Big data analytics for smart cities: Review

Honar M. Taha*[1] Shahab Wahhab Kareem[2]

[1]College of Engineering, Software & Informatics Department, University of Salahaddin, Iraq

[2]College of Engineering, Technical Information System, Engineering Department, Erbil Polytechnic University, Iraq

*Email of corresponding author: honar.taha@su.edu.krd

https://doi.org/10.31972/iceit2024.019

Abstract

Data science is driving major technological and operational improvements in cities by providing the necessary automation and intelligence to municipal operations. This work focuses on "Smart City Data Science," This is the process of searching for patterns and insights in municipal data from sensors and internet-connected devices to enhance decision-making and offer residents better services. Machine learning analytical modeling enhances the computer process's intelligence and actionability for real-world services, allowing for a deeper knowledge of municipal data. The report also includes a list of ten open research questions for the creation of data-driven smart cities. The key concepts, assessment methods, instruments, measurements, algorithm kinds, benefits, and drawbacks are also examined. Lastly, some of the difficulties, unsolved problems, and promising directions for further study are highlighted in large data handling strategies in smart cities.

Keywords: IoT, Big data, Smart City, Data analytics, ML, DL, data-driven.

1. Introduction

Several countries are supporting the idea of smart cities as a way to address the problems of ongoing urbanization, increasing population density, and providing a higher quality of life (Khan et al., 2022). The idea of "smart cities" as we know it today evolved from sustainable urban development. According to Khan (2022) it is the inventive and creative integration of information and communication technologies (ICTs) for intelligent and productive use of resources. Globally, more than 50% of people live in cities, and within the next 30 years, that number is expected to rise to 68% (Atitallah et al., 2020). The UN predicts that by 2050, there will be 2.5 billion more people on the planet. Due to their rapid growth, urban areas will face several issues, such as managing and developing sustainably and preserving high living standards for their citizens (Atitallah et al., 2020).

The development of smart cities is generally viewed as an application of the Internet of Things (IoT), which is characterized as a network of interconnected, heterogeneous components that sense, gather, transmit, and analyze data for intelligent services and systems(Iqbal H Sarker, 2022) .

IoT-connected smart devices have the ability to share and access permitted data to facilitate intelligent decision-making. IoT thus offers the fundamental building blocks for smart cities, including application handling for smart city services, data analytics, and intelligent decision-making (Iqbal H Sarker, 2022).

IoT generates massive amounts of data. These findings are sometimes referred to as big data which describes massive amounts of data that necessitate new technologies and architectures for data management (gathering and processing) in order to enable value extraction for improved understanding and decision-making (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review, 2021). Artificial Intelligence concepts are used to make intelligent decisions about huge data. Artificial intelligence possesses the capability to examine vast quantities of information, emulate cognitive processes, and even simulate human conduct (Zhang et al., 2021).

Despite having distinct analytical meanings, the terms "smart cities" and "big data" (or data analytics) are commonly used interchangeably to describe technological and digital expenditures made in metropolitan areas (Löfgren and Webster, 2020). As illustrated in Fig. 1, the creation of digital smart cities encompasses big data from a variety of fields, covering public safety, infrastructure, social services, business, government, finance, environmental sanitation, health, hygiene, culture, tourism, parks, entertainment, and water resources (Li et al., 2022).

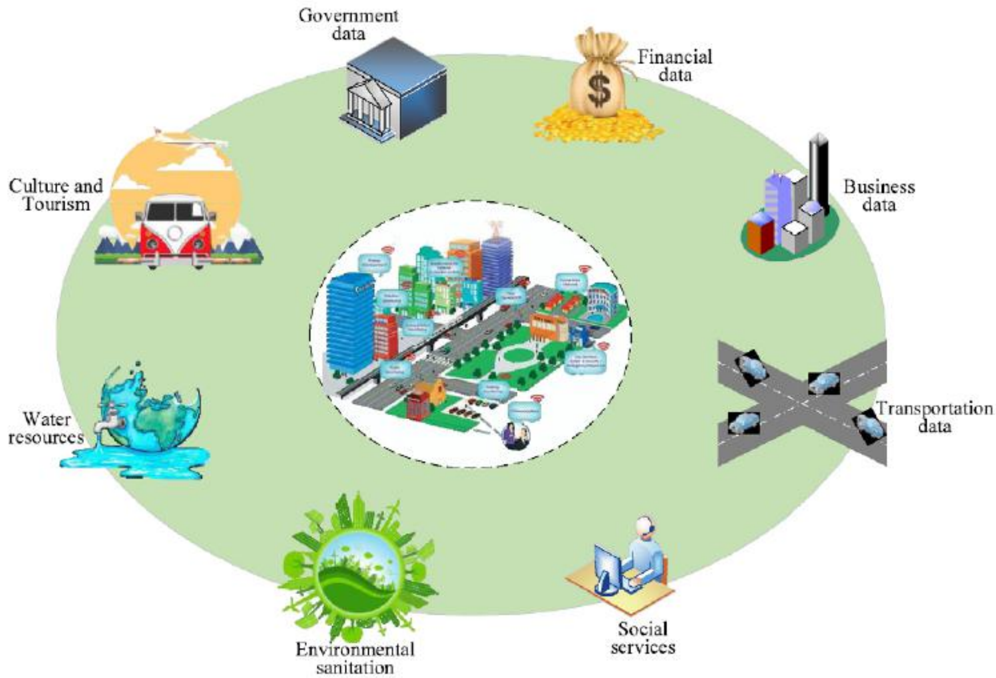

Fig 1(Smart City) Taken form (Li *et al.*, 2022)

Smart-environment applications have been greatly enhanced by the Internet of Things (IoT), but there are still many obstacles to be overcome in terms of gathering, storing, processing, and evaluating massive amounts of data. These applications require real-time processing across high-speed data streams, and users find it challenging to interpret and derive insights from big data due to the exponential increase in new sensors (Moustaka, Vakali and Anthopoulos, 2019). Fig 2 displays the primary tools utilized in this field.

