Power Quality Event Classification in Distribution Grids Using Machine Learning

Thet Paing Tun
School of Computing, Engineering and Digital Technologies
Teesside University
Middlesbrough, UK
Email: W9057615@tees.ac.uk

Gobind Pillai
School of Computing, Engineering and Digital Technologies
Teesside University
Middlesbrough, UK
Email: g.g.pillai@tees.ac.uk

***Abstract*— With the penetration of non-linear loads, renewables and distributed generation with power electronic converters, solutions for maintaining good power quality have become a major concern for the stakeholders of electrical power systems. In this paper, a machine learning based model for power quality event classification is developed and tested. 16 categories of the most commonly occurring power quality events are classified by means of wavelet transform and select machine learning based methods to evaluate the best performing machine learning model. The outcome of classifications and effectiveness of machine learning methods is evaluated using the 'Classification Learners' application in MATLAB. The selected machine learning model is implemented in Simulink for test distribution grid circuits. The results obtained from simulation showed acceptable accuracy and performance and demonstrated the efficiency of the model in different operating conditions.**

Keywords—power quality, machine learning, wavelet transform, classification.

I. Introduction

As the penetration rate of non-linear loads like electric vehicles and renewable based distributed generation integrating power electronic devices increase in modern electricity grids, power quality problems such as voltage violation, dip, swell, flicker and harmonics are increasing becoming common. Such power quality events are a threat to the electrical assets and devices connected along the grid and it may further lead to negatively affect the reliability of power system and the safety of network operators and electricity users [1, 2].

In addition to the existing techniques such as Fast Fourier Transform (FFT), intelligent machine learning techniques are starting to take a prominent role in power quality analysis and monitoring. According to the author of [3], the research outputs concerning the theme of power quality of electricity networks are still low compared to other smart grid themes. This research is aimed at contributing to the addressing of this research gap. The aim of this work is to demonstrate and evaluate the effectiveness of intelligent machine learning algorithms in classifying power quality events and monitoring power quality at distribution grids.

II. Methodology

This research is mainly composed of four main stages, namely: (i) Data preparation: selection of power quality events (event selection) and feature extraction; (ii) selecting suitable machine learning algorithms, their testing and validation; and (iii) testing the efficiency of the machine learning algorithm in monitoring power quality events by simulating it in test Simulink distribution grid models. The machine learning method selection approach of the work is summarised in Fig. 1. The overall research methodology of this work is summarised in Fig 2.

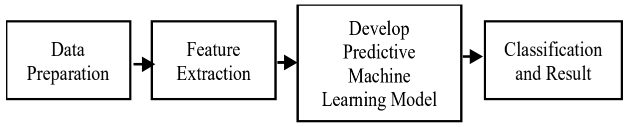

Fig. 1. Machine learning method seletion approach

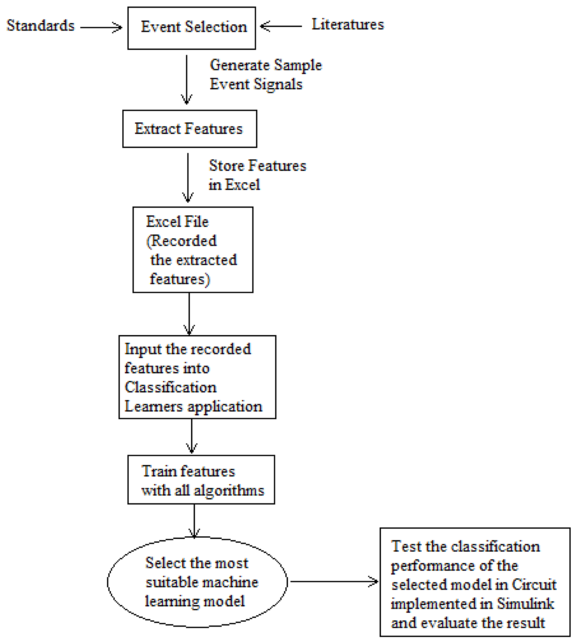

Fig. 2. Flowchart summarising the overall research methodology.

The IEEE standard on monitoring power quality [4] and available literature [5-7] were reviewed to select power quality events that is studied in this work. Table I provides the lists of events selected. Datasets with synthetic signals of individual events were then created in MATLAB for training and validation of machine learning algorithms.

