Open Access Research Journal of
Engineering and Technology
Journals home page: https://oarjpublication/journals/oarjet/
ISSN: 2783-0128 (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(REVIEW ARTICLE)

Check for updates

Review of machine learning applications to power systems studies

Omorogiuwa Eseosa * and Ashiathah Ikposhi

Electrical/Electronic Engineering, Faculty of Engineering University of Port Harcourt, Rivers State, Nigeria.

Open Access Research Journal of Engineering and Technology, 2021, 01(01), 021–031

Publication history: Received on 15 February 2021; revised on 16 March 2021; accepted on 20 March 2021

Article DOI: https://doi.org/10.53022/oarjet.2021.1.1.0101

Abstract

The complexity of electric power networks from generation, transmission and distribution stations in modern times has resulted to generation of big and more complex data that requires more technical and mathematical analysis because it deals with monitoring, supervisory control and data acquisition all in real time. This has necessitated the need for more accurate analysis and predictions in power systems studies especially under transient, uncertainty or emergency conditions without interference of humans. This is necessary so as to minimize errors with the aim targeted towards improving the overall performance and the need to use more technical but very intelligent predictive tools has become very relevant. Machine learning (ML) is a powerful tool which can be utilized to make accurate predictions about the future nature of data based on past experiences. ML algorithms operate by building a model (mathematical or pictorial) from input examples to make data driven predictions or decisions for the future. ML can be used in conjunction with big data to build effective predictive systems or to solve complex data analytic problems. Electricity generation forecasting systems that could predict the amount of power required at a rate close to the electricity consumption have been proposed in several works. This study seeks to review machine learning applications to power system studies. This paper reviewed applications of ML tools in power systems studies.

Keywords: Machine Learning; Big Data; Data Analytic; Power systems

1. Introduction

The electric power system is one of the most complex systems ever built by mankind. Economy and security of supply are among the factors to be managed for the successful operation of the network. Simply put, the security of a power system is descriptive of the power system's capability to ensure the continuity of supply within the likelihood of a diversity of disturbances (variations in consumer demand, internal failures, external perturbations like lightning, storms, etc.). Security is usually achieved through two complementary strategies [1]. Preventive control and emergency control. The former is carried out by human operators in order to maintain the system in a state where it can withstand disturbances while the latter acts automatically after a disturbance has occurred in order to minimize its consequences. Since disturbances are intrinsically random, preventive control is essentially aimed at balancing the economic cost of normal operation against the risk of instability/insecurity. On the other hand, emergency control is aimed at reducing the severity of instabilities. Worthy of note is the fact that due to increased competitive pressure in the electricity industry and the possibilities offered by modern communication and computing technologies, the trend in power systems is to rely more on emergency control. One of the challenges in the design of emergency controls is defining appropriate criteria against which a prediction in real time whether a system is in the process of losing stability or not can be done. This implies the selection of appropriate real-time measurements (among a multitude of possible ones) and combining these in order to formulate detection rules. Emergency controls design process essentially relies on numerical simulation of the power system under various conditions likely to drive it towards instability. Due to the rapidly growing amounts of computing power, techniques such as the Monte-Carlo sampling method can be used to achieve screening of very large samples (several thousands) of large-scale simulations, yielding large data bases of

* Corresponding author: Omorogiuwa Eseosa
Electrical/Electronic Engineering, Faculty of Engineering University of Port Harcourt, Rivers State, Nigeria.

Copyright © 2021 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution Liscense 4.0.

simulation results. These data can then be exploited using automatic learning such as machine learning with a view to extracting useful information.