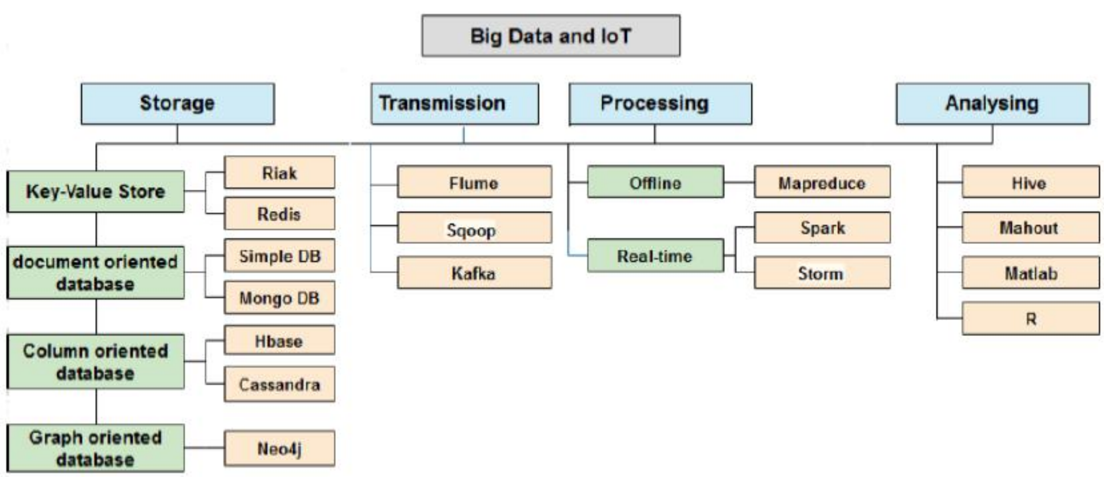

Fig 2(Primary tools utilized in IoT and Big data) Taken form (Moustaka, Vakali and Anthopoulos, 2019)

2. Research Methodology

This research employed a systematic approach to review the current literature on smart cities and big data. The initial search yielded over 520,000 publications. We applied strict criteria to filter these articles, focusing on those published from 2018 onwards, from reputable sources such as Springer and IEEE. Articles were selected based on their relevance to smart city data analytics, and a comprehensive analysis was conducted on the methodologies and challenges in this field.

This review is the result of our research interests and the heated debate in academic and commercial circles surrounding the smart city. The study utilized a methodology that sought to grasp how scientists approached the creation, gathering, and examination of data related to smart cities in order to offer comprehension and insights into the related cutting-edge.

Articles Selection: A preliminary search of the internet yielded 520,000 publications containing relevant studies. It was thought that a methodical approach would be required to help restrict the first number of publications using stringent criteria because the subject has a lot of references (Moustaka, Vakali and Anthopoulos, 2019).

2.1 Published Year

We applied a filter to obtain only articles no older than 2018 as in Fig 3

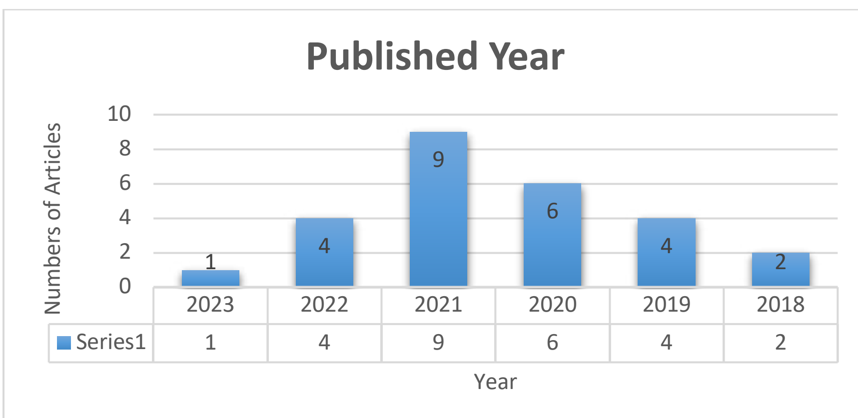

Fig 3 (Articles published between 2018 and 2022)

2.2 Taken From

The filter is also included to have most articles from well-known sources like Springer, IEEE, DirectScience, and others as in Fig 4.

