<Conference abbreviation>
<Conference Series name> <Volume number and Year> <DOI Number>

Integration of Deep Learning and Edge-Cloud for Big Data Analytics in IOT

[1]Kushwant Kaur, [2]Kirat Kaur, [3]Nishu Bansal

[1,2,3]*Department of Computer Science and Engineering, Chandigarh University, Mohali*
[1]*khushi.rai1@gmail.com*, [2]*kirat.e12999@cumail.in*, [3]*nishubl@gmail.com*

Abstract.

The Internet of Things (IOT) produces large data voumes from wearable devices and wireless sensors. The data may be unstructured, disorganized, and unregulated. The massive amount of data gathered is useless unless it is processed and used to acquire insights in manufacturing, smart cities, healthcare, and agriculture. Big IOT Data must be kept in the cloud for its proper utilization, increasing network traffic and delays. Edge computing, which is growing, can solve traffic problems by placing computing jobs to local network edge nodes or IOT devices. Edge computing cannot compute complicated computations on huge unstructured data. Deep Learning handles difficult datasets. This article suggests integrating Deep Learning and Edge Computing. It describes how this integration may reduce latencies and maximize big data use to boost efficiency and performance.

Keywords. IOT, Big Data Analytics, Deep Learning, Edge Computing, Machine Learning, Cloud Computing.

1. Introduction

The exponential rise in the count of sensor-based devices, such as cell phones, wristbands, medical devices, and other smart devices, has led to an explosion in data growth in recent years[1]. There are more devices connected via internet than the total population of the world and which is expected to hike more in future. The rapid convergence of wireless connectivity, embedded systems, and micro-electro-mechanical systems has marked the beginning of IOT and data explosion[2].The Internet of Things comprises of gadgets with limited resources, such as radio frequency identification (RFID), sensors that are powered by batteries[3]. IOT technological innovations lead to ingenious applications such as, smart cities, health care, traffic management, agricultural advancements, e-commerce and supply chain management etc. This may result in the large collection of data in an unfavorable way [4]. Such IOT datasets should be analyzed for enhancing smart decision making and future predictions[5].

In IOT Analytics, Cloud Computing is a traditional approach that transports datasets to the cloud for processing and consistent storing. After processing data, the results are delivered back to IOT devices or software applications[6],[7]. Transferring data to cloud(miles away from the IOT devices) for processing leads to problems such as, increased data transfer latencies and network congestion. On the other hand, analyzing the data at the device itself is also possible; but most of the IOT devices have limited computation facility to perform complex data analytics and limited storage capacity. Edge Computing(EC) is an emerging technology in the applications that requires fast response with limited bandwidth[8]. It minimizes latencies by performing computations on the device itself or on an edge node nearer to the data source. Hence, limited data-volumes are moved to cloud for further processing. Also, it is challenging to handle huge amount of heterogeneous data that is either semi-structured or un-structured in nature. Deep Learning, an emerging technique of machine learning, has recently shown its potential in handling large data quantities collected from IOT devices. It aids learning by converting unstructured input into hierarchical abstract representations that allow for the acquisition of desirable characteristics[10].

This paper proposes an integrated approach of Deep Learning, Edge and Cloud Computing for analyzing the IOT data. Section 2 presents the significance of data analysis, framework, and how the big data and sensors data complement each other. Section 3 discusses the various Learning models and architectures that are useful. Additionally, it emphasizes the importance of DL in analytical models. Section 4 delineates the commonalities and major distinctions between edge and fog computing. Section 5 proposed Edge-Cloud and Deep Learning based integration to boost efficiency of data analytics. Section 6 ends the report and offers guidance for future research endeavors.

2. IOT BASED ANALYTICS

Handling vast amounts of dispersed data is crucial to the increasing rise of IOT data analytics in smart decision making [11]. Web-connected device data is complicated and unstructured, requiring lightning-fast analytics using modern technology. With IOT advancements, we have tens of billions of linked devices[7]. Tons of data from linked devices are transmitted across several channels to the database for consistent storage before processing. Big data is processed with data mining and ML methodologies to maximize profits.

