Received May 17, 2022, accepted June 3, 2022, date of publication June 9, 2022, date of current version June 14, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3181718

When Smart Cities Get Smarter via Machine Learning: An In-Depth Literature Review

SHAHAB S. BAND[1], SINA ARDABILI[2], MEHDI SOOKHAK[3], ANTHONY THEODORE CHRONOPOULOS[4,5], SAID ELNAFFAR[6], MASSOUD MOSLEHPOUR[7,8], MAKO CSABA[9], BERNAT TOROK[9], HAO-TING PAI[10], AND AMIR MOSAVI[11,12]

[1]Future Technology Research Center, College of Future, National Yunlin University of Science and Technology, Douliou 64002, Taiwan
[2]Department of Informatics, J. Selye University, 94505 Komárom, Slovakia
[3]Department of Computer Science, Texas A&M University at Corpus Christi, Corpus Christi, TX 78412, USA
[4]Department of Computer Science, The University of Texas at San Antonio, San Antonio, TX 78249, USA
[5]Department Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece
[6]Faculty of Engineering, Applied Science and Technology, Canadian University Dubai, Dubai, United Arab Emirates
[7]Department of Business Administration, College of Management, Asia University, Taichung 413, Taiwan
[8]Department of Management, California State University, San Bernardino, CA 92407, USA
[9]Institute of the Information Society, University of Public Service, 1083 Budapest, Hungary
[10]International Graduate Institute of Artificial Intelligence, National Yunlin University of Science and Technology, Douliou 64002, Taiwan
[11]John von Neumann Faculty of Informatics, Óbuda University, 1034 Budapest, Hungary
[12]Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology in Bratislava, 81107 Bratislava, Slovakia

Corresponding authors: Amir Mosavi (amir.mosavi@uni-obuda.hu), Massoud Moslehpour (writetodrm@gmail.com), and Hao-Ting Pai (htpai@yuntech.edu.tw)

This research is supported by the European Union's Horizon 2020 Research and Innovation Programme under the Programme SASPRO 2 COFUND Marie Sklodowska-Curie under Grant 945478. The authors would like to acknowledge the contribution of the Slovak Research and Development Agency under the project APVV-20-0261.

ABSTRACT The manuscript represents a comeprehensive and systematic literature review on the machine learning methods in the emerging applications of the smart cities. Application domains include the essential aspects of the smart cities including the energy, healthcare, transportation, security, and pollution. The research methodology presents a state-of-the-art taxonomy, evaluation and model performance where the ML algorithms are classified into one of the following four categories: decision trees, support vector machines, artificial neural networks, and advanced machine learning methods, i.e., hybrid methods, ensembles, and Deep Learning. The study found that the hybrid models and ensembles have better performance since they exhibit both a high accuracy and low overall cost. On the other hand, the deep learning (DL) techniques had a higher accuracy than the hybrid models and ensembles, but they demanded relatively higher computation power. Moreover, all these advanced ML methods had a slower processing speed than the single methods. Likewise, the support vector machine (SVM) and decision tree (DT) generally outperformed the artificial neural network (ANN) for accuracy and other metrics. However, since the difference was negligible, it can be concluded that using either of them is appropriate.

INDEX TERMS Smart city, big data, machine learning, ensemble, artificial intelligence, deep learning, data science, smart grid.

NOMENCLATURE

ABC	Artificial Bee Colony.
ACO-RR	Ant Colony Optimization Ridge Regression.
AE	AutoEncoder.
ANFIS	Adaptive Neuro-Fuzzy Inference System.
ANN	Artificial Neural Network.
ARIMA	Autoregressive Integrated Moving Average.
BN	Bayesian Network.
CART	Classification And Regression Tree.
CNN	Convolutional Neural Network.
DBN	Deep Belief Network.
DT	Decision Tree.
DL	Deep Learning.
ELM	Extreme Learning Machine.
IoT	Internet of Things.
KELM	Kernel Extreme Learning Machine.
MSE	Mean Square Error.

The associate editor coordinating the review of this manuscript and approving it for publication was Seifedine Kadry.

MAPE	Mean Absolute Percentage Error.
RAE	Relative Absolute Error.
LR	Logistic Regression.
LSTM	Long Short-Term Memory.
MAD	Median Absolute Deviation.
MAE	Mean Absolute Error.
ML	Machine Learning.
MLP	Multi-Layer Perceptron.
MLR	Multi-Linear Regression.
MSE	Mean Squared Error.
EFK	Extended Kalman Filter.
FFNN	Feed-Forward Neural Network.
GA	Genetic Algorithm.
GBRT	Gradient Boosting Regression Tree.
GI	Gini Index.
SA-PSO	Self-Adaptive Particle Swarm Optimization.
GPS	Global Positioning System.
GT	Game Theory.
GWO	Grey Wolf Optimizer.
ICA	Independent Component Analysis.
IG	Information Gain.
INS	Inertial Navigation Systems.
NARX	Nonlinear AutoRegressive eXogenous.
NRMSE	Normalized Root Mean Square Error.
PSO	Particle Swarm Optimization.
RBM	Restricted Boltzmann Machines.
RF(R)	Random Forest (Regressor).
GP	Genetic Programming.
SLR	Simple Linear Regression.
SMO	Sequential Minimal Optimization.
SVM	Support Vector Machine.
RMSE	Root Mean Square Error.
VAF	Variance Account Factor.
WT	Wavelet Transform.

I. INTRODUCTION

Due to abundance of resources, facilities, and welfare, half of the world's population lives in cities [1]. The lack of a specific definition of what makes a smart city smart [2] led to many cities around the world tagging themselves smart [3]. In this paper, we define the smart city as a city that utilizes the various Information and Communication Technologies (ICT) [4], [5] to improve the lives of its citizens, to solve problems (e.g., pollution, traffic, crime, etc.) [6], and to preserve its natural resources [7]. Conceptually, smart cities might be the answer to goals such as improving living standards, provisioning more services and facilities, and attaining social sustainability [4], [6], [8], [9]. Consequently, numerous technologies, such as the Internet of Things (IoT) [7], Big Data, and Cloud Computing technologies [10] have been among the tools used to support smart cities and the goals behind constructing them [11]. IoT-based devices help to optimize the decisions to enhance the performance of the city services to citizens [12]. However, adopting IoTs in smart cities can have its toll on lifestyles and undesirable impacts such as the increase in energy consumption [13], and the increase in the pollution levels in the air, soil, and water resources [14]. As a result, several studies emerged to mitigate such cons. For example, Ghahramani *et al.* evaluated an intelligent technique for routing recommendations in an IoT-based waste management complex [15]. Ghahramani *et al.* provided a unified topic modeling technique to disclose urban green space characteristics using artificial intelligence (AI) techniques [16]. Alsamhi *et al.* [17], in a study, proposed Green IoT as an environmentally friendly solution for the future use of IoTs. Almalki *et al.* [18] also presented a low-cost platform to monitor environmental parameters by employing flying IoT of real-time applications. Figure 1 portrays a future smart city which shows that the concept of smart cities integrates with all aspects of human life with a variety of ICT technologies. The adoption of ICT inevitably induces a big volume of data that a machine can learn and discover latent patterns from them. Big data is another technology that will help us analyze smart cities' data efficiently and at a higher degree of scalability [19]. Figure 1. represents a holistic representation of smart cities application [20].

FIGURE 1. An outlook on smart cities.

The literature is rich with studies investigating the role of AI and ML-based techniques in smart city applications. For example, Ullah *et al.* [19] reviewed recent trends in the

TABLE 1. **The description of the conducted survey studies.**

Ref.	Highlights	Database information	Systematic review	Statistical analysis	Probable gap
[19]	Recent trends in the application of artificial intelligence techniques in smart cities	N.A.	☒	☒	Database information and subject review interval
[21]	A survey on the applications of data mining and single ML techniques against complex datasets	The most cited methods and datasets	☒	☑	Evaluation interval
[22]	The use of ML and DL techniques in smart cities for prediction, planning, and uncertainty analysis	Database from web of science (WoS) and Scopus	☒	☑	Subject review interval
[23]	IoT-based ML techniques in healthcare, smart grids, and vehicular communications	N.A.	☒	☒	Database information and subject review interval
[23]	ML and internet-of-thing-(IoT) techniques used in healthcare, and smart grids	N.A.	☒	☑	Database information
[24]	ML data mining techniques in smart city applications	Database from web of science (WoS) and Scopus	☑	☒	Evaluation and comparison interval
[25]	The relationship between AI and smart cities	N.A.	☒	☒	Database information and subject review interval
[26]	Big data in smart city applications from an ML techniques point of view	N.A.	☒	☒	Database information and evaluation interval
[27]	Qualitative analysis of DL-based techniques for smart city applications	N.A.	☒	☒	Database information and evaluation interval
[28]	Deep reinforcement learning and clustering in smart city applications	N.A.	☒	☒	Database information and evaluation interval

application of AI techniques in smart cities but limited their analysis to ML and reinforcement learning and a selected set of applications (i.e., transportation, cyber-security, smart grids, unmanned air vehicles, and healthcare). The study lacks a comparison between the performance of the different ML techniques. Shafiq *et al.* [21] presented a survey on the applications of data mining and single ML techniques to have sustainable smart cities. The study discussed the performance of these techniques against complex datasets. Nosratabadi *et al.* [22] reviewed the use of ML and deep learning techniques in smart cities for prediction, planning, and uncertainty analysis. Din *et al.* [23] studied IoT-based ML techniques in some aspects of smart cities such as healthcare, smart grids, and vehicular communications. Similarly, Din *et al.* [23] surveyed single ML and internet-of-thing-(IoT) techniques used in healthcare, smart grids, and vehicular communications. Souza *et al.* [24] surveyed ML data mining techniques and their role in smart city applications using the arrangement method [24] and the e VOS viewer [24]. Its aim and purpose were focused on statistical perspectives, not comparing performance or recommending certain techniques for smart city applications.

Batty [25] discussed the relationship between AI and smart cities and proposed ML techniques for real-time city functions. Mohammadi and Al-Fuqaha [26] shed light on the challenge of big data in smart city applications from a machine learning point of view. The study focused on deep reinforcement learning and how it was used to handle the cognitive aspect of smart city services. Bhattacharya *et al.* [27] developed a qualitative study for discussing the future of DL-based techniques for smart city applications. Kolomvatsos and Anagnostopoulos [28] studied the application of deep reinforcement learning and clustering for query controller application in smart cities as a comparative analysis. Table 1 presents the study's strengths and weaknesses to generate the central research gap. This table compares the conducted studies with the criteria of the present study. Despite the abundance of the conducted studies, they still have shortcomings and limitations that warrant further investigation and study. Specifically, they do not provide a classification for the ML and DL techniques used or do not categorize their roles and functionality in smart cities. In addition, researchers in the field may be challenged by the scarcity of reviews that contrast the performance of ML techniques and analyze their suitability to solve different problems. Currently the literature lacks a comprehensive review that categorizes ML algorithms and their applications to smart cities. Such a study would guide researchers in the field of smart cities to use the right tool for a given problem. Managing a significant amount of data in review articles can help in the successful implementation of smart cities for future planning and policy-making [29]. Our analysis in this study bridges the gap by providing a taxonomy of the ML algorithms and their contributions to improving smart cities. Furthermore, we provide a quantitative analysis of the performance of these ML algorithms to select the most likely effective one in a given field. We evaluate these algorithms concerning efficiency, accuracy, and computational complexity. Our contribution in this paper aims to introduce a novel taxonomy that focuses on the type of ML algorithms and approaches rather than the type of applications in smart cities. The proposed taxonomy may help researchers, policy makers, and practitioners to enhance the living standards in smart cities by leveraging the right ML tools. The rest of the manuscript is organized as follows. Section II explains the methodology we used to carry out this literature review. Section III surveys the literature, describes the role of state-of-the-art ML algorithms in solving problems in smart cities and presents the taxonomy of the AI and Ml-based techniques for application in smart city concepts.

Section IV evaluates the surveyed algorithms by comparing their performance results throughout applications. An evaluation of the ML methods and discussion are presented in Section V. In Section VI, we highlight some of the open issues and challenges, and in Section VII we conclude the review.

II. METHODOLOGY

It is challenging to search and identify all studies in which ML algorithms have supported smart cities due to the abundance of such algorithms and their variations. The simple search queries for "smart city" and "machine learning" may not provide a comprehensive list of relevant literature. The search phrase "smart city" is not the only one that we would solely bank on because other search phrases that bear close semantics, such as "intelligent city," "smart urban planning," "smart urban mobility," etc., should not be neglected. The complexity notably increases when we compound the query with the names of many ML algorithms. We relied on the main algorithms discussed in textbooks and in surveys such as [30] for the names of the ML algorithms. In this research, the Scopus database has been used as the primary repository as it indexes the major authenticated publishers.

Our review ultimately aims to identify, organize, and classify the ML techniques that have been used to serve smart cities into one of the four architecture categories: single models, hybrid models, ensemble models, and DL. Figure 2 depicts our review methodology which consists of four stages. In the first stage, an initial set of relevant articles is identified based on the search queries: "smart city" and "machine learning methods". For each ML method, we applied a new search query taking into consideration the specifics of each ML method and its variations. In the second and third stages of the review methodology, we analyzed and classified the ML algorithms based on how each algorithm was applied in smart cities, the datasets used, and the results attained. Finally, in the fourth stage, the ML models are classified into the four aforementioned categories. Overall, our search has generated more than 430 relevant documents. During the second stage, we have carefully analyzed these documents to discern the most relevant ones (i.e., those belonging to the fields depicted in Figure 2) and thus we narrowed the search pool down to 100 relevant papers. In the third stage, the papers pool was further refined so that we ended up with 80 core papers to review. There was a considerable increase in the number of articles that used ML methods over the last ten years (2010 to 2020) (Figure 3).

A. LIMITATION

Research work on smart cities dates back to 2010. However, the research has progressed very rapidly in terms of the number of papers published after 2016. Additionally, the popularity of ML applications in smart city technologies has also been recognized since 2016 with significant growth of publications in the last years as shown in Figure 3. Consequently, and for the sake of staying current and relevant, the focus of this survey has been confined to papers published in 2016 or after.

III. SMART CITIES AND MACHINE LEARNING

The concept of the smart city has been used in literature since the early 90s [31]. However, the term "smart city/cities" has

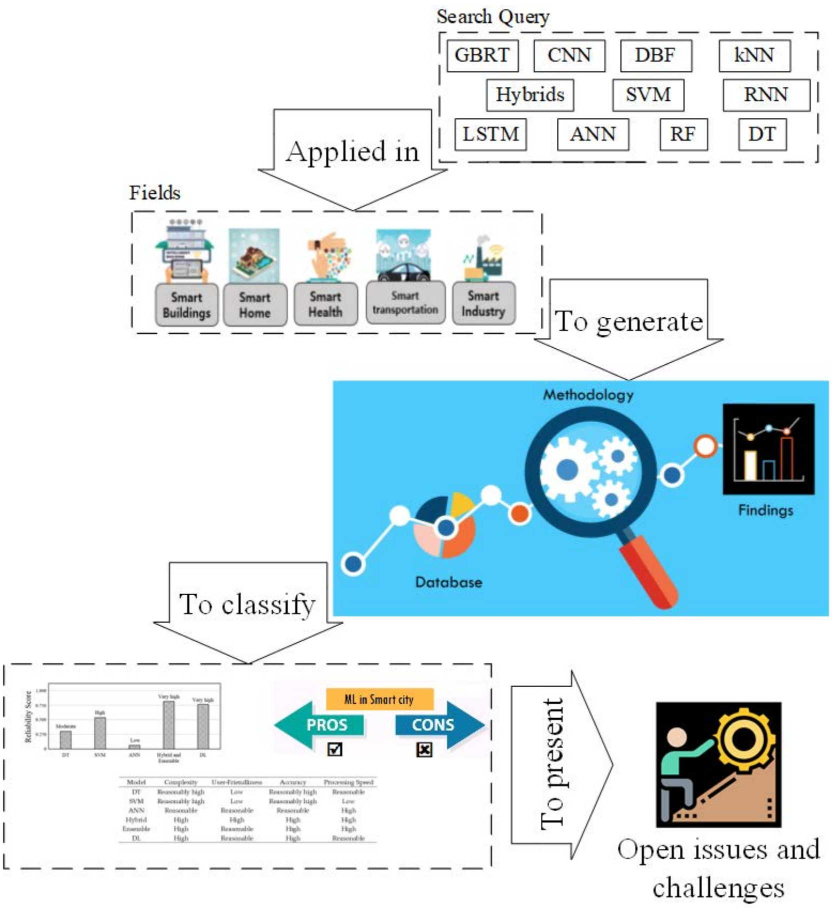

FIGURE 2. The methodology of our review.

