IoT Revolutionized: How Machine Learning is Transforming Data, Applications, and Industries

Taiwo Abdulahi Akintayo [1], Raphael Aduramimo Olusola [2], Ewemade Cornelius Enabulele [3], Ayodele Oyesanya [4], Samuel Ayanwunmi Olanrewaju [5], Moyosore Owoeye Celestina [6], Balogun Oluwaseyi Sulaimon [7], Ogechukwu Ada Lorretta Anoliefo [8], Ayanwunmi Victor Olumide [9], Olamilekan Jamiu Ridwan [10], Adedokun Seyi Adediran [11]

[1] *National Centre of Artificial Intelligence and Robotics*
No 28, Port Harcourt Crescent, Off Gimbiya Street, P. M. B 564, Area 11, Garki, Abuja, Nigeria

[2] *Olusegun Agagu University of Science and Technology*
Km. 6, Okitipupa-Igbokoda Road, P. M. B. 353, Okitipupa, Ondo State, Nigeria

[3] *The Federal University of Technology Akure*
P. M. B. 704, Akure, Ondo State, Nigeria

[4] *Federal University of Technology, Minna*
Gidan kwanu, P.M.B 65, Minna, Niger State, Nigeria

[5] *Yaba College of Technology*
443 Herbert Macaulay Wy, Abule ijesha, Lagos State, Nigeria

[6] *Federal University, Oye-Ekiti*
P. M. B 373, Km 3 Oye – Afao Road, Ekiti State, Nigeria

[7] *Tai Solarin University of Education*
Ijagun Road, Ijebu Ode, Ogun State Nigeria

[8] *East Tennessee State University*
1276 Gilbreath Drive 109 Burgin E. Dossett Hall Johnson City, Tennessee, USA

[9] *University of Ilorin*
Ilorin, Kwara State, Nigera

[10] *Florida State University*
222 South Copeland Street, Suite 424, Tallahassee, Florida, USA

[11] *Ladoke Akintola University of Technology*
Old Oyo/ Ilorin Rd, 210214, Ogbomosho, Oyo-State, Nigeria

DOI: 10.22178/pos.105-30

LCC Subject Category: T58.5-58.64

Received 21.05.2024
Accepted 25.06.2024
Published online 30.06.2024

Corresponding Author:
Taiwo Abdulahi Akintayo
taiwoabdulahi15@gmail.com

Abstract. Integrating machine learning (ML) with the Internet of Things (IoT) reveals hidden patterns and insights from extensive sensor data, enabling IoT to become omnipresent and make intelligent decisions without explicit programming. ML is essential for IoT to meet the future needs of businesses, governments, and individuals. IoT aims to sense its environment and automate decision-making through intelligent methods, emulating human decisions. This paper reviews and categorises existing literature on ML-enabled IoT from three perspectives: data, applications, and industries. We examine advanced methods and applications by reviewing various sources, emphasising how ML and IoT work together to create more innovative environments. We also discuss emerging trends such as the Internet of Behavior, pandemic management, autonomous vehicles, edge and fog computing, and lightweight deep learning. Furthermore, we identify challenges to

IoT in four categories: technological, individual, business, and societal. This paper aims to leverage IoT opportunities and address challenges for a more prosperous and sustainable future.

Keywords: Internet of Things (IoT); Machine Learning (ML); Sensor Data; Intelligent Decision-Making; Data Analysis; Smart Environments; Internet of Behavior.

INTRODUCTION

The Internet of Things (IoT) has transformed how we live and work, generating vast amounts of data from connected devices [1]. Machine Learning (ML) is a critical technology that enables IoT systems to learn from this data and make intelligent decisions without explicit programming [2]. Integrating ML and IoT can potentially revolutionise various industries, including healthcare, manufacturing, and transportation [3]. As IoT devices become increasingly ubiquitous, the amount of sensor data generated continues to grow exponentially [4–7]. ML algorithms can be applied to this data to uncover hidden patterns and insights, enabling IoT systems to become more intelligent and autonomous [8]. However, challenges are associated with integrating ML and IoT, including data quality issues, privacy concerns, and the need for more advanced analytics tools [9].

This paper reviews and categorises existing literature on ML-enabled IoT from three perspectives: data, applications, and industries. We examine advanced methods and applications by reviewing various sources, emphasising how ML and IoT work together to create more innovative environments.

Figure 1 depicts application areas of IoT: smart homes, warning systems, smart shopping, smart gadgets, smart cities, intelligent roads, health care, fire systems, threat-identification systems, tracking, and surveillance.

RESULTS AND DISCUSSION

Data production. The Internet of Things (IoT) provides the technological infrastructure to monitor and sense various activities and events in our surroundings.

As a result, IoT is expected to generate vast data from diverse vendors, leading to the concept of 'data as a service.' To fully harness the potential of smart cities and societies, data sharing and collaboration will be essential for providing sustainable and ubiquitous applications and services. Integrating different data types and formats will enhance data quality and decision-making in ubiquitous environments.

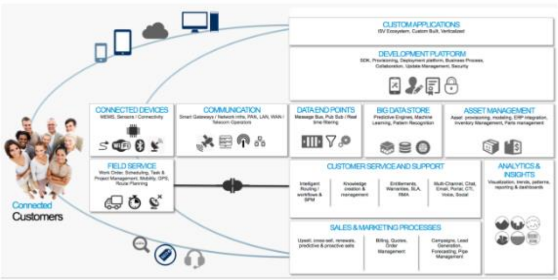

Figure 1 - Application of Internet of Things

This process, known as data fusion, involves combining data from various sensors and sources into a unified format. The timely analysis of big data from IoT sensor networks enables accurate and reliable decision-making. However, managing ubiquitous environments poses a significant challenge for IoT, which can be addressed using advanced sensors and intelligent algorithms.

Machine Learning and Iot. Machine Learning (ML) enables IoT to understand human preferences, predict events, and manage various scenarios effectively. IoT relies on ML algorithms (MLAs) to comprehend the vast amounts of data that millions of interconnected objects produce. ML's integration into IoT is essential for achieving its goal of ubiquity, ensuring accessibility from anywhere. By extracting insights from the extensive data generated by connected devices, ML enhances the usefulness of IoT devices, thereby promoting their true ubiquity. Embedded Intelligence (EI) forms the core of IoT's capabilities, combining product functionality with intelligent features to enhance automation, efficiency, productivity, and connectivity.

Whether in the physical or virtual realm, intelligence is acquired through learning processes. Machine Learning (ML) excels in discovering patterns, potentially laying the groundwork for human-like intelligence. ML enhances our world understanding by generalising these patterns into valuable insights and trends. In IoT, ML's primary objective is to achieve full automation by enhancing learning capabilities, thereby enabling brighter objects. ML empowers IoT systems to emulate human decision-making through data-driven training, continually improving environmental awareness. Information visualisation is crucial in enhancing human data comprehension, leveraging the human visual system's capabilities to extract insights more effectively. This visualisation facilitates deeper understanding without extensive data analysis and harnesses cognitive abilities to interpret information. Ultimately, IoT aims to replace costly, conventional systems with affordable sensor-based ML systems, revolutionising various industries. The impact of IoT extends beyond technological advancements to fostering progress and prosperity in our world. IoT simplifies and connects multiple aspects of our lives daily, supported by ML, which imbues intelligence and pervasiveness into IoT systems. IoT's heterogeneous nature encompasses diverse devices, network technologies, data types, and applications, presenting unique challenges for ML.

These challenges include:

1) Managing the vast amounts of big data generated by IoT, addressing issues of data validity biases, and determining the worthiness of processing all data.

2) Adapting ML algorithms to efficiently learn from both abundant and sparse data sources in different IoT applications;

3) Dealing with the inherent inaccuracies and unreliability of sensing devices through outlier detection and data imputation.

4) Tailoring ML approaches to suit the specific data properties inherent to different IoT application domains (Figure 1).