2.1. *Framework of IOT Data Analytics*

Data Analytics consists of discovering new patterns and knowledge by exploring and analyzing large volumes of IOT data for making profitable decisions and achieving business intelligence in diverse application domains. Fig. 2.1 represents the framework in which IOT data analytics is performed.

Firstly, IOT devices produce huge amounts of datasets that are integrated into data warehouse. The data is then analyzed either at cloud or device itself. Data Analytics involves various operations on data, such as preprocessing, integration, transformation and reduction [12], [13]. Once the data is cleaned, pre-processed and reduced, it is analyzed by using machine learning algorithms for knowledge discovery. The final output is meaningful information that helps in smart decision-making various sectors.

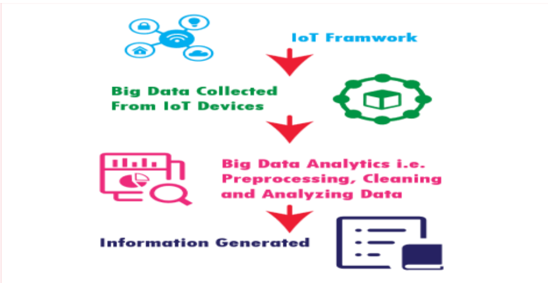

Figure 2.1. Framework of IOT Data Analytics

2.2. Big Data and IOT are distinct but Complementary

IOT and Big Data are distinct but complementary terms. Both refer to collecting substantial quantities of data that is unstructured and complex in nature. Big Data and IOT are interrelated terms in which Big Data is ocean of information and IOT devices are series of streams and rivers feeding into these oceans. IOT working together with Big Data analytics can show tremendous results in various fields. Table 2.1 represents the comparison of Internet Of Things and Big Data based on common goals and distinct features. The factors for comparison has been chosen on the basis of studies[1,3,5,8].

TABLE 2.1. COMPARISON OF IOT AND BIG DATA

Factors for Comparison	IOT	BIG DATA
Support for Intelligent Decision Making	Provides interconnection of devices to create smart environment for decision making	Performs large collection of Data
Nature of Work	Finds insight into data collected from IOT devices for decision making	Refers to collection of huge amount of Data
Application Areas	Smart Cities, Industries, Healthcare, Transportation, Sales &Marketing and Agriculture etc.	Smart Cities, Industries, Healthcare, Transportation, Sales &Marketing and Agriculture etc
Time of Analysis	IOT focus on immediate analysis of data on the device itself or by cloud services	Big Data tools focus on collection of data and data analysis is performed afterwards

Big Data and the Internet of Things (IoT) are distinct but complimentary notions. These have a vital interaction that will strengthen as technology improves. Companies seeking maximum advantages from IOT devices must carefully select IOT data analysis tools and methods.

3. DEEP LEARNING IN IOT DATA ANALYTICS

Deep learning is a cutting-edge technique implements multiple layers by cascading them, with each successive layer experiencing a non-linear transition[14]. By applying feature extraction, it transforms data into abstract and meaningful hierarchical representations[6]. Deep Learning is invaluable analytical tool for dealing with unsupervised, unstructured and semi-structured data [10].

3.1. *Deep Learning Architectures*

DL architectures are broadly classified into two types as follows: 1) Discriminative Model, that supports supervised learning and 2) Generative Model, supports unsupervised learning[15]. The most popular architectures are as follows:

