2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)

An Exploratory analysis of Machine Learning adaptability in Big Data Analytics Environments: A Data Aggregation in the age of Big Data and the Internet of Things

Ratchana Rajendran
Assistant Professor
International School ofManagement Excellence
Bengaluru, India.
r.ratchana@gmail.com

Priyanka Sharma
Assistant Professor
Faculty of Computer Applications
MRIIRS
Faridabad, India.
priyankasharmaiitd2@gmail.com

Nitin Kumar Saran
Sr. Technical Manager
SPV-M- Project Execution Division
Central Electronics Limited
Ghaziabad, Uttar Pradesh, India.
nitin.saran@gmail.com

Samrat Ray
PHD student, Economics
The Institute ofIndustrial Management
Economics and Trade
Peter The Great Saint Petersburg Polytechnic University
Russia.
samratray@rocketmail.com

Joel Alanya-Beltran
Professor
Electronic Department
Universidad Tecnológica del Perú
C18121@utp.edu.pe

Korakod Tongkachok
Professor
Faculty of Law
Thaksin University
Thailand.
tokorakod@tsu.ac.th

***Abstract*— The paper discusses a new concept combining the potentialities of Big Data processing as well as machine learning developed for security monitoring of mobile Internet of Things. The structure of the security monitoring system is considered as a most effective and useful element to create a new viewpoint of mobile IoT. This article focuses implementation of machine learning in online education. Thus mobile IoT has found successful application in few areas such as security monitoring in public places, transport management,medicine, smart houses, industrial production, electrical consumption, and robotics. All the mathematical foundations along with issues related to this have been considered in this study. In order to solve the classification task, several machine learning mechanisms have been mentioned in this paper. Large organizations are incorporating data-driven actions, and decision making in organizational function. The role of data aggregation is effective here achieving the business objectives. Vast amount of raw data can be processed linearly through data aggregation. This article describes the interaction of data aggregation through wireless networking assuming its effectiveness in online education. Data aggregation in machine learning is highlighted based on evidence based data.**

The purpose of this research article is to investigate the machine learning adaptability in big data analytics environments with the approach of IoT. In order to collect accurate data, the researcher has taken the help of a secondarydata collection method. It has helped the researcher to find out the valid information about mobile IoT. In addition, qualitativemethods have been adapted to malaise the collected data within a systematic way. Moreover, this study will help the readers to understand the value of mobile IoT helping in machine learning adaptability in big data analytics.

Keywords*— *Big Data, IoT, Machine learning, data analytics,data aggregation, integration of ML, ML adaptability

I. INTRODUCTION

Presently, the entire society is affected by a sudden health emergency, and it has changed the way of living. Covid-19 has explored fragility in all the fields including heath, economy, and also in the education sector. Important tools and techniques that are helping to combat the situation are ICT (Information and Communication technologies). It has allowed us to carry out most of the work securely, and remotely. In this article the integration of ML, AI, and Big data analysis in online education is illustrated focusing on student learning. ML as the part of AI is incorporated in the education model to manage the performance of the students. Key objective of this study is to emphasize Machine learning and big data analysis improving academic activities. The study describes machine learning adaptability in the education sector identifying the IoT environment. Different phases of the ML implementation can be useful to assume its importance in the upcoming days.

II. LITERATURE REVIEW

A. Integration of Machine Learning, and Big Data

For system integration, and new technology a particular model can be followed (shown in Figure 1) that represents the circulation of large amounts of data emphasizing interaction, and different activities in the academic sector. It is not direct; therefore it is common there providing information in big data analysis. This model shows the activetime on the platform.

978-1-6654-6643-1/22/$31.00

2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM) | 978-1-6654-6643-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICIPTM54933.2022.9753921

Along with that it helps to note the usual connection schedule of the students. These data are stored within the database of different academic and administrative systems. This crucial information allows us toacquire the analysis covering greater variables including big data architecture in the processing charge [1].

