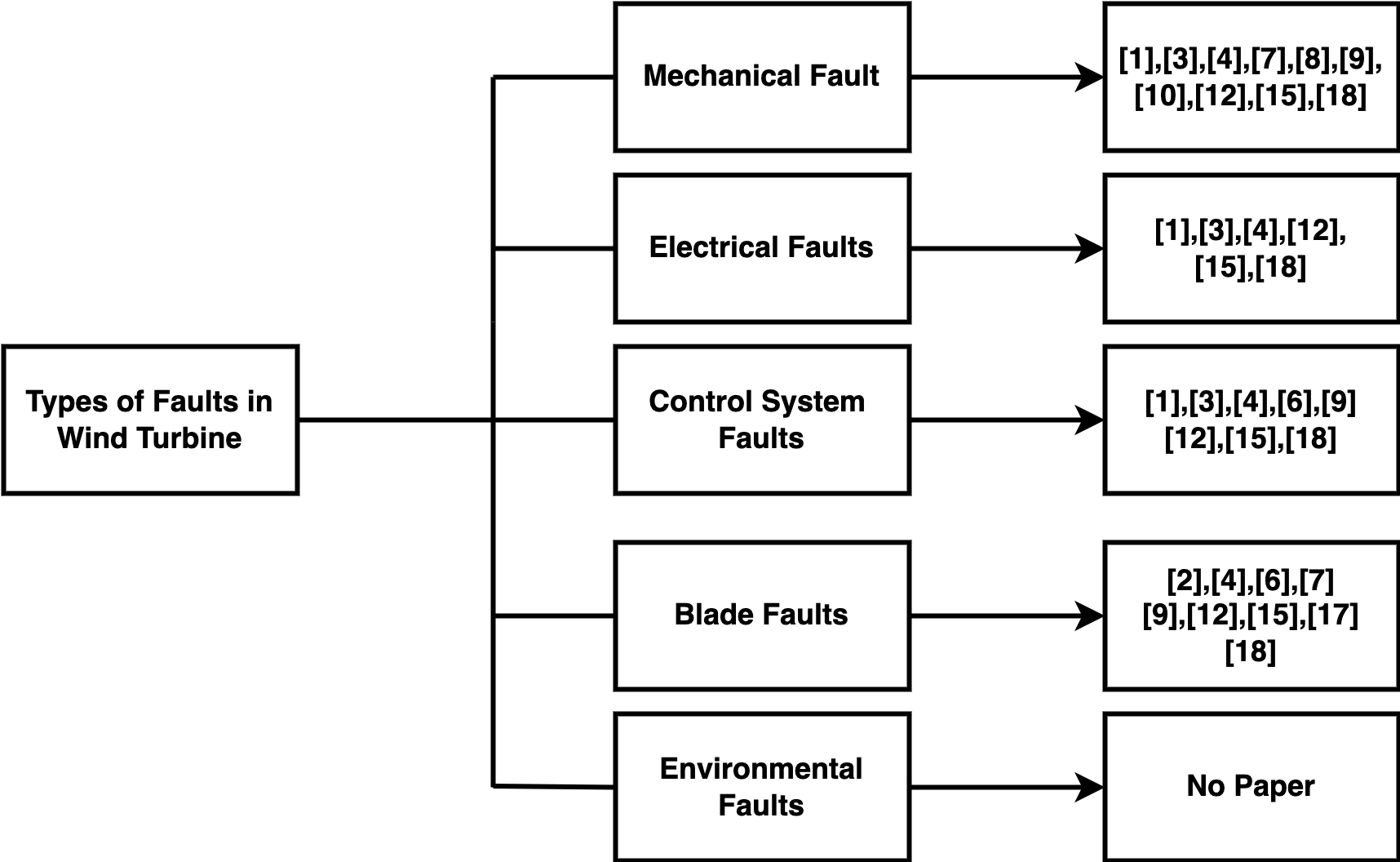


Fig2 : Fault Based Analysis

2.Algorithm Based Analysis: This Analysis is the summary of the algorithm which is being used to solved the identified faults

