Locating Faults in Thyristor-Based LCC-HVDC Transmission Lines Using Single End Measurements and Boosting Ensemble

Aleena Swetapadma 1, Shobha Agarwal 2, Satarupa Chakrabarti 1, Soham Chakrabarti 1, Adel El-Shahat 3, and Almoataz Y. Abdelaziz 4

1 School of Computer Engineering, KIIT Deemed to Be University, Bhubaneswar 751024, India; aleena.swetapadma@gmail.com (A.S.); chakrabartisatarupa@gmail.com (S.C.); 1605488@kiit.ac.in (S.C.)
2 Department of Higher Technical Education and Skill Development, Govt. of Jharkhand, Ranchi 834002, India; shobhaagarwal@rediffmail.com
3 Energy Technology Program, School of Engineering Technology, Purdue University, West Lafayette, IN 47907, USA
4 Faculty of Engineering and Technology, Future University in Egypt, Cairo11835, Egypt; almoataz.abdelaziz@fue.edu.eg

* Correspondence: asayedah@purdue.edu

Abstract: Most of the fault location methods in high voltage direct current (HVDC) transmission lines usethemethods which require signals from both ends. It will be difficult to estimate fault location if the signal recorded is not correct due to communication problems. Hence a robust method is required which can locate fault with minimum error. In this work, faults are located using boosting ensembles in HVDC transmission lines based on single terminal direct current (DC) signals. The signals are processed to obtain input features that vary with the fault distance. These input features are obtained by taking maximum of half cycle current signals after fault and minimum of half cycle voltage signals after fault from the root mean square of DC signals. The input features are input to a boosting ensemble for estimating the location of fault. Boosting ensemble method attempts to correct the errors from the previous models and find outputs by combining all models. The boosting ensemble method has been also compared with the decision tree method and the bagging-based ensemble method. Fault locations are estimated using three methods and compared to obtain an optimal method. The boosting ensemble method has better performance than all the other methods in locating the faults. It also validated varying fault resistance, smoothing reactors, boundary faults, pole to ground faults and pole to pole faults. The advantage of the method is that no communication link is needed. Another advantage is that it allows reach setting up to 99.9% and does not exhibit the problem of over-fitting. Another advantage is that the percentage error in locating faults is within 1% and has a low realization cost. The proposed method can be implemented in HVDC transmission lines effectively as an alternative to overcome the drawbacks of traveling wave methods.

Keywords: line commutated converter (LCC); LCC-HVDC transmission system; bagging ensemble; decision tree; boosting ensemble; fault location; single end measurement

1. Introduction

Owing to the boost in energy demand at the moment, bulk power transmission is essential. One of the ideas is to implement HVDC transmission lines as it is cost effective and emerging in the power sector. The advantages of using HVDC transmission lines over alternating current (AC) lines are that they can carry power to long distances, but locating faults in HVDC lines is not an easy task due to the long transmission line length whose impedance is very small, as well as severe environmental conditions. Most of the HVDC transmission lines use traveling waves which have drawbacks such as the need for communication links and both end measurements. Hence, it is essential to design a scheme to locate the faults with minimum error in all operating conditions.
To predict fault location, traveling-wave method based on current data has been proposed in [1]. In [2], a fault direction comparison method using both end measurements based on traveling wave has been proposed. The amplitude of the first backward traveling wave is greater than that of forward traveling wave for a forward fault that is used for direction comparison. In [3], using Prony algorithm, locations are predicted in DC lines. In [4], a traveling wave-based protection has been suggested, but the sampling rate is very high, making it complicated to implement. A traveling wave based directional protection has been proposed for DC transmission lines in [5]. In [6], a distance protection method with traveling wave based on wavelet transform has been suggested. The drawback of the above-mentioned methods [1–6] is that they are based on traveling waves in which different positions have different characteristics.

In [7], a method has been proposed for multi-terminal using electromagnetic time reversal. In [8], a method using fault currents from two ends has been proposed. Using the fault component current at both converter stations, faults are identified, but this cannot detect high impedance fault. A current differential protection method has been suggested in [9] which remove the effects of capacitor current. The sum of the current calculated with each station’s data has been used to detect faults. However, the methods suggested in [7–9] need data from both ends and communication link. A current differential protection scheme for current source converters (CSC) based HVDC transmission lines has been proposed in [10]. However, this method takes more time to operate to avoid the impact of capacitive current. In [11], an impedance-frequency characteristic is used for fault identification. In this approach, transient energy within a specific frequency band has been used to identify internal faults, but it does not emphasize the estimation of fault location. In [12] a wavelet-based approach has been suggested for fault detection, pole identification and fault location, but a drawback of the method is that it cannot estimate location varying resistance.

In recent studies, artificial intelligence-based techniques are also used for locating faults. In [13], a support vector machine (SVM) based method has been proposed to locate the fault. In [14], a deep learning approach called long short-term memory network has been used. In [15], a discrete wavelet transform (DWT) and SVM based technique has been projected for fault location, but it has used a very high sampling frequency and the percentage error is maximum up to 7%. In [16], a deep learning approach has been proposed that can directly learn the fault conditions based on feature. It has been tested on a modular multilevel converter based four-terminal high-voltage direct current system. In [17], a fault location method for a modular multilevel converter (MMC)-based HVDC grid based on dynamic state estimation and gradient descent uses data window of 5 ms after the occurrence of the fault. In [18], a novel fault location approach based on characteristic-harmonic measured impedance (CHMI) has been proposed. In this method, fault branch is identified by the amplitude ratio of the two CHMI. In [19], current limiting reactor power based primary protection scheme and variation tendency of current based the backup protection has been proposed. In [20], a backup protection method for bipolar HVDC lines using two-end synchronized sampled currents has been proposed. It can be applied to detect and recognize the faulted poles without long time delays. In [21], internal faults are partitioned into four zones along the transmission line. Then, the polarities and arrival times of traveling-waves are used for zone partition.

The research gap from the above discussion can be outlined as follows:

- Most of the methods for used are traveling wave-based method. In a traveling wave method, sometimes estimating fault location is difficult because the signal recorded is not strong;
- Another problem with traveling wave methods is that if one end does not record the signal correctly then it cannot estimate the location of fault;
- As traveling waves are an online method, if there is a problem in the communication link between terminals then also it will not be able to estimate the location;
Another drawback of traveling wave method is that if the obstacle in the wave front path is irregular then it will provide wrong result for the location of fault. In these scenarios, the maintenance crew has to patrol the line, which is not economical and takes time, ultimately affecting uninterrupted power flow.