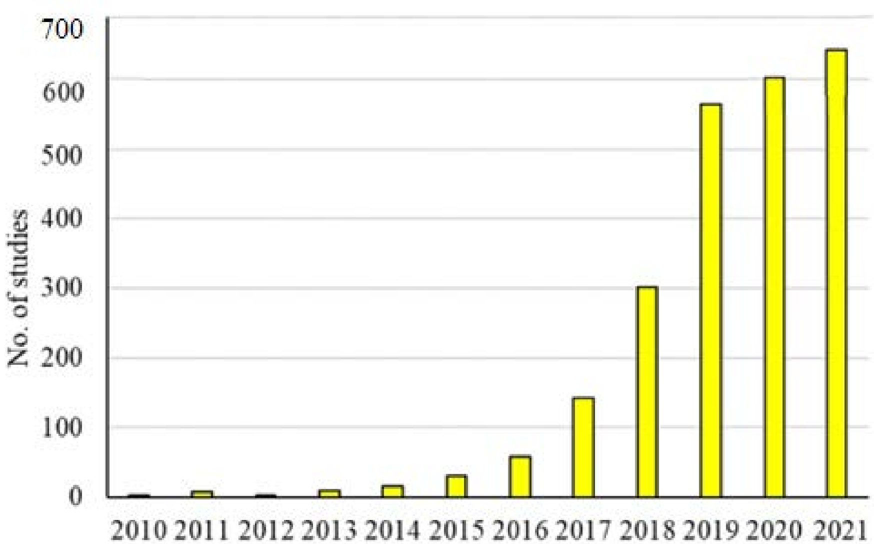

FIGURE 3. The number of studies that used ML in smart cities has doubled annually between 2011 and 2021.

been used only in a limited number of articles until 2011, when the concept started to be widely popular. Additionally, the importance of ML methods has exponentially grown over the past few years (see Figure 3). In reverse chronological

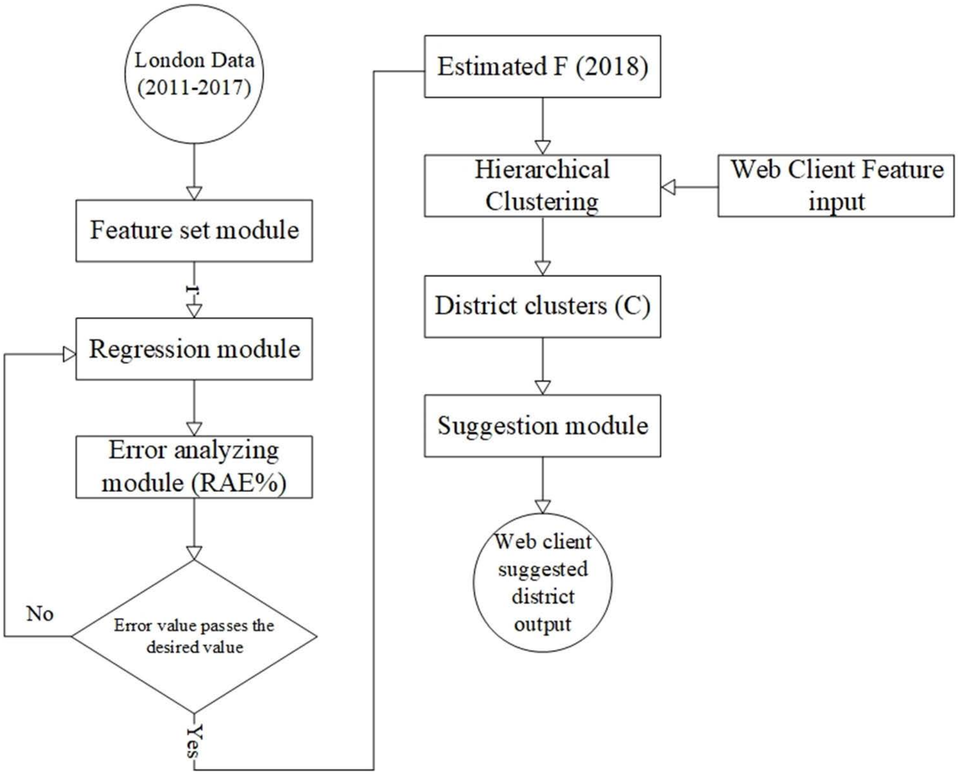

FIGURE 4. An algorithm to estimate business locations in smart cities.

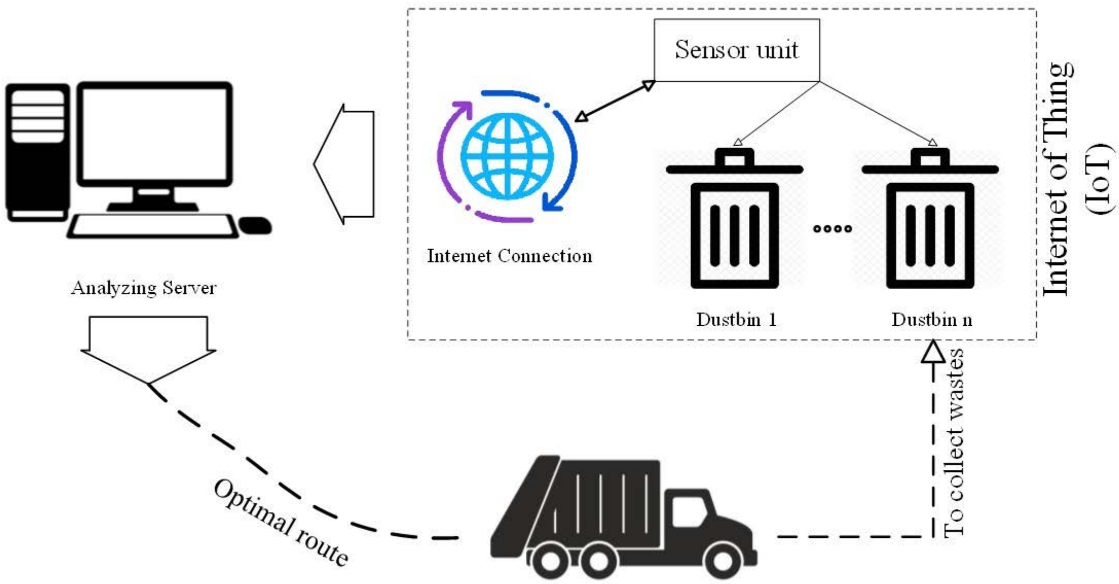

FIGURE 5. IoT trash collection mechanism.

order, Table 2 provides a summary of the most important studies in which ML algorithms were used in smart cities. Next, we discuss these studies in more detail.

Elsaeidy *et al.* [32] used Restricted Boltzmann Machines (RBM) as the ML technique to detect distributed denial of service attacks in smart cities. The use of RBM

TABLE 2. Notable ML methods used in smart-city studies.

References	Year	Description of Machine Learning Approach	Smart city applications
[42]	2021	ML methods for IoT applications in smart cities, smart homes, and smart healthcare	Prediction and clustering for enhancing the IoT applications
[43]	2021	ML-based techniques for the prediction of Energy consumption in smart cities	Estimation of the Energy consumption
[44]	2021	ML-based techniques for handling the nanogenerators towards smart cities	Trend recognition
[45]	2020	ML-based techniques for enhancing the unmanned aerial vehicles efficiency	Process optimization
[46]	2020	Supervised and unsupervised ML-based techniques for handling electric vehicles in a smart city	Charging behavior analysis
[32]	2019	RBM technique for handling Distributed Denial of Service attacks related to smart cities	Detection of distributed denial of service
[33]	2019	RF algorithm used in IoT-based systems for detecting compromised IoT devices	Intelligent anomaly detection
[34]	2019	RF-based approach for estimating global solar radiation in comparison with other ML techniques	Estimation of global solar radiation
[35]	2018	MLP and MLR techniques for detecting and estimating the location of a business	Estimation of business location
[36]	2018	ML based-IoT system for waste management as a case study in smart cities	Waste management
[37]	2018	ML techniques for detecting the criminal patterns in the presence of historical data	Detection of criminal patterns
[38]	2018	IoT-based SVM for classification of vehicular traffic in smart cities as a case study	Classification of vehicular traffic
[39]	2018	RF, k-NN, and Bagging ML techniques for forecasting air pollution in smart cities	Forecasting air pollution
[40]	2018	ML techniques in combination with IoT devices for prediction of air pollution in smart cities	Forecasting air pollution
[41]	2017	ML-based IoT system used to develop personalized services by leveraging weather data	Developing personalized services

was justified by the high number of features in the datasets. Evaluation results showed that the approach can cope with the attack-detection task as it showed high accuracy and reliability scores. Alrashdi *et al.* [33] used the IoT-based Random Forest (RF) technique to intelligently detect anomalies in a smart city. In comparison with other several techniques, the authors found that RF gives the most reliable and accurate results for detecting compromised IoT-based systems at distributed fog nodes. Similarly, Meenal and Selvakumar [34] found that the RF technique is promising for the detection of global solar radiation when compared with other ML techniques. Bilen *et al.* [35] tackled the problem of estimating business locations in smart cities using the Multi-Layer Perceptron (MLP) and Multi-Linear Regression (MLR) techniques. They justified the use of these techniques to the large number of features involved and to the need for high accuracy. The method was to import London data to the main algorithm for a Feature set module. The next step was to develop the regression module followed by error analysis using Relative Absolute Error (%). If the calculated error value (%) is higher than the desired value, the algorithm returns to the regression module to perform the modeling operation again. Then, the estimated value was imported by the clustering module in parallel with the Web Client Feature Input for Hierarchical Clustering. This module generates district clusters as the Web client suggested district. Figure 4 presents the related algorithm reproduced from Bilen *et al.* [35].

Bakhshi and Ahmad [36] combined the ML algorithms with the IoT techniques to manage waste in smart cities. Figure 5 presents the algorithm of the mechanism. According to the mechanism represented in the figure reproduced from [36], there is an IoT-based unit for collecting the information of all the dustbins. This system monitors whether the Dustbin_i is full or empty by using sensors implemented. Then transfers the information to the analyzing server by using an internet connection. This unit forms the IoT unit. The next step is to generate the optimal route for the garbage truck to collect the wastes. This is a brief description of the mechanism. To increase safety and security in smart cities, Lourenço *et al.* [37] used ML techniques to detect criminal patterns based on historical data to increase safety and security in smart cities. The main mechanism can be found in Figure 6 which is reproduced from [37]. The Citizen from the client side communicates with the Data Center Module. Data were imported by the Sci-Cumulus workflow engine on the server-side. The existing ML-based technique in this unit employs external sources and communicates with the knowledge base unit. The next step is to export the information to the analytical module. This module, as a decision-making system, communicates with the police. The ML techniques showed promising prediction results in comparison with other non-ML tools. Reid *et al.* [38] focused on one of the crucial issues in smart cities, namely traffic jams. The authors found the Support Vector Machine (SVM) showing high accuracy for classifying vehicular traffic in their attempt to mitigate air and noise pollution and optimize fuel consumption. Martínez-España *et al.* [39] experimented with RF and compared it with k-NN and Bagging ML techniques for forecasting air pollution in smart cities. Results, evaluated using the RMSE and correlation coefficient values, showed that RF provides the highest accuracy among the considered ML techniques. In another study by Chung and Jeng [40], ML techniques were also used for the prediction of air pollution and to determine the factors that affect air quality. In another weather-related problem, Chin *et al.* [41] developed a proper personalized service using an ML-based IoT system that correlates weather data (i.e. rainfall and temperature) with short journeys made by cyclists. Alsamhi *et al.* [42] provided a classification of ML-based techniques for enhancing the applications of IoT-based technologies in a smart city. Carrera *et al.* [43] employed a meta-XGBoost model integrated with meta-regression to generate energy data to enhance the prediction accuracy of the energy production. Alagumalai *et al.* [44] also used ML-based techniques to assess the trends of using nano generators in smart cities. Ullah *et al.* [45] analyzed the different applications of ML-based techniques employed for enhancing unmanned aerial vehicles' efficiency. Shahriar *et al.* [46] discussed supervised and unsupervised ML-based techniques for handling electric vehicles in a smart city. By analyzing the above studies, we noticed that two main motives compelled the use of ML techniques in smart cities. First, most of the tackled problems have high dimensionality datasets (the number of features is big). Second, accuracy and reliability were a priority in most of the studies to have a sustainable ecosystem in smart cities. Next, we briefly describe each ML technique used in smart cities.

A. MACHINE LEARNING ALGORITHMS

This section describes the commonly used ML algorithms in the smart-city-related literature, such as Decision Trees (Sec. 4.1), Support Vector Machines (Sec. 4.2), and Artificial Neural Networks (Sec. 4.3). Finally, Section 4.4 describes advanced ML approaches for smart-city applications based on the hybrid ML techniques, ensembles of ML algorithms, and the Deep Learning paradigm.

1) DECISION TREES

The first regression tree was initially proposed and implemented in 1963 by Morgan and Sonquist [47]. Then, the first work on Decision Trees (DTs) was published in 1966 by Hunt in the psychology field [48].

The DT algorithm is a supervised learning method [49] that can be employed for classification and regression tasks. More specifically, a DT leverages a tree-based data structure in which the samples are recursively partitioned based on the selected feature whose values most effectively split to maximize a purity measure [51, 52]. Figure 4 presents a simple DT algorithm schematic diagram with leaf nodes as attributes.

As seen in Figure 7, a DT algorithm collects the outcome of each node and decides the final results reproduced from [55]. Attribute selection is one of the most important challenges in constructing a DT. The value of attributes is measured by two functions: Information Gain (IG) and Gini Index (GI) [52]–[54]. IG computes the entropy changes in the whole mechanism of DT based on Eq. 1:

$$Gain\,(S.A) = Entropy\ (S) - \sum\nolimits_{v\epsilon Values\,(A)} \frac{|S_v|}{|S|} \cdot Entropy(S_v) \quad (1)$$

where S, A, and S_V, define the set of instances, attributes, and instances in the V^{th} attribute, respectively, whereas the entropy characterizes the impurity of an arbitrary collection.

On the other hand, the GI determines the frequency of incorrect identification for a randomly chosen element, which leads to favoring an attribute with a lower GI. Eq. 2 shows GI's formula:

$$GI = 1 - \sum_j p_j^2 \quad (2)$$

where p refers to the probability of the event occurring.

In addition to a single DT, the RF approach is constructed by considering an ensemble of multiple DTs, constituting a "forest" of simpler estimators. Each tree is built on different portions of the training set to minimize the error between the predictions and the actual values. Figure 4 presents a simple flowchart for a decision-making purpose in smart city applications by DT. Table 3 presents the most important studies in smart cities that leveraged the DT-based techniques:

Connected vehicles in a smart city are a hot topic due to their security [69] and control aspects (e.g., platooning) [70]. DT was used in [56] to estimate the traffic classification in comparison with other ML-based techniques. Results have been evaluated using the accuracy metric (99.18% for DT). In [57], DT was used for pandemic prediction and compared to other ML-based techniques. According to the findings,

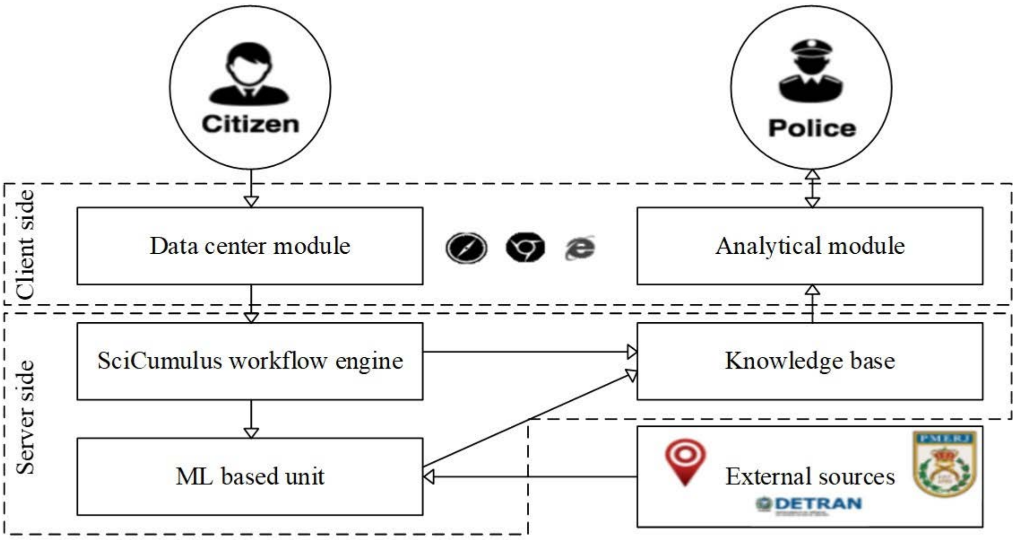

FIGURE 6. Using ML techniques to detect criminal patterns.

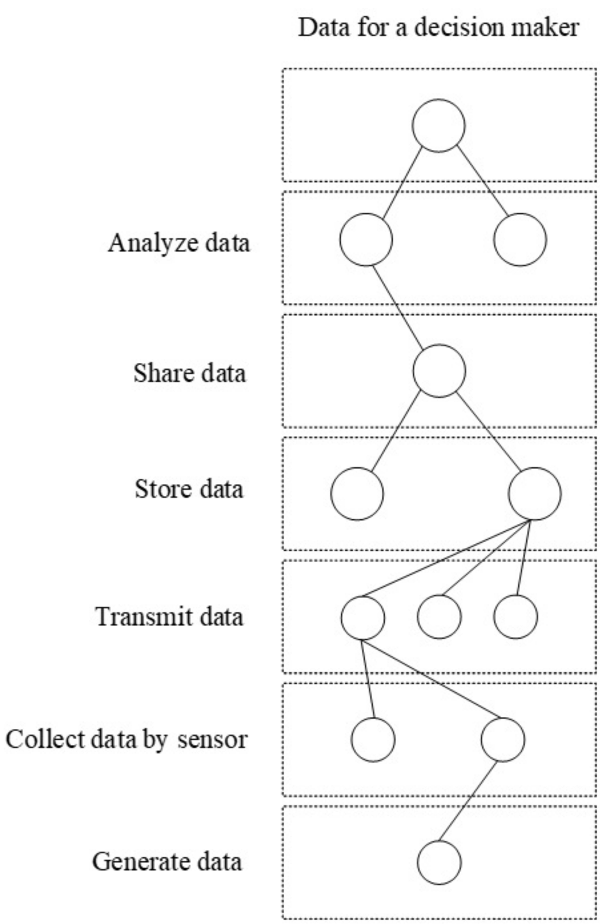

FIGURE 7. Decision trees are among the ML techniques used in smart cities.

DT provided a good accuracy (about 99%) for the estimation task. Balta *et al.* [58] employed DT integrated with a fuzzy approach for the optimization of the traffic signals in a smart city. Accordingly, nearly 15% to 17% performance improvement was obtained using the proposed technique.

The study of Aloqaily *et al.* [60] deemed transportation one of the important fields in smart cities when they investigated how to detect connected vehicles using Deep Belief Networks (DBN) and DT. The performance of the proposed technique was evaluated using accuracy and detection rate.

In the telecommunication field, Manzanilla-Salazar *et al.* [61] detected failures in the LTE infrastructures using the DT and SVM techniques and compared them. Early detection of failures in the LTE infrastructure can be a big cost saver. The study showed that the proposed DT technique can increase the accuracy and detection rate of failures. To protect smart cities from cyber-security, Alrashdi *et al.* [33] developed a system for the detection of attack points using an IoT-based RF whose accuracy reached 99.34% on real datasets.

Solar radiation is one of the vital issues in smart cities that captured the attention of Meenal and Selvakumar [34] when they found that the RT technique outperformed other ML techniques when tried on empirical data collected in Tamil Nadu. Furthermore, in the field of electrical cards, the RF technique showed another success in smart cities when the authors in [43] tried to predict the charging demands of electric vehicles. Detecting and locating road anomalies is a significant aspect of smart cities. To that end, El-Wakeel *et al.* [63] used the DT algorithm with great success. Education and predicting student performance were the focus of Gomede *et al.* [64], who relied on an RF-based technique to do so.

Orlowski *et al.* [65] presented a DT-based IoT model for increasing the performance of building business models. The paper discussed sustainable decision-making processes

TABLE 3. DT-based studies in smart cities.

References	Year	Description	Method	Findings	Application	Pros. and Cons.
[56]	2021	DT in comparison with other ML-based techniques to estimate network traffic classification	DT	DT provided the highest accuracy (about 99.18%) in comparison with other ML-based techniques	Traffic classification	High accuracy in classification
[57]	2021	DT in comparison with other ML-based techniques for pandemic prediction in a smart city	DT	DT provided the highest accuracy (about 99.20%) in comparison with kNN and LR	Pandemics prediction	Fast detection and rapid screening
[58]	2020	DT integrated by fuzzy for the traffic signaling in a smart city	DT-fuzzy	The proposed method improved the situation by about 15 to 17%	optimization of the traffic signal	Complexity of algorithm
[59]	2020	DT for predicting warming climate-induced energy stress in a smart city	DT	About 80 % lower cooling energy stress can be achieved	Energy stress mitigation	Better classification capability
[60]	2019	DT and DBN for detection of connected vehicles	DT + DBN	The proposed technique could increase the accuracy	Vehicles' connections	Better for complex tasks but is not easy to use
[61]	2019	DT is compared with linear SVM for failure detection in LTE infrastructures	DT and Linear SVM	DT increased accuracy score when compared to SVM	LTE-related failure detection	DT provides better classification
[33]	2019	IoT-based RF technique for detection of cyber-attack in smart cities	RF	High accuracy (up to 99.34%) in cyber-attack detection	Cyber-attack detection	Better detection ability
[34]	2019	RF technique compared with other ML techniques for the prediction of solar radiation	RF	RF provided the highest prediction performance	Solar radiation prediction	Complex algorithm
[62]	2018	RF technique for predicting the charging demand of electrical vehicles	RF	The proposed system effectively optimizes the charging demand	Sharing electrical vehicles	Complex algorithm
[63]	2018	DT employed for detection and localization of the road surface anomalies	DT	The proposed technique can successfully detect road anomalies	Classification of road surface condition	Better detection ability
[64]	2018	RF-based classification for the development of knowledge profile of students and related decision making support	RF	The proposed model is suitable for the prediction of student performance	Improving educational system	Reasonable results
[65]	2018	DT-based IoT technique for increasing the performance of building business models	DT	The proposed method is proven to be effective in building business models	Business models for IoT nodes measuring air quality	Better detection ability

TABLE 3. *(Continued.)* **DT-based studies in smart cities.**

[66]	2017	DT and GT-based techniques for handling travel information for passengers in smart cities	DT + GT	The proposed method provided acceptable performance	Handling the travelers' process	Better for the complex tasks
[67]	2017	RF for prediction of air pollution in smart cities	RF	RF provided high prediction accuracy ranging from 70% to 90%	Prediction of air pollution	Other ML-based can be better for prediction
[68]	2017	DT for classification of CCTV cameras based on their applications and locations in smart cities	DT	The proposed technique provided 87.96% accuracy	Classification of CCTV cameras	Reasonable but not good in classification

in smart cities and highlighted the importance of DTs and business models for making decisions. Validation was performed via a case study on air quality. Similarly, in [66], Mei *et al.* proposed a Rule-based Incentive Framework utilizing a DT along with a Game Theory (GT)-based technique that was evaluated in terms of decision-making accuracy for handling traveling information of passengers in a smart city. Simulations showed that the proposal constitutes an effective way to incentivize travelers to change travel routes, proving to be an essential smart city service.

Air pollution in smart cities was also investigated by Benedict [67] who built a prediction framework based on RF for estimating air pollution, which is considered one of the most urgent challenges in smart cities. Using real validation data, the accuracy ranged between 70% and 90%. Pribadi *et al.* [68] developed a DT-based decision-making mechanism for handling CCTV cameras in smart cities. The performance evaluation showed an accuracy of 87.96%.

2) SUPPORT VECTOR MACHINES

Support Vector Machine (SVM) algorithm was firstly developed by Vladimir *et al.* in 1963 [71], whereas Boser *et al.* provided an approach to employ non-linear classifiers using the kernel trick in 1992 [72]. SVM is one of the most frequently used supervised ML algorithms and employs the related learned model to handle both classification and regression tasks. In detail, the SVM represents the training samples as points in the feature space to find a set of hyperplanes that provide the best class separation, whereas new points are classified or predicted according to the portion of space they belong.

The input-output formulation of an SVM is formally described by $f(x)$ given in Eq. 3:

$$f(x) = w^T \varphi(x) + b \tag{3}$$

where w^T denotes the transposed vector related to the output layer, $\varphi(\mathrm{x})$ represents the kernel function, and b the bias.

Overall, the matrix $\boldsymbol{X}$ has $N \times n$ dimensions in which n and N refer to the number of input parameters and data points, respectively. The following cost function is optimized to evaluate w^T and b parameters [73]:

$$\textit{cost function} = \frac{1}{2} w^T + C \sum_{k=1}^{N} \left(\xi_k - \xi_k^*\right) \tag{4}$$

which is constrained by Eq. 5:

$$\begin{cases} y_k - w^T \varphi(x_k) - b \le \varepsilon + \xi_k, & \text{for } k = 1, 2, \ldots, N \\ w^T \varphi(x_k) + b - y_k \le \varepsilon + \xi_k^*, & \text{for } k = 1, 2, \ldots, N \\ \xi_k . \xi_k^* \ge 0 \end{cases} \tag{5}$$

in which X_k and Y_k are the k^{th} input and output, respectively, whereas ε represents the fixed precision of the estimation; the slack variables (ξ_k, ξ_k^*) are also in charge to determine the acceptable error margin.

The following Lagrangian optimization is applied to minimize the cost function:

$$\begin{aligned} L\left(a, a^*\right) = &-\frac{1}{2} \sum_{k.l=1}^{N} \left(a_k - a_k^*\right)\left(a_l - a_l^*\right) K\left(x_k . x_l\right) \\ &- \varepsilon \sum_{k=1}^{N} \left(a_k - a_k^*\right) + \sum_{k=1}^{N} y_k \left(a_k - a_k^*\right) \end{aligned} \tag{6}$$

$$\sum_{k=1}^{N} \left(a_k - a_k^*\right) = 0 \; a_k . a_k^* \in [0.c] \tag{7}$$

$$K\left(x_k . x_l\right) = \varphi\left(x_k\right)^T \varphi\left(x_l\right), \quad \text{for } k = 1, 2, \ldots, N \tag{8}$$

where a_k, a_k^* are the Lagrangian multipliers. In the last step, the $f(x)$ of the SVM is given as follows:

$$f(x) = \sum_{k.l=1}^{N} \left(a_k - a_k^*\right) K\left(x . x_k\right) + b \tag{9}$$

Table 4 presents notable studies that employed SVM-based techniques in smart cities for different purposes. References are ordered by year in descending order. More details on each are given hereinafter.

Recently, Manogaran *et al.* [74] adopted an approach that integrates the SVM with the shared Adaptive Computing Model for a traffic management system that provided an improved platform by increasing the decision reliability and reducing the computing time compared to the SVM alone. In [75], the SVM was compared to other ML-based

TABLE 4. Some of the researchers that used SVM techniques in smart cities.

References	Year	Description	Method	Findings	Application	Pros. and Cons.
[74]	2021	SVM integrated by shared Adaptive Computing Model for traffic management system	SVM	The integrated model improved the decision reliability and computing time in comparison with the SVM	Smart connected vehicles	Higher complexity wait time
[75]	2020	SVM in comparison with other ML-based techniques for cyber-attack detection in a smart city	SVM	Ensemble-based classifiers showed better performance	cybersecurity	Low-cost detection
[76]	2020	To analyze the priority data in a smart city-transportation sector using the SVM	SVM	SVM provided a proper and acceptable performance score	Smart transportation	Proper classification score
[77]	2019	SVM-based model for securing blockchain IoT data in smart cities	SVM	The proposed technique provided high efficiency and confidentiality	Secure and reliable IoT data sharing platform	Higher efficiency
[78]	2019	SVM is employed for energy management of electrical vehicles in smart cities	SVM	SVM successfully improved the system efficiency and robustness	Power management of electrical vehicles	Higher efficiency
[79]	2019	SVM is used for the classification and detection of violence in social media	SVM	SVM showed up to 97% accuracy	Classification and detection tasks in social media	Reasonable classification score
[80]	2019	SVM compared with ML and hybrid techniques for estimating the heating load of buildings in smart cities	SVM	SVM provided a moderate accuracy	Estimation of building heating load	Reasonable estimation score
[81]	2018	Hybrid GA-SVM method for optimization of smart city energy consumption	GA-SVM	The proposed method improved performance by more than 21%	Optimization of energy consumption	Requires hybrid algorithm
[82]	2018	SVM is employed for the detection of anomalies in a laboratory environment	SVM	SVM provided high accuracy and reliability despite the complex configuration	Anomaly detection	Higher complexity
[83]	2018	SVM for low-cost detection and prediction of vehicle position	SVM	The proposed method provided up to 94% accuracy in position prediction	Prediction of vehicle position	Low-cost detection
[84]	2018	SVM-based clinical decision support system for heart failure detection in smart cities	SVM	SVM successfully predicted heart failures with 76.9% sensitivity	Heart failure detection	lower detection score

techniques for cyber-attack detection in smart cities but the performance was not promising. Shen *et al.* [77] devised a secure and privacy-preserving SVM using blockchain-based encrypted IoT data. Results reported the accuracy and confidentiality of the proposed technique showing that it could successfully cope with the considered task and ensure the confidentiality of sensitive data. The SVM is leveraged also by Aymen and Mahmoudi [78] that presented a methodology for management and optimization of power status in electrical vehicles for smart cities. The evaluation used energy consumption and charge state of batteries and showed that the SVM attains high performance and

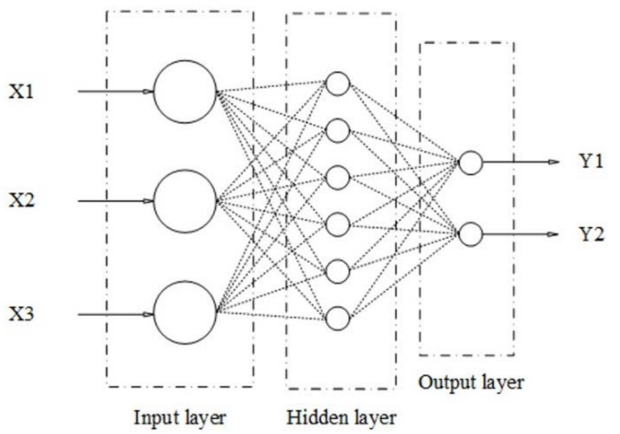

FIGURE 8. The architecture of a multi-layer perceptron.

robustness. Differently, Pujol *et al.* [79] developed an SVM-based system to detect and classify violence types in social media. This system monitors social media space and decides about observations by using a set of terms and rules. The accuracy measure of the proposed system exhibited acceptable performance between 85% and 97%. Le *et al.* [80] developed a platform for predicting and estimating building heating load in smart cities using ML methods, including SVM and RF, and a hybrid technique based on particle swarm optimization and extreme gradient boosting machine (PSO-XGBoost). Evaluations were performed using the Root Mean Square Error (RMSE) and correlation coefficient measures. Results demonstrated that the best method (i.e. SVM) generates predictions with moderate accuracy values but also emphasized the capability of hybrid techniques able to outperform single models (cf. later Sec. 4.4). Likewise, Chui *et al.* [81] presented a study aimed at the optimization of energy consumption in smart cities. The proposed method employed a Genetic Algorithm (GA) to construct a hybrid GA-SVM technique that was compared with other single ML techniques in terms of specificity, sensitivity, and accuracy. The proposed technique improved the performance by more than 21%, thanks to the presence of the GA optimizer.

Garcia-Font *et al.* [82] tested an SVM method for anomaly detection in a laboratory that reproduces a real smart city use case with heterogeneous devices, algorithms, protocols, and network configurations. Results indicated the high reliability and accuracy attained by the proposed method for anomaly detection despite possible technical difficulties in configuring and implementing ML models in such environments. Belhajem *et al.* [83] developed an estimation platform for vehicle position using SVM and Extended Kalman Filter. The dataset was gathered via the Global Position System (GPS) and Inertial Navigation System (INS). This technique is aimed at low-cost detection of vehicle position. Experimental results showed an improvement of up to 94% in position prediction in case of GPS failures compared to related baselines. In another study, Aborokbah *et al.* [84] devised and evaluated a platform for a clinical decision support system based on SVM. The latter was developed with the RBF kernel function and leveraged to detect heart failures. Performance has been evaluated using the sensitivity measure and demonstrated that the SVM could provide a sensitivity of 76.9%.

3) ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANNs) were first developed by Warren McCulloch *et al.* in 1943 [85]. This work simulated a simple neural network with electrical circuits to investigate the performance of neurons in learning tasks. The ANN is an initial and simple way to design an intelligent learning system inspired by the biological neurons that constitute brains. This system uses a training stage related to a certain task that extracts knowledge from a training dataset without the need to be programmed by task-specific rules [86]. Indeed, the basic idea of ANNs is performing tasks without any prior knowledge about the nature of phenomena. Consequently, ANNs can generate identifying characteristics (i.e., extracting discriminative features) from the data that are given as input [87]. ANNs can be considered as a comprehensive modeling framework to process complex datasets. Recently, ANNs have been employed for forecasting, regression, and curve-fitting purposes [86]. In an ANN model, neurons represent the fundamental components that employ transfer functions for generating the output values. The most important advantage of ANNs is that they are simple and cost-effective methods for handling large datasets [88]. Multilayer Perceptron (MLP) is one of the simplest and most frequently used variants of feedforward ANNs. MLP is characterized by 3-layer, or more, architecture, as shown in Figure 8 [89]. The first layer is the input layer, the intermediate layers are the hidden layers, and the last layer is the output layer [90]. An MLP can have multiple hidden layers. In that case, we refer to it as a "deep" MLP (cf. Sec 4.4).

Based on Figure 5, in a simple feedforward ANN, input layers and hidden layers are linked by the input weight matrix, whereas hidden layers and output layers are linked by the output weight matrix. Both matrices are learned during the training phase of the model. Hence, the ANN is characterized by the following equation (Eq. 10) for generating the output values [86]:

$$O = f(B + \sum_{i=1}^{n} w_i x_i) \tag{10}$$

where w is referred to as weight values that control the propagation value x from input to output with n being the number of layers, whereas O is referred to as the output value from each node to be modified by the bias B value.

Table V presents notable papers employing ANNs for smart city applications. Works are ordered by the year starting from the most recent ones. In the following, we provide details on each study.

Alsamhi *et al.* [91] developed a platform using an ANN to predict the signal strength of a drone. The independent variables were drone altitude and path loss. Results have been

TABLE 5. Notable smart-city studies that used ANN-based techniques.

