Fadi Al-Turjman · Satya Prakash Yadav ·
Manoj Kumar · Vibhash Yadav ·
Thompson Stephan *Editors*

Transforming Management with AI, Big-Data, and IoT

Springer

Transforming Management with AI, Big-Data, and IoT

Fadi Al-Turjman • Satya Prakash Yadav
Manoj Kumar • Vibhash Yadav
Thompson Stephan
Editors

Transforming Management with AI, Big-Data, and IoT

Springer

Editors
Fadi Al-Turjman
Department of Artificial Intelligence Engineering
AI and Robotics Institute, Near East University
Mersin, Turkey

Manoj Kumar
School of Computer Science
University of Petroleum and Energy Studies
Dehradun, Uttar Pradesh, India

Thompson Stephan
Department of Computer Science and Engineering, Faculty of Engineering and Technology
M. S. Ramaiah University of Applied Sciences
Bangalore, Karnataka, India

Satya Prakash Yadav
Department of Computer Science and Engineering
G. L. Bajaj Institute of Technology and Management (GLBITM), Affiliated to AKTU
Greater Noida, India

Vibhash Yadav
Department of Information Technology
Rajkiya Engineering College
Banda, Uttar Pradesh, India

ISBN 978-3-030-86748-5 ISBN 978-3-030-86749-2 (eBook)
https://doi.org/10.1007/978-3-030-86749-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book discusses the effect that artificial intelligence (AI) and Internet of Things (IoT) have on industry. The authors start by showing how the application of these technologies has already stretched across domains such as law, political science, policy, and economics and how it will soon permeate areas of autonomous transportation, education, and space exploration, only to name a few. The authors then discuss applications in a variety of industries. Throughout the volume, the authors provide detailed, well-illustrated treatments of each topic with abundant examples and exercises. This book provides relevant theoretical frameworks and the latest empirical research findings in various applications. The book is written for professionals who want to improve their understanding of the strategic role of trust at different levels of the information and knowledge society, that is, trust at the level of the global economy, of networks and organizations, of teams and work groups, of information systems, and, finally, of individuals as actors in the networked environments. It presents research in various industries and how artificial intelligence and Internet of Things are changing the landscape of business and management. It presents new and innovative features in artificial intelligence and IoT use and in raising economic efficiency at both micro and macro levels. Moreover, it examines case studies with tried-and-tested approaches for resolution of typical problems in each application of study.

Mersin, Turkey — Fadi Al-Turjman
Greater Noida, India — Satya Prakash Yadav
Dehradun, Uttar Pradesh, India — Manoj Kumar
Banda, Uttar Pradesh, India — Vibhash Yadav
Bangalore, Karnataka, India — Thompson Stephan

Contents

Artificial Intelligence for Smart Data Storage in Cloud-Based IoT

Pushpa Singh, Narendra Singh, P. Rama Luxmi, and Ashish Saxena

1 Introduction

Artificial intelligence (AI), the Internet of Things (IoT), and cloud computing are buzzwords in the modern technological era. AI is the technology that aims at making computers or machines equivalent to the human brain and thus capable of learning and problem-solving [27]. AI-based applications can be integrated easily with other emerging technologies like IoT, cloud, Big Data, and Blockchain [12]. IoT states a system of interrelated, connected objects or things that can collect and transfer data via the Internet. A substantial number of physical things are being associated with the Internet at an exceptional rate recognizing the concept of the IoT. Reports and recent trends show that there are more than 30 billion IoT connections, almost four IoT devices per person on average by the year 2025 (https://iot-analytics.com/state-of-the-iot-2020-12-billion-iot-connections-surpassing-non-iot-for-the-first-time/). These IoT applications generate massive data, and cloud

P. Singh (✉)
Department of Computer Science and Information Technology, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
e-mail: pushpa.kiet@kiet.edu

N. Singh
Department of Management Studies, GL Bajaj Institute of Management and Research, Greater Noida, India
e-mail: narendra.singh@glbimr.org

P. R. Luxmi
School of Electrical Engineering, Vellore Institute of Technology, Chennai, India
e-mail: sriramalakshmi.p@vit.ac.in

A. Saxena
Department of Management Studies, Sharda University, Greater Noida, India
e-mail: ashish.saxena2@sharda.ac.in

F. Al-Turjman et al. (eds.), *Transforming Management with AI, Big-Data, and IoT*, https://doi.org/10.1007/978-3-030-86749-2_1

computing delivers a way for those generated data to travel to their endpoint [5]. The adoption of cloud computing is recognized as a data-processing and storage facility. All real-time applications connected with IoT need just-in-time processing and quick action over the clouds. AI and IoT-based data have obtained much attention from researchers, academicians, and industrialists in health care, agriculture, telecommunication, e−/m-commerce, and transportations. Nowadays, AI-based approaches amplify the role of IoT in business monitoring, health-care monitoring, disease prediction, bioinformatics, research and development, stock market prediction, social network analysis, weather analysis, agriculture, transportation, and resource optimization. Implementation of these applications requires data storage and computational capacity generally provided by cloud-based services [13]. AI techniques are used to process the stored data in a high-precision and just-in-time manner. The cloud is a powerful tool for transmitting data through the traditional Internet channels as well as via a devoted direct link. IoT becomes the source of generating huge data, and the clouds become crucial for data storage [9]. Hence, the IoT and clouds are closely integrated to offer commercial business services and generally referred to as cloud-based IoT. Businesses like Amazon Web Services (AWS), Google, and Microsoft have become certain cloud-based IoT services leaders, making the challenge even more worthwhile. Further, cloud-based IoT is used to connect a wide range of smart things in various applications.

AI, IoT, and cloud computing play significant roles in various aspects in the present and in the future too. AI methods aim to gather data from various industries to process and collect the data generated from cloud-based IoT. Integration of AI, IoT, and cloud has transformed the overall storage capacity and digital world [24] and hence has become a hot topic for all researchers and academicians. This chapter aims to emphasize on the role of AI in cloud and IoT-based data storage.

The remainder of the chapter is systematized as follows: Section 2 focuses on cloud-based data storage. Section 3 discusses the role of IoT in clouds. Further, the role of AI in IoT and cloud data storage is introduced in Sect. 4. Section 5 explains the applications of AI, IoT, and clouds in various sectors, and Sect. 6 concludes the chapter.

2 Cloud-Based Data Storage

Cloud storage is an Internet-based storage system in which data are transmitted on remote storage systems. The data generated from IoT and other devices are stored, maintained, managed, backed up, and accessible to users via the Internet. Users usually pay according to their consumption of cloud storage on a monthly basis. The cloud-based primary services are database services, computing services, and storage services. There are four basic types of cloud storage: public, private, hybrid, and community cloud data storage, as shown in Fig. 1.