Once the data samples created above are ready, features from those data samples are extracted. In this case, the statistical features described in Table II (see section below on Data Preparation) are selected by reviewing Wavelet Analyzer Toolbox and related literature such as [5], [8] and [9]. After the feature extraction is done, those features are imported into 'Classification Learners' application in MATLAB using which the most suitable machine learning model is selected as shown below in the flow chart of Fig 3.

The optimal machine learning model is obtained is then tested using IEEE-5 bus system and electric arc furnace [10] test distribution grid circuits implemented in Simulink. In testing phase, the scenarios of power quality events are created based on the methods proposed by [10].

TABLE I. SELECTED POWER QUALITY EVENTS

Event Index	**Event Description**
S1	Normal Condition
S2	Voltage Sag
S3	Voltage Swell
S4	Voltage with Harmonics
S5	Voltage Flicker
S6	Voltage Interruption
S7	Oscillatory Transient
S8	Impulsive Transient
S9	Voltage Sag with Flicker
S10	Voltage Swell with Flicker
S11	Voltage Flicker with Harmonics
S12	Voltage Sag with Harmonics
S13	Voltage Swell with Harmonics
S14	Voltage Interruption with Harmonics
S15	Voltage Sag with Transient
S16	Voltage Swell with Transient

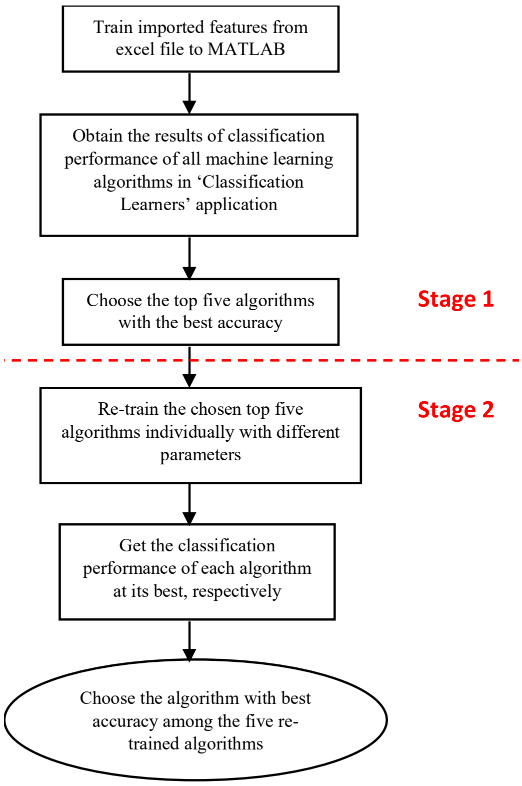

Fig. 3. Processing of Choosing Machine Learning Algorithm

III. DATA PREPARATION

Data PreparationRelevant groups of sample data are required to input into the machine learning model for training. Fig. 4 summarises the overall process of data preparation.

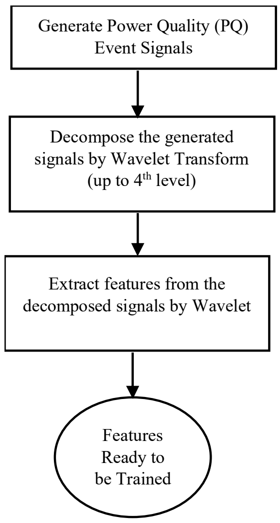

Fig. 4. Process flow of Data Preparation

A. Power Quality Signal Generation

180 samples were generated per event for all 16 individual power quality events categories described in Table I. The power quality events were created in MATLAB with a sampling frequency of 512 samples per cycle which is equivalent to 25.6kHz as recommended by IEEE 1159-2019 [4] to cover higher frequency events like oscillatory transients.

B. Signal Decomposition

The samples generated above are decomposed by being band-passed in a discrete wavelet transform for up to level-4 decomposition. Because of the lower frequency, a higher level of decomposition does not lead to a significant difference in the features utilised for classification. According to [5], the Daubechies 4 (DB4) wavelet performs much better than other types of wavelets. Hence, it was used in this study.

Normally, abnormal electrical power signals caused by fast electromagnetic transients are non-periodic with high-frequency components [11]. Authors of [12] recommend reconstructing the original signal for fast and reliable performance in power quality analysis applications. By removing noise, the specific signal can be analysed without ambiguity resulting in high performance when used by an intelligence system.

