Online-Journals.org

JFSE Journal for Future Society and Education

JFSE | eISSN: 3078-6010 | Vol. 2 No. 4 (2025) | OPEN ACCESS

https://doi.org/10.3991/jfse.v2i4.59259

PAPER

Service-Oriented Architecture for Big Data Analytics in Smart Cities

Shanza Iftikhar[1](✉), Bushra Hamid[1], Sidra Tahir[1], Kashif Mehmood[1], Ahthasham Sajid[2], Sabitha Banu[3]

[1]PMAS – Arid Agriculture University, Rawalpindi, Pakistan

[2]Fazaia Bilquis College of Education for Women, Rawalpindi, Pakistan

[3]PSGR Krishnammal College for Women Coimbatore, Coimbatore, India

shanzaiftikhar14@gmail.com

ABSTRACT

The rapid urbanization and the increased development of the Internet of Things (IoT) technologies have produced the unprecedented amounts of heterogeneous data in smart cities. The paper suggests a scalable and modular service-oriented architecture (SOA) that is combined with the big data analytics (BDA) to improve the decision-making process, resource management, and citizen services within the smart urban settings. The system integrates both real-time stream and batch-based processing model, which is interoperable, secure, and reliable through simple object access protocol (SOAP)-based services. The main elements are the tendency to analyze the requirements with the stakeholders, the layered systems design, and intelligent data management. Benchmark dataset experimental analysis of mobility, energy, and urban indices shows numerous (35, 40, 25, and 91) times lower latency, throughput, resource usage, and data integration accuracy than traditional monolith models. The paper brings to the fore the potential of SOA when it comes to solving the smart city problem of data heterogeneity, scalability, and real-time responsiveness. Future studies will investigate the combination of AI, edge computing, and blockchain to optimize the architecture further in terms of various urban areas.

KEYWORDS

smart cities, service-oriented architecture (SOA), big data analytics (BDA), Internet of Things (IoT), real-time processing

1 INTRODUCTION

Smart cities refer to cities that make the best use of resources by employing the latest technology and data analysis with the view of improving the quality of life in a particular city and ensuring sustainability. Because smart cities generate huge volumes of data generated by a variety of sources (sensors, devices, social media, and other IoT-enabled systems), the issue of data processing and analytics becomes the most significant [1]. One of the most important ingredients to creating valuable insights out of this information flood is big data analytics (BDA), which allows making informed decisions and taking real-time actions [2].

Iftikhar, S., Hamid, B., Tahir, S., Mehmood, K., Sajid, A., Banu, S. (2025). Service-Oriented Architecture for Big Data Analytics in Smart Cities. *Journal for Future Society and Education (JFSE)*, 2(4), pp. 20–42. https://doi.org/10.3991/jfse.v2i4.59259

Article submitted 2025-07-02. Revision uploaded 2025-10-10. Final acceptance 2025-10-20.

The fast urbanization has been associated with problems of resource management, sustainability, and the quality of life. A solution to such problems has been suggested to be smart cities through the use of advanced technologies and data-driven approaches to enhance lives in the city [3]. The key element to this revolution is the ability to collect, process, and analyze great amounts of data created by heterogeneous systems and devices in a city [4]. According to many, this colossal amount of data that is also often referred to as big data has tremendous potential in terms of improving decision-making, resource optimization, and innovative urban services. However, it has data integration, scalability, and real-time processing complications [5].

The core of the unlocking of the potential of smart cities is occupied by the BDA, which helps extract actionable information out of vast and complex amounts of data. These data come in large numbers based on a wide range of sources such as the Internet of Things (IoT), social media, cities, and citizen services [6]. An example is traffic data in sensors, energy data in grids, health care data in systems, and environmental data in monitors that are stated to be analyzed to produce information about city planning, managing resources, and citizen engagement [7]. Despite the potential game-changer BDA in smart cities, heterogeneous data formats, large volumes of data and strict security and privacy requirements are presented as challenges [8].