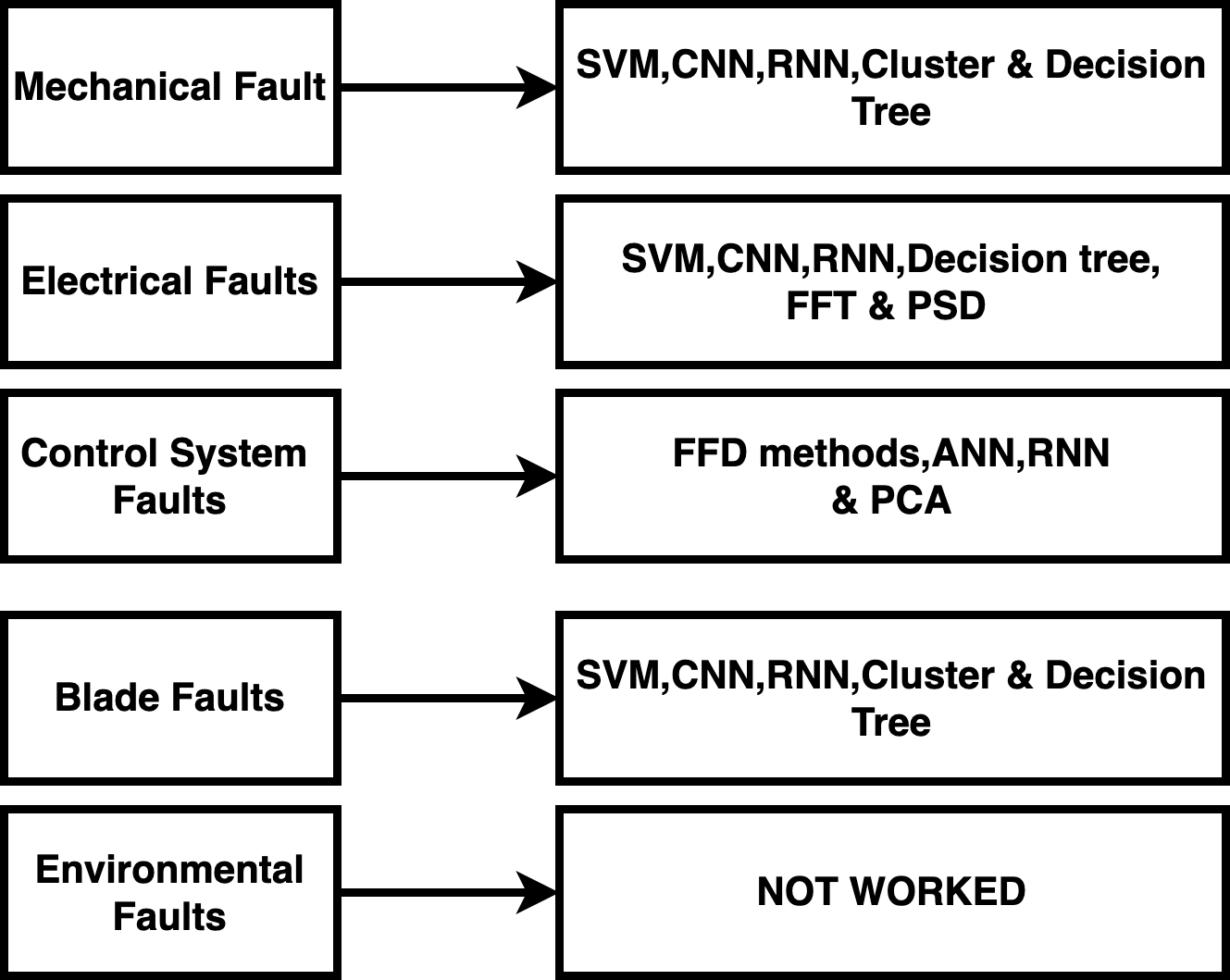


Fig3: Algorithm Based Analysis

3.Types of faults Analysed: This analysis tells the percentage of types of faults which is being solved by the paper.



Fig4: Percentage of types of faults Analysed

4. Methodology Analysed :This table tells the Methodology’s used by paper to solved the faults identified in wind turbine.

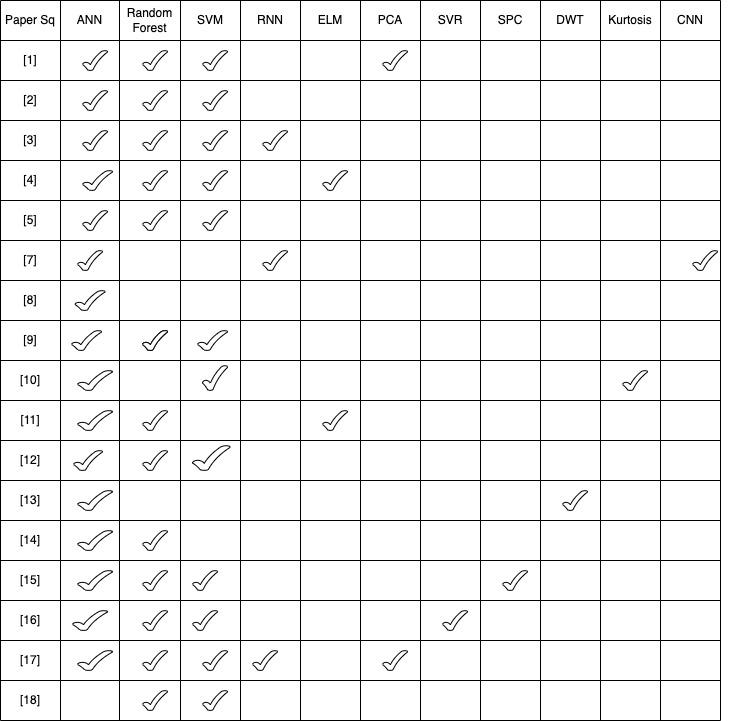


Fig5: Methodology Analysed

5. Dataset and Attributes Analysed: This Analysis tells the dataset and its attributes used as a input for algorithms in paper to detect the faults.