References	Year	Description	Method	Findings	Application	Pros. and Cons.
[91]	2021	ANN to estimate the signal strength from a drone	ANN	The solution provided a proper agreement for the prediction of Target values	Signal estimation	Simple and reliable method
[92]	2020	ANN for the prediction and estimation of arsenic zones in the groundwater environment	ANN	ANN provided a reasonable accuracy for the estimation task	Land use pattern	Simple and easy to use
[93]	2019	Single and hybrid ANNs for forecasting building load heating	ANN GA-ANN PSO-ANN	ANN-GA had the highest determination coefficient (0.9) and the lowest RMSE (1.625)	Forecasting building heating load	The ability to integrate ANN with optimizers
[94]	2019	ANN-based smart system for detection of lighting in smart cities	ANN	ANN provided the highest accuracy (92.6%)	Lighting system detection	Simple in detection
[95]	2019	ANN-based monitoring system for drainage handling in a smart city	ANN	ANN could successfully predict urban drainage with up to 99% accuracy	Smart drainage monitoring system	High accuracy in small datasets
[96]	2019	ANN for mapping air pollution data in smart cities	ANN	ANN could successfully predict air pollution in a target prototype	Mapping air pollution	Simplicity of ANN
[97]	2017	ANN-based technique for managing the urban bus transportation paths in smart cities	ANN	ANN provided accurate estimations of arrival times	Urban bus path management	Reasonable accuracy
[98]	2018	Data-driven learning algorithms based on ANNs to investigate and evaluate their application to social welfare, fairness, and privacy	ANN	Provided guidelines depending on requirements and privacy constraints of the smart application	Social welfare, fairness, privacy	Processing time
[99]	2016	Real-time managing console for public transportation systems employing an ANN-based monitoring system	ANN	Fleet management console used by administrators as a real-time monitoring system for buses	Public transportatio n system	Real-time implementation
[100]	2017	Intelligence platform for traffic and flash flood monitoring system using wireless technology	ANN	Real-time platform with high reliability and accuracy	Traffic flow and flash flood monitoring	Real-time implementation

analyzed using a determination coefficient. Findings showed reasonable validation accuracy determination coefficient equal to 0.96 and 0.98, respectively, for varying heights and distances. Singh *et al.* [92] employed ANN to estimate the arsenic vulnerable zones with reasonable accuracy.

Le *et al.* [93] developed different types of ANNs combined with optimization techniques (e.g., GA, PSO, etc.) to estimate building heating loads in smart cities and optimize energy efficiency. Results evaluated the RMSE and determination coefficient measures. They showed that hybrid ANN-GA

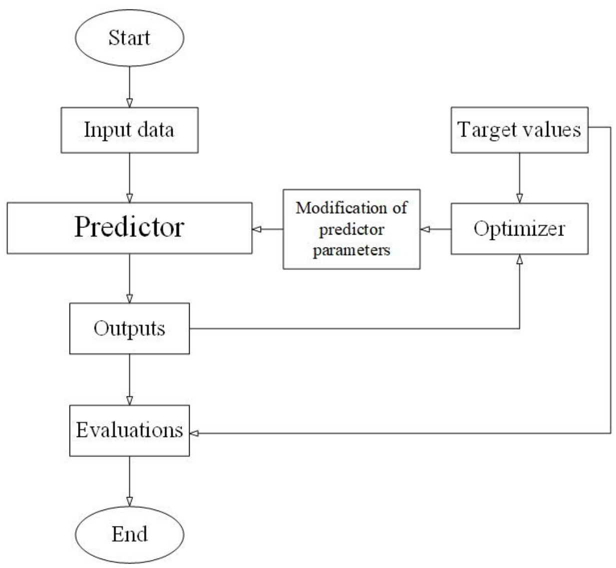

FIGURE 9. A flow-chart of hybrid methods.

provides the highest determination coefficient and the lowest RMSE equal to 0.9 and 1.625, respectively. Similarly, Ullah *et al.* [94] developed an ANN-based smart system to detect lighting in smart cities (i.e. a classification task). The evaluation was performed by employing accuracy and reported that the proposed ANN provides an accuracy value of 92.6%. Recently, Keung *et al.* [95] developed an ANN-based monitoring system for drainage handling in a smart city. Monitoring and prediction of urban drainage employed data of IoT sensors and ANN capability, respectively. Results showed that the proposed ANN could successfully perform drainage prediction since it obtained 99% accuracy on the testing dataset related to Hong Kong. Banach *et al.* [96] developed a platform for mapping air pollution in smart cities using an ANN-based system. Specifically, the ANN was trained and tested for the required prediction task and integrated into a laboratory target prototype. The evaluation was based on using accuracy and showed that the ANN has acceptable performance and could be implemented and tested in a real operational scenario.

Sharad *et al.* [97] devised an ANN-based technique for solving the time problems in reaching destinations for bus drivers. The technique managed the urban bus transportation paths in smart cities and monitored them to find the shortest path. The authors demonstrated then the ANN could provide an accurate estimation of the arrival time effectively reducing the delays. Bennati *et al.* [98] employed various data-driven learning algorithms based on ANNs to investigate and evaluate their application to social welfare, fairness, and privacy in smart cities. The algorithms were evaluated through computer simulations based on real-world data (i.e. smart-meter readings and participatory sensing) and considering two implementation scenarios (i.e. smart grid and traffic congestion information system). The authors identified algorithm trade-offs and provided a set of guidelines depending on the requirements and privacy constraints of the specific smart-city scenario and application. Differently, Sharad *et al.* [99] developed a real-time managing console for public transportation systems in smart cities employing an ANN-based monitoring system. The latter computed the shortest path to reach a destination and provided that information to the bus driver. In addition, the ANN was used to estimate the arrival time for the commuters accurately. Based on the findings of a real-time implementation, the authors demonstrated that the proposed technique could successfully provide a fleet management console to administrators use as a real-time monitoring system in buses. Jiang and Claudel [100] implemented an intelligence platform for wireless technology for traffic/flash flood monitoring systems. The platform worked in real-time and provided high reliability and accuracy on complex problems arising in smart cities (e.g., traffic flow monitoring, machine-learning-based flash flood monitoring, and Kalman-filter-based vehicle trajectory estimation). More specifically, for flash food monitoring, the authors employed an ANN that learned the variations of the air temperature profile in function of the ground and air temperature inputs measured by passive sensors.

B. ADVANCED MACHINE LEARNING APPROACHES: HYBRID, ENSEMBLES, AND DEEP LEARNING

This section presents hybrid approaches, ensembles, and DL-based techniques that we have categorized as advanced ML methods.

1) HYBRID APPROACHES

Hybrid approaches refer to integrating two or more (ML-based) methods for jointly exploiting their advantages in solving learning tasks (e.g., joint prediction and optimization) [101]. Figure 9 reports an example flowchart that illustrates the application of a hybrid method. Hereinafter, we provide a brief explanation of hybrid method development and goal. As depicted in Figure 9, input data are fed to the predictor component which in turn produces the output values. The latter values are given as input to the optimizer component that compares them with target values (i.e. the ground truth) to optimize a compound cost function. Depending on the specific optimization task, the cost function can be either minimized or maximized. In detail, this optimization procedure aims to tune predictor parameters and the cycle continues until achieving the desired performance. This is obtained by comparing the output of the optimized predictor and target values and computing the related evaluation metrics.

2) ENSEMBLE METHODS

Similarly, ensemble methods jointly employ different ML techniques (usually called ''weak learners'') but for different purposes, such as decreasing variance and bias or increasing prediction performance. They are based on the assumption

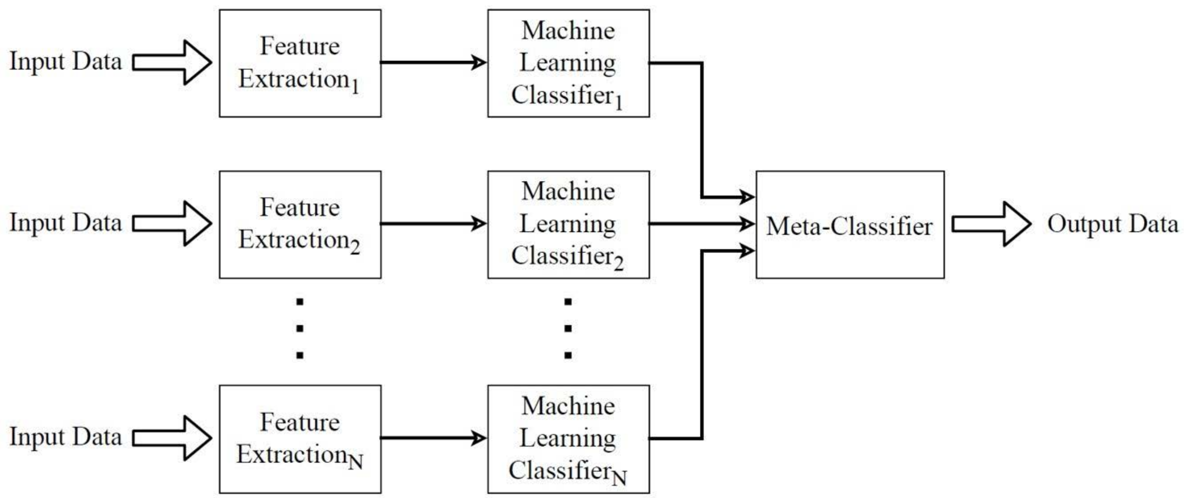

FIGURE 10. An ensemble model encompassing N ML-based classifiers.

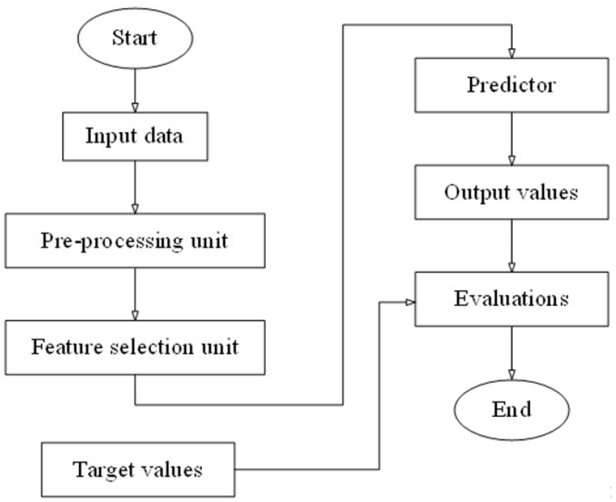

FIGURE 11. A flow-chart of the general ensemble method.

that combining multiple models to solve the same problem can produce a model with better performance. Figure 10 depicts the general structure of an ensemble model encompassing N ML-based "weak classifiers" whose outputs are combined via a meta-classifier. It should be noted that each ML-based classifier can be fed with a different set of features (viz., inputs).

Bagging, stacking, and boosting are common meta-algorithms for obtaining an ensemble of ML-based algorithms [102]. Bagging employs homogeneous weak learners trained in parallel and combines their outcomes using deterministic averaging. Bagging is frequently used to successfully improve the performance of DTs used as weak learners in RF [103]. On the other hand, boosting also uses homogenous weak learners trained sequentially in an adaptive fashion (i.e., there is a dependence between each model and the previous one) that are deterministically combined. Finally, stacking considers weak learners trained in parallel that are combined by training a meta-algorithm (i.e., a meta-classifier) that provides a prediction by intelligently combining the "base" models (see, e.g., Figure 7). Advanced combination techniques can exploit both hard decisions and soft outputs of base models [49].

Figure 11 sketches the flowchart of an example application of an ensemble method. In this workflow, the ensemble predictor is shown as a black box ("Predictor" in the Figure), and its role is independent of the specific ensemble meta-algorithm adopted. First, input data enter the pre-processing component that performs dataset cleaning and normalization. Then the data are passed to the feature selection unit. In more detail, the former component makes input values suitable to feed the ensemble predictor, whereas the latter aims to select the most informative features to improve ensemble performance that is assessed by comparing the output values with the target ones.

3) DEEP LEARNING

Among advanced ML techniques, DL has emerged as a possible disruptive breakthrough allowing the automatic design of inference systems that can distill complex dependencies among input data limiting human-expert need in designing accurate features. The term "deep" refers to the usage of multiple transformation steps to create these features, which is reflected in computations performed by a neural network encompassing many "hidden" layers placed between the input layer (passing input data to the first hidden layer) and the output layer (producing the output variables).

A wide variety of practical and robust methods are comprised within this subset of ML techniques. The most common DL architectures that we have found in the literature fall within the families of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Auto

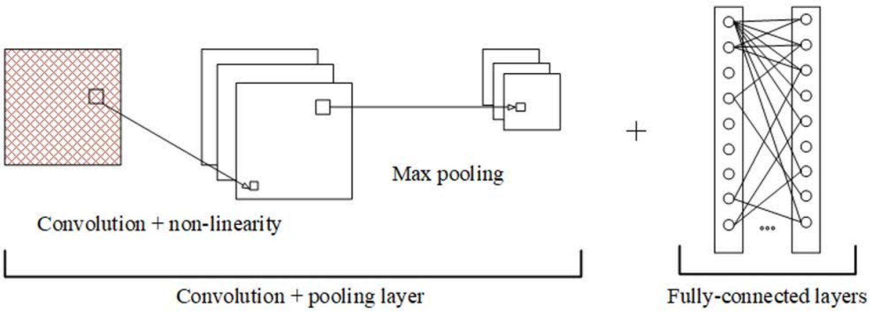

FIGURE 12. Example of a 2D-CNN architecture adopted by.

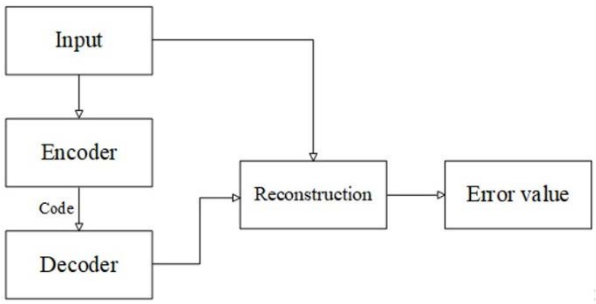

FIGURE 13. Example workflow of an AE.

Encoders (AEs) [104]. These DL techniques are widely used for multiple purposes, such as audio and speech processing, computer vision, network traffic analysis, social network filtering, pattern recognition, and big data applications. The parameters of DL networks are learned iteratively via the stochastic gradient descent optimization algorithm that finds the minimum of a cost (or loss) function. Specifically, an estimate of the gradients is calculated from a random subset of the training data. Also, the backpropagation algorithm is leveraged to efficiently compute the gradient of the loss function [79]. We briefly describe the most common variants of DL networks in the following.

A *CNN* architecture is inspired by the visual functioning of living creatures and is one of the most popular DL techniques, finding applications, especially in computer vision [105]. Figure 12 depicts an example of a bi-dimensional CNN (briefly a 2D-CNN). From a macroscopic viewpoint, the CNN architecture encompasses two main parts. The former is a chain of convolutional layers that employ transition-invariant filters—whose dimensionality depends on the input nature (e.g., bi-dimensional in the case of images)—that extract the features from a given input region within their receptive field by convolving with the input data. Commonly, each convolutional layer is followed by a pooling layer (e.g., a max-pooling in Figure 12 adapted from [106]) that performs down sampling of intermediate convolutional representation to reduce complexity and avoid overfitting.

The latter part consists of a series of fully connected layers that generate the proper output values depending on the considered task (e.g., classification vs. regression).

An *RNN* architecture presents neuron connections forming direct cycles and is usually employed to recall temporal information via a state vector. It has as input a vector sequence and outputs either its final state or its entire time-evolution. *Long Short-Term Memory (LSTM)* is one of the most common variants of RNNs and presents special neurons (called cells) that can store and model dynamic temporal behaviors with long-term dependencies. An LSTM cell is made of three main gates (i.e. internal mechanisms that operate with sigmoid and hyperbolic-tangent activation functions and sum and product operations of vector variables), namely input, output, and forget gates, which control the input and output of the cell, regulate the information flow, and decide which information is relevant to recall or forget [107].

The *AE* is a type of ANN commonly used for (unsupervised) feature learning, whose aim is to (ideally) output a reconstruction of the input by learning a compressed data representation. Figure 13 reports the example architecture of an AE. Specifically, the first AE block adopted from [108]. (i.e. the encoder) provides a lower-dimensional data representation (via a hidden layer of neurons), whereas the second block (i.e., the decoder) tries to reconstruct the data from the compressed representation [108] . The AE is commonly trained via fast, optimized backpropagation algorithms like the conjugate gradient [109]. Several studies have demonstrated the higher capability of advanced ML approaches (i.e., hybrid, ensemble, and DL techniques) in designing accurate models compared to traditional ML approaches. Ardabili *et al.* [110] presented a comparative study among single and hybrid Extreme Learning Machine (ELM) techniques for predicting and optimizing ethyl and methyl esters production, claiming that hybrid ELM techniques provided higher accuracy and optimized efficiency performance compared with that of single ELM. Jesús Cuenca-Jara *et al.* [111] proposed a novel data-driven methodology employing a fuzzy classifier based on volunteer geographic information to label spatial-temporal trajectories. Results were evaluated considering real-time detection of tourists and local citizens' flows. Comparisons were performed regarding classification accuracy with a well-established trajectory classifier used as a baseline, proving that the proposed solution is suitable for coping with the task.

TABLE 6. Notable smart city studies that used advanced ML methods.

References	Year	Description	Method	Findings	Application	Pros. and Cons.
[112]	2021	Using an RNN-based technique to detect cyber-attacks in a smart city	LSTM	The proposed technique provided an accuracy higher than 90%	Cyber-Resilient	Higher ability in general datasets
[113]	2021	Using an RNN-based technique for traffic prediction in a smart city	LSTM	The proposed technique provided a higher accuracy	Noise pollution analyses	Higher accuracy in time-series data
[114]	2020	Broad DRL technique for improving and supporting traffic services in a smart city	broad reinforcement learning	The proposed technique improved the accuracy by about 11.7%.	Traffic services	Time-consuming processing
[115]	2020	To employ Bi-LSTM for recognition of duplicacy within the medical community sites	Bi-Directional LSTM	The proposed technique provided an accuracy of 86.375%	healthcare communication	High reliability and confidence
[116]	2019	Platform using LSTM for the prediction of load and cost of an electricity grid system	LSTM	LSTM showed a lower MAE and NRMSE than ANN and ELM	Load and cost of electricity grids	Higher prediction score
[117]	2019	Hybrid ACO-RR algorithm for optimally reusing existing resources and systems in small and medium smart cities	ACO-RR SVM DT ANN	ACO-RR provided higher accuracy than single ML-based techniques	Resource management	Higher complexity
[118]	2018	Composition of RNN and CNN (RCNN) for the prediction of criminal acts	RCNN	RCNN had higher accuracy than single DL algorithms but higher training time	Prediction of criminal acts	Higher adoptability with the general dataset
[119]	2018	Hybrid reasoning model (CBR) for handling big data is aimed at obtaining a healthy environment in smart cities	CBR SVM DT k-NN BN LR	Hybrid CBR outperforms single ML-based methods by proper management of big data	Health	Higher complexity

TABLE 6. ***(Continued.)*** **Notable smart city studies that used advanced ML methods.**

[120]	2017	Survey discussing the application of DL techniques for handling data generated by smart connected health systems	CNN RNN DBF	DL techniques proved to be effective in handling data of smart health systems	Health	Higher complexity and need for a robust processor
[121]	2017	Ensemble method of ML-based classifiers to predict home-care hours	LR RF	The ensemble method reached an AUC value up to 0.715	Home-care	Reasonable accuracy
[122]	2017	Hybrid ANN-ARIMA technique to predict the real-time vehicle position in smart cities based on GPS data	ANN-ARIMA	ANN-ARIMA showed 95% accuracy in predicting vehicle position	Vehicle position	Reasonable accuracy in real-time application
[123]	2017	Ensemble technique based on GBRT for the prediction of car parking availability with fifteen-minute intervals	GBRT	The ensemble GBRT had higher performance than SVM and DT	Car parking availability	Higher accuracy and complexity

Recent research has shown that advanced ML techniques have become more and more popular due to their applicability in different research fields and higher performance when compared to traditional ML approaches. Smart city application is one of the most relevant fields that has found benefit from the appropriate usage of advanced ML methods. Table 6 presents notable papers—starting from the most recent ones—that have leveraged advanced ML methods for smart city applications.

In [112], an RNN-based LSTM platform was proposed to detect cyber-attacks in a smart city. The proposed technique provided an accuracy of more than 90%. In [113], the RNN-based LSTM technique was employed for preparing a platform to estimate traffic using noise pollution analyses in a smart city. The proposed technique provided a higher accuracy. In [115], Kumar employed Bi-LSTM to recognize the duplicity within the medical community sites. The obtained results suggested that the proposed technique provided an accuracy of 86.375%.

Yin *et al.* [117] proposed a hybrid Ant Colony Optimization Ridge Regression (ACO-RR) algorithm, a smart-city evaluation method based on ridge regression, exploited to help construct small and medium-sized smart cities intelligently reusing existing resources and systems. Experimental evaluation is performed considering real smart-city datasets spanning over different years and coming from the evaluation report on the development level of China's smart cities. The results showed that the hybrid ACO-RR technique provides higher accuracy compared to SVM, DT, and ANN, thus proving to be more reliable than single ML approaches in the evaluation of smart cities. Kwon *et al.* [119] developed a hybrid reasoning model via a combination of crowd knowledge extracted from open source data and collective knowledge (CBR) for handling huge amounts of data aimed at obtaining a healthy environment (i.e. diagnosing wellness levels in patients suffering from stress or depression) in smart cities. The empirical evaluation demonstrated that the proposed approach performs better than traditional ML-based methods (e.g., SVM, DT, k-NN, Bayesian Network, Logistic Regression) due to the ability of hybrid CBR to properly manage big data (and possible class imbalance). Belhajem *et al.* [122] presented a study on the real-time prediction of vehicle positions in a smart city using a hybrid approach based on ANN and Autoregressive Integrated Moving Average (ARIMA) techniques. The ANN-ARIMA model is trained with GPS data to jointly learn both linear and non-linear dependencies in vehicle positions. Results showed up to 95% accuracy in predicting vehicle position during GPS outages compared to the Extended Kalman Filter.

Besides, a group of works applied ensemble methods in smart cities. Hansen *et al.* [121] presented an ensemble method of ML-based classifiers exploiting both Logistic Regression (LR) and RF for forecasting home-care hours in a smart city. Experimental results are carried out considering data of Copenhagen citizens receiving home care from 2013 to 2017 and showed that the proposed method reaches an Area Under Curve (AUC) value of 0.715. The authors claimed that the proposed methodology can properly predict large increases in home-care hours, which is one of the major health expenses in a smart city. Alajali *et al.* [123] developed an ensemble technique based on Gradient Boosting Regression Trees (GBRT) for the prediction of car parking availability in smart cities. The method exploited data from multiple sources (i.e., car parking, pedestrian, and car traffic data) for extracting the relationship between pedestrian volume and car parking demand to predict parking availability at fifteen-minute intervals. The authors compared the proposed ensemble method with traditional SVM and DT inaccuracy and error probability. Experimental results demonstrated that the proposed ensemble technique has higher performance than single ML-based techniques, presenting an error probability of 0.029. Finally, DL techniques have also been widely applied in smart cities, as discussed here. Mujeeb *et al.* [116] employed LSTM for developing a prediction platform for the load and cost of an electricity grid system in the presence of data generated in smart cities. The proposed DL-based method was compared with ANN, and ELM techniques in terms of Mean Absolute Error (MAE) and Normalized Root Mean Square Error (NRMSE) measures. The results demonstrated that the LSTM outperformed compared forecasting methods in terms of accuracy, proving the efficiency of the proposed method for electricity price and load prediction. Indeed, the LSTM showed an MAE of 1.95 and an NRMSE of 0.08 for price forecasting on the ISO NE (Independent System Operator, New England) dataset, while an MAE of 2.9 and an NRMSE of 0.087 for load forecasting, showing better performance than ANN and ELM. Chackravarthy *et al.* [118] employed a DL architecture as a composition of an RNN with a CNN to predict criminal acts (e.g., assault detection, car theft, etc.) in smart cities. The proposed system aims to overcome the limitation of single DL techniques in analyzing video stream data playing criminal acts. The results showed higher accuracy compared to single DL algorithms at the cost of higher training time, thus allowing the implementation of an effective crime detection system that can reduce the workload of supervising officials in smart cities. Obinikpo *et al.* [120] presented a survey discussing the application of DL techniques for handling data generated by connected smart health systems. Specifically, they considered how these techniques can be exploited to improve the prediction of data sensed by IoT devices and to help decision-making in smart health services. The authors focused on both architectures (e.g., CNN, RNN, DBF, etc.) and methods for data collection using different sensor types, studying also challenges and open issues for identifying future directions for the application of DL techniques in smart health systems (e.g., medical imaging, bioinformatics, and predictive analysis).

IV. EVALUATION OF THE ML METHODS

As explained in the previous section, many ML algorithms tackled various challenges in smart cities. Therefore, we trust it is useful to augment our study by evaluating each ML technique (Section 5.2) based on different performance metrics (Section 5.1).

A. OVERVIEW OF PERFORMANCE EVALUATION METRICS

Several evaluation criteria were used to evaluate the ML algorithm used throughout the tens of papers that we reviewed in this study. Figure 14 shows the commonly used evaluation criteria and the frequency distribution of their use across nearly 40 case studies that adequately analyzed and reported on their model's performance. The figure shows that accuracy, precision, and recall are the most common metrics, followed by error-related metrics (i.e., MAE and RMSE) and the correlation coefficient. Other metrics are less common since they either complement the above metrics (e.g., MAPE and MSE) or are simply specific to the case study (e.g., sensitivity and specificity related to binary classification tasks).

B. EXPERIMENTAL RESULTS REPORTED IN THE SURVEYED CASE STUDIES

Next, we briefly explain the most common metrics (i.e. accuracy, precision, recall, RMSE, and correlation coefficient) and compare the experimental results reported in the studies we reviewed in this paper using them.

1) ACCURACY

Accuracy has a positive correlation with the performance of ML methods and a negative correlation with RMSE (in general with error-related metrics). Per Equation (11), accuracy is the fraction of correctly classified samples among the total number of samples:

$$Accuracy = \frac{True_p + True_n}{True_p + True_n + False_p + False_n} \quad (11)$$

where $True_p$ denotes the true positives, $True_n$ the true negatives, $False_p$ the false positive, and $False_n$ the false negatives.

Figure 15 depicts the accuracy values of different ML models used in surveyed works. The horizontal axis reports the methods developed in each research, while the vertical axis indicates each method's accuracy values (in percentage). We can notice that when comparing different techniques exploited for the same smart-city applications, single ML models (e.g., ANN, SVN, and Feed Forward Neural Networks) provided the lowest accuracy values (see e.g.,

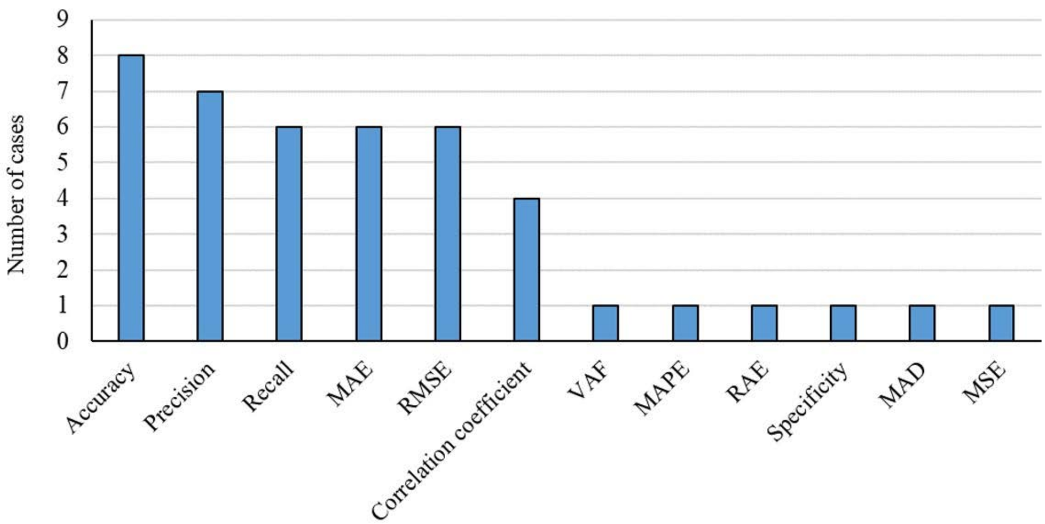

FIGURE 14. The frequency distribution of evaluation metrics in the surveyed literature.

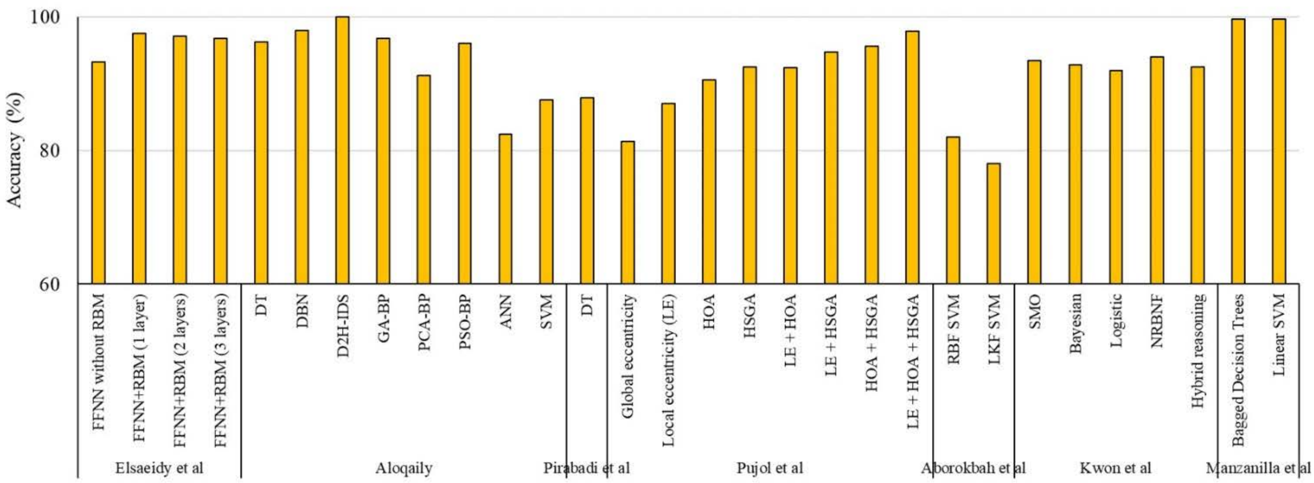

FIGURE 15. Comparison of accuracy values reported in the surveyed studies.

Elsaeidy *et al.* [32] and Aloqaily *et al.* [60]). On the other hand, the highest accuracy values are attained when an advanced approach such as hybrid, ensemble, and DL are adopted. The results provided by Manzanilla *et al.* [61], Pujol *et al.* [79], and Kwon *et al.* [119] support this conclusion.

This could be justified by the hypothesis that the inference power increases when we combine multiple predictors or voters which help optimize the final performance (see e.g., [124] and [30]).

2) RECALL

Recall (also known as sensitivity, particularly in binary classification) is a metric that measures the relevance of a model. Equation (12) shows the formal definition of the recall metric, defined as the fraction of relevant instances of a class that are correctly classified (i.e., the class-conditional accuracy):

$$Recall = \frac{True_p}{True_p + False_n} \tag{12}$$

where $True_n$ denotes the true positives and $False_n$ the false negatives. Figure 16 compares the results in terms of recall values as reported in reviewed studies. The horizontal axis reports the methods employed, while the vertical axis represents the associated recall values.

Again, single methods (e.g., SMO, NRBNF, LR) more often provided lower recall values as shown in Kwon *et al.* [119], whereas DTs and hybrid methods reached higher recall values based on the finding reported, for instance, by Elsaeidy *et al.* [32] and Kwon *et al.* [119].

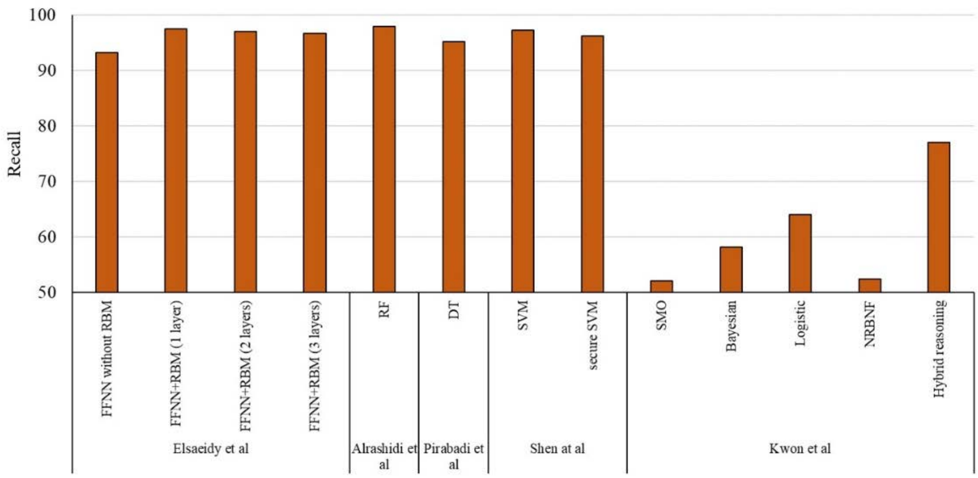

FIGURE 16. Comparison of recall values provided in the surveyed studies.

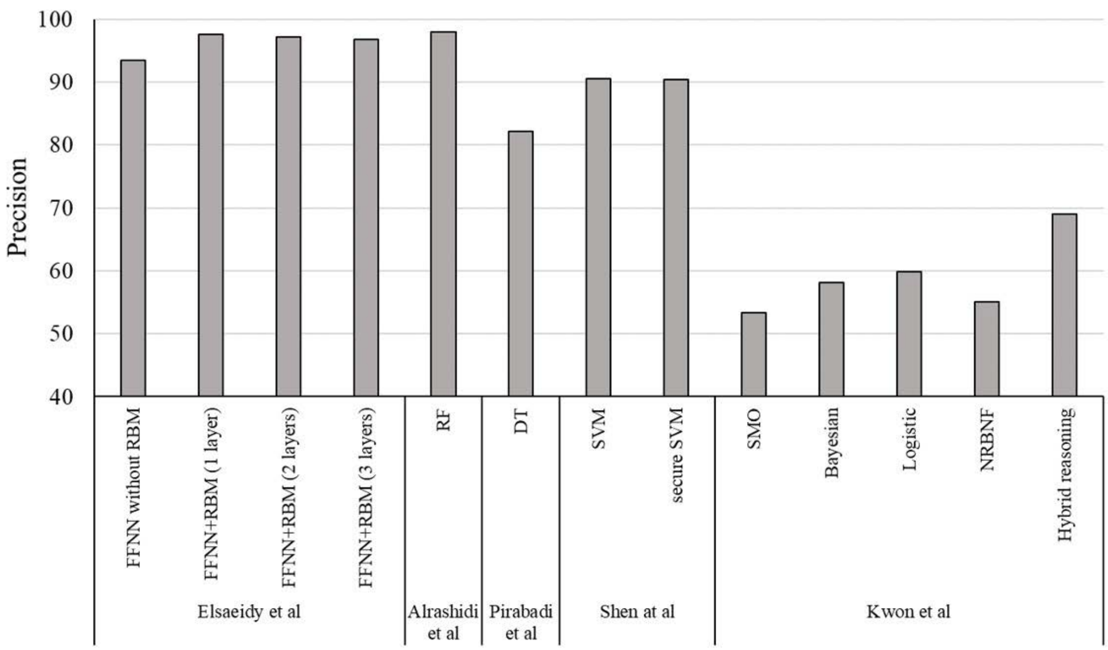

FIGURE 17. Comparison of precision values reported in the surveyed studies.

