energies

MDPI

Article

Generator Fault Diagnosis with Bit-Coding Support Vector Regression Algorithm

Whei-Min Lin

School of Mechanical and Electrical Engineering, Tan Kah Kee College, Xiamen University, Zhangzhou 361005, China; wmlin@xujc.com

Abstract: Generator fault diagnosis has a great impact on power networks. With the coupling effects, some uncertain factors, and all the complexities of generator design, fault diagnosis is difficult using any theoretical analysis or mathematical model. This paper proposes a bit-coding support vector regression (BSVR) algorithm for turbine generator fault diagnosis (GFD) based on a support vector machine (SVM) capable of processing multiple classification problems of fault diagnosis. The BSVR can simplify the design architecture and reduce the processing time for detection, where m classifier is needed for m class problems compared to the $[m(m-1)]/2$ size of the original multi-class SVM. Compared with conventional methods, numerical test results showed a high accuracy, good robustness, and a faster processing performance.

Keywords: generator fault diagnosis (GFD); support vector machine (SVM); support vector regression (SVR); bit-coding support vector regression (BSVR)

Citation: Lin, W.-M. Generator Fault Diagnosis with Bit-Coding Support Vector Regression Algorithm. *Energies* **2023**, *16*, 3582. https://doi.org/10.3390/en16083582

Academic Editor: Abu-Siada Ahmed

Received: 16 February 2023
Revised: 17 April 2023
Accepted: 18 April 2023
Published: 20 April 2023

1. Introduction

Shifting to eco-friendly power generation over the last two decades, gas has replaced 4% of the coal energy and 15% of the oil energy in Taiwan. However, the total energy generation from renewables is only 5.8%, which includes 4.1% of the old hydro plants. This can be seen as a typical example of a country shifting to renewables or clean energy. The power generation from the conventional synchronous turbine-generator (TG) remains high, and the large turbine-generator sets will last for a long time.

The generator's fault is that it is often accompanied by vibrations, which not only damage the generator but also cause outages and even an economic threat to society. This paper discusses the fault detection for conventional synchronous machines, and the same idea can be extended to machines with a similar shaft and generator.

With the coupling effects, some uncertain factors, and all the complexities of generator design, highly non-linear and sometimes uncertain vibration features can occur, and fault diagnosis is difficult using any theoretical analysis or conventional mathematical model [1,2]. With the recent emphasis on renewables, the fault detection of wind generators has become a popular topic, and there are many publications in this field. In addition to the generator, the unique structure of the windmill needs attention. Common problems encountered are from the drive train [3,4], the gearbox [5], the bearing [6], the imbalance [7,8], the braking system [9], and sometimes the sensor and the actuator [10]. Methods used are conventional current models [3,6], frequency analysis [11], etc. However, with advances in artificial intelligence (AI), an optimized solution becomes possible, such as the use of an artificial neural network (ANN) [12–14], the fuzzy logic (FL) theorem [15,16], the fuzzy neural network (FNN) [17], and the expert system [18]. Diagnosis that considers the vibration feature can gain a higher accuracy [19,20]. Combining vibration features with AI could also provide promising results, such as the ANN based method [19], wavelet techniques [21], and the FL-based scheme [22,23].

The ANN technique uses error-back propagation to adjust weighting parameters, achieving the desired diagnosis through a nonlinear mapping relationship. Slower learning

processes, weight interferences among fault patterns, and the local minimum are the major drawbacks. The wavelet neural network (WNN) is successful for pattern recognition problems, such as classifying voices and images [24]. The WNN application on TG fault diagnosis appeared in [21]. Initialization of many parameters is important to the success of the complicated WNN network, but the learning time is too long for practical applications. On the other hand, the fuzzy based method strongly depends on experienced experts. The defuzzifier and inference rules must be continuously maintained and revised.

The SVM [25–35] technique is based on a linear machine, nonlinearly related to the input space in a high dimensional feature space, known for fast training even with a large number of input variables and big training sets [32]. The SVM for classification (SVC) is strong for binary classifications. The standard SVC can combine with the one-versus-one (OVO) or one-versus-rest (OVR) approach to solve multi-class problems [36,37]. However, these methods may sometimes suffer from the huge network size, complicated training data preparations, or the heuristic solution scheme.

This paper improves the SVC by the use of the bit-coding approach, combining SVM with a regression technique called bit-coding support vector regression (BSVR), where training becomes an easy and fast regression problem, capable of processing multiple classification problems, including generator fault diagnosis (GFD). The classification problem converges and is coded into a single output bit. BSVR simplifies the design architecture and reduces the processing time. Only m SVR is needed for m classification problems, in comparison with the $[m(m-1)]/2$ size of the original SVM. Numerical results show that BSVR can effectively process multiple GFD, which is not easily attainable with other methods.

2. Problem Description

A turbine generator consists of three parts: the turbine, the generator, and the exciter. The turbine can have sections with high, low, and intermediate pressure sections, girdled by the bearings for diagnostic information [38].

Three faults are generally considered:

(1) The electrical fault.
(2) Mechanical vibration fault.
(3) The cooling system fault.

The first two are the common incipient faults in the generator. The electrical fault could be the stator-winding ground fault, the rotor excitation short circuit, and the stator-winding short circuit involving the three-phase fault or line-to-line fault. The rotor fault causes the unbalanced magnetic pull, and vibration could occur. The major mechanical vibration faults studied in the paper are:

(1) Rotor unbalance: unbalanced rotor weighting or the poor base.
(2) No orderliness: shaft straight line or insufficient warming.
(3) Oil membrane oscillation: fault lubricating system or pump.

Physically, (1) and (2) may twist the shaft and deform the turbine, and (3) could cause the bearing to burn. Shaft vibration regulated with a pre-set limit is the restriction to shut down and trip the machine automatically. The displacement meter and accelerometer are generally set at the bearing position. The power spectra data of the vibration signals are important pieces of information to detect the machinery conditions.

