An ABC-SVM based Fault Prognosis of Wind Turbines using SCADA Data

L.V.Revathi [a,*], M.Sudalaimani [b], S.Jeyadevi [c]
Kamaraj College of Engineering and Technology

** Corresponding Author:*
Contact: lvrevathi@gmail.com, +91-9360668744 (L.V.Revathi)
msmani.be4@gmail.com +91-9443428572 (M.Sudalaimani)

Abstract: In modern world, Fault prognosis play an essential role in the safe and reliable operation of wind turbines (WT). An efficient fault prognosis method will help in the earlier identification of WT faults and failures, thereby reducing WT maintenance costs and improving their operating time. WT's are often controlled by Supervisory Control and Data Acquisition (SCADA), which, apart from controlling, provides very rich data pertaining to WT's working parameters which could be used for fault prognosis of WT's. Most of the WT prognosis systems make use of Support Vector Machines (SVMs) in conjunction with SCADA data to predetermine the faults that might occur in the turbines in the near future. In these models, proper selection of SVM parameters is essential for precise fault classification. In this work, an optimised methodology using Artificial Bee Colony Optimisation (ABC) is proposed to find an optimal penalty factor and kernel function parameter that guarantees better classification accuracy for the SVM model. Based on the real time SCADA data availed from a wind farm, it is observed that the proposed ABC-SVM fault prognosis model has a quick convergence rate and good accuracy compared to the other GA-SVM, PSO-SVM, and ACO-SVM based models.

Keywords: Fault Diagnosis, Wind Turbines, Artificial Bee Colony (ABC), Support Vector Machine (SVM), Parameter Optimization

1 INTRODUCTION

In recent years, the gradual depletion of fossil fuels and their environmental impact, as proposed by Tao Li [1], have led to a manifold increase in research efforts towards Renewable Energy Sources (RES) like Wind Energy Systems (WES), Photovoltaic Systems (PV), etc. Among the various RES systems, as proposed by Haas [2] and Wretling [3] - WES is considered an alternative to traditional fossil-based energy generation because of its clean nature and efficiency, and is gradually replacing fossil-based power sources. WES in general, Wind Turbines (WT) are thus gaining importance in recent days as the world is recognizing their significance in attaining the Paris Agreement's goal of limiting global warming to 1.5 degrees Celsius by 2100 [4].

This necessitates the development of more accessible and reliable wind energy solutions that provide guaranteed renewable energy generation. The remote and dispersed layout of WT's, along with their persistent exposure to unpredictable situations, contributes to an increase in the frequency of WT failures. In order to assure continuous power generation and dependable operation from WT's, various studies have focused on WT safety and preventive maintenance. Hossain [5] has proposed methodologies to assure a continuous and safe operation of WT's. Liu [6] has proposed the need for preventive maintenance of WT rotors, gearboxes, blades, generator, etc. To optimize the preventive maintenance interval and to ensure that maintenance is carried out only when it is really necessary, real time Condition Monitoring (CM) of WT's is necessary.

As proposed by Wen [7], among the various CM techniques, condition monitoring based on sensor data and, in general, SCADA are drawing a lot of research interest. The ability to use SCADA data for fault diagnosis with no additional expense as opposed to other CM techniques has made SCADA data based CM an ideal choice for condition monitoring of WT's. Though data analytic methods are proposed for WT Level and Wind Farm Level CM, WT level CM is commonly adopted as proposed by Jin [8].

Sequeira et al. [9] demonstrated the correlation between SCADA data and turbine faults and have proposed methods to predict the WT fault based on the SCADA data. Qiu [10] proposed a WT fault diagnosis model based on the Dempster–Shafer evidence theory, which correlates and cross-correlates the available SCADA data like maintenance records and alarm data. Encalada [11] proposed the use of SCADA data alone for the fault prognosis of WT's.

However, the direct application of ML algorithms to the WT SCADA dataset is a little bit time consuming owing to the number of parameters/features stored in the SCADA dataset. To overcome the above difficulty, Velandia [12] and Wen [13] proposed the use of feature dimensionality reduction algorithms like PCA and Relief PCA to remove the unnecessary and irrelevant features from the dataset, thereby reducing the dataset's features, speeding up the calculation time, and providing an accurate result.

As reviewed by Leite [14], SCADA based CM of WT's are mainly centered on Artificial Neural Networks (ANNs) and support vector machines (SVMs) owing to their accuracies. Yue [15], Kong [16] and Xiang [17] have proposed an ANN and CNN based CM model for WT to detect the gearbox failure at an earlier stage using the SCADA data. ANN and CNN besides its popularity in other applications finds a very limited application in WT - CM systems due to its complexity and increased training time with the increase in the number of samples and fault classes.

SVM-based CM as proposed by Hu [17] and Leahy [18] on the other hand is less complex and is highly preferred in CM of WT's because of its low sample demand, strong generalization ability, and high diagnostic accuracy. However, the performance of SVM based CM too depends on the proper selection of the kernel function and penalty factor "C". Several optimization methods, like Grid Search GS, Bio Inspired Algorithms like Particle Swarm Optimization (PSO), Ant Colony Optimisation (ACO), Sparrow Search Algorithm (SSA) & Genetic Algorithms (GA) have been proposed in the literature to aid in the optimal parameter selection for the SVM-based CM.

Li [19] proposed using GS to find the optimal tuning parameters for the SVM model. GS is an inch-by-inch search algorithm, that often opts for low-dimensional data and is often characterised by sluggish search rates, with parameter optimisation often depending on the user experience. PSO based parameter optimisation as proposed by Zhang [20], Yerui [21] and ACO based parameter optimisation as proposed by Li [22], Yanjun[23] on the other hand, have a sluggish convergence rate and often get trapped in local optimal points, failing to find an optimal parameter. To overcome the sluggish behaviour of PSO and ACO based parameter optimisation, Xue [24] and Tuerxun [25] has proposed an SSA based PSO parameter optimisation. SSA, apart from having a fast convergence rate, has a very meagre ability to search for a global optimum and often ends up at a local optimum.