Mathematically, the wavelet signal is expressed as [11]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} . \psi\left(\frac{t-b}{a}\right) \text{.........} \quad (1)$$

$$where\ \psi_{a,b}(t) = Continuous\ Wavelet\ Signal$$

$$\frac{1}{\sqrt{a}} = Normalizing\ Constant, a$$

$$= Scaling\ Paramter,$$

$$b = Translation\ Paramter$$

$$With\ \ CWT_{(a,b)}\ continouous\ wavelet\ transform\ =$$

$$CWT_{(a,b)} = \int_{-\infty}^{+\infty} x(t).\psi_{a,b}(t) = \frac{1}{\sqrt{a}}.\int_{-\infty}^{+\infty} x(t)\,.\,\psi\left(\frac{t-b}{a}\right) dt \text{ ....} (2)$$

$$where\ x(t)$$

$$= signal\ to\ be\ decomposed\ by\ wavelet\ transform$$

However, the continuous wavelet transform mentioned above can create a burden for computational performance because of the huge amount of redundant information. Therefore, the discretised wavelet transform (DWT) [11] with the discretised scale and translation factors as described below is used:

$$\psi_{m,p} = \frac{1}{\sqrt{a_0^m}} . \psi\left(\frac{t - pb_0 a_0^m}{a_0^m}\right) \ldots\ldots\ldots (3)$$

$$DWT_{m,k} = \frac{1}{\sqrt{a}} . \sum_n x[n] . g\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \ldots\ldots\ldots. (4)$$

$$\begin{aligned} where\ \psi_{m,p} &= Discrete\ Wavelet\ Signal \\ m &= Discrete\ Scaling\ Parameter \\ n &= Discrete\ Translation\ Parameter \\ a_0 &= Discrete\ Scaling\ Factor \\ b_0 &= Discrete\ Translation\ Factor \\ g &= Mother\ Wavelet, x[n] = Discrete\ Signal \end{aligned}$$

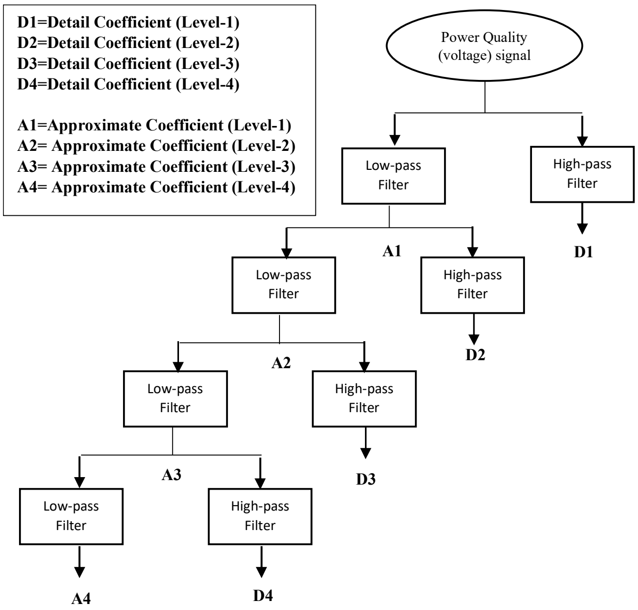

Fig. 5. Wavelet Decomposition up to 4th Level

C. Feature Extraction

After decomposing the total 2880 signal samples of the chosen 16 categories of power quality events, the features of every wavelet-decomposed signal (from level-1 to level-4 of detail coefficients and level-4 of approximate coefficients) must be extracted as described in Table II.