Considering the drawbacks of the existing methods, a scheme for locating faults has been suggested based on single terminal signals using a boosting ensemble learning method. An advantage of the proposed method compared to other methods is that it does not need a communication link and is simple, reliable and robust. Compared with the existing methods which require communication links leading to high cost and low reliability, the proposed method uses signals from one terminal. The proposed method does not need a complex algorithm to improve the location accuracy. Compared to a traveling wave that needs high sampling frequency, the proposed method provides a better result with a low sampling frequency. Hence the proposed method is robust and consistent. The manuscript is organized as follows: Section 2 contains the materials and methods, Section 3 contains the results, Section 4 contains the discussion, Section 5 contains the implantation of proposed method and Section 6 contains the conclusion.

2. Materials and Methods

The fault line can be located offline in DC lines. An effective fault location method is very important because it helps to send the maintenance crew to the fault location which will reduce the repair time and maintenance costs. The proposed fault location estimation scheme has various steps, as shown in Figure 1. It is based on a training module which is designed using three methods, namely decision tree regression (DTR), bagging ensemble regression (BGER) and the boosting ensemble regression (BTER) technique, to correlate input features with fault locations. DC current and voltage signals are recorded at the rectifier end using a sampling rate of 1 kHz. The previously trained module takes the input data and estimates the fault location.
2.1. System under Study

Bipolar LCC-HVDC transmission lines have various advantages, hence are implemented more frequently. Power transmission line is 1100 km long, designed using MATLAB/SIMULINK [22]. As the HVDC lines are long lines, fault occurrence is more common. Power system network used in this study is shown in Figure 2.
i. Obtain the signals from rectifier end;
ii. Obtain the root mean square (rms) value of signals;
iii. Take half cycle samples after occurrence of fault;
iv. Find the maximum of current and minimum of voltage from the half cycle samples that are the input features given to the fault location estimation module.

Step i and step ii are shown in Figure 3 for pole 1 to ground (P1G) fault at 150 km. Figure 3a–g show the current of pole 1, current of pole 2, voltage of pole 1 and voltage of pole 2, respectively. Figure 3e–h show the rms current of pole 1, current of pole 2, voltage of pole 1 and voltage of pole 2, respectively. It is observed that half cycle of the samples after the occurrence of a fault has certain variation with respect to location.

**Figure 3.** Signals obtained during P1G fault at 150 km in 100 ms time in transmission lines.
Step iii is shown in Figure 4, i.e., half cycle signals obtained for P1G fault at various locations. Figure 4a shows that the half cycle current is different for each location. From Figure 4b it can be observed that the half cycle voltage signals are different for each location. Step iii is also shown in Figure 5, i.e., half cycle rms signals obtained during P1G fault at 1080 km for various fault resistance. From Figure 4a it can be observed that the half cycle current signals are different for each fault resistance. From Figure 5b it can be observed that the half cycle voltage signals are different for each fault resistance. Figures 4 and 5 show the correlation between signals obtained after fault and fault distance. So, these features or any variation of these features can be used for fault location estimation.

![Figure 4](image_url1)

**Figure 4.** (a) Half cycle rms current and (b) Half cycle rms voltage during P1G fault at various locations.

![Figure 5](image_url2)

**Figure 5.** (a) Half cycle rms current and (b) Half cycle rms voltage during P1G fault at 1080 km for various fault resistance.
Step iv is shown in Figure 6, i.e., input features obtained for various locations (21 km to 1079 km in step of 2 km) of both poles during P1G fault. From Figure 6a it can be seen that the current features of pole 1 are varying according to location of fault, as the fault is P1G. From Figure 6c it can be observed that the voltage features of pole 1 are also varying according to location of fault as the fault is P1G. Pole2 current and voltage features are shown in Figure 6b and 6d, respectively.

![Figure 6. Input features obtained during P1G fault varying different fault location.](image)

### 2.3. Design of Proposed Method

Machine learning is one of the solutions when any conventional or hard computing method fails to deliver a solution for a problem. Machine learning provides an approximate solution from uncertain problems with partial truths. Fault location estimation of transmission line fits into the area where machine learning can be a solution. Among the various types of regression machine learning method, some methods have many advantages over other methods, such as DTR, BGER and BTER.

Learning methods such as DTR, BGER and BTER are used, which first learns from the example samples before it predicts the outcome when unknown samples are given. Learning based methods have two phases, the training phase and the testing phase. To design the training modules, appropriate features are required. Features given to the training modules for locating the faults are discussed in the above section. Various fault cases used to design the training network and testing. Different cases used for designing the training network are 4 fault type (P1G, pole 2 to ground (P2G), pole to pole to ground (P1P2G) and pole to pole (P1P2)), 1099 location, 40 fault resistance, 5 smoothing reactors. Test cases used are those cases which are not given in training. Fault location estimation module is designed using three methods, namely DTR, BGER and BTER, which will be discussed in the subsection below.

#### 2.3.1. Design of DTR Based Method

DTR is part of the supervised learning algorithms that can solve regression problems. The main aim of the DTR is to train and create a model that can predict different values using decision rules. A set of nodes and edges forms the basic structure of a decision tree. There is the presence of one root node followed by the internal and leaf nodes. In a decision tree, the foremost challenge lies in attribute selection, or identifying the attribute that would
represent the root node and nodes at each level [23]. The choice of strategic split affects the performance of the tree. Entropy for more than one class is represented in (2). Information gain for an attribute (A) is represented in (3). Intrinsic information about an attribute is given in (4). Gain Ratio is defined in (5), Gini Index is calculated in (6) and average Gini Index can be written as in (7)

\[ \text{Entropy} E(S) = \sum_{i=1}^{c} - p_i \log_2 p_i \]  

(1)

where \( S \) is the current state and \( p_i \) is the probability of an event \( i \) of the state \( S \).

\[ E(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 \ldots - p_n \log_2 p_n = -\sum_{i=1}^{n} p_i \log_2 p_i \]  

(2)

\[ \text{Gain}(S, A) = E(S) - I(S, A) = E(S) - \sum_i \frac{|S_i|}{|S|} \cdot E(S_i) \]  

(3)

\[ \text{IntI}(S, A) = -\sum_i \frac{|S_i|}{|S|} \log \left( \frac{|S_i|}{|S|} \right) \]  

(4)

\[ \text{GR}(S, A) = \frac{\text{Gain}(S, A)}{\text{IntI}(S, A)} \]  

(5)

\[ \text{Gini}(S) = 1 - \sum_{i} p_i^2 \]  

(6)

\[ \text{Gini}(S, A) = \sum_i \frac{|S_i|}{|S|} \cdot \text{Gini}(S_i) \]  

(7)

Variance reduction is represented as,

\[ \text{Variance} = \frac{\sum (X - \bar{X})^2}{n} \]  

(8)

where \( X \) is the actual value, \( \bar{X} \) is the mean and \( n \) representing the number of values. Mathematically, Chi-Square is defined as,

\[ \chi^2 = \sum \frac{(O - E)^2}{E} \]  

(9)

where \( O \) is the observed score and \( E \) is the expected score.