Service-oriented architecture (SOA) provides a flexible, modular, and scalable solution to these issues [9]. By providing a structure of software assets in the form of interoperable services, SOA makes integration of software elements smooth, efficient, and optimal for using data. SOA can also be used to integrate big data analytics services to support different applications in a smart city, including traffic control, energy conservation, citizen safety, and citizen engagement [10].

The chapter explains that SOA is an underpinning paradigm to integrate the big data analytics into the smart city infrastructure [11]. The combination of the scalability of big data technologies and the modularity and interoperability of SOA can help the smart city to build strong systems capable of addressing the multiplicity of data sources, impart real-time insights, and support new solutions to the city [12]. The study also determines the barriers and prospects of implementing an SOA-based architecture to big data analytics in smart cities, resulting in efficient and sustainable urban development [13].

Service-oriented architecture provides a solid structure to deal with these issues as it structures software elements into interoperable and reusable services [14]. SOA is a design approach and makes use of modules of functions where functionality is incorporated in services and they communicate with each other through well-documented interfaces. Through this architecture, it is feasible to combine disparate, heterogeneous information sources and applications of analytics together in a seamless manner and provides a platform that can be scaled and flexible to analyze big data in smart cities [15].

The ability to handle heterogeneous data in smart cities has become one of the key advantages of SOA. Since data may be obtained in various fields such as transportation, healthcare, energy, and the security of the people, an architecture must exist that facilitates interoperability and standardized communication [16]. SOA achieves this by establishing certain protocols and interfaces in order that dissimilar systems can communicate freely with one another. As an example, the traffic management systems may be integrated with the public transport and weather services in order to provide combined knowledge about the city mobility [17].

The other important necessity of BDA in smart cities is scalability. As cities are being decentralized and the numbers of IoT devices are becoming increasingly concentrated, the volume of information generated is skyrocketing [18]. The introduction of new services or upgrading of old services is facilitated by SOA without any impact on the whole system. This allows the architecture to grow in line with the evolving needs of smart cities, be it the addition of new sources of data or the execution of complex analytics tools [19].

One of the pillars of the majority of the smart city use cases entails real-time analytics, such as emergency response, traffic management, and energy management. SOA enables real-time processing by enabling orchestration of services that can handle high-velocity data streams. As an example, surveillance cameras within the framework of the traffic sensors and emergency centers can be processed in real-time in order to detect the events and organize the same to be responded to in the quickest manner possible [20]. This skill enhances the effectiveness and efficiency of the city services and makes the city a safer and more resilient place.

Besides its technical advantages, SOA promotes reusability and flexibility, which are critical to efficient and cost-effective projects of a smart city. Reuse of services can occur in a different context, assuming that they have been defined in one application, thereby saving time and money [21]. An environmental data service that processes environmental data to monitor the air quality can also be used to monitor the water quality or track weather patterns. This service reuse enables the cities to maximize their technology and infrastructure investment [22].

Smart city BDA implementation based on SOA is not a painless process despite its many benefits. The complexity of creating and managing a service-based system can be a nightmare, particularly in a rapidly evolving urban environment where the needs and priorities are highly likely to undergo a transformation at any given time [23]. The high-performance and real-time support that is required in the conditions of large amounts of data demand excellent infrastructure and the most advanced technologies. In addition, the issues of data governance, security, and privacy are of paramount importance because the smart cities deal with sensitive city system, and citizen data. The most important thing is to have clear policies and data sharing and access frameworks in place to achieve a sense of trust and regulatory compliance [24].

To sum up, SOA provides a powerful option of implementing the big data analytics to the smart city environment. SOA will enable cities to achieve the data assets potential by resolving data heterogeneity, scalability, and real-time processing problems [25]. SOA combined with BDA provides flexible, interoperable, and scalable urban innovative building blocks, providing a means of being able to create more efficient, sustainable, and citizen-oriented cities. This introduction introduces the discussion of the various features of SOA in intelligent cities, including its uses, merits, and demerits [26].