TABLE II. LIST OF FEATURES BEING USED

Feature Name	Description of Feature
Mean	$\frac{1}{N} . \sum_{k=1}^{N} s_k \ldots\ldots(5)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Median	$\underset{k=1,2,\ldots,N}{\text{median}}\ s_k \ldots\ldots (6)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Mode	$\underset{k=1,2,\ldots,N}{\text{mode}}\ s_k \ldots.. (7)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Standard Deviation	$\sqrt{\frac{\sum_{k=1}^{N}(s_k - Mean(S))^2}{N-1}} \ldots\ldots\ldots.(8)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$ $S = Set\ of\ sample = [s_1, s_2, \ldots, s_N]$
Mahattan Distance (L1 Norm)	$\sum_{k=1}^{N} \lvert s_k \rvert \ldots\ldots\ldots (9)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Euclidean Distance (L2 Norm)	$\sqrt{\sum_{k=1}^{N} \lvert s_k \rvert^2} \ldots\ldots.(10)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Entropy	$-\sum_{k=1}^{N} p_k . \log(p_k) \ldots\ldots.. (11)$ $where\ p_k = probability\ of\ occurence\ of\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Skewness	$\frac{\sum_{k=1}^{N}(s_k - Mean(S))^3}{(N-1).\sigma} \ldots\ldots.(12)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$ $S = Set\ of\ sample = [s_1, s_2, \ldots, s_N]$ $\sigma = Standard\ Deviation$
Kurtosis	$\frac{1}{N-1} . \sum_{k=1}^{N} (\frac{s_k - Mean(S)}{\sigma})^4 \ldots. (13)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$ $S = Set\ of\ sample = [s_1, s_2, \ldots, s_N]$ $\sigma = Standard\ Deviation$
Maximum	$\max_{k=1,2,\ldots,N}\ s_k \ldots\ldots. (14)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Minimum	$\min_{k=1,2,\ldots,N}\ s_k \ldots\ldots.(15)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$
Mean Absolute Deviation	$\frac{1}{N} . \sum_{k=1}^{N} \lvert s_k - Mean(S) \rvert \ldots.. (16)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$ $S = Set\ of\ sample = [s_1, s_2, \ldots, s_N]$
Median Absolute Deviation	$\underset{k=1,2,\ldots,N}{\text{median}}\ \lvert s_k - Median(S) \rvert \ldots\ldots.(17)$ $where\ s_k = signal\ value\ at\ kth\ sample$ $N = total\ number\ of\ samples\ in\ a\ single\ segment$ $S = Set\ of\ sample = [s_1, s_2, \ldots, s_N]$

IV. Selecting Machine Learning Model And Training

Following feature extraction, all samples extracted above are then imported into the 'Classification Learners' application in MATLAB for training. Fig. 6 shows the algorithm used for model training. In this work, the 5-fold cross-validation method described in Fig. 7 is used to validate the data samples and identify the best performing machine learning methods. The models that performed best for the imported data are based on the methods: (i) Ensemble Bagged Trees, (ii) Support Vector Machine (SVM) with cubic kernel, (iii) SVM with quadratic kernel, (iv) SVM with linear kernel and (v) Fine Tree. These methods are briefly described as below.