2. Literature review

An efficient forecasting system for the generation of electricity based on machine learning with big data has been proposed by [2]. In the study, it is shown that there is a direct correlation between the quantity of power generated and the quantity of resources (such as coal, gas, nuclear, petroleum, oil, and renewable energy) that is used to generate electricity. Studies show that predicting power generation could be used in providing vague information about power demand and probably the need to increase the quantity to be imported from neighbouring countries. The prediction is challenging as a result of the accuracy requirement. The prediction is even more cumbersome when the data sets are enormous and have high noise and high volatility. Several forecasting methods using different types of ML algorithms have been proposed to deal with electricity forecasting problems. These algorithms include Fuzzy Neural Networks (FNN), Gray Algorithm (GA), Gray Markov Model (GMM), and Support Vector Regression (SVR) [2]. It has been investigated that these models have shown impressive results in terms of forecasting. The large penetration of renewable energy resources (RES) such as wind and solar types has increased the uncertainty of generation and has increased forecasting of power generation in order to allocate resources that produce the power as well as the estimation of the quantity to be imported from neighbouring areas. This is of altmost importance. In [3], it has been reported that power transformers are one of the most expensive and critical equipment in a power network and consequently, it is necessary to develop effective techniques for the monitoring and diagnosing of the transformer insulation system. Furthermore it was put forth that in recent years, a number of techniques such as Dissolved Gas Analysis (DGA), Polarization and Depolarization Currents (PDC) measurements, and Frequency Domain Spectroscopy (FDS) have been adopted across utilities for transformer diagnosis. However, there are still considerable challenges remaining in correlating measured data to actual transformer insulation condition. The study developed Machine Learning (ML) algorithms namely; Self-Organising Maps (SOM) and Support Vector Machines (SVM) for automatically analysing measurement data and making diagnosis on transformer insulation systems. The key advantages of these algorithms is their capability of acquiring the knowledge of underlying statistical dependency between archived data and the conditions of corresponding transformers and making use of such knowledge to assist in transformer insulation diagnosis. In [4], ML based power transformer lifetime prediction is proposed. It is reported that the prediction of the remaining life of high voltage power transformers is important for energy companies because of the need for planning maintenance and capital expenditures. ML models were used for the analysis of the transformers as well as for predicting lifetime of the transformers. It is demonstrated that the integration of ML models with experimental models improved transformer lifetime estimation. A parametric lifetime model is used to predict the lifetime distribution of the individual transformers. A statistical procedure is developed for computing the remaining life of individual transformers currently in use. By using ML algorithms, the power transformer loss values are delivered to the end users via email or any other means of communication before damage to the transformers can occur. In so doing, end users can be alerted about the remaining lifetime of the transformer in a bid to avoid failure of the transformer by providing proper maintenance in advance. In [5], a new type of reinforcement learning algorithm known as “fitted Q iteration” is considered for the design of some intelligent agents for power system control. The main characteristic of the algorithm is to formulate the reinforcement learning problem as a sequence of standard supervised learning problems. It is reported that “fitted Q iteration” has the potential to address real world power system control problems due to its ability to generalize information. Two problems were identified to be used in circumventing the problems associated with the use of “fitted Q algorithm”. Encouraging simulation results are obtained from a strategy intended to control in real-time a Thyristor-Controlled Series Capacitor (TCSC) installed in a 4-machine power system. In [6], it is reported that modern power systems with deeper penetration of renewable energy generation as well as higher level of demand side participation is faced with increasing degrees of complexities and uncertainties. It is also reported that reliable operation of the grid calls for improved techniques in system modelling, assessment, and decision. While it is now possible for system operators to have access to fine-grained electricity data through the use of smart meters and advanced sensing technologies, there is an urgent need for efficient and near-real time algorithms to analyse and make better use of these available data. Recent advances in machine learning (ML) algorithms especially the giant leaps in deep learning makes ML a good tool for solving a series of data driven problems in power systems. For instance, ML methods such as Recurrent Neural Networks (RNN) can find straight forward applications in wind, solar power and building load forecasting. ML has been applied to power grid outage detection. High Voltage Alternating Current (HVAC) control and grid protection policy formulation problems can also solved using ML approaches. In [7], it is reported that increased use of renewable energy liberalization of energy markets and most importantly the integration of various monitoring, measuring, and communication infrastructure into modern power system networks offers the opportunity for building a resilient and efficient power grid network at various voltage levels. Also of concern are various threats of instability and insecurity in the form of cyber-attacks, voltage instability, power quality (PQ) disturbances amongst others to the complex networks. Furthermore, ML techniques such as Artificial Neural Networks (ANNs), Decision Trees

(DTs), and Support Vector Machines (SVMs) have been used in effective decision-making and control actions in the secure and stable operations of the power system. The paper presents a comprehensive review of the most recent studies where machine learning technologies (MLTs) were developed for power system security and stability especially in cyber-attack detection, PQ disturbance studies, and dynamic security assessment studies. The aim of the study is to highlight the methodologies, achievements, and more importantly the limitations in the classifier design, data set, and test systems employed in reviewed publications. A brief review of Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) approaches to transient stability assessment is also presented. The research gap highlighted states that despite the enormous accomplishments in terms of power system investigations in the areas of security and stability studies, a number of challenges still remain unresolved. The prediction and detection accuracy of MLTs are known to depend majorly on the quality and quantity of the data set and test systems employed. However, sometimes due to the non-availability and inadequacy of realistic power systems, data from real power stations and field devices, scholars and researchers have been restricted to the use of simulated data sets and the development of scalable test beds which have shown inconsistency in predictions and classification. Apart from the number of input data sets, another important factor that is peculiar to ML applications is the tuning of the parameters. The rigorous events performed in tuning the parameters so as to achieve the desired results means MLTs approaches require a high level of expert interaction. Also, MLTs approaches can sometimes be time-consuming. Furthermore, most articles in the literatures usually assume that PMU data are complete, accurate and available for online use. In practice, the measurements may not always be available due to jamming, malfunctioning, and attacks. Finally, it was suggested that future research work on MLT based approach to power system security and stability studies should focus on detailed validation of the approaches using large-scale test systems which have similar characteristics as modern power systems. In [8], it was found that modern power systems required real time monitoring and fast control to be protected from faults on power transmission lines. As the Smart Grid becomes a reality, the installation of high-quality sensors such as remote terminal units, phasors measurement units, smart meters and other measuring devices tend to generate considerable amount of heterogeneous data required for the operational control and performance analysis of the grid. Conventional time-domain techniques had a tendency to be inefficient from the computational point of view and had a possibility of not meeting real-time application specification. However, with the aid of ML algorithms it is possible to learn without directly programming the data and once exposed to new data, can respond independently. ML approaches such as Artificial Neural Networks (ANNs), Decision Trees (DTs), Deep Learning Models (DLM) etc. are capable of providing interesting information on safety in power systems. The paper proposed a classification and detection of faults in power systems based on machine learning. The study gave a list of ML techniques for fault classification. These are enumerated;

- Support Vector Machine
- Bayesian Learner
- Sequential Minimal Optimization
- Logistic Regression
- Decision Tree
- K. Nearest Neighbour

