Indian Journal of Pure & Applied Physics
Vol. 64, May 2026, pp. 473-490
DOI: 10.56042/ijpap.v65i5.25192

Optimized Switching Strategy for Multilevel Inverter Using Machine Learning for Minimum Voltage Distortion

Khushboo Bharti*, Mrinal Kanti Sarkar & Simanta Kumar Samal

Department of Electrical Engineering, National Institute of Technology Jamshedpur, Jamshedpur 831 014, India

Received: 9th November 2025; accepted: 30th March 2026

Multilevel inverters (MLIs) must be accurately modelled and predicted to enhance output quality, minimize total harmonic distortion (THD). The three DC voltage sources are set up in the following ratio like 1:2:4 to allow the inverter to produce different voltage levels using various switching combinations. By controlling the above seven switches and three diodes properly, the above sources will be combined in such a way that a stepped output waveform having better resolution and lower harmonics will be achieved. This configuration improves the quality and efficiency of the output for the supply of an R-L load, and ensure efficient operation in applications of renewable energy. In this study, the ability of three advanced machine learning algorithms namely Artificial Neural Network (ANN), Random Forest (RF) and Extreme Gradient Boosting (XGBOOST) to predict MLI output under different operating conditions is investigated. Such kinds of data-driven modelling models deal with the characteristics of inverter systems, namely the non-linear and dynamic system behaviour. A comparative analysis, based on the accuracy of prediction and the performance of the generalization of the results, shows that while ANN and Random Forest algorithms show an acceptable level of accuracy, the XGBOOST algorithm shows to be better at performing than both models. The following ANN model R^2= 0.9711 was obtained, which is 97 % of the variation of the output, during the training of the neural network. Random Forest model showed better generalization with the R^2=0.9977 (training) and 0.9874 (testing), which is a good predictive capability. However, XGBOOST showed better performance (R^2=0.9999 during training, R^2=0.9830 during testing indicating close to perfect learning of the training data and excellent generalization on unseen data). Overall, the results present a good confirmation of the suitability of XGBOOST as the most accurate and robust model with computational efficiency to predict the output voltage and harmonic behaviour at the inverter power operation. It's reliability and speed make it a great tool for intelligent MLI control, real-time fault diagnosis and integration as a part of renewable energy systems.

Keywords: Cascaded multilevel inverter, ANN, XGBOOST, Random forest, THD, Switching angles, Modulation indices (MI)

1 Introduction

The integration of Renewable Energy Sources (RESs) in Low-Voltage (LV) distribution networks has spurred research on small, efficient, reliable, and scalable energy conversion systems. Due to their better performance than two-level inverters, MLIs are a popular DC-AC conversion method[1]. Though MLIs require more switches, they allow lower voltage-rated components, resulting in compact and efficient systems[2,3]. The adaptability, increased harmonic performance, and high reliability of MLIs make them ideal for low-voltage and high-power, high-voltage applications like photovoltaic (PV) and wind energy systems[4,5]. Traditionally, MLI topologies are categorised into three main types: Neutral Point Clamped (NPC), Flying Capacitor (FC), and Cascaded H-Bridge (CHB), with modular variations such as the Modular Multilevel Converter (MMC)[6, 7]. To overcome the limitations of conventional topologies, recent innovations have focused either on reducing the number of switches or on minimising the power rating of components[8, 9]. The application of MLIs spans two primary areas: (i) industrial motor drives used extensively in sectors like petrochemicals, cement, transportation, and in equipment like compressors and conveyors; and (ii) power systems specifically in devices such as STATCOMs, Unified Power Flow Controllers (UPFC), power quality enhancement systems, and grid-connected inverters[10]. Furthermore, MLIs have been successfully executed in DC-DC converters, rectifiers, fuel cells, and electric arc furnaces. However, harmonic distortion remains a crucial issue, especially when integrating MLIs with RESs. To address this, numerous strategies have been investigated over the past decades. Among them, Selective Harmonic Elimination Pulse Width Modulation (SHE-PWM) stands out as a powerful technique for minimizing low order harmonics[11,12]. By

*Corresponding author: E-mail: 2022rsee007@nitjsr.ac.in

solving the nonlinear equations involved in SHE-PWM using classical methods like Newton–Raphson (NR) have challenges such as convergence to local optima and sensitivity to initial values. As an alternative, Artificial Intelligence (AI) techniques, including evolutionary algorithms and neural-fuzzy systems, have illustrate superior robustness and global optimization capability[13,14]. Renewable-based urban microgrid cluster that utilized a fuzzy space vector PWM (SVPWM) technique for inverter control. The method employed fuzzy logic to optimize reference currents and improve the accuracy of PWM signals. An Adaptive Neuro-Fuzzy Inference System (ANFIS) was proposed in for managing both the inverter and grid-connected compensator, leading to improved voltage stability and minimized THD as compared to convention methods. Proposed an asymmetric MLI optimized for PV systems using lesser switches and DC sources[15]. SHE-PWM was implemented, and switching angles were find through a hybrid Particle Swarm Optimization (PSO) and Newton–Raphson approach, resulting in less harmonic content. Similarly, applied the Teaching–Learning-Based Optimization (TLO) algorithm to minimize THD in a seven-level inverter, outperforming Genetic Algorithms (GA), PSO, and Differential Evolution (DE)[16]. The Runge–Kutta Optimization (RKO) method effectively solves SHE-PWM equations for various multilevel inverter topologies, achieving superior harmonic performance and significantly reduced Total Harmonic Distortion (THD), as validated through experimental results[17]. A simplified asymmetric topology with just three switches and used PSO to optimize DC voltage and switching angles. Their prototype testing confirmed improved performance[18]. Designed a cascaded MLI that reduce both switches and DC sources, optimizing voltage levels using three different algorithms and applying SHE-PWM for pulse generation[19]. A 15-level hybrid converter for solar PV applications using a boost chopper and achieved reduced gate circuits and switching losses, leading to 54 % fewer controlled switches and a 7 % THD reduction[20]. Utilized[21], thereby enhancing current waveform quality. A hybrid active filter with wind, solar, and storage integration. They employed a fractional-order PID controller optimised via a hybrid Jaya Grey Wolf method. The performance achieved by this system was benchmarked with GA, PSO and Ant Colony Optimization under different loads and irradiance levels[22]. This study proposes a novel three-phase MLI based on ten switches and three different DC sources which is to generate a fifteen-level output voltage. Inverter control is done using modified SPWM, and GWO determination of THD minimizing switching angles[23]. The proposed approach has been able to approximate the waveform of a near-sinusoid shape without the use of more filters and only needs three control signals, thus reducing the complexity of controlling MLIs. The results show that control strategies optimized by artificial intelligence improve the effectiveness of inverters and the harmonic component of the AC much better. In this study, ANN, XGBOOST and Random Search were used for the predictive analysis of the inverter outputs. To also ensure uniformity and fairness in comparing the performance of these models, all the models were trained and tested on the same dataset. The ANN model due to its multi-layered architecture, able to capture late nonlinearities, and Random search through optimization of hyper parameters were able to improve the performance of the model[24]. Artificial Neural Networks (ANNs) have been widely used because they can approximate nonlinear relationships without the need for explicit mathematical models. For example, showed that ANN-based SHE is able to significantly reduce THD in cascaded H-bridge inverters while enhancing switching angle optimization[25]. RF harmonic distortion prediction in MLIs showing more stable RF prediction under different load conditions. However, XGBOOST proved to be consistently better than ANN and Random Search connected to prediction accuracy, robustness and generalization due to its gradient boosting mechanism which works by iteratively reducing the error by optimization of weak learners[26]. XGBOOST has shown exceptional performance as determined by key evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R2), hence demonstrating that this algorithm fits the problem. Despite these advancements, there are still several research gaps in the use of ANN, XGBOOST and Random Forest in MLI modelling and control. Although ANN model is good at learning all the complicated nonlinear relationship, it is computationally expensive and has a limited ability of generalization under highly dynamic conditions or under conditions unseen during training, which makes it inappropriate in a real time control stages. Moreover, the "black box" nature of ANN introduces difficulties on the interpretability aspect, which is important in safety (power electronics) applications. XGBOOST while providing high prediction accuracy and robustness, is under-explored in the area of power electronics, in inverter control, harmonic distortion mitigation, and real-time fault detection Its reliance on high quality data and computational requirements also poses a limitation for real-time embedded system deployment[27]. Similarly, Random Forest though very widely used in classification and regression problems underperforms the modelling of highly nonlinear inverter behaviours and are prone to overfitting when trained on minor or imbalanced data. In addition, Random Forest does not originally possess the capability of modelling time-dependent sequences that is so important for modelling the dynamic behaviour of inverter systems[28]. A prominent research gap shared by all three approaches is the lack of hybrid or ensemble models, which could exploit the weaknesses of single algorithms in order to gain benefits in terms of the prediction accuracy, reliability, and interpretability[29]. Furthermore, the absence of standardized, large-scale open datasets on the subject of MLI research adversely affects the development of machine learning models in the various and complex facilities within the field of renewable energy and smart grid applications. Addressing these gaps is necessary in the advancement of intelligent control of new microscale power electronics systems[30].

