Analysis on IoT and Machine learning fusion

Rohini B Khanderay[1], Swati Khokale [2], Radha Sali[3]

[1]Rohini B Khanderay, (Computer Science and engineering, Sandip University, Nashik, India
[2]Swati Khokale (IT department, Sandip Institute of technology and research, Nashik, India
[3]Radha Sali (Computer Science and engineering, Sandip University, Nashik, India

---***---

Abstract: *In the present era of rapidly growing mature technologies and their inter-connection with hardware devices and software applications plays an important role for the emergence of different sensor devices that are inter-connected through internet in order to determine interaction with the physical objects within the world. The technology with the interconnection of various devices through internet is termed as internet of things (IOT). IOT generates massive amount of knowledge with reference to various characteristics and qualities of knowledge. Machine learning fusion with IOT ensures the pervasive development to increase the intelligence of the IOT devices and applications. The exposure of different smart IOT applications with machine learning helps in observation, systematic analysis, processing and smart usages of the large volume of knowledge in several fields. Many industries are using the machine learning and more specifically, the Machine learning-as –a-Service (MLaaS) to take advantage of the IOT's potential. This present paper consists of machine learning basic introduction, machine learning algorithms, reviews of various researcher's study, various sensor devices and therefore the various applications of machine learning algorithms with IOT. And the last section of the paper consists of dialogue and conclusion.*

*Keyword*s — **Internet of things(IoT), MLaaS, Machine learning**

I. INTRODUCTION

From the last decades, as world is witness for real time flow of different sorts of structured and unstructured data received from social media, transportation, communication, devices and sensors. As per information, International Data Corporation forecasts that nearby 180 Zeta Byes of knowledge are going to be generated up to 2025. This huge explosion of knowledge has given rise to a new economy to the planet that's data economy. In today's environment of digital world, data are often considered because the new oil which is pretty important but having worth when it's preceded by cleaned and processed stage. This incremental growth in data economy has given the hike within the usage of concept of IOT or Internet of Things and also incites the existence of knowledge science. IOT may be a hybrid combination between the physical world objects and embedded devices that are connected through the internet to determine the communication. As per day by day involvement of IOT in lifestyle also as in business and industry, the more and more amount of knowledge is generated with the help of sensors. instead of the entire generated newly data it's better to extract the smart data from the IOT data that can be as its best representative. Data Science is additionally continuously moving ahead in direction to achieve a replacement paradigm that's AI which makes possible to show the machines by learning through the data and propel towards the range of its effective insights. Generally, the large rise in amount of knowledge to exchange makes it impossible to research using traditional techniques. So, machine learning intense towards the computation and wishes an enormous amount of knowledge for training that has repetitive training so as to refine the power of learning also as decision-taking regarding the applied algorithms. Machine learning helps the IOT devices and applications in making the more intelligent decisions with reference to the smart data.

II. MACHINE LEARNING

The discipline of Machine Learning is as a subset of Artificial Intelligence which cares with the potential to the computer systems or machines so as to enhance the performance automatically throughout its experience. Along with the addition of excessive data, Machine Learning refines the automotive learning process through training and lead towards adaptation of its algorithm. ML algorithms are wont to implement the various models. Basically machine learning are often categorized as: Guided, Unguided, Semi-guided and reinforcement learning. A. Guided Learning: When learning is applied on labelled data or the specified outcomes is understood called as guided learning. for instance, Amazon's Recommendation System, voice assistant, Weather Apps, Gmail Spam Filtration etc. It also helps in prediction of future results for unseen data. B. Unguided Learning: If the training is applied on unlabelled datasets or the info isn't known beforehand then this sort of learning is understood as unguided learning. for instance , NASA uses this learning approach to create the various clusters of heavenly bodies each of which consists of comparable nature objects.

C. Semi-guided or Semi-Supervised learning approach may be a hybrid approach as a mixture of guided and unguided learning with few labelled and unlabelled data. For example, automatically detection of face book photos for the multiple photos of a same person from a family function photo cluster (clustering- unguided learning) by

naming that person once (guided learning) and afterwards it'll automatically attached the tag thereto person in all photos.