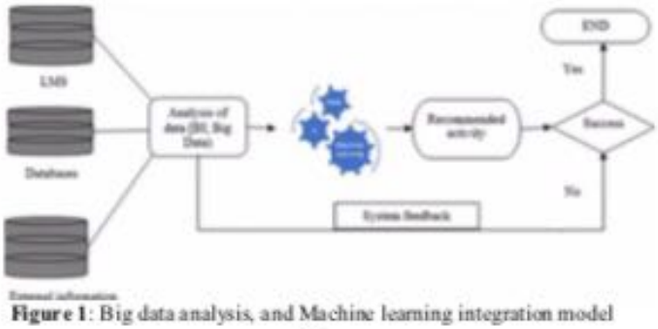

Figure 1: Big data analysis, and Machine learning integration model [3]

The framework of big data analysis is responsive inthe initial phase extracting data from multiple sources, and these obtained data may be unstructured, and structured.Once it collects all the necessary data it is possessed in a useful way to obtain the knowledge that creates artificial intelligence through ML. Machine learning is found responsible in recognizing important patterns of its analysis,and then it performs in the classification. Those patterns are represented as significant characteristics in each group.

Major objective is to find the strategies, and technique knowing the requirements of each group, and the capable system improving the way of presented activities. Furthermore, it helps to enhance the learning ways recommending the activities to academicians, and students according to their requirements. After giving the important recommendations, ML enters into an important stage of result analysis. In order to accomplish this system emphasizes to analyse the grades that the participants obtain with the recommended activities in a disciplined way. If it represents a positive outcome, the process ends at that stage,and it goes back to the initial stage of the entire system [4]. Ifthe system recognizes that there is no significant improvement in the performance it integrates the data into the analysis phase where it starts to process again to obtain asatisfactory result.

Machine learning core consist specific learning algorithmthat evolves continuously to improve the assigned tasks. By accurate structure, and proper data, these eventually produce important data of recognition, and the pattern of predictive modelling. The machine learning procedure the data procession is just like exercise that is in tune. There are multiple programming language of machine learning such as Python, Scala, Java that are used as comprehensive tools of machine learning. Big data is the data consisting observation, measurement, words, or numbers. It indicates vast set of data including both unstructured, and structured. Cloud network, web scraping, and mix initiate systems are several measures to integrate machine learning with big data analytics.

B. Online education model applying ML and Big data

The development of big data analytics has explored countless possibilities to hold educational attainment that creates the best opportunities for the people to access qualityeducation regardless of the time, and place of the student. Indeed, it incorporates the access alternatives in finding the gap between the teaching and learning procedure. The application of ML and Big data analysis has set the model inthe education sector concerning the situation of pandemic. Online education displays the modality of distance education developing the digital education, and now it has come to be known as the virtual classroom to the students accessing through the internet connection and using various technological tools in the learning, and teaching processes [3].

It is helpful to create the acronym us learning model that includes the interaction time. Online education has emerged from the unexpected social distancing. Though it was prevalent before this time, the huge necessity is assumed in this period. Before implementing the ML, and big data analysis in the education sector the core characteristics of online education must be known. Interactive model through application of ML creates best integration between the teachers, and fellow students including the required content into it. The best feature of ML application in the online education model is that it creates the education network regardless, the time and place of the education, and this is the key focus of the implementation here. Asynchronous, and synchronous, it allows numbers of students to participate in a particular activity at the same time. Online education with significant implementation of ML accesses the available resources without the need of physical appearance.

C. Machine Learning adaptability in Online Education

ML is the simulation, and maintenance of human intelligence through the machine. Generally, the bestdiscipline tries to set a capable system of reasoning, and learning power as a common human being. This systemfocuses on the learning process from the experience of having the capabilities to solve the problems within certain conditions. Contrasting the information it carries out the logical tasks. Typically, Machine learning is capable of analyzing huge volumes of data through pattern, and trend identification, therefore it helps to formulate the predictions automatically, accurately, and quickly.

ML makes the learning procedure smarter by integrating predictive analysis, and other techniques within its application depending on the daily use pattern. It helps in language processing, suggest tagging on specific terms, recognizing the images, or algorithms, and also it provides recommendations that can be important for better system

integration. Real time navigation and optimization of the traffic are important parts of machine learning. With various important techniques it is going to be used in the education sector widely. As the acute part of the artificial intelligence machine learning creates the actual function in the computer giving it a special learning ability. Computational area is used to store important knowledge for strategic purposes [9]. Deeper layers denote more significant strategies to adapt the ML for online education. The possibilities of the machine learning implementation within online education is greater, and the demand for it is increasing day-by-day. ML is the combination of systematic computer learning that relies on memorizing, learning, and communicating operations.

