IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 December 2023, accepted 12 December 2023, date of publication 19 December 2023,
date of current version 28 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3344669

SURVEY

Survey on Blockchain-Enhanced Machine Learning

OZGUR URAL AND KENJI YOSHIGOE, (Senior Member, IEEE)
Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA
Corresponding author: Ozgur Ural (uralo@my.erau.edu)

ABSTRACT The convergence of blockchain and Machine Learning (ML) promises to reshape technological innovation by enhancing security, efficiency, and transparency in ML systems. This survey explores the transformative potential of integrating these two technologies. We outline the foundational principles of blockchain and ML, clarifying their capabilities and synergies. We examine how blockchain strengthens ML as a secure, immutable platform for data sharing, model validation, and executing tasks. We emphasize the opportunities for heightened data security, improved model validation, and decentralized, privacy-preserving systems. However, challenges exist like scalability, energy-wise, and the need for new tailored consensus mechanisms. We provide insights based on recent research at this intersection. Additionally, we explore emerging trends and future directions, like blockchain's application in federated learning for secure, transparent data sharing and model validation. We also investigate privacy-preserving systems such as Proof of Learning, where blockchain enables secure execution while maintaining data privacy. Moreover, we examine the potential for decentralized AI systems leveraging blockchain to deploy and execute models. This survey offers a comprehensive overview of the evolving landscape at the intersection of blockchain and ML, highlighting opportunities and challenges while suggesting future research directions.

INDEX TERMS Blockchain, decentralized AI, energy-wise, federated learning, machine learning, privacy-preserving, proof of learning, scalability.

I. INTRODUCTION

The digital landscape is rapidly evolving, with blockchain and Machine Learning (ML)[1] emerging as pivotal elements shaping the future of various sectors [1], [2]. Individually, these technologies have revolutionized numerous fields. However, their intersection presents a unique opportunity to enhance ML systems [3], [4]. This survey explores this potential, focusing on how blockchain can improve various aspects of ML, from model training and validation to ensure the privacy and security of data [5], [6], [7], [8], [9].

Blockchain, initially conceived to underpin cryptocurrency networks [10], [11], has evolved into a versatile platform with significant implications beyond finance [2], [12]. Blockchain offers a decentralized, secure, and immutable record of transactions underpinned by distributed ledger technology. [11]. Its features present an innovative approach to ensuring data integrity, traceability, and transparency [2], [12]. These are particularly beneficial for ML applications, as they can enhance data security, provide a layer of security for computer vision systems, and facilitate efficient and secure data sharing [6], [9].

ML plays a similarly transformative role across countless sectors. With its ability to learn from and make predictions or decisions based on data, ML is revolutionizing myriad sectors, from healthcare to finance and transportation to entertainment [1], [13], [14].

ML can benefit from integration with blockchain through enhanced privacy-preserving predictive modeling, improved transaction confirmation time prediction, decentralized Proof Of Learning, and decentralized Federated Learning. While numerous studies have explored the advantages of integrating blockchain with ML, such as privacy-preserving predictive modeling, secure computer vision systems, and efficient

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

[1]Complete words of abbreviations are spelled out when they first appear in this manuscript. However, a list of abbreviations is available in the appendix for convenience.

transaction confirmation time prediction [5], [6], [7], [8], [9], there remains a gap in understanding the specific benefits of decentralization that blockchain can bring to ML applications.

Despite the promising opportunities presented by integrating these technologies, challenges remain. For instance, current training on ML models generally requires large amounts of data, which are often unavailable in practice or limited due to the high cost of collection [15]. Filtering out "bad data" is a constant battle with spammers or malicious contributors, who can submit low-effort or nonsensical data and still receive rewards for their work [16]. It is also hard to generalize ML models to reflect the future due to out-of-date training [17]. Concerns about privacy and leakage still exist in fields such as the Industrial Internet of Things [18]. However, with blockchain, these problems can be efficiently solved. We will also address the limitations of current research and outline their potential future directions.

The survey's objectives include:

1) **Review of the Intersection of Blockchain and ML:** Thoroughly review the intersection of blockchain and ML, including integration approaches, benefits, and challenges. For instance, we will explore how blockchain can enhance data security and privacy in ML applications, a topic explored in various studies [3], [6], [15].
2) **Review of Existing Consensus Mechanisms for ML:** Examine existing consensus mechanisms for ML, analyzing their strengths and limitations, such as Proof Of Learning and deep learning. We aim to understand their weaknesses and explore potential improvements. The goal is to understand the landscape of consensus mechanisms and their role in blockchain and ML integration [19], [20], [21], [22], [23].
3) **Analyzing Blockchain Use Cases in ML Applications:** Explore use cases applying blockchain to enhance ML applications. This includes exploring the specific benefits of these use cases and the challenges encountered. We will focus on use cases such as proof of learning and federated learning. For instance, we will examine how blockchain has been used for secure and privacy-preserving federated learning [24], federated learning for autonomous vehicles [25], and healthcare applications [26]. We will also explore the novel blockchain consensus mechanism, Proof of Learning, based on ML competitions [27].
4) **Future Directions and Recommendations:** We will explore areas where this integration could provide significant benefits but has not yet been extensively studied, such as the potential for blockchain to support trustless ML contracts [28] and secure IoT data for e-health applications [9], [29]. We will identify gaps in the existing research, emphasizing the need for a more holistic understanding of the potential of this integration. Then, we will recommend areas where further research could lead to significant advancements, such as using blockchain to incentivize data sharing and penalize dishonest behavior in federated learning [9], [30].

Throughout the rest of this survey, we will delve into the intricate relationship between blockchain and ML, emphasizing their synergies, challenges, and prospects [31], [32]. The paper is structured as follows:

- **Section II: Background** - In this section, we provide a comprehensive overview of the fundamental principles of blockchain. We explore various aspects of blockchain, including its distributed ledger architecture and the different layers that constitute its framework. Additionally, we delve into the core principles of ML, covering topics like supervised learning, unsupervised learning, reinforcement learning, and the application of ML in Proof-of-Useful Work. We also discuss the role of smart contracts in machine-learning contexts.
- **Section III: Related Works** - Reviews relevant literature and differentiates this survey from existing work, summarizing related surveys' contributions and establishing this study's uniqueness.
- **Section IV: Applications and Innovations** - In this section, we investigate the potential for integrating blockchain and ML. Our focus is on enhancing security and privacy in ML models. We delve into real-world applications, including securing ML models and exploring innovative marketplaces that incentivize data sharing. We also examine blockchain-enhanced federated learning systems, the concept of Sharing Updatable Models (SUM) on the blockchain, and platforms like DeepChain and LearningChain Marketplace.
- **Section V: Challenges and Future Research Directions** - This section comprehensively examines the current obstacles and limitations in integrating blockchain with ML. We delve into the technical and practical challenges hindering this convergence and discuss how they can be addressed. Additionally, we forecast emerging trends and potential research avenues, mapping out the future landscape of blockchain and ML integration. This exploration provides a roadmap for future developments and innovations in the field.
- **Section VI: Conclusion** - Concludes the paper by summarizing key findings, insights, and the overall implications for the field.

II. BACKGROUND

A. BLOCKCHAIN

Blockchain is a form of distributed ledger technology [10]. Fundamentally, a blockchain is a decentralized and distributed digital ledger that records transactions across a network of computers in a manner that prohibits retroactive alterations to the registered transactions [2]. This characteristic imbue the system with inherent resistance to data modification, thereby ensuring the immutability and transparency of data. These features of blockchain make it an ideal candidate for integration with ML applications,

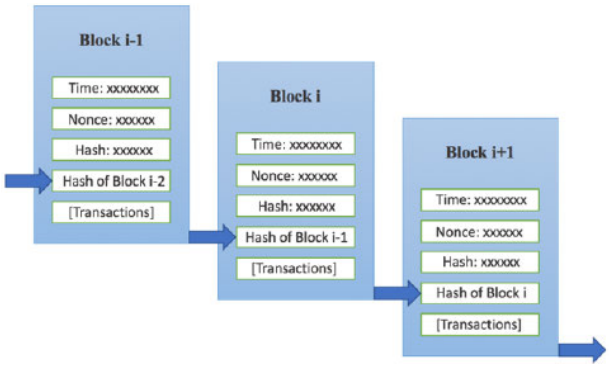

FIGURE 1. Blockchain architecture: components of a block including timestamp, nonce, hash of the block's contents, and transaction list [4].

where data security, privacy, and integrity are of paramount importance [15].

Blockchain technology employs cryptographic methods to safeguard data security and integrity. Typically, each block in a blockchain contains a hash of its contents, a timestamp marking when the data was recorded, a nonce (a one-time number), and the previous block's hash, along with a list of transactions [10]. The nonce, computed during mining in systems like Proof of Work (PoW), ensures the block hash adheres to the network's difficulty criteria [33]. This mechanism is integral to securing new transactions and upholding blockchain integrity. The preceding block's hash links each block sequentially, forming an unalterable chain. While PoW involves mining, other systems might use different consensus methods like Proof of Stake (PoS) to validate transactions and add new blocks, each ensuring network-wide agreement and decentralization. This design thwarts data tampering within a block, as it would require altering all subsequent blocks. The decentralized consensus mechanisms enhance the system's robustness and security, as depicted in Figure 1 [4].

Recent advancements in blockchain technology have further extended its applications into cybersecurity, particularly in mitigating Distributed Denial-of-Service (DDoS) attacks. Innovative frameworks such as Cochain-SC offer a blockchain-based approach combining software-defined networks (SDN) and smart contracts for effective intra- and inter-domain DDoS mitigation [34]. Similarly, Co-IoT utilizes blockchain and SDN for collaborative DDoS mitigation in IoT environments [35]. Other notable developments include ChainSecure, which provides a proactive solution for protecting blockchain applications from DNS amplification attacks in SDN contexts [36], and BrainChain, which employs machine learning techniques within SDN to safeguard permissioned blockchain nodes from similar threats [37]. These studies underscore the growing role of blockchain in enhancing cybersecurity measures, showcasing its versatility beyond traditional transaction recording.

Based on the level of access and data validation permissions, blockchains can be categorized into public, private, and consortium chains, as outlined by Zheng et al. in 2018 [38]. A comparison of these three distinct blockchain types is presented in Table 1.

The principles of blockchain can be summarized as follows:

1) **Distributed ledger:** Blockchain maintains a distributed ledger of all transactions across the network nodes. This ledger is replicated among all participants; no central authority manages it.
2) **Immutability:** Immutability in blockchain refers to the permanent and unchangeable nature of data once it has been written to the blockchain. After adding a block to the chain, altering the data within it would require changing it and all subsequent blocks. This is due to the cryptographic linking of blocks, where each block contains the previous block's hash, creating a dependency chain. Altering any block's data would invalidate the hashes in all following blocks, which is computationally impractical due to the design of blockchain networks. This immutable nature of blockchain provides a tamper-proof and enduring record of all transactions, ensuring the integrity and authenticity of the data stored on the blockchain.
3) **Cryptography:** Blockchain uses cryptographic techniques like hashing and digital signatures to link blocks together and verify transactions. This provides security and authenticity.
4) **Consensus mechanism:** As there is no central authority, blockchain uses a consensus mechanism to validate transactions and add new blocks to the chain. This ensures all nodes have the exact copy of the ledger.
5) **Transparency:** All transactions stored in the blockchain are publicly visible to all participants. This provides transparency and trust.
6) **Irreversibility:** Irreversibility in blockchain pertains to the non-reversible nature of transactions once they have been confirmed and recorded on the blockchain. Blockchain transactions are final once validated and added to a block. This irreversible aspect is fundamental to the security of blockchain networks, as it prevents double-spending and other fraudulent activities. It contributes to the overall trustworthiness and reliability of the blockchain system.

With these critical principles providing a foundation, blockchain shows promise for numerous applications. However, there remain challenges to blockchain's widespread adoption, including issues like scalability, energy inefficiency, and regulatory issues, which need to be addressed for its widespread adoption [39].

1) DISTRIBUTED LEDGER

Unlike traditional ledgers, which are usually controlled by a single entity, a distributed ledger is spread across numerous nodes in a network, and each node holds a copy of the complete ledger. This distribution ensures that no single entity has absolute control over the data, enhancing the

TABLE 1. Comparison of three types of blockchains [15].

Attribute	Public	Private	Consortium
Who run/manage the chain	All miners	One organization/user	Selected users
Permission to Access	No	Yes	Yes
Security	Nearly impossible to fake	Could be tampered	Could be tampered
Efficiency	Low	High	High
Centralized	No	Yes	Partial
Example	Bitcoin, Ethereum	IBM HyperLedger	ConsenSys Quorum

security of the ledger [40]. In the context of private and consortium blockchains, decentralization can prevent a single point of failure and protect data from unauthorized access. However, in public blockchains, while the data is secured against tampering, it is accessible for verification and reading by any participant in the network. This distinction highlights the varying degrees of data privacy and accessibility across blockchain architectures. This benefits ML applications where data privacy and security are crucial [15]. Transactions on the ledger are grouped into blocks, and each block references the one before it, creating a chain of blocks - hence the term blockchain [41].

Distributed ledgers use consensus mechanisms to validate transactions and ensure that all network nodes agree on the ledger's state. These mechanisms ensure that all transactions are validated, recorded, transparent, and immutable, enhancing the system's trustworthiness [33]. The decentralized nature of distributed ledger technology has been recognized for its potential benefits in various sectors, including government and construction, where it can improve information sharing and enable smart contracts, respectively [42]. Building on the concept of distributed ledgers, blockchain systems comprise multiple architectural layers, as explored next.

2) BLOCKCHAIN ARCHITECTURE OVERVIEW

In line with the research presented in [43] and [44], a standard blockchain system can be broken down into six primary layers, as depicted in Figure 2. These layers, which include the data, network, consensus, incentive, contract, and application layers, serve distinct roles and are elaborated upon below.

a: DATA LAYER

The data layer is chiefly concerned with transactions and blocks that store transactional data from various applications. Each block consists of multiple transactions and is connected to its predecessor, forming a sequential list of blocks. As illustrated in Figure 1, a block is divided into a header and main data. The header contains metadata such as the block version, hash pointers to previous and current blocks, timestamp, and Merkle root [46]. The main data section holds all the executed transactions, the nature of which depends on the blockchain service. In Directed Acyclic Graph networks, transactions directly reference previous transactions, eliminating the need for blocks.

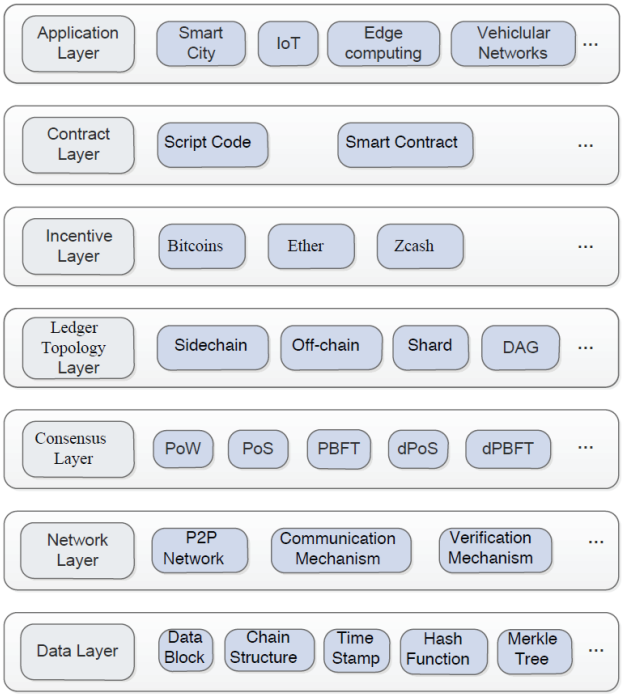

FIGURE 2. Blockchain hierarchical layers: from data to application layer, with key functions and examples [45].

b: NETWORK LAYER

The network layer is responsible for the specific networking mechanisms employed in blockchain to distribute, verify, and audit data generated by the data layer. Typically, this layer operates as a Peer-to-Peer (P2P) network, facilitating decentralized data distribution.

c: CONSENSUS LAYER

The consensus layer specifies the consensus algorithm used to reach an agreement among untrusted parties in decentralized systems. There are various consensus protocols, such as PoW, PoS, and Byzantine Fault Tolerance (BFT) protocols like Practical Byzantine Fault Tolerance (PBFT). The choice of consensus protocol varies depending on the

type of blockchain. For instance, public blockchains often use incentive-based schemes like PoW, while private blockchains may use BFT protocols like PBFT.

d: LEDGER TOPOLOGY LAYER

This layer defines the ledger structure for storing data produced by the consensus layer. While most blockchain applications are built upon a traditional chain of blocks, they are increasingly being complemented by alternative structures such as Directed Acyclic Graphs (DAG), sidechains, and off-chain solutions to address scalability issues. These structures do not replace the fundamental chain of blocks but work alongside it to enhance performance and scalability.

e: INCENTIVE LAYER

The incentive layer introduces economic incentives to encourage nodes to verify data, which is crucial in maintaining a decentralized system without centralized control.

f: CONTRACT LAYER

The contract layer adds programmability to blockchain systems. It employs various script codes and smart contracts to facilitate more complex transactions. Smart contracts, written in Turing-complete languages, extend transaction semantics and implement intricate business rules [47].

g: APPLICATION LAYER

The application layer encompasses a wide range of applications, including but not limited to IoT, smart cities, and edge computing. These applications can transform their respective fields by offering efficient, secure, decentralized solutions.

3) CONSENSUS PROTOCOLS FOR ML APPLICATIONS

A key aspect of blockchain is the consensus mechanism. Given that the blockchain is a distributed system with no central authority, there needs to be a method for validating transactions and agreeing on the current state of the blockchain. This process is achieved through various consensus mechanisms such as PoW, PoS, and others [48]. These mechanisms ensure that all nodes in the blockchain network agree on the validity and order of transactions, maintaining the integrity and consistency of the distributed ledger [49]. Consensus mechanisms, central to blockchain technology as exemplified by Bitcoin, are crucial in preventing double-spending and ensuring the validity of transactions in various applications, including machine learning. These mechanisms facilitate the recording and validation of transactions, contributing to the integrity and reliability of the system. While attributes like transparency and immutability can vary depending on the specific blockchain architecture, they generally enhance trustworthiness across diverse applications [15].

The PoW consensus mechanism involves nodes solving a primary hash function and generating a nonce, thereby proving the work done to validate transactions [49]. However, PoW is criticized for its energy inefficiency and low scalability. On the other hand, the PoS consensus mechanism, used by cryptocurrencies like Ethereum, selects validators deterministically based on their stake in the network, which can reduce energy consumption and improve scalability [48]. In the context of ML applications, PoW can prevent Sybil attacks in distributed ML, while PoS can incentivize honest participation in federated learning [15].

Despite these advancements, consensus mechanisms in blockchain still face challenges, such as the risk of centralization and the need for a balance between security and performance. When applied to ML applications, these challenges can include the need for real-time decision-making and the handling of large datasets [15]. Therefore, further research is needed to develop more efficient and secure consensus mechanisms for blockchain [48], [49].

In addition to consensus protocols, blockchain systems increasingly incorporate smart contracts self-executing code with various applications.

4) SMART CONTRACTS

Smart contracts, first proposed by Nick Szabo in 1994, are self-executing contracts with the terms of the agreement between buyer and seller directly written into code. These contracts automatically execute transactions following predetermined rules when certain conditions are met, without needing a trusted intermediary [50]. In the context of ML, smart contracts can be used to automate the process of data sharing in federated learning or to enforce privacy-preserving protocols, enhancing the efficiency and security of ML applications [15].

The use of smart contracts extends beyond simple transaction processing. They can be used to create decentralized applications (DApps) that run on the blockchain, providing a wide range of services without needing a centralized authority [11]. These decentralized applications can create decentralized ML platforms, enabling users to train and deploy ML models in a decentralized and privacy-preserving manner [4].

Blockchain's ability to provide a transparent, immutable, and secure transaction environment, coupled with its decentralized nature, makes it a potentially transformative tool for numerous industries [2]. One such industry is healthcare, as illustrated in Figure 3, where AI techniques excel at classifying and analyzing large datasets but face data integrity and trustworthiness challenges in decision-making [51]. Integrating blockchain and AI technologies can address these issues by securely storing data on a decentralized, immutable patient record. Smart Contracts can validate various stages such as diagnosis, analytics, and critical decision-making [51].

Similarly, in supply chain management, smart contracts can facilitate complex multi-step processes, where each process step can be automatically triggered by completing the previous step. In ML, smart contracts can be used to automate complex ML workflows, enhancing the efficiency and scalability of these workflows [15].

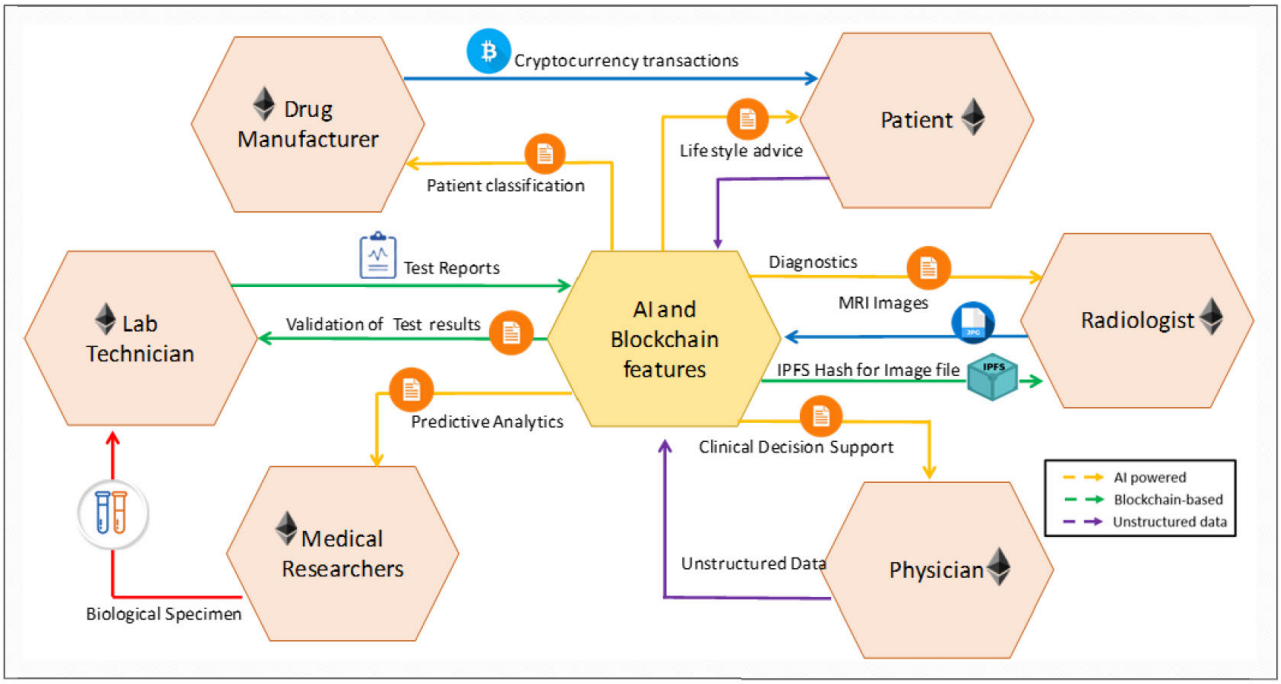

FIGURE 3. Decentralized healthcare through collective intelligence [51].

Having reviewed core blockchain concepts, we now turn to ML and its foundational paradigms.

B. ML

ML leverages algorithms and statistical models to evolve system performance over time based on experiential data. This adaptive approach contrasts with rule-based programming, as ML models autonomously identify patterns and make decisions, reducing the need for human intervention [52].

ML's utility is further amplified when integrated with blockchain technology, which is widely applied across various sectors such as cybersecurity, smart cities, healthcare, e-commerce, and agriculture. This synergy enhances the reliability and effectiveness of ML applications [15].

A significant aspect of contemporary ML research focuses on securing training processes. Innovations in this domain aim to authenticate the derivation of ML model parameters, ensuring they result from computational training efforts. This is crucial for addressing model ownership and integrity, especially in distributed training scenarios [53].

Recent critiques, however, such as "Proof-of-Learning is Currently More Broken Than You Think" by Fang et al., shed light on the vulnerabilities within PoL mechanisms. These studies call for more robust security solutions in ML training processes, indicating an area ripe for further research [54].

Furthermore, ML's capacity to process blockchain data for pattern recognition, fraud detection, and predictive analytics enhances blockchain's utility. Conversely, blockchain's decentralized architecture is instrumental for secure data sharing, which is necessary for ML model training. This mutual benefit is bolstered by blockchain's distributed computing capabilities and smart contracts and consensus mechanisms, which collectively ensure the validity and integrity of ML models [15].

In summary, the integration of ML and blockchain technologies improves existing applications and opens new frontiers in research and development [15].

The role of AI in addressing blockchain challenges is detailed in Table 2.

TABLE 2. Prospects of AI in Blockchain Problems. [55] *Note:* Tick (✓) indicates AI solution probable; Tick (×) indicates AI solution improbable.

S/N	Blockchain Weakness	AI Solution Applicable
1	Smart contracts	✓
2	Authorization and verification	✓
3	Complex cryptography functions	×
4	Weak hashing functions	×
5	Mining decision-making process	✓
6	Vulnerabilities to attacks and threats	✓
7	Transactions delays	✓
8	Byzantine generals' problem	✓
9	Identity privacy and security leakages	✓
10	Secrecy of public-private keys	✓
11	Validation of transactions	✓
12	Endorsement of transactions	✓
13	Network control and management	✓
14	Scalability and interoperability	×
15	Data sharing ineffectiveness	✓
16	Consensus mechanisms and protocols	✓
17	Transparency and openness	×
18	Anonymization of PII	✓

The key advantages of integrating Blockchain and AI technologies are outlined in Table 3. These benefits range from enhanced data security to decentralized intelligence and high-efficiency AI [56].

ML can be categorized into three types: supervised learning, where models are trained using labeled data; unsupervised learning, where models identify patterns in unlabeled data; and reinforcement learning, where models learn to make decisions by interacting with their environment and receiving feedback [1]. Each of these types of ML can benefit from integration with blockchain, for example, through enhanced data security and privacy and the ability to learn from decentralized datasets [15]. Here is a more detailed overview of each type of ML and how it can integrate with blockchain:

1) SUPERVISED LEARNING

In supervised learning, the ML model is trained on a labeled dataset, i.e., a dataset where the target outcome is known. This process involves mapping input data (features) to known outputs (labels). The model uses this known input-output pair to learn the underlying function that governs the data. Once the function is learned, it can predict the output for unseen input data [57].

Supervised learning is commonly applied in regression, classification, and forecasting tasks. For instance, it has been used in classifying astronomical objects, where features such as brightness and color are used to classify objects like stars, galaxies, and quasars [58]. Supervised learning has been used in education to predict students' performance based on features like attendance, participation, and previous grades [59]. In medicine, features such as patient age, blood pressure, and cholesterol levels have been used in supervised learning models to predict the likelihood of heart disease [60].

Integrating supervised learning with blockchain capitalizes on both technologies' strengths to enhance data security and privacy in these applications. Blockchain's decentralized structure and immutable ledger ensure data integrity and protect against unauthorized access. This integration also provides a transparent and immutable record of the supervised learning process, reinforcing trust and verifiability in applying these technologies [15].

2) UNSUPERVISED LEARNING

Unlike supervised learning, unsupervised learning operates on datasets where the target outcome is unknown. These algorithms uncover the data's inherent structure by identifying patterns, correlations, and similarities within the input data. The primary objective of unsupervised learning is to discover the intrinsic distribution of the data [61].

Clustering, a typical application of unsupervised learning, involves grouping similar data points. This technique is widely used in various fields. For example, in marketing, clustering can segment customers into different groups based on their purchasing behavior, which can help businesses tailor their marketing strategies to different customer segments. In computer vision, clustering can be used for image segmentation, where an image is divided into multiple segments that share similar characteristics [62].

Another typical application of unsupervised learning is dimensionality reduction, which aims to reduce the number of random variables under consideration by obtaining a set of principal variables. This technique is beneficial in dealing with high-dimensional data, as it can help alleviate issues such as overfitting and make the data more accessible to visualize [63].

Unsupervised learning methods have also been employed to analyze molecular simulation data in material science, solid-state physics, biophysics, and biochemistry. These methods include feature representation, density estimation, and kinetic models, among others [64].

Integrating unsupervised learning with blockchain technology leverages the strengths of both fields to enhance the security and privacy of data in these applications. Blockchain's immutable ledger and cryptographic security ensure the integrity and confidentiality of data used in unsupervised learning processes. Additionally, the transparent nature of blockchain provides an immutable record of the learning process, fostering trust and verifiability in unsupervised learning applications [15].

3) REINFORCEMENT LEARNING

Reinforcement learning is a type of ML where an agent learns to behave in an environment by performing specific actions and receiving rewards or penalties in return. It is an approach to train an agent to take steps based on the current state to maximize the cumulative reward. The exploration-exploitation tradeoff is a critical challenge in reinforcement learning. This involves the decision of whether to take the best action based on current knowledge (exploitation) or to try a new step in hopes of finding a better one (exploration) [65].

A recent study highlights an interesting connection between two regularization techniques in offline reinforcement learning - actor and critic regularization. The study suggests that under certain conditions, these two approaches can be equivalent, providing valuable insights into the design and efficiency of RL algorithms [66]. This finding contributes to the ongoing discussion on the optimal regularization methods in RL, especially in scenarios with limited data.

Reinforcement learning has been applied in various domains, including game playing. For instance, the AlphaZero algorithm, developed by DeepMind, taught itself to play Go, chess, and shogi (a Japanese version of chess) and beat state-of-the-art programs specializing in these three games. This ability of AlphaZero to adapt to various game rules is a notable step toward achieving a general game-playing system [67].

The integration of reinforcement learning with blockchain technology contributes to enhancing data security and

TABLE 3. Key features and benefits of Blockchain integration with AI [56].

Blockchain	AI	Integration Benefits
Decentralized	Centralized	Enhanced Data Security
Deterministic	Changing	Improved Trust on Robotic Decisions
Immutable	Probabilistic	Collective Decision Making
Data Integrity	Volatile	Decentralized Intelligence
Attacks Resilient	Data, Knowledge, and Decision-centric	High Efficiency

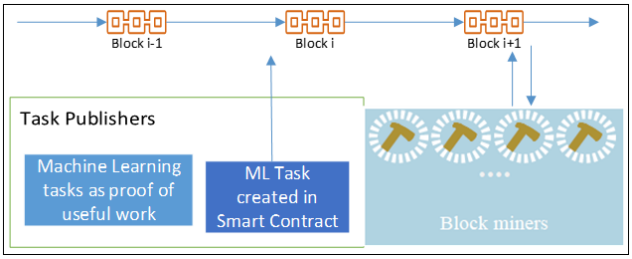

FIGURE 4. Blockchain maintained ML competition algorithm that uses PoUW.

privacy. Blockchain's immutable and transparent record-keeping complements the adaptive decision-making process in reinforcement learning, ensuring the data and learning process are securely stored and verifiable. This synergy offers robust protection against data tampering and unauthorized access while maintaining transparency in the learning outcomes [15].

In addition to the core types of ML, there are other important applications of ML in blockchain contexts:

4) ML IN PROOF-OF-USEFUL-WORK

ML is integral to numerous domains. As computing power has rapidly increased alongside growing research interests, the performance of ML models has significantly improved. However, achieving high-performance models demands substantial computational power [68]. Furthermore, accurate models necessitate ML experts to fine-tune them through multiple iterations of training and evaluation with varying hyperparameters. Consequently, a high-quality model comes with the considerable expense of computational resources.

Several Proof-of-Useful-Work (PoUW) mechanisms, including Primecoin [19], PoX [69], Privacy-Preserving Blockchain Mining [70], Coin.AI [71], WekaCoin [72], and Proof of Deep Learning (PoDL) [73], require miners to perform valuable tasks. Figure 4 illustrates a Blockchain Maintained ML Competition Algorithm that employs the PoUW mechanism, showcasing how ML tasks can be integrated into the mining process.

5) SMART CONTRACTS FOR ML

Smart contracts, self-executing contracts where the terms are written into code and stored on the blockchain, can automate various tasks in ML processes. For instance, smart contracts can facilitate the automatic validation of ML models, trigger actions based on the performance of the models, or handle incentives for data providers in decentralized ML ecosystems [74].

In the context of ML, smart contracts can create a decentralized marketplace for data, where data providers are incentivized to share their data for ML model training [75]. This can lead to more robust and diverse ML models, as they can be trained in a broader range of data.

Moreover, smart contracts can also ensure the fair and transparent evaluation of ML models. Encoding the evaluation criteria into a smart contract can automatically validate an ML model's performance and trigger actions based on the results, such as releasing payment to the model developer. Integrating smart contracts with ML can lead to more efficient and transparent ML processes [74].

C. DEEP LEARNING

Deep learning, a specialized branch within AI and ML, is adept at discerning the underlying representations of fundamental data types, including images, text, and speech signals. This capability allows for executing various applications with accuracy comparable to or exceeding human performance. Examples of deep learning applications span image classification, object detection [76], and autonomous driving [77].

The training of deep learning models heavily depends on extensive labeled datasets, with data quality being crucial for accurate decision-making in ML algorithms, especially with time series data. Insufficient data detail can compromise the model's effectiveness, regardless of algorithm sophistication. While feature engineering may aid data reconstruction by extracting features from raw data [78], deep learning algorithms autonomously extract high-level latent features. These models have multiple layers; lower layers handle essential elements, and higher layers manage more abstract aspects. Notably, the number of layers in these models impacts their accuracy and security; increased complexity, as indicated by more layers, can make the models more susceptible to adversarial attacks, posing potential security risks [79].

Tuning hyperparameters in a deep learning model is crucial for optimizing performance for a specific task [80]. When considering scalability, the choice of network topology becomes a significant factor. For instance, a client-server

architecture offers scalability as the model is trained at the server end, with each node receiving the trained representative [81]. If retraining is needed, it can be conducted at the server end and broadcast to the entire network. This approach simplifies the tuning of hyperparameters and allows for generalization across the network. In contrast, a P2P network may not be as scalable, as each node must individually train or retrain its model instance, which can be computationally intensive.

Deep learning also has applications in cybersecurity, where it analyzes internet traffic patterns to identify threats. Current challenges in security include malware, data breaches, social engineering, phishing, Denial-of-Service (DOS), and insider attacks. Deep learning can aid in detecting and preventing these threats, analyzing data traffic, and verifying transaction signatures for intrusion detection. It can also recognize suspicious activities within the system, using Natural Language Processing, a subtype of deep learning, to detect threats related to social engineering and data theft [82].

1) BLOCKCHAIN-ENHANCED DEEP LEARNING

Blockchain's ability to enable the reusability and secure sharing of deep learning models is a vital necessity. This technology also supports audibility, data verification, result attestation, provenance, ownership traceability, usage monitoring, and fairness assurance, all crucial for integrating blockchain with deep learning. Deep learning models, trained on diverse datasets, depend on the quality of this data to learn and generate accurate predictions.

In addressing the data management challenge in deep learning applications, recent studies have highlighted the effectiveness of leveraging off-chain storage solutions like the InterPlanetary File System (IPFS). For instance, healthcare research has demonstrated a scalable blockchain model using off-chain IPFS storage for patient health records, significantly enhancing data scalability. Similarly, an integrated blockchain and IPFS storage network have been proposed for Electronic Health Records (EHR), focusing on creating a patient-centric access model [83]. This integration effectively addresses scalability issues when storing all data types directly on the blockchain.

Moreover, using IPFS in blockchain systems can drastically reduce the block size, leading to more lightweight and efficient blockchain networks. This reduction is achieved by storing only data indices on-chain, while the actual data resides off-chain, thereby speeding up data replication among network nodes and enhancing overall system efficiency [83]. A method combining Named Data Network (NDN) technology with a distributed blockchain and IPFS has also been proposed to ensure safe storage and efficient sharing of copyrighted files, demonstrating the utility of IPFS in secure data storage and distribution [84].

Furthermore, a secure data storage solution has been developed where only the encrypted hash of the data is stored on the blockchain. This solution maintains a consistent hash size on the chain, regardless of the volume of raw data stored off-chain, offering a scalable and secure data sharing scheme [85]. These advancements underscore the potential of blockchain and off-chain storage solutions like IPFS in enhancing deep learning applications, especially in managing large-scale data efficiently while ensuring data integrity and security.

TABLE 4. A summary of the deep learning and blockchain features that assist in improving deep learning-based applications [56].

Blockchain	Deep learning	Potential outcomes
Immutable	Scalable	Flexibility in learning strategies
Transparent	Layered	Collaborative model update
Integrity	Resource intensive	Enhanced scalability
Cybersecurity	Data-intensive	Upgraded data security

The blockchain, as a decentralized and verifiable global database, thus empowers network nodes to hold securely and exchange data. It enhances the efficiency and effectiveness of deep learning applications by ensuring data integrity, quality, and secure sharing. The key features of blockchain in deep learning are summarized in Table 4, and Figure 5 showcases the primary aspects, scenarios, and categories benefiting from this integration.

In our pursuit to present a thorough analysis of blockchain-enhanced deep learning, we have incorporated Table 5, derived from the seminal work of Shafay et al. [86]. This work is a cornerstone in understanding the multifaceted interaction between blockchain technologies and deep learning methodologies. The table presents a comprehensive comparison and analysis of state-of-the-art blockchain-based deep learning frameworks. By including this table, we aim to give our audience a broad perspective on the advancements and innovations in this evolving field.

The table meticulously categorizes various deep learning applications associated with different types of blockchain technologies and their corresponding consensus protocols. This categorization is crucial as it sheds light on the diverse approaches adopted in the field and how they align with specific blockchain characteristics. For instance, using private blockchains in specific applications emphasizes the requirement for privacy and controlled access, while public blockchains are chosen for their transparency and broader reach.

Furthermore, the table explores the deep learning methods employed in these studies, ranging from traditional neural networks to more advanced techniques like LSTM and CNN. This exploration is vital for understanding the complexity and suitability of different deep learning techniques in blockchain environments. It also helps in assessing the compatibility of these methods with the inherent properties of blockchain technology, such as immutability and decentralization.

Each entry in the table also provides an insightful analysis of the strengths and limitations inherent in each study. This

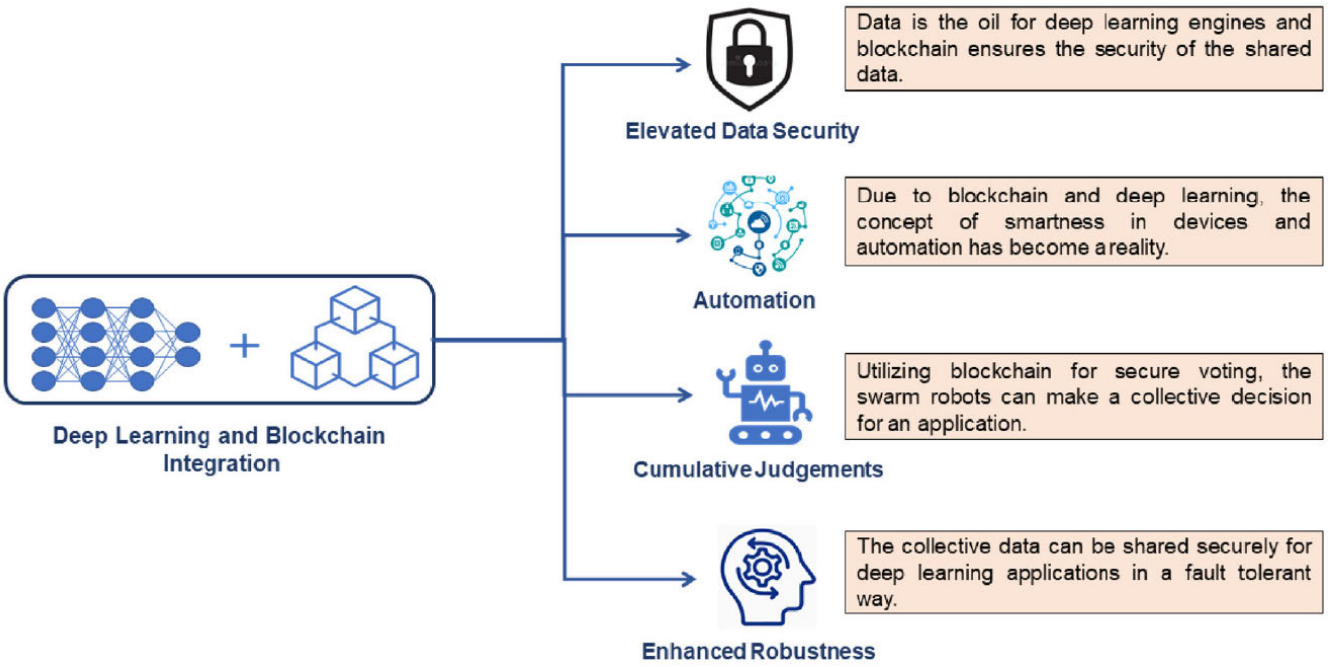

FIGURE 5. Advantages arising from the convergence of deep learning and blockchain [86].

analysis is instrumental in identifying the current challenges faced in the domain, such as scalability issues, computational demands, and data privacy concerns. It also highlights the achievements in the field, like enhanced security measures, improved data integrity, and innovative application scenarios.

This examination of various blockchain-based deep learning frameworks opens avenues for future research. It underscores the need for developing more efficient, secure, and scalable solutions. The insights gained from this table are academic and have profound practical implications, paving the way for more robust, reliable, and versatile applications in the healthcare and finance sectors.

Blockchain and deep learning, when integrated, offer complementary benefits. Blockchain's decentralized, immutable ledger provides robust security for data, which is essential for deep learning models that rely on large datasets for training and accurate predictions. However, storing large datasets for deep learning directly on the blockchain is not feasible due to size and cost constraints. Instead, blockchain can effectively keep critical aspects such as data hashes, model parameters, and metadata, ensuring data integrity and traceability. Off-chain storage solutions like IPFS can be employed for most training data.

Integrating blockchain and deep learning automates tasks requiring robust data handling and security. Blockchain's stability, permanence, and decentralization offer a secure framework to manage data, while deep learning algorithms analyze this reliable data to extract insights and make predictions. This synergy is paving the way for innovative applications across various industries, merging blockchain's data security with the analytical power of deep learning.

A summary of the advantages stemming from the fusion of blockchain with deep learning algorithms includes:

- **Data Security:** Blockchain's decentralized nature ensures robust security for information. Private blockchain platforms are utilized to handle confidential data, and the private keys of the nodes, essential for accessing blockchain data, must be kept secret. Deep learning algorithms can leverage the stable data from the blockchain, leading to more credible, precise, and dependable decision-making [86].
- **Automated Decision Making:** Recognized for processing transactions on a P2P basis, blockchain simplifies the verification of decisions made by deep learning models through its traceability feature. This also guarantees the integrity of documents during human-aided auditing phases [97].
- **Collective Judgments:** In certain situations, autonomous digital agents make decisions based on specific scenario-related data. Deep reinforcement learning and swarm robotics exemplify agent-based decision-making systems. A voting-based method can guide robot decision-making, utilizing data from swarm robotics on the blockchain [98].
- **Increased Robustness:** In some instances, the decision accuracy of deep learning models exceeds human capability, enhancing stakeholder trust. The decentralized nature of the technology further ensures the

TABLE 5. A comparison and analysis of state-of-the-art blockchain-based deep learning frameworks [86].

Category	Blockchain type	Consensus protocol	Deep learning method	Dataset used	Study strengths	Study limitations
Ovarian cancer prediction [87]	N/A	N/A	One-shot Learning	Human Protein Atlas	Requires less training time as only one example per class is required	Performance is not as good as DNN
Data exchange [5]	Private	Proof-of-Information	Incremental Learning	N/A	The considered model can learn new classes on a pre-trained network	Possibility of happening of catastrophic forgetting
EHR prediction [88]	N/A	N/A	LSTM	EHR-based dataset	Superior performance on sequential data	Longer training time and excessive memory requirement
Arrhythmia classification [89]	N/A	N/A	SDA + Sigmoid	MIT-BIH Dataset	Deals well with noise and random variations in the data	Requires a large amount of data for better results
Miner Node Selection [90]	N/A	Zero-Knowledge Proof	Deep Boltzmann Machine	N/A	Data labelling is not required	Expensive in terms of memory and CPU cycles
Communication security [91]	Private	N/A	DNN+ Reinforcement Learning	N/A	Implemented method can solve complex problems that conventional methods cannot	Not suitable for simple problems
Securing blockchain [92]	N/A	Proof-of-Work	N/A	Game Theory-based Utility Function	Data is fully secured	This approach is not practical
Traffic jam prediction [93]	Public	Proof-of-Authority	ANN + LSTM	Historical Traffic Data + Custom Dataset	ANN are fault-tolerant as information is distributed over all the nodes	Black box behavior makes it impossible to develop a relation between dependent and independent variables
Traffic flow prediction [94]	Consortium	Delegated PBFT	GRU	N/A	Lesser expensive in terms of memory requirement as compared to LSTM	Lesser learning ability compared to LSTM
Incident prediction [95]	N/A	N/A	CNN	Custom Dataset	Excellent feature extraction capability and computationally efficient solution	Large dataset is required for training and noise in data can cause misclassification
GPS correction [96]	Public	Delegated PoS	DNN	Custom Dataset	Reduced need of feature engineering	Not prone to data redundancy and inconsistency

system's robustness. The merger of deep learning with blockchain proves valuable in business contexts, allowing parties to operate in a trustless and automated environment [99].

III. RELATED WORKS

A. SUMMARY OF CONTRIBUTIONS OF RELATED SURVEY PAPERS

This survey paper builds upon previous blockchain and ML work. Several survey papers have explored the integration of blockchain and ML, each with unique perspectives and contributions. The titles, publication dates, and main focus of these papers are listed in Table 6.

B. WHAT MAKES OUR SURVEY UNIQUE IN THE BLOCKCHAIN AND ML LANDSCAPE

Our survey paper distinctively contributes to the blockchain and machine learning (ML) landscape, offering insights and perspectives that set it apart from other surveys listed in Table 6.

- **Comprehensive Overview of Blockchain-ML Convergence:** Our survey provides a detailed exploration of how blockchain enhances ML, focusing on secure data sharing, model validation, and task execution, a perspective not extensively covered in surveys such as [25] or [100].
- **In-Depth Focus on Decentralization and Privacy-Preserving Techniques:** We delve into decentralized ML methods like federated learning and privacy-preserving systems like Proof of Learning, offering a more focused analysis than general reviews like [56] or [101].
- **Balanced Analysis of Opportunities and Challenges:** Unlike papers that primarily focus on benefits or challenges, like [102] and [103], our survey presents a balanced view of the potentials and limitations of integrating blockchain in ML.
- **Exploration of Emerging Trends and Future Directions:** We investigate the future of blockchain in federated learning and decentralized AI, offering insights into trends not thoroughly explored in surveys like [86] and [104].
- **Unique Emphasis on Scalability and Energy-Wise Approaches:** Building on foundations laid by surveys

TABLE 6. Summary of Survey Papers For Blockchain And ML.

Title of the Paper	Year	Main Focus of the Paper
Survey on the Convergence of ML and Blockchain [25]	2023	The paper explores the convergence of ML and blockchain, focusing on how these technologies can be combined to solve existing problems efficiently.
ML in/for blockchain: Future and challenges [100]	2020	The paper reviews the integration of ML and blockchain, emphasizing their potential for secure and efficient data sharing and analysis.
Blockchain for AI: Review and Open Research Challenges [56]	2019	The paper provides a detailed review of blockchain applications for AI, discussing the limitations of current research and suggesting future directions.
Comprehensive Survey of IoT, ML, and Blockchain for Health Care Applications: A Topical Assessment for Pandemic Preparedness, Challenges, and Solutions [101]	2021	The paper provides an extensive survey of emerging IoT technologies, ML, and blockchain for healthcare applications, focusing on their adaptability and performance.
Survey on IoT Security: Challenges and Solution using ML, Artificial Intelligence, and Blockchain [102]	2020	The paper reviews ML, Artificial intelligence, and Blockchain technologies to address security issues in IoT, identifying primary security issues and solutions.
Blockchain-based Federated Learning: A Comprehensive Survey [103]	2021	The paper conducts a comprehensive survey of literature on blockchain-based federated learning, discussing how blockchain can be applied to federated learning from the perspective of system composition and mechanism design.
Leveraging ML and blockchain in E-commerce and Beyond benefits, models, and application [104]	2023	The paper summarizes state-of-the-art research on the combination of blockchain and ML across various applications, including e-commerce, discussing the challenges, benefits, and limitations of integrating these technologies.
Blockchain for deep learning: review and open challenges [86]	2023	The paper reviews the integration of blockchain with deep learning, exploring the importance of this integration and presenting the state-of-the-art blockchain-based deep learning frameworks.
Blockchain and ML: A Critical Review on Security [105]	2023	The paper critically reviews the current state of blockchain and ML about security, highlighting the importance of these technologies in enhancing security and discussing future research directions.
Artificial Intelligence for Demystifying Blockchain Challenges: A Survey of Recent Advances [55]	2022	The paper provides a comprehensive survey of recent advances in addressing blockchain problems by leveraging artificial intelligence approaches, providing valuable information and guidance on the design of Blockchain-based systems.
A Survey on Secure and Private Federated Learning Using Blockchain: Theory and Application in Resource-constrained Computing [106]	2023	The paper reviews the challenges, solutions, and future directions for the successful deployment of blockchain in resource-constrained federated learning environments, providing a comprehensive review of various blockchain mechanisms suitable for federated learning.
Blockchain and ML for Communications and Networking Systems [94]	2020	The paper reviews the integration of blockchain and ML for decentralized, secure, intelligent, and efficient network operation and management, exploring the impacts of these technologies on development.
A Survey on Blockchain-Based Federated Learning and Data Privacy [107]	2023	The paper compares the performance and security of various data privacy mechanisms adopted in blockchain-based federated learning architectures, providing an in-depth overview of blockchain-based federated learning, its essential components, principles, and potential applications.
ML in/for blockchain: Future and challenges [15]	2021	Focused on using blockchain for secure and privacy-preserving ML applications. They highlighted the potential of blockchain in providing a secure and decentralized infrastructure for training ML models, protecting sensitive data from breaches, and ensuring the integrity and transparency of ML processes.
Blockchain-Enabled Federated Learning: Challenges and Opportunities [108]	2023	Discussed the challenges and opportunities of blockchain-enabled federated learning. Their work provides a comprehensive overview of how blockchain can enhance the privacy and efficiency of distributed ML.
Survey on the Convergence of ML and Blockchain [4]	2022	Surveyed collaborative ML and blockchain convergence. They investigated different ways of combining these two technologies and examined their fields of application.
Proof of Deep Learning: Approaches, Challenges, and Future Directions [109]	2023	Explores the enhancement of deep learning by increased computational power and its use in blockchain via PoDL. Surveys PoDL approaches their benefits, applications, challenges, and prospects.
Digital Transformation through Blockchain and AI: Emerging Paradigms and Technological Progress [110]	2019	Examines blockchain, AI, ML, and deep learning convergence in the "AI first" shift. Highlights advancements in cloud computing, voice/vision recognition, AI in applications, shift to AI-driven data centers, and edge computing. Focuses on blockchain platform maturation and deep learning efficiency through neural networks.

like [55] and [105], we uniquely focus on scalability and energy-wise considerations in blockchain-ML integration.

- **Critical Evaluation of Security and Data Privacy:** Our survey provides a critical analysis of security and data privacy issues in blockchain-ML convergence, extending the discussions found in works like [94] and [106].

IV. APPLICATIONS AND INNOVATIONS

As the introduction highlights, blockchain offers significant potential when integrated with ML applications. Blockchain provides a decentralized infrastructure [111]. It offers a trustless, automated, and decentralized framework that can streamline how businesses operate, much like how the internet revolutionized the collaborative economy [112]. Figure 6 shows the taxonomy of blockchain for AI [56]. This research

categorized these works regarding the decentralization of AI methodologies and operations, a blockchain infrastructure and types, and the underlying consensus protocols utilized for distributed decentralized transaction validations across underlying networks [56].

In line with this taxonomy, novel approaches combining blockchain and ML continue to emerge, including federated learning, PoL, and PoDL, which can lead to more secure and privacy-preserving ML applications [113]. One of the novel approaches in blockchain and ML integration is the Proof of Learning consensus algorithm. This algorithm channels computational power to train neural network models, empowering ML with consensus building on blockchains [19]. Additionally, federated learning enables decentralized training across multiple devices and can be further secured and optimized using blockchain [114]. The PoDL is another emerging concept that leverages blockchain to ensure the integrity and authenticity of deep learning models [115].

The following sections will delve into state of the art in the confluence of Blockchain and ML.

A. SECURING ML MODELS

In the context of ML, blockchain presents a robust framework for enhancing the security and privacy of ML models and their training data. This section explores various dimensions of how blockchain can be pivotal in securing ML models.

1) ENSURING DATA INTEGRITY

Blockchain can be harnessed to ensure the integrity of training data and the trained models themselves [112]. By recording all transactions related to the model's training process and data on a distributed ledger, blockchain provides a tamper-evident system that safeguards the trustworthiness of ML models [113]. This becomes especially critical in domains like healthcare, where the reliability of ML models holds significant implications [116].

2) PRESERVING DATA PRIVACY

In ML applications, blockchain can preserve data privacy and security. Blockchain's decentralized nature allows data to be stored across a network of nodes, mitigating the risks associated with centralized databases [112]. Employing cryptographic techniques within blockchain ensures that data remains confidential and can only be accessed by authorized entities [117]. This is particularly valuable in domains like healthcare, where patient data privacy is paramount [118], [119].

3) MITIGATING SECURITY THREATS

ML models are susceptible to various security threats, including data poisoning and model inversion attacks [120]. Blockchain can fortify ML models against such vulnerabilities. The ledger records all transactions related to the model's training process and data, ensuring the integrity and security of both the model and the data [112]. The tamper-evident nature of blockchain makes any unauthorized modifications readily detectable, further enhancing the security of ML models [113].

4) PRIVACY-PRESERVING ML AND SECURE DATA SHARING

Privacy-preserving ML methods, such as federated learning, have gained prominence in the heightened data privacy concerns era. Blockchain plays a crucial role in these techniques by enabling data to remain on users' devices, with only model parameters shared and updated on the blockchain [100]. This approach guarantees that raw data remains confidential, safeguarding user privacy [121].

5) SECURE DATA SHARING

In ML, data is a valuable and sensitive resource. Blockchain ensures secure and controlled data sharing, empowering data owners to maintain control over who accesses their data [117]. This controlled data sharing can be pivotal in scenarios where data privacy regulations restrict the sharing of raw data [117].

Integrating blockchain with ML promises enhanced data security and privacy and creates decentralized ML models. These models can leverage data from multiple sources without necessitating the sharing of raw data, thereby preserving data privacy, especially in contexts where regulations limit essential data sharing. While this integration holds excellent potential, addressing its challenges and fully realizing its benefits requires further research and exploration [32].

Challenges around efficiency, incentives, and complexity must be addressed to fully realize the benefits of secure data sharing between blockchain and ML.

In addition to securing ML models, blockchain can enable marketplaces that incentivize data and model sharing. This section explores blockchain-based platforms facilitating secure, transparent data and model exchange.

B. MARKETPLACE THAT INCENTIVIZES DATA SHARING

Several blockchain-based platforms create marketplaces where participants can share data and get rewarded. These decentralized data-sharing ecosystems promote open and diverse datasets for training ML models.

1) BLOCKCHAIN-ENHANCED FEDERATED LEARNING SYSTEM

The system under discussion [122] is designed to focus on the secure sharing of data generated from connected devices within the Industrial Internet of Things (IIoT) paradigm. It introduces an innovative approach to privacy protection by sharing data models instead of revealing the data itself. This ensures that the raw data, owned by the contributors, are stored locally, eliminating the risk of data leakage.

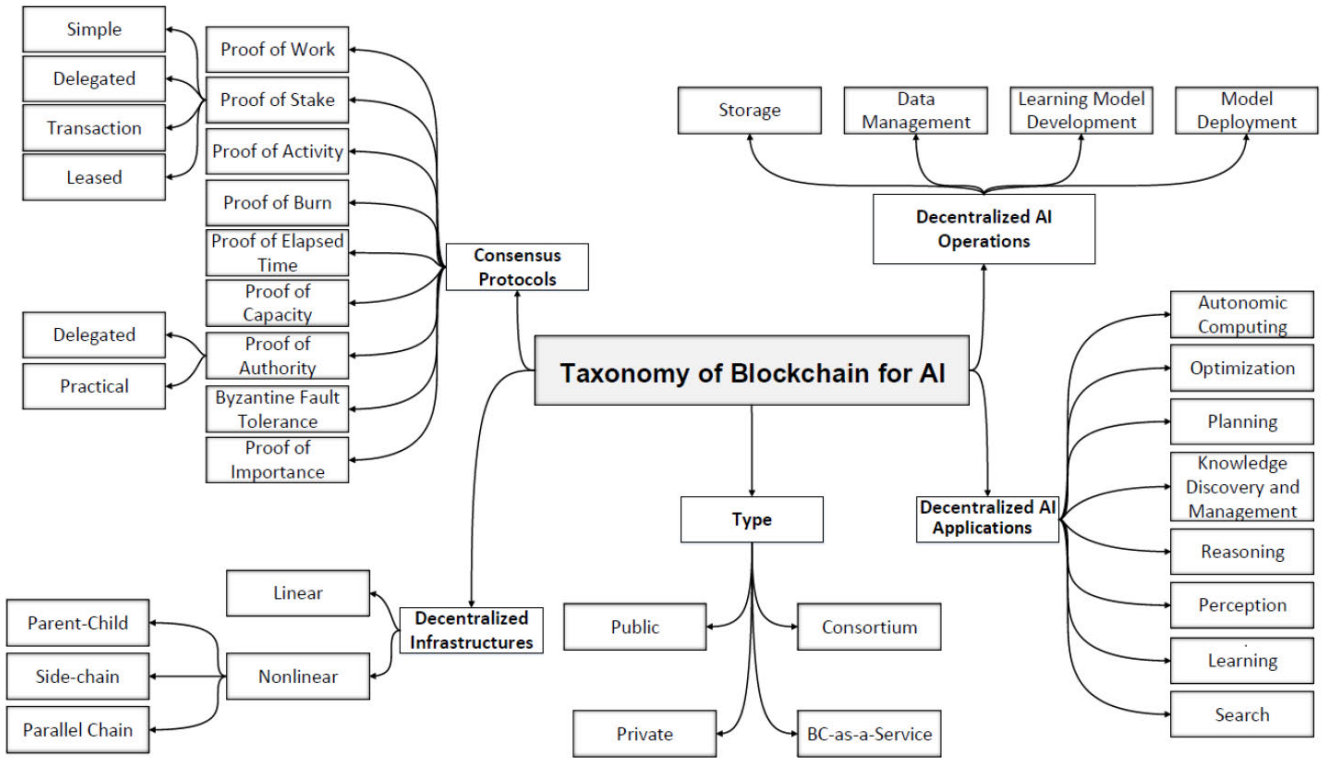

FIGURE 6. Taxonomy of blockchain for AI [56].

As depicted in Fig. 7, the proposed system comprises two primary modules: the permissioned blockchain module and the Federated Learning module. Initially, the Data requester submits a requirement for data sharing to the permissioned blockchain. The blockchain encrypts the information and records it on the blocks. Multiple parties then train the model with new data, and the Data requester obtains the federated data model, which is also logged into the blockchain. A Cached model is employed to verify if the request has been previously processed, thereby preventing redundant operations.

a: PERMISSIONED BLOCKCHAIN MODULE

Positioned at the core, the permissioned blockchain establishes secure connections among all the end IoT devices through its encrypted records. It meticulously manages data accessibility, recording retrieval transactions, data-sharing actions, and all related data-sharing events. Importantly, it does this without recording the raw data, enhancing security.

b: FEDERATED LEARNING MODULE

Situated at the bottom, the FL module facilitates sharing the federated data model learned across multiple decentralized parties. A unique consensus mechanism, known as the Proof of Training Quality (PoQ), is introduced, reducing computing costs and minimizing communication resources.

The evaluation results from the developers of this system reflect the efficacy of the blockchain-empowered data-sharing scheme in enhancing the secure data-sharing process. Moreover, integrating FL into the consensus process of the permissioned blockchain has led to significant improvements in both the utilization of computing resources and the efficiency of the data-sharing process.

Combining Federated Learning with blockchain presents a promising and innovative way to ensure data privacy in data sharing. This approach aligns with the growing need for secure and efficient data handling in the IIoT landscape, and it sets a precedent for future developments in this field.

2) SHARING UPDATABLE MODEL (SUM) ON BLOCKCHAIN

Similarly, frameworks have been proposed for collaboratively sharing updatable ML models on the blockchain. The framework under discussion [123] emphasizes the collective sharing of data from various participants to train an ML model utilizing smart contracts. Specifically, the ML model, trained on the IMDB reviews dataset for sentiment classification, can enhance its performance using collaboratively built datasets. A depiction of the SUM structure can be seen in Figure 8. One of the standout features of this SUM on a blockchain is incorporating incentive mechanisms within a smart contract. These mechanisms promote higher-quality data submission, eliminate incorrect or ambiguous data, and uphold the model's accuracy.

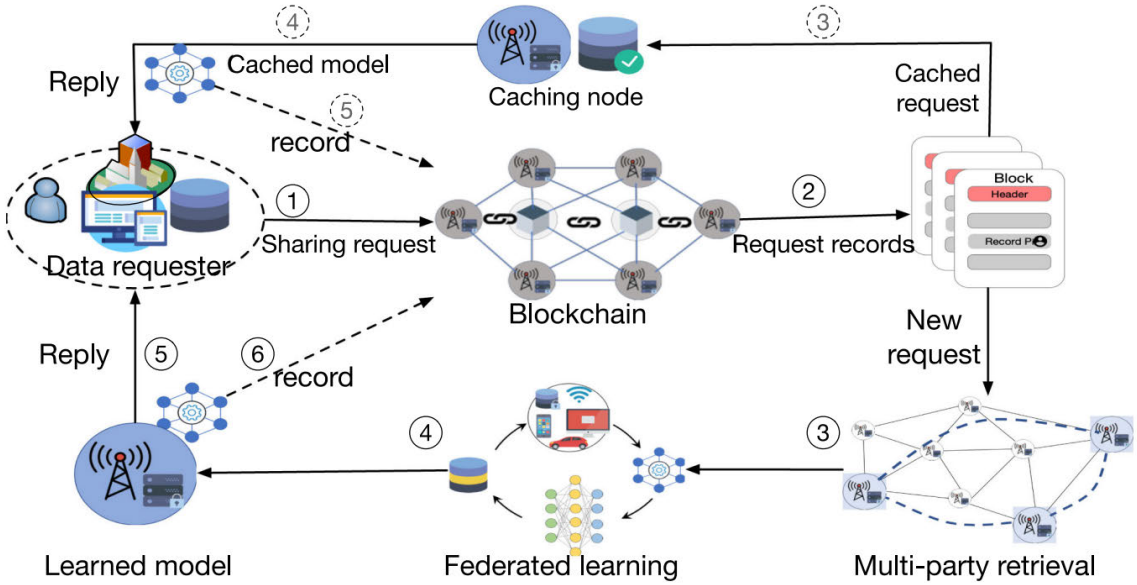

FIGURE 7. Secure data sharing architecture: process from data request to model training and verification, recorded on permissioned blockchain [122].

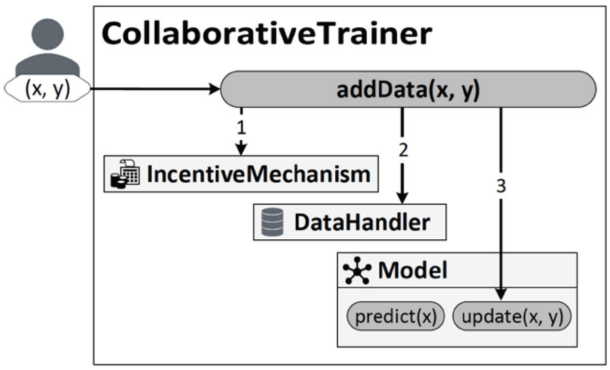

FIGURE 8. SUM structure: three steps from incentive validation to ML model update, including data recording and model training [123].

The SUM framework introduces three distinct incentive mechanisms to foster collaboration and ensure data accuracy:

- **Gamification:** This approach leverages non-financial incentives such as points and badges to recognize and reward reliable data contributors.
- **Rewards Mechanism Based on Prediction Markets:** This monetary reward system incentivizes submitting accurate data. It consists of three phases:
  1) *Commitment Phase:* The provider deposits the reward, defines a loss function $L(h, D)$, uploads test datasets, and sets an end condition.
  2) *Participation Phase:* Participants contribute their datasets and train the ML model. The trained model updates are sent to the smart contract for aggregation.
  3) *Reward Phase:* The smart contract updates each participant's balance b_t through the equation $b_t = b_{t-1} + L(h_{t-1}, D) - L(h_t, D)$. Rewards or penalties are assigned based on data quality.
- **Deposit, Refund, and Take: Self-Assessment:** This mechanism enforces a deposit when contributing data to penalize bad submissions. It consists of four phases:
  1) *Deployment Phase:* An initial ML model h is trained.
  2) *Deposit Phase:* Participants deposit currency d when providing data, influenced by the time interval between updates, as in $d \propto f(\text{time interval})$.
  3) *Refund Phase:* Participants whose data agrees with h have their deposit returned after time t, with $t \propto f(P(hx = y))$, where P is the probability of correctness. The relationship is subject to $t \geq 7$ days.
  4) *Validation Phase:* The smart contract validates data, rewards good contributors, and takes a portion of the deposit from those whose validation result is $hx \neq y$.

Figure 9 illustrates the simulation results of the third incentive mechanism. Good Agents, who contribute quality data, are rewarded, while Bad Agents, who upload incorrect data, eventually deplete their balance. The smart contract effectively mitigates the adverse impact of bad data, maintaining the stability of the model. The SUM framework offers a practical platform for collaborative dataset building, utilizing smart contracts to maintain a continuously updated ML model. Both financial and non-financial incentives contribute to the provision of quality data and the preservation of model accuracy.

3) DEEPCHAIN

Building upon these data-sharing models, the DeepChain prototype tackles security issues in federated learning through blockchain incentives. In this approach [32], data privacy is safeguarded, and a protocol that promotes data sharing is set

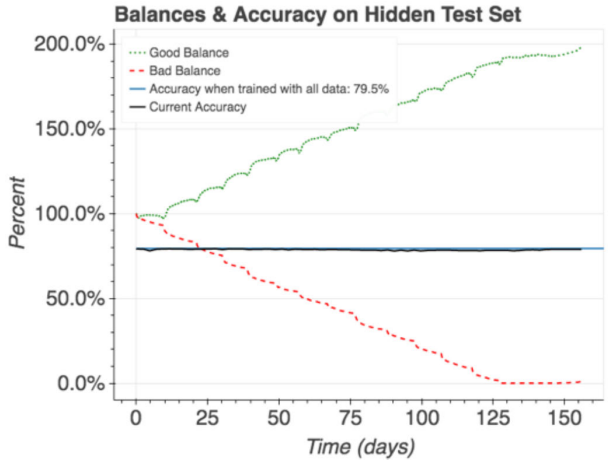

FIGURE 9. SUM simulation experiment: balancing ML model accuracy with reward mechanism for data quality [123].

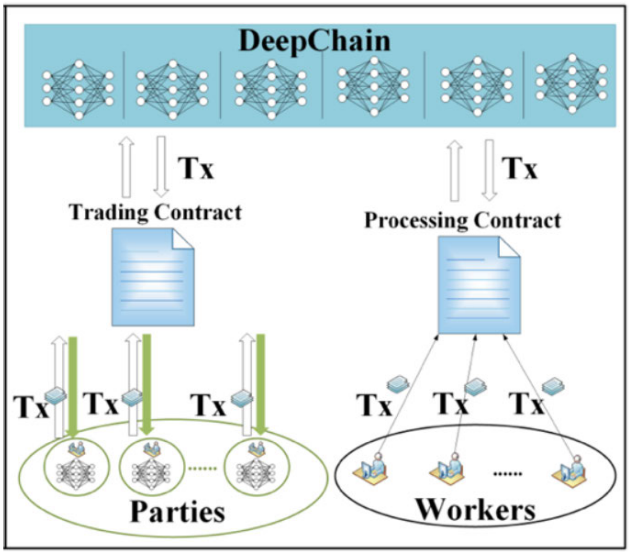

FIGURE 10. DeepChain prototype overview: stakeholder parties, trading contract, and worker roles in secure training with smart contracts [4].

up. The authors of the DeepChain prototype argue that their framework, which is both distributed and secure, tackles a variety of security issues commonly ignored in Federated Learning. Moreover, it incorporates a blockchain-based incentive mechanism to ensure participants act appropriately. Figure 10 provides a schematic representation of this framework. Parties represent stakeholders with similar objectives in this framework but cannot train a model independently. The Trading Contract enables these Parties to upload their local gradients to DeepChain. Workers log transaction data onto the blockchain and receive incentives determined by the Processing Contract. The Trading Contract and Processing Contract are smart contracts within DeepChain, jointly overseeing a secure training procedure. Here, TX stands for transactions.

The research team also constructed a functional prototype of DeepChain. They employed the decentralized ledger Corda V3.0 [124] and the MNIST dataset [125] to create a blockchain for simulation testing.

During the evaluation stage, the model was trained on DeepChain in a multi-party configuration. Various metrics were examined, such as cipher size, throughput, training accuracy, and the total time cost. The results revealed that the accuracy improves as more parties participate in training. However, throughput diminishes with an increase in the number of gradients, and the time required for training also escalates as the number of participating parties increases.

DeepChain's primary innovations include: First, it introduces an incentive structure to promote collaborative involvement in training deep learning models and disseminating the resulting local gradients. Second, DeepChain ensures these local gradients' confidentiality while providing a transparent and auditable training workflow.

4) LEARNINGCHAIN MARKETPLACE

Other decentralized marketplaces like LearningChain illustrate blockchain's potential for enabling secure and private data sharing for ML. The demand for robust ML algorithms for large-scale, distributed data is pressing in the extensive data landscape. LearningChain emerges as a solution that transcends the limitations of traditional master-worker distributed ML algorithms, which rely on a central trusted server. The LearningChain Framework [126] is decentralized, privacy-preserving, and secure, supporting both linear and nonlinear learning models. It employs a decentralized Stochastic Gradient Descent (SGD) algorithm, integrates differential privacy mechanisms, and deploys an l-nearest aggregation algorithm to counteract Byzantine attacks.

The framework operates in three pivotal phases. First, the "Blockchain Initialization" phase sets up a P2P network involving nodes and data owners. Second, the "Local Gradient Calculation" phase provides data owners with model replicas for local gradient computation, adding a noise factor for privacy. Lastly, the "Global Gradient Aggregation" phase identifies a winning node through a PoW challenge and updates the global model.

LearningChain maintains a trade-off between user privacy and model accuracy. Reduced privacy correlates with increased test errors, affirming the framework's efficiency and effectiveness in utilizing blockchain.

The architecture of LearningChain, as shown in Figure 11, is flexible and does not impose specific blockchain requirements. The participants are data owners, like coin owners in cryptocurrency systems or computing nodes, similar to miners. These computing nodes assist data owners in learning the model and are compensated based on their contributions. The roles are fluid; a computing node can also be a data owner if it has relevant data, and vice versa.

Beyond architectures for data sharing, blockchain also shows promise for marketplaces that incentivize model sharing.

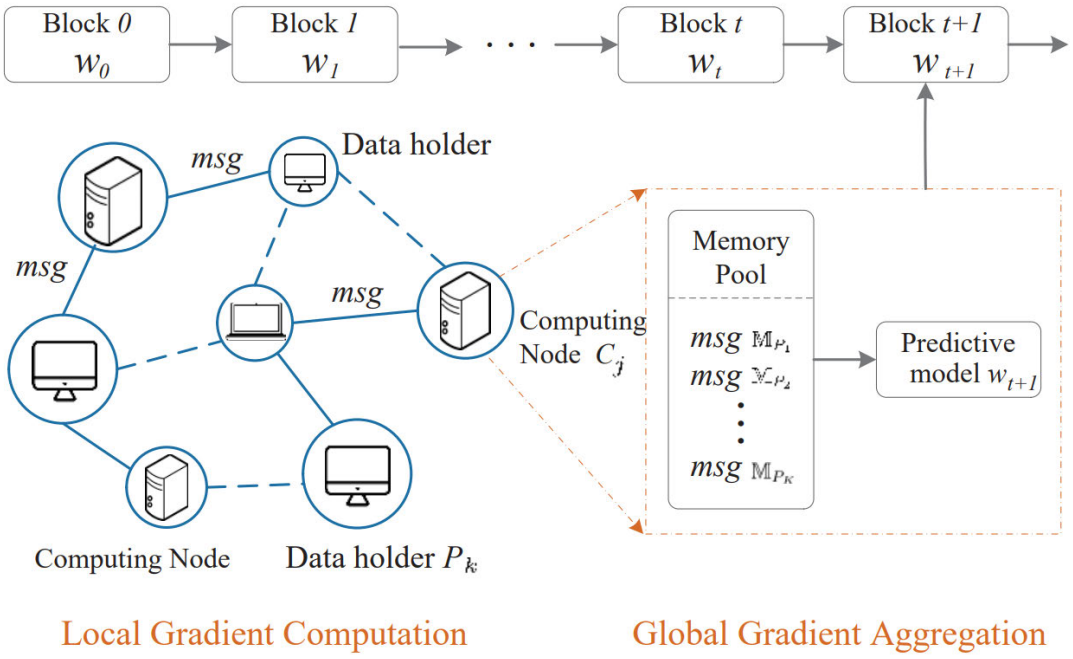

FIGURE 11. LearningChain architecture [126].

C. MARKETPLACE THAT INCENTIVIZES MODEL SHARING

Blockchain systems can also incentivize the sharing of pre-trained ML models. Platforms that enable model exchange allow developers to build on existing models rather than training from scratch. This section discusses projects that use blockchain for model sharing and exchange.

1) DANKU CONTRACT

Built upon an Ethereum-blockchain-based marketplace, this original and classic protocol is designed to evaluate and exchange ML models, serving as a trustless platform for supply and demand [28]. The demander (or organizer), possessing the dataset, poses a question, while the supplier (or submitter) provides a well-trained ML model to address the issue. In the final stage, the supplier with the top-performing model receives a payout from the demander. To ensure data privacy, hashed data is disclosed through a contract, and all transactions are recorded on the blockchain. Figures 12 to 16 depict the structure of this marketplace, outlining five phases that ensure its successful operation.

Several innovative features characterize this marketplace:

- **Automation and Anonymity**: Unlike the Ethereum blockchain, which necessitates reputations, the ML model exchange here is automated and anonymous. This is achieved through cryptographic verification enforced by the protocol.
- **No Requirement for a trusted 3rd party**: The smart contract's ability to automatically validate solutions submitted by submitters removes any uncertainty regarding the correctness of the solution. Consequently, organizers can solicit AI solutions globally, and submitters can directly monetize rewards with their ML models.
- **Data Privacy**: The DanKu contract assists the organizer in creating a cryptographic dataset for training and testing, utilizing the sha3-keccak hashing function, thereby ensuring data privacy [28].

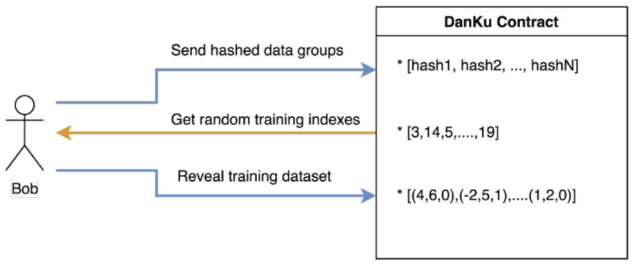

FIGURE 12. Initialization phase: contract formulation with problem definition, dataset allocation, and ethereum wallet integration [28].

The developers emphasize in their paper that the Smart Contract fosters a market where individuals skilled in solving ML problems can monetize their abilities. Simultaneously, any organization or software agent needing an AI solution can seek global solutions. This approach encourages the creation of superior ML models and enhances AI accessibility for various companies and software agents [28], marking a significant contribution to ML's advancement.

Beyond evaluating and exchanging models, novel consensus mechanisms like Proof of Learning demonstrate how blockchain can enable decentralized development and selection of ML models.

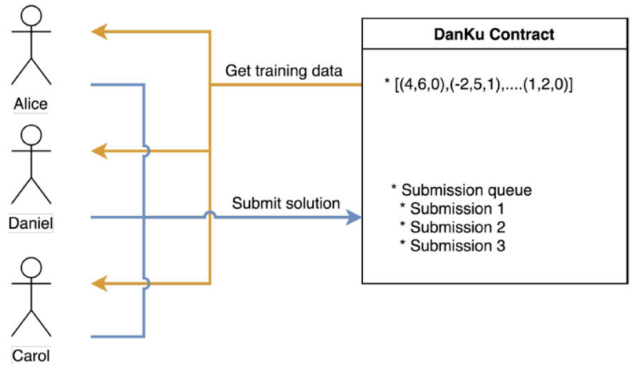

FIGURE 13. Submission phase [28]: during this stage, submitters present their trained ml models as potential solutions to the problem.

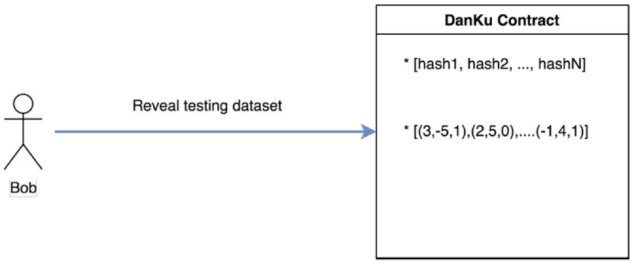

FIGURE 14. Test dataset disclosure phase [28]: The organizer unveils the hashed testing data sets at this stage.

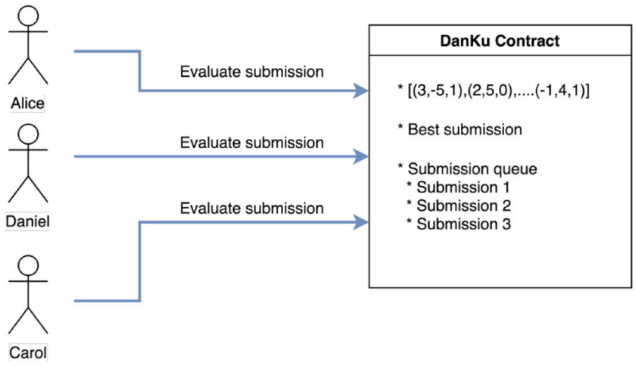

FIGURE 15. Evaluation phase: model assessment using evaluation function and marking of top-performing or first-evaluated models [28].

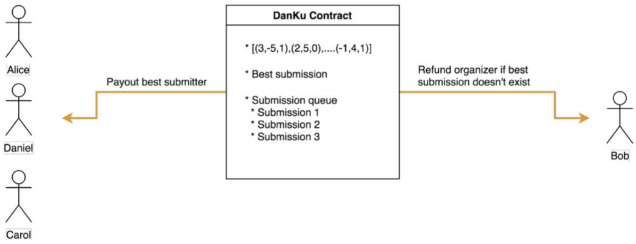

FIGURE 16. Finalization phase: reward allocation to best model submitter or return to organizer if no model passes evaluation [28].

2) PROOF OF LEARNING

Another innovative application of ML in the blockchain is the concept of proof of learning. This distributed consensus protocol ranks ML systems for a given task. This protocol aims to reduce the computational waste of hashing-based puzzles and create a publicly distributed and verifiable database of state-of-the-art machine-learning models and experiments. ML can significantly augment blockchain, enhancing its security, efficiency, and utility in various applications. In this mechanism, the right to create a block is granted based on the performance of ML tasks. This provides a practical alternative to the energy-consuming Proof of Work and promotes the development and application of ML models [127].

In the Proof of Learning mechanism, the nodes in the blockchain network participate in ML competitions. The node that achieves the best performance in the competition earns the right to create the next block. This approach ensures the block creation process's fairness and encourages the nodes to improve their ML capabilities [127].

Proof-of-learning can be extended to selecting optimal machine-learning models through Smart Contracts. In this setup, ML competitions are held where the participants compete to develop the most accurate model. The performance of these models is then evaluated and validated through smart contracts, ensuring a transparent and fair model selection process. This approach not only incentivizes the development of high-quality machine-learning models but also promotes the use of blockchain in machine-learning applications.

Proof of Learning (PoL) adds a unique dimension to the ecosystem of consensus mechanisms by specifically incentivizing the development of machine learning (ML) models within a decentralized framework. While traditional mechanisms like Proof of Work (PoW), Proof of Stake (PoS), and Delegated Proof of Stake (DPoS) primarily focus on securing blockchain networks and ensuring their resistance to tampering, PoL aims to democratize access to ML. It fosters innovation by rewarding the creation and improvement of ML models, thereby aligning the interests of participants with the advancement of ML technologies. This approach not only stimulates growth in the field of ML but also offers an alternative to the computationally intensive tasks typically associated with blockchain consensus mechanisms, potentially reducing computational waste and focusing resources on productive ML model development.

a: PROOF-OF-LEARNING AND SMART CONTRACTS IN ML MODEL SELECTION

Integrating blockchain and ML has opened up new possibilities for optimizing model selection in ML, mainly through smart contracts and novel consensus mechanisms such as Proof of Learning [127].

In a typical blockchain network, miners compete to solve a computational puzzle, a process known as PoW [10]. However, Proof of Learning replaces this process with ML model training [127]. Miners train models and submit their performance values. Full nodes in the network then validate these performance values. The miner with the best-performing model can write the next block in

the blockchain. This approach promotes the efficient and sustainable use of computational resources and energy for ML training [127].

Smart contracts, self-executing contracts with the terms of the agreement directly written into code [47], play a crucial role in this process. In the context of PoL, smart contracts can be used to define the rules of the model training competition, including the evaluation metrics and the reward distribution.

When a miner believes they have a model that meets the required performance criteria, they can submit it to the network. The smart contract then automatically verifies the model's performance and, if it meets the requirements, updates the blockchain and awards the miner their reward.

Proof of Learning provides several advantages as a consensus mechanism tailored for ML but also poses some challenges in implementation. It provides a transparent, auditable, and fair process for model selection [127]. It also allows for the decentralized execution of ML tasks, increasing security and reducing reliance on a single central authority.

b: ADVANTAGES OF PROOF OF LEARNING

Proof of Learning offers several potential benefits for decentralized systems due to its blockchain nature, which ensures transparency, security, and privacy:

1) Incentivizing participation: By rewarding participants for their contributions to developing high-quality ML models, Proof of Learning encourages increased network participation and expertise sharing [128].
2) Improving model quality: As participants compete to produce the best ML models, they are motivated to put forth their best efforts, resulting in accurate and generalizable models [127].
3) Decentralizing ML: Proof of Learning enables the development and deployment of ML models decentralized, promoting a more distributed and democratic approach to ML [128].
4) Reducing computational waste: By using Proof of Useful Work instead of Proof of Work, Proof of Learning minimizes computational waste, directing resources toward valuable outcomes [127].
5) Supplying computational power to AI-based systems: Proof of Learning can channel the computational energy dedicated to mining activities into supporting AI-based systems, facilitating the advancement of artificial intelligence [127].

c: DISADVANTAGES OF PROOF OF LEARNING

1) Complexity of implementation: Proof of Learning introduces more sophisticated mechanisms than traditional consensus algorithms, making it more difficult to implement and maintain in decentralized systems [128].
2) Scalability issues: The competition for accurate ML models may require significant computational resources, potentially leading to scalability concerns as the network grows [127].
3) High storage requirements: Storing ML models, training data, and test data within the blockchain may impose substantial storage burdens, impacting overall system efficiency [128].
4) Network latency: The submission of ML models, training parameters, and test data can introduce network delays, potentially affecting the overall performance of the blockchain [127].
5) Heterogeneous models and datasets: As participants compete with diverse ML models and datasets, achieving a consensus on the best model can be challenging, potentially resulting in longer block validation times [127].
6) Potential centralization risks: As powerful entities may have access to better hardware, data, and expertise, there is a risk that the Proof of Learning system could inadvertently centralize control over the network [128].
7) Quality assurance challenges: Ensuring the quality and integrity of the models and data used in the Proof of Learning system can be difficult and resource-intensive, requiring constant monitoring and validation [127].

d: COMPARISON OF PROOF OF LEARNING WITH THE OTHER CONSENSUS ALGORITHMS

Proof of Learning is a consensus algorithm that fosters the decentralized development of ML models. Unlike traditional consensus algorithms that rely on various mechanisms like computational challenges or stake ownership to primarily secure blockchain networks, Proof of Learning uniquely utilizes the ML process, employing models' training as a basis for achieving consensus within the network [127].

Table 7 illustrates a comparison of Proof of Learning with other consensus algorithms:

Systems like WekaCoin illustrate how Proof of Learning can be applied, using blockchain to incentivize model training and validation.

e: WEKACOIN AND PROOF-OF-LEARNING

WekaCoin [127] operates on a peer-to-peer network utilizing a unique blockchain framework. This framework incorporates distinct roles: ''trainers'' develop ML models using new datasets provided by ''suppliers,'' ensuring that the training involves previously unseen data. ''Validators'' then assess these models based on supplier-defined performance metrics and contribute to the blockchain by validating new blocks. The highest-performing trainers are rewarded with WekaCoins, while validators receive compensation through transaction fees and the issuance of new WekaCoins. This structure ensures the integrity of transaction blocks and builds a verifiable database with a comprehensive record of transactions, ML models, and related experiments.

The paper introduces the concept of Proof of Learning, which harmonizes two distinct tasks: validating transactions in a distributed ledger and storing ML models and experiments in a decentralized database. Proof of Learning

TABLE 7. Comparison of proof of learning with other consensus algorithms.

Consensus Algorithm	Energy Wise Resource Use	Security	Scalability	Transaction Speed	Byproduct
Proof of Work (PoW)	Low	High	Low	Low	Cryptocurrency
Proof of Stake (PoS)	Medium	Medium	Medium	Medium	Cryptocurrency
Proof of Authority (PoA)	High	Low	High	High	N/A
Delegated Proof of Stake (DPoS)	Medium	Medium	High	High	Voting Power
Practical Byzantine Fault Tolerance (PBFT)	High	High	Low	Medium	N/A
Proof of Elapsed Time (PoET)	Medium	Medium	Medium	Medium	Random Leader Selection
Proof of Capacity (PoC)	Medium	Medium	High	Low	Data Storage
Proof of Burn (PoB)	High	Medium	High	High	Coin Burning
Proof of Authority with Identity (PoA-ID)	High	Medium	High	High	Identity Verification
Proof of Learning (PoL)	Low	Medium	Medium	Medium	AI Models
Proof of Deep Learning (PoDL)	Low	Medium	Medium	Medium	Trained Deep Learning Models
Proof of Importance (PoI)	High	Low	High	High	Reputation Score
Proof of Reputation (PoR)	High	Low	High	High	Reputation Score
Proof of Weight (PoWt)	High	Low	High	High	Node Weight
Proof of Time (PoT)	High	Medium	High	High	Time Ownership
Proof of Activity (PoAcy)	High	Medium	High	High	Account Activity
Proof of Authority Round-Robin (PoA-RR)	High	Low	High	High	Round Robin Selection
Proof of Believability (PoB)	High	Low	High	Medium	Believability Score
Proof of Personhood	Medium	High	Medium	Low	Identity Credentials
Proof of Space (PoSpace)	High	Medium	High	Medium	Storage Utilization
Proof of Storage (PoStorage)	High	High	Medium	Medium	Data Availability
Proof of Capacity (PoC)	High	Medium	High	Medium	Storage Efficiency
Proof of Space-Time (PoST)	High	High	Medium	Medium	Time-Stamped Storage

is designed to efficiently combine these tasks, drawing inspiration from mechanisms such as reCAPTCHA.

The Proof of Learning workflow involves three primary actors:

- **Suppliers:** These entities host ML competitions.
- **Trainers:** They are responsible for training models and submitting solutions for available tasks.
- **Validators:** Validators evaluate models on test data, reach a consensus on winning models, and propose new blocks to the blockchain.

WekaCoin implements three distinct types of transactions:

- **Standard Transactions:** These transactions facilitate the transfer of WekaCoins between users.
- **Task Publication Transactions:** Suppliers propose ML competitions using these transactions.
- **Model Transactions:** Trainers employ these transactions to submit solutions for specific tasks.

Suppliers can publish ML tasks, including details such as training data, rewards, performance metrics, and testing release schedules. A task publication transaction also includes a standard transaction for transferring the reward and a task hosting fee. Suppliers who fail to release the testing data as scheduled face penalties in the form of lost tokens.

WekaCoin transactions are validated by the Proof-of-Learning process, illustrated in Figure 17. There are three types of actors involved in this process: 1) suppliers who host ML competitions; 2) trainers who train and submit models for any available task; and 3) validators who evaluate the models on the test data, reach consensus about the winning system, and propose new blocks to the chain.

Trainers select tasks and train ML models using training datasets. Models are submitted through model transactions, which include essential information like the model's hash, a timestamp, and performance scores based on the training

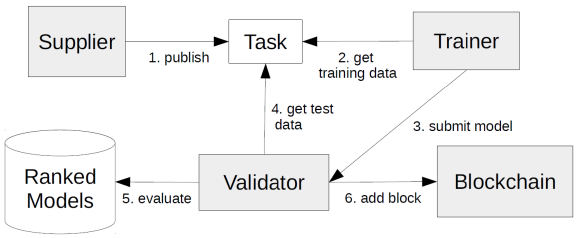

FIGURE 17. Illustration of wekacoin transaction validation process [127].

data. While trainers can submit multiple models, a nominal participation fee discourages spam submissions. Models are considered fully submitted after the release of the test data.

Validators, selected through a cryptographic sortition method, are crucial in proposing new blocks by evaluating models for an assigned task. Models must meet specific criteria, including successful download, execution, and correct performance scores. Validators create candidate rankings based on model performance and share them among themselves.

Genuine validators reach a consensus on the proposed block and rankings. Transactions are subsequently added to the block, including:

- Reward transfers to the owners of the top-performing models.
- Distribution of task hosting fees to valid validators.
- Redemption of participation fees for trainers with top-performing models.
- Compensation for validators involved in handling disqualified models.
- Distribution of newly minted WekaCoins among valid validators.

Metadata for validation purposes is included in each block, digitally signed by corresponding validators. This ensures

transparency and facilitates the verification of the block's validation process.

WekaCoin combines blockchain with ML tasks, allowing decentralized model training, validation, and rewards. Proof of Learning aligns two unrelated tasks, creating a distributed repository of ML models and datasets while securing the blockchain. Validators play a crucial role in maintaining the system's integrity, and transactions ensure fair compensation for participants.

As an extension of Proof of Learning, Proof of Deep Learning is an emerging concept leveraging blockchain's capabilities for optimizing and securing deep learning models [129].

3) PROOF OF DEEP LEARNING

One of the related papers on this topic, titled ''Proof of Learning (PoLe): Empowering ML with Consensus Building on Blockchains,'' [19] presents a novel consensus mechanism, PoLe, that leverages blockchain to optimize neural networks (NN) while achieving consensus. The paper introduces PoLe, a consensus mechanism that directs computational power spent on consensus toward optimizing NN. PoLe involves releasing training/testing data to the blockchain network, where consensus nodes train NN models as PoLe. The paper compares PoLe with Proof of Work and demonstrates its benefits regarding stability in block generation rates and efficient transaction processing. A cheating prevention mechanism, the Secure Mapping Layer (SML), has also been introduced.

The proposed system comprises a decentralized peer-to-peer network comprising data and consensus nodes. Data nodes initiate ML tasks, including training datasets, model specifications, accuracy requirements, and rewards. These tasks are broadcast to the network and added to the global task list. Consensus nodes, or miners, compete in training models that meet data nodes' requirements and receive rewards. The blockchain also functions as a decentralized data store. Figure 18 depicts an overview of this system.

Data nodes commission ML tasks, including encrypted training datasets, model specifications, accuracy requirements, and time limits. They sign timestamps to prevent forgery and release test data only after receiving trained models. This ensures the test data's integrity and prevents malicious nodes from using it prematurely.

Consensus nodes, or miners, supply computational power to the network and compete for tasks issued by data nodes. The PoLe consensus algorithm directs their behavior. Miners select the highest-value task from the task list, initialize model parameters, and create the SML based on the current block hash. They optimize the specified ML model and broadcast a new block when the minimum training accuracy is achieved (see Algorithm 1 in Figure 19).

The winning block and ommer blocks are added to the miner's blockchain. Blocks consist of headers and bodies, with headers containing block ID, winner's ID, selected task, previous block hash, and more. Block bodies store data in a Merkle Tree structure, including uncompleted tasks, newly collected tasks, encrypted data, transactions, and test datasets (see Figure 20).

The paper introduces an encryption mechanism to prevent nodes from starting training prematurely. This mechanism employs inner-product functional encryption and query vectors generated based on the previous block's hash. The inner product between data feature vectors and query vectors is used for decryption.

The PoLe design encourages accurate estimation of training time by data nodes, as overestimation leads to lower task priority. Consensus nodes receive rewards from data nodes and additional rewards by referring to ommer blocks. The SML prevents malicious nodes from starting training early (see Algorithm 2 in Figure 21). PoLe also enhances security by making it costly for attackers to manipulate data.

This paper introduces PoLe as a novel consensus mechanism that optimizes neural networks while achieving consensus on a blockchain network. It provides a detailed system architecture, tools for data node interactions, consensus node behaviors, and secure data storage. Additionally, it discusses incentives and security measures to ensure the system's integrity.

4) HAWK

Hawk [130] primarily addresses the issue of transactional privacy, often overlooked in existing systems where transactional data is openly and publicly recorded on the blockchain. This transparency poses a significant barrier to adopting smart contracts in sectors like finance, where transactional data is susceptible. As depicted in Figure 22, a Hawk program consists of a private section designed to secure participants' data and monetary exchanges and a public section that does not interact with data or funds. The private area conceals the flow and amounts of money by transmitting encrypted data to the blockchain. The manager and the user can execute the protocol's rules to maintain financial integrity. While the user can contribute regular data and currency to the protocol, the manager oversees the transaction without influencing its results and faces penalties for dishonest actions. The protocol is divided into a private section, which conceals financial and other sensitive data through encryption, and a public section, which includes non-sensitive elements like incentive mechanisms.

In [130], the authors introduce a decentralized smart contract system that keeps financial transactions private, not storing them in plaintext on the blockchain. Hawk allows programmers to create private smart contracts that establish rules for economic fairness without requiring manual cryptography implementation, as Hawk automatically generates the necessary cryptographic protocols. Both users and managers (or monitors) can execute Hawk contracts. Notably, the manager does not require trust, as penalties are imposed for contract abortions, compensating the user.

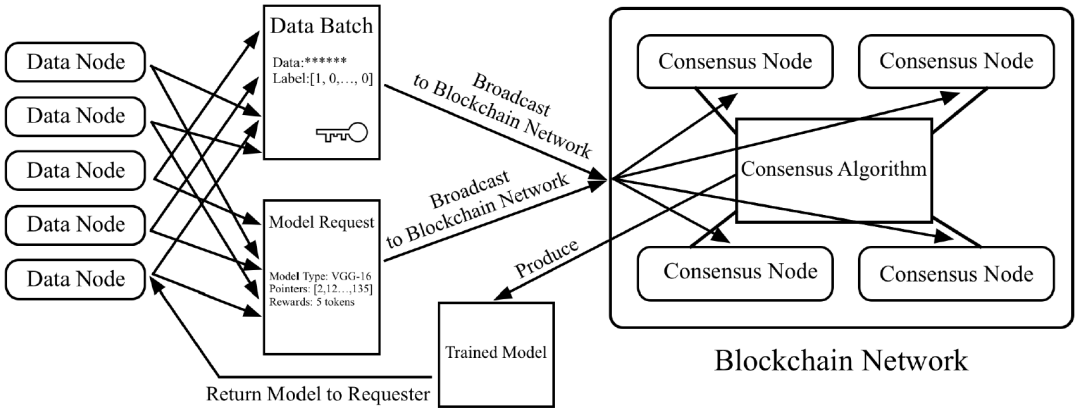

FIGURE 18. The overview of the proposed system in the PoLe [19].

Algorithm 1: The PoLe Consensus Algorithm

Input: task_list: the task list stored in previous block
blk_chain: the blockchain
PHS: the hash value of the previous block

```
task ← PopMostValuable ( task_list )
train_data ← CollectData ( task.data_pointers )
SMLayer ← CreateSMLayer(PHS)
sm_model ← InsertLayer(task.model, SMLayer)
received_blks ← []
while t < time_max do
    train sm_model for one step
    calculate train_accuracy
    if train_accuracy ≥ task.required_accuracy then
        blk ← CreateBlock(sm_model)
        broadcast blk to other consensus nodes
        Append(received_blks, blk)
        break
    end
    if received a new solution blk then
        Append(blk_chain, blk)
        break
    end
    if received the test data block then
        break
    end
    t++
end
blk ← CreateBlock(sm_model)
broadcast blk to other consensus nodes
new_blks ← WaitAndReceiveBlocks()
received_blks ← Append(received_blks, new_blks)
received_blks ← Append(received_blks, blk)
sort received_blks in descending order of test_accuracy
for each blk in received_blks do
    if VerifyBlock(blk, PHS, train_data, test_blk) =
     True then
        Append(blk_chain, blk)
        return
    end
end
```

FIGURE 19. Algorithm 1: The PoLe consensus algorithm [19].

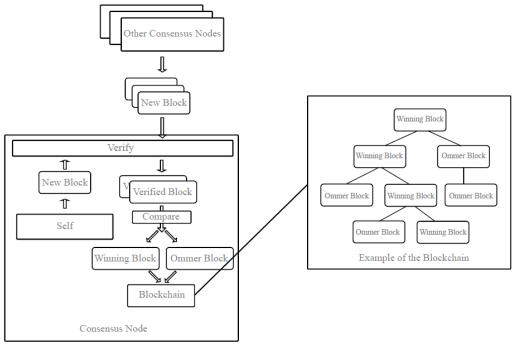

FIGURE 20. Validating and adding blocks to the blockchain in the PoLe consensus algorithm [19].

Algorithm 2: VerifyBlock (blk, PHS, train_data)

Input: blk: received new block
train_data: the training dataset
train_blk: the test data block
PHS: the hash value of the previous block

Output: verified: True or False

```
SMLayer ← CreateSMLayer(PHS)
sm_model ← InsertLayer(task.model, SMLayer)
train_accuracy ← CalcAccuracy(sm_model, train_data)
if train_accuracy ≥ required_accuracy then
    if blk.secure_mapping = PHS and blk.timestamp <
     test_blk.timestamp then
        return verified = True
    end
end
return verified = False
```

FIGURE 21. Algorithm 2: VerifyBlock (blk, PHS, train_data) in the PoLe Consensus Algorithm [19].

According to the authors, Hawk is the first system to define a cryptographic model for blockchain formally.

5) DINEMMO

The architecture, known as Decentralized Incentivization for Enterprise Marketplace Models (DInEMMo) [131], is an extensive marketplace that facilitates data and model sharing. It operates at the intersection of decentralized AI and blockchain technologies. This platform allows users to upload new ML models or improve existing ones by contributing their datasets. Fig. 23 depicts a diagram illustrating this marketplace. The paper demonstrates the marketplace's utility through a medical diagnostic case involving data from two hospitals. The DInEMMo framework comprises three primary modules—BLOCKCHAIN, KEY STORE, and ML SERVICE. Initially, the Controller collects users and activates the Model Manager, which collaborates with the

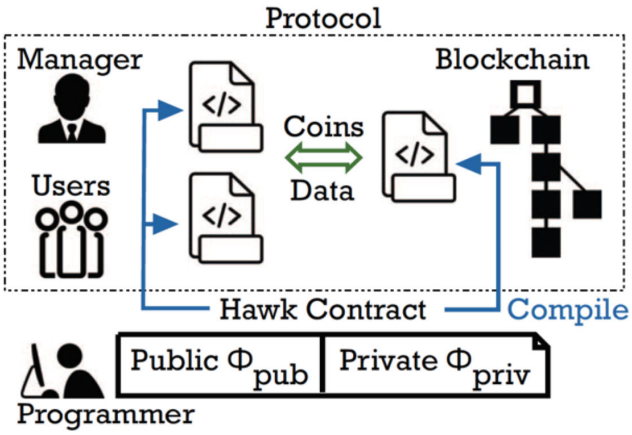

FIGURE 22. Hawk architecture: roles of manager and user in protocol execution with private and public sections for data security and transparency [130].

Marketplace to identify a fitting model through a validation process. Next, the Service Adapter communicates with the ML SERVICE module to acquire the privacy-preserving key of the selected model from the KEY STORE module. The Service Adapter then updates the newly trained ML model, and the Model Manager uploads it to the Model Store for subsequent utilization. The Incentive Engine calculates the rewards for those who contributed data, while the Pricing Engine establishes the model's price.

In this context, DInEMMo offers a secure platform for hospitals to exchange data and collaboratively refine ML models without exposing sensitive patient information. Recall, accuracy, and precision metrics determine the model's cost. The architecture supports several key functionalities:

1) Users can choose the most suitable model for their data using a built-in validation mechanism.
2) Contributors who enhance the performance of an ML model are rewarded based on a unique Incentive Engine (IE), differing from the DanKu contract's first-place-only reward system.
3) Users looking to solve an ML problem can purchase an appropriate model from the marketplace, with the Pricing Engine (PE) determining the model's cost.

The creators of DInEMMo claim it to be the first platform to reward ML model contributors equitably. It considers both domain-specific properties and ML attributes when allocating rewards. The architecture also features configurable smart contracts that serve multiple purposes, such as representing ML and use case attributes, generating new or enhanced ML models based on user contributions, determining model pricing based on user policies, and calculating incentives for model owners and co-contributors.

V. CHALLENGES AND FUTURE RESEARCH DIRECTIONS

The integration of blockchain and ML technologies presents a myriad of opportunities and challenges. This section discusses the potential benefits and hurdles associated with the fusion of these two technologies.

A. CHALLENGES

Despite the promising opportunities, integrating blockchain and ML also presents several challenges that must be addressed to exploit these technologies' potential.

1) TECHNICAL CHALLENGES

a: SCALABILITY AND PERFORMANCE

Both blockchain and ML systems can be resource-intensive, and their integration can exacerbate scalability issues. Developing scalable solutions that can handle large datasets and complex learning tasks without compromising performance is a crucial challenge [131], [132], [133].

b: CONSENSUS MECHANISMS FOR ML

Traditional blockchain consensus mechanisms may not be suitable for ML applications, especially those that require real-time decision-making. Developing new consensus mechanisms that can support the specific needs of ML is a significant area of research [30], [132].

c: INTEROPERABILITY AND TECHNICAL COMPLEXITY

ML models and blockchain systems are often developed independently using different protocols and standards. Ensuring interoperability between these systems when integrating them is a technical challenge [9], [134].

2) DATA SECURITY AND PRIVACY

a: SECURITY

Security is a primary concern in any technology, and the intersection of blockchain and ML is no exception. While blockchain can provide security to ML models, the underlying data used by these models can still be vulnerable to attacks. Research is ongoing in developing robust privacy-preserving ML algorithms that can work effectively in a blockchain environment [135]. Ensuring the security of the integrated system requires addressing the vulnerabilities in both blockchain and ML components.

b: PRIVACY

The significance of safeguarding data privacy is paramount. Data contributors may hesitate to disclose their information or transactions on a public blockchain, where any user possessing the key can access the data. Several solutions to this challenge are as follows:

1) As suggested by the authors in [123], one approach is to avoid directly submitting data to the smart contract. Instead, contributors could submit encrypted inputs or provide inputs to a concealed model that operates behind an Application Program Interface (API), thereby not being publicly accessible.
2) Utilizing blockchain effectively can standardize and monitor the usage of stored data, thereby preventing

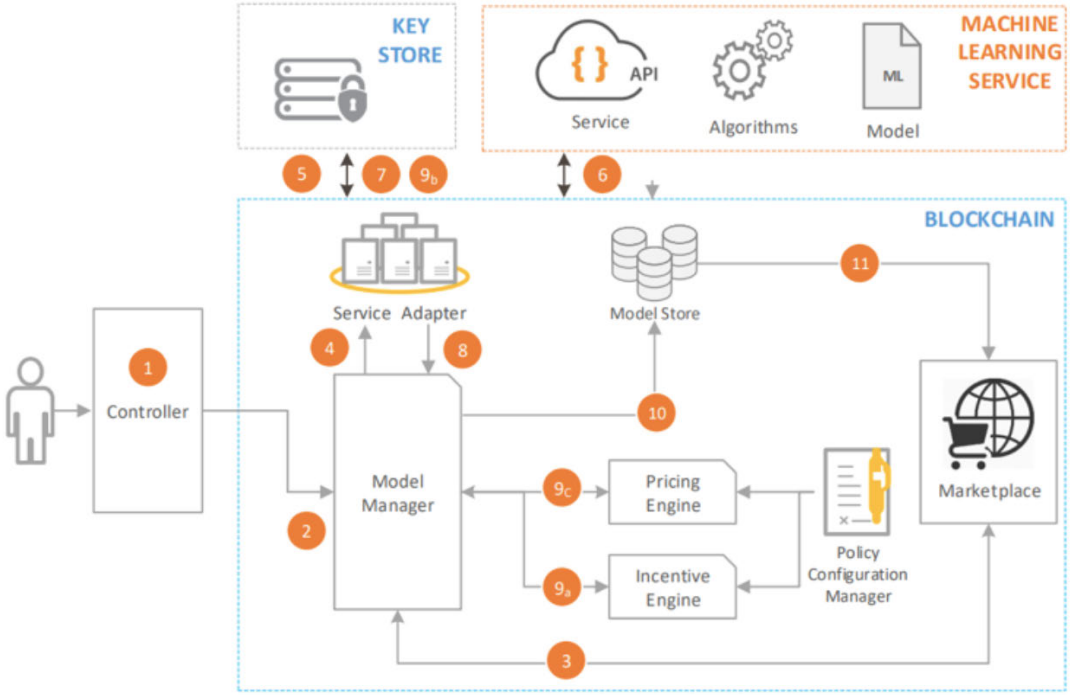

FIGURE 23. DInEMMo framework: interplay of BLOCKCHAIN, KEY STORE, ML SERVICE modules, with focus on model selection, training, incentive calculation, and pricing [131].

misuse and privacy breaches. Smart contracts can be engineered to disclose hashed data or data models rather than raw data, as demonstrated in [28] and [122].

3) As noted in [28], the model weights in DanKu contracts are not fully anonymized. To address this, homomorphic encryption techniques could be integrated into the protocol to ensure the anonymity of the models submitted to the smart contracts.
4) Most existing protocols and marketplaces have yet to address transaction privacy. Further research in this area is warranted, as indicated in [130].

3) REGULATORY AND ETHICAL ISSUES

Integrating blockchain and ML raises regulatory and ethical issues, such as data ownership, the right to explanation, and algorithmic bias. Addressing these issues is crucial for the responsible deployment of these technologies [39], [136].

4) DATA MANAGEMENT

a: DATA ACCESSIBILITY CHALLENGES

Access to relevant data sets for training ML models can be limited, especially when the ideal data is proprietary. This restriction can lead to high costs in data collection or hinder the development of effective models. Conversely, data set holders may struggle to find suitable ML models. A potential solution is a marketplace that facilitates data exchange and ML models, bridging the gap between data availability and model accessibility. Such a platform could foster collaboration and efficiency in ML development [4].

b: MALICIOUS CONTRIBUTORS AND ERRONEOUS DATA

Even with access to updatable data for the ML model, not all contributors may be trustworthy. Spammers or malicious providers might upload incorrectly labeled data, leading to suboptimal training results. The updated ML model may perform poorly without properly filtering this erroneous information [4].

5) ENERGY WISE RESOURCE UTILIZATION

Optimizing energy-wise utilization of resources is a critical challenge in blockchain and ML systems. Task scheduling is essential for efficient resource usage but becomes more complex when considering energy demands. While ML has been applied for energy-aware scheduling, most techniques predict resource needs rather than optimize schedules [137].

Furthermore, proof-of-work consensus in blockchain networks is computationally intensive, raising sustainability concerns [2]. Integrating ML in the consensus process provides opportunities for more energy-wise alternatives. For instance, ML can predict energy consumption in cloud data centers and optimize resource allocation accordingly.

However, existing energy management limitations in cloud computing warrant innovative solutions to enhance performance while minimizing energy demands. As ML and blockchain integration mature, a core focus should be leveraging ML to optimize task scheduling, resource usage, and consensus protocols for improved energy efficiency.

Studies have also explored the accuracy vs. energy trade-offs in ML models. Significant energy savings can be

achieved with minimal loss in model accuracy by careful tuning and optimization [138]. This highlights the importance of considering sustainability alongside model performance.

A key research direction is developing ML techniques to optimize resource usage in blockchain networks and cloud data centers for energy-efficient operation.

6) FINANCIAL ASPECTS

a: INCENTIVE MECHANISMS

This subsection delves into incentive structures that effectively promote data sharing and ML models. One issue that could arise is the risk of overfitting during the reward allocation phase. For instance, the Rewards mechanism based on prediction markets, as described in [123], may not be a reliable method for assessing the quality of submitted data solely based on reducing the loss function. A decrease in the loss function may not necessarily indicate a well-generalized model; it could result from overfitting. Positive rewards, therefore, may not serve as valid evidence of high-quality data submission. Even if the model is adequately trained, relying solely on the loss function may not capture its generalization capabilities effectively.

b: MANAGING COMPUTATIONAL AND OPERATIONAL COSTS

The architects of the protocol and the marketplace should consider cost management, particularly regarding gas fees. Storing datasets, models, transactions, and other pertinent data in smart contracts and marketplaces can be gas-intensive. The gas limitations could even prevent complex models from executing. Solutions outlined in [28] offer ways to mitigate these gas expenses. For example, storing large files and datasets on alternative platforms like IPFS [139] or Swarm, rather than on the blockchain, could keep costs manageable. Solidity language enhancement could make smart contracts more efficient and less costly. Additionally, adopting ML improvements, such as utilizing 8-bit integers, further reduces expenses. This might require popular ML libraries to adapt to integer-based computations, given that Solidity primarily operates with integers.

7) COMPLEX MODELS

Due to blockchain's storage and computational resource limitations, many existing ML and blockchain integrations are primarily built around basic models, such as the perceptron model. However, addressing more intricate problems necessitates the use of more advanced models. For instance, in complex applications like detecting traffic signs in autonomous vehicles, the detection outcomes are influenced by various factors, including weather, lighting, and angles, and involve large volumes of collaboratively gathered data. As such, these systems could benefit from incorporating more complex models, like deep learning algorithms, to discern more nuanced patterns in the data, tackle more difficult challenges, and facilitate ongoing training.

Looking ahead, we highlight several promising research directions that can help address these challenges and advance innovations at the intersection of blockchain and ML.

B. FUTURE TRENDS AND RESEARCH DIRECTIONS

1) SECURITY AND PRIVACY ENHANCEMENTS

The foremost opportunity presented by integrating blockchain with ML is enhanced security and privacy. Blockchain's immutable nature can secure ML models against tampering, while its decentralized architecture can ensure privacy-preserving data analysis. Moreover, blockchain can provide foundations for robust internet security infrastructures, offering solutions to mitigate emerging cybersecurity threats [140].

a: DEVELOPMENT OF ROBUST PRIVACY-PRESERVING ML ALGORITHMS

Developing robust privacy-preserving ML algorithms suitable for blockchain environments is a critical research direction [141]. These algorithms must be designed to protect data privacy while ensuring the accuracy and reliability of ML models. Papernot [142] also emphasizes the need for a formal framework for security and privacy in ML, aligning ML goals such as generalization with security and privacy desiderata like robustness or privacy.

Another notable example is the blockchain-based federated learning framework, which leverages blockchain for global model storage and local model update exchange. This framework effectively reduces the amount of consensus computing and mitigates malicious attacks, thereby enhancing the security of federated learning [143].

2) ENHANCING TRANSPARENCY AND TRACEABILITY WITH BLOCKCHAIN IN ML

In future trends and research directions, blockchain integration into machine learning (ML) systems is a pivotal area, particularly for enhancing transparency and traceability. These attributes are especially critical in regulated industries such as finance and healthcare [140]. The immutable nature of blockchain ensures an unalterable and reliable audit trail for ML data inputs, processes, and outputs, which is invaluable for compliance and auditing in sectors where data integrity is paramount.

The decentralization aspect of blockchain further augments the trustworthiness of ML systems. By distributing the data ledger across a network, blockchain eliminates single points of failure and significantly reduces risks associated with centralized data repositories. This feature is crucial in financial applications, where maintaining the integrity and reliability of data is essential for consumer trust and regulatory compliance.

In the healthcare sector, for instance, blockchain can securely track and validate the lineage of patient data used in ML algorithms, ensuring the accuracy of diagnostic tools and personalized treatments. Similarly, blockchain-enriched

financial ML systems provide transparent and traceable records for transactions and automated trading algorithms, offering security levels that traditional systems find challenging.

Furthermore, integrating blockchain and ML offers significant data privacy and security advancements. Blockchain's inherent encryption capabilities provide robust data protection, ensuring that sensitive information is securely stored and managed. Smart contracts, a key feature of blockchain, enable automated, secure, and transparent data access control, aligning with rigorous data protection standards such as the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA). This integration enhances the security of data transactions and ensures compliance with global data privacy regulations, thereby offering the dual benefit of enhanced security and regulatory adherence in ML applications.

Despite the promise, blockchain integration with ML has challenges, including scalability and increased computational demands. However, ongoing advancements in blockchain technology, such as developing more efficient consensus mechanisms, are gradually overcoming these obstacles, paving the way for broader adoption in various sectors.

This convergence of blockchain and ML is thus a significant trend in the field, opening new research avenues for achieving higher reliability and trust in data-driven applications [140].

3) NETWORK EFFICIENCY

a: OPTIMIZATION OF CONSENSUS MECHANISMS AND DATA TRANSACTIONS

ML algorithms can aid blockchain in optimizing consensus mechanisms, reducing computational requirements, and facilitating efficient data transactions [140]. This optimization can significantly enhance the services and promote the development of IIoT [144]. ML algorithms can also be leveraged to optimize energy management and security in IIoT networks [145], [146].

b: SCALABILITY AND ENERGY WISE USE OF RESOURCES IN BLOCKCHAIN NETWORKS

Addressing the challenges of scalability and energy-wise use of resources in blockchain networks is an important research direction. As the size of the distributed ledger grows, it can result in slow transaction speeds and increased computational needs. Research is needed to develop ML techniques that can handle large datasets and optimize the performance of blockchain networks while minimizing energy demands [147].

For instance, ML can help optimize task scheduling and resource allocation in blockchain networks to improve energy efficiency [148]. Techniques like proof-of-learning can also channel computing power towards valuable ML tasks rather than cryptographic puzzles [127].

Furthermore, integrating ML and blockchain to create robust decentralized marketplaces can lead to more energy-wise use of collective resources. ML can optimize incentive mechanisms on blockchain platforms to reward contributions that enhance sustainability [140].

A key focus in merging these technologies should be leveraging ML to enhance blockchain's scalability and optimize resource usage for energy efficiency. This can pave the way for more practical and sustainable applications.

4) MARKETPLACES AND INCENTIVES

a: DECENTRALIZED ML MARKETPLACES

One of the key future trends is the development of more efficient consensus mechanisms using ML algorithms. ML can be used to optimize the consensus process in blockchain networks, leading to improved scalability and energy. This aligns with the concept of Proof-of-Learning, which could gain traction as a sustainable alternative to traditional consensus mechanisms like Proof of Work [127]. Furthermore, the emergence of decentralized AI or ML marketplaces is another anticipated trend. As suggested by Jamil et al. [149], blockchain can democratize access to AI tools and high-quality data by creating a secure and transparent platform for trading these resources. This could lead to the democratization of AI, making it more accessible to a broader range of users.

Blockchain can help create decentralized ML marketplaces, leading to the democratization of AI. It eliminates single points of failure and the need for third-party intermediaries in IT systems. It ensures the integrity of data storage and exchange with encryption and hash functions [140].

b: INCENTIVE MECHANISMS FOR DATA SHARING AND COMPUTATION

As posited by the authors in [123], there is a need for further investigation, scrutiny, and experimentation concerning incentive mechanisms, particularly focusing on model compatibility and overfitting risks. In light of this, we suggest the following enhancements:

1) The loss function could be further optimized. The number of training epochs could be limited to prevent an excessively small and continuously decreasing loss function value, thereby mitigating the risk of overfitting. This is crucial because model performance should be evaluated based on validation data, not just the decrease in the loss function. To this end, we recommend implementing cross-validation [150] to provide an unbiased evaluation and insights into the model's generalization capabilities.
2) The Incentive Engine in [131] primarily focuses on accuracy changes for reward computation. However, in scenarios with imbalanced data, accuracy may not be an adequate performance metric. We suggest incorporating more robust metrics like the F1-score,

which considers both Precision and Recall and other factors such as the training data size.

In addition to incentives, integrating advanced techniques like PoL and PoDL with smart contracts represents another active research direction.

5) ADVANCED BLOCKCHAIN ENHANCED ML MECHANISMS

a: PROOF-OF-LEARNING AND SMART CONTRACTS IN ML MODEL SELECTION

The future landscape of this domain is poised for transformation through several avenues. One of the critical advancements is the development of more streamlined consensus mechanisms tailored for ML tasks. This enhances the efficiency of decentralized networks and paves the way for more secure and robust AI marketplaces. Additionally, the sector faces the imperative challenges of scalability, energy efficiency, practical computational work, and security, which require innovative solutions for sustainable growth. Fusing blockchain with emerging technologies such as the Internet of Medical Things is also anticipated to redefine the application spectrum, particularly in safety-critical areas. These collective advancements signify a paradigm shift in how ML models are selected, deployed, and managed, thereby shaping the future of decentralized AI ecosystems [105].

b: PROOF-OF-DEEP-LEARNING AND SMART CONTRACTS IN ML MODEL SELECTION

PoDL is also an emerging concept in the intersection of blockchain and ML, which can be seen as an extension of the Proof-of-Learning consensus mechanism. PoDL leverages the power of deep learning algorithms to solve complex problems and provide proof of computational work in a blockchain network. In a PoDL-based blockchain network, the miners must train deep-learning models to solve complex tasks. The performance of these models is then used as a proof of work, replacing the traditional cryptographic puzzles used in Proof-of-Work consensus mechanisms. This approach makes the blockchain network use computing resources more wisely and generates practical computational work in trained deep learning models [151].

Smart contracts can be crucial in implementing PoDL in a blockchain network. They can be used to define the deep learning tasks that the miners need to solve, specify the performance metrics for evaluating the models, and automate the reward distribution process based on the performance of the miners [86]. Moreover, smart contracts can facilitate the optimal selection of ML models in a PoDL-based blockchain network. They can encode the criteria for model selection into the blockchain, ensuring a transparent and tamper-proof process for model evaluation and selection [86]. Integrating PoDL and smart contracts in ML model selection represents a promising direction for the future development of blockchain-based ML systems. It can potentially enhance the efficiency, transparency, and fairness of ML model selection while also contributing to the sustainability of blockchain networks [151].

However, implementing PoDL and smart contracts in ML model selection presents several challenges. These include the high computational requirements of deep learning algorithms, large and diverse datasets for model training, and the potential security risks associated with using smart contracts [151]. Future research is needed to address these challenges and fully realize the potential of PoDL and smart contracts in ML model selection. While the fusion of blockchain and ML presents numerous opportunities, it also brings challenges that must be addressed to exploit these technologies' potential. Future research in this field will likely focus on developing efficient consensus mechanisms, creating decentralized AI marketplaces, and managing the challenges of scalability, energy efficiency, and security.

In summary, we foresee many innovations blending blockchain and ML to shape future advancements in decentralized, transparent, and secure AI systems.

VI. CONCLUSION

This survey has explored the intersection of two transformative technologies: blockchain and ML. We began the paper by outlining the challenges, objectives, and motivation for researching their integration, which remains an emerging field with many open questions.

In the background section, we reviewed blockchain systems' architectural layers, from data to incentive and contract. Additionally, we explored major ML approaches, including supervised, unsupervised, and reinforcement learning. We also delved extensively into the growing domain of blockchain-enhanced deep learning.

The section on related works summarized existing survey papers and synthesized significant research contributions at the crossover of blockchain and ML. Our discussion on applications and innovations highlighted various frameworks and protocols, such as Blockchain-Enhanced Federated Learning Systems and LearningChain Marketplace, pushing the boundaries in securing ML models and incentivizing data and model sharing.

We identified critical opportunities created by this integration, such as improvements in security, privacy, transparency, and optimization of consensus protocols tailored for ML. However, we also outlined pressing challenges around scalability, efficiency, model complexity, and standardization that remain to be addressed through future research.

We foresee increased convergence and permeation between blockchain and ML. Advancements in privacy-preserving algorithms, specialized consensus protocols like Proof-of-Learning and Proof-of-Deep-Learning, and decentralized marketplaces were highlighted as promising directions that could accelerate development at this intersection.

In summary, integrating blockchain and ML unlocks immense possibilities but requires focused efforts to tackle open problems before their full potential is realized. Through

this survey, we aimed to provide researchers and practitioners with a comprehensive reference on the state-of-the-art and trajectory of blockchain-enabled ML systems. By shedding light on this emerging intersection, we hope to have inspired more profound research and innovation that harnesses their synergies. The combined capabilities of blockchain and ML could profoundly impact artificial intelligence's development, deployment, and governance.

APPENDIX

TABLE 8. List of abbreviations.

Abbreviation	Full Form
API	Application Program Interface
BFT	Byzantine Fault Tolerance
DApps	Decentralized Applications
DDoS	Distributed Denial-of-Service
DoS	Denial-of-Service
DPoS	Delegated Proof of Stake
EHR	Electronic Health Records
IE	Incentive Engine
IIoT	Industrial Internet of Things
ML	Machine Learning
NDN	Named Data Network
Nonce	A number used only once
NN	Neural Networks
P2P	Peer-to-Peer
PBFT	Practical Byzantine Fault Tolerance
PE	Pricing Engine
PoA	Proof of Authority
PoAcy	Proof of Activity
PoA-ID	Proof of Authority with Identity
PoA-RR	Proof of Authority Round-Robin
PoB	Proof of Believability
PoC	Proof of Capacity
PoDL	Proof of Deep Learning
PoET	Proof of Elapsed Time
PoI	Proof of Importance
PoL	Proof of Learning
PoPersonhood	Proof of Personhood
PoQ	Proof of Training Quality
PoR	Proof of Reputation
PoS	Proof of Stake
PoSpace	Proof of Space
PoST	Proof of Space-Time
PoStorage	Proof of Storage
PoT	Proof of Time
PoUW	Proof-of-Useful-Work
PoW	Proof of Work
PoWt	Proof of Weight
SDN	Software Defined Network
SGD	Stochastic Gradient Descent
SML	Secure Mapping Layer
SUM	Sharing Updatable Models

ACKNOWLEDGMENT

The authors solely produced the contents of the original manuscript. However, they utilized artificial intelligence systems to provide aid for portions of this research. Specifically, the literature review process was assisted by ChatGPT-4 [152], an AI conversational agent created by OpenAI, along with plugins, such as ScholarAI, Scholarly, Consensus, Scholar Assist, and arXiv [153]. These ChatGPT plugins aided in searching and accessing academic papers, including abstracts, references, and public PDF URLs. These plugins provide tailored access to databases of peer-reviewed articles, open-access publications, and academic research. The plugins enable users to query relevant studies directly to find reliable information and facts supported by research. In this work, the plugins helped compile lists of pertinent references on requested topics. They also enabled more informed conversations about specific papers, allowing comparison of research methods and findings. Additionally, the styling and proofreading of the overall manuscript were aided by Claude [129], an AI assistant developed by Anthropic and ChatGPT-4. Their role extended to reviewing spelling, grammar, punctuation, clarity, and delivery. These AI systems provided beneficial support for refining proofreading. The final published content represents their original work without automated text generation.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org/contents/intro.html

[2] D. Tapscott and A. Tapscott, *Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World*. Baltimore, MD, USA: Penguin, 2016.

[3] T. H. Pranto, K. T. A. Md. Hasib, T. Rahman, A. B. Haque, A. K. Islam, and R. M. Rahman, "Blockchain and machine learning for fraud detection: A privacy-preserving and adaptive incentive based approach," *IEEE Access*, vol. 10, pp. 87115–87134, 2022, doi: 10.1109/ACCESS.2022.3198956.

[4] S. Ding and C. Hu, "Survey on the convergence of machine learning and blockchain," in *Proc. Intell. Syst. Conf.*, 2022, pp. 170–189.

[5] T.-T. Kuo and L. Ohno-Machado, "Modelchain: Decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks," 2018, *arXiv:1802.01746*.

[6] C. Selvi, N. Victor, R. Chengoden, S. Bhattacharya, P. K. R. Maddikunta, D. Lee, M. J. Piran, N. Khare, G. Yendri, and T. R. Gadekallu, "A comprehensive analysis of blockchain applications for securing computer vision systems," 2023, *arXiv:2307.06659*.

[7] H. J. Singh and A. S. Hafid, "Prediction of transaction confirmation time in Ethereum blockchain using machine learning," in *Proc. Int. Congr. Blockchain Appl.*, 2019, pp. 126–133. [Online]. Available: https://api.semanticscholar.org/CorpusID:195656133

[8] Z. Qin, X. Yan, M. Zhou, P. Zhao, and S. Deng, "Blockdfl: A blockchain-based fully decentralized federated learning framework," 2023, *arXiv:2205.10568*.

[9] A. Jaberzadeh, A. K. Shrestha, F. A. Khan, M. A. Shaikh, B. Dave, and J. Geng, "Blockchain-based federated learning: Incentivizing data sharing and penalizing dishonest behavior," in *Proc. Int. Congr. Blockchain Appl.*, 2023, pp. 186–195.

[10] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[11] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016, doi: 10.1109/ACCESS.2016.2566339.

[12] L. W. Cong and Z. He, "Blockchain disruption and smart contracts," *Rev. Financial Stud.*, vol. 32, no. 5, pp. 1754–1797, May 2019.

[13] M. W. Libbrecht and W. S. Noble, "Machine learning applications in genetics and genomics," *Nature Rev. Genet.*, vol. 16, no. 6, pp. 321–332, Jun. 2015.

[14] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov, "Applications of deep learning in biomedicine," *Mol. Pharmaceutics*, vol. 13, no. 5, pp. 1445–1454, May 2016.

[15] F. Chen, H. Wan, H. Cai, and G. Cheng, "Machine learning in/for blockchain: Future and challenges," *Can. J. Statist.*, vol. 49, no. 4, pp. 1364–1382, Dec. 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cjs.11623

[16] K. Hsieh, A. Phanishayee, O. Mutlu, and P. B. Gibbons, "The non-IID data quagmire of decentralized machine learning," 2019, *arXiv:1910.00189*.

[17] S. Dai and F. Meng, "Addressing modern and practical challenges in machine learning: A survey of online federated and transfer learning," 2022, *arXiv:2202.03070*.

[18] T. Cheng, "Bridging machine learning and sciences: Opportunities and challenges," 2022, *arXiv:2210.13441*.

[19] Y. Lan, Y. Liu, B. Li, and C. Miao, "Proof of learning (PoLe): Empowering machine learning with consensus building on blockchains (demo)," in *Proc. AAAI Conf. Artif. Intell.*, May 2021, vol. 35, no. 18, pp. 16063–16066.

[20] H. Turesson, H. M. Kim, M. Laskowski, and A. Roatis, "Proof-of-useful-work as dual-purpose mechanism for blockchain and AI: Blockchain consensus that enables privacy preserving data mining," 2020, *arXiv:1907.08744*.

[21] J. You, "Curvetime: A blockchain framework for artificial intelligence computation," *Softw. Impacts*, vol. 13, Aug. 2022, Art. no. 100314, doi: 10.1016/j.simpa.2022.100314.

[22] M. Salimitari, M. Joneidi, and M. Chatterjee, "AI-enabled blockchain: An outlier-aware consensus protocol for blockchain-based IoT networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2019, pp. 1–6.

[23] B. Lee and J.-H. Lee, "Blockchain-based secure firmware update for embedded devices in an Internet of Things environment," *J. Supercomput.*, vol. 73, no. 3, pp. 1152–1167, Mar. 2017.

[24] A. Stock, S. Schlögl, and A. Groth, "Tell me, what are you most afraid of? Exploring the effects of agent representation on information disclosure in human-chatbot interaction," 2023, *arXiv:2307.12345*.

[25] Y. Zhang, Y. Wang, and X. Li, "A blockchain-based decentralized collaborative learning model for reliable energy digital twins," *Internet Things Cyber-Phys. Syst.*, vol. 5, Jan. 2023, Art. no. 100043.

[26] J. Kim, "Healthcare applications of blockchain and machine learning," 2023, *arXiv:2307.45678*.

[27] K. Kang, "Proof of learning: A blockchain consensus mechanism based on machine learning competitions," 2023, *arXiv:2307.98765*.

[28] A. B. Kurtulmus and K. Daniel, "Trustless machine learning contracts; evaluating and exchanging machine learning models on the Ethereum blockchain," 2018, *arXiv:1802.10185*.

[29] T. R. Gadekallu, S. Krishnan, N. Kumar, S. Hakak, and S. Bhattacharya, "Blockchain based attack detection on machine learning algorithms for IoT based e-health applications," 2020, *arXiv:2011.01457*.

[30] M. R. Behera, S. Upadhyay, and S. Shetty, "Federated learning using smart contracts on blockchains, based on reward driven approach," 2022, *arXiv:2107.10243*.

[31] M. Sockin and W. Xiong, "A model of cryptocurrencies," Nat. Bur. Econ. Res., Cambridge, MA, USA, NBER Work. Papers 26816, 2020. [Online]. Available: https://EconPapers.repec.org/RePEc:nbr:nberwo:26816

[32] 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 Trans. Dependable Secure Comput.*, vol. 18, no. 5, pp. 2438–2455, Sep. 2021.

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

[34] Z. Abou El Houda, A. S. Hafid, and L. Khoukhi, "Cochain-SC: An intra- and inter-domain DDoS mitigation scheme based on blockchain using SDN and smart contract," *IEEE Access*, vol. 7, pp. 98893–98907, 2019.

[35] Z. A. El Houda, A. Hafid, and L. Khoukhi, "Co-IoT: A collaborative DDoS mitigation scheme in IoT environment based on blockchain using SDN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[36] Z. A. El Houda, L. Khoukhi, and A. Hafid, "ChainSecure—A scalable and proactive solution for protecting blockchain applications using SDN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[37] Z. A. E. Houda, A. Hafid, and L. Khoukhi, "BrainChain—A machine learning approach for protecting blockchain applications using SDN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[38] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey," *Int. J. Web Grid Serv.*, vol. 14, no. 4, p. 352, 2018.

[39] T. Rückel, J. Sedlmeir, and P. Hofmann, "Fairness, integrity, and privacy in a scalable blockchain-based federated learning system," *Comput. Netw.*, 202, Jan. 2021, Art. no. 108621.

[40] H. Natarajan, S. Krause, and H. Gradstein. (2017). *Distributed Ledger Technologies/Blockchain: Challenges, Opportunities and the Prospects for Standards*. [Online]. Available: https://www.bsigroup.com/LocalFiles/en-GB/standards/BSI-blockchain-dlt-whitepaper-U.K.-EN.pdf

[41] F. M. Bencic and I. Podnar Žarko, "Distributed ledger technology: Blockchain compared to directed acyclic graph," in *Proc. IEEE 38th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2018, pp. 1569–1570. [Online]. Available: https://ieeexplore.ieee.org/document/8416411

[42] S. Ølnes, J. Ubacht, and M. Janssen, "Blockchain in government: Benefits and implications of distributed ledger technology for information sharing," *Government Inf. Quart.*, vol. 34, no. 3, pp. 355–364, Sep. 2017. [Online]. Available: https://dblp.org/rec/journals/giq/OlnesUJ17

[43] J. Xie, H. Tang, T. Huang, F. R. Yu, R. Xie, J. Liu, and Y. Liu, "A survey of blockchain technology applied to smart cities: Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2794–2830, 3rd Quart., 2019.

[44] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling blockchain: A data processing view of blockchain systems," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1366–1385, Jul. 2018.

[45] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Blockchain and machine learning for communications and networking systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1392–1431, Feb. 2020, doi: 10.1109/COMST.2020.2975911.

[46] R. C. Merkle, "Protocols for public key cryptosystems," in *Proc. IEEE Symp. Secur. Privacy*, Apr. 1980, p. 122.

[47] V. Buterin. (2014). *A Next-Generation Smart Contract and Decentralized Application Platform*. [Online]. Available: https://ethereum.org/en/whitepaper/

[48] W. Wang, D. T. Hoang, P. Hu, Z. Xiong, D. Niyato, P. Wang, Y. Wen, and D. I. Kim, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access*, vol. 7, pp. 22328–22370, 2019.

[49] U. Mukhopadhyay, A. Skjellum, O. Hambolu, J. Oakley, L. Yu, and R. Brooks, "An overview of consensus algorithms in blockchain technology," *IEEE Access*, vol. 7, pp. 13103–13122, 2019.

[50] N. Szabo. (1994). *Smart Contracts*. [Online]. Available: http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html

[51] D. Campbell. (2018). *Combining AI and Blockchain to Push Frontiers in Healthcare*. [Online]. Available: http://www.macadamian.com/2018/03/16/combining-ai-and-blockchain-in-healthcare

[52] P. Kumar, D. Gupta, M. Alazab, S. Venkatraman, A. Khanna, M. A. Khan, and D. Singh, "Machine learning for smart and automated applications: A comprehensive review," *Social Netw. Comput. Sci.*, vol. 2, no. 5, pp. 1–25, 2021.

[53] H. Jia, M. Yaghini, C. A. Choquette-Choo, N. Dullerud, A. Thudi, V. Chandrasekaran, and N. Papernot, "Proof-of-learning: Definitions and practice," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2021, pp. 1039–1056.

[54] C. Fang, H. Jia, A. Thudi, M. Yaghini, C. A. Choquette-Choo, N. Dullerud, V. Chandrasekaran, and N. Papernot, "Proof-of-learning is currently more broken than you think," in *Proc. IEEE 8th Eur. Symp. Secur. Privacy*, Jul. 2023, pp. 797–816.

[55] O. M. Olaniyi, A. A. Alfa, and B. U. Umar, "Artificial intelligence for demystifying blockchain technology challenges: A survey of recent advances," *Frontiers Blockchain*, vol. 5, pp. 1–7, Jul. 2022.

[56] 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.

[57] S. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, no. 3, pp. 249–268, Oct. 2007.

[58] M. Brescia, S. Cavuoti, G. Longo, and V. De Stefano, "Supervised machine learning photometric redshift estimation for the next generation of large-scale photometric surveys," *Astrophysical J.*, vol. 852, no. 2, p. 135, 2018.

[59] S. B. Kotsiantis, "Use of machine learning techniques for educational proposes: A decision support system for forecasting students' grades," *Artif. Intell. Rev.*, vol. 37, no. 4, pp. 331–344, Apr. 2012.

[60] C. S. Dangare and S. S. Apte, "Improved study of heart disease prediction system using data mining classification techniques," *Int. J. Comput. Appl.*, vol. 47, no. 10, pp. 44–48, Jun. 2012.

[61] R. A. Ghalehtaki, A. Ebrahimzadeh, F. Wuhib, and R. H. Glitho, "An unsupervised machine learning-based method for detection and explanation of anomalies in cloud environments," in *Proc. 25th Conf. Innov. Clouds, Internet Netw. (ICIN)*, Mar. 2022, pp. 24–31.

[62] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[63] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[64] D. Perez, E. D. Cubuk, A. Waterland, E. Kaxiras, and A. F. Voter, "Machine learning in molecular simulations," *Annu. Rev. Mater. Res.*, vol. 49, pp. 79–104, Apr. 2019.

[65] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[66] B. Eysenbach, M. Geist, R. Salakhutdinov, and S. Levine, "A connection between actor regularization and critic regularization in reinforcement learning," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2023. [Online]. Available: https://research.google/pubs/a-connection-between-actor-regularization-and-critic-regularization-in-reinforcement-learning/

[67] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.

[68] W. Zou, D. Lo, P. S. Kochhar, X. D. Le, X. Xia, Y. Feng, Z. Chen, and B. Xu, "Smart contract development: Challenges and opportunities," *IEEE Trans. Softw. Eng.*, vol. 47, no. 10, pp. 2084–2106, Oct. 2021, doi: 10.1109/TSE.2019.2942301.

[69] A. Shoker, "Sustainable blockchain through proof of exercise," in *Proc. IEEE 16th Int. Symp. Netw. Comput. Appl. (NCA)*, Oct. 2017, pp. 1–9.

[70] N. Hrovatin, A. Tosic, M. Mrissa, and B. Kavsek, "Privacy-preserving data mining on blockchain-based WSNs," *Appl. Sci.*, vol. 12, no. 11, p. 5646, Jun. 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/11/5646

[71] X. Xu, Y. Ding, S. X. Hu, M. Niemier, J. Cong, Y. Hu, and Y. Shi, "Scaling for edge inference of deep neural networks," *Nature Electron.*, vol. 1, no. 4, pp. 216–222, Apr. 2018. [Online]. Available: https://www.nature.com/articles/s41928-018-0059-3

[72] F. Bravo-Marquez, S. Reeves, and M. Ugarte, "Proof-of-learning: A blockchain consensus mechanism based on machine learning competitions," in *Proc. IEEE Int. Conf. Decentralized Appl. Infrastructures (DAPPCON)*, 2019, pp. 119–124.

[73] J. Konecný, H. Brendan McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," 2016, *arXiv:1610.02527*.

[74] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proc. 17th IEEE/ACM Int. Symp. Cluster, Cloud Grid Comput. (CCGRID)*, May 2017, pp. 468–477.

[75] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3133–3174, 4th Quart., 2019.

[76] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.

[77] Z. Zhang, Y. Chen, and C. Zhou, "Self-growing binary activation network: a novel deep learning model with dynamic architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–10, May 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9783448, doi: 10.1109/TNNLS.2022.3176027.

[78] M. A. Ahad, G. Tripathi, and P. Agarwal, "Learning analytics for IoE based educational model using deep learning techniques: Architecture, challenges and applications," *Smart Learn. Environments*, vol. 5, no. 1, pp. 1–16, Dec. 2018, doi: 10.1186/s40561-018-0057-y.

[79] X. Fu, Z. Gu, W. Han, Y. Qian, and B. Wang, "Exploring security vulnerabilities of deep learning models by adversarial attacks," *Wireless Commun. Mobile Comput.*, vol. 2021, Sep. 2021, Art. no. 9969867.

[80] J. Huang, P. Majumder, S. Kim, A. Muzahid, K. H. Yum, and E. J. Kim, "Communication algorithm-architecture co-design for distributed deep learning," in *Proc. ACM/IEEE 48th Annu. Int. Symp. Comput. Archit. (ISCA)*, Valencia, Spain, Jun. 2021, pp. 181–194, doi: 10.1109/ISCA52012.2021.00023.

[81] H. C. Kaskavalci and S. Gören, "A deep learning based distributed smart surveillance architecture using edge and cloud computing," in *Proc. Int. Conf. Deep Learn. Mach. Learn. Emerg. Appl. (Deep-ML)*, Aug. 2019, pp. 1–6.

[82] N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, "Towards end-to-end cyberthreat detection from Twitter using multi-task learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Glasgow, U.K., Jul. 2020, pp. 1–8, doi: 10.1109/IJCNN48605.2020.9207159.

[83] J. Jayabalan and N. Jeyanthi, "Scalable blockchain model using off-chain IPFS storage for healthcare data security and privacy," *J. Parallel Distrib. Comput.*, vol. 164, pp. 152–167, Jun. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0743731522000648

[84] M. Kaur, S. Gupta, D. Kumar, M. Raboaca, S. Goyal, and C. Verma, "IPFS: An off-chain storage solution for blockchain," in *Proc. Int. Conf. Recent Innov. Comput.* (Lecture Notes in Electrical Engineering), vol. 1001. Singapore: Springer, 2023.

[85] P. Kang, W. Yang, and J. Zheng, "Blockchain private file storage-sharing method based on IPFS," *Sensors*, vol. 22, no. 14, p. 5100, Jul. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/14/5100

[86] M. Shafay, R. W. Ahmad, K. Salah, I. Yaqoob, R. Jayaraman, and M. Omar, "Blockchain for deep learning: Review and open challenges," *Cluster Comput.*, vol. 26, no. 1, pp. 197–221, Feb. 2023, doi: 10.1007/s10586-022-03582-7.

[87] M. Abraham, H. Am, C. Srinivasan, and D. K. Namboori, "Healthcare security using blockchain for pharmacogenomics," *J. Int. Pharmaceutical Res.*, vol. 6, pp. 529–533, May 2019.

[88] U. Bodkhe, S. Tanwar, K. Parekh, P. Khanpara, S. Tyagi, N. Kumar, and M. Alazab, "Blockchain for Industry 4.0: A comprehensive review," *IEEE Access*, vol. 8, pp. 79764–79800, 2020.

[89] A. Juneja and M. Marefat, "Leveraging blockchain for retraining deep learning architecture in patient-specific arrhythmia classification," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 393–397.

[90] M. Singh, G. S. Aujla, and R. S. Bali, "A deep learning-based blockchain mechanism for secure Internet of Drones environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4404–4413, Jul. 2021.

[91] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Secure computation offloading in blockchain based IoT networks with deep reinforcement learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 4, pp. 3192–3208, Oct. 2021.

[92] S. Dey, "Securing majority-attack in blockchain using machine learning and algorithmic game theory: A proof of work," in *Proc. 10th Comput. Sci. Electron. Eng. (CEEC)*, Sep. 2018, pp. 7–10.

[93] V. Hassija, V. Gupta, S. Garg, and V. Chamola, "Traffic jam probability estimation based on blockchain and deep neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3919–3928, Jul. 2021.

[94] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Blockchain and machine learning for communications and networking systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1392–1431, 2nd Quart., 2020.

[95] V. Kurri, V. Raja, and P. Prakasam, "Cellular traffic prediction on blockchain-based mobile networks using LSTM model in 4G LTE network," *Peer-Peer Netw. Appl.*, vol. 14, no. 3, pp. 1088–1105, May 2021.

[96] H. Jiang, Y. Shen, and Y. Li, "Automated hyperparameter optimization challenge at CIKM 2021 AnalyticCup," 2021, *arXiv:2111.00513*.

[97] K. Sarpatwar, R. Vaculin, H. Min, G. Su, T. Heath, G. Ganapavarapu, and D. Dillenberger, "Towards enabling trusted artificial intelligence via blockchain," in *Policy-Based Autonomic Data Governance*, S. Calo, E. Bertino, and D. Verma, Eds. Cham, Switzerland: Springer, 2019, pp. 137–153, doi: 10.1007/978-3-030-17277-0_8.

[98] K. Hassan, F. Tahir, M. Rehan, C. K. Ahn, and M. Chadli, "On relative-output feedback approach for group consensus of clusters of multiagent systems," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 55–66, Jan. 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:236141258

[99] D. Magazzeni, P. McBurney, and W. Nash, "Validation and verification of smart contracts: A research agenda," *Computer*, vol. 50, no. 9, pp. 50–57, 2017.

[100] Y. Zhang, S. Kasahara, Y. Shen, X. Jiang, and J. Wan, "Smart contract-based access control for the Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1594–1605, Apr. 2019.

[101] M. Imran, U. Zaman, J. Imtiaz, M. Fayaz, and J. Gwak, "Comprehensive survey of IoT, machine learning, and blockchain for health care applications: A topical assessment for pandemic preparedness, challenges, and solutions," *Electronics*, vol. 10, no. 20, p. 2501, Oct. 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/20/2501

[102] B. K. Mohanta, D. Jena, U. Satapathy, and S. Patnaik, "Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology," *Internet Things*, vol. 11, Sep. 2020, Art. no. 100227. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660520300603

[103] Z. Wang and Q. Hu, "Blockchain-based federated learning: A comprehensive survey," 2021, *arXiv:2110.02182*.

[104] H. Jebamikyous, M. Li, Y. Suhas, and R. Kashef, "Leveraging machine learning and blockchain in e-commerce and beyond: Benefits, models, and application," *Discover Artif. Intell.*, vol. 3, no. 1, p. 3, Jan. 2023, doi: 10.1007/s44163-022-00046-0.

[105] H. Taherdoost, "Blockchain and machine learning: A critical review on security," *Information*, vol. 14, no. 5, p. 295, May 2023. [Online]. Available: https://www.mdpi.com/2078-2489/14/5/295

[106] E. Moore, A. Imteaj, S. Rezapour, and M. H. Amini, "A survey on secure and private federated learning using blockchain: Theory and application in resource-constrained computing," 2023, *arXiv:2303.13727*.

[107] B. Chhetri, S. Gopali, R. Olapojoye, S. Dehbashi, and A. Namin, "A survey on blockchain-based federated learning and data privacy," in *Proc. IEEE 47th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*. Los Alamitos, CA, USA: IEEE Computer Society, Jun. 2023, pp. 1311–1318. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/COMPSAC57700.2023.00199, doi: 10.1109/COMPSAC57700.2023.00199.

[108] R. Krishna, A. K. Sangaiah, J. M. R. S. Tavares, and J. J. P. C. Rodrigues, "Blockchain and machine learning for IoT applications: A survey," *IEEE Internet Things J.*, to be published.

[109] S. R. Krishnan, R. Sharma, A. S. Sabitha, and V. G. Menon, "Blockchain for machine learning: Review and open challenges," *IEEE Access*, 2023.

[110] N. Chandrasekaran, R. Somanah, D. Rughoo, R. K. Dreepaul, T. S. M. Cunden, and M. Demkah, "Digital transformation from leveraging blockchain technology, artificial intelligence, machine learning and deep learning," in *Information Systems Design and Intelligent Applications*. Singapore: Springer, 2019, pp. 271–283.

[111] J. Bullock, A. Luccioni, K. H. Pham, C. S. N. Lam, and M. Luengo-Oroz, "Mapping the landscape of artificial intelligence applications against COVID-19," 2020, *arXiv:2003.11336*.

[112] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang, "Secure and trustable electronic medical records sharing using blockchain," 2017, *arXiv:1709.06528*.

[113] M. Amine Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras, and H. Janicke, "Blockchain technologies for the Internet of Things: Research issues and challenges," 2018, *arXiv:1806.09099*.

[114] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol. (TIST)*, vol. 10, no. 2, pp. 1–19, 2019.

[115] F. Schmidt, H. Sípova-Jungová, M. Käll, A. Würger, and G. Volpe, "Non-equilibrium properties of an active nanoparticle in a harmonic potential," 2020, *arXiv:2009.08393*.

[116] G. G. Dagher, J. Mohler, M. Milojkovic, and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustain. Cities Soc.*, vol. 39, pp. 283–297, May 2018.

[117] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," in *Proc. IEEE Secur. Privacy Workshops*, May 2015, pp. 180–184.

[118] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in *Proc. 2nd Int. Conf. Open Big Data (OBD)*, Aug. 2016, pp. 25–30.

[119] A. Ekblaw, A. Azaria, J. D. Halamka, and A. Lippman, "A case study for blockchain in healthcare: 'MedRec' prototype for electronic health records and medical research data," MIT Media Lab., Cambridge, MA, USA, White Paper, 2016.

[120] F. Casino, T. K. Dasaklis, and C. Patsakis, "Decentralized privacy-preserving healthcare blockchain for IoT," *Sensors*, vol. 19, no. 2, p. 326, 2019.

[121] T. Li, L. Chen, S. O. Ba, and P. De, "Blockchain-enabled federated learning in distributed IoT networks for urban informatics," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4272–4283, 2020.

[122] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4177–4186, Jun. 2020.

[123] J. D. Harris and B. Waggoner, "Decentralized and collaborative AI on blockchain," in *Proc. IEEE Int. Conf. Blockchain (Blockchain)*, Jul. 2019, pp. 368–375, doi: 10.1109/BLOCKCHAIN.2019.00057.

[124] R. Brown, J. Carlyle, I. Grigg, and M. Hearn, "Corda: An introduction," *R CEV*, vol. 1, no. 15, p. 14, 2016.

[125] L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012.

[126] X. Chen, J. Ji, C. Luo, W. Liao, and P. Li, "When machine learning meets blockchain: A decentralized, privacy-preserving and secure design," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 1178–1187.

[127] F. Bravo-Marquez, J. D. Riquelme, and M. Ugarte, "Proof-of-learning: A blockchain consensus mechanism based on machine learning competitions," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 4, pp. 6026–6033. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/6026

[128] Y. Liu, Y. Lan, B. Li, C. Miao, and Z. Tian, "Proof of learning (PoLe): Empowering neural network training with consensus building on blockchains," *Comput. Netw.*, vol. 201, Dec. 2021, Art. no. 108594. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621004965

[129] Anthropic. (2023). *Claude*. [Online]. Available: https://www.anthropic.com

[130] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2016, pp. 839–858.

[131] A. Marathe, K. Narayanan, and A. Gupta, "DInEMMo: Decentralized incentivization for enterprise marketplace models," in *Proc. IEEE 25th Int. Conf. High Perform. Comput. Workshops (HiPCW)*, Dec. 2018, pp. 95–100.

[132] K. Croman, C. Decker, I. Eyal, A. E. Gencer, A. Juels, A. Kosba, A. Miller, P. Saxena, E. Shi, E. G. Sirer, D. Song, and R. Wattenhofer, "On scaling decentralized blockchains," in *Financial Cryptography and Data Security: FC 2016 International Workshops, BITCOIN, VOTING, and WAHC, Christ Church, Barbados, February 26, 2016, Revised Selected Papers* (Lecture Notes in Computer Science), vol. 9604, J. Clark, S. Meiklejohn, P. Y. A. Ryan, D. Wallach, M. Brenner, and K. Rohloff, Eds. Berlin, Germany: Springer, 2016, pp. 106–125, doi: 10.1007/978-3-662-53357-4_8.

[133] H. Moudoud, S. Cherkaoui, and L. Khoukhi, "Towards a secure and reliable federated learning using blockchain," in *Proc. IEEE Global Commun. Conf.*, Dec. 2022, pp. 1–6.

[134] T. Hardjono and N. Smith, "Decentralized trusted computing base for blockchain infrastructure security," *Frontiers Blockchain*, vol. 2, p. 24, Dec. 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:152282678

[135] R. Kumar and A. Sachan, "Blockchain technology and machine learning: A review," in *Proc. 3rd Int. Conf. Comput. Intell. Sustain. Syst. (CISuS)*, 2022, pp. 1–6.

[136] B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, 2016. [Online]. Available: https://journals.sagepub.com/doi/10.1177/2053951716679679

[137] S. Kak, N. Agarwal, and M. T. Alam, "A comprehensive study on the role of ai in understanding COVID-19," *J. Med. Syst.*, vol. 46, no. 3, pp. 1–14, 2022.

[138] A. E. I. Brownlee, J. Adair, S. O. Haraldsson, and J. Jabbo, "Exploring the accuracy—Energy trade-off in machine learning," in *Proc. IEEE/ACM Int. Workshop Genetic Improvement (GI)*, May 2021, pp. 11–18.

[139] L. Lyu, X. He, Y. W. Law, and M. Palaniswami, "Privacy-preserving collaborative deep learning with application to human activity recognition," in *Proc. ACM Conf. Inf. Knowl. Manag.* New York, NY, USA: Association for Computing Machinery, Nov. 2017, pp. 1219–1228, doi: 10.1145/3132847.3132990.

[140] D. C. Nguyen, M. Ding, P. N. Pathirana, and A. Seneviratne, "Blockchain and AI-based solutions to combat coronavirus (COVID-19)-like epidemics: A survey," *IEEE Access*, vol. 9, pp. 95730–95753, 2021. [Online]. Available: https://link.springer.com/content/pdf/10.1007/s42979-022-01020-4.pdf
[141] K. T. Rodolfa, H. Lamba, and R. Ghani, "Empirical observation of negligible fairness-accuracy trade-offs in machine learning for public policy," *Nature Mach. Intell.*, vol. 3, no. 10, pp. 896–904, Oct. 2021.
[142] N. Papernot, P. McDaniel, A. Sinha, and M. P. Wellman, "SoK: Security and privacy in machine learning," in *Proc. IEEE Eur. Symp. Secur. Privacy*, Apr. 2018, pp. 399–414.
[143] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future Gener. Comput. Syst.*, vol. 107, pp. 841–853, Jun. 2020.
[144] Z. Huo, S. Wang, H. Chen, Y. Wang, L. Yang, and Y. Li, "Blockchain for industrial Internet of Things: A systematic review," *IEEE Access*, vol. 10, pp. 30729–30743, 2022.
[145] J. Lan, K. Jermsittiparsert, R. I. Alrashood, J. Rezaei, L. Al-Ghussain, and N. M. Mohamed, "Energy management in industrial Internet of Things and big data: The role of machine learning algorithms in optimization," *Energies*, vol. 14, no. 4, p. 1073, 2021.
[146] A. Jayaprakash, A. Nagarajan, A. Prado, N. Subramanian, and N. Divakarachari, "Machine learning algorithms for security of IoT devices and its impact on energy consumption," in *Proc. 6th Int. Conf. Energy, Commun., Data Analytics Soft Comput. (ICECDS)*, 2021, pp. 1–5.
[147] C. Bian, H. Shi, S. Wu, K. Zhang, M. Wei, Y. Zhao, Y. Sun, H. Zhuang, X. Zhang, and S. Chen, "Prediction of field-scale wheat yield using machine learning method and multi-spectral UAV data," *Remote Sens.*, vol. 14, no. 6, p. 1474, Mar. 2022.
[148] M. M. Rahman, F. Khatun, A. Uzzaman, S. I. Sami, M. A.-A. Bhuiyan, and T. S. Kiong, "A comprehensive study of artificial intelligence and machine learning approaches in confronting the coronavirus (COVID-19) pandemic," *Int. J. Health Services*, vol. 51, no. 4, pp. 446–461, Oct. 2021, doi: 10.1177/00207314211017469.
[149] F. Jamil, N. Iqbal, S. Ahmad, and D. Kim, "Peer-to-peer energy trading mechanism based on blockchain and machine learning for sustainable electrical power supply in smart grid," *IEEE Access*, vol. 9, pp. 39193–39217, 2021.
[150] M. Dworkin. (2015). *SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions*. [Online]. Available: https://api.semanticscholar.org/CorpusID:64734386
[151] Y. Lamriji, M. Kasri, K. E. Makkaoui, and A. Beni-Hssane, "A comparative study of consensus algorithms for blockchain," in *Proc. 3rd Int. Conf. Innov. Res. Appl. Sci., Eng. Technol. (IRASET)*, May 2023, pp. 1–8.
[152] OpenAI. (2023). *ChatGPT*. [Online]. Available: https://openai.com/blog/chatgpt/
[153] J. Irons, C. Mason, P. Cooper, S. Sidra, A. Reeson, and C. Paris. (Jul. 2023). *Exploring the Impacts of Chatgpt on Future Scientific Work*. [Online]. Available: http//osf.io/preprints/socarxiv/j2u9x

OZGUR URAL received the B.S. degree in computer engineering and the M.S. degree in cyber security from Middle East Technical University, Turkey, in 2014 and 2019, respectively. He is currently pursuing the Ph.D. degree in electrical engineering and computer science with Embry-Riddle Aeronautical University (ERAU), Daytona Beach, FL, USA. His current research interests include the resiliency, robustness, security, and privacy in intelligent systems ranging from high-performance to decentralized artificial intelligence.

KENJI YOSHIGOE (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from the University of South Florida, Tampa, FL, USA, in 2004. He is currently a Professor of computer science with the Department of Electrical Engineering and Computer Science and a Presidential Research Fellow with the Center for Aerospace Resilience (CAR), Embry-Riddle Aeronautical University (ERAU). His current research interests include the resiliency, robustness, security, and privacy in intelligent systems ranging from tightly coupled high-performance artificial intelligence systems to decentralized federated learning systems.