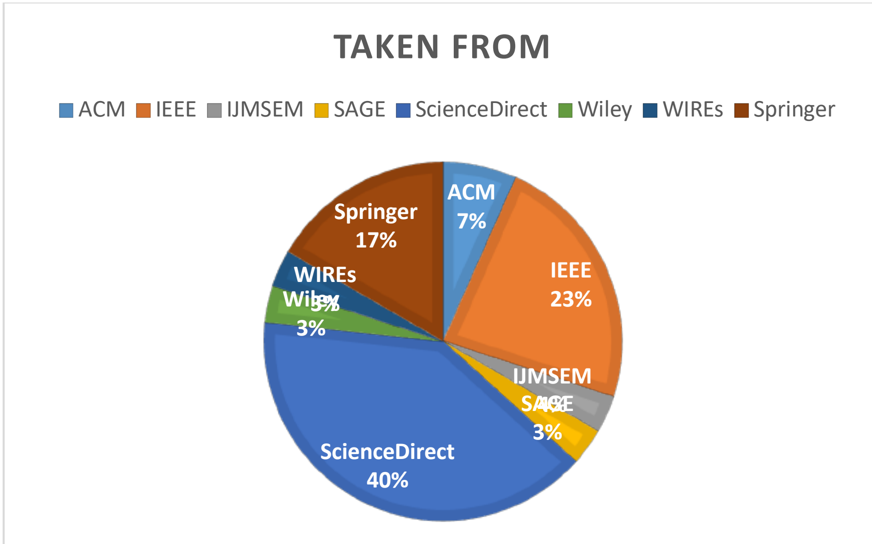

Fig 4 (Articles taken from)

3. Main Body

3.1 Big data technologies for smart cities

IoT devices collect data, which must be retained and analyzed to obtain pertinent information that can aid in decision-making. Data analytics is the process of transforming data from its most basic forms into insights and actions. As illustrated in Fig. 5, there are three categories of data analytics: descriptive, predictive, and prescriptive (Atitallah et al., 2020).

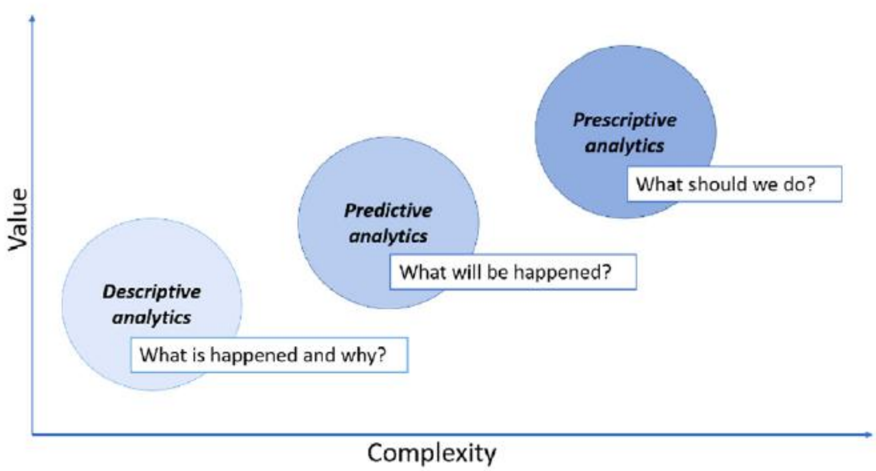

Fig 5 (Data Analytics types taken from (Atitallah et al., 2020))

The 7Vs—Volume, Velocity, Variety, Veracity, Variability, Value, and Visibility—define the characteristics of big data.

The format in which the data is collected can include descriptive metadata in addition to text, numbers, photos, audio, and video. Depending on the type of data source, metadata might include information about time, measurements, records, unique characteristics (ID, MAC address), social media content (posts, links, photos), and other data. Finally, according to Moustaka, Vakali, and Anthopoulos (2019), there are three methods to collect data: offline (for surveys and statistics), periodically (for sensors, actuators, and RFID), or continuously (for systems and applications).

Many machine learning approaches, which provide actionable insights or a deeper understanding of the data 130 and make the computing process smart, intelligent, and automated, can be used to achieve effective data science modeling.

The following is a list of key terminology for data-driven smart cities (Iqbal H. Sarker, 2022).

Internet of Things (IoT)

Internet of Things (IoT)-connected devices, including sensors, lights, meters, and more, are used to collect and analyze city data. Subsequently, the cities employ the data to enhance public utilities, infrastructure, services, and various functions (Sarker et al., 2022) .

Machine Learning (ML)

Machine learning applies knowledge gleaned from the collected municipal data to automate the process of building analytical models. Various machine-learning approaches can be used, depending on the type of data, to construct data science models for smart cities (Sarker, 2021b).

Deep Learning (DL)

Using the acquired city data as a source, deep learning also automates the process of building analytical models. However, it gradually pulls out more advanced features from the unprocessed city data by using several layers. The construction of data science models for smart cities can employ a variety of deep learning methods, contingent upon the type of data(Sarker, 2021a) .

Artificial Intelligence (AI)

The primary objective of AI (artificial intelligence) is to give machines the appearance of human intelligence. As illustrated in Fig. 6, artificial intelligence (AI) approaches such as the previously discussed ML (machine learning) and DL (deep learning) automate the creation of analytical models utilizing city data in the context of smart cities (Iqbal H. Sarker, 2022).

