





Research Article

The Applications of Blockchain in Artificial Intelligence

Ruonan Wang ^{1,2} **Min Luo** ¹ **Yihong Wen**,³ **Lianhai Wang**,⁴
Kim-Kwang Raymond Choo ⁵ and **Debiao He** ^{1,2}

¹School of Cyber Science and Engineering, Wuhan University, Wuhan, China

²Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin, China

³The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, China

⁴Shandong Provincial Key Laboratory of Computer Networks, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China

⁵Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA

Correspondence should be addressed to Min Luo; mluo@whu.edu.cn

Received 21 June 2021; Accepted 8 September 2021; Published 28 September 2021

Academic Editor: Wenxiu Ding

Copyright © 2021 Ruonan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There has been increased interest in applying artificial intelligence (AI) in various settings to inform decision-making and facilitate predictive analytics. In recent times, there have also been attempts to utilize blockchain (a peer-to-peer distributed system) to facilitate AI applications, for example, in secure data sharing (for model training), preserving data privacy, and supporting trusted AI decision and decentralized AI. Hence, in this paper, we perform a comprehensive review of how blockchain can benefit AI from these four aspects. Our analysis of 27 English-language articles published between 2018 and 2021 identifies a number of research challenges and opportunities.

1. Introduction

Artificial intelligence (AI), an important branch of computer science, underpins the research and development of the theory(ies), method(s), technology(ies), and application(s) for simulating, extending, and expanding human intelligence. While AI was first proposed in 1956, the interest in AI probably increased significantly after AlphaGo (an AI-based computer program) defeated Lee Sedol, the world Go champion. AI has been applied in diverse settings, ranging from healthcare [1–4] to drug discovery [5] to (medical) image recognition [6–9] to automated driving [10, 11] and so on. McKinsey Global Institute, for example, predicted that the AI market will grow to 13 trillion dollars by 2030 [12].

There are three key aspects in AI technology, namely, data, algorithm, and computing power, in the sense that significant data is required for training the algorithm to obtain a classification model, and the training process requires significant computing power. In our big data era, data can come from different sources (e.g., sensor systems,

Internet of Things (IoT) devices and systems, and social media platforms) and/or owned by different stakeholders. This may lead to some challenges.

One of the key challenges is isolated data islands, where data from one source/stakeholder is not accessible to others or the training of the AI model or it is too expensive or impractical to collect the large volume of distributed data for centralized processing and training [13, 14]. There is also the risk of being a single point of failure in centralized architectures [15], which may lead to data tampering.

In addition, data from different sources can be unstructured and vary in quality. It can also be difficult to determine the source and authenticity of the data. There is also the risk of invalid or malicious data. All these limitations can impact on the accuracy of the prediction [16]. For example, in 2017, a group of researchers from MIT demonstrated how one can trick Google's AI classification to classify a 3D printed turtle as a rifle [17]. It has also been shown that fake biometric features can be used to taint recognition models, to facilitate impersonation and fraud

[18]. Such attacks are also referred to as adversarial machine learning in the literature, and it is an ongoing research topic [19–22]. However, it is difficult for conventional AI architectures to screen data effectively and/or track malicious data providers. Conventional centralized AI architectures may also lead to privacy disclosure and data abuse. For example, AI training tasks may need to handle sensitive user data, such as users' medical data, which may be divulged or even tampered with during the training process. There are also privacy regulations that may limit the sharing of user data, even for model training. This motivates the design of data sharing/trading platforms [23].

In practice, AI models are created, trained, and used by different entities. The training process is opaque to the users, and users may not fully trust the model they are using. In addition, as AI algorithms become more and more complex, it is difficult for people to understand how the training result is obtained. Hence, a recent trend is to move away from centralized AI approaches to decentralized AI approaches.

Compared with AI, blockchain is a relatively young technology as it was first proposed by Nakamoto and Wright in 2008 [24]. Blockchain, a peer-to-peer distributed system, ensures tamper-proofing through the underlying hash algorithm and time stamp technology. The privacy of data stored on the blockchain is guaranteed by using some cryptographic algorithms. Through the use of smart contracts, the program can be executed automatically to ensure the credibility of the execution results. Through consensus mechanism and distributed ledger technology, all nodes can participate in bookkeeping and verify the transactions. The market capitalization of Bitcoin as of February 20, 2020, was approximately 175 billion dollars [25]. As shown in Table 1, these characteristics of blockchain may overcome the challenges faced by AI; hence, it forms the focus on this paper.

There has been some research on how blockchain and AI can be combined. In [26], the authors described how the integration of blockchain and AI can develop a new ecosystem for a decentralized economy in terms of decentralized data storage and management, infrastructure, and AI applications. However, there is a lack of discussion on how the specific technology of blockchain is applied and how privacy is protected. The works in [27–29] focused on the integration of blockchain and artificial intelligence and their mutual reinforcement. In [30], the authors discussed the feasibility and benefits of combining blockchain and artificial intelligence for energy cloud management. But it does not categorise the literature available for discussion, while the application scenarios are limited.

Many companies have also done a lot of exploration of blockchain-based AI applications. SingularityNet [31] utilized blockchain technology to build distributed AI trading marketplaces. TraneAi [32] has created a blockchain-based AI platform to accelerate the artificial intelligence training process in a decentralized manner. Neureal [33] is a distributed open source platform for artificial intelligence that provides peer-to-peer super AI computing. In general, most of the applications are aimed at constructing distributed ecosystems and infrastructures via blockchain.

In this paper, we survey the existing literature focusing on the applications of blockchain in AI. Specifically, we searched using keywords such as (blockchain and AI) on major academic databases (e.g., IEEE Xplore, ScienceDirect, ACM Digital Library, and SpringerLink) for articles published in English between Feb, 2018 and Jan, 2021. We located over 500 articles, and we excluded articles that are not directly relevant. Eventually, we included 27 articles for discussion in this paper.

The rest of the paper is organized as follows. In the next two sections, we briefly introduce blockchain and AI, prior to presenting our review of the applications of blockchain in AI (i.e., data sharing, privacy protection, trusted AI decision, and decentralized AI) in Section 4. Finally, we discuss the findings, research challenges, and identified research opportunities in the last section.

2. Blockchain Technology

The architecture of blockchain mainly comprises the data layer, network layer, consensus layer, incentive layer, contract layer, and application layer, see also Figure 1.

The data layer mainly focuses on the data structure, including the hash function, digital signature, Merkle tree, asymmetric encryption, and other technologies. The most important structure of the data layer is the block, and the block structure is shown in Figure 2. A block consists of both the block head and the block body. The block header contains the Merkle root, timestamp, and hash value of the current block and previous block. The block body mainly includes transaction information and Merkle tree. Each transaction is signed by the transaction's initiator and then processed and verified by the miner. The verified transaction is embedded in the block. The hash value of every transaction is combined in pairs to calculate the hash, and then, the resulting hash value is combined in pairs to calculate the hash value again until the Merkle root, which is recorded in the block header. Every change to the information about every transaction stored on the blockchain affects the Merkle root. In this way, the tamper-proofing of blockchain can be realized. Every block additionally stores the hash value of the previous block and timestamp, resulting in a time sorted chain.

The network layer mainly contains P2P network, design of the data communication mechanism, and data verification mechanism. There is no centralized server in the blockchain. All messages propagate between nodes in a peer-to-peer manner. All nodes maintain the blockchain together. One node generates a new block and transmits it to the other nodes. Other nodes store the copy of the block after verification. Subsequent blocks will also be generated on the basis of this block. In this way, all nodes can maintain a bottom ledger.

The consensus layer mainly includes various consensus algorithms. The consensus algorithm is used to determine which node can add new blocks to the main chain. Common consensus algorithms include PoW, PoS, and PBFT.

The incentive layer mainly includes some incentive measures. There is no centralized server in the blockchain, so the safe operation of the blockchain depends on the active

TABLE 1: Blockchain for AI.

	Blockchain	AI	Blockchain for AI
Data	(i) Trust (ii) Security	(i) Ensure high quality data (ii) Secure data sharing	Blockchain supports AI in terms of facilitating trustworthy data and secure data sharing
Algorithm	(i) Automation (ii) Immutability (iii) Traceability	(i) Need a credible training process	Model training is automatically executed by the smart contract, which greatly improves the credibility of the training results
Computer power	(i) Decentralization (ii) Trust (iii) Traceability	(i) Centralized AI-high computing cost	The decentralized blockchain structure provides distributed computing power for AI

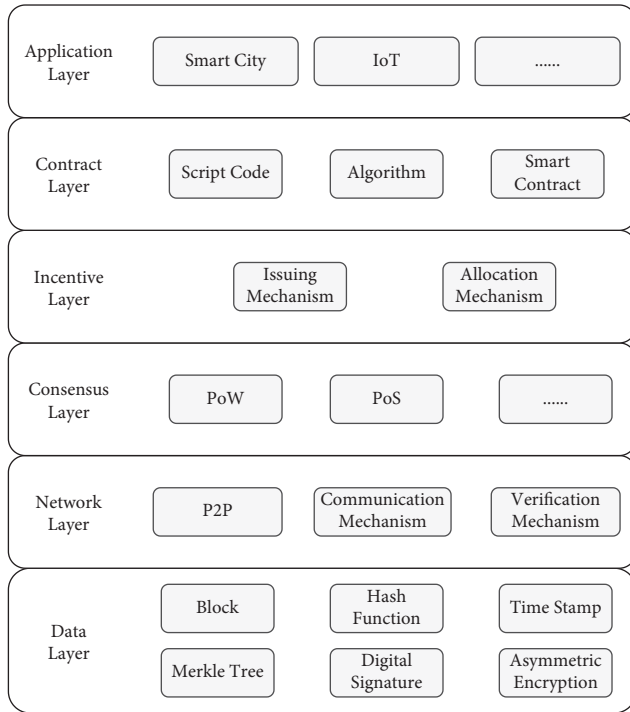


FIGURE 1: Blockchain architecture.

participation of each node. At present, the commonly used incentive measures include (1) the corresponding reward for each block's bookkeeping right and (2) the service charge for each transaction. With the development of blockchain, the design of the incentive layer of blockchain in the future is not only limited to economic rewards but also to achieve common goals.

The contract layer encapsulates several scripts, algorithms, and smart contracts to support the programmable features of blockchain. Through the preset rules and conditions, it can be automatically executed without a third party. This is the foundation of blockchain trust. The last part is the application layer. It contains all types of blockchain applications, including finance, law, audit, and health care.

2.1. Smart Contract. Smart contract was first defined by Szabo in 1994 [34]. Smart contracts have not been adopted on a large scale because of the lack of a reliable execution environment before the introduction of blockchain

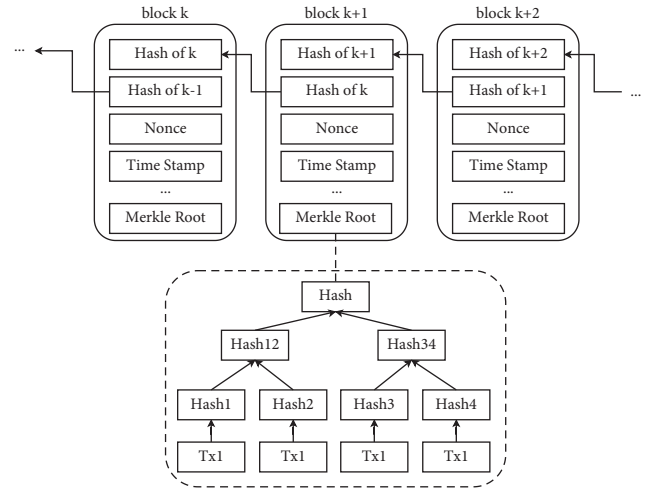


FIGURE 2: Structure of blocks.

technology. With the rapid development of distributed ledger technology, especially the large-scale deployment and application of blockchain, people pay more attention to smart contract, which is the key and important part of distributed ledger technology.

Smart contract is a kind of computer protocol that can self execute, self enforce, self verify, and self restrict the execution of its instructions. It allows transactions to be executed between untrusted or anonymous parties without the need for a trusted third party. These transactions are traceable and irreversible. Smart contract consists of value, address, function, and state. The transaction is taken as input, the corresponding code is executed, and an output event is triggered; then, the state changes according to the functional logic [35]. All parties agree on the details of the smart contract in advance, including scenarios that trigger contract execution, state transition rules, and responsibility for breach of contract. Then, the smart contract is deployed on the blockchain in the form of a code. After that, when the requirements are satisfied, the smart contract will be triggered and automatically executed.

Ethereum is the most popular platform for the development of smart contract [36]. The code of Ethereum smart contract is written in stack-based bytecode language and runs on Ethereum virtual machine (EVM). Solidity and Serpent are often used to write smart contracts. Hyperledger

fabric can also deploy smart contract, which is called “chain code.” It is the single interaction channel with the blockchain and the only source for transaction generation. Chain code is usually developed using Go or Java.

2.2. Consensus Mechanism. There is no trust relationship between each node in the blockchain. It is necessary to coordinate each independent node in order to share information in such a network. Therefore, the network system will decide who is the next bookkeeper through relevant protocols so as to reach a consensus of which is the consensus mechanism. The essence of the consensus mechanism is to solve the problem of decentralized trust [37]. It is an essential technology for the independent operation of the blockchain. In the steady functioning of a blockchain system, a good consensus mechanism plays a very significant part. The blockchain can successfully negotiate and construct a consistent blockchain structure using an effective consensus mechanism.

There are two sorts of consensus algorithms. One is for non-Byzantine fault, such as RAFT and Paxos. The other is for the Byzantine general problem [38], such as PoW, PoS, DPoS, and PBFT. There are two ways to deal with Byzantine fault. One is to limit the probability of malicious behavior by increasing the cost of doing evil, such as PoW and PoS. The PoW algorithm will count computing power as cost, and the PoS algorithm will count stakes as cost. Another way is to design certain rules. Even if there are certain malicious nodes, all other nodes can still reach a consensus, such as the practical Byzantine fault tolerant algorithm. Several common consensus algorithms are described as follows:

- (1) **Proof of Work:** the PoW algorithm was first proposed in Bitcoin, and its core idea is the competition of node computing power. The miner can have the bookkeeping right by consuming a lot of computing power to calculate a hash value that meets the requirements [39]. The block will be added to blockchain after the other nodes validating it. Then, the bookkeeping node will be rewarded. In the PoW consensus mechanism, it takes a lot of resources for malicious nodes to destroy the system (control more than 50 percent of nodes). Therefore, it can limit the malicious behavior of malicious nodes. PoW can be decentralized, and nodes can enter and leave freely. But obviously, it will cause a waste of resources and low efficiency.
- (2) **Proof of Stake:** the PoS algorithm is an alternative to solve the waste of resources in the PoW algorithm. It reduces the difficulty of mining due to the number and time of tokens taken by each node. To a certain extent, it shortens the time to reach consensus and avoids a lot of waste of resources in the PoW algorithm. But, at the same time, PoS benefits wealthy miners and may lead to near monopoly. Therefore, blockchain projects using the PoS algorithm usually need to run the PoW consensus algorithm for a period of time and then convert to PoS to prevent a

large number of stakes accumulating in a small number of nodes.

- (3) **Delegated Proof of Stake:** the DPoS consensus algorithm is improved on the basis of the PoS algorithm. The consensus process no longer requires all participating nodes to fight for the bookkeeping rights, but pick some representatives through voting. It greatly improves the efficiency of consensus.

There are also some other consensus algorithms. The comparison of consensus algorithms is shown in Table 2.

2.3. Taxonomy. Generally, blockchain can be divided into three types according to the access level of blockchain data: public blockchain, private blockchain, and consortium blockchain [40]. The comparison of these three types of blockchain is shown in Table 3.

2.3.1. Public Blockchain. All records stored in the public blockchain are open and transparent to the public, and all nodes can join and leave the blockchain network freely. Everyone can check and verify the transaction and also compete for the rights to bookkeeping. Bitcoin and Ethereum are both public blockchains.

2.3.2. Private Blockchain. The private blockchain is completely controlled by an organization. Not every node is allowed to participate in the blockchain. Only those nodes from specific organizations are allowed to join the competition for bookkeeping rights. It has strict authority management for data access.

2.3.3. Consortium Blockchain. It is a combination of public blockchain and private blockchain. The nodes with permission are selected in advance to participate in the consensus process of the consortium chain. Other nodes can participate in the transactions, but cannot obtain the bookkeeping rights. The data in the blockchain can be public or private. The consortium chain can be seen as partially decentralized. Hyperledger fabric is a consortium blockchain.

3. Artificial Intelligence

The research of AI covers a wide range of topics, including machine learning, computer vision, and natural language processing. Among them, machine learning is an important technology that allows AI to imitate human thought and behavior, and most current AI programs are based on it. Machine learning has been developed over a long period of time, now has a relatively complete technical framework and mature algorithms, and has developed techniques such as deep learning, reinforcement learning, and federated learning.

3.1. Machine Learning. Machine learning was first defined by Samuel [41] in 1959 as “the field of study that gives computers the ability to learn without being explicitly

TABLE 2: Comparison of consensus algorithms.

Characteristics	PoW	PoS	DPoS	PBFT	Raft
Crash fault tolerance (%)	50	50	50	33	50
Byzantine fault tolerance (%)	50	50	50	33	NA
Verification speed (s)	>100	<100	<100	<10	<10
Scalability	High	High	High	Low	Low
Throughput	Very low	Low	Low	Average	Very high
Consistency	Probabilistic	Probabilistic	Probabilistic	High	High
Energy consumption	Very high	Low	Low	Low	Low

TABLE 3: Comparison of blockchain.

Property	Public blockchain	Consortium blockchain	Private blockchain
Governance type	Public	Managed by a set of participants	Managed by a single organization
Read permission	Public	Could be public or restricted	Could be public or restricted
Consensus process	Without permission	With permission	With permission
Data immutability	Nearly impossible to tamper	Could be tampered	High
Transactions throughput	Low	High	High
Centralized	No	Partial	Yes
Network scalability	High	Low to medium	Low to medium

programmed.” As shown in Figure 3, the typical workflow of machine learning involves training and testing. In the training phase, the original data is preprocessed first. Then, feature extraction and model training are carried out based on these data. In the test phase, data preprocessing and feature extraction are required for the test dataset, and then, the test data is analysed and categorised by the training model.

Machine learning can usually be divided into supervised learning, semisupervised learning, and unsupervised learning. Supervised learning uses the labeled data to train the model, which is used to predict. *K*-nearest neighbor, decision tree, neural network, and SVM are all supervised learning algorithms. Unsupervised learning uses training data set with no labels. The key of unsupervised learning is to analyze the hidden structure of data and find out whether there is a divisible set. Semisupervised learning combines supervised learning with unsupervised learning, using a few labeled data and a large amount of unlabeled data for training and classification.

3.2. Federated Learning. The model training of machine learning needs a large number of sensitive data, and data privacy is a very important issue. At the same time, data is distributed in different organizations. These decentralized data are usually heterogeneous and unbalanced, so it is difficult to combine data. Google first proposed federated learning in 2016 [42], which combines machine learning with distributed computing. As shown in Figure 4, the data owners train their local data to get their local submodel. Then, they will upload the updated parameters to the coordinator, which aggregates the local submodel into the federated model. In federated learning, participants only need to share their own training model parameters and do not need to share the original data, which can protect the data privacy to a certain extent.

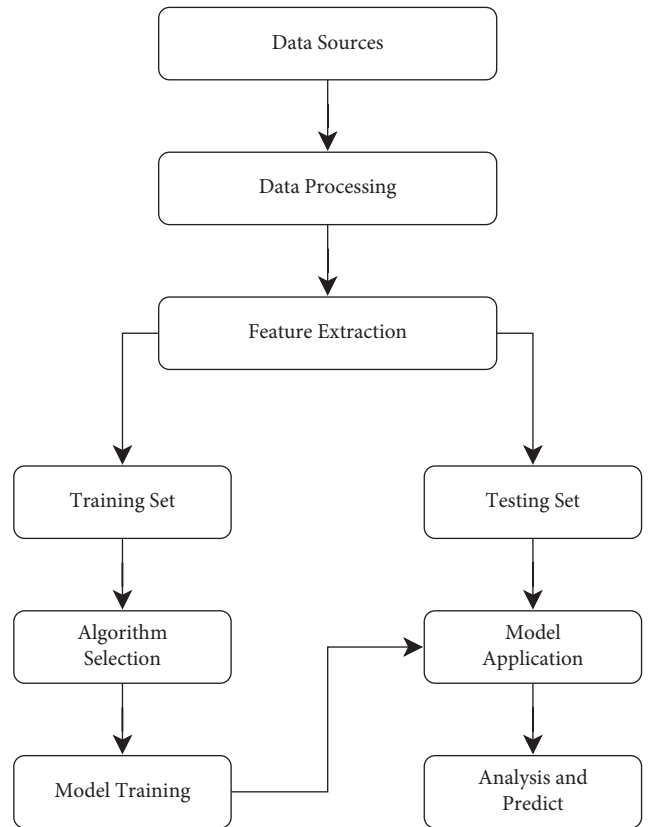


FIGURE 3: A workflow of machine learning.

4. Applications of Blockchain in AI

4.1. Data Sharing. Data is the most important resource of AI. The quantity and quality of data directly affect the accuracy of AI classification results. But, in the process of sharing data, there are some problems. First, the data needed for training is controlled by different stakeholders, and they

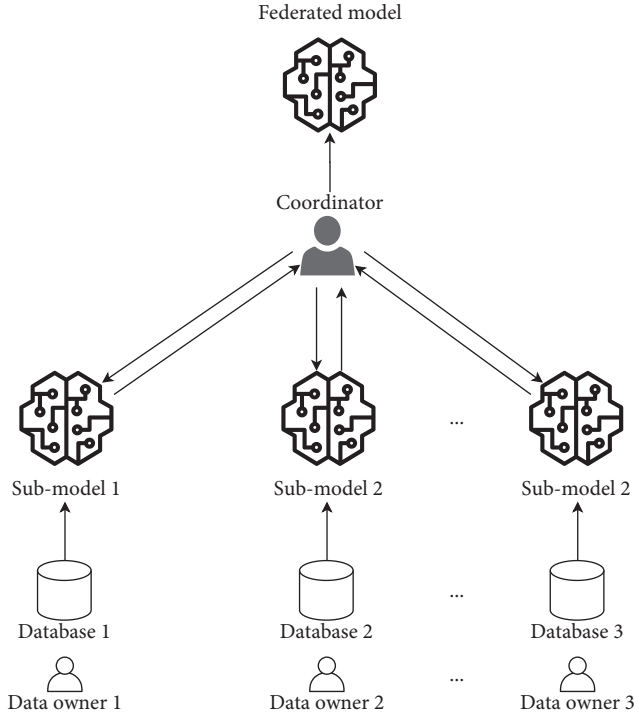


FIGURE 4: The logic of federated learning.

cannot trust each other. It is difficult to authorize or verify the data. Second, there may be malicious users sharing malicious data for certain purposes. There are already some blockchain-based solutions to these problems. Detailed comparison is shown in Table 4.

The work in [43] proposed a blockchain-based, decentralized, and untrusted data market. IoTs' equipment providers and AI solution providers can carry out transparent interaction and cooperation through this platform and realize user registration, data upload, data search, purchase, payment, and feedback through smart contract.

The authors in [44] proposed a SecNet architecture, which realizes secure data sharing and ownership protection through blockchain. A data storage module and an access control module are included in each SecNet node. The use of the private data center (PDC) [48] enables users to monitor any operation of their own data to achieve fine-grained management of data access behavior. Any data ready to be shared is registered in the blockchain, and the access of other parties to the data is also verified and recorded in the blockchain. The economic incentive of sharing data or exchanging security services between different entities is realized through smart contracts.

Singh et al. [46] designed an IoT architecture based on blockchain and AI, called BlockIoTIntelligence. As shown in Figure 5, there are four layers in this framework: device intelligence (DI), edge intelligence (EI), fog intelligence (FI), and cloud intelligence (CI). Each intelligent device in DI is a node in the blockchain. Through the blockchain, data is transmitted between various IoT devices in a distributed manner. And, IoT devices share their data with the EI. The EI transfers the processed data and the underlying computing

tasks to the FI in a distributed manner. In FI, AI technology is mainly used to train models and make decisions. A number of fog nodes that support AI are connected with the blockchain to share intermediate parameters or architecture information to CI. Data center, the core component of the CI, and support AI are linked with blockchain to provide the service of decentralized and secure big data analysis for IoTs' applications.

The work in [47] proposed a cognitive manufacturing mining process based on blockchain. This paper improves the traditional blockchain network and uses the distributed consensus blockchain network based on sidechain to improve the problem of limited storage space of traditional blockchain. This method stores the data of smart devices in a separate database and then maps the data in the sidechain transactions of blocks. Based on the traceability and tamper-proofing of blockchain, the data loss can be effectively avoided at any stage of the cognitive production process.

Robots are used more and more widely nowadays, and they need to interact with each other to improve their abilities. Therefore, the security data sharing between each robot is also very important. In [45], the authors proposed a data model sharing framework called RoboChain, which is used for secure data sharing between robots located in different locations. Robots constantly learn to improve their interaction ability to monitor and improve human health. All operations are performed locally. Personal data will not leave the local hub. After the model is updated, the data will be deleted. At the same time, the local repository will publish a "change" to the hub. The hub will announce the update to the whole network. Robots on other sites will know that there are available updates on the network through their own local hub. After each model update, they need to send a model update transaction to the blockchain. The transaction contains the parameters of the transaction update and the information of the robots participating in the consensus so that other participants can verify how the model is created.

4.2. Privacy Preserving. Privacy preserving is also a key issue. The protection of such personal sensitive data during the sharing process is difficult, which will prevent users from sharing their data. Additionally, the data should be completely controlled by the owner, but now users need to send their own data to the service provider when using the service, resulting in the abuse of personal data by some big companies. Detailed comparison of the solutions for privacy preserving is shown in Table 5.

The work in [49] proposed a distributed multilayer ledger named DeepLinQ, which may enable privacy-preserving data sharing. Taking medical data as an example, a multilayer blockchain architecture is proposed in this paper, which combines the advantages of blockchain (such as complete decentralization, consensus mechanism, and user anonymity) with the current status and actual needs of Electronic Health Records (EHRs), such as EHRs has been centralized; due to the requirement of efficiency, it is impossible to use PoW, and there is a need to introduce jury and verification committee. The underlying blockchain

TABLE 4: Data sharing.

Ref.	Use case	Technologies	Contributions
[43]	IoT	Blockchain smart contract	Propose a blockchain-based data marketplace, where providers of IoT devices and providers of AI/ML solutions will cooperate and trade with each other
[44]	Medical data	Blockchain smart contract PDC	Propose the SecNet, an architecture that can safely store, compute, and share data in the large-scale internet environment
[45]	Robots	Blockchain transaction	Propose RoboChain, a learning framework that enables secure, decentralized, and efficient data and model sharing among multiple robot units in different locations
[46]	IoT	Blockchain	Propose a blockchain-enabled intelligent IoT architecture that enables the effective combination of blockchain and AI for IoT
[47]	Cognitive manufacturing	Sidechain	Propose a subject mining method for cognitive production based on blockchain networks

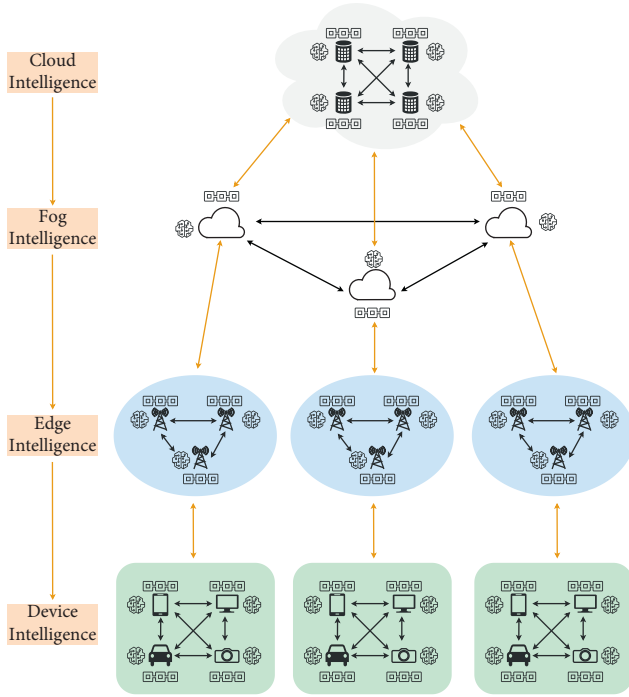


FIGURE 5: BlockIoTIntelligence architecture.

maintains the current properties of Ethereum, and the branch layer meets these actual needs through the design of functions. Medical information is kept off-chain in the architecture proposed in this paper, and the blockchain holds pointers that can find stored information off-chain. This paper presents four ways to design branch layers: (1) add trustworthy advanced validators; (2) use subgroup signatures; (3) create trusted branches; (4) employ efficient consensus protocols. This proposed architecture reflects the management of patients on their own health records, eliminates users' concerns about privacy disclosure, and promotes data sharing between different hospitals.

In [50], the authors proposed to run machine learning algorithms on the blockchain, where different nodes in the blockchain each calculate a part of the machine learning algorithm and cooperate to complete the whole machine learning task. In the smart home environment, the data of IoTs devices is collected to predict users' activities. For example, a device is automatically opened when a specific

user enters a room. The configuration file is generated for each user and device through association rule mining and calculation of personalized parameters. The configuration file is stored in IPFS, and the hash of the configuration file is stored on the blockchain. At the same time, the blockchain calculates another hash value based on the transaction information and stores it on a smart hub. When the user enters the room, the smartphone logs into a smart hub. The transaction information in the blockchain is queried by the hash value stored in the smart hub. The user configuration file is obtained on the IPFS through the hash value stored in the transaction to determine the setting of a device. The confidentiality and authenticity of users' data are guaranteed through the blockchain and smart contract.

Kuo et al. [51] proposed to spread part of the model and other meta information by the metadata in the transactions of private blockchain. Each node can initialize, update, evaluate, and forward the model. Both the model and the hash of the model are included in the update transaction. Other transactions contain simply the hash of model to decrease the blockchain size. The private chain used in this paper does not provide mining reward, and the motivation for block mining comes from the use of cross-agency data in a privacy-protected way to improve the accuracy of predictive models. At the same time, this paper proposed a proof-of-information algorithm based on the PoW algorithm. When a site trains the model with local data and publishes it, other sites, respectively, evaluate the model with their own local data. The site which has the highest error wins the "information bid" and then trains the model with its own local data. When a site updates the model and discovers itself to be in the greatest mistake, the online learning process ends. The current model is the consensus model. The proof-of-information algorithm can use all patients' data to train the model, but it does not need to transmit personal protected health information, which effectively achieves the goal of privacy preserving.

In addition to designing the blockchain architecture and consensus mechanisms, cryptography is also often used to achieve data privacy preserving in AI. The work in [52] proposed a blockchain-enabled learning data management method. The tamper-proofing of blockchain can help prevent the forgery and ensure the integrity of data. The original data and the hash of the original data are stored in the blockchain according to the schemes proposed in this paper.

TABLE 5: Privacy preserving.

Ref.	Use case	Technologies	Contributions
[49]	Healthcare	Blockchain Smart contract	Propose a multilayer blockchain architecture to support privacy-preserving distributed data sharing
[50]	Smart home	Blockchain Smart contract IPFS	Propose a method to run machine learning models in a decentralized manner on the blockchain; different nodes cooperate to complete a training task
[51]	Healthcare	Blockchain Consensus	Propose ModelChain, where partial models are transmitted through transactions of the blockchain and design a proof-of-information algorithm
[52]	AI algorithms	Blockchain Hash function	Propose a blockchain-based solution to prevent data integrity from being compromised in the learning environment
[53]	IoT	Blockchain Paillier	Propose a blockchain-based data sharing platform and construct a secure privacy-preserving SVM algorithm
[54]	Accelerate biomedical Healthcare	Blockchain Threshold encryption	Propose to apply blockchain combined with threshold encryption to healthcare to deal with data privacy issues

The AI model verifies the data integrity by comparing the hash value of the received original data with the hash value stored in the blockchain.

In [53], blockchain has been used to construct a trustworthy and secure data platform across numerous data sources, and IoT data is stored on the blockchain through Paillier encryption. IoT data providers' sensitive information and SVM model parameters are confidential. Data providers can update the gradient without knowing the model parameters through homomorphic encryption. In this way, collusion between data providers and data analysts can be avoided. Combined with the tamper-proofing of blockchain, a secure SVM training algorithm is established to address data integrity and privacy issues.

In healthcare, both consumers and research companies require personal data to train their deep neural networks. Data security and privacy are two of the key difficulties they encounter. Mamoshina et al. [54] proposed to apply blockchain to healthcare to deal with data privacy issues. Users can store and sell their own biological data on the blockchain architecture presented in this article and only allow organizations that have paid for them to access the data. Data validators are used to verify the data to assure the quality of the data given by users. All interactions (user uploads the data, data validator validates the data, and customer buys the data) are carried out through transactions on the blockchain, so as to all data use activities are tracked fairly. This study also offers a threshold encryption scheme to assure data security and user privacy. Users utilize symmetric encryption to encrypt data and then divide the key and distribute it to the blockchain nodes, which act as the key managers. The customers obtain enough keys from the key managers to decrypt the data after they buy the data. In long-term data storage, the threshold encryption scheme can effectively address the single point of failure. The leakage of memory and single key manager will not lead to data leakage, which effectively ensures data security.

4.3. Trusted AI Decision. Different organizations create, train, and use models of machine learning and AI. The entities that train the model are different from the entities

that provide the data. Failure to pay attention to data used in training the model might lead to improper outcomes. Meanwhile, the trained model might have some restrictions if we employ biased data. All of the above operations are opaque to users, and users cannot trust the model they are using. Therefore, we need a mechanism to record the entire process of AI (model creation, training data, and training process), and these records cannot be changed or forged. Blockchain, a platform that enables numerous participants to trustfully share data, has the characteristics of tamper-proofing and transparency, which are very suitable for recording the whole process of machine learning. Detailed comparison of the solutions for trusted AI decision is shown in Table 6.

The work in [55] proposed to manage swarm robots using blockchain to deal with the problem of Byzantine robots (the robots that show arbitrary errors or malicious behavior). In the scheme proposed in this paper, each robot may be utilized as a node or a miner on the blockchain. *Registerrobot*, *applyStrategy*, and *vote* are implemented by smart contract. Robots register in the blockchain through *registeRobot*, publish their own opinions through *Vote*, and then obtain the views of other robots through *applyStrategy*. *applystrategy* achieves consensus of robot views through certain strategy functions. Finally, the experiment demonstrated that the use of blockchain can detect and reject Byzantine robots from the group. Thus, the decision-making results with higher accuracy are obtained.

In [56], a trusted routing scheme for wireless sensor networks based on blockchain and reinforcement learning was proposed. The routing information (destination address, transmitted data, etc.) between various routing nodes is stored on the blockchain. Through the tamper-proofing of blockchain, the credibility of routing information is improved. Firstly, different data packets are represented by different blockchain tokens; secondly, when the routing node joins the blockchain network, it needs to register through the smart contract; Thirdly, before sending the data packets to the next hop route, each routing node has to confirm the routing information stored in the blockchain,

TABLE 6: Trusted AI decision.

Ref.	Use case	Technologies	Contributions
[55]	Robots	Blockchain Smart contract	Propose to manage swarm robots using blockchain to deal with the problem of Byzantine robots
[56]	WSNs	Blockchain Smart contract	Propose a blockchain and reinforcement learning-based routing information management platform to guarantee trusted and efficient routing decisions
[57]	Military platforms Model	Blockchain Smart contract	Propose a framework combining AI, machine learning, and private blockchain to provide decision support for operational centers
[58]	marketplace Coalition operations	Blockchain Transaction	Propose to utilize blockchain to record important processes of AI in an invariable and verifiable way
[59]	Machine learning	Blockchain Smart contract	Propose a trusted machine learning analysis framework which utilizes smart contract to realize automatic machine learning
[60]	DNN applications	Blockchain Edge computing Node.js	Propose a novel DNN architecture combined with edge computing and blockchain technology to overcome the limitation of availability, scalability, and integrity
[61]	Neural networks	Blockchain Edge computing	Propose to use edge computing to reduce the pressure of the server and utilize blockchain to cope with overtraining

and the server node confirms the information through the consensus system and publishes it in the blockchain. Then, each routing node's learning model collects information from blockchain. According to the information, a routing scheduling scheme based on reinforcement learning is proposed to help routing nodes make better routing decisions. Due to the tamper-proofing, traceability, and consensus mechanism of blockchain, the credibility of routing information is enhanced, and the credibility of routing decision is enhanced.

In military operational centers, commanders usually need to make timely decisions based on information from multiple intelligence sources. The work in [57] proposed a framework combining AI, as shown in Figure 6, machine learning and private blockchain to provide decision support for operational centers. This paper proposed to synthesize multiple data sources through multiple AI agents to predict and evaluate the current decision. Blockchain plays two roles in this framework. First, different AI agents can verify whether they are operating in the same blockchain state as other agents, to guarantee that all agents are analyzing the same dataset. It can better compare and analyze the decision results. Secondly, the AI agents are rewarded by the blockchain record to motivate the AI agents to provide decision support.

The work in [58] took federated learning as an example and proposed to utilize blockchain to record important processes of AI in an invariable and verifiable way. Every process is recorded as transactions in the blockchain.

Wang [59] proposed a trusted ML analysis framework based on blockchain. The automated execution of ML algorithms is achieved through smart contracts, and the credibility of machine learning is realized by storing data permanently. Model initialization, training, validation, scoring, evaluation, and reward allocation are automatically executed by the preset smart contract, which greatly improves the credibility of the training results.

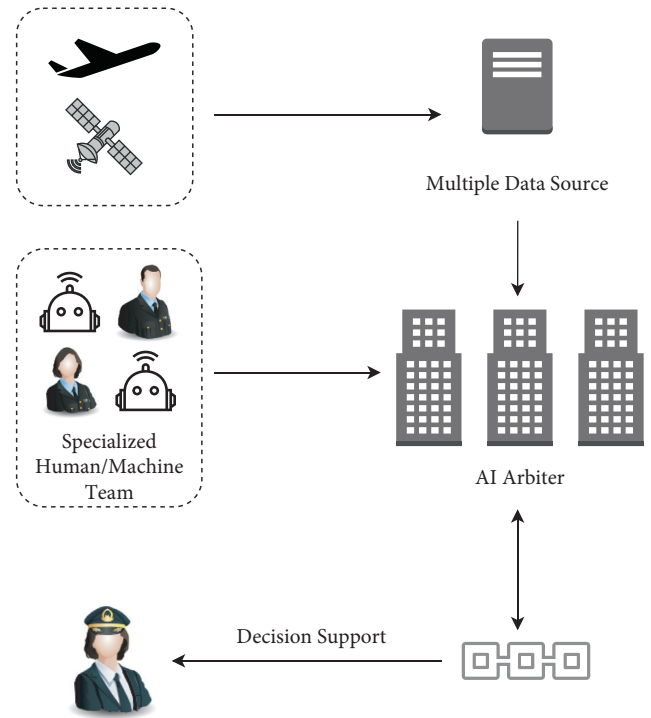


FIGURE 6: Blockchain for real-time decision support.

In the process of AI training, a lot of computing power needs to be consumed. Therefore, in [60], the authors proposed to unload a large amount of computation of the deep neural network from the cloud server to the edge server and then utilized blockchain to accomplish two things: first, encourage the edge server to accept and complete the computation; second, ensure the credibility of edge computing results. Both the embedded device and the edge computing server need to pay a deposit to the blockchain, and then, the embedded device publishes the computing task. After the

TABLE 7: Decentralized intelligence.

Ref.	Use case	Technologies	Contributions
[62]	Distributed ML	Blockchain	Propose a blockchain-based decentralized security learning model LearningChain
[63]	Multiparty learning	Blockchain	Propose BEMA, a novel blockchain-based secure multiparty learning system and design “off-chain sample mining” and “on-chain mining” schemes
[64]	Deep learning Federated learning	Blockchain Consensus	Propose a secure and decentralized privacy-preserving deep learning framework based on blockchain, which encourages participants to engage in training
[65]	IoT	Blockchain Smart contract Consensus	Propose a knowledge market of IoT based on edge-AI, which uses knowledge blockchain for knowledge management and trading
[66]	IoT	Blockchain Transaction	Propose a collective Q-learning combined with blockchain and deep Q-learning, which is used to allocate computing network resources to users
[67]	Federated learning	Blockchain Fabric channel	Propose FLchain to enhance the security of federated learning and implemented on fabric
[68]	Federated learning	Blockchain Incentive mechanism	Propose to utilize consortium blockchain to manage workers’ reputation in a decentralized manner
[69]	Collaborative AI	Blockchain Smart contract	Propose a framework in which everyone may openly view the model’s forecast and input data to enhance the model and utilize smart contracts to host a continuously updated model
[70]	Machine learning	Blockchain Smart contract	Propose to use blockchain technology and smart contract to create an automatic anonymous machine learning model trading market

edge servers have completed its computing tasks, it gets the reward from the embedded device to motivate the edge servers to calculate. At the same time, the credibility of the edge computing results of all the above steps is guaranteed by the smart contract, thus ensuring the reliability of the deep neural network training model. In addition, this paper proposed to use Node.js instead of EVM to deal with smart contracts and perform some complex calculations. The changed program status and results are returned to the blockchain for storage after Node.js execution.

Winnicka and Keşik [61] also proposed to use edge computing to reduce the pressure of the server when training very large amounts of data. In addition, they also mentioned another problem of current artificial neural network training. If the server receives enormous numbers of data in a short space of time and processes them, there may be a problem of overtraining. For this problem, this paper presented to apply blockchain idea to AI. The scheme supports the re-training of classifiers in case of urgent need through setting the priority of every task. All tasks are sent through blockchain transactions, and each task has a priority. The higher the priority, the greater the reward. When miners receive tasks with higher priority, they can choose to interrupt the current task in order to achieve urgent needs. The interrupted task would be broadcasted again in the future, and their training would be executed again until the task is completed, i.e., the stop condition is reached. At the same time, the paper proposed that the more rewards the device gets, the more tasks it performs, which represents the consumption of the device to a certain extent, so the device is more likely to fail.

4.4. Decentralized Intelligence. A vast quantity of IoT data has been created with the fast evolution of the IoTs. Through the AI service, we can obtain the learning results and models

from the massive IoT data. In order to perform complex model training tasks, collaboration with multiple devices is normally needed owing to the distribution of IoT devices and edge computing devices. There are two ways to collaborate here. First, different IoT devices or edge devices need to share data for complete data analysis and prediction (such as intelligent monitoring, monitor in different regions needs to share data). Second, different IoT devices or edge devices share their own learning models and then aggregate the models, that is, federated learning. Detailed comparison of the solutions for decentralized intelligence is shown in Table 7.

Federated learning is a distributed machine learning technology with privacy preserving. A large number of nodes utilize their own local data to train their own local model in a distributed manner. Besides, they only need to share their models instead of sharing original data, which may prevent the leakage of sensitive data.

The work in [62] proposed a decentralized security learning model LearningChain. In the learning chain network, there are two kinds of participants, data holder and computing node. The computing node helps the holder to train the learning model, which is paid by the data holder. Finally, different data holders work together to train a global model. The whole training process is divided into three sections. The first step is the initialization of blockchain and the establishment of the P2P network. The second step is local gradient computation. The third step is global gradient aggregation. The winning node of the PoW adopts an aggregation algorithm, updates the model, and then creates a new block containing the local model and global model information and adds it to the blockchain. The pseudoidentity is utilized in the scheme proposed in this paper to prevent the disclosure of the data owner’s identity information. The data holders can constantly transform the pseudoidentity information in multiple generations to

further protect the identity. When a computing node creates a new block, it checks the validity of the previous block. If the block is not valid, it checks the validity of the previous block until the correct block is found to resist the attack of the Byzantine node.

In [63], the authors proposed a secure multiparty learning system called BEMA based on blockchain. This paper mentioned two forms of Byzantine attacks: (1) malicious participants broadcast a malicious local model to other parties for changing the outcomes of categorization; (2) malicious participants send malicious calibration information to specific parties to mislead the update process of the local model. Through the system proposed in this paper, the first type of Byzantine attack can be effectively prevented, and the second type of Byzantine attack can be controlled in an acceptable range. BEMA consists of system initialization, off-chain sample mining, and on-chain mining. In the initialization phase, each participant broadcasts its local model parameters and stores them in the blockchain. Off-chain sample mining encourages every participant to test the learning model on the blockchain with their own local data. Then, they may find data samples that can calibrate these models and broadcast them to the blockchain. In the on-chain mining phase, miners will update the corresponding model with received calibration data samples after verification.

Although federated learning does not need to share the original data, the privacy of training data cannot still be totally protected. Some research shows that the important information about original training data may be calculated through the intermediate gradient. We need to combine the knowledge of cryptography to overcome these difficulties. For example, homomorphic encryption is utilized to preserve data privacy in [71]. And, secret sharing and symmetric encryption are utilized to preserve data privacy in [72].

The work in [64] proposed a secure and decentralized privacy-preserving deep learning framework. In this framework, N participants initially reached a consensus on the initial parameters of the cooperation model. The parameters were encrypted and recorded in the blockchain through the transaction Tx_{co}^0 signed by all parties. For any participant P_j , the local data is trained to get the intermediate gradient $w_{i,j}$, which is encrypted and recorded in the blockchain through the transaction Tx_{P_j} . Then, the workers will download these transactions to obtain the intermediate gradient of the participants and calculate the new collaboration gradient w_{i+1} at round $i + 1$ through the smart contract. The idea of homomorphic encryption is used here. It is not necessary to decrypt the previous cooperation gradient and the intermediate gradient of the participants when calculating the new cooperation gradient. We can get the encrypted weight at round $i + 1$ by computing $C(w_{i+1}) = (1/N) \cdot C(w_i) \cdot \prod_{j=1}^N C_{P_j}(\Delta w_{i,j})$. Then, $C(w_{i+1})$ will be attached to Tx_{co}^{i+1} . This article also proposed a consensus protocol called blockwise-ba. Three steps are included in the consensus protocol: (1) a leader is chosen randomly to generate a new block is by utilizing cryptographic sortition mentioned in [73], (2) the new block is verified and accepted through executing a Byzantine

agreement protocol by a committee, which comprises of participants whose transactions are included in the new block, and (3) members of the committee inform their neighboring participants of the new blocks via gossip protocol. Through these three steps, all participants reach a consensus.

With the continuous growth of data generated by IoT, data processing and analysis need to be transferred to the edge computing devices to reduce the burden of the cloud. The work in [65] proposed a knowledge market of IoT based on edge-AI, which uses knowledge blockchain (k-chain) for knowledge management and trading. The knowledge market consists of edge-AI nodes and knowledge aggregators. AI algorithms are deployed in the edge-AI nodes who receive the IoT data, analyze the data, and obtain knowledge, and they can be either a buyer or a seller in this market. Edge-AI nodes need to upload the encrypted knowledge to the nearby KAGs. KAG is an enhanced base station (BS) with better resources, which aggregate the knowledge of edge-AI nodes in the corresponding region.

K-chain is separated into two subchains: knowledge management chain (KM-chain) and knowledge trading chain (KT-chain). KAGs collect the uploaded knowledge of its coverage, generate knowledge blocks, and store them in KM-chain. At the same time, knowledge management smart contract (KMSC) is deployed in KM-chain to realize automatic knowledge management. The consensus algorithm used in KM-chain is called proof of capacity (PoC). That is, KAG who has contributed the most storage capacity in the past period of time can be a leader and broadcast the new knowledge block. KT-chain is used to record knowledge trading. All the trading process is completed through knowledge trading smart contract (KTSC) to ensure trading efficiency and fairness. Besides, this paper proposed a new consensus algorithm Proof of Trading (PoT) which combines PoW and PoS. For a KAG, the total number of knowledge trading currency (KC) it owns is taken as its stake. Then, the difficulty of the hash puzzle to be solved in the consensus process will be adjusted dynamically according to the stakes owned by KAGs. The more the stake KAG owns, the easier the hash puzzle it has to solve. PoT cannot only avoid the waste of resources compared with PoW but also resist some attacks compared with PoS.

The work in [66] also deployed AI at the edge of network. The authors combined blockchain and deep Q-learning and proposed a collective Q-learning, which is used to allocate computing network resources to users. First, each edge node learns locally and trains DNNs. Then, the parameters of the local model and other required records are encapsulated into the transaction. The learning results are shared through the blockchain. This enables multiple nodes to work together to complete a complex task.

FLchain is proposed in [67] to enhance the security of federated learning implemented on fabric. The scheme proposed in this paper utilizes the concept of fabric channel. The channel is used to achieve communication isolation between at least two peers. Entities outside the channel are unable to obtain the information in the channel, so as to achieve the privacy of transactions. Every

global model is trained on a different channel in FLchain. The specific steps are as follows: (1) the device inquires the available channel and obtains the channel list; (2) the device selects a channel and registers on channel; (3) the device downloads the current global model of the channel from the blockchain; (4) the device calculates the local model with local data; (5) the device sends the local model to the blockchain; (6) after a period of time, members on channel update the global model through the consensus algorithm and generate new blocks.

Federated learning can effectively protect privacy because it does not need to share original data. However, there may be some malicious behavior, which will influence the model's quality. The choice of dependable workers is highly vital for task publishers. For this problem to be solved, the authors in [68] proposed to utilize consortium blockchain to manage workers' reputation. The architecture of the system in this paper suggested that the task publisher evaluates the model quality and generates reputation opinions after receiving the local model uploaded by all workers. The reputation opinions are stored in the block and maintained by the consortium blockchain after being verified by the miners of consortium blockchain. Due to the decentralization and tamper-proofing of blockchain, reputation opinions stored in blocks cannot be changed. After that, task publishers can choose reliable workers through reputation blockchain to improve the quality of the training model.

In addition to federated learning, there are some other schemes for decentralized intelligence. The work in [69] proposed a framework in which everyone may openly view the model's forecast and input data to enhance the model. That is, the model is trained by many contributors in collaboration. Verification of the incentive mechanism, data upload, and model training are accomplished through the execution of smart contract. This paper introduces in detail the incentive mechanism used to motivate participants to provide useful data. As shown in Figure 7, the model provider saves a deposit as a reward and defines a loss function $L(h, D)$, where h represents the model and D represents the dataset. Each participant needs to deposit 1 unit of currency as a deposit. The smart contract initially pays $L(h_0, D)$ to the first person t_1 , and t_1 updates the model to h_1 with its own data. Then, t_1 need to pay $L(h_1, D)$ to the second person t_2 , until the final person pays $L(h_t, D)$ to the smart contract. So, the reward for the t th person is his own deposit plus $L(h_{t-1}, D) - L(h_t, D)$. The better the model h_t trained by each participant t is, the less the amount it needs to pay to the next participant. This can motivate participants to provide useful data.

Kurtulmus and Daniel [70] proposed to use blockchain technology and smart contract to create an automatic anonymous machine learning model trading market. The entire procedure consists of four phases. The first phase is initialization. Bob, the organizer who wants to obtain the machine learning model, creates an Ethereum smart contract and provides the hash value for the data to the smart contract to trigger the randomization function to generate the index of training and test data. Then, Bob sends the

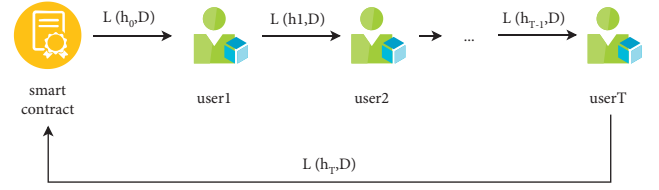


FIGURE 7: Collaborative AI.

training data and nonce to the smart contract. The data can be verified by the hash value provided previously; the second stage is the submission stage. Different scheme providers can obtain training data from smart contract and submit their own training model; the third stage is the evaluation stage. The organizer Bob sends the test data to the smart contract. If Bob does not upload the test data within the specified time, the training data is used for testing. The submitting party calls the evaluation function to submit the evaluation score. The fourth stage is the finalize stage. The organizer Bob pays rewards to the best model provider. If there is no best model, the deposit will be returned to Bob. This scheme helps users to obtain machine learning models at a certain cost and automatically trains, evaluates, and trades models through smart contracts. Organizers can obtain different models trained by decentralized committers and select the best model, which greatly improves the efficiency and credibility of model trading.

5. Conclusion and Future Research

We surveyed the existing literature to understand the potential applications of blockchain in AI. For example, we explained how the different characteristics of blockchain can be used in supporting data sharing, privacy preserving, trusted AI decision, and decentralized intelligence.

- (i) First, as a decentralized platform, blockchain enables data owners and data users to share or trade data in a peer-to-peer manner [23]. Because blockchain is transparent and immutable, it can minimize the potential for fraud in distributed data sharing or transaction.
- (ii) In addition, the underlying cryptographic algorithms (hash algorithms, homomorphic encryption, threshold encryption, etc.) used to process data stored on the blockchain help ensure the confidentiality, integrity, and authenticity of sensitive data.
- (iii) The use of smart contracts to automate model creation, training, sharing, decision-making, and traceability on blockchain helps ensure the credibility of decision results.
- (iv) Incentive mechanisms can be designed on blockchain to promote the cooperation of all participants in completing the AI training tasks.

In addition, we also identified a number of existing and emerging challenges, which will hopefully guide future research agenda.

5.1. Identity Privacy. Blockchain privacy preserving can be either identity privacy or transaction privacy. Identity privacy preserving guarantees that an attacker cannot match an address on the blockchain to a user's actual identity, and transaction privacy preserving guarantees sensitive data from being stolen or tampered. Most existing schemes use blockchain and decision-making process to, respectively, record the training data and protect the privacy of sensitive training data. However, identity privacy preserving is generally ignored.

Bitcoin and Ethereum provide anonymity by using pseudonyms instead of real names for managing and verifying transactions. However, the user's real identity can still be inferred by monitoring the user's transactions [74,75]. Androulaki et al. [76], for example, demonstrated how one can use behavior-based clustering to analyze Bitcoin transactions and consequently match 40 percent of student identities to the associated blockchain addresses, even when users had adopted the privacy measures recommended by Bitcoin. Monroe uses a number of methods to achieve anonymized transactions, such as stealth-address and ring confidential transactions (RingCT) [77]. However, the number of transactions is limited due to the use of ring signatures.

In addition, it can be challenging to achieve blockchain identity privacy preserving, when one takes into consideration blockchain/privacy regulation, deployment difficulty, and robustness of the corresponding architecture for privacy preserving schemes and the impact of privacy preserving schemes on performance.

This reinforces the importance of designing lightweight privacy-preserving schemes.

5.2. Performance. The scalability of blockchain can be considered from both data storage and transaction rate. In AI applications, significant storage space is needed to record the training data and the generated transactions. However, because of the restricted blockchain storage space, it is impractical to store the complete training data. Some existing schemes use sharding [78], sidechain (see [79] for an overview of sidechains), and some other ways to mitigate the storage limitations in blockchain. In addition, the throughput of most public blockchains is very limited. For example, Bitcoin can only handle 7 transactions per second, while Ethereum can handle 7–15 transactions per second. Such rate does not meet the needs of time-sensitive tasks, for example in a smart grid environment.

One possible solution is to design a more efficient consensus mechanism, for example by designing blockchain-based AI applications that utilize private blockchains or consortium blockchains (which can effectively improve throughput) or by designing incentive mechanisms to motivate nodes in the network to participate in the consensus (which can improve the efficiency).

5.3. Security of Smart Contracts. Most blockchain-based AI applications rely on smart contracts to automate the training process. There may be errors and loopholes in smart

contracts [80–82]. For example, vulnerabilities in DAO smart contract built on the Ethereum platform were exploited in an attack in 2016, which resulted in the loss of 3.6 million ethers [83].

Hence, designing secure smart contracts is a topic of ongoing importance [80, 81, 84]. For example, can we also design AI-based approaches to identify and repair vulnerabilities in smart contracts?

Data Availability

The telemetry data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported by the Shandong Provincial Key Research and Development Program (no. 2020CXGC010107), Blockchain Core Technology Strategic Research Program of Ministry of Education of China (no. 2020KJ010301), National Natural Science Foundation of China (nos. 61972294, 61932016, and 62172307), Special Project on Science and Technology Program of Hubei Province (no. 2020AEA013), Natural Science Foundation of Hubei Province (no. 2020CFA052), Wuhan Municipal Science and Technology Project (no. 2020010601012187), and Foundation of Guangxi Key Laboratory of Trusted Software (no. kx202001). The work of K. K. R. Choo was supported only by the Cloud Technology Endowed Professorship.

References

- [1] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of echocardiograms using deep learning," *NPJ digital medicine*, vol. 1, no. 1, pp. 1–8, 2018.
- [2] A. Choudhury and O. Asan, "Role of artificial intelligence in patient safety outcomes: systematic literature review," *JMIR medical informatics*, vol. 8, no. 7, Article ID 18599, 2020.
- [3] H. Zerouaoui and A. Idri, "Reviewing machine learning and image processing based decision-making systems for breast cancer imaging," *Journal of Medical Systems*, vol. 45, no. 1, pp. 1–20, 2021.
- [4] S. Secinaro, D. Calandra, A. Secinaro, V. Muthurangu, and P. Biancone, "The role of artificial intelligence in healthcare: a structured literature review," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–23, 2021.
- [5] E. Gawehn, J. A. Hiss, and G. Schneider, "Deep learning in drug discovery," *Molecular informatics*, vol. 35, no. 1, pp. 3–14, 2016.
- [6] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification?," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017–5032, 2015.
- [7] J. Frank, T. Eisenhofer, L. Schönherr, A. Fischer, D. Kolossa, and T. Holz, "Leveraging frequency analysis for deep fake image recognition," in *Proceedings of the 37th International*

- Conference on Machine Learning, ICML, Vienna, Austria*, 13-18 July 2020.
- [8] C. M. Dourado, S. P. P. Da Silva, R. V. M. Da Nóbrega, P. P. Rebouças Filho, K. Muhammad, and V. H. C. De Albuquerque, "An open ioh-based deep learning framework for online medical image recognition," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 541–548, 2020.
 - [9] O. J. Hénaff, "Data-efficient image recognition with contrastive predictive coding," in *Proceedings of the 37th International Conference on Machine Learning, ICML, Vienna, Austria*, 13-18 July 2020.
 - [10] M. Bojarski, D. Del Testa, D. Dworakowski et al., "End to End Learning for Self-Driving Cars," 2016, <https://arxiv.org/abs/1604.07316>.
 - [11] Y. Xing, C. Lv, X. Mo, Z. Hu, C. Huang, and P. Hang, "Toward safe and smart mobility: energy-aware deep learning for driving behavior analysis and prediction of connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4267–4280, 2021.
 - [12] J. Bughin, M. Chui, R. Joshi, J. Manyika, and J. Seong, *Notes from the AI Frontier – Modeling the Impact of AI on the World Economy*, McKinsey Global Institute, Brussels, Belgium, 2018.
 - [13] V. S. Verykios, E. Bertino, I. N. Fovino, L. P. Provenza, Y. Saygin, and Y. Theodoridis, "State-of-the-art in privacy preserving data mining," *ACM Sigmod Record*, vol. 33, no. 1, pp. 50–57, 2004.
 - [14] S. Gupta, W. Zhang, and F. Wang, "Model accuracy and runtime tradeoff in distributed deep learning: a systematic study," in *Proceedings of the IEEE 16th International Conference on Data Mining (ICDM)*, vol. 1, pp. 171–180, IEEE, Barcelona, Spain, December 2016.
 - [15] T. N. Dinh and M. T. Thai, "Ai and blockchain: a disruptive integration," *Computer*, vol. 51, no. 9, pp. 48–53, 2018.
 - [16] Y. Qi and J. Xiao, "Fintech," *Communications of the ACM*, vol. 61, no. 11, pp. 65–69, 2018.
 - [17] A. Athalye, L. Engstrom, A. Ilyas, and K. Kwok, "Synthesizing robust adversarial examples," in *Proceedings of the International conference on machine learning*, pp. 284–293, February 2018.
 - [18] B. Biggio, L. Didaci, G. Fumera, and F. Roli, "Poisoning attacks to compromise face templates," in *Proceedings of the 2013 International Conference on Biometrics (ICB)*, pp. 1–7, IEEE, Madrid, Spain, 4-7 June 2013.
 - [19] S. Lu, L.-M. Duan, and D.-L. Deng, "Quantum adversarial machine learning," *Physical Review Research*, vol. 2, no. 3, Article ID 033212, 2020.
 - [20] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing connected & autonomous vehicles: challenges posed by adversarial machine learning and the way forward," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 998–1026, 2020.
 - [21] L. Demetrio, A. Valenza, G. Costa, and G. Lagorio, "Waf-amole: evading web application firewalls through adversarial machine learning," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pp. 1745–1752, Brno, Czech Republic, March 2020.
 - [22] E. Quiring, D. Klein, D. Arp, M. Johns, and K. Rieck, "Adversarial preprocessing: understanding and preventing image-scaling attacks in machine learning," in *Proceedings of the 29th USENIX Security Symposium* *USENIX Security 20*, pp. 1363–1380, Boston, MA, USA, August 2020.
 - [23] W. Dai, C. Dai, K.-K. R. Choo, C. Cui, D. Zou, and H. Jin, "SDTE: a secure blockchain-based data trading ecosystem," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 725–737, 2020.
 - [24] S. Nakamoto and C. Wright, "Bitcoin: a peer-to-peer electronic cash system," vol. 2008, 2008.
 - [25] I. Shaikh, "Policy uncertainty and bitcoin returns," *Borsa Istanbul Review*, vol. 20, no. 3, pp. 257–268, 2020.
 - [26] 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.
 - [27] Z. Zhang, X. Song, L. Liu, J. Yin, Y. Wang, and D. Lan, "Recent advances in blockchain and artificial intelligence integration: feasibility analysis, research issues, applications, challenges, and future work," *Security and Communication Networks*, vol. 2021, Article ID 9991535, 2021.
 - [28] E. Karger, "Combining blockchain and artificial intelligence - literature review and state of the art," in *Proceedings of the 41st International Conference on Information Systems, ICIS 2020, Making Digital Inclusive: Blending the Local and the Global*, J. F. George, S. Paul, R. De', E. Karahanna, S. Sarker, and G. Oestreicher-Singer, Eds., Association for Information Systems, Hyderabad, India, December 13-16, 2020.
 - [29] Y. Wu, Z. Wang, Y. Ma, and V. C. M. Leung, "Deep reinforcement learning for blockchain in industrial iot: a survey," *Computer Networks*, vol. 191, Article ID 108004, 2021.
 - [30] A. Kumari, R. Gupta, S. Tanwar, and N. Kumar, "Blockchain and AI amalgamation for energy cloud management: challenges, solutions, and future directions," *Journal of Parallel and Distributed Computing*, vol. 143, pp. 148–166, 2020.
 - [31] "SingularityNet," 2021, <https://singularitynet.io/>.
 - [32] "TranceAI," 2021, <https://github.com/TraneAI>.
 - [33] "Neureal," 2021, <https://neureal.net/>.
 - [34] N. Szabo, "Smart contracts: building blocks for digital markets," *EXTROPY: The Journal of Transhumanist Thought*, vol. 18, no. 2, 1996.
 - [35] B. K. Mohanta, S. S. Panda, and D. Jena, "An overview of smart contract and use cases in blockchain technology," in *Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–4, IEEE, Bengaluru, India, 10-12 July 2018.
 - [36] S. Wang, Y. Yuan, X. Wang, J. Li, R. Qin, and F.-Y. Wang, "An overview of smart contract: architecture, applications, and future trends," in *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 108–113, IEEE, Changshu, China, June 2018.
 - [37] C. Zhang, C. Wu, and X. Wang, "Overview of blockchain consensus mechanism," in *Proceedings of the 2020 2nd International Conference on Big Data Engineering*, pp. 7–12, Shanghai, China, May 2020.
 - [38] L. Lamport, R. Shostak, and M. Pease, "The Byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, 1982.
 - [39] M. Vukob, "The quest for scalable blockchain fabric: proof-of-work vs.," *BFT replication*, vol. 9591, pp. 112–125, 2016.
 - [40] I.-C. Lin and T.-C. Liao, "A survey of blockchain security issues and challenges," *IJ Network Security*, vol. 19, no. 5, pp. 653–659, 2017.
 - [41] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229, 1959.
 - [42] H. B. McMahan, E. Moore, D. Ramage, and B. A. y Arcas, "Federated Learning of Deep Networks Using Model Averaging," 2016, <https://arxiv.org/abs/1602.05629>.
 - [43] K. R. Özyilmaz, M. Doğan, and A. Yurdakul, "Idmob: Iot Data Marketplace on Blockchain," in *Proceedings of the 2018 Crypto*

- valley Conference on Blockchain Technology (CVCBT), pp. 11–19, IEEE, Zug, Switzerland, June 2018.
- [44] K. Wang, J. Dong, Y. Wang, and H. Yin, “Securing data with blockchain and ai,” *IEEE Access*, vol. 7, pp. 77981–77989, 2019.
 - [45] E. C. Ferrer, O. Rudovic, T. Hardjono, and A. Pentland, “Robochain: a Secure Data-Sharing Framework for Human-Robot Interaction,” 2018, <https://arxiv.org/abs/1802.04480>.
 - [46] S. K. Singh, S. Rathore, and J. H. Park, “BlockIoTIntelligence: a blockchain-enabled intelligent IoT architecture with artificial intelligence,” *Future Generation Computer Systems*, vol. 110, pp. 721–743, 2020.
 - [47] K. Chung, H. Yoo, D. Choe, and H. Jung, “Blockchain network based topic mining process for cognitive manufacturing,” *Wireless Personal Communications*, vol. 105, no. 2, pp. 583–597, 2019.
 - [48] H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing, “Hyperconnected network: a decentralized trusted computing and networking paradigm,” *IEEE Network*, vol. 32, no. 1, pp. 112–117, 2018.
 - [49] E. Y. Chang, S.-W. Liao, C.-T. Liu et al., “Deeplinq: distributed multi-layer ledgers for privacy-preserving data sharing,” in *Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 173–178, IEEE, Taichung, Taiwan, December 2018.
 - [50] K. Singla, J. Bose, and S. Katariya, “Machine learning for secure device personalization using blockchain,” in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 67–73, IEEE, Taichung, Taiwan, 10–12 Dec.2018.
 - [51] T.-T. Kuo and L. Ohno-Machado, “Modelchain: decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks,” 2018, <https://arxiv.org/abs/1802.01746>.
 - [52] J. Kim and N. Park, “Blockchain-based data-preserving ai learning environment model for ai cybersecurity systems in iot service environments,” *Applied Sciences*, vol. 10, no. 14, p. 4718, 2020.
 - [53] 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 of Things Journal*, vol. 6, no. 5, pp. 7702–7712, 2019.
 - [54] P. Mamoshina, L. Ojomoko, Y. Yanovich et al., “Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare,” *Oncotarget*, vol. 9, no. 5, pp. 5665–5690, 2018.
 - [55] V. Strobel, E. C. Ferrer, and M. Dorigo, “Managing byzantine robots via blockchain technology in a swarm robotics collective decision making scenario,” in *Proceedings of the International Foundation for Autonomous Agents and Multiagent Systems*, pp. 541–549, Richland, SC, USA, July 2018.
 - [56] J. Yang, S. He, Y. Xu, L. Chen, and J. Ren, “A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks,” *Sensors*, vol. 19, no. 4, p. 970, 2019.
 - [57] M. Blowers, S. Scraftford, and J. Williams, “Blockchain technologies and distributed ledger systems as enablers for real time decision support,” in *Proceedings of the Disruptive Technologies in Information Sciences II*, vol. 11013, Article ID 110130L, May 2019.
 - [58] K. Sarpatwar, R. Vaculin, H. Min et al., “Towards Enabling Trusted Artificial Intelligence via Blockchain,” in *Policy-Based Autonomic Data Governance*, pp. 137–153, Springer, Berlin, Germany, 2019.
 - [59] T. Wang, “A Unified Analytical Framework for Trustable Machine Learning and Automation Running with blockchain,” in *Proceedings of the IEEE International Conference on Big Data (Big Data)*, pp. 4974–4983, IEEE, Seattle, WA, USA, 10–13 Dec.2018.
 - [60] J.-Y. Kim and S.-M. Moon, “Blockchain-based edge computing for deep neural network applications,” in *Proceedings of the Workshop on INTElligent Embedded Systems Architectures and Applications*, pp. 53–55, New York, NY, USA, October 2018.
 - [61] A. Winnicka and K. Kesik, “Idea of using blockchain technique for choosing the best configuration of weights in neural networks,” *Algorithms*, vol. 12, no. 8, p. 163, 2019.
 - [62] X. Chen, J. Ji, C. Luo, W. Liao, and P. Li, “When machine learning meets blockchain: a decentralized, privacy-preserving and secure design,” in *Proceedings of the IEEE International Conference on Big Data (Big Data)*, pp. 1178–1187, IEEE, Seattle, WA, USA, 10–13 Dec.2018.
 - [63] Q. Wang, Y. Guo, X. Wang, T. Ji, L. Yu, and P. Li, “Ai at the edge: blockchain-empowered secure multiparty learning with heterogeneous models,” *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9600–9610, 2020.
 - [64] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, and W. Luo, “Deepchain: auditable and privacy-preserving deep learning with blockchain-based incentive,” *IEEE Transactions on Dependable and Secure Computing*, vol. 2018, p. 679, 2019.
 - [65] X. Lin, J. Li, J. Wu, H. Liang, and W. Yang, “Making knowledge tradable in edge-ai enabled iot: a consortium blockchain-based efficient and incentive approach,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6367–6378, 2019.
 - [66] C. Qiu, X. Wang, H. Yao, J. Du, F. R. Yu, and S. Guo, “Networking integrated cloud-edge-end in iot: a blockchain-assisted collective q-learning approach,” *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12694–12704, 2020.
 - [67] U. Majeed and C. S. Hong, “Flchain: federated learning via mec-enabled blockchain network,” in *Proceedings of the 2019 20th asia-pacific network Operations and management symposium (APNOMS)*, pp. 1–4, IEEE, Matsue, Japan, 18–20 Sept.2019.
 - [68] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, “Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory,” *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10700–10714, 2019.
 - [69] J. D. Harris and B. Waggoner, “Decentralized and collaborative ai on blockchain,” in *Proceedings of the IEEE International Conference on Blockchain (Blockchain)*, pp. 368–375, IEEE, Atlanta, GA, USA, 14–17 July 2019.
 - [70] A. B. Kurtulmus and K. Daniel, “Trustless Machine Learning Contracts; Evaluating and Exchanging Machine Learning Models on the Ethereum Blockchain,” 2018, <https://arxiv.org/abs/1802.10185>.
 - [71] Y. Aono, T. Hayashi, L. Wang, S. Moriai, and L. T. Phong, “Privacy-preserving deep learning via additively homomorphic encryption,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1333–1345, 2017.
 - [72] K. Bonawitz, V. Ivanov, B. Kreuter et al., “Practical secure aggregation for privacy-preserving machine learning,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1175–1191, Dallas, Texas, USA, October 2017.
 - [73] J. Chen and S. Micali, “Algorand: a secure and efficient distributed ledger,” *Theoretical Computer Science*, vol. 777, pp. 155–183, 2019.

- [74] D. K. Soni, H. Sharma, B. Bhushan, N. Sharma, and I. Kaushik, "Security issues & seclusion in bitcoin system," in *Proceedings of the 2020 IEEE 9th International Conference On Communication Systems And Network Technologies (CSNT)*, pp. 223–229, IEEE, Gwalior, India.
- [75] M. Bhargavi, S. M. Katti, M. Shilpa, V. P. Kulkarni, and S. Prasad, "Transactional data analytics for inferring behavioural traits in ethereum blockchain network," in *Proceedings of the IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 485–490, IEEE, Cluj-Napoca, Romania, 3-5 Sept. 2020.
- [76] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin, Financial Cryptography and Data Security," in *Proceedings of the International Conference On Financial Cryptography And Data Security*, pp. 34–51, Springer, Okinawa, Japan, April 1-5, 2013.
- [77] A. Averin, A. Samartsev, and N. Sachenko, "Review of Methods for Ensuring Anonymity and De-anonymization in Blockchain," in *Proceedings of the 2020 International Conference Quality Management, Transport And Information Security, Information Technologies (IT&QM&IS)*, pp. 82–87, IEEE, Yaroslavl, Russia, 7-11 Sept. 2020.
- [78] H. Dang, T. T. A. Dinh, D. Loghin, E.-C. Chang, Q. Lin, and B. C. Ooi, "Towards scaling blockchain systems via sharding," in *Proceedings of the 2019 International Conference on Management of Data*, pp. 123–140, Amsterdam, Netherlands, June 2019.
- [79] A. Singh, K. Click, R. M. Parizi, Q. Zhang, A. Dehghantanha, and K. R. Choo, "Sidechain technologies in blockchain networks: an examination and state-of-the-art review," *Journal of Network and Computer Applications*, vol. 149, 2020.
- [80] Y. Zhuang, Z. Liu, P. Qian, Q. Liu, X. Wang, and Q. He, "Smart contract vulnerability detection using graph neural network," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, C. Bessiere, Ed., pp. 3283–3290, Yokohama, Japan, July 2020.
- [81] X. L. Yu, O. Al-Bataineh, D. Lo, and A. Roychoudhury, "Smart contract repair," *ACM Transactions on Software Engineering and Methodology*, vol. 29, no. 4, pp. 1–32, 2020.
- [82] D. He, Z. Deng, Y. Zhang, S. Chan, Y. Cheng, and N. Guizani, "Smart contract vulnerability analysis and security audit," *IEEE Network*, vol. 34, no. 5, pp. 276–282, 2020.
- [83] D. Siegel, "Understanding the dao attack," vol. 13, Article ID 2018, 2016.
- [84] E. Zhou, S. Hua, B. Pi et al., "Security Assurance for Smart Contract," in *Proceedings of the 2018 9th IFIP International Conference On New Technologies, Mobility And Security (NTMS)*, pp. 1–5, IEEE, Paris, France, 2018.