Many techniques were used to find the fault symptoms, such as frequency analysis, phase analysis, precession, vibration waveform, and probability density analysis [39]. In addition, FFT is effective at extracting time-domain signals and converting to the frequency-domain for analysis. The frequency spectra of the vibration will continue to increase as incipient faults continue to grow. Table 1 shows the data [38,39] by studying many 50 MW generators, where vibration signals are the high and low limits. Many methods were used for diagnosis [40,41]. In [38,39], typical vibration frequencies (VFs) used are $1f$, $2f$, $3f$, plus three bands $<0.4f$, $0.4f \sim 0.5f$, $>3f$, where f is the rotor frequency, f_i is the i-th vibration

frequency, and A_{fi} is the amplitude of f_i. For VF, the high and low values of the power spectrum would indicate the probable fault symptoms, as shown in Table 1 [38]. It took a lot of effort and resources to engage in the field survey. As generator fault records are not easy to come by, extensive field work should be conducted to collect new data. According to the field data, diagnostic information can be generated to monitor mechanical conditions by using the spectrum of the vibration signals. The fault classes chosen were based on the available database, originally built for the study of vibrations.

Table 1. Limits of Power Spectra in the Frequency Domain [38].

Fault Type	Rotor Vibration Frequency (VF, in Hz)		A_{fi} Amplitude Range (μm) (10^{-6} m)	
			Lower Value	Upper Value
Oil Membrane Oscillation	f_1	<0.4 f	2.70	6.50
	f_2	1 f	11.00	19.00
	f_3	2 f	1.10	4.90
	f_4	3 f	0.80	2.40
	f_5	>3 f	0.50	3.80
Unbalance (Imbalance)	f_1	<0.4 f	0.54	2.70
	f_2	1 f	38.00	54.50
	f_3	2 f	2.70	6.80
	f_4	3 f	0.54	4.10
	f_5	>3 f	0.00	2.70
No Orderliness	f_1	<0.4 f	0.54	1.90
	f_2	1 f	22.00	30.00
	f_3	2 f	22.00	26.50
	f_4	3 f	14.00	19.50
	f_5	>3 f	5.40	16.20
Normal Condition	f_1	<0.4 f	0.00	0.54
	f_2	1 f	0.00	8.60
	f_3	2 f	0.00	3.30
	f_4	3 f	0.00	3.30
	f_5	>3 f	0.00	2.7

3. Fundamental Theory

Support Vector Regression (SVR) is a regression technique based on SVM. It can be designed to work in the high dimensional feature space when dealing with highly complicated data by mapping the N-dimensional input vector into a K-dimensional feature space (K > N) through a kernel function. The typical kernel functions used are the radial basis function, and the polynomial or sigmoid functions.

Let $\Re^p$ be the p-dimensional real space, and the input and output sets be $\mathbf{X} \subseteq \Re^n$ and $\mathbf{Y} \subseteq \Re$. The example of a training set becomes

$$\mathbf{Tr} := \{\mathbf{x}_i, y_i\} \subseteq \mathbf{X} \times \mathbf{Y},\ i = 1, \ldots,\ l \tag{1}$$

Let linear regression be

$$f_R(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T\mathbf{x} + b = \langle \mathbf{w}, \mathbf{x} \rangle + b \tag{2}$$

Vapnik introduced an error function, with ε-insensitivity. We can isolate a zone for the target output $y \in \mathbf{Y}$ [25,35] by

$$|y - f_R(\mathbf{x}, \mathbf{w})_\varepsilon| = \begin{cases} 0 & if|y - f_R(\mathbf{x}, \mathbf{w})| \leq \varepsilon \\ |y - f_R(\mathbf{x}, \mathbf{w})| - \varepsilon & otherwise \end{cases} \tag{3}$$

The empirical risk of training error is

$$R^{\varepsilon}_{Emp}(\mathbf{w}, b) = \frac{1}{l}\sum_{i=1}^{l}\left|y_i - \mathbf{w}^T\mathbf{x_i} - b\right|_\varepsilon \tag{4}$$

For regression, the objective is to minimize R containing R^{ε}_{Emp} with the estimation error $\|\mathbf{w}\|^2$ [25,35]. That is,

$$R = 2^{-1}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}\left|y_i - \mathbf{w}^T\mathbf{x_i} - b\right|_\varepsilon \tag{5}$$

From Figure 1, all training data outside ε tube satisfy

$$|y - f_R(\mathbf{x}, \mathbf{w})| - \varepsilon = \xi, \text{ for data "above" an } \varepsilon \text{ tube} \tag{6}$$

$$|y - f_R(\mathbf{x}, \mathbf{w})| - \varepsilon = \eta, \text{ for data "below" an } \varepsilon \text{ tube} \tag{7}$$

Minimizing risk R in (5) is equivalent to minimizing

$$R(\mathbf{w}, \xi, \eta) = 2^{-1}\|\mathbf{w}\| + C\left\{\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\eta_i\right\} \tag{8}$$

subject to

$$y_i - [\langle \mathbf{w}, \mathbf{x}\rangle + b] \leq \varepsilon + \xi_i,\ \xi_i \geq 0 \tag{9}$$

$$[\langle \mathbf{w}, \mathbf{x}\rangle + b] - y_i \leq \varepsilon + \eta_i,\ \eta_i \geq 0, \text{ for } i = 1, 2, \ldots, l \tag{10}$$

where C is a user defined trade-off between the risk and the estimation. ξ_i and η_i are slack variables of the measurement of upper and lower bounds of outputs.

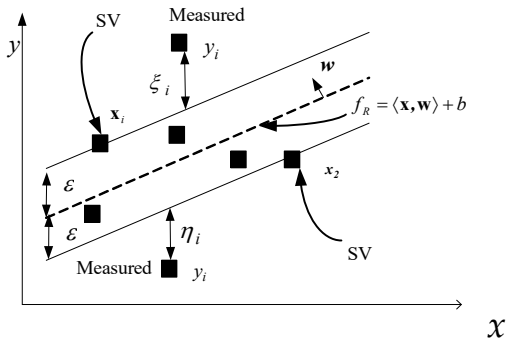

Figure 1. The slack variables and linear function of SVR.

