The Scientific Temper (2024) Vol. 15 (4): 3087-3098

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.16

E-ISSN: 2231-6396, ISSN: 0976-8653

https://scientifictemper.com/

RESEARCH ARTICLE

Power quality assessment in solar-connected smart grids *via* hybrid attention-residual network for power quality (HARN-PQ)

S. Dhivya*, S. Prakash

Abstract

The aim of the proposed method is to solve the difficulties associated with anomaly detection and real-time data processing in complex network systems. The process begins by collecting data from internet of things (IoT) devices and smart grid sensors. Advanced interpolation techniques are used in pre-processing methods to deal with missing data, while the Isolation Forest algorithm is used to find outliers. Ensures data normalization through robust scaling, reducing the impact of outliers. Higher-order statistics such as skewness, kurtosis, and entropy measures, as well as various statistical metrics such as mean absolute deviation (MAD), interquartile range (IQR), and coefficient of variation (CV) are extracted in the feature extraction process. A unique method called hybrid horse-based zebra optimization (HHZO) is used to select features. It combines the zebra optimization algorithm (ZOA) and the horse herd optimization algorithm (HHO). Weighted ensemble energy quality residual attention network (WEARN-PQ) architecture is proposed for deep learning-based detection, which integrates extended recurrent neural networks (Stack-RNN) and stack-gated recurrent units (GRU) with attention layers and convolutional neural networks (CNN) with residual connections and attention mechanisms. To ensure reliability, split-sampling K-Fold cross-validation is used during training and validation.

Keywords: Smart grid sensors, Hybrid Horse based Zebra optimization, Weighted ensemble based attention-residual network, Power quality, Stacked gated recurrent units, K-Fold cross-validation.

Introduction

Solid-state device employment in a power system may lead to power delivery to end users that are not balanced and may exhibit harmonics, flashing, notches, spikes, and voltage swings. For generating, transmission, and distribution systems, the primary goal is to guarantee a steady, uninterrupted sinusoidal voltage with balanced sinusoidal currents. For sensitive and important loads, smooth, continuous sinusoidal voltage with regulated currents and consistent frequency and magnitude is required (Mohamed, 2022; Sekhar, 2022; Nandagopal *et al*., 2023). A protective system malfunction could result in major losses of data, time, product quality, and services if incompetent power providers are unable to supply such highly qualified electricity. Clear power quality standards have been created by the International Electrotechnical Commission (IEC) and the Institute of Electrical and Electronics Engineers (IEEE) in order to guarantee uniformity in power quality (Ratnakaran *et al*., 2023; Ravi and Kumar, 2023).

There is a sharp increase in the usage of distributed generation (DG) technologies, such as fuel cells, solar power plants, wind turbines, and others. Diverse economic, environmental, and technical advantages come with DG. On the other hand, DG integration with electric utilities systems is fraught with difficulties (Afzal *et al*., 2022; Akpolat *et al*., 2023; Rajagopalan *et al*., 2022). The main DG control parameter is reactive power compensation. The energy consumption patterns of large loads can frequently be changed, causing voltage sags and swells in the system that can change the true power demand. Reactive power that has not been adjusted can also have an impact on distributed generation (DG) systems' efficiency, power factor (PF), and active power capacity (Ayalew *et al*., 2022; Khalkhali *et al*., 2022; Thentral *et al*., 2022). When connecting DG systems to the utility grid, power electronic converters (PECs) must be used in order to guarantee the equipment's

Department of Electrical and Electronics Engineering, Bharath Institute of Higher Education and Research, Chennai, India.

***Corresponding Author:** S. Dhivya, Department of Electrical and Electronics Engineering, Bharath Institute of Higher Education and Research, Chennai, India., E-Mail: sdhivya955@gmail.com

How to cite this article: Dhivya, S., Prakash, S. (2024). Power quality assessment in solar-connected smart grids via hybrid attention-residual network for power quality (HARN-PQ). The Scientific Temper, **15**(4):3087-3098.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.16

Source of support: Nil

Conflict of interest: None.

Received: 10/10/2024 **Accepted:** 16/11/2024 **Published** : 20/12/2024

safe operation and enable source switching. Though this can lead to a number of power quality (PQ) problems (Eristi and Eristi, 2022), reactive power loads, imbalances, flicker, interruptions, neutral current, impulse transients, voltage and current harmonics, and voltage sag/swell are all included in this spectrum (Krishna *et al*., 2022; Kumar *et al*., 2022).

Power electronics switching devices are widely used, which has led to a number of PQ issues with voltage and current in transmission systems (Sarita *et al*., 2023). PQ issues in networked power systems, such as low power factor, reactive power demand, load imbalances, and excessive current harmonics, can be categorized as issues pertaining to utilities or customers (Rao *et al*., 2023). Utility-related difficulties, on the other hand, include things like flicker, notches, voltage sag/s well, uneven loads, and voltage distortions (Ravi and Kumar, 2023). PQ issues can drastically decrease. These PQ issues may result in decreased power transmission efficiency as well as harm to distribution system-connected equipment. The main contributions of the paper are as follows:

- A unique optimization approach is proposed for selecting the most relevant features from the dataset. This approach combines two algorithms, the horse herd optimization algorithm (HHO) and the zebra optimization algorithm (ZOA), to enhance feature selection.
- A specialized architecture called WEARN-PQ is introduced for deep learning-based anomaly detection. This architecture integrates various deep learning techniques such as CNNs, GRUs, and Stack-RNNs.
- Training and validation of the models are performed using K-Fold cross-validation with stratified sampling. This technique helps ensure the robustness and reliability of the models by systematically validating them on different subsets of the data.

The study is organised as follows: section 2 discusses current existing works. Section 3 explains the suggested method for identifying migraines, section 4 compares the findings of our proposed model to existing methodologies, and section 5 concludes the study with a conclusion.

Literature Review

The power quality in integrated solar energy systems was enhanced by using a modular multilevel converter (MMC) as the basis of a unified power quality conditioner (UPQC) to improve the electrical reliability of medium-voltage and high-voltage solar power systems connected to a PV grid. The proposed MMC-UPQC is highly standardized and features low DC link voltage, excellent harmonic isolation, and improved main system voltage regulation. Since the fuzzy controller deals with uncertainties and nonlinearities in the system, it is used as a DC voltage regulator. The system can be implemented in scenarios where complex system dynamics occur because it uses a set of fuzzy rules to help map input signals to output signals (Garikapati *et al*., 2023).

The frequency deviation identification and fault effect analysis in a smart solar-connected grid were evaluated using several new islanding detection techniques for grid-connected PV systems. As per the discourse, the anti-islanding research trend can be broadly classified into two categories: passive approaches, which rely on measuring system parameters like frequency and voltage. Unfortunately, not all load circumstances can be trusted using passive approaches, and setting thresholds might be challenging because of their size. At the time of islanding, the injected signal causes a terminal voltage variation (Tripathi and Pachori, 2023).

Artificial neural network (ANN) controller with an ideal design and performance study was created for a UPQC that was connected to both solar and batteries. A reduction in the mean square error (MSE) is the main objective of the FF-ANNC suggestion. In addition, this will reduce sag, swell, and other problems and increase power factor (PF). It will also keep the DC link capacitor voltage (DLCV) constant during load and radiation fluctuations. To assess how well the suggested FF-ANNC performed, five test studies with various load kinds and source voltage balancing/unbalancing circumstances were employed (Ramadevi *et al*., 2023).