MLAs deployed in IoT perform various tasks, including data quality improvement and pattern recognition (Figure 2). They predict outcomes from extensive datasets and enhance data quality by identifying outliers and imputing data before training, improving overall learning effectiveness.

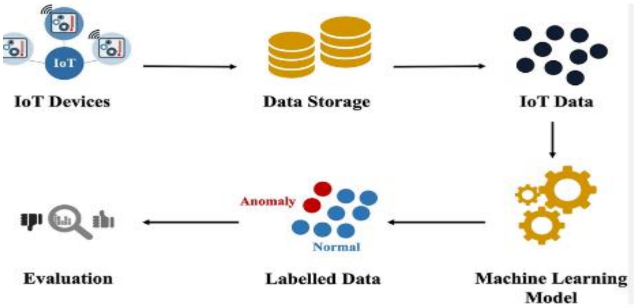

Figure 2 – Machine Learning Data and Pattern Recognition

Data perspective. The data perspective in IoT is a multifaceted concept that encompasses the collection of data from various sources, including sensors, cameras, and smart devices, handling different data types such as temperature, vibration, and electrical metrics, storing data in cloud-based platforms like AWS, IBM, Oracle, and Microsoft, addressing data issues like outliers, missing values, and feature selection, and enhancing data quality and pattern recognition using Machine Learning Algorithms (MLAs) to extract valuable insights and support informed decision-making.

Data add value to the IoT paradigm and are collected using various sensors; IoT has both cheap and expensive sensors in its arsenal. For example, the temperature detection sensor is more affordable than lidar, which is too costly. The type of sensor used largely depends on the type of application of that data. Wild animal tracking sensors will have lifelong battery life, as replacing batteries in wild animal tracking applications is hard. In contrast, sensors like lidar, cameras, and radar continuously need a power supply. Also, low-cost sensor data have such issues as outliers and missing values as their hardware quality could be improved.

On the other hand, vision sensors bring many features, and selecting only the best feature is challenging. In the proceeding subsections, we discuss IoT data sources, data storage platforms, and the three types of data challenges, specifically focusing on machine learning. Data Sources One of the major applications of IoT is sensing our surroundings and communicating that data to intelligent applications, which will be used to predict and forecast using machine learning algorithms. Further, the learning outcome is used to develop AI for making decisions. Later, the decision is transformed into mechanical output using actuators. Today, billions of devices with sensors surround our daily lives. IoT produces and will produce an enormous amount of data that needs to be stored, processed, and archived for future needs. IoT infrastructures still need to be implemented thoroughly, even in developed economies. Developing economies like India, Malaysia, etc., are slowly working on mega-smart-city projects using IoT infrastructure. As depicted in Figure 4, three significant issues with sensed data, outliers, missing values, and feature selection are discussed in the proceeding sections.

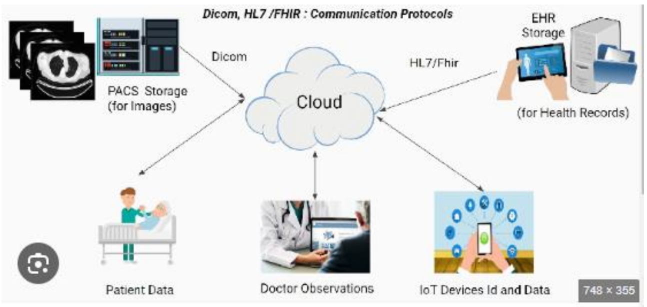

Figure 4 – Internet of Things Data

Outlier detection. Outlier detection is crucial in IoT data preprocessing to identify and remove anomalous data points affecting machine learning model accuracy. Outliers are defined as data points significantly different from the rest, and there are two types: point and contextual outliers. Various detection methods include statistical, density-based, and machine-learning techniques, such as z-score, DBSCAN, LOF, and One-Class SVM. However, IoT data is often noisy, making distinguishing outliers from legitimate data challenging. These algorithms can be applied in various IoT applications, including anomaly detection in sensor data, fraud detection in IoT networks, and predictive maintenance in industrial IoT. A detailed illustration is given in Figure 5.