In [9], four powerful and popular ML techniques (Bagging Classifier, Boosting Classifier, Radial Basis Function Classifier (RBF), Naïve Bayes Classifier (NBC)) for identifying and locating faults over a 600 km long power transmission line was introduced. In the paper, eleven (11) faults are found to be detected, predicted, and located. The results from the experiment suggest that RBF, Bagging Classifier, and NBC techniques could be used for fault type prediction as high prediction accuracy is achieved. For fault location prediction experiment, the attained prediction results were not as accurate as in the case of fault type prediction. However, the RBF, NBC and the Bagging Classifier achieved the highest prediction accuracy. Finally, it was reported that ML techniques could be used for identifying the transmission line faults. Results for location prediction accuracy may need to be improved upon in order to achieve precise fault location. In [10], comprehensive review of fault detection, classification, and location in transmission lines is presented. Before introducing methods used for fault detection, classification, and location an overview of feature extraction methods was presented. The ground work for fault identification algorithms, various transforms, along with dimensionality reduction techniques was discussed. Newly developed ideas and their comparisons with some noteworthy aspects regarding fault detection were also discussed. It is presented that ML methods are widely employed by researchers for fault type classifications. However, in addition to Support Vector Machines (SVM), Fuzzy Inference System (FIS), Artificial Neural Networks (ANNs), Decision Trees (DT), Deep Learning based algorithms such Convolutional Neural Networks (CNN) and Restricted Boltzmann Machine Learning (RBML) are recommended for fault classification. Fault location finding algorithms were discussed alongside Artificial Intelligence (AI)-based methods. Deep learning methods were recommended for future fault location finding methods due to increased involvement of communication and computation in transmission systems. In [11], the focus was on detecting, classifying, and locating faults in a power system using Artificial Neural Networks (ANNs). Feed-forward Neural Networks were employed and trained with back propagation algorithms. The WSCC 9 bus test system was modelled in MATLAB/SIMULINK and used to validate the proposed fault detection system. Results from the simulation indicate that ANNs can be used for detecting, classifying and locating faults. Also, results showed that ANNs could relatively detect, locate, and classify faults even for places it was not trained for. Depending on the type of fault, networks with different number of neurons in the hidden layer can be used. Single phase-to-ground fault can be detected and located with the smallest number of neurons in the hidden layer using only five (5) Networks for the detection, location, and classification of two-phase and two-phase-to-ground fault require a minimum of 10 neurons in the hidden layer. Three-phase faults require neural networks with 30 to 35 neurons in the hidden layer. It was reported that further increase in neurons in the hidden layer did not lead to an improvement in the results. In [12], it is reported that driven by the large amounts of data involved, the application of deep learning frameworks was extended to performing automatic disturbance classification. In order to achieve this, a set of measurements from several Phasors Measurement Units (PMUs) installed in low voltage sections of an interconnected system was used from which representative patterns are extracted so as to endow a classifier of knowledge related to system disturbances. In particular, the strategies adopted consists of the application of Multi-Layer Perceptron (MLP), Deep Belief Networks and Convolutional Neural Networks (CNN), the latter having out-performed the others in terms of classification accuracy. Additionally, these architectures were implemented in both the CPU and the GPU to ascertain the resulting gains in speed. In [13], the application of ANNs for the detection and classification of faults on a three-phase transmission line system is presented. The method developed utilizes both the three-phase voltages as well as three-phase currents as inputs to the neural network. The inputs were normalized with respect to their pre-fault values respectively. The results obtained were for line-to-ground faults only. The ANNs studies adopted in this work used the back propagation neural network architecture. The simulation results obtained prove that satisfactory performance was achieved by all the proposed neural networks and are practically implementable. The importance of choosing the most appropriate ANN configuration in order to get the best performance is emphasised in the study. The following important conclusions that can be drawn from the research are;

- ANNs are a reliable and effective method for an electrical power system transmission line fault classification and detection in view of the increasing dynamic connectivity of the modern electrical power transmission systems.
- The performance of an ANN should be analysed properly. (In particular, a neural network structure and learning algorithm should be analysed properly before choosing it for a practical application).
- Back propagation neural networks deliver good performance when trained with large training data sets which is easily available in power systems.

In [14], a method based on a combination of wavelet singular value and Fuzzy Logic (FL) is presented for fault detection and fault classification in power transmission systems. The results show that the proposed indices for FL are sensitive to variations. The method is robust to parameter variations such as fault type, fault inception location, fault resistance and power angle and can properly detect faults. The proposed algorithm has proven to be a convenient and rapid method for fault detection and fault classification in different conditions and is able to detect and classify faults and determine a healthy phase from a faulty phase in less than 10ms after fault inception. In [15], a reliable scheme for the detection, classification, and location of faults on transmission lines is developed. The scheme combines the feature extraction capability of the discrete wave transform (DWT) and the intelligent classification capability of the Adaptive-Neuro Fuzzy Inference System (ANFIS). The developed DWT-ANFIS model is tested and the results compared with Impedance-ANFIS model. Faults are detected within 8ms that is less than one complete cycle from fault inception to prevent equipment damage and prolonged power outage. In [16], a wavelet transform-based approach to detect and classify different shunt faults that may occur in transmission lines is presented. The algorithm is based mainly on calculating the RMS values of the wavelet coefficients of current signals at both ends of the transmission lines over a moving window length of half cycle. The current signals are analysed with “dB4” wavelet to obtain detail coefficients and compared with threshold values to detect and classify the faults. To illustrate the effectiveness of the proposed technique, extensive simulations using PSCAD/EMTDC and MATLAB have been carried out for different types faults considering wide variations of resistances, inception angle, and loading levels. The study proposes that the techniques investigated are well suited for implementation in digital distance protection schemes. In [17], an endeavour aimed at the automation of power system fault identification using information conveyed by the wavelet analysis of power system transients is proposed. Probabilistic Neural Network (PNN) is also used in the study. The focus of the study is on the identification simple power system faults. Also performed was a wavelet transform of the transient disturbance caused as a result of the occurrence of a fault. The detail coefficient for each type of simple fault is characteristic in nature. PNN was used in distinguishing the detailed coefficients and hence the fault. The application of wavelet transform to determine the type of fault and its automation incorporating PNN could achieve an accuracy of 100% for all type of faults. Back propagation algorithm was limited in distinguishing the entire phase-to-ground and double line-to-ground fault. In [18], the study says that transmission line relaying involves three major tasks.