3) PRECISION

Precision is a metric that measures the overall performance stability of a model. Equation (13) outlines the formal definition of the precision metric, defined as the share of classifier decisions for a certain class that is correct:

$$Precision = \frac{True_p}{True_p + False_p} \tag{13}$$

where $True_p$ denotes the true positives and $False_p$ the false positives. Figure 17 reports the precision values of different ML techniques used in surveyed studies. By visually comparing Figures 16 and 17, we will notice that recall and precisions are correlated for the same studies they were reported in.

Like other performance metrics, we notice that DTs and hybrid methods provided a higher precision value compared to other methods based on the results reported in Elsaeidy *et al.* [32] and Kwon *et al.* [119]. Ensemble techniques (e.g., RF used in Alrashdi *et al.* [25]) also showed high precision values (i.e. >95%) when applied to smart city applications.

4) RMSE

RMSE is an error-related metric that measures the difference between actual and predicted values. In general, increasing the difference between actual and predicted values reduces the accuracy and increases the error metrics such as the

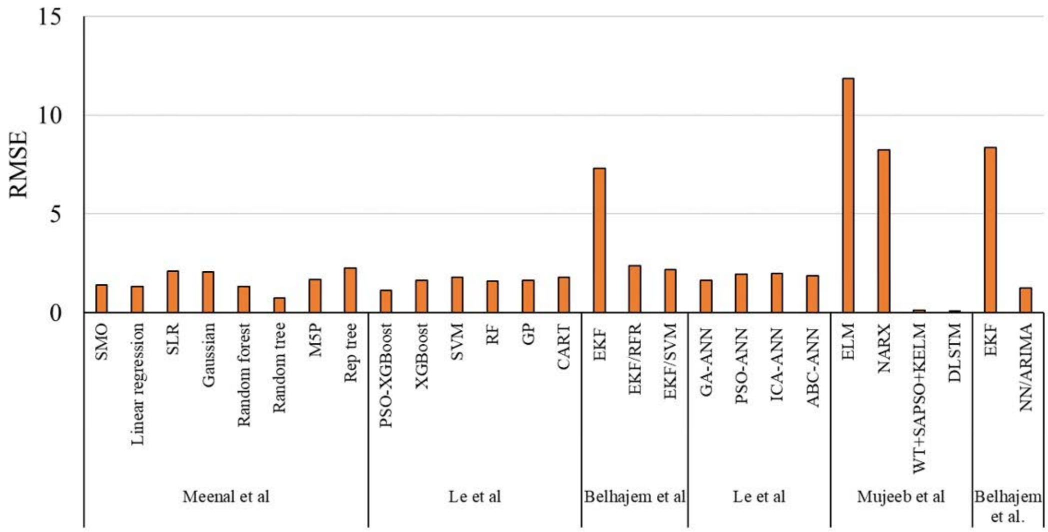

FIGURE 18. Comparison of RMSE values provided in the surveyed studies.

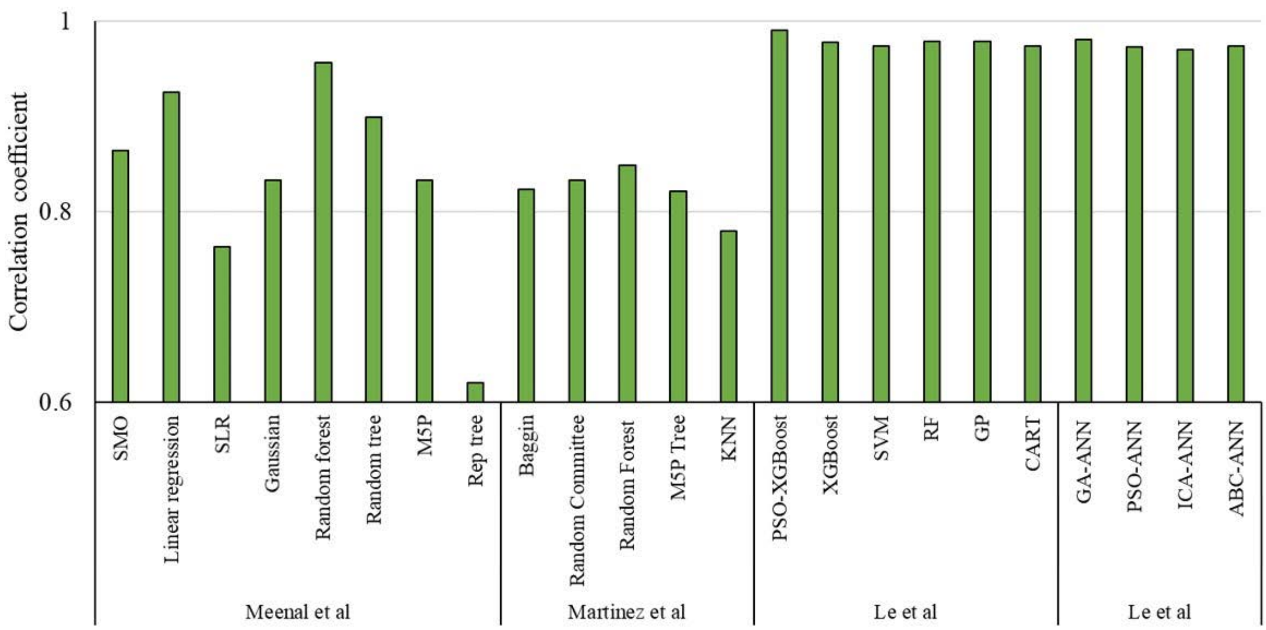

FIGURE 19. Comparison of correlation coefficient values reported in the surveyed studies.

RMSE. Equation 14 defines the RMSE formally as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)}^{2} \quad (14)$$

where N denotes the total number of samples, x_i the actual samples, and $\hat{x}_i$ the predicted samples. Figure 18 depicts the RMSE values obtained in the related studies using different ML techniques. Single ML techniques, also, provided the highest RMSE values (i.e. the lowest performance). Particularly, SVM, DT, and RF are the best performing ML methods, attaining the lowest RMSE values as reported in Le *et al.* [80] and Meenal *et al.* [34]. Moreover, from Figure 12 it is evident that the best models reaching the minimum RMSE are DL-based and hybrid techniques. For instance, the DLSTM and the hybrid WT+SAPSO+KELM compared in Mujeeb *et al.* [85], the hybrid GA-ANN, PSO-ANN, and ABC-ANN employed in Le et al [66], and the PSO-XGboost proposed in Le *et al.* [19] spawned significantly lower RMSE values compared with the single method baselines.

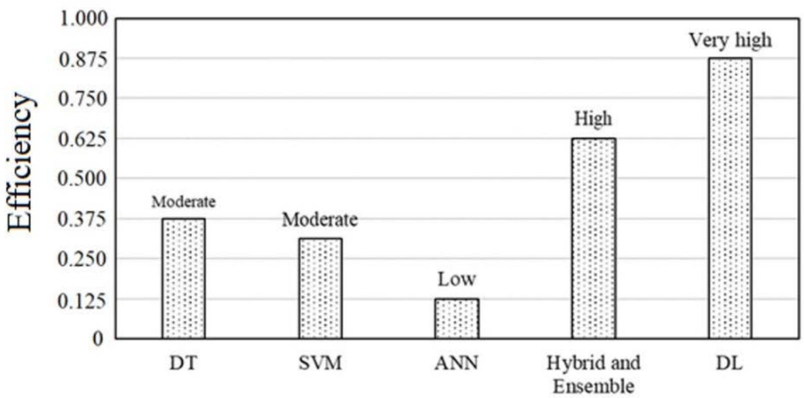

FIGURE 20. Processing time score of ML-based techniques used in smart cities applications.

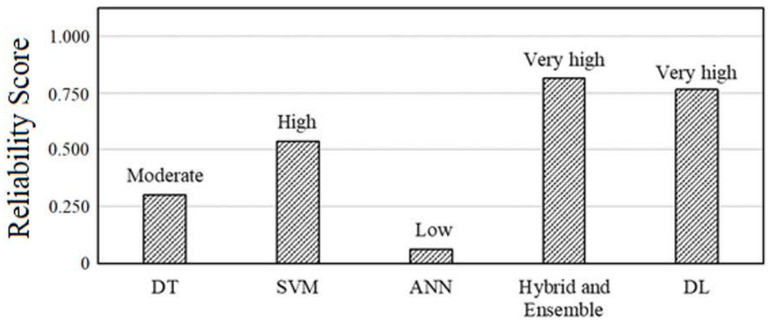

FIGURE 21. Reliability score of the ML-based techniques used in smart cities applications.

5) CORRELATION COEFFICIENT

The correlation coefficient measures the (linear) statistical relationship between actual and predicted values. In particular, a higher correlation between target and output values increases the overall accuracy and reduces total error. Equation (15) shows the formulation for calculating the correlation coefficient:

$$Correlation\ Coefficient = \frac{Cov(x, \hat{x})}{\sigma_x \sigma_{\hat{x}}} \tag{15}$$

where x refers to actual samples, $\hat{x}$ to predicted samples, $Cov(x, \hat{x})$ to the covariance between x and $\hat{x}$, and σ to the standard deviation (calculated for both x and $\hat{x}$). The correlation coefficient ranges between -1 and $+1$. A negative number indicates a negative correlation, whereas a positive number denotes a direct correlation between target and output values: the closer the coefficient to 1, the higher the resulting correlation as well as the accuracy. Figure 19 presents the comparison of the correlation coefficient obtained by different ML methods in reviewed studies. First of all, we can notice that the values of the correlation coefficient are always positive, indicating a direct correlation and thus the suitability of all proposed models for applications in smart cities. Again, Figure 19 demonstrates that single techniques provided a lower correlation coefficient than hybrid techniques with the notable exception of DT-based ones that had comparable performance (e.g., the RF in Meenal *et al.* [34]). Specifically, the hybrid techniques PSO-XGboost proposed in Le *et al.* [80] and GA-ANN and PSO-ANN presented in Le *et al.* [93] confirm this claim.

V. ANALYSIS AND DISCUSSION

In this section, we discuss ML methods used in smart cities from different perspectives. Based on our survey, we analyze how these methods compare to each other for efficiency (processing time), reliability (accuracy of results), and other performance aspects.

A. EFFICIENCY (PROCESSING TIME) ANALYSIS

For the processing time, Figure 20 sketches the processing time score, the lower this score the faster the ML algorithm is. As seen, the x-axis of the chart lists the ML algorithms while the y-axis represents the *processing time score*. These scores are normalized using min-max normalization by applying Equation 16.

$$X_N = \frac{\frac{Absolute\ processing\ time(s)}{\#\ data\ samples} - X_{min}}{X_{max} - X_{min}} \tag{16}$$

where X_N denotes the normalized processing time score, and X_{min} and X_{max} are the parameters used for the min-max normalization and depend on the specific dataset employed. This ensures having a range of scores between 0 and 1. For better interpretation, we further categorized this score into four zones:

(i) *Low* if $0 \leq X_N < 0.25$;
(ii) *Moderate* if $0.25 \leq X_N < 0.5$;
(iii) *High* if $0.5 \leq X_N < 0.75$;
(iv) *Very high* $0.75 \leq X_N \leq 1$.

The lower the score means that the ML algorithm is faster. Therefore, we noticed that the ANN is the fastest model, whereas DL and hybrid/ensemble are slower due to their complex computational architecture.

1) RELIABILITY ANALYSIS

When it comes to reliability, Figure 21 compares the accuracy of the output of each ML algorithms used in the smart city studies that we reviewed. The x-axis lists the ML algorithms, and the y-axis indicates the reliability score which is computed based on normalizing the performance metric (i.e. accuracy, precision, recall, RMSE, and correlation coefficient) used in the relevant work. To make these metrics comparable we normalized them using the min-max normalization as shown in Equation 17 at the bottom of the page, where Y_N denotes the normalized reliability score, and Y_{min} and Y_{max} the parameters (depending on the specific metric reported) used for the min-max normalization. This ensures having a range of scores between 0 and 1. For better interpretation, we further categorized this score into four zones:

(i) *Low* if $0 \leq Y_N < 0.25$;
(ii) (ii) *Moderate* if $0.25 \leq Y_N < 0.5$;
(iii) (iii) *High* if $0.5 \leq Y_N < 0.75$;
(iv) (iv) *Very high* $0.75 \leq Y_N \leq 1$.

Based on the reliability analysis, we conclude that that the ANN was the least reliable while DL and hybrids/ensemble methods are the highest. Among the single ML category, we noticed that the SVM had shown better performance (High) than the DT (Moderate).

2) OVERALL ANALYSIS

To deepen the comparative analysis of the studied ML methods, Table 12 gives a comprehensive comparison of the single ML-based, hybrid, ensemble, and DL-based models. The table describes the complexity, user-friendliness, accuracy, and processing speed of models used in smart city applications using the following categories: Low, Reasonable, Reasonably high, and High.

We can notice that hybrid models and ensembles are the best performers since they exhibit both high accuracy and not-costly complexity. On the other hand, and despite that the DL techniques had higher accuracy than the hybrid models and ensembles, but they demanded relatively higher computation power. Moreover, all these advanced ML methods had a slower processing speed than the single methods. Likewise, the SVM and DT generally outperformed the ANN for accuracy and other metrics. However, since the difference is negligible, we can conclude that using any one of them is appropriate (cf. Sec. 5).

The summary of Table 7 suggests that the advanced ML methods are the best candidates to use in mart cities based on accuracy and efficiency. Nevertheless, it is not uncommon to use the ANN and SVM as they have a simpler design, faster, and with acceptable accuracy.

3) PROS AND CONS

Table 8 highlights the pros and cons of each method and a recap of the discussion presented herein. Based on this report, we may claim that advanced ML models have superior performance to the single ML techniques, but given their higher complexity, they can still be used successfully in specific applications.

4) APPLICATION SHARE

Figure 22 depicts the relative share of each ML method under different smart city applications such as vehicles and transportation, mobile communications, building, energy, health care, data management, public safety, management of (IoT) sensors, pollution monitoring, and reduction, etc.

As illustrated in Figure 22, the ANN, DT, and SVM were predominantly used in smart transportation, mobile communication, IoT sensors, smart energy, smart education, smart building, and air pollution monitoring. On the other hand, the advanced ML models are commonly used in more complex applications, for instance, with those that have big data. Consequently, the hybrid, ensemble, and DL methods are more popular in smart health and transportation systems applications and manage open big data and resources in smart environments.

VI. OPEN ISSUES AND CHALLENGES

Smart city applications have been faced with ML-based techniques as a new paradigm in this area. ML-based techniques are introduced as the vital element of smart cities, but the developed studies have not sufficiently and comprehensively considered these techniques. This part of the study discusses some open issues and challenges that can be targeted for future studies. For example, smart-city-based datasets are big and used by time-sensitive applications that demand real-time or semi-real-time analytics. This highlights the need for a new analytic platform that supports big data analytics with fast/streaming data analytics. Furthermore, in applying the ML-based methods for smart city applications, the system's

$$Y_N = \frac{f(Accuracy,\ Precision,\ Recall,\ RMSE,\ Correlation\ Coefficient) - Y_{min}}{Y_{max} - Y_{min}} \tag{17}$$

TABLE 7. Comparative analysis of ML models applied in smart cities.

Model	Complexity	User-Friendliness	Accuracy	Processing Speed
DT	Reasonably high	Low	Reasonably high	Reasonable
SVM	Reasonably high	Low	Reasonably high	Low
ANN	Reasonable	Reasonable	Reasonable	High
Hybrid	High	High	High	High
Ensemble	High	Reasonable	High	High
DL	High	Reasonable	High	Reasonable

TABLE 8. Summary of advantages and disadvantages of ML-based models applied in smart cities.

Model	Advantages	Disadvantages
DT	Successfully applied in various smart city contexts	Moderate-to-high complexity and low user-friendliness
SVM	Frequently exploited for control-related applications	Low processing speed and user-friendliness
ANN	Successfully applied in smart cities with moderate-to-high accuracy	Low reliability in managing huge datasets
Hybrid	High accuracy, reliability, and user-friendliness	High complexity and moderate-to-low processing speed
DL	Very high accuracy and reliability and suitable to manage huge datasets	High complexity and low user-friendliness
Ensemble	High accuracy, reliability, and user-friendliness	High complexity and moderate-to-low processing speed

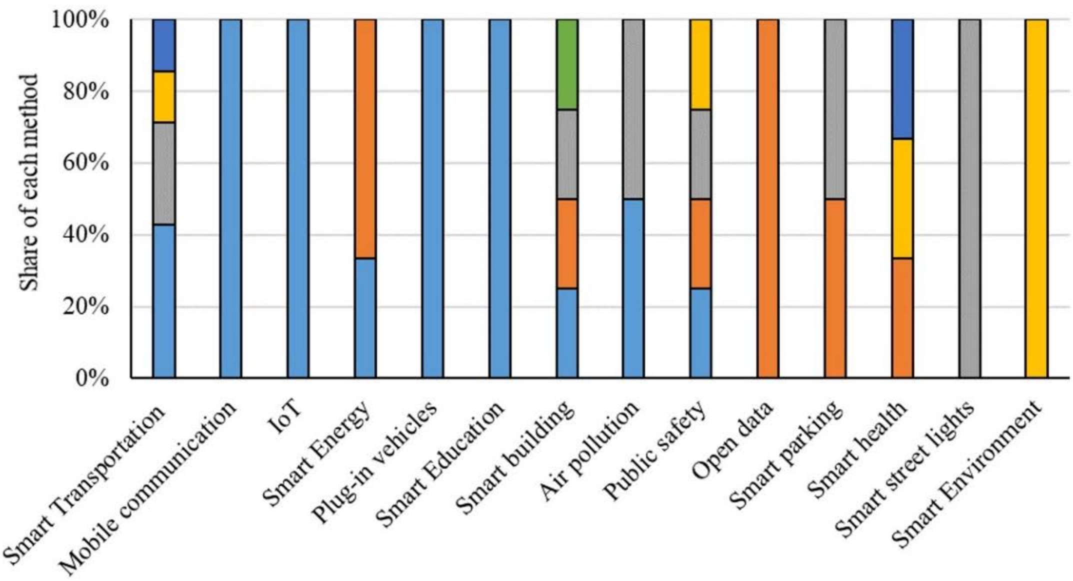

FIGURE 22. Share of each ML method under different types of smart city applications.

validity is closely related to the accuracy and precision of the data. On the other hand, data availability is a major challenge from the point of view of copyright issues and ethics. Furthermore, due to the nature of the data required for smart city applications, many performance domains can be easily rendered inaccessible if the results with large volumes of data for simulation are not confirmed. Therefore, the success of ML-based techniques in smart city applications depends on overcoming these challenges and excelling over them. Furthermore, due to the real-time applications of smart cities, the need for an ML-based technique that can provide high accuracy while providing a high operating speed and light platform can improve system reliability, stability, sustainability and availability.

A. ADVANCED ML AND DL TECHNIQUES

The number of applications in smart cities and their complexity will keep increasing due to the increase in the human population, the advent of new technologies every day, and the complexity of orchestrating all these systems together. This situation will continually generate big data that require more computational power and smarter algorithms. Handling this massive amount of data will remain a constant challenge for scientists to tackle by introducing more efficient and reliable ML algorithms that can be practically used in smart cities. The current advancements in ML and DL-based technologies rendered the concept of Smart Cities a reality. Nevertheless, more improvements will remain in demand if we aim to have smarter cities in all fields such as healthcare, security, transportation, traffic congestion, parking, pollution, etc.

B. IoT IN SMART CITY APPLICATIONS

The presence of IoT in smart city applications can be a game-changer in applications. Many open issues related to security, healthcare, safety, transportation, waste management, etc., can benefit from the IoT sensors. Combining those sensors and the data they collect with ML algorithms can foster the development of smart cities and make them more efficient and sustainable. This alliance between the sufficient datasets collected by IoT sensors and more powerful ML algorithms can provide practical solutions for the serious challenges that smart cities encountering today.

C. ML FOR SECURITY AND PRIVACY

Nowadays, cities are changing to evolve as smart cities all over the world. Accordingly, they need to collect and analyze huge volumes of data for different applications like automating processes, enhancing service quality, improving marketing services for users, and making better decisions. One of the main challenges of the creation of smart cities is to increase the quality of life for humans using digital interconnectivity, leading to increased efficiency and accessibility in cities. This leads Smart cities to move towards the enhancement of privacy and security to ensure the participation of citizens because the existence of security and privacy in society guarantees the satisfaction of citizens and the stability of the society. Therefore, one of the most important challenges of a smart city is ensuring security and privacy. Security and privacy challenges in the smart city include different subsections, which are described as follows:

1) CYBER RISKS

Smart City covers several advantages and benefits. IoT-based technologies in smart city applications can successfully enhance critical infrastructures. But there are arrangements required for preventing cyber risks to smart cities, such as threats that endanger the safety of citizens and the continuation of operations and services. Also, these arrangements have to prevent personal privacy reliance on rapid data sharing and data mining techniques. A smart city is integrated with a database to store data securely. In the meantime, employing ML-based techniques can successfully prevent cyber-attacks and strengthens security infrastructure. In fact, ML techniques benefit pattern recognition ability, estimation of behaviors, organizing a huge volume of files, recognizing potentially dangerous ones, and blocking perceived threats.

2) PUBLIC SAFETY

The requirement for public safety in a smart city is a growing challenge in general. New digital technologies to enhance the efficiency of different applications of smart cities are followed by urban population growth. The 5G technology, AI and IoT are the basis of smart cities. Progress in all of these increases the sense of need for public safety. Increasing public safety in smart cities increases trust and confidence in the system. The progress in AI and ML-based techniques for smart city applications can successfully enhance public safety by concentrating insights into the IoT networks to be monitored, analyzed, and acted upon in real-time. Data-based systems are one of the fields in which ML-based techniques have been successfully tested and employed to increase the system's abilities. Public safety in smart city applications can be considered one of the data-based systems that can be integrated with ML-based techniques for enhancing the system's efficiency. Several applications of ML-based techniques like Image processing, speech recognition, and efficient monitoring algorithms can be considered as elements to enable roaming ML-based techniques around public safety in smart city applications. ML-based techniques are considered a collection of intelligence technologies to provide considerable benefits to the criminal justice, Law Enforcement, Corrections, Courts, homeland security, and public safety domains such as Fire and Emergency Management Services.

3) MONITORING AND SENSOR-BASED TECHNOLOGIES

Monitoring in smart city applications is an innovative and significant open issue that can also be known as an effective challenge. Monitoring needs to be equipped with powerful information technology enabling ML-based transformation of big data into a wide range of custom services to monitor and control complex urban processes in real-time. Monitoring provides a holistic vision and transparency of the complex processes in the urban area as a practical system

TABLE 9. ML techniques that support the main security and privacy challenges in smart city applications.

Method	Cyber risks	Public privacy	Public safety	Monitoring and sensor-based technologies	References
DT			☑	☑	[63], [65], [68]
RF	☑			☑	[33], [67]
SVM		☑	☑	☑	[77], [83], [84]
ANN		☑		☑	[95]-[100]
RCNN			☑		[118]
kNN			☑		[119]
CNN			☑		[120]
RNN		☑			[121]
ARIMA				☑	[122]

in real-time applications. Accordingly, the stakeholders can enhance the efficiency and quality of Local and Regional Management and Governance. Accordingly, the quality of life and community transparency will promote new business models. Monitoring can be employed for traffic, public transportation, and natural hazards monitoring systems like flash floods and air pollution monitoring systems. The rise of smart city applications toward monitoring systems causes a considerable growth opportunity for sensor makers. This growth supports technologies such as 5G, robots, AI, and edge computing for smart city applications. The electronic, infrared, thermal, and proximity sensors are sensor technologies for smart city applications. As it is clear, the future of smart cities is intertwined with new technologies in the sensor industry, and we must wait for tremendous progress in this area. Table 9 summarizes the studies and ML-based techniques which support the security and privacy challenges in the smart city applications and identifies which challenges and ML-based techniques require high-level studies and experimental work for future perspectives.

VII. CONCLUSION

In this work, we present a comprehensive, systematic review of machine learning algorithms in smart city applications. As a result, we can conclude that the ML algorithms can fall into one of the following four categories: decision trees, support vector machines, artificial neural networks, and advanced machine learning methods (i.e., hybrid methods, ensembles, and Deep Learning techniques).

We give a theoretical description for each ML algorithm and demonstrate how it was used across many applications in the smart city context. Furthermore, we evaluate all reviewed ML algorithms concerning efficiency (computational speed), reliability (accuracy of the output), and the pros and cons of each. Among the many important observations we encountered through our analysis, we found that hybrid methods, ensembles, and deep learning techniques can outperform single methods at the cost of higher complexity and processing time. With this analysis and comparisons, we hope to guide researchers, practitioners, and policymakers to select the appropriate ML tool for the right problem. Many challenges and issues are still open for smart cities. We believe that coupling IoT with more powerful and reliable ML algorithms that can process a massive amount of data collected from the sensors will be the trend in the coming years. This might result in solutions for important problems typically associated with urban cities such as traffic, healthcare, pollution, education, etc.

REFERENCES

[1] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Sensing as a service model for smart cities supported by Internet of Things," *Trans. Emerg. Telecommun. Technol.*, vol. 25, no. 1, pp. 81–93, Jan. 2014.
[2] J. Desdemoustier, N. Crutzen, and R. Giffinger, "Municipalities' understanding of the smart city concept: An exploratory analysis in Belgium," *Technol. Forecasting Social Change*, vol. 142, pp. 129–141, May 2019.
[3] J. Edelenbos, F. Hirzalla, L. van Zoonen, J. van Dalen, G. Bouma, A. Slob, and A. Woestenburg, "Governing the complexity of smart data cities: Setting a research agenda," in *Smart Technologies for Smart Governments*. Cham, Switzerland: Springer, 2018, pp. 35–54.
[4] R. S. Farias, R. M. de Souza, J. D. McGregor, and E. S. de Almeida, "Designing smart city mobile applications: An initial grounded theory," *Empirical Softw. Eng.*, vol. 24, no. 6, pp. 3255–3289, Dec. 2019.
[5] R. Carli, M. Dotoli, and R. Pellegrino, "ICT and optimization for the energy management of smart cities: The street lighting decision panel," in *Proc. IEEE 20th Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2015, pp. 1–6.
[6] K. H. Law and J. P. Lynch, "Smart city: Technologies and challenges," *IT Prof.*, vol. 21, no. 6, pp. 46–51, Nov. 2019.
[7] R. Giffinger and H. Gudrun, "Smart cities ranking: An effective instrument for the positioning of the cities?" *Archit., City Environ.*, vol. 4, no. 12, pp. 7–26, 2010.
[8] V. P. Shrivastava and J. Singh, "'Smart city' definition ITS practice in India," *Int. J. Recent Technol. Eng.*, vol. 8, no. 4, pp. 43–53, 2019.
[9] H. Vasudavan, S. S. Gunasekaran, and S. Balakrishnan, "Smart city: The state of the art, definitions, characteristics and dimensions," *J. Comput. Theor. Nanosci.*, vol. 16, no. 8, pp. 3525–3531, Aug. 2019.
[10] A. A. Omar, A. K. Jamil, A. Khandakar, A. R. Uzzal, R. Bosri, N. Mansoor, and M. S. Rahman, "A transparent and privacy-preserving healthcare platform with novel smart contract for smart cities," *IEEE Access*, vol. 9, pp. 90738–90749, 2021.
[11] D. van den Buuse and A. Kolk, "An exploration of smart city approaches by international ICT firms," *Technol. Forecasting Social Change*, vol. 142, pp. 220–234, May 2019.
[12] O. B. Mora-Sanchez, E. Lopez-Neri, E. J. Cedillo-Elias, E. Aceves-Martinez, and V. M. Larios, "Validation of IoT infrastructure for the construction of smart cities solutions on living lab platform," *IEEE Trans. Eng. Manag.*, vol. 68, no. 3, pp. 899–908, Jun. 2020.
[13] S. H. Alsamhi, F. Afghah, R. Sahal, A. Hawbani, M. A. A. Al-qaness, B. Lee, and M. Guizani, "Green Internet of Things using UAVs in B5G networks: A review of applications and strategies," *Ad Hoc Netw.*, vol. 117, Jun. 2021, Art. no. 102505.

[14] F. Almalki, S. H. Alsamhi, R. Sahal, J. Hassan, A. Hawbani, N. S. Rajput, A. Saif, J. Morgan, and J. Breslin, "Green IoT for eco-friendly and sustainable smart cities: Future directions and opportunities," *Mobile Netw. Appl.*, vol. 11, pp. 1-25, Aug. 2021.
[15] M. Ghahramani, M. Zhou, A. Molter, and F. Pilla, "IoT-based route recommendation for an intelligent waste management system," *IEEE Internet Things J.*, early access, Dec. 2, 2021, doi: 10.1109/JIOT.2021.3132126.
[16] M. Ghahramani, N. J. Galle, C. Ratti, and F. Pilla, "Tales of a city: Sentiment analysis of urban green space in Dublin," *Cities*, vol. 119, Dec. 2021, Art. no. 103395.
[17] S. H. Alsamhi, O. Ma, M. S. Ansari, and Q. Meng, "Greening Internet of Things for greener and smarter cities: A survey and future prospects," *Telecommun. Syst.*, vol. 72, no. 4, pp. 609–632, Dec. 2019.
[18] F. A. Almalki, B. O. Soufiene, S. H. Alsamhi, and H. Sakli, "A low-cost platform for environmental smart farming monitoring system based on IoT and UAVs," *Sustainability*, vol. 13, no. 11, p. 5908, May 2021.
[19] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of artificial intelligence and machine learning in smart cities," *Comput. Commun.*, vol. 154, pp. 313–323, Mar. 2020.
[20] N. Cvar, J. Trilar, A. Kos, M. Volk, and E. S. Duh, "The use of IoT technology in smart cities and smart villages: Similarities, differences, and future prospects," *Sensors*, vol. 20, no. 14, p. 3897, Jul. 2020.
[21] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: A survey," *Sustain. Cities Soc.*, vol. 60, Sep. 2020, Art. no. 102177.
[22] S. Nosratabadi, A. Mosavi, R. Keivani, S. Ardabili, and F. Aram, "State of the art survey of deep learning and machine learning models for smart cities and urban sustainability," in *Proc. Int. Conf. Global Res. Educ.*, 2019, pp. 228–238.
[23] I. U. Din, M. Guizani, J. P. C. Rodrigues, S. Hassan, and V. V. Korotaev, "Machine learning in the Internet of Things: Designed techniques for smart cities," *Future Gener. Comput. Syst.*, vol. 100, pp. 826–843, Nov. 2019.
[24] J. Souza, A. Francisco, C. Piekarski, and G. Prado, "Data mining and machine learning to promote smart cities: A systematic review from 2000 to 2018," *Sustainability*, vol. 11, no. 4, p. 1077, Feb. 2019.
[25] M. Batty, *Artificial Intelligence and Smart Cities*. London, U.K.: Sage, 2018.
[26] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: Approaches and challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.
[27] S. Bhattacharya, S. R. K. Somayaji, T. R. Gadekallu, M. Alazab, and P. K. R. Maddikunta, "A review on deep learning for future smart cities," *Internet Technol. Lett.*, vol. 1, no. 5, p. e187, 2020.
[28] K. Kolomvatsos and C. Anagnostopoulos, "Reinforcement learning for predictive analytics in smart cities," *Informatics*, vol. 4, no. 3, p. 16, Jun. 2017.
[29] A. Al-Rahamneh, J. J. Astrain, J. Villadangos, H. Klaina, I. P. Guembe, P. Lopez-Iturri, and F. Falcone, "Enabling customizable services for multimodal smart mobility with city-platforms," *IEEE Access*, vol. 9, pp. 41628–41646, 2021.
[30] A. Dineva, A. Mosavi, S. Faizollahzadeh Ardabili, I. Vajda, S. Shamshirband, T. Rabczuk, and K.-W. Chau, "Review of soft computing models in design and control of rotating electrical machines," *Energies*, vol. 12, no. 6, p. 1049, Mar. 2019.
[31] V. Shetty, "A tale of smart cities," *Commun. Int.*, vol. 24, no. 8, p. 16, 1997.
[32] A. Elsaeidy, K. S. Munasinghe, D. Sharma, and A. Jamalipour, "A machine learning approach for intrusion detection in smart cities," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–5.
[33] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, and H. Ming, "AD-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2019, pp. 305–310.
[34] R. Meenal and A. I. Selvakumar, "Assessment of solar energy potential of smart cities of Tamil Nadu using machine learning with big data," in *Advances in Big Data and Cloud Computing*, vol. 750, J. D. Peter, A. H. Alavi, and B. Javadi, Eds. Singapore: Springer-Verlag, 2019, pp. 27–36.
[35] T. Bilen, M. Erel-Ozcevik, Y. Yaslan, and S. F. Oktug, "A smart city application: Business location estimator using machine learning techniques," in *Proc. IEEE 20th Int. Conf. High Perform. Comput. Commun., IEEE 16th Int. Conf. Smart City, IEEE 4th Int. Conf. Data Sci. Syst. (HPCC/SmartCity/DSS)*, Jun. 2018, pp. 1314–1321.
[36] T. Bakhshi and M. Ahmed, "IoT-enabled smart city waste management using machine learning analytics," in *Proc. 2nd Int. Conf. Energy Conservation Efficiency (ICECE)*, Oct. 2018, pp. 66–71.
[37] V. Lourenco, P. Mann, A. Guimaraes, A. Paes, and D. de Oliveira, "Towards safer (smart) cities: Discovering urban crime patterns using logic-based relational machine learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–8.
[38] A. R. Reid, C. R. C. Pérez, and D. M. Rodríguez, "Inference of vehicular traffic in smart cities using machine learning with the Internet of Things," *Int. J. Interact. Design Manuf.*, vol. 12, no. 2, pp. 459–472, May 2018.
[39] R. Martínez-España, A. Bueno-Crespo, I. M. Timon-Perez, J. Soto, A. Muñoz, and J. M. Cecilia, "Air-pollution prediction in smart cities through machine learning methods: A case of study in Murcia, Spain," *J. Univers. Comput. Sci.*, vol. 24, no. 3, pp. 261–276, 2018.
[40] C.-C. Chung and T.-S. Jeng, "Information extraction methodology by web scraping for smart cities–using machine learning to train air quality monitor for smart cities," in *Proc. CAADRIA*, 2018, pp. 515–524.
[41] J. Chin, V. Callaghan, and I. Lam, "Understanding and personalising smart city services using machine learning, the Internet-of-Things and big data," in *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2017, pp. 2050–2055.
[42] S. H. Alsamhi, F. A. Almalki, H. Al-Dois, S. Ben Othman, J. Hassan, A. Hawbani, R. Sahal, B. Lee, and H. Saleh, "Machine learning for smart environments in B5G networks: Connectivity and QoS," *Comput. Intell. Neurosci.*, vol. 2021, Sep. 2021, Art. no. 6805151.
[43] B. Carrera, S. Peyrard, and K. Kim, "Meta-regression framework for energy consumption prediction in a smart city: A case study of Songdo in South Korea," *Sustain. Cities Soc.*, vol. 72, Sep. 2021, Art. no. 103025.
[44] A. Alagumalai, O. Mahian, M. Aghbashlo, M. Tabatabaei, S. Wongwises, and Z. L. Wang, "Towards smart cities powered by nanogenerators: Bibliometric and machine learning–based analysis," *Nano Energy*, vol. 83, May 2021, Art. no. 105844.
[45] Z. Ullah, F. Al-Turjman, U. Moatasim, L. Mostarda, and R. Gagliardi, "UAVs joint optimization problems and machine learning to improve the 5G and beyond communication," *Comput. Netw.*, vol. 182, Dec. 2020, Art. no. 107478.
[46] S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou, and M. Nijim, "Machine learning approaches for EV charging behavior: A review," *IEEE Access*, vol. 8, pp. 168980–168993, 2020.
[47] J. N. Morgan and J. A. Sonquist, "Problems in the analysis of survey data, and a proposal," *J. Amer. Stat. Assoc.*, vol. 58, no. 302, pp. 415–434, Jun. 1963.
[48] E. B. Hunt, J. Marin, and P. J. Stone, *Experiments in Induction*. Boston, MA, USA: Academic, 1966.
[49] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Multi-classification approaches for classifying mobile app traffic," *J. Netw. Comput. Appl.*, vol. 103, no. 1, pp. 131–145, Feb. 2018.
[50] W. Du and Z. Zhan, "Building decision tree classifier on private data," in *Proc. IEEE Int. Conf. Privacy, Secur. Data Mining*, vol. 14, Jan. 2002, pp. 1–8.
[51] C. H. Gladwin, *Ethnographic Decision Tree Modeling*. Newbury Park, CA, USA: Sage, 1989.
[52] Y. Freund and L. Mason, "The alternating decision tree learning algorithm," in *Proc. ICML*, 1999, pp. 124–133.
[53] G. M. Cramer, R. A. Ford, and R. L. Hall, "Estimation of toxic hazard—A decision tree approach," *Food Cosmetics Toxicol.*, vol. 16, no. 3, pp. 255–276, Jan. 1976.
[54] M. A. Friedl and C. E. Brodley, "Decision tree classification of land cover from remotely sensed data," *Remote Sens. Environ.*, vol. 61, pp. 399–409, Sep. 1997.
[55] L. G. Anthopoulos, "The smart city in practice," in *Understanding Smart Cities: A Tool for Smart Government or an Industrial Trick*. Cham, Switzerland: Springer, 2017, pp. 47–185.
[56] R. M. AlZoman and M. J. F. Alenazi, "A comparative study of traffic classification techniques for smart city networks," *Sensors*, vol. 21, no. 14, p. 4677, Jul. 2021.
[57] D. Ngabo, W. Dong, E. Ibeke, C. Iwendi, and E. J. M. Masabo, "Tackling pandemics in smart cities using machine learning architecture," *Math. Biosci. Eng.*, vol. 18, no. 6, pp. 1–19, 2021.

