A REPORT ON

Machine Learning Methods for Wind Turbine Condition Monitoring

BY

Name of Student	-	ID Number
Srikar Lankamalla		*2018A4PS1035H*
Vishwanath Guruvayur		*2018A4PS1048H*
Niladri Nilamadhab		*2018A4PS0229H*
Akhil Annapaneni		*2018A4PS1213H*
Ackashh Dhevaraju		*2018A4PS1210H*
Ankit Ramesh Kaviath		*2018A4PS1208H*

ME F483 - Wind Energy

April 2021

Contents:

Introduction:

Wind Energy is one of the major aspects when it comes to renewable energy production. The process of harvesting wind energy is done by utilizing turbines that can be placed onshore or offshore. A major challenge when it comes to offshore wind turbines or farms is the maintenance and the amount of funding that is needed to achieve this. Over the past few years, we have seen a rapid development in techniques and analysis related to Machine Learning (ML). Condition Monitoring (CM) is another important feature that is used to boost the maintenance and reliability-related factors. In order to help reduce costs and improve feasibility, there can be an integrated system which has a Machine Learning based platform that can be utilized for Condition Monitoring of Wind turbines.

Offshore Wind Turbines

Condition Monitoring Overview:

Condition Monitoring (CM) is a key part of maintenance, operations and management and it includes inspecting the sub-components of a wind turbine in order to identify possible faults. Condition Monitoring has a drastic effect on reducing the cost and expenses of running these wind turbines. Maintenance is an aspect of Condition Monitoring that is essential to the structural integrity and overall reliability of the turbine and its components.

Maintenance can be classified into two parts:

1. Reactive: This method of maintenance is where procedures are carried out once a fault is identified, therefore it is done as a process to correct or fix the faults that are discovered. It is a more expensive method and it does not involve condition monitoring as components are being used till failure and they are fixed or replaced only once the fault is present.

2. Preventive: This method of maintenance is where procedures are carried out on a regular basis in order to check on components and their reliability. This type of scheduled maintenance helps reduce costs and also ensures minimum failures. It includes a basis of Condition Monitoring and it informs us in advance about components that have a risk of failure.

Condition Monitoring can be done at different sub-levels. We can choose to monitor components of the wind turbine and analyze them individually. This would be one of the smaller levels of Condition Monitoring. We could also choose to monitor the wind farm as a whole unit and analyze the overall working condition. This would be a higher level of Condition Monitoring.

The two types of Condition Monitoring:

1. <u>Intrusive Monitoring:</u> This type of Condition Monitoring includes methods such as vibration analysis, shock pulsation methodologies, etc. These methodologies cause slight wear & tear on the sub-components that are being monitored and hence it is termed intrusive monitoring.
2. <u>Non-Intrusive Monitoring:</u> This type of Condition Monitoring includes methods such as ultrasonic techniques, inspection, thermography, etc. These methods do not affect the components unlike the intrusive monitoring methods.

With respect to timelines, we can split Condition Monitoring into a further two types:

1. <u>Fault Detection (Present timeline)</u>: This is where Condition Monitoring is done to find faults after they have already occurred. This is similar to a reactive model of Condition Monitoring.
2. <u>Fault Prediction (Future timeline)</u>: This is where Condition Monitoring is done to predict and foresee faults that may occur in the future timeline. Here patterns in data are analyzed to figure out when a fault may occur. This is similar to the Preventive model of Condition Monitoring.

Another key aspect of Condition Monitoring is the prioritization of different components based on their failure rates and the likelihood of a fault. Different components can have different effects on an overall system and one component may be more likely to fail than another. Based on this, we should assign priorities to the components that may have drastic effects on the system if they are to fail.

Machine Learning Overview:

Machine learning is an implementation of AI that empowers machines to consequently enhance and develop from past occurrence without being explicitly instructed by an expert. Machine learning deals with the creation of software programs that can receive information and gain insights on its own.

The analysis starts with predictions or data, such as illustrations, actual experience, or training, to look for data and make predictions about the future based on indicators we supply. The prime aim is for machines to learn automatically without outside involvement or additional help and then alter their actions correspondingly.

Machine learning algorithms are frequently classified as either supervised or unsupervised.

Supervised Learning - Supervised machine learning algorithms can forecast future events by applying what they have managed to learn in the past to new data using labelled examples.