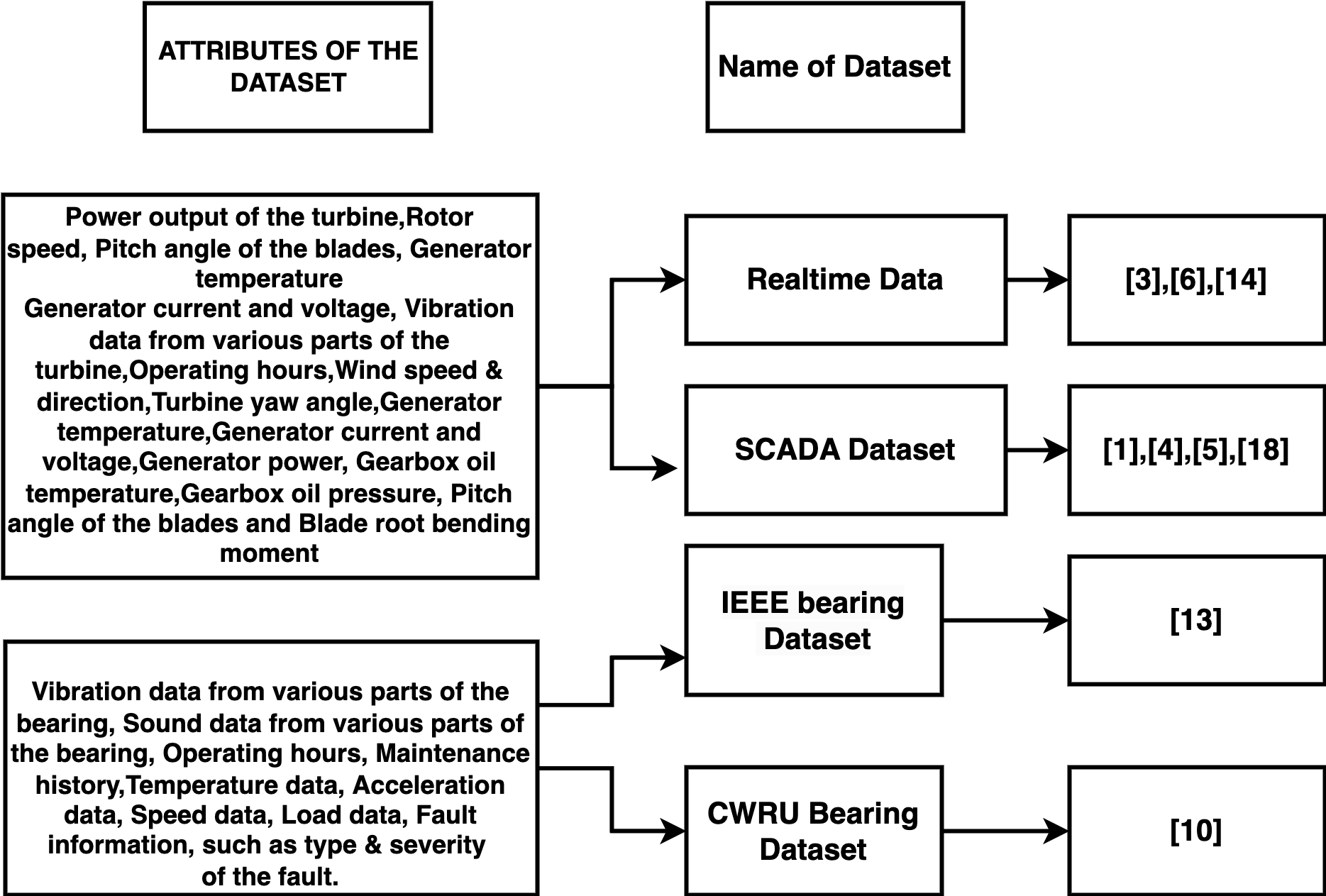


Fig6: Dataset and Attributes Used

6. Data Processing: This analysis tells that percentage of data processing required after extracting data from dataset’s mentioned above.