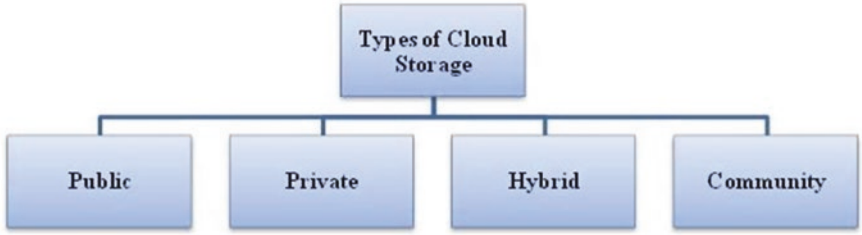

Fig. 1 Types of cloud storage

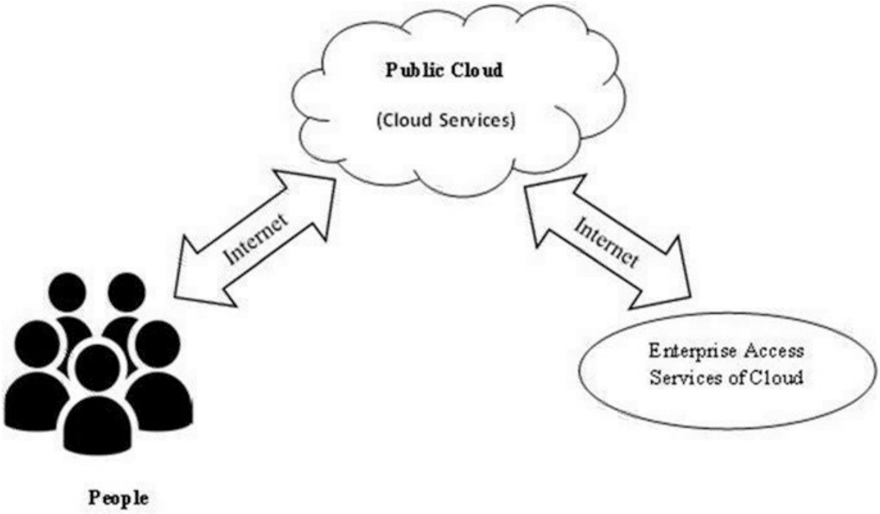

Fig. 2 Public cloud storage

2.1 Public Cloud Storage

Public cloud storage, also called Software as Services (SaaS), is provisioned for open use by the general public. Public cloud storage may be owned, managed, and functioned by a commercial, a government, or an academic. It offers data storage on a pay-per-use basis. In public cloud storage, the communication links can be assumed to be realized over the public Internet, as shown in Fig. 2. Any number of people, clients, organizations, or businesses can access cloud services by the mean of the public Internet. Examples of public cloud storage are Google App Engine, Microsoft Windows Azure, IBM Smart Cloud, Amazon Elastic Compute Cloud (EC2), etc.

Advantages

1. The provider is responsible for establishment, maintenance, technical support, and maintenance of the storage infrastructure and its associated costs.
2. Provider holds overall control of cloud environment. The subscriber's workload or data can be migrated at any time.
3. The workload can be transferred to data centers where the cost is low.
4. Public clouds potentially have a high degree of flexibility.
5. It is suitable for individual users and mid-sized organizations.

Limitation

- Data are at a high security risk due to the public domain and overall control of providers.

2.2 Private Cloud Storage

Private cloud storage [31] is a scalable and redundant storage solution where data are stored on distant servers devoted to an individual user, as shown in Fig. 3. Private cloud storage is safer than public cloud storage since it can be placed in the office/ premises of the company's data center or in another company's data center. Private clouds are divided into two parts:

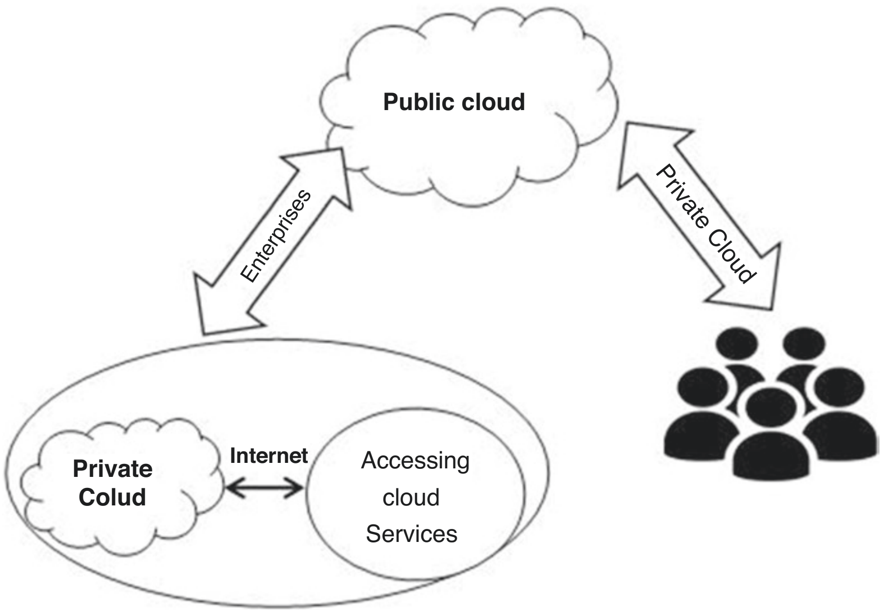

Fig. 3 Private cloud storage

On-Site Private Cloud—This type of private cloud is implemented at the user's locations.

Outsourced Private Cloud—This type of private cloud is employed at the server side, which is applied to a hosting business.

Advantages of Private Cloud Storage

1. The main benefit of private cloud storage is that it permits the user to have complete control over their data.
2. It offers high security, scalability, and reliability.
3. It is suitable for large enterprises or organizations.

Limitation

- It is expensive compared with public cloud storage.