Unsupervised - Unsupervised machine learning algorithms, on the other hand, are used when the training data is not classified or labelled.

A series of steps can be used to represent the ML process:

- Data collection and Preprocessing: Data collection is the method of collecting and measuring information on specific variables in an existing framework, allowing one to answer pertinent questions and shows the results. The more the amount of data the better the model will be developed. Data preprocessing: Mathematical transformations on Data to make it Machine readable improving its quality and reducing the dimensionality Column Normalization is values are squished to a unit hyper-cube [0, 1], this gets rid of scales of each feature Column

Standardization: In a data matrix, with rows as data points and columns as features. Numerical features are sometimes encoded in one hot encoding or hash encoding. Categorical features are encoded in methods like Bag of Words or Tf-idf vectorizer. Data is cleaned of outliers. Imbalanced data sets are made balanced by up sampling or downsampling or introducing weights. Missing values are filled by data imputation using mean, median or any other aggregated methods.

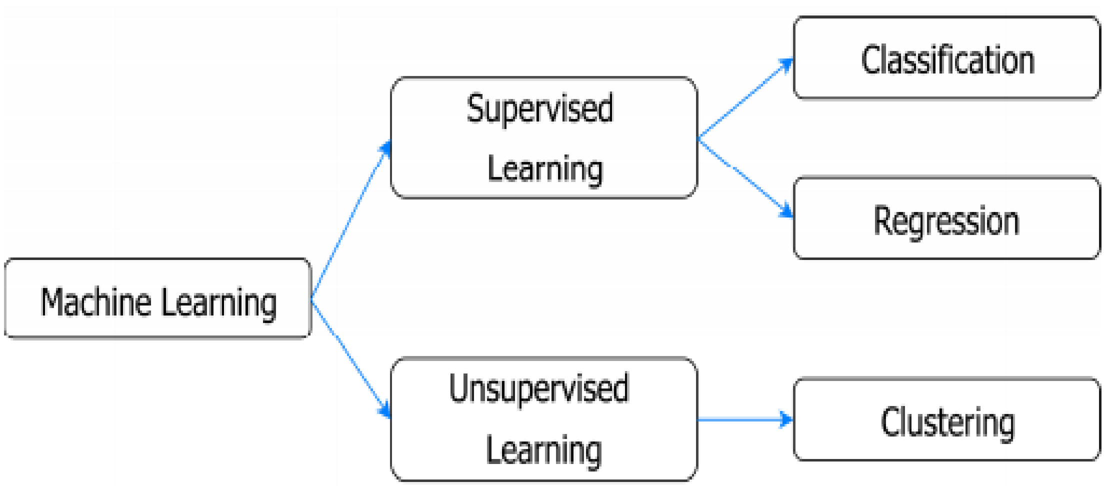
- Feature Selection and Extraction: Important features are Features that are useful for the machine learning model in making predictions; this improves model interpretability; feature importance allows us to understand a model better. Forward Feature Selection is one of the methods that can be used to select features.
- Choosing a ML model: Different ML models such as Logistic Regression, SVMs can be used to test out the data with hyperparameter tuning to find out the best hyperparameter and different scoring methods such as accuracy, f1-score, roc-auc curve can be used to determine the best ML Model.
- Validation of the model: Methods like K-fold cross validation can be used which would make efficient use of data and mostly give robust results.

Neural Networks (NN) and Support vector machines (SVM) are two popular ML methods for diagnostics and prognostics.

In the wind energy sector, NNs are being extensively used for predicting (e.g., forecasting of wind velocity), controlling aspects (wind turbine power control, etc.), recognition, & assessment (e.g. Fault Diagnosis.) Neural Networks are an important Machine Learning Modelling framework. Neural Networks handle outliers by using batch training and robust cost functions that include regularization. Neural Networks are so flexible that they have a drawback in terms of hyperparameters which can be tweaked. A simple MLP can have lots of hyperparameters: Number of hidden layers, number of neurons in these hidden layers, type of activation function to activate the inputs of neurons in each of these hidden layers, weights initialization logic, and so on.

One method is trying different combinations of hyperparameters and observing which combination of these parameters has higher accuracy on the created validation set. Connectionist models, such as NNs, are made up of simple activation based computational units called as neurons that interact with and transfer data to one another via synapses.

