Applied Sciences Research Periodicals
August 2023, Vol. 1, No. 5, pp. 15-39

Power Quality Engineering Evaluation and Generalization of Deep Learning

Ifeoluwa Adeloye, Benjamen Agbo, and Victor Bamboleo

1. Federal School of Engineering Research Institute

Abstract:
This paper aims to introduce deep learning to the power quality community by reviewing the latest applications and discussing the open challenges of this technology. Publications covering deep learning to power quality are stratified in terms of application, type of data, and learning technique. This work shows that the majority of the deep learning applications to power quality are based on unrealistic synthetic data and supervised learning without proper labelling. Some applications with deep learning have already been solved by previous machine learning methods or expert systems. The main barriers to implementing deep learning to power quality are related to lack of novelty, low transparency of the deep learning methods, and lack of benchmark databases. This work also discusses that even with automatic feature extraction by deep learning methods, power quality expert knowledge is still needed to implement and analyses the results. The main research gaps identified in this work are related to the applications of semi-supervised learning, explainable deep learning and hybrid approaches combining deep learning with expert systems. Providing a stronger level of collaboration between grid stakeholders and academia to monitor power quality events, properly labeling and enlarging datasets for deep learning methods, outlining the end-to-end decision-making of deep learning methods, and offering open-access databases for comparison purposes are some suggestions for overcoming the current limitations.

Keywords: Power quality, Deep learning, Data analysis

INTRODUCTION

The electric power sector is continuously modernizing to become more environmentally friendly, economical, and reliable. The modernization goals are reached by, for instance, integrating renewable energy, installing new devices in both supply and demand ("smart grid equipment"), the deregulation of the sector, and the advancement in measuring infrastructures. However, the ongoing increase of equipment based on power electronics has an impact on the probability of interference through changes in emission, immunity, and transfer of power quality (PQ) disturbances [1,2]. Many of the power grid stakeholders perform continuous PQ monitoring to obtain information on the supply and equipment performance [3,4]. Long-term PQ measurements result in a large amount of data. For instance, [5] shows that a two-year PQ measurement campaign at five different locations resulted in 250 GB of data. The true value of PQ monitoring depends on its ability to analyses and interpret voluminous raw data obtained from that monitoring. Manual analysis of this data type is possible; however, it is time consuming. Proper analytical tools are needed to accelerate the "PQ big data" interpretation process.

The term "PQ big data" refers to a large amount of data resulting from continuous PQ monitoring. The term "big data" itself is not very clearly defined [6]. For instance, in terms of size, big data for the internet is measured in terms of exabytes 1018 and zettabytes 1021 [7] while PQ data is in MB or GB [6]. However, the consensus definition is that big data is a massive amount of data with specific complexities, the so-called 4V's: Volume, Variety, Velocity, and Veracity [8]. Although PQ big data is smaller than internet big database, both types of data are complex and difficult to process by employing traditional methods.

Artificial intelligence (AI) has been applied in many fields to handle analytics in big data. The first applications of AI were based on expert systems making decisions through rules defined by human expertise [9]. However, since the 1970s, a new subset of AI called machine learning (ML) has made computers capable of learning without explicit programming [10]. By employing ML to analyse and learn large amounts of existing data, computers can find patterns, predictions, and judgments to assist humans in making decisions [11].

Driven by the vast improvements in computer processing, a subset of ML based on artificial neural networks (ANNs) has been developed to tackle increasingly complex problems without human intervention. The initial ANNs still required significant human involvement for selecting and defining suitable features. However, the so-called deep learning (DL) applications can automatically extract optimal features from raw data [8]. DL approaches are usually implemented in pattern recognition systems due to the capability of DL to extract optimal information from high-dimensional data. DL methods have shown very promising results, especially in computer vision and image analysis, and performance is shown to be comparable to or even surpassing the conventional ML methods that use handcrafted features defined by human experts [8,12]. Most of the review papers on this subject only cover signal processing and AI techniques to classify and identify PQ disturbances [13–16]. Moreover, most of the existing review papers were published before DL popularity, as in [15] or, even later, they do not fully address DL methods [16,17]. This paper is the first to consider a comparative overview of the literature applying DL to PQ data. The paper also introduces big data from a PQ perspective and used the information from the literature to propose a DL workflow for PQ. After this introduction, the basic concepts concerning artificial intelligence and big data are introduced in Section 2. Section 3 distinguishes between the informatics-related terms used in this paper: AI, ML, and DL. Both Sections 2 and 3 aim to provide the readers from the PQ community with the terminology used in the AI community. Section 4 provides a workflow for applying DL to PQ. Section 4 also describes briefly the most common DL methods and their applications to PQ data. Section 5 provides a literature overview of AI techniques applied for the processing of PQ data, emphasizing the present applications with DL. Section 6 presents a critical discussion and recommendations regarding applications of DL to PQ. Finally, Section 7 concludes the paper.

POWER QUALITY BIG DATA

Concepts of Big Data

The term big data was first used by John Mashey [18] to refer to handling and analysing massive datasets. The concept of big data gained strength in the early 2000s when Doug Laney [19] defined big data by the "3V's": volume, velocity, and variety. By this definition, "big data" refers to a large amount of data that increases fast and is difficult or even impossible to handle by traditional methods. The most common way of defining big data nowadays is by the "4V's" which adds veracity to the "3V's" [20]. The list below details each of the "4V's" based on [21] and [22]:

a. Volume: refers to the amount, size, and scale of the data. The amount of data reaches such a level that it cannot be managed without dedicated analytic tools. The size can be defined either vertically by the number of samples in a dataset or horizontally by the number of features.

b. Velocity: refers to the speed by which the data is generated and how fast the data should be processed.
c. Variety: refers to the heterogeneity of the data. Big data often comes from different sources, which can be diverse in type, format, semantics, and volume.
d. (d)Veracity: refers to the quality of the collected data. It is related to biases, noise, and abnormality in data. The accuracy of any analytic process applied to the data depends greatly on the veracity of the source data.

Power Quality Big Data

The raw data of PQ monitoring consists of voltage and current samples in the time domain. The PQ monitors pre-process such raw data to detect and extract events such as voltage dips and transients. Besides, the raw data is pre-processed by mathematical transformations to obtain indexes to characterize waveform distortion, deviations from the ideal voltage magnitude, and other variations. Both raw and pre-processed PQ data can be considered big data because it contains the 4V's complexities. The following details each of the 4V's of PQ data.

Volume:

PQ monitoring results in a large amount of data. For instance, waveform measurements contain 256 samples per cycle, considering a sampling frequency of 12.8 kHz in 50 Hz. For each hour of measurement, this results in 1 080 000 samples (3 voltages and 3 currents). This volume complexity also holds for the pre-processed data. For example, harmonic values are obtained by Fourier transform and are aggregated every 10 min to simplify the data analysis. However, even with the aggregated values, one year of monitoring results in about 31 million data points per location considering 39 harmonics and 40 inter- harmonics (3 voltages, 3 currents, 10 min values).

Velocity:

The velocity of data is related to the set sampling frequency in the PQ monitor. Considering the previous example of 256 samples per cycle, a new sample in every voltage and current channel is obtained every 78.125 µs. The velocity complexity is also related to the PQ data analysis. streams: offline and online analysis. Offline analyses are mainly suitable for system performance evaluation, problem characterisation, and systems diagnosis. For online analysis, results are convenient when actions must be taken immediately.

Variety:

Even though the raw data comes from only two signals (voltage and current), it is diverse in terms of measurement sources. PQ monitors are installed in distinct voltage levels and locations of a power system to obtain enough data for understanding and characterizing PQ phenomena. Besides, PQ data is collected from many other monitoring devices on the system (intelligent relays, smart meters, digital fault recorders, phasor measurement units, etc.) [23]. In addition to the multiple data sources, the pre-processed PQ data is heterogeneous. Depending on the monitor settings, the pre-processed data can contain, for example, har- monic values and THD obtained each 10 min, voltage dips and transient waveforms with 256 samples per cycle, and rms voltage over 150 cycles of the power-system frequency. Although there are standardized methods to extract the data, as IEC 61000-4-30 [24], PQ monitors can also be configured with different settings, which results in data with different sampling, formats, and sizes. To reduce this complexity, the IEEE 1159.3 PQDIF Task Force has developed a standard format called Power Quality Data Interchange Format (PQDIF) [25].

Veracity:
According to the standard IEC 61000-4-30 [24], PQ measurement devices must comply with specific accuracy requirements (class A requirements). For harmonics, accuracy requirements are defined in IEC 61000-4-7 [26]. This accuracy is influenced by many factors, which include external sensors (e.g., instrument transformers, Rogowski coils), A/D converters, or measurement algorithms (e.g., aliasing, leakage) [27]. The requirements in [24] for PQ measurements point out the veracity complexity of PQ data.