The Artificial Bee Colony Algorithm (ABC) is an evolutionary algorithm based on the foraging behavior of bees. As proposed by Karaboga [26], Yang [27], and Wang [28] unlike other metaheuristic algorithms, ABC considers both local and global optimums at each iteration. The aforesaid feature of ABC increases the likelihood of finding the ideal solution while substantially reducing precocity, which makes ABC preferable over other metaheuristic optimization algorithms. In this work, a SVM based fault prognosis model with SVM parameters optimized through the ABC algorithm is proposed.

The rest of the manuscript is structured as follows: Section 2 provides a short introduction to the SVM model, ABC algorithm, and ABC based SVM parameter selection. Section 3 provides a short description of the dataset used and the pre-processing techniques. Section 4 gives the experimental validation and feasibility of implementation of the proposed ABC-SVM on real-time data, followed by the comparison of the results obtained with PSO-SVM, ACO- SVM and GS-SVM models. Finally, Section 5 concludes this manuscript with concluding remarks.

2 ABC – SVM ALGORITHM

2.1 SVM

SVM tries to solve any classification problem by building an ideal hyper plane that categorizes fault data from normal data. SVM classifies both linear and nonlinear data. SVM tries identifying Support Vectors (SV) to plot the boundary line. SV are data points that lie on either side of the hyperplane and are closest to it.

Consider a generalized training dataset T defined by

$$T = \begin{cases} (x_{11,} x_{12,} \ldots x_{1d}), y_1 \\ (x_{21,} x_{22,} \ldots x_{2d}), y_2 \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ (x_{n1,} x_{n2,} \ldots x_{nd}), y_n \end{cases} \quad (1)$$

where $(x_{i1,} x_{i2,} \ldots x_{id})$ corresponds to the i^{th} data point in the given dataset with its corresponding label denoted by y_i. ‘n’ and “d” represents the number of data points available in the dataset and the dimension or number of features available in the dataset.

SVM tries to find an optimal hyperplane that classifies the fault data from normal data based on margin maximization, resulting in the formation of the following optimization problem.

$$\min_{w,\in,b} \quad \frac{1}{2}\|w\|^2 + C\sum_{j=1}^{n} \in_j \quad (2)$$

subject to the condition

$$y_j\left(\left(w.\sigma(x_j)\right) + b\right) > 1 - \epsilon_j \quad (3)$$

$$\epsilon_j \geq 0, i = 1,2, \ldots . n \quad (4)$$

where “w” and “b” are the normal support vector and offset Value that correspond to the separating hyperplane. “*C*” is the penalty factor for the error term ϵ and it corresponds to the interval size and classification accuracy weight. $\sigma(.)$ is the mapping function that maps any given input sample x_j onto a higher dimensional space. Equation 2 to Equation 4 could be translated to

$$\min_{\alpha} \quad \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \left(\sigma(x_i).\sigma(x_j)\right) - \sum_{j=1}^{n} \alpha_j \quad (5)$$

Subject to the condition given by Equation 6

$$\sum_{i=1}^{n} y_i \alpha_i = 0 \quad (6)$$

Solving Equation 5 & Equation 6 results in the formation of the classification decission function which is given by Equation 7

$$f(x) = sign\left(\sum_{i=1}^{n} \alpha_i y_i \left(\sigma(x_i).\sigma(x_j)\right) + b\right) \quad (7)$$

where $k = \left(\sigma(x_i).\sigma(x_j)\right)$ forms the kernel function. Among the various kernel functions, radial basis kernel function provides a wide adaptable range and hence the radial basis kernel function given by Equation 8 is used

$$k\left(\sigma(x_i).\sigma(x_j)\right) = exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right), \delta > 0 \qquad (8)$$

From the Equation 8, it is evident that the accuracy of SVM depends on the selection of Kernel Function type *"k"*, Kernel Function Parameter *"δ"*, and penalty factor *"C"*.

Penalty factor "*C*" on approaching *"0"*, results in lower classification accuracy and in under fitting. Though the classification accuracy increases with the increasing value of C, higher values of C on the other hand, result in over fitting. Similarly, a smaller value of "*δ*" results in over fitting and a higher value results in under fitting.

An optimal selection of kernel function Type *"k"*, kernel function parameter "δ", and penalty factor *"C"* will mutually improve the classification accuracy of SVM model. To aid in the selection of optimal parameters for SVM, and thereby improve the model classification accuracy, an ABC algorithm based optimisation of SVM parameters (penalty factor *"C"*, kernel function parameter *"δ"*) is proposed in the upcoming sections.

2.2 Artificial Bee Colony Algorithm

ABC algorithm an evolutionary algorithm based on the foraging behaviour of bees, tries to arrive at an optimal solution for any given problem by finding the fitness value (amount of food source and nectar in that specific source or solution) at each possible source or solution and iteratively trying to find the best optimal source or solution based on the foraging behaviour of bees.

ABC makes use of the foraging behavior of Employee, Onlooker and Scout bees to find the optimum food source. In the ABC algorithm, employee bees are assigned the task of finding food resources and sharing information with other bees. Onlooker bees are assigned the task of honey collection based on the information shared by the employee bees. Scout bees are assigned the task of finding a new food source after the existing one is depleted.

Like any other swarm algorithm, the ABC algorithm is iterative. The optimisation process starts with the initialization of the number of bees in each category and the location of the food source. The foraging behaviour of the employee, scout, and onlooker bee is iteratively carried out to find the optimal food source or solution to the problem.

In the ABC algorithm, each potential solution is given by the position of each nectar source, and its fitness value is represented by the amount of nectar present in the particular source. Like any other swarm optimization algorithm, ABC algorithm starts with the initialization phase given by Equation 9,

$$x_{i,j} = l_{min,j} + rand(0,1)\left[l_{max,j} - l_{min,j}\right] \qquad i = 1,2,3,\ldots\ldots NB \qquad (9)$$

where NB is the Scout bee population. $l_{max,j}$ and $l_{min,j}$ are the lower and upper bounds of $x_{i,j}$. Each food source defined by $x_{*,j}$ is a potential solution vector having *"n"* variables that needs to be optimized to find the optimal solution.