The first phase of DTR method is to design the DTR based training module. The input and target, i.e., the corresponding location of the input features are given to the DTR based training network. During designing the training module, it is important to select the optimal value of parameters that are required to locate faults correctly. The choice of strategic split affects the performance of the decision tree. After various trials and errors, it has been found that the Gini index method gives better results than all other methods. The tree pruning was completed after designing the DTR and it has been found that the highest level of the tree is 659. The optimal parameters used for designing the DTR based training network are shown in Table 1. The DTR based training network is tested and results are discussed in the next section.
Table 1. Training parameters used.

| AI based Method | Parameters                      | Value               |
|-----------------|---------------------------------|---------------------|
| DTR             | Splitting Criterion             | Categorical Predictor Split |
|                 | Model                           | Gini index          |
|                 | Number of Levels                | Regression Tree     |
|                 | Splitting Criterion             | Interaction-Curvature |
|                 | Model                           | Bagging             |
| BGER            | Number of trees                 | 659                 |
|                 | Quality of split                | Interaction-Curvature |
|                 | Method                          | Bagging             |
| BTER            | Learn Rate                      | 300                 |
|                 | Learning Cycles                 | Mean square error   |
|                 |                                 | LSBoost             |

2.3.2. Design of BGER Based Method

With the gradual growth and improvement in the field of machine learning, different newer methods have come into the spotlight. Among them, one framework that has garnered substantial attention is ensemble learning framework. The pivotal priority of achieving better results has led to the notion of merging models and recent trends show a renewed interest in the combination of models, mainly in the area of machine learning [23]. Although it is seen from various works that ensemble function better for classification problems but that does not limit it from being used in regression problems [24]. The ensemble learning framework has given rise to different new techniques such as boosting and bagging. Generally, these methods vary in handling the training data that is used for learning the base model. Parameter settings also differ in the training of the base model such that various ensemble components can bring their diversity to the framework [25]. Since their successful implementation, ensemble systems are now used in addressing different other problems related to machine learning such as error correction, selection of features, class imbalanced data, confidence estimation and so on [26].

There are different aggregation methods in ensemble learning, and bagging is one of them, which is used in a predictive model to reduce variance. Bagging ensemble regression (BGER) method is a learning-based method used to perform regression by constructing multiple trees at the time of training and returning the mean regression (prediction) as output. The problem of over-fitting of the training dataset in decision trees is handled in BGER. Here multiple decision trees are constructed by keeping the variance in control. Deep decision trees over-fit the training dataset by learning extremely irregular patterns, thus having low bias and large variance. Although the bias increases to a very small extent, the high variance is ultimately reduced in BGER and thus increasing the performance of the end model. In BGER, the stability and accuracy are improved by bagging the learning trees in the training algorithm. Let in the training dataset $X = x_1, \ldots, x_n$ be the predictor variables and $Y = y_1, \ldots, y_n$ are the target variables, then by bagging $B$ number of times, a random sample with replacement of the training dataset is selected and the trees are then fitted to the selected samples [27,28]. When the training is completed, values of unseen samples $X'$ are predicted by taking the average of all predictions from every individual regression tree on $X'$ as shown in (10)

$$\hat{f} = \frac{1}{B} \sum_{b=1}^{B} f_b(X')$$

(10)

Here the feature set used in this work is $X$ (current of pole 1, current of pole 2, voltage of pole 1 and voltage of pole 2) and approximate fault location is the response variable. By bootstrapping, the trees are de-correlated by showing them different training datasets, thus decreasing the variance without increasing the bias, which ultimately boosts the performance of the end model. Figure 7 shows a simple BGER algorithm diagrammatically.
BGER based method is designed using the input features. In the training or learning process of BGER, a random feature subset is used at every candidate split where p is the total number of features. Various parameters used for training the fault location modules with random forest are number of trees, quality of split, impurity split, etc. The error in location has been checked for a varying number of trees. The optimal number of trees chosen to design the training module is 300. Parameters of best estimator used for training are 300 numbers of trees, the function used to measure the quality of split is MSE, minimum impurity split (threshold used for early stopping in tree growth) is 10–7. The optimal parameters used for designing the BGER based training network are shown in Table 1. The BGER based training network is tested and the results are discussed in the next section.

2.3.3. Design of BTER-based Method

BTER method helps in converting a group of weak learners into an aggregated strong model. Another method that quite overt and uses the association between boosting and optimization techniques is the gradient boosting machine. The principal step is ensemble generation, a step where the set of models are generated. This is followed by the second step called ensemble pruning where redundant models generated in the first step are eliminated or pruned. Lastly in the final step of ensemble integration is to combine the models. This approach is generally used later for obtaining the prediction results of other new problems [23]. Figure 8 shows the simple architecture of BTER method. The final model can be represented as a sum of all the base learners or predictors denoted as \( b_{an}(x) \):

\[
f(x) = \sum_{n=1}^{N} b_{an}(x) \tag{11}
\]

In forward stage-wise additive modeling, \( f(x) \leftarrow 0 \) and for \( n = 1, 2, \ldots, N \).

\[
b_{an} \leftarrow \arg \min_{ba} \sum_{(\tilde{x}, \tilde{y})}^{n} L(f_{n-1}(\tilde{x}) + ba(\tilde{x}), \tilde{y}) \tag{12}
\]
The idea is to decrease the loss function and unless it uses least square loss, the optimization would not change into a least square problem.

In Gradient boosting, \( f_0(x) \leftarrow 0 \) and for \( n=1,2,\ldots,N \) and for all \((\hat{x}, \hat{y})\)

\[
r(\hat{x}, \hat{y}) \leftarrow -\frac{d}{df}L(f_{n-1}(\hat{x}), \hat{y})
\]

(14)

\[
ba_n \leftarrow \arg\min_{ba} \sum_{(\hat{x}, \hat{y})} (ba(\hat{x}) - r(\hat{x}, \hat{y}))^2
\]

(15)

\[
f_n \leftarrow f_{n-1}(x) + vba(x)
\]

(16)

In this work, ensemble learning method has been used to locate faults.