D. Reinforcement Learning may be a machine learning which allows the efficient observation of surroundings and consistent learning behaviour to a learning system so as to enhance the frequency of cumulative incentives or rewards. it's also referred to as reward-based learning system. for instance, Robot during a manufacturing unit.

III. RECENT IOT DEVICES AND TECHNOLOGY

Arduino Device: Arduino device is an embedded microcontroller kit used for implementing the digital devices which will sense the objects of the physical world and do the control over them to supply the accessibility and establish communication between those physical objects. Currently, Arduino is one among the popular IOT board utilized in market that's manufactured with sets of digital and analog I/O Pins that can even be interfaced with other circuits and should include USB Port Connection to attach with PC's. 1. Intel Galileo: The 2nd generation board of Intel Galileo contains Intel Quark Super computer processor, 256 MB RAM and different parts so as to provide support to Arudino Devices. 2. Raspberry – Pi: hottest IOT board as Pi 3 that provides the more compact and standalone computer systems by involving the Bluetooth, built -in wifi facility and also ready to include different programming packages like python, LAMP Stack etc. 3. Samsung Gear Fit: it's a fitness tracker feature that is present in market as Samsung IOT product. It has curved display, dust also as water proof and longlasting battery.

IV. IOT PLATFORM

IOT Platform is an integrated service which is employed to achieve the slot or as an interface between IOT connected devices (Sensors) and IOT applications (Network) and provides offer to bring world's physical objects on-line. Various IOT Platforms which facilitates to deploy IOT
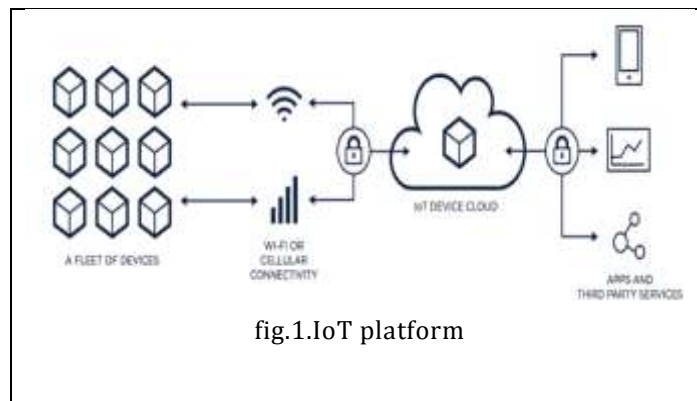
fig.1.IoT platform

Applications are as:

• AWS IOT Platform: Amazon web Services IOT Platform performs the info collection from the connected devices and apply it to the important world scenario by maintaining the security criteria.
• Microsoft Azure IOT Platform: Azure IOT Platform provides the bidirectional communication between connected devices and platforms with strong security mechanism, scalability and straightforward integration with systems. It also included azure streaming analytics which may process the large information that's generated by sensors in world .
• Google Cloud Platform: This platform provides some additional features like cloud IOT Core, Speed-Up IOT Device, Cloud Publisher and Subscriber and Cloud Machine Learning Engine. IBM Watson IOT Platform and Artik Cloud IOT Platform (SAMSUNG).

V. LITERATURE STUDY

IOT is emerging because the challenging area of data and Technology (IT) and digital world that enlighten the glimpse of further research opportunities within the environment of internet connected devices and applications. Different researchers has been already illustrated their research during this field with some Machine Learning algorithms so as to form proper utilization of giant amount of generated data and to supply some specific point of views.