The power quality issues in solar PV integrated power systems are mitigated by ANFIS-controlled MMC-UPQC. To improve power quality, the endeavor seeks to limit changes in voltage and current, remove harmonics produced by the VSC, manage the DC link voltage, and compensate for reactive power requirements from the PV system and load. The adaptive neuro-fuzzy inference system is the controller utilized for MMC-UPQC control in DC voltage regulation (ANFIS). Numerous dynamic situations are considered, including changes in demand, irradiance fluctuations for the PV system, and voltage sag and swell at the point of common coupling (PCC) (Garikapati *et al*., 2023).

An innovative fuzzy logic controller with MLIUPQC improving the quality of PV-BESS photovoltaics was used in this study. A battery energy storage system (BESS) with a fuzzy logic controller (FLC) is provided for a solar-powered multilayer UPQC inverter. It was simpler to resolve power quality issues in the distribution system by utilizing the new topology. Shunt and series regulators linked to the distribution system's common PCC enable UPQC regulation. More improvements have been made to system stability and power quality than with UPQC's traditional two-level inverter and PI controller architecture (Tejakrishna and Ramu 2022).

UPQC and the updated multilevel modular converter work together to improve power quality. To address the power quality challenges related to voltage and current, this study discusses the implementation of a modified UPQC based on a modular multilevel converter. Moreover, MMC improves system performance compared to traditional methods and also has a very simple modular design. The proposed design provides voltage-related compensation using a seven-level MMC and current compensation using a low-voltage source inverter with four switches (Thentral *et al*., 2022).

Power quality using solar PV fed UPQC was enhanced using an updated generalized integrator-based control approach. This study discusses a modified generalized integrator with DC offset elimination capabilities that control a UPQC with SPVA (UPQC-SPVA) system. The extracted basic element of the load current is processed by the shunt active filter of the UPQC-SPVA system to produce reference currents. This control method improves power quality (PQ) in a number of ways, including reactive power adjustment, harmonic elimination, grid voltage sag reduction, and voltage swell mitigation. By adding a solar PV array at the UPQC DC connection, this solution offers benefits for clean energy and increases PQ (Chandrakala Devi *et al*., 2020).

Problem statement

Although solar photovoltaic (PV) systems are being integrated into power grids at an increasing rate, power quality problems still exist and pose threats to the stability and dependability of the grid. These problems include fluctuations in voltage and current, harmonic distortions, spikes and drops in voltage, and demands for reactive power. Because these issues are dynamic in nature, traditional power quality mitigation strategies are frequently insufficient to handle them, particularly in solar-integrated systems where uncertainties and nonlinearities are common. There is still a need for dependable, effective, and flexible solutions that can successfully mitigate power quality issues in solar PV integrated power systems, despite the fact that numerous research efforts have proposed solutions like UPQCs and advanced control methods like fuzzy logic, ANNs, and fuzzy-based modular control.

When it comes to dealing with dynamic system variables such as load changes and solar radiation changes, existing systems often face obstacles such as complex control design, difficulties in defining thresholds and island detection limits. Therefore, to ensure grid stability and reliability in the presence of renewable energy sources, there is an urgent need for innovative research and development to improve the power quality of integrated photovoltaic systems. Research and development efforts should focus on developing advanced control strategies, efficient hardware applications, and robust island detection techniques.

Proposed Methodology

The aim of the proposed methodology is to address the challenges of anomaly detection and real-time data processing in complex network systems. Smart networked sensors and Internet of Things (IoT) devices are used to collect and track a continuous stream of data. Data integrity is ensured through advanced pre-processing techniques such as interpolation and outlier detection. Robust measurement strategies standardize data and reduce the impact of outliers. Time series analysis and higher-order statistics are used in the feature extraction phase to extract important insights. Feature selection is performed using the unique hybrid HHZO algorithm. Deep learning-based anomaly detection is enabled by the WEARN-PQ architecture, which combines CNN, SGRU and Stack-RNN networks. Figure 1 shows the workflow of writing the article.

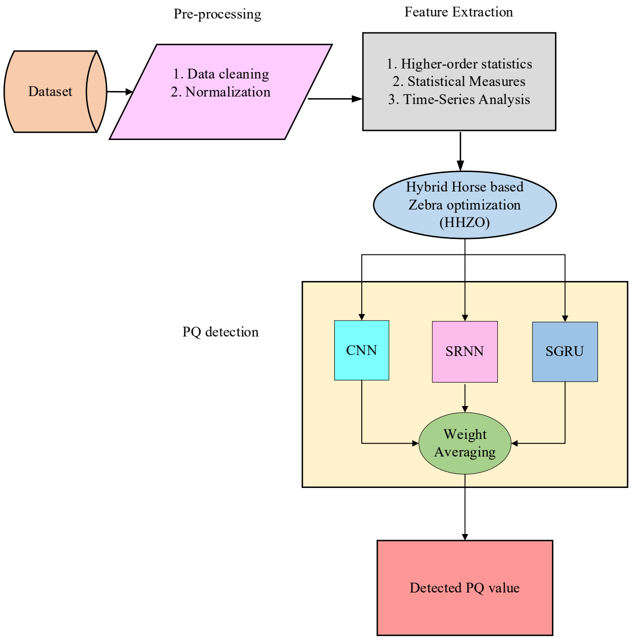

Figure 1: Block diagram of the proposed PQ assessment model

Pre-processing

The pre-processing phase of this technique involves outlier selection using the isolation forest algorithm to find and remove unusual data points. To ensure the completeness of the dataset, interpolation techniques are used to fill in missing data. Robust scaling is used to standardize the data and reduce the impact of outliers.

Data cleaning

- *Isolation tree*

The procedure for building an isolation tree (iTree) in an isolated forest (iForest) is based on the idea of making random partitioning decisions. There are two sections of the IF. The first step is to create an isolated forest using iTrees, and the second is to determine the degree of skewness for each study sample.

IF is used for the calculation and evaluation of abnormal scores for any study sample. This means that you should walk through every iTree in the forest one at a time, extract the sample's node depth from each iTree, and then figure out the sample's average depth across the forest. Considering a subsampled set of Φ instances, the sample's average path length is,

$$c(\varphi) = \begin{cases} 2H(\varphi - 1) - \frac{2(\varphi-1)}{\varphi}, & \varphi > 2 \\ 1, & \varphi = 2 \\ 0, & \varphi < 2 \end{cases} \tag{1}$$

Where $H(k)$ is the harmonic number, which is calculated using the Euler's constant, $ln(k) + 0.5772156649$. The average of $h(\eta)$ is $c(\phi)$, which use to normalize $h(\eta)$, where is the depth of instance η in the iTree node. The definition of the instance η's anomaly score is,

$$s(\eta) = 2^{-\frac{E(h(\eta))}{c(\varphi)}} \tag{2}$$

Where the average of $h(\eta)$ from a group of iTrees is denoted by $E[h(\eta)]$. The probability that the sample η is an anomalous point increases with the distance between $s(\eta)$ and 1. It is more likely to be a normal sample the closer it is to 0.