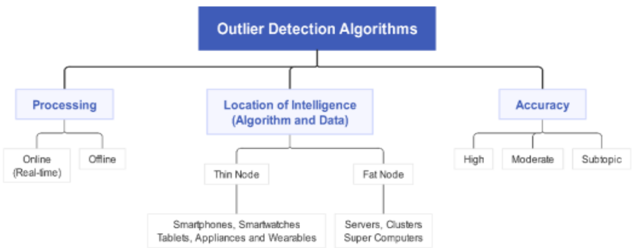

Figure 5 – Outlier Detection Algorith

Data imputation. Data imputation is a crucial step in handling missing data in IoT datasets, involving replacing missing or corrupted data with estimated values. There are three types of data imputation: mean/median imputation, regression imputation, and multiple imputation. Various methods are used, including KNN imputation, linear regression imputation, decision tree-based imputation, and MVE-GAN. Evaluation metrics such as MAE, MSE, and RMSPE are used to assess the performance of data imputation methods. Data imputation is essential in IoT applications to ensure data quality and accuracy and to prevent biased or inaccurate analysis results. Effective data imputation methods can help improve decision-making, predictive maintenance, and overall IoT system reliability [10].

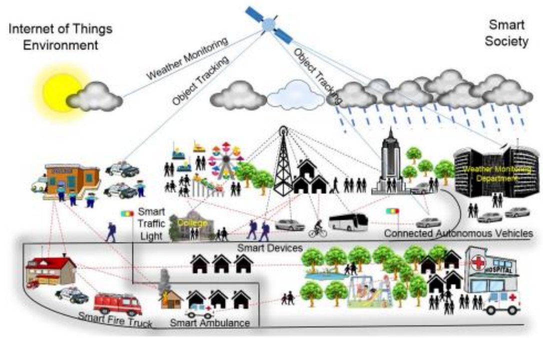

Figure 6 – Overview of the Application of IoT

Application perspective. The application perspective of IoT is vast and diverse, transforming various industries and aspects of our lives. Here are some examples:

1) Industrial Automation: IoT enables real-time monitoring, predictive maintenance, and optimised production processes [11].

2) Smart Homes: IoT devices control temperature, lighting, security, and entertainment systems, enhancing comfort and energy efficiency [12].

3) Healthcare: IoT wearables and sensors monitor vital signs, track fitness, and detect health anomalies, improving patient care and outcomes [4].

4) Transportation: IoT sensors and cameras optimise traffic flow, predict maintenance needs, and enhance vehicle safety [5].

5) Agriculture: IoT sensors monitor soil moisture, temperature, and crop health, enabling data-driven farming decisions [6].

Emerging trends. In the last few years, IoT has had some interesting emerging trends, such as edge computing, fog computing, deep learning, and connected autonomous vehicles. Also, in the previous few months, we have seen IoT used successfully in managing and controlling the COVID-19 pandemic. The following subsections discuss all the above-mentioned emerging trends.

Internet of Behavior (IoB). IoT is a fusion of sensors, actuators, and connectivity technologies, whereas the Internet of Behavior (IoB) is a fusion of IoT, intelligence, and behavioural science. IoB can be seen as an extension of IoT. Its goal is to understand better data to facilitate better product development and promotion, focusing more on evolving human psychology. Javaid et al. posited that the inception of IoB can change the dynamics of product or service design, marketing, and customer services due to its ability to understand and modify consumer behaviours based on their comportment, tastes, and imaginations. Authors [13] analysed the power of various "things" in IoT products to enhance purchase intention by improving the functional and emotional experience. Stary stressed that IoB would transform the business and organisation with choreographic intelligence. IoB is now in its infancy, and its success coincides with large-scale IoT deployments with a high level of user acceptance.