- Fault Detection
- Fault Classification
- Fault Location

These three tasks must be done as fast and as accurate as possible so as to de-energize the faulted line. Against this background and others, the study proposed a novel method for transmission line fault detection and classification using oscillographic data. The fault detection and its clearing were determined based on a set of rules obtained from the current waveform analysis in time and wavelet domains. The method is able to single out faults from other power quality disturbances such as voltage sags and oscillatory transients which are common in power system operation. An artificial neural network (ANN) classifies the fault from the current and voltage waveform pattern recognition in the time domain. It is reported that the method was used for fault detection and classification from real oscillographic data of a Brazilian utility company with excellent results obtained. In [19], the focus of the paper was on developing a single artificial neural network (ANN) to detect and classify a fault on Nigeria's 33 kV electric power transmission lines. The study employed a feed-forward artificial neural network with back propagation algorithm in developing a fault detector-classifier. Simulation results have been provided to demonstrate the efficiency of the developed intelligent systems for fault detection and classification on 33 kV lines. The performance of the detector-classifier is evaluated using the mean square error (MSE) and the confusion matrix. The system achieved an acceptable MSE of 0.00004279 and an accuracy of 95.7% showing that the performance of the developed intelligent system is satisfactory. In comparison with other systems in the literature concerning Nigeria's transmission lines, the developed system in this work is adjudged to be better. In [20], the focus of the study was on discussing the possibility of using deep learning architecture using convolutional neural network (CNN) for real-time power system fault classification. The work studied fault classification only and not about localization and was aimed at classifying power system voltage signal samples in real-time and determining whether it belong to faulted or non-faulted state. The data is produced by simulating a simple two-bus power system with three-phase balanced load. The voltage signal is measured at the beginning of the line between the two buses while the fault occurs at half of the line length between the two buses. In the first step, wavelet transform is used to extract the fault harmonics using dB4 daubechies mother wavelets. A sample window of fixed size is slid over the wavelet detail at decomposition level 4 which seems to be a suitable choice. After normalization, the generated training samples are fed into the convolutional neural network (CNN) for learning procedure. The CNN learns fault features of the power system through training by faulted and non-faulted samples to finally classify samples from a test set. In conclusion, it was shown that CNNs could successfully learn power system fault's features and classify those correctly. For certain training scenarios (only faulted test samples) a per phase testing accuracy of over 85% is achieved. This has been validated by a simulation of a two-bus power system with balanced load. In [21], it was reported that in consideration of machine learning applications, it has become easier to handle complex power system challenges. The traditional techniques are not computationally promising solutions since they are limited in capacity to manage the massive amounts of data (including chunks of heterogeneous datasets) coming from measurement units such as smart meters, and phasors measurement units (PMUs). The study said there was in existence several advanced, efficient, and intelligent learning algorithms that have been developed to improve the accuracy of solutions to many real-world problems in a diversity of areas such as voltage stability, power flow management, rotor system diagnosis to mention a few. The study indicates that supervised machine learning classifications are more in use compared to other methods. What this means is that classification algorithms yield more benefits to problems than others. The study concludes by saying that it can be inferred that by applying machine learning to electrical engineering problems, difficult issues are not only simplified but also results secured are also reliable and precise. In [22], Machine learning (ML) techniques for power system security assessment is presented. It is reported that modern electricity grids continue to be vulnerable to large scale blackouts. It is also reported that as all states leading to large scale blackouts are unique, there is no algorithm to identify pre-emergency states. Moreover, numerical conventional methods are computationally expensive which makes it difficult for them to be used for online security assessment. Machine learning techniques with their pattern recognition, learning capabilities, and high-speed identification of potential breach of security boundaries can offer an alternative approach. The study put forth that during the last 10 years, events in the North American continent, in Europe, and in Asia has clearly demonstrated an increasing likelihood of large blackouts. This indicates that the security monitoring and control of power systems need to be improved. The paper presents a novel method for online security assessment using machine learning techniques. Multiple machine learning techniques such as Artificial Neural Networks (ANNs), Support Vector Machines (SVM), Decision Trees (DTs) are first trained online using the resampling cross validation method. Resampling the training samples allows to know when a poor choice of values of the machine learning tuning parameters is being made. The best model of the ML technique is selected based on its performance. For the online application, the final of the best of the best ML is used as the candidate technique with the best performance. If required the final ML is checked and updated in order to account for new changing system states as accurately as possible. The results obtained in the work showed that the proposed approach can identify potentially dangerous states with high accuracy, and if required the final ML model can produce an alarm for triggering emergency and protection systems. In [23], an exploration of how

Reinforcement Learning (RL) as applied to the control of power systems is presented. Also presented is a description of some challenges in power system control as well as how some of those challenges can be overcome using RL techniques. The difficulties associated with application of RL methods to power system control as well as the strategies to overcome them are also presented. Two RL modes are considered.

- The online mode in which interaction occurs with the real power system.
- The offline mode in which interaction occurs with a simulation of the model of the real power system.

Two case studies made on a 4-machine power system model are presented where in the first case, the design of a dynamic brake controller by means of RL algorithms used in the offline mode is considered. The second case concerns RL methods used in the online mode when applied to control a thyristor-controlled series capacitor aimed at damping power system oscillations. The RL methods can reveal themselves to be an interesting tool for power system agents' design for reasons enumerated;

- RL methods do not make any strong assumptions on the system dynamics. In particular they can cope with partial information, nonlinear and stochastic behaviours. They can therefore be applied to the design of many practical types of control schemes.
- This method learn closed loop control laws which is ascertained to be robust. This aspect is important notably when the real power system is faced with situations that were not accounted for in the simulation model.
- RL methods open avenues to adaptive control since the RL driven agents learn continuously and can adapt to changing operating conditions.
- The method can be used in combination with traditional control methods to improve performances.

In [24], it was presented that power system protection includes the process of identifying and correcting faults before fault currents cause damage to utility equipment or customer property. In distribution systems where the number of measurements is increasing, there is an opportunity to improve upon fault classification techniques. Fault classification using machine learning (ML) techniques and quarter-cycle signatures is presented. Separate voltage and current-based feature vectors are defined using multi-resolution analysis are input to a two-stage classifier. The classifier is trained and tested on experimental fault data using a Reconfigurable Distribution Automation and Control software/hardware laboratory. Results show Non-linear and even Non-contiguous decision regions on a fault plane using a phase voltage-based feature. An accurate classifier for determining the grounding status of multiphase faults using a neutral current-based feature. In [25], the focus is on the concept of using reinforcement learning to control the power system's unit commitment and economic dispatch problem. The idea of reinforcement learning strives to present an ever-optimal system even when there are load fluctuations. This is done by training the agent (system) thereby enriching its knowledge base which ensures that even without manual intervention, all the available resources are judiciously used. Also, the agent learns to reach the long-term objective of minimizing cost by autonomous optimization. A model-free reinforcement learning method called "Q Learning" is used to find the cost at various loadings and is compared with the conventional priority list method and the performance improvement due to Q learning is proved. For future directions, it was presented that single agent reinforcement learning can be extended to multi-agent reinforcement learning to accommodate other types of renewable energy resources such as solar and wind along with thermal units. In [26], an active machine learning (ML) technique for monitoring the voltage stability in transmission systems is presented. It is shown that ML algorithms may be used to supplement the traditional simulation approach. However, they suffer from difficulties associated with online ML model update and offline training data preparation. An active learning solution to enhance existing ML applications by actively interacting with the online prediction and offline training processes is presented. The method identifies operating points where ML predictions based on power system measurements contradict with actual system conditions. By creating the training set around the identified operating points, it is possible to improve the capability of ML tools to predict future power system states. The method also accelerates the offline training process by reducing the amounts of simulations on a detailed power system model around operating points where correct predictions are made. Experiments show a significant advantage in relation to the training time, prediction time, and number of measurements that need to be queried to attain high prediction accuracy. In [27], a summary of artificial intelligence (AI) and its increasingly widespread application control is presented. Compared with traditional technologies, AI technologies possess obvious advantages. The application of intelligent technology in electrical automation control systems can reflect the basic characteristics of high precision, high efficiency, and high coordination. The application of AI while achieving automatic control can greatly improve the operating efficiency quality of the control system. Intelligent control can also realize optimal allocation of resources, reduce resource cost investment, improve the economic benefits of related companies, and promote the sustainable development of a nation's electrical industry. In [28], the study addresses the on-going work of the application of machine learning (ML) to the dynamic security assessment of power systems. It lists several methods which have been applied to the Greek power system. These methods include;

- Offline Supervised Learning (Radial Basis Function Neural Network (RBFNN), Support Vector Machine (SVM), Decision Trees (DT)).
- Offline Unsupervised Learning (Self-Organizing Maps (SOM))
- Online Supervised Learning (Probabilistic Neural Network (PNN))

The results from the application of the ML methods show the accuracy and versatility of the methods. RBFNN and SVM perform not only classification of the system states but also regression and gives an estimation of the security criterion (voltage and/or frequency) value. This is very important as it can be used as a measure of the security margin of the system. Decision Trees on the other hand provide explicit rules to the system operator while inverse reading of the rules can also establish load shedding schemes when the safety of the system is jeopardized. The advantage of the SOM in comparison to the offline supervised learning methods is that its construction is independent of the security criterion applied. This means that when the classification criterion is change, the only modification required is the straightforward recalculation of the security indices for each of the map's node. On the contrary, a change in the classification would require the reconstruction of any of the offline supervised learning methods. The advantage of the online learning method such as PNN is that it can deal with changes in the structure of the power system without the need to completely retrain the PNN. In [29], it is presented that with increasing complexity, uncertainty, and data dimensions in power systems, conventional methods often meet with bottlenecks when trying to solve decision and control problems. Data-driven methods aimed at solving the decision and control problems are currently being extensively studied. Deep Reinforcement Learning (DRL) is one of such data driven methods and is regarded as real artificial intelligence (AI). DRL is a combination of deep learning (DL) and reinforcement learning (RL). DRL has achieved rapid development in solving sequential decision-making problems around theoretical, methodological, and experimental fields. In particular, DL obtains an objects attributes or characteristics from the environment while RL makes decisions with regards to the control strategies according to the information. Therefore, DRL can solve problems in large, high dimensional states, and action spaces. As power systems evolve, to reflect the smart grid, there is in existence new challenges such as the integration of renewable energy and liberalization of the electricity market which offer difficulties to traditional techniques trying to solve problems in these areas. AI methods such as DRL can solve problems arising from these. In [30], a new control architecture for future power distribution system protective relay setting is envisioned. With increased penetration of distributed energy resources, at the end-user level, it has been recognized as a key engineering challenge to redesign the protective relays in the future distribution system. Conceptually, these protective relays are the discrete ON/OFF control devices at the end of each branch and node in a power network. The key technical difficulty lies in how to set up the relay control logic so that the protection could successfully differentiate heavy loads and faulty operating conditions. The study proposes a new nested reinforcement learning approach to take advantage of the structural properties of distribution networks and develop a new set of training methods for tuning protective relays.

3. Machine learning

Machine learning is the study of computer algorithms that improve automatically with time through experience by the use of data obtained for the study of interest. These algorithms are used to build a model (mathematical, pictorial etc.) based on sample data known as "training data" in order to make predictions or decisions without being explicitly programmed to do so (i.e. it can easily be adopted to solve a particular task). Machine learning involves computer learning from data provided so that they can carry out certain specific tasks. For simple task assigned to computers, it is possible to programme algorithms notifying the machine on how to execute all steps required to solve the problems at hand; On the computer's part no learning is needed. For more advanced tasks, it can be challenging for a human to manually create the needed algorithms. In practice, it can turn out to be more effective to help the machine develop its own algorithm rather than having human programmers specify every needed step. Machine learning (ML) is a subfield of artificial intelligence which deals with the study of algorithms which learn from past experiences. The goal of machine learning is to give the attribute of intelligence to computers and ultimately machines. Machine learning involves computers discovering how to carry out tasks without being explicitly programmed to do so. It involves computers learning from data provided so that they can carry out certain task. For simple tasks assigned to computers, it is possible to programme algorithms telling the machines how To solve the problem at hand. On the computer's part no learning is needed. For more advanced tasks, it can be challenging for a human designer to manually create the needed algorithms. In practice it can be more effective to help the machine develop its own algorithm rather than having human programmers specify every needed step. Machine learning has overlapping similarities with expert systems in which a computer system emulates the decision-making ability of a human expert. Expert systems are designed to solve complex problems by reasoning through bodies of knowledge. An expert system is composed of two subsystems: the inference engine and the knowledge base. The knowledge base represents facts and rules. The inference engine applies the rules to the known facts to determine new facts.

3.1. Classification of machine learning algorithms

Machine learning techniques are conventionally divided into three broad categories depending on the nature of the "signal" of "feedback" available to the learning system. These techniques are;

3.2. Supervised learning

Supervised learning algorithms build a mathematical model of a set of data that contains both the input and the desired outputs given by a "teacher". The data is known as the training data, and consists of a set of training examples [31]. Each training example has one or more inputs and the desired outputs, also known as the "supervisory" signal. In the mathematical model, each training example is represented by an array or vector, sometimes called a feature vector, and the training data is represented by a matrix. Through iterative optimization of an objective function, supervised learning algorithms learn a function that can be used to predict the output associated with new inputs. An optimal function will allow the algorithm to correctly determine the outputs for inputs that were not part of the training data. An algorithm that improves the accuracy of its output or predictions over time is said to have learned to perform that task. Types of supervised learning algorithms include active learning, classification, and regression.

3.3. Unsupervised learning

Unsupervised learning algorithms take a set of data that contains only inputs and find structure in the data like grouping or clustering of data points. The algorithms therefore learn from test data that has not been labelled, classified, or categorized. Instead of responding to feedback, unsupervised learning algorithms identify commonalities in the data and react based on the presence or absence of such commonalities in each new piece of data.

3.4. Semi-supervised learning

Semi-supervised learning falls between unsupervised (without any labelled training data) and supervised learning (with completely labelled training data). It has been found that unlabelled when used in conjunction with a small amount of labelled data, can produce a considerable improvement in learning accuracy.

3.5. Reinforcement learning

This is an area of ML concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. In ML, the environment is typically represented as a Markov Decision Process (MDP). Reinforcement learning algorithms do not assume knowledge of an exact mathematical model of the MDP and are used when exact models are infeasible.

4. Models

These models provide a mathematical framework for learning. A model is human derived and is based on human observation and experiences. Figure 1 shows a single neuron model. Different models have different mathematical and pictorial representations.