- **Artificial Neural Networks (ANN):** It is a three-tiered design with several sets of perceptrons/neurons at each tier. Each layers learns some weight. It is useful for Content knowledge, tabular data, and digital images.
- **Convolutional Neural Networks (CNN):** The CNN employs a back propagation technique to determine the receptive fields of basic units. These are common in image processing activities and perform dimensionality reduction when the hidden layer's number of neuron connections is limited to only a few input neurons[16].
- **Autoencoder (AE):** These are ANN based unsupervised models that learns efficient data coding. Its objective is to discover a representation for a set of data, typically to reduce its dimensionality. The AE consists of encoder and decoder. The encoder decodes the input's representative characteristics and converts them to other latent features, while the decoder receives the encoder's latent features and attempts to reconstruct the input data in its original form while minimizing the reconstruction error.[6,8,17].
- **Deep Neural Networks(DNN):** It is one of the most promising strategy for Time Series Classification (TSC) in big data mining[18]. A DNN is a type of stacked neural networks, i.e., networks composed of multiple layers between input and output layers[19].
- **Recurrent Neural Networks(RNN):** RNN is a network with looping system. It is widely used in Time-series, text and Audio Data analysis[18].

3.2. Need of Machine Learning for IOT

IOT devices are computationally restricted. Machine Learning for IOT analytics may enhance IOT device efficiency, accuracy, production, and cost [20]. Combining ML and IOT improves communication, computation, control, and decision-making. Advanced ML methodologies like Deep Learning may mine enormous volumes of unstructured and complicated IOT sensory data to improve real-world problem-solving and decision-making[20].

4. EDGE CLOUD COMPUTING & FOG CLOUD COMPUTING

This Section discusses the role of Edge and Fog Computing for analytics.

4.1. *Edge Computing*

Cloud technology has good scalability, reliability, pay-as-you-go pricing, and inexpensive startup costs. [8]. Applications with low latencies and bandwidth constraints favour Edge Computing (EC). EC enables calculations near data sources. It computes on the cloud-edge or IOT device (with sensors) to decrease data transit before reduction [5,18]. Simpler communications, collaborative computing, enhanced performance, uptime, and reduced latencies are EC advantages. [21]. Although edge computations are beneficial, IOT devices' scalability and interpretability limitations make them incompatible with some cloud services.

4.2. *Fog Computing*

In fog computing, nodes or Gateway nodes process data. It computes near IOT devices like edge computing. The computing infrastructure is decentralized. The Fog Node sits between the cloud and the data source, processing and storing data.

4.3. *Comparison of Edge-Fog Cloud Computing*

Both Fog and edge computing gather, examine, and process data more efficiently than traditional cloud computing. After performing various studies as presented in[9,20,22,23],we have tried to summarize the key features of Fog computing & Edge Computing in Table 4.1.

TABLE 4.1 COMPARISON OF FOG COMPUTING & EDGE COMPUTING

Parameters/ Features	Edge Computing	Fog Computing
Data Processing Location	Edge Node or on IOT devices	Fog Node between Data source and Cloud on LAN itself
Amount of Data saved in cloud	Reduced	Reduced
Network Traffic[6,17]	Reduced	Reduced
Transfer Latency[17,24]	Reduced	Reduced
Support Real-time Data Processing [24]	Moderate	High
Volume of data Processed	Less as compared to Fog	More as compared to Edge

The following table shows that Fog and edge Computing offer similar benefits. Their architecture makes them different in a few ways. Edge computing focuses on intelligence, computational power, and communication on an edge gateway or directly on IOT devices, while fog computing processes data in a fog node or gateway of IOT appliances on the local area network.

5. INTEGRATION OF EDGE-CLOUD AND DEEP LEARNING

Integration consists of combining features of both the technologies and maximizes performance and accuracy of the devices. Because linked devices have limited computational capabilities, data analytics calculation cannot be performed solely on them[8]. This section discusses combining Edge Computing with Deep Learning for overcoming the traffic problems and henceforth, enhancing Quality of IOT Services in various areas.