Based on the empirical risk, we can derive the regression formula. Applying Karush-Kuhn-Tucker condition (KKT), the dual Lagrangian $L_d(\alpha, \beta)$ function to maximize is

$$\begin{aligned} L_d(\alpha, \beta) = & \sum_{i=1}^{l} (\alpha_i - \beta_i) y_i - \varepsilon \sum_{i=1}^{l} (\alpha_i + \beta_i) \\ & -2^{-1} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \beta_i)(\alpha_j - \beta_j) \langle \boldsymbol{x_i}, \boldsymbol{x_j} \rangle \end{aligned} \tag{11}$$

This is subject to

$$\sum_{i=1}^{l} (\alpha_i - \beta_i) = 0,\ 0 \le \alpha_i \le C,\ 0 \le \beta_i \le C,\ i = 1,\ 2,\ \ldots, l \tag{12}$$

where α_i and β_i will be nonzero for training samples "above" and "below" an ε tube. Let

$$K(\mathbf{x}, \mathbf{z}) = \langle \phi(\mathbf{x}), \phi(\mathbf{z}) \rangle,\ \mathbf{x}, \mathbf{z} \in \mathbf{X} \tag{13}$$

where ϕ is the feature map designed by mapping the input space X to the feature space F equipped with the inner product $\langle \cdot, \cdot \rangle$. In non-linear cases, we can use $K(\mathbf{x_i}, \mathbf{x_j})$ to replace $\langle \mathbf{x_i}, \mathbf{x_j} \rangle$ in (11).

We have

$$\begin{aligned} L_d(\alpha, \beta) = & \sum_{i=1}^{l} (\alpha_i - \beta_i) y_i - \varepsilon \sum_{i=1}^{l} (\alpha_i + \beta_i) \\ & -2^{-1} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \beta_i)(\alpha_j - \beta_j) K(\boldsymbol{x_i}, \boldsymbol{x_j}) \end{aligned} \tag{14}$$

This abides by the same constraints in (12). Now, (14) can be expressed in the matrix notation by [33]

$$L_d(\boldsymbol{A}) = \mathbf{0.5} \boldsymbol{A^T} \mathbf{H} \boldsymbol{A} + \boldsymbol{c^T} \boldsymbol{A} \tag{15}$$

where $c = [\varepsilon - Y\ \varepsilon + Y]^T$, $\boldsymbol{A} = [\boldsymbol{\alpha}\ \boldsymbol{\beta}]^T$, $(\boldsymbol{\alpha})_i = \alpha_i$, $(\boldsymbol{\beta})_i = \beta_i$. $\mathbf{H}$ denotes the Hessian matrix with

$$\mathbf{H} = \begin{bmatrix} \mathbf{G} & -\mathbf{G} \\ -\mathbf{G} & \mathbf{G} \end{bmatrix} \tag{16}$$

where $\mathbf{G}_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$.

If ${\alpha_i}^o$ and ${\beta_i}^o$ ($i = 1,\ 2,\ \ldots,\ l$) can be found to solve the problem, define ${\gamma_i}^o := {\alpha_i}^o - {\beta_i}^o$, and then the training points with ${\gamma_i}^o \neq 0$ are the SVs and (14) depends entirely on the SVs. We have

$$\mathbf{w}^o = \sum_{i=1}^{l} (\alpha_i^o - \beta_i^o) \phi(\mathbf{x_i}) = \sum_{i \in SVs} \gamma_i^o \phi(\mathbf{x_i}) \tag{17}$$

From KKT, choosing any $0 < \alpha_k^o < C$, we have $\xi_k^o = 0$

$$0 = \varepsilon + \xi_k^o - y_k + \langle \mathbf{w^o}, \mathbf{x_k} \rangle + b^o\ = \varepsilon - y_k + \langle \mathbf{w^o}, \mathbf{x_k} \rangle + b^o \tag{18}$$

We can solve

$$b^o = y_k - \varepsilon - \sum_{i \in SVs} (\alpha_i^o - \beta_i^o) K(\mathbf{x_i}, \mathbf{x_k}) = y_k - \varepsilon - \sum_{i \in SVs} \gamma_i^o K(\mathbf{x_i}, \mathbf{x_k}) \tag{19}$$

We can then find the optimal regression function used in this paper as

$$\begin{aligned} f_R^o(\mathbf{x}) & = \langle \mathbf{w}^o, \mathbf{x} \rangle + b^o \\ & = \sum_{i \in SVs} (\alpha_i^o - \beta_i^o) K(\mathbf{x_i}, \mathbf{x}) + b^o \\ & = \sum_{i \in SVs} \gamma_i^o K(\mathbf{x_i}, \mathbf{x}) + b^o \end{aligned} \tag{20}$$

In Figure 2, we can see that a SVR is represented as a feed forward neural network, where the number of SVs can determine the number of hidden units. One SVR structure is built for one fault class with its dataset $\mathbf{X}_i$ and its target y_i, as in (1). Training has become an easy and fast optimization problem by computing parameters. When training is finished, parameters b^0 and ${\gamma_i}^o \neq 0$ are found and frozen, and we can implement the optimal regression function (20) with kernel $\mathbf{G}_{ij} = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$, as in Figure 2.

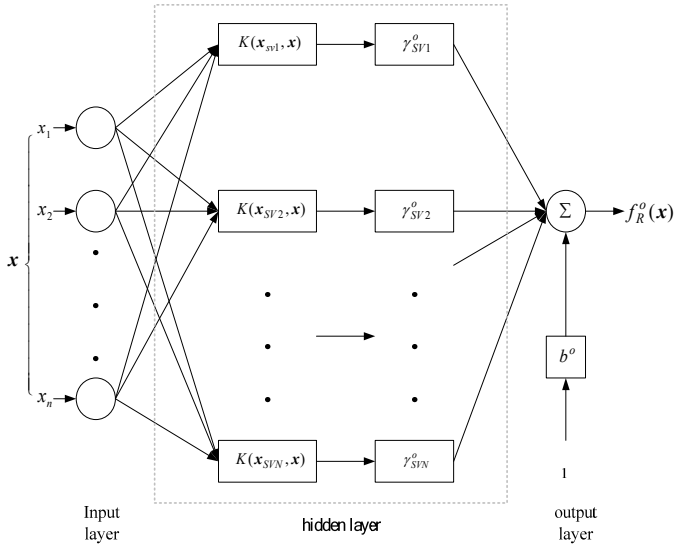

Figure 2. Structure of the SVR network.