After the initialization process, the employee bees forage randomly for nectar or food sources, and the foraging behaviour is given by Equation 10.

$$v_{in} = x_{in} + \emptyset(x_{in} - x_{kn}) \qquad (10)$$

where v_{in} corresponds to the food source value corresponding to the i^{th} iteration and *"n"* is the selected dimension or the number of optimisation parameters, x_{kn} denotes the neighbouring nectar source.

After the foraging process, the nectar source fitness of each employee bee is calculated using the Equation 11.

$$fit_i = \begin{cases} \frac{1}{1+f_i}, & f_i \geq 0 \\ 1 + abs(f_i), & otherwise \end{cases} \qquad (11)$$

where f_i is the fitness value of the i^{th} nectar source, the objective function value that needs to be optimised. When the new nectar source fitness $v_i = [v_{i1}, v_{i2}, \ldots, v_{id}]$ is better than x_i, greedy selection method is used to replace x_i with v_i

After the completion of the foraging process, the entire employee bee population shares the nectar information with the onlooker bees, along with the amount of nectar. Onlooker bees in turn evaluate the employee bee's nectar information and select a nectar source that needs to be foraged with a probability of P_i using Equation 12 the selection probability for each solution.

$$P_i = \frac{fit_i}{\sum_{m=1}^{Sm} fit_m} \qquad (12)$$

The Scout Bee phase gets started when nectar source updation is not happening with successive iterations. The scout bee then selects a random food source using Equation 13, thereby replacing and abandoning the current food source.

$$x_{ij} = x_j^{min} + rand(0,1)(x_n^{max} - x_n^{min}) \quad (13)$$

2.3 ABC based SVM Model:

The steps to find the optimal parameters for the SVM model using the ABC algorithm is given as a flow chart in Figure 1, and the steps are narrated below.

Step 1: Import the SCADA data and perform common pre-processing like data normalisation and outlier removal.

Step 2: Using PCA, reduce the dimensionality or the number of features available in the data to speed up the computation and to prevent over fitting.

The optimal number of parameters for PCA reduction is derived from the SCREE plot of the given dataset.

Step 3: Split the given dataset into training and test data sets, with 70% of the data used as the training data set and 30% as the test data set

Step 4: Initialize the SVM and ABC algorithm parameters

Initialize the SVM parameters - penalty factor "C", kernel function Parameter "δ" that need to be optimized by ABC with some initial values. Also initialise the ABC algorithm parameters like bee population, lower and upper limit of solutions and dimension.

Step 5: Calculate the new optimal nectar source/solution using the ABC algorithm and compute the fitness value of the new nectar source. The current fitness value corresponds to the SVM classification accuracy and is updated in real time with the maximum fitness value as the iterations progress.

If the current fitness value of a nectar source is greater than the previous nectar source, the original fitness value is replaced with the current value. Once an optimal nectar source is found, the combination of parameters (C, δ) that corresponds to the current optimal nectar value is saved.

Step 6: Save and pass on the corresponding SVM parameter (C, δ) for SVM model fitting. Calculate the classification accuracy. If the accuracy is met, save the model and exit; else go to step 5 until an optimal SVM parameter (C, δ) that guarantees the desired classification accuracy is found.

3 DATASET DESCRIPTION

3.1 Dataset Description

The experimental data used in this work is obtained from a single bearing, 3 blade, 2 Megawatt WT located in the southern part of Tamil Nadu, India. The SCADA unit deployed to control the WT collects data at a 10-minute interval, and the data from January 2022 to December 2022 is utilized for this work.

The dataset encompasses normal data along with the data for 4 different faults that occurred during the aforementioned period. Because of data security and a non-disclosure agreement, the names of faults and the exact location of the WT are not disclosed here. The frequency of occurrence of each fault type during the period is given in Table 1.

3.2 Dataset Pre-Processing

Initial SCADA data received from the WT had 185 parameters with numerical and non-numerical data like relay and switch positions. In general, a larger number of features drastically increases the computation time and might result in overfitting of the data.

To overcome the above difficulty, as a first step in feature dimension reduction, all non-numeric features like relay and safety switch positions and parameters which don't carry much information and are of little use, are removed. After the removal of the non-salient features, 72 useful parameters were left, and the same are used further in this work.

3.3 PCA based Dimensionality Reduction

To further reduce the number of features and thereby improve the computation time, Relieff-PCA-based feature dimension reduction is carried out. Relief PCA combines the principles of ReliefF and Principal Component Analysis (PCA) to reduce the data dimension and complexity. The relevance and information value of each feature is calculated using the ReliefF algorithm and once the relevance and information value of each feature has been calculated, Relief PCA uses PCA to select the features that are most relevant and thereby making the selected features linearly non-correlated. The feature with the highest data variance forms the first principal component, and the feature with the next largest variance forms the second principal component, and the process repeats until the desired number of principal components is found. A Scree plot, which gives the proportion of variance with respect to the number of parameters, is used to select the optimal number of parameters.

Figure 2 and Figure 3 give the SCREE plot between cumulative variance vs. number of parameters and Eigen value vs. number of parameters for the dataset considered.

Based on the SCREE plots, it is found that a total of 23 parameters seem to be an ideal threshold choice for PCA reduction, and is PCA is used to further reduce the SCADA dataset dimension to 23 principle features. Table 2 gives the list of principal features selected from the given dataset that would be used for the experimental model verification process.

4 EXPERIMENTAL VERIFICATION

4.1 Performance Metrics

To evaluate the classification accuracy of the proposed ABC-SVM fault diagnosis model, computational experiments are carried out on the real time SCADA data using the proposed ABC-SVM model, and its performance is compared with the effectiveness of GA-SVM, PSO-SVM, and ACO-SVM based models.

To evaluate the performance of the proposed algorithm with respect to other models, the following performance metrics are used

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$
$$Precision = \frac{TP}{TP+FP}$$
$$Recall = \frac{TP}{TP+FN}$$
$$F1\ Score = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \qquad (12)$$

Where

TP = True Positive, number of correct predictions of a particular Fault
FP = False Positive, number of wrong predictions of the particular Fault
FN = False Negative, incorrect false prediction of positive Faults
TN = Total score excluding the other three values TP,TN & FP.

The above performance metrics are selected as they provide different perspectives of the model performance in identifying and classifying faults. Precision, a metric that corresponds to the actual correct positive predictions is considered to minimize false positives as false positives can lead to unnecessary downtime and maintenance costs. Recall indicating the proportion of actual positive is considered to minimize false negatives as false negatives might lead to safety risks and financial losses. F1 score a harmonic mean of precision and recall, providing an equal weightage to precision and recall is considered here as it indicates model's ability to identify both upcoming faults and avoid predicting unnecessary maintenance, thereby reducing the overall cost of maintaining WT's.

4.2 Selection of SVM Activation Function

The classification accuracy and rate of convergence of SVM models are highly dependent on the activation function used. SVM models commonly employ the Sigmoid, Tanh or Relu as the activation function.

To select an optimal activation function for the classification problem considered, the loss function value of the SVM model with the aforesaid three activation functions is calculated and plotted. Figure 4 shows the plot of the loss function for the different activation functions vs. iterations.

From Figure 4, it is clear that the SVM model with Sigmoid activation functions converges very slowly to a loss value of 0.32 under the 70th iteration, which is much larger than the other two activation functions. Though the loss curve under Tanh and ReLU activation functions remains almost similar, the convergence rate of the ReLU activation function is little bit faster. Also, the SVM model with the Relu activation function converges to a loss value of 0.18 within 20 iterations, with the final loss value reaching out to 0.11, which is far better than the loss value obtained through the other activation functions. Hence, a ReLU based activation function is used in the SVM model, which is proposed in the upcoming sections.

4.3 Initialization Parameters

Any metaheuristic optimization algorithm like PSO, ACO, and ABC starts with the initialization of parameters like herd size, number of iterations, and tuning parameters. Table 3 gives the initialization parameters that are used in this study for PSO, ACO, and ABC-based model optimization methods.

The optimal range for the penalty factor C and kernel function parameter σ is set as [0.001, 100] for all the above metaheuristic algorithms.

4.4 ABC – SVM Diagnostic Model Performance Comparison

Each metaheuristic algorithm is iteratively tuned for a fixed number of iterations on the test data to find the optimal hyper parameter. The hyper parameter, which provides the best optimal fitness value, is used along with the SVM model for fault detection and classification. Figure 5 gives the plot of fitness value achieved by each of the models considered (GA-SVM, PSO-SVM, ACO-SVM, and ABC-SVM) with respect to the number of iterations.

From Figure 5, it is clear that compared to the other models, the ABC-SVM model reached the best fitness value, thereby providing an optimal parameter value for the SVM model in a very less number of iterations. For the SCADA data considered, the proposed ABC-SVM model reached the optimal parameter by the 20th iteration, whereas the other counterparts reached the optimal parameters by the 30th iteration. Also, for the data considered, the ABC-SVM model was able to provide a better fitness value of nearly 97% compared to the other models. A better fitness value, in turn, improves the performance of the SVM classifier.

For further performance visualization of the proposed ABC-SVM model, the confusion matrix for the models considered is generated and provided in Table 4. From a total of 410 reported faults, ABC-SVM was able to classify 405 faults accurately, with an accuracy of 98.6%. The ABC-SVM classification accuracy is relatively higher compared to the classification accuracy of other models like PSO-SVM, which has an accuracy of 94.4%.

Table 5 gives the consolidated performance metrics of the ABC-SVM model along with those of the GS-SVM, GA-SVM, PSO-SVM, and ACO-SVM models considered. From the results obtained, it is evident that the classification accuracy and overall performance metrics of the proposed ABC-SVM model are better than those of its counterpart models. The F1 score for ABC-SVM stands at 99.38%, which indicates the ABC-SVM model's ability to accurately classify the faults even with an unbalanced dataset.

Figure 6 gives the classification accuracy of the proposed ABC-SVM model and other models considered, considering only a single type of fault exists at a time in the dataset provided. From Figure 6, it is evident that the ABC-SVM model classification accuracy remains almost constant irrespective of their frequency of occurrence, even in an unbalanced data. ABC- SVM model is able to classify faults with an accuracy of 99%, irrespective of the faults considered, which is far better compared to the other models considered.

Figure 7 gives the classification accuracy of the proposed ABC-SVM model and other models, considering multiple faults at a time. From the Figure 7, it is evident that the classification accuracy of the models drops considerably with the increase in fault number; on the other hand, the proposed ABC-SVM model accuracy almost remains constant at around 98%, irrespective of the number of faults considered.

To ensure the generalization ability of the proposed model and due to the limited public availability of WT fault data, 5-fold Cross validation is carried out on the available fault data as proposed by Tiwari [29]. The dataset is divided into five parts with four parts used for training while the remaining one part is used as a validation data and carrying out the process over 5 iterations as shown in Figure 8.

Table 6 gives the performance indices of the proposed ABC-SVM model across each fold of cross-validation. From Table 6, it is evident that the proposed ABC-SVM model almost provides a consistent accuracy with the each fold validation accuracies given by 98.31%, 98.17%, 98.03%, 98.01%, and 97.92%, with an average cross-validation accuracy of 98.09% ensuring the generalization ability of the proposed ABC-SVM.

From above the analysis, compared to the Grid Search algorithm (GS), the Genetic Algorithm (GA) and other metaheuristic algorithms like Particle Swarm Optimization (PCO), Ant Colony Optimization (ACO) Algorithm, in Artificial Bee Colony optimization (ABC) both global search and local search takes place at each iteration, thus greatly increasing the probability to find the optimal hyper parameter for the SVM model which guarantees the best classification accuracy and results in a more robust classification model.