5.1. Quality Focus in IOT applications

To gain maximum benefit from IOT services in various arenas like smart cities, healthcare, and agriculture; ensuring Quality of services (QOS) is a critical task. Different metrics for analyzing quality of services are available but our study focuses on reducing latencies and network traffic. Both Deep Learning and Edge Cloud Computing have their own pros and cons. If we want to ensure quality and efficiency in IOT analytics then these two techniques can be merged to utilize best features of both. Network performance can be improved with intrusion detection system (IDS) using ANN and DL[25].

5.2. *Edge-Cloud Integrated approach*

By relocating computation to the network's edges and data sources, EC has emerged as a method for addressing latency issues. Fog computing[7] has a lot of similarities EC. Edge Computing is focused on nodes that are nearer to IOT devices[30], whereas fog might comprise of resources between the end device and the cloud. EC is helpful in reducing the network traffic by moving computing near the source of data. The edge can perform high-performance computations in addition to cloud computing, reducing network traffic and latency.

5.3. *Edge-Cloud and Deep Learning*

It has been presented in [15], that with increased volumes of datasets, accuracy of Deep Learning algorithms also improves. DL is a class of ML, consisting of numerous computational layers for abstract data representations[14]. By limiting the quantity of data that has to be transferred to the cloud, Edge-Cloud and DL integration enables data reduction at the edge, which lowers network traffic.For classification purposes, reduced data from all Edge nodes is used. The reduced data can be processed by Deep Learning techniques for extracting features from the data that can be further used for decision making.

6. CONCLUSION AND FUTURE SCOPE

Agriculture and healthcare researchers face dataset shortages. They must create datasets in various situations, which takes time and effort. IoT devices produce significant amounts of data to address this difficulty. Data analytics may help you make wise decisions and find new information. Business intelligence relies on data analytics as IoT grows. Big data, complicated and unstructured, is created by the growth of connected devices. Big data are studied to get information, but sophisticated machines are needed to handle them. Cloud computing facilitates the storage and processing of extensive data sets; nevertheless, transmitting data to the cloud and receiving results impedes communication speed. Communications networks would be strained. This paper used Deep Learning with Edge Computing to analyze data near devices to prevent delay. This integration paradigm might be used in healthcare, agriculture, and other smart devices with various sensors and applications.

7. REFERENCES

[1] M. Ge, H. Bangui, and B. Buhnova, "Big Data for Internet of Things: A Survey," *Futur. Gener. Comput. Syst.*, vol. 87, no. May, pp. 601–614, 2018, doi: 10.1016/j.future.2018.04.053.

[2] E. Ahmed *et al.*, "The role of big data analytics in Internet of Things," *Comput. Networks*, vol. 129, pp. 459–471, Dec. 2017, doi: 10.1016/j.comnet.2017.06.013.

[3] I. Batra *et al.*, "Hybrid logical security framework for privacy preservation in the green internet of things," *Sustain.*, vol. 12, no. 14, 2020, doi: 10.3390/su12145542.

[4] M. A. Amanullah *et al.*, "Deep learning and big data technologies for IoT security," *Computer Communications*, vol. 151. Elsevier B.V., pp. 495–517, Feb. 01, 2020, doi: 10.1016/j.comcom.2020.01.016.

[5] B. Vijayalakshmi *et al.*, "An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainable smart city," *Int. J. Commun. Syst.*, vol. 34, no. 3, pp. 1–14, 2021, doi: 10.1002/dac.4609.

[6] A. M. Ghosh and K. Grolinger, "Deep Learning: Edge-Cloud Data Analytics for IoT," *2019 IEEE Can. Conf. Electr. Comput. Eng. CCECE 2019*, pp. 1–7, 2019, doi: 10.1109/CCECE.2019.8861806.

[7] M. R. Anawar, S. Wang, M. Azam Zia, A. K. Jadoon, U. Akram, and S. Raza, "Fog Computing: An Overview of Big IoT Data Analytics," *Wireless Communications and Mobile Computing*, vol. 2018. Hindawi Limited, 2018, doi: 10.1155/2018/7157192.