![Diagram of the BTER method for prediction.](image)

**Figure 8.** BTER method for prediction.

In BTER-based method, LSboost has been used because it has shown better performance to locate faults. A weak learner in the form of a decision tree has been used to create the bagged ensemble. The minimum value of the leaf size has been kept as five, which is the default value for regression. This helps in building deep trees that are ideal for the prognostic power of the ensemble. Another important parameter that is necessary at every split is the number of predictors. This is randomly selected for each split and the default value for regression is one-third of the predictors present. The number of learning cycles has been varied before fixing it at 400 for the trained ensemble module. In addition, the learning rate used in the trained module has been fixed at one. The optimal parameters used for designing the BTER-based training network are shown in Table 1. The BTER-based training network is tested and results are discussed in the next section.

### 3. Implementation of Proposed Method

The proposed method is a fault location estimation method which can be implemented offline, as shown in Figure 9. The proposed method consist of two parts, i.e., creating a training fault location module and testing of faulted signals across training module when fault occurs to estimate the location of fault. The training module can be designed by simulating the power system network and recording the signals of fault cases. The signals will be then processed and features are extracted. The extracted features are then given to
the BTER method along with the target to design the training module. The parameters of the BTER method have been changed and an optimal training module is decided based on better accuracy. This training module will be kept and will be used later when the fault occurs in the system. When fault occurs in the HVDC transmission line, the signals will be measured at the rectifier end. The signals will be then imported to the computer where a training module is designed. The signals will be then processed and input features will be extracted. The input features will be then tested against the trained BTER module and fault location will be estimated. The repair crew then can be sent to the location for possible repair of the line.

**Figure 9.** Implementation of proposed method.
4. Results

To analyze the accuracy, it is validated with various fault cases. The fault cases are generated by varying different system parameters and tested. The estimated fault location after the testing has been analyzed. The percentage error is calculated using (17),

$$E = \left[ \frac{|ActualLocation - EstimatedLocation|}{LineLength} \right] \times 100$$ (17)

4.1. Fault Resistance

It is possible that when a fault occurs it may go through a high resistance medium, and fault resistance is unknown. The fault location algorithm should be designed in such a way that it can locate fault irrespective of fault resistance. Table 2 shows the results of fault occurring at different levels of resistance. DTR based method has more error in location estimation than the other two methods. BTER method has less error in locating the fault than the BGER method, as shown in Table 2. Considering all the three methods used, BTER method has a maximum error of up to 3 km while the other two methods may have maximum error of up to 10 km during high resistance faults. Figure 10 shows the estimated fault location with fault resistance 100 \(\Omega\) during P1G fault. Hence, the BTER method is better for locating faults than the BGER and DTR methods in HVDC transmission lines.

| Fault Resistance (\(\Omega\)) | Actual Location (km) | Estimated Location (km) | Error (%) |
|-------------------------------|----------------------|-------------------------|-----------|
| 1                             |                      | DTR  BGER  BTER         | DTR  BGER  BTER |
| 54.2                          | 58.757               | 55.972                  | 54.831    |
| 94.2                          | 93.854               | 93.515                  | 94.629    |
| 134.2                         | 137.527              | 134.383                 | 135.137   |
| 194.2                         | 197.052              | 193.939                 | 193.991   |
| 274.2                         | 282.000              | 274.366                 | 274.266   |
| 314.2                         | 323.448              | 324.275                 | 314.463   |
| 364.2                         | 362.297              | 364.104                 | 364.144   |
| 414.2                         | 407.812              | 415.205                 | 414.547   |
| 474.2                         | 465.800              | 473.710                 | 473.406   |
| 524.2                         | 521.727              | 525.222                 | 524.189   |
| 564.2                         | 570.074              | 564.794                 | 564.818   |
| 614.2                         | 613.459              | 614.760                 | 614.129   |
| 784.2                         | 779.808              | 782.983                 | 783.703   |
| 874.2                         | 873.285              | 874.814                 | 874.126   |
| 954.2                         | 951.360              | 952.910                 | 954.030   |
| 1014.2                        | 1018.212             | 1019.349                | 1014.972  |

Table 2. Performance Varying Fault Resistance.

Figure 10. Estimated fault location during P1G fault in 100 \(\Omega\).
4.2. Far End and Near End Faults

Conventional fault location schemes are not capable of locating faults near a boundary. Faults may occur near the relay location or at the far end of the lines. Few results of boundary faults are given in Table 3. Table 3 shows that the fault location of near and far boundary faults are estimated correctly. BTER method has less error in locating the fault than other two methods. The %error is within 1% for the BTER method. It can be concluded that the BTER method is unaffected by boundary faults.

Table 3. Actual and estimated fault location.

| Boundary | Actual Location (km) | Estimated Location (km) | Error (%) |
|----------|----------------------|-------------------------|-----------|
|          |                      | DTR         | BGER      | BTER      | DTR       | BGER       | BTER      |
| Near     | 0.1                  | 3.834       | 2.192     | 1.051     | 0.339     | 0.190      | 0.086     |
|          | 0.5                  | 3.286       | 2.825     | 1.123     | 0.253     | 0.211      | 0.056     |
|          | 1.5                  | 4.356       | 4.329     | 1.980     | 0.259     | 0.257      | 0.043     |
|          | 2.5                  | 4.834       | 4.738     | 3.195     | 0.212     | 0.203      | 0.063     |
|          | 3.5                  | 5.974       | 5.439     | 2.895     | 0.224     | 0.176      | 0.055     |
|          | 4.5                  | 6.346       | 5.914     | 5.391     | 0.167     | 0.128      | 0.081     |
|          | 5.5                  | 6.327       | 6.382     | 6.442     | 0.075     | 0.080      | 0.085     |
|          | 6.5                  | 7.377       | 7.122     | 7.348     | 0.079     | 0.056      | 0.077     |
| Far      | 1085.5               | 1087.53     | 1085.50   | 1085.03   | 0.184     | 0.000      | 0.042     |
|          | 1087.5               | 1085.02     | 1087.44   | 1087.14   | 0.225     | 0.005      | 0.032     |
|          | 1089.5               | 1086.31     | 1089.48   | 1089.54   | 0.290     | 0.001      | 0.003     |
|          | 1091.5               | 1089.18     | 1091.22   | 1092.61   | 0.210     | 0.025      | 0.100     |
|          | 1093.5               | 1095.25     | 1093.14   | 1093.79   | 0.159     | 0.032      | 0.026     |
|          | 1095.5               | 1093.32     | 1095.36   | 1094.85   | 0.198     | 0.012      | 0.059     |
|          | 1097.5               | 1095.61     | 1096.56   | 1095.61   | 0.171     | 0.085      | 0.171     |
|          | 1099.5               | 1097.73     | 1097.60   | 1099.03   | 0.160     | 0.172      | 0.042     |

4.3. Smoothing Reactors

Smoothing reactors are connected in series with the converters for smoothing the DC currents. As the fault location algorithm in the proposed work uses current signals, it will have some impact on the proposed method. Hence, smoothing reactors may affect the accuracy of the fault location method. Table 4 shows results of all three methods with %error. The results show that the error is within 1% for all the three methods used but BTER method has less error than other two methods.