5 CONCLUSION

The gradual depletion of fossil fuels and their environmental impacts has led to a manifold increase in the use of renewable energy sources and, in general, wind energy systems. The increased use of WT's necessitates the need for continuous condition monitoring and troubleshooting of WT's so as to assure continuous power generation and dependable operation from WT. As SVM-based classification model is commonly deployed for the fault diagnosis of WT's, an ABC algorithm-based optimal parameter selection for the SVM model is proposed in this paper.

The proposed model's effectiveness is verified by comparing the performance metrics of the proposed model with those of traditional models like GS-SVM, GA-SVM, PSO-SVM, and ACO-SVM on an experimental SCADA dataset provided by a WT. On comparing the results, the ABC-SVM based model is capable of accurately classifying multiple faults with an accuracy of 98.6%, which is better when compared with the other models. F1 Score of the proposed ABC-SVM model is around 99.38%, which indicates its ability to classify the faults though the data is unbalanced. Also, the classification accuracy of the proposed model remains almost constant, thus providing an opportunity to deploy the proposed ABC-SVM model for multiple fault classification without losing accuracy. Though the data considered in this study were obtained from a single WT, the proposed approach is transferable and generalizable to multiple wind farms.

Reference:

[1] Tao Li, Ang Li, Xiaopeng Guo, "The sustainable development-oriented development and utilization of renewable energy industry - A comprehensive analysis of MCDM methods", *Energy, Elseiver* Volume 212, 2020,118694 (2020). https://doi.org/10.1016/j.energy.2020.118694

[2] Haas, R., Panzer, C., Resch, G., Ragwitz, et al., "A historical review of promotion strategies for electricity from renewable energy sources in EU countries", *Renewable and sustainable energy reviews*, 15(2), pp.1003-1034 (2011). https://doi.org/10.1016/j.rser.2010.11.015

[3] Wretling, V., Balfors, B. and Mörtberg, U, "Balancing wind power deployment and sustainability objectives in Swedish planning and permitting", *Energy, Sustainability and Society*, 12(1), p.48 (2022). https://doi.org/10.1186/s13705-022-00376-y

[4] United Nation Climate Change. In Proceedings of the COP26: *The Glasgow Climate Pact in United Nation Climate Change Conference UK*, Glasgow, UK, 31 October–12 (2021)

[5] Hossain, M.L.; Abu-Siada, A.; Muyeen, S.M., "Methods for Advanced Wind Turbine Condition Monitoring and Early Diagnosis: A Literature Review", *Energies*, 11, 1309 (2018). https://doi.org/10.3390/en11051309.

[6] Liu, Zepeng, and Long Zhang, "A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings", *Measurement* 149 (2020): 107002 (2020). https://doi.org/10.1016/j.measurement.2019.107002

[7] Wen, Yuxin, Md Fashiar Rahman, et al., "Recent Advances and Trends of Predictive Maintenance from Data-Driven Machine Prognostics Perspective", *Measurement*, 187, 110276 (2022). https://doi.org/10.1016/j.measurement.2021.110276

[8] Jin, Xiaohang, Zhuangwei Xu, and Wei Qiao, "Condition Monitoring of Wind Turbine Generators using SCADA data analysis", *IEEE Transactions on Sustainable Energy*, 12, no. 1, 202-210 (2020). doi: 10.1109/TSTE.2020.2989220

[9] Sequeira, C., Pacheco, A., Galego, P., et al., "Analysis of the Efficiency Of Wind Turbine Gearboxes Using The Temperature Variable", *Renewable Energy*, 135, 465–472 *(*2019). https://doi.org/10.1016/j.renene.2018.12.040

[10] Qiu, Y.; Feng, Y.; Infield, D., "Fault Diagnosis of Wind Turbine With SCADA Alarms Based Multidimensional Information Processing Method", *Renewable. Energy*, 145, 1923–1931 (2020). https://doi.org/10.1016/j.renene.2019.07.110

[11] Encalada-Dávila, Ángel, Bryan Puruncajas, et al., "Wind Turbine Main Bearing Fault Prognosis Based Solely On SCADA Data", *Sensors* 21, no. 6: 2228 (2021). https://doi.org/10.3390/s21062228

[12] Velandia-Cardenas, C., Vidal, Y., & Pozo, F. "Wind Turbine Fault Detection using Highly Imbalanced Real SCADA data". *Energies*, 14(6), 1728 (2021). https://doi.org/10.3390/en14061728

[13] Wen, X., & Xu, Z.. "Wind Turbine Fault Diagnosis based on ReliefF-PCA and DNN". *Expert Systems with Applications*, 178, 115016 (2021). https://doi.org/10.1016/j.eswa.2021.115016

[14] G.d.N.P.Leite, A.M.Araújo, and P.A.C. Rosas, "Prognostic techniques applied to maintenance of wind turbines: a concise and specific review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1917–1925, 2018 (2018). https://doi.org/10.1016/j.rser.2017.06.002

[15] Cui, Yue, Pramod, et al.,"A fault detection framework using recurrent neural networks for condition monitoring of wind turbines." *Wind Energy* 24.11: 1249-1262 (2021). https://doi.org/10.1002/we.2628

[16] Kong, Ziqian, et al. "Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units." *Renewable Energy* 146: 760-768 (2020). https://doi.org/10.1016/j.renene.2019.07.033

[17] Xiang, Ling, et al. "Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism." *Measurement* 175: 109094 (2021). https://doi.org/10.1016/j.measurement.2021.109094

[18] Leahy, K., Hu, R. L., Konstantakopoulos, I. C., Spanos, C. , et al., "Diagnosing and Predicting Wind Turbine Faults from Scada Data Using Support Vector Machines", *International Journal of Prognostics and Health Management*, *9*(1) (2018). https://doi.org/10.36001/ijphm.2018.v9i1.2692

[19] B. Li, M. Liu, and P. Yang, Application of EWT and GS-"SVM in Mechanical Fault Diagnosis of Circuit Breakers", *Entropy*., vol. 39, no. 8, pp. 1422–1430 (2018). doi: 10.3390/e20060448

[20] Zhang, X., Han, P., Xu, L., et al., "Research on bearing fault diagnosis of wind turbine gearbox based on 1DCNN-PSO-SVM". *IEEE Access*, *8*, 192248-192258 (2020). doi: 10.1109/ACCESS.2020.3032719.