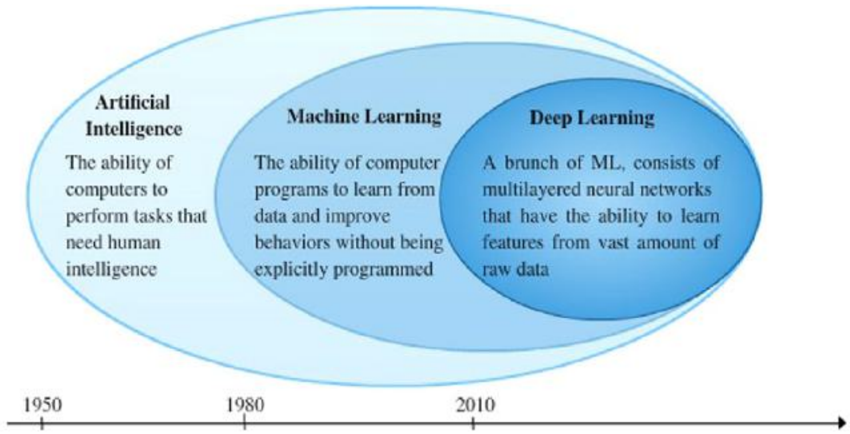

Fig 6 (A Venn Diagram of Deep Learning, Machine Learning and AI)

Cybersecurity

Smart cities are more vulnerable to different cyberattacks and threats as a result of technological advancement. Cybersecurity, then, is a concern because it refers to the process of defending sensitive data and important systems against cyberattacks, for which data science modeling may be a useful defense (Sarker et al., 2022).

After examining the ways in which IoT and big data were applied in many settings, we concluded that the most often used ecosystem across all of these researches was Hadoop. Other services, though, might also be made use of.

Apache Hadoop: The foundation of any big data infrastructure is this open-source project from the Apache Foundation. This design is meant for massively parallel processing and distributed data storage on commodity hardware computer clusters, which is a lot less expensive to run than a dedicated data center. A variety of frameworks and tools are available in the Hadoop ecosystem to manage different aspects of heterogeneous data handling, such as searching, data analysis, data aggregation, processing resource management, and storing (structured, semi-structured, and unstructured) (IEEE Staff, 2017). The two main subprojects that makeup Apache Hadoop's core are MapReduce for processing and HDFS for storage. To begin, Hadoop divides a big task into many data blocks, which it then distributes throughout the cluster's nodes (Shah et al., 2019). After that, MapReduce distributes the code bundle among the nodes so they can process it concurrently. Every block of data can be handled by a separate node as long as the prior process does not require identical nodes (Kannan et al., no date).

Apache Flume: enormous amounts of data streams, encompassing events and log files, can be efficiently collected, transferred, and aggregated using this distributed service into Hadoop. Data can be gathered from numerous real-time servers using the flexible, distributed, dependable, and available flume service. Ingesting data into Hadoop is the primary use case for this service.

Apache Kafka: This is a publish-subscribe distributed messaging system with excellent throughput. It is the most widely used technique in Apache Storm to retrieve data. Kafka is more resilient to failure, performs better than other conventional systems, and has an unchangeable and unrepeatable classification. Furthermore, it makes data usable for a variety of purposes and is appropriate for large-scale message processing (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review, 2021).

Apache Sqoop: A parallel load of data is provided by this open-source application. Bulk data is transmitted using the sqoop tool between HDFS and structured data stores, or relational databases (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review, no date).

Apache Spark: This open-source fast clustering computing engine is designed to process large amounts of data quickly. Spark is believed to be quicker than Hadoop while being built on top of its features. These include its in-memory capacity, interactive queries, and stream processing. minimal latency Additional concepts presented by Spark include high-level APIs in several languages, including as Python, R, Scala, Java, and Scala, as well as a variety of high-level tools (Shah et al., 2019).

Apache Storm: This is a strong foundation for networked computing in real time. Storm is an open-source program that is primarily utilized for extensive data analysis, much like Hadoop. While both frameworks function well together, there are several areas in which they differ, such as the fact that Storm is unstable but can accomplish all tasks, and Hadoop works well with MapReduce but has latency in real-time computations. Storm is interoperable with numerous programming languages, including Python and Scala, and can handle a large number of records per node per second on a moderately sized cluster (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review, 2021).

Apache Hive: This data warehouse project is built on Hadoop and is open-source. Hive is a tool for querying data stored in various databases and file systems that interface with Hadoop via a SQL-based interface (Babar and Arif, 2017). Big data analysis and complex query response via SQL-like queries (e.g., HiveQL) are made quick and easy with Hive.

Cloud Computing: Rather than controlling applications on local or personal servers, this type of computing uses shared remote servers that are accessible via the internet for data administration, storage, and processing (Dupont et al., 2017). Technologies related to big data and cloud computing can be thought of as "one body and two sides," with cloud computing serving as the platform for cloud computing applications and big data serving as the fundamental layer of computing resources and enabling the higher layer of big data processing (like Hadoop). Both cloud computing and IoT have developed swiftly and separately. Though these two technologies differ, they are complementary in many ways, enabling IoT to transcend its technological limits (storage, processing, etc.) by utilizing cloud computing's resources and almost infinite capabilities (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic

Review, no date). To meet user needs, the cloud provides three primary services: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Finally, there are the three varieties of cloud computing—public, private, and hybrid—each of which meets a particular set of user requirements. The most widely used cloud computing platforms for environmental applications are shown in Fig 7.