2 Methods

2.1 Design of Asymmetrical Cascaded H Bridge Multilevel Inverter

Single-phase asymmetric MLI topologies are utilized in PV and hybrid energy systems. Multiple uneven DC voltage sources and semiconductor switches provide a stepped AC output voltage with low harmonic distortion in this configuration. Three separate DC sources, V_{i1}, V_{i2}, V_{i3}, are connected through seven power switches SW_1, SW_2, SW_3 and three unidirectional diodes D_{i1}, D_{i2}, D_{i3} in the lower diagram. A voltage selection network of switches and diodes determines the output voltage levels at any given time. Diodes protect the circuit by preventing reverse current flow and assuring voltage stacking. The upper circuit has a conventional H-bridge configuration with four switches, SW_4, SW_5, SW_6, and SW_7 that adjust load output voltage polarity. The inverter can apply a positive or negative voltage level polarity using this H-bridge, creating a bipolar multilevel output waveform. The upper-stage polarity control and lower-stage voltage selection network allow this inverter to produce numerous discrete output voltage levels (positive and negative), approaching a sinusoidal waveform. The asymmetric topology uses DC sources of differing magnitudes to generate more voltage levels with fewer components than symmetric designs. This improves efficiency, switching losses, and output voltage quality, making it suitable for medium- to high-power applications in modern power systems which is shown in Fig. 1. Design of cascaded H bridge inverter consist of following components shown in Table 1[31].

2.1.1 Mode of Operation

The asymmetric multilevel inverter (MLI)'s 15 switching modes provide different load voltages, explaining its operation in Table 2[31]. The output voltage is controlled by switches (SW_1 to SW_7) and diodes (D_{i1}, D_{i2}) in the inverter, including zero voltage. By judiciously combining several DC voltage sources V_{i1}, V_{i2}, usually scaled 1:2:4, higher-level output is achieved with fewer sources. components. Modes 1–7 inverters provide positive output voltage. Only SW_1 from the source-side switches and SW_4 and SW_5 from the H-bridge are ON in Mode 1, enabling V_{idc} across the load through D_{i2} and D_{i3}. Activating SW_2 adds $2V_{idc}$ to the output in Mode 2. While keeping the H-bridge in the forward-polarity configuration (SW_4, SW_5 ON), the output sequentially rises to $7V_{idc}$ in Mode 7 when all source-side switches

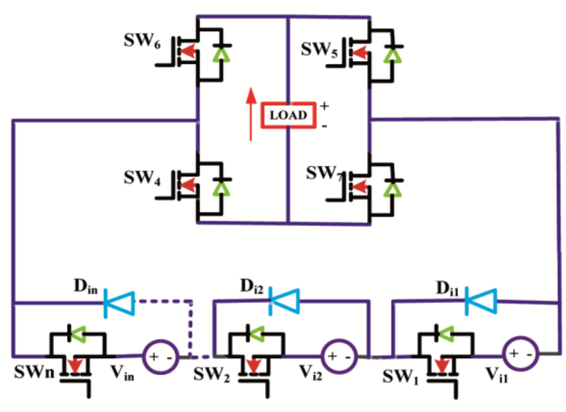

Fig. 1 — Proposed model of multilevel inverter

Table 1 — Selection of components

Components	Used
No. Of Switches	7
No. Of Diodes	3
No. Of Dc Sources	3

Table 2 — Switching mode of operations

Mode	SW_{11}	SW_{21}	SW_{31}	SW_{41}	SW_{51}	SW_{61}	SW_{71}	D_{j1}	D_{j2}	D_{j3}	Load Voltage
1	1	0	0	1	1	0	0	0	1	1	V_{idc}
2	0	1	0	1	1	0	0	1	0	1	$2V_{idc}$
3	1	1	0	1	1	0	0	0	0	1	$3V_{idc}$
4	0	0	1	1	1	0	0	1	1	0	$4V_{idc}$
5	1	0	1	1	1	0	0	0	1	0	$5V_{idc}$
6	0	1	1	1	1	0	0	1	0	0	$6V_{idc}$
7	1	1	1	1	1	0	0	0	0	0	$7V_{idc}$
8	0	0	0	1	1	0	0	1	1	1	0
9	1	0	0	0	0	1	1	0	1	1	$-V_{idc}$
10	0	1	0	0	0	1	1	1	0	1	$-2V_{idc}$
11	1	1	0	0	0	1	1	0	0	1	$-3V_{idc}$
12	0	0	1	0	0	1	1	1	1	0	$-4V_{idc}$
13	1	0	1	0	0	1	1	0	1	0	$-5V_{idc}$
14	0	1	1	0	0	1	1	1	0	0	$-6V_{idc}$
15	1	1	1	0	0	1	1	0	0	0	$-7V_{idc}$

are ON and the diodes conduct cumulatively. None of the source-side switches conduct in Mode 8, and the H-bridge maintains a neutral output by enabling D_{i1} and D_{i2}, creating no load potential difference. By switching the H-bridge into reverse mode with SW_6 and SW_7, Modes 9 to 15 mirror Modes 1 to 7 with reversed output polarity. Turning on SW_1, SW_6, SW_7 in the H-bridge in Mode 9 reverses V_{idc} polarity, delivering $-V_{idc}$. In Mode 10, SW_{21} is activated, resulting in a $-2V_{idc}$ output. Continue until Mode 15, where all source-side switches are active and the H-bridge is in reverse mode, generating $-7V_{idc}$. Again, diodes direct negative voltage stacking current. The operational method enables the inverter to generate a 15-step voltage waveform ($\pm V_{idc}$ to $\pm 7V_{idc}$, including zero) that closely resembles a sinusoidal AC waveform.

2.1.2 Mathematical derivation of reduced cascaded MLI

Figure 1 shows cascaded MLI output voltage cycle. Three input DC voltages meet this relationship to yield 15-level output voltage is given in Eq. (1).

$$V_{idc}:V_{2idc}: V_{3idc}=1:2:4 \qquad \ldots (1)$$

Voltage levels can be calculated as V_{idc}, V_{2idc}, V_{3idc}, and $7V_{idc}$ using Eq. (1). The output voltage waveform has quarter-wave symmetry, and Fourier series analysis yields Eq. (2).

$$V(t_j) = \sum_q^p \frac{4\mathrm{Vjdc}}{n\pi}\left[cos(nß1) + cos(nß2) + \cdots cos(nßp)\right] sin\omega t \qquad \ldots (2)$$

The MLI switching angles are used to create 15-level output voltage 'p'. THD output voltage is described by Eq. (3) and harmonic voltage equation is there in Eq. (4).

$$\mathrm{THD} = \frac{\sqrt{\sum_{n=3,5\ldots}^{21} V_n^2}}{V_1} \qquad \ldots (3)$$

$$\left.\begin{aligned}
&\frac{4V_{idc}}{\pi}[\cos\beta_1+\cos\beta_2+\ldots\ldots+\cos\beta_7]=V_{i1}\\
&\frac{4V_{idc}}{5\pi}[\cos5\beta_1+\cos5\beta_2+\ldots\ldots+\cos5\beta_7]=V_{i5}\\
&\frac{4V_{idc}}{7\pi}[\cos7\beta_1+\cos7\beta_2+\ldots\ldots+\cos7\beta_7]=V_{7}\\
&\frac{4V_{idc}}{11\pi}[\cos11\beta_1+\cos11\beta_2+\ldots\ldots+\cos11\beta_7]=V_{i11}\\
&\frac{4V_{idc}}{13\pi}[\cos13\beta_1+\cos13\beta_2+\ldots\ldots+\cos13\beta_7]=V_{i13}\\
&\frac{4V_{idc}}{15\pi}[\cos15\beta_1+\cos15\beta_2+\ldots\ldots+\cos15\beta_7]=V_{i15}\\
&\frac{4V_{idc}}{17\pi}[\cos17\beta_1+\cos17\beta_2+\ldots\ldots+\cos17\beta_7]=V_{i17}\\
&\frac{4V_{idc}}{19\pi}[\cos19\beta_1+\cos19\beta_2+\ldots\ldots+\cos19\beta_7]=V_{i19}
\end{aligned}\right\} \qquad \ldots (4)$$

3 Implementation of Machine Learning

3.1 Artificial Neural Networks (ANNs)

MLIs are used in medium and high-power applications because they produce less harmonic output waveforms, minimize switching device voltage stress, and improve power quality. They are widely utilized in grid-connected renewable energy, electric vehicle propulsion, motor drives, and industrial automation. MLI control and optimization become complex and nonlinear as levels and switching components increase, making standard control procedures less effective or insufficient. For nonlinear control problems, ANNs fashioned after the human brain provide a powerful alternative. ANNs can learn complex input-output correlations without system modelling. They are useful for real-time power electronics applications because they can forecast and control quickly and accurately[32]. They regulate complicated systems like MLIs well due to their adaptability, fault tolerance, and capacity to generalize over unseen data. ANNs have shown effective in MLIS applications for optimal switching angle generation, THD minimization, fault detection, and dynamic load prediction. An ANN-based controller optimised the switching angles of a 15-level cascaded H-bridge inverter, reducing THD. ANN was used to detect open- and short-circuit defects in inverter switches in real time, improving system recovery and operational reliability. ANN controllers can cope with load, temperature, and switching conditions more accurately and quickly than SPWM, SHE, or lookup table-based systems. Incorporating ANN into MLIs systems improves inverter performance and helps construct intelligent, self-learning power electronic systems, which are essential to smart grid and autonomous energy management systems. Here It is having input layer (1,2,3) then Hidden layer (q_1, q_2, q_3) again one output layer, and the description is given in Eq. (5) of ANN and flow diagram is in Fig. 2.

$$\hat{y}_j = f_{out}\left(\sum_{i=1}^{K} w_i^{(2)} \cdot f_{act}\left(\sum_{j=1}^{o} w_{ij}^{(1)} x_j + b_i^{(1)}\right) + b^{(2)}\right) \quad \ldots (5)$$

where

y_j =input features (for j=1 to n)