[58] M. Balta and İ. Özçelik, "A 3-stage fuzzy-decision tree model for traffic signal optimization in urban city via a SDN based VANET architecture," *Future Gener. Comput. Syst.*, vol. 104, pp. 142–158, Mar. 2020.
[59] R. Bardhan, R. Debnath, J. Gama, and U. Vijay, "REST framework: A modelling approach towards cooling energy stress mitigation plans for future cities in warming global south," *Sustain. Cities Soc.*, vol. 61, Oct. 2020, Art. no. 102315.
[60] M. Aloqaily, S. Otoum, I. A. Ridhawi, and Y. Jararweh, "An intrusion detection system for connected vehicles in smart cities," *Ad Hoc Netw.*, vol. 90, Jul. 2019, Art. no. 101842.
[61] O. Manzanilla-Salazar, F. Malandra, and B. Sanso, "ENodeB failure detection from aggregated performance KPIs in smart-city LTE infrastructures," in *Proc. 15th Int. Conf. Design Reliable Commun. Netw. (DRCN)*, Mar. 2019, pp. 51–58.
[62] T. Xu, H. Sun, G. Han, C. Ma, and L. Jiang, "A deployment model of charging pile based on random forest for shared electric vehicle in smart cities," in *Proc. 14th Int. Conf. Mobile Ad-Hoc Sensor Netw. (MSN)*, Dec. 2018, pp. 49–54.
[63] A. S. El-Wakeel, J. Li, M. T. Rahman, A. Noureldin, and H. S. Hassanein, "Monitoring road surface anomalies towards dynamic road mapping for future smart cities," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2017, pp. 828–832.
[64] E. Gomede, F. Gaffo, G. Briganó, R. de Barros, and L. Mendes, "Application of computational intelligence to improve education in smart cities," *Sensors*, vol. 18, no. 1, p. 267, Jan. 2018.
[65] C. Orlowski, A. Sarzyński, K. Karatzas, and N. Katsifarakis, "*Decision processes based on IoT data for sustainable smart cities* (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11290. Cham, Switzerland: Springer-Verlag, 2018, pp. 136–146.
[66] H. Mei, S. Poslad, and S. Du, "A game-theory based incentive framework for an intelligent traffic system as part of a smart city initiative," *Sensors*, vol. 17, no. 12, p. 2874, Dec. 2017.
[67] S. Benedict, "Revenue oriented air quality prediction microservices for smart cities," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 437–442.
[68] A. Pribadi, F. Kumiawan, M. Hariadi, and S. M. S. Nugroho, "Urban distribution CCTV for smart city using decision tree methods," in *Proc. Int. Seminar Intell. Technol. Appl. (ISITIA)*, Aug. 2017, pp. 21–24.
[69] A. Dainotti, A. Pescapé, and G. Ventre, "A cascade architecture for DoS attacks detection based on the wavelet transform," *J. Comput. Secur.*, vol. 17, no. 6, pp. 945–968, 2009.
[70] A. Petrillo, A. Pescapé, and S. Santini, "A collaborative approach for improving the security of vehicular scenarios: The case of platooning," *Comput. Commun.*, vol. 122, pp. 59–75, Jun. 2018.
[71] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1055–1064, Sep. 1999.
[72] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, 1992, pp. 144–152.
[73] J. A. K. Suykens, T. Van Gestel, and J. De Brabanter, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.
[74] G. Manogaran, J. J. Rodrigues, S. A. Kozlov, and K. Manokaran, "Conditional support-vector-machine-based shared adaptive computing model for smart city traffic management," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 1, pp. 174–183, Feb. 2021.
[75] M. M. Rashid, J. Kamruzzaman, M. M. Hassan, T. Imam, and S. Gordon, "Cyberattacks detection in IoT-based smart city applications using machine learning techniques," *Int. J. Environ. Res. Public Health*, vol. 17, no. 24, p. 9347, Dec. 2020.
[76] S. Ftaimi and T. Mazri, "Handling priority data in smart transportation system by using support vector machine algorithm," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 5, no. 6, pp. 1422–1427, Dec. 2020.
[77] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani, "Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7702–7712, Oct. 2019.
[78] F. Aymen and C. Mahmoudi, "A novel energy optimization approach for electrical vehicles in a smart city," *Energies*, vol. 12, no. 5, p. 929, Mar. 2019.
[79] F. A. Pujol, H. Mora, and M. L. Pertegal, "A soft computing approach to violence detection in social media for smart cities," *Soft Comput.*, vol. 24, no. 15, pp. 11007–11017, Aug. 2020.
[80] R. Carli, M. Dotoli, and R. Pellegrino, "A multi-period approach for the optimal energy retrofit planning of street lighting systems," *Appl. Sci.*, vol. 9, no. 5, p. 1025, Mar. 2019.
[81] K. T. Chui, M. D. Lytras, and A. Visvizi, "Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption," *Energies*, vol. 11, no. 11, 2018, Art. no. 2869.
[82] V. Garcia-Font, C. Garrigues, and H. Rifà-Pous, "Difficulties and challenges of anomaly detection in smart cities: A laboratory analysis," *Sensors*, vol. 18, no. 10, p. 3198, Sep. 2018.
[83] I. Belhajem, Y. B. Maissa, and A. Tamtaoui, "Improving vehicle localization in a smart city with low cost sensor networks and support vector machines," *Mobile Netw. Appl.*, vol. 23, no. 4, pp. 854–863, Aug. 2018.
[84] M. M. Aborokbah, S. Al-Mutairi, A. K. Sangaiah, and O. W. Samuel, "Adaptive context aware decision computing paradigm for intensive health care delivery in smart cities—A case analysis," *Sustain. Cities Soc.*, vol. 41, pp. 919–924, Aug. 2018.
[85] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.
[86] S. Amid and T. M. Gundoshmian, "Prediction of output energies for broiler production using linear regression, ANN (MLP, RBF), and ANFIS models," *Environ. Prog. Sustain. Energy*, vol. 36, no. 2, pp. 577–585, Mar. 2017.
[87] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, and A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," *Sol. Energy*, vol. 84, no. 8, pp. 1468–1480, Aug. 2010.
[88] I. Yılmaz and A. G. Yuksek, "An example of artificial neural network (ANN) application for indirect estimation of rock parameters," *Rock Mech. Rock Eng.*, vol. 41, no. 5, pp. 781–795, Oct. 2008.
[89] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *J. Pharmaceutical Biomed. Anal.*, vol. 22, no. 5, pp. 717–727, 2000.
[90] E. S. Elmolla, M. Chaudhuri, and M. M. Eltoukhy, "The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process," *J. Hazardous Mater.*, vol. 179, nos. 1–3, pp. 127–134, Jul. 2010.
[91] S. H. Alsamhi, F. Almalki, O. Ma, M. S. Ansari, and B. Lee, "Predictive estimation of optimal signal strength from drones over IoT frameworks in smart cities," *IEEE Trans. Mobile Comput.*, early access, Apr. 22, 2021, doi: 10.1109/TMC.2021.3074442.
[92] A. Singh, A. K. Patel, J. P. Deka, A. Das, A. Kumar, and M. J. G. Kumar, "Prediction of arsenic vulnerable zones in the groundwater environment of a rapidly urbanizing setup, Guwahati, India," *Geochemistry*, vol. 80, no. 4, 2019, Art. no. 125590.
[93] L. T. Le, H. Nguyen, J. Dou, and J. Zhou, "A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning," *Appl. Sci.*, vol. 9, no. 13, p. 2630, Jun. 2019.
[94] I. Ullah, M. N. R. Baharom, H. Ahmad, F. Wahid, H. M. Luqman, Z. Zainal, and B. Das, "Smart lightning detection system for smart-city infrastructure using artificial neural network," *Wireless Pers. Commun.*, vol. 106, no. 4, pp. 1743–1766, Jun. 2019.
[95] K. L. Keung, C. K. M. Lee, K. K. H. Ng, and C. K. Yeung, "Smart city application and analysis: Real-time urban drainage monitoring by IoT sensors: A case study of Hong Kong," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2018, pp. 521–525.
[96] M. Banach, T. Talaśka, J. Dalecki, and R. Długosz, "New technologies for smart cities–high-resolution air pollution maps based on intelligent sensors," *Concurrency Comput., Pract. Exper.*, vol. 32, no. 13, 2019, Art. no. e5179.
[97] S. Sharad, P. B. Sivakumar, and V. A. Narayanan, "The smart bus for a smart city—A real-time implementation," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst.*, Nov. 2017, pp. 1–6.
[98] S. Bennati, I. Dusparic, R. Shinde, and C. Jonker, "Volunteers in the smart city: Comparison of contribution strategies on human-centered measures," *Sensors*, vol. 18, no. 11, p. 3707, Oct. 2018.
[99] S. Sharad, P. B. Sivakumar, and V. A. Narayanan, "The smart bus for a smart city—A real-time implementation," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, Nov. 2016, pp. 1–6.

[100] J. Jiang and C. Claudel, "A high performance, low power computational platform for complex sensing operations in smart cities," *HardwareX*, vol. 1, pp. 22–37, Apr. 2017.
[101] T. Shon and J. Moon, "A hybrid machine learning approach to network anomaly detection," *Inf. Sci.*, vol. 177, no. 18, pp. 3799–3821, Sep. 2007.
[102] I. Syarif, E. Zaluska, A. Prugel-Bennett, and G. Wills, "Application of bagging, boosting and stacking to intrusion detection," in *Proc. Int. Workshop Mach. Learn. Data Mining Pattern Recognit.*, 2012, pp. 593–602.
[103] J. Dou, A. P. Yunus, D. T. Bui, A. Merghadi, M. Sahana, Z. Zhu, C.-W. Chen, Z. Han, and B. T. Pham, "Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan," *Landslides*, vol. 17, no. 3, pp. 641–658, Mar. 2020.
[104] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 2, pp. 445–458, Feb. 2019.
[105] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[106] A. Deshpande, *A Beginner's Guide to Understanding Convolutional Neural Networks*, vol. 31. Boston, MA, USA: Academic, Mar. 2016.
[107] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep learning for solar power forecasting—An approach using AutoEncoder and LSTM neural networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 2858–2865.
[108] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, Sep. 2014.
[109] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
[110] S. F. Ardabili, B. Najafi, M. Alizamir, A. Mosavi, S. Shamshirband, and T. Rabczuk, "Using SVM-RSM and ELM-RSM approaches for optimizing the production process of methyl and ethyl esters," *Energies*, vol. 11, no. 11, p. 2889, Oct. 2018.
[111] J. Cuenca-Jara, F. Terroso-Sáenz, M. Valdés-Vela, and A. F. Skarmeta, "Classification of spatio-temporal trajectories from volunteer geographic information through fuzzy rules," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105916.
[112] M. Mohammadpourfard, A. Khalili, I. Genc, and C. Konstantinou, "Cyber-resilient smart cities: Detection of malicious attacks in smart grids," *Sustain. Cities Soc.*, vol. 75, Dec. 2021, Art. no. 103116.
[113] F. M. Awan, R. Minerva, and N. Crespi, "Using noise pollution data for traffic prediction in smart cities: Experiments based on LSTM recurrent neural networks," *IEEE Sensors J.*, vol. 21, no. 18, pp. 20722–20729, Sep. 2021.
[114] J. Tang, X. Wei, J. Zhao, and Y. Gao, "Semi-supervised double duelling broad reinforcement learning in support of traffic service in smart cities," *IET Intell. Transp. Syst.*, vol. 14, no. 10, pp. 1278–1285, Oct. 2020.
[115] A. Kumar, "Using cognition to resolve duplicacy issues in socially connected healthcare for smart cities," *Comput. Commun.*, vol. 152, pp. 272–281, Feb. 2020.
[116] S. Mujeeb, N. Javaid, M. Ilahi, Z. Wadud, F. Ishmanov, and M. Afzal, "Deep long short-term memory: A new price and load forecasting scheme for big data in smart cities," *Sustainability*, vol. 11, no. 4, p. 987, Feb. 2019.
[117] Q. Yin, K. Niu, N. Li, X. Peng, and Y. Pan, *ACO-RR: Ant Colony Optimization Ridge Regression in Reuse of Smart City System*, vol. 11602, X. Peng, A. Ampatzoglou, and T. Bhowmik, Eds. Cham, Switzerland: Springer-Verlag, 2019, pp. 204–219.
[118] S. Chackravarthy, S. Schmitt, and L. Yang, "Intelligent crime anomaly detection in smart cities using deep learning," in *Proc. IEEE 4th Int. Conf. Collaboration Internet Comput. (CIC)*, Oct. 2018, pp. 399–404.
[119] O. Kwon, Y. S. Kim, N. Lee, and Y. Jung, "When collective knowledge meets crowd knowledge in a smart city: A prediction method combining open data keyword analysis and case-based reasoning," *J. Healthcare Eng.*, vol. 2018, pp. 1–15, Oct. 2018.
[120] A. Obinikpo and B. Kantarci, "Big sensed data meets deep learning for smarter health care in smart cities," *J. Sensor Actuator Netw.*, vol. 6, no. 4, p. 26, Nov. 2017.
[121] C. Hansen, C. Hansen, S. Alstrup, and C. Lioma, "Smart city analytics: Ensemble-learned prediction of citizen home care," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Nov. 2017, pp. 2095–2098.
[122] I. Belhajem, Y. Ben Maissa, and A. Tamtaoui, "A hybrid machine learning based low cost approach for real time vehicle position estimation in a smart city," in *Proc. Int. Symp. Ubiquitous Netw.*, vol. 397, M. Benjillali, R. El-Azouzi, D. S. Menasche, E. Sabir, and F. De Pellegrini, Eds. Singapore: Springer-Verlag, 2017, pp. 559–572.
[123] W. Alajali, S. Wen, and W. Zhou, "On-street car parking prediction in smart city: A multi-source data analysis in sensor-cloud environment," in *Proc. Int. Conf. Secur., Privacy Anonymity Comput., Commun. Storage*, vol. 10658, G. Wang, Z. Yan, K. K. R. Choo, and M. Atiquzzaman, Eds. Cham, Switzerland: Springer-Verlag, 2017, pp. 641–652.
[124] A. Mosavi, M. Salimi, S. Faizollahzadeh Ardabili, T. Rabczuk, S. Shamshirband, and A. Varkonyi-Koczy, "State of the art of machine learning models in energy systems, a systematic review," *Energies*, vol. 12, no. 7, p. 1301, Apr. 2019.