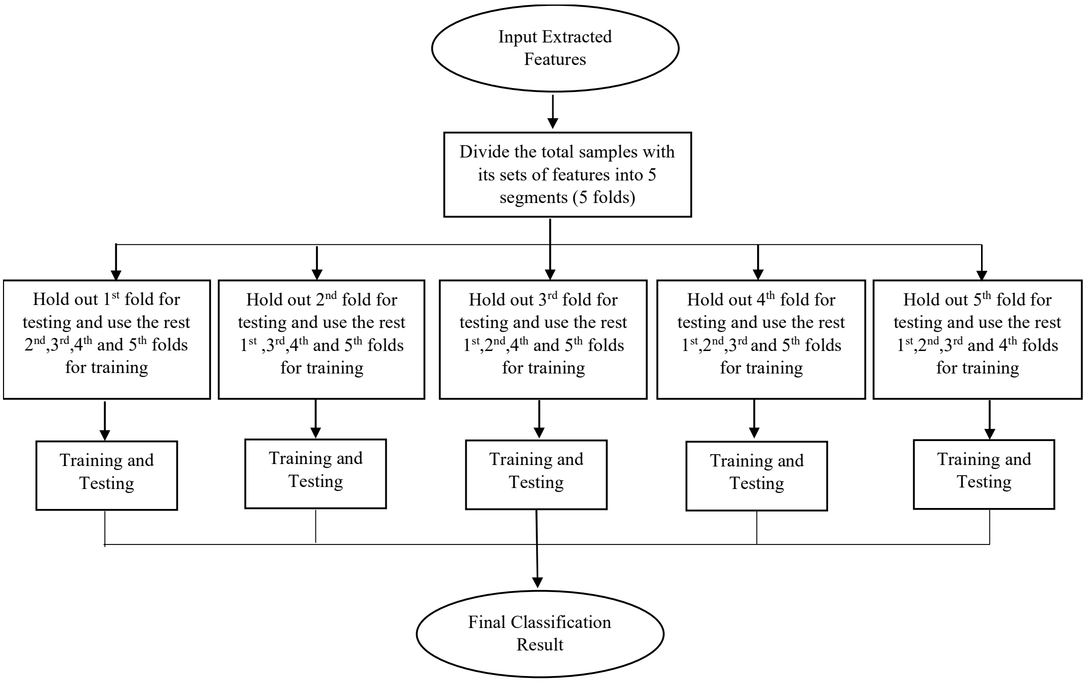

Fig. 6. Algorithm for Training Input Features

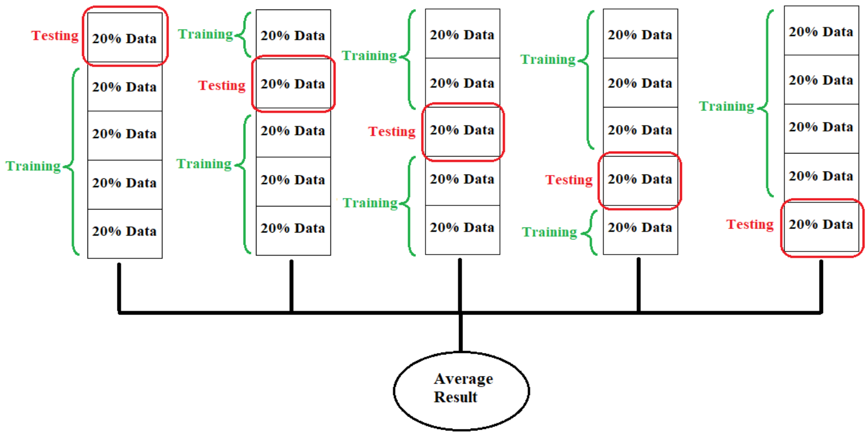

Fig. 7. Five-fold Cross-validation

A. Ensemble Bagged Trees

Ensemble Bagged Trees is a branch of the 'Ensemble Bagging' or 'Bootstrapping' method in which the decision tree is applied as a single classification model. The prediction results from all individual learners (trees) are combined and averaged to output the final outcome of classification. A voting system for the overall practically reduces variance and overfitting [13].

Mathematically, the ensemble bagging algorithm can be expressed as:

$$f_{bag} = f_1 + f_2 + \cdots + f_b \text{ ....(18)}$$

$where\ f_{bag} = prediction\ from\ ensemble\ bagging\ learner$
$f_1, f_2, \dots, f_b = prediction\ from\ all\ individual\ 'b'\ number\ of\ learners$

B. Support Vector Machine (SVM)

SVM does the classification function by inserting a hyperplane as the decision boundary separating different classes on the feature dataset. Its key objective is to find the optimal hyperplane with the maximal box constraint margin among all the classes to be classified while mapping the input data samples to the appropriate output classes. There are different types of kernels available to set up the hyperplane to classify the input data [14].

$$\text{Linear Kernel} \rightarrow W^T x + b,\ x \in R^d \text{ ....(19)}$$

$$\text{Polynomial} \rightarrow \left(W^T x + b\right)^d,\ \ \sigma > 0 \text{ ....(20)}$$

$where\ x = input\ data,$
$W^T = dimensional\ vector$
$= W_0, W_1, \dots, W_m\ (for\ m - dimensional\ vector)$
$b = biased\ term$
$d = degree\ of\ polynomail = 2,3, \dots.$

In this case, the box constraint is a parameter to adjust the fitting of the model to input data.

C. Fine Tree

Fine Tree is a highly flexible model type of Decision Tree method that has a larger number of leaves compared to other tree sub-categories like 'Medium Tree' and 'Coarse Tree'. A Decision Tree is composed of the main components: root, branch and leaf in which the root nodes contain the input data or features to be trained and the branch nodes contain the decision rules to train the data. Meanwhile, the leaf nodes show the output classes or labels resulted from training and the maximum number of splits that can be set in the application to control the accuracy of the model [15].

The top five methods described above are trained again with different parameters until optimal results are obtained. Table III shows the results.

TABLE III. RESULTS OF TOP FIVE MODELS TRAINED WITH OPTIMAL PARAMETERS

Machine Learning Model	Parameters	Accuracy
Ensemble Bagged Trees	Number of Learners = 120	98%
SVM Quadratic Kernel	Box Constraint Level = 12	95.1%
SVM Cubic Kernel	Box Constraint Level = 20	94.5%
SVM Linear Kernel	Box Constraint Level = 25	94.5%
Fine Tree	All	92.8%

As seen in the results of Table III, the Ensemble Bagged Trees model had the best accuracy and is selected as machine learning model for the next phase of work. Fig. 8 shows the confusion matrix for the selected model.