Table 4. Actual and estimated fault location varying smoothing reactor.

| Smoothing Reactors (mH) | Actual Location (km) | Estimated Location (km) | Error (%) |
|-------------------------|----------------------|-------------------------|-----------|
|                         |                      | DTR         | BGER      | BTER      | DTR       | BGER       | BTER      |
| 545                     | 271.1                | 273.263     | 271.597   | 271.743   | 0.196     | 0.045      | 0.058     |
|                         | 311.1                | 307.681     | 309.189   | 311.092   | 0.310     | 0.173      | 0.000     |
|                         | 341.1                | 341.058     | 342.466   | 342.295   | 0.003     | 0.124      | 0.108     |
|                         | 401.1                | 400.000     | 399.073   | 399.422   | 0.100     | 0.184      | 0.152     |
|                         | 461.1                | 463.600     | 461.453   | 461.342   | 0.227     | 0.032      | 0.022     |
|                         | 501.1                | 506.210     | 500.749   | 500.632   | 0.464     | 0.031      | 0.042     |
| 555                     | 541.1                | 542.866     | 540.141   | 540.740   | 0.160     | 0.087      | 0.032     |
|                         | 601.1                | 595.900     | 602.146   | 601.982   | 0.472     | 0.095      | 0.080     |
|                         | 641.1                | 637.500     | 640.205   | 640.851   | 0.327     | 0.081      | 0.022     |
|                         | 691.1                | 695.428     | 691.228   | 690.768   | 0.393     | 0.011      | 0.030     |
|                         | 721.1                | 727.000     | 720.442   | 720.428   | 0.536     | 0.059      | 0.061     |
|                         | 771.1                | 770.428     | 770.899   | 770.549   | 0.061     | 0.018      | 0.050     |
|                         | 831.1                | 827.526     | 831.266   | 831.178   | 0.324     | 0.015      | 0.007     |
|                         | 871.1                | 872.677     | 871.507   | 871.904   | 0.143     | 0.037      | 0.073     |
|                         | 921.1                | 924.500     | 922.414   | 922.387   | 0.309     | 0.119      | 0.117     |
|                         | 1011.1               | 1015.240    | 1012.834  | 1012.343  | 0.376     | 0.157      | 0.113     |
4.4. Sampling Frequency

Sampling frequency has an effect on the accuracy of fault location method of DC transmission lines to locate the fault. In this work, the proposed method has been tried with various sampling frequencies to choose the optimal sampling frequency for the location estimation. Figure 11 shows the performance of all three methods in locating the faults for various sampling frequencies. Figure 11a shows the performance of the DTR based method which has fewer fault cases for 500 Hz in error range <1% than other sampling frequencies, and all the other sampling frequencies used have almost the same fault cases for error range <1%. So lower sampling frequencies should be chosen, i.e., 1 kHz which will be easier to realize. The same scenario occurred for BGER and BTER methods and is shown in Figure 11b and 11c, respectively.

![Performance varying sampling frequency](image1)

![Performance varying sampling frequency](image2)

![Performance varying sampling frequency](image3)

**Figure 11.** Performance varying sampling frequency (a) DTR (b) BGER (c) BTER.

4.5. Single Pole to Ground Fault

The most commonly occurring faults are pole to ground faults. The method has been checked with varying fault locations during P1G and P2G fault. Table 5 shows the results of different tested fault cases for all three methods. Table 5 shows that percentage error in location of the tested fault cases is within 1%. BTER method has less error in locating the fault than BGER method, as shown in Table 5. Figure 12 shows the performance of all three methods. BTER method has less error than the DTR and BGER methods. Hence, the proposed method estimates the location of single pole to ground faults accurately.
Table 5. Actual and estimated fault location for single pole to ground fault.

| Fault Type | Actual Location (km) | Estimated Location (km) | Error (%) |
|------------|----------------------|-------------------------|-----------|
|            |                      | DTR | BGER | BTER | DTR | BGER | BTER |
| P1G        |                      |     |      |      |     |      |      |
| 51.5       | 58.757               | 51.885 | 0.659 | 0.067 | 0.035 |
| 121.5      | 122.835              | 121.802 | 0.121 | 0.084 | 0.027 |
| 271.5      | 267.153              | 272.346 | 0.395 | 0.166 | 0.076 |
| 371.5      | 373.454              | 372.444 | 0.177 | 0.114 | 0.085 |
| 441.5      | 446.404              | 442.289 | 0.445 | 0.000 | 0.071 |
| 561.5      | 568.200              | 561.130 | 0.609 | 0.031 | 0.033 |
| 661.5      | 654.931              | 662.358 | 0.597 | 0.052 | 0.078 |
| 751.5      | 753.760              | 750.600 | 0.205 | 0.125 | 0.081 |
| 851.5      | 857.837              | 850.746 | 0.576 | 0.017 | 0.068 |
| 981.5      | 989.849              | 982.034 | 0.759 | 0.166 | 0.048 |
| P2G        |                      |     |      |      |     |      |      |
| 43.5       | 40.500               | 44.17 | 0.272 | 0.084 | 0.083 |
| 163.5      | 163.000              | 164.491 | 0.045 | 0.037 | 0.090 |
| 263.5      | 264.000              | 263.283 | 0.045 | 0.055 | 0.019 |
| 373.5      | 371.000              | 373.612 | 0.227 | 0.005 | 0.010 |
| 463.5      | 463.000              | 463.997 | 0.045 | 0.025 | 0.045 |
| 573.5      | 570.500              | 573.366 | 0.272 | 0.021 | 0.012 |
| 673.5      | 677.500              | 673.247 | 0.363 | 0.008 | 0.023 |
| 783.5      | 786.000              | 783.185 | 0.227 | 0.039 | 0.028 |
| 873.5      | 877.000              | 873.522 | 0.318 | 0.019 | 0.002 |
| 983.5      | 985.000              | 982.778 | 0.136 | 0.115 | 0.065 |

Figure 12. Actual and estimated fault locations for all the three methods during P1G fault.