4. The Proposed BSVR Classifier

4.1. Design Architecture and Bit-Coding Approach

According to the field data sorted, diagnostic information can be generated. The major vibration problems concerned are rotor unbalance, rubbing, rotor crack, and oil membrane oscillation, i.e., the purposes of the original field investigation for conventional generators. BSVR contains an input layer, SVR layer, and the output layer. Figure 3 shows the BSVR structure, where one SVR unit is designed for one fault class, i.e., m classifier is needed for m class problems against the original $[m(m-1)]/2$ sized SVM problem. This shows a greatly simplified structure of BSVR.

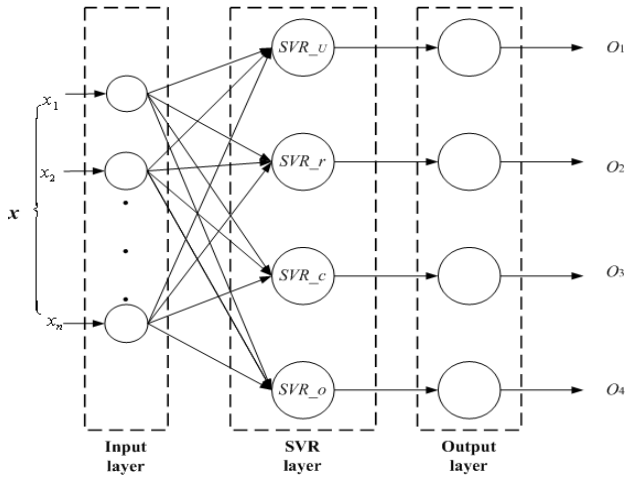

Figure 3. Structure of the proposed BSVR.

In our example with four faulty classes, only four nonlinear SVRs are needed with the bit-coding approach to solve the nonlinear GFD problem, which shows a greatly simplified

classification problem. The network can easily expand with more fault classes without any limitation.

As in Table 2, the targets of each SVR are encoded by "1" or "0", with y_i set to "1" for fault, and "0" otherwise. The GFD problem can be processed as follows:

(1) Construct the diagnostic system.
(2) Train the network with known fault patterns.
(3) Use the diagnostic system to diagnose the fault.

Outputs [O_1 O_2 O_3 O_4] denote the detected fault types, including the rotor unbalance, rubbing, rotor crack, or oil membrane oscillation, respectively. The physical meaning of "1" means "accurate detection," and 0 means "accurate non-detection."

Table 2. The Targets of each SVR in the Training Set.

SVR Type	Targets [O_1 O_2 O_3 O_4]
SVR_u	rotor unbalance: 1; otherwise: 0
SVRM_r	rubbing: 1; otherwise: 0
SVR_c	rotor crack: 1; otherwise: 0
SVM_o	oil membrane oscillation: 1; otherwise: 0

4.2. BSVR-Based GFD System (BGFDS)

The process of running the proposed BGFDS are shown below, where BGFDS works with existing equipment without extra measurement devices. We have:

1. Data acquisition to collect sampled vibration data periodically at a regular interval.
2. Data sent to the Data Processor for interpretation.
3. FFT to analyze data in order to acquire the frequency spectrum.
4. Data processed by the fault diagnosis processor with two portions of
   (1) sampled data from field be constructed in EXCEL Workspace, and
   (2) data analysis and storage of BSVR be manipulated on this database.

A fair threshold of 0.5 is preselected to separate "*fault*" from "*no fault*". We have $O_m \geq 0.5$ for when component m is faulty, and $O_m < 0.5$ for when component m is not faulty.

A flow chart is provided in Figure 4 to show the process. The fault diagnosis processor includes: (1) the sampling data from field measurements in EXCEL Workspace, and (2) the data analysis and storage for BSVR. The diagnostic algorithms can be realized in a PC-based device with the virtual instrument and the hardware device. With the simple structure proposed, the database of EXCEL Workspace is sufficient to support the application. It is simple and easy to use, and it is readily available on most PCs.

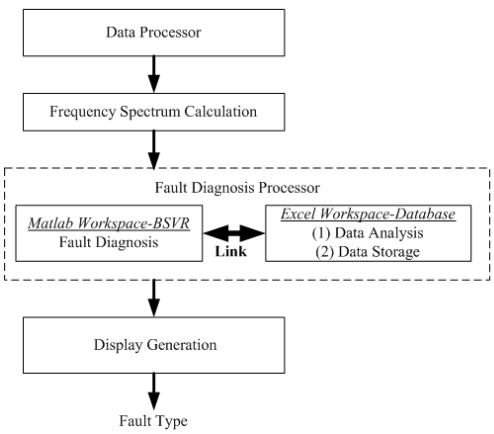

Figure 4. PC-based *BGFDS* for generator fault detection.

5. Simulation and Tests

5.1. Training Patterns Creation

Let the rotor frequency be f, with typical fault features described by swing ratios in nine bands, i.e., (0~0.39) f, (0.40~0.49) f, 0.5 f, (0.51~0.99) f, 1f, 2 f, (3~5) f, odd f, and >5f [23]. The BSVR was trained with 80 samples of faults, 20 samples each category as:

- Rotor unbalance (F_1): [1 0 0 0] with 20 instances.
- Rubbing (F_2): [0 1 0 0] with 20 instances.
- Rotor crack (F_3): [0 0 1 0] with 20 instances.
- Oil membrane osc. (F_4): [0 0 0 1] with 20 instances.

Every fault class has its dataset $\mathbf{X}_i$ and target y_i for training. Training will finish with parameters b^0 and ${\gamma_i}^o \neq 0$ found. We can obtain optimal f^o as in (20) for output detection.