- *Spline interpolation*

A smooth interpolation result is ensured by spline interpolation, which uses cubic interpolation to create piecewise polynomials with continuous second-order derivatives. The shape of a spline is modeled for $n + 1$ pair of observations $\{(t_i, x_i): i = 0,1,\ldots n\}$ by using polynomials given in Eq. (6) to interpolate between all the pairings of observations $(t_i - 1, x_i - 1)$ and (t_i, x_i).

$$\boldsymbol{x = q_i(t), i = 1, 2, \ldots, n} \tag{3}$$

Normalization using robust scaling

A technique called robust scaling uses statistics that are resistant to outliers to scale features. By removing the median, this scaler quantiles the data (IQR: Interquartile Range is the default).

$$X_{scaled} = \frac{X - X_{median}}{IQR} \tag{4}$$

Where IQR stands for interquartile range, X_{median} for X's median, X_{scaled} for the scaled feature, and X for the original feature. Strong in the face of outliers, which makes it appropriate for features with skewed distributions or outliers.

Feature Extraction

The goal of the feature engineering stage is to extract useful details from the pre-processed data. This involves collecting higher-order statistics like kurtosis and skewness and performing time-series analysis to find dependencies and trends in the data across time.

Higher-order statistics

The features like skewness, kurtosis, and entropy measures are extracted from this stage.

- *Kurtosis*

The degree to which the data deviate from a normal distribution is indicated by a statistic known as kurtosis. Conversely, huge outliers or heavy tails are typically present in data sets with a high kurtosis. Light tails and a few outliers are common features of data sets with low kurtosis. The worst-case situation would be a uniform distribution.

$$Kurtosis = \frac{4^{th}\, Moment}{4^{th}\, Moment^2} \tag{5}$$

- *Skewness*

The degree to which the data deviate from a normal distribution is indicated by a metric known as kurtosis. Similarly, datasets with high kurtosis usually feature heavy tails or notable outliers. Data sets with low kurtosis frequently have light tails and no outliers. The worst-case scenario would be a uniform distribution.

$$Skewness = \frac{3(\text{Mean}-\text{Median})}{SD} \tag{6}$$

- *Entropy measures*

The flow elements' entropy is calculated for every time slot. The randomness of a set of data is measured by entropy. There is an inverse relationship between entropy and data randomness. For a given random variable X, its entropy equals.

$$H(X) = \sum_{i=1}^{N} p(x_i) \log_2\left(\frac{1}{p(x_i)}\right) \tag{7}$$

Where the range of values of X is denoted by $x_1, x_2, \ldots x_N$, and the probability of X has the value x_i is represented by $p(x_i)$. Among all probability distributions, the uniform distribution $\log_2\left(\frac{1}{N_0}\right)$ entropy is the biggest. The quantity of unique xi values in a time slot is denoted by N0.

Statistical measures

In statistical measures, features like MAD, IQR and CV are extracted.

- *Mean absolute deviation (MAD)*

The data's dispersion can be inferred from this dependable measure of data point variability around the median value. The standard deviation and variation are two other measurements of variability or dispersion, but the MAD is more reliable since it doesn't change when the extreme values change. Determined by computing the absolute deviation from the given data's median using the MAD values.

$$MAD = median[|s_i - \tilde{s}_i|] \tag{8}$$

$$\tilde{s}_i = median(s_i) \tag{9}$$

Where $\tilde{s}_i$ denotes an individual observation in the provided data, and $median(s_i)$ denotes the data's median.

- *Interquartile range (IQR)*

It calculates the range that contains the middle 50% of the data. Variability is the characteristic that divides a dataset into quartiles or regions that contain the majority of the value. The quartiles of the dataset are the first, second, third, and fourth. The symbols SQ1, SQ2, and SQ3, respectively, indicate the three quartile boundaries that separate the dataset's equal halves. The first half of the dataset's SQ1 represents the middle value; the second half's SQ2 represents the median value; and the third half's SQ3 represents the middle value. Following is the definition of the interquartile range.

$$IQR = SQ3 - SQ1 \tag{10}$$

- *Coefficient of variation (CV)*

The CV provides a methodical approach to evaluating the dispersion of a probability or frequency distribution in statistics. Lower values of the coefficient of variation indicate higher stability and less fluctuation in the data. The coefficient of variation formula is provided below.

$$C_v = \frac{\sigma}{\mu} * 100 \tag{11}$$

Where C_v is the coefficien t of variation, σ denotes the variance and μ represents the mean value.

Time-series analysis

- *Autocorrelation coefficients*

The correlation similarity between data as a function of the time lag between them is represented by the autocorrelation function (ACF). It is a time domain metric for the stochastic process memory, not revealing anything about the frequency content of the process. For an error data e_t, the ACF is defined generally as follows,

$$\rho_k = \frac{Cov(e_t, e_{t+k})}{\sqrt{Var(e_t)Var(e_{t+k})}} \tag{12}$$

The preceding ACF formulation reduces to for a stable stochastic process with variance σ^2.

$$\rho_k = \frac{Cov(e_t, e_{t+k})}{\sigma^2} \tag{13}$$

In this instance, time is irrelevant. A white noise mechanism's autocorrelation function is zero for all lags except lag zero, for which a value of unity denotes total uncorrelatedness.

Feature Selection Using HHZO Algorithm

The input for the feature selection step is the features that have been retrieved using the various procedures. The HHZO algorithm, which combines the HHO and ZOA algorithms, is used to choose the best characteristics from among these features. The defense strategy of the HHO is incorporated within the defense property of the ZOA optimization to improve the performance of the zebra.

Initialization

Since zebras are a part of its population, HHZO is an optimizer based on populations. The plain where the zebras are located is the problem's search space, and each zebra stands for a potential mathematical solution. The placements of each zebra (search agent) inside the search zone define the values for the selection parameters. As a result, each zebra may be represented as a member of the HHZO by means of a vector, the values of which correspond to the variables under consideration. A matrix may be used to analytically model the zebra population. The zebras' starting places are chosen at random inside the search area. The HHZO population matrix is given by Eq. (14).

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,M} \end{bmatrix}_{N\times m} \tag{14}$$

Where X represents the zebra population, X_i is the i^{th} zebra, m is the number of option variables, N is the number of population members (zebras), and $x_{i,j}$ is the value for the j^{th} problem variable that the i^{th} zebra provided. Each zebra represents a possible fix for the optimization problem. As a result, using each search agent's suggested values for the issue variables, the objective function may be evaluated. Equation (15) is utilized to provide the values acquired for the goal function as a vector.

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N\times 1} \tag{15}$$

Where F_i is the fitness function value attained for the i^{th} search agent and F represents the vector of values for the objective function. The candidate solution that best suits the given problem is identified by comparing the values of the objective function. This process successfully evaluates the quality of the candidate solutions that are linked with the given problem. Members of the HHZO population are, therefore, updated in two distinct periods during each repetition.

- *Phase 1: Foraging behavior*

Population members are updated in the first phase using zebra behavior models during foraging searches. Within HHZO, the most exceptional individual within the population is recognized as the pioneer zebra, guiding the neighbor members of the population in the direction of its predetermined destination inside the search area. Therefore, it is possible to statistically model how zebras' locations vary throughout the foraging phase using Eqs. (16) and (17).

$$x_{i,j}^{new,P1} = x_{i,j} + r.(PZ_j - I.x_{i,j}) \tag{16}$$

$$X_i = \begin{cases} X_i^{new,P1}, & F_i^{new,P1} < F_i \\ X_i, & else \end{cases} \tag{17}$$

Where PZ is the top member and the pioneering zebra; PZ_j is its j^{th} dimension; r is a random number in the interval

[0, 1]; $I = round\ (1 + rand)$, where rand is a random value in the interval [0, 1]; and $x_{i,j}^{new,P1}$ is its j^{th} dimension value. Based on the first phase, this represents the new status of the i^{th} zebra. As a result, $I \in \{1, 2\}$ and if parameter $I = 2$, then the population movement varies significantly more.

- *Phase 2: Techniques for defending against predators*

The position of HHZO individuals in the search region is updated in the second phase through the simulation of zebra defense systems against predator assaults. Zebras have several defense strategies depending on the kind of predator they face. Zebras defend themselves against lion attacks by running in a zigzag manner and occasionally turning sideways.

It is expected in the HHZO design that one of the following two scenarios is occur with an equal chance:

- The zebra decides on an escape route after the lion assaults it.
- When another predator attacks the zebra, it is decided how to go on the offensive.

The first tactic involves the zebras attacking the lions and running away from the attack in the area around their current location. This strategy can be formally represented by the mode S_1 in Eq. (18). As a predator attacks a zebra, the herd's closer members go in that direction as part of the second technique, which aims to confuse and frighten the attacker by forming a protective structure. In mathematics, the mode S_2 in Eq. (18) represents the zebras' approach. If the objective function has a higher value at the new site after the zebras are relocated, then they accept the change. Using Eq. (19), this update condition is modeled.

$$x_{i,j}^{new,P2} = \begin{cases} S_1: -\, d_m^{Iter,AGE}\left[\left(\frac{1}{qN}\sum_{i=1}^{pN} X_i^{(Iter-1)}\right) - X^{(Iter-1)}\right], & P_s \leq 0.5 \\ S_2: x_{i,j} + r.\left(AZ_j - I.x_{ij}\right), & else \end{cases} \tag{18}$$

$$X_i = \begin{cases} X_i^{new,P2} \\ X_i, \quad else \end{cases}, F_i^{new,P2} < F_i \tag{19}$$

The other behavior used in HOA is called the horses' protection mechanism. It is characterized by the horses running from horses who show inadequate reactions. The defense system is described by Eq. (20) and Eq. (21).

Where

$$AGE = \alpha, \gamma, and \gamma \tag{20}$$

$$D_m^{Iter,AGE} = d_m^{(Iter-1),AGE} \times \omega_d \tag{21}$$

The escape vector of the i^{th} horse from the average of some horses with worst placements, which are illustrated by the X vector, is indicated by $D_m^{Iter,AGE}$ in the equations above. qN is the number of horses with the worst placements. It is advised to fix the value of q to 20% of the total number of horses. The reduction factor for dieter is represented by ω_d. AZ denotes the attacking zebra's status, whereas AZ_j represents its j^{th} dimension value. The objective function value is $F_i^{new,P2}$ and its j^{th} dimension value is $x_{i,j}^{new,P2}$.

WEARN-PQ -Based Detection

The WEARN-PQ model represents an ensembled approach to anomaly detection in power quality data within grid systems. Leveraging a weighted ensemble approach, it combines predictions from multiple sub-models, each with diverse architectures or configurations, to enhance overall detection performance. Different elements of the incoming data are constantly highlighted and weighted using attention techniques, enabling the model to concentrate on the time steps or attributes that are most essential. Residual connections within the neural network architecture facilitate effective information flow and alleviate the vanishing gradient problem, enabling the model to learn complex representations of the data. The integration of CNNs for spatial feature extraction, stacked GRUs with attention layers for modeling sequential dependencies, and Stack-RNNs for capturing hierarchical patterns ensures comprehensive analysis of both spatial and temporal aspects of the power quality measurements. This combined architecture enables the WEARN-PQ model to accurately detect anomalies in power quality data, contributing to the proactive management and maintenance of grid infrastructure. Figure 2 shows the comparison of accuracy for different k -folds.

A CNN equipped with an attention function and residual connection

To shorten the training period and lower the training parameters, it employs residual connections. Additionally, by taking into account only the most crucial data, present a hybrid attention technique that enhances the network focusing on those more crucial regions on the feature maps. To be more precise, w utilize inRA-Conv (outRA-Conv) to denote the location of a hybrid attention module within (between) the residual convolution. As a result, the classifiers are called classifier-inRAC and classifier-outRAC, respectively. Through the use of channel and spatial attention mechanisms, the cascaded hybrid attention module is able to compute complementary attention by concentrating on channel and attention, respectively. Let F in particular, represent the input 2D feature map of the channel attention sub-module. According to Eq. (22) the 1D output F is further processed by the spatial attention sub-module to create a more refined 2D feature map M.

$$M = f_s(f_c(F) \otimes F) \otimes F \tag{22}$$

Where f_C and f_S stand for the channel and spatial functions, respectively and are provided by Eq. (23) and Eq. (24). The element-wise multiplication is given in Eq. (23) and Eq. (24).

$$f_c(F) = \sigma\left(MLP\left(AvgPool(F)\right) + MLP\left(MaxPool(F)\right)\right) \tag{23}$$

$$f_s(F) = \sigma\left(f^{k \times k}\left(AvgPool(F)\right) + \left(MaxPool(F)\right)\right) \tag{24}$$

Where $f^{k \times k}$ is a convolutional function having a kernel size of $k \times k$, MLP is the multi-layer perceptron, AvgPool (MaxPool) is the average (max) pooling, and σ is the sigmoid function.

SoftMax layer

The SoftMax function is utilized by the network output nodes to determine the total number of unordered classes. Eq. (25) defines a SoftMax function.

$$f(x_j) = \frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}} \tag{25}$$

Where n is the number of output nodes, x_j is the input of the network for the j^{th} output node, and $f(x_j)$ is the score of the j^{th} output node. Actually, the total of all the output values $f(x)$ is 1, as each one represents a probability between 0 and 1.

SGRU modeling

A number of GRU units make up the SGRU. The input sequence $\{e_1, e_2, \dots, e_t\}$ for time sequence t initially enters hidden layers $\{h_1^1, h_2^1, \dots, h_t^1\}$ to gather all of the data from previous time steps. Subsequently, the outputs from the bottom hidden layers at an identical time step are used as input by the top hidden layers to extract further features. In particular, $\{h_1^2, h_2^2, \dots, h_t^2\}$ are the higher layers of the hidden layers. As can be observed in Eq. (29), the update gate, reset gate, and candidate value for each layer's hidden state h_t^i are found using Eq. (26), Eq. (27), and Eq. (28). The embedding vector e_t is placed in the first layer. Utilize the hidden state from the most recent time step in place of et in Eq. (26), Eq. (27), and Eq. (28), starting from the second layer upward in the past layer h_{t-1}^i.

$$u_t^i = \sigma(W_u^i h_{t-1}^i + U_u^i e_t + b_u^i \tag{26}$$

$$r_t^i = \sigma(W_r^i h_{t-1}^i + U_r^i e_t + b_r^i \tag{27}$$

$$\tilde{C}_t^i = \tanh(W_c^i . [r_t^i * h_{t-1}^i + U_c^i e_t + b_c^i \tag{28}$$

$$h_t^i = u_t^i * \tilde{C}_t^i + (1 - u_t^i) * h_{t-1}^i \tag{29}$$

SRNN network

An input layer, an output layer, and a hidden layer with a recurrent time-delayed link constitute an RNN. The transmission of information over time is made possible by the recurrent link. When given a series of tokens, an RNN predicts the probability y_t of the next symbol by using the one-hot encoding (x_t) of the current token as input. Additional details about the tokens viewed in the sequence prior to this one is stored in a hidden layer consisting of m units. More specifically, using the encoding x_t of the current token and its previous state h_{t-1}, the hidden layer ht is updated at each time t, as stated in the Eq. (30):

$$h_t = \sigma(Ux_t + Rh_{t-1} \tag{30}$$

Where R is the $m \times m$ matrix of recurrent weights, U is the $d \times m$ token embedding matrix, and $\sigma(x) = 1/(1 + exp(-x))$ is the sigmoid activation function applied coordinate wise. The network then outputs the probability vector y_t of the subsequent token based on the state of these hidden units, as shown by the Eq. (31):

$$y_t = f(Vh_t) \tag{31}$$

Where V is the $m \times d$ output matrix, d is the number of distinct tokens, and f is the softmax function. This design can pick up on quite complicated patterns that resemble those that N-grams are able to record. The RNNs are now fascinating for language modeling, but it's possible that they can't understand how algorithmic patterns are created. The ability to add an external memory to RNNs, which can theoretically learn basic algorithmic patterns.

- *Pushdown network*

The pushdown automata, or automaton that uses a stack, served as the model for the basic structured memory. The top element of a stack is the only way to access this kind of persistent memory. A stack is capable of three basic operations: NO-OP does nothing, PUSH adds a new element to the top of the stack, and POP removes the top element. Initially, examine a basic design in which the model is limited to selecting either a POP or a PUSH at every time step. It considers that a 2-dimensional variable at which the value of the hidden variable h_t determines this decision:

$$a_t = f(S\, h_t) \tag{32}$$

Where f is a SoftMax function and A is a $2 \times m$ matrix, where m is the hidden layer's size. It indicates the likelihood of the PUSH action by $a_t[PUSH]$, and the likelihood of the POP action by $a_t[POP]$. Assume that the stack is kept in a vector st of size p at time t. Keep in mind that p is not set and can be changed as needed, allowing the model's capacity to expand. Position 0 stores the top element, which has the value $s_t[0]$:

$$S_t[0] = a_t[PUSH]\sigma(Dh_t) + a_t[POP]s_{t-1}[1] \tag{33}$$

Where the matrix D is $1 \times m$. The value below takes the place of the top element if $a_t[POP] = 1$ (all values are shifted by one position up in the stack structure). All values in the stack are moved down and a value is added to the top if $a_t[PUSH]$ equals 1. Likewise, the following update rule for an element in the stack that is stored at a depth of $i > 0$:

$$s_t[i] = a_t[PUSH]s_{t-1}[i-1] + a_t[POP]s_{t-1}[i+1] \tag{34}$$

The information is transferred to the concealed layer at the subsequent time step using the stack. A stack that is empty has s_t set to –1. The concealed layer has been updated as:

$$h_t = \sigma\left(Ux_t + Rh_{t-1} + Ps_{t-1}^k\right) \tag{35}$$

Where s_{t-1}^k are the k top-most elements of the stack at time $t - 1$, and P is a $m \times k$ recurrent matrix ($k = 2$).

- *Stack with a no-operation*

By making a small modification to the stack update rule, the NO-OP action enables the stack to maintain the same value at the top. In lieu of Eq. (36):

$$s_t[0] = a_t[PUSH]\sigma(Dh_t) + a_t[POP]s_{t-1}[1] + a_t[NO-OP]s_{t-1}[0] \tag{36}$$

Ensemble model

Ensemble modeling is the process of integrating and weighing individual outcomes to make a final decision. Through the training of several independent classifiers and their combination to increase the model's overall predictive capacity, the deep learning models' accuracy has increased because to these tactics. They use a number of classifiers like CNN, SGRU and SRNN and combine them by averaging the weights of the different classifiers. When the model is subjected to this process, its overall accuracy is higher than when it is subjected to a single classifier.

Result and Discussion

In this section, the results obtained for the proposed power quality management is discussed with the existing methods. The historical data is used in this work for predicting the power quality. The total data is divided into training (70%) and testing (30%). The implementation is performed using the Python platform. The performance metrics like accuracy, precision, recall, F1-score, recall, MCC, MAE, MSE and RMSE are evaluated for the comparison.

Overall Comparison of the Proposed Model by Varying the K-fold

The original sample is divided into k subsamples at random in k-fold cross-validation. For the classifier's test, one subsample is kept as validation data, while the rest k – 1 are utilized as training data. The cross-validation procedure is then carried out k times, using one instance of the test data for each of the k subsamples. After that, an average of the k fold findings is obtained to generate a single performance estimate.

The Table 1 compares the performance of different machine learning techniques – RNN, LSTM, Fuzzy Logic, ANN and the proposed method – on several k values. The proposed method outperforms other methods in terms of highest precision (0.9857), precision (0.9856), recall (0.9856), F1 score (0.9856), R2 (0.9623) and MCC (0.9809), indicating overall correctness, validity and balance, efficiency. It also achieves the lowest errors, with MAE (0.0243), MSE (0.0483), and RMSE (0.2198) showing the lowest mean and squared prediction errors. This means that the proposed technique is the most effective and reliable for the given data and task.

The Table 2 compares the performance of different machine learning techniques – RNN, LSTM, fuzzy logic, ANN and the proposed method – on several k=2 metrics. With the best precision (0.9923), precision (0.9924), recall (0.9923), and F1-score (0.9923), the proposed strategy consistently outperforms others. Additionally, it has the largest MCC (0.9898) and R2 (0.9815), which means it can generate accurate binary rankings. Moreover, the proposed approach is characterized by the smallest mean and squared prediction errors, as evidenced by the MAE (0.0123), MSE

Table 1: Comparison of the proposed and existing techniques for k=1

Performance metrics	*RNN*	*LSTM*	*Fuzzy logic (Garikapati et al., 2023)*	*ANN (Ramadevi et al., 2023)*	*Proposed*
Accuracy	0.9690	0.9597	0.9503	0.9753	0.9857
Precision	0.9690	0.9595	0.9503	0.9753	0.9856
Recall	0.9689	0.9599	0.9501	0.9754	0.9856
F1	0.9690	0.9596	0.9502	0.9754	0.9856
R2	0.9180	0.8796	0.8628	0.9401	0.9623
MCC	0.9587	0.9462	0.9338	0.9671	0.9809
MAE	0.0527	0.0707	0.0837	0.0397	0.0243
MSE	0.1053	0.1467	0.1703	0.0743	0.0483
RMSE	0.3246	0.3830	0.4127	0.2726	0.2198