Big Data Analytics
Big data analytics is the process of extracting information and detecting patterns from datasets with the 4V's complexity [22]. Conventionally, statistical methods, data mining, and visualization techniques were the most used tools for big data analytics. Recently, ML methods (artificial intelligence) have gained attention in analyzing big data. In the context of big data analytics, the exploration of time-varying data remains a key challenge in informatics [28]. Time-varying data is defined as spatiotemporal volumetric data, which means that each variable exhibits different values at particular time intervals. Several methods have been applied to extract patterns of time-varying data, such as visualization tools, clustering, and feature extraction [28].

PQ Monitoring and Analytics of PQ Big Data
PQ monitoring is the process of gathering, analyzing, and interpreting voltage and current measurements into useful information [52,53]. The continuous PQ monitoring allows the network operators to obtain information about the performance of utilities and customer facilities [54]. Moreover, the analysis of continuous measurements from PQ monitoring allows researchers to obtain knowledge of the PQ phenomena. Several tools are available for reducing large data volumes from PQ monitoring; this is mainly done through indices and reporting formats, like the ones defined in IEC 61000-4-30 [24] and the recommendations by CIGRE C4.112 [29]. The IEC document defines 10 min values, for among others, harmonics and inter harmonics. With many monitors and several years of data, this could still result in large amounts of data.

Reports employing classical statistical techniques in 10 min values might hide important information about variations with time. In some cases, shorter time scales should be employed instead of the 10 min values [30]. For instance, 10 min values of the harmonic voltage might not be proper for PV installations because the distortion varies over shorter time scales, depending, among others, on the solar irradiance and dynamic changes caused by fast cloud passages [45–48]. A more recent example is the charging of electric buses, where the charging is rarely longer than 10 min. The resulting data can be even larger if shorter time scales are employed instead of the 10 min values. In some cases, shorter time scales should be considered to proper addressing PQ. Analysis of PQ variations over time is essential to identify, for instance, excessive waveform distortion on power systems as a function of load and system characteristics (e.g., resonance conditions) [23]. Many approaches have been proposed to analyse and visualize time-varying events, such as Kalman filter [31], Parseval's [32], Hilbert–Huang Method [33], S-transform [34], and Wavelets [35]. An overview of such methods can be found in [36, 37]. In addition, graphical methods [38], expert systems [39], ML, and DL [40–43] have been applied to PQ measurements to extract additional information and provide better visualization of the raw data. One of the remaining challenges is detecting the dynamic changes and selecting important time intervals on a large amount of PQ measurements.

ARTIFICIAL INTELLIGENCE: EXPERT SYSTEMS, MACHINE LEARNING, AND DEEP LEARNING

The term AI refers to the entire universe of computing technology that exhibits anything resembling human intelligence [44]. In other words, AI is the enterprise of constructing an intelligent artefact. The most accepted definition of AI is one by Alan Turing, a computer can be said to possess AI if it can mimic human responses under specific conditions [45]. Since 1956 [46], AI has been considered an interdisciplinary subject to meet human intelligence. The period between the 1950s and 1960s is known as the first AI wave when many theoretical developments were made. However, the computers were not mature enough to process large neural networks, and not enough data was available for developing purposes. The following period is known as the first winter of AI. A second wave started in the 1980s with the development of expert systems that were based on rules defined by human experts. AI faced the second winter in the late 1980s and early 1990s mainly because expert systems required specialized hardware, and commercial vendors failed to provide an ample assortment of applications. In the same period, the development of the backpropagation algorithm played an important role in the progress of training multilayer neural networks [47]. A major milestone of AI is related to IBM's Deep Blue in 1997 [48], an AI system designed specifically for playing chess. Deep Blue was the first machine to beat a world champion in chess. The development of Deep Blue inspired researchers to create AI approaches that could tackle other complex problems. Since then, AI has faced a new wave associated with three main factors: availability of large amounts of data, advancements in computer processing, and massive investment from the industry [49]. AI can be grouped into six main categories [11]: game theory, decision-making algorithms, statistical models, search/optimization methods, expert systems, and learning methods (ML and DL). AI also covers or shows overlap with other fields such as robotics, sentimental analysis, and artificial emotion. In this work, we limit ourselves to those AI algorithms that are or can be applied to PQ. To this point, Fig. 1 illustrates these AI techniques for PQ, showing that ML is a subset of AI while DL is a subset of ML. The first applications of AI were based on expert systems which make decisions through rules based on expert knowledge [9]. On the other hand, ML and its subset DL have been made computers capable of learning without explicit programming [10].

Expert Systems

Expert systems are a type of AI that relies on expert knowledge and an inference engine. The basic idea behind expert systems is that the expert knowledge is transferred to a computer program through an inference engine. The inference engine is a set of "if-then" rules: if some condition is true, then a specific inference can be made, or action can be taken [50]. Due to the inference engine, expert systems are also called rule-based systems. The rules can be based on Boolean logic or Fuzzy logic. In Boolean logic, the rules incorporate only two values: "0" or "1". Fuzzy logic presents a gradual transition from "0" to "1" by expressing a set of values between the two logic states [51]. The main benefit of expert systems over ML is the explanation facility. A decision taken from an expert system can be explainable through the rules. The main limitations of expert systems are in creating inference rules; experts do not always agree, they are not always able to explain their reasoning, and some rules may be difficult to implement in terms of computational costs. Section 5.1 lists the applications that have employed expert systems in PQ field.

Machine Learning

ML using a large amount of data is a small subset within AI [10]. Conventional ML methods rely on human expertise to design the best features from the data for various tasks, such as classification, prediction, and regression. The general idea behind most ML applications is that a computer learns to perform a task by studying a training set of examples. Two main strategies can be employed in ML: supervised and unsupervised learning. Two other categories can also be incorporated into ML: semi-supervised learning and reinforcement learning.

Supervised Learning:
In supervised learning, the training set contains data and the correct output of the task with that data [10]. Supervised learning can be employed by logic regression, support vector machines, classification trees, random forests, artificial neural networks (ANNs), among others.

Unsupervised Learning:
The training set in unsupervised learning contains data but not the outputs, which means that the computer must find the solutions independently [10]. Unsupervised learning includes clustering algorithms as k-means, dimensional reduction techniques as principal components analysis, and ANNs.

Semi-Supervised Learning:
Semi-supervised learning is a type of ML that is between supervised and unsupervised. Its training combines a small amount of labelled data with a large amount of unlabeled data. To this point, semi-supervised learning is convenient when the ground-truth labels are available only for part of the data. Furthermore, the combination of labelled and Artificial intelligence groups and examples of some techniques.

Reinforcement Learning
Reinforcement learning is a type of ML that can be trained by interacting with a real-time environment. A solution is found by the computer employing trial and error to a given problem. The learning system is referred to as an agent, which can observe the environment, select and perform actions, and get rewards or penalties in return. The agent's objective is to maximize the total reward, and the best action is called policy [52]. This type of ML has been applied to robotics, game theory, and data science [53].

Deep Learning
Driven by the huge improvements in computer processing, a subset of ML-based on ANNs has been developed to tackle evermore-complex problems without human intervention. These so-called DL applications can perform a specific task by automatically extracting essential features from raw data [8]. A deep learning method is composed of a multilayer stack of simple models that maps non-linearly its output by its input [54]. Each layer further refines the previous layer's outputs that enables to increase both the selectivity and the invariance of the representation [54]. The mapping is learned from the input data by adapting the weights of each neuron by using an algorithm called backpropagation [63]. The weights are adapted based on the gradient of the error at the output. The backpropagation algorithm calculates the gradient of the error and distributed it back in terms of weights to the previous layers. The objective is to minimize the error in the output, and this weight adapting process is repeated over many iterations.

DL approaches are usually implemented in pattern recognition systems due to the power of DL to extract abstract concepts from high- dimensional data. Section 4.4 provides an overview of the main DL architectures used in PQ.

DEEP LEARNING WORKFLOW FOR POWER QUALITY

This session summarizes the steps in applying DL to a PQ problem. The workflow provided in this session should not be considered a strict practice. Instead, this workflow aims to give a guideline for the useful implementation of DL for PQ. The workflow proposed here is inspired by DL guidelines presented for other fields such as computer science [55], space weather [56], and medicine [57].

Problem Formulation

The first step in applying DL is to formulate the problem and map the specific needs that require DL. Then, the use of DL should be justified by indicating the limitations of traditional methods for solving this specific problem. In this stage, the need for data should be addressed; where no or insufficient data is available, there is a need to perform new PQ measurements. Traditional methods of presenting and analyzing PQ measurements can help to provide directions for selecting a DL method. For instance, it can help verify the type of training that suits better for a DL algorithm. Supervised learning is suitable if labels are either available for the data or if manual labelling is possible. On the other hand, unsupervised learning is the best choice when no labels are available. Semi-supervised learning is an option if part of the data set contains labels. Linking the measurements with other data sets can be used for labelling. For instance, labels or sequences can be correlated by the simultaneity among PQ events or variations with their causes and ef-facts. The algorithm should be selected once a decision is made between supervised, semi-supervised, and unsupervised.

Data Pre-Processing

Once the raw data from the PQ measurements are available, they should be pre-processed to transform them into useful data features for the DL application. Pre-processing consists of three main steps: cleaning, normalization, and splitting.