5.2. Simulation Results

The design architecture of BSVR is shown in Table 3. With the normalized input signals [42], ε-insensitivity was set in the range ($10^{-3} - 10^{-2}$). In this paper, $\varepsilon = 0.01$ is chosen and the Gaussian Radial Basis Function (GRBF) kernel is used with

$$K(\boldsymbol{x}, \boldsymbol{z}) = \exp\left[\sum_{i=1}^{9} \frac{(x_i - z_i)^2}{2\rho^2}\right], \rho = 0.8, \ \boldsymbol{x}, \boldsymbol{z} \in \Re^9 \tag{21}$$

Table 3. Architecture of the BSVR.

Network Size			Number of Training Sets	C
I	**S**	**O**		
9	4	4	80	50

Note: I-S-O: Input Layer–SVR Layer–Output Layer.

This research also developed a back-propagation neural network (BPNN) for comparison. The popular BPNN was designed to have three layers, trained with the back-propagation learning algorithm [43], using the same number of Input and Output nodes as BSVR. BPNN has one hidden layer, and the number of hidden nodes is from the experience formulas shown in [44–46].

Table 4 shows the test configuration of the BPNN. Proper parameters of BSVR and BPNN were determined by using the trial-and-error process. The training parameters are set with C = 50 and L = 0.2. To show the performance and effectiveness of the proposed BGFDS, many tests were conducted and a few cases were chosen for demonstration.

Table 4. Architecture of the BPNN.

Network Size			Number of Training Sets	Learning Rate (L)
I	**H**	**O**		
9	7	4	80	0.2

Note: I-H-O: Input Layer–Hidden Layer–Output Layer.

For the accuracy test, the percentage root mean square error (PRMSE) is adopted for each output node by

$$PRMSE = \sqrt{\frac{1}{Q}\sum_{p=1}^{Q} (E(p) - O(p))^2} \times 100\% \tag{22}$$

where

$E(p)$ Target value of the p-th sample
$O(p)$ Real output value of the p-th sample
Q Total number of test samples

5.2.1. Generalization Ability Test

After training finished, this study chose 14 fault samples for tests, as shown in Table 5. These samples do not belong to the group of training data, showing the power spectra of some field data in each frequency band. Data analysis shows the fault type of each sample, e.g., sample No.1~3 are rotor unbalance faults. Bold italic type represents the items with additional noises added in the robustness tests. Noises were added by adjusting the amplitude of A_{fi} (10^{-3} m) as expained in Table 1 by either increasing or decreasing the amplitude, as indicated in the last row. The data were fed to the trained BSVR and BPNN respectively. The detection and non-detection outputs are shown in Figures 5 and 6. In Figure 5, we can see that both methods can detect the rotor unbalance fault with values above 0.5. For example, sample No.1 has a BSVR value of 0.9247, closer to 1, compared to 0.6895 of BPNN. BSVR provides aggregated data pointed along the expectation line, while BPNN produced scattered points with higher errors.

Table 5. Test Dataset for Simulation.

Sample No.	Input Data									AFT
	(0.01~0.39) *f*	(0.4~0.49) *f*	0.5 *f*	(0.51~0.99) *f*	1 *f*	2 *f*	(3~5) *f*	Odd *f*	>5 *f*	
1	***0.00256***	0.00122	***0.00993***	***0.01826***	0.81123	***0.07904***	0.04958	0.04958	0.00397	F_1
2	0.05129	0.00267	0.00227	***0.01846***	***0.7578***	0.09388	0.03373	***0.03373***	0.00596	F_1
3	0.00494	***0.00162***	0.00131	***0.01049***	***0.84174***	0.05299	**0.01962**	0.01962	0.00322	F_1
4	0.11805	0.01598	***0.00831***	0.12527	0.56643	***0.01725***	0.04653	0.02356	***0.07864***	F_2
5	0.03012	0.01275	0.02175	***0.16904***	***0.61279***	0.01977	0.05657	0.02518	***0.052***	F_2
6	0.1167	***0.00545***	***0.00523***	0.17401	0.56365	0.02107	0.05358	***0.01288***	0.05344	F_2
7	0.00344	0.00344	0.00553	0.00723	***0.54074***	0.15488	0.12893	0.12893	0.02687	F_3
8	***0.00178***	0.00178	0.00323	***0.00566***	0.58058	0.15624	***0.11422***	0.11422	0.0223	F_3
9	0.0132	0.00261	***0.00281***	0.00642	***0.63413***	0.14974	0.07667	***0.07667***	0.0377	F_3
10	0.02475	0.18273	0.39201	0.19642	0.05736	0.09657	0.02254	***0.02254***	***0.0051***	F_4
11	0.00482	0.24	0.50575	0.07214	***0.08549***	0.03526	***0.02535***	0.02535	***0.0059***	F_4
12	0.02363	0.14473	***0.53938***	0.10211	0.05216	***0.091177***	0.02069	0.02069	0.00548	F_4
13	0.00755	0.26129	0.4818	0.0761	***0.08415***	***0.03498***	0.02331	***0.02331***	0.0055	F_4
14	0.01321	***0.23394***	0.488	***0.06358***	0.09938	0.03841	***0.02777***	***0.02777***	0.00791	F_4
Noise	−40%	20%	−10%	+10%	+10	−20%	+30%	+20%	−20%	

Note: 1. AFT: Actual fault types; 2. NS: Noise; 3. F_1: rotor unbalance; F_2: rubbing; F_3: rotor crack; F_4: oil membrane oscillation; 4. Bold italic type represents the items with the last row noises added in the robustness test.

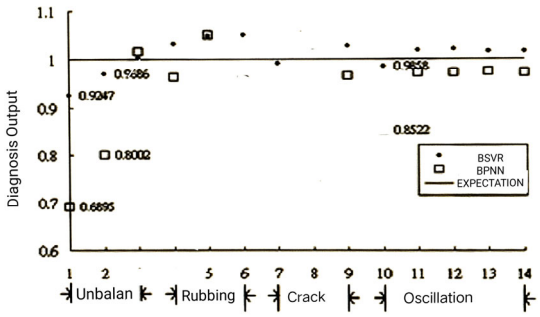

Figure 5. BSVR and BPNN output of detection.

