Rituraj Rituraj – Dániel T. Várkonyi – Amir Mosavi – József Pap – Annamária R. Várkonyi-Kóczy – Csaba Makó

MACHINE LEARNING IN SMART GRIDS: A SYSTEMATIC REVIEW, NOVEL TAXONOMY, AND COMPARATIVE PERFORMANCE EVALUATION

Rituraj Rituraj, John von Neumann Faculty of Informatics, Óbuda University Budapest, Hungary

Dániel T. Várkonyi, Faculty of Informatics, Eötvös Loránd University Budapest, Hungary

Amir Mosavi, Research Fellow, Ludovika University of Public Service, Institute of the Information Society, mosavi.amirhosein@uni-nke.hu

József Pap, Széchenyi University Doctoral School of Management (SzEEDSM), Győr, Hungary

Annamária R. Várkonyi-Kóczy, Professor, Óbuda University, Institute of Software Design and Software Development, varkonyi-koczy@uni-obuda.hu

Csaba Makó, Professor, Ludovika University of Public Service, Institute of the Information Society, mako.csaba@uni-nke.hu

This article presents a state-of-the-art review of machine learning (ML) methods and applications used in smart grids to predict and optimise energy management. The article discusses the challenges facing smart grids, and how ML can help address them, using a new taxonomy to categorise ML models by method and domain. It describes the different ML techniques used in smart grids as well as examining various smart grid use cases, including demand response, energy forecasting, fault detection, and grid optimisation, and explores how ML can improve these cases. The article proposes a new taxonomy for categorising ML models and evaluates their performance based on accuracy, interpretability, and computational efficiency. Finally, it discusses some of the limitations and challenges of using ML in smart grid applications and attempts to predict future trends. Overall, the article highlights how ML can enable efficient and reliable smart grid systems.

Studies •

INTRODUCTION

The smart grid (SG) is an upgraded type of electrical grid that improves reliability, security, and efficiency using advanced technology, facilitating real-time communication for managing power supply and demand. It promotes the integration of renewable energy sources and supports electric vehicles and distributed energy resources, reducing reliance on fossil fuels. It also enhances grid resilience and security, potentially transforming the electricity sector into a more sustainable and dependable energy system. In Ahmad et al.'s (2007) review,[1] they highlight the unique challenges arising from the growing integration of energy storages and renewable energy sources into the conventional power systems and AI. This shift requires forward-thinking investments in SG technologies, integrating advanced measurement equipments, controllable transmission assets, and software control systems. Kwak and Heo (2007) stress the importance of creating a resilient and adaptable infrastructure capable of responding to both internal and external changes, given the intricate interconnectedness of modern infrastructure systems, which can magnify the impact of local disruptions into broader cascade failures.[2] The vision for the SG, as presented by Bari et al. (2014),[3] is of a profound transformation in the electric power sector. This transformation centres on the integration of bidirectional power and information flows, addressing critical factors like capacity, efficiency, reliability, sustainability, consumer engagement, and the ever-growing energy demand. It promotes a range of generation and storage solutions, and advocates for environmentally responsible practices.[4] Ardito et al. (2013) add that the development of the SG entails enhancing the existing network with new features and services while preserving the core physical infrastructure, marking a significant stride toward a more resilient and adaptable power system.[5] The conventional electricity grid, often referred to as the legacy or analogue grid, functions as a one-way system in which electricity is centrally generated and transmitted to consumers via long-distance lines. In contrast, the SG is a modernised grid that harnesses advanced digital technology to enhance the power system's reliability, security, and efficiency.

---

[1]   Ahmad et al. 2022.
[2]   Kwak–Heo 2007.
[3]   Bari et al. 2014.
[4]   Mei–Chen 2013.
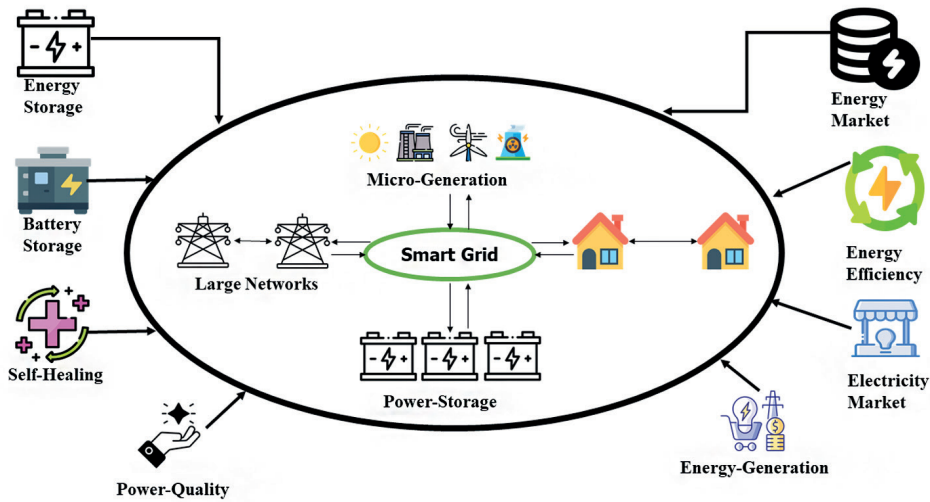[5]   Ardito et al. 2013.

Figure 1: Overview of an SG implementation
Source: compiled by the authors based on an adaptation from AHMAD et al. 2022.

The SG stands out due to four key features: bidirectional communication, facilitating real-time monitoring and management of electricity supply and demand; integration of RESs such as solar and wind; support for electric vehicles and distributed energy resources. It reduces dependency on fossil fuels, and enhances environmental sustainability. Moreover, robust security measures provide protection against potential threats. Figure 1 visually demonstrates the information and energy flow within the SG infrastructure, a concept absent in the traditional power system.

In the context of modern SGs, Berghout et al. (2022) emphasise the critical importance of condition monitoring, which is facilitated by cutting-edge computing technology and secure cyber-physical connectivity.[6] ML, and particularly deep learning (DL), has rapidly advanced and demonstrated exceptional performance in various SG-related activities, as highlighted by Xu et al. (2022).[7] The transition from traditional power distribution systems, which relied on human intervention, to more robust SGs has played a crucial role in ensuring reliable power delivery, as discussed by Elbouchikhi et al. (2021).[8] Effectively processing the extensive data within SGs, required for tasks such as power flow optimisation and system monitoring, necessitates dynamic energy management, as elucidated by Hossain et al. (2019).[9] ML and DL techniques offer valuable tools for SG development, encompassing three key phases: the constituting element phase, the process phase, and the power

---

[6]   BERGHOUT et al. 2022.
[7]   XU et al. 2022.
[8]   ELBOUCHIKHI et al. 2021.
[9]   HOSSAIN et al. 2019.

Studies •

converter stage, as discussed by Behara and Saha (2022).[10] However, developments such as the integration of prosumers into SGs, power system decarbonisation using blockchain and artificial intelligence (AI), and the functionalities of various SG applications represent both challenges and opportunities, as highlighted by Hua et al. (2022).[11] Applications of SG include such aspects as demand response initiatives, automated data processing for SCADA systems, voltage stability assessment, smart city planning, and home automation, as noted by Chaurasia and Kamath (2022).[12] The possible sources of disruption, including harmonic production, load variations, and wiring and grounding issues pose risks to the electricity supply system, as explained by Rangel-Martinez et al. (2021).[13] Widespread implementation of advanced technologies like 5G and specialised algorithms plays a pivotal role in developing ML-based sustainable non-industrial energy management applications, as outlined by Omitaomu and Niu (2021).[14] Security issues in SGs, particularly the threat of false data attacks, have substantial implications and are a subject of concern, as addressed by Cui et al. (2020).[15]
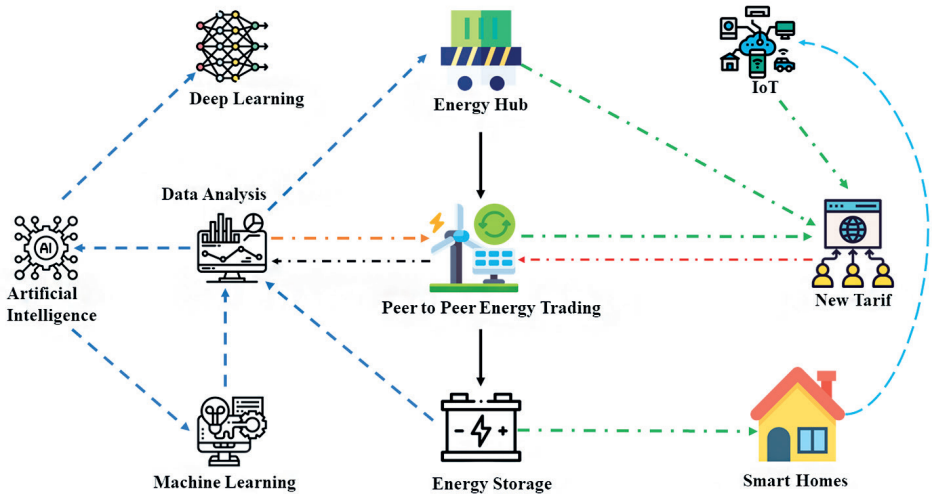

Figure 2: Machine learning's role in SG
Source: compiled by the authors based on an adaptation from Ahmad et al. 2022.

---

10   Behara–Saha 2022.
11   Hua et al. 2022.
12   Chaurasia–Kamath 2022.
13   Rangel-Martinez et al. 2021.
14   Omitaomu–Niu 2021.
15   Cui et al. 2020.

Additionally, concerns about cybersecurity emerge due to the heavy reliance of SGs on communication technologies, while the management of extensive data volumes for privacy and analytical purposes presents an additional obstacle. Enhancing grid resilience to natural disasters, achieving regulatory standardisation, fostering consumer engagement, addressing the high initial infrastructure costs, and seamlessly integrating diverse energy resources constitute critical areas for enhancement. Making sufficient progress in these domains requires collaborative efforts among stakeholders and advancements in technology to facilitate the widespread adoption of sustainable electricity grids. In this regard, Raza and Khosravi (2015) review techniques for developing SGs and building load demand forecasting based on AI, with Figure 2 illustrating the various applications of AI, computational intelligence, and ML across different aspects of SG.[16]

METHODOLOGY

This section presents an overview of the classification and research methodology employed in the comprehensive review study. It first explicitly delineates the primary taxonomy of the article, followed by an explanation of the process utilised for information gathering and conducting related investigations. The section is structured into six main segments, including an introduction, details of the study's execution, its primary objectives, a review of studies focusing on energy demand forecasting, the introduction of evaluation criteria, and a presentation of research findings. The central focus of this review is the examination of machine learning applications in smart grids (SGs). The study follows the PRISMA standard for data collection, which consists of four essential phases: identification, screening, eligibility, and inclusion. Initially, a total of 550 articles were identified, and following a rigorous evaluation, 89 articles were chosen for in-depth analysis, constituting the qualitative and quantitative foundation of the study. Figure 3 offers a visual representation of the flowchart illustrating the PRISMA technique employed in constructing the study's database.
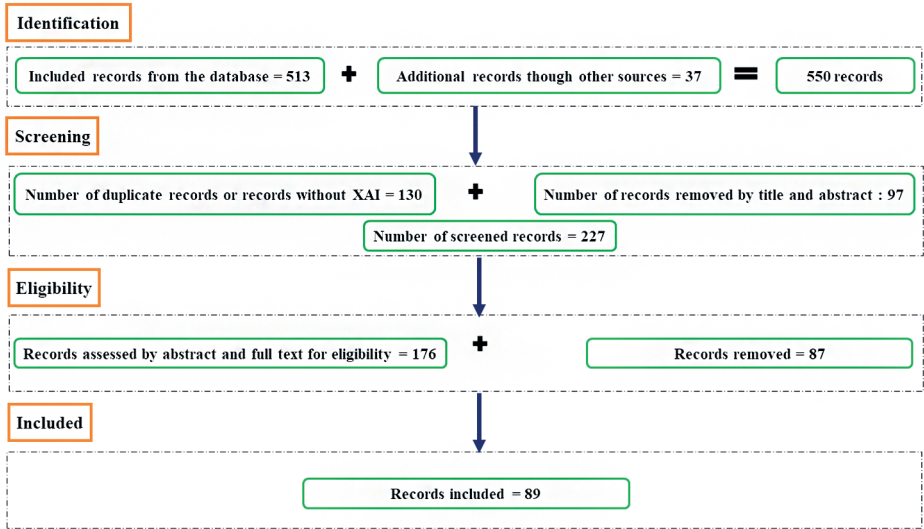
---

[16]   Raza–Khosravi 2015.

Figure 3: The PRISMA flowchart illustrating the standard search and screening
Source: compiled by the authors

RESULTS

Artificial Neural Network

Artificial Neural Networks (ANNs) are ML algorithms inspired by the structure of the human brain and are thus well-suited for pattern recognition and feature extraction tasks in SGs. They are pivotal in load forecasting, fault detection, and energy management within SGs, and have been widely adopted by researchers and practitioners. ANNs predict electricity demand based on factors such as weather and time, optimising grid operations and reducing the need for additional energy generation or storage. They also excel at identifying power transmission line faults, enhancing grid performance, and preventing outages through sensor data analysis. In energy management, ANNs optimise the operation of distributed energy resources, such as photovoltaic panels and wind turbines, using historical data and predictive analysis, and thus significantly improving SG efficiency, reliability, and sustainability. The creation of ANN networks involves a crucial training process with connections and nodes, detailed in Equation (1).[17]

$$I_j(t) = \sum_i O_i(t) w_{ij} + w_{oj} \qquad (1)$$

---

[17]   ZAIN et al. 2012.

The equation provided describes the input value (Ij) from neuron i to neuron j, with Oj representing the output value of neuron i. The weight value is denoted as wij, and the related bias for neuron j is woj. A simplified representation of a basic ANN approach in the presence of related components is depicted in Figure 4. The development of an ANN approach typically involves three stages: training, testing, and validation.
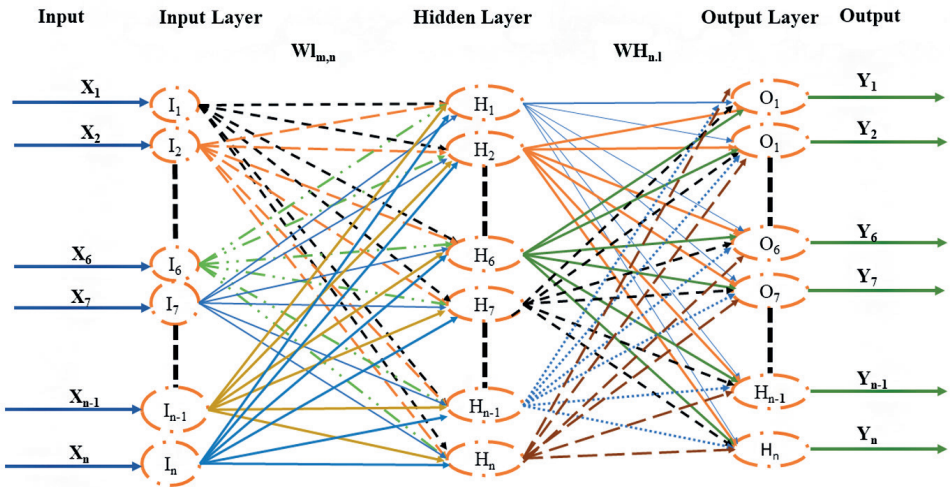

Figure 4: The architecture of ANN
Source: compiled by the authors based on an adaptation of a standard ANN.

ANNs find application in SG contexts, such as material selection for dye-sensitised solar cells, based on Bhagya Raj et al. (2022).[18] These networks have demonstrated their prowess at capturing nonlinear relationships even without prior knowledge and exhibit high degree of accuracy, with metrics like root mean square error (RMSE), high correlation coefficients, and low relative deviation, as exemplified in Li et al.'s (2021) photovoltaic fault detection study.[19] Recurrent Neural Networks (RNN), incorporating feedback loops, excel at tasks such as time-series forecasting but may face challenges with long-term dependencies. Convolutional Neural Networks (CNN) are adept at grid-structured data tasks like image recognition but may be less effective for non-grid-structured data. Autoencoders efficiently handle data compression and noise removal in SG applications, but their effectiveness depends on data quality. Common ANN configurations include multi-layer perceptrons and the implementation of the back-propagation learning technique. ANNs, as described by Jawad et al. (2021), are a ML technique inspired by biological neurons, sharing computational parallels with human brain learning processes.[20] Feedforward Neural Networks are a fundamental type, with unidirectional interconnected nodes

---

[18] Bhagya Raj – Dash 2022.
[19] Li et al. 2021.
[20] Jawad et al. 2021.

suitable for basic SG tasks like regression and classification. However, FNNs may struggle with complex relationships in nonlinear data. ANNs aim to create links between input and output variables through data-driven learning procedures. Additionally, a study by Dong et al. (2003) underscored the extensive application of radial basis function networks and multi-layer perceptron's in function approximation, contributing to the development of Support Vector Machines (SVMs).

SUPPORT VECTOR MACHINES

SVMs are versatile supervised learning algorithms applied in SGs for both classification and regression tasks, addressing such challenges as load forecasting, fault detection, and power quality event classification. SVMs excel at handling high-dimensional power systems data, being robust to noise, and can perform well even with limited training data. They effectively model nonlinear relationships between features, enhancing system accuracy. SVM integration in SG applications has the potential to improve power system efficiency, reliability, and security when using advanced ML. In SVM, data analysis for regression and classification creates reliable prediction models, assigning new instances during training. The closest data points to the hyperplane, referred to as support vectors, influence both the hyperplane's position and orientation. It is essential to maximise the margin, which is the distance between the support vectors and the hyperplane, when selecting the hyperplane. Even a slight alteration in the position of these support vectors can change the hyperplane. Nu-SVM introduces the "nu" parameter, providing flexibility in controlling support vectors and errors. Despite offering a more intuitive trade-off, selecting an appropriate "nu" value involves difficulties, with interpretations varying across datasets. Equation (2) represents any hyperplane as the set of desired points, as explained by Blanco et al. (2022).[21]

$$\vec{\omega} . \vec{x}_1 - b = 0, \tag{2}$$

Here, $\vec{\omega}$ is the standard vector to the hyperplane. $\frac{b}{\vec{\omega}}$ represents the offset value of the hyperplane from the origin along the normal vector.

Within SVM, two types of margins are considered: soft margins and hard margins. The concept behind the soft margin is to permit SVM to make minor errors, thus enabling the widest possible margin to correctly classify other data points. Linear-SVM seeks optimal class separation through a hyperplane, maximising the margin with support vectors. It excels at high-dimensional spaces for approximately linear relationships but is less effective with complex, nonlinear data and is sensitive to outliers. This approach leads to a distinct optimisation problem, as shown in equation (3).[22]

---

[21]   BLANCO et al. 2022.
[22]   RANGANATHAN et al. 2011.

$$\left[\frac{1}{n}\sum_{i=1}^{n}max(0,1-y_i(\vec{\omega}.\vec{x}_1-b))\right]+\lambda||\vec{\omega}^2|| \tag{3}$$

In this context, λ represents the trade-off between the margin size increase and the requirement for xi to remain on the current side of the margin. Nonlinear-SVM with Kernels extends linear SVM, effectively capturing complex relationships through higher-dimensional transformations. It is versatile in scenarios with nonlinear boundaries and its performance relies on kernel selection, which raises challenges and increased computational costs. Support Vector Regression adapts SVM for regression, effectively minimising deviations within a specified margin and handling nonlinear relationships. It is robust to outliers. It demands careful parameter selection, particularly with large datasets. Various studies have highlighted the versatility of SVMs in different applications. Diana et al. (2019) emphasised the adaptability and high accuracy of SVMs for electromyographic signal classification.[23] Multiclass-SVM expands binary SVM for multiple classes, utilising one-vs-one or one-vs-all strategies. It is versatile for scenarios with more than two classes, while its performance is strategy-dependent, with heightened computational costs for additional classes.

EXTREME LEARNING MACHINE

Extreme Learning Machine (ELM) is a widely used ANN in SGs due to its quick training and high accuracy. ELM plays a pivotal role in predicting energy consumption, improving load forecasting, and enhancing grid stability in SGs. ELM functions as a single-hidden layer feedforward NN, featuring randomly assigned input weights and biases, and exclusively learning output weights. It emphasises swift training with a fixed hidden layer. It can optimise grid operations by forecasting energy demand based on weather and time. ELM can also detect system failures, such as power line faults, preventing outages and improving overall grid performance. Additionally, ELMs are valuable for real-time network retraining, although they may not match the accuracy of CNNs, as shown in equation (4).[24]

$$f_l(x)=\sum_{i=1}^{L}\beta_i g_i(x)=\sum_{i=1}^{L}\beta_i g(\omega_i*x_j+b_i), j=1,\ldots N \tag{4}$$

L is the number of hidden or covert units in the NN. N is the number of training samples or data points. β represents the scaled vector between the hidden layer and the output layer, ω represents the scaled vector between the input layer and the hidden layer, g is an activation function applied to the output of the hidden layer, b is a bias vector typically added to the output of the hidden layer, while x is the input vector to the NN. Zheng et al. (2022)

---

[23]   Diana et al. 2019.
[24]   Mishra et al. 2022.

conducted a review of ELM in data stream classification, emphasising its effectiveness, universal approximation capabilities, and simplicity.[25] Mohanty et al. (2021) proposed a hybrid model combining kernel ELM and an autoencoder for financial market prediction, which exhibited improved profitability analysis compared to traditional methods.[26] ELM has proven to be computationally efficient for extensive datasets, particularly in big data applications. Its advantages include reduced overfitting through hidden layer randomisation and faster training times compared to traditional neural networks. Yaseen et al. (2019) used ELM for river flow forecasting and water resource management, enhancing prediction accuracy through orthogonal decomposition.[27] ELM has limitations such as potential constraints on model interpretability and potential performance gaps in complex tasks requiring deep hierarchical feature learning. Additionally, the absence of iterative tuning in the hidden layer may restrict its adaptability to specific data types. Kariminia et al. (2016) applied ELM to analyse visitors' thermal comfort in public spaces, achieving precise predictions and reducing training time.[28]

Dou et al. (2022) identified a vulnerability in power system state estimators with a new cyberattack called a false data injection (FDI) attack, using online sequential ELM.[29] Dewangan et al. (2022) proposed an enhanced ELM model for forecasting the stability of cyber-physical systems in SGs, considering technical and socioeconomic factors that could impact their stability.[30] Zhang et al. (2020) improved intrusion detection in SGs by applying a genetic algorithm ELM, enhancing accuracy and reducing false positives.[31] Naz et al. (2019) enhanced the recurrent-ELM model's scalability and prediction accuracy, especially with larger datasets.[32] Duo et al. (2019) employed a hybrid model combining Variational Mode Decomposition and online sequential ELM for FDI attack detection in SGs.[33] Xue et al. (2019) introduced a predictive recovery technique for addressing incorrect data using geographical power data correlation to enhance system resilience.[34] Li et al. (2018) proposed an intrusion detection system using online sequential extreme learning, optimised with the artificial bee colony-differential evolution algorithm, achieving a 95.3% accuracy rate for detecting bogus data injection attacks.[35]

[25] Zheng et al. 2022.
[26] Mohanty et al. 2021.
[27] Yaseen et al. 2019.
[28] Kariminia et al. 2016.
[29] Dou et al. 2022.
[30] Dewangan et al. 2022.
[31] Zhang et al. 2020.
[32] Naz et al. 2019.
[33] Dou et al. 2019.
[34] Xue et al. 2019.
[35] Li et al. 2018.

DECISION TREE

Decision Trees (DT) are widely used in SGs due to their simplicity, interpretability, and versatility in handling various data types. DTs have diverse applications in SGs, including load forecasting, fault detection, and energy management. They predict electricity demand based on factors such as weather and time, allowing grid operations to be optimised, and reducing the need for extra energy generation or storage. DTs also help detect power line faults, improving grid performance. In energy management applications, DTs optimise the operation of distributed energy resources, maximising their contribution. When used in SGs, DTs can enhance efficiency, reduce costs, and enhance grid reliability and sustainability. Figure 6 illustrates the DT workflow, and DT algorithms apply equations like (5) for classification and (6) for regression.[36]

$$\sum_{i=1}^{c} - f_i \log(f_i) \tag{5}$$

$$\frac{1}{N} \sum_{i=1}^{N} (y_i - \mu)^2 \tag{6}$$

Where, f_i is the frequency of label i at a node and c is the number of unique labels, y_i is the label for an instance, N is the number of instances and μ is the mean given by $\frac{1}{N} \sum_{i=1}^{N} y_i$.

Jena and Dehuri (2020) explored the role of data mining in prediction, encompassing regression and classification.[37] DTs are preferred for their simplicity and effectiveness, but their complexity grows with larger datasets, often requiring advanced algorithms. Zekić-Sušacand and Knežević (2021) investigated the Classification and Regression Tree Algorithm (CART) and its relevance to energy cost.[38] Kadiyala and Kumar (2018) used Python to assess various DT-based boosting algorithms and found extreme gradient boosting to perform best.[39] Subramaniam et al. (2017) compared CARTs and Conditional Inference Trees (CTree) in terms of subgroup identification and prediction accuracy.[40] Peng et al. (2017) emphasised CART's significance in cocaine reward research.[41] Heidari et al. (2013) introduced a hybrid DT and discrete wavelet transform model to facilitate the detection of islanding in distributed systems.[42] DT models have also been applied for the transient security assessment (TSA) of power systems.[43]

Turanzas et al. (2022) introduced an event detection algorithm for the SG focused on identifying the status and location of attacked devices. They employed two DTs to improve event detection accuracy, achieving 80.59% accuracy for status prediction and

---

[36]   ACHLERKAR et al. 2016.
[37]   JENA–DEHURI 2020.
[38]   ZEKIĆ-SUŠACAND–KNEŽEVIĆ 2021.
[39]   KADIYALA–KUMAR 2018.
[40]   SUBRAMANIAM et al. 2017.
[41]   PENG et al. 2017.
[42]   HEIDARI et al. 2013.
[43]   NIAZI et al. 1999.

maintaining 79.39% for location prediction using the nested FDI algorithm.[44] Da Cunha et al. (2022) employed the DT algorithm to analyse power system stability in SGs, achieving a remarkable degree of accuracy of 93% when identifying small signals within the system.[45] Tehrani et al. (2020) utilised DTs, Random Forest (RF), and gradient boosting algorithms to detect non-technical losses in power consumption data, achieving accuracy rates of 87%, 88.1%, and 88.6% for DT, RF, and gradient boosting, respectively.[46] Taghavinejad et al. (2020) stressed the significance of IoT security in SGs and proposed a hybrid DT method with 83.14% accuracy in detecting technical losses, outperforming DT, KNN, and SVM with accuracy rates of 80.90%, 79.12%, and 78.52%, respectively. Eissa et al. (2019) compared DT's role in load optimisation to other methods, noting its efficient management of local heat ventilation and air conditioning units.[47] Radoglou-Grammatikis et al. (2019) introduced an intrusion detection system based on DT for safeguarding advanced metering infrastructure, achieving high accuracy and a true positive rate.[48] Wang and Kong (2019) enhanced air quality predictive modelling in a "weather-smart grid" through DT-based techniques.[49] Singh and Govindarasu (2018) presented an intelligent remedial action scheme for detecting cyber-attacks and physical disturbances in the SG.[50] Steer et al. (2012) demonstrated the effectiveness of DT ensembles, referred to as "forests", in delivering near-ideal control strategies in real-time.[51]

RANDOM FOREST

RF, a widely used ML algorithm, is prominent in the SG industry for its high accuracy and ability to handle complex datasets. In SG-related applications, RF excels at load forecasting, fault detection, and energy management, optimising grid operations and reducing the need for additional energy generation. It also enhances grid performance by detecting and diagnosing power transmission line faults and improving sustainability by optimising distributed energy resources. The integration of DTs into the RF algorithm to make predictions involves the combination of independent DT models through majority voting. Gini Importance, as implemented in Scikit-learn, evaluates feature relevance in each DT, as illustrated in equations (7) and (8). An overview of the RF algorithm is provided in Equation (9).[52]

[44] Turanzas et al. 2022.
[45] Da Cunha et al. 2022.
[46] Tehrani et al. 2020.
[47] Eissa et al. 2019.
[48] Radoglou-Grammatikis–Sarigiannidis 2019.
[49] Wang–Kong 2019.
[50] Singh–Govindarasu 2018.
[51] Steer et al. 2012.
[52] Das et al. 2022.

$$ni_j = w_j c_j - w_{left(j)} c_{left(j)} - w_{right(j)} c_{right(j)} \tag{7}$$

The following formula is used to determine each feature's relevance on a DT:

$$fi_i = \frac{\sum_{j:node\ j\ splits\ on\ features\ i} ni_j}{\sum_{k \in all\ nodes} ni_k} \tag{8}$$

$$normfi_i = \frac{fi_i}{\sum_{j \in all\ features\ of} fi_j} \tag{9}$$

At the RF level, the ultimate feature relevance is determined by finding its average over all trees. Calculating the relevance of each attribute for each tree, then dividing that total by the number of trees, yields:

$$RFfi_i = \frac{\sum_{f_{j \in all\ tress}} normfi_{ij}}{T} \tag{10}$$

Where, ni sub(j) = the value of node j, w sub(j) = scaled number of samples reaching node j, C sub(j) = the value of the impurity node j, left(j) = toddler node from the left split on node j, right(j) =+ toddler node from right split on node j, fi sub(i) = the importance of feature i, ni sub(j) = the importance of node j, RFfi sub(i) = the importance of feature i calculated from all trees in the RF model, normfi sub(ij) = the normalised feature importance for i in tree j, T = total number of trees.

Recently, RF has been successfully applied in water resource applications, offering predictive power with simplicity and speed. RF aids in urban surface thermal environment analysis[53] and enhances high-voltage circuit breaker diagnosis through a hybrid RF and stacked Autoencoder model.[54] Its interpretability supports applications requiring understanding, and its efficiency and minimal parameter adjustment make it a valuable tool.[55]

LOGISTIC REGRESSION

Logistic Regression (LR) is a widely employed statistical model in the field of SGs, being particularly well-suited for classification tasks focused on predicting event probabilities based on predictor variables. In SG applications, LR is effective in tasks such as load forecasting, fault detection, and energy management, and has gained favour among practitioners and researchers. For example, LR can predict the likelihood of power outages based on factors including weather and infrastructure age, thus contributing to enhanced grid reliability. Furthermore, LR aids in optimising the operation of distributed energy resources such as solar panels and wind turbines by predicting their capacity to meet

---

[53]   Xu et al. 2021.
[54]   Ma et al. 2019.
[55]   Tyralis et al. 2019.

Studies •

electricity demand. In summary, LR enhances efficiency, cost-effectiveness, and grid reliability and sustainability in SG. LR algorithms are applied for both classification and regression techniques, while Equation (11) embodies the core equation for evaluating the LR ML algorithm.[56]

$$y = \frac{e^{(b_0 + b_1 x)}}{1 + e^{(b_0 + b_1 x)}} \tag{11}$$

Where x = input value, y = predicted output, b_0 is bias or intercept term, b_1 is the coefficient for input (x).

Drilling engineering, a crucial aspect of gas and oil exploration, involves significant investments and technological complexities. Deng et al. (2021) used LR algorithms to predict the rate of penetration, contributing to the field.[57] Hewett et al. (2020) proposed a five-factor maximum model for risk prediction with LR in drilling engineering.[58] Sun et al. (2018) focused on environmental ecosystem monitoring, employing an early-warning LR model for timely alerts.[59] Additionally, Bashir et al. (2021) found LR to outperform other classification algorithms, achieving high precision and accuracy in experiments.[60]

ANALYSIS AND DISCUSSION

The reliability of various ML algorithms in SG applications can be assessed using a dependability score based on performance metric normalisation. Equation (12), developed by Band et al. (2022), demonstrates the min-max normalisation process for these metrics, ensuring comparability with scores ranging from 0 to 1.[61]

$$Y_N = \frac{f(Accuracy, Precision, Recall, RMSE, Correlation\ Coefficient) - Y_{min}}{Y_{max} - Y_{min}} \tag{12}$$

The dependability scores have been divided into four zones for easier interpretation:
1. Low if $0 \leq YN < 0.25$
2. Moderate if $0.25 \leq YN < 0.5$
3. High if $0.5 \leq YN < 0.75$
4. Very high if $0.75 \leq YN < 1.0$

---

56   Bashir et al. 2022.
57   Deng et al. 2021.
58   Hewett et al. 2020.
59   Sun et al. 2018.
60   Bashir et al. 2021.
61   Band et al. 2022.

According to the reliability analysis, the most reliable ML algorithm for SG applications is ANN. Additionally, ELM, RF, and LR exhibit comparatively high levels of reliability within the SG context.

Efficiency analysis

Figure 5 illustrates an efficiency analysis using processing time as a metric for the surveyed ML methods, employing min-max normalisation for score comparability shown in Equation (13).[62]

$$X_N = \frac{\frac{Absolute\ processing\ time(s)}{data\ samples} - X_{min}}{X_{max} - X_{min}} \tag{13}$$

The equation defines the normalisation of processing time scores using the min-max method, yielding scores between 0 and 1, and categorises them into four zones for clarity.
1. Low if $0 \leq X_N < 0.25$
2. Moderate if $0.25 \leq X_N < 0.5$
3. High if $0.5 \leq X_N < 0.75$
4. Very high if $0.75 \leq X_N < 1.0$

A lower score signifies faster ML algorithms, with ANN demonstrating the fastest performance, while DL and hybrid/ensemble models exhibit reduced speed due to their intricate computational structure.

Table 1: Novel studies based on ML-based method in SG

Reference	Contribution	Application	Source	ML-based method
ZHENG et al. 2021	Addressing the load shifting associated with demand response	Customer incentive pricing	IET SG	Sliding time window technique, genetic algorithm
JARMOUNI et al. 2021	Ensuring the safety of electricity for consumers while maximising the integration of RESs	Multiple-source energy management	International Journal of Power Electronics and Drive Systems	Multilayer perceptron network
LIU & SHU 2021	Implementing a hybrid ML model to minimise false attacks in SGs	Security	Computers and Security	Gradient-based and population-based algorithm
GUPTA et al. 2021	Suggesting a cyber detection method for recognising cyber intrusions in the SG	FDI system	International Journal of Engineering, Transactions B: Applications	Intelligent Loop Based-ANN

---

62   BAND et al. 2022.

Studies •

Reference	Contribution	Application	Source	ML-based method
Teekaraman et al. 2022	Identifying regression losses within the SG	Loss identification in large-scale system	Wireless Communications and Mobile Computing	LSTM, heuristic algorithm, Adaptive ARIMA, Linear regression
Bao et al. 2022	Optimising the dispatch of the SG through a multi-objective approach	Forecasting the actual load	Mobile Information Systems	Multi-particle swarm optimisation
Wang et al. 2022	Real-time assessment of regulatory system risks	Communication and data security	Journal of Shenyang University of Technology	CNN-SVM classification model
Dou et al. 2022	Evaluating the detection accuracy and assessing the impact of attack intensity and environmental noise on the performance of the SG	FDI attack	CSEE Journal of Power and Energy Systems	Variational Mode Decomposition, OS-ELM
Dewangan et al. 2022	Introducing a Genetic Algorithm-ELM for forecasting the stability of cyber-physical systems	Stability prediction	Electric Power Systems Resiliency: Modelling, Opportunity, and Challenges	Hessenberg Decomposition ELM, GA-based ELM
Zhang et al. 2020	Enhances the accuracy, detection rate, and precision of intrusion detection while minimising the false positive rate	Intrusion detection	Energies	Online sequential-ELM, GA-ELM
Singh & Govindarasu 2018	Reducing the impact on both system reliability and economic factors	Anomaly detection	IEEE Power and Energy Society General Meeting	Special Protection Scheme, Intelligent Remedial Action Scheme
Steer et al. 2012	Implementing a receding horizon controller to dynamically adjust the output of the SG in real-time	Online operation of SG	Energy Conversion and Management	C4.5 algorithm, Particle swarm optimisation
Ganesan et al. 2022	Regularly monitoring the operational status of wearable sensing devices in the SG	Human activity recognition	Mathematical Problems in Engineering	K-means ++ algorithm, RF
Chen et al. 2021	Enhancing energy efficiency and promoting the use of clean energy	Energy optimisation	IOP Conference Series: Earth and Environmental Science	RF
Lin et al. 2020	Enhancing the precision, recall rate of fault prediction, and reducing the rate of negative samples	Fault prediction	Enterprise Information Systems	Voted RF algorithm, NSGA algorithm

| Reference	Contribution	Application	Source	ML-based method
Moldovan 2021	Categorising the stability status of an SG	Stability	Lecture Notes in Networks and Systems	Improved Kangaroo Mob Optimisation, LR
Manoharan et al. 2020	Identifying optimal solutions for energy consumption, distance, and cost	Monitoring	Energies	Binary LR
Mukherjee et al. 2020	Developing a cost-effective solution for interconnecting electrical and electronic devices	Lightweight sustainable intelligent LF	Sustainable Computing: Informatics and Systems	KNN-regressor model, SVM
Naz et al. 2019	Ensuring the effectiveness of load scheduling and reducing prices	Short-term electric load and price forecasting	Energies	Enhanced-LR, Enhanced Recurrent ELM

Source: compiled by the authors
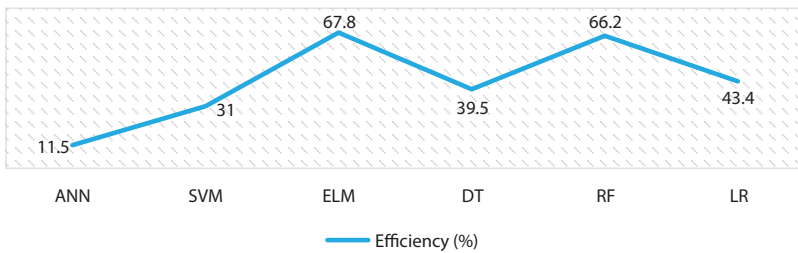
Figure 5: The score for processing speed of ML-based methods used to SG applications
Source: compiled by the authors

Figure 6 shows the evaluation criteria used in the reviewed papers for ML methods, where such metrics as accuracy, precision, recall, correlation coefficient, and various error-related measurements are prevalent, while specialised or complementary metrics are less frequently employed.

Obtaining precise accuracy percentages for different ML algorithms in SG data security is challenging due to factors including application, training data quality, problem complexity, and hyperparameters. Typically, more complex algorithms, such as ANNs and SVMs, tend to achieve higher accuracy compared to simpler ones like DT and KNN. However, actual accuracy varies depending on the specific application and data quality. It is important to consider factors beyond accuracy, including computational resources, model interpretability, and generalisation capabilities. Figure 6 depicts accuracy values, highlighting the lower accuracy produced by single ML models (e.g. ANN, DE, binary LR, Feed Forward NNs) across various SG applications.
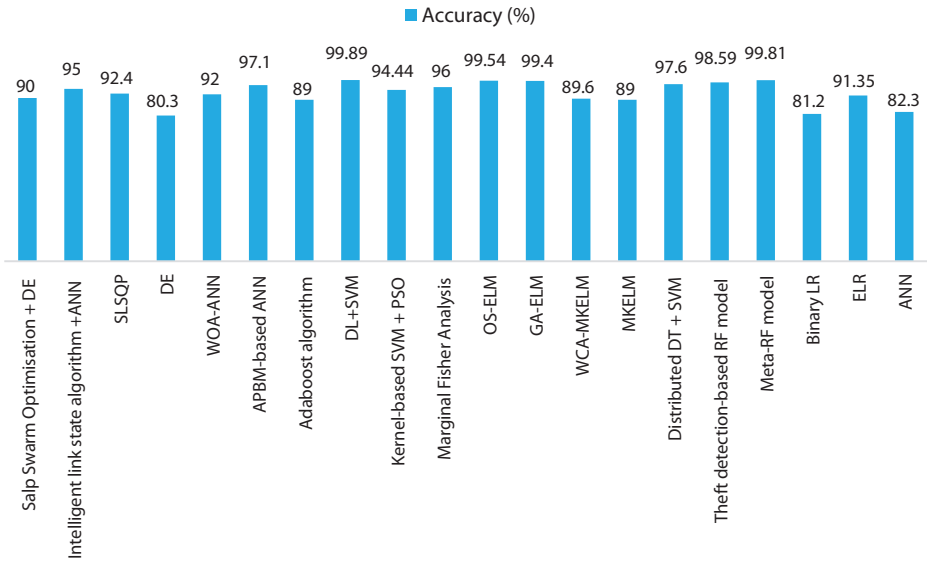
Studies •

Figure 6: Comparison of the models' accuracies
Source: compiled by the authors

Multiple ML algorithms are useful for SG applications, including LR for continuous value predictions, DTs for adaptable classification, RF for ensemble learning, ANN for various types of parameter forecasting, SVM for its trade-off between high accuracy and processing speed, and ELM for rapid scalability. Selection of the most appropriate algorithm for a task hinges on the specific challenges and data attributes within SG contexts.

Figure 7 depicts the allocation of ML methods across diverse SG applications, with ANN and RF models being prevalent in security. Additionally, ELM and ANN excel at FDI, whereas ELM is predominant in customer incentive pricing, while LR and ANN dominate energy management. RF emerges as a suitable choice for real-time voltage stability monitoring in the SG. ML algorithms play a crucial role in optimising Demand Response by accurately predicting energy demand in the Grid, thereby enhancing program efficiency and grid management. Google's DeepMind utilises ML for renewable energy integration, specifically in wind farms, improving their overall efficiency through advanced wind forecasting and turbine optimisation. Predictive analytics powered by ML contribute to grid stability and predictive maintenance, proactively identifying potential equipment failures. ML is applied for energy theft detection, analysing consumption patterns to prevent unauthorised usage. ML models are employed to improve the grid's resilience to natural disasters, simulating and strategising for potential impacts post-events. ML is also leveraged for electric vehicle integration, by optimising charging infrastructure and alleviating strain on distribution networks. ML is utilised for grid anomaly detection, swiftly identifying and responding to cybersecurity threats and unexpected grid behaviour in real-time.
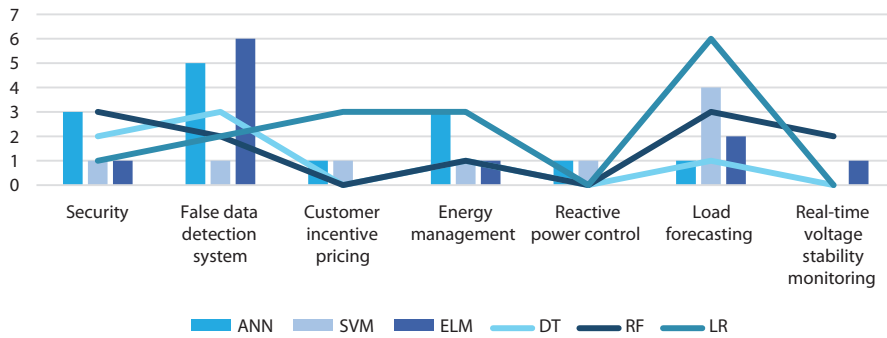
Figure 7: Distribution of the utilisation percentages of various ML techniques in this study among diverse application categories within the field of SG
Source: compiled by the authors

Limitations, challenges, and future trends

The definition of AI varies by time. Therefore, this study instead of reviewing the AI methods, enforced limitation to only investigate standard ML methods. The deep learning methods had been excluded from this study and can be studied in a future research. There are some limitations in the usefulness of ML for SG Applications: a) data quality and variability – ML models face difficulties due to the quality and variability of SG data, affecting their performance as they struggle with inaccuracies and fluctuations; b) interpretability –some ML models exhibit inherent complexity, reducing their interpretability. This limitation hinders a clear understanding of decision-making processes, which is crucial in critical infrastructure like SGs; c) scalability – the growth of SG systems raises concerns about scalability, particularly when handling extensive data processing and resource-intensive ML algorithms. The challenges arising from the use of ML for SG applications included: A) privacy and ethical concerns – striking a balance between the demand for data-driven insights and privacy considerations presents a significant challenge. Careful implementation of privacy-preserving measures and ethical considerations is thus essential; B) regulatory compliance – continuous adherence to evolving regulations and ensuring compliance with legal frameworks governing ML applications in SGs poses an ongoing challenge for system operators and developers; C) cybersecurity risks – ML models and the data they handle are susceptible to cyber threats, demanding the implementation of robust cybersecurity measures to safeguard against potential attacks and breaches. The likely future trends in ML for SG applications may include: A) explainable AI- anticipated future ML models in SGs will prioritise enhancing interpretability through explainable-AI techniques, fostering transparency and trust in decision-making processes; B) edge computing for real-time processing: The integration of edge computing with ML in SGs is set to revolutionise real-time data processing at the edge, reducing latency, and

Studies •

enhancing overall system responsiveness; C) federated learning – the future may witness the rise of federated learning, emphasising decentralised model training across devices. This approach is expected to address privacy concerns effectively by keeping sensitive data localised; D) hybrid models – an emerging trend is the combination of traditional physics-based models with ML techniques, leading to more accurate and robust SG models that leverage the strengths of both approaches.

CONCLUSION

This article studied the ML applications in SGs, conducted a systematic review, introducing a novel taxonomy, and carried out a thorough comparative performance evaluation. The well-structured taxonomy established a foundation for categorising and comprehending various ML approaches within SG context. The performance evaluation not only illuminates the strengths and weaknesses of the various ML models, but also furnishes invaluable insights for practitioners aiming to implement effective solutions in real-world SG scenarios. Navigating the intricate landscape of contemporary energy distribution, the integration of ML not only becomes a technological imperative but also a transformative catalyst, amplifying efficiency, resilience, and sustainability in SG operations. This research not only contributes to the academic debate, but can also serve as a pragmatic guide for engineers, system operators, and policymakers striving to advance and optimise SG infrastructures. This research found that the applications of ANNs, ELM and RF methods are popular. Looking forward, the findings presented underscore the dynamic nature of the field, fostering continual research, development, and collaboration to propel innovation and tackle the evolving challenges in SG management.

TABLE OF ACRONYMS

RESs	Renewable Energy Resources
SG	Smart Grid
ML	Machine Learning
KNN	K-Nearest Neighbours
DL	Deep Learning
AI	Artificial Intelligence
ANN	Artificial Neural Network
NN	Neural Network
SVM	Support Vector Machine
ELM	Extreme Learning Machines
DT	Decision Tree
CART	Classification and Regression Tree Algorithm
CTree	Conditional Inference tree

RF	Random Forest
LR	Logistic Regression
FDI	False Data Injection attack
LSTM	Long-Short Term Memory
CNN	Convolutional Neural Network
NB	Naive Bayes
IoT	Internet of Things

REFERENCES

Achlerkar, Pankaj D. – Samantaray, S. R. – Manikandan, M. Sabbarimalai (2016): Variational Mode Decomposition and Decision Tree-based Detection and Classification of Power Quality Disturbances in Grid-Connected Distributed Generation System. *IEEE Transactions on Smart Grid,* 9(4), 3122–3132. Online: https://doi.org/10.1109/TSG.2016.2626469

Ahmad, Tanveer – Zhu, Hongyu – Zhang, Dongdong – Tariq, Rasikh – Bassam, A. – Ullah, Fasih – AlGhamdi, Ahmed, S. – Alshamrani, Sultan S. (2022): Energetics Systems and Artificial Intelligence: Applications of Industry 4.0. *Energy Reports,* 8, 334–361. Online: https://doi.org/10.1016/j.egyr.2021.11.256

Ahmed, Saeed – Lee, Youngdoo – Hyun, Seung-Ho – Koo, Insoo (2019): Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest. *IEEE Transactions on Information Forensics and Security,* 14(10), 2765–2777. Online: https://doi.org/10.1109/tifs.2019.2902822

Aly, Hamed H.H. (2020): A Proposed Intelligent Short-Term Load Forecasting Hybrid Models of ANN, WNN and KF Based on Clustering Techniques for Smart Grid. *Electric Power Systems Research,* 182, 106191. Online: https://doi.org/10.1016/j.epsr.2019.106191

Ardito, Luca – Procaccianti, Giuseppe – Menga, Giuseppe – Morisio, Maurizio (2013): Smart Grid Technologies in Europe: An Overview. *Energies,* 6(1), 251–281. Online: https://doi.org/10.3390/en6010251

Areo, Oluwafadekemi S. – Gershon, Obindah – Osabuohien, Evans (2020): Improved Public Services and Tax Compliance of Small and Medium Scale Enterprises in Nigeria: A Generalised Ordered Logistic Regression. *Asian Economic and Financial Review,* 10(7), 833–860. Online: https://doi.org/10.18488/journal.aefr.2020.107.833.860

Banga, Alisha – Ahuja, Ravinder – Sharma, S. C. (2022): Accurate Detection of Electricity Theft Using Classification Algorithms and Internet of Things in Smart Grid. *Arabian Journal for Science and Engineering,* 47(8), 9583–9599. Online: https://doi.org/10.1007/s13369-021-06313-z

Band, Shahab S. – Ardabili, Sina – Sookhak, Mehdi – Chronopoulos, Theodore A. – Elnaffar, Said – Moslehpour, Massoud – Makó, Csaba – Török, Bernát – Pai, Hao-Ting – Mosavi, Amir (2022): When Smart Cities Get Smarter via Machine

Learning: An In-depth Literature Review. *IEEE Access,* 10, 60985–61015. Online: https://doi.org/10.1109/ACCESS.2022.3181718

Bao, Man – Zhang, Hongqi – Wu, Hao – Zhang, Chao – Wang, Zixu – Zhang, Xiaohui (2022): Multiobjective Optimal Dispatching of Smart Grid Based on PSO and SVM. *Mobile Information Systems,* 2051773. Online: https://doi.org/10.1155/2022/2051773

Bashir, Ali K. – Khan, Suleman – Prabadevi, B. – Deepa, N. – Alnumay, Waleed S. – Gadekallu, Thippa R. – Maddikunta, Praveen K. R. (2021): Comparative Analysis of Machine Learning Algorithms for Prediction of Smart Grid Stability†. *International Transactions on Electrical Energy Systems,* 31(9), e12706. Online: https://doi.org/10.1002/2050-7038.12706

Bashir, Hayat – Lee, Seonah – Kim, Kyong H. (2022): Resource Allocation through Logistic Regression and Multicriteria Decision Making Method in IoT Fog Computing. *Transactions on Emerging Telecommunications Technologies,* 33(2), e3824. Online: https://doi.org/10.1002/ett.3824

Bari, Ataul – Jiang, Jin – Saad, Walid – Jaekel, Arunita (2014): Challenges in the Smart Grid Applications: An Overview. *International Journal of Distributed Sensor Networks,* 10(2), 974682. Online: https://doi.org/10.1155/2014/974682

Behara, Ramesh K. – Saha, Akshay K. (2022): Artificial Intelligence Methodologies in Smart Grid-Integrated Doubly Fed Induction Generator Design Optimization and Reliability Assessment: A Review. *Energies,* 15(19), 7164. Online: https://doi.org/10.3390/en15197164

Berghout, Tarek – Benbouzid, Mohamed – Muyeen, S. M. (2022): Machine Learning for Cybersecurity in Smart Grids: A Comprehensive Review-based Study on Methods, Solutions, and Prospects. *International Journal of Critical Infrastructure Protection,* 38, 100547. Online: https://doi.org/10.1016/j.ijcip.2022.100547

Bhagya Raj, G. V. S. – Dash, Kshirod K. (2022): Comprehensive Study on Applications of Artificial Neural Network in Food Process Modeling. *Critical Reviews in Food Science and Nutrition,* 62(10), 2756–2783. Online: https://doi.org/10.1080/10408398.2020.1858398

Blanco, Victor – Japón, Alberto – Puerto, Justo (2022): A Mathematical Programming Approach to SVM-based Classification with Label Noise. *Computers & Industrial Engineering,* 172, 108611. Online: https://doi.org/10.1016/j.cie.2022.108611

Chaurasia, Kiran – Kamath, H. Ravishankar (2022): Artificial Intelligence and Machine Learning Based: Advances in Demand-Side Response of Renewable Energy-Integrated Smart Grid. In Somani, Arun K. – Mundra, Ankit – Doss, Robin – Bhattacharya, Subhajit (eds.): *Smart Systems: Innovations in Computing.* Singapore: Springer, 195–207. Online: https://doi.org/10.1007/978-981-16-2877-1_18

Chen, Lingxuan – Xie, Pengyang – Ma, Shuaihua – Yang, Chen (2021): Evaluation of Low-Carbon Benefits of Smart Grid Based on Random Forest Algorithm. *IOP Conference Series: Earth and Environmental Science,* 804(3), 032001. Online: https://doi.org/10.1088/1755-1315/804/3/032001

Chen, Zhenyu – Yuan, Shuai – Wu, Longfei – Guan, Zhitao – Du, Xiaojiang (2022):
False Data Injection Attack Detection Based on Wavelet Packet Decomposition and
Random Forest in Smart Grid. In *2021 IEEE 23rd International Conference on High
Performance Computing and Communications, 7th International Conference on Data
Science and Systems, 19th International Conference on Smart City and 7th International
Conference on Dependability in Sensor, Cloud and Big Data Systems and Applications,*
HPCC-DSS-SmartCity-DependSys Haikou, Hainan, China, 1965–1971. Online: https://
doi.org/10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00294

Cui, Lei – Qu, Youyang – Gao, Longxiang – Xie, Gang – Yu, Shui (2020): Detecting False
Data Attacks Using Machine Learning Techniques in Smart Grid: A Survey. *Journal
of Network and Computer Applications,* 170, 102808. Online: https://doi.org/10.1016/j.
jnca.2020.102808

Da Cunha, Guilherme L. – Fernandes, Ricardo A. S. – Fernandes, Tatiane C. C.
(2022): Small-Signal Stability Analysis In Smart Grids: An Approach Based on
Distributed Decision Trees. *Electric Power Systems Research,* 203, 107651. Online:
https://doi.org/10.1016/j.epsr.2021.107651

Das, Ratnakar – Mishra, Jibitesh – Mishra, Sujogya – Pattnaik, P. K. – Das,
Subhalaxmi (2022): Mathematical Modeling using Rough Set and Random Forest
Model to Predict Wind Speed. In *2022 9th International Conference on Computing for
Sustainable Global Development* (INDIACom), New Delhi, India, 207–213. Online:
https://doi.org/10.23919/INDIACom54597.2022.9763275

Deng, Shuang – Wei, Minghui – Xu, Mingze – Cai, Wei (2021): Prediction of the Rate of
Penetration Using Logistic Regression Algorithm of Machine Learning Model. *Arabian
Journal of Geosciences,* 14(21), 1–13. Online: https://doi.org/10.1007/s12517-021-08452-x

Dewangan, Fanidhar – Biswal, Monalisa – Patnaik, Bhaskar – Hasan,
Shazia – Mishra, Manohar (2022): Smart Grid Stability Prediction Using Genetic
Algorithm-Based Extreme Learning Machine. In Bansal, Ramesh C. – Mishra,
Manohar – Sood, Yog Raj (eds.): *Electric Power Systems Resiliency: Modelling,
Opportunity and Challenges.* [s. l.]: Academic Press, 149–163. Online: https://doi.
org/10.1016/B978-0-323-85536-5.00011-4

Dou, Chunxia – Wu, Di – Yue, Dong – Jin, Bao – Xu, Shiyun (2019): A Hybrid Method
for False Data Injection Attack Detection in Smart Grid Based on Variational Mode
Decomposition and OS-ELM. *CSEE Journal of Power and Energy Systems,* PP (99),
118195762. Online: https://doi.org/10.17775/CSEEJPES.2019.00670

Dou, Chunxia – Wu, Di – Yue, Dong – Jin, Bao – Xu, Shiyun (2022): A Hybrid Method
for False Data Injection Attack Detection in Smart Grid Based on Variational Mode
Decomposition and OS-ELM. *CSEE Journal of Power and Energy Systems,* 8(6), 1697–
1707. Online: https://doi.org/10.17775/cseejpes.2019.00670

Durairaj, Danalakshmi – Wróblewski, Łukasz – Sheela, A. – Hariharasudan,
A. – Urbański, Mariusz (2022): Random Forest Based Power Sustainability and Cost
Optimization in Smart Grid. *Production Engineering Archives,* 28(1), 82–92. Online:
https://doi.org/10.30657/pea.2022.28.10

Studies •

Eissa, M. M. – Ali, A. A. – Abdel-Latif, K. M. – Al-Kady, A. F. (2019): A Frequency Control Technique Based on Decision Tree Concept by Managing Thermostatically Controllable Loads at Smart Grids. *International Journal of Electrical Power and Energy Systems,* 108, 40–51. Online: https://doi.org/10.1016/j.ijepes.2018.12.037

Elbouchikhi, Elhoussin – Zia, Muhammad F. – Benbouzid, Mohamed – El Hani, Soumia (2021): Overview of Signal Processing and Machine Learning for Smart Grid Condition Monitoring. *Electronics,* 10(21), 2725. Online: https://doi.org/10.3390/electronics10212725

Ganesan, Anusha – Paul, Anand – Seo, HyunCheol (2022): Elderly People Activity Recognition in Smart Grid Monitoring Environment Mathematical Problems in Engineering, 2022, 9540033. Online: https://doi.org/10.1155/2022/9540033

Gupta, Praveen K. – Singh, Neeraj K. – Mahajan, Vasundhara (2021): Intrusion Detection in Cyber-Physical Layer of Smart Grid Using Intelligent Loop Based Artificial Neural Network Technique. *International Journal of Engineering, Transactions B: Applications,* 34(5), 1250–1256. Online: https://doi.org/10.5829/ije.2021.34.05b.18

Guyot, Dimitri – Giraud, Florine – Simon, Florian – Corgier, David – Marvillet, Christophe – Tremeac, Brice (2019): Overview of the Use of Artificial Neural Networks for Energy-Related Applications in the Building Sector. *International Journal of Energy Research,* 43(13), 6680–6720. Online: https://doi.org/10.1002/er.4706

Hafeez, Ghulam – Alimgeer, Khurram S. – Qazi, Abdul B. – Khan, Imran – Usman, Muhammad – Khan, Farrukh A. – Wadud, Zahid (2020): A Hybrid Approach for Energy Consumption Forecasting with a New Feature Engineering and Optimization Framework in Smart Grid. *IEEE Access,* 8, 9057473. 96210–96226. Online: https://doi.org/10.1109/access.2020.2985732

Heidari, M. – Seifossadat, G. – Razaz, M. (2013): Application of Decision Tree and Discrete Wavelet Transform for an Optimized Intelligent-Based Islanding Detection Method in Distributed Systems with Distributed Generations. *Renewable and Sustainable Energy Reviews,* 27, 525–532. Online: https://doi.org/10.1016/j.rser.2013.06.047

Hewett, Timothy E. – Webster, Kate E. – Hurd, Wendy J. (2019): Systematic Selection of Key Logistic Regression Variables for Risk Prediction Analyses: A Five Factor Maximum Model. *Clinical Journal of Sport Medicine,* 29(1), 78–85. Online: https://doi.org/10.1097/jsm.0000000000000486

Hossain, Eklas – Khan, Imtiaj – Un-Noor, Fuad – Sikander, Sarder S – Sunny, M. Samiul (2019): Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review. *IEEE Access,* 7, 13960–13988. Online: https://doi.org/10.1109/access.2019.2894819

Hua, Weiqi – Chen, Ying – Qadrdan, Meysam – Jiang, Jing – Sun, Hongjian – Wu, Jianzhong (2022): Applications of Blockchain and Artificial Intelligence Technologies for Enabling Prosumers in Smart Grids: A Review. *Renewable and Sustainable Energy Reviews,* 161, 112308. Online: https://doi.org/10.1016/j.rser.2022.112308

Jarmouni, Eziitouni – Mouhsen, Ahmed – Lamhammedi, Mohammed – Ouldzira, Hicham (2021): Integration of Artificial Neural Networks for Multi-Source Energy Management in a Smart Grid. *International Journal of Power Electronics and Drive Systems,* 12(3), 1919–1927. Online: https://doi.org/10.11591/ijpeds.v12.i3.pp1919-1927

Javaid, Nadeem – Qasim, Umar – Yahaya, Adamu S. – Alkhammash, Eman H. – Hadjouni, Myriam (2022): Non-Technical Losses Detection Using Autoencoder and Bidirectional Gated Recurrent Unit to Secure Smart Grids. *IEEE Access,* 10, 56863–56875. Online: https://doi.org/10.1109/access.2022.3171229

Jawad, Jasir – Hawari, Alaa H. – Zaidi, Syed J. (2021): Artificial Neural Network Modeling of Wastewater Treatment and Desalination Using Membrane Processes: A Review. *Chemical Engineering Journal,* 419, 129540. Online: https://doi.org/10.1016/j.cej.2021.129540

Jena, M. – Dehuri, S. (2020): An Experimental Study on Decision Tree Classifier Using Discrete and Continuous Data. In Mallick, P. – Balas, V. – Bhoi, A. – Chae, GS. (eds.): *Cognitive Informatics and Soft Computing.* Advances in Intelligent Systems and Computing, vol 1040. Singapore: Springer. Online: https://doi.org/10.1007/978-981-15-1451-7_35

Kariminia, Shahab – Shamshirband, Shahaboddin – Motamedi, Shervin – Hashim, Roslan – Roy, Chandrabhushan (2016): A Systematic Extreme Learning Machine Approach to Analyze Visitors' Thermal Comfort at a Public Urban Space. *Renewable and Sustainable Energy Reviews,* 58, 751–760. Online: https://doi.org/10.1016/j.rser.2015.12.321

Kwak, Moon K. – Heo, Seok (2007): Active Vibration Control of Smart Grid Structure by Multi-input and Multioutput Positive Position Feedback Controller. *Journal of Sound and Vibration,* 304(1–2), 230–245. Online: https://doi.org/10.1016/j.jsv.2007.02.021

Li, B. – Delpha, C. – Diallo, D. – Migan-Dubois, A. (2021): Application of Artificial Neural Networks to Photovoltaic Fault Detection and Diagnosis: A Review. *Renewable and Sustainable Energy Reviews,* 138, 110512. Online: https://doi.org/10.1016/j.rser.2020.110512

Li, Huifang – Li, Yidong – Dong, Hairong (2017): A Comprehensive Learning-Based Model for Power Load Forecasting in Smart Grid. *Computing and Informatics,* 36(2), 470–492. Online: https://doi.org/10.4149/cai_2017_2_470

Li, Xiaonuo – Yi, Shiyi – Cundy, Andrew B. – Chen, Weiping (2022): Sustainable Decision-Making for Contaminated Site Risk Management: A Decision Tree Model Using Machine Learning Algorithms. *Journal of Cleaner Production,* 371, 133612. Online: https://doi.org/10.1016/j.jclepro.2022.133612

Li, Yuancheng – Qiu, Rixuan – Jing, Sitong (2018): Intrusion Detection System using Online Sequence Extreme Learning Machine (OSELM) in Advanced Metering Infrastructure of Smart Grid. *PLoS ONE,* 13(2), e0192216. Online: https://doi.org/10.1371/journal.pone.0192216

Lin, Rongheng – Pei, Zixiang – Ye, Zezhou – Wu, Budan – Yang, Geng (2020): A Voted Based Random Forests Algorithm for Smart Grid Distribution Network Faults

Studies ●

Prediction. *Enterprise Information Systems,* 14(4), 496–514. Online: https://doi.org/10.1080/17517575.2019.1600724

Liu, Fen – Yu, Zheng – Wang, Yixi – Feng, Hao – Zha, Zhiyong – Liao, Rongtao – Zhang, Ying (2020): Falsified Data Filtering Method for Smart Grid Wireless Communication Based on SVM. *International Journal of Internet Protocol Technology,* 13(4), 177–183. Online: https://doi.org/10.1504/ijipt.2020.10028657

Liu, Tian – Shu, Tao (2021): On the Security of ANN-Based AC State Estimation in Smart Grid. *Computers and Security,* 105, 102265. Online: https://doi.org/10.1016/j.cose.2021.102265

Liu, Xinchun (2021): Empirical Analysis of Financial Statement Fraud of Listed Companies Based on Logistic Regression and Random Forest Algorithm. *Journal of Mathematics,* 2021(Special Issue). Online: https://doi.org/10.1155/2021/9241338

Ma, Suliang – Chen, Mingxuan – Wu, Jianwen – Wang, Yuhao – Jia, Bowen – Jiang, Yuan (2019): High-Voltage Circuit Breaker Fault Diagnosis Using a Hybrid Feature Transformation Approach Based on Random Forest and Stacked Autoencoder. *IEEE Transactions on Industrial Electronics,* 66(12), 9777–9788. Online: https://doi.org/10.1109/tie.2018.2879308

Manoharan, Hariprasath – Teekaraman, Yuvaraja – Kirpichnikova, Irina – Kuppusamy, Ramya – Nikolovski, Srete – Baghaee, Hamid R. (2020): Smart Grid Monitoring by Wireless Sensors Using Binary Logistic Regression. *Energies,* 13(15), 3974. Online: https://doi.org/10.3390/en13153974

Mei, ShengWei – Chen, LaiJun (2013): Recent Advances on Smart Grid Technology and Renewable Energy Integration. *Science China Technological Sciences,* 56(12), 3040–3048. Online: https://doi.org/10.1007/s11431-013-5414-z

Mishra, Manohar – Nayak, Janmenjoy – Naik, Bignaraj – Patnaik, Bhaskar (2022): Enhanced Memetic Algorithm-Based Extreme Learning Machine Model for Smart Grid Stability Prediction. International Transactions on Electrical Energy Systems, 2022. Online: https://doi.org/10.1155/2022/8038753

Moldovan, Dorin (2021): Improved Kangaroo Mob Optimization and Logistic Regression for Smart Grid Stability Classification. In Silhavy, R. (ed.): *Artificial Intelligence in Intelligent Systems.* CSOC 2021. Lecture Notes in Networks and Systems, vol 229. Cham: Springer, 469–487. Online: https://doi.org/10.1007/978-3-030-77445-5_44

Mukherjee, Amartya – Mukherjee, Prateeti – Dey, Nilanjan – De, Debashis – Panigrahi, B. K. (2020): Lightweight Sustainable Intelligent Load Forecasting Platform for Smart Grid Applications. *Sustainable Computing: Informatics and Systems,* 25, 100356. Online: https://doi.org/10.1016/j.suscom.2019.100356

Mohanty, D. K. – Parida, Ajaya K. – Khuntia, Shelly S. (2021): Financial Market Prediction under Deep Learning Framework using Auto Encoder and Kernel Extreme Learning Machine. *Applied Soft Computing,* 99, 106898. Online: https://doi.org/10.1016/j.asoc.2020.106898

Nayak, Janmenjoy – Vakula, Kanithi – Dinesh, Paidi – Naik, Bighnaraj – Pelusi, Danilo (2020): Intelligent Food Processing: Journey from Artificial Neural Network to Deep Learning. *Computer Science Review,* 38, 100297. Online: https://doi.org/10.1016/j.cosrev.2020.100297

Naz, Aqdas – Javed, Muhammad U. – Javaid, Nadeem – Saba, Tanzila – Alhussein, Musaed – Aurangzeb, Khursheed (2019): Short-Term Electric Load and Price Forecasting Using Enhanced Extreme Learning Machine Optimization in Smart Grids. *Energies,* 12(5), 866. Online: https://doi.org/10.3390/en12050866

Niazi, K. R. – Arora, C. M. – Surana, S. L. (1999): Transient Security Evaluation of Power Systems Using Decision Tree Method. *Journal-Institution of Engineers India Part El Electrical Engineering Division,* 76–79.

Omitaomu, Olufemi A. – Niu, Haoran (2021): Artificial Intelligence Techniques in Smart Grid: A Survey. *Smart Cities,* 4(2), 548–568. Online: https://doi.org/10.3390/smartcities4020029

Radoglou-Grammatikis, Panagiotis I. – Sarigiannidis, Panagiotis G. (2019): An Anomaly-Based Intrusion Detection System for the Smart Grid Based on CART Decision Tree. *2018 Global Information Infrastructure and Networking Symposium (GIIS),* Thessaloniki, Greece, 1–5. Online: https://doi.org/10.1109/giis.2018.8635743

Ranganathan, Raghuram – Qiu, Robert – Hu, Zhen – Hou, Shujie – Pazos-Revilla, Marbin – Zheng, Gang – Zhe, Chen – Guo, Nan (2011): Cognitive Radio for Smart Grid: Theory, Algorithms, and Security. *International Journal of Digital Multimedia Broadcasting,* 2011. Online: https://doi.org/10.1155/2011/502087

Rangel-Martinez, Daniel – Nigam, K. D. – Ricardez-Sandoval, Luis A. (2021): Machine Learning on Sustainable Energy: A Review and Outlook on Renewable Energy Systems, Catalysis, Smart Grid and Energy Storage. *Chemical Engineering Research and Design,* 174, 414–441. Online: https://doi.org/10.1016/j.cherd.2021.08.013

Raza, Muhammad Q. – Khosravi, Abbas (2015): A Review on Artificial Intelligence-based Load Demand Forecasting Techniques for Smart Grid and Buildings. *Renewable and Sustainable Energy Reviews,* 50, 1352–1372. Online: https://doi.org/10.1016/j.rser.2015.04.065

Roger, Coleen – Lasbleiz, Adèle – Guye, Maxime – Dutour, Anne – Gaborit, Bénédicte – Ranjeva, Jean-Philippe (2022): The Role of the Human Hypothalamus in Food Intake Networks: An MRI Perspective. *Frontiers in Nutrition,* 8, 1191. Online: https://doi.org/10.3389/fnut.2021.760914

Singh, V.K. – Govindarasu, M. (2021): Cyber Kill Chain-Based Hybrid Intrusion Detection System for Smart Grid. In Haes Alhelou, H. – Abdelaziz, A. Y. – Siano, P. (eds.): *Wide Area Power Systems Stability, Protection, and Security.* Cham: Springer. Online: https://doi.org/10.1007/978-3-030-54275-7_22

Steer, Kent C. B. – Wirth, Andrew – Halgamuge, Saman K. (2012): Decision Tree Ensembles for Online Operation of Large Smart Grids. *Energy Conversion and Management,* 59, 9–18. Online: https://doi.org/10.1016/j.enconman.2012.01.010

Studies •

Sun, Li-Yan – Miao, Cheng-lin – Yang, Li (2018): Ecological Environmental Early-Warning Model for Strategic Emerging Industries in China Based on Logistic Regression. *Ecological Indicators,* 84, 748–752. Online: https://doi.org/10.1016/j.ecolind.2017.09.036

Taghavinejad, Seyedeh M. – Taghavinejad, Mehran – Shahmiri, Lida – Zavvar, M. – Zavvar, Mohammad H. (2020): Intrusion Detection in IoT-Based Smart Grid Using Hybrid Decision Tree. *2020 6th International Conference on Web Research* (ICWR), Tehran, Iran. 152–156. Online: https://doi.org/10.1109/ICWR49608.2020.9122320

Tahir, Aroosa – Khan, Zahoor A. – Javaid, Nadeem – Hussain, Zeeshan – Rasool, Aimen – Aimal, Syeda (2019): Load and Price Forecasting Based on Enhanced Logistic Regression in Smart Grid. In Barolli, L. – Xhafa, F. – Khan, Z. – Odhabi, H. (eds.): *Advances in Internet, Data and Web Technologies. EIDWT 2019.* Cham: Springer, 221–233. Online: https://doi.org/10.1007/978-3-030-12839-5_21

Teekaraman, Yuvaraja – Kirpichnikova, Irina – Manoharan, Hariprasath – Kuppusamy, Ramya – Angadi, Ravi V. – Thelkar, Amruth R. (2022): Diminution of Smart Grid with Renewable Sources Using Support Vector Machines for Identification of Regression Losses in Large-Scale Systems. *Wireless Communications and Mobile Computing,* 6942029. Online: https://doi.org/10.1155/2022/6942029

Tehrani, Soroush O. – Moghaddam, Mohammad H. Y. – Asadi, Mohsen (2020): Decision Tree based Electricity Theft Detection in Smart Grid. *Proceedings of the 4th International Conference on Smart Cities, Internet of Things and Applications* (SCIoT), Mashad, Iran, 46–51. Online: https://doi.org/10.1109/SCIOT50840.2020.9250194

Tomar, Divya – Agarwal, Sonali (2015): Twin Support Vector Machine: A Review from 2007 to 2014. *Egyptian Informatics Journal,* 16(1), 55–69. Online: https://doi.org/10.1016/j.eij.2014.12.003

Turanzas, J. – Alonso, M. – Amaris, H. – Gutierrez, J. – Pastrana, S. (2022): A Nested Decision Tree for Event Detection in Smart Grids. *Renewable Energy and Power Quality Journal,* 20, 353–358. Online: https://doi.org/10.24084/repqj20.308

Tyralis, Hristos – Papacharalampous, Georgia – Langousis, Andreas (2019): A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. *Water,* 11(5), 910. Online: https://doi.org/10.3390/w11050910

Venkatasubramaniam, Ashwini – Wolfson, Julian – Mitchell, Nathan – Barnes, Timothy – JaKa, Meghan – French, Simone (2017): Decision Trees in Epidemiological Research. *Emerging Themes in Epidemiology,* 14(11), 1–12. Online: https://doi.org/10.1186/s12982-017-0064-4

Wang, N. – Tian, J.-Y. – Dong, N. – Han, M. – Chen, Y. (2022): Real-Time Risk-Assessment and Early-Warning Technology of Smart Grid Regulation System Based on Improved SVM. *Journal of Shenyang University of Technology,* 44(1), 7–13. Online: https://xb.sut.edu.cn/EN/10.7688/j.issn.1000-1646.2022.01.02

Wang, Yuanni – Kong, Tao (2019): Air Quality Predictive Modeling Based on an Improved Decision Tree in a Weather-Smart Grid. *IEEE Access,* 7, 8917641. 172892–172901. Online: https://doi.org/10.1109/access.2019.2956599

Xu, Chongchong – Liao, Zhicheng – Li, Chaojie – Zhou, Xiaojun – Xie, Renyou (2022): Review on Interpretable Machine Learning in Smart Grid. *Energies,* 15(12), 4427. Online: https://doi.org/10.3390/en15124427

Xu, Nuo – Deng, Fan – Liu, Bingqi – Li, Caixia – Fu, Hancong – Yang, Huan – Zhang, Jiahua (2021): Changes in the Urban Surface Thermal Environment of a Chinese Coastal City Revealed by Downscaling MODIS LST with Random Forest Algorithm. *Journal of Meteorological Research,* 35(5), 759–774. Online: https://doi.org/10.1007/s13351-021-0023-4

Xue, Dongbo – Jing, Xiaorong – Liu, Hongqing (2019): Detection of False Data Injection Attacks in Smart Grid Utilizing ELM-based OCON Framework. *IEEE Access,* 7, 31762–31773. Online: https://doi.org/10.1109/access.2019.2902910

Yang, Liqun – Li, Yuancheng – Li, Zhoujun (2017): Improved-ELM Method for Detecting False Data Attack in Smart Grid. *International Journal of Electrical Power and Energy Systems,* 91, 183–191. Online: https://doi.org/10.1016/j.ijepes.2017.03.011

Yaseen, Zaher M. – Sulaiman, Sadeq O. – Deo, Ravinesh C. – Chau, Kwok-Wing (2019): An Enhanced Extreme Learning Machine Model for River Flow Forecasting: State-of-the-art, Practical Applications in Water Resource Engineering Area and Future Research Direction. *Journal of Hydrology,* 569, 387–408. Online: https://doi.org/10.1016/j.jhydrol.2018.11.069

Zain, Azlan M. – Haron, Habibollah – Qasem, Sultan N. – Sharif, Safian (2012): Regression and ANN Models for Estimating Minimum Value of Machining Performance. *Applied Mathematical Modelling,* 36(4), 1477–1492. Online: https://doi.org/10.1016/j.apm.2011.09.035

Zekić-Sušac, M., Has, A., & Knežević, M. (2021): Predicting energy cost of public buildings by artificial neural networks, CART, and random forest. Neurocomputing, 439, 223-233. https://doi.org/10.1016/j.neucom.2020.01.124

Zhang, Ke – Hu, Zhi – Zhan, Yufei – Wang, Xiaofen – Guo, Keyi (2020): A Smart Grid AMI Intrusion Detection Strategy Based on Extreme Learning Machine. *Energies,* 13(18), 4907. Online: https://doi.org/10.3390/en13184907

Zhang, Yanyu – Zeng, Peng – Zang, Chuanzhi (2014): Review of Home Energy Management System in Smart Grid. *Power System Protection and Control,* 42(18), 144–154.

Zheng, Xiulin – Li, Peipei – Wu, Xindong (2022): Data Stream Classification Based on Extreme Learning Machine: A Review. *Big Data Research,* 30, 100356. Online: https://doi.org/10.1016/j.bdr.2022.1003566

Zheng, Yingying – Celik, Berk – Suryanarayanan, Siddharth – Maciejewski, Anthony A. – Siegel, Howard J. – Hansen, Timothy M. (2021): An Aggregator-Based Resource Allocation in the Smart Grid using an Artificial Neural Network and Sliding

Time Window Optimization. *IET Smart Grid,* 4(6), 612–622. Online: https://doi.org/10.1049/stg2.12042

Zidi, Salah – Mihoub, Alaeddin– Mian Qaisar, Saeed – Krichen, Moez – Abu Al-Haija, Qasem (2022): Theft Detection Dataset for Benchmarking and Machine Learning Based Classification in a Smart Grid Environment. *Journal of King Saud University – Computer and Information Sciences,* 35(1), 13–25. Online: https://doi.org/10.1016/j.jksuci.2022.05.007

Zubia, I. – Arrambide, I. – Azurza, O. – García, P. M. – Ugartemendia, J. J. (2013): Rural Smart Grids: Planning, Operation and Control Review. *Renewable Energy and Power Quality Journal,* 1(11), 1099–1104. Online: https://doi.org/10.24084/repqj11.544

Rituraj Rituraj is a PhD candidate at the Óbuda University Doctoral School of Applied Informatics and Applied Mathematics. His research focuses on the applications of AI in smart grids. He proposed methodologies and surveys to improve predictive maintenance, load forecasting, and renewable energy integration in smart grids, which were published in various IEEE venues.

Dániel T. Várkonyi is employed as a Research Assistant Fellow at the Eötvös Loránd University (ELTE), Department of Data Science and Engineering. He has an MSC Degree in Computer Science at ELTE, and a specialised degree in economics at the University of Szeged. His industrial work experience for more than 12 years is connected to BI, Data science and database management in the banking, ICT and heavy industry sector. He changed to the academic in 2019, and is working as a researcher. He is specialised in Datamining on audio signal, in the field of precision livestock farming (PLF). As a university researcher, he is involved in many research projects addressed to audio research, PFL and IoT, and has an interest in biomedical engineering. His research activity and publications are connected to the fields of industry and PLF.

Amir Mosavi is a Research Fellow at the Ludovika University of Public Service Institute of the Information Society. His research interests include applied artificial intelligence. He was an Alexander von Humboldt fellow and the recipient of the Marie Skłodowska-Curie and the European Research Consortium for Informatics and Mathematics fellowships.

József Pap is a PhD candidate at the Széchenyi István University, Győr, Hungary and Head of Supply Chain Excellence & Process Management at Nokia Solutions and Network/Cloud Network Solutions. His research interest is the future of work in the new hybrid reality, from the perspective of organising, managing and enhancing work

by the outcomes of the digital revolution. His latest major project is 'Democracy at Work Through Transparent and Inclusive Algorithmic Management' (INCODING, EU, 2021–2024).

Annamária R. Várkonyi-Kóczy, PhD, DSc, Fellow of the IEEE, Elected member of the Hungarian Academy of Engineers, Vice President of the Hungarian Doctoral Council, formal Vice President of the Hungarian Fuzzy Association. Professor of the Institute of Software Design and Software Development at Óbuda University, Hungary. Since 2013, she is also a full professor of the Department of Informatics at Selye János University, Slovakia. Her research interests include digital image and signal processing, uncertainty handling, machine intelligence and intelligent computing. She is author/contributor of 50 books, 79 journal papers, and more than 260 conference papers. She was founding chair of the IEEE-WISP and IEEE-SOFA symposium series, general chair of Interacademia'2004, 2008, 2012, and 2019, IEEE MeMeA'2012, IEEE CIVEMSA'2016, IPC chair of SAMI'2004, SAMI'2005, SAMI'2006, IEEE-ICCA'2005, ICONS'2008, I2MTC'2010 and 2012, and MeMeA'2021.

Csaba Makó is a Professor at the Ludovika University of Public Service Institute of Information Society. His research interest is international comparative analysis of the social – economic institutions/regulations of the game-changing technological (digitisation, smart automatisation, artificial intelligence and algorithmic management) and organisational innovations. His special focus is related to quality of work, skills formation and employment status. His latest major project is 'Democracy at Work Through Transparent and Inclusive Algorithmic Management' (INCODING, EU, 2021–2024).

Studies •