IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 8, 2019, accepted December 16, 2019, date of publication December 23, 2019, date of current version January 2, 2020.

*Digital Object Identifier 10.1109/ACCESS.2019.2961372*

# Machine Learning Adoption in Blockchain-Based Smart Applications: The Challenges, and a Way Forward

**SUDEEP TANWAR[1], QASIM BHATIA[1], PRUTHVI PATEL[1], APARNA KUMARI[1], PRADEEP KUMAR SINGH[2], AND WEI-CHIANG HONG[3], (Senior Member, IEEE)**

[1]Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad 382481, India
[2]Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat 173234, India
[3]Department of Information Management, Oriental Institute of Technology, New Taipei 220, Taiwan

Corresponding author: Wei-Chiang Hong (samuelsonhong@gmail.com)

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 108-2410-H-161-004.

**ABSTRACT** In recent years, the emergence of blockchain technology (BT) has become a unique, most disruptive, and trending technology. The decentralized database in BT emphasizes data security and privacy. Also, the consensus mechanism in it makes sure that data is secured and legitimate. Still, it raises new security issues such as majority attack and double-spending. To handle the aforementioned issues, data analytics is required on blockchain based secure data. Analytics on these data raises the importance of arisen technology Machine Learning (ML). ML involves the rational amount of data to make precise decisions. Data reliability and its sharing are very crucial in ML to improve the accuracy of results. The combination of these two technologies (ML and BT) can provide highly precise results. In this paper, we present a detailed study on ML adoption for making BT-based smart applications more resilient against attacks. There are various traditional ML techniques, for instance, Support Vector Machines (SVM), clustering, bagging, and Deep Learning (DL) algorithms such as Convolutional Neural Network (CNN) and Long short-term memory (LSTM) can be used to analyse the attacks on a blockchain-based network. Further, we include how both the technologies can be applied in several smart applications such as Unmanned Aerial Vehicle (UAV), Smart Grid (SG), healthcare, and smart cities. Then, future research issues and challenges are explored. At last, a case study is presented with a conclusion.

**INDEX TERMS** Blockchain, machine learning, smart grid, data security and privacy, data analytics, smart applications.

## I. INTRODUCTION

From the past few decades, data has become an essential source of intelligence and carries new opportunities to the real-life problems such as wireless communications, bioinformatics [1], agriculture [2], and finance [3] through smart applications. These applications are data-driven and incorporate actionable insights into user experience, which enables individuals to complete the desired task more efficiently. It operationalizes insights, personalizes the customer experience, optimizes customer interactions, improves operational efficiency, and enable new business model. There are various smart applications such as SG, UAV, Smart Cities, which makes the life of an individual easier. These applications generate a huge amount of data, and storage of this ever-evolving data in databases is a problem, and its communication also raises security concerns. To handle these issues, BT can be used, which has a distributed database network. It was coined by Satoshi Nakamoto in the year 2008 and contained a time-stamped series of tamper-proof records, which are managed by a cluster of distributed computers. It comprises of a chain of blocks that are connected using cryptographic primitives. The three mainstays of BT are immutability, decentralization, and transparency. These three characteristics opened its door for a wide variety of applications, for example, digital currency existence (currency with no physical existence) and analysis on the suitability of it in smart applications [4]. Although BT ensures security and privacy

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

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

issues, some vulnerabilities also started appearing after its implementation. For instance, the nature of attacks began to be increasingly sophisticated such as majority attacks (51% attack) that control voting, Sybil attacks for fake identity generation to control the consensus [5]. To handle the aforementioned issue, a robust Intrusion Detection System (IDS) is required in place because the traditional methods use a signature-based approach to detect specific patterns. But, to detect intrusions and attack patterns, one of the emerging technology known as ML can be used to analyze the data traffic. Thus, designing efficient and effective algorithms to analyze this massive amount of data is in dire need of handling the blockchain-based smart applications. Hence, ML is highly prevalent today and uses a dozen times a day without even realizing it. ML encompasses computers to study, think, and act without intervention of humans. It is considered to be one of the applications of Artificial Intelligence (AI). ML provides computers the competency to learn without being programmed it explicitly. Its basic idea is to build an efficient algorithm that can accept input data and, with the help of statistical analysis, make a prediction, and update the outputs. A substantial amount of data can be analyzed by ML to create data-driven decisions.

In a communication network of blockchain-based smart applications, there is layer-wise handling of security issues. Some security issues are handled at the network layer, such as malicious packets and some at the application layer such as malware [6]. At the network layer, malicious packets can be used to impose the network to establish fraudulent consensus. A naive solution to this problem can be to use a firewall to ensure that packets meet pre-defined security criteria [7]. Though, the attacks are becoming more sophisticated with unseen patterns to bypass a firewall. To prevent this issue, packets header data can be analyzed using ML models [8] in real-time using historical data. This analysis helps to detect new and changing patterns. Similarly, ML techniques can be used to classify malware to end-point such as servers, mobile, or workstations. Further, several blockchain-based smart applications such as UAV [9], Data Trading [10], SG build trust between data exchangers [11]. It is very crucial in any smart application at the same time; data should be secure. BT ensures data security but to build confidence, and ML techniques are used to predict untrustworthy nodes based on past patterns. Similarly, UAVs have significantly different network topology compared to the conventional blockchain network topology [9]. It includes communication using satellites and various ground stations. For UAV, BT is used to securely store coordinates and other relevant data to maintaining graph integrity for the vehicles. In subsequent sections, we explore the recent research work on ML adoption in the blockchain-based smart application.

### A. SCOPE OF THIS SURVEY

Several survey work has been published till date by the different researchers on the various aspects of the adoption of ML in blockchain-based smart applications [5]. As per our knowledge, most of the surveys have focused on specific areas, fields, or applications where it requires both ML and blockchain. The proposed survey covers all the fundamental aspects of ML to be applied BT based applications, for instance, intrusion detection. A review conducted by Meng *et al.* [12] describes how blockchain help to meet intrusion detection. The authors mitigate trust issues by establishing a collaborative IDS using smart contracts [17]. Further, Conti *et al.* [13] surveyed on security and privacy issues of bitcoin. They discuss various categories of attacks, such as double-spending attacks, client-side security threats, and mining pool attacks. To handle the aforementioned issue, Rahouti *et al.* [5] discuss specific ML-based solutions and discuss the handling of specific social issues such as human trafficking and drug sales through cryptocurrency. Then, Ucci *et al* [14] explored malware analysis using ML techniques. Features of malware were thoroughly discussed, and a detailed taxonomy has been proposed. Salah *et al.* [15] and Casino *et al.* [16] conducted a review on blockchain-based applications. To clarify the main difference between other surveys paper and this survey paper, a comprehensive comparison has been shown in Table 1. It includes objectives, merits, and demerits of peer surveys concerning numerous parameters such as architecture, applications, open issues, taxonomy, and merits, demerits of the existing approaches. A master taxonomy of ML for BT is summarized in Figure 3.

### B. RESEARCH CONTRIBUTIONS

Though several research works exist to address the ML usages for the blockchain-based applications, but not been exploited to its full potential. In this paper, we investigated the ML usages for blockchain based smart applications. Following are the research contribution of this paper.

- A brief discussion on how ML and blockchain can be used together in smart applications with a proposed architecture.
- To develop taxonomy covering ML techniques required for BT based environment. In each part of the taxonomy, existing work has been discussed in detailed to handle several issues, such as preventing and predicting attacks on the blockchain network.
- A case study is presented to demonstrate the usage of ML techniques in blockchain-based smart applications, such as SG, UAV, etc.

### C. ORGANIZATION

The rest of the paper is organized as follows: Section 2 provides a background of Blockchain, ML, and list all benefits while applying ML for BT based application. Then, we proposed an architecture. Survey procedures and taxonomy presented in Section 3. Then, in Section 4-7, we present existing advancements in ML approaches for BT-based smart applications. A discussion on future research issues and challenges is presented in Section 8. Further, section 9 covers a case study on the SG system, and finally, we concluded the paper.

**TABLE 1.** Comparison of the existing surveys with the proposed survey.

| Authors | Year | Objectives of Survey | Merits | Demerits | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|
| Meng *et al.* [12] | 2018 | To present use of blockchain in intrusion detection | Scope of application of blockchain was discussed | Discusses only data sharing and trust management issues of collaborative intrusion detection | X | ✓ | ✓ | X | ✓ |
| Conti *et al.* [13] | 2018 | To discuss various security and privacy issues in bitcoin | A Comprehensive review of possible attacks on bitcoin and provided countermeasures | Blockchain issues are not highlighted | ✓ | X | ✓ | X | ✓ |
| Rahouti *et al.* [5] | 2018 | Survey on ML security solutions for bitcoin | In-depth and wide classification of major threats and extensive explanation of the role of ML | Other applications of blockchain are missing | X | X | ✓ | ✓ | ✓ |
| Ucci *et al.* [14] | 2018 | To study ML techniques for malware analysis | Time and space complexity for various methodologies has been described in detail | Lacks discussion on uses of these techniques in a blockchain environment | X | ✓ | ✓ | ✓ | X |
| Salah *et al.* [15] | 2019 | Discuss applications, platforms, and protocols in blockchain specifically for AI | The decentralization feature of blockchain is explained with a specific view of AI | Discussion on privacy is not covered in detail | ✓ | ✓ | ✓ | ✓ | X |
| Casino *et al.* [16] | 2019 | Review blockchain-based applications and identify open issues | Prerequisites for blockchain applications are thoroughly discussed | Focused on applications, not the open issues | X | ✓ | ✓ | ✓ | X |
| Proposed Survey | - | To survey how ML can be used in blockchain-based smart applications | Discusses architecture and technology at a fundamental level and bridges the gap between two technology | - | ✓ | ✓ | ✓ | ✓ | ✓ |

Parameters- 1:Architecture, 2:Application, 3:Open Issues and Challenges, 4:Taxonomy, 5:Security
Notations- ✓: considered, and X: not considered.