Cleaning:

Failures in the measurements can result in missing data points. Malfunctions in the measurement systems can also produce outliers, which are erroneous values in the data set. A common practice in pre- processing is to fill in the missing values and remove or minimize the effects of outliers. For DL applications, it is essential to assess beforehand if the method for handling missing data and measurement errors could impact the results. Recommendations for handling both missing data and outliers can be found in [58].

Normalization:

Data normalization is the pre-processing step that transforms the data into a common range, making an equal contribution to each feature. The main objective of normalization is to minimize the effect of features with higher numerical contributions than others. In case the importance of the features is not known, the importance of the features will be assumed equally distributed. Different normalization methods can be implemented, and mathematical details and recommendations are presented in [59].

Splitting and Reshaping:

The data is usually split into three sets for supervised learning: training, validation, and testing. The DL learning model is trained with the training data set. The validation data set is the one to proceed with evaluating the DL method during the DL model training. The testing data set is used to verify the performance of the trained and validated DL model.

In this stage, the imbalance of the dataset should be verified, i.e., if some classes have more data instances than others. A common approach is to correct the imbalance in the training data to reduce the biases to- ward the predictions. The different methods for correcting the imbalance in data sets are summarized in [60]. For unsupervised learning, data reshaping can help provide the most proper input for the DL model. Data reshaping consists of rearranging the data form without changing the content of the dataset. An example is the transformation of yearly time-series into daily time-series [42].

Algorithm Selection:

The algorithm choice should be based on the class of problem addressed, i.e., supervised, unsupervised, or semi-supervised. A comprehensive overview of the most appropriate algorithms for each class of problem is presented in [61]. Methods employed for supervised learning are deep neural network (DNN), convolutional neural network (CNN), deep belief network (DBN), recurrent neural network (RNN), and its variants such as long short-term memory (LSTM). For unsupervised learning, common algorithms are deep autoencoder (DAE), generative adversarial network (GAN), self-organizing map (SOM), restricted Boltzmann machine (RBM), and deep belief network (DBN). For semi-supervised learning, GAN is often applied as a method to estimate unknown labels. The forthcoming sections contain a brief description of the most common DL methods and their applications to PQ data: DNN, CNN, DBN, RNN, LSTM, GAN, and DAE. For more details and other techniques, the reader is referred to [54, 61].

Short Illustration Table: 1

1. Concepts of Big Data
  1.1 Definition of Big Data
    - 3V's: Volume, Velocity, Variety
    - 4V's: Adding Veracity
  1.2 Explanation of Each V
    - Volume: Amount, size, scale of data
    - Velocity: Data generation speed, processing speed
    - Variety: Data heterogeneity from different sources
    - Veracity: Data quality, accuracy, biases

2. Power Quality Big Data
  2.1 Raw Data and Pre-processing
    - Voltage and current samples
    - Event detection and extraction
    - Mathematical transformations for indexes
  2.2 The 4V's of PQ Data
    - Volume: Large amount of data from measurements
    - Velocity: Sampling frequency and data analysis speed
    - Variety: Diverse sources, monitoring devices
    - Veracity: Compliance with accuracy requirements

3. Big Data Analytics
  3.1 Definition and Challenges
    - Extracting information from 4V's data
    - Traditional methods vs. Machine Learning (ML)
    - Challenge of time-varying data analysis
  3.2 Methods for Time-varying Data
    - Visualization, clustering, feature extraction

- Kalman filter, Parseval's, Hilbert-Huang, etc. - AI methods: ML, Deep Learning (DL) 4. PQ Monitoring and Analytics 4.1 Purpose of PQ Monitoring - Gathering, analyzing, interpreting measurements - Utility and customer facility performance 4.2 Handling Large Data Volumes - Indices, reporting formats - IEC 61000-4-30, CIGRE C4.112 recommendations - Need for shorter time scales in some cases 4.3 Analyzing Time-varying Events - Methods for identifying dynamic changes - Kalman filter, Parseval's, ML, DL

Deep Neural Network (DNN):
DNN are essentially neural networks with multiple hidden layers, each of which further refines the previous layer's outputs [62]. The term deep comes from the fact that the network contains more layers (is deeper) than conventional neural networks [63]. Because of this, conventional ANNs are also called shallow ANNs. DNN is applied to power quality in classification of disturbances [64] and microgrids dynamic stability [65].

Convolutional Neural Network (CNN):
CNN is a multichannel input DL structure composed of learnable weight and bias. The term "convolutional" indicates using convolution instead of matrix multiplication in at least one layer. The simplest CNN architecture contains one convolutional and one pooling layer, option- ally followed by a fully connected layer for supervised prediction. Pooling layers are applied to reduce the data dimensions by combining the outputs of a group of neurons at one layer into a single neuron in the next layer. CNN is applied in PQ mostly for classification of PQ events [66–85], voltage dip classification [40,86,87], recognition of voltage dip causes [88,89], prediction of harmonics [41, 90], and control of voltage unbalance in microgrids [91].

Deep Belief Network (DBN):
DBN are composed of multilayers of Restricted Boltzmann Machine (RBM). In turn, RBM is a type of ANN that can learn a probability distribution from its input [92]. DBN are applied in PQ for classification of PQ events [93,94]. According to [93,94], DBN avoids global fine-tuning and improves the accuracy of power quality disturbance classification compared to traditional ML methods.

Long-Short-Term Memory (LSTM):
LSTM is a variant of recurrent neural network (RNN), composed of a cell, an input gate, and a forget gate. RNN is a DL structure similar to a feedforward neural network that allows exhibiting temporal dynamic behavior. However, unlike a feedforward neural network, RNNs can use their memory to process variable-length sequences of inputs. The unit is called a long short-term memory block because the program uses a structure founded on short-term memory processes to create longer- term memory. Each layer of the LSTM categorizes some level of information, refines it, and passes it along to the next layer. It uses long, short- term memory blocks to provide context for how the program receives inputs and creates outputs. The LSTM block is a complex unit with various components such as weighted inputs, activation functions, in- puts from

previous blocks, and eventual outputs. LSTM is applied to classification of PQ events [95,71–74], recognition of voltage dip causes [96], voltage dip classification [97], harmonic prediction [41,98,99], and islanding detection in microgrids [100,101].

Generative Adversarial Network (GAN):
A GAN is a network that consists of a generator and a discriminator. The generator learns certain data distributions and generates synthetic data according to the distributions. The discriminator distinguishes between the true and synthetically generated data. GAN is applied to the classification of PQ events [75,76] and voltage dip labelling [102].

Deep Autoencoder (DAE):
DAE is an unsupervised learning architecture composed of three main parts: encoder, coding and decoder. The encoder takes the input data and transforms it into a smaller dataset in the coding layer. The decoder takes the dimensional reduced dataset and reconstructs a similar representation of the encoder input and its output. DAE is trained to reduce reconstruction errors. The encoder and decoder contain several fully connected layers instead of one; the number of layers and the number of neurons is symmetric between the encoder and decoder. The main application of AE and DAE is to obtain the principal features of a dataset. The applications in PQ are for feature extraction as part of the classification of PQ events [103–105]; and unsupervised feature learning for clustering of daily harmonics variations at a single location [42], multiple locations [106], spectral data [107], and analytics of waveform distortion in railway installations [108].

Tuning the Hyperparameters
Once the method is chosen, the DL architecture (number of layers and the number of neurons per layer) and the hyperparameters should be selected, depending on the chosen method and application. Hyper- parameters are values that control the learning process, such as the number of epochs, learning rate, and batch size. Although there are benchmark methods for some types of data, such as Alex Net for images [109], there is no standard way to determine the DL architecture and hyperparameters. Therefore, a common practice is manually tuning the method by trial and error [110]. Some recommendations for defining this task are found in [111]. In addition, automatic tuning based on optimization methods has started being proposed in the literature [112].

Training, Validating, and Testing
During the training stage, the values of the DL model are adjusted to fit the training data. The adjustable parameters (weights) define the input-output function of the DL model. A typical DL model contains millions of adjustable weights. Optimization methods based on stochastic gradient descent are often employed to update the weights to minimize the error in estimating the outputs. During the training, validation plays a role in providing an unbiased evaluation of the DL model. After training, the performance of the DL model is evaluated through the testing data set. The test aims to generalize the ability of the DL model to produce outputs based on inputs that were not seen in the training stage. A DL model with good performance in the training stage can present worse performance on estimating outputs with unseen instances. This is called overfitting, and it often occurs due to insufficient amount of data. The opposite phenomenon is underfitting when the DL model fits neither the training set nor the testing set. Underfitting is mainly a consequence of the low complexity of the DL model. Both overfitting and underfitting should be avoided. Overfitting can be solved either by reducing the complexity of the DL model

by using a lower number of neurons and layers or by using regularization techniques such as dropout. Under- fitting can be avoided by increasing the complexity of the DL method by adding more neurons and layers.

Deployment

This step defines how the model will be deployed as an application tool. Providing transparency to the DL model is important in this step; in other words, ensuring that the results can be interpretable, and the DL method does not appear as a closed box. A possibility is to explain the features extracted by the DL models and their role in estimating the outputs. In this step, it is also important to define how the tool would be delivered to users, in this case, typically PQ engineers, and how they can use it.

