International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 8, 2025
https://theaspd.com/index.php

High Impedance Fault Detection In Radial Feeder Under Different Micro Climateic Conditions Using Optimised Support Vector Machine Algorithms

Trivedh Yadav Guttimari[1*] and Gowri Manohar T[2]
[1]Sri Venkateswara University College of Engineering, Sri Venkateswara University, Tirupati
tejatrivedh@gmail.com, https://orcid.org/0009-0009-0309-2282
[2]Sri Venkateswara University College of Engineering, Sri Venkateswara University, Tirupati
gowrimanohart@gmail.com, https://orcid.org/0000-0003-0441-6022

Abstract
Detection of high impedance faults (HIF) in radial feeders is a challenging task. HIF faults occurs mostly due to single line to ground faults in rural and semi urban areas in Andhra Pradesh, India of about 80%. Existing methods detect the single line to ground faults (SLGF) and not classifies the subtypes of SLGF faults such as small-impedance faults (SIFs), medium-impedance faults (MIFs), high-impedance faults (HIFs), arc grounding faults (AGFs), intermittent grounding faults (IGFs), and transient grounding faults (TGFs). SLGF subtype classification helps for selective tripping, Minimizes Outages, Limits Prolonged Fault Operation, Prevents Cascading Failures and improves the system monitoring. In this paper, optimized Support Vector Machines (SVM) algorithms such as (i)cubic-SVM, (ii) PSO-SVM and (iii) BO-SVM are proposed for the detection and classification of the HIF under different environmental i.e. microclimatic conditions. From the results, BO-SVM provides the better accuracy of 97% in HIF detection and subtype classification for 4 - node and 13 -node IEEE radial test feeders. The proposed algorithms are compared with the traditional algorithms and evaluated the performance analysis.
Keywords *Evaluation metrics, High Impedance Fault, Linear Regression, Support Vector Machines, microcliamtic conditions.*

INTRODUCTION

In radial feeder distribution system, the energized conductor contacts the ground surface and leads to high impedance fault (HIF). HIF cause life threatening for living beings due to transfer of power in and around the fault premises. Above 80 percent of the faults are occurring due to single line to ground (L-G) fault. When the energized conductor falls on the surfaces such as asphalt, sand, grass etc. starts conducting through them and surfaces have low impedance value. This low impedance leads to small current, which is equivalent to load current value and could never be detected through conventional protection systems such as over current relays.

In [1], HIF fault locations are detected using SVM algorithm for dc microgrids. In [2], HIF detection and isolation method in power distribution network were performed through Principal Component Analysis (PCA), Fisher Discriminant Analysis, Binary and Multiclass SVM. Multiclass SVM provides high accuracy than traditional methods and ensures reliability, security, and dependability of distribution network. In [3], identified and categorized the LL and LG faults on the DC side of PV systems. The authors used hierarchical method to decrease the training dataset for the learning process, and applied the machine learning algorithms. The hierarchical method had achieved an accuracy of about 96.66% and 91.66%. In [4], single ended travelling wave-based fault location method for a hybrid transmission line was developed. Discrete Wavelet Transform (DWT) is used for the transient extraction and SVM is used for the fault identification. The developed method is used for different fault identifications such as different Fault Inception Angles (FIA), fault resistances, Non-Linear High Impedance Faults (NLHIF). In [5], HIFs is classified in the distribution systems. Hermite Transform (HT) is used and compared with DWT, improved performance the of HIFs detection. In [6], HIF detection system based on Kernel Extreme Learning Machine (KELM) has been used for the detection of HIFs. DWT and Hilbert-Haung Transform

(HHT) are used to characterize the HIF. XG Boost-based fault feature selection is utilized for characterization of HIFs. The KELM provides high accuracy in HIF detection.

In [7], HIF detection is performed based on incremental learning for non-stationary data streams with changing distributions. The Stationary Wavelet Transform (SWT) is extracts the fault characteristics, Back Propagation Neural Network (BPNN) is utilized for HIF detection. The suggested approach maintains a high degree of detection accuracy for current cycle data and improves the reliability in HIF detection. In [8], power system analysis tool pandapower is used for the fault detection using time series calculations. In [9], radial test feeders are utilized for fault detection. In [10], Multi Resolution Analysis (MRA) and Adaptive Neuro Fuzzy Inference System (ANFIS) are used for the discrimination of HIF such as load switching and capacitor switching. The proposed method has high reliability in differentiating HIF during different transient conditions.

In [11], authors developed differential protection of distribution line of 5-bus system, which contain a wind farm in the microgrid for HIF detection through different machine learning algorithms such as Decision Tree (DT), Random Forest (RF), Gradient Boost (GB), Multilayer perceptron, Naïve Bayes (NB), K-Nearest Neighbor (KNN), and SVM. SVM is used for the purpose of maintaining the security and dependability of microgrid protection. In [12], authors analysed the status of distribution networks using supervised learning algorithms such as Feed Forward Neural Network (FFNN), LR, and SVM for IEEE 34 and 37 node test feeders. The impact of errors in measurement studied using FFNN and SVM algorithms. In [13], HIF detection in solar PhotoVoltaic (PV) integrated power system was performed using recurrent neural network-based Long Short-Term Memory (LSTM) in IEEE 13 bus system, which was interconnected with 300 KW solar PV plants. DWT was applied for feature extraction and compared with KNN, SVM, J48 DT and NB algorithms. LSTM outperforms all the performance indices (PI) compared to the other techniques. In [14], deployed a HIF model and the combination of DWT and SVM was presented for HIF detection. An Eskom network was utilized for this study. In [15], ML based algorithm was used for the detection of arc fault in experimental testbed. Udoo X86 ultra microcontroller was utilized for implementing the proposed method and the results were tested in distribution system. In [16], discrete wavelet packet transform (DWPT) and machine learning were used for identification of feeder Malfunction. In [17], conventional autoencoder framework for HIF detection (CAE-HIFD) was used in IEEE 13-node feeder for HIF detection. In [18], pattern recognition method was used for the detection of High Impedance Arcing Faults (HIAFs). Training data were extraction from 25KV microgrid, which includes of distributed generation (DG) and renewable energy sources. The developed methodology was compared with the Pearson Correlation Coefficient (PCC) and PCA method for rapid identification of HIAF in a microgrid.

In [19], automatic detection and diagnosis algorithm framework using three different ML classifiers such as DT based on different split criteria, K-NN based on different distance and weighting functions, SVM based on different kernel functions and multi classifier approaches, and Bayesian optimization are applied to tune hyperparameters to classifiers errors. In [20], SVM-based fault detection technique for islanded microgrids (MGs) has developed. The fault site is discovered using an inter harmonic injection approach. The approach is tested on 9-bus meshed MG, and it is capable of accurately detecting faults with high fault impedances. In [21], a novel data-driven approach for fault classification has been developed and identified the problem at hierarchical levels in a distribution system. Feeders are monitored through the sensors; smart meters are with sound classifiers. The classifier ML algorithms are ANN (Artificial neural network), SVM (support vector machines) BT and Adaptive Boosting (AdaBoost). A 13-bus test system is taken for the case study and developed method is used for the LG-type fault analysis. The developed strategy is evaluated for their performance under system uncertainties such as incomplete information of observable state measurements and different DG penetration levels. A K-means clustering algorithm is used for prediction of missing data, which improved the overall accuracy.