2.3 Hybrid Cloud

Hybrid cloud combines private cloud (on-premises or off-premises) and public cloud as represented in Fig. 4. They have substantial deviations in performance, reliability, and security properties depending upon the type of cloud chosen to build

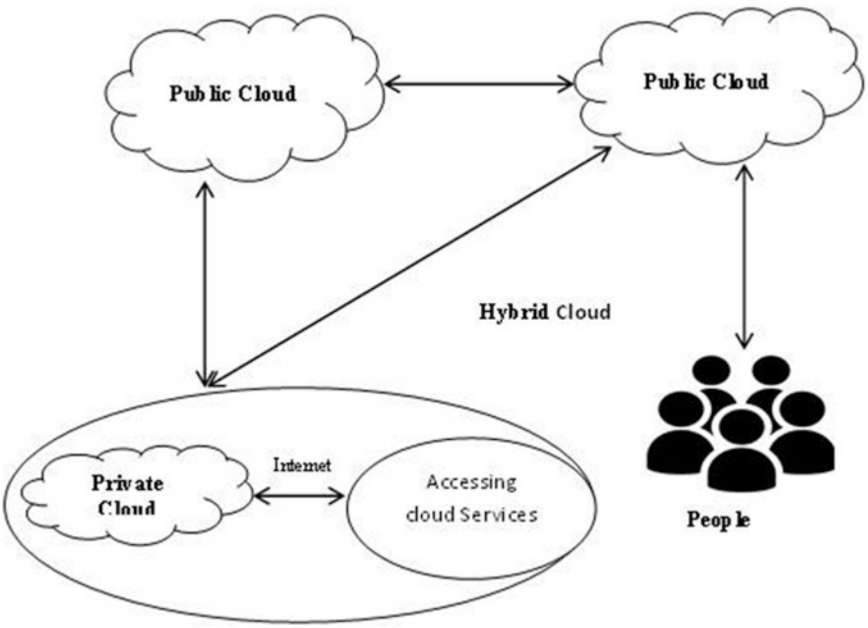

Fig. 4 Hybrid cloud storage

a hybrid cloud. Examples of hybrid clouds are Windows Azure, VMware, and vCloud.

Advantages

1. Customization is easy.
2. It is secure and reliable.
3. It is suitable for small and mid-size companies.

Limitation

- It consists of the limitation of both public and private clouds.

2.4 Community Cloud

Community cloud storage is the deviation of private cloud storage and exclusively useful for a particular community of users from organizations that have common concerns. Any data is stored in the community cloud owned as private cloud storage to manage the community's security. It offers a unique chance for businesses to work on common assignments. Google Apps for Government and Microsoft Government Community Cloud are some well-known community clouds.

Advantages

1. Community clouds are flexible as they permit to modification of the properties according to the user's needs.
2. Community clouds enable organizations to interact with their remote employees and provision of diverse heterogeneous devices.
3. Community clouds offer the user block facility. The service provider can block specific users from inserting, deleting, modifying, and downloading certain data sets and services.

Limitation

- It needs specialized security concerns as various organizations access and control the infrastructure.

3 Role of IoT in Cloud Storage

IoT has practically taken many industries such as health care [17], agriculture [18], transportation [6], telecommunication [28], real estate, etc. IoT offers best-connected environment for devices referred as "things" and generates huge amounts of data to be processed, analyzed, and communicated for the cloud. The public cloud storage assists the IoT services by offering third-party access to the organization. The integration of IoT and cloud can support IoT data or computational components operating over IoT devices [1]. Cloud-based IoT framework is represented in Fig. 5.

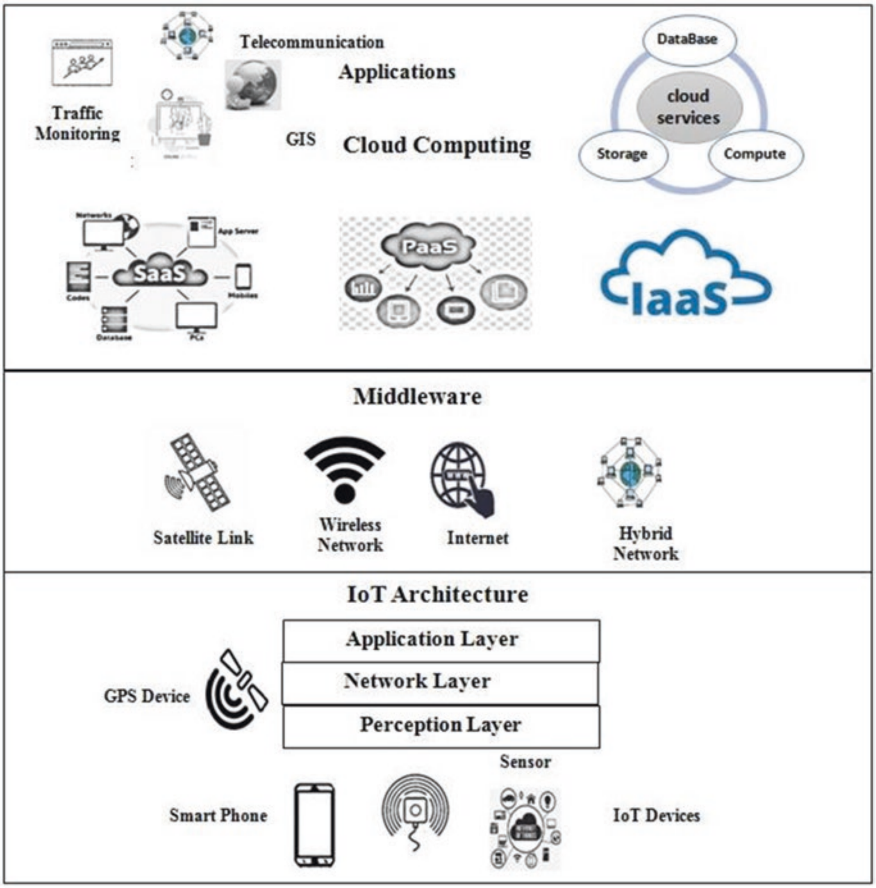

Fig. 5 Cloud and IoT-based environments

In Fig. 5, there are three components of cloud-based IoT storage. The first component provides IoT-based infrastructure where devices of different applications are connected with clouds using the second component, referred as middleware, which consists of communication technology such as 5G, Internet, Satellite Network, Wi-Fi, etc. The third component is the cloud infrastructure.

IoT-based Infrastructure: IoT-based infrastructure is based on a three-layered architecture, that is, perception layer, network layer, and application layer [7, 22].

- *The perception layer* is the physical layer or sensor layer of the architecture. The sensors and actuators are collecting data from things and transmit data for further processing [32]. It senses specific physical parameters or identifies other smart objects in the location.
- *The network layer* is liable for linking to objects (things), network devices, and servers.

- The *application layer* is the uppermost layer where IoT can be deployed. The application layer is accountable for conveying application-specific facilities to the users such as smart homes, smart transportation, smart agriculture, and smart health.

Cloud-based infrastructure offers cloud-based services such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) [14] for extending the application smarter. Users usually pay according to their usage, such as central processing unit (CPU) per hour, gigabyte (GB) storage per hour, IP usage per hour, etc.