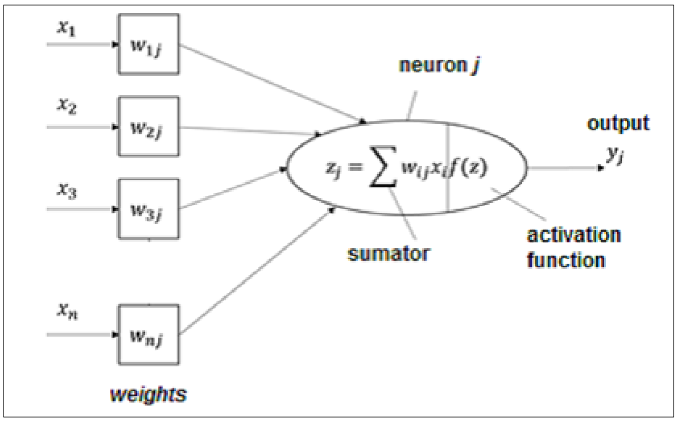

Figure 1 Model of a Single Neuron [11]

Performing ML endeavour involves creating a model, which is trained on some training data and then can process additional data to make predictions. Some of these models include;

- Artificial Neural Network: it is also called neural networks (NNs) as it is based on nodes connections that models the neurons used to transmit signals to other neurons formed in the same manner. This signal is inform of real numbers as input and the output is then computed as a non-linear function summing up its inputs.
- Decision Trees: it is a tool that supports decision making using the model that tree-like in nature.it considers consequences of taking certain decision while putting certain factors such as cost of resources and utility as well as possibilities of outcome of events and summarily shows an algorithm that gives statements that are conditional.
- Support Vector Machine: This is a kind of computer algorithm that learns using examples to assign labels to objects and it has been successfully applied in solving some technical problems in the areas of power systems studies [32]. In other words, it can be defined as a formulated mathematical algorithm for maximizing a specific mathematical expression as a function of a given set of collected data. Sequential Minimal Optimization is an optimization tool for training support vector machine which requires a very large optimization programme.
- Regression Analysis: This consist of a set of machine learning methods that enables prediction of an output variable based on the values of one or more predictor variable(s).it is termed machine learning because its task is to give an estimated value based on certain predictive features. It use test sets to validate its accuracy of prediction.
- Bayesian Network: It is defined as a probabilistic graphical model consisting of two parts (structure and parameters).the structure part is known as the directed acyclic graph for dependent and independent conditional expressions. Probability is used to represent all uncertainties within the model and it used to obtain the posterior probabilities based as well as additional recent information.
- Genetic Algorithms: It is a stochastic search algorithms that has found vast applications in the area of ML to power systems studies.it has cross over, mutation and fitness selection as its three components.it is used to obtain solutions required for optimization and search problems through inspiration obtained biologically with respect to these three components.
- Deep Learning: This is a sub-set of ML and it is mathematically a complex improvement of ML algorithm.it analyse data logically and then draws intelligent conclusion the way humans will make. This is achieved through supervised or unsupervised training pattern and it uses layer structure to achieve this.it is worthy of note that with the current high cloud computing and transfer learning ability of ML, the training time is now very fast and more accurate.
- K. Nearest Neighbour (KNN): It is used for analysing large data, thus it can accommodate very large training data. In KNN, a set of variable characterise each data point.it is an algorithm developed to store all cases of existing variables and then make appropriate classification based on similarities of the previous cases. Its advantage is that no assumption is about the data used, thus making the final results obtained to be very accurate.

5. Conclusion

The applications of ML to power system studies are increasing as the size and complexities of the system continue to expand. This review as considered various ML strategies as they apply to power systems studies including the critical identification of their knowledge gaps as investigated by each researcher. Moreso, ML techniques models and algorithms as well as their applications to various sectors of the power network (generation, transmission and distribution) were reviewed. Various subdivisions of ML which include supervised, unsupervised and reinforcement learning have been discussed alongside their associated algorithms as applied to power system problems .Furthermore, advantages and disadvantages of some ML techniques were also considered. Various models used to achieve ML was also considered. These include Artificial Neural Network (ANN), Decision Tree, Support Vector Machine, Regression Analysis, Bayesian Network (BN), Genetic Algorithm (GA), Deep Learning (DL) and the K-Nearest neighbour(KNN).

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest.

References

[1] P Geurts, L Wehenkel. Early Prediction of Electric Power System Blackout by Temporal Machine Learning, A.A.A.I Technical Report, WS-98-07. 1998.

[2] MA Rahman, A Ismailpour, J Zhao. Machine Learning with Big Data: An Efficient Electricity Generation Forecasting System, Published in Big Data Research. 2016; 5: 9-15.

[3] H Ma, TK Saha, C Ekanayake. Machine Learning Techniques for Power Transformer Diagnosis. 1-6.

[4] G Pooja, M Preethi, B Jeevabavatharani, IG Hemanandhini. Power Transformer Lifetime Prediction Through Machine Learning, Published in the International Journal of Creative Research Thoughts. 2020; 8(7): 1667-1670.

[5] D Ernst, M Glavic, P Geurts, L Wehenkel. Approximate Value Iteration in the Reinforcement Learning Context: Application to Electric Power System Control", Published in the International Journal of Emerging Electric Power System. 2005; 3(1): 1066, 1-35.

[6] Y Chen, Y Tan, D Deka. Is Machine Learning in Power Systems Vulnerable?", Published in the 2018 IEEE International Conference on Communications, Control and Computing Technologies for Smart Grids. 2018.