III. Prepare Your Paper Before Styling

For development of this article the necessary informationhas been collected from secondary sources. This article is focused on the application of Machine learning in online education in India. Secondary qualitative analysis helps to specify the environment of machine learning, and big data analysis application and integration in a disciplinary way. It is considered helpful to determine the way that it is ideal to integrate the technologies. As another instance the data analysis is important to understand the conjugation of ML, big data analysis in IoT implementation. The participation of the secondary data in this article has emerged as crucial tutorials in finding the relevance and evaluation of ICT focusing ML, AL, and Big data analysis.

IV. Discussion

A. Identification of the IoT environment

In this article two modalities of the study are described- the first is face-to-face that meets with the modalities of the traditional way. In it the learning depends on the experience, and methods applied by the responsible academic person/teacher. Students emerge as the spectator of their own learning, and it is complied with the previously prepared schedule [6]. Additionally, the teacher appears as the entity to help in determining learning content, and learning procedures. The next model is the online education observed over a decade. The evaluation has taken place through the virtual education model integrating through information technology. Technological structure of the machine learning implementation is crucial here. Platform of online education is moulded by technological implementation considering relevant information. Table 1 explores the analysis about the increasing use of machine learning in the near future. Data volume will continue to be increased, and migrated to the cloud. It shows that 60% of the machine learning process, and big data analysis is going to be applied by the year 2025[2]. It is developed based on the world population, and categorizing it in public cloud, consumer, and in enterprise. Therefore individuals have the potential to follow the growth in the near future. The percentage of clouds in different categories indicates the way of data procession through machine learning. Physical growth of own data is signified providing agility, ease of use, and scalability. The future of machine learning can be assumed through it.

TABLE I. Exponential Increase Of Data By Funding (Source Inspired By [5])

Funding in different categories	Amount (Billion US Dollar)
Application of ML	28.5
ML platforms	14.4
Smart robots	7.5
Platforms of computer vision	7.4
Language processing	6.5
Computer vision (CV) application	4.4
Recommendation engines	2.8
Virtual assistants	2.4
Speech recognition	1.1
Gesture control	0.7

Not only presently, ML, and AI applications are associated with the open-structure program, and it provides affordable solutions. From the amount of funding in different categories, the potential increase can be predicted easily. Following these mentioned categories, the inclusion of Machine learning in the education sector is trying to find the ability with significant advice of the experts [4].

TABLE II. Percentage Of Skills Of Different Iot Techniques

Different Skills	Percentage of Skills
Application of ML	28.5
ML platforms	14.4
Smart robots	7.5
Platforms of computer vision	7.4
Language processing	6.5
Computer vision (CV) application	4.4
Recommendation engines	2.8
Virtual assistants	2.4
Speech recognition	1.1
Gesture control	0.7

Table 2 indicates the high demand of CDOs, and data scientists, though the market is already high concerning the crisis in pandemic time. Skills of big data analytics are 44% that will be helpful for the skill gap in the education sector. Technological structure will be helpful carrying the mobility of machine learning. It will be helpful to find the best solution regarding ML application.

B. ML implementation phases

Before including the technological solution, it is crucial to indicate the objective that would be solved through machine learning. It includes diverse goals such as, reduction of churn, improvement of conversation, enhancement of the consumer satisfaction. Elements of optimization should be clarified in order to fulfil the original goal. Figure 2 depicts different phases of machine learning, and its interaction method with each other.

Figure 2: Different phases of ML implementation [3]

C. *IoT based data aggregation*

Data aggregation and integration with the existing objects of the environment get the amity by IoT. With technological advancement use of electronic devices such as cell phones, or similar apparatuses associated with IoT [10]. In the education sector the data can be aggregated through IoT to find the best response. It is primarily based on the data analysis process identifying different patterns, and it helps to modify the learning teaching behaviour through application of IoT in a disciplined manner [11]. Knowledge engineering of ML depends on the necessary use to expertise the system. Scalability, error tolerance, transforming network regulation reduces the redundancy in wireless sensor networks.

Through wireless networking the efficiencies of the data integration can be improved consisting of infinite smart sensors.

D. *Automated Exploratory Data Analysis in ML*

Exploratory data analysis is the process of data exploration to understand different aspects of information. It is a crucial step before performing machine learning tasks. In this case, it helps to identify clear mistakes and it distinguishes exceptions in datasets. Along with that, it can uncover significant components and discover designs inside information. As per the views of Zhai and Chen (2018), exploratory data analysis helps to improve understanding of variables by extracting averages, minimum, mean, and maximum values, and others. In the machine learning process, EDP has a major contribution to provide real-time data and analyze appropriate information. In this case, EDATK (Automated EDA Toolkit) can increase the value of EDA in ML process.