REVIEW OF AI APPLICATIONS FOR PQ

Fig. 2 shows that the number of publications related to AI and PQ increased a lot from 1992 to 2021 in the Scopus Database. The legend of Fig. 2 shows the logical structures used for the search. The blue color in Fig. 2 indicates the number of occurrences of references that included the term "power quality" together with any of the AI terms: "deep learning", "machine learning", "artificial intelligence", "expert system". The red color indicates only the number of references that included the terms "power quality" and "deep learning". The search considers the keywords and abstract of the publication.

Although most AI methods have been developed before, the application of AI to PQ started only after 2000.

(b). The text data mining from [113] produces a visualizing tool that groups words from the title, abstract, and keywords of the different publications. If a group of words appears at least five times, it is considered as a cluster. Each cluster is represented by a color and il- lustrates the most correlated word in the analyzed periods. In this perspective, the keywords are related to the diverse AI and signal-processing techniques that have been implemented to detect, extract, and analyses PQ variations and events. Section 5.1 details the techniques and applications before 2017 and Section 5.2 after 2017. Section 5.3 details and compares the applications in the literature of DL to PQ.

Techniques and Applications from 1992 to 2017

The blue cluster in Fig. 3(a) represents the first applications that were based on expert systems that make decisions through rules based on human expertise [114]. Such methods were applied to PQ data to perform the classification and recognition of PQ events until 2010 [9,39, 115–117]. Signal processing methods (magenta and yellow clusters in Fig. 3(a)) play an important role in signal decomposition and feature extraction. The two techniques appearing most were Wavelet and S-transform; these were mentioned in Section 2.4 to analyses and visualize time-varying events. Moreover, signal-processing methods are combined with expert systems or ML tools. To classify PQ disturbances, for example, [118] uses wavelet and fuzzy support vector machines; [119] uses S-transform and ML; and [120] uses S-transform for the feature extraction stage and support vector machines for the pattern recognition problem of non-intrusive load monitoring.
Learning algorithms have been applied for similar applications as the expert system, i.e., for the classification and recognition of different events [121–126]. Most algorithms are based on supervised learning that requires pre-labeled data, e.g., shallow neural networks [93–95, 98], and

support vector machines [96,97]. Few works apply unsupervised learning to find patterns in PQ data without prior.

knowledge: [127] applies principal component analysis and [128] k-means clustering. The work in [128] was the first attempt to find patterns of PQ big data by applying an unsupervised method. However, [128] did not extract the principal features for clustering, and the analysis is limited to the correlation among harmonics. Moreover, [128] did not allow to obtain the typical daily patterns of time-varying distortion and did not result in a user-friendly orientation to further analyses the data Optimization algorithms also appear as a trend for AI applications to PQ. This cluster, represented in red in Fig. 3(a), is related to optimization methods such as genetic algorithms mostly for decision-making on the placement of capacitor banks [129] and active filters [130].

Techniques and Applications from 2018 to 2021

For the period between 2018 and 2021, the techniques explained in Section 5.1 also appear in the clusters of Fig. 3(b). However, those techniques appear primarily in hybrid approaches. For example, the red cluster contains approaches that apply fuzzy logic combined with optimization, mainly for control systems [131–133]. Moreover, methods combining optimization with ML are found in the green cluster. For instance, [134] combines metaheuristics with ANNs and support vector machines to detect, segment, and classify voltage dips. In addition to the green cluster, the group of learning algorithms presents two more sub- groups in Fig. 3(b): ML in magenta, DL in blue and in yellow (the latter with the keyword “convolutional neural network”). In the context of DL (blue), the keywords "classification", "disturbance classification", and "power quality disturbances" are found together with "deep learning".

Applications of Deep Learning for PQ

The analysis by the text mining in Fig. 4 gave a general overview of the DL applications to PQ. To obtain more detailed information, Table 1 stratifies 46 references concerning the type of application, type of data, learning technique, and DL method. The publications were found by searching "power quality" and "deep learning" in the Scopus database and refining the engineering/energy fields search. The references in Table 1 cover the period from January 2018 to September 2022. Table 1 shows that most of the publications are related to the classification and recognition of PQ disturbances. This points out that, even with the possibility of automatic feature extraction, the applications of DL in the literature are still largely the same as the applications of expert systems or the earliest ML tools. Moreover, most of the published studies are based on synthetically generated data and on supervised techniques. Other applications are voltage dip classification, voltage dip estimation, recognition of voltage dip causes, prediction of harmonics, recognition of patterns in daily harmonic variations, identification of spectral patterns, control of voltage unbalance in microgrids, islanding detection in microgrids, and dynamic stability in microgrids.

Table: 2

<table>
<tr><td colspan="2">AI Applications for PQ</td></tr>
<tr><td colspan="2">DL Method</td></tr>
<tr><td colspan="2">Description and Applications</td></tr>
<tr><td>DNN</td><td>CNN</td></tr>
<tr><td>DBN</td><td>LSTM</td></tr>
<tr><td>GAN</td><td>DAE</td></tr>
</table>

Hyperparameter Tuning
Training, Validation, and Testing Deployment
Review of AI Applications
Techniques and Applications (1992-2017)
Techniques and Applications (2018-2021)
Applications of Deep Learning for PQ

This table provides a structured overview of the discussed topics and their applications in the field of power quality using artificial intelligence and deep learning techniques. You can use this as a reference to create a visual diagram using diagramming software.

DISCUSSIONS AND RECOMMENDATIONS

Lack of Novelty

Many publications have proposed the automatic classification of PQ records by DL. This trend was also identified for PQ applications with signal processing [135]. However, the classification of PQ disturbances does not present any novelty when the classes are individual PQ disturbances such as swell, sag, interruption, harmonics, and transients. Standards such as IEEE 1159, IEC 61000-4-30 and EN 50160 define methods for classifying PQ events based on characteristics such as spectral content, duration, and voltage magnitude. This type of application is already present in PQ monitors and their commercial computing platforms; there is no need for DL applications to replace classification methods defined in standards.

Training Based on Synthetic Data Sets

A substantial part of the approaches within the literature is based on synthetically generated and often non-realistic data. There is a risk of overfitting in case synthetic data is used, and in this case, the test error seems high even for low training error [6].

There are several reasons for using synthetic data in supervised DL by PQ researchers. One is the non-availability and inadequacy of real or realistic power system data, for example, from field devices. With PQ data obtained from measurements, the class labels from PQ data sequences needed to develop a classification system are typically unknown. The use of synthetic data does not suffer from this disadvantage as the labels for the data are known according to the process by which the data sequences are generated. A synthetic set would be better in covering all classes, but it would not be an appropriate statistical representation of the population of disturbances.

The performance of grid measurements in AI methods trained by synthetic data has been evaluated for ML and DL. In the context of ML for PQ, [137] has tested a support vector machine by employing measurements from two different power networks and synthetic data. The classifier presented a good detection rate when trained by data from one power network and tested with measurements from another network. However, the training employed with synthetic data did not produce acceptable results when tested with measured data. Combining measurements and synthetic data improved the performance, but the detection rate was still low compared to the classifier trained with measurements. About DL for PQ, the training of a LSTM for classification of voltage dip types with synthetic data and testing with real data has been done in [138]. It was concluded by [138] that the DL model had a

Table 3: Overview and comparison of recent publications covering the application of DL to power quality big data.

	DAE+CNN [77]	CNN [86,87]	CNN [40]	LSTM [97]
Voltage dip or fault classification	X	X	X	X
Voltage dip estimation	X		X	
Recognition of voltage dip or fault causes	X	X	X	X
Pre-detection/Prediction/Estimation of Harmonics	X	X	X	X
CNN+LSTM [41]	X	X	X	

Unsupervised Semi-supervised low performance when trained with synthetic data and tested with real data.

Enlarging Datasets by Adding Gaussian Noise

An example of non-realistic data generation is the application of noise to the datasets. A large group of publications adds noise to the data to evaluate the robustness of the methods or to enlarge the input datasets.

[67,71,72,75,77–80,85,93,96,103,104,139–141]. However, the noise is Gaussian and does not correspond to real noise in the electrical power grid [142]. Such Gaussian noise might comprise the decision-making of DL algorithms due to the addition of frequency components that do not appear in reality. Studies considering the effect of Gaussian noise instead of real noise are needed in the DL context. Efforts to make available open datasets with real noise can benefit the PQ research community.

Lack of Benchmark Databases

One of the challenges of DL in PQ applications is comparing the performance of the different algorithms. The lack of a standard database that can be used as a benchmark has been reported in [143]. Although some efforts have been made to provide public databases, they are limited in terms of application. Most of the available data are without labels. An example of a raw database is the one provided by the IEEE PES Subcommittee on Big Data & Analytics for Power Systems [144] which contains 1380 files of current, voltage, and active power measurements. In addition, some databases are suitable for studies of only one type of disturbance. For instance, [145] presents a collection of PQ real-life impulsive events, and [22] is a data set that contains waveforms of voltage dips. Also, some databases are for a specific type of application; for example, [146] contains measured voltage and current waveforms of railways, and [147] has harmonic measurements for an installation with electric vehicles. In sum, the public databases do not follow a standard, presenting different disturbances, sampling, and applications.

