Transformation of Smart Grid using Machine Learning

Salahuddin Azad
School of Engineering and Technology
Central Queensland University
Melbourne, Australia
s.azad@cqu.edu.au

Fariza Sabrina
School of Engineering and Technology
Central Queensland University
Sydney, Australia
f.sabrina@cqu.edu.au

Saleh Wasimi
School of Engineering and Technology
Central Queensland University
Melbourne, Australia
s.wasimi@cqu.edu.au

***Abstract*— With the advent of distributed and renewable energy sources, maintaining the stability of power grid is becoming increasingly difficult. Traditional power grid can be transformed into a smart grid by augmenting it with information and communication technologies, and machine intelligence. Machine learning and artificial intelligence can enable smart grid to make intelligent decisions and respond to sudden changes in customer demands, power outages, sudden drops and rises in renewable energy output or any other catastrophic events. Machine learning can also help capture customer consumption patterns, forecast energy demand and power generation of intermittent sources, and predict equipment failures. Reinforced learning can aid in making energy dispatch decisions and activate demand management signals in order to maintain balance of power supply and demand. The usage of wireless technologies in smart grid renders it vulnerable to cyber security threats. With the increase in data volume, it is now possible to employ machine learning for the detection and prevention of anomalous behaviour, intrusion, cyber-attacks, and malicious activities as well as data authentication. This paper reviews the application of different machine learning approaches that aims at enhancing the stability, reliability, security, efficiency and responsiveness of smart grid. This paper also discusses some of the challenges in implementing machine learning solutions for smart grid.**

Keywords—smart grid, machine learning, transformer loss of life, power quality, energy dispatch decision, electricity market operation, smart grid security

I. INTRODUCTION

With the introduction of distributed and renewable energy sources, it is becoming increasingly difficult to maintain the balance between demand and supply of power as well as the quality of power in the electricity network. As the traditional power grid is not designed to handle bidirectional power flow, electricity networks are struggling to handle backflow of power from distributed generation sources such wind and solar. The intermittent nature of the renewable energy sources is also making it difficult to maintain stable power flow in the electricity network. The traditional power grid can be transformed into an intelligent, automated and responsive power grid by augmenting it with information and communication technologies, and machine intelligence. This type of grid is often termed as *smart grid*. The motivation behind developing smart grid is to ensure stable, reliable, efficient, economical and sustainable generation, distribution and usage of conventional and renewable power [1]. Smart grid can offer the following benefits to the generators, transmission and distribution operators, and consumers.

- *Demand management*: Smart grid allows customers and network operators to manage electricity consumption to reduce peak demand and avoid network overloading.
- *Network stability*: Smart grid maintains the balance of supply and demand and regulates the voltage by varying generation and managing demand.
- *Network reliability*: Smart grid allows network operators to anticipate and locate faults and outages remotely that could lead to quicker restoration of power.
- *Empowering consumers*: Smart grid makes real-time consumption of data available so that customer can control their consumption to take advantage of the off-peak price and reduce power bill.
- *Integration of renewable energy sources*: Smart grid facilitates integration of intermittent renewable energy sources by making energy dispatch decisions and controlling loads in real-time.

Machine learning is a branch of artificial intelligence where machines learn automatically from data without the direct involvement of humans [2]. Machine learning algorithms embody the techniques that look for patterns and interrelations in the data and build a model that represent those patterns and interrelations. The learning experience enables the machine to predict future events based on past examples. Machine learning can be *supervised* or *unsupervised* [3]. In the supervised machine learning, the algorithm is provided with training data that has outcomes/labels. The algorithm builds a model/function from the pattern buried in the data. When presented with new data, the algorithm infers outcomes based on the model/function already built. In unsupervised learning, there is no outcome/label provided with the data. The algorithm finds the unknown structure or pattern in the data without any information on how to name or label it.

The supervised learning algorithms generally fall into two categories – *classification* and *regression*. The classification problem separates data into predetermined labels of outcome. In classification problem, the outcome variable has a finite number of categories (such as male/female, low/moderate/high). The regression problem deals with predicting the real value of an outcome (such as age, price). Unlike classification problem, the outcome variable in regression problem has continuous values. The most common supervised learning algorithms are artificial neural network (ANN), support vector machine (SVM), decision tree, *k*-nearest neighbours (KNN) etc.

The unsupervised learning algorithms also generally fall into two categories – *clustering* and *association*. In clustering problem, the algorithm endeavours to discover the intrinsic grouping of data based on the similarity of data points (such as grouping of patients based on symptoms). Popular clustering algorithms are *k*-means clustering, hierarchical clustering, DBSCAN, OPTICS etc. In association problem, the co-occurrence of events with high frequency in large dataset is discovered and expressed as association rules (such as customers who buy bread also buy egg). The most frequently used association rule algorithms are AIS, SETM, Apriori etc.

There is another kind of machine learning algorithm that fall between supervised and unsupervised machine learning algorithms, and thus called *semi-supervised* learning algorithms [3]. This sort of learning algorithms is used when vast majority of data are not labelled, only a small portion of the data is labelled. This situation arises when labelling of data requires high expertise and/or the labelling is resource-intensive.

Reinforcement learning is about learning to how to take the best action in a certain environment so that the cumulative reward is maximised [4]. In this type of learning, an agent learns through trial and error based on the reward it receives as a result of its actions. The difference between supervised learning and reinforcement learning is that the labels in supervised learning represent the answer, while in reinforcement learning, the agent has to decide the best action by correlating the delayed return with the action. Reinforced learning can be formulated as a *Markov Decision Process* (MDP) [5] with a finite set of states, a fixed set of actions and rewards. An agent in MDP selects an action to make transition from current state to another state and earn rewards. MDP enables agents to choose the best action in each state so that it can earn the maximum reward. The collection of optimal actions by an agent constitutes a policy. *Q-learning* is a reinforcement learning that learns an action-value function which yields the expected utility of a specified action in a specified state [6].

Unlike statistical analysis methods, machine learning requires vast amount of data to learn the interrelations and patters in the data. While most statistical methods require certain assumptions to hold about the distribution of data, machine learning algorithms can be applied to any data. However, data quality is very crucial for machine learning to build a useful model from the data. In many situations, the available data cannot be directly fed into the machine learning algorithms. The data has to go through a preprocessing stage where it is cleaned, transformed, and enriched before being fed into the machine learning algorithm. The number of attributes or features in the data plays an important role in the design of a machine learning pipeline. When the number of features is too high, the number of data points required in the training phase would be exceedingly high and the model could be exceptionally complex. This is often referred to as the *curse of dimensionality* [7] for machine learning. To decrease the complexity of the model and the volume of data required, the dimensionality needs to be reduced. There are two well-known approaches to reduce the dimensionality of data: *feature selection* and *feature extraction* [8]. Feature selection selects a subset of features that are most relevant for model construction. Feature selection doesn't transform the features, only drops the less important features. The features extraction by contrast, similar to principal component analysis, transforms the features into a smaller set of features that have fewer features than the original set of features.

Machine learning and artificial intelligence can enable smart grid to make intelligent decisions and respond to sudden changes in customer demands, power outages, sudden drops and rises in renewable energy output or any other catastrophic events. The smart meters installed in the customer premises, and other measurement devices in the power system, generate a huge amount of data that could be fed into the machine learning algorithms to capture customer consumption patterns. The use of machine learning to forecast energy demand and power generation of intermittent sources are among the most common applications of machine learning in the power grid. Machine learning can be used to predict any equipment failure by analysing industry-wide equipment failure data, thus helping in asset maintenance and portfolio management. Besides supervised machine learning, unsupervised machine learning can have applications in the smart grid such as creating energy demand profiles through clustering. Reinforced learning algorithms can be used to make energy dispatch decisions and activate demand management signals in order to maintain balance of power supply and demand.