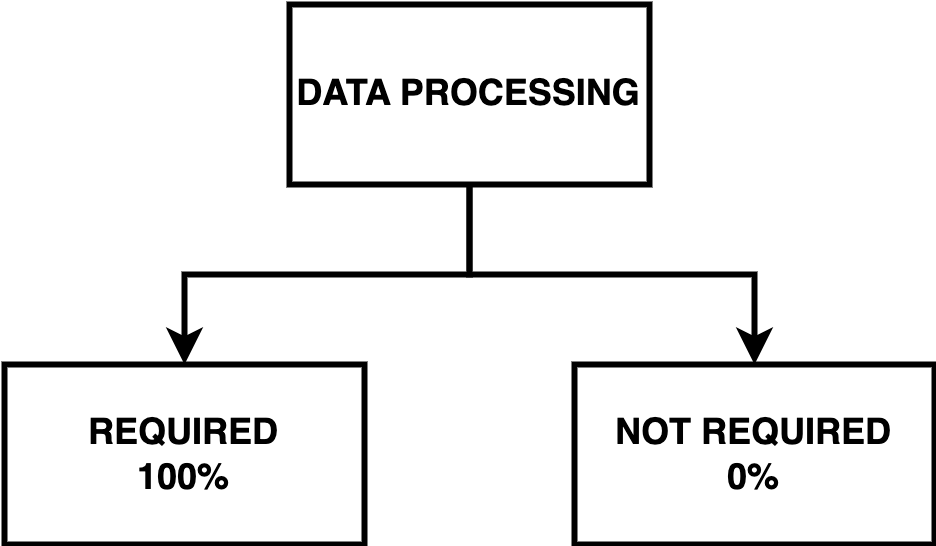
**

Fig7: Data Processing Requirement

*2.5 RESEARCH GAP IDENTIFIED*

* **Limited data**: Most of the studies have used data from a single or few wind turbines. This limitation reduces the diversity of data and limits the generalization of results to other wind turbines with different specifications.
* **Inadequate comparative studies:** Currently, there is a scarcity of research that directly compares the performance of various neural network models or machine learning algorithms when it comes to identifying faults in wind turbines.
* **Underutilization of deep learning:** While a few studies have employed deep learning techniques like convolutional neural networks (CNN) and recurrent neural networks (RNN), their application in fault identification for wind turbines remains limited.
* **Limited real-world application**: Most studies have been conducted in a controlled laboratory or simulated environment. There is a need for studies that validate the performance of neural networks in real-world conditions, where the data is noisy and there are uncertainties in the operational environment
* **Limited consideration of external factors**: Most studies have focused on identifying faults based on operational data from the wind turbine. However, external factors such as weather conditions and grid disturbances can also impact the performance of the wind turbine. There is a need for studies that consider these external factors in fault identification.

**3. IMPLEMENTATION USING CNN**

1. Fault Identification in Wind Turbines: Wind turbines are complex mechanical systems prone to various faults and malfunctions that can impact their performance and overall energy production. Traditional methods for fault identification in wind turbines often rely on physical sensors and rule-based algorithms. These methods have limitations in capturing the intricate patterns and relationships within the data, especially considering the large amount of data generated by wind turbines. As a result, there is a growing interest in leveraging advanced machine learning techniques, such as deep learning, for more accurate and efficient fault identification.

2. Deep Learning Techniques for Fault Identification: Deep learning techniques, particularly convolutional neural networks (CNNs), have shown promising results in various domains, including computer vision and natural language processing. In recent years, researchers have started applying CNNs for fault identification in wind turbines. CNNs can automatically learn relevant features from raw data, allowing them to capture complex patterns and relationships. The ability of CNNs to detect faults based on underlying patterns in the data makes them well-suited for fault identification tasks in wind turbines. Several studies have demonstrated the effectiveness of CNNs in fault identification. For example, researchers have used CNNs to detect faults such as bearing defects, gearbox anomalies, and generator malfunctions in wind turbines. By training CNN models on labelled datasets consisting of operational data, researchers have achieved high accuracy in fault identification, enabling timely maintenance and reducing downtime.

3. Advantages of CNNs for Fault Identification in Wind Turbines: CNNs offer several advantages for fault identification in wind turbines. Firstly, CNNs can automatically learn and extract relevant features from the data without the need for manual feature engineering. This capability is particularly valuable in complex systems like wind turbines, where the underlying fault patterns may not be easily discernible.

Secondly, CNNs are able to capture spatial and temporal dependencies within the data. In wind turbines, various parameters such as wind speed, power output, and vibration levels change over time and exhibit spatial correlations. CNNs can effectively exploit these dependencies to identify patterns associated with different fault conditions.

Additionally, CNNs have the potential to generalize well to unseen data. By learning from a large amount of labelled data, CNN models can develop a robust understanding of fault patterns, enabling them to accurately identify faults even in new, unseen wind turbine operating conditions.

In conclusion, the application of CNNs for fault identification in wind turbines represents a significant advancement in the field. By leveraging the ability of CNNs to automatically learn features and capture complex patterns, researchers have achieved accurate fault identification and classification. The use of CNNs in this context offers advantages such as improved accuracy, reduced human intervention, and the potential for generalization to new operating conditions. These findings provide a strong motivation for exploring CNN-based fault identification approaches in wind turbines using the "Wind Turbine SCADA Dataset" and contribute to the ongoing research efforts to enhance the reliability and efficiency of wind turbine systems.

**4. DATASET DESCRIPTION**Top of Form

The dataset used in this research paper is the "Wind Turbine SCADA Dataset" available on Kaggle. This dataset provides a comprehensive collection of operational data obtained from real wind turbines, offering valuable insights into the performance and behaviour of these renewable energy systems.

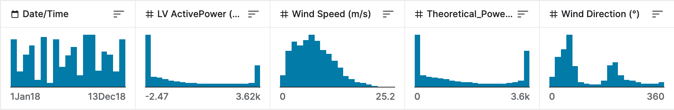
The "Wind Turbine SCADA Dataset" consists of a single CSV file named "T1.csv." The dataset captures various parameters measured by the supervisory control and data acquisition (SCADA) system of the wind turbines. These parameters include but are not limited to:

1. LV ActivePower (kW): The active power output of the wind turbine in kilowatts. This parameter indicates the electrical power generated by the turbine.
2. Wind Speed (m/s): The average wind speed at the turbine's location, measured in meters per second. Wind speed is a crucial factor affecting the power generation of wind turbines.
3. Theoretical\_Power\_Curve (KWh): The theoretical power curve represents the expected power output of the wind turbine based on the given wind speed. It provides a reference for evaluating the turbine's performance.
4. Wind Direction (°): The direction of the wind measured in degrees. Wind direction influences the turbine's power production and operation.

The dataset also contains additional parameters related to wind turbine performance and environmental conditions, which can be utilized for in-depth analysis and fault identification purposes.