Access Technology: Strong, dependable, and energy-efficient data transmission is the cornerstone of any smart environment system. From a communication perspective, access technologies for any smart environment application should ensure consistent network connectivity and optimal services for effective data transmission between data collection sensors and back-end servers. It is crucial to guarantee data flow, network connectivity, and component safety, especially in emergency situations (like environmental monitoring) and warning scenarios (like disaster management). Therefore, a combination of several communication protocols and networks is required to provide a full network architecture for data transit and to help connect disparate data sources (Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review, 2021).

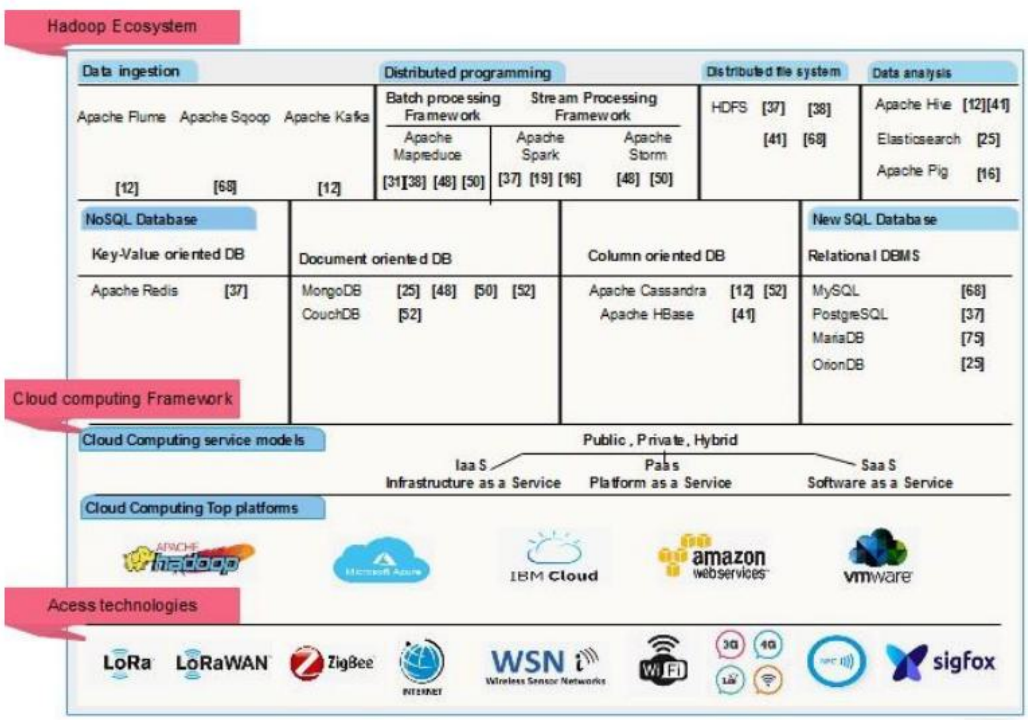

Fig 7 (big data and IoT tools/technologies for Smart city applications)

3.2 Applications of big data-driven in smart cities

This section highlights the significance of data-driven decision-making within the context of smart cities by examining data-driven smart city services and systems in a range of potential application domains as shown in Fig 8.

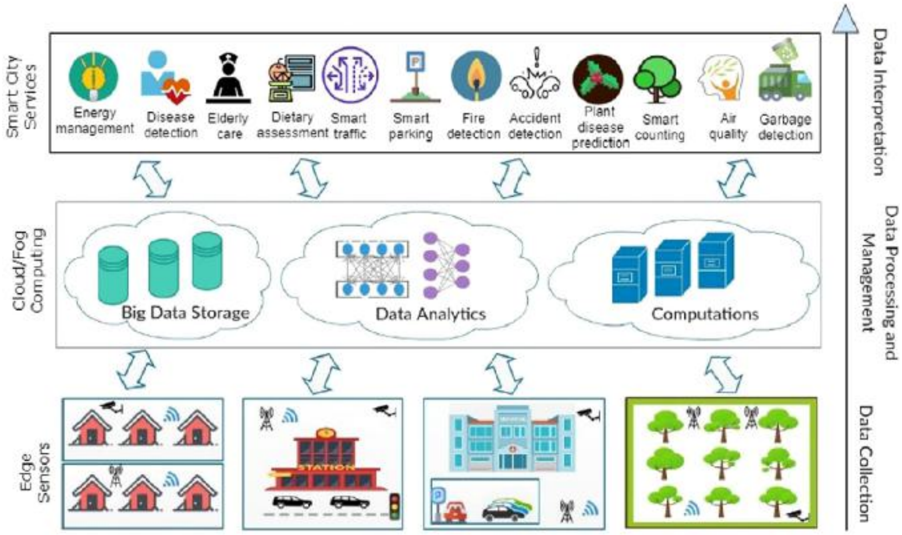

Fig 8 (Commonly used IoT and big data tools and technologies for applications in Smart Cities)

We have categorized them into ten groups:

Smart People or Citizen: Smart cities primarily serve as consumers for smart services and devices, requiring residents to engage in conversation and interaction to share online experiences and physical space. Data science modeling can help analyze historical data and inform decision-making. For instance, a crowdsourced weather application can evaluate current and upcoming weather events by combining manual data with automatic sensor readings from cell phones(Iqbal H Sarker, 2022).