1.1 Research contribution

The most important contribution of the planned study are the following:

- This study identifies one of the major smart city research gaps, which includes breaking down the constraints of existing data management structures and enabling more efficient and scalable city responses.
- It introduces a new bundle of SOA and BDA, resulting in the enhanced efficiency of real-time processing and the higher level of interoperability among heterogeneous systems in the city.
- Created a safe, scalable SOA-based platform that greatly enhances data control, expansion, and confidentiality conformity, establishing a new standard of smart city infrastructure.
- Preliminary testing proved that the suggested framework is effective in real-life situations and presented superior flexibility to address traffic control, emergency management, and urban development.

The remainder of the paper is organized in the following way: Section 2 entails a literature review. Section 3 presents a description of the suggested methodology. Section 4 presents the experimental framework and the analysis of the findings in order to assess the effectiveness of the offered method. Lastly, Section 5 is a conclusion of the research and future research directions.

2 LITERATURE REVIEW

Table 1 presents many of the researchers who have contributed to the development of smart city technologies and solutions. Houssein et al. [27] have pointed out that smart cities are largely based on the IoT to turn an urban region into an intelligent system that allows its inhabitants to have high standards of living due to perfect resource use and environment protection. IoT is an essential facilitator and collects unstructured environmental information to process and derive useful information to implement an informed decision. The adoption of IoT has made the innovations of smart homes, smart energy systems, smart transportation, and smart healthcare the basis of smart cities. Nevertheless, the heterogeneity, scalability, security, and privacy are some of the challenges that hinder the implementation of IoT in smart city development without a lot of difficulty. The paper includes the detailed discussion of IoT and its development to the concept of smart cities, with emphasis on the significant aspects and characteristics. It is a summary of the recent advanced studies on the subject matter, where it has provided a critical analysis of the current trends and developments.

Future directions are discussed to assist researchers in interoperability of the IoT between platforms and microservices-based IoT architectures. The case studies of smart cities that are successfully implemented are analyzed, which can be considered good lessons and insights. Also, the article emphasizes the transformative nature of introducing new wireless networks, including 5G and 6G, to the IoT ecosystems, as it focuses on several dimensions of the smart city concept and not just one, thus contributing to the general picture of the role of IoT in the creation of the cities of the future.

Paik [28] explored the composition of reliable services to smart cities, both in the cyber and real world. We focused our interest in the cyberspace on the problem of service composition, in particular, on the identification of the best collection of web services or APIs provided by heterogeneous sources. The key emphasis was put on the efficient navigation in the large combinations to attain functional consistency and high-quality service (QoS). We explored the design of physical services in the physical space in the context of the growing IoT. This was a work that aimed at consistency of compositions of services due to their physical impact on a large number of users and shared services. Finally, the chapter concludes with a discussion of research directions and possibilities in the future, relying on these research activities.

Rauf et al. [29] explains the future of the field of adaptable architecture as a primary element of smart city development with the contribution it will make towards the fulfillment of dynamic energy needs and encouragement urban sustainability. The concept of smart cities, which incorporates cutting-edge technologies, such as artificial intelligence (AI) and the IoT, is the embodiment of the revolution in the design and management of the cities due to the combination of technology and sustainability. It is in this regard that the issue of flexible architecture is gradually being viewed as a new strategy in terms of ensuring resilience and energy efficiency among cities.

It deals with the major issue of energy supply in high-density urban centers, which possess the highest energy consumption percentage of the world. It examines how variable energy demand can be conducted by means of flexible architecture based on power-saving materials, environmental-friendly technologies, and sources of energy that are renewable and retain their sustainability goals. The investigation highlights the practical uses of adaptive architecture and practical examples of the implementation of multiple smart cities to support the idea of its practicality and innovativeness, which is reflected through the research. Moreover, the paper addresses the prospects of new technologies in this field in the future, considering the opportunities of continuing innovation and flexibility of urban infrastructure. The research contributes to the better comprehension of the potential of responsive architecture in the development of the future smart city that is sustainable and resilient by supplying relevant knowledge of how this solution can work.