4.6. Double Pole to Ground Fault

In DC lines, double pole to ground fault does not occur more frequently but still needs to be located if it occurs. Table 6 shows the results of different cases during double pole to ground faults. From the results, it has been observed that the percentage error is within 1%. BTER method has less error in locating faults than BGER and DTR method. Hence, the proposed fault location estimation method is capable of locating P1P2G faults accurately.

Table 6. Actual and estimated fault location for double pole to ground fault.

| Actual Location (km) | Estimated Location (km) | Error (%) |
|----------------------|-------------------------|-----------|
|                      | DTR | BGER | BTER | DTR | BGER | BTER |
| 42.5                 | 39.000 | 43.660 | 43.046 | 0.318 | 0.105 | 0.049 |
| 142.5                | 139.500 | 142.889 | 142.607 | 0.272 | 0.035 | 0.009 |
| 242.5                | 241.500 | 243.578 | 243.299 | 0.090 | 0.098 | 0.072 |
| 362.5                | 359.500 | 362.545 | 362.719 | 0.272 | 0.004 | 0.019 |
| 452.5                | 452.500 | 453.203 | 452.797 | 0.000 | 0.063 | 0.027 |
| 542.5                | 545.000 | 543.200 | 543.224 | 0.227 | 0.063 | 0.065 |
4.7. Pole to Pole Fault

In DC lines, pole to pole faults are not common but still a possibility. The proposed method has been tested during pole to pole (P1P2) faults. Table 7 shows the results of different tested fault cases for P1P2 fault. The percentage error in location is within 1% for all three methods. However, BTER method has better results in locating the fault than the other 2 methods. Hence, the fault location estimation method locates P1P2 faults accurately.

Table 6. Cont.

| Actual Location (km) | Estimated Location (km) | Error (%) |
|----------------------|-------------------------|-----------|
|                      | DTR         | BGGER     | BTER      | DTR         | BGGER     | BTER      |
| 672.5                | 671.000     | 672.940   | 673.103   | 0.136       | 0.040     | 0.054     |
| 742.5                | 738.500     | 744.303   | 743.237   | 0.363       | 0.163     | 0.067     |
| 852.5                | 849.000     | 852.277   | 852.466   | 0.318       | 0.020     | 0.003     |
| 962.5                | 966.000     | 962.554   | 962.471   | 0.318       | 0.004     | 0.002     |

Table 7. Actual and estimated fault location for pole to pole fault.

| Actual Location (km) | Estimated Location (km) | Error (%) |
|----------------------|-------------------------|-----------|
|                      | DTR         | BGGER     | BTER      | DTR         | BGGER     | BTER      |
| 47.1                 | 48.500      | 47.290    | 47.169    | 0.127       | 0.017     | 0.006     |
| 157.1                | 156.500     | 158.190   | 157.874   | 0.054       | 0.099     | 0.070     |
| 267.1                | 271.500     | 268.214   | 267.678   | 0.400       | 0.101     | 0.052     |
| 367.1                | 371.000     | 367.193   | 367.449   | 0.354       | 0.008     | 0.031     |
| 467.1                | 470.500     | 467.653   | 467.779   | 0.309       | 0.050     | 0.061     |
| 587.1                | 587.500     | 587.307   | 587.202   | 0.036       | 0.018     | 0.009     |
| 667.1                | 670.500     | 667.217   | 666.979   | 0.309       | 0.010     | 0.011     |
| 777.1                | 777.500     | 776.479   | 776.421   | 0.036       | 0.056     | 0.061     |
| 867.1                | 869.500     | 867.882   | 867.558   | 0.218       | 0.071     | 0.041     |
| 977.1                | 976.500     | 976.945   | 976.888   | 0.054       | 0.014     | 0.019     |

5. Discussion

In HVDC transmission lines, fault location is estimated using traveling wave in most of the real HVDC power networks. In traveling wave method, fault location is measured from time taken by the reflected and refracted wave fronts of wavelets multiplied with the speed of light. However, this method suffers from serious drawbacks as it cannot be assumed to there to always be a smooth earth surface. When fault occurs, the earth surface is not uniform for all locations so the time taken by the reflected wave front cannot be assumed to be the same in proportion to the distances. In addition, the severity of fault is also not the same if the faulty phase during a ground fault has more bending or irregularity; in addition, in this instance time taken by the reflected wave front will not be the same. For various fault impedances of the same location, the reflected wave front may have more secondary wave fronts. Hence the propagation time to reach the measuring station cannot be the same. Refracted wave front time is also considered for fault location, but this wave front will not reach the measuring station—if the conductor experiences phase fault and the medium is different, the wave front will travel in the different dielectric mediums where it strikes, but cannot be refracted back to the same media from where it travels. So, considering this parameter fault location estimation will lead to inconsistency.

A reflected wave front formed after a fault and followed by new wave front that is formed will act as the secondary source of disturbances and this process continues successively. However, there are some wavelets that will be coherent and some that will be incoherent, and the resultant wave interference will be destructive or constructive depending on the fault—sampling of data on these parameters is neglected. In diffraction, due to waves striking the earth or an obstacle in the path, the waves will bend near the corner and a fringe pattern will be formed resulting in wave fronts. Path difference under
these circumstances will not be the same and so high sampling of data is not reliable. The surface of earth and the nature of the obstacle are very irregular, and under these circumstances using a conventional method to determine the fault location is not accurate.