[7] OA Alimi, K Ouhada, AM Abu-Mahfouz. A Review of Machine Learning Approaches to Power System Security and Stability, Published in IEEE Access. 8: 113512-113531.

[8] D Baskar, P Selvam. Machine Learning Framework For Power System Fault Detection and Classification, Published in the International Journal of Scientific and Technological Research. 2020; 9(2).

[9] AN Hassan, PSP Eboule, B Twala. The Use of Machine Learning Techniques to Classify Power Transmission Line Fault Types and Location", Published in the 2017 International Conference on Optimization of Electrical and Electronics Equipment. 2017; 221-226.

[10] A Raza, A Benrabah, T Alquthami, M Akmal. A Review of Fault Diagnosing Methods in Power Transmission Systems", Published in the Journal of applied Sciences. 2020; 10: 1-27.

[11] A Karic, T Konjic, A Jahic. Power System Fault Detection, Classification, and Location Using Artificial Neural Networks", Published in the International Symposium on Advanced Electrical Power Systems.

[12] PEA Cardoso. Deep Learning Applied to PMU Data in Power Systems, Master of Engineering Thesis Submitted to the Department of Electrical & Computer Engineering, Faculty of Engineering, University of Porto.

[13] M Jamil, SK Karma, R Singh. Fault Detection and Classification in Electrical Power Transmission System Using Artificial Neural Network", Published in Springer Open Journal. 2015; 1-13.

[14] M Nayeripour, AH Rajaei, MM Ghanbarian, M Deghani. Fault Detection and Classification in Transmission Lines Based on a Combination of Wavelet Singular Values and Fuzzy Logic", Published in the Cumhuriyet University Faculty of Science, Science Journal. 2015; 36: 69-82.

[15] FO Ogban, KM Udofia, CN Kalu. Fault Detection, Classification and Location on 132 kV Transmission Line Based on DWT and ANFIS", Published in the Journal of Multidisciplinary Engineering Science and Technology. 2020; 7(6): 12367-12376.

[16] S Chakraborty, S Singh, A Bhalla, P Saxena, R Padarla. Wavelet Transform Based Fault Detection and Classification in Transmission Lines, Published in the International Journal of Research in Engineering and Applied Sciences. 2012; 2(5): 67-74.

[17] KH Kashyap, UJ Shenoy. Classification of Power System Faults Using Wavelet Transforms and Probabilistic Neural Networks.

[18] KM Silva, BA Silva, NSD Brito. Fault Detection and Classification in Transmission Lines Based on Wavelet Transform and Artificial Neural Networks, Published in the IEEE Transactions on Power Delivery. 21(4).

[19] PO Mbamaluikem, AA Awelewa, IA Samuel. An Artificial Neural Network-Based Intelligent Fault Classification System for the 33 kV Nigerian Transmission Line, Published in the International Journal of Applied Engineering Research. 2018; 13(2): 1274-1285.

[20] F Rudin, L Guo-Jie, K Wang. An Algorithm for Power System Fault Analysis Based on Convolutional Deep Learning Neural Networks, Published in the International Journal of All Research Education and Scientific Methods. 2017; 5(9): 11-18.

[21] SM Miraftabzadeh, M Pasetti. A Survey of Machine Learning Applications for Power System Analytics. 2019.

[22] NV Tomin, VG Kurbatsky, DN Sidorov, AV Zhukov. Machine Learning Techniques for Power System Security Assessment, Published in IFAC-Papers Online. 2016; 49-12: 445-450.

[23] D Ernst, M Glavic, L Wehenkel. Power System Stability Control: Reinforcement Learning Framework, Published in the IEEE Transaction on Power Systems. 2004; 1-9.

[24] NS Coleman, C Schegan, KN Miu. A Study of Power Distribution System Fault Classification with Machine Learning Techniques, Proceedings at the 2015 North American Power Symposium. 2015.

[25] L Raju, RS Milton, S Suresh, S Sankar. Reinforcement Learning in Adaptive Control of Power System Generation", Published in Procedia Computer Science. 2015; 46: 202-209.

[26] V Mulbasa, C Zheng, C Po-Chen, T Popovic, M Kezunovic. Voltage Stability Prediction Using Active Machine Learning. Proceedings from the IEEE transactions on Smart Grids. 2017; 1-8.

[27] C Jiang, X Xiong, T Zhu, J Cao, J Yu. Research on Application of Artificial Inteligence Technology in Electrical Automation Control, Published in the Journal of Physics. 2020; 1-6.

[28] EM Voumvoulakis, AE Gavoyiannis, ND Hatziargyriou. Application of Machine Learning on Power System Dynamic Security Assessment, Proceedings from 14th International Conference on Intelligent System Applications to Power System, ISAP. 2007; 118-123.

[29] Z Zhang, D Zhang, RC Qiu. Deep Reinforcement Learning for Power System Applications: An Overview, Published in the CSEE Journal of Power and Energy Systems. 2020; 6(1): 213-225.

[30] D Wu, Y Zheng, Q Qian. Nested Reinforcement Learning-Based Control for Protective Relays in Power Distribution Systems, Proceedings from the 2019 IEEE 58th Conference on Decision and Control. 2019; 1925-1930.

[31] https://en.wikipedia.org/wiki/Machine_learning.

[32] WS. Noble What is a Support Vector Machine, National Publishing Group. 2006.