Table 2: Comparison of the proposed and existing techniques for k=2

Performance metrics	*RNN*	*LSTM*	*Fuzzy logic (Garikapati et al., 2023)*	*ANN (Ramadevi et al., 2023)*	*Proposed*
Accuracy	0.9757	0.9687	0.9353	0.9820	0.9923
Precision	0.9756	0.9688	0.9353	0.9819	0.9924
Recall	0.9757	0.9686	0.9353	0.9820	0.9923
F1	0.9756	0.9687	0.9353	0.9820	0.9923
R2	0.9370	0.8960	0.8205	0.9440	0.9815
MCC	0.9675	0.9582	0.9138	0.9760	0.9898
MAE	0.0390	0.0590	0.1107	0.0320	0.0123
MSE	0.0763	0.1297	0.2247	0.0687	0.0237
RMSE	0.2763	0.3601	0.4740	0.2620	0.1538

Table 3: Comparison of the proposed and existing techniques for k=3

Performance metrics	*RNN*	*LSTM*	*Fuzzy logic (Garikapati et al., 2023)*	*ANN (Ramadevi et al., 2023)*	*Proposed*
Accuracy	0.9637	0.9643	0.9420	0.9773	0.9900
Precision	0.9637	0.9643	0.9420	0.9774	0.9900
Recall	0.9635	0.9644	0.9419	0.9774	0.9900
F1	0.9636	0.9643	0.9419	0.9774	0.9900
R2	0.9115	0.9026	0.8424	0.9357	0.9682
MCC	0.9516	0.9524	0.9226	0.9698	0.9867
MAE	0.0580	0.0603	0.0973	0.0397	0.0183
MSE	0.1107	0.1223	0.1960	0.0823	0.0397
RMSE	0.3327	0.3498	0.4427	0.2869	0.1992

(0.0237) and RMSE (0.1538) values. This means that after considering k=2, the proposed method outperforms RNN, LSTM, fuzzy logic and ANN in terms of efficiency and reliability.

Table 3 uses different performance metrics to compare the proposed method with existing methods (RNN, LSTM, fuzzy logic and ANN) for k=3. With the highest values of precision, precision, recall, F1-score (0.9900), R-squared (0.9682), and Matthews correlation coefficient (0.9867), the proposed method outperforms all other methods in terms of performance. It also has the fewest errors, with an RMSE of 0.1992, an MSE of 0.0397, and an MAE of 0.0183.

Accuracy

The proposed method outperforms RNN (0.9690), LSTM (0.9597), fuzzy logic (0.9503) and ANN (0.9753) with a maximum accuracy of 0.9857 for k=1. Similarly, for k=2, the proposed approach achieves a maximum accuracy of 0.9923, out performing ANN (0.9820), RNN (0.9757), LSTM (0.9687), and fuzzy logic (0.9353). Moreover, with a maximum accuracy of 0.9900 for k=3, the proposed method remains superior to RNN (0.9637), LSTM (0.9643), fuzzy logic (0.9420), and ANN (0.9773) (Figure 1).

Precision

The proposed method outperforms RNN (0.9690), LSTM (0.9595), fuzzy logic (0.9503) and ANN (0.9753) at k=1. The maximum accuracy is 0.9856. The proposed approach outperforms RNN (0.9756), LSTM (0.9688), fuzzy logic (0.9353) and ANN (0.9819) with the highest accuracy of 0.9924 for k = 2. In the same vein, the proposed method further sequence outperforms RNN (0.9637), LSTM (0.9643), fuzzy logic (0.9420) and ANN (0.9774) at k = 3. Figure 3 shows the comparison of precision for different k-folds.

F1-score

When k = 1, the proposed technique achieves the highest F1 score of 0.9856, outperforming RNN (0.9690), LSTM (0.9596), fuzzy logic (0.9502), and ANN (0.9754). For k=2, the proposed method again achieves the highest F1 score of 0.9923, outperforming RNN (0.9756), LSTM (0.9687), fuzzy logic (0.9353), and ANN (0.9820). Similarly, with k = 3, the proposed technique maintains an excellent F1 score of 0.9900, outperforming RNN (0.9636), LSTM (0.9643), fuzzy logic (0.9419), and ANN (0.9774). Figure 4 shows the comparison of F-measure for different k- folds.

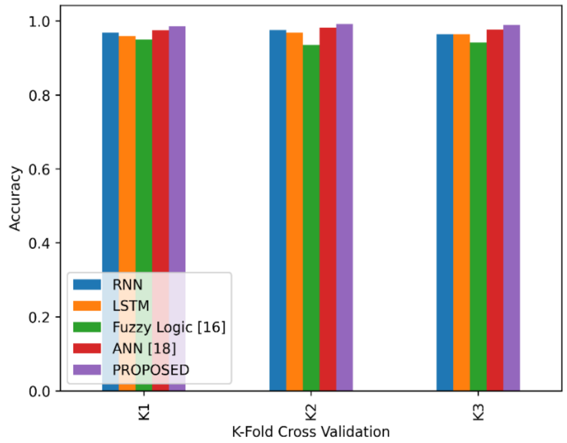

Figure 2: Comparison of the accuracy for different K- folds

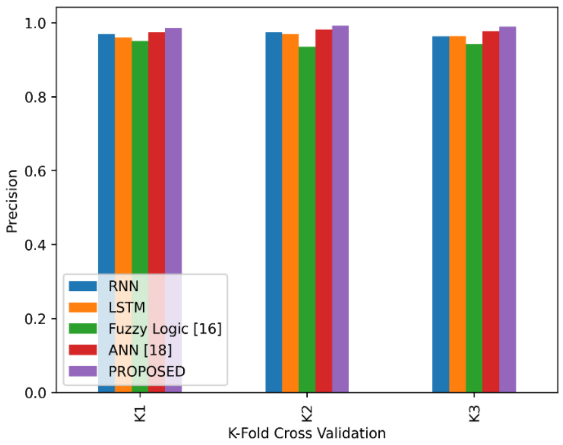

Figure 3: Comparison of the precision for different K- folds

R2 error

Comparing the R2 error measure in Tables 1, 2, and 3, for k=1, the proposed technique achieves the highest R2 value of 0.9623, outperforming RNN (0.9180), LSTM (0.8796), fuzzy logic (0.8628) and ANN (0.9401). For *k*=2, the proposed method again leads with the highest *R*2 value of 0.9815, surpassing RNN (0.9370), LSTM (0.8960), fuzzy logic (0.8205), and ANN (0.9440). Similarly, with *k*=3, the proposed technique maintains its superior *R*2 value at 0.9682, exceeding RNN (0.9115), LSTM (0.9026), fuzzy logic (0.8424), and ANN (0.9357). Figure 5 shows the MAE for different K- folds.

MCC

For *k*=1, the proposed technique achieves the highest MCC of 0.9809, outperforming RNN (0.9587), LSTM (0.9462), fuzzy logic (0.9338), and ANN (0.9671). For *k*=2, the proposed method again leads with the highest MCC of 0.9898, surpassing RNN (0.9675), LSTM (0.9582), fuzzy logic (0.9138), and ANN (0.9760). Similarly, with *k*=3, the proposed technique maintains its superior MCC at 0.9867, exceeding RNN (0.9516), LSTM (0.9524), fuzzy logic (0.9226), and ANN (0.9698). Figure 6 shows the MCC for different K-folds.