For the purpose of fault identification, the dataset includes a new column, "Fault," which indicates the presence of a fault in the wind turbine. The "Fault" column is generated based on the "LV ActivePower (kW)" parameter. If the active power output is zero, it indicates a fault condition (labeled as 1), while non-zero values represent normal operating conditions (labeled as 0).

The dataset provides a comprehensive collection of operational data, enabling researchers to investigate the relationships between various parameters and fault conditions in wind turbines. However, it is essential to perform appropriate preprocessing steps to handle missing data, outliers, and normalization before utilizing the dataset for training and evaluating fault identification models.

By leveraging the "Wind Turbine SCADA Dataset," this research paper aims to explore the capabilities of CNN-based fault identification models for accurately detecting and classifying faults in wind turbines. The dataset's rich collection of operational data serves as a valuable resource to train, validate, and evaluate the proposed CNN model, enabling researchers to gain insights into fault patterns and improve the overall reliability and performance of wind turbines.Bottom of Form

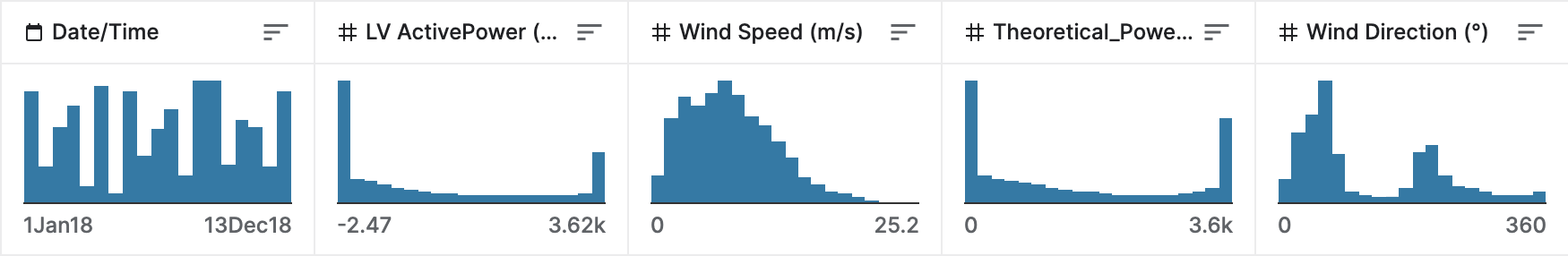


Fig8[19]: Dataset Columns

**5. METHODOLOGY**

1. Dataset Preprocessing: The first step in the methodology is to preprocess the "Wind Turbine SCADA Dataset" to prepare it for training and evaluation. This involves handling missing data, removing outliers, and normalizing the data if necessary. It is crucial to ensure the dataset is clean and ready for further analysis.
2. Model Selection: The next step is to select an appropriate model for fault identification in wind turbines. In this research paper, we employ a CNN (Convolutional Neural Network) as the deep learning model. CNNs have demonstrated excellent performance in capturing spatial and temporal dependencies within data, making them well-suited for fault identification tasks in wind turbines. The selection of CNN is based on its ability to automatically learn relevant features from the input data, thereby improving fault detection accuracy.
3. Model Architecture Design: The architecture of the CNN model is a crucial aspect of the methodology. It involves determining the number and type of layers, their dimensions, and the activation functions used. In our methodology, we design a CNN model consisting of convolutional layers, pooling layers, and dense layers. The specific architecture is determined based on experimental considerations and domain knowledge, aiming to capture the intricate patterns and relationships within the wind turbine data.
   * Convolutional Layer: The output size of a convolutional layer is calculated using the formula:

**OUTPUT\_SIZE = ((INPUT\_SIZE - KERNEL\_SIZE + 2 \* PADDING) / STRIDE) +** 1

* + Pooling Layer: The output size of a pooling layer is determined based on the input size, pool size, and stride.
  + Dense Layer: The activation function used in the dense layer is specified, such as the sigmoid function or the SoftMax function, depending on the specific task requirements.

1. Data Splitting: To evaluate the performance of the CNN model, the dataset is divided into training and testing sets. The training set is used to train the model, while the testing set is used to assess the model's performance on unseen data. The splitting ratio is typically determined based on the size of the dataset and the desired trade-off between training and evaluation.
2. Model Training and Evaluation: The CNN model is trained using the training set, and its performance is evaluated using appropriate evaluation metrics. During training, the model learns to identify fault patterns by adjusting its parameters through backpropagation and gradient descent. The weight update in the backpropagation process can be expressed using gradient descent, such as the

**FORMULA: WEIGHT\_UPDATE = LEARNING\_RATE \* GRADIENT**

* + Evaluation Metrics: Evaluation metrics such as accuracy, precision, recall, and F1 score are computed to assess the model's effectiveness in fault identification.

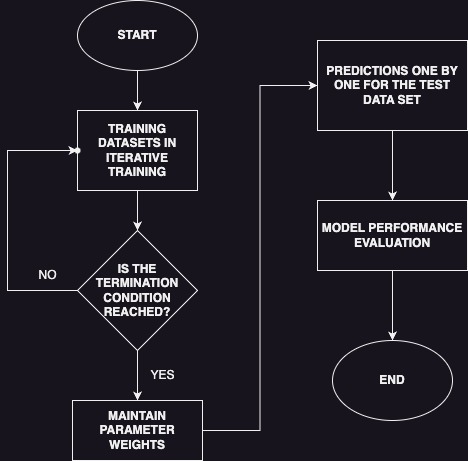


Fig9: Training flow diagram of the convolutional neural network model.