Smart Transportation: Big cities with dense populations experience heavy traffic, which can harm roads and other infrastructure in addition to wasting a lot of fuel and time. One result of using DL models with IoT data is smart transportation. There are now a number of applications available that offer cutting-edge traffic management and transportation services. Enhancing the city's parking and traffic management, as well as enhancing individual safety and network coordination, are all made possible by improved transportation systems (Atitallah et al., 2020).

Smart Healthcare: Governments and the general public need to receive efficient digital health services in smart cities. Optimized health information and recommendations are provided by personalized health services, which use sophisticated Big Data analytics algorithms to profile individuals and populations. This supports the provision of individualized therapy and treatment plans, the promotion of good health, and the modification of dietary, physical activity, and stress-reduction practices. Healthcare costs are reduced as a result of people's increased ability to better control their lifestyle and health. These services are provided and monitored using IoT, m-Health, cloud-based, and personal applications (Iqbal et al., 2020).

Smart Energy: To raise the standard of living in smart cities, energy is necessary by lowering emissions, pollution, and energy use. Energy systems are becoming more intelligent thanks to developments in IoT and cloud computing, which are utilizing operational data to make better decisions (Soomro et al., 2019). Cities can benefit from an energy monitoring system since it provides officials and citizens with a clear image of the energy needed for different services. Big data analytics can be used in cities to achieve this. As a result, determining the primary energy-consuming sources and prioritizing tasks to improve behavior will be simple. This is in line with the European directive for future improvements in energy efficiency.

Smart governance: Smart governance, which prioritizes efficient public services, A few essential components of smart cities are continual innovation, mobile working, and community leadership. Data-driven decision-making allows for the intelligent provision of services based on citizen preferences. Governments can better understand trends and demands by using data science modeling to create policies and implement astute administrative practices. Four primary criteria are used to assess government actions: effectiveness, efficiency, transparency, and collaboration. Data-driven decision-making is used in smart governance to establish an atmosphere that is open, inclusive, collaborative, communication-based, and sustainable for both citizens and governments (Iqbal H Sarker, 2022).

Smart surveillance: Surveillance is the act of keeping an eye on and controlling behavior and activities. Automatic video analysis technology is used in smart surveillance systems to safeguard people and property against threats like fires and crimes. As such, intelligent monitoring has the potential to enhance municipal governance. Several practical applications have been made available to lessen the impact of anomalous occurrences like fires, floods, accidents, and medical crises (Atitallah et al., 2020).

Waste Management: Waste management is a major concern in many contemporary cities due to landfill storage issues as well as service costs. Adopting ICT solutions in this field may have significant positive effects on the economy, the environment, and savings. Recycling quality may be improved and waste collection costs may be decreased, for example, by using intelligent waste containers that can determine the amount of load and optimize the route for the collection vehicles (Karimi et al., 2021).

Smart Education: Since smart education services and applications can engage people in active learning environments to adapt to society's and the environment's rapid changes, they are often flexible and intelligent for citizens' lifelong learning. Efficient and productive educational processes can be achieved through the use of well-processed data, which can raise knowledge levels and improve teaching-learning resources. Enhancing a country's competitiveness and advancing the development of a knowledge-based society can be achieved through leveraging data science modeling expertise in smart cities (Iqbal H Sarker, 2022).

Air Quality: The 20-20-20 Renewable Energy Directive aims to reduce climate change by achieving a 20% decrease in energy consumption, a 20% reduction in greenhouse gas emissions, and a 20% increase in renewable energy consumption by 2020. Urban big data analytics can help monitor air quality in crowded areas and connect health applications on joggers to infrastructure, enabling people to choose the healthiest path (Karimi et al., 2021).

Smart lighting: It is thought that optimizing street lighting efficiency is a crucial component for complying with the 20-20-20 directive. The primary factors that this service can optimize are the time of day, the weather, and the number of people present. Streetlights must be integrated into the smart city infrastructure for this service to function effectively.

4. Discussion

The process of planning and making decisions regarding services in smart cities including smart parking, weather forecasting, public safety, transportation management, and surveillance is significantly impacted by Big data analytics. Big data analytics provides strong support for these purposes, hence implementing it is crucial to creating a sustainable smart city. Nonetheless, there are other obstacles to Big data analytics adoption, and this study examines them. The obstacles relating to urban planning are the most significant ones that prevent Big data analytics adoption.

4.1 Challenges

As smart cities evolve, several future directions are crucial for further research. The integration of IoT analytics with 5G and 6G technologies will be essential for improving data processing speeds and connectivity. Blockchain technology also offers promising solutions for enhancing the security and scalability of IoT systems in smart cities. Moreover, maintaining spatial and temporal dependencies in data integrity will be critical as the number of IoT devices continues to grow.

Creating a smart city ecosystem necessitates taking into account several novel difficulties. these difficulties can be categorized into two parts business and technical:

4.1.1 Business

Based on its application and consumers, IoT may be divided into three categories:

Industrial IoT: This subgroup is associated with industrial applications, including robotic production, wastewater systems, and additional uses for networked devices and systems.

Commercial IoT: Things utilized for commercial capacity, such as inventory controls, medical devices, and device trackers, are included in this class.

Consumer IoT: This subgroup comprises all IoT devices utilized by individual customers for amusement, personal usage, or business. These devices may include laptops, smartphones, smart automobiles, and entertainment systems.

4.1.2 Technical

Security and privacy: The issue of security is one of the primary difficulties that various IoT applications connected to smart cities face. Threats to deep learning models have been identified recently. The performance of these models is impacted, and their validity, accuracy, and reliability are all reduced. False data injection (Atitallah *et al.*, 2020) is one such attack that uses IoT devices and sensors located in various locations to deliver erroneous measurements and data. False data injection attempts to deceive the analytics process, producing inaccurate recommendations, forecasts, and results. Furthermore, one of the main issues with the smart city ecosystem is privacy. Sensors that are intended to collect data every second are deployed in various public spaces for use in Internet of Things applications. These data, which will be transferred to the cloud or fog for processing and analysis, may include images of people, speeches, or actions. Sensitive data sent to the cloud raises several privacy issues, such as images of individuals whose faces or actions are captured in public spaces, patient health information, etc.). Additionally, a third party without authorization may misuse or even get access to the data that has been gathered in the cloud or fog.

IoT big data challenges: Many IoT devices generate enormous amounts of data, which presents several challenges for data storage, communication, elaboration, and analytics. Since the created data cannot be managed using standard database administration tools, long-term storage of the data creates a significant difficulty. Large volumes of structured and unstructured data require specialized infrastructures to handle. Furthermore, certain technologies are needed for the analytics of IoT big data to extract useful insights and information. High-performance processors, cloud computing services that are scalable, fog and edge paradigms, and big data analytics applications, such as Hadoop, Apache, etc. Additionally, the generation of massive data increases the difficulty of quickly and well analyzing this data (Thakuriah, Tilahun and Zellner, 2017).

Quality of service: The quality of service (QoS) of the offered applications is a crucial requirement that needs to be verified to create a successful smart city ecosystem. Various Quality of Service (QoS) indicators, such as response time, availability, scalability, and dependability, are used to assess the applications' and systems' quality. Cloud services, storing frameworks, and processing platforms are some of the technologies that are used in conjunction to provide smart city services. To guarantee a flexible, strong, and dependable smart city ecosystem, it is crucial to guarantee the quality of service offered by these necessary technologies (Atitallah *et al.*, 2020).

Case Study: Smart City Implementation in Barcelona

Barcelona has been a leader in implementing smart city initiatives. By using IoT devices to monitor traffic, waste management, and energy usage, the city was able to reduce traffic congestion by 21%, lower carbon emissions by 15%, and optimize waste collection routes. This case demonstrates both the potential benefits and challenges faced in real-world smart city applications, showcasing how IoT and big data can improve urban living conditions.

Case Study: Smart City Implementation in Singapore

Singapore is another prime example of a smart city that has utilized big data analytics to transform urban living. Through its 'Smart Nation' initiative, Singapore integrates IoT sensors across public services to optimize transportation, energy, and water usage. One of the most notable achievements is the reduction of traffic jams by 30% through predictive traffic analysis and smart road pricing, improving both efficiency and quality of life.

Table 1: Comparison of Smart City Implementations

City	Key Technologies	Benefits	Challenges
Barcelona	IoT, Big Data, Smart Traffic Systems	21% reduction in traffic, 15% reduction in carbon emissions	Integration of systems, public acceptance

4.2 Future directions

Below are some future directions need to be reviewed more by researchers:

4.2.1 Transfer Learning:

The new model of learning known as transfer learning applies prior knowledge to solve new issues by first applying and transferring it. The main advantage of this type of learning is that the model training phase may be managed with minimal resources and in a shorter amount of time than many deep learning models that employ conventional methods of learning supervised/unsupervised/reinforcement (Atitallah *et al.*, 2020).

4.2.2 Integration of IoT analytics with 5G, and 6G

Mobile devices are widely used as data collection platforms, capturing daily activities and supporting embedded communication sensors. Data processing using deep learning (DL) approaches can enhance smart city services. As mobile data usage grows, communication networks need to be updated to support new issues. 5G is a new wireless network with high data rates, low latency, and reliability, surpassing current 4G networks. Integrating software with 5G can improve sensing, speed, and intelligence, creating opportunities for IoT and smart city development. 6G wireless networks are being researched to address high demands from mobile data traffic, transforming human life into a smart world.

4.2.3 Blockchain technology for secure and safe Internet of Things

Blockchain technology, defined as a sequence of blocks holding transaction records, has gained attention from researchers in various fields. It was first introduced to assist security methods, but it has since found use in a variety of industries and applications.

4.2.4 Reliable and scalable Internet of Things analytics

in smart city applications require efficient safety measures and systems capable of detecting malicious attacks. Log traces can be analyzed using deep learning approaches to identify weaknesses, enhance recovery processes, and prevent failures.

4.2.5 Maintaining Spatial and Temporal dependencies in data integrity and quality

The number of IoT devices in smart cities is expected to reach 41.6 billion by 2025, generating trillions of gigabytes of data. Infrastructures need to become scalable and adaptable; standards and technologies need to be updated, and this has to be managed. Particularly in mobile traffic scenarios that need for event location and device localization, accurate spatiotemporal data analysis is essential.