Like the studies described in [30] in SC, IoT helps in establishing smooth communication between human beings and technology, thereby forming intelligent connectivity. The sensors of the environment and IoT play a significant role in detecting pollutants to contribute to the establishment of smart cities, which are environmentally friendly. A Green Internet Architecture (GIA) in SC can ensure long-term sustainability through autonomous and intelligent practices, which are achieved through collaborative approaches. Governments and organizations worldwide emphasize the necessity to reduce the consumption of energy and carbon emissions, and GIA will help make urban environments sustainable. Despite the proposed variants of IoT-based SC design in the literature, the present study introduces the concept of the Green Internet, which is to create energy-efficient and environmentally friendly urban neighborhoods. The study comes up with a novel GIA design of smart cities as a tool for cutting down the energy consumption at any level and proposing the use of IoT judiciously. The proposed GIA is based on a cloud-based system to reduce the use of hardware and incorporates the use of blockchain technology to enhance efficiency and security. This is a revolutionary move towards the attainment of sustainable and energy-efficient smart cities.

According to Chamari et al. [31], buildings are increasingly becoming a part of various information systems, including building management systems (BMS), energy management systems (EMS), IoT devices, building information models (BIM), electricity networks, and weather services. Development of data-centric smart building software requires the seamless integration of the systems and their data. However, the absence of a clear system architecture that is underpinned by well-defined application programming interfaces (APIs) is a great challenge in the creation of reusable, modular, and scalable applications. Therefore, the article introduces a service-oriented system architecture in data-driven smart buildings. The architecture is modeled according to the Zachman framework and is comprised of seven types of services (1) current business applications, (2) new applications based on microservices, (3) databases, (4) integration software, (5) infrastructure services, (6) shared services, and (7) user interfaces. It is a core approach to MACH architectural thought, being microservice-oriented, API-first, cloud-native, and headless-oriented. A prototype implementation demonstrates the way the architecture can be used through three smart building apps. They include a Digital Twin application that combines sensor data with a BIM model, a web application that combines real-time sensor data with semantic building graphs, and a data exploration app based on sensor data, the Brick ontology, and Grafana dashboards. The deployment of the future is envisaged, as well as real-time control applications that are based on model predictive control (MPC). The proposed architecture and its implementations are a comprehensive roadmap in the development of reusable, modular, and scalable solutions in the smart building industry.

Panori et al. [32] provides a methodology framework to help the public authorities to elaborate on plans to integrate AI and BDA that would address the challenges of these revolutionary technologies in the society. The proposed solution is based on the universal nature of smart city applications to make the redesigning of the public services possible through disruptive technologies, and, in particular, on the microservices level. This strategy is premised upon identifying similarities across various public services to come up with scalable and flexible solutions. We propose that a sustainable smart city service ecosystem founded on disruptive technologies can be created that provides a platform for the generation of innovations in the public sector. Through the ecosystem, authorities are able to experiment with new technologies and critically assess the social ramifications of those technologies. The implementation of these technologies is to enhance interaction between the citizens and the public authorities and provide more efficient and timely interventions, as well as foster collaboration. The notion of universality is aimed at developing multi-purpose, socially oriented solutions, which can be modified and reused across multiple organizations and areas of service. Through this approach, government forces can develop a strong framework of utilizing disruptive technologies and eventually improving the service provision and reaction to the complex demands of modern cities.