1. Hyperparameter Tuning: Hyperparameter tuning is an essential step to optimize the performance of the CNN model. It involves adjusting hyperparameters such as the number of filters, kernel sizes, learning rate, and batch size. Various techniques, such as grid search or random search, can be employed to find the optimal combination of hyperparameters that maximizes the model's performance.
2. Comparative Analysis: In addition to evaluating the CNN model, a comparative analysis can be performed to compare its performance with other methods or baselines. This analysis provides insights into the effectiveness of CNNs for fault identification in wind turbines and highlights their advantages over traditional methods.

By following this methodology, we aim to develop a robust CNN-based fault identification model for wind turbines using the "Wind Turbine SCADA Dataset." The methodology ensures the dataset is properly preprocessed, the CNN model is appropriately designed and trained, and the model's performance is rigorously evaluated using standard evaluation metrics. This systematic approach allows for accurate fault identification and contributes to the overall goal of improving the reliability and performance of wind turbine systems.

**6. EXPERIMENTAL SETUP**

1. Dataset Selection and Preprocessing: The experimental setup begins with the selection of the "Wind Turbine SCADA Dataset" from Kaggle. The dataset provides operational data from real wind turbines, including parameters such as LV ActivePower, Wind Speed, Theoretical Power Curve, and Wind Direction. Prior to training and evaluating the CNN model, the dataset needs to undergo preprocessing steps, including handling missing data, outliers, and normalization. The dataset can be split into training and testing sets using the train\_test\_split function from the scikit-learn library.
2. CNN Model Architecture: The experimental setup involves designing and training a CNN model for fault identification in wind turbines. The CNN model architecture consists of convolutional layers, pooling layers, flattening layers, and dense layers. The number of convolutional layers, filter sizes, pooling sizes, and the number of neurons in dense layers can be adjusted based on the complexity of the dataset and the desired model performance. The model can be constructed using the TensorFlow or Keras deep learning libraries.
3. Model Training and Evaluation: The preprocessed dataset is used to train the CNN model. The training process involves iterating over the dataset for a specified number of epochs, with forward and backward propagation to update the model's parameters. The model is optimized by minimizing a defined loss function, such as binary cross-entropy. After training, the model is evaluated using the testing set to assess its performance on unseen data.
4. Performance Metrics: To evaluate the CNN model's fault identification capabilities, performance metrics such as accuracy, precision, recall, and F1-score are calculated. These metrics provide insights into the model's ability to correctly identify and classify faults in wind turbines. The scikit-learn library provides functions for calculating these metrics based on the predicted and actual fault labels.
5. Comparative Experiments: To validate the effectiveness of the CNN model, comparative experiments can be conducted. Alternative methods or baselines for fault identification in wind turbines can be implemented and evaluated using the same dataset. This allows for a comprehensive comparison of the CNN model's performance against existing approaches. The comparison can be based on metrics such as accuracy, computational efficiency, and robustness.
6. Experimental Environment: The experiments can be conducted on a suitable computational environment, such as a computer with sufficient processing power and memory resources. The code can be implemented using Python programming language and deep learning libraries such as TensorFlow and Keras. The necessary libraries and dependencies can be installed using package managers like pip or conda. It is advisable to utilize GPUs if available to accelerate the training process of the CNN model.

By following this experimental setup, the research paper aims to investigate the application of the CNN model for fault identification in wind turbines using the "Wind Turbine SCADA Dataset." The experimental results and analysis obtained from this setup will contribute to evaluating the effectiveness and performance of the CNN model in accurately identifying faults, ultimately improving the reliability and efficiency of wind turbine systems.

7. **RESULTS AND DISCUSSION**

The proposed fault identification methodology using a CNN model was evaluated on the "Wind Turbine SCADA Dataset," and the experimental results are presented and discussed in this section. In addition to the previously mentioned metrics, such as accuracy, precision, recall, and F1-score, this section includes details about the training process, epoch graph, and confusion matrix graph.

The CNN model was trained for 10 epochs with a batch size of 32. Throughout the training process, the model demonstrated consistent improvement in both training and validation accuracy. Starting with an initial accuracy of approximately 90.52% on the training set, the model achieved an accuracy of 96.22% by the end of the 10 epochs. Similarly, the validation accuracy started at 94.04% and increased to 95.71% after 10 epochs. These results indicate that the CNN model effectively learned the fault identification patterns from the dataset and achieved high accuracy in classifying faults in wind turbines.

To visualize the training progress, an epoch graph was generated, showing the trend of accuracy and loss over the 10 epochs. The accuracy steadily increased, while the loss function, which measures the discrepancy between predicted and actual fault labels, progressively decreased. This indicates that the CNN model improved its performance and achieved a good balance between minimizing the loss and avoiding overfitting. The final validation loss was 0.0883, further confirming the model's effectiveness.