Pandemic Management. Today's world is witnessing the devastation of the COVID-19 pandemic. The WHO categorically said several times that the world response could be far better than what we did and are doing. Most healthcare systems must be prepared and overburdened, whether in developed countries like Italy and the US or developing countries like India and Brazil. From the prism of sophisticated technologies, IoT can be used effectively to monitor and control the COVID-19 pandemic. IoT infrastructure coupled with intelligence can address challenges during the lockdowns, such as social distancing, contact tracing, healthcare monitoring, prescreening, remote meetings, accessibility anytime and anywhere, etc. IoT can play a significant role in providing virtual healthcare (contactless) tools and telemedicine to the masses, which will eventually help in achieving the goals of Healthcare 4.0. For example, Smart Field Hospital in Wuhan used IoT and AI-based applications to help healthcare workers relax. Robots and IoT devices have helped perform contactless body temperature monitoring, cleaning, disinfecting, etc. On the other hand, IoT sensors can help to track infection by forming a web of human nodes and their connections.

However, some serious privacy concerns must be addressed. Why has the world messed up in managing and controlling COVID-19? The answer is that human decision-making could be faster and more balanced. Some authors proposed iResponse, an intelligent IoT-enabled system for autonomous COVID-19 pandemic management, to address this issue. The authors demonstrated through iResponse that the fusion of IoT and intelligence could help break the chain of infection, cure development, treatment, resource planning, pandemic analytics, and decision-making. Still, we must deploy IoT infrastructure on a large scale worldwide to exploit its benefits. Connected Autonomous Vehicles IoT will connect AVs and help develop the driving cognitive. However, connected AV development is in its infancy, and a lot depends on how the adopter's willingness to accept the change and pricing of these vehicles as well as such issues as pedestrian detection, intersection navigation, communications, collision avoidance, and security. The core concept used in TAAWUN can also use IoT infrastructure in the future and sense data for prediction. AVs have recently suffered deadly crashes during the testing phase. This shows that ML algorithms used by AVs still need to mature for real-world challenges. After TAAWUN, a few words will be used to examine the benefits of connected AVs. Some authors critically discussed recent advances in connected AVs, focusing on five major areas: intersection navigation, pedestrian detection, collision avoidance, communications, and security. Safety is one of the foremost goals of any autonomous technology.

Edge and Fog Computing. Edge and fog computing push data intelligence and its processing closer to the nodes from where data is sensed or required, as depicted in Edge computing, which brings computational power closer to the sensing data than sending it to a remote cloud. This results in efficient speed and enhanced data-transport performance of devices and applications. Fog computing, an emergent architecture, can be termed a subset of the edge computing paradigm. Fog computing enables the cloud to be closer to the intelligent objects that generate data and actuators that act on sensed data. It defines the standards for edge computing, data transfer, storage, computation, and networking. Edge and fog computing are enabling technologies that will help IoT infrastructure to assist innovative applications and serve the giant objective of smart cities around the globe.

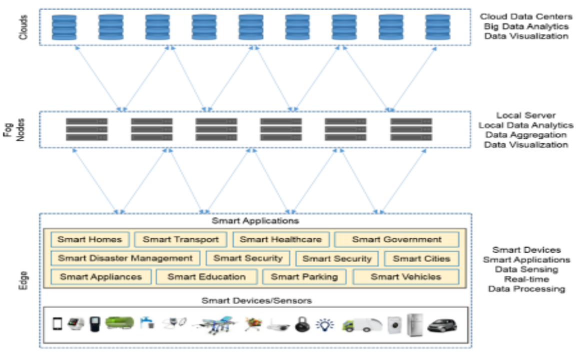

Figure 7 – Overview of Edge Computing

CONCLUSIONS

As a society, we need a prosperous and sustainable living environment. IoT can help significantly by providing actionable decision-making support. However, the critical question is: Are we equipped today as a society to use IoT? And are we ready to accept the cognitive changes it will bring to our daily lives? These essential questions will shape public acceptance of IoT applications and services. It will increase demand, encouraging industries to put serious efforts into the large-scale deployment of IoT, its applications, and services.

Further, the digital divide is another major challenge today. Our societies suffer due to the imbalance in global economic growth. If we want to exploit IoT to its full potential, we must address the issues mentioned above.