An attempt to establish a reference dataset for DL classifiers is presented in [143] by proposing an open-source software to enable the generation of synthetic data. However, the tool is appropriate only for the applications that concern events classification in terms of disturbance type. Furthermore, the authors of [143] also emphasize that such a dataset should be explicitly used to compare PQ disturbances classification algorithms. In other words, this tool might be proper for testing, not for training algorithms. The grid stakeholders and academia should make efforts to provide one or more reference databases to compare DL applications to PQ. Suggestions for databases are waveforms of PQ events with labels, long- term measurements of

PQ variations, and sub-10 min values of PQ variations [30] for different power installations and voltage levels.

Learning Strategies and Labelling

Most of the applications of DL to PQ are based on supervised learning, which requires labelled data. This trend is also observed for ML, deep or not, in most applications in different research fields [54]. Controversially, most of the available PQ data is non-labelled. Non-labelled data sets are suitable for unsupervised learning but not for supervised learning. Only a few works have explored unsupervised learning [41, 106, 107, 98], which are also among the few ones that applied measurements instead of synthetic data. The supervised training trend can be again due to the non-availability of field measurements for researchers and scholars. One, yet unsolved, issue is the verification/testing of unsupervised learning of PQ variations. Most DL models for PQ are based on supervised learning with synthetic data. As discussed in Section 6.2, synthetic data is easier to label as the process of synthetic data generation is known. However, the class choices for supervised learning should represent reality in measurements, but that is often not the case. In addition to the classes of individual PQ disturbances, some works propose classes for combined disturbances, such as swell with harmonics, dip with harmonics, dip with transients, and so on. With field measurements, events such as voltage dips always occur in combination with harmonics and transients [148]. Moreover, events can exist of multiple stages: the transient can be the starting, the dip can present many segments, and even it can occur as the initiating event for an interruption [149]. Instead of event classes, the labelling for supervised learning should be based on the origins and consequences of the event. The main limitation against precise PQ classification is the need for ground-truth and notated data [102]. No large ground-truth labelled datasets are available, and it is even more complicated with many measurements [102]. There is an uncertainty in the labelling process, which can be associated with incorrect labelling by power systems operators, or to the automatic method used for labelling, samples close to the borders, samples polluted with noise, and so on [102]. The grid stakeholders should make an effort to keep track of the events and correctly annotate the data.

Annotated data can serve as a basis also for semi-supervised learning. In semi-supervised learning, the algorithms can deal with partially labelled data. In the literature, only [102] has applied DL semi-supervised learning for automatic labelling of voltage dip sequences by using a small set of ground-truth labelled data. The method based on GAN, described in Section 5.4.5, was applied to a large set of measured dips; it presented 83% average accuracy with a 3.2% false alarm rate. Generating synthetic data has not been used yet for unsupervised learning. However, the application of synthetic data might assist in understanding the feature extraction in unsupervised learning.

Low Transparency

Users of ML and DL algorithms, like experts in PQ, might find it hard to trust the results from such an algorithm due to low transparency in the DL process. The core obstacle to the practical use of DL is that most DL algorithms appear like closed boxes. With DL, human experts do not select the features, but those features are obtained by using a learning procedure. It is different from an open box as physical models and expert systems, in which the process of defining an output is explainable in terms of mathematical equations and logic rules, respectively. The explanation of the decision-making process of DL methods can make them transparent. By increasing the explain ability of DL, human experts can benefit through new rules in a decision-making problem.

According to [150], one of the difficulties in interpreting deep neural networks is related to the activation process of neurons. The activation of some neurons can occur for a few data instances, whereas the activation of other neurons can be more globally. In this way, the output is a function of both local and global effects. Therefore, this makes it difficult to map an equivalent function that explains the prediction by DL for the data.

The lack of explanation in the DL decision-making has been pointed out as the main obstacle in sensitive fields such as precision medicine, law, and financial sector [150]. For power systems, this discussion has been considered in the context of ML [151], emphasizing the major barrier of providing trust in ML models such as neural networks and adopting them in practice. A first trial for measuring the explainability and trustworthiness of DL methods for PQ disturbances is presented in [152]; however, the discussion covers only the classification in terms of the type of disturbances. So far, there is no well-defined optimal method for explaining DL [153]. Some suggestions for increasing the explain- ability of DL, in general, are presented in [153]. It includes indexes and methods to evaluate the contribution of features to the predictions. Further efforts are needed to make DL more transparent to be more broadly applied in the PQ field.

Expert Knowledge and Development of Hybrid Approaches

In most applications, even with the possibility of automatic feature extraction and settings, PQ knowledge is still needed to interpret correctly and provide appropriate solutions based on DL results. As discussed in [43], the results provided by DL can be used to decide about further manual analysis of the PQ data. However, there is no standard method available yet for the additional manual analyses. Reference [43] suggests developing a hybrid method that combines the results given by DL with an expert system. The logic rules in the expert system could be defined to take some of the decision-making away from the human expert. However, some level of expert involvement will remain needed, as every case is unique. Approaches that combine DL and expert systems are not yet explored in the literature. Therefore, it is suggested to develop logic-based rules to make the interpretation of DL results at least partly automatic.

CONCLUSIONS

This review has covered the latest applications of DL to PQ big data. DL can be a solution to turn raw PQ measurements into a much more valuable asset. However, a large group of publications has limitations related to innovation and applicability. Even with the possibility of automatic feature extraction, most proposed DL algorithms execute the same task as expert systems or the early ML tools (classification, recognition of events, and underlying causes). By stratifying the publications in terms of application, learning technique, and type of data, this paper has demonstrated that: Most publications apply DL to problems solved by expert systems and early ML tools. Although such works still contribute to the knowledge on DL methods, the practical applicability is limited. A large number of works apply synthetic datasets that do not represent the reality in PQ data, limiting the practical applicability of those works.

Few works have provided new solutions to perform analytics to PQ big data; some examples of new solutions are predicting PQ data, extracting daily patterns, and automatic data labelling. In most applications, even with the possibility of automatic feature extraction and settings, PQ knowledge is still needed to correctly interpret the results and provide appropriate solutions to PQ issues.

REFERENCES

[1] S. Ronnberg, M. Bollen, Power quality issues in the electric power system of the future, Electr. J. 29 (10) (2016) 49–61.

[2] F. Z. (convenor), "CIGRE technical brochure 719, power quality and EMC issues with future electricity networks," no. March. 2018.

[3] J. Milanovic et al., "CIGRE/CIRED JWG C4. 112: guidelines for power quality monitoring–measurement locations, processing and presentation of data," 2014.

[4] S. Elphick, V. Gosbell, V. Smith, S. Perera, P. Ciufo, G. Drury, Methods for harmonic analysis and reporting in future grid applications, IEEE Trans. Power Deliv. 32 (2) (2016) 989–995.

[5] A. Mariscotti, Direct measurement of power quality over railway networks with results of a 16.7-Hz network, IEEE Trans. Instrum. Meas. 60 (5) (2010) 1604–1612.

[6] Y. Zhang, T. Huang, E.F. Bompard, Big data analytics in smart grids: a review, Energy Inform. 1 (1) (2018) 1–24.

[7] S. Kaisler, F. Armour, J.A. Espinosa, W. Money, Big data: issues and challenges (6) (2018) 2794–2802 moving forward, in: 2013 46th Hawaii international conference on system sciences, 2013, pp. 995–1004.

[8] M.M. Najafabadi, F. Villanustre, T.M. Khoshgoftaar, N. Seliya, R. Wald, E. Muharemagic, Deep learning applications and challenges in big data analytics, J. big data 2 (1) (2015) 1–21.

[9] W.R.A. Ibrahim, M.M. Morcos, Artificial intelligence and advanced mathematical tools for power quality applications: a survey, IEEE Trans. Power Deliv. 17 (2) (2002) 668–673.

[10] P. Louridas, C. Ebert, Machine learning, IEEE Softw 33 (5) (2016) 110–115.

[11] L. Cheng, T. Yu, A new generation of AI: a review and perspective on machine learning technologies applied to smart energy and electric power systems, Int. J. Energy Res. 43 (6) (2019) 1928–1973.

[12] C. Ge, Machine Learning Methods for Image Analysis in Medical Applications, from Alzheimer's Disease, Brain Tumors, to Assisted Living, Chalmer University of Technology, 2020. https://research.chalmers.se/en/publication/517576.

[13] S. Khokhar, A.A.B. Mohd Zin, A.S.B. Mokhtar, M. Pesaran, A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances, Renew. Sustain. Energy Rev. 51(2015) 1650–1663.

[14] G.S. Chawda, et al., Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration, IEEE Access 8 (2020) 146807–146830.

[15] W.R. Anis Ibrahim, M.M. Morcos, Artificial intelligence and advanced mathematical tools for power quality applications: a survey, IEEE Trans. Power Deliv. 17 (2) (2002) 668–673.