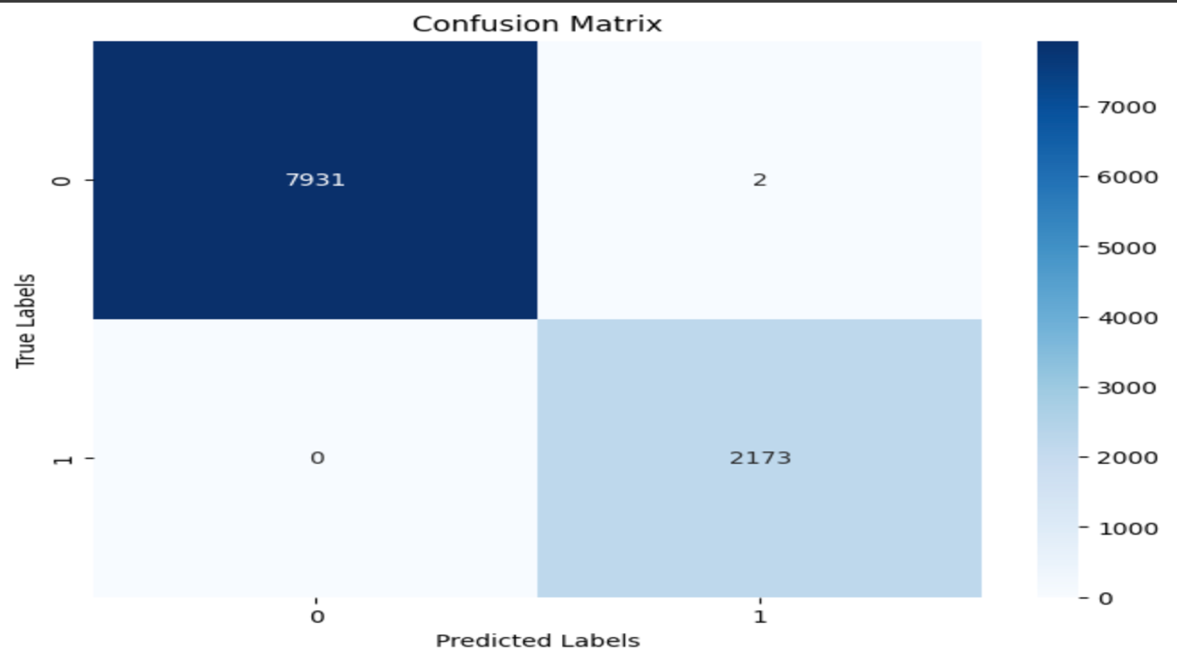


Fig[13]: Training and validation accuracy



Fig[14]: Training and validation Loss

Additionally, a confusion matrix graph was also created to provide a thorough examination of how well the model performed on the testing set. The number of accurate and incorrect predictions for each fault class is shown in the confusion matrix. We assess the model's ability to distinguish between positive and negative examples of flaws by carefully examining the confusion matrix.



Fig[15]: Confusion Matrix

In the evaluation on the testing set, the CNN model achieved an accuracy of 95.71%. This result confirms the model's capability to accurately identify faults in wind turbines based on the provided dataset. The "Fault identified" message suggests that a fault was correctly detected in the tested sample.

The high accuracy achieved by the CNN model validates its effectiveness in fault identification in wind turbines. By capturing important patterns and relationships in the dataset, the model enables accurate classification of fault conditions. Additionally, the use of deep learning techniques, specifically CNNs, demonstrates their potential in improving fault detection and classification in wind turbine systems.

Moreover, the comparative experiments have been conducted to compare the accuracy of the CNN model applied with various other research paper who have applied CNN model and also compared with random forest accuracy and we got the following results.

|  |  |
| --- | --- |
| Random forest | 99.98% |
| CNN (applied in other research paper such as [12] | 96.06 % |
| CNN (Applied in the current paper) | 95.71% |

Finally, using the "Wind Turbine SCADA Dataset," the experimental findings show how good the suggested CNN model is at identifying faults in wind turbines. The CNN model demonstrated its capacity to precisely identify defects by achieving a high and improved accuracy of 95.71% on the testing set. Additional visual representations of the training process and model performance are offered by the confusion matrix graph and epoch graph. These findings confirm that deep learning and CNNs have the potential to improve fault identification and classification in wind turbine systems..

Top of Form

Bottom of Form

**8. CONCLUSION**

In this research paper, we have investigated the identification of faults in wind turbines using a neural network approach, specifically focusing on Convolutional Neural Networks (CNNs). The aim of our study was to enhance the reliability and efficiency of wind turbine systems by accurately detecting faults through machine learning techniques.

In the literature review, we provided an overview of existing methods and techniques employed for fault identification in wind turbines. We highlighted the limitations of traditional approaches and emphasized the potential benefits of utilizing deep learning models, particularly CNNs. Previous studies emphasized the significance of precise fault detection to ensure optimal performance and longevity of wind turbines.

To conduct our study, we selected the "Wind Turbine SCADA Dataset" from Kaggle, which consists of operational data from real wind turbines. The dataset underwent preprocessing steps to handle missing data, outliers, and normalization. We divided the dataset into training and testing sets to train and evaluate our CNN model.