- *Infrastructure as a service (IaaS)*: IaaS offers virtualized computing infrastructure over the Internet. In IaaS cloud-based services, it can be paid for services like storage, networking, servers, and virtualization. The example of IaaS is AWS EC2, Rackspace, Google Compute Engine (GCE), etc.
- *Platform as a service (PaaS)*: PaaS is a model in which cloud vendors offer developers a platform for building apps. PaaS provides a run-time environment to create, test, run, and deploy the application. Examples of PaaS are AWS Elastic Beanstalk, Heroku, and Windows Azure. Windows Azure is the most commonly used PaaS.
- *Software as a service (SaaS)*: By SaaS, one can generally access the software from any device, at any time, without installation of any software on your computer via the Internet. Examples of SaaS are BigCommerce, Google Apps, Salesforce, Dropbox, etc.

The foremost benefit of placing the IoT system in a cloud is that it offers more flexibility, scalability, and reliability to connected applications. Sometimes businesses don't need consistent requirements of data storage; it may be occasional. For example, online sales are increased particularly at the time of Diwali or Christmas; hence, only for 1–2 weeks, organizations require extra storage and computing infrastructure to meet all businesses' online requirements.

4 Role of AI in Smart Data Storage

Due to digitalization and recent technological advancements, data are generated enormously. About 90% of the world's data are generated just during the last couple of years [15]. These generated data can produce interesting patterns, meaningful information, and correlation if stored and processed efficiently. Earlier, these data were stored in data centers. In a data center, data are mostly stored on the premises of the business organization. Some data centers may be in a distributed location and not accessible in an efficient manner. Here, the cloud-based solution comes into the picture. The cloud is entirely off-premises, and entire data are accessible from anywhere and at any place via the Internet. IoT and clouds are making data storage as "smart." The IoT produces vast amounts of data, and cloud computing conveys a pathway for those data to travel to their destination. Data storage requirements of big industries are also high. These industries mostly rely on hybrid clouds due to the potential advantages of hybrid clouds such as elasticity, agility, cost-efficiency, etc. [20].

Before cloud storage, AI task was very expensive due to large data, hardware, and software requirements. The potential of cloud computing makes AI capabilities in a highly accessible manner. AI-based technologies and algorithms focus on the data to discover patterns or models that can explain or inform and predict. AI offers a new business dimension with cloud computing (cloud data storage) and assists corporations/businesses to organize their data, discover interesting patterns and correlation, deliver customer experiences, and improve workflows. Just imagine about a driverless car where a high level of accuracy is driven by AI and required to process data in terabytes just for a single autonomous car, but there is no provision to store our terabyte to exabyte data before processing. Here, cloud comes into the picture. But, if it does not have the ability to calculate accurate precision on time, then it is impossible to get exact route information timely for the driverless autonomous car. Here, AI takes the leading role. Vehicle-to-vehicle information is also required to avoid accidents, and here, it is important to realize the role of IoT. Overall, a driverless car or self-driving car cannot work without coordinating all these three buzzwords, that is, AI, cloud, and IoT.

It is significant to deal and lead with data before applying AI techniques for businesses. The following are the benefits of AI with cloud computing.

1. AI tools are used to enhance data management. Nowadays, businesses generate and collect massive repositories of data, and AI tools can process data management.
2. AI tools can support and streamline the way data are ingested, updated, and managed to provide precise real-time data to users, predict fraudulent, and identify the risk areas.
3. The customer relationship management (CRM) platform Salesforce and its Einstein AI tool are used to manage and enhance the CRM.
4. Within a cloud environment, AI requires historical data to identify patterns and trends and makes better recommendations for the customer.
5. Leveraging AI with clouds is a cost-efficient initiative. The cloud permits organizations to purchase only the storage they actually need and when they need.

AI and cloud-based storage system are transforming the business at every level. AI with cloud services enhance the AI based application and practices into the business so that the service providers can respond quickly, according to the market's competitive environment in advance. IoT-based cloud assists in renovating business and changing the world with AI.

5 Applications of AI, IoT, and Cloud in Various Sectors

In the recent age, emerging technologies such as AI, IoT, and clouds are at the bleeding edge of cost optimization and improved quality. AI and IoT have already proven themselves in different areas and sectors like health care, agriculture, transportation, e-commerce, and telecommunication [29] as shown in Fig. 6. In order to work with real-time applications, cloud computing is extended as fog computing that can satisfy the need of time-critical applications [30]. The applications of the AI with IoT and clouds in various sectors are accelerated the implementation of Industry 4.0 [34]. These emerging technologies are ready to adapt to the industry changes in real time and maximize the industry turnover.

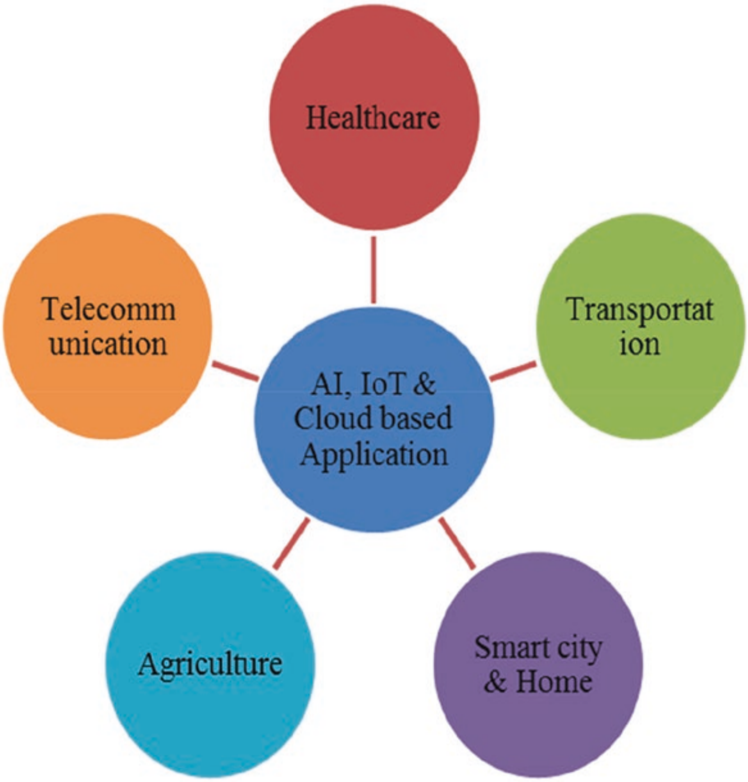

Fig. 6 Various AI, IoT, and cloud-based applications

5.1 Health-Care Sector

The health-care industry is generating enormous data that are stored in the cloud. Remote patients are connected via IoT sensor devices such as body temperature sensor, an electrocardiogram (ECG) sensor, blood pressure sensor, activity tracker, continuous glucose monitor (CGM), etc. [11]. These devices generate huge data from different locations. Hence, these devices need a cloud for continuous monitoring of specific attributes like blood glucose level, heartbeat, etc., for numerous days at a time by taking evaluations at regular intervals. AI and related techniques like machine learning and deep learning techniques are serving in timely analysis and prediction of diagnosis and prognosis of the disease [25]. Cloud assists in securing patient data and permitting health-care providers to continue conveying advanced technological care. A diversity of public, private, and hybrid cloud platforms supports getting better access to the patient's records and facilitating coordination between the doctor and patient.