MAE

When *k*=1, the proposed technique achieves the lowest MAE of 0.0243, outperforming RNN (0.0527), LSTM (0.0707), fuzzy logic (0.0837), and ANN (0.0397). For *k*=2, the proposed method again leads with the lowest MAE of 0.0123, surpassing RNN (0.0390), LSTM (0.0590), fuzzy logic (0.1107), and ANN (0.0320). Similarly, with *k*=3, the proposed technique maintains its superior MAE at 0.0183, exceeding RNN (0.0580), LSTM (0.0603), fuzzy logic (0.0973), and ANN (0.0397). Figure 7 shows the comparison of R2 for different k-folds.

MSE

When *k*=1, the proposed technique achieves the lowest MSE of 0.0483, outperforming RNN (0.1053), LSTM (0.1467), fuzzy logic (0.1703), and ANN (0.0743). For *k*=2, the proposed method again leads with the lowest MSE of 0.0237, surpassing RNN (0.0763), LSTM (0.1297), Fuzzy Logic (0.2247), and ANN (0.0687). Similarly, with *k*=3, the proposed technique maintains its superior MSE at 0.0397, exceeding RNN (0.1107), LSTM (0.1223), fuzzy logic (0.1960), and ANN (0.0823). Figure 8 shows the comparison of recall for different K-folds.

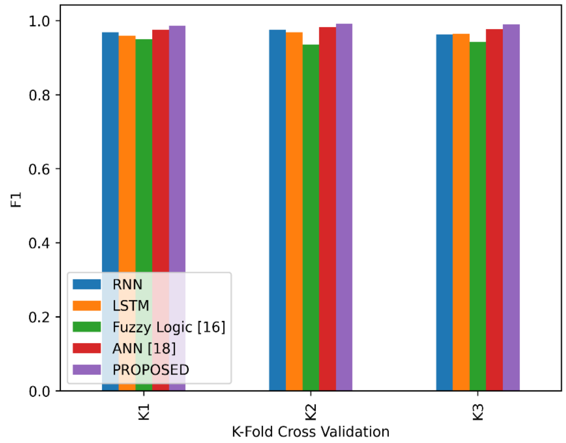

Figure 4: Comparison of the F-measure for different K- folds

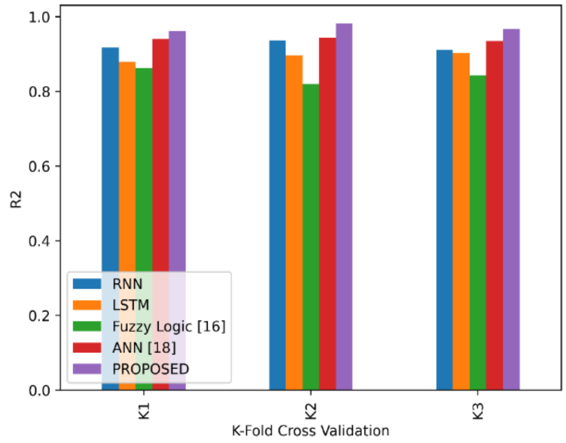

Figure 5: Comparison of the MAE for different K- folds

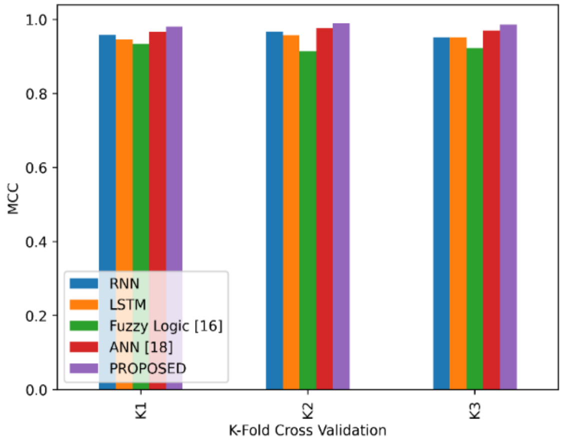

Figure 6: Comparison of the MCC for different K- folds

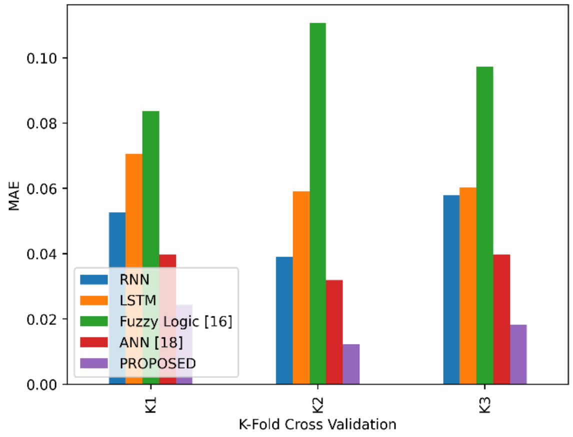

Figure 7: Comparison of the R2 for different K- folds

RMSE

When *k*=1, the proposed technique achieves the lowest RMSE of 0.2198, outperforming RNN (0.3246), LSTM (0.3830), fuzzy logic (0.4127), and ANN (0.2726). For *k*=2, the proposed method again leads with the lowest RMSE of 0.1538, surpassing RNN (0.2763), LSTM (0.3601), fuzzy logic (0.4740), and ANN (0.2620). Similarly, with *k*=3, the proposed technique maintains its superior RMSE at 0.1992, exceeding RNN (0.3327), LSTM (0.3498), fuzzy logic (0.4427), and ANN (0.2869). The comparison of RMSE values is displayed in Figure 9.

The AUC curve of the proposed model is shown in Figure 10.