Lei Zhang and et al. (2012) introduced an approach based on an application for tracking the situation i.e. Sens Track which is used with smart phones embedded with Wi-Fi facility in order to scale back the usage of GPS thanks to its availability at high cost with negative impacts on battery in very short period. Sens Track operated the GPS sample by using stored information and may switch the situation and to re-build the track route from recorded location Gaussian Process Regression approach is followed. Mohammed S. Alam and Son T. Vuong (2013) has been used random forest algorithm of supervised machine learning technique to classify the Android features and extract the application anomaly as malicious or benign from the given dataset so as to take care of the accuracy during classification. This paper has main focused approach on testing the random forest parameters for the given dataset. During experiment the experimental release of Weka used over the stable release 3.6.9 because newly released version of WEKA provided the multithreading support so as to attenuate the time-complexity level. It also has limitation that WEKA doesn't implement the computation regarding Feature importance just in case of random forest classification thanks to which that's not computed within the paper. Hyo-Sik Ham and et al.(2014) has introduced an technique for the detection of malware in an android platform by using Linear-SVM (Support Vector Machine) on dataset collected by monitoring the resources on android platform to detect the various malwares that weren't handled in previous studies and also perform comparative analysis with different labeled classifiers. This paper provides the remedial effect to overcome the disadvantage of behaviour- based approach for detecting the malwares.

Divyansh Khanna and et al. (2015) has been presented a comparative study for (multi) classification based methods such that Linear-Support Vector Machine, Logistic Regression and Neural Network for accurate prediction of heart diseases where easier techniques SVM with linear kernel and logistic regression provided more accuracy for 5-folds 87.6% and 86.8% respectively. They used the publicly available Cleveland Heart Disease dataset to use the classification techniques.

Tri Doan and Jugal Kalita (2015) proposed an approach to allot the ranking to the various machine learning algorithms in ranking list with by default setup of parameters by using linear Regression. This list of algorithms ranking helped the Auto-WEKA for selection of feasible algorithms for further execution. But with a touch drawback that a minimum of four actual datasets features are required that's overcome in their future work in the sector of tongue Processing and extending the work for big-data challenge. Vishwajeet Hari Bhide and dr Sanjeev Wagh (2015) proposed a completely unique technique so as to supply the higher assimilation of surrounding situations in home with lesser interference of human and also ready to automatically detect the fault or problem in any device. albeit, addition of Naïve Bayes algorithm increased the extent of intelligence. Further use case of this technique can also ready to manage and supply lesser cost factor, efficient amount of energy and adaptability to the smart homes. This paper also included the longer term scope of their work to incorporate the sms alerts, email notifications or voice alerts with the extended feature to reinforce the safety receptionentrance door by motion detection. Alberto M. C. Soza and Jose` R. Amazonas (2015) has been implemented as Principal Component Analysis (PCA) based clustering algorithm for fault detection that used Hadoop Framework and Mahout implementation. This algorithm integrated with IOT architecture implemented by the LinkSmart middleware. Proposed implementation and architecture increased the potential and functionality of IOT LinkSmart middleware. Arijit Ukil and et al. (2016) depicted a ubiquitous health monitoring system based upon (unguided learning) anomaly detection of diseases at early stage so as to supply the essence of life. The accuracy of anomaly detection varied with the SNR rate (more accurate with low SNR). This paper also described a use-case of anomaly detection for cardiac with the help of affordable android smart phones with the attitude of early detection of heart related diseases that avoid leading the severe condition. Hardi Desai and et al. (2017) has proposed a vision to implement a reasonable and compatible IOT based wireless sensor network so as to monitoring and analysing the grocery levels at supermarkets also as at homes. This technique also provides an immense to use as future scope within the kitchens and to watch the various storage places to manage the commodities in smooth manner. Shuai Zhao and et al. (2017) has been proposed a multidimensional model-based approach so as to gauge the similarities between different IOT services. On the idea of clustering algorithm for dynamically discover, match making and exchange are going to be performed in efficient manner. Evaluation of experiments on proposed system and results are promising and supported real world. it's drawback that within the present paper conducted the sensible on single dataset with limited size of test data set.