5.2 Agriculture Sector

Automation of the agricultural system has enhanced the soil's gain and strengthened the fertility of the soil [3]. AI and IoT-based agriculture systems have various nodes such as soil, farmer, weather, irrigation, fertilizer, and crop management. These nodes or agents are like to distribute node that is connected through IoT, stored in a cloud environment for AI analytics. The various sensors are used for monitoring agricultural system. For examples, DHT11 is a low-cost digital temperature and humidity sensor, applied in the agricultural systems, and soil moisture sensor is applied for determining the soil moisture level. Cloud-based IoT is very supportive of integrating all agricultural-related data, such as soil-related data, weather data, crop-related data, farmers, supply chain, fertilizers, and retailers in the cloud [8]. The overall objective is to increase the productivity, prediction, and estimation of farming parameters to enhance economic efficiency.

5.3 Transportation Sector

Currently, transportation automation is the foremost area and emergent topic. IoT, AI, and machine learning techniques are profound for smart lighting systems and smart parking applications [35], as shown in Fig. 7. Moreover, route optimization, parking, and accident/detection seem to be the most popular applications. These applications produce and require a huge amount of data to be stored and processed. Cloud storage fulfills all these requirements of the smart transportation sector. Integration of AI, IoT, and clouds in transportation sectors makes this smart and also referred to as intelligent transport system (ITS) [16].

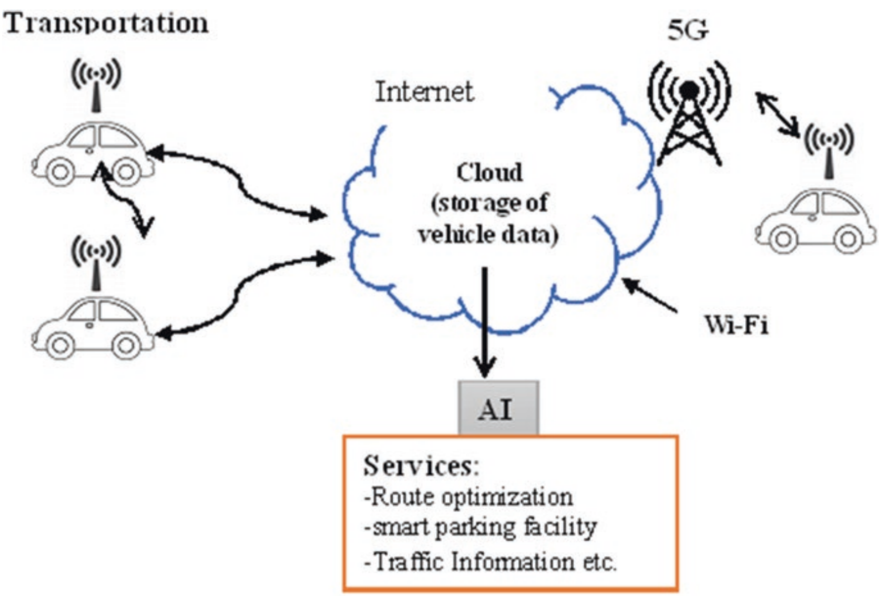

Fig. 7 Transportation with cloud, IoT, and AI

5.4 Telecommunication Sector

Telecommunication is itself facing a problem of poor quality of service (QoS), network selection criteria among multiple options and customer churning problems [4]. Telecommunication and mobile users want to always have the best-connected network. In fact, network infrastructure is a primary aspect of all emerging technologies. Telecommunication sector provides fastest network infrastructure to transmit the huge amount of data generated by AI based application. Emerging technologies also assist telecommunication infrastructure in searching best network, intelligent network selection, resource optimization, and user retention mechanism. Telecommunication network also leveraged with AI for intelligent network selection, resource optimization, and on-demand storage facility offered by cloud computing [23].

5.5 Smart City and Home

Smart city and smart home play a significant role in innovation trends [33]. AI and IoT are combinable with clouds and make the city and home smart [10]. Per and Skouby (2016) suggested an IoT architecture that incorporates smart homes and smart cities via the Cloud of Things (CoT) [19]. CoT virtualizes the IoT that offers monitoring and control [2]. AI, IoT, and clouds will contribute to the smart city/

home services developments. Concepts of the smart city, smart hospitals, smart education, smart governance, smart banking, and smart business [21] rely on AI, IoT, and cloud-based data storage. The collection vending machine (CVM) for e-waste management is associated with AWS cloud in order to make a city smart [26].

6 Conclusions

This chapter outlined the potential of AI and IoT in cloud-based data storage system. Recently, IoT-based applications are created by billions of connected devices. These devices are generating huge data, demanding huge data, and processing massive data. Various types of clouds such as public, private, hybrid, and community clouds fulfill the requirement of storage, processing, and maintenance of multiple applications. Further, the role of AI is also explained for analyzing and predicting the response to the IoT-based applications. In addition, cloud provides a place where one can implement the AI-based techniques for data analytics to make predictions about real-time behaviors of various applications such as health care, agriculture, transportation, telecommunication, and smart city/home.

References

1. Atlam, H. F., Alenezi, A., Alharthi, A., Walters, R. J., & Wills, G. B. (2017, June). Integration of cloud computing with internet of things: Challenges and open issues. In *2017 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData)* (pp. 670–675). IEEE.
2. Alohali, B., Merabti, M., & Kifayat, K. (2014, September). A cloud of things (cot) based security for home area network (han) in the smart grid. In *2014 eighth international conference on next generation mobile apps, services and technologies* (pp. 326–330). IEEE.
3. Al-Sakran, H. O. (2015). Framework architecture for improving healthcare information systems using agent technology. *International Journal of Managing Information Technology, 7*(1), 17.
4. Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data, 6*(1), 28.
5. Bhardwaj, A., Al-Turjman, F., Kumar, M., Stephan, T., & Mostarda, L. (2020). Capturing-the-invisible (CTI): Behavior-based attacks recognition in IoT-oriented industrial control systems. *IEEE Access*, 1. https://doi.org/10.1109/ACCESS.2020.2998983
6. Bui, K. H. N., & Jung, J. J. (2019). ACO-based dynamic decision making for connected vehicles in IoT system. *IEEE Transactions on Industrial Informatics, 15*(10), 5648–5655.
7. Burhan, M., Rehman, R. A., Khan, B., & Kim, B. S. (2018). IoT elements, layered architectures and security issues: A comprehensive survey. *Sensors, 18*(9), 2796.
8. Choudhary, S. K., Jadoun, R. S., & Mandoriya, H. L. (2016). Role of cloud computing technology in agriculture fields. *Computing, 7*(3), 1–7.
9. Chithaluru, P., Al-Turjman, F., Kumar, M., & Stephan, T. (2020). I-AREOR: An energy-balanced clustering protocol for implementing green IoT in smart cities. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2020.102254

10. Dubey, S., Singh, P., Yadav, P., & Singh, K. K. (2020). Household waste management system using IoT and machine learning. *Procedia Computer Science, 167*, 1950–1959.
11. Dias, D., & Paulo Silva Cunha, J. (2018). Wearable health devices—Vital sign monitoring, systems and technologies. *Sensors, 18*(8), 2414.
12. Gill, S. S., Tuli, S., Xu, M., Singh, I., Singh, K. V., Lindsay, D., … Pervaiz, H. (2019). Transformative effects of IoT, Blockchain and artificial intelligence on cloud computing: Evolution, vision, trends and open challenges. *Internet of Things, 8*, 100118.
13. Greco, L., Percannella, G., Ritrovato, P., Tortorella, F., & Vento, M. (2020). Trends in IoT based solutions for health care: Moving AI to the Edge. *Pattern Recognition Letters, 135*, 346–353.
14. Iqbal, S., Kiah, M. L. M., Anuar, N. B., Daghighi, B., Wahab, A. W. A., & Khan, S. (2016). Service delivery models of cloud computing: Security issues and open challenges. *Security and Communication Networks, 9*(17), 4726–4750.
15. Laura Drechsler. (2019, September). *The future of technology and innovation*. Online. Accessed from: https://www.experian.com/blogs/insights/2019/09/the-future-of-technology-and-innovation/
16. Nikitas, A., Michalakopoulou, K., Njoya, E. T., & Karampatzakis, D. (2020). Artificial intelligence, transport and the smart city: Definitions and dimensions of a new mobility era. *Sustainability, 12*(7), 2789.
17. Pandey, R. S., Upadhyay, R., Kumar, M., Singh, P., & Shukla, S. (2020). IoT-based HelpAgeSensor device for senior citizens. In *International conference on innovative computing and communications* (pp. 187–193). Springer.
18. Prathibha, S. R., Hongal, A., & Jyothi, M. P. (2017). IoT based monitoring system in smart agriculture. In *2017 international conference on recent advances in electronics and communication technology (ICRAECT)*. IEEE.
19. Per, L., & Skouby, K. (2016). Complex IoT systems as enablers for smart homes in a smart city vision. *Sensors, 16*, 1840.
20. Philip, A. T. (2018). *Designing and building a hybrid cloud*. O'REILLY. Online. Accessed from https://cdw-prod.adobecqms.net/content/dam/cdw/on-domain-cdw/brands/nutanix/o--reilly-ebook-designing-and-building-a-hybrid-cloud.pdf
21. Radu, L. D. (2020). Disruptive technologies in smart cities: A survey on current trends and challenges. *Smart Cities, 3*(3), 1022–1038.
22. Sethi, P., & Sarangi, S. R. (2017). Internet of things: Architectures, protocols, and applications. *Journal of Electrical and Computer Engineering, 25*, Article ID 9324035. https://doi.org/10.1155/2017/9324035.
23. Shane Wang. (2019). *Artificial intelligence applications in telecommunication clouds*. Online. Accessed from https://01.org/blogs/qwang10/2019/artificial-intelligence-applications-telecommunication-clouds.
24. Siebel, T. M. (2019). *Digital transformation: Survive and thrive in an era of mass extinction*. RosettaBooks.
25. Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Diagnosing of disease using machine learning. In *Machine learning and the internet of medical things in healthcare* (pp. 89–111). Academic Press.
26. Singh, K., Arora, G., Singh, P., et al. (2021). IoT-based collection vendor machine (CVM) for E-waste management. *Journal of Reliable Intelligent Environments*. https://doi.org/10.1007/s40860-020-00124-z
27. Singh, P., Singh, N., & Deka, G. C. (2020). Prospects of machine learning with Blockchain in healthcare and agriculture. In *Multidisciplinary functions of Blockchain technology in AI and IoT applications* (pp. 178–208). IGI Global.
28. Singh, P., & Agrawal, R. (2018). A customer centric best connected channel model for heterogeneous and IoT networks. *Journal of Organizational and End User Computing, 30*(4), 32–50.
29. Singh, P., & Singh, N. (2020). Blockchain with IoT and AI: A review of agriculture and healthcare. *International Journal of Applied Evolutionary Computation (IJAEC), 11*(4), 13–27.

30. Singh, P., & Agrawal, R. (2020). An overloading state computation and load sharing mechanism in fog computing. *Journal of Information Technology Research (JITR), 14*(3).
31. Shankar, A., Pandiaraja, P., Sumathi, K., Stephan, T., & Sharma, P. (2020). Privacy preserving E-voting cloud system based on ID based encryption. *Peer-to-Peer Networking and Applications*. https://doi.org/10.1007/s12083-020-00977-4
32. Stephan, T., Al-Turjman, F., Joseph, K. S., Balusamy, B., & Srivastava, S. (2020). Artificial intelligence inspired energy and spectrum aware cluster based routing protocol for cognitive radio sensor networks. *Journal of Parallel and Distributed Computing*. https://doi.org/10.1016/j.jpdc.2020.04.007
33. Talari, S., Shafie-Khah, M., Siano, P., Loia, V., Tommasetti, A., & Catalão, J. P. (2017). A review of smart cities based on the internet of things concept. *Energies, 10*(4), 421.
34. Wan, J., Yang, J., Wang, Z., & Hua, Q. (2018). Artificial intelligence for cloud-assisted smart factory. *IEEE Access, 6*, 55419–55430.
35. Zantalis, F., Koulouras, G., Karabetsos, S., & Kandris, D. (2019). A review of machine learning and IoT in smart transportation. *Future Internet, 11*(4), 94.

Big Data Analytics and Big Data Processing for IOT-Based Sensing Devices

Pawan Kumar Pal, Charu Awasthi, Isha Sehgal, and Prashant Kumar Mishra

1 Introduction

The specialized advancements and brisk converging of radio transmission, arithmetical personal computer (PC) gadgets, and miniature micro-electromechanical system (MEMS) instruments, that is, electromechanical framework instruments, lead to the improvement in Internet of Things (IoT); predictable with the Cisco report [1], the Web's associated measure of items has outstripped the measure of populace in the world. Such objects, which are associated with the Internet, incorporate PCs, cell phones, tablets, Wi-Fi-empowered sensors [29], wearable gadgets, and family apparatuses, and structure IoT, are given in Fig. 1.

According to the reports, the measure of gadgets associated with the Internet has increased usage from 22.9 billion (in 2016) to 50 billion (in 2020) as deduced from Fig. 2. Almost all IoT applications do not just represent considerable authority in checking discrete occasions, but also the information collected by IoT objects is mined. Almost all data assortment devices in the IoT climate are gadgets that have sensors fitted and need custom conventions like the following:

- Message Queuing Telemetry Transport Protocol (MQTT)
- Data dispersion administration (DDS)

P. K. Pal · C. Awasthi (✉) · P. K. Mishra
Department of Computer Science and Engineering, PSIT, Kanpur, Uttar Pradesh, India
e-mail: charu.awasthi@psit.ac.in

I. Sehgal
Department of Information Technology, PSIT College of Engineering,
Kanpur, Uttar Pradesh, India

F. Al-Turjman et al. (eds.), *Transforming Management with AI, Big-Data, and IoT*, https://doi.org/10.1007/978-3-030-86749-2_2

Fig. 1 Big data sources in IoT

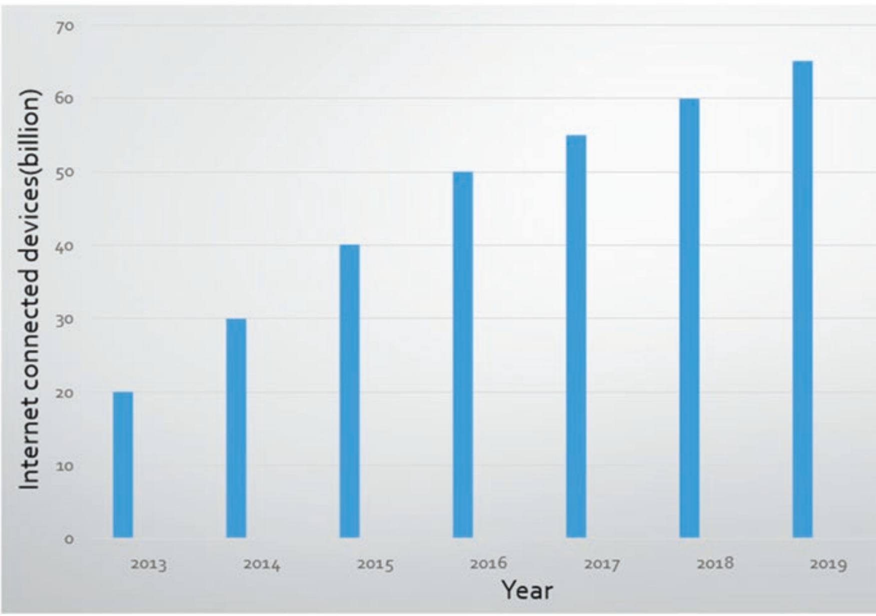

Fig. 2 Internet Connected Devices (Billion)

However, sensors are recommended in virtual businesses, IoT is anticipated to gracefully colossal measure of the information. Information created from the IoT gadgets is frequently incorporated for discovering:

- Potential examination patterns
- Exploration of the effect of specific occasions or choices

Such data are handled using different logical instruments [2].

The advancement of enormous data and IoT is rapidly influencing all territories like innovations and organizations by intensifying unions' preferences (associations) and people. The extension of information made through IoT [] has presumed a genuine element on the enormous data set. To help with comprehension of quite a gigantic idea, data researchers at IBM advocated the four "V's" of massive data: volume, variety, velocity, and veracity [3]:

1.1 Data Volume

Enterprises have ever-developing data, all things considered, effectively accumulating terabytes, even petabytes; for example, changing 350 billion yearly meter readings to raise and foresee utilization. Data sets are regularly enormous to such an extent that they can't fit on one worker and must instead be dispersed between a few stockpiling areas. Data examination programming like Hadoop [34] is made to oblige the need for circulated capacity and conglomeration.

1.2 Data Variety

For time-delicate cycles like getting fakes, huge data must be utilized since it streams into the undertaking. The present advanced data cannot be regularly corralled into conventional structures. Ground-breaking examination programming seems to saddle unstructured data, like pictures and recordings, and blend it in with more direct data streams to gracefully extra bits of knowledge.

1.3 Data Velocity

Big Data [32] comprises a wide range of information—organized and unstructured data like content, sensor data, sound, video, log records, and so on. Presently, data are gathered at an awesome pace of two 0.5 quintillion bytes for each day. From a large number of Web-based media that present on every 5 billion Google search for each day, gathered data are gushing into workers at a formerly exceptional speed.

1.4 Data Veracity

Data veracity alludes to honesty or exactness of a specific arrangement of the information, which has the assessment of information source:

- Is it dependable?
- Or would it not lead the investigators free?

As helpless data quality costs around $3.1 trillion per annum in the United States, so seeking after veracity is crucial. There are incorporations in it, leading to the following:

- Wipe out duplication
- Limit predisposition
- Cycle data in the manners by which it accumulates for the apparatus or vertical

The overall acknowledgment of IoT has made huge data examinations fascinating due to the preparation as well as an assortment of data by a particular sensor encompassed by the IoT biological system. The International Data Corporation (IDC) article showed that the gigantic data commercial center led to an increase of over US$125 billion by the year 2019 [4]. Figure 3 shows the strategy for information assortment, checking, and data investigation [5].

Large data examination enables data diggers and researchers to investigate gigantic measures of unstructured data that will be used for utilizing customary apparatuses [6] along with huge data examination that woks to immediately remove educated data-utilizing and data-preparing procedures that help in:

- Creating expectations
- Distinguishing ongoing patterns
- Finding concealed data
- Settling on the choices [7]

Procedures in data preparation are broadly sent for both issue explicit techniques and summed up data investigation. As needs be, measurable and AI techniques are used. IoT data are unique in relation to ordinary huge data gathered by means of frameworks as far as attributes due to different sensors as well as articles required during data assortment. It incorporates:

- Heterogeneity
- Commotion
- Assortment
- Fast development

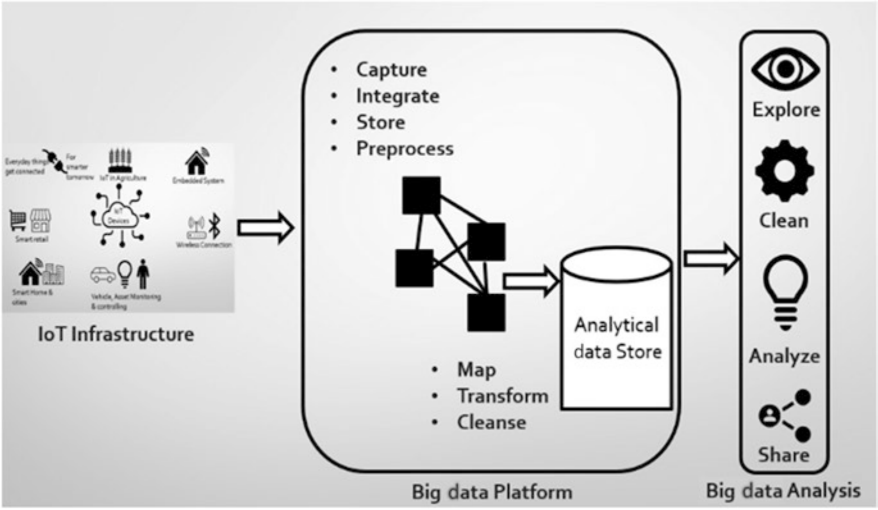

Fig. 3 Strategy for information assortment, checking, and data investigation

2 Outline of IoT Along with Big Data

2.1 Internet of Things (IoT)

IoT bids a stage for sensors as well as gadgets to talk flawlessly in a reasonable climate and empowers data sharing across stages in an advantageous way. The ongoing variation in different remote innovations places IoT on the grounds that the following progressive innovation by making the most of the total open doors offered by the Web's novelty. IoT has witnessed its ongoing appropriation in the shrewd urban areas and interest in creating keen frameworks, similar to office, retail, shrewd agribusiness, keen water, keen transportation, and smart medical services [30], and great energy [8].

Numerous specialized gadgets in the IoT worldview are implanted into sensor gadgets in the world. Data putting away contraptions sense measurements and move these data utilizing installed correspondence machines. The range of gadgets along with articles is interlocked through a spread of the correspondence arrangements, such as Bluetooth, Wi-Fi, Zigbee, and Global System for Mobile Communication (GSM). Such specialized gadgets transmit data and get orders from the distantly controlled gadgets. These orders license direct reconciliation with the actual world through PC-created frameworks to upgrade expectations for everyday comforts.

2.2 Big Data

The large volume of the information is formed by sensors, gadgets, Web-based media, medical care applications [31], temperature sensors, as well as different other programming applications along with advanced gadgets that persistently produce lots of organized, or semi-organized and unstructured, data, which is emphatically increasing. Such large data set prompts "enormous data" [9]. Enormous data-related investigations illustrate that the following outskirts for development, rivalry, and profitability. A few proclaim volume as a significant element of large data without offering an unadulterated definition [10]. Nonetheless, different scientists presented extra attributes for goliath data, similar to veracity, worth, changeability, and multifaceted nature [11]. The 4V's model, or its inductions, is the commonest portrayals of the expression "enormous data."

3 Big Data Analytics

It includes cycles of exploring through a database, mining, and breaking down data dedicated for improving organization execution [12]. It is the strategy for analyzing huge data sets that contain a spread of information types [13] to discover concealed examples like masked relationships, market patterns, client preferences, and other useful business data. Most of the target of Big Data analytics is to aid business relationships to get improved comprehension of information and settle on skillful choices. Big Data analytics requires advancements and apparatuses that will change an outsized measure of "organized, unstructured, and semi-organized data" in the form of a more reasonable data and metadata design for logical cycles. These scientific instruments' calculations should find examples, patterns, and connections throughout a spread of timelines in the data [14]. These instruments envision the discoveries in tables, diagrams, and spatial graphs for effective choosing after investigating the data. Subsequently, Big Data examination might be a genuine test for a few applications because of data intricacy and the adaptableness of basic calculations that support such cycles [15].

4 Variety of Data Types

The time associated with the origin of Big Data has delivered spread of different data sets from various sources in a few areas. These data sets include various modes, every one of which features a particular depiction, dispersion, scale, and thickness.

4.1 Network Data (Online)

One of the most focal points related to organization's Big Data is online informal community (online social networks [OSNs]) data, such as Facebook [16] and Second Life [17]. Center has expanded with headways in data investigation. Numerous examinations are performed with online informal communities utilizing information about delegate attributes at the full-scale level, for instance, little world highlights. Nonetheless, factors for the highlights of potential miniature cycles aren't very much spoken to in these investigations.

4.2 Mobile and IoT Data

The most common pattern of organization's Big Data is the investigation of portable and IoT data. With the occasion of 5G innovation, joined portable organizations have brought about critical upgrades in machine-to-machine interchanges execution. Joined versatile networks uncover relinquished reach groups in cellulite organizations, as Long-Term Evolution-Advanced, by utilizing intellectual radio innovation.

4.3 Geography Data

One examination tends to the gauntlets of significant kinds of innovation for three-dimensional (3D) association, and volume-delivering innovation upheld graphics processing unit (GPU) innovation. This work investigates visual programming for the hydrological climate upheld data direction. Also, it produces sea plans and forms planning of surfaces, component field planning, and dynamic recreation of the predominant field [18]. To bring present highlights up in space and accomplish continuous overhauling of an outsized measure of hydrological climate data, the examination builds hubs on the spot to control math to acknowledge dynamic planning of high properties.

4.4 Spatial–Temporal Data

Data are characterized in numerous classes upheld highlights and contrasts. Since the distinctions in data decide the accomplishment of the examination, they assume a significant job. Various highlights likewise are applied to search for a comparable highlight. Data in huge databases are regularly recovered by data preparing. In the instance of time-evolving data, when time becomes associated, the information is mined regarding both realities. The investigation of information mining as far as both existences has impacted the investigation of cell phones' investigation [19]. For the most part, spatial Big Data is engaged with vector data and raster data for the most part, and spatial Big Data is engaged with vector data, raster data, and organization data. The issue with utilizing databases from the disposition of room is that there are numerous obstructions at the security level.

4.5 Streaming as well as Real-Time Data

Amid the expansion for Internet Web-based features, network Big Data has advanced from spatial–worldly data to constant spatial–transient data. Gathering overviews for the most part needs consistent measurements' study owing to the constant reclamation of reports and insights over huge sum of data streams. The age of Big Data has started, and far of the information are wont to investigate the dangers of a spread of business applications. There are innovative preliminaries in the assortment of Big Data during a complex indoor modern climate. Indoor remote sensor organization (wireless sensor network [WSN]) innovation can conquer such limitations by getting together with Big Data from source hubs.

4.6 Visual Data

Within the period of Big Data, truly expanding measures of picture data should have presented critical difficulties to current picture investigation and recovery. It is vital to record pictures productively and adequately with semantic catchphrases, especially when gone up against the on the Web's quick developing properties.

4.7 Data's Associated Challenges

Every extraordinary data space increases difficulty, which, appropriately tended to, may crucially affect cutting-edge Big Data frameworks. In the primary spot, online organization data remain anticipating better models, with an expanded help from