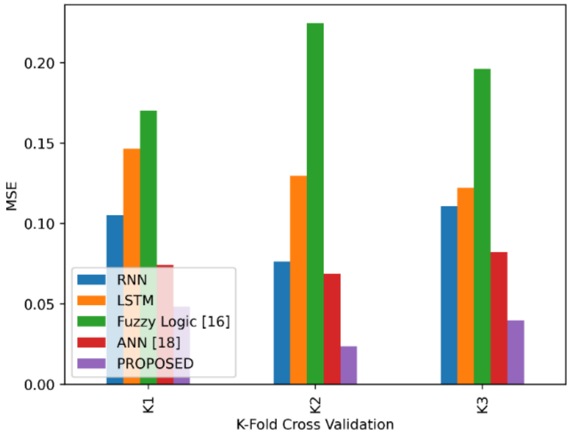

Figure 8: Comparison of the recall for different K- folds

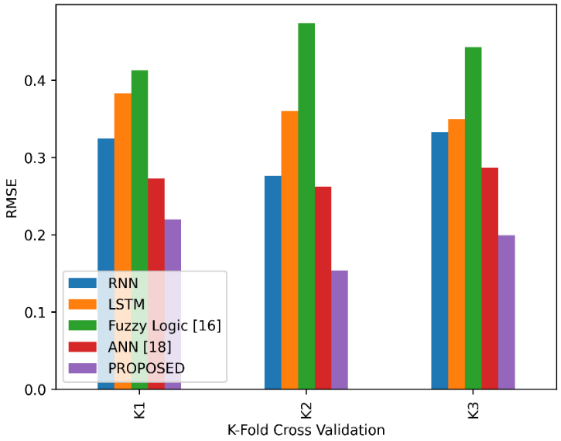

Figure 9: Comparison of the RMSE values

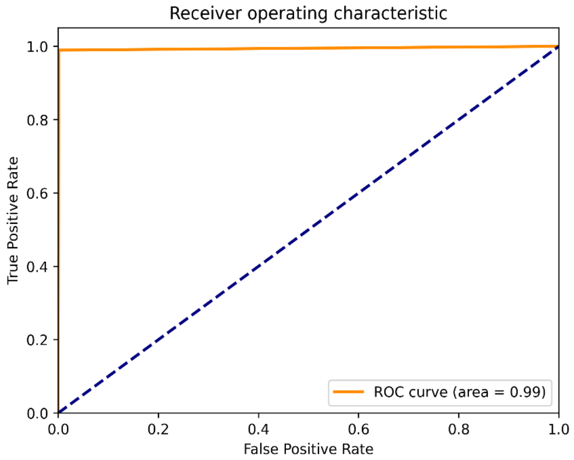

Figure 10: AUC curve of the proposed model

Conclusion

In conclusion, the proposed methodology presents a holistic approach to address the intricate challenges of real-time data analysis and anomaly detection within complex grid systems. By integrating smart grid sensors and IoT devices and employing advanced pre-processing techniques, the methodology ensures the integrity and completeness of the data. Feature extraction and a novel hybrid optimization approach facilitate the extraction of meaningful insights and the selection of relevant features. The WEARN-PQ architecture, with its integration of various deep learning techniques, offers a powerful framework for anomaly detection. Through rigorous training and validation using K-Fold cross-validation with stratified sampling, the methodology ensures the robustness and reliability of the models. Overall, this methodology provides a comprehensive solution to enhance the efficiency and reliability of grid systems through proactive anomaly detection and timely intervention.

Acknowledgments

The authors would like to thank the Deanship of Bharath Institute of Higher Education and Research for supporting this work.

References

Afzal, M.Z., Aurangzeb, M., Iqbal, S., Rehman, A.U., Kotb, H., AboRas, K.M., Elgamli, E. and Shouran, M., 2022. A resilience-oriented bidirectional anfis framework for networked microgrid management. Processes, 10(12), p.2724.

Akpolat, A.N., Dursun, E. and Yang, Y., 2023. Performance Analysis of a PEMFC-Based Grid-Connected Distributed Generation System. Applied Sciences, 13(6), p.3521.

Ayalew, M., Khan, B., Giday, I., Mahela, O.P., Khosravy, M., Gupta, N. and Senjyu, T., 2022. Integration of renewable based distributed generation for distribution network expansion planning. Energies, 15(4), p.1378.

Chandrakala Devi, S., Singh, B. and Devassy, S., 2020. Modified generalised integrator-based control strategy for solar PV fed UPQC enabling power quality improvement. IET Generation, Transmission & Distribution, 14(16), pp.3127-3138.

Eristi, B. and Eristi, H., 2022. Classification of power quality disturbances in solar PV integrated power system based on a hybrid deep learning approach. International Transactions on Electrical Energy Systems, 2022.

Garikapati, R., Kumar, S.R. and Karthik, N., 2023. ANFIS Controlled MMC-UPQC to Mitigate Power Quality Problems in Solar PV Integrated Power System. Journal of Advanced Research in Applied Sciences and Engineering Technology, 36(1), pp.102-130.

Garikapati, R., Kumar, S.R. and Karthik, N., 2023. Power Quality Improvement in Solar Integrated Power Systems Using Fuzzy-Based MMC UPQC. International Journal of Intelligent Systems and Applications in Engineering, 11(6s), pp.120-131.

Khalkhali, H., Oshnoei, A. and Anvari-Moghaddam, A., 2022. Proportional Hysteresis Band Control for DC Voltage Stability of Three-Phase Single-Stage PV Systems. Electronics, 11(3), p.452.

Krishna, D., Sasikala, M. and Ganesh, V., 2022. Adaptive FLC-based UPQC in distribution power systems for power quality problems. International Journal of Ambient Energy, 43(1), pp.1719-1729.

Kumar, C., Ghosh, P. and Chatterjee, S., 2022. Enhancement of Power Quality by mitigating of sag and swell problem in power system using DVR. IFAC-PapersOnLine, 55(1), pp.131-137.

Mohamed, E.L.B.A.R., 2022. Power Quality Enhancement Techniques in Smart Grid (Doctoral dissertation, Djelfa University).

Nandagopal, V., Damodhar, T.S., Vijayapriya, P. and Thamilmaran, A., 2023. Improving Power Quality by DSTATCOM Based DQ Theory with Soft Computing Techniques. Intelligent Automation & Soft Computing, 36(2).

Rajagopalan, A., Swaminathan, D., Alharbi, M., Sengan, S., Montoya, O.D., El-Shafai, W., Fouda, M.M. and Aly, M.H., 2022. Modernized planning of smart grid based on distributed power generations and energy storage systems using soft computing methods. Energies, 15(23), p.8889.

Ramadevi, A., Srilakshmi, K., Balachandran, P.K., Colak, I., Dhanamjayulu, C. and Khan, B., 2023. Optimal design and performance investigation of artificial neural network controller for solar-and battery-connected unified power quality conditioner. International Journal of Energy Research, 2023.

Rao, J.V.R., Balasubramanian, M., Gangabhavani, C.S., Devi, M.A. and Devi, K.D., 2023. A TLBO algorithm-based optimal sizing in a standalone hybrid renewable energy system. *The Scientific Temper*, *14*(03), pp.629-636.

Ratnakaran, R., Rajagopalan, G.B. and Fathima, A., 2023. Artificial ecosystem optimized neural network controlled unified power quality conditioner for microgrid application. Energy Informatics, 6(1), p.45.

Ravi, T. and Kumar, K.S., 2023. Detection and classification of power quality disturbances using stock well transform and improved grey wolf optimization-based kernel extreme learning machine. IEEE Access.

Sarita, N.S., Reddy, S.S. and Sujatha, P., 2023. Control strategies for power quality enrichment in Distribution network using UPQC. Materials Today: Proceedings, 80, pp.2872-2882.

Sekhar, D.C., 2022. Hybrid UPQC for Power Quality Management of Renewable Energy Penetrated System. International Transactions on Electrical Engineering and Computer Science, 1(2), pp.75-103.

Tejakrishna, D. and Ramu, M., 2022. An Innovative fuzzy logic controller for power quality Optimization using MLI-UPQC with solar PV-BESS.

Thentral, T.M., Palanisamy, R., Usha, S., Vishnuram, P., Bajaj, M., Sharma, N.K., Khan, B. and Kamel, S., 2022. The improved unified power quality conditioner with the modular multilevel converter for power quality improvement. International Transactions on Electrical Energy Systems, 2022.

Thentral, T.T., Palanisamy, R., Usha, S., Bajaj, M., Zawbaa, H.M. and Kamel, S., 2022. Analysis of Power Quality issues of different types of household applications. Energy Reports, 8, pp.5370-5386.

Tripathi, S. and Pachori, A., 2023. Fault Effect Analysis and Frequency Deviation Detection in Smart Solar Connected Grid.