## II. BACKGROUND

### A. MACHINE LEARNING

ML is the field of study that focuses on building applications that learn through experience. It is the ability to teach a computer without programming it explicitly [18]. ML encompasses its work from a diverse set of disciplines, including philosophy, information theory, probability and statistics, control theory, psychology and neurobiology, computational complexity, and artificial intelligence [19]. ML algorithms are used in many application and benefited it as listed below:

- In Data mining, large databases contain different patterns that can be discovered automatically by using ML techniques to analyze outcomes, for instance, medical treatments of a patient from health record databases or to identify the creditworthiness of a person from financial databases.
- ML applies in areas where a deterministic algorithm is not promising, such as human face recognition from images.
- Application domains where the adaptable programming is required, for instance, controlling manufacturing processes as per the demand of the customer and adapting to the varying reading interests of readers.

ML algorithm eare application specific and depends on the output required by the system. There are several ML algorithms, such as Supervised ML, Semi-Supervised ML, and Unsupervised ML. (i) Supervised ML uses statistical models to predict output in numerical data and classify the correct label [20]. Here, the most commonly known algorithms include the regression approach and decision trees. (ii) Unsupervised ML does not have label data. Here, data samples are grouped into clusters depending on their similarity or dissimilarity [21] using a different approach. For example, K means clustering and association rules algorithms. (iii) Semi-Supervised ML is also a category of interest [22]. It involves a mixture of supervised and unsupervised ML techniques. Unsupervised learning may be applied to discover the structure of input variables, following which it is used to make best guess predictions for the unlabeled data. It feeds that predicted data back into the supervised ML algorithm as training data and use the model to make predictions on unseen data.

### B. BLOCKCHAIN

Blockchains are an immutable set of records that are cryptographically linked together for audit [23]. It is similar to an accounting ledger. Here, previous records in accounting ledger cannot be changed, and new records need to be verified by a trusted party. The only difference between these two is that new blocks (set of records) checked by a decentralized structure of nodes that have a copy of the ledger. There is no centralized party to verify the records. Blockchain is formed by linking valid blocks together; the current block contains the hash of the previous block, and so on, as shown in FIGURE 1. This makes blockchain traceable and resistant to change [24]. Older blocks cannot be modified, in case they are changed in any way; their hash would change. This emphasis to link hash in all subsequent blocks to make the blockchain network valid again. A copy of the blockchain is available with every individual within the network; henceforth, any changes can be cross verified by the other users. These copies of the blockchain are updated with the addition of a new block. Then, everyone can see the block, depending on the permissions assigned by the administrator. BT uses a cryptographic secure hash algorithm (SHA) such as SHA-256 and SHA-512 to maintain the data integrity within the block. Each block has a unique hash value. For instance, Ethereum uses Keccak-256 and Keccak-512, while Bitcoin uses double SHA-256. This SHA is a collision-resistant algorithm, where no two different input data could produce the same output (hash value). Henceforward SHA can be used to check if the data is the same or not. There are various SHA algorithms

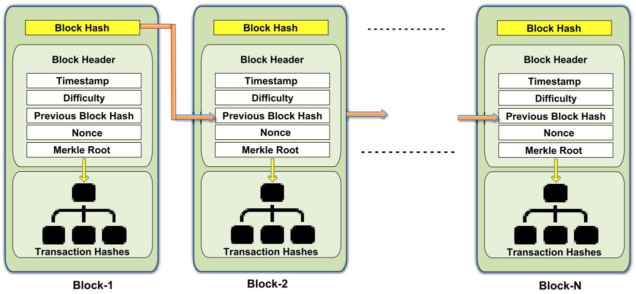


**FIGURE 1.** Blockchain structure.

developed by the NSA and NIST and belongs to the SHA-2 family [25]. SHA was initially envisioned as a fragment of the Digital Signature Standard (DSS) to produce the signature. Further, consensus algorithms are used to determine block validity. The selection of algorithms depends on the type of blockchain, such as public, private, and consortium blockchain. The selected algorithm should ensure the consensus among nodes [26]. It must be able to use resources efficiently and tolerate a degree of safety during the event of attacks.

Further, a smart contract is a program (set of codes) that runs on the blockchain and adds blocks whenever certain conditions are met. It is defined by translating actual legal contracts into programs to enforce the legal contract onto the blockchain [27]. It is similar to stored-procedures in relational databases. It is stored as scripts in the blockchain and executed according to the data fed to them to produce outputs that are expected from the original contract [28]. It governs transactions either executed fully or partially based on the current input. The primary purpose of a smart contract is to provide superior security with a reduction in cost and delays associated with traditional contracts.

### C. INTEGRATION OF MACHINE LEARNING IN BLOCKCHAIN-BASED APPLICATIONS

The learning capabilities of ML can be applied to blockchains based applications to make them smarter. By using ML security of the distributed ledger may be improved. ML may also be used to enhance the time taken to reach consensus by building better data sharing routes. Further, it creates an opportunity to build better models by taking advantage of the decentralized architecture of BT. We proposed architecture for ML adoption in BT-based smart application, as shown in Figure 2. Here, the smart application collects data from different data sources such as sensors, smart devices, and Intenet of Things (IoT) devices. Data collected from these devices get processed as part of smart applications. The blockchain work as an integral part of these smart applications. Then, ML can be applied to these application's data for analysis (Data analytics and real-time analytics) and prediction. The data sets used by ML models could be stored on a blockchain network. This reduces errors in the data such as duplication, missing data value, errors, and noise. Blockchains are focused on the data, and hence data-related issues in ML models may be eliminated. ML-models can be based on specific segments of the chain rather than the entire data-set. This could give custom models for different applications, such as fraud detection and identity theft detection. Few of the benefit are listed beneath when ML is applied:

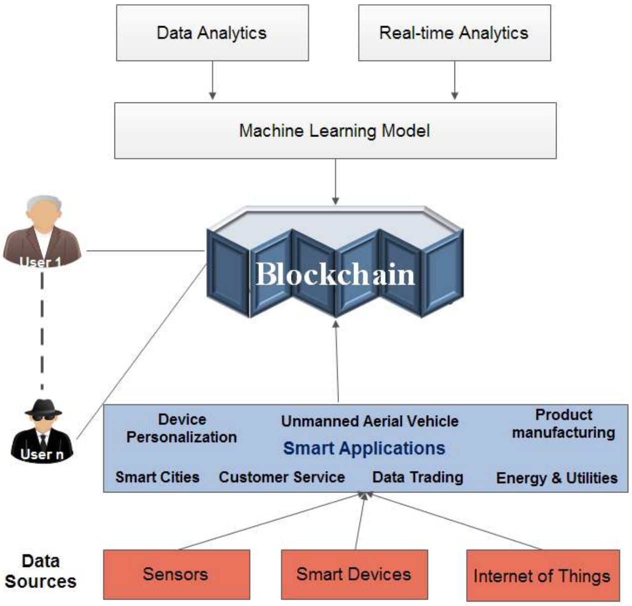


**FIGURE 2.** Proposed architecture for Machine Learning adoption in Blockchain-based Smart applications.

- User authentication as a legitimate user for requesting or performing any transaction in the blockchain network.
- BT provides a high level of security and trust.
- Blockchain integrates public ML models into smart contracts to ensure that the conditions and terms which were previously agreed are sustained.
- BT helps in the reliable implementation of an incentive-based system; thus, it encourages users/customers to contribute data. This huge data will help to improve ML model performance.
- ML models can be updated on-chain environment of BT with a small fee and off-chain, locally on an individual's device without any costs
- Good data contributions can happen from users/ customers, these data consistently computed, and rewards can be given to the users.
- Tamper-proof smart contracts can be evaluated by different machines (having different hardware configuration), ML models will not diverge from their potential and produce results exactly as it is supposed to do.
- Payments processed in real-time with trust on a blockchain environment.
- Blockchains tools, for instance, Ethereum, deal with thousands of decentralized machines all over the world. This guaranteed users that it is never completely unreachable or offline.

## III. SURVEY PROCEDURES AND TAXONOMY

In this section, we refer to the procedures used to perform this study; for example, our search approach and inclusion criteria for the final set of papers. Likewise, we present a detailed taxonomy based on our literature review.

### A. SYSTEMATIC LITERATURE REVIEW

We used standard databases (for example IEEEXplore, ACM Digital Library, ScienceDirect, and Springerlink) and Google Scholar to search existing research work, using keywords such as "Machine Learning for Blockchain-based Smart Application," ("Blockchain for Smart Applications" AND "Machine Learning for Smart applications"). In the initial phase, publications emphasis essentially on blockchain approaches for smart applications. Then, in the next phase of the search, we concentrated on an ML-based solution for blockchain-based smart applications. Based on the results of these searches, we removed duplicate articles and obtained our first set of more than 350 publications. Further, we steered the search procedure with several magazines, journals, and conferences dedicated to the parent field, ML, smart applications, and Blockchain. Based on that, we found 130 articles. Then, we studied the different sections of the articles like abstract, conclusion, and introduction. Then we categorized these articles as "relevant" or non-relevant" to ML for blockchain-based smart applications. Lastly, we only selected 60 publications to present the taxonomy, as shown in Figure 3. Here, each layer is color-coded to tailor each level of the taxonomy. For example, the root represents the level-0 of the classified taxonomy and presented in blue color. In a similar way, the four major dimensions are representing as level-1 in light-green color and so onwards.