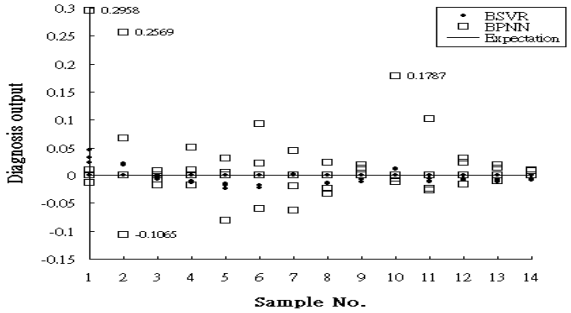

Figure 6. BSVR and BPNN output of non−detection.

Figure 6 shows the non-detection with range −0.15~0.3 blown up for more detail. It has the "detection" output of Figure 5 placed on expectation line "0" for reference. Another three outputs of each method are scattered around. For example, sample No.1 of BPNN has the worst value of 0.2958 for rotor unbalance fault non-detection. BSVR provides aggregated data points along the expectation line. The error indices PRMSE are in Table 6 to show that BSVR has better accuracy.

Table 6. The PRMSE of Each Output Node.

Method	The PRMSE (%) Value			
	O_1	O_2	O_3	O_4
BSVR	2.408	2.449	1.897	1.673
BPNN	10.752	6.746	10.817	4.754

5.2.2. Robustness Test

To test the robustness, 10% to 40% of noise was added to the amplitudes of the original field data in Table 5, as noted in bold italic type. The output results with detection and non-detection are shown in Figures 7 and 8. We can see that BSVR has a high noise rejection capability, with output values near the expectation line. Figure 7 shows a range of 0.9~1.06 for detection. Figure 8 blows up the range −0.03~0.06 with increment 0.01 for more detail, since the points tend to cluster with high accuracy.

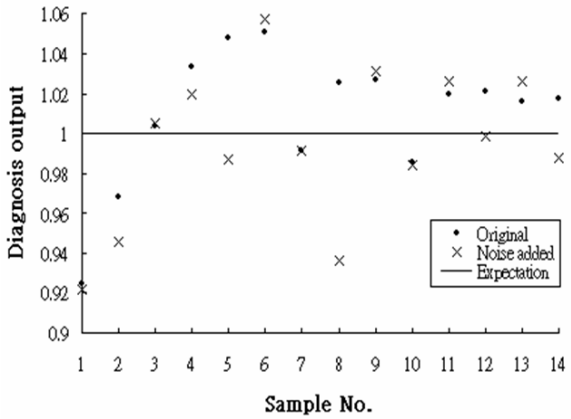

Figure 7. The output of detection for robustness test.

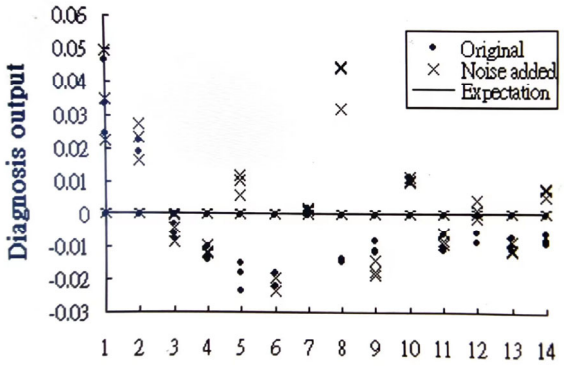

Figure 8. The output of non−detection for robustness test.

5.2.3. Consistency Test

To check the consistency of BSVR and BPNN, this test retrains and retests both methods to a specific n number of times. After testing for the given number of times, the mean value μ and the standard deviation σ were recorded for each output O_i.

Part of the results for consistency tests are shown in Table 7. From the tables, we can see that BSVR has the same mean value and "zero" standard deviation for each retrain and retest. The important consistent machine learning character of BSVR shows a more robust nature than BPNN.

Table 7. The Consistency Test of the Two Methods (mean value and standard deviation).

Method	n	**SN**	**Mean Value of Each Output Node O_1~O_4**			
			μ_1	μ_2	μ_3	μ_4
BSVR	12	1	0.9247	0.0331	0.0466	0.0242
	24	1	0.9247	0.0331	0.0466	0.0242
	12	2	0.9686	0.0190	0.0000	0.0226
	24	2	0.9686	0.0190	0.0000	0.0226
BPNN	12	1	0.7922	−0.0117	0.2260	−0.0079
	24	1	0.8085	−0.0215	0.2212	−0.0089
	12	2	0.8596	−0.0060	0.1491	0.0028
	24	2	0.8598	0.0195	0.1217	0.0023
Method	n	SN	Standard deviation of each output node O_1~O_4			
			σ_1	σ_2	σ_3	σ_4
BSVR	12	1	0	0	0	0
	24	1	0	0	0	0
	12	2	0	0	0	0
	24	2	0	0	0	0
BPNN	12	1	0.0831	0.0317	0.0649	0.0186
	24	1	0.0833	0.0466	0.0605	0.0225
	12	2	0.0788	0.0740	0.1031	0.0252
	24	2	0.0932	0.0831	0.1022	0.0417

Note: n number of test times; SN sample no. in Table 5.

5.2.4. Performances Test

The learning performance of BSVR and BPNN are shown in Table 8. The BPNN was one of the tools developed in house for much research [45,46]. To compare a common base with any hardware, the training and the testing time were normalized with BSVR as the bases. When training finishes, all weights are frozen. From the table, we can see

that the training time of BSVR outperformed BPNN substantially, while the testing time of both methods are close to each other. BSVR has a fast learning process, and requires no estimation for the number of layers or for the number of hidden nodes. With similar training data, the proposed BSVR yields a better overall performance than BPNN.

Table 8. Performance Comparison of BSVR and BPNN.