[8] A. M. Ghosh and K. Grolinger, "Edge-Cloud Computing for Internet of Things Data Analytics: Embedding Intelligence in the Edge with Deep Learning," *IEEE Trans. Ind. Informatics*, vol. 17, no. 3, pp. 2191–2200, Mar. 2021, doi: 10.1109/TII.2020.3008711.

[9] S. Ben Atitallah, M. Driss, W. Boulila, and H. Ben Ghezala, "Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions," *Computer Science Review*, vol. 38. Elsevier Ireland Ltd, Nov. 01, 2020, doi: 10.1016/j.cosrev.2020.100303.

[10] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, p. 1, 2015, doi: 10.1186/s40537-014-0007-7.

[11] M. Marjani *et al.*, "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges," *IEEE Access*, vol. 5, pp. 5247–5s1, 2017, doi: 10.1109/ACCESS.2017.2689040.

[12] V. Agarwal, "Research on Data Preprocessing and Categorization Technique for Smartphone Review Analysis," *Int. J. Comput. Appl.*, vol. 131, no. 4, pp. 30–36, 2015, doi: 10.5120/ijca2015907309.

[13] D. S. Kumar, "A Comparative Study of Various Data Transformation Techniques in Data Mining," *Int. J. Sci. Eng. Technol.*, vol. 4, no. 3, pp. 146–148, 2015, doi: 10.17950/ijset/v4s3/305.

[14] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[15] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," Dec. 2017, [Online]. Available: http://arxiv.org/abs/1712.04301.

[16] A. P. Singh et al., "A Novel Patient-Centric Architectural Framework for Blockchain-Enabled Healthcare Applications," in IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5779-5789, Aug. 2021, doi: 10.1109/TII.2020.3037889.

[17] F. Wang, M. Zhang, X. Wang, X. Ma, and J. Liu, "Deep Learning for Edge Computing Applications: A State-of-the-Art Survey," *IEEE Access*, vol. 8, pp. 58322–58336, 2020, doi: 10.1109/ACCESS.2020.2982411.

[18] J. Azar, A. Makhoul, R. Couturier, and J. Demerjian, "Robust IoT time series classification with data compression and deep learning," *Neurocomputing*, vol. 398, pp. 222–234, Jul. 2020, doi: 10.1016/j.neucom.2020.02.097.

[19] A. Hussain et al., "A Resource Efficient hybrid Proxy Mobile IPv6 extension for Next Generation IoT Networks," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2021.3058982.

[20] E. Adi, A. Anwar, Z. Baig, and S. Zeadally, "Machine learning and data analytics for the IoT," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 16205–16233, 2020, doi: 10.1007/s00521-020-04874-y.

[21] A. Ahmed and E. Ahmed, "A survey on mobile edge computing," *Proc. 10th Int. Conf. Intell. Syst. Control. ISCO 2016*, no. July, 2016, doi: 10.1109/ISCO.2016.7727082.

[22] S. A. Osia *et al.*, "A Hybrid Deep Learning Architecture for Privacy-Preserving Mobile Analytics," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4505–4518, May 2020, doi: 10.1109/JIOT.2020.2967734.

[23] H. Alshammari, S. A. El–Ghany, and A. Shehab, "Big IoT Healthcare Data Analytics Framework Based on Fog and Cloud Computing," *J. Inf. Process. Syst.*, vol. 16, no. 6, pp. 1238–1249, 2020, doi: 10.3745/JIPS.04.0193.

[24] A. Jabbar, P. Akhtar, and S. Dani, "Real-time big data processing for instantaneous marketing decisions: A problematization approach," *Ind. Mark. Manag.*, vol. 90, pp. 558–569, Oct. 2020, doi: 10.1016/j.indmarman.2019.09.001.

[25] P. Rani, S. Verma, and G. Nhu Nguyen, "Mitigation of Black Hole and Gray Hole Attack Using Swarm Inspired Algorithm With Artificial Neural Network," doi: 10.1109/ACCESS.2020.3004692.