IJFEAT
INTERNATIONAL JOURNAL FOR ENGINEERING APPLICATIONS AND TECHNOLOGY

Use of Machine learning in Healthcare System using Blockchain Technology

Dr. Avinash P.Jadhao
Assistant Professor
Department of Computer Sci & Engineering
J.D.I.E.T, Yavatmal, India
apjadhao@gmail.com

Dr.Dinesh N.Chaudhari
Professor
Department of Computer Sci & Engineering
J.D.I.E.T, Yavatmal, India
dnchaudhari2007@rediffmail.com

Abstract:

In today's world information systems and computerization of business processes by organizations have led to a faster, secure, easier and more accurate data analysis and accuracy. The machine learning techniques have been used increasingly in the analysis of data in various fields from medicine to organization, education and energy applications. Machine learning techniques make it possible to deduct meaningful further information from those data processed by using Different Method. Such meaningful and significant information helps organizations to establish their future policies and able to provide security. This study applies classification machine learning techniques to Process the data also survey of active learning regarding selection methods, query strategies, applications and security. For Security we are using BlockChain Technology.

Blockchain technology is rapidly gaining traction in healthcare industry as one of the most exciting technological developments. In particular, blockchain technology presents numerous opportunities for healthcare industry such as reduced transaction costs, increased transparency for regulatory reporting, efficient healthcare data management and healthcare records universality as well as able to access data from any location. In the context of smart health, blockchain may provide distinct benefits, particularly from a context-aware perspective where efficient and personalized solutions may be provided to citizens and the society in general. In this paper, we are going to discuss relationship between Machine learning and blockchain related to smart health care system. In addition, we discuss several challenges for actually implementing machine learning using blockchain based secure applications in the healthcare industry along with several opportunities for future research directions.

***Keyword*:** Machine Learning, Blockchain, Security.

I. Introduction:

In this world information systems and computerization of business organizations focuses mainly on future predication by consider the historical data. Machine Learning is an approach or subset of Artificial Intelligence that is based on the idea that

machines can be given access to data along with the ability to learn from it. It learns from examples and experience, without being explicitly programmed. Instead of writing code, you feed data to the generic algorithm, and it builds logic based on the data given. In a very layman manner, Machine Learning (ML) can be explained as automating and improving the learning process of computers based on their experiences without being actually programmed i.e. without any human assistance [1]. The process starts with feeding good quality data and then training our machines (computers) by building machine learning models using the data and different algorithms. The choice of algorithms depends on what type of data do we have and what kind of task we are trying to automate [2].

Need of Machine Learning:

Machine Learning is a field which is raised out of Artificial Intelligence (AI). Applying AI, we wanted to build better and intelligent machines. But except for few mere tasks such as finding the shortest path between point A and B, we were unable to program more complex and constantly evolving challenges.

There was a realization that the only way to be able to achieve this task was to let machine learn from itself. This sounds similar to a child learning from its self. So machine learning was developed as a new capability for computers. And now machine learning is present in so many segments of technology, that we don't even realize it while using it. Finding patterns in data on planet earth is possible only for human brains [2]. The data being very massive, the time taken to compute is increased, and this is where Machine Learning comes into action, to help people with large data in minimum time. If big data and cloud computing are gaining importance for their contributions, machine learning as technology helps analyze those big chunks of data, easing the task of data scientists in an automated process and gaining equal importance and recognition. The techniques we use for data mining have been around for many years, but they were not effective as they did not have the competitive power to run the algorithms. If you run deep learning with access to better data, the output we get will lead to dramatic breakthroughs which is machine learning.

Blockchain Technology:

Blockchain Technology Distributed Ledger Technology (DLT) has attracted widespread attention in recent years. DLT is a transparent, distributed, secure data storage and transfer technology that works without any centralized trusted third party. A distributed ledger is a decentralized database that is maintained by several nodes over a peer-to-peer network. The ledger is verified and replicated by each node. Blockchain is one form of DLT [3]. The blockchain organizes data into blocks, which are chained together using an append-only structure. The chainbased block structure is the most popular data structure of DLT, but it is not the only one. There are other data structures to implement DLT, such as Directed Acyclic Graph (DAG). The DAG-based DLT can be divided into two categories: blockDAG and Transaction DAG (TDAG). BlockDAG is a DAG structure, where each block is allowed to reference multiple previous blocks. Inclusive BlockDAG and Spectre are two examples of blockDAG systems. In the TDAGbased DLT, transactions are directly added to a graph, forming a graph of transactions. Each transaction references multiple previous transactions.