Method	tr	te	*C*	*L*	Tr_ep	Tr_t (*sec*)	Te_t (*sec*)
BSVR	80	14	50	-	-	1	1
BPNN	80	14	-	0.2	50 k	13.89	1

Note 1: tr: Number of training data; te: Number of testing data. *Note* 2: Tr_t: Training time; Te_t: Testing time; Tr_ep: Training epochs. *Note* 3: *L*: Learning rate of BPNN. *Note* 4: Training and testing time of BSVR is set to 1.

6. Conclusions

In this paper, the BSVR-based Generator Fault Diagnosis System (BGFDS) was proposed. BSVR was designed with a simple network architecture to shorten the processing time. In a turbine generator set, vibration signals are available through data acquisition (DA), and BSVR requires no extra devices. The proposed architecture could effectively detect machine faults with information provided by the frequency spectrum of vibration signals. The BGFDS has many advantages:

- BSVR integrates a bit-coding approach, simple SVR, and a small number of training data for problems that are not linearly separable as GFD.
- SVR trains SVs with standard quadratic optimization technique, which has a unique solution and is globally optimal, requiring much less computation time.
- SVR needs no determination for the size of hidden layers. The number of SVs determines the number of hidden units automatically.
- The training and testing of BSVR are very fast compared with other ANNs.
- A minimum sized network is built with simple learning algorithms.
- Only m SVR is needed for the m classification problem in comparison with $[m(m-1)]/2$ size for traditional multi-class SVM (MSVM).
- The design architecture can use existing devices without adding extra measurement devices.
- BSVR has good classification capability, performance, consistency, noise rejection ability, and robustness, i.e., many good characteristics for machine learning.
- The proposed diagnostic algorithms can be realized in a portable device for convenient application.

Test results show that BSVR-based GFDS are precise, very effective, expandable, and easy to work with.

Funding: This research was funded by Tan Kah Kee College, Xiamen University, grant number JG2021SRF01.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to give special thanks to Chien-Hsien Wu for the very useful data and material when writing this manuscript.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Doorwar, A.; Bhalja, B.; Malik, O.P. A New Internal Fault Detection and Classification Technique for Synchronous Generator. *IEEE Trans. Power Deliv.* **2019**, *34*, 739–7492. [CrossRef]
2. Bouzid, M.B.K.; Champenois, G. An Efficient Simplified Physical Faulty Model of a Permanent Magnet Synchronous Generator Dedicated to Stator Fault Diagnosis Part II: Automatic Stator Fault Diagnosis. *IEEE Trans. Ind. Appl.* **2017**, *53*, 2762–2771. [CrossRef]

3. Gong, X.; Qiao, W. Current-based mechanical fault detection for direct-drive wind turbines via synchronous sampling and impulse detection. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1693–1702. [CrossRef]
4. Siegel, D.; Zhao, W.; Lapira, E.; AbuAli, M.; Lee, J. A comparative study on vibration-based condition monitoring algorithms for wind turbine drive trains. *Wind. Energy* **2014**, *17*, 695–714. [CrossRef]
5. Zhang, Z.; Verma, A.; Kusiak, A. Fault analysis and condition monitoring of the wind turbine gearbox. *IEEE Trans. Energy Convers.* **2012**, *27*, 526–535. [CrossRef]
6. Gong, X.; Qiao, W. Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals. *IEEE Trans. Ind. Electron.* **2013**, *60*, 3419–3428. [CrossRef]
7. Gong, X.; Qiao, W. Imbalance fault detection of direct-drive wind turbines using generator current signals. *IEEE Trans. Energy Convers.* **2012**, *27*, 468–476. [CrossRef]
8. Gardels, D.J.; Qiao, W.; Gong, X. Simulation studies on imbalance faults of wind turbines. In Proceedings of the IEEE PES General Meeting, Minneapolis, MN, USA, 25–29 July 2010; pp. 1–5.
9. Entezami, M.; Hillmansen, S.; Weston, P.; Papaelias, M.P. Fault detection and diagnosis within a wind turbine mechanical braking system using condition monitoring. *Renew. Energy* **2012**, *47*, 175–182. [CrossRef]
10. Wei, X.; Verhaegen, M. Sensor and actuator fault diagnosis for wind turbine systems by using robust observer and filter. *Wind Energy* **2011**, *14*, 491–516. [CrossRef]
11. Lu, D.; Qiao, W. Frequency demodulation-aided condition monitoring for drivetrain gearboxes. In Proceedings of the 2013 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA, 16–19 June 2013; pp. 1–6.
12. Lin, W.M.; Yang, C.D.; Lin, C.H.; Tsay, M.T. A Fault Classification Method by RBF Neural Network with OLS Learning Procedure. *IEEE Trans. Power Deliv.* **2001**, *16*, 473–477. [CrossRef]
13. Zhang, Y.; Ding, X.; Liu, Y.; Griffin, P.J. An Artificial Neural Network Approach to Transformer Fault Diagnosis. *IEEE Trans. Power Deliv.* **1996**, *11*, 1836–1841. [CrossRef]
14. Wan, S.; Li, H.; Zhaofeng, X. A new method of turbine-generator vibration fault diagnosis based on correlation dimension and ANN. In Proceedings of the International Conference on Power System Technology, Kunming, China, 13–17 October 2002; Volume 3, pp. 1655–1659.
15. Islam, S.M.; Wu, T.; Ledwich, G. A Novel Fuzzy Logic Approach to Transformer Fault Diagnosis. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 177–186. [CrossRef]
16. Huang, Y.-C.; Yang, H.-T.; Huang, C.-L. Developing a new transformer fault diagnosis system through evolutionary fuzzy logic. *IEEE Trans. Power Deliv.* **1997**, *12*, 761–767. [CrossRef]
17. Wu, C.Z.; Yan, H.; Ma, J.F. Method Research of Noise Diagnosis Based on Fuzzy Neural Network. In Proceedings of the ICSP'98. 1998 Fourth International Conference on Signal Processing (Cat. No. 98TH8344), Beijing, China, 12–16 October 1998; pp. 1370–1373.
18. Lee, H.-J.; Ahn, B.-S.; Park, Y.-M. A Fault Diagnosis Expert System for Distribution Substations. *IEEE Trans. Power Syst.* **2000**, *15*, 92–974.
19. Wan, S.; Li, H.; Li, Y. Adaptive Radial Basis Function Network and Its Application in Turbine-Generator Vibration Fault Diagnosis. In Proceedings of the International Conference on Power System Technology, Kunming, China, 13–17 October 2002; Volume 3, pp. 1607–1610.
20. Zhang, J.; Li, R.X.; Han, P.; Wang, D.F.; Yin, X.C. Wavelet packet feature extraction for vibration monitoring and fault diagnosis of turbo-generator. In Proceedings of the 2003 International Conference on Machine Learning and Cybernetics, Xi'an, China, 5 November 2003; Volume 1, pp. 340–344.
21. Xie, D.M.; Song, X.; Zhou, H.L.; Guo, M.W. Fuzzy vibration fault diagnosis system of steam turbo-generator rotor. In Proceedings of the International Conference on Machine Learning and Cybernetics, Beijing, China, 4–5 November 2002; Volume 1, pp. 411–415.
22. Ghazali, M.H.M.; Rahiman, W. Vibration-Based Fault Detection in Drone Using Artificial Intelligence. *IEEE Sens. J.* **2022**, *22*, 8439–8448. [CrossRef]
23. Mitchell, J.S. *Machinery Analysis and Monitoring*; Pennwell Publishing Company: Tulsa, OK, USA, 1981.
24. Szu, H.; Brian, T. Neural Network Adaptive Wavelets for Signal Representation and Classification. *Opt. Eng.* **1992**, *31*, 1907–1916. [CrossRef]
25. Kecman, V. *Learning and Soft Computing*; MIT Press: Cambridge, MA, USA, 2001; pp. 11–298.
26. Vapnik, V. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.
27. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
28. Schölkopf, B.; Smola, A.J. *Learning with Kernel*; MIT Press: Cambridge, MA, USA, 2001.
29. Schölkopf, B.; Smola, A.J.; Willianson, R.; Bartlett, P. New support vector algorithms. *Neural Comput.* **2000**, *12*, 207–1245. [CrossRef]
30. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
31. Hsu, C.-W.; Lin, C.-J. A comparison of methods for multi-class support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425.
32. Moulin, L.S.; da Silva, A.P.A.; Sharkawi, M.A.E.I.; Mark, R.J., II. Support Vector Machines for Transient Stability Analysis of Large-Scale Power Systems. *IEEE Trans. Power Syst.* **2004**, *19*, 818–825. [CrossRef]