[16] Rahul, Review of signal processing techniques and machine learning algorithms for power quality analysis, Adv. Theory Simul. 3 (10) (2020).

[17] A. Eslami, M. Negnevitsky, E. Franklin, S. Lyden, Review of AI applications in harmonic analysis in power systems, Renew. Sustain. Energy Rev. 154 (2022), 111897. Feb.

[18] J.R. Mashey, Big data and the next wave of InfraStress problems, solutions, opportunities, in: 1999 USENIX Annual Technical Conference, 1999.

[19] D. Laney, 3D data management: controlling data volume, velocity and variety, META Gr. Res. Note 6 (2001).

[20] J. Gantz and D. Reinsel, "EXtracting value from chaos," IDC iview, vol. 1142, no. 2011, pp. 1–12, 2011.

[21] A. L'heureuX, K. Grolinger, H.F. Elyamany, M.A.M. Capretz, Machine learning with big data: challenges and approaches, Ieee Access 5 (2017) 7776–7797.

[22] H. Hu, Y. Wen, T.-S. Chua, X. Li, toward scalable systems for big data analytics: a technology tutorial, IEEE Access 2 (2014) 652–687.

[23] M.F. McGranaghan, S. Santoso, Challenges and trends in analyses of electric power quality measurement data, EURASIP J. Adv. Signal Process. 2007 (2007) 1–5.

[24] I.-I. E. Commission and others, "IEC 61000-4-30," 2015.

[25] "IEEE recommended practice for power quality data interchange format (PQDIF)," IEEE Std 1159.3-2019 (Revision IEEE Std 1159.3-2003), pp. 1–185, 2019.

[26] I. E. C. Standard, "IEC 61000-4-7: testing and measurement techniques – general guide on harmonics and interharmonics measurements and instrumentation, for power supply systems and equipment connected thereto," 2009.

[27] P.F. Ribeiro, C.A. Duque, P.M. Ribeiro, A.S. Cerqueira, Power Systems Signal Processing for Smart Grids, John Wiley & Sons, 2013.

[28] K. Lu, C. Wang, K. Wu, M. Gong, Y. Wang, A unified framework for exploring time-varying volumetric data based on block correspondence, Vis. Inform. 3 (4) (2019) 157–165.

[29] Joint working group CIGRE/CIRED C4.112, "Guidelines for power quality monitoring - measurement locations, processing and presentation of data," CIGRE TB 596, 2014.

[30] A. Gil-de-Castro, M.H.J. Bollen, S.K. Ro¨nnberg, Variations in harmonic voltage at the sub-10-minute time scale, Electr. Power Syst. Res. 195 (2021), 107163.

[31] K.K.C. Yu, N.R. Watson, J. Arrillaga, An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking, IEEE Trans. Power Deliv. 20 (2) (2005) 1577–1584.

[32] C.S. Moo, Y.N. Chang, P.P. Mok, A digital measurement scheme for time-varying transient harmonics, IEEE Trans. Power Deliv. 10 (2) (1995) 588–594.

[33] N. Senroy, S. Suryanarayanan, P.F. Ribeiro, An improved Hilbert–Huang method for analysis of time-varying waveforms in power quality, IEEE Trans. Power Syst. 22 (4) (2007) 1843–1850.

[34] R.G. Stockwell, L. Mansinha, R.P. Lowe, Localization of the complex spectrum: the S transform, IEEE Trans. Signal Process. 44 (4) (1996) 998–1001.

[35] P.M. Silveira, P.F. Ribeiro, Visualization of time-varying waveform distortions with wavelets. Time-Varying Waveform Distortions in Power Systems, Wiley Online Library, 2009, pp. 176–186.

[36] P.F. Ribeiro, Time-Varying Waveform Distortions in Power Systems, Wiley Online Library, 2009.

[37] J. Xu, N. Senroy, S. Suryanarayanan, P. Ribeiro, Some techniques for the analysis and visualization of time-varying waveform distortions, in: 2006 38th North American Power Symposium, 2006, pp. 257–261.

[38] N. Nakhodchi, R.A. de Oliveira, M.H.J. Bollen, Graphical methods for presenting time-varying harmonics, in: CIRED 2021 Conference, 2021.

[39] E. Styvaktakis, M.H.J. Bollen, I.Y.H. Gu, EXpert system for classification and analysis of power system events, IEEE Trans. Power Deliv. 17 (2) (2002) 423–428. Apr.

[40] A. Bagheri, I.Y.H. Gu, M.H.J. Bollen, E. Balouji, A robust transform-domain deep convolutional network for voltage dip classification, IEEE Trans. Power Deliv. 33

[41] E. Balouji, K. Ba¨ckstro¨m, T. McKelvey, O. Salor, Deep-learning-based harmonics and interharmonics predetection designed for compensating significantly time- varying eaf currents, IEEE Trans. Ind. Appl. 56 (3) (2020) 3250–3260.

[42] C. Ge, R.A. de Oliveira, I.Y.-H. Gu, M.H.J. Bollen, Deep feature clustering for seeking patterns in daily harmonic variations, IEEE Trans. Instrum. Meas. 70 (2021) 1–10.

[43] R.A. de Oliveira, V. Ravindran, S.K. Ro¨nnberg, M.H.J. Bollen, Deep learning method with manual postprocessing for identification of spectral patterns of waveform distortion in PV installations, IEEE Trans. Smart Grid (2021) 1.

[44] I. Wachsmuth, The concept of intelligence in AI. Prerational Intelligence: Adaptive Behavior and Intelligent Systems Without Symbols and Logic, Volume 1, Volume 2 Prerational Intelligence: Interdisciplinary Perspectives on the Behavior of Natural and Artificial Systems, Volume 3, Springer, 2000, pp. 43–55.

[45] A.M. TURING, I.—computing machinery and intelligence, Mind LIX (236) (1950) 433–460. Oct.

[46] C.E. Shannon, J. McCarthy, others, Automata Studies, 11, Princeton University Press, Princeton, NJ, 1956.

[47] P.J. Werbos, Backpropagation through time: what it does and how to do it, Proc. IEEE 78 (10) (1990) 1550–1560.

[48] M. Campbell, A.J. Hoane, F. Hsu, Deep blue, Artif. Intell. 134 (1) (2002) 57–83.

[49] E. Strickland, the turbulent past and uncertain future of AI: is there a way out of AI's boom-and-bust cycle? IEEE Spectr 58 (10) (2021) 26–31.

[50] Shu-Hsien Liao, EXpert system methodologies and applications—a decade review from 1995 to 2004, EXpert Syst. Appl. 28 (1) (2005) 93–103. Jan.

[51] Y. Liao, J.-B. Lee, A fuzzy-expert system for classifying power quality disturbances, Int. J. Electr. Power Energy Syst. 26 (3) (2004) 199–205. Mar.

[52] T.T. Nguyen, N.D. Nguyen, S. Nahavandi, Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications, IEEE Trans. Cybern. 50 (9) (2020) 3826–3839. Sep.

[53] A. G´eron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media, 2019.

[54] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444. May.

[55] C. Garbin, X. Zhu, O. Marques, Dropout vs. batch normalization: an empirical study of their impact to deep learning, Multimed. Tools Appl. 79 (19–20) (2020) 12777–12815. May.

[56] E. Camporeale, The challenge of machine learning in space weather: nowcasting and forecasting, Sp. Weather 17 (8) (2019) 1166–1207. Aug.

[57] E. Montagnon, et al., Deep learning workflow in radiology: a primer, Insights Imaging 11 (1) (2020) 22. Dec.

[58] Matlab, "Handling missing data and outliers," 2022. [Online]. Available: https://www.mathworks.com/help/ident/ug/handling-missing-data-and-outliers. html.

[59] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, Appl. Soft Comput. 97 (2020), 105524.

[60] N. Rout, D. Mishra, M.K. Mallick, Handling imbalanced data: a survey. International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications, Springer, 2018, pp. 431–443.

[61] I.H. Sarker, Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions, SN Comput. Sci. 2 (6) (2021) 420. Nov.

[62] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Netw. 61(2015) 85–117.

[63] R. Dechter, "Learning while searching in constraint-satisfaction problems," 1986.

[64] H. Liu, et al., Signal processing and deep learning techniques for power quality events monitoring and classification, Electr. Power Components Syst. 47 (14–15) (2019) 1332–1348.

[65] V. Anantha Krishnan, P. Balamurugan, an efficient DLN2-CRSO approach based dynamic stability enhancement in micro-grid system, Appl. Energy 322 (2022), 119432. Sep.

[66] W. Chen, J. He, X. Pei, Classification for power quality disturbance based on phase-space reconstruction and convolution neural network, Dianli Xitong Baohuyu Kongzhi/Power Syst. Prot. Control 46 (14) (2018) 87–93.

[67] S. Wang, H. Chen, A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network, Appl. Energy 235 (2019) 1126–1140.