IoT is far more mature now. More IoT applications are in practical use. Individuals, governments, and businesses have shown a keen interest in leveraging IoT's opportunities. An important question remains: How will IoT learn and think to provide a high degree of automation? The answer comes from other branches of computer science that understand and act like humans with the help of ML. In this paper, we highlighted the importance of ML for IoT's success and diverse ML-powered IoT applications rather than doing a classical literature review. We classify ML developments in IoT from three perspectives: data, application, and industry. The literature reviewed is wholly or partially applicable to the IoT ecosystem.

Further, we identified and discussed emerging IoT trends, including Internet of Behavior (IoB), pandemic management, edge and fog computing, connected autonomous vehicles, and lightweight deep learning, primarily focusing on machine learning to develop futuristic and sustainable solutions. Despite IoT's ability to transform our present-day societies into smarter and more sustainable ones, it has to overcome challenges, e.g., technological, individual, business, and those related to our societies. We conclude that ML developments in IoT will revolve around currently available and well-established ML methods, at least in the short term. However, in the future, we can see a fully autonomous IoT ecosystem with embedded intelligence capabilities that will be a tricky development from an ML point of view regarding device data and processing abilities. With the help of this work, the reader can see what ML means to IoT, how ML is used with IoT, and the prospects of ML in IoT.

REFERENCES

1. McKinsey Global Institute. (2015, June). *The Internet of Things: Mapping the value beyond the hype*. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Industries/Technology%20Media%20and%20Telecommunications/High%20Tech/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/Unlocking_the_potential_of_the_Internet_of_Things_Executive_summary.ashx

2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer.

3. Mahdavinejad, M. S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018). Machine learning for internet of things data analysis: a survey. *Digital Communications and Networks, 4*(3), 161–175. doi: 10.1016/j.dcan.2017.10.002

4. Quraishi, S. J., & Yusuf, H. (2021). Internet of Things in Healthcare, A Literature Review. *2021 International Conference on Technological Advancements and Innovations (ICTAI)*. doi: 10.1109/ictai53825.2021.9673369

5. Kaur, M., Kaur, H., & Singh. A. (2020). Smart Transportation Sytem using Internet of Things (IOT) : A Review. *International Journal of Advanced Science and Technology*, *29*(10s), 2293-2298.

6. Gupta, S. K., Thakur, P., Kumar, R., & Bharadwaj, P. (2024). IoT in Agriculture: A Review. *Recent Evolutions in Energy, Drives and e-Vehicles*, 73–83. doi: 10.1007/978-981-97-0763-8_7

7. Singh, R., & Singh, N. S. (2020). Use of IoT and Machine Learning for Efficient Power Management Through Smart Grid: A Review. *International Journal of Advanced Science and Technology, 29*(4), 8982–8990.

8. Piyush, Kumar, R., Goomer, N., Rana, H. S., Chhabra, S., & Singh, A. (2023). Predictive Maintenance for Industrial Equipments Using ML & DL. *2023 International Conference on Advanced Computing Communication Technologies (ICACCTech)*. doi: 10.1109/icacctech61146.2023.00071

9. Li, J., Othman, M. S., Chen, H., & Yusuf, L. M. (2024). A critical review of feature selection methods for machine learning in IoT security. *International Journal of Communication Networks and Distributed Systems, 30*(3), 264–312. doi: 10.1504/ijcnds.2024.138214

10. Van Buuren, S. (2018). *Flexible Imputation of Missing Data* (2nd ed.). Retrieved from https://stefvanbuuren.name/fimd/

11. Liao, Y., de Freitas Rocha Loures, E., & Deschamps, F. (2018). Industrial Internet of Things: A Systematic Literature Review and Insights. IEEE Internet of Things Journal, 5(6), 4515–4525. doi: 10.1109/jiot.2018.2834151

Singh, S., Anand, S., & Satyarthi, M. (2023). A Comprehensive Review of Smart Home Automation Systems. *Advances in Computer Science and Information Technology, 10*(2), 61–66.

13. Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks, 54*(15), 2787–2805. doi: 10.1016/j.comnet.2010.05.010