5. Conclusion

The importance of smart city data science in data-driven, intelligent decision-making in smart city services and systems is examined in this study. The goal is to get insights from city data, starting with study design and on to recommendations for fixes. In data-driven smart cities, various deep learning and machine learning approaches are used to create analytical models that cater to people's demands. Real-world services like smart environments, public safety, education, healthcare, transportation, energy management, and cybersecurity are all included in the idea of data-driven smart cities. By securing the availability of city resources in social, economic, and environmental aspects, a data-driven smart city can raise the standard of living for its residents. The study sheds light on the conception, reasoning, modeling, and processing of data science in smart cities and emphasizes the suitability of machine learning techniques for data-driven, intelligent decision-making in smart city services. Researchers and academics are able to undertake additional research in particular areas because the report reveals problems and open research issues. Researchers, scholars, and smart city designers have a bright future ahead of them with data-driven smart city studies, which will boost participation in government, business, and academia.

References

Atitallah, S. Ben *et al.* (2020) 'Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions', *Computer Science Review*. Elsevier Ireland Ltd. Available at: https://doi.org/10.1016/j.cosrev.2020.100303.

Babar, M. and Arif, F. (2017) 'Smart urban planning using big data analytics based internet of things', in *UbiComp/ISWC 2017 - Adjunct Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM*

International Symposium on Wearable Computers. Association for Computing Machinery, Inc, pp. 397–402. Available at: https://doi.org/10.1145/3123024.3124411.

Big Data and IoT-based Applications in Smart Environmental Fields: A Systematic Review (no date).

Dupont, C. *et al.* (2017) *International Information Management Corporation*. Available at: www.IST-Africa.org/Conference2017.

IEEE Staff (2017) *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*. IEEE.

Iqbal, R. *et al.* (2020) 'Big data analytics: Computational intelligence techniques and application areas', *Technological Forecasting and Social Change*, 153. Available at: https://doi.org/10.1016/j.techfore.2018.03.024.

Kannan, S. *et al.* (no date) *A Smart Agricultural Model by Integrating IoT, Mobile and Cloud-based Big Data Analytics*. Available at: https://www.researchgate.net/publication/323277477.

Karimi, Y. *et al.* (no date) *Leveraging big data in smart cities: A systematic review*.

Khan, M.A. *et al.* (2022) 'Investigation of Big Data Analytics for Sustainable Smart City Development: An Emerging Country', *IEEE Access*, 10, pp. 16028–16036. Available at: https://doi.org/10.1109/ACCESS.2021.3115987.

Khan, S. (2022) 'Barriers of big data analytics for smart cities development: a context of emerging economies', *International Journal of Management Science and Engineering Management*, 17(2), pp. 123–131. Available at: https://doi.org/10.1080/17509653.2021.1997662.

Li, X. *et al.* (2022) 'Big data analysis of the Internet of Things in the digital twins of smart city based on deep learning', *Future Generation Computer Systems*, 128, pp. 167–177. Available at: https://doi.org/10.1016/j.future.2021.10.006.

Moustaka, V., Vakali, A. and Anthopoulos, L.G. (2019) 'A systematic review for smart city data analytics', *ACM Computing Surveys*. Association for Computing Machinery. Available at: https://doi.org/10.1145/3239566.

Sarker, I.H. (2021a) 'Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions', *SN Computer Science*. Springer. Available at: https://doi.org/10.1007/s42979-021-00815-1.

Sarker, I.H. (2021b) 'Machine Learning: Algorithms, Real-World Applications and Research Directions', *SN Computer Science*. Springer. Available at: https://doi.org/10.1007/s42979-021-00592-x.

Sarker, Iqbal H. (2022) 'AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems', *SN Computer Science*, 3(2). Available at: https://doi.org/10.1007/s42979-022-01043-x.

Sarker, I.H. *et al.* (2022) 'Internet of Things (IoT) Security Intelligence: A Comprehensive Overview, Machine Learning Solutions and Research Directions'. Available at: https://doi.org/10.20944/preprints202203.0087.v1.

Sarker, Iqbal H (2022) 'Smart City Data Science: Towards Data-Driven Smart Cities with Open Research Issues'. Available at: https://doi.org/10.20944/preprints202204.0044.v1.

Shah, S.A. *et al.* (2019) 'The rising role of big data analytics and IoT in disaster management: Recent advances, taxonomy and prospects', *IEEE Access*, 7, pp. 54595–54614. Available at: https://doi.org/10.1109/ACCESS.2019.2913340.

Soomro, K. *et al.* (2019) 'Smart city big data analytics: An advanced review', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. Wiley-Blackwell. Available at: https://doi.org/10.1002/widm.1319.

Thakuriah, P. (Vonu), Tilahun, N.Y. and Zellner, M. (2017) 'Big data and urban informatics: Innovations and challenges to urban planning and knowledge discovery', in *Springer Geography*. Springer, pp. 11–45. Available at: https://doi.org/10.1007/978-3-319-40902-3_2.

Zhang, Y. *et al.* (2021) 'Big data and artificial intelligence based early risk warning system of fire hazard for smart cities', *Sustainable Energy Technologies and Assessments*, 45. Available at: https://doi.org/10.1016/j.seta.2020.100986.