[68] F.M. Arrabal-Campos, A. Alcayde, F.G. Monloya, J. Martinez-Lao, J. Castillo- Marline, R. Ban˜os, Robust-PCA deep learning for PQ disturbances classification using synchrosqueezing wavelet transform, Renew. Energy Power Qual. J. 19 (2021) 546–551.

[69] K. Cai, T. Hu, W. Cao, G. Li, classifying power quality disturbances based on phase space reconstruction and a convolutional neural network, Appl. Sci. 9 (18) (2019) 3681.

[70] K. Cai, W. Cao, L. Aarniovuori, H. Pang, Y. Lin, G. Li, Classification of power quality disturbances using Wigner-Ville distribution and deep convolutional neural networks, IEEE Access 7 (2019) 119099–119109.

[71] M. Cao, Y. Zhang, Classification for power quality disturbances based on CNN- LSTM network, Dianli Xitong Baohu yu Kongzhi/Power Syst. Prot. Control 48 (2) (2020) 86–92.

[72] D. Gu, Y. Gao, Y. Li, Y. Zhu, C. Wu, A novel label-guided attention method for multilabel classification of multiple power quality disturbances, IEEE Trans. Ind. Inform. (2021).

[73] G. Bayrak, A. Küçüker, A. Yılmaz, Deep learning-based multi-model ensemble method for classification of PQDs in a hydrogen energy-based microgrid using modified weighted majority algorithm, Int. J. Hydrogen Energy (2022). Jun.

[74] K. Sekar, K. Kanagarathinam, S. Subramanian, E. Venugopal, C. Udayakumar, an improved power quality disturbance detection using deep learning approach, Math. Probl. Eng. 2022 (2022) 1–12. May.

[75] X. Jian, X. Wang, A novel semi-supervised method for classification of power quality disturbance using generative adversarial network, J. Intell. Fuzzy Syst. 40 (3) (2021) 3875–3885.

[76] D. Gu, Y. Gao, Y. Li, Y. Zhu, C. Wu, A novel label-guided attention method for multilabel classification of multiple power quality disturbances, IEEE Trans. Ind.Inform. 18 (7) (2022) 4698–4706. Jul.

[77] K. Nandi, A.K. Das, R. Ghosh, S. Dalai, B. Chatterjee, Hyperbolic window S- transform aided deep neural network model-based power quality monitoring framework in electrical power system, IEEE Sens. J. 21 (12) (2021) 13695–13703.

[78] H. Liu, F. Hussain, Y. Shen, S. Arif, A. Nazir, M. Abubakar, Complex power quality disturbances classification via curvelet transform and deep learning, Electr. Power Syst. Res. 163 (2018) 1–9.

[79] X. Xiao, K. Li, Multi-label classification for power quality disturbances by integrated deep learning, IEEE Access 9 (2021) 152250–152260.

[80] H. Sindi, M. Nour, M. Rawa, O¨ ztürk, and K. Polat, A novel hybrid deep learning approach including combination of 1D power signals and 2D signal images for power quality disturbance classification, EXpert Syst. Appl. 174 (2021).

[81] H. Sindi, M. Nour, M. Rawa, S¸. O¨ ztürk, K. Polat, an adaptive deep learning framework to classify unknown composite power quality event using known single power quality events, EXpert Syst. Appl. 178 (2021).

[82] L. Ramalingappa, A. Manjunatha, Power quality event classification using complex wavelets phasor models and customized convolution neural network, Int. J. Electr. Comput. Eng. 12 (1) (2022) 22–31.

[83] I. Topaloglu, Deep learning based a new approach for power quality disturbances classification in power transmission system, J. Electr. Eng. Technol. (2022). Jul.

[84] B. Eristi, H. Eristi, A new deep learning method for the classification of power quality disturbances in hybrid power system, Electr. Eng. (2022). Jun.

[85] J. Tian, B. Wang, J. Li, Z. Wang, Adversarial attacks and defense for CNN based power quality recognition in smart grid, IEEE Trans. Netw. Sci. Eng. (2021).

[86] S. Turizo, G. Ramos, D. Celeita, Voltage sags characterization using fault analysis and deep convolutional neural networks, IEEE Trans. Ind. Appl. 58 (3) (2022) 3333–3341. May.

[87] D.-H. Yoon, J. Yoon, Deep learning-based method for the robust and efficient fault diagnosis in the electric power system, IEEE Access 10 (2022) 44660–44668.

[88] M. Zanoni, R. Chiumeo, L. Tenti, M. Volta, Automated tool based on deep learning to assess voltage dips validity: integration in the QuEEN MV network monitoring system, Renew. Energy Power Qual. J. 19 (2021) 235–240.

[89] H. Liu, S. Liu, J. Zhao, T. Bi, X. Yu, Dual-channel convolutional network-based fault cause identification for active distribution system using realistic waveform measurements, IEEE Trans. Smart Grid (2022) 1. –1.

[90] N. Severoglu, O. Salor, Statistical models of EAF harmonics developed for harmonic estimation directly from waveform samples using deep learning framework, IEEE Trans. Ind. Appl. 57 (6) (2021) 6730–6740.

[91] S. Bera, S. Chakraborty, D. Kumar, N. Ali, M. Lehtonen, Optimal deep learning-based aggregation of TCLs in an inverter fed stand-alone microgrid for voltage unbalance mitigation, Electr. Power Syst. Res. 210 (2022), 108178. Sep.

[92] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput 18 (7) (2006) 1527–1554. Jul.

[93] Y. Gao, Y. Li, Y. Zhu, C. Wu, D. Gu, Power quality disturbance classification under noisy conditions using adaptive wavelet threshold and DBN-ELM hybrid model, Electr. Power Syst. Res. (2021).

[94] Z. Chen, M. Li, T. Ji, Q. Wu, Real-time recognition of power quality disturbance- based deep belief network using embedded parallel computing platform, IEEJ Trans. Electr. Electron. Eng. 15 (4) (2020) 519–526. Apr.

[95] I˙. O¨ zer, S.B. Efe, H. O¨ zbay, CNN/Bi-LSTM-based deep learning algorithm for classification of power quality disturbances by using spectrogram images, Int. Trans. Electr. Energy Syst. (2021).

[96] Z. Zheng, L. Qi, H. Wang, M. Zhu, Q. Chen, Recognition method of voltage sag causes based on Bi-LSTM, IEEJ Trans. Electr. Electron. Eng. 15 (3) (2020) 418–425.

[97] R. Turovi´c, et al., An end-to-end deep learning method for voltage sag classification, Energies 15 (8) (2022) 2898. Apr.

[98] L. Zjavka, Power quality statistical predictions based on differential, deep and probabilistic learning using off-grid and meteo data in 24-hour horizon, Int. J. Energy Res. (2021).

[99] L. Zjavka, Power quality 24-hour prediction using differential, deep and statistics machine learning based on weather data in an off-grid, J. Franklin Inst. (2022). Jul.

[100] Y. Xia, F. Yu, X. Xiong, Q. Huang, Q. Zhou, A novel microgrid islanding detection algorithm based on a multi-feature improved LSTM, Energies 15 (8) (2022) 2810. Apr.

[101] A.K. O¨ zcanlı, M. Baysal, A novel multi-LSTM based deep learning method for islanding detection in the microgrid, Electr. Power Syst. Res. 202 (2022).

[102] A. Bagheri, I.Y.H. Gu, M.H.J. Bollen, Generative adversarial model-guided deep active learning for voltage dip labelling, in: 2019 IEEE Milan PowerTech, 2019, pp. 1–5.

[103] H. Liu, F. Hussain, S. Yue, O. Yildirim, S.J. Yawar, Classification of multiple power quality events via compressed deep learning, Int. Trans. Electr. Energy Syst. 29 (6) (2019).

[104] A.D. Gonz´alez, V. Martinez, M.Delgado Prieto, J. Saucedo, R. Osornio, Power quality monitoring and disturbances classification based on autoencoder and neural network for electrical power supply, in: International Conference on Renewable Energies and Power Quality, 2020, pp. 1–5.

[105] P.K. Dash, E.N.V.D.V. Prasad, R.K. Jalli, S.P. Mishra, Multiple power quality disturbances analysis in photovoltaic integrated direct current microgrid using adaptive morphological filter with deep learning algorithm, Appl. Energy 309 (2022), 118454. Mar.

[106] C. Ge, R.A.D. Oliveira, I.Y.H. Gu, M.H.J. Bollen, Unsupervised deep learning and analysis of harmonic variation patterns using big data from multiple locations, Electr. Power Syst. Res. 194 (2021).

[107] R.A. De Oliveira, V. Ravindran, S.K. Ronnberg, M.H.J. Bollen, Deep learning method with manual post-processing for identification of spectral patterns of waveform distortion in PV installations, IEEE Trans. Smart Grid 12 (6) (2021) 5444–5456.

[108] R.S. Salles, R.A. de Oliveira, S.K. Ronnberg, A. Mariscotti, Analytics of waveform distortion variations in railway pantograph measurements by deep learning, IEEE Trans. Instrum. Meas. 71 (2022) 1–11.

[109] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM 60 (6) (2017) 84–90. May.