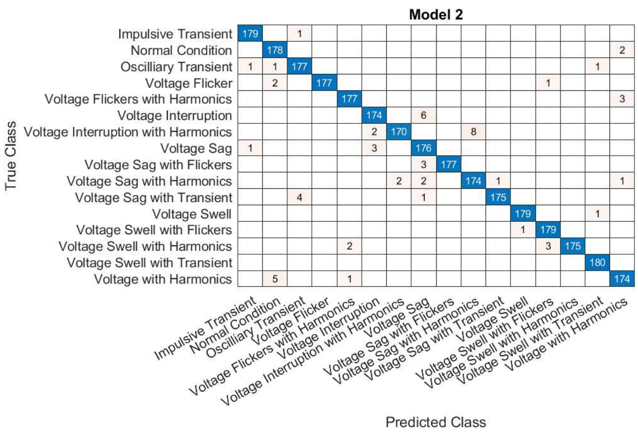

Fig. 8. Confusion Matrix of Selected 'Ensemble Bagged Tree'

V. MODEL TESTING AND RESULTS

A. Test with distribution grid circuit based on IEEE 5-bus system

The selected machine learning model is tested on a distribution grid circuit based on IEEE 5-bus architecture implemented in Simulink (Fig. 9) by creating events. The simulation is run for 13 seconds. Occurrence of different events at all 3 phases of the individual buses, for 1-second segments, are simulated as indicated in Table IV. The machine learning model selected reads in the voltages and then classifies the waveforms for every time-segment. The results showed acceptable performances but contains some minor ambiguities in classification results.

For all events in the time segments other than 6-7 second, the machine learning model classified all the occurred events correctly (S1, S2 etc.) as expected from Table IV. In the 6-7 second segment, all phases of MV1, MV3, LV3 and LV4 are correctly classified as 'S3' which is 'voltage swell'. In MV2, MV4, MV5 and LV1, the phase-3 of those buses were ambiguously classified as having S16. This is because of the voltage variation that occurred at these buses being far from the zero-crossing point. Although it is acceptably low to be classified as an ordinary voltage swell, the model was ambiguous in classifying it as swell with transient. Similarly, phase-2 of LV2 in this time segment were mis-classified as swell with transient (S16) for the same reason.

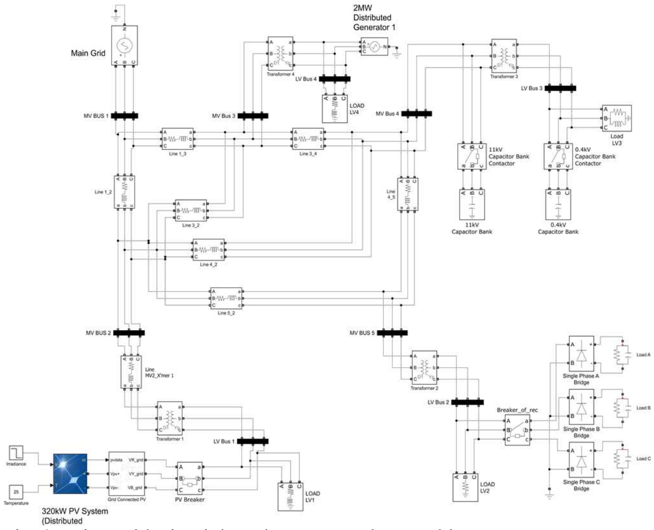

Fig. 9. Microgrid Circuit based on IEEE 5-bus Architecture

TABLE IV. EVENTS OCCURRING AT ALL 3 PHASES OF EVERY MEDIUM VOLTAGE AND LOW VOLTAGE BUS

	Time Segments (s)												
	0 - 1s	1 - 2s	2 - 3s	3 - 4s	4 - 5s	5 - 6s	6 - 7s	7 - 8s	8 - 9s	9 - 10s	10 - 11s	11 - 12s	12 - 13s
MV1	S1	S2	S1	S1	S1	S2	S3	S1	S1	S2	S3	S6	S1
MV2	S1	S2	S1	S8	S1	S2	S3	S1	S1	S2	S3	S6	S1
MV3	S1	S2	S1	S8	S1	S2	S3	S1	S1	S2	S3	S6	S1
MV4	S1	S2	S1	S8	S1	S2	S3	S1	S1	S2	S3	S6	S1
MV5	S1	S2	S1	S8	S1	S2	S3	S1	S1	S2	S3	S6	S1
LV1	S4	S12	S1	S1	S1	S2	S3	S1	S1	S2	S3	S6	S1
LV2	S1	S2	S1	S1	S1	S2	S3	S1	S4	S12	S13	S14	S4
LV3	S1	S2	S1	S1	S7	S15	S16	S1	S1	S2	S3	S6	S1
LV4	S1	S2	S1	S1	S1	S2	S3	S1	S1	S2	S3	S6	S1