33. Gunn, S.R. Support Vector Machines for Classification and Regression. *ISIS Tech. Rep.* **1998**, *14*, 5–16.
34. Smola, A.; Schölkopf, B. *A Tutorial on Support Vector Regression*; NeuroCOLT Tech. Rep. NC-TR-98-030; Royal Holloway University: London, UK, 1998; Available online: http://www.kernel-machines.org/ (accessed on 15 February 2023).
35. Ertekin, S.; Bottou, L.; Giles, C.L. Nonconvex Online Support Vector Machines. *IEEE Trans. Pattern Anal. Mach. Intelligence* **2011**, *33*, 368–381. [CrossRef]
36. Deng, J.-L. Control Problems of Grey Systems. *Syst. Control. Lett.* **1982**, *1*, 288–294.
37. Deng, J.-L. Introduction to Grey System Theory. *J. Grey Syst.* **1989**, *1*, 1–24.
38. Wang, M.-H. Application of Extension Theory to Vibration Fault Diagnosis of Generator sets. *IEE Proc. Gener. Transm. Distrib.* **2004**, *151*, 503–508. [CrossRef]
39. Li, H.; Sun, C.X.; Hu, X.S.; Yue, G.; Wang, K. The fuzzy inputting and outputting method in vibration fault diagnosis of steam turbine-generator set. *J. Chongqing Univ.* **1999**, *22*, 36–41. (In Chinese)
40. Kawada, M.; Yamada, K.; Yamashita, K.; Isaka, K. Fundamental Study on Vibration Diagnosis for Turbine Generators Using Wavelet Transform. In Proceedings of the IEEE PES 2004 Power System Conference and Exposition, New York, NY, USA, 10–13 October 2004; Volume 13, pp. 1215–1220.
41. Lin, C.-H. Classification Enhancible Grey Relational Analysis for Cardiac Arrhythmias Discrimination. *Med. Biol. Eng. Comput.* **2006**, *44*, 311–320. [CrossRef]
42. Lin, C.-H.; Huang, P.-Z.; Wu, C.-H.; He, C.-Z. Window Based Assistant Tool for Oil-Immersed Transformer Fault Diagnosis Using Grey Clustering Analysis. In Proceedings of the 27th Symposium on Electrical Power Engineering, Tsing-Hua University, Hsin-Chu, Taiwan, 22–23 December 2006; pp. PD2.20.1–PD2-20.5.
43. Liu, S.F.; Zhao, L.; Wang, Z.Y.; Yi, L. A New Method for Venturous Capital Pricing. *Chin. J. Manag. Sci.* **2001**, *9*, 22–26.
44. Hong, Y.-Y.; Chen, Y.-C. Application of Algorithms and Artificial Intelligence Approach for Locating Multiple Harmonics in Distribution System. *IEE Proc. Gener. Transm. Distrib.* **1999**, *146*, 325–329. [CrossRef]
45. Lin, W.-M.; Lin, C.-H.; Tu, K.-P.; Wu, C.-H. Multiple harmonic source detection and equipment identification with cascade correction network. *IEEE Trans. Power Deliv.* **2005**, *20*, 2166–2173. [CrossRef]
46. Lin, W.-M.; Lin, C.-H.; Sun, Z.-C. Adaptive Multiple Fault Detection and Alarm Processing for Loop System with Probabilistic Network. *IEEE Trans. Power Deliv.* **2004**, *19*, 64–69. [CrossRef]