[21] Fan, Yerui, et al. "A bearing fault diagnosis using a support vector machine optimised by the self-regulating particle swarm." *Shock and Vibration*: 1-11 (2020). https://doi.org/10.1155/2020/9096852

[22] Li, X., Zhang, X., Li, C., & Zhang, L. "Rolling element bearing fault detection using support vector machine with improved ant colony optimisation". *Measurement*, *46*(8), 2726-2734 (2013). https://doi.org/10.1016/j.measurement.2013.04.081

[23] Yang, Yanjun, et al. "Fault early warning of wind turbine gearbox based on multi input support vector regression and improved ant lion optimization." Wind Energy 24.8: 812-832 (2021). https://doi.org/10.1002/we.2604

[24] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," *Systems ence & Control Engineering An Open Access Journal*, vol. 8, no. 1, pp. 22-34, 202 (2020) https://doi.org/10.1080/21642583.2019.1708830

[25] Tuerxun, W., Chang, X., Hongyu, G., et al., "Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm". *IEEE Access*, *9*, 69307-69315 (2021). https://doi/10.1109/ACCESS.2020.3018320.

[26] Karaboga D, Basturk B, "A Powerful and Efficient Algorithm for Numerical Function Optimisation: Artificial Bee Colony (ABC) Algorithm", *Journal of Global Optimisation*, 39(3):459-471 (2007). https://doi.org/10.1007/s10898-007-9149-x

[27] Yang, Dalian, Yilun Liu, Songbai Li, et al., "Gear Fault Diagnosis Based On Support Vector Machine Optimized by Artificial Bee Colony Algorithm", *Mechanism and Machine Theory*,90, 219-229 (2015).

https://doi.org/10.1016/j.mechmachtheory.2015.03.013
[28] Wang, Haiquan, et al. "An improved bearing fault detection strategy based on artificial bee colony algorithm" *CAAI Transactions on Intelligence Technology* 7.4 (2022): 570-581.
https://doi.org/10.1049/cit2.12105
[29] Tiwari, V., Joshi, R. C., & Dutta, M. K. "Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images". *Ecological Informatics*, 63, 101289 (2021).
https://doi.org/10.1016/j.ecoinf.2021.101289

Biographies

L.V.Revathi is a Master of Engineering holder in Control and Instrumentation Engineering and she received her B.E., Electronics and Instrumentation Engineering Degree from Madurai Kamaraj University. She has worked as a Research Assistant in NIT Trichy and has a vast Teaching Experience. Her research interests include Non Destructive Testing, Fault Identification, Controller Design and Implementation

M.Sudalaimani received his Ph.D., degree in Electrical Engineering, Master of Engineering in Control and Instrumentation Engineering from Anna University and he received his B.E., Electrical and Electronics Engineering Degree from Madurai Kamaraj University. He has a vast Teaching & Industrial Experience. Her research interests include Power Electronics, Optimization Techniques, Controller Design and Implementation, Mainframe Technologies

Dr. S. Jeyadevi received the Ph.D. degree on 2011 from Anna University, Chennai. She has around 19 years of academic experience. She is currently working as Professor in Mangayarkarasi college of Engineering, Madurai, Tamilnadu, India. Her research interests include Power system optimization, Artificial intelligence and Smart grid.

Figure Captions:

Table Captions:

