Abstract

The importance of big data in machine learning cannot be overemphasized in recent times. Through the evolution of big data, most scientific technologies that relied heavily on enormous data in solving complex issues in human lives gained grounds; machine learning is an instance of these technologies. Various machine learning models that yield groundbreaking throughputs with high efficiency rates in predicting, detecting, classifying, discovering and acquiring in-depth knowledge about events that would otherwise be very difficult to ascertain have been made possible due to big data. Although big data has undoubtedly helped in the field of machine learning research, over the years, its mode of acquisition has posed great challenge in industries, education and other agencies that obtained them for various purposes. This is because these large quantities of data cannot be stored on personal computers with limited storage capability but required the use of high storage capacity servers for effective storage. These servers may be owned by a group of companies or individuals who had the singular privilege to modify the data in their possession as and when deemed relevant thus the creation of a centralized data storage environment. These were mostly refered to as the Third Parties
(TP) in the data acquisition process. For the services they rendered, these trusted parties priced data in their possession expensively. The adverse effect is a limitation on various researches that could help solve a number of problems in human lives. It is worth mentioning that the security of these data being purchased expensively cannot be even assured limiting various researches that thrive on secured data. In order to curb these occurrences and have better machine learning models, the incorporation of Blockchain Technology databases into machine learning. This paper discusses the concept of big data, Machine Learning and Blockchains. It further discusses how Big data has impacted the Machine learning Community, the significance of Machine Learning and how the BlockChain Technology could be used similarly impact the Machine Learning Community. The aim of this paper is to encourage further research in incorporating the BlockChain Technology into Machine Learning.

References

8. Jiawei Han, Jian Pei, and Micheline Kamber. Data mining: concepts and techniques. Elsevier, 2011.

**Index Terms**

Computer Science

Information Sciences

**Keywords**

Big Data, Machine Learning, Blockchains, Data Preprocessing