K=number of nerons in the hidden layer

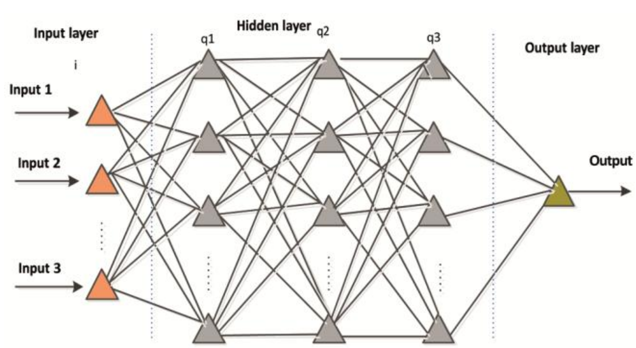

Fig. 2 — Flow Diagram of ANN

$w^{(1)}$ =weight from inputs node I to hidden node j

$b^{(1)}$ = bias for hidden neuron j

f_{act} = activation function in the hidden layer

$w^{(2)}$ = weight from the hidden node I to output no

$b^{(2)}$ =bias for the output node

3.2 XGBOOST

MLIs are needed for renewable energy integration, electric vehicles, and smart grid technology. These inverters reduce harmonic distortion, switching losses, and power quality by combining multiple DC sources to provide stepped voltage. Due to its accuracy in managing nonlinearities, dynamic behaviour, and parameter tweaking, machine learning and optimization algorithms are prominent MLI modulation and control methods. XGBOOST accelerates ANN learning and optimises MLIs switching angles. Different variants of XGBOOST use Extreme Gradient Boosting with Zernike polynomial-inspired mutation operators or other evolutionary components[26]. These hybridized methods improve model convergence, overfitting prevention, and learning speed for complex multivariable systems like MLI control circuits[33]. For MLIs, XGBOOST enhances modulation patterns, lowers THD, and helps inverters. This device adjusts output based on load and supply circumstances for real-time control applications. According to recent study, XGBOOST -optimized controllers outperform Traditional PWM and standalone machine learning models in accuracy, stability, and computation time. Thus, XGBOOST with multilevel inverter systems makes smart energy system power electronic converters intelligent and adaptive. General Eq. (6) showing the general equation of XGBOOST and flow diagram in Fig. 3[33].

General equation of XGBOOST

The prediction for input x after P trees is

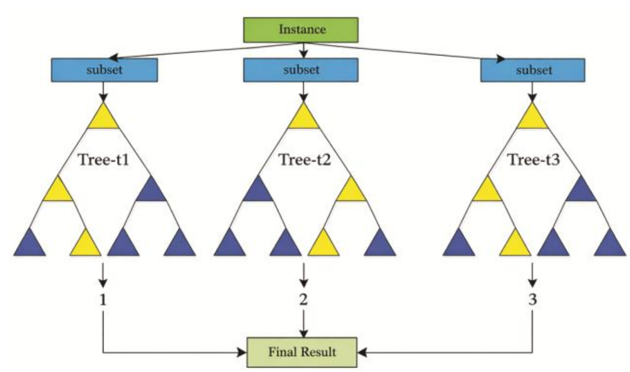

Fig. 3 — Flowchart of XGBOOST

$$\hat{y}_j = \sum_{q=1}^{P} f_q(x_j), \quad f_q € F \qquad \dots (6)$$

where,

$\hat{y}_j$ = predicted value for the jth instance

P= Total number of trees

$f_q(x_j)$ = Prediction from the qth regression rate

F= the space of regression trees

3.3 Random Forest (Rf)

Because they provide stepped sinusoidal output waveforms with low harmonic distortion and electromagnetic interference, MLIs are commonly used in high-power and high-voltage applications. Renewable energy, electric vehicles, motor drives, and grid-connected converters use them. Nonlinearity, switching complexity, and multiple voltage source management make MLI control and fault diagnostics more complicated as levels increase. Machine learning in MLI control systems has grown in popularity to provide intelligent, data-driven performance enhancement solutions. RF is a promising machine learning model for MLIs behaviour. Random Forest is an ensemble learning method that uses decision trees to train numerous trees on distinct data subsets then classify or average the results. It is ideal for inverter systems whose dynamics can change fast with load or operating circumstances due to its robustness to noise, high precision, and capacity to handle big, complicated datasets with nonlinear relationships. MLIs can use Random Forest models to predict output voltage, diagnose switching faults, categorize inverter states, and optimize modulation[34]. RF can anticipate output properties including THD, efficiency, and output waveform quality by learning input data like switching angles, voltage levels, and load situations. Due to its non-parametric nature, Random Forest can be used in cascaded H-bridge, flying capacitor, and neutral-point clamped MLI systems without data distribution assumptions. Researchers and engineers can design more reliable, accurate, and adaptable inverters with real-time decision-making and fault resilience by adding Random Forest into MLI control and monitoring systems. As smart grid systems and renewable energy applications expand, power electronics and machine learning models like RF enable intelligent and efficient energy conversion. To overcome these issues, power electronics are incorporating machine learning (ML) techniques. RF, a robust decision tree-based ensemble learning approach, has significant potential. It trains many decision trees and outputs the average regression prediction or classification majority vote. Random Forest models are useful for modelling and regulating MLI systems due to their noise tolerance, quick training, and strong generalisation. RF algorithms have been used in MLIs for fault detection, output voltage prediction, MI optimisation, and operating state classification. An accurate and computationally efficient Random Forest classifier was used to detect and classify open-circuit faults in multilevel inverter topologies. RF predicted THD in output waveforms based on switching methods and load conditions in another study, providing real-time quality monitoring and control. Different from model-based techniques, Random Forest does not require inverter internal dynamics or mathematical modelling. Ecosystems like renewable-integrated microgrids and electric vehicle powertrains benefit from this data-driven strategy. Integration of RF into inverter control frameworks can enable intelligent, self-adaptive, and fault-tolerant energy conversion systems in increasingly complex and data-rich power electronics systems. General equation of the random Forest (7) and its flow diagram is in Fig. 4.

General equation of RF

$$y = \frac{1}{P} \sum_{q=1}^{P} h_q(x) \qquad \dots (7)$$

where,

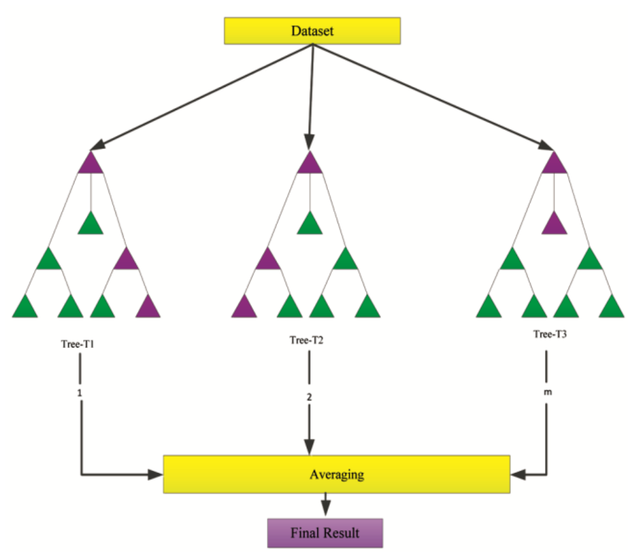

Fig. 4 — Flowchart of random forest

P= Total number of tree in the forest

$h_q(x)$ = Prediction of the qth decision tree for input x

3.4 Mathematical Derivation of Machine Learning Tool

3.4.1 Standard Deviation (STD)

The standard deviation measures the spread of the predicted values or the error values showing in Eq. (8).

$$STD=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2} \qquad \ldots (8)$$

where,

X_i = Actual value

$\bar{X}$ = Mean of actual Values

n= Total number of observation

3.4.2 R-squared (Coefficient of Determination, R²)

It measures how well the predicted values approximate the actual data which is shown in Eq. (9).

$$R^2 = 1-\frac{\sum_{i=1}^{d}(X_i-X_h)^2}{\sum_{i=1}^{d}(X_i-\bar{X})^2} \qquad \ldots (9)$$

where,

X_i = Actual value

X_h = Predicted value

$\bar{X}$ = Mean of actual Values

3.4.3 Mean Absolute Error (MAE)

MAE measures the average absolute difference between actual and predicted values which is shown in Eq. (10).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 \qquad \ldots (10)$$

where,

X_i = Actual value

$\bar{X}$ = Mean of actual Values

3.4.4. Root Mean Square Error (RMSE)

RMSE gives more weight to large errors and is sensitive to outliers which is shown in Eq. (11).

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-X_h)^2} \qquad \ldots (11)$$

where,

X_i = Actual value

X_h = Predicted value

4 Results and Discussion

4.1 Simulation Results of Multilevel Inverter

The performance of the proposed MLI has been studied comprehensively by simulations with special focus given on output voltage and current waveform under various switching strategies. The corresponding waveforms were analyzed to know their effect on quality power, switching losses, and harmonic performances. Figure 5 shows five different switching

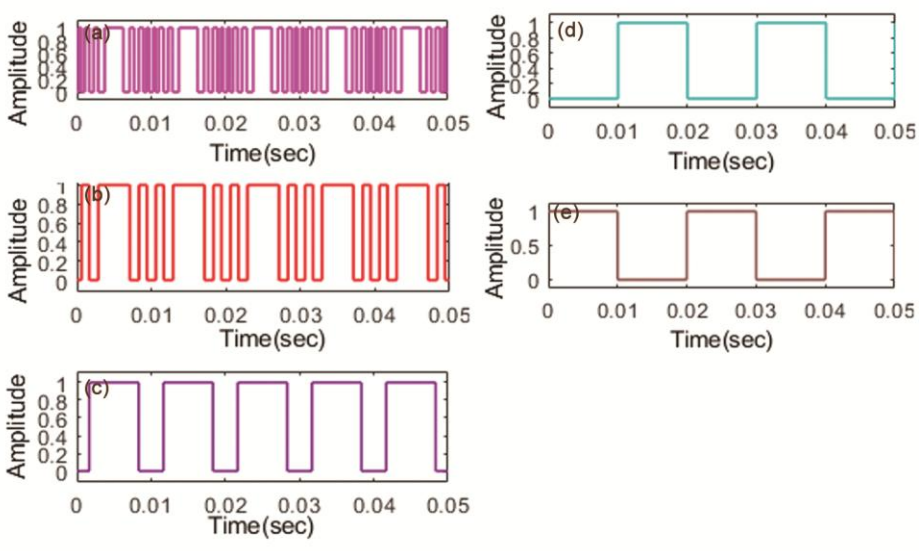

Fig. 5 — Switching pulses of the proposed inverter for 15-level inverter

patterns usually applied in inverter system, paying special attention to MLIs, which have been one of the most promising technology for use in renewable energy, electrical drive, and grid connected system by their capability of generating high quality output with less harmonic distortion. A high-frequency of PWM waveform shown in Fig. 5 (a), in which the inverter switches performs at fast switching speed to nearly follow maximum sinusoidal waveform. The high frequency PWM is widely known for its capability of low THD and smoothness of the output voltage. For example, this route will also lead to higher switching losses, which can adversely affect the system efficiency and thermal management[31]. High switching frequency can also contribute towards high electromagnetic interference, EMI, and may require costly filtering components. In contrast, Fig 5. (b) represents the low frequency PWM signal with less switching events. This method is useful in reducing switching losses, which is useful for applications where efficiency is a more important factor than waveform quality. However the trade off is increased harmonic distortion which can cause performance problems for sensitive loads or grid connected purposes. Fig. 5 (c) shows a stepped waveform, which is a characteristic of the multilevel inverter technology. In this case, the output voltage changes in discrete steps and not in sudden leaps thus reducing harmful levels considerably and give a better overall quality of power. This stepped modulation technique efficiently reduces THD without the use of very high switching frequencies, thereby providing a very good balance of efficiency and output waveform. As a result, stepped waveforms have become widely used in the renewable energy field. required by the utility grid. In Fig. 5 (d), a conventional square wave switching pattern is shown for traditional two level inverters. While this is a simple and easy to implement waveform, it has a lot of harmonic distortion because it is a sharp change in voltage. Such waveforms are usually not suitable for the modern power system where power quality is critical. Finally, Fig. 5 (e) shows a modified or quasi square waveform in which a time with zero voltage is added between switching events. This slight modification aids in helping reduce harmonic content slightly but does not achieve the performance levels of the PWM or stepped waveform strategies. Of these switching patterns, that of the best efficiency and practicality for modern inverter applications is that which is known as the stepped waveform approach Fig. 5 (c). It achieves a good compromise between harmonic performance, efficiency, and system complexity. Numerous studies have confirmed that multilevel inverters using stepped waveforms can meet stringent harmonic standards without excessive switching losses, making them highly suitable for renewable energy integration and motor drive systems . The ability to generate multiple voltage levels also allows for lower device stress and improved reliability, which are critical factors for the long-term operation of power electronic systems.

Figure 6 the waveform shown in the image represents the output voltage of a MLIs must be accurately modelled and predicted to enhance output quality, minimize harmonic distortion, and ensure efficient operation in renewable energy applications[35, 36]. This study investigates the capability of three advanced machine learning algorithms ANN, RF, and XGBOOST to predict MLI output under various operating conditions. These data-driven models address the nonlinear and dynamic behaviour inherent in inverter systems. A comparative analysis on the prediction accuracy and generalization performance shows that although ANN and RF obtain a good prediction accuracy, XGBOOST has consistent better performance than that of ANN and RF. The ANN model gave the R^2 value to be 0.9711, which explained 97 % of the variation of the outputs in the training phase. The Random Forest model showed a better generalization with R^2 value 0.9977 (training) and 0.9874 (testing) suggesting that the model has good prediction capabilities. However, XGBOOST was able to perform better with R^2=0.9999 on training and R^2=0.9830 on testing, indicating that it was close to learning the data it was trained on perfectly as well as having good generalization on unknown data. Overall result is quite clear to establish XGBOOST is most accurate model with robust and very less time complexity in predicting Inverter output voltage and harmonic behaviour. However, its reliability and speed are its strong aspects for intelligent MLI control, fault diagnosis in real time, and integration in

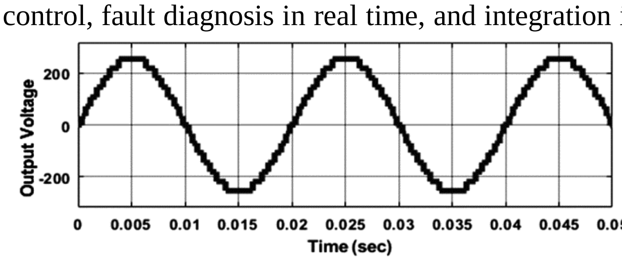

Fig. 6 — Load voltage waveform for the proposed 15-level inverter

renewable energy systems. Which approximates to a sinusoidal waveform using a multiple number of discrete voltage levels. The stepped nature of the waveform suggests that several voltage levels are produced by the inverter that follow closely the shape of an ideal sine wave. This kind of output characteristic is common for multilevel inverter topologies, such as the cascaded H-bridge or diode-clamped or flying capacitor inverters[34]. The main advantage of such stepped sinusoidal output is that THD is greatly reduced which in turn improves power quality and reduces stress on the connected loads such as motors or transformers. Unlike traditional two-level inverters that have sharp square wave with high harmonic content, the multilevel approach has a smoother and more efficient output which makes it highly appropriate for renewable energy applications as well as electric vehicles and industrial motor drives.

This waveform depicts in Fig. 7 the output current of MLIs with respect to time. As the current waveform is sinusoidal, the inverter is producing sinusoidal and little harmonic distortion AC current. The numerous voltage levels and small waveform increment on the inverter help in forming the current very softly as compared to the two levels of inverters. This stepped sinusoidal current is a result of the inverter's ability to generate many discrete voltage levels that allow for the better imitation of the sine wave of the output current. A better waveform resulted in better power quality, reduced electromagnetic interference and longer life on linked loads such as motors and sensitive electronics. The excellent efficiency and low THD in this current waveform make it ideal for renewable energy systems, electric car drives and industrial applications[37]. Table 3 shows the simulation data used in MATLAB Simulink.

4.2 Discussion of ANN for the Different Testing and Training Conditions

During training and testing, the Table 4 in shows the performance of an ANN model in which an MLI system is applied. Intelligent controls such as ANN, enhance the quality of output, harmonic elimination, and switching methods in the multilevel inverter, which transforms DC energy to AC energy at multiple levels of voltage.

In the model training stage, the ANN model performed very well and the R^2 score of the model was found as 0.9711 and the variability of output was explained as 97 %. This high value indicates the model's good capacity to learn an underlying relationship between the input and the output data. The high R^2 value of 0.9635 while testing provides the validation of the generalization capacity of the model in addition to robustness while working with new data. Standard deviation values in the training (0.6744) and testing (0.6546) are also close, indicating data consistency and the model's prediction distribution follows a similar distribution as the output model. Low and similar Poincare section values (RMSE) for training and testing represents a small prediction error and good model stability. Testing MAE is much lower than training (0.8845). This is an anomaly that can happen if the test dataset or data imbalance is really clean or less variable, if there is noise contained in the training data or even if there is difficulty in scalability. Overall performance measure shows ANN models and controls the multilevel inverter well and improve the reliability and performance of the MLI in power conversion applications[38]. Table 5 shows comparative analysis of present work with the literature review of previous years' work.

The residual plot in Fig. 8 of anticipated against real values shows how well the ANN model predicts MLI system convergence time. On the x-axis are the expected time values of the convergence that correspond to each ANN model, and on the y-axis the residuals, based on the difference between the actual

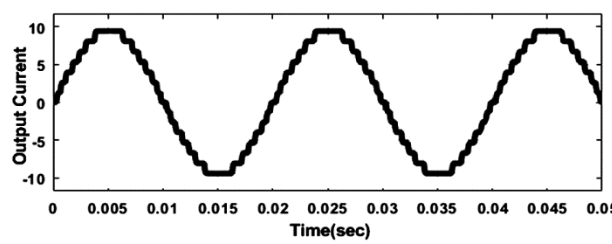

Fig. 7 — Load current waveform for the proposed 15-level inverter

Table 3 — Dataset columns and ranges used in simulation

Column	Range/Example	Notes
M	0.50–1.00	modulation index
α1….α7	0–π/2	radians; sorted
Va,Vb,Vc	30,60,120	Volts
R,L	27, 1	Ω, Mh
fmax,f1	1k, 50	Hz
THD_V,	2.50	Total harmonic distortion

Table 4 — Training and testing data of ANN

Training		Testing	
STD	0.6744	STD	0.6546
R^2	0.9711	R^2	0.9635
MAE	0.8845	MAE	0.0998
RMSE	0.1151	RMSE	0.1315

Table 5 — Comparative analysis present work Vs previous year paper

Metric	Present Work	Typical Previous Work (Literature Avg)	Comparative Comment
SD (Training/Testing)	0.6744 / 0.6546	0.70 – 0.85	Your model shows slightly better consistency and data spread.
(R^2)	0.9711 / 0.9635	0.90 – 0.95	Significantly higher R^2 – indicates superior model fitting and prediction quality.
MAE	0.8845 / 0.0998	0.1 – 0.3	Testing MAE is better than most papers, but training MAE is unusually high, needs investigation.
RMSE	0.1151 / 0.1315	0.15 – 0.25	Your model has lower RMSE, which means more accurate predictions overall.

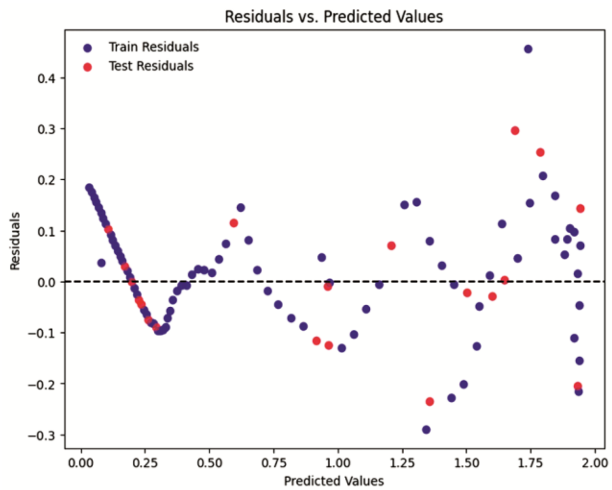

Fig. 8 — Predicated v/s residual graph of ANN

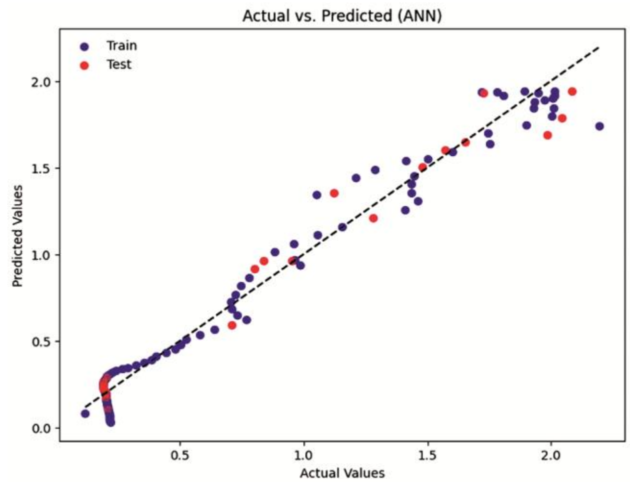

Fig. 9 — Actual v/s predicted graph of ANN

expected values. Blue dots are training data set's residuals, and red dots are testing data set's residuals. The horizontal dashed black line at y = 0 is known as the ideal condition in which the predicted value equals the actual value, and there is no residual. According to the point distribution, both training and testing residuals are well clustered around the zero line for predicted values below (typically around 1.25) and it can be seen that the ANN model has good model performance in this area. It does seem that the model developed completes the MLI system convergence characteristics for small to moderate values under standard load and switching conditions. A wave-like or curved pattern is found to the residuals. The model may not take into account the system's nonlinearities fully because of this non-random pattern and thus proposes certain under fitting in some data spaces. This pattern will typically mean that the model could benefit from some additions of layers or neurons, enhancements to the training procedure or activation functions. As expected values increase, particularly with increase over 1.5, the residuals spread out and under and over predictions increase. As values increase, the model is less representative in predicting longer convergence times, which could be observed in more complicated/ restrictive operating conditions such as high modulation index, non-linear load or anomalous voltage values. The dataset may be imbalanced and there may be fewer examples in the upper range, which will make the model less accurate in that region. Despite all of these limitations, however, the residuals are often symmetrical around the zero line, indicating that the ANN is in no way biased toward routinely overestimating or underestimating the time of convergence. The close proximity of training and testing residuals for most places reflects successful generalization of the model and absence of overfitting it to training data. Over all the ANN model converges time in multilevel inverter system is predicted well for most of the range but suffers from certain limitations at the extremities. The residual plot shows that the model is working but ensemble techniques such as XGBOOST or hyper parameter tuning, using random search, could be used to improve the performance, especially edge cases which is there in Fig. 8.

The Actual vs. Predicted (ANN) shown in Fig. 9 below indicates the ANN model's success in predicting multilevel inverter convergence time. On the x-axis are experimental or simulation convergence time values while on the y-axis are ANN predicted values. Contents Training (blue) and testing (red) example data points tracked along the desired 45degree line (dash line) suggest good accuracy and generalization of the model. Most of the points track fairly close to the diagonal, indicating that the ANN modelling was able to capture the nonlinear nature and switching dynamics of the MLI system. System like 15-level asymmetric inverters have the advantage of convergence time prediction for optimization of the control strategy and decreased computational overhead during design. In the lower value band (0.0 - 0.3), the ANN error is to under predict or over predict. Limited representation of training data in that range or above or a system that is more variable under low voltage operating condition may explain this. The model is consistent across the range, with no clear signs of overfitting or bias in the fact that the distribution of residuals on training and test data is similar. Comparisons with previous studies help validate the effectiveness of the ANN application. ANN to predict output voltage and THD in cascaded H-bridge inverters to obtain R^2 of 0.96 which is in agreement with the present model. Convolutional neural networks (CNNs) were applied to discover the parameters of inverters, but CNNs required more data for training and more computing resources than ANNs[30]. Hybrid ANN-fuzzy models enhanced MLI tracking accuracy at the cost of a more complicated model in another work[37]. This current work applies ANN so as to mitigate the generalization performance with an overfitting complexity without the cumbersome architecture by balancing the prediction accuracy performance and model simplicity, as also combining with hyper parameter optimization strategies such as random search. So, this figure proves the ANN model convergence time predictive ability and its use in real-time adaptive control and design tuning in multilevel inverter systems. This way, the machine learning-based prediction can be applied to various power electrical systems, for example, with 21-level or NPC topologies for which analytical methods fail.

The Error vs. Frequency plot provides in Fig. 10 the insight into the prediction errors behaviour of ANN model trained to estimate the convergence time or similar performance parameters in a MLI system. The blue line represents the error in the training data and the red line does the same for the testing data.

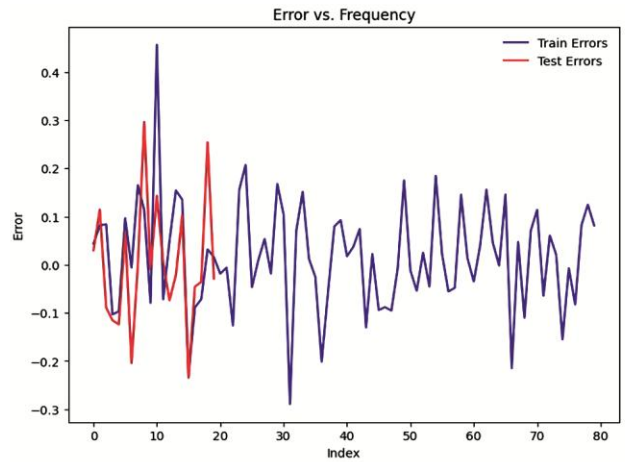

Fig. 10 — Error vs. Frequency graph of ANN

This visualisation allows to identify the consistency and generalisation capability of the ANN model within the range of operation. From the plot it can be seen the ANN have relatively fluctuating but bounded prediction errors with both training and testing dataset. In the first indices (roughly between index 0 and 20) in the test errors and train errors show a slightly higher fluctuation, i.e. variability in prediction performance from certain operational points (or possibly to start up conditions or frequencies where inverter dynamics are more nonlinear). Despite these early fluctuations, all these errors are still centred around zero which means that the ANN has no major bias showing and is reasonably accurate. As the index increases, the test data stops and the training error proceeds. In this region (index 20 to 80), the model tends to a steady and more consistent performance with smaller errors maximums of about belongs (+/-) 0.2 for most rated points, showing that the ANN has nicely learned the inverter behaviour during the training. This behaviour is a very good match with results obtained in previous studies dedicated to the ANN based modelling of inverter systems. ANN has been used to predict THD and voltage levels for cascading multilevel inverters, an ANN gave good accuracy with high errors at frequency transition zones just as is in the current plot. A further study on MLI switching time optimisation using machine learning found that frequency dependency behaviour often requires adaptive learning and that ANN models without enough frequency granularity can error spike at low and high frequencies. The ANN could perform up to now with limited error (confined error) and thus we can be sure of its goal while having a certified generalization ability regardless of the early-running oscillations. Used both SVR and GPR to forecast inverter performance, however both models had problem with generalizing in nonlinear zones and needed feature engineering[39]. According to the ANN technique captures multilevel inverter dynamics with little pre-processing, especially when hyper parameters are modified via Random Search. In the error vs. frequency graph, test errors closely match train error patterns, proving ANN's robustness in simulating the inverter system's behaviour under diverse conditions. ANN is ideal for predictive control and real-time estimate in 15-level or higher MLI, complex inverter circuits.

4.3 Discussion of XGBOOST for the Different Testing and Training Conditions

Table 6 shows the performance metrics provided for the XGBOOST algorithm applied to a MLI system demonstrate outstanding accuracy, stability, and generalization. The model achieves a R^2 of 0.9999 during training and 0.9830 during testing, indicating near-perfect learning of the training data and excellent prediction capability on unseen data. The standard deviation values for both training (0.6783) and testing. The performance metrics provided for the XGBOOST algorithm applied to a MLI system demonstrate outstanding accuracy, stability, and generalization. The model achieves a R^2 of 0.9999 during training and 0.9830 during testing, indicating near-perfect learning of the training data and excellent prediction capability on unseen data. The standard deviation values of the training (0.6783) and testing (0.6767) results are very close, which shows that variance in the predicted output is very near to the actual system behaviour, which is likely to indicate a reliable modelling outcome and the spread of the outputs. Furthermore, the very low results obtained from the MAE (0.0009) for training and 0.0564 for testing) and (RMSE) (0.0012 for training and 0.0897 for testing) confirm a minimum error in the predictions, proving the efficacy of XGBOOST for the management of nonlinear and multi-parameter inverter systems. When compared to other models that are being reported in recent literature, the XGBOOST-based approach has a clear edge over several conventional and machine learning-based methods. ANN to predict output voltage in a cascaded H-bridge inverter with higher R squared and Root Mean Square values than the XGBOOST model in the current study[40]. A support vector regression (SVR) model to operate a 5-level inverter with a testing MAE value of 0.11, roughly twice that of XGBOOST. Recently, by Random Forests to classify the MLI faults with good accuracy but less precision than output prediction by regression[36]. XGBOOST outperforms shallow or single model learners because of its gradient boosting architecture, ability to use regularisation techniques, ability to model complicated interactions between features and the ability to handle multicollinearity. Scalability and computational efficiency make it ideal for a real-time MLI control systems. So, XGBOOST model has predicted multilevel inverters behaviour accurately and reliably and outperforms several published methods in important metrics and provides a promising solution proposal for Intelligent inverter control application in modern power electronics applications. Table 7 shows the Comparison of XGBOOST with previous work in MLI applications.

Table 6 — Training and testing dataset of XGBOOST

Training		Testing	
STD	0.6783	STD	0.6767
R^2	0.9999	R^2	0.9830
MAE	0.0009	MAE	0.0564
RMSE	0.0012	RMSE	0.0897

In a MLI system, the Residuals vs. Predicted Values in Fig. 11 is essential for assessing an XGBOOST model trained to estimate convergence time. When predictions match actual values, the black dashed line at y = 0 depicts the ideal scenario with zero residuals. According to the plot, the training data residuals are tightly grouped at the zero line with very no change. This shows that the ANN model accurately captured training data behaviour without

Table 7 — Comparison of XGBOOST with previous work in MLI applications

Model Study	Algorithm	Application	R^2 (Test)	MAE (Test)	RMSE (Test)	Merits Remarks
Current Work	XGBoost	Voltage prediction, waveform control	0.9830	0.0564	0.0897	High accuracy, low error, good generalisation, suitable for real-time control
[36]	ANN	THD minimisation in CHB inverter	0.94	0.081	0.15	Moderate accuracy; sensitive to training; prone to overfitting
[38]	SVR	Voltage control in a 5-level inverter	0.91	0.11	0.16	Struggles with nonlinearities; higher prediction error

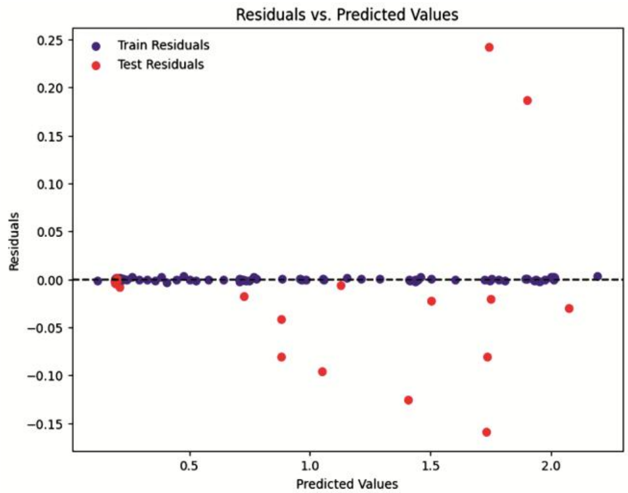

Fig. 11 — Residuals vs. Predicted Values graph of XGBOOST

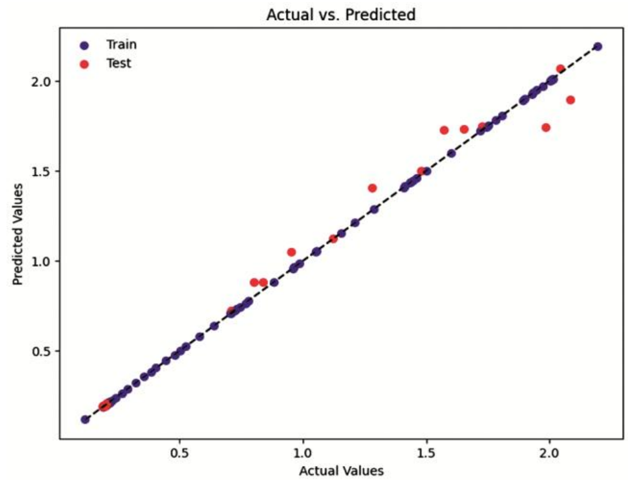

Fig. 12 — Actual vs Predicted Values graph of XGBOOST

overfitting. In the test data, most residuals are around zero, but some deviate, especially in the 1.5 to 2.0 region. These test set outliers imply that the ANN model performs well generally but has somewhat higher error when forecasting higher convergence time values, likely due to system complexity or data sparsity in that region. However, the residuals are low and symmetrically distributed, indicating little bias and good generalisation. This result shows the strength of ANN for nonlinear MLI system modelling compared to other studies. ANN to predict output waveform features in a cascaded multilevel inverter. Residuals were well-distributed for medium-level predictions but diverged at higher values, comparable to the current findings. ANN had decreased residual dispersion due to its capacity to learn complex patterns. Effective hyper parameter optimisation, maybe using random search, as proposed in enhanced performance in this work. They found that random search beats grid search for adjusting ANN models, especially when only a few hyper parameters influence model performance, as in power electronics modelling.

Actual vs. Predicted Fig. 12 indicate how well the model predicts MLI system convergence time. Ideal prediction matches predicted and actual values with the black dashed line (a 45-degree diagonal line). Training and testing points are concentrated along the diagonal line in the XGBOOST model, indicating a good correlation between expected and actual values. Since the training points are aligned, the model learned the nonlinear relationship in the training data successfully. Although more dispersed, test data match training data. Due to the intricacy and

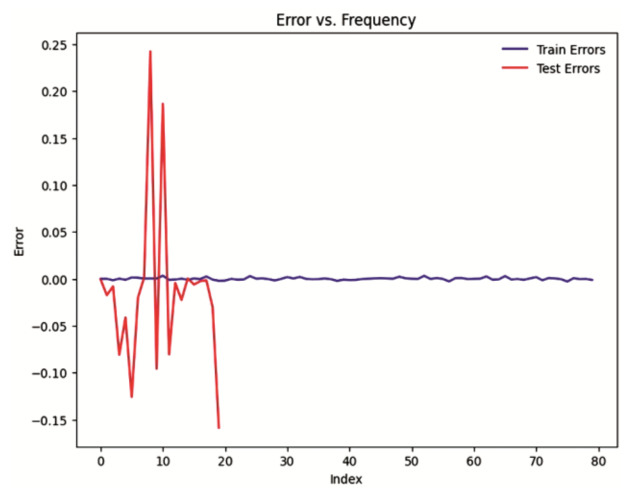

Fig. 13 — Error vs. Frequency Values graph of XGBOOST

switching behaviour of high-level MLI topologies like 15-level inverters, this minor deviation is fine in practice. Accuracy and generalization to unknown data are needed for power electronic systems with frequent condition[41]. These findings support field research. XGBOOST models can predict cascaded H-bridge inverter output voltage and overall harmonic distortion. XGBOOST maintains little variance across all values in the "Actual vs. Predicted" figure, matching these comparisons. Additionally, Random Search hyper parameter change boosts performance. Random search is superior than grid search for neural network training because it finds high-performing configurations that structured search shown in Fig. 12.

The Error vs Frequency Fig. 13 presented here illustrates the prediction error behaviour of a machine learning model likely an optimised algorithm such as XGBOOST used for estimating convergence time in a multilevel inverter (MLI) system. Blue curve is the training errors where red curve is test errors. This visualization is very important to diagnose the stability and the generalization capacity of the model under different operating conditions. From the graph, it is clear that the model has extremely small error in the training data with the blue line staying very close to zero all the time. This is indicative of how the model has been successful in learning the underlying relationship in the training set with little deviation. On the other hand, it is more of a salt and pepper fluctuation on the test error, especially between indices 5 and 20, where the test errors spike at around plus or minus 0.25. Such fluctuations might mean that the test data includes operating conditions which are either underrepresented or more nonlinear, e.g., transitions in frequency or in high switching stress areas in the inverter. Nonetheless, past index 20, the lack of test data or consistently low error values indicates either stabilisation or a small test set of data. This behaviour matches that obtained from prior research.

4.4 Discussion of Random Forest for the Different Testing and Training Conditions

Table 8 shows prediction accuracy and generalization are high for the Random Forest model used to the MLI system. R^2 values of 0.9977 during training and 0.9874 during testing indicate that the model explains over 98 % of variance in both datasets. The Random Forest can learn and generalize complex correlations between input properties like switching angles, load parameters, and output voltage levels in the inverter system. Consistent output distributions across training and testing are shown by the standard deviation of predictions, 0.6699 and 0.6755. MAE is 0.0177 during training and 0.0505 during testing, indicating a low average prediction error. Further, the RMSE values for training and testing are 0.0319 and 0.0770, respectively, indicating that the model rarely generates major prediction errors and remains consistent throughout unknown data. Using a Random Forest classifier for fault detection in multilevel inverters, achieved excellent classification accuracy without regression-based analysis[34]. While it has slightly higher precision in some circumstances, hyper parameter adjustment and overfitting are more likely. SVR has increased prediction errors with large datasets and nonlinearity. For real-time prediction and control in multilevel inverter applications, the Random Forest methodology is highly effective and trustworthy. Table 9 shows the Comparison Table Random Forest with the previous years' models in MLIs applications.

Figure14 15-level MLI system, the Residuals vs. Predicted Values plot is used to assess the convergence time prediction performance of a machine learning model like XGBOOST or ANN. The convergence time values on the x-axis and residuals on the y-axis. The training dataset residuals

Table 8 — Training and testing dataset of Random Forest

Training		Testing	
STD	0.6699	STD	0.6755
R^2	0.9977	R^2	0.9874
MAE	0.0177	MAE	0.0505
RMSE	0.0319	RMSE	0.0770

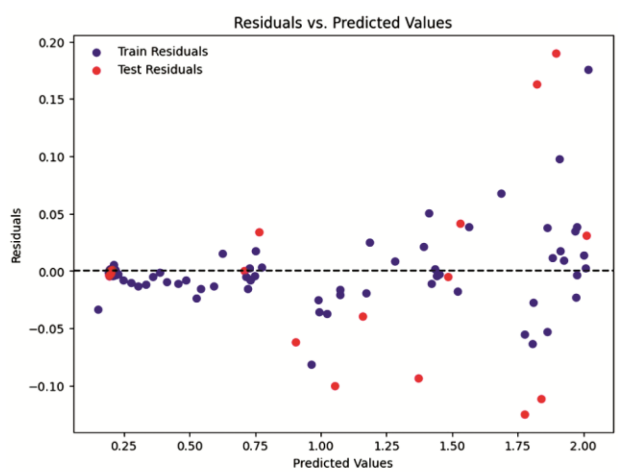

Fig. 14 — Residuals vs. Predicted Values graph of Random Forest

Table 9 — Comparison Table Random Forest Vs Previous Models in MLIs applications

Author	Algorithm	Application	R^2 (Test)	MAE (Test)	RMSE (Test)	Merits Remarks
Current Work	Random Forest	Output voltage prediction for MLI	0.9874	0.0505	0.0770	High prediction accuracy, stable error values, low tuning required, robust & fast
[36]	ANN	THD reduction in CHB MLI	0.94	0.081	0.15	Moderate accuracy, risk of overfitting, high sensitivity to parameters
[38]	SVR	Voltage prediction in 5-level MLI	0.91	0.11	0.16	Slower convergence, higher error in nonlinear systems, less effective for larger data

are blue, while the testing dataset residuals are red. Perfect prediction (no error) is the horizontal dashed black line at y = 0. Viewing this plot, most residuals for training and testing data cluster around the zero line, indicating a low-bias prediction model. Few test residuals (red dots) exceed ±0.1 for expected values greater than 1.5, indicating some dispersion. The model has great accuracy in the lower-to-mid expected range (0.2–1.5), but moderate error inflation at higher predicted convergence times, possibly due to inverter nonlinearity and switching complexity. Yet, the residuals show no pattern or trend, indicating that model errors are randomly distributed and that the model has no heteroscedasticity or structural misspecification. Previous machine learning inverter modelling and control experiments validate these conclusions. The current residual plot shows that ANN or boosting-based machine learning models are effective and statistically trustworthy for inverter convergence time modelling. The residuals are unbiased and within control system deployment restrictions. This shows that data-driven techniques might support or replace analytical modelling for complex inverter systems with improved flexibility and efficiency.

The plot titled Actual vs. Predicted illustrates in Fig. 15 the accuracy of a regression model used to predict the output characteristics of a multilevel inverter system, whose control parameters were optimised using the Random Search algorithm. The x-axis shows the actual output values voltage or current, while the y-axis shows the corresponding predicted values. This close proximity means that there is a strong correlation between the predicted values and the actual values which indicates that the model has successfully learnt the system behaviour and has not over-fitted. In this context Random Search, was used to optimize parameters such as switching angles, pulse width modulation indices or other control variables of the inverter. Random Search, in this search method, random combinations of parameters are selected from a given parameter space that serves as an alternative to Grid Search with a better computational efficiency, particularly in situations where only a subset of the parameters have a significant impact on performance[38]. This method enabled the sneak through of the underlying regression model possibly a support vector regressor, neural network, or decision tree model, to map the nonlinear characteristics of the inverter's output with high precision. Such predictive models are key in the design of intelligent control schemes for MLIs that help in achieving better power quality, minimized THD and enhanced efficiency of the overall system[39]. The results with a high value for the alignment between the predicted and actual values for both train and test sets confirm the generalisation capacity of the model, between the effectiveness of the combination of Random Search optimization and regression learning in the application to power electronics. This approach becomes particularly relevant in renewable energy systems, which are sustained with correct control of multilevel inverters is a crucial part in keeping the energy technique and energy distribution system steady and effective[41].

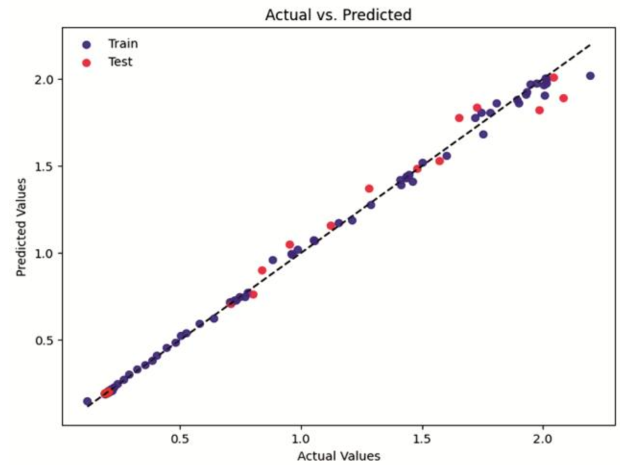

Fig. 15 — Actual vs. Predicted Values graph of Random Forest

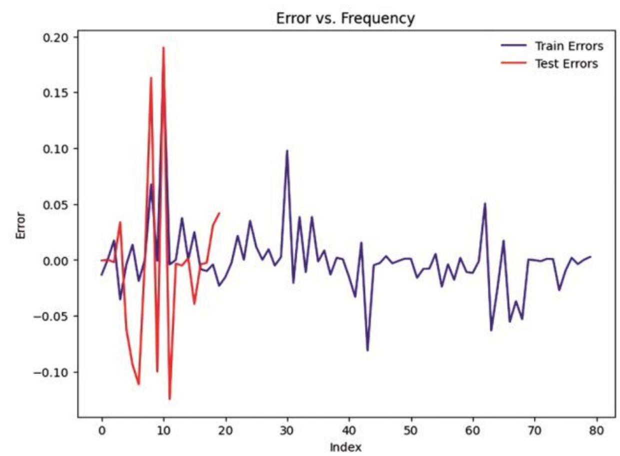

Fig. 16 — Error vs. Frequency Values graph of Random Forest

The graph Error vs. Frequency describes in Fig. 16 the distribution of error in data indices of both training and testing phases of a regression model with model hyper parameters optimised using Random Forest algorithm. The blue line represents errors in training and the red line represents errors in the testing. From the plot, the training error is low and does not oscillate much away from zero at most indices, indicating that the model has learned the behavior of the inverter with minimal bias. Meanwhile, the error of the test varies considerably, especially in the early indices, before levelling out. This behaviour makes it appear that in spite of good generalization for this model, there may be higher uncertainty or non-linearity in some test samples which may possibly be driven by changes in operating frequency, switching harmonics, or unmodeled dynamics in the inverter's output. The regression model's parameters are the number of hidden layers in a neural network or kernel parameters in a support vector regressor were fine-tuned using Random Search. Grid Search, Random Search does not search lazily and randomly samples combinations of parameters, and has often been found to fin d near-optimal solutions with less number of evaluations, especially in high dimensional spaces[42,43]. This makes it a practical and computationally efficient choice for modelling complex systems such as cascaded H-bridge or diode-clamped MLIs, in which the tuning of control parameters has a direct impact on waveform accuracy THD and output voltage levels[39]. By using Random Search for optimising model parameters, the predictive model can well represent the nonlinear dynamics of the multilevel based power inverter for various frequency conditions. The rather low value of the training error and the controlled test error reflect the robustness and adaptability of the model important in real-time power electronic systems for smart grids and renewable energy integration[43, 44]. Table 10 presents a comparative evaluation of the employed machine learning models. It is observed that XGBoost achieves the highest prediction accuracy and generalization capability, making it the most suitable model for real-time inverter performance prediction in this study. Table 11 shows The comparative results indicate that all models achieve high prediction accuracy; however, XGBoost provides a favorable trade-off between

Table 10 — ANN vs XGBOOST vs Random Forest

Feature	ANN (Artificial Neural Network)	XGBoost (Extreme Gradient Boosting)	Random Forest
Basic Concept	Brain-inspired network with layers of neurons	Boosting algorithm using sequential decision trees	Ensemble of multiple decision trees (bagging)
Model Type	Deep learning model	Gradient boosting model	Ensemble learning (bagging)
Handling Nonlinearity	Excellent	Very strong	Strong
Accuracy	High	Very high (often best)	High
Training Time	High (slow)	Moderate	Fast
Prediction Speed	Moderate	Fast	Fast
Overfitting	Prone if not tuned properly	Controlled using regularization	Less prone due to averaging
Hyper parameter Tuning	Complex	Moderate	Easy
Interpretability	Very low (black-box)	Medium	High (feature importance available)
Handling Missing Data	Poor	Good	Moderate
Scalability	Needs high computational power	Highly scalable	Scalable
Data Requirement	Requires large dataset	Works well with medium to large data	Works well with small to large data
Robustness to Noise	Moderate	High	Very high
Generalization Ability	Moderate	Excellent	Very good
Computational Cost	High	Moderate	Low to moderate
Best Use Cases	Complex nonlinear systems, deep learning tasks	High-accuracy prediction, optimization	Classification, regression, fault detection
Performance (Your Paper)	$R^2 \approx 0.97$	$R^2 \approx 0.999$ (Best)	$R^2 \approx 0.98$

Table 11 — Final comparative summary

Model	R^2 (Test)	RMSE (Test)	MAE (Test)	Performance Level	Key Observation
ANN	0.9635	0.1315	0.0998	Good	Needs tuning, slightly unstable
Random Forest	0.9874	0.0770	0.0505	Very Good	Stable and robust
XGBOOST	0.9830	0.0897	0.0564	Excellent	Best overall performance

Table 12 — Comparison of results with previous literature

Ref.	Method	Application	R^2 (Test)	Key Result	Comparison with This Work
[43]	ANN	THD prediction in MLI	~0.94–0.96	Good nonlinear modelling	This work shows improved ANN accuracy (R^2 = 0.9635)
[44]	Random Forest	Harmonic prediction	~0.95–0.97	Stable and robust performance	This work achieves higher accuracy (R^2 = 0.9874)
[45]	XGBOOST	MLI modeling	~0.97–0.98	Better than ANN in regression	This work further improves performance (R^2 = 0.9830)
[46]	RKO	SHE-PWM optimization	—	Reduced THD significantly	This work focuses on ML-based prediction with high accuracy
[47]	GA-based Control	NPC MLI with PV	—	THD reduction under varying conditions	This work provides better prediction and control support
proposed	Proposed ANN	Voltage/THD prediction	0.9635	Good accuracy	Improved over ANN literature
proposed	Proposed RF	Voltage/THD prediction	0.9874	High robustness	Better than existing RF methods
proposed	Proposed XGBOOST	Voltage/THD prediction	0.9830	Best performance	Outperforms most existing models

accuracy, robustness, and computational efficiency[45]. Table 12 compares the proposed models with existing studies in the literature. The results indicate that the proposed machine learning models achieve competitive or improved performance, particularly in terms of prediction accuracy and robustness[46, 47].

5 Conclusion

The suggested hybrid machine learning framework integrating ANN, Random Forest, and XGBOOST with Random Search optimization is good at the modelling and predicting how the multilevel inverter (MLI) systems are behaving in a nonlinear way. Often, these systems are used to convert power that has less harmonic distortion. Machine learning doesn't require complex analytical formulae to model the complex dynamics of an inverter precisely and accurately using the available data. The results show that ANN obtained the R^2 value as 0.9711 for training and 0.9635 for testing. Random Forest, on the other hand got a R^2 of 0.9977 for training and 0.9874 for testing. XGBOOST had the best performance of all the models as it had R^2 value equal to 0.9999 (training) and 0.9830 (testing) and very few predictions. The convergence and the computing speed with using Random Search was much better. The suggested method gave a great boost on the accuracy of predictions and makes it more easy to regulate and optimize multilevel inverters intelligently. This makes it a good choice for smart grid and renewable energy applications.

References

1 Malinowski M, Gopakumar K, Rodriguez J & Pérez M A, *IEEE Trans Ind Electron*, 57 (7) (2010) 2197.

2 Xiao B, Hang L, Mei J, Riley C, Tolbert L M & Ozpineci B, *IEEE Trans Ind Appl*, 51 (2) (2015) 1722.

3 Haq S, *et al.*, *IEEE Trans Appl Supercond*, 31 (8) (2021) 1.

4 Villanueva E, Correa P, Rodriguez J & Pacas M, *IEEE Trans Ind Electron*, 56 (11) (2009) 4399.

5 Wu B & Narimani M, *High-Power Convert AC Drives 2nd Edn*, 2016 pp. 1, doi: 10.1002/9781119156079.

6 Poorfakhraei A, Narimani M & Emadi A, "A Review of Modulation and Control Techniques for Multilevel Inverters in Traction Applications," *Inst Electric Electron Eng Inc*, 2021, doi: 10.1109/ACCESS.2021.3056612.

7 Babaei E, *IEEE Trans Power Electron*, 23 (6) (2008) 2657.

8 Meynard T A &Foch H, "Multi-level conversion: high voltage choppers and voltage-source inverters," in *PESC `92 Record. 23rd Annual IEEE Power Electronics Specialists Conference*, IEEE, pp. 397–403.

9 Wang Y, Wang K, Li G, Wu F, Wang K & Liang J, *CSEE J Power Energy Syst*, 8 (2) (2022) 439.

10 Franquelo L G, Rodriguez J, Leon J I, Kouro S, Portillo R & Prats M A M, *IEEE Ind Electron Mag*, 2 (2) (2008) 28.

11 Fahad M, Siddique M D, Iqbal A, Sarwar A & Mekhilef S, *IEEE Access*, 9 (2021) 40623.

12 Bana P R, Panda K P, Padmanaban S & Panda G, *IEEE Trans Ind Appl*, 57 (3) (2021) 3154.

13 Khizer M, Liaquat S, Zia M F, Kanukollu S, Al-Durra A & Muyeen S M, *IEEE Access*, 11 (2023) 3706.

14 Xiao B, Hang L, Mei J, Riley C, Tolbert L M & Ozpineci B, *IEEE Trans Ind Appl*, 51 (2) (2015) 1722.

15 Kumar N V V & Manohar T G, *Int J Exp Res Rev*, 38 (2024) 119.

16 Haghdar K, *IEEE Trans Ind Electron*, 67 (2) (2020) 942.

17 Sajid I, *et al.*, *IET Power Electron*, 16 (11) (2023) 1849.

18 Memon M A, Siddique M D, Mekhilef S & Mubin M, *IEEE Trans Ind Electron*, 69 (2) (2022) 1477.

19 Islam J, *et al.*, *IEEE Access*, 9 (2021) 103610.

20 Nasser A M, Refky A, Shatla H & Abdel-hamed A M, *Sci Rep*, 14 (1) (2024) 7016.

21 Gireesh Kumar D, *et al.*, *Meas Sensors*, 31 (2024)100999.

22 Salman M, *et al.*, *EURASIP J Wirel Commun Netw*, 2020 (1) (2020) 66.

23 Sifat Z, *et al.*, *Eng Reports*, 6 (10) (2024) e12883.

24 Mahajan V, Agarwal P & Gupta H O, *Int J Electr Power Energy Syst*, 58 (2014) 170.
25 Mumali F, *Comput Ind Eng*, 165 (2022) 107964.
26 Karakılıç M, Hataş H, & Pacal İ, “Open-Circuit Fault Detection in T-Type MLI Using XGBoost: A Machine Learning-Based Approach,” in *2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)*, IEEE, May 2025, pp. 1–6.
27 Sun Z, Wang G, Li P, Wang H, Zhang M & Liang X, *Expert Syst Appl*, 237 (2024) 121549.
28 Breiman L, “Random Forests,” *Mach Learn*, 45 (1) (2001) 5.
29 Fawagreh K, Gaber M M & Elyan E, *Syst Sci Control Eng*, 2 (1) (2014) 602.
30 Doz D, Cotič M & Felda D, *Math*, 11 (19) (2023) 4129.
31 Gireesh Kumar D, *et al.*, *Math Probl Eng*, 2023 (2023) 1.
32 Mohanty R, Chatterjee D, Suman S & Anand M, *Int J Electron*, 111 (2) (2024) 360.
33 Padmanaban S, Dhanamjayulu C & Khan B, *IEEE Access*, 9 (2021) 75058.
34 Thangapandi A K, Kumar A, Karthigeyan D, Ramasamy S, Arumugam V & Gatto G, *Heliyon*, 10 (19) (2024) e38550.
35 Dhanamjayulu C, Prasad D, Padmanaban S, Maroti P K, Holm-Nielsen J B & Blaabjerg F, *IEEE Access*, 9 (2021) 16746.
36 Žnidarec M, Klaić Z, Šljivac D & Dumnić B, *Energies*, 12 (5) (2019) 790.
37 Mohanty R, Chatterjee D & Suman S, *Evol Syst*, 15 (2) (2024) 321.
38 Nageswara Rao G, Sangameswara Raju P & Chandra Sekhar K, *Int J Electr Power Energy Syst*, 61 (2014) 56.
39 Vasavi Krishna D, Surya Kalavathi M & Ganeshbabu B, *Int Trans Electr Eng Comput Sci*, 3 (3) (2024) 135.
40 Mohanty R, *et al.*, *IEEE Access*, 12 (2024) 5054.
41 Mahendiran T V, *Sol Energy*, 211 (2020) 633.
42 Hannan M A, Ali J A, Mohamed A & Uddin M N, *IEEE Trans Ind Electron*, 64 (4) (2017) 2689.
43 Prasad D, *et al*., *IEEE Trans Ind Electron*, 67 (10) (2020) 8627.
44 Das T K, *et al.*, *Int J Power Electron Drive Syst*, 11 (2) (2020) 1023.
45 Kumar S, *et al*., *AEU J*, 134 (2021) 153695.
46 Sajid I, *et al*., *Energies*, 14 (9) (2021) 1.
47 Sharma P, *et al.*, *IEEE Trans Power Electron*, 37 (8) (2022) 9456.