Misclassification in one or more phases

No Misclassification

In the 8-9 second segment, the machine learning model classified the events at the buses from MV1 to MV5, LV1, LV3 and LV4 correctly as mentioned in Table IV. At LV2, the non-linear loads were starting to be energised. At that instant, phase-1 of LV2 started at the zero-crossing point, while phase-2 and -3 were at non-zero point. As a result, it deviated from the original steady-state waveform. Phase-1 of LV2 is correctly classified by the machine learning model as 'S4' (voltage with harmonics) while phase-2 and -3 as interruption with harmonics and sag with harmonics, respectively.

B. Test with Electric Arc Furnace Model

The electric arc furnace model by [10] was another circuit implemented in Simulink (Fig. 10) that was used to test the performance of the developed machine learning model in classifying the voltage fluctuation events. To test the conditions of flicker with sag, swell and harmonics, their mathematical logics as described in Table V were simulated as indicated in the block diagram of Fig. 11 and added to the output signal of the arc furnace in Fig. 10. It was found out that all the events of 'Voltage Flicker', 'Voltage Sag with Flickers', 'Voltage Swell with Flickers' and 'Voltage Flickers with Harmonics' are classified correctly for all 3 phases of voltage at every bus in the network.

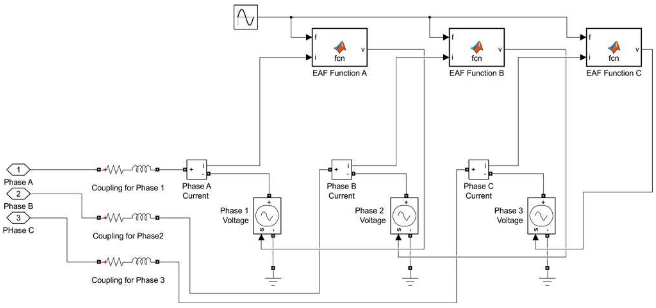

Fig. 10. Electric Arc Furnace Model [10]

TABLE V. ADDITIONAL DISTURBANCES DURING FLICKER CONDITION

Name	Mathematical Expression
Voltage Sag with Flicker	$\{1-\alpha_1(u(t-t1)-u(t-t2))\}\times V_{flicker}$ ... (21) $where\ \alpha_1 = sag\ magnitude$
Voltage Swell with Flicker	$\{1+\alpha_2(u(t-t1)-u(t-t2))\}\times V_{flicker}$ ... (22) $where\ \alpha_2 = swell\ magnitude$
Voltage Flicker with Harmonics	$\{h_1.\sin(\omega t)+h_3.\sin(3\omega t)+h_5.\sin(5\omega t)+$ $h_7.\sin(7\omega t)\}\times V_{flicker}$ ……..(23) $where\ h_1 = fundamental\ harmonic$ $h_3 = 3rd\ harmonic$ $h_5 = 5th\ harmonic$ $h_7 = 7th\ harmonic$

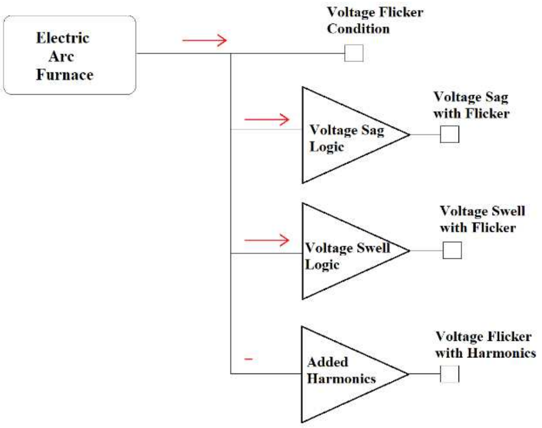

Fig. 11. Outputs of Electric Arc Furnace

VI. CONCLUSIONS

In this paper, 16 categories of most commonly occurring power quality events are classified by means of wavelet transform and machine learning based methods. The outcome of classifications and effectiveness of machine learning methods for the proposed application is evaluated to identify the best performing machine learning method. The ‘Classification Learners’ application in MATLAB is used for analysing the performance. The selected model namely Ensembled Bagged Trees was then implemented in Simulink for test distribution grid circuits. The results obtained from simulation demonstrated the efficiency of the method and showed acceptable accuracy and performance. Future work will look to implement an intelligent decision-making system for power network operators based on the power quality event classification model developed here.