### B. TAXONOMY

In this paper, we studied the papers that are predominantly based on either ML, blockchain, or both. The presented survey has been divided into four dimensions: Goal Oriented, Layer Oriented, Countermeasures, and Smart Applications. Figure 3 shows the taxonomy of ML adoption for blockchain-based smart applications.

## IV. GOAL ORIENTED

In this dimension, research work has been included from the following aspects: (i) preventing, (ii) predicting, (iii) monitoring, (iv) detection, and (v) response of blockchain. For instance, [29] discussed the detection of intrusion in a collaborative manner, [30] predicts bitcoin prices using Bayesian regression method and [31] monitors the blockchain network in a distributed manner.

### A. DETECTION

This subsection includes the detection of the attacks to handle data security issues. Meng *et al.* [12] has proposed a Collaborative IDS (CIDS) with a data-sharing agreement on a blockchain-based environment. Data privacy issues of CIDS can be addressed with the use of ML classifiers on BT. These classifiers run locally by the data owner, and the results shared with other users within the network. To handle the trust computation issue in CIDS, Alexopoulos *et al.* [29] proposed an approach to handle the alerts and data distribution among the participants. Then, it verifies the data and executes a suitable consensus algorithm to add a new data block within the blockchain network. These data alerts could be encrypted with keys distributed to selected parties. Another approach was to keep them on a separate blockchain while still being part of the network. Further, an anomaly detection system (ADS) was implemented based on the assumption that a similar kind of attack may occur, but on different nodes in the blockchain network [32]. This system does not discard information about orphans and forks that are usually done by other ADS. The attacked nodes shared this information with other neighbor's node within the network. In the experimental analysis, the system successfully prevented the same kind of attack (on other nodes) with negligible overhead.

### B. MONITORING

A blockchain monitoring system has been prototyped using Self Organising Maps (SOMs) [33]. It has modeled the dataset without external control. Here, large vectors map to smaller and lower dimension vectors using SOMs. The system analyzed blockchain data using Kohonen and SOMbrero libraries (in the R programming environment). The result shows the effective key attributes monitoring of the blockchain nodes and finds the emerging patterns. More, another approach included the distributed pattern recognition system with the concept of Graph Neurons (GNs) [31] to monitor the blockchain system. GNs are scalable and could recognize patterns from similar or incomplete patterns. The GN communicated with adjacent nodes to detect events within the network by using input data. Preliminary results showed object detection by the GN was accurate; still, further work is needed in the area.

### C. PREDICTION

The ML models are mostly used for prediction. A good prediction model helps to make the right decision making and analytics. Added to this, an ML model to predict the bitcoin prices has been proposed by Valenkar *et al.* [30]. This model uses bayesian regression and random forest with several features such as block size, total bitcoins, day high, number of transactions, and trade volume. The trained dataset was normalized using log, z-score, and box-cox normalization techniques. Further, a price prediction study has been done for several cryptocurrencies such as Ripple, Litecoin, Dash, Bitcoin, and Ethereum cryptocurrencies [34]. Here, correlation matrices for feature selections used and reported the general trends within the network. The proposed model used multiple regression techniques on bitcoin. The system showed the 0.9944 accuracies for price prediction.

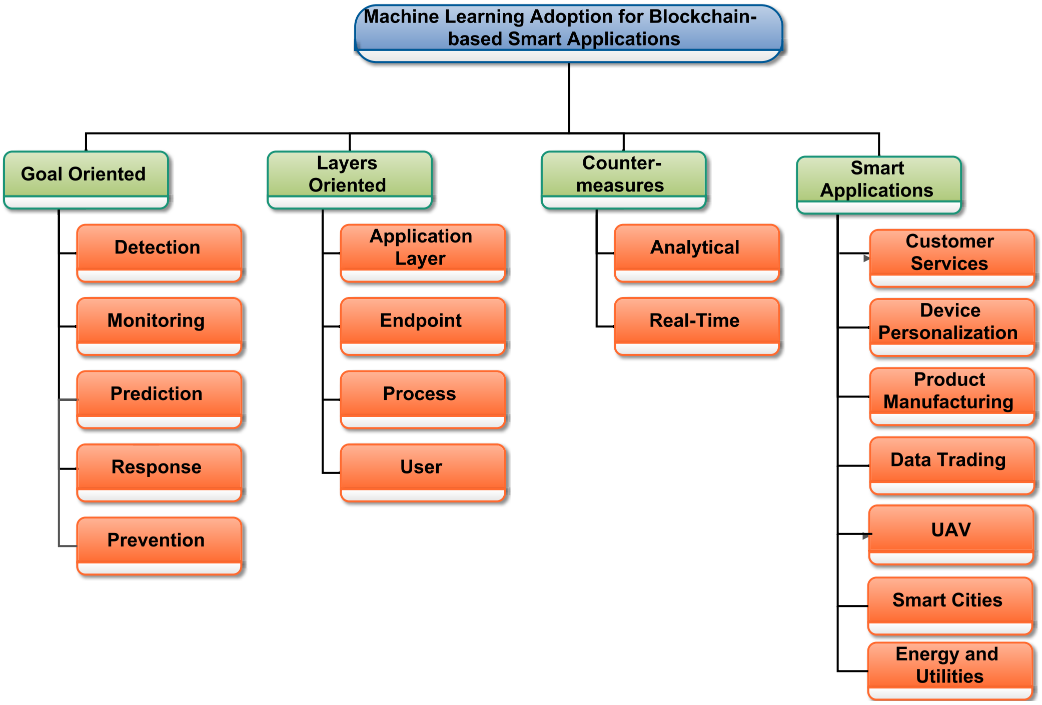


**FIGURE 3.** Proposed Taxonomy for Machine Learning adoption in Blockchain-based Smart Applications.

## D. RESPONSE

Tsolakis *et al.* [35] presented an approach for the secure exchange of energy data between a cluster of end-users and a Virtual Node. This approach uses a blockchain-based architecture to give a demand response solution. This solution included Fog-Enabled Intelligent Devices (FEID) [36] at the user side (act as a blockchain node) and managed smart contracts with the energy producer [36] in cloud computing platform [37], [38]. Then, a group of energy users combined into a centralized Virtual Node. All predictions and forecasting about the energy consumption have done on the Virtual Nodes. The end-user learned itself from each experience that helps to improve the accuracy of data passed as the next iteration to the virtual node.

## E. PREVENTION

One of the based uses of ML in BT based application is the prevention of an issue; for instance, employee verification by a future employer. A mechanism proposed to reduce the time taken to verify work history details provided by a prospective employee [39]. The previously worked organization has to compile details of an employee (date of joining, leaving, post, etc.), and a public key is generated to encrypt the data. Subsequently, a new smart contract is created for the employee. Then, the address of the smart contract is entered into the organization's database, and later it verified by the future employer. In a future organization, work history is fetched from the smart contract of employees and decrypted using the key. Then, each work history entry compared with the database of the previous organization. This ensures consistency and integrity of data as well as the authenticity of the sender.

A similar idea to tracking student's learning history has been proposed [40]. The main aim was to maintain the learning histories of the students and ensure access control, privacy, and security. To ensure privacy and access control, three contracts were suggested: Learner - Learning Provider Contract (LLPC), Registrar - Learning Provider Contract (RLPC), and Index Contract (IC). The role of RLPC is to specify all conditions that describe request access by the provider within the network. LLPC shows that a learner's data exists on the provider's database. The IC map the learners and providers and the corresponding learning history. Here, specialized Learning Blockchain Application Program Interfaces (LB APIs) has been proposed to encourage the adoption of this technology. Initially, the system is started with a boot node in the network. A new provider is added to the network as per the conditions described in RLPC. A new learner account is created and noted LLPC. In case, the new provider wants access to learners' data, it has to request the learner using the LLPC, and upon granting, data will be accessible.

## F. COMPARISON OF EXISTING APPROACHES FOR GOAL ORIENTED DIMENSION

A detailed comparison of the approaches discussed above is shown in Table 2. This comparison of works included several parameters such as trust management, ML, algorithm, blockchain, detection, merits, and demerits.

**TABLE 2.** Comparative analysis of goal oriented approaches.

| Authors | Year | Objective | Merits | Demerits | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|
| Meng *et al.* [12] | 2018 | To present use of blockchain in Intrusion Detection | Scope of application of blockchain was discussed | Discusses only data sharing and trust management issues of collabrative intrusion detection | $X$ | ✓ | ✓ | ✓ | ✓ |
| Alexopoulos *et al.* [29] | 2018 | Improvement in trust and accountability in Collaborative Intrusion System using blockchain | Design considerations thoroughly explained | Use of ML not covered | $X$ | ✓ | ✓ | $X$ | ✓ |
| Signorini *et al.* [32] | 2018 | to provide a blockchain based anomaly detection system | extremely less bandwidth overhead of about 0.248% | Only basic testing carried out consisting of only 3 node | ✓ | ✓ | ✓ | $X$ | $X$ |
| Chawathe *et al.* [33] | 2018 | self-organizing maps to monitor blockchain data in real-time | Running time for the number of iterations, attributes and instances shows a linear relationship | Security issues are not highlighted in storing bitcoin data in a database | ✓ | ✓ | $X$ | $X$ | $X$ |
| Hudaya *et al.* [31] | 2018 | distributed pattern recognition for event monitoring within IoT-blockchain network | use of graph neuron approach enables distributed pattern recognition | integration of IoT-blockchain not covered in depth | ✓ | ✓ | $X$ | ✓ | $X$ |
| Saad *et al.* [34] | 2018 | predicting bitcoin prices with high accuracy using ML | user activity dynamics and its effects discussed extensively | not enough focus on other factors other than supply demand | ✓ | ✓ | $X$ | $X$ | $X$ |
| Tsolakis *et al.* [35] | 2018 | demand response system using blockchain | security concerns addressed appropriately | ML aspects not covered | ✓ | ✓ | $X$ | $X$ | $X$ |
| Sarda *et al.* [39] | 2018 | To prevent work-history related frauds using blockchain | Extensive use of blockchain to share and verify work history | No intersection of blockchain and ML was found, discussion on security issues was not carried out | $X$ | ✓ | $X$ | ✓ | $X$ |
| Liang *et al.* [41] | 2019 | Present micro-blockchain based dynamic intrusion detection | Scope of application of blockchain was discussed | Discusses attacks and trust management issues only | $X$ | ✓ | ✓ | ✓ | ✓ |
| Sgantzos *et al.* [42] | 2019 | Present genetic algorithm implementation on blockchain | Discussed framework and future applications | Natural Language interaction with blockchian based system need to be explored | ✓ | ✓ | ✓ | $X$ | ✓ |

Parameters- 1:ML, 2:Blockchain, 3:Detection, 4:Algorithm, 5:Trust Management
Notations- ✓: Considered, and $X$: Not Considered.

## V. LAYERS ORIENTED

At this moment, BT is very young, and understanding of layer division is not enough. Considering the BT as a solitary layer is similar to enduring everything among the physical layers and transport layers into a single layer. While separating the BT into multiple layers, we can understand various properties of BT which are needed to be implemented. Some of the properties are: (i) Security: Nodes that do not control rare resources (commonly computing power) majorly cannot convince others for a different version of the ledger. (ii) Liveness: Here, new blocks can be added to the blockchain with suitable latency. (iii) Stability: Nodes within the blockchain network should not amend their belief of the consensus ledger except rare cases. (iv) Accuracy: Blocks added to the ledger must signify valid transactions such as they imitate to a description of how new blocks relate to previous blocks. This dimension discusses the adoption of ML techniques in various layers of a blockchain network, such as endpoints, application layer, process, network (intrusion), and finally, user level. The proposed categorization into layers recognizes each property. Security is accomplished at each layer, Liveness is realized at the user layer, Stability is achieved at the network and application layer, and accuracy is achieved at the endpoints, wherever blocks have significance.

### A. APPLICATION LAYER

The application layer helps to accomplish the security, liveliness within the network. An approach to detect malware (in the form of a portable executable file) has been proposed using DL methods [43]. The portable executable file was converted to a grayscale image to feed to the Deep belief network. In this solution, the receiving node broadcast the file to other nodes within the network. The file executed locally in the local detection model of each node. The probability value of being malware is appended with the block. A weighted average for the trust of the node is applied to the appended probabilities. The weighted average calculated on the condition that the file is malware or not. The model used Restricted Boltzmann Machines (RBM) techniques with 3000 hidden units. The results showed good accuracy for malware detection.

### B. ENDPOINT

Endpoints included the small computing devices which are participating in the BT based application network. To handle the issue of fast computing and reduction in computation power, emergent technology Edge Computing comes into the picture. The Edge Computing Service Provider (ECSP) can meet the requirements, and an approach has been proposed to maximize the revenue of ECSP [44]. The ECSP allocates the resource unit to the highest bidder who participated in the bidding process.

Similarly, Kim *et al.* [45] proposed an architecture to offload resources computations in Deep Neural Networks to edge computing. Within this architecture, the embedded device (the computation device at the edge server) and the edge server had to make initial deposits. After that, the embedded device sends the computations to edge servers. Then, the edge server returns the results, which are verified by other nodes within the network. The server is rewarded for its work after validation. The typical Ethereum architecture had been modified in this implementation by replacing Ethereum virtual machine by V8 javascript virtual machine. The blocks generation part had been separated from the virtual machine to ensure that smart contract execution cannot

**TABLE 3.** Comparative analysis for Layer-Oriented approaches.

| Author | Year | Objective | 1 | 2 | 3 | 4 | 5 | Pros | Cons |
|---|---|---|---|---|---|---|---|---|---|
| Raje *et al.* [43] | 2017 | To design a decentralized firewall using blockchain and deep learning | ✓ | X | ✓ | X | X | Discussed Deep Belief Networks and training procedure in detail and explored various architectures | Number of devices in experimental setup was small and working of blockchain not discussed |
| Portnoff *et al.* [46] | 2017 | To identify human traffickers using classified ads and bitcoin transactions | ✓ | X | X | ✓ | ✓ | A detailed discussion on data analysis | Bitcoin transaction linking was not perfect, false positives were high, blockchain aspect not covered |
| Wasim *et al.* [47] | 2017 | Proposed a Law as a service architecture to monitor Contract Breaches and issue injunctions | ✓ | X | X | X | ✓ | Results and system model defined clearly, proposed unsupervised ML algorithm | Blockchain not included in the model description |
| Luong *et al.* [44] | 2018 | Edge computing usage for mining applications and ML for resource allocation | ✓ | ✓ | ✓ | X | X | ML algorithms thoroughly discussed, well-documented results | Focus on optimizing revenue but no optimal allocation of resources |
| Kim *et al.* [45] | 2018 | Proposed deep neural networks architecture for blokchain based edge computing application | ✓ | ✓ | ✓ | X | X | Executes complex programs on ethereum blockchain using virtual machines | Experimental results not discussed in detail, Deep Neural Network applications not covered |

Parameters- 1:Blockchain, 2:Edge Computing, 3:Neural Network, 4:Supervised Learning, 5:Unsupervised Learning
Notations-✓: considered, and *X*: not considered.

affect the rate of generation of blocks. The system performance shows better results compare to the state-of-art approaches.

### C. PROCESS

This subsection comprises ML techniques to the process involves in the blockchain-based applications. An ML classifier has been used to detect the human traffickers using adult classified ads [46]. It included ads dataset from Backpage (a website for online classifieds). It used a supervised learning model and used logistic regression for classification. The labeled dataset with identifiers like email and phone number has been used for testing. This approach takes two ads at a time as an input to the model. If the ads are from the same author, the output will be same. The trained model showed an 89.54% true positive rate. Further, this approach used to build a graph and found links between ads and their related bitcoin transactions. This approach showed a high false-positive rate.

### D. USER

Users are the individual or system that uses the system functionality seating at one end of the application. Wasim *et. al* [47] proposed law as a service ML-based architecture to monitor contract breaches. Previously, contract breaches were dealt with or without the use of courts. The proposed system used an unsupervised ML algorithm called Probability-based Factor Model to issue injunctions. This model simulated using three service providers Redis [48], MongoDB [49], and Memcached Servers [50]. Services monitor the contracts for breaches. The results showed that services that perform complex operations are more likely to breach contracts.

### E. COMPARISON OF EXISTING APPROACHES FOR LAYER ORIENTED USE CASES

A detailed comparison of the approaches is discussed in Table 3. This comparison is made based on several parameters such as edge computing, blockchain, neural network, supervised learning, pros, and cons of the proposed approach.

## VI. COUNTERMEASURES

The ML approached in response to a threat or monitoring a system can be classified as real-time analytics or analysis based on historical data [33]

### A. ANALYTICAL

In today's world, every enterprise generates huge amounts of data from different sources such as social media, smartphones, IoT, and other computing devices. These data are tremendously valuable to organizations. The overall technique to find a meaningful pattern from these data is called data analytics. It is a process to convert data from foresight to insight. It describes what happened in the past, draw awareness about the present, and make predictions (with some ML techniques) about the future. ML techniques can be categories as supervised, unsupervised, and reinforcement learning to analyses the data. The supervised learning included classification and regression originated on the idea of example-based learning. Next, unsupervised learning techniques perform clustering, dimensionality reduction, and recommendation of a system based on the dataset. This used to recognize hidden patterns or focus on the well-educated behavior of the machine. The reinforcement learning approach helps to reward maximization. In the next decades, our society will be driven by new technological developments in ML and BT. Further, a traceability algorithm has been proposed for bitcoin mining by using the ANN approach [51]. It aimed to remove irrelevant data in mining and introduce traceability to the system. This was helpful in a distributed architecture and decrease traceability time.

DL is a subset of ML that is originated on definitive algorithms influenced by the overall structure and work of the neural network in the human brain. DL rationalizes tasks to perform speech recognition, image recognition, make insightful decisions on natural language processing. DL network consisted of input, output, and hidden multi-layer. It accepted a new block and a history of the previous block as an input in BT based applications. It used a state-transition algorithm on the features such as hash value, nonce, address, and transaction data. An ML-based classification approach is used to identify the cyber-crime in Bitcoin [52]. The results provided a glimpse of the size of cybercrime in the Bitcoin ecosystem.

### B. REAL-TIME

ML enables real-time analytics for all types of data, such as social - accessible, transactional, and operational. It uses in-memory computing though leaving unremittingly updated data securely. It improves analytics accuracy and accelerates the predictive behavior of ML models. It has been noticed that BT makes real-time cross edge transactions in monetary and payment frameworks. Several fintech innovators and banks are currently investigating blockchain due to fast and real-time settlement of massive amounts independent of geographic hindrances. Similarly, associations that require real-time analytics of information on a huge scale can approach an ML and blockchain-empowered framework to accomplish the goal. With ML and BT, financial institutions and other associated organizations across the globe can trace the data changes to make quick and fast business decisions regardless of irregular activities or suspicious transactions. Moreover, the performance of the blockchain-based Software-Defined Vehicular Network (SDVN) has been improved using deep Q-learning methodology [53], [54]. This approach used a permissioned blockchain with Byzantine fault tolerance as a consensus protocol. Results showed that this scheme managed network and computing resources better and gave the best throughput in the SDVN [55]. Further, Liu et al. [22] recommended a data collection framework for Industrial-IoT (IIoT) applications. This framework combines the use of deep reinforcement learning (DRL) and Ethereum blockchain. To store and share data, Ethereum nodes were categories into two categories: mining nodes and nonmining nodes. The proposed DRL algorithm used three plain components, actions (moving path and remoteness of mobile terminals), states (environment description), and rewards (amount achieved). The proposed algorithm shows the 34.5% increase in geographical fairness compared to a random solution.

## VII. SMART APPLICATIONS

The last dimension covers the adoption of ML techniques to BT-based smart applications such as data trading, UAV, product manufacturing, Medical, and Healthcare [56], Smart Cities, automated Customer Service, and Device Personalization as shown in Figure 4. ML and BT are revolutionizing up-to-date technologies by transforming customer experiences, behaviors, and business models [57]. Both are making strides in the major smart application:

### A. CUSTOMER SERVICE

With the upsurge of the customer, the Customer service has to be more efficient and automated to meet rising customer needs. One of the best ways is to automate the process to increase a company's capabilities. Wang *et al.* [58] proposed an AutoML framework for the blockchain-based application. It consists six layers: (i) Organization Layer (Includes entities such as shops, malls, online stores), (ii) Customer Layer, (iii) Application Layer (applications provide services to the consumers), (iv) Data conversion layer(Data masking and unmasking), (v) AutoML layer (Consists various ML models such as Linear regression, SVM and logistic regression for predictions), (vi) Blockchain Layer (Secure storage of data and ensures the safe data exchange) This framework helps organizations to keep their data safe, automate their processes and share data with other organizations in a mutually beneficial and safe way.

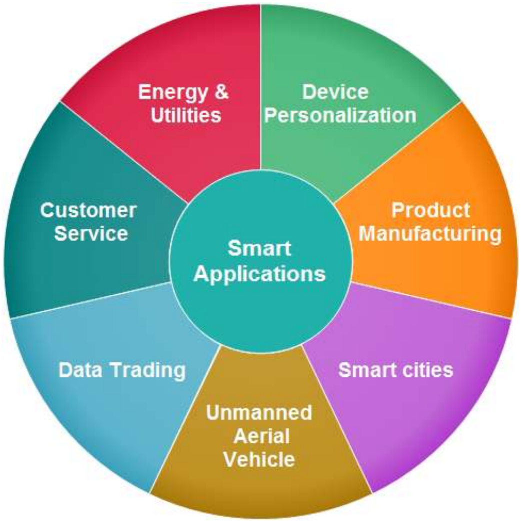


FIGURE 4. A list of smart applications.

### B. DEVICE PERSONALIZATION

Device Personalization (DP) is a component that uses predictions across smart devices to improves the quality of service (QoS) such as actions in the launcher, smart text selection during writing on text pad. DP Services uses system permissions to provide smart predictions. In a smart home environment, and ML model-based single DP framework has been implemented [59]. Here, smart devices connected to a smart hub. Every time a user uses a device, a log is generated with user data, device data, and other parameters for that device. For example, uses of air conditioner at home. This log data help to adjust the operation of the device.

### C. PRODUCT MANUFACTURING

As a feature of the manufacturing process, organizations have started trusting blockchain-based procedures to empower production, security, transparency, and compliance checks. Instead of following fixed schedules of machine maintenance, ML algorithms are being used to make flexible plans at specific periods. Product testing and quality control have also automated increasingly, with versatile algorithms. It successfully detects faulty and good products, particularly in profoundly delicate situations. For example, Porsche (a car manufacturing company) is one of those early adopters of ML and BT technologies; to improve automobile safety and increase capabilities. The organization utilizes blockchain innovation to transfer data more safely and rapidly, offering its clients peace of mind, regardless of parking, charging, and third-party access to their car.

### D. DATA TRADING

DataTrading is an innovative platform that makes advanced trading possible for retail traders from all over the world. An Ethereum based data trading framework has presented that succeeded in preventing single-point failure and preserving privacy at the same time [10]. The framework consists of three entities, data provider, a data consumer, and a market manager. Once the network is set up, the data consumer and data provider registered themselves with the market manager. The data provider needed to deposit an amount with the manager greater than the amount to be paid by the consumer. A list indicating topics of data available with the provider is published. The client referred it and requests some encrypted data blocks for content validation using a distance metric learning technique. After successful validation, the client responded, and the provider sent a signature, and the client published a smart contract to the network. The provider then sent another signature, and these two signatures, the client, decrypted the data. There are different protocols such as setup protocol, register protocol, payment and acknowledgment protocol, and query protocol that facilitate the communication between provider, consumer, and the manager.

### E. UNMANNED AERIAL VEHICLE

A UAV or drone is an aircraft that runs without a pilot (human) aboard. Kuzmin *et al.* [9] has proposed blockchain-based UAVNet model includes different devices such as a network of satellites, cellular base stations, and ground control stations. Here, BT served multiple purposes, such as preserving the integrity of the data and for distributed graphs computation. The communication between satellites and base stations is prone to electromagnetic jamming, hence a blockchain-based system enables UAVs to store relevant coordinate's data and operate autonomously within the jamming zone. This solution used a proof-of-graph consensus algorithm with a simplified memory bounded algorithm of any existing shortest path to validating a new transaction.

### F. SMART CITIES

Smart cities improve the living experiences of individuals. ML and blockchain emergent technologies play a crucial role in the innovation of smart cities to provide critical services and infrastructure components such as healthcare [60], city administration, smart homes, education, transportation, real estate, and utilities [61]. (i) Smart Homes: Smart homes can be monitored, and DP can help to improve the quality of livelihood. (ii) Smart Parking System: Here, arrival and departure of vehicles can be tracked for different parking lots available in a decentralized manner within the smart city [61]. Consequently, smart parking lots should be designed by seeing the number of cars in each region. Furthermore, new parking lots must be recognized spontaneously to benefits in dealers and vehicle owner's daily life. (iii) Smart Weather and Water Systems: Here, the system can use some sensors to generate appropriate data such as temperature, wind speed, rain, and pressure. The analysis of these data through ML techniques can contribute to improving the density of smart cities. (iv) Smart Vehicular Traffic: Vehicular traffic data with a suitable analysis will benefit the government and citizens to a great extent [62]. Everyone can decide the arrival time to a destination by using these data. (v) Surveillance Systems: Physical security is an utmost important concern for citizens anywhere they live. To address this issue, smart technology such as ML and BT can be configured for it. Consequently, collecting and analyzing data and identifying crimes is one of the challenging tasks. (vi) Smart Healthcare: It includes the accessibility of the caregivers and doctors, identification of nearest medical stores, and clinic. Hence, ML and BT play a vital role while data is getting generated [63], [64]. (vii) Smart Governance and Smart Education: Smart governance can maintain a city smartly. A smart city includes a different way of education. It captures the data of the students and employees in educational and government institutions. Predictions and analytics required to keep the stuff up to standards.

### G. ENERGY AND UTILITIES

In the Energy industry, BT is assisting to simplify energy exchanges. For example, IOTA enables smart transformation across the entire energy industry by implementing BT [65]. It uses the concept of peer-to-peer (P2P) energy production and consumption. Smart energy microgrids are progressively developing fame as a technique of making ecological energy resources. LO3 Energy (a Newyork-based organization) is also using a blockchain-based development to develop energy generation, transmission, conservation, and exchanging within neighborhood networks. The technology uses microgrid and smart meters, together with smart contracts, to manage and track energy transactions. GE Digital, together with Evolution Energie (a startup), has created an application that encourages the tracking of renewable energy in SG and uses blockchain to give energy sources certificate. The idea is allowing organizations and individuals to trade renewable energy sources without the involvement of third-party [36].

Globally, most of the industries are dogged by third-party, who increase the business cost. BT has already interrupted that model by facilitating the P2P model for customers. The revolution works even better when combined with the ML and BT together. We have been witnessed the financial revolution by using the cryptocurrency world; similarly, we are going to observe more innovation disruptions as a result of combining ML and BT.

## VIII. FUTURE RESEARCH ISSUES AND CHALLENGES

Researchers are looking forward to these technologies across the globe, but still, various obstacles resist the integration of BT and ML [66]. Their integration is still in its infancy.

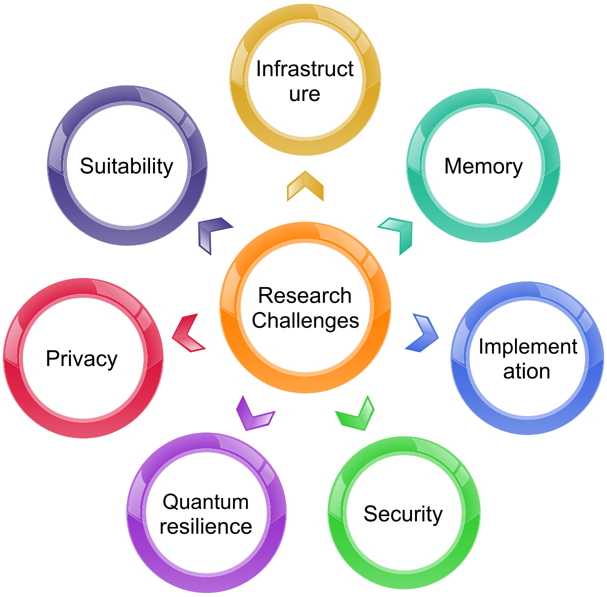


**FIGURE 5. Future research issues and challenges.**

Many open issues and challenges are yet to be addressed. Here, we discuss the futuristic open issues and challenges of ML adoption in BT for secure communication, as shown in Figure 5. We highlighted the challenges as Suitability, Infrastructure, Privacy, security, Memory, Implementation, and Quantum resilience.

### A. SUITABILITY

Blockchain is a viable solution if the source of data cannot be trusted, and several entities are high in the distributed environment. If performance is required, then a simple database is a better option. Therefore, the architecture of blockchain must be understood before its use in any application [67].

### B. INFRASTRUCTURE

Blockchain specific hardware and network infrastructure improve the performance of many blockchain-based applications. These could include network administration, mining hardware, decentralized storage, and communication protocols [30]. However, products tailored to use in blockchain are still under investigation (involving major tech companies and financial institutions).

### C. PRIVACY

Data generated by devices to be stored on the blockchain is available to the entire blockchain nodes [33]. This leads to a potential privacy concern for data that needs to be kept either private or confidential. Such issues could be resolved by the use of private blockchains, controlled access, and encryption. However, the ML models adoption on these limited data imposed barriers for predictions and analytics.

### D. MEMORY

The size of blockchain keeps growing as new blocks are added to it. Consequently, the entire chain must be stored by all nodes, which creates a significant memory constraint on these devices. The performance is also affected by an increase in chain size. Besides that, the storage of irrelevant data also wastes computational resources. Blockchains are immutable, and hence storage management is a major issue in most implementations.

### E. IMPLEMENTATION

A large- scale blockchain network will require an equivalently large number of transactions. Implementation of this size of blockchain invites potential issues; for instance, high demand for internet bandwidth, which not easy to reduce, transactions will be a burden on the network. Hence, the addition of blocks and transactions needs to be decreased to meet the inevitable demand.

### F. SECURITY

Blockchains are decentralized, and they are prone to security issues [68]. The most common concern is that due to attacks, the consensus protocol may be compromised, such that the mining power of a few farms will control which blocks are added to the network. This particular concern is present in public blockchains. Private versions are unaffected by this attack as they have each node identified, and an appropriate consensus protocol is in place.

### G. QUANTUM RESILIENCE

The hashing algorithms used by blockchains could soon be broken with quantum computers. A blockchain is predominantly at risk from this because it uses one-way functions for encryption (only protection in digital signature). This would cause all the features that make blockchains a viable storage structure (obsolete). Luckily, quantum computing offers opportunities to boost the performance and security of blockchains. Quantum cryptography can reinforce the security of the blockchain network as quantum communication is authenticated inherently (users cannot mimic another user). It can encrypt entire P2P communications and replace classical digital signatures in the blockchain network. Research is underway to design blockchain with quantum computing [69].

## IX. CASE STUDY

In order to demonstrate the proposed architecture for ML adoption in blockchain-based smart applications, we present a case study, as shown in Figure 6. In this case study, we study a blockchain-based SG system for energy transactions using cryptocurrency [70], [71]. This framework uses a blockchain and DL approach to complete an energy transaction. It is a reliable P2P energy system. It is based on the Byzantine fault tolerance algorithm to produce high system throughput. This approach consists of five phases (i)setup phase, (ii) agreement phase, (iii) consensus-making phase, (iv) block creation phase, and (v) a change view phase. Here, blocks are generated using hash functions and a short signature. It is an IDS that works on recurrent neural networks to detect

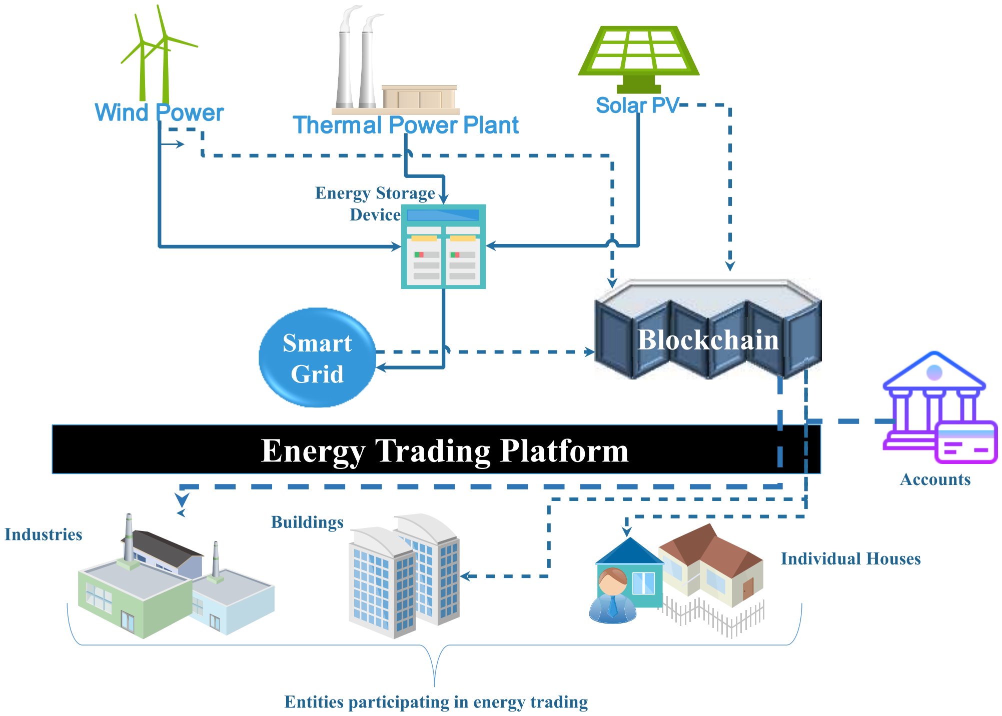


**FIGURE 6.** An energy trading system.

fraudulent transactions and network attacks in blockchain-based energy applications. The performance of this IDS has been studied on different energy datasets. It consists of four entities for communications: (a) Energy Buyer, (b) Energy Seller, (c) Blockchain, and (d) IDS [41]. Energy buyer trade energy with the energy seller. An energy buyer shows that he has sufficient energy money (cryptocurrency) that fulfills the minimum asset requirement of energy sellers. An energy buyer will be a person or commercial building or industry. Energy buyers can be located in a Home Area Network (HAN), Building Area Network (BAN), or Neighborhood area Network (NAN). Energy seller entities demonstrate that it has sufficient energy to sell. Energy sellers can be located in HAN, BAN, NAN, or energy companies. Energy seller (entities) produces energy from renewable sources such as wind energy, solar energy, and biomass. The seller can be a neighbor, local society with renewable energy resources, utility provider, or SG. Blockchain entity is a distributed digital ledger, which is encompassing all energy transactions in the SG system.

Once an energy seller produces energy from renewable energy sources and uses that energy for his own use. After that, if the seller left with energy, then he publishes his per-unit prices per kilowatt energy on the blockchain network. In this P2P energy trading system, the needed energy buyer will look upon the published price and unit of energy. First of all, the energy buyer checks his account that he has sufficient cryptocurrency balance for energy trading. Then, if the energy demand of energy buyers and per-unit price matched his requirement of energy at that moment, he sends a purchase request to the energy seller through the blockchain network. After validation of the buyer on the blockchain network seller sells the energy to the buyer and receives cryptocurrency in his account. This P2P energy trading system required further data analysis to identify the frequent buyer and seller for the recognition of malicious transactions within the network. A malicious block or node can impact the P2P system and compete for energy trading. IDS can help to figure out the deceitful transactions and network attacks on energy trading applications [72]. The result shows the good performance of the system compares to the state-of-art approaches.

## X. CONCLUSION

The recent advancements in Blockchain and ML have made them path-breaking technologies. The distributed ledger has the possibility to work as the backbone of various smart applications such as smart cities, UAV, SG, data trading. In this paper, we have presented detailed information on BT and ML, along with their usages in smart applications and proposed an ML-BT based architecture. This architecture can be used to design and deploy an ML-BT based data analysis system. A discussion and comparison of various existing surveys are presented. Then, we presented ML-BT solution taxonomy, focusing on goal oriented, layer oriented, countermeasures, and smart application dimensions. A comparative analysis of available methodologies and approaches is presented in each dimension. Then, we have listed several research challenges being faced during ML adoption in BT-based systems, which require solutions. We also emphasized a number of research prospects such as infrastructure availability, quantum resilience, and privacy issues that can serve as future research directions in this field. Then, we presented a case study on the energy trading system to verify the effectiveness of the proposed architecture and concluded the paper at last.

## REFERENCES

[1] S. Kaneriya, J. Vora, S. Tanwar, and S. Tyagi, "Standardising the use of duplex channels in 5G-WiFi networking for ambient assisted living," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.

[2] S. Tyagi, M. S. Obaidat, S. Tanwar, N. Kumar, and M. Lal, "Sensor cloud based measurement to management system for precise irrigation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[3] U. Bodkhe, P. Bhattacharya, S. Tanwar, S. Tyagi, N. Kumar, and M. S. Obaidat, "Blohost: Blockchain enabled smart tourism and hospitality management," in *Proc. Int. Conf. Comput., Inf. Telecommun. Syst. (CITS)*, Aug. 2019, pp. 1–5.

[4] S. Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2009. [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[5] M. Rahouti, K. Xiong, and N. Ghani, "Bitcoin concepts, threats, and machine-learning security solutions," *IEEE Access*, vol. 6, pp. 67189–67205, 2018.

[6] A. P. Namanya, A. Cullen, I. U. Awan, and J. P. Disso, "The world of malware: An overview," in *Proc. IEEE 6th Int. Conf. Future Internet Things Cloud (FiCloud)*, Aug. 2018, pp. 420–427.

[7] S. Krit and E. Haimoud, "Overview of firewalls: Types and policies: Managing windows embedded firewall programmatically," in *Proc. Int. Conf. Eng. MIS (ICEMIS)*, May 2017, pp. 1–7.

[8] G. Betarte, E. Gimenez, R. Martinez, and A. Pardo, "Improving Web application firewalls through anomaly detection," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 779–784.

[9] A. Kuzmin and E. Znak, "Blockchain-base structures for a secure and operate network of semi-autonomous unmanned aerial vehicles," in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat. (SOLI)*, Jul. 2018, pp. 32–37.

[10] Y. Zhao, Y. Yu, Y. Li, G. Han, and X. Du, "Machine learning based privacy-preserving fair data trading in big data market," *Inf. Sci.*, vol. 478, pp. 449–460, Apr. 2019.

[11] C. Pop, T. Cioara, M. Antal, I. Anghel, I. Salomie, and M. Bertoncini, "Blockchain based decentralized management of demand response programs in smart energy grids," *Sensors*, vol. 18, no. 1, p. 162, Jan. 2018.

[12] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: A review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018.

[13] M. Conti, E. S. Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3416–3452, 4th Quart., 2018.

[14] D. Ucci, L. Aniello, and R. Baldoni, "Survey of machine learning techniques for malware analysis," *Comput. Secur.*, vol. 81, pp. 123–147, Mar. 2019.

[15] K. Salah, M. Rehman, N. Nizamuddin, and A. Al-Fuqaha, "Blockchain for AI: Review and open research challenges," *IEEE Access*, vol. 7, pp. 10127–10149, 2019.

[16] F. Casino, T. Dasaklis, and C. Patsakis, "A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telematics Inform.*, vol. 36, pp. 55–81, Mar. 2018.

[17] W. Xiong and L. Xiong, "Smart contract based data trading mode using blockchain and machine learning," *IEEE Access*, vol. 7, pp. 102331–102344, 2019.

[18] T. M. Mitchell, *Machine Learning*, 1 ed. New York, NY, USA: McGraw-Hill, 1997.

[19] P. Louridas and C. Ebert, "Machine learning," *IEEE Softw.*, vol. 33, no. 5, pp. 110–115, May 2016.

[20] R. Saravanan and P. Sujatha, "A state of art techniques on machine learning algorithms: A perspective of supervised learning approaches in data classification," in *Proc. 2nd Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Jun. 2018, pp. 945–949.

[21] V. Vats, L. Zhang, S. Chatterjee, S. Ahmed, E. Enziama, and K. Tepe, "A comparative analysis of unsupervised machine techniques for liver disease prediction," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2018, pp. 486–489.

[22] C. Liu, X. Xu, and D. Hu, "Multiobjective reinforcement learning: A comprehensive overview," *IEEE Trans. Syst., Man, Cybern, Syst.*, vol. 45, no. 3, pp. 385–398, Mar. 2015.

[23] Z. Zheng, S. Xie, H. Dai, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *Proc. IEEE Int. Congr. Big Data, Big Data Congr.*, Jun. 2017, pp. 557–564.

[24] G. B. Mermer, E. Zeydan, and S. S. Arslan, "An overview of blockchain technologies: Principles, opportunities and challenges," in *Proc. 26th Signal Process. Commun. Appl. Conf. (SIU)*, May 2018, pp. 1–4.

[25] W. Penard and T. van Werkhoven, "On the secure hash algorithm family," *Cryptogr. Context*, pp. 1–18, 2008.

[26] M. Vukolić, "The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication," in *Open Problems in Network Security*, J. Camenisch and D. Kesdoğan, Eds. Cham, Switzerland: Springer, 2016, pp. 112–125.

[27] S. Thompson, P. L. Seijas, and D. Adams, "Scripting smart contracts for distributed ledger technology," Tech. Rep., Dec. 2016. [Online]. Available: https://kar.kent.ac.uk/61162/

[28] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[29] N. Alexopoulos, E. Vasilomanolakis, N. R. Ivánkó, and M. Mühlhäuser, "Towards blockchain-based collaborative intrusion detection systems," in *Critical Information Infrastructures Security*, G. D'Agostino and A. Scala, Eds. Cham, Switzerland: Springer, 2018, pp. 107–118.

[30] S. Velankar, S. Valecha, and S. Maji, "Bitcoin price prediction using machine learning," in *Proc. 20th Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2018, pp. 144–147.

[31] A. Hudaya, M. Amin, N. M. Ahmad, and S. Kannan, "Integrating distributed pattern recognition technique for event monitoring within the iot-blockchain network," in *Proc. Int. Conf. Intell. Adv. Syst. (ICIAS)*, Aug. 2018, pp. 1–6.

[32] M. Signorini, M. Pontecorvi, W. Kanoun, and R. Di Pietro, "BAD: Blockchain anomaly detection," 2018, *arXiv:1807.03833*. [Online]. Available: https://arxiv.org/abs/1807.03833

[33] S. Chawathe, "Monitoring blockchains with self-organizing maps," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 1870–1875.

[34] M. Saad and A. Mohaisen, "Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2018, pp. 704–709.

[35] A. C. Tsolakis, I. Moschos, K. Votis, D. Ioannidis, T. Dimitrios, P. Pandey, S. Katsikas, E. Kotsakis, and R. García-Castro, "A secured and trusted demand response system based on blockchain technologies," in *Proc. Innov. Intell. Syst. Appl. (INISTA)*, Jul. 2018, pp. 1–6.

[36] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and J. J. Rodrigues, "Fog computing for smart grid systems in the 5G environment: Challenges and solutions," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 47–53, Jun. 2019.

[37] J. Vora, P. DevMurari, S. Tanwar, S. Tyagi, N. Kumar, and M. S. Obaidat, "Blind signatures based secured e-healthcare system," in *Proc. Int. Conf. Comput., Inf. Telecommun. Syst. (CITS)*, Jul. 2018, pp. 1–5.

[38] S. Tanwar, J. Vora, S. Kaneriya, and S. Tyagi, "Fog-based enhanced safety management system for miners," in *Proc. 3rd Int. Conf. Adv. Comput., Commun. Autom. (ICACCA)*, Sep. 2017, pp. 1–6.

[39] P. Sarda, M. J. M. Chowdhury, A. Colman, M. A. Kabir, and J. Han, "Blockchain for fraud prevention: A work-history fraud prevention system," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 1858–1863.

[40] P. Ocheja, B. Flanagan, and H. Ogata, "Connecting decentralized learning records: A blockchain based learning analytics platform," in *Proc. 8th Int. Conf. Learn. Anal. Knowl.*, New York, NY, USA, 2018, pp. 265–269.

[41] H. Liang, J. Wu, S. Mumtaz, J. Li, X. Lin, and M. Wen, "MBID: Micro-blockchain-based geographical dynamic intrusion detection for V2X," *IEEE Commun. Mag.*, vol. 57, no. 10, pp. 77–83, Oct. 2019.

[42] K. Sgantzos and I. Grigg, "Artificial intelligence implementations on the blockchain. Use cases and future applications," *Future Internet*, vol. 11, no. 8, p. 170, 2019.

[43] S. Raje, S. Vaderia, N. Wilson, and R. Panigrahi, "Decentralised firewall for malware detection," in *Proc. Int. Conf. Adv. Comput., Commun. Control (ICAC)*, Dec. 2017, pp. 1–5.

[44] N. C. Luong, Z. Xiong, P. Wang, and D. Niyato, "Optimal auction for edge computing resource management in mobile blockchain networks: A deep learning approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[45] J.-Y. Kim and S.-M. Moon, ''Blockchain-based edge computing for deep neural network applications,'' in *Proc. Workshop Intell. Embedded Syst. Archit. Appl. (INTESA)*, New York, NY, USA, 2018, pp. 53–55.

[46] R. S. Portnoff, D. Y. Huang, P. Doerfler, S. Afroz, and D. McCoy, ''Backpage and bitcoin: Uncovering human traffickers,'' in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, New York, NY, USA, 2017, pp. 1595–1604.

[47] M. U. Wasim, A. A. Z. A. Ibrahim, P. Bouvry, and T. Limba, ''Law as a service (LAAS): Enabling legal protection over a blockchain network,'' in *Proc. 14th Int. Conf. Smart Cities, Improving Qual. Life Using ICT IoT (HONET-ICT)*, Oct. 2017, pp. 110–114.

[48] *Redis*. Accessed: Dec. 2019. [Online]. Available: https://redis.io/

[49] *Mongodb*. Accessed: Nov. 4, 2019. [Online]. Available: https://www.mongodb.com/

[50] *Memcached*. Accessed: Nov. 11, 2019. [Online]. Available: https://memcached.org/

[51] R.-Y. Chen, ''A traceability chain algorithm for artificial neural networks using T–S fuzzy cognitive maps in blockchain,'' *Future Gener. Comput. Syst.*, vol. 80, pp. 198–210, Mar. 2018.

[52] H. S. Yin and R. Vatrapu, ''A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning,'' in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 3690–3699.

[53] C. Qiu, F. R. Yu, F. Xu, H. Yao, and C. Zhao, ''Blockchain-based distributed software-defined vehicular networks via deep Q-learning,'' in *Proc. 8th ACM Symp. Design Anal. Intell. Veh. Netw. Appl. (DIVANet)*, New York, NY, USA, 2018, pp. 8–14.

[54] J. Bhatia, R. Dave, H. Bhayani, S. Tanwar, and A. Nayyar, ''Sdn-based real-time urban traffic analysis in vanet environment,'' *Comput. Commun.*, vol. 149, pp. 162–175, Oct. 2019.

[55] J. Bhatia, Y. Modi, S. Tanwar, and M. Bhavsar, ''Software defined vehicular networks: A comprehensive review,'' *Int. J. Commun. Syst.*, vol. 32, no. 12, p. e4005, 2019.

[56] J. Vora, A. Nayyar, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and J. J. P. C. Rodrigues, ''Bheem: A blockchain-based framework for securing electronic health records,'' in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6.

[57] S. K. Singh, M. M. Salim, M. Cho, J. Cha, Y. Pan, and J. H. Park, ''Smart contract-based pool hopping attack prevention for blockchain networks,'' *Symmetry*, vol. 11, no. 7, p. 941, 2019.

[58] W. M. Wang, H. Guo, Z. Li, Y. Shen, and A. V. Barenji, ''Towards open and automated customer service: A blockchain-based automl framework,'' in *Proc. 2nd Int. Conf. Comput. Sci. Appl. Eng. (CSAE)*, New York, NY, USA, 2018, pp. 27:1–27:6.

[59] K. Singla, J. Bose, and S. Katariya, ''Machine learning for secure device personalization using blockchain,'' in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2018, pp. 67–73.

[60] R. Gupta, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and B. Sadoun, ''Habits: Blockchain-based telesurgery framework for healthcare 4.0,'' in *Proc. Int. Conf. Comput., Inf. Telecommun. Syst. (CITS)*, Aug. 2019, pp. 1–5.

[61] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, M. Maasberg, and K.-K. R. Choo, ''Multimedia big data computing and Internet of Things applications: A taxonomy and process model,'' *J. Netw. Comput. Appl.*, vol. 124, pp. 169–195, 2018.

[62] A. Mewada, S. Tanwar, and Z. Narmawala, ''Comparison and evaluation of real time reservation technologies in the intelligent public transport system,'' in *Proc. 5th Int. Conf. Parallel, Distrib. Grid Comput. (PDGC)*, Dec. 2018, pp. 800–805.

[63] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, ''Fog computing for healthcare 4.0 environment: Opportunities and challenges,'' *Comput. Elect. Eng.*, vol. 72, pp. 1–13, Nov. 2018.

[64] S. Tanwar, K. Parekh, and R. Evans, ''Blockchain-based electronic healthcare record system for healthcare 4.0 applications,'' *J. Inf. Secur. Appl.*, vol. 50, 2020, Art. no. 102407.

[65] *Iota*. Accessed: Nov. 4, 2019. [Online]. Available: https://www.iota.org/

[66] S. K. Singh, S. Rathore, and J. H. Park, ''Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence,'' *Future Gener. Comput. Syst.*, to be published, doi: 10.1016/j.future.2019.09.002.

[67] K. Salah, M. H. U. Rehman, N. Nizamuddin, and A. Al-Fuqaha, ''Blockchain for ai: Review and open research challenges,'' *IEEE Access*, vol. 7, pp. 10127–10149, 2019.

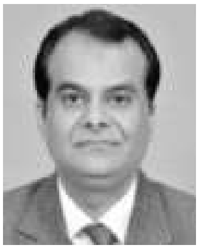
[68] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, ''Verification and validation techniques for streaming big data analytics in Internet of Things environment,'' *IET Netw.*, vol. 8, no. 2, pp. 92–100, 2018.

[69] A. K. Fedorov, E. O. Kiktenko, and A. I. Lvovsky, ''Quantum computers put blockchain security at risk,'' *Nature*, vol. 563, pp. 465–467, Nov. 2018.

[70] M. A. Ferrag and L. Maglaras, ''DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids,'' *IEEE Trans. Eng. Manage.*, to be published.

[71] N. Wang, X. Zhou, X. Lu, Z. Guan, L. Wu, X. Du, and M. Guizani, ''When energy trading meets blockchain in electrical power system: The state of the art,'' *Appl. Sci.*, vol. 9, no. 8, p. 1561, Apr. 2019.

[72] N. Kabra, P. Bhattacharya, S. Tanwar, and S. Tyagi, ''Mudrachain: Blockchain-based framework for automated cheque clearance in financial institutions,'' *Future Gener. Comput. Syst.*, vol. 102, pp. 574–587, 2020.

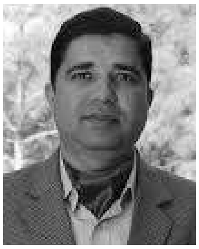


**SUDEEP TANWAR** received the B.Tech. degree from Kurukshetra University, India, in 2002, the M.Tech. degree (Hons.) from Guru Gobind Singh Indraprastha University, Delhi, India, in 2009, and the Ph.D. degree in wireless sensor network, in 2016. He is currently an Associate Professor with the Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. He is also a Visiting Professor with Jan Wyzykowski University, Polkowice, Poland, and also with the University of Pitesti in Pitesti, Romania. He has authored or coauthored more than 100 technical research articles published in leading journals and conferences from the IEEE, Elsevier, Springer, and Wiley. Some of his research findings are published in top cited journals such as the IEEE Transactions on TVT, the IEEE Transactions on Industrial Informatics, *Applied Soft Computing*, the *Journal of Network and Computer Application*, *Pervasive and Mobile Computing*, the *International Journal of Communication System*, *Telecommunication System*, *Computer and Electrical Engineering*, and the IEEE Systems Journal. He has also published three edited/authored books with International/National Publishers. He has guided many students leading to M.E./M.Tech. and guiding students leading to Ph.D. His current interest includes wireless sensor networks, fog computing, smart grid, the IoT, and blockchain technology. He has been awarded best research paper awards from IEEE GLOBECOM 2018, IEEE ICC 2019, and Springer ICRIC-2019.He was invited as a Program Chair, a Publications Chair, a Publicity Chair, and a Session Chair in many International Conferences held in North America, Europe, Asia, and Africa. He is an Associate Editor of IJCS, Wiley and *Security and Privacy* Journal, Wiley. He was invited as a Guest Editors/Editorial Board Members of many International Journals, invited for keynote Speaker in many International Conferences held in Asia.

**QASIM BHATIA** is currently pursuing the degree with Nirma University, Ahmedabad, India. His research interest includes machine learning, network security, fog computing, and cloud computing.

**PRUTHVI PATEL** is currently pursuing the degree with Nirma University, Ahmedabad, India. His research interest includes big data analytics, fog computing, and cloud computing.

**APARNA KUMARI** received the M.Tech. degree from Jawaharlal Nehru University, Delhi, India, in 2012. She is currently pursuing the Ph.D. degree from the Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India. Her research interest includes big data analytics, smart grid, blockchain technology, and cloud computing.



**PRADEEP KUMAR SINGH** received the M.Tech. degree (Hons.) in CSE from GGSIPU, New Delhi, India, and the Ph.D. degree in computer science and engineering from Gautam Buddha University (State Government University), Greater Noida, India. He is currently working as an Assistant Professor (Senior Grade) with the Department of CSE, Jaypee University of Information Technology (JUIT), Waknaghat. He has published nearly 85 research articles in various International Journals and Conferences of repute. He has received three sponsored research projects grant from Governement of India and Government of HP worth Rs 25 Lakhs. He has edited total eight books from Springer and Elsevier and also edited several special issues for SCI and SCIE Journals from Elsevier and IGI Global. He has Google scholar citations 401, H-index 12 and i-10 Index 15 in his account. He is having life membership of Computer Society of India (CSI), a Life Member of IEI and promoted to Senior Member Grade from CSI and ACM. He is an Associate Editor of the *International Journal of Information Security and Cybercrime* (IJISC) a scientific peer reviewed journal from Romania.

**WEI-CHIANG HONG** (M'04–SM'10) is currently a Professor with the Department of Information Management, Oriental Institute of Technology, Taiwan. His research interests mainly include computational intelligence (neural networks and evolutionary computation), and application of forecasting technology (ARIMA, support vector regression, and chaos theory), and machine learning algorithms. He serves as the program committee for various international conferences including premium ones such as IEEE CEC, IEEE CIS, IEEE ICNSC, IEEE SMC, IEEE CASE, and IEEE SMCia. In May 2012, his article had been evaluated as Top Cited Article 2007-2011 by Elsevier Publisher (Netherlands). In September 2012, once again, his article had been indexed in ISI Essential Science Indicator database as Highly Cited Articles, in the meanwhile, he also had been awarded as the Model Teacher Award by Taiwan Private Education Association. He is indexed in the list of Who's Who in the World (25th–30th Editions), Who's Who in Asia (2nd Edition), and Who's Who in Science and Engineering (10th and 11th Editions). He has Google scholar citations 5424, H-index 38, and i-10 Index 64 in his account. He is a Senior Member of IIE. He is currently appointed as the Editor-in-Chief of the *International Journal of Applied Evolutionary Computation*, in addition, he serves as a Guest Editor for the *Energies*, and is appointed as an Associate Editor of the *Neurocomputing*, *Forecasting*, and the *International Journal of System Dynamics Applications*.