Support Vector Machines are frequently used in detection of faults and, more broadly, or complex data sets. This is a very popular Machine Learning Model for classification and regression problems. The objective will be to separate the data points with a hyper plane which is as far as possible from the data points.



DATA SET:

What is a data set: - In most situations, a wind turbine or a wind farm requires many different forms of data. So, a data set, quite literally, is a collection of data points. A data set lists the values for different variables which can be taken under consideration.

Different variables as part of the data set: -

- **Volume: -** Volume is the amount of data which is being produced by the turbines. Generally, a single turbine can produce up to 200 MB of data per second.
- **Velocity: -** Every time a turbine rotates, a certain amount of data is produced. This process has a fluctuating frequency which changes according to the wind speed in the area. The frequency is known as velocity.
- **Variety: -** The different types of data which can be taken as data points. Some examples of this are action data, image data and video data.

Veracity: - In certain cases, the above three forms of data are not enough to process, or the data provided could be very inconsistent. Veracity is the process of making sure that the accuracy of data is being maintained.

What is feature selection: - Feature selection is a process where we can individually select the specific variables and data which we want to study for the data set. There are many methods of feature selection. This can be done as an automatic method instead of manually. Some of the most used automatic methods are listed below.

Feature selection and Extraction:

Methods: -

- **Wrapper Method: -** This method is used specially for ML based algorithms. The method sees the components as small blackboxes and provides a variety of different subsets of the variables required for our output. One minor drawback is that the algorithm and the performance criterion should be specified before the process begins for this method to function properly.

- **Filter Method: -** This method operates independently of the model itself. By conducting a test between the signals and outcomes and ranking them in an order, this method selects the individual significance of them. The number of features is generally selected by the user before beginning.
- **Embedded Method: -** This method is commonly used in conjunction with the filter method stated above to increase the rate of success. This method conducts the selection by taking into consideration the relative importance of each node of the decision tree.

What is feature extraction? : - It is the process used to reduce the high dimensional time data by removing unnecessary parts of the data (such as additional background noise) and making sure that the main data is not changed. This can make the entire training process faster and improve the output. This is generally used for time series analysis. The most used methods are listed below.

Methods: -

- **Statistics: -** The statistical method is the easiest and most straightforward method of feature extraction. This includes data points such as mean, impulse factor, minimum and maximum, crest factor, etc. In many cases, this method is also the most efficient one present.
- **Fitted Model's Parameters: -** Coefficients of Fitted Models and Auto-correlation statistics.
- **Time-Frequency Properties: -** This is a method used to transform the signals from a time domain-based system into a frequency based domain. Some methods of this include the Wavelet transform and Fourier transform.

For many years, the main method of feature extraction in wind turbine research was time-frequency strategies. Wavelet transform has been considered as a better method and approach over Fourier transform. This is due to Fourier only being able to provide the frequency components by decomposing the time domain signal. On the other hand, Wavelet can provide both the time domain and frequency localization.

Developing the Machine Learning Models:

Once the step of Feature Extraction is completed, the next step is to develop the Machine Learning Model in order to conduct the conditional monitoring using the previously known information in Supervised Learning Methodology.

There are 2 types of Models based on the type of desired output variable from the model:

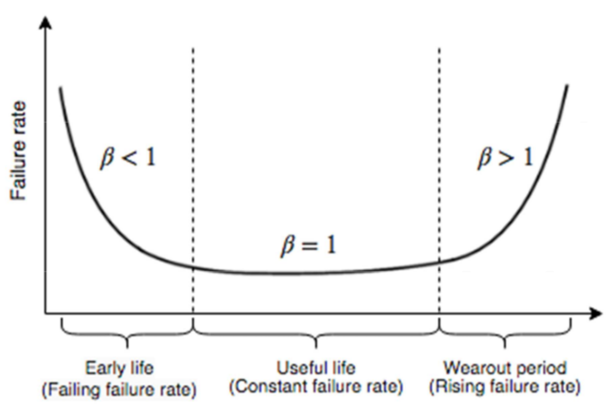
1) Regression Based Models (Numerical Output)
2) Classification Based Models (Categorical Output)

We will discuss these two Machine Learning Modelling types.

1) REGRESSION BASED MODELS

Regression Based Models are used for predicting a specific desired output parameter which is of numerical format depending on the model trained with prior acquired data.
In this type of modelling the main task is to understand the manner in which the different signals and output features are correlated. This relation between signals and output features is established with the data acquired during the peak performance times of the Wind Turbine.

The Mechanical and Electrical components of the Wind Turbine are generally follow a Failure Rate known as the Bathtub Curve as shown in the figure below. The regression based model is trained on the data when the Turbine is in the "Useful Life" Time when the components have the least failure rate.



Once the model is developed and is well trained for the Turbine's peak healthy state timeline, the training process is completed and the next step is to Test the model on new incoming data from a desired Wind Turbine Establishment. This new data is fed into the ML Model in a timely manner. Using the model, the output feature is predicted for the specific newly input data.

If the predicted output variable varies more than a specified threshold for many consecutive iterations, an alarm is raised for inspection of the specific component of which the data is showing the large deviation.

The most important Regression Based Model is the model which relates the Power Output of the Turbine versus the Wind Velocity around the turbine. This type of modelling is considered to be of the highest level of conceptuality, as in this, the whole turbine setup is recognised as a Black Box.

There are two types of Power Curve Modelling:

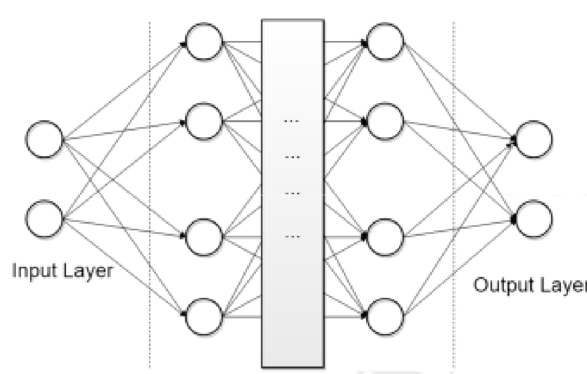
A] **Parametric Modelling Techniques** : Models are constructed by considering Parametric Vectors made up of a finite set of Parameters.

For Eg: Linear Segmented Model (2), 4-Parameter Logistic Function, Polynomial Curves, etc.

B] **Non-Parametric Modelling Techniques**: Models are constructed by Parametric Vectors that are not bounded with respect to their length.

For Eg: Neural Networks, Fuzzy Logic Methods, Cubic Spline Interpolation, etc.

In Neural Networks, the number of hidden computational neuron layers and the number of neurons in each of these layers can be varied and these parameters will directly affect the Vector Length.



2) CLASSIFICATION BASED MODELS

Classification based analysis involves the finding of a relationship between various different input variables which are generally represented in the form of vectors and one of categorical variables (output) is identified by labels. While training the model, the model is iteratively given an input vector of variables along with the label/class of the state.

Once the model is trained with appropriate amounts of data, the model is used as a prediction model for the Condition Monitoring of Wind Turbine Models based on the various parameters at the location of establishment as a function of time.

As this method comes under supervised Machine Learning, the output labels/classes to be given signifying exactly which training data points they belong to. This is a highly time consuming process and it is prone to a lot of errors because of which there is a probability of resulting in input vectors with an imbalance in the number of classes/labels. This is a very frequently occurring issue when performing classification type analysis.

Ways of getting rid of such problems is to undergo methods like:

● Under-Sampling : We have to exclude data-points that belong to the majority class/label.

● Oversampling: We have to include more data-points from minority class/label.

● Tomek-links: We have to remove such data-points in the majority class which are termed as redundant/ noise/ borderline/ etc.

Most commonly used classification methods are Support Vector Machines (SVM), Boosting Tree Algorithm (BTA), etc.

Support Vector Machine (SVM) is a supervised machine learning algorithm capable of performing classification, regression and even outlier detection. The support vector machine method's goal is to find a hyperplane in an N-dimensional space (N = the number of features) that clearly classifies the data points. There are numerous hyperplanes that could be used to distinguish two classes of data points.

Validation:

The validity of the generated Machine Learning models can be estimated using a range of specific methods in conjunction with techniques like k-fold cross validation method, which determines how accurately the model's output will make assumptions with data different from the training data.

Why should we use cross validation?

We split the data set into training and testing set based on a splitting ratio eg:70:30 ; We train the model using training set and compute accuracy using test set; For every hyperparameter we determine test accuracy and select the hyperparameter that gives best accuracy on test set; Using test dataset to determine best hyperparameter or best hyper parameters is not right; Thus we split the data set into train, cross validate and test datasets; we want the model to have well generalization ability on future unseen data; We train the model parameters compute nearest neighbors on training set, we determine best hyper parameters on cross validation dataset and we evaluate best ML models on test set; this will help ML models have good generalization ability

Method of K-FOLD CROSS VALIDATION

Data is split into train, cross validation and test set; While test set should not be touched, cross validation data becomes untouched while training which will lead to loss of information; We can use k-fold cross validation to incorporate cross validation data during training; We are trying to get the information from cross validation set for ML model training

1. Dataset is split randomly into train and test set

2. Train set is further divided randomly into k’ equal sized parts

3. For each hyper parameter in the ML model, we will use k’-1 parts of train set for training and the remaining 1 part of train set as cross validation data set, we then compute accuracy or model performance on cross validation data set; we will roll around the parts of train data set to get k’ accuracies for different values of hyperparameter, we will then average this accuracies; as a result we will have average of k’ accuracies of the train set; this is called as k’-fold cross validation, we will repeat the k’ fold cross validation for all hyper parameter k value choices

4. We pick best k from best average k’ cross validation accuracies

5. Then apply the best hyper parameter for measuring performance on test set

Integrating ML CM models for decision support

Finally, integrating these developed ML models for real time information will help in decision making for the operators.ML CM models also provide an advantage of looking over how the conclusions are reached.There are some of the “White Box” models such as Decision Trees which can be easily interpreted and there are also some “Black Box” models such as Neural Networks which provide no information.Appending confidence levels for the predicted values can help in decision strategy. Though some models like Logistic Regression are probabilistic , it can also be added for some other models.Sensors collecting data also must be regularly updated as this may lead to degradation in performance measures.

Conclusion:

This study has helped us to understand Machine Learning models- regression based and classification based, its fusion with Conditional monitoring and their integration with decision systems. Integrating these developed ML models for real time information will help in decision making for the operators. Similarly, these models have a number of different advantages related to Condition Monitoring , as some of them are easy to interpret, but fail in accuracy, however some are accurate , but are difficult to interpret. Decision strategy can be improved through different predicted values that can be obtained from the models. Therefore, integrating ML models with a robust Condition Monitoring system can have huge benefits towards the overall maintenance and management of a wind farm or turbine.

References:

1. A. Verma and A. Kusiak, “Fault Monitoring of Wind Turbine Generator Brushes: A Data-Mining Approach,” J. Sol. Energy Eng., vol. 134, no. 2, p. 021001, May 2012.
2. H. Offshore, “Offshore Windfarm Condition Monitoring and Maintenance (Position Paper),” 2017.
3. P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. SlaouiHasnaoui, T. A. Tameghe, and G. Ekemb, “Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges,” Energies, vol. 7, no. 4, pp. 2595–2630, 2014.
4. W. Qiao and D. Lu, “A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis —Part II: Signals and Signal Processing Methods,” IEEE Trans. Ind. Electron., vol. 62, no. 10, pp. 6546–6557, 2015.
5. H. Dias, M. De Azevedo, A. M. Araújo, and N. Bouchonneau, “A review of wind turbine bearing condition monitoring: State of the art and challenges,” Renew. Sustain. Energy Rev., vol. 56, pp. 368–379, 2017.
6. R. K. Ibrahim, J. Tautz-Weinert, and S. J. Watson, “Neural networks for wind turbine fault detection via current signature analysis,” 2016.
7. Géron, Hands-On Machine Learning with ScikitLearn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O’Reilly Media, 2017.
8. R. Kohavi and G. H. John, “Wrappers for feature subset selection,” Artif. Intell., vol. 97, pp. 273–324, 1997.
9. A.L. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” Artif. Intell., vol. 97, no. 1–2, pp. 245–271, 1997.
10. D. Kateris, D. Moshou, X.-E. Pantazi, I. Gravalos, N. Sawalhi, and S. Loutridis, “A machine learning approach for the condition monitoring of rotating machinery,” J. Mech. Sci. Technol., vol. 28, no. 1, pp. 61–71, 2014.