Guangie Han and et al. (2018) focused on critical issue regarding the safety in WSNs to take care of location privacy and introduced the novel protection scheme based upon cluster-based location privacy (KCLP). This particular scheme enhanced the safety and minimizes the delay in expense of energy consumption at minor level which will be reduced further in future work. K. Gai, M. Qiu (2018) proposed a completely unique technique for optimal allocation of resources in cybernetics physical environment (IOT) based upon the reinforcement learning paradigm. Reinforcement method is employed to get the high satisfactory level that's QOE (Quality- of Experience), more accuracy and lesser cost in resource allocation. Reinforcement learning considered QOE level as reward

parameter and therefore the resulted value as optimal solution. Srijanee Mookherji and Suresh Sankaranarayanan (2019) has been introduced the IOT Automated Traffic Signaling System so as to dynamically regulate the traffic in congested areas. For data security analysis purpose addressed to Man- In – Middle Attack (MITM) and Support Vector Machine (SVM) supported supervised (guided) learning method is used to deploy at edge to classify the traffic data as raw dataset. This system is implemented on IOT Board Raspberry- Pi3 using Scikit. Jin Ren Ng and et al. (2019) has introduced a motive in perspective to gather the info and with reference to which analyzed conditions of road surfaces in Malaysia using IOT Sensor devices and used supervised (guided) classification algorithm (KNN, Random Forest or SVM). It basically provided the value effective approach so as to watch the roads condition which successively the faster repairing of roads. It helped to supply the higher roads that reduce the probabilities of fatal road accidents thanks to poor maintenance of roads. Wenjuan Li and et al. (2019) has introduced a combination of design for multi-view based classification of emails for IOT system and therefore the disagreement based semisupervised learning paradigm. so as to mitigate this major issue of suspicious emails and security a standard solution is proposed by taking the 2 datasets as internal feature dataset and external feature dataset. Basically semi supervised learning paradigm approached both labeled and unlabeled data. The availability of multi-view dataset for email classification provided the more accuracy compared to single view email classification which will be improved further with the assistance of connecting with different IOT systems.

VI. DISCUSSION AND CONCLUSION

Machine Learning has an explosive potential to be the key technology for the pervasive environment of IOT. The recent radiations of success of Machine Learning along side IOT prove their fusion because the wonderful source within the field of knowledge science. This paper discussed the various applications of machine learning algorithms for IOT and covers the recent advanced technologies utilized in IOT environment. This paper also emphasized the thorough study on recent researches regarding the applications of machine learning algorithms for IOT. during this cutting-edge of artifacts based intelligence of upcoming technologies, the new inventions are reaching at the pinnacle with the attitude of building inter-connections in between the physical world objects that gives advanced, smart and reliable life to living being.

REFERENCES

[1] https://www.seagate.com/in/en/our-story/data-age-2025/

[2] Laizhong Cui, Shu Yang and et al. "A survey on application of machine learning for Internet of Things". International Journal of Machine
Learning and Cybernetics, August 2018, Vol. 9, Issue 8, pp 1399–1417.

[3] Luigi at zori, Antonio Iera, Giacomo Morabito. "The Internet of Things: A survey". Computer Networks, 2010. Vol. 54, Issue 15, pp 2787-2805, ISSN 1389-1286.

[4] Chen, Feng and et al. "Data Mining for the Internet of Things: Literature Review and Challenges." International Journal of Distributed Sensor Networks, August 2015.

[5] 13] Ukil, S. Bandyoapdhyay, C. Puri and A. Pal, "IoT Healthcare Analytics:The Importance of Anomaly Detect ion". IEEE 30th International Conference on Advanced Information Networking and Applications(AINA), Crans-Montana, 2016. pp 994-997.

[6] Desai, Hardi, et al. "IoT based grocery monitoring syst em". Fourteenth International Conference on Wireless and Optical Communications Networks (WOCN). IEEE, 2017.

[7] Ng, Jin Ren, et al. "Identification of Road Surface Conditions using IoTSensors and Machine Learning." Computational Science andTechnology, 2019. pp 259-268.

[8] Li, Wenjuan, et al. "Design of multi-view based email classificat ion for IoT systems via semi-supervised learning." Journal of Network and Computer Applications, 2019. Vol. 12, Issue 8, pp 56-63.

[9] https://iot.intersog.com/blog/iot-platforms-overview-arduino-raspberrypi-intel-galileo-and-others/

[10] https://www.i-scoop.eu/internet-of-things-guide/iot technology-stackdevices-gateways-plat forms