Various AI methods have been implanted in transmission line fault location such as artificial neural network [29,30], support vector machine [29], k-nearest neighbor [31], decision tree [29], random forest (bagging ensemble) [30], boosting ensemble [30], deep learning [14] and fuzzy [12]. Artificial neural networks have no specific rule for determining the structure of network. The suitable network is achieved through trial and error. Support vector machine algorithm is not suitable for large data sets, and it does not perform sound when the data set has more noise. Accuracy of k-nearest neighbor depends on the quality of the data and prediction stage might be slow for large data. It requires high memory to store all the training data, and can be computationally expensive. In decision tree, small changes in the data cause large changes in the structure of the decision tree that will cause instability and so it may have a problem of over fitting. The over fitting problem of decision tree has been solved by using random forest and boosting ensembles. Deep learning method have no specific rule for determining the structure of network and has high hardware implementation cost. Although fuzzy based method is simple, it is unable to locate high resistance fault. Ensemble learning has been applied to AC transmission lines, this does not guarantee that it will work for DC transmission lines, as the signals of AC and DC are different. Hence, the ensemble learning method has been used in this work for locating DC transmission line faults, considering its simplicity and robustness. All the methods implemented in this work have been compared in terms of the percentage error obtained. A total of nearly 10,000 fault cases are used for testing the fault location estimation algorithms designed for HVDC transmission lines. Percentage error and percentage of fault cases in each error range have been calculated and compared, as shown in Figure 13. BTER method has more accuracy in locating the faults than BGER and DTR based methods. Hence, BTER-based method can be used in HVDC transmission lines for locating the faults correctly.

![Figure 13. Comparison of various techniques used.](image)

In recent years, various methods have been proposed for locating faults. Table 8 shows the comparisons which are made based on different criteria. In [1], multiple signal classification method has been used to obtain the natural frequency. Traveling-wave velocity and reflect coefficient of natural frequency has been used to locate faults. However, since the sampling rate is very high that is not easy to realize. In [3], the value of the natural frequency of VSC-HVDC transmission line is related to fault distance that is used to locate fault, but it also needs a higher sampling rate. In [5], the ratio coefficient of wave velocities between different frequencies and digital band-pass filter has been used. A drawback of this method is that it requires a communication link and has high sampling rate. In [7], both end measurements of the faulted line have been time-reversed and further calculated to identify the fault. This also requires a communication link. A drawback of the method, suggested in [12], is that it is unable to locate high resistance faults. In [13], rectifier side AC RMS voltage, DC voltage and current on both the poles have been used.
to detect the fault. However, the effect of fault resistance has not been studied in this work. In [14], faults are detected, poles are identified, and faults are located using one end measurement, but this has a high hardware implementation cost. Some of the methods use traveling waves for fault location estimation. High sampling frequency is the drawback of the traveling wave method, while the proposed method works with only 1 kHz sampling frequency. If the sampling frequency is high, then it will be difficult to implement. Some of the methods are not capable of locating faults with high fault resistance. Some of the methods overlooked the reach setting of the proposed method which is also important. Some of the methods are designed for monopolar transmission lines which may or may not work for bipolar transmission lines. Table 8 shows that the proposed method uses only one end measurement, has a low sampling frequency, more reach settings, and locates high resistance faults. Hence, the proposed BTER-based method is more suitable for fault location estimation to locate faults.

| Methods                 | Techniques Used                      | Measurement Required | Sampling Frequency (kHz) | Reach Setting (%) | Maximum Error   |
|-------------------------|-------------------------------------|----------------------|--------------------------|------------------|-----------------|
| He et al. [1]           | Traveling-wave                       | Single end           | 100                      | 90               | Not mentioned   |
| Song et al. [3]         | PRONY algorithm                      | Single end           | 50                       | 90               | 0.7%            |
| Yuansheng et al. [5]    | Traveling wave theory with Bergeron  | Both end             | 1000                     | -                | 0.777 km        |
| Zhang et al. [7]        | Electromagnetic time-reversal method | Both end             | 50                       | -                | Not mentioned   |
| Swetapadma et al. [12]  | Wavelet transform                    | Single end           | 1                        | -                | More than 10%   |
| Johnson et al. [13]     | SVM                                  | Single end           | 1                        | -                | 0.4831%         |
| Livani et al. [14]      | DWT and SVM                          | Single end           | >100                     | -                | 7.1%            |
| Proposed scheme         | Boosting ensemble learning           | Single end           | 1                        | 99.9             | 1.0%            |

If conventional methods are compared to the method suggested in this work, the suggested work is more superior for fault location, and has many advantages. In comparison to all these methods, the suggested BTER method uses extracted features that are not time dependent. It depends on the pattern formed after a fault and averaging of maximum intensity signal and minimum intensity signal due to resultant interference of different waves formed. Methods used in this work are learning methods which have the capability of adapting to various scenarios. Hence, the BTER method can locate fault in natural conditions and adverse conditions effectively.

6. Conclusions

A novel method has been proposed using BTER method for DC lines to locate faults. The BTER-based method has been tested with the CIGRE benchmark model. The proposed fault location estimation scheme uses only one end signal. The BTER method has been tested using various fault cases. The results of the proposed scheme are promising and can be implemented in real systems. Considering the growing amount of data in recent years, the proposed method will be a good tool for locating faults. The BTER method uses post fault samples and a lower sampling rate. The BTER method uses one end signals only. The proposed BTER-based fault location estimation method has many advantages:
i. The proposed method has fault location error within 1%;
ii. The proposed method has reach setting of 99.9%;
iii. The proposed method has sampling frequency of 1 kHz;
iv. No need communication link;
v. The proposed method is simple, robust and reliable.
The proposed fault location estimation method is an offline procedure. Hence, it can be implemented easily with a PC set up once the measurement from substation is directed to the PC. The signal processing can be carried out in any open source software available,
such as Python. After the features are obtained, the BTER-based fault location method can be implemented using scikit learn of python to estimate the location offaults.

**Author Contributions:** A.S., Conceptualization, Formal analysis, Data curation, Methodology, Writing—original draft; S.A., Investigation, Writing—original draft; S.C. (Soham Chakrabarti), Data curation, Resources, Software, Writing—original draft; S.C. (Satarupa Chakrabarti), Software, Visualization, Writing—original draft; A.E.-S., Methodology, Project administration, Supervision, Writing—review & editing; A.E.-S., A.Y.A., Methodology, Project administration, Supervision; Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. He, Z.; Liao, K.; Li, X.; Lin, S.; Yang, J.; Mai, R. Natural Frequency-Based Line Fault Location in HVDC Lines. *IEEE Trans. Power Deliv.* 2014, 29, 851–859. [CrossRef]

2. Zhang, Y.; Tai, N.; Xu, B. Travelling wave-based pilot direction comparison protection for HVDC line. *Int. Trans. Electr. Energy Syst.* 2013, 23, 1304–1316.

3. Song, G.; Chu, X.; Cai, X.; Gao, S.; Ran, M. A fault-location method for VSC-HVDC transmission lines based on natural frequency of current. *Int. Trans. Electr. Energy Syst.* 2014, 63, 347–352. [CrossRef]

4. Liu, X.; Osman, A.H.; Malik, O.P. Real-Time implementation of a hybrid protection scheme for bipolar HVDC line using FPGA. *IEEE Trans. Power Deliv.* 2011, 26, 101–108. [CrossRef]

5. Liang, Y.; Wang, G.; Li, H. Time-Domain Fault-Location Method on HVDC Transmission Lines under Unsynchronized Two-End Measurement and Uncertain Line Parameters. *IEEE Trans. Power Deliv.* 2015, 30, 1031–1038.

6. Long, C.S.; Chun, S.H.; Zhi, W.Y.; Fang, Z. Distance protection scheme with travelling wave for UHVDC transmission line based on wavelet transform. In Proceedings of the Third International Conference Electric Utility Deregulation and Restructuring and Power Technologies, Nanjing, China, 6–9 April 2008; pp. 2162–2164.

7. Zhang, X.; Tai, N.; Wang, Y.; Liu, J. EMT based fault location for DC line in VSC-MTDC system using high-frequency currents. *IET Gen. Trans. Dist.* 2017, 11, 2499–2507. [CrossRef]

8. Gao, S.; Song, G.; Ma, Z.; Jin, X. Novel pilot protection principle for high voltage direct current transmission lines based on fault component current characteristics. *IET Gener. Transm. Distrib.* 2015, 9, 468–474. [CrossRef]

9. Gao, S.; Suonan, J.; Song, G. A new current differential protection principle for HVDC transmission lines. *Autom. Electr. Power Syst.* 2010, 30, 45–49.

10. Kong, F.; Hao, Z.; Zhang, B. Improved differential current protection scheme for CSC-HVDC transmission lines. *IET Gener. Transm. Distrib.* 2017, 11, 978–986. [CrossRef]

11. Dai, Z.; Liu, N.; Zhang, C.; Pan, X.; Wang, J. A Pilot Protection for HVDC Transmission Lines Based on Transient Energy Ratio of DC Filter Link. *IEEE Trans. Power Deliv.* 2020, 35, 1695–1706. [CrossRef]

12. Swetapadma, A.; Agarwal, S.; Ranjan, A.; Abdelaziz, A.Y. A Novel Fault Distance Estimation Method for Voltage Source Converter-Based HVDC Transmission Lines. *Electr. Power Comp. Syst.* 2021, 48, 1760–1772. [CrossRef]

13. Johnson, J.; Yadav, A. Complete protection scheme for fault detection, classification and location estimation in HVDC transmission lines using support vector machines. *IET Sci. Meas. Technol.* 2017, 11, 279–287. [CrossRef]

14. Swetapadma, A.; Chakrabarti, S.; Abdelaziz, A.Y.; Alhelou, H.H. A Novel Relaying Scheme Using Long Short Term Memory for Bipolar High Voltage Direct Current Transmission Lines. *IEEE Access* 2021, 9, 119894–119906. [CrossRef]

15. Livani, H.; Evrenosoglu, C. A single-ended fault location method for segmented HVDC transmission line. *Electr. Power Syst. Res.* 2014, 107, 190–198. [CrossRef]

16. Yang, Q.; Li, J.; Santos, R.; Huang, K.; Igić, P. Intelligent fault detection and location scheme for modular multi level converter multi terminal high voltage direct current. *High Volt.* 2020, 6, 1–13. [CrossRef]

17. Wang, B.; Liu, Y.; Lu, D.; Yue, K.; Fan, R. Transmission Line Fault Location in MMC-HVDC Grids Based on Dynamic State Estimation and Gradient Descent. *IEEE Trans. Power Deliv.* 2021, 36, 1714–1725. [CrossRef]

18. Lin, S.; Liu, L.; Sun, P.; Lei, Y.; Teng, Y.; Li, X. Fault Location Algorithm Based on Characteristic-Harmonic Measured Impedance for HVdc Grounding Electrode Lines. *IEEE Trans. Inst. Meas.* 2020, 69, 9578–9585. [CrossRef]

19. Li, S.; Chen, W.; Yin, X.; Chen, D.; Teng, Y. A Novel Integrated Protection for VSC-HVDC Transmission Line Based on Current Limiting Reactor Power. *IEEE Trans. Power Deliv.* 2020, 35, 226–233. [CrossRef]

20. Saber, A. A Backup Protection Algorithm for Bipolar Line-Commutated Converter HVDC Lines. *IEEE Syst. J.* 2021, 15, 1172–1178. [CrossRef]

21. Yang, S.; Xiang, W.; Wen, J. An Improved DC Line Fault Detection Scheme Using Zone Partition for MTDC Wind Power Integration Systems. *IEEE Trans. Power Deliv.* 2021. [CrossRef]

22. MATLAB. *MATLAB User’s Guide: R2016a Documentation*; MathWorks Inc.: Natick, MA, USA, 2016.
23. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2011.

24. Frayman, Y.; Rolfe, B.F.; Webb, G.I. Solving regression problems using competitive ensemble models. *Aust. Joint Conf. Artif. Intell.* **2002**, *2557*, 511-522.

25. Alhamdoosh, M.; Wang, D. Fast decorrelated neural network ensembles with random weights. *Inf. Sci.* **2014**, *264*, 104-117. [CrossRef]

26. Mendes-Moreira, J.; Soares, C.; Jorge, A.M.; Sousa, J.F.D. Ensemble approaches for regression: A survey. *ACM Comp. Surv.* **2012**, *45*, 1-10. [CrossRef]

27. Zhang, C.; Ma, Y. *Ensemble Machine Learning: Methods and Applications*; Springer Science & Business Media: Boston, MA, USA, 2012.

28. Bowles, M. *Machine Learning in Python: Essential Techniques for Predictive Analysis*; John Wiley & Sons: Indianapolis, IN, USA, 2015.

29. Swetapadma, A.; Yadav, A.; Abdelaziz, A.Y. Intelligent schemes for fault classification in mutually coupled series-compensated parallel transmission lines. *Neur. Comput. Appl.* **2020**, *32*, 6939-6956. [CrossRef]

30. Swetapadma, A.; Chakrabarti, S.; Abdelaziz, A.Y. Feasibility study of intelligent fault location estimation methods for double-circuit transmission lines. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e13198. [CrossRef]

31. Swetapadma, A.; Yadav, A. A novel single-ended fault location scheme for parallel transmission lines using k-nearest neighbor algorithm. *Comput. Electr. Eng.* **2018**, *69*, 41-53. [CrossRef]