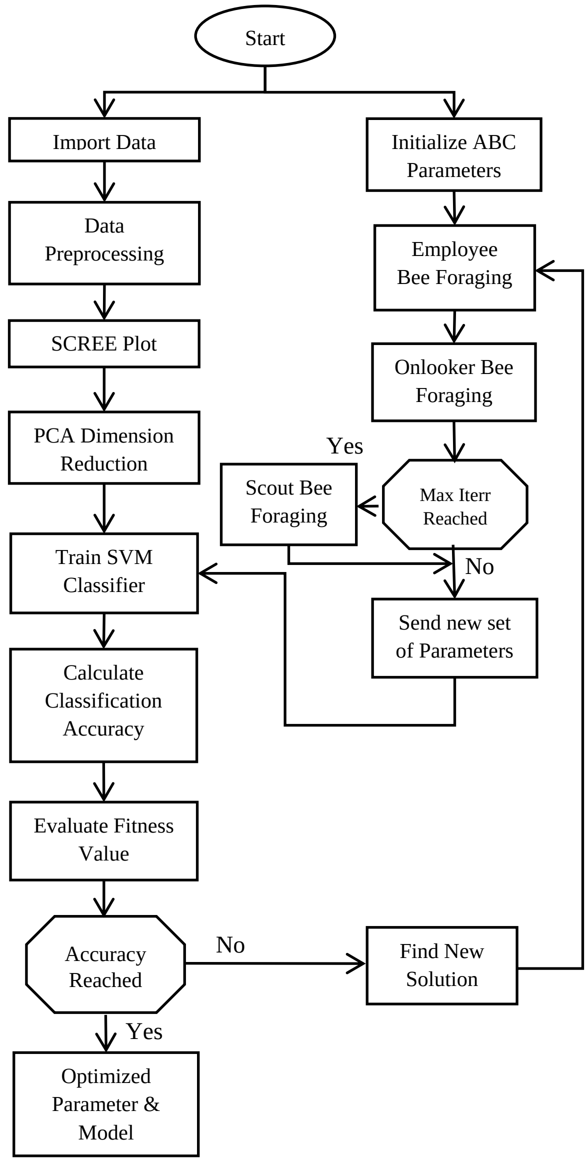


Figure 1 ABC – SVM Model Flow Chart

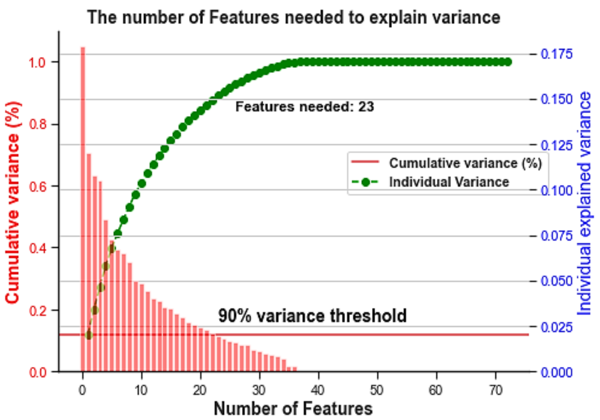


Figure 2 Feature Variance Distributions

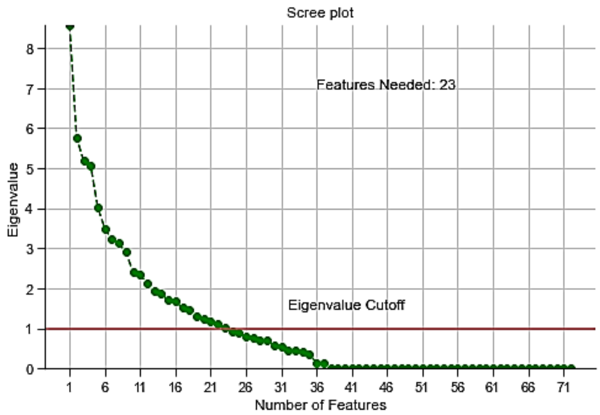


Figure 3 Feature Eigen Value Distributions

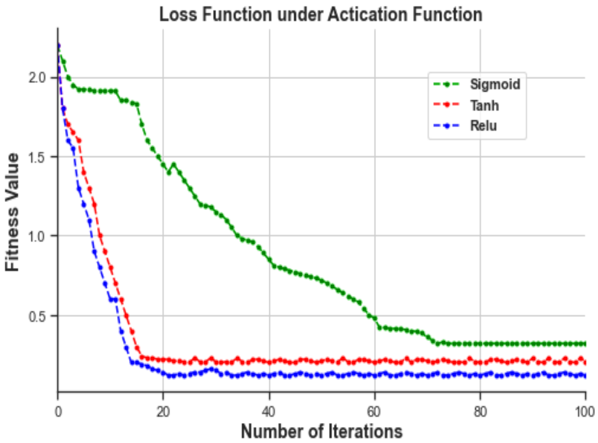
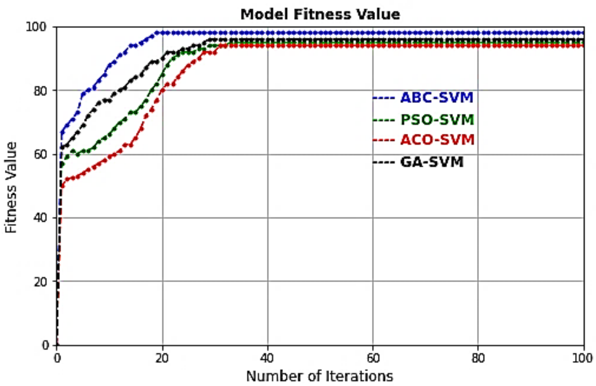
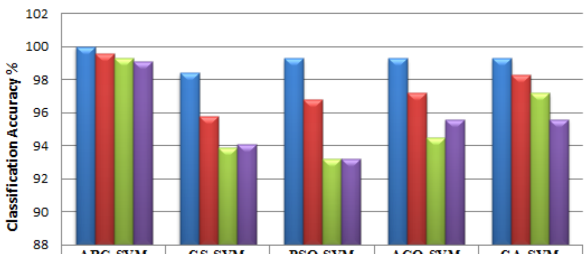


Figure 4 Loss Function under various Activation Function





	ABC-SVM	GS-SVM	PSO-SVM	ACO-SVM	GA-SVM
Fault_1	100	98.42	99.3	99.3	99.3
Fault_2	99.6	95.8	96.8	97.2	98.3
Fault_3	99.3	93.9	93.2	94.5	97.2
Fault_4	99.1	94.1	93.2	95.6	95.6

Figure 5 Fitness Value Vs Iterations

Figure 6 Model Performances under Single Faults

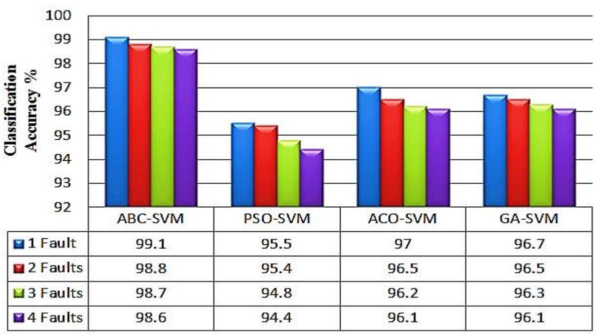


Figure 7 Model Performance under Multiple Faults

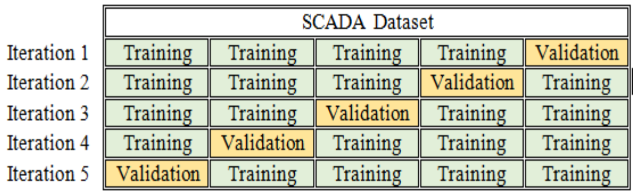


Figure 8 K fold Cross Validation

Table 1 Frequency of Fault Occurrence

S.No	Fault Number	Frequency of Occurrence
1	Fault_1	80
2	Fault_2	92
3	Fault_3	116
4	Fault_4	121