There are some features of EDATK and those are it can reduce the issues of usage process. As per the words of Drosos *et al.* (2020), running automated exploratory data analysis over a pandas data frame is just one line of code. Along with that, EDATK, through giving a folder location, can build an HTML report that represents tables and other visuals in a clean manner. In addition to that, through target column parameter, it specifies to include visual layers in a supervised Machine Learning Setup.

V. Conclusion And Future Direction

This article provides significant data about machine learning, and big data emphasizing deep learning. Discussion exemplifies a broad set of ML and Big data analysis in virtual, online, or hybrid models of education. This article allows ML in becoming ideal assistance for both-the teachers, and the students. Discussion of this study will help to develop the learning model, improving it in a desirable way. The future of data aggregation and wireless networking in IoT will be increased concerning its growing usage in different fields such as financial sector, educational sector, and health sector. Therefore significant responsibilities of the machine learning implementation can be found through responsible activities here. The research would be able to give better understanding to the readers, and those related participants of the research work.

References

[1] Staegemann, D., Volk, M., Daase, C. and Turowski, K. Discussing Relations Between Dynamic Business Environments and Big Data Analytics. *Complex Systems Informatics and Modeling Quarterly*, (23), pp. 58-82, 2020.

[2] A. Jain, A. K. Pandey. "Modeling and Optimizing of Different Quality Characteristics in Electrical Discharge Drilling Of Titanium Alloy (Grade-5) Sheet" *Material Today Proceedings*, 18, 182-191, 2019.

[3] Villegas-Ch, W., Román-Cañizares, M. and Palacios-Pacheco, X. Improvement of an online education model with the integration of machine learning and data analysis in an LMS. *Applied Sciences*, 10(15), p.5371, 2020.

[4] Wamba, S.F., Dubey, R., Gunasekaran, A. and Akter, S. The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism. *International Journal of Production Economics*, 222, p.107498, 2020.

[5] A. Jain, A. K. Pandey. "Multiple Quality Optimizations in Electrical Discharge Drilling of Mild Steel Sheet" *Material Today Proceedings*, 8, 7252-7261, 2019.

[6] Adi, E., Anwar, A., Baig, Z. and Zeadally, S. Machine learning and data analytics for the IoT. *Neural Computing and Applications*, 32(20), pp. 16205-16233, 2020.

[7] Dubey, R., Gunasekaran, A., Childe, S.J., Bryde, D.J., Giannakis, M., Foropon, C., Roubaud, D. and Hazen, B.T. Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics*, 226, p.107599, 2020.

[8] A. Jain, A.K.Yadava& Y. Shrivastava. "Modelling and optimization of different quality characteristics in electric discharge drilling of titanium alloy sheet" *Material Today Proceedings*, 21, 1680-1684, 2019. https://doi.org/10.1016/j.matpr.2021.03.642

[9] Kozjek, D., Vrabič, R., Rihtaršič, B., Lavrač, N. and Butala, P. Advancing manufacturing systems with big-data analytics: A conceptual framework. *International Journal of Computer Integrated Manufacturing*, 33(2), pp. 169-188, 2020.

[10] López Belmonte, J., Segura-Robles, A., Moreno-Guerrero, A.J. and Parra-González, M.E. Machine learning and big data in the impact literature.*A bibliometric review with scientific mapping in Web of science*. Symmetry, 12(4), p.495, 2020.

[11] V.Panwar, D.K. Sharma, K.V.P.Kumar, A. Jain & C. Thakar. "Experimental investigations and optimization of surface roughness in turning of EN 36 alloy steel using response surface methodology and genetic algorithm" *Materials Today: Proceedings*, 2021.

[12] Zhai, B. and Chen, J. Development of a stacked ensemble model for forecasting and analyzing daily average PM2. 5 concentrations in Beijing, China. *Science of The Total Environment*, *635*, pp.644-658, 2018.

[13] Drosos, I., Barik, T., Guo, P.J., DeLine, R. and Gulwani, S. April. Wrex: A unified programming-by-example interaction for synthesizing readable code for data scientists. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-12), 2020.