Blockchain is the most widely used distributed ledger technology, so in this paper, we mainly focus on blockchain technology. Blockchain systems are typically classified into three categories: public blockchain, consortium blockchain and private blockchain . The public blockchain is permissionless blockchain, while both consortium blockchain and private blockchain are permissioned blockchain. In the public blockchain, anyone is allowed to join the network, participate in the consensus process, read and send transactions, and maintain the shared ledger. Most crypto currencies and some open-source blockchain platforms are permissionless blockchain systems. Bitcoi and Ethereum are two representative public blockchain systems. Bitcoin is the most famous crypto currency that is created by Satoshi Nakamoto in 2008. Ethereum is another representative public blockchain that supports extensive decentralized applications using its Turing-complete smart contract programming languages [4].

Need of Blockchain:

The blockchain mainly used for Security purpose for securing the data. The data sent to each block on the distributed ledger is based on encrypted Merkle Trees, which is a technical way of saying that no fraudulent transactions can be recorded. If any transaction that does not follow protocol rules is detected by the network nodes, it is expelled immediately [7]. This inherently secure nature of distributed blockchain technology means that it prevents damage to the entire blockchain shared database and can cut off a hacking attempt at one block [8]. Due to this secure nature blockchain mainly used where more Protection is needed to Database.

Categories of BlockChain:

Blockchain structures fall into three categories:

- Public blockchain architecture

  A public blockchain architecture means that the data and access to the system is available to anyone who is willing to participate (e.g. Bitcoin, Ethereum, and Litecoin blockchain systems are public).
- Private blockchain architecture

  As opposed to public blockchain architecture, the private system is controlled only by users from a specific organization or authorized users who have an invitation for participation.
- Consortium blockchain architecture

  This blockchain structure can consist of a few organizations. In a consortium, procedures are set up and controlled by the preliminary assigned users [9].

The Three Pillars of Blockchain Technology:

The three main properties of Blockchain Technology which has helped it gain widespread acclaim are as follows:

- Decentralization
- Transparency
- Immutability

II. Healthcare System

Currently, there are several gaps in the medical care industry related to health records, particularly on how healthcare related data from different sources is accessed, processed, and analyzed from the various healthcare players and shareholders. For instance, there

are currently no universal health records to easily transfer lab tests, imaging, or a patient's medications between visits. Even though paper records for most medical practices have evolved to digital, all of that data currently resides in silos, where consumers attempt to reconcile data among their providers and health payers.

This can be challenging, as there is no single source that identifies where all of the health data of an individual's reside, let alone the order in which it was recorded. Interoperability and varying data standards are significant challenges that have vexed stakeholders for many years. Years of heavy regulation and bureaucratic inefficiency have slowed innovation for EHRs. In addition, there is a strong case of double spending in healthcare that might dwarf the case for finance (duplicate tests such as blood test and MRI scans). These unnecessary and expensive tests are wasteful, while simultaneously put the patient at risk. Prescription fraud also faces a similar problem. From the proliferation of digital health data comes another challenge: that of keeping the data secure. The past few years have seen an explosion of data breaches and medical identity theft. Consequently, health care providers are looking for more effective ways to secure the personal health information.

Generally, there are four fundamental challenges in the healthcare industry related to managing EHRs.

1) Various regulations, while valuable and necessary, make it extremely hard to efficiently combine and share data between people and various stakeholders,

2) Patients often have limited access to their own data and their management.

3) All data not available on any decentralized system for accessing.

4) Patient Not able to get time to time updation related to health.

III. Supervised & Unsupervised Method in Medical Field

Supervised learning:

Supervised learning starts with the goal of predicting a known output or target. In machine learning competitions, where individual participants are judged on their performance on common data sets, recurrent supervised learning problems include handwriting recognition (such as recognizing handwritten digits), classifying images of objects (e.g. is this a cat or a dog?), and document classification (e.g. is this a clinical trial about heart failure or a financial report?). Notably, these are all tasks that a trained person can do well and so the computer is often trying to approximate human performance. Supervised learning focuses on classification, which involves choosing among subgroups to best describe a new data instance, and prediction, which involves estimating an unknown parameter (such as the temperature in San Francisco tomorrow afternoon).

What might be some examples of supervised learning in medicine? Perhaps the most common example seen by a cardiologist is the automated interpretation of the EKG, where pattern recognition is performed to select from a limited set of diagnoses (i.e. a classification task). In radiology, automated detection of a lung nodule from a chest X-ray would also represent supervised learning. In both these cases, the

computer is approximating what a trained physician is already capable of doing with high accuracy.

Supervised learning is often used to estimate risk. The Framingham Risk Score for coronary heart disease (CHD) may in fact be the most commonly used instance of supervised learning in medicine. Such risk models exist across medicine, and include guiding antithrombotic therapy in atrial fibrillation and implantation of automated implantable defibrillators in hypertrophic cardiomyopathy. In modeling risk, the computer is doing more than merely approximating physician skills but finding novel relationships not readily apparent to human beings.

Unsupervised learning

In contrast, in unsupervised learning, there are no outputs to predict. Instead, we are trying to find naturally occurring patterns or groupings within the data. This is inherently a more challenging task to judge and often the value of such groups learned through unsupervised learning is evaluated by its performance in subsequent supervised learning tasks (i.e. are these new patterns useful in some way?).

When might such approaches be used in medicine? Perhaps the most compelling opportunity represents the "precision medicine" initiative. Frustrated by the inherent heterogeneity in most common diseases, there is a growing effort to redefine disease according to pathophysiologic mechanisms, which could, in turn, provide new paths to therapy. But identifying such mechanisms for complex multifactorial diseases will not be easy. Let us think about how one might apply unsupervised learning in cardiac disease towards that end, taking a heterogeneous condition like myocarditis. One can start with a large group of apparently similar individuals with unexplained acute systolic heart failure. One can then perform myocardial biopsies on them, and characterize the cellular composition of each sample with a technique such as immunostaining [20]. For example, one would have a tally of T lymphocytes, neutrophils, macrophages, eosinophils, etc. One could then see if there are recurring patterns of cellular composition, which, in turn, might suggest mechanism and guide therapies to explore. A similar approach, albeit focused on genomics, led to identifying an eosinophilic subtype of asthma, which uniquely responds to a novel therapy targeting the eosinophil-secreted cytokine IL-13. Note the contrast with supervised learning – there is no predicted outcome – we are only interested in identifying patterns in the data. In fact, treating this as a supervised learning problem – such as developing a model of mortality in myocarditis and classifying patients by risk – might miss such subgroups completely, thereby losing a chance to identify novel disease mechanisms.

Data Science of healthcare data analytics:

There has been information explosion of big data in the healthcare field. Traditional technologies adopted earlier to analyze genomics, DNA, and cancer with trial and methods through Human Genome Project have taken more than a decade to understand and analyze the composition of DNA and the patterns of the data. Big Data Analytics introduced revolutionary tools and techniques to analyze the chronic diseases for prevention and cure. Genome sequencing has been used to understand the potential root causes of tumor growth causing cancer. The data has grown exponentially from

terabytes to exabytes [14]. The healthcare data from X-Rays, CT scan and MRI has increased by leaps and bounds concerning the volume of the big data. The advanced technologies of medicine through big data analytics allowed to diagnose the patients records and perform a comparison to a global population to separate the noises from the signal to understand the trends of the tumor growth which was not possible earlier and speed up the diagnosis and treatment. Though there are several theories and techniques that can be applied for the diagnosis of the illnesses, this paper briefly reviews some of the key techniques.

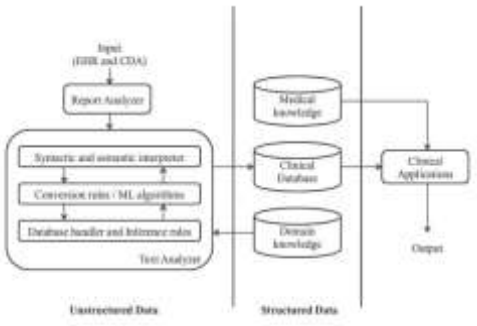

Figure 01: How Healthcare related data Process

One of the most exciting and practical goals of combining healthcare with technology is to mine large quantities of data to discover what, if anything, has eluded researchers—either through a lack of sufficiently large datasets or a lack of human ability to notice unlikely relationships. Unsupervised learning is a promising avenue for pursuing this goal, because unsupervised machine learning techniques do not require existing human knowledge to generate new insights about structure within datasets [15]. This video, designed for learners with a basic understanding of statistics and computer programming, provides a detailed introduction to three specific types of unsupervised learning: cluster analysis, association analysis, and principal components analysis, as applied to health data sets both at the individual and population levels. Examples will be introduced in both Python and R.

IV. Conclusion:

Today, it is inevitable to consider and use Machine Learning in view of the ever-increasing amount of computerized business processes and the huge amount of data to be analyzed in parallel. It is possible to make accurate estimations or predictions about future results through applying machine learning techniques to the data made available for analysis. Also we have discuss how the data can be secure. This study used various classification techniques and Blockchain technology to a group of individuals who are in the process for accessing the Data. Since using machine learning techniques and blockchain method in classification studies results in accurate outcomes accompanied with significant saving in terms of time , cost and Security, it is highly recommended to make use of machine learning techniques and blockchain method in Data Processing and for providing the security to access those data. Due to it patient can able to access data from anywhere with proper authorization which may save the time & cost.

REFERENCES:

1. Sabyasachi Dash, Sushil Kumar Shakyawar , Mohit Sharma and Sandeep Kaushik, “Big data in healthcare: management, analysis and future prospects”,

https://doi.org/10.1186/s40537-019-0217-0,springer 2019.

2. Karim Abouelmehdi, Abderrahim Beni‑Hessane and Hayat Khaloufi, "Big healthcare data: preserving security and privacy", https://doi.org/10.1186/s40537-017-0110-7,2018.
3. Kashif Saleem, Xiaodong Yang, Abdelouahid Derhab, Mehmet A. Orgun, Waseem Iqbal, Imran Rashid, And Asif Yaseen "Privacy Preservation in e-Healthcare Environments: State of the Art and Future Directions" , Received August 25, 2017, accepted October 11, 2017, date of publication October 30, 2017, date of current version February 14, 2018. Digital Object Identifier 10.1109/ACCESS.2017.2767561.
4. Tareq Ahram, Arman Sargolzaei, Saman Sargolzaei, Jeff Daniels, and Ben Amaba, "Blockchain Technology Innovations", 978-1-5090-1114-8/17/$31.00 ©2017 IEEE.
5. Pinyaphat Tasatanattakool, Chian Techapanupreeda, "Blockchain: Challenges and Applications", 978-1-5386-2290-2/18/$31.00 ©2018 IEEE.
6. Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends", 978-1-5386-1996-4/17 $31.00 © 2017 IEEE DOI 10.1109/BigDataCongress.2017.85.
7. Mohamed Amine Ferrag, Makhlouf Derdour, Mithun Mukherjee, Member, IEEE, Abdelouahid Derhab, "Blockchain Technologies for the Internet of Things: Research Issues and Challenges", 2327-4662 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/ publications_standards/ publications/rights/index.html for more information.
8. P. Otte, M. de Vos, and J. Pouwelse, "TrustChain: A Sybil-resistant scalable blockchain," Futur. Gener. Comput. Syst., sep 2017.
9. Q. Wang, B. Qin, J. Hu, and F. Xiao, "Preserving transaction privacy in bitcoin," Futur. Gener. Comput. Syst., sep 2017.
10. J.-H. Lee, "BIDaaS: Blockchain Based ID As a Service," IEEE Access, vol. 6, pp. 2274–2278, 2018.
11. K. Fan, Y. Ren, Y. Wang, H. Li, and Y. Yang, "Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G," IET Commun., vol. 12, no. 5, pp. 527–532, Mar. 2018.

12. Q. Lin, H. Yan, Z. Huang, W. Chen, J. Shen, and Y. Tang, "An IDbased linearly homomorphic signature scheme and its application in blockchain," IEEE Access, pp. 1–1, 2018.
13. Mohamed Elhoseny , Gustavo Ramírez-González, Osama M. Abu-Elnasr , Shihab A. Shawkat, Arunkumar N, And Ahmed Farouk, "Secure Medical Data Transmission Model for iot-Based Healthcare Systems" , 2169-3536 2018 IEEE. Translations and content mining are permitted for academic research only.
14. P.K. Kavitha, P. Vidhya Saraswathi, "A Survey on Medical Image Encryption" , ICASCT2501 | ICASCT | March-April-2017 [(3) 5 : 01-08]
15. Javaria Tahir ,Nadeem Javaid, "Bootstrapping a Blockchain Based Ecosystem for Big Data Exchange", 978-1-5386-1996-4/17 $31.00 © 2017 IEEE DOI 10.1109/BigDataCongress. 2017.67
16. Eranga Bandara , Wee Keong Ng , Kasun De Zoysa ,Newton Fernando , Supun Tharaka P. Maurakirinathan , Namal Jayasuriya, "Mystiko Blockchain Meets Big Data" , 978-1-5386-5035-6/18/$31.00 ©2018 IEEE.
17. Massoud Sokouti, Ali Zakerolhosseini and Babak Sokouti, "Medical Image Encryption: An Application for Improved Padding Based GGH Encryption Algorithm" , The Open Medical Informatics Journal, 2016, 10, 11-22.
18. Alexandru Soceanu, Maksym Vasylenko, "Managing the Privacy and Security of ehealth Data", 978-1-4799-1780-8/15 $31.00 © 2015 IEEE ,DOI 10.1109/CSCS.2015.76.
19. Vidhya Ramani, Tanesh Kumar, An Braeken, Madhusanka Liyanage, Mika Ylianttila, "Secure and Efficient Data Accessibility in Blockchain based Healthcare Systems", December 2018 DOI: 10.1109/GLOCOM.2018.8647221.
20. Jingwei Liu_, Xiaolu Li_, Lin Yey, Hongli Zhangy, Xiaojiang Duz, and Mohsen Guizanix, "BPDS: A Blockchain based Privacy-Preserving Data Sharing for Electronic Medical Records", 978-1-5386-4727-1/18/$31.00 ©2018 IEEE.
21. Pinyaphat Tasatanattakool, "Blockchain: Challenges and Applications", 978-1-5386-2290-2/18/$31.00 ©2018 IEEE.