Table 2 Principle Features

	Parameter	Description
1.	Inverter_Current	Inverter Line Current (A)
2.	Blade_1_Motor_Current	Blade 1 Motor Current (A)
3.	Blade_1_Motor_Temp	Blade 1Temperature (^{o}C)
4.	Blade_2_Motor_Current	Blade 2 Motor Current (A)
5.	Blade_2_Motor_Temp	Blade 2 Temperature (^{o}C)
6.	Blade_3_Motor_Current	Blade 3 Motor Current (A)
7.	Blade_3_Motor_Temp	Blade 3 Temperature (^{o}C)
8.	Bearing_Temperature	Temperature of Bearing (^{o}C)
9.	Hub_Temperature	Hub Temperature (^{o}C)
10.	Amb_Temperature	Atmospheric Temperature (^{o}C)
11.	Inv_Temperature	Inverter Temperature (^{o}C)
12.	Pitch	Rotor Pitch
13.	Rotor_Acceleration	Rotor Acceleration (m/sec2)
14.	Pitch_Speed	Actual Pitch Speed
15.	Current_RPM	Current Speed of Wind Turbine
16.	Nacelle_Temperature	Nacelle Temperature (^{o}C)
17.	Vibration_Longitudinal	Longitudinal Axis Vibration
18.	Vibration_Transversal	Vibration along Transversal Axis
19.	Windspeed	Current Wind Speed (m/sec)
20.	Voltage	Grid Voltage (Volt)
21.	Current	Grid Current (A)
22.	Grid_Power	Grid Power (kW)
23.	Reactive Power	Generated Reactive Power (kVAR)

Table 3 Initialization Parameters

PSO	Number of Birds		50
	Moment of Inertia, ω		0.8
	Maximum Number of Iterations		100
	Tuning Parameters	C_1	0.2
		C_2	2.0
ACO	Number of Ants		50
	Maximum Number of Iterations		100
	Pheromone Evaporation Ratio		0.5
	Tuning Parameters	α	0.9
		β	0.1
ABC	Bee Colony Size		100
	Number of Employee Bees		50
	Number of Onlooker Bees		50
	Scout Bee Ratio		0.5
	Maximum Number of Iterations		100
	Tuning Parameters	α	0.9
		β	0.1

Table 5 Performance Metrics of Models

Model	Evaluation Metrics	Fault Types			
		Fault_1	Fault_2	Fault_3	Fault_4
ABC-SVM	Precision	100%	97.85%	98.28%	99.17%
	Recall	98.77%	100%	98.28%	98.35%
	F1 - Score	99.38%	100%	99.13%	99.17%
	Overall Accuracy : 98.60%				
GS-SVM	Precision	95.23%	93.52%	91.18%	92.27%
	Recall	96.13%	94.8%	92.28%	92.25%
	F1 - Score	97.38%	96.12%	95.13%	95.11%
	Overall Accuracy : 92.60%				
PSO – SVM	Precision	96.3%	93.62%	93.86%	94.17%
	Recall	97.50%	95.65%	92.24%	93.39%
	F1 - Score	98.73%	97.78%	95.96%	96.58%
	Overall Accuracy : 94.40%				
ACO – SVM	Precision	98.73%	96.70%	95.65%	94.35%
	Recall	97.50%	95.65%	94.83%	96.69%
	F1 - Score	98.11%	97.17%	96.74%	97.70%
	Overall Accuracy : 96.10%				

GA-SVM	Precision	98.73%	96.81%	98.25%	95.90%
	Recall	97.50%	98.91%	96.55%	96.69%
	F1 – Score	98.11%	98.82%	97.63%	97.70%
	Overall Accuracy : 97.30%				

Table 4 Confusion Matrix of Models

ABC-SVM		**Fault _1**	**Fault _2**	**Fault _3**	**Fault _4**	**Precision**
	Fault_1	80	0	0	0	100.00%
	Fault_2	0	92	2	0	97.85%
	Fault_3	0	0	114	2	98.28%
	Fault_4	1	0	0	119	99.17%
	Recall %	98.77	100	98.28	98.35	**98.6%**
	F1Score %	99.38	100	99.13	99.17	
	Overall Accuracy					

PSO - SVM		**Fault _1**	**Fault _2**	**Fault _3**	**Fault _4**	**Precision**
	Fault_1	78	0	3	0	96.30%
	Fault_2	0	88	2	4	93.62%
	Fault_3	0	3	107	4	93.86%
	Fault_4	2	1	4	113	94.17%
	Recall %	97.50	95.65	92.24	93.39	**94.40%**
	F1Score %	98.73	97.78	95.96	96.58	
	Overall Accuracy					

ACO-SVM		**Fault _1**	**Fault _2**	**Fault _3**	**Fault _4**	**Precision**
	Fault_1	78	0	0	1	98.73%
	Fault_2	0	88	2	1	96.70%
	Fault_3	0	3	110	2	95.65%
	Fault_4	2	1	4	117	94.35%
	Recall	97.50	95.65	94.83	96.69	96.10%
	F1-Score	98.11	97.17	96.74	97.70	
	Overall Accuracy					

GA-SVM		**Fault _1**	**Fault _2**	**Fault _3**	**Fault _4**	**Precision**
	Fault_1	78	0	0	1	98.73%
	Fault_2	0	91	2	1	96.81%
	Fault_3	0	0	112	2	98.25%
	Fault_4	2	1	2	117	95.90%
	Recall	97.50	98.91	96.55	96.69	**97.30%**
	F1-Score	98.11	98.82	97.63	97.70	
	Overall Accuracy					

Table 6 Performance Metrics of ABC-SVC across 5 Folds

	Training Scores in %				**Validation Scores in %**			
	Precision	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**Accuracy**
Fold I	99.32	98.67	98.8	98.91	98.14	98.27	98.82	98.31
Fold II	98.85	99.23	99.41	98.63	98.13	99.21	99.31	98.17
Fold III	98.18	98.38	99.13	98.43	97.88	97.81	98.23	98.03
Fold IV	98.92	98.25	98.67	98.69	98.47	97.92	98.34	98.01
Fold V	98.93	98.91	99.2	98.42	98.23	98.02	98.72	97.92
Average	**98.84**	**98.68**	**99.04**	**98.62**	**98.17**	**98.25**	**98.67**	**98.09**