This paper endeavours to review the application of different machine learning approaches that aims at enhancing the stability, reliability, security, efficiency and responsiveness of smart grid. The rest of the paper is organised as follows: Section 2 reviews the use of machine learning to estimate transformer loss of life. Section 3 reviews the machine learning methods to detect power quality events. Section 4 reviews the prospect of machine learning to make optimal energy dispatch decisions aiming at balancing power supply and demand. Section 5 reviews the efficacy of using machine learning in improving electricity market operations. Section 6 reviews the machine learning techniques to detect vulnerabilities and threats in smart gird. Section 7 discusses some of the challenges in implementing machine learning solutions for smart grid. Finally, Section 8 concludes the paper.

II. Estimating Transformer Loss of Life

Proactive maintenance of power system equipment involves prediction of equipment failure and replace or repair those equipment before the failure occurs, thus helping to minimise the unplanned downtime and unexpected disruption of power supply. However, prioritising the equipment for replacement or repair so that limited resources are spent towards the equipment that requires the most urgent attention is a challenging task. The machine learning models can predict the *mean time between failures* (MTBF) and rank feeders accordingly so that the feeder which is most likely to fail gets replaced or repaired first.

Transformers are vital part of the electricity network and failure of transformers can jeopardise the reliability of the network. The repair and maintenance of transformers are costly as well as time consuming, and therefore, it pivotal to estimate the remaining life and replace the ageing transformer before the failure happens. The estimation of transformer loss of life is an important part of transformer

asset management. The insulation of a transformer is the most vulnerable component, and hence, the life of a transformer is dictated by the condition of its insulation. The load profile and ambient temperature have an effect on the aging of transformer insulation. The prediction of load profile and temperature using historical data can help estimate the time to insulation failure for the transformer [9].

Majzoobi at al. [10] developed a static model for hourly estimation of transformer loss of life based on *adaptive network-based fuzzy inference system* (ANFIS) which is a combination of ANN based machine learning and fuzzy inference system. Mahoor and Khodaei [11] further integrated ANFIS and *radial basis function* (RBF) to improve the accuracy of estimation even more. The data fusion techniques that were used to integrate the outputs of ANFIS and RBF are *ordered weighted averaging* (OWA) and *sequential Kalman filter*.

III. Power Quality Event Detection

Machine learning can be used to predict outage or power quality events (PQE) in the power grid by using validated models [12]. Self-healing grids can be developed by automatically detecting and mitigating fault events with the aid of machine learning. Knowledge of fault signatures and their development combined with machine learning can facilitate the detection of fault events in the power grid [13]. Traditionally power system states are estimated using models consisting of high dimensional nonlinear equations; however, the computational complexity involved makes it very time consuming [14]. Consequently, these models are unable to predict the system state (such as frequency) soon enough after a disturbance has happened. Simplified models have been developed that work much faster than the original models, but these ignore some of the factors affecting the system states, and hence, produce unreliable results in some instances. Machine learning appears to be a viable solution for system state prediction as it is not as unreliable as the simplified models, and on the other hand, it doesn't need to perform complex calculations every time it needs to estimate the system state. The machine learning models need to be trained with historical time series data in order to enable it to distinguish fault events from normal situations. The training process may be time consuming, but once trained, the machine learning models can reliably predict the system state within a short period of time.

Ucar at al. [15] proposed a method to detect power quality events using *extreme learning model* (ELM). ELM is a single feedforward neural network which consists of a single layer of hidden nodes. The weights between the input and hidden layers are randomly assigned and remain fixed, while weights between the hidden and output layer is updated during the training phase. Due to having a single layer, the complexity of ELM is very low. Therefore, it can be trained extremely fast, while exhibiting acceptable performance.

Fault detection is a research area where machine learning has gained prominence. SVM and ANN have been extensively used to classify the faults. The features from fault-induced current and voltage signals are generally extracted using *discrete wavelet transform* (DWT) which can extract the characteristics of a signal in a time window. *Fast discrete orthonormal S-transform* (FDOST) can also be used to extract features of current and voltage signals [16]. Other machine learning algorithms that are used for fault detection are decision tree, *k*-nearest neighbours, ELM, bagged tree ensemble, fuzzy logic etc. [17]. Machine learning techniques can also be used to determine fault locations. Support vector regressions (SVR), ANN, back propagation ANN, *k*-nearest neighbours are commonly used machine learning algorithms for this purpose. The features are extracted using *fast Fourier transform* (FFT), *wavelet packet transform* (WPT).

Complex event processing (CEP) is a Big Data analytics technique which analyses trends, patterns and correlation in the data to identify complex situations. CEP can be quite useful for smart grid as it would allow smart grid to analyse complex situations and respond in real-time. An architectural framework based on *Lambda architecture* for complex event processing in smart grid was presented in [18].

IV. Making Energy Dispatch Decisions

The power grid operators rely heavily on forecasting of the electricity demand to adjust power generation to meet the demand and to avoid power system overload. The introduction of renewable and distributed energy sources is making it increasingly difficult to match the power generation with demand. Nowadays, it is essential to forecast both electricity demand and distributed generation to know how much power needs to be produced by the conventional generators to meet the net demand.

Battery storage plays an important role in smoothing out the fluctuations in energy demand and distributed generations. Batteries can store power during periods of excess energy generation, and release it later to fill the gap during the period of generation deficit. A consumer can choose to consume power from any of the three sources – power grid, battery or PV, depending on the grid electricity price and availability of power from the battery or PV. The consumer can sell excess energy from PV or battery when the feed-in tariff is higher and store excess energy from PV in the battery or consume it when the feed-in tariff is lower. Similarly, the consumer can consume electricity from the power grid during off-peak period when the price is low and use up stored energy from battery during peak period when the price is high.

A huge number of heuristic algorithms (such as particle swarm optimisation, genetic algorithm) and game theory based methods have been developed to make optimal energy dispatch decisions in smart grid [19],[20],[21]. Reinforcement learning based energy management algorithms can also make optimal energy dispatch decisions to reduce the cost of energy for the consumers. A Q-learning based energy management algorithm was developed in [22], which learns to make better energy dispatch decisions through experiences without having any prior knowledge. Simulation results using real-life data suggest that the proposed algorithm can significantly reduce the total energy cost based only on current information, while heuristic and game theory based methods rely on future energy demand forecast to make optimal energy dispatch decisions.

V. Electricity Market Operations

Maintaining balance between supply and demand is one of the key goals of smart grid. As mentioned earlier,

increasing penetration of renewable and distributed sources has made it harder to find the balance. Various demand management and dynamic pricing schemes have complicated the prediction of energy usage patterns [23]. Electricity market operators match the demand for electricity with supply from generators in five minutes, 30 minutes or one hour period. A more efficient and flexible approach is to introduce automated electricity brokers. These brokers can operate in a localised market to eliminate the inefficiencies arising from wide-area transmissions. Automated brokers can effectively trade in two different but interconnected markets to buy electricity from producers and sell electricity to retailers. The challenge faced by a broker is to balance the demand and supply of power and make profit while competing with other brokers.

Q-learning was successfully used in [24] to make an automated broker learn an effective strategy to choose an optimal action in each state in a tariff market. As the price of electricity in real market is continuous, there is unlimited possibility of prices. To limit the state space, the price range was divided into some discrete values. Also, the energy portfolio was kept limited to a small set of states to control the state space.

Recurrent *deep Q-learning* (DQN) can address continuous state space and discrete action space. A DQN based multiagent autonomous broker was employed in [25] to trade power in a local tariff market. As the tariff market prices are inherently continuous and temporal, a *long short-term memory* (LTSM) based DQN was employed to enable a multiagent broker to learn pricing strategies. LTSM has excellent capability in modelling sequential data. The paper clustered customers according to usage patterns to limit the state space. Due to multiagent nature of the broker, a negative reward of a suboptimal action by an individual agent may be concealed by the global reward. The authors proposed a reward reshaping method to identify the implication of actions taken by each agent so that each agent is aware of the negative consequence of its suboptimal action.

VI. Securing Smart Grid

Smart grid has two aspects of infrastructure - cyber and physical infrastructure and consists of several complex components which need to be interconnected for smooth operation [26]. One of the essential components of smart grid is smart meter which collects energy consumption data and sends to the service provider. Smart meters are designed to communicate with the service provider using an *advanced metering infrastructure* (AMI) network. AMI provides facilities to automate metering, monitoring, and controlling of power distribution/outage management through a wireless network [27]. In the AMI network, millions of smart meters communicate with the local utility service provider/control centre using a bidirectional network which can be a mesh, hierarchical, or hybrid topology. *Internet of Things* (IoT) could also be integrated in all major components of smart grid such as power generation, transmission, distribution, and utilisation.

Seamless communication through the usage of wireless technology is a core feature of today's smart grid. This communication involves device to device, device to cloud, device to gateway communication to name a few [28]. The usage of wireless technologies in smart grid makes it more vulnerable to cyber security threats. For reliable operation of smart grid, security and privacy need to be ensured both at physical and cyber level [29]. The security vulnerabilities of smart grid could be categorised as process control security, smart meter security, smart grid communication security, cyber security etc [26]. The potential security threats to smart grid include hacking, meter data tampering or meter fraud, injection of false information/data/commands, data stealing, privacy breaches, denial-of-service (DDoS) attacks, man-in-the-middle attacks, meter viruses and compromise of physical security [27],[30]. To develop a secure and reliable smart grid, it is essential to have a security system that can prevent different types of cyber-attacks. With the increase in data volume, machine learning could be efficiently used for the detection and presentation of anomalous behaviour, intrusion, cyber-attacks, and malicious activities (including virus) as well as data authentication.

Parvez et al. [27] proposed a two-level security scheme for smart meters where a localization-based key management system was used for data encryption and KNN algorithm was used for meter authentication. In this system, a *maximum likelihood estimator* (MLE) based on *received signal strength* (RSS) of radio signal was used for the localization of a meter. This system creates a local positioning map based on the signal strength, where each meter has its own coordinate. Each meter has a secret key associated with its coordinates and this key is used along with a random key index for data encryption to prevent potential data hacking.

Kurt et al. [31] proposed an online attack/anomaly detection framework using a model-free reinforcement learning approach for partially observable MDP. Using this approach, an unknown attack type can be detected without requiring the knowledge of any previous attack model. The simulation results demonstrate that the proposed solution can efficiently detect cyber-attacks, but some further improvement could be done by developing sophisticated memory management techniques, and incorporating linear/non-linear function approximation techniques and deep reinforcement learning to enhance performance.

Zhang et al. [32] investigated different machine learning algorithms that could be used for efficient detection and defense of DDoS attacks. Features that are typically used to detect anomalies in the traffic are number of packets, average packet size, time interval variance, packet size variance, packet rate and bit rate. The authors suggest that machine learning algorithms such as *random forest* (RF) and *naive Bayes* (NB) could be used for achieving superior performance in classifying malicious and normal traffic.

Karimipour et al. [33] proposed a real-time anomaly detection method for *false data injection* (FDI) attacks. In their work, the authors took an unsupervised approach based on statistical correlation between measurements to detect anomalies in smart grid. *Symbolic dynamic filtering* (SDF) was used for feature extraction to discover usual interactions among different smart grid subsystems. A *Dynamic Bayesian network* (DBN) based learning model was employed to detect unobservable cyber-attacks using free energy as an anomaly index.

Ozay et al. [34] employed supervised and semi-supervised machine learning algorithms for FDI attack detection. Four

different types of learning algorithms such as perceptron, KNN, SVM, and sparse logistic regression were used as supervised learning methods for attack detection. Semi-supervised SVM was used as a semi-supervised learning method for attack detection. The authors used decision and feature level fusion algorithms as well as both batch and online learning methods for classification of measurements. The experimental results suggest that the state-of-the-art supervised machine learning algorithms can detect attacks more efficiently than state vector estimation (SVE) based algorithms for detection of different types of attacks. Perceptron is less sensitive and KNN is more sensitive to the system size than other algorithms. Moreover, KNN performs better in small scale systems whereas SVM performs better in large scale systems than other algorithms. However, SVM has challenges as it is sensitive to kernel type and sparsity of the system. The authors also found that semi-supervised learning algorithms are more robust to the sparsity of data than the supervised learning algorithms. The fusion algorithms - Adaboost and *multiple kernel learning* (MKL) make the learning models more robust in terms of change in system size and data sparsity.

Zhang et al. [35] addressed a new security issue in the wireless communication channels by proposing a *distributed intrusion detection system for smart grids* (SGDIDS). The proposed system embeds a number of analysing modules (AM) in each layer of the smart grid. Each AM deploys SVM and *artificial immune system* (AIS) for efficient detection and classification of malicious data and potential cyber-attacks.

Ahmed et al. [36] proposed a supervised machine learning based scheme to detect a new type of cyber-attack named *covert cyber deception assault* (CCD) where the attacks (data tampering) are initiated by hackers with a good knowledge of the system and hence, it is difficult for bad data detectors in traditional state estimators to detect such attacks. This work used a *genetic algorithm* (GA) based feature selection technique for selecting the optimal features, and then used SVM for classifying the CCD attacks. The feature selection technique can identify the most discriminative features and hence, reduces the computational complexity of the attack detection mechanism. Moreover, it reduces the delay, and increases the accuracy of attack detection.

VII. Challenges

The supervised learning algorithms require fully labelled data for training purposes; however, fully labelled data is either limited or difficult to obtain [37]. In case of predicting power outage events, the scarcity and imbalance of event data poses a significant problem as power grid failures are quite rare. For example, distribution feeders have different kinds of failures and there are few training examples available for each kind [38]. The failure pattern may change rapidly and the model for prediction may be obsolete after some time. If deep learning algorithms are used for modelling problems in smart grid, the efficacy of modelling will depend on the availability of vast amount of quality data. It is possible to generate high resolution synthetic data for training by modelling and sampling, but this is not without challenges. The challenges mentioned in [39] are: (1) high dimensionality, (2) large number of observations, (3) non-Gaussian marginal probability of individual variable, and (4) complex nonlinear dependency among variables.

Efficient classification/characterisation of threat and source of threat is very important for accurate detection of security attacks in smart grid [40]. With new types of attacks evolving over time, the classification might need to be regularly updated. Hence, the challenge is to develop a robust classification mechanism that can detect different types of evolving attacks reliably. As mentioned earlier, unavailability of past labelled data for using supervised learning-based approaches is a challenge, which makes anomaly detection in smart grid a challenging task. In such cases, careful consideration should be given to contextual information and other events (rather than depending only on one type of event) for successful characterisation of anomalous behaviour [41]. With the huge amount of data being produced by smart grids, the real-time detection of attacks has become an overwhelming task. Efficient feature selection techniques should be used to reduce the time delay and computational complexity of real-time attack detection methods [34]. Another challenge for security solutions is that the smart meters in smart grid have low memory and computational capability, so the security solutions need to be light weight but robust [27].

VIII. Conclusion

This paper has reviewed the application of different machine learning techniques that could be used to enhance the stability, reliability, security, efficiency and responsiveness of smart grid. The findings of the paper show that machine learning algorithms could be efficiently used for estimating transformer loss of life, detecting power quality events and faults, making optimal energy dispatch decisions to reduce cost of energy, efficient electricity market operations, and securing data and preventing attacks on smart grid. This paper has also discussed some of the challenges in implementing machine learning solutions for smart grid. These include limitation of finding past labelled data, evolution of new types of attacks, rapid changes in failure patterns, issues with generating high resolution synthetic data for training, finding efficient feature selection techniques, low memory and computational capability of smart meters. With the integration of new energy sources and technologies, smart grid is becoming increasingly complex and vulnerable. Although many machine learning based smart grid solutions have been proposed in the literature, there are still a lot of opportunities for improvement. Deep Learning and Big Data can play a vital role in solving problems of smart grid in future.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid—The new and improved power grid: A survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 944–980, 2011.

[2] *Machine Learning: The Power and Promise of Computers That Learn by Example*, The Royal Society, 2017.

[3] I. Salian, *SuperVize Me: What's the Difference between Supervised, Unsupervised, Semi-Supervised and Reinforcement Learning*, 2018. Accessed on: 5 August 2019. [Online]. Available: https://blogs.nvidia.com/blog/2018/08/02/supervised-unsupervised-learning/

[4] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Massachusetts: The MIT Press.

[5] M. van Otterlo and M. Wiering, "Reinforcement learning and Markov decision processes," in *Reinforcement Learning. Adaptation, Learning, and Optimization*, vol. 12. M. Wiering and M.van Otterlo, Eds. Berlin: Springer, 2012.

[6] S. Paul, *An introduction to Q-Learning: Reinforcement Learning*, 2019. Accessed on: 5 August 2019. [Online]. Available: https://blog.floydhub.com/an-introduction-to-q-learning-reinforcement-learning/

[7] V. Spruyt, *The Curse of Dimensionality in Classification*, 2014. Accessed on: 5 August 2019. [Online]. Available: https://www.visiondummy.com/2014/04/curse-dimensionality-affect-classification/

[8] A. Desarda, *Getting Data ready for modelling: Feature engineering, Feature Selection, Dimension Reduction* (*Part two*), 2018. Accessed on: 5 August 2019. [Online]. Available: https://towardsdatascience.com/getting-data-ready-for-modelling-feature-engineering-feature-selection-dimension-reduction-39dfa267b95a

[9] K.T. Muthanna, A. Sarkar, K. Das, and K. Waldner, "Transformer insulation life assessment*," IEEE Transactions on Power Delivery*, vol. 21, no. 1, 2006.

[10] A. Majzoobi, M. Mahoor, and A. Khodaei, "Machine learning applications in estimating transformer loss of life," *IEEE Power and Energy Society General Meeting*, pp.1-5, 2017.

[11] M. Mahoor and A. Khodaei, "Data fusion and machine learning integration for transformer loss of life estimation," *IEEE/PES Transmission and Distribution Conference and Exposition* (T&D), 2018.

[12] J. McClelland, *Connected Assets: How Machine Learning Will Transform the Utilities Industry*, 2018. Accessed on: 5 August 2019. [Online]. Available: https://www.digitalistmag.com/iot/2018/02/26/connected-assets-how-machine-learning-will-transform-the-utilities-industry-05921360

[13] B. Torsæter, *Preliminary Results Show that Machine Learning Can Predict Faults and Disturbances in the Power System*, 2019. Accessed on: 5 August 2019. [online]. Available: https://blog.sintef.com/sintefenergy/machine-learning-can-predict-faults-and-disturbances-in-the-power-system/

[14] Y. Tang, H. Cui, and Q. Wang, "Prediction model of the power system frequency using a cross-entropy ensemble algorithm," *Entropy*, vol. 19, no. 10, 2017.

[15] F. Ucar , O. Alcin, B. Dandil, and F. Ata, "Power quality event detection using a fast extreme learning machine," *Energies*, vol. 11, no.1, 2018.

[16] P. K. Mishra, A. Yadav, and M. Pazoki, "A novel fault classification scheme for series capacitor compensated transmission line based on bagged tree ensemble classifier," *IEEE Access*, vol. 6, pp. 27373–27382, 2018.

[17] V. Venkatesh, "Fault classification and location identification on electrical transmission network based on machine learning methods," *Master of Science Thesis*, Virginia Commonwealth University, 2018.

[18] G. Liu, W. Zhu, C. Saunders, F. Gao, and Y. Yu, "Real-time complex event processing and analytics for smart grid," *Procedia Computer Science*, vol. 61, 2015.

[19] S. A. Azad, A. M. T. Oo, and M. F. Islam, "A low complexity residential demand response strategy using binary particle swarm optimization," in *Proceedings of Australasian Universities Power Engineering Conference*, 2012.

[20] M. Awais, N. Javaid, N. Shaheen, Q. Z. Iqbal, G. Rehman, K. Muhammad, and I. Ahmad, "An efficient genetic algorithm based demand side management scheme for smart grid," in Proc. of *18th IEEE International Conference on Network-Based Information Systems* (NBiS), 2015. doi: 10.1109/NBiS.2015.54

[21] N. Hajj and M. Awad, "A game theory approach to demand side management in smart grids," in *Intelligent Systems'2014. Advances in Intelligent Systems and Computing*, vol 323. D. Filev et al. Eds. Cham: Springer, 2014.

[22] S. Kim and H. Lim, "Reinforcement learning based energy management algorithm for smart energy buildings," *Energies*, vol. 11, no. 8, 2018.

[23] M. Peters, W. Ketter, M. Saar-Tsechansky, and J. Collins, "A reinforcement learning approach to autonomous decision-making in smart electricity markets," *Machine Learning*, vol. 92, no. 1, 2013.

[24] P. Reddy and M. Veloso, "Strategy learning for autonomous agents in smart grid markets," in *Proc. of the Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[25] Y. Yang, J. Hao, M. Sun, Z. Wang, C. Fan, and G. Strbac, "Recurrent deep multiagent Q-learning for autonomous brokers in smart grid," in Proc. of *Joint Conference on Artificial Intelligence* (IJCAI), 2018.

[26] S. Ahmad, "Anomaly modeling and detection for smart grid communication infrastructures," *PQDT – Global*, 2013.

[27] I. Parvez, A. I. Sarwat, L. Wei, and A. Sundararajan, "Securing metering infrastructure of smart grid: A machine learning and localization based key management approach," *Energies*, vol. 9, no. 9, 2016.

[28] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of big data and machine learning in smart grid, and associated security concerns: A Review," in *IEEE Access*, vol. 7, pp. 13960-13988, 2019. doi: 10.1109/ACCESS.2019.2894819

[29] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," in *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210-224, 2012. doi: 10.1109/JPROC.2011.2165269

[30] S. Bisoi and A. Dash, "The role of utilities in securing a smart grid," *Electric Light and Power*, vol. 89, no. 6, pp. 70-72, 2011.

[31] M. N. Kurt, O. Ogundijo, C. Li, and Z. Wang, "Online cyber-attack detection in smart grid: A reinforcement learning approach," *IEEE Transactions on Smart Grid*, 2018. doi: 10.1109/TSG.2018.2878570

[32] B. Zhang, T. Zhang, and Z. Yu, "DDoS detection and prevention based on artificial intelligence techniques," 3rd *IEEE International Conference on Computer and Communications* (ICCC), pp.1276–1280, 2017.

[33] H. Karimipour, A. Dehghantanha, R. Parizi, K. Choo, and H. Leung, "A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids," *IEEE Access*, vol. 7, no. 99, 2019.

[34] M. Ozay. I. Esnaola, F. T. Vural, S. J. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773–1786, 2016.

[35] Y. Zhang, L. Wang, W. Sun, R. C. Green II, and M. Alam. "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 796–808, 2011.

[36] S. Ahmed, Y. Lee, S Hyun, and I. Koo, "Feature selection–based detection of covert cyber deception assaults in smart grid communications networks using machine learning," *IEEE Access*, vol. 6, pp. 27518–27529, 2018.

[37] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Partial knowledge data-driven event detection for power distribution networks," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, 2018.

[38] C. Rudin et al. "Machine learning for the New York city power grid," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, 2012.

[39] M. Sun, I. Konstantelos, and G. Strbac, "A deep learning-based feature extraction framework for system security assessment," *IEEE Transactions on Smart Grid*, 2018. doi:10.1109/TSG.2018.2873001

[40] A. O. Otuoze, M. W. Mustafa, and R. M. Larik, "Smart grids security challenges: Classification by sources of threats," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 3, pp. 468–483, 2018.

[41] B. Rossi, S. Chren, B. Buhnova, and T. Pitner , "Anomaly detection in smart grid data: An experience report," *IEEE International Conference on Systems, Man, and Cybernetics* (SMC), pp. 002313–002318, 2016.