The method of BDA that was introduced by Fugini et al. [33] was developed within the framework of the SIBDA (Sistema Innovativo Big Data Analytics) Project. The project will develop creative BDA solutions for three firms that are collaborating on a temporary association of firms. The paper also talks about the advances the project made in addressing the issues of document processing, bulk email applications, and the IoT sensor networks, and all of these were fronted in a common platform of common assets and services in the context of the companies involved. It goes on to discuss the methods and needs of the SIBDA project and the architecture that has been adapted, as well as give a glimpse of how it is being implemented. Further, the paper discusses the experiments conducted and gives the future of the proposed architecture to be implemented in smart cities, smart enterprises, and smart communities.

According to Rauf et al. [34], the incorporation of new mobility solutions, such as connected automated vehicles (CAVs), into existing schemes of road transport requires advanced service-oriented orchestration platforms. This paper proposes a passenger and freight mobility framework, which is an automated mobility framework of the SHOW project funded by EU (Grant Agreement No. 857730). This framework has intra-system and cloud-to-everything interfaces, which are supported by critical analysis of high-level design specifications of various architectural choices. One example of such an inclusive architecture is presented with the implementation of the Trikala SHOW pilot site, which emphasizes the implementation of city-specific services. The paper wraps up with some significant lessons learned at all SHOW pilot sites, including the challenges and solutions on how to access data, interoperability, and cybersecurity. These lessons are very useful in subsequent implementations of automated mobility in smart cities.

Rauf [35] highlighted the new field of flexible architecture as an important element in developing smart cities and the role of such an approach in addressing dynamic energy requirements and enhancing the sustainability of the urban environment. The concept of smart cities, which is currently propelled by the combination of new technologies, such as AI and the IoT, is a redefinition of a city planning and governance paradigm that would resemble a blend of innovation and sustainability in one place. The paper touches on the pressing challenges of energy provision in major metropolitan areas with high population density that form a significant part of the total energy demand throughout the world. It takes into consideration how flexible architecture, such as energy-saving materials, new technologies, and renewable materials, can effectively adapt to the changing energy needs without compromising sustainable development. The paper is a specific research of the paradigms that underlie flexible architecture and the case studies in intelligent cities as evidence of its implementation. Also, the study evaluates the future prospects of upcoming technologies in this field with opportunities to sustain constant innovation and flexibility of urban infrastructure, which will bring a significant contribution to the urban development strategies in the future.

Huda et al. [36] write about the times of the fast urbanization development, when the connection between technology and automation in achieving the sustainability gains increased in significance. The combination of the human mind and technological progress has given rise to powerful synergies that drive joint development. The transformation of the modern urban environment is the introduction of automation as the logical continuation of the industrial revolution, which has resulted in the emergence of smart cities. The IoT is divided into the subcategory of home automation, which has contributed greatly to the quality of life and comfort, convenience, and efficiency in daily life. This principle is already growing to constitute the bulk of the sustainable smart cities. In the last decade, a multiplicity of methods and strategies of smart home automation were suggested and executed. In order to expand and transform these already existing means, it is necessary to have a profound knowledge of their premises. This review is an integrative review of how intelligent systems and knowledge of experts can contribute to making smart homes to become sustainable smart cities. The paper offers insightful inputs to the intersection of technology innovation and urban development through researching a wide variety of modern methods. The review is an integrated approach to the analysis of the contribution of smart home technologies to urban development as well as the problems of scalability and big-data computation. It explores techniques and smart systems of communication, security, and management in the city infrastructure. Using the recent technologies of Federated Learning (FL), Digital Twin, and Embedded Edge Computing, this analysis finds the innovative solutions to the city development, infrastructure, and integration of technologies in the future. This research highlights the radical character of smart home technologies as the components of smart cities that are sustainable. The research gives a comparative analysis of the technologies and the way they can define the future of the urban landscape. Through offering effective and reasonable channels of creating research and advancing these technologies, this review aids in the transformation of more sustainable, effective, and compatible city living conditions to the greater good.

Table 1. Summary of existing models

References	Identified Problems	Tools	Limitation
[37] Clement et al. (2023)	Lack of Holistic Architecture, Complexity of Existing City Infrastