IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Decentralized Artificial Intelligence on Blockchain

Yatrik Buch [1*], Mosin Hasan[2], Prashant Swadas [3]

[1,2,3]Computer Engineering Department, BVM Engineering College, Vallabh Vidyanagar, India

[*]*Corresponding Author: yatrik007@gmail.com, Tel.: +91-9898572748*

DOI: https://doi.org/10.26438/ijcse/v7i2.844847 | Available online at: www.ijcseonline.org

Accepted: 20/Feb/2019, Published: 28/Feb/2019

***Abstract*—** Blockchain's integration with various domains beyond cryptocurrency has produced exciting results and innovative products. Blockchain's capabilities of being distributed peer-to-peer network, scalability, reliability and security has potential to solve many challenges faced by AI and at the same time add exciting features to it. This article analyses and reviews many existing research where Blockchain is integrated with AI. These research works have produced many encouraging results for data privacy, security, distributed processing and trustless collaboration. Blockchain has also contributed to enhance security of the AI system and bring trust among various AI systems. Cryptocurrency has also added trading capabilities to AI and encouraged models like AI as Service. This article also discusses scope for future research work. With more maturity and feature fullness, Blockchain is poised to become one of the most suitable platforms for decentralized AI applications.

***Keywords*—** Blockchain, Artificial Intelligence, Machine Learning, Deep Learning, Robotics, Decentralized Applications

I. INTRODUCTION

With vision to replace traditional banking system, Satoshi Nakamoto in 2008 developed a new platform called "Blockchain" and new currency called "Cryptocurrency". The objective was to provide an open public ledger system and a digital currency which can be accepted worldwide. These objectives resulted in certain inherent properties in Blockchain architecture like cryptographically secured structure, peer-to-peer network and collaborative processing. In 2013, Vitalik Buterin proposed a new Blockchain framework, Ethereum, which was capable of storing and executing user defined code in the Blockchain. This development made it suitable for all kinds of applications beyond just cryptocurrency.

Artificial Intelligence is also on top of the list for technological research and futuristic application development. AI has found its scope of application in almost every aspect of human existence. AI is being used for classification, recommendations, prediction, optimization etc. However AI methodologies are having certain inherent challenges. Researchers are integrating Blockchain and AI to solve these challenges and build innovative systems.

This paper studies and analyses research work carried out for integration of AI with Blockchain. The remaining part of the paper is organized into three sections. Section 2 explains and analyzes existing research work and already released applications. Section 3 describes scope of new work. Section 4 summarizes the paper.

II. EXISTING RESEARCH

This section reviews and analyzes the work that has already been done in the research and industry. The work majorly focuses on decentralized machine learning and deep learning, where training data management, training data processing, exchange of trained models and some other applications are in focus. The review starts with a research paper of Melanie Swan [1] where she represents the very basic ideas of decentralized AI and ends with reviewing couple of white papers where ideas have already taken shape of actual products.

A. Brain as DAC

Melanie Swan in 2015 [1], proposed a research titled "Blockchain Thinking: The Brain as a DAC (Decentralized Autonomous Organization)". It was one of the early research works where implementation of AI on top of Blockchain was conceptualized. She proposed a concept where the system implemented on Blockchain simulates brain and performs thinking. In her definition of thinking she described it as input-processing-output paradigm. She also proposed architecture to implement it. However she did not provide the implementation. In her architecture she describes memory, storage and files as inputs. She proposes three layers of memory - layer 1 has IPFS, layer 2 has blockchain and layer 3 has GitHub as memory. Active processing happens on blockchain based memory in decentralized way. She also proposes a new consensus algorithm called "Proof of Intelligence" [1] where she argues need of consensus for different application can be different so industry standard of

PoW [2] or PoS [3] should be replaced with Proof of Intelligence. Output of the system can be in terms of action, feedback loop or notification. This paper provides very high level of concept and does not carry out any experiment to generate results. However at conceptual level it talks about creating a system where data is distributed in memories at various level and AI algorithms may run on top of them in decentralized and parallel fashion to generate outputs.

B. Decentralized, Trustless and Collaborative AI

In 2018, Mendis et al [4], proposed implementation of Deep Learning on blockchain for distributed data processing and collaborative development. Their work focuses on two major problems in Deep Learning that is availability of the data and availability of the processing power. Availability of data is affected by ownership and privacy control of the data. Processing also needs costly hardware systems like clusters of GPUs. Mendis et al proposed an innovative architecture based on Blockchain to solve these challenges. In their proposal they consider different actors connected by a blockchain network. The actors are - Problem Initiator [4] who initiates deep learning model, Processing Contributor [4] who trains the model on their own set of data and Verification Contributor [4]who verifies the trained models with their own test data set. The data for training is stored on its owner's node. Hence it is distributed and remains private to the owner. The processing is also distributed as it happens where the data is stored. Initiator creates the deep learning model. Processing contributors train the model on their own data and store the training parameters into IPFS. Smart contracts triggers events and invokes verification contributor who then tests the deep learning model. These models then go through fusion process. The researchers have proposed two fusion models. First model is based on probabilistic softmax function and second one uses gradual model fusion. In their simulation they found the second fusion model giving slightly better results than first one. The overall accuracy is at par with centralized learning models.

The availability and privacy of training data is of utmost importance to any Machine Learning model. Researchers Galtier et al in 2017 [5], proposed a new model of traceable machine learning that keeps training data hidden from the users. The proposed system uses Blockchain technology to enable traceability of learning. They proposed architecture which makes use of transfer learning where contributors can contribute training data that is kept encrypted in the database. Key of the encryption algorithm is stored in client application on contributor's machine and in a blockchain on server. Machine learning model fetches the data from database and key from blockchain to perform training. Apart from machine learning algorithm, other actors of the system do not have access to decryption key. Also while performing training, the algorithm creates logs which contain information about data and model. These logs are stored in blockchain. Blockchain's temper proof and traceable data storage makes this training process traceable. However the model uses centralized architecture to store and process the data which will face empirical challenge of having high amount of storage and processing power.

When multiple AI or Machine Learning models are created for a specific purpose by different parties, there is scope of utilizing knowledge created by other parties. However this needs trust among different parties sharing their knowledge with each other. Researchers at IBM research and University of Illinois at Urbana-Champaign in 2018 [6], proposed using Blockchain for establishing trust among multiple machine learning and digital simulation models [6]. They used OpenMalaria framework for their experiment. Their proposed architecture has several elements which provide external interface, processes training data, suggests relevant models and validates it. This architecture logs all the processing information, results and validation results in blockchain. Blockchain being temper proof can be considered source of trusted information. Researchers have implanted this system as set of microservices with each component deployed on different server. Experiment involves 144 users using OpenMalaria models and each user generating 8 OpenMalaria simulations. Results generated by these simulations are validated to identify the dishonest workers. Thus this model can be used to establish trust among different users / parties using machine learning model and also identify and eliminate the dishonest users.

Collaborative Machine Learning is increasingly becoming popular as the data and its processing is becoming distributed and owned by different parties. However such systems based on centralized architecture are prone to many security breaches. Researchers from University of British Columbia, Shayan et al, in 2018 [7], proposed a peer-to-peer machine learning model which is based on Blockchain. This model is claimed to be more secure then the centralized model. It takes care of poisoning attack [7] and information leakage attack [7] . The proposed model is based on Stochastic Gradient Decent (SGD) [7] algorithm. Every peer in the Blockchain network has data to contribute to the learning process. In each iteration, local SGD is calculate by peer who then sent it to the network but before sending a noise is added to the result to mask the data. The noise is generated using Variable Random Function (VFR) [7] by a group of other peers. The masked updates are then verified by another group of peers to prevent poisoning attack. On successful verification each peer in verification group signs the update. Majority peers in this group must sign the update to consider it valid. The updates are then sent to third group of peers called "aggregation committee" [7], who then aggregate the updates. In their research the researchers found that this proposed system is able to tolerate up to 48% of the peers being malicious. The method not only performs distributed learning on distributed data, but also defends learning process from malicious attacks.

Blockchain's original usecase of cryptocurrency can add financial ability to the AI market. Kurtulmas et al, in 2018 [8] proposed creating a Blockchain based market place where AI models can be traded. In real world the people who own data does not have AI capabilities and the developers who can develop AI systems does not have access to data. Proposed platform brings these two parties on a common platform. Data owner who needs a machine learning model submits the data to the platform along with a criteria for evaluation of the model and a reward in terms of cryptocurrency. The developers of machine learning model gets the data from platform develop a model and submit it back to blockchain. The verification data is kept separate to make sure the developers does not overfit the model to data. It happens at later point of time with separate test dataset and criteria set by the data owner while initiating the model. On successful verification, the developer receives their reward in form of cryptocurrency. This system is implemented using Ethereum smart contracts. As ethereum do not support full math library activation functions like Sigmoid is difficult to implement. Hence the solution uses ReLU as activation function. Also storing large data in Ethereum is costly as it generates more amount gas [9]. Hence alternative like IPFS should be explored.

Prediction has extremely important role to play in business decision making. Kuo et al in 2018 [10], created a prediction model for healthcare. Using Blockchain they brought multiple healthcare institutions on a common platform. A patient who has already been to multiple institutions has his data distributed. Aggregation of data may not be possible due to legal and privacy issues. Hence researchers developed a Blockchain based model to achieve distributed and privacy preserving machine learning. They apply online machine learning on top of Blockchain and a newly developed consensus algorithm called "Proof of Information" [10]. This algorithm determines the order of the machine learning and improves accuracy of the model in incremental manner. This proposed models support distributed data with distributed processing and making sure that the privacy of the data is preserved by not moving data away from owner.

Swarm systems are discipline of robotics which has found its application in variety of use cases from agriculture to food delivery. The robots have unique capability of not just having global behaviors but also developing local behaviors by communicating with each other. If they communicate using centralized architecture then there will be challenges of availability, security and scalability. Eduardo Castelló Ferrer in 2017 [11], proposed a new communication architecture for swarm robots based on Blockchain. As Blockchain is peer-to-peer network it more appropriate for this type of usecase. Here different robots can be part of a blockchain and perform task of transmitting and receiving messages from one peer to another. On this kind of network, the message may have to travel less and can be sent or received without intermediate point of contact like server. In his proposal Ferrer assigns a public-private pair of keys to each robot. The robots are identified on the blockchain network with their public key. Now if robot A wants to send message to robot B then the message will be encrypted using B's public key and broadcasted on the network. When robot B receive message, it will decrypt it using its own private key. Hence the communication using this method will be more secure, efficient and reliable then centralized model. Also as the messages are encrypted they are secured from the malicious robots in the network.

C. Whitepapers

Most of the AI systems today either work in isolation or in closed group. SingularityNET [12] build a protocol to interconnect different AI systems which can then coordinate with each other using this platform. SingularityNET works on ERC20 [12] protocol using which they have created an opensource platform. Anyone can add AI or Machine Learning service to SingularityNET which will be then consumed by others on the network. This creates market place for host of coordinated AI services which are decentralized on Blockchain. Hence SingularityNET aims to become a network of many AI services hosted in decentralized way by different developers. Anyone who has ERC20 compliant blockchain account can join the network and host of consume any AI service.

For Deep Learning biggest challenge is to have huge cluster of GPUs so that the training of the models can happen in reasonable time. Now blockchain miners traditionally use hardware which has clusters of GPUs. DeepBrain Chain [13] is a product which makes use of these mining GPUs for deep learning purpose. DeepBrain has created a huge cloud of GPUs that can be used for training deep learning models. Owners of the GPUs receive a rent for leveraging their hardware in form of cryptocurrency. The product is built on ERC20 standard. DeepBrain Chain makes sure that the owners of the GPUs do not get access to the data hence maintaining privacy.

III. DISCUSSION ON SCOPE OF FUTURE WORK

All the research work done for integrating AI with Blockchain takes benefits from many of the inherent properties of Blockchain like security, traceability, decentralized storage and processing etc. However there are certain properties like consensus, redundancy etc have potential to contribute to futuristic AI systems.

AI expert system can be developed where consensus among different AI models can act as opinions of different experts in real life. Consensus algorithm can be developed in such a way that opinion of each AI model is considered in a specific way to build final decision.

Consensus can also be used to encourage competition among different AI models. PoW is one such competition which puts a cryptographic challenge in front of participants. Nature of this puzzle can be changed in such a way that participants need AI capabilities to solve. Such consensus algorithm will allow only the smartest to commit the work rather than the fastest.

Another important property of Blockchain is redundancy. All the data is stored on each node on the network. If training data is stored in blockchain, it will become possible to train multiple models on the same data in parallel and then compare them and choose the best one. This can accelerate development process of the AI models.

Many of these ideas are currently difficult to implement as they may generate huge amount of gas or require special features. Current Blockchain frameworks do not have sufficient capabilities to implement AI systems. This is due to the fact that they were not developed with this expectation. It will not be exaggerating to expect Future Blockchain frameworks to have better support to applications beyond cryptocurrency.

IV. CONCLUSION

Blockchain's contribution to areas like AI has produced encouraging results in research. Blockchain's peer-to-peer networking and highly collaborative nature, has made it suitable for decentralized AI applications. This paper has reviewed 8 research papers and 2 white papers. The existing researches have tried to address many challenges faced by AI and contribute additional features to it.

Challenges like storage and processing of data, security, collaboration with other AI systems have been addressed by these research works. The experiments conducted had produced encouraging results. This paper has reviewed research works where significant improvements in distributed processing and distributed data storage are achieved. Also certain research works has focused on achieving trustless collaboration among different AI applications. Security enhancement has been inherent feature in all works.

Considering developments in Blockchain as technology and its proposed updates, its contribution to AI will be much more than its current state. Blockchain technology frameworks are also undergoing active development and with more features to offer, Blockchain is poised to become one of the most preferred platforms for all types of decentralized application and AI will be no exception.

REFERENCES

[1] M. Swan, "Blockchain Thinking: The Brain as a DAC(Decentralized Autonomous Organization)," in Texas Bitcoin Conference, 2015.

[2] "Proof of Work," Wikipedia, 16 01 2019. [Online]. Available: https://en.wikipedia.org/wiki/Proof-of-work_system. [Accessed 21 01 2019].

[3] "Proof of Stake," wikipedia, 18 01 2019. [Online]. Available: https://en.wikipedia.org/wiki/Proof-of-stake. [Accessed 21 01 2019].

[4] G. J. Mendis, M. Sabounchi, J. Wei and R. Roche, "Blockchain as a Service: An Autonomous, Privacy Preserving, Decentralized Architecture for Deep Learning," 05 07 2018. [Online]. Available: https://arxiv.org/abs/1807.02515. [Accessed 21 01 2019].

[5] M. Galtier and C. Marini, "Morpheo Traceable Machine Learning on Hidden data," [Online]. Available: https://arxiv.org/abs/1704.05017.

[6] N. K. Bore, R. K. Raman, I. M. Markus, S. L. Remy, O. Bent, M. Hind, E. K. Pissadaki, B. Srivastava, R. Vaculin, K. R. Varshney and K. Weldemariam, "Promoting Distributed Trust in Machine Learning and Computational Simulation via a Blockchain Network," [Online]. Available: https://arxiv.org/abs/1810.11126. [Accessed 20 01 2019].

[7] M. Shayan, C. Fung, C. J. Yoon and I. Beschastnikh, "Biscotti: A Ledger for Private and Secure Peer-to-Peer Machine Learning," 27 11 2018. [Online]. Available: https://arxiv.org/abs/1811.09904. [Accessed 21 01 2019].

[8] A. B. Kurtulmus and K. Daniel, "Trustless Machine Learning Contracts; Evaluating and Exchanging Machine Learning Models on the Ethereum Blockchain," 2018.

[9] T.-T. Kuo and L. Ohno-Machado, "ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks," [Online]. Available: https://arxiv.org/abs/1802.01746. [Accessed 20 01 2019].

[10] E. C. Ferrer, "The blockchain: a new framework for robotic swarm systems," in FTC 2018: Proceedings of the Future Technologies Conference (FTC) 2018, 2018.

[11] SingularityNET Foundation, "SingularityNET: A decentralized, open market and inter-network for AIs," [Online]. Available: https://public.singularitynet.io/whitepaper.pdf. [Accessed 20 01 2019].

[12] DeepBrain Chain Foundation, "DeepBrain Chain Whitepaper," [Online]. Available: https://www.deepbrainchain.org/assets/pdf/DeepBrainChainWhitepaper_en.pdf. [Accessed 20 01 2019].

[13] "Blockchain," Wikipedia, 02 12 2018. [Online]. Available: https://en.wikipedia.org/wiki/Blockchain. [Accessed 03 12 2018].

[14] "Machine Learning," Wikipedia, 30 11 2018. [Online]. Available: https://en.wikipedia.org/wiki/Machine_learning. [Accessed 03 12 2018].

[15] Wikipedia, "Ethereum," Wikipedia, 19 01 2019. [Online]. Available: https://en.wikipedia.org/wiki/Ethereum. [Accessed 21 01 2019].