REFERENCES

[1] C. S. Lam, Y. Yang, Q. Zhong, Y. Chen and K. W. Lao, “Advanced Power Electronic Converters and Power Quality Conditioning,” in Hindawi: Journal of Electrical and Computer Engineering, vol. 2018, pp. 1-2, November 2018, doi: https://doi.org/10.1155/2018/9701013

[2] A. Sharma, B. S. Rajpurohit and S. N. Singh, “A review on economics of power quality: Impact, assessment and mitigation,” in Elsevier: Renewable Sustainable Energy Reviews, vol. 88, pp. 1-5, 2018, doi: https://doi.org/10.1016/j.rser.2018.02.011

[3] T. S. Bomfim, "Evolution of Machine Learning in Smart Grids," *2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE)*, 2020, pp. 82-87, doi: 10.1109/SEGE49949.2020.9182023.

[4] "IEEE Recommended Practice for Monitoring Electric Power Quality," in *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)* , vol., no., pp.1-98, 13 Aug. 2019, doi: 10.1109/IEEESTD.2019.8796486.

[5] S. Dash and U. Subudhi, "Multiple Power Quality Event Detection and Classification using Wavelet Transform and Random Forest Classifier," *2018 International Conference on Applied Electromagnetics, Signal Processing and Communication (AESPC)*, 2018, pp. 1-5, doi: 10.1109/AESPC44649.2018.9033313.

[6] P. Kanirajan and V. S. Kumar, “Power quality disturbance detection and classification using wavelet and RBFNN” in Elsevier: Applied Soft Computing, vol. 35, pp. 470-481, July 2015, doi: http://dx.doi.org/10.1016/j.asoc.2015.05.048 1568-4946/

[7] L. Zhao and M. Gai, “Power Quality Disturbances Identification Method Based on Novel Hybrid Kernel Function” in *Journal of Information Processing Systems*, vol. 15, no. 2, pp. 422-432, April 2019, doi: 10.3745/JIPS.04.0112

[8] F. Hafiz, A. Swain, C. Naik, S. Abecrombie and A. Eaton, “Identitfication of power quality events: selection of optimum base wavelets and machine learning algorithm,” in *IET Science, Measurement & Technology*, vol. 13, no. 2, pp. 240-271, December 2018, doi: 10.1049/iet-smt.2018.5044

[9] F. Uçar, Ö. F. Alçin, B. Dandil and F. Ata, "Machine learning based power quality event classification using wavelet — Entropy and basic statistical features," *2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR)*, 2016, pp. 414-419, doi: 10.1109/MMAR.2016.7575171.

[10] R. H. G. Tan and V. K. Ramachandaramurthy, “A Comprehensive Modeling and Simulation of Power Quality Disturbances Using MATLAB/SIMULINK”, in *Power Quality Issues in Distributed Generation*, J. Luszcz, Eds. United Kingdom: IntechOpen, 2015, pp. 83-107.

[11] M. O. Oliveira, J. H. Reversat, and L. A. Reynoso, “Wavelet transform analysis to applications in electric power systems,” in *Wavelet Transform and Complexity*, D. Baleanu, Eds. United Kingdom: IntechOpen, 2019, pp. 1-7.

[12] L. Saribulut, A. Teke, M. B. Latran and M. Tumay, “Fundamentals and literature review of wavelet transform in power quality issues”, in *Journal of Electrical and Electronics Engineering Research*, vol. 5, no. 1, pp. 9-22, May 2013, doi: 10.5897/JEEER2013.0436

[13] B. Boehmke and B. M. Greenwell, Hand-on Machine Learning with R, United States: CRC Press, 2020.

[14] R. Amami, D. B. Ayed and N. Ellouze, “Practical Selection of SVM Supervised Parameters with Different Feature Representation for Vowel Recognition”, in *International Journal of Digital Content Technology and its Applications (JDCTA)*, vol. 7, no. 9, pp. 418-424, May 2013, doi:10.4156/jdcta.vol7.issue 9.

[15] H.H. Patel and P. Prajapati, “Study and Analysis of Decision Tree Based Classification Algorithms”, in *International Journal of Computer Science and Engineering,* vol. 6, no. 10, pp. 74-78, October 2018, doi: 10.26438/ijcse/v6i10.7478