In [22], authors developed an intelligent strategy for HIF identification that combines Extended Kalman Filter (EKF) and SVM. The suggested technique was trained on 300 datasets and evaluated on 200 datasets, demonstrated its speed, accuracy, and robustness for detecting HIF in distribution feeders. In [23], developed an SVM-based single-pole-to-ground fault location method for MMC-HVDC transmission lines. To achieve an accurate location for various grounding resistance values, the developed location method requires a single ended fault voltage waveform recorded at a fault recording frequency of 20KHZ. In [24], ML algorithm is used for the prediction of HIF on a distribution network. The location identification is done using SVM, NN and LSTM with LSTM and shown high accuracy in HIF detection.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 8, 2025
https://theaspd.com/index.php

PROPOSED METHODOLOGY

Linear Regression (LR):

Linear Regression method [25] uses labelled data for learning and correlating data points with best-fitting linear functions and predicted the future data. A data set whose corresponding target value is already known is called labelled data.LR algorithm calculates a linear relationship between the dependent variable and one or more independent features through fitting a linear equation to the data obtained, as shown in Figure 1. Simple Linear Regression has one dependent variable and one independent variable as in equation (1).

$$Y = \theta_0 + \theta_1 X \qquad (1)$$

Where, Y represents dependent variable, $\boldsymbol{X}$ represents independent variable, θ_0 represents intercept, θ_1 represents slope

Multiple Linear regression consists of one or more independent variable and one dependent variable.

$$Y = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \cdots + \theta_n X_i \qquad (2)$$

Where, Y represents dependent variable, $X_1, X_2, .., X_i$ are independent variables θ_0 represents intercept, $\theta_1, \theta_2, \ldots, \theta_n$ are slopes.

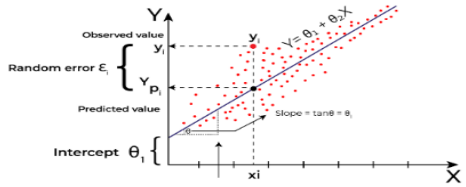


Figure 1: Linear Regression representation

SUPPORT VECTOR MACHINES(SVM):

SVM is a supervised learning technique used for both classification and regression problems [26]. SVM separates n-dimension space that contains different types of data points for classifying new data point into categories. The hyperplane is the specific boundary. SVM selects support vectors and points to draw the hyperplane. support vector concentration technique finds their parametric boundaries that are essential. The figure 2 shows two distinct categories that are categorized using a decision boundary or hyperplane:

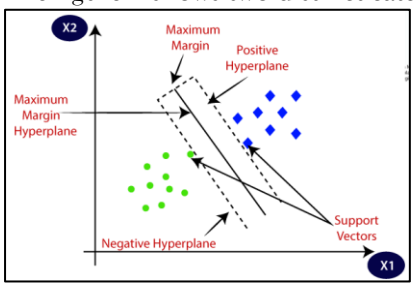


Figure 2: SVM categories

Support vector is the data point or vector, which is closest to the hyperplane and influences its position. These vectors are called support vectors because they support the hyperplane.

More than one line or decision boundary may be required to distinguish classes in an n-dimensional space. The decision boundary is useful in classification of the data points. The dimensions of the hyperplane are based on the dataset features. For two attributes, the hyperplane's dimensions will be a straight line. If hyperplane has two-dimensional plane with three features and maximum distance between data points are included, also known as the maximum margin, when creating a hyperplane.

DESIGN CONSIDERATIONS

IEEE 4-NODE TEST SYSTEM

The single line diagram for the IEEE 4-node test system is shown in the Figure.3.

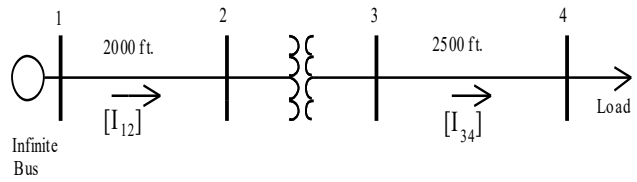


Figure 3: IEEE 4 - node test system

The line and load data of 4 - node test system carry out the short circuit calculations as explained below. The distance between each node has been tabulated as shown in the Table 1.

Table 1: distance between buses for 4-node system

S. No.	From bus	To bus	Distance in km
1	0 or slack bus	1	0.6096
2	1	2	0.001
3	2	3	0.762

The resistance and reactance of the lines are tabulated in Table 2.
Table 2: resistance and reactance data for 4-node system

S. No.	Bus number	R in ohm per km	X in ohm per km	R0 in ohm per km	X0 in ohm per km
1	0	0.1902	0.418	0.367	1.855
2	1	0.01	0.06	0.02	0.12
3	2	0.1902	0.418	0.367	1.855

The load data is stated in table 3.
Table 3: load data for 4-node system

S. No	Bus No	Pa_MW	Pb_MW	Pc_MW	Qa_MVAR	Qb_MVAR	Qc_MVAR
1	3	1.275	1.8	1.275	0.79	0.871	0.78

IEEE 13 node radial test feeder is used in this study.
The network diagram of 13 node radial test feeder is shown in the Figure 4.

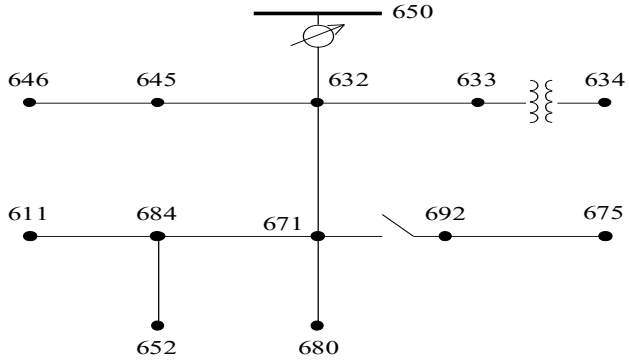


Figure 4: IEEE 13-node test system
For easy computing for the 13-node test system the nodes 650, 632, 633, 634, 645, 646, 671, 692, 675, 684, 911, 652, 680 are written as 0 or slack bus, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 respectively. The line data and load data for the IEEE 13- bus test system is tabulated in Table 4 to Table 6.
Table 4: Line data of 13 bus test system

S. No.	From bus	To bus	Distance in ft	Distance in km
1	Slack bus	1	2000	0.6096
2	1	2	500	0.1524
3	2	3	0	0.001

4	1	4	500	0.1524
5	4	5	300	0.09144
6	1	6	2000	0.6096
7	6	7	0	0.001
8	7	8	500	0.1524
9	6	9	300	0.09144
10	9	10	300	0.09144
11	9	11	800	0.24384
12	6	12	1000	0.3048

Table 5: Resistance and Reactance data between buses

S. No.	From bus	To bus	R (ohm/km)	X (ohm/km)
1	0	1	0.5493	1.6649
2	1	2	1.1964	1.9488
3	2	3	0.011	0.02
4	1	4	2.1299	2.1832
5	4	5	2.1299	2.1832
6	1	6	0.5493	1.6649
7	6	7	0.001	0.001
8	7	8	1.2843	0.7181
9	6	9	2.139	2.1675
10	9	10	2.1387	2.1681
11	9	11	2.1601	0.8244
12	6	12	0.5493	1.6649

Table 6: Load data for each phase at different buses in test system

S. No.	Load	Phase a		Phase b		Phase c	
	Bus no.	MW	MVAr	MW	MVAr	MW	MVAr
1	1	0.01648	0.00945	0.06645	0.03806	0.11697	0.06705
2	3	0.04263	0.02018	0	0	0	0
3	4	0	0	0.17053	0.12509	0	0
4	5	0	0	0.23022	0.13197	0	0
5	6	0.38370	0.21995	0.38370	0.21995	0.38370	0.21995
6	7	0	0	0	0	0.17053	0.15138
7	8	0.48602	0.18907	0.06821	0.06055	0.28991	0.21265
8	10	0	0	0	0	0.17053	0.08074
9	11	0.12790	0.08579	0	0	0	0

Steps involved in short circuit analysis and ML implementation:

Step 1: Initialize the model:

Firstly, calling the required functions and designing the system model in Jupyter Notebook is done in this phase.

Step 2: Calculating SC currents:

By executing the design model, the SC currents are obtained as Ikss in Kilo Amperes (KA). The values are saved in .csv format excel file for easy computation.

Step 3: training and testing the model:

In this step, the obtained data is split into train and test data for training the model. The predictions of the fault bus are carried out. All the results are saved in both .csv and .png files for validation.

System requirements:

The minimal requirements are to have a system which can run python and scikit learn libraries and matplotlib functions.

RESULTS AND DISCUSSION

For 4-node test system:

The short circuit calculations have been carried out on pandapower [27,28] and the SC current (Ikss) in KA for different values of resistances as per IEC 60909 standards are tabulated for both min and max conditions in Table 7 and Table 8 for 4-node test system.

Table 7: short circuit result max fault condition for 4-node system

S. No.	Bus number	Ikss (KA) for 5 ohm	Ikss (KA) for 10 ohm	Ikss (KA) for 15 ohm	Ikss (KA) for 20 ohm
1	0	1.563468	0.786872	0.525714	0.39471
2	1	1.50589	0.773552	0.519975	0.391534
3	2	1.505883	0.773551	0.519975	0.391534
4	3	1.420943	0.754692	0.512086	0.387251

Table 8: short circuit result min fault condition for 4-node system

S. No.	Bus number	Ikss (KA) for 5 ohm	Ikss (KA) for 10 ohm	Ikss (KA) for 15 ohm	Ikss (KA) for 20 ohm
1	0	1.378731	0.708188	0.47531	0.357521
2	1	1.300513	0.691014	0.468197	0.353684
3	2	1.300504	0.691012	0.468197	0.353683
4	3	1.203192	0.668495	0.458866	0.348674

The obtained metrics are tabulated in table 9.

Table 9: evaluation metrics for training and testing cases

S. No.	Evaluation metrics	For Training Dataset		For Testing Dataset	
		LR	SVM	LR	SVM
1.	MAE	0.3006	0.2939	0.3332	0.2313
2.	MSE	0.1354	0.1424	0.1446	0.1193
3.	R2_score	0.8862	0.8803	0.8734	0.8955

Few values of actual fault bus number and the predicted fault bus number for training and testing cases for IEEE 4-node test system individually for both LR and SVM methods are tabulated in table 10,11,12, and 13 respectively.

Table 10: LR train values and their predicted values

S. No.	Actual value	Predicted value	S. No.	Actual value	Predicted value
1	2	1.487517	14	3	2.9346
2	3	3.064344	15	1	1.459887
3	2	1.48848	16	2	1.511997
4	2	1.500594	17	1	1.50048
5	2	1.460005	18	0	0.088601
6	3	3.084425	19	3	3.114139
7	0	0.187936	20	2	1.485897
8	0	0.051493	21	3	3.115338
9	1	1.488362	22	1	1.487403
10	2	1.37243	23	3	3.072602
11	3	3.023319	24	0	0.000118
12	1	1.351785	25	1	1.485779
13	0	0.182468			

Table 11: LR test values and their predicted values

S. No.	Actual value	Predicted value
1	1	1.372309
2	0	0.23982
3	3	3.047087
4	0	0.263565
5	2	1.3519
6	1	1.511884

7	0	0.249988

Table 12: SVM train values and their predicted values

S. No.	Actual value	Predicted value	S. No.	Actual value	Predicted value
1	2	1.356032	14	3	2.928393
2	3	3.046317	15	1	1.396755
3	2	1.430048	16	2	1.385371
4	2	1.380446	17	1	1.380306
5	2	1.396895	18	0	-0.10005
6	3	3.058818	19	3	3.093154
7	0	0.09952	20	2	1.433546
8	0	-0.01884	21	3	3.100263
9	1	1.429909	22	1	1.355892
10	2	1.331163	23	3	3.048238
11	3	3.01836	24	0	0.067202
12	1	1.244286	25	1	1.433406
13	0	0.100263			

Table 13: SVM test values and their predicted values

S. No.	Actual value	Predicted value
1	1	1.331022
2	0	0.026457
3	3	3.017911
4	0	0.055094
5	2	1.244427
6	1	1.385232
7	0	0.048093

The actual and predicted fault bus number for both methods is represented graphically as shown in figure 5 and 6.

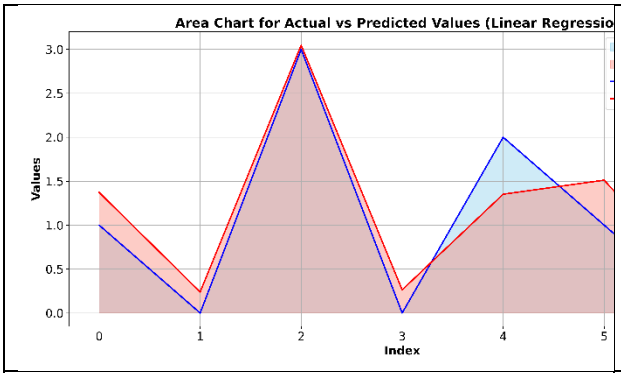


Figure 5: actual vs predicted for linear regression area chart for 4 node test system

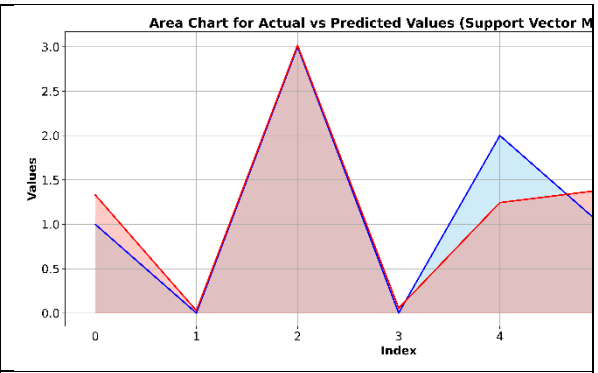


Figure 6: actual vs predicted for support vector machines area chart for 4 node test system

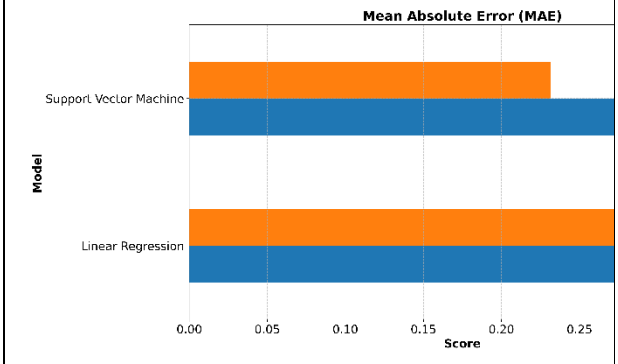


Figure 7: MAE Train vs Test for 4 node test system

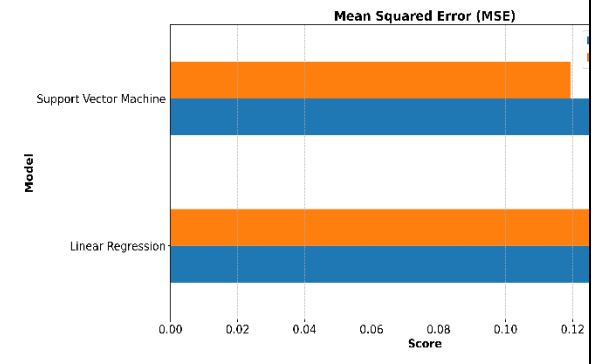


Figure 8: MSE Train vs Test for 4 node test system

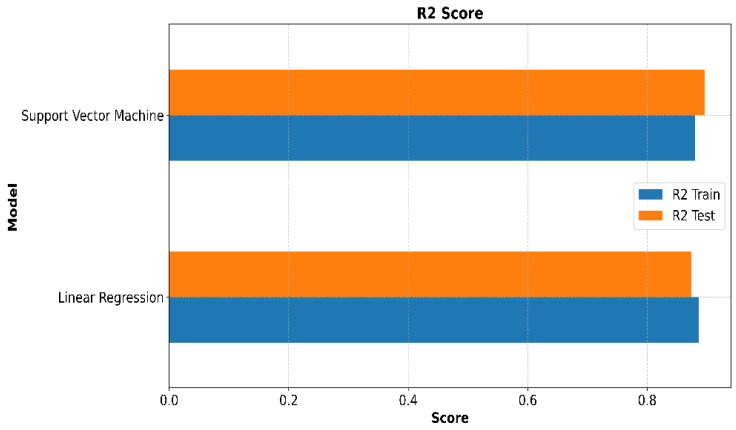


Figure 9: R2 score Train vs Test for 4 node test system

The metrics result plots for MAE, MSE and R2_score for 4-node test system is seen in figure 7,8 and 9 for training and testing the model.

For IEEE 13 bus test system:

Similarly, the same analysis is carried out for IEEE 13 node test system also. The SC results are tabulated in table 14 and 15 for both min and max fault conditions.

Table 14: short circuit result for max fault condition and different ground impedances for IEEE 13-node test system

S. No.	Bus number	Ikss (KA) for 5 ohm	Ikss (KA) for 10 ohm	Ikss (KA) for 15 ohm	Ikss (KA) for 20 ohm
1	0	1.383196	0.695112	0.464184	0.348428
2	1	1.229788	0.659543	0.448966	0.340055
3	2	1.159602	0.641367	0.440958	0.335592
4	3	1.159602	0.641367	0.440958	0.335592
5	4	1.123243	0.6301	0.435632	0.332511
6	5	1.066632	0.613349	0.427885	0.328086
7	6	1.064429	0.61887	0.431779	0.330748
8	7	1.064429	0.61887	0.431779	0.330748
9	8	1.018197	0.603426	0.424377	0.326453
10	9	1.010667	0.601854	0.423832	0.326205
11	10	0.961792	0.585603	0.416105	0.321751
12	11	0.911939	0.56541	0.405615	0.315397
13	12	0.988031	0.597818	0.422715	0.325833

Table 15: short circuit result for min fault condition and different ground impedances for IEEE 13-node test system

S. No.	Bus number	Ikss (KA) for 5 ohm hIf	Ikss (KA) for 10 ohm	Ikss (KA) for 15 ohmHIF	Ikss (KA) for 20 ohm
1	0	1.269681	0.634986	0.423349	0.317521
2	1	1.111112	0.597453	0.40714	0.308554
3	2	1.040357	0.57855	0.398698	0.303814
4	3	1.040356	0.578549	0.398698	0.303814
5	4	1.002386	0.566572	0.392992	0.300499
6	5	0.945985	0.549274	0.384864	0.295816

7	6	0.951638	0.556635	0.389534	0.298902
8	7	0.951638	0.556635	0.389534	0.298902
9	8	0.904848	0.540571	0.381732	0.294339
10	9	0.898489	0.539213	0.381255	0.294122
11	10	0.850791	0.522738	0.373266	0.289463
12	11	0.799364	0.501596	0.36217	0.282698
13	12	0.880268	0.536086	0.380463	0.293904

The evaluation metrics result is tabulated in table 16 for training and testing datasets.

Table 16: evaluation metrics for training and testing datasets for 13-node test system

S. No.	Evaluation metrics	For Training Dataset		For Testing Dataset	
		LR	SVM	LR	SVM
1.	MAE	1.04608	1.02808	0.93737	0.65555
2.	MSE	1.72795	2.16279	1.38780	0.7357
3.	R2_score	0.87896	0.84851	0.89243	0.9429

Few values of actual fault bus number and the predicted fault bus number for training and testing cases for IEEE 13-node test system individually for both LR and SVM methods are tabulated in table 17,18,19, and 20 respectively.

Table 17: LR train values and their predicted values

S. No.	Actual value	Predicted value	S. No.	Actual value	Predicted value
1	5	5.388028	14	5	5.500816
2	8	8.009742	15	3	4.185817
3	10	9.877814	16	7	7.284487
4	7	7.400074	17	12	9.859625
5	6	7.396686	18	0	-2.47572
6	1	2.450467	19	12	9.807983
7	9	8.281916	20	4	4.323301
8	1	2.662984	21	11	10.54494
9	0	-2.49545	22	8	8.127972
10	6	7.35107	23	12	9.550667
11	3	3.99517	24	1	2.658319
12	1	2.600603	25	10	9.756605
13	6	7.284487			

Table 18: LR test values and their predicted values

S. No.	Actual value of fault bus	Predicted value Bus no	S. No.	Actual value	Predicted value
1	6	7.269125	14	5	5.316361
2	8	7.886219	15	10	9.481333
3	12	9.440644	16	4	4.378633
4	10	9.889817	17	1	2.372711
5	2	4.185616	18	5	5.316361
6	9	8.582058	19	11	10.15113
7	5	5.504597	20	3	4.060776
8	2	3.995066	21	1	2.701993
9	1	2.481133	22	8	7.782989

10	5	5.425619	23	11	10.10045
11	10	9.481333	24	3	4.176188
12	4	4.378633	25	8	8.131647
13	1	2.372711			

Table 19: SVM train values and their predicted values

S. No.	Actual value	Predicted value	S. No.	Actual value	Predicted value
1	5	5.038316	11	5	4.953925
2	8	8.259974	12	3	3.463688
3	10	9.936503	13	7	7.581239
4	7	7.606514	14	12	10.44234
5	6	7.495657	15	0	-3.70923
6	1	1.769081	16	12	10.44212
7	9	8.731627	17	5	4.953925
8	1	1.973282	18	3	3.463688
9	0	-3.95708	19	7	7.581239
10	6	7.611618	20	12	10.44234

Table 20: SVM test values and their predicted values

S. No.	Actual value	Predicted value	S. No.	Actual value	Predicted value
1	6	7.596954	11	10	9.906241
2	8	8.147374	12	4	3.725824
3	12	10.4031	13	1	1.924184
4	10	9.856712	14	5	5.029825
5	2	3.463677	15	11	10.03705
6	9	8.772508	16	3	3.642477
7	5	5.065575	17	1	1.948003
8	2	3.614059	18	8	8.251057
9	1	1.934997	19	11	10.03876
10	5	5.042325	20	3	3.652288

The actual and predicted fault bus number for both the methods is represented graphically as shown in figure 10 and 11.

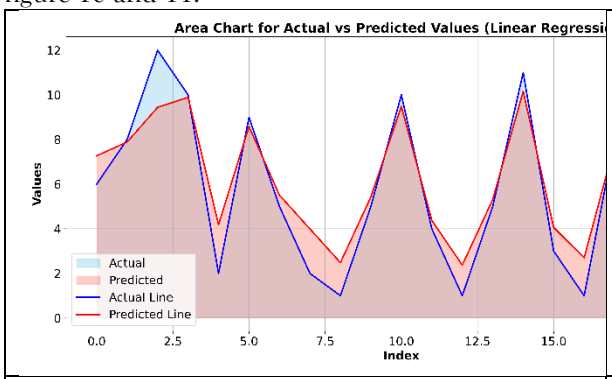


Figure 10: actual vs predicted for linear regression area chart for 13 node test system

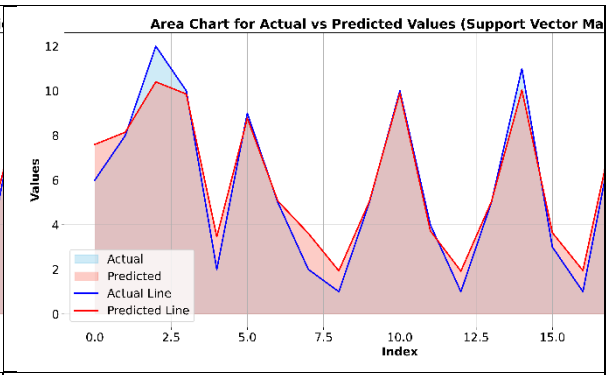


Figure 11: actual vs predicted for support vector machines area chart for 13 node test system

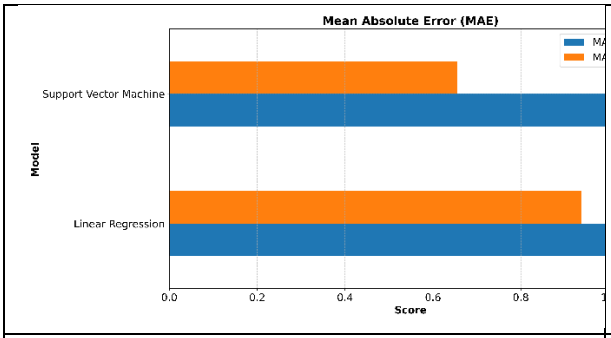


Figure 12: MAE Train vs Test for 13 node test system

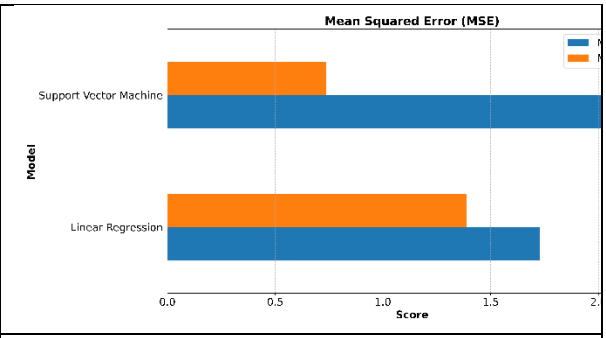


Figure 13: MSE Train vs Test for 13 node test system

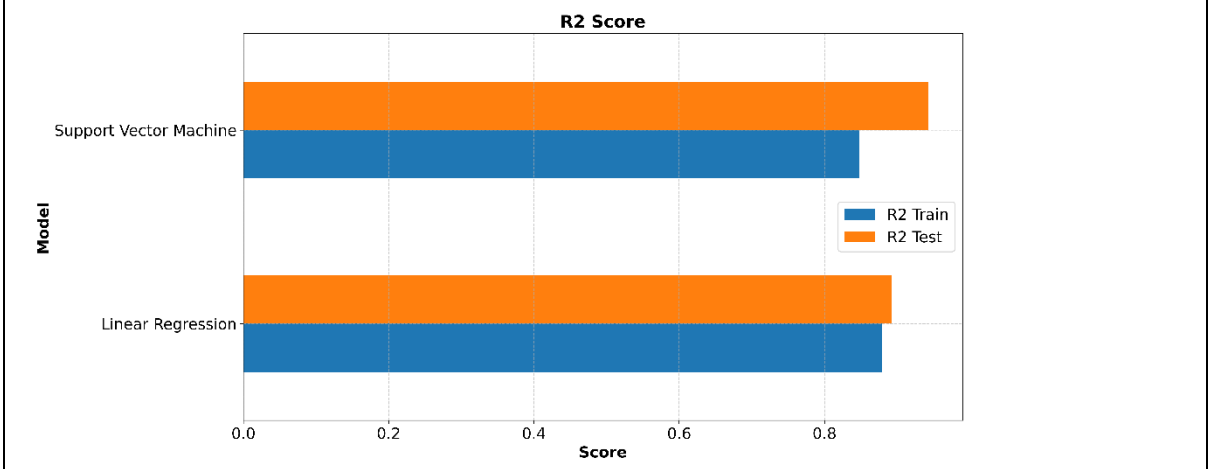


Figure 14: R2 Score Train vs Test for 13 node test system

The metrics result plots for MAE, MSE and R2_score for 13 -node test system is seen in figure 12,13,14 for training and testing the model.

MICROCLIMATIC BASED IGF AND TGF FAULT PREDICTION USING cubic-SVM, and optimized SVM

Table 21 dataset for IGF AND TGF faults in different microclimatic conditions

Row	Bus No.	Fault Type	Cycle On (ms)	Cycle Off (ms)	Peak Current (kA)	Fault Resistance (Ω)	Temp. (°C)	Humidity (%)	Wind Speed (m/s)
1	0	IGF	14.36	34.47	0.39	108.3	32.8	61	2.7
2	1	IGF	28.77	15.58	0.65	90.7	32.8	61	2.7
3	2	IGF	23.3	21.69	0.32	174.31	32.8	61	2.7
4	3	IGF	19.97	24.65	0.94	103.51	32.8	61	2.7
5	4	IGF	8.9	28.24	0.48	92.14	32.8	61	2.7
6	5	IGF	8.9	41.41	0.76	131.4	32.8	61	2.7
7	6	IGF	6.45	17.99	0.52	71.14	32.8	61	2.7
8	7	IGF	26.65	30.57	0.66	170.33	32.8	61	2.7
9	8	IGF	20.03	33.7	0.68	61.18	32.8	61	2.7
10	9	IGF	22.7	11.86	0.43	198.03	32.8	61	2.7

11	10	TGF	19.85	21.34	0.68	152.1	32.8	61	2.7
12	11	TGF	25.22	12.45	0.82	140.2	32.8	61	2.7
13	12	TGF	30.13	17.28	0.75	160.8	32.8	61	2.7
1	13	TGF	27.89	19.77	0.85	132.5	32.8	61	2.7
5	14	TGF	21.5	15.98	0.72	148.7	32.8	61	2.7
6	15	TGF	24.44	18.66	0.79	145.3	32.8	61	2.7
7	16	TGF	29.12	14.78	0.83	138.6	32.8	61	2.7
8	17	TGF	22.98	16.22	0.77	155.4	32.8	61	2.7
9	18	TGF	26.71	13.87	0.81	142.9	32.8	61	2.7
10	19	TGF	28.34	15.45	0.84	150.6	32.8	61	2.7

Table 22 Dataset of Single Line to Ground Faults (SLGF) in different microclimatic condition

S. No.	Bus No.	ISSS (kA)	Temperature (°C)	Humidity (%)	Wind Speed (m/s)	Fault Type-Different types of SLGF Faults
1	0	1.383196	32.8	61.5	2.7	ISSS (5Ω) – SIFs
2	1	1.227978	32.8	62.30	2.7	ISSS (5Ω) – SIFs
3	2	1.159602	32.8	61.14	2.7	ISSS (5Ω) – SIFs
4	3	1.140000	32.8	61	2.7	ISSS (5Ω) – SIFs
5	4	1.123243	32.8	61	2.7	ISSS (5Ω) – SIFs
6	5	0.695112	32.8	61	2.7	ISSS (10Ω) – MIFs
7	6	0.659543	32.8	61	2.7	ISSS (10Ω) – MIFs
8	7	0.641367	32.8	61	2.7	ISSS (10Ω) – MIFs
9	8	0.461484	32.8	61	2.7	ISSS (15Ω) – HIFs
10	9	0.448966	32.8	61	2.7	ISSS (15Ω) – HIFs
11	10	0.348428	32.8	61	2.7	ISSS (20Ω) – AGFs/HIFs
12	11	0.330055	32.8	61	2.7	ISSS (20Ω) – AGFs/HIFs
13	12	0.335892	32.8	61	2.7	ISSS (20Ω) – AGFs/HIFs

(i) CUBIC-SVM

A cubic Support Vector Machine (SVM) is a classification model that uses a third-degree polynomial kernel to separate data that cannot be divided by a straight line. Unlike a linear SVM, which only considers simple boundaries, a cubic SVM maps the input features into a higher-dimensional space where relationships up to the third order are taken into account. This allows the model to capture more complex interactions, such as combinations of variables influencing the output together. The cubic kernel function is expressed as $K(x, x') = (\gamma\, x^T x' + r)^3$, where γ adjusts the influence of each feature and r is an offset. By raising the inner product to the power of three, the model introduces flexibility while still avoiding the extreme complexity of higher-degree polynomials. In practice, cubic SVMs are often applied to datasets where moderate nonlinear patterns exist, and they generally perform better than linear models when such relationships are present. However, they also require careful tuning of hyperparameters like C, γ, and r, since poor choices can lead to overfitting or underfitting.

(ii) PSO-SVM

PSO-SVM provides highest classification accuracy or lowest validation error. In PSO-SVM, each particle in the swarm represents set of SVM hyperparameters, such as C and gamma. The position of a particle is a vector [C, gamma]. The fitness of a particle is evaluated by training an SVM with those parameters on training data and measuring its performance on validation data. The swarm iteratively updates particle positions and velocities to find better solutions over generations. This process continues until convergence criteria are met. Manual tuning is labor-intensive and often suboptimal, especially when the parameter

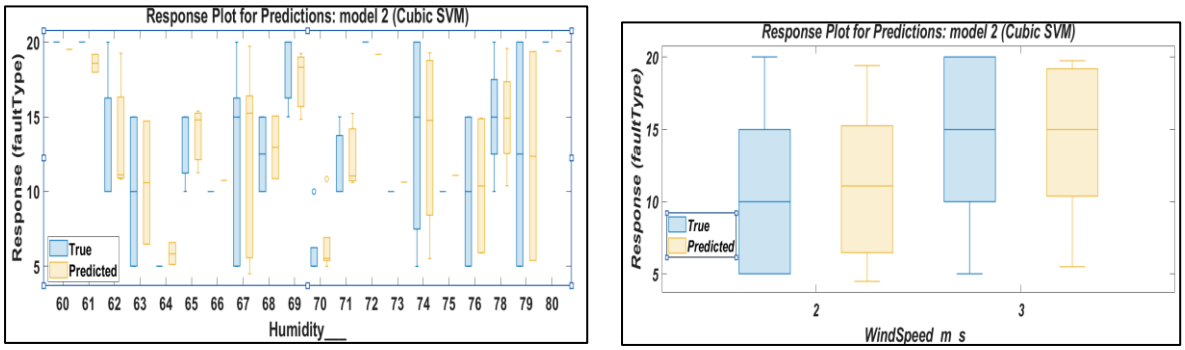
space is large or contains complex interactions. PSO automates through trial-and-error approaches and ensures more systematic search for optimal parameters.

(iii) BO-SVM

Bayesian Optimization selects SVM hyperparameters and increases the predictive accuracy. BO-SVM performs efficient evaluations and handles noisy objective measurements. BO-SVM avoids local minima and scalable for higher dimensional hyperparameter spaces. BO-SVM Controls trade-off between margin maximization and classification error through C , which is a regularization parameter and Controls how far the influence of a single training point reaches through γ (gamma in RBF kernel). These tuning avoids underfitting or overfitting. BO-SVM find the values of hyperparameters such as C, γ, that minimizes an error metric. For the prediction of the faults in the IGF and TGF, the proposed algorithms such as (i) Cubic-SVM (ii) PSO-SVM (iii) BO-SVM are applied.

COMPARISON OF CUBIC-SVM, PSO-SVM, AND BO-SVM IN FAULT PREDICTION FOR IGF AND TGF SYSTEMS

For predicting the type of faults such as IGF and TGF, Cubic-SVM, PSO-SVM, and BO-SVM algorithms are used and evaluated the key metrics such as precision, accuracy, and recall for assessment of their effectiveness.The Cubic-SVM method has robust decision boundaries, demonstrated reliable accuracy and balanced recall. However, Particle Swarm Optimization (PSO) and Bayesian Optimization (BO) are integrated with SVM, and provides higher precision and recall, minimizing false alarms and missed detections in fault prediction scenarios. BO-SVM method has achieved the highest accuracy and maintains strong precision and recall. Cubic-SVM provides high accuracy in fault detection. Cubic_SVM leads to more reliable and accurate fault detection such as IGF and TGF and supports for the maintenance decisions and system reliability.



(a) Humidity (b) WindSpeed

Figure 15 Cubic-SVM based fault type and (a) Humidity relation, (b) Windspeed.

Cubic-SVM achieves balanced accuracy and recall, and detects initial fault detection tasks. Cubic-SVM performance may plateau without further optimization, especially in complex or noisy datasets. PSO-SVM fine-tunes SVM parameters dynamically and provides higher precision and recall, which reduces both false positives and false negatives in fault prediction. PSO-SVM minimizes missed detections and false alarms. PSO enhances the SVM's adaptability to data patterns. BO-SVM uses probabilistic modeling to optimize SVM hyperparameters efficiently. BO-SVM leads to the highest accuracy and maintains strong precision and recall. The systematic search for optimal parameters ensures robust performance, even as data characteristics shift or become more complex. Cubic-SVM provides a solid starting point with reliable, balanced results. PSO-SVM enhances detection by adaptively tuning model parameters, leading to better fault identification and fewer errors. Cubic-SVM is effective, significantly boosts predictive accuracy and reliability for fault detection in IGF and TGF systems. This leads to more informed maintenance and operational decisions, ultimately improving system reliability and safety.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 8, 2025
https://theaspd.com/index.php

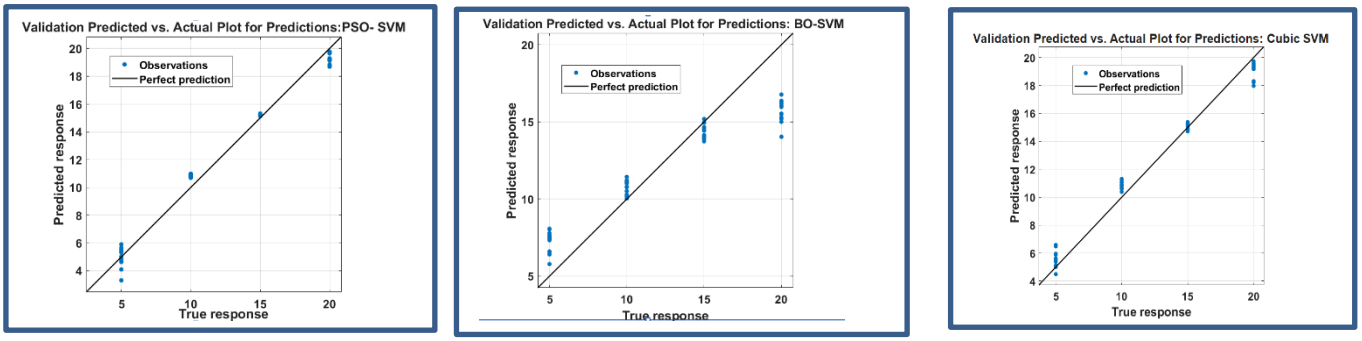


Figure 16 Cubic-SVM, PSO-SVM and BO-SVM Fault type prediction response in different types of microclimatic conditions.

Table 23 shows the IGF and TGF fault detection in different microclimatic condition. **Table 24 shows Single Line to Ground Faults (SLGF) in different microclimatic condition.**

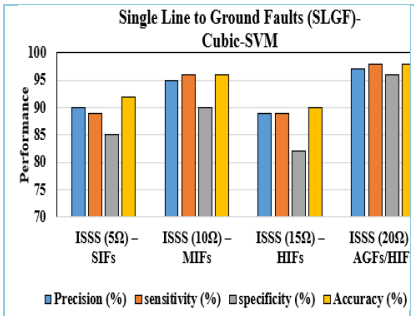
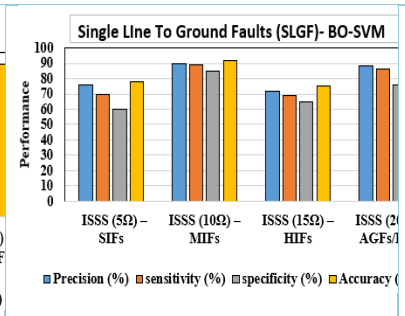
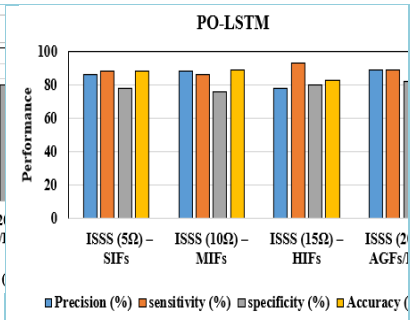
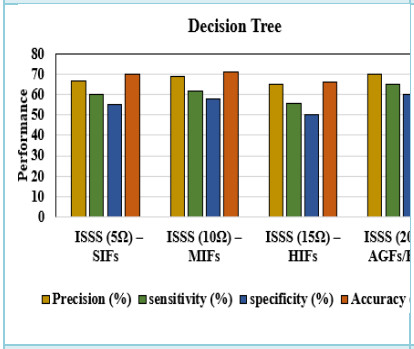
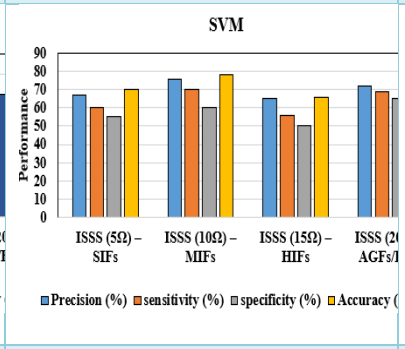
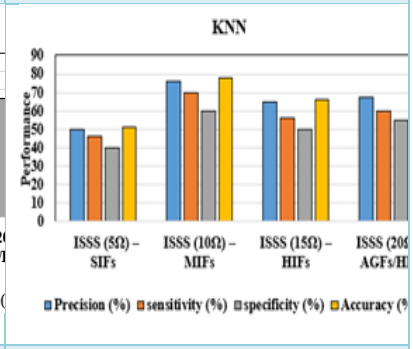
		
Proposed	**Proposed**	**Ref(29)**
		
Ref(30)	**Ref(31)**	**Ref(32)**

Table 23 IGF and TGF fault detection in different microclimatic condition

Algorithm	Precision (%)	Accuracy (%)	Recall (%)	Key Advantage in
Cubic-SVM	95	97	96	Efficient optimization, best overall accuracy
PSO-SVM	82	87	89	Adaptive parameter tuning, fewer false alarms

BO-SVM	85	90	93	Robust baseline performance

Table 24 Single Line to Ground Faults (SLGF) in different microclimatic condition

Algorithm	Precision (%)	Accuracy (%)	Recall (%)	Key Advantage in Single Line to Ground Faults (SLGF) detection
Cubic-SVM	96	98	97	Efficient optimization, best overall accuracy
PSO-SVM	84	83	86	Adaptive parameter tuning, fewer false alarms
BO-SVM	95	92	94	Robust baseline performance

CONCLUSION

From the analysis, the Cubic-SVM and BO-SVM algorithm attained better results when compared to linear regression model. The cubic-SVM evaluation metric values of MAE of 0.2313, 0.65555 and MSE of 0.1193, 0.7357 and R2_score of 0.8955, 0.9429 when compared to that of LR method with MAE of 0.3332, 0.65555 and MSE of 0.1446, 0.7357 and R2_score of 0.8734, 0.9429 are obtained for IEEE 4-node and 13 – node radial test systems for detecting the HIF accurately. The cubic-SVM method has been suggested as best trained model with less error. Table 24 shows the overall performance of the proposed algorithm such as Single Line to Ground Faults (SLGF) IGF and TGF fault in different microclimatic condition and without including the microclimatic data in the algorithm**.**

Table 25. Comparison of proposed algorithm with and without microclimatic data for SLGF types and IGF and TGF fault predictions

Single Line to Ground Faults (SLGF) in different microclimatic condition			
Algorithm	**Precision**	**Accuracy**	**Recall**
Without Microclimatic Data			
Cubic-SVM	90	92	91
BO-SVM	80	91	90
With Microclimatic Data			
Cubic-SVM	95	97	96
BO-SVM	85	90	93
IGF and TGF fault detection in different microclimatic condition			
Algorithm	**Precision**	**Accuracy**	**Recall**
Without Microclimatic Data			
Cubic-SVM	90	92	91
BO-SVM	80	91	90
With Microclimatic Data			
Cubic-SVM	95	97	96
BO-SVM	85	90	93

In future, the implementation of combination of such ML techniques may give better metrics when compared with other conventional methods.

REFERENCES

[1] A. Navid Bayati, Ebrahim Balouji, Hamid Reza Baghaee, Amin Hajizadeh, Mohsen Soltani, Zhengyu Lin, Mehdi Savaghebi, Locating high-impedance faults in DC microgrid clusters using support vector machines, Applied Energy, Volume 308, 2022, 118338, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2021.118338.
[2] Sarwar, M., Mehmood, F., Abid, M., Khan, A. Q., Gul, S. T., & Khan, A. S. (2020). High impedance fault detection and isolation in power distribution networks using support vector machines. Journal of King Saud University - Engineering Sciences, 32(8), 524–535. https://doi.org/10.1016/j.jksues.2019.07.001.
[3] A. Eskandari, J. Milimonfared and M. Aghaei, "Fault Detection and Classification for Photovoltaic Systems Based on Hierarchical Classification and Machine Learning Technique," in IEEE Transactions on Industrial Electronics, vol. 68, no. 12, pp. 12750-12759, Dec. 2021, doi: 10.1109/TIE.2020.3047066.
[4] H. Livani and C. Y. Evrenosoglu, "A Machine Learning and Wavelet-Based Fault Location Method for Hybrid Transmission Lines," in IEEE Transactions on Smart Grid, vol. 5, no. 1, pp. 51-59, Jan. 2014, doi: 10.1109/TSG.2013.2260421.
[5] D. Guillen, J. Olveres, V. Torres-García and B. Escalante-Ramírez, "Hermite Transform Based Algorithm for Detection and Classification of High Impedance Faults," in IEEE Access, vol. 10, pp. 79962-79973, 2022, doi: 10.1109/ACCESS.2022.3194525.
[6] Wanxing Sheng, Keyan Liu, Dongli Jia, Yao Wang, An improved high-impedance fault identification scheme for distribution networks based on kernel extreme learning machine, International Journal of Electrical Power & Energy Systems, Volume 155, Part B, 2024, 109543, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2023.109543.
[7] Mou-Fa Guo, Meitao Yao, Jian-Hong Gao, Wen-Li Liu, Shuyue Lin, An incremental high impedance fault detection method under non-stationary environments in distribution networks, International Journal of Electrical Power & Energy Systems, Volume 156, 2024, 109705, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2023.109705.
[8] L. Thurner et al., "Pandapower—An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems," in IEEE Transactions on Power Systems, vol. 33, no. 6, pp. 6510-6521, Nov. 2018, doi: 10.1109/TPWRS.2018.2829021.
[9] W. H. Kersting, "Radial distribution test feeders," in IEEE Transactions on Power Systems, vol. 6, no. 3, pp. 975-985, Aug. 1991, doi: 10.1109/59.119237.
[10] Suliman, Mohammed Yahya; Alkhayyat, Mahmood Taha. Archives of Electrical Engineering; Warsaw Vol.70, Iss.4, 2021, 873-886. DOI:10.24425/aee.2021.138267.
[11] P. K. Bera, V. Kumar, S. R. Pani and V. Bargate, "Detection of High Impedance Faults in Microgrids using Machine Learning," 2022 IEEE Green Energy and Smart System Systems (IGESSC), Long Beach, CA, USA, 2022, pp. 1-5, doi: 10.1109/IGESSC55810.2022.9955330.
[12] Hong and Y. -S. Kim, "Supervised Learning Approach for State Estimation of Unmeasured Points of Distribution Network," in IEEE Access, vol. 8, pp. 113918-113931, 2020, doi: 10.1109/ACCESS.2020.3003049.
[13] V. Veerasamy et al., "LSTM Recurrent Neural Network Classifier for High Impedance Fault Detection in Solar PV Integrated Power System," in IEEE Access, vol. 9, pp. 32672-32687, 2021, doi: 10.1109/ACCESS.2021.3060800.
[14] K. Moloi, J. A. Jordaan and Y. Hamam, "High impedance fault detection technique based on Discrete Wavelet Transform and support vector machine in power distribution networks," 2017 IEEE AFRICON, Cape Town, South Africa, 2017, pp. 9-14, doi: 10.1109/AFRCON.2017.8095447.
[15] V. Le and X. Yao, "Ensemble Machine Learning Based Adaptive Arc Fault Detection for DC Distribution Systems," 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), Anaheim, CA, USA, 2019, pp. 1984-1989, doi: 10.1109/APEC.2019.8721922.
[16] X. -D. Zeng, M. -F. Guo and D. -Y. Chen, "Machine-learning-based single-phase-to-ground fault detection in distribution systems," 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 2017, pp. 1-6, doi: 10.1109/EI2.2017.8245233.
[17] Rai, K.; Hojatpanah, F.; Badrkhani Ajaei, F.; Grolinger, K. Deep Learning for High-Impedance Fault Detection: Convolutional Autoencoders. Energies 2021, 14, 3623. https://doi.org/10.3390/en14123623.
[18] Mostafa Eslami, Mohsen Jannati, S. Sepehr Tabatabaei, An improved protection strategy based on PCC-SVM algorithm for identification of high impedance arcing fault in smart microgrids in the presence of distributed generation, Measurement, Volume 175, 2021, 109149, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2021.109149.
[19] M. M. Badr, M. S. Hamad, A. S. Abdel-Khalik, R. A. Hamdy, S. Ahmed and E. Hamdan, "Fault Identification of Photovoltaic Array Based on Machine Learning Classifiers," in IEEE Access, vol. 9, pp. 159113-159132, 2021, doi: 10.1109/ACCESS.2021.3130889.
[20] A. Forouzesh, M. S. Golsorkhi, M. Savaghebi and M. Baharizadeh, "Fault Location Identification in Meshed AC Microgrids Using Interharmonic Injection," 7th Iran Wind Energy Conference (IWEC2021), Shahrood, Iran, 2021, pp. 1-6, doi: 10.1109/IWEC52400.2021.9467007.
[21] H. M. Mesbah Maruf, F. Müller, M. S. Hassan and B. Chowdhury, "Locating Faults in Distribution Systems in the Presence of Distributed Generation using Machine Learning Techniques," 2018 9th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), Charlotte, NC, USA, 2018, pp. 1-6, doi: 10.1109/PEDG.2018.8447728.

[22] S. R. Samantaray and P. K. Dash, ‘‘High impedance fault detection in distribution feeders using extended Kalman filter and support vector machine,’’ Eur. Trans. Electr. Power, vol. 20, no. 3, pp. 382–393, Apr. 2010, doi: 10.1002/etep.321.
[23] J. -Y. Wu, S. Lan, S. -J. Xiao and Y. -B. Yuan, "Single Pole-to-Ground Fault Location System for MMC-HVDC Transmission Lines Based on Active Pulse and CEEMDAN," in IEEE Access, vol. 9, pp. 42226-42235, 2021, doi: 10.1109/ACCESS.2021.3062703.

[24] K. V. Shihabudheen et al., "Detection of High Impedance Fault using Machine Learning Techniques," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, pp. 2117-2122, doi: 10.1109/TENCON.2019.8929365.
[25] https://www.geeksforgeeks.org/ml-linear-regression/
[26] https://www.geeksforgeeks.org/support-vector-machine-algorithm/
[27] https://www.pandapower.org/
[28] Lavanya, S., Prabakaran, S. and Ashok Kumar, N. 2022. A Deep Learning Technique for Detecting High Impedance Faults in Medium Voltage Distribution Networks. Engineering, Technology & Applied Science Research. 12, 6 (Dec. 2022), 9477–9482. DOI:https://doi.org/10.48084/etasr.5288.
[29] Lin, G., Abdel-salam, M., Hu, G. and Jia, H., 2025. Adaptive Differentiated Parrot Optimization: A Multi-Strategy Enhanced Algorithm for Global Optimization with Wind Power Forecasting Applications. *Biomimetics*, *10*(8), p.542.
[30] Rahman, T., Hasan, T., Ahammad, A., Ahmed, I. and Rakhaine, N., 2025. MLPNN and Ensemble Learning Algorithm for Transmission Line Fault Classification. International Transactions on Electrical Energy Systems, 2025(1), p.6114718.
[31] Guo, Z., Deng, Y., Zhang, T. and Huang, Z., 2025. A novel directional pilot protection method for EHV transmission lines based on S-transform and SVM. *Discover Applied Sciences*, 7(4), pp.1-19.
[32] Xiao, X., Yin, Y., Zhou, H. and Zhou, Q., 2025, March. Research and application of transmission line fault location using k-nearest neighbor algorithm. In *International Conference on Physics, Photonics, and Optical Engineering (ICPPOE 2024)* (Vol. 13552, pp. 429-436). SPIE.

AUTHORS PROFILE

Guttimari Trivedh Yadav is a student member of IEEE, completed his B.Tech from sri venkateswara college of engineering, and M.Tech from sri vidyanikethan engineering college, tirupati. Currently he is pursuing his Ph.D in Sri Venkateswara University college of engineering under the supervision of Professor T Gowri Manohar in the area of power systems. His areas of interest are power systems, renewable energy sources etc.

T Gowri Manohar is a senior member of IEEE, currently working as a Professor in Sri Venkateswara University college of engineering, tirupati. He has 25 years of experience in teaching and research. His areas of interest are power system operation and control, reactive power control, power quality, renewable energy sources etc.