[110] F. Hutter, J. Lücke, L. Schmidt-Thieme, Beyond manual tuning of hyperparameters, KI-Künstl. Intell. 29 (4) (2015) 329–337.

[111] "How to grid search hyperparameters for deep learning models in python with keras," 2018.

[112] A.H. Victoria, G. Maragatham, Automatic tuning of hyperparameters using Bayesian optimization, Evol. Syst. 12 (1) (2021) 217–223. Mar.

[113] N. J. Van Eck and L. Waltman, "Text mining and visualization using VOSviewer," arXiv Prepr. arXiv1109.2058, 2011.

[114] W.E. Kazibwe, H.M. Sendaula, EXpert system targets power quality issues, IEEE Comput. Appl. Power 5 (2) (1992) 29–33.

[115] M.B.I. Reaz, F. Choong, M.S. Sulaiman, F. Mohd-Yasin, M. Kamada, EXpert system for power quality disturbance classifier, IEEE Trans. Power Deliv. 22 (3) (2007) 1979–1988.

[116] P.K. Dash, S. Mishra, M.A. Salama, A.C. Liew, Classification of power system disturbances using a fuzzy expert system and a Fourier linear combiner, IEEE Trans. Power Deliv. 15 (2) (2000) 472–477. Apr.

[117] B.R. Klingenberg, P.F. Ribeiro, Fuzzy logic application for time-varying harmonics, Time-Varying Waveform Distort. Power Syst 6 (2009) 197.

[118] G.-S. Hu, J. Xie, F.-F. Zhu, Classification of power quality disturbances using wavelet and fuzzy support vector machines, in: 2005 International Conference on Machine Learning and Cybernetics 7, 2005, pp. 3981–3984.

[119] H. Ericsti, O¨. Yildirim, B. Ericsti, Y. Demir, Automatic recognition system of underlying causes of power quality disturbances based on S-transform and extreme learning machine, Int. J. Electr. Power Energy Syst. 61 (2014) 553–562.

[120] Y. Jimenez, C. Duarte, J. Petit, G. Carrillo, Feature extraction for nonintrusive load monitoring based on S-Transform, in: 2014 Clemson University Power Systems Conference, 2014, pp. 1–5.

[121] Z.-L. Gaing, Wavelet-based neural network for power disturbance recognition and classification, IEEE Trans. Power Deliv. 19 (4) (2004) 1560–1568.

[122] S. Mishra, C.N. Bhende, B.K. Panigrahi, Detection and classification of power quality disturbances using S-transform and probabilistic neural network, IEEE Trans. Power Deliv. 23 (1) (2007) 280–287.

[123] M. Valtierra-Rodriguez, R. de Jesus Romero-Troncoso, R.A. Osornio-Rios, A. Garcia-Perez, Detection and classification of single and combined power quality disturbances using neural networks, IEEE Trans. Ind. Electron. 61 (5) (2013) 2473–2482.

[124] H. Eri\csti, A. Uçar, Y. Demir, Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines, Electr. Power Syst. Res. 80 (7) (2010) 743–752.

[125] P.G. V AXelberg, I.Y.-H. Gu, M.H.J. Bollen, Support vector machine for classification of voltage disturbances, IEEE Trans. Power Deliv. 22 (3) (2007) 1297–1303.

[126] S. Santoso, E.J. Powers, W.M. Grady, A.C. Parsons, Power quality disturbance waveform recognition using wavelet-based neural classifier. II. Application, IEEE Trans. Power Deliv. 15 (1) (2000) 229–235.

[127] E. Gursoy, D. Niebur, Independent component analysis for harmonic studies, Time-Varying Waveform Distort. Power Syst 6 (2009) 217.

[128] A. Asheibi, D. Stirling, D. Sutanto, analyzing harmonic monitoring data using supervised and unsupervised learning, IEEE Trans. power Deliv. 24 (1) (2008) 293–301.

[129] M.A.S. Masoum, M. Ladjevardi, A. Jafarian, E.F. Fuchs, Optimal placement, replacement and sizing of capacitor Banks in distorted distribution networks by genetic algorithms, IEEE Trans. Power Deliv. 19 (4) (2004) 1794–1801.

[130] W. Lenwari, M. Sumner, P. Zanchetta, The use of genetic algorithms for the design of resonant compensators for active filters, IEEE Trans. Ind. Electron. 56 (8) (2009) 2852–2861.

[131] O. Maroufi, A. Choucha, L. Chaib, Hybrid fractional fuzzy PID design for MPPT- pitch control of wind turbine-based bat algorithm, Electr. Eng. 102 (2020) 2149–2160.

[132] F. Wu, J. Wang, Z. Sun, T. Wang, L. Chen, X. Han, An optimal wavelet packets basis method for cascade hydro-pv-pumped storage generation systems to smooth photovoltaic power fluctuations, Energies 12 (24) (2019).

[133] A. Naderipour, et al., Optimal designing of static var compensator to improve voltage profile of power system using fuzzy logic control, Energy 192 (2020), 116665.

[134] E.A. Nagata, et al., Real-time voltage sag detection and classification for power quality diagnostics, Measurement 164 (2020), 108097.

[135] M.H.J. Bollen, et al., Bridging the gap between signal and power, IEEE Signal Process. Mag. 26 (4) (2009) 12–31.

[136] Y. Deng, L. Wang, H. Jia, X. Zhang, X. Tong, A deep learning method based on bidirectional WaveNet for voltage sag state estimation via limited monitors in power system, Energies 15 (6) (2022) 2273. Mar.

[137] M.H.J. Bollen, I.Y.H. Gu, P.G.V. AXelberg, E. Styvaktakis, Classification of underlying causes of power quality disturbances: deterministic versus statistical methods, EURASIP J. Adv. Signal Process. 2007 (1) (2007), 079747. Dec.

[138] R. Turovic, et al., Training an LSTM voltage sags classificator on a synthetic dataset, in: 2021 21st International Symposium on Power Electronics (Ee), 2021, pp. 1–6.

[139] C. Ge, et al., A new convolutional network structure for power quality disturbance identification and classification in micro-grids, IEEE Access 9 (3) (2019) 3250–3260.

[140] O. Cortes-Robles, E. Barocio, A. Obushevs, P. Korba, F.R.S. Sevilla, Fast-training feedforward neural network for multi-scale power quality monitoring in power systems with distributed generation sources, Meas. J. Int. Meas. Confed. 170 (2021).

[141] A.-D. Gonzalez-Abreu, M. Delgado-Prieto, R.-A. Osornio-Rios, J.-J. Saucedo- Dorantes, R.-J. Romero-Troncoso, A novel deep learning-based diagnosis method applied to power quality disturbances, Energies 14 (10) (2021).

[142] A. Tjader, I.Y.H. Gu, M.H.J. Bollen, S.K. Ronnberg, Performance evaluation for frequency estimation of transients using the ESPRIT: measured noise versus white noise, in: 2008 13th International Conference on Harmonics and Quality of Power, 2008, pp. 1–8.

[143] R. Machlev, A. Chachkes, J. Belikov, Y. Beck, Y. Levron, Open-source dataset generator for power quality disturbances with deep-learning reference classifiers, Electr. Power Syst. Res. 195 (2021).

[144] I. P. S. on B. D. & A. for P. Systems, "Power quality data set.".

[145] J. M. S.-F. J. J. Gonza´lez-de-la-Rosa, A. Agüera-P´erez, J. C. P. Salas, "Real-life power quality transients," 2018.

[146] A. Mariscotti, Data sets of measured pantograph voltage and current of European AC railways, Data Br 30 (2020), 105477. Jun.

[147] D. Di DomenicoPetrKorba, "One-year recordings of electric vehicle charging fleet.".

[148] M.H. Bollen, Understanding Power Quality Problems: Voltage Sags and Interruptions, IEEE Press, 2000.

[149] E. Styvaktakis, "Automating power quality analysis," Chalmers University of Technology (Sweden), 2020.

[150] J. Choo, S. Liu, Visual analytics for explainable deep learning, IEEE Comput. Graph. Appl. 38 (4) (2018) 84–92. Jul.

[151] A. Venzke, S. Chatzivasileiadis, Verification of neural network behaviour: formal guarantees for power system applications, IEEE Trans. Smart Grid 12 (1) (2021) 383–397. Jan.

[152] R. Machlev, M. Perl, J. Belikov, K. Levy, Y. Levron, measuring explainability and trustworthiness of power quality disturbances classifiers using XAI - explainable artificial intelligence, IEEE Trans. Ind. Inform. (2021) 1. –1.

[153] W. Samek, G. Montavon, S. Lapuschkin, C.J. Anders, K.-R. Muller, Explaining deep neural networks and beyond: a review of methods and applications, Proc. IEEE 109 (3) (2021) 247–278. Mar.

[154] Y. Lai, A comparison of traditional machine learning and deep learning in image recognition, J. Phys. Conf. Ser. 1314 (1) (2019), 012148. Oct.

[155] S. Ryu, H. Choi, H. Lee, H. Kim, Convolutional autoencoder based feature extraction and clustering for customer load analysis, IEEE Trans. Power Syst. 35 (2) (2020) 1048–1060. Mar.