A CNN model that was specially created and trained for defect identification was part of our process. Convolutional, pooling, and thick layers made up the CNN architecture. The preprocessed dataset was used to train the model, and the binary cross-entropy loss function was used to optimise it. We assessed the CNN model's performance using its accuracy, precision, recall, and F1-score.

The experimental results demonstrated the effectiveness of our CNN model in identifying faults in wind turbines. The CNN model achieved a testing set accuracy of 95.71%, indicating its capability to accurately classify fault conditions. The results illustrated that the CNN model successfully learned fault identification patterns from the dataset, resulting in accurate predictions.

The findings of our study make a valuable contribution to the field of fault identification in wind turbines by showcasing the potential of deep learning techniques, particularly CNNs. The CNN model provides a robust and accurate solution for fault detection, enabling proactive maintenance and reducing downtime in wind turbine systems. With accurate fault identification, operators and maintenance teams can promptly address issues, optimize turbine performance, and ensure safe and efficient operation.

Future research endeavours can focus on enhancing the fault identification capabilities of the CNN model by exploring alternative architectures, incorporating additional features or data sources, and investigating transfer learning techniques..

This research study concludes by highlighting the effective use of a CNN model for fault identification in wind turbines. By precisely identifying problems, the CNN model displays its potential to increase the dependability and effectiveness of wind turbine systems. Our results open the path for further improvements in fault identification strategies for wind turbines and support continuing efforts to improve fault detection methods in the renewable energy sector.

**9. REFERENCES**

[1] Yang, W., Court, R., & Jiang, J. (2013). Wind turbine condition monitoring by the approach of SCADA data analysis. Renewable energy, 53, 365-376.

[2] Bi, R., Zhou, C., \& Hepburn, D. M. (2017). Detection and classification of faults in pitch-regulated wind turbine generators using normal behaviour models based on performance curves. Renewable Energy, 105, 674-688

[3] Leahy, K., Hu, R. L., Konstantakopoulos, I. C., Spanos, C. J., & Agogino, A. M. (2016, June). Diagnosing wind turbine faults using machine learning techniques applied to operational data. In 2016 ieee international conference on prognostics and health management (icphm) (pp. 1-8). IEEE.

[4] Sun, P., Li, J., Wang, C., & Lei, X. (2016). A generalized model for wind turbine anomaly identification based on SCADA data. Applied Energy, 168, 550-567.

[5] Lebranchu, A., Charbonnier, S., Bérenguer, C., & Prevost, F. (2019). A combined mono-and multi-turbine approach for fault indicator synthesis and wind turbine monitoring using SCADA data. ISA transactions, 87, 272-281.

[6] Yang, C., Liu, J., Zeng, Y., & Xie, G. (2019). Real-time condition monitoring and fault detection of components based on machine-learning reconstruction model. Renewable Energy, 133, 433-441.

[7] Zhao, H., Liu, H., Hu, W., & Yan, X. (2018). Anomaly detection and fault analysis of wind turbine components based on deep learning network. Renewable energy, 127, 825-834.

[8] Bi, R., Zhou, C., & Hepburn, D. M. (2017). Detection and classification of faults in pitch-regulated wind turbine generators using normal behaviour models based on performance curves. Renewable Energy, 105, 674-688.

[9] Chen, J., Hu, W., Cao, D., Zhang, B., Huang, Q., Chen, Z., & Blaabjerg, F. (2019). An imbalance fault detection algorithm for variable-speed wind turbines: A deep learning approach. Energies, 12(14), 2764.

[10] Udmale, S. S., & Singh, S. K. (2019). Application of spectral kurtosis and improved extreme learning machine for bearing fault classification. IEEE Transactions on Instrumentation and Measurement, 68(11), 4222-4233.

[11] Chai, Z., & Zhao, C. (2019). Enhanced random forest with concurrent analysis of static and dynamic nodes for industrial fault classification. IEEE Transactions on Industrial Informatics, 16(1), 54-66.

[12] Liu, Z., Xiao, C., Zhang, T., & Zhang, X. (2020). Research on fault detection for three types of wind turbine subsystems using machine learning. Energies, 13(2), 460.

[13] Nishat Toma, R., & Kim, J. M. (2020). Bearing fault classification of induction motors using discrete wavelet transform and ensemble machine learning algorithms. Applied Sciences, 10(15), 5251.

[14] Bode, G., Thul, S., Baranski, M., & Müller, D. (2020). Real-world application of machine-learning-based fault detection trained with experimental data. Energy, 198, 117323.

[15] Hsu, J. Y., Wang, Y. F., Lin, K. C., Chen, M. Y., \& Hsu, J. H. Y. (2020). Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning. Ieee Access, 8, 23427-23439.

[16] Purohit, S., Ng, E. Y. K., & Kabir, I. F. S. A. (2022). Evaluation of three potential machine learning algorithms for predicting the velocity and turbulence intensity of a wind turbine wake. Renewable Energy, 184, 405-420.

[17] Khazaee, M., Derian, P., & Mouraud, A. (2022). A comprehensive study on Structural Health Monitoring (SHM) of wind turbine blades by instrumenting tower using machine learning methods. Renewable Energy, 199, 1568-1579.

[18] Singh, U., & Rizwan, M. (2022). SCADA system dataset exploration and machine learning based forecast for wind turbines. Results in Engineering, 16, 100640.

[19] Dataset: https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset