Fault Detection in Wind Turbines

December 2021

1 Problem Statement and Motivation.

Wind turbines are primarily located in secluded areas to avoid wind interference near cities. Due to this allocation, a problem arises in increasing the costs of operation, maintenance, And the high costs of transportation and installation. The locations in which these wind turbines are installed, such as coastal areas and more commonly offshore, make them more prone to damage due to unpredictable weather conditions, thus demanding more frequent operation and maintenance (OM). Since the costs of transportation and installation are unavoidable, it seemed that the most efficient way to lower the costs of wind farms as a whole was to reduce the amount of OM needed for each wind turbine. As a counter to this problem, predict the faults in wind turbines before they occur and maintain them accordingly. That will allow the wind farms to operate sufficiently and reach maximum utilization, thus encouraging clean energy moving forward and hopefully reducing the pollution caused by fossil fuels[3].

2 Current Solutions

Research in this field has prevailed, where most traditional machine learning methods have already been studied and experimented with, such as SVM and random forest algorithms. More recently, research into deep-learning methods has been taken up to improve the results garnered by machine learning solutions. CNN, RNN, and LSTM have taken the lead in being the most promising techniques. Reihane Rahimilark[2] utilized two CNN models with one convolutional layer and two convolutional layers. In another paper by Jian Fu[1], CNN-RNN is used for feature extraction while LSTM takes this reduced data as input and gearbox bearing temperature as output. These researches have shown promising results compared to traditional models such as ARMA (Autoregressive moving average model) and DBN (Deep belief network) but still have space for improvement.

3 Sytem Operation in Real Life

On-site engineers will access the system through a web application that alerts them of impending failures and the expected timing of their occurrence to allow for timely maintenance, along with reports of wind power generation predictions.

4 High-level Architecture and Software Overview

The dataset used for this system is time-series; SCADA collected data from a wind farm in the West African Gulf of Guinea in 2016. It is divided into 3 main components, a data log containing the types of failures that might occur. The faults and their timings in this failure log are the third signal data consisting of 83 features, including RPM, Generator bearing temperature, and wind direction. The proposed system has a few primary steps: data acquisition and selection, processing, training model, and finally using a model on new data to detect faults and predict power. As shown in figure1, the forecasted data and the SCADA data are compared. When a significant difference between the forecasted and acquired data occurs, the data log is then checked when the difference occurred. The model is then validated as to whether a fault was detected before occurrence.

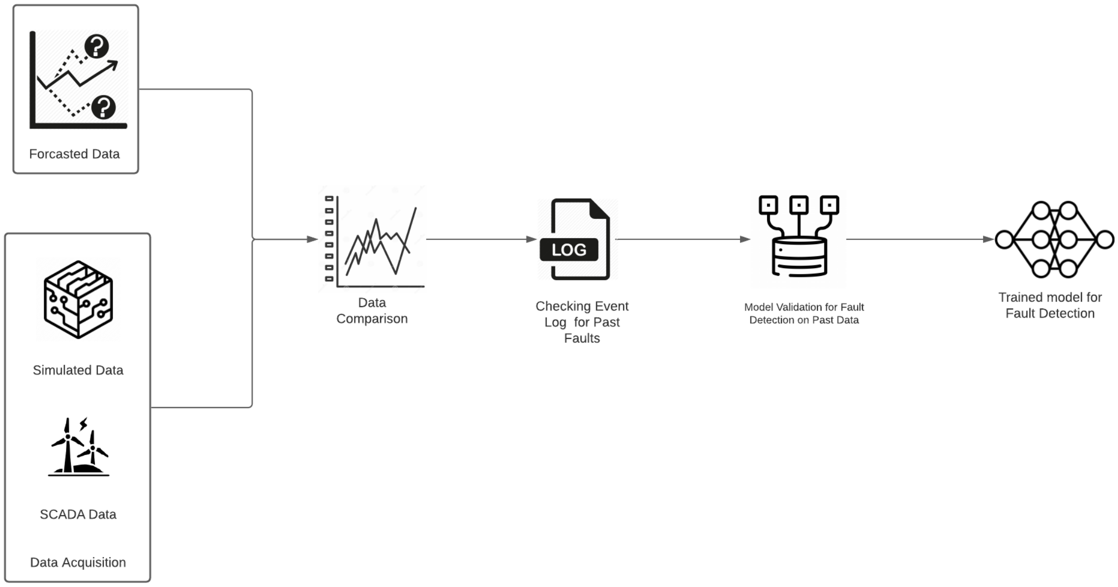

Figure 1: Data Preprocessing

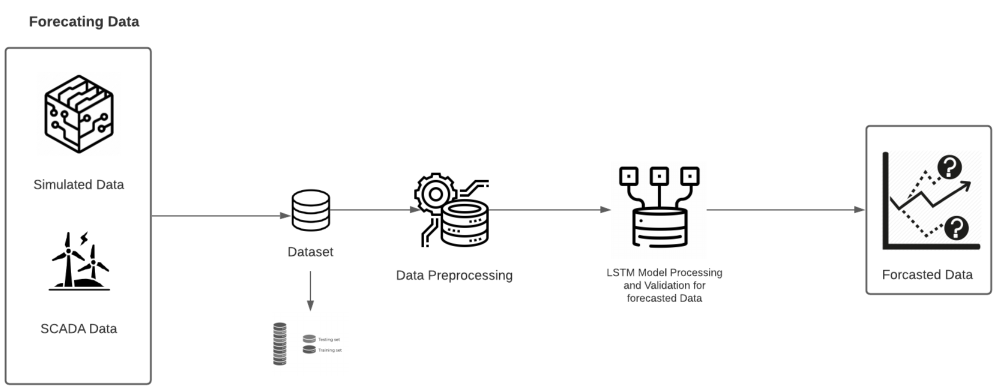

Figure 2: System Overview

As forecasting, since the SCADA data is time-series data, an LSTM model is built for forecasting features. As shown in figure 2, data is split into train and testing, and then preprocessing that includes normalization takes place, and finally, LSTM is used for forecasting. Data prepossessing helps ensure the model will achieve the best result, and that is done through data cleaning and data transformation. Data is then split into training and testing sets. Since the proposed system works with is time-series data, the LSTM network is most likely to be used for the model. As a contingency plan, we are working on different approaches parallel to LSTM (CNN - Autoencoders) that might not produce similar results to LSTM. However, we will hold up on the final product. A software accessible to on-field engineers would display the results reached by the model.

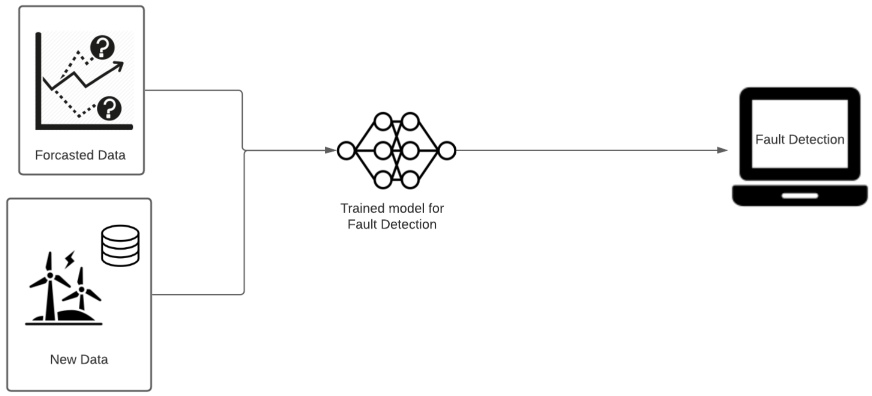

Figure 3: Output Application

For this system, acquired signals from SCADA system will be regularly forecasted and the model shall be used for fault detection as shown in figure 3.

5 Languages Used

For front-end stack PHP, CSS and JavaScript will be used for the web application interface, while Python will be used for implementing LSTM model.

6 New Ideas and Solution

Since the research into LSTM solutions for both fault detection and power generation prediction is still in its early stages, it's our plan to pick up where it drops off, aiming for better accuracies. Since it hasn't been done yet in our region, we aim to present a fully functioning, accessible and easy to use system that can be used to detect the anomalies and hence predict and specify the faults with enough of a time interval so preventative methods could be taken. This system should be able to operate on any form of SCADA data with the same precision.

7 System Development Methodology

Our team follows the Spiral model as it falls in line with our milestones time plan, in which the four phases are completed for each milestone. Testing is rotational between team members, where test cases based on analysis of the bias of the neural network used.

8 Timeplan

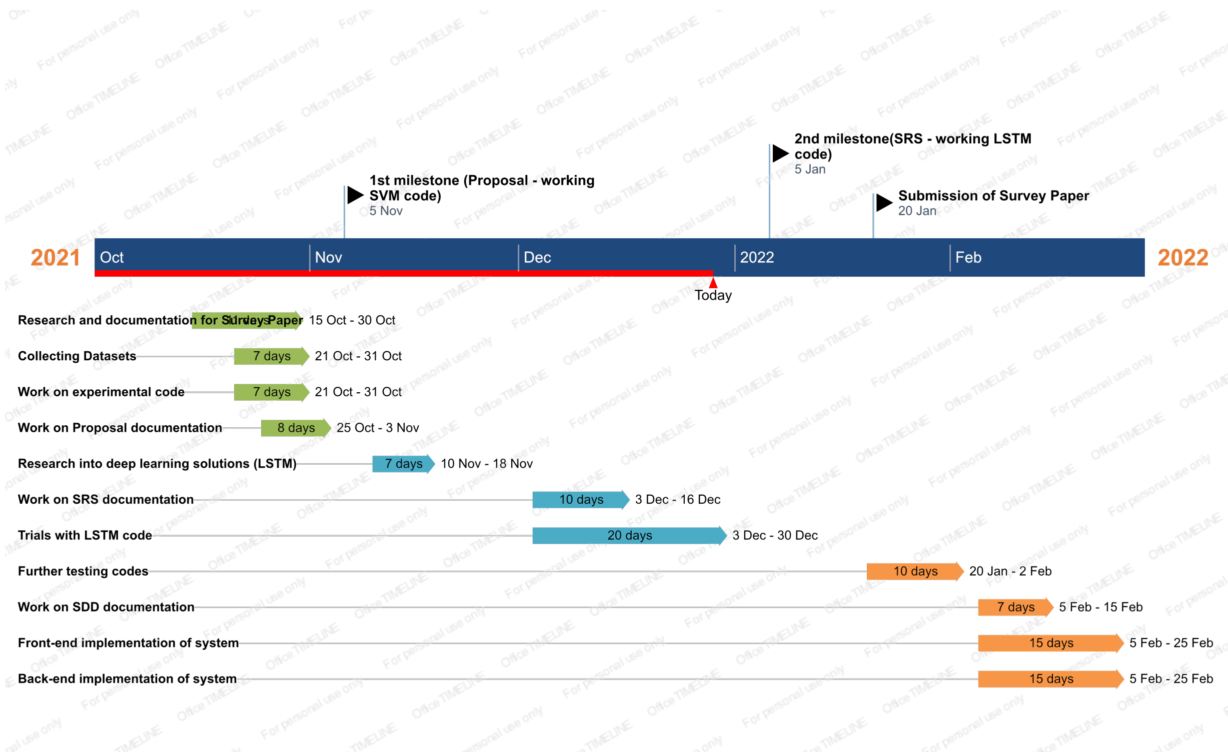

Figure 4: Time plan

Table 1: Project time plan

Id	Task	Start Date	Days	Team Member
1	Research and documentation for Survey Paper	15/10/2021	15	All group Members
2	Collecting Datasets	21/10/2021	10	All group members
3	Work experimental code	21/10/2021	10	Mahi Ayman, Mariam Othman, Nour Mahmoud
4	Work on Proposal documentation	25/4/2020	15	Nour Mahmoud, Zeina Tamer
5	1st milestone (Proposal - working SVM code)	5/11/2021	-	All group members
6	Research into deep learning solutions (LSTM)	10/11/2021	7	All group members
7	Work on SRS documentation	3/12/2021	10	All group members
8	Trials with LSTM code	3/12/2021	20	All group members
9	2nd milestone(SRS - working LSTM code)	5/1/2022	-	All group members
10	Submission of Survey Paper	20/1/2022	-	All group members
11	Second phase of trials on implemetation	20/1/2022	10	All group members
12	Work on SDD documentation	5/2/2022	-	All group members
13	3rd milestone (SDD - updated implementation)	23/2/2022	-	All group members
14	Front-end implementation of system	28/2/2022	15	Mahi Ayman, Mariam Othman
15	Back-end implementation of system	28/2/2022	15	Nour Mahmoud, Zeina Tamer
16	4th/final milestone (complete system)	25/4/2022	-	All group members

Figure 5: Time plan

References

[1] Jian Fu et al. "Condition Monitoring of Wind Turbine Gearbox Bearing Based on Deep Learning Model". In: *IEEE Access* 7 (2019), pp. 57078–57087. DOI: 10.1109/ACCESS.2019.2912621.

[2] Reihane Rahimilarki et al. "Time-series Deep Learning Fault Detection with the Application of Wind Turbine Benchmark". In: *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*. Vol. 1. 2019, pp. 1337–1342. DOI: 10.1109/INDIN41052.2019.8972237.

[3] Ana Rita Nunes, Hugo Morais, and Alberto Sardinha. "Use of Learning Mechanisms to Improve the Condition Monitoring of Wind Turbine Generators: A Review". In: *Energies* 14.21 (2021). ISSN: 1996-1073. DOI: 10.3390/en14217129. URL: https://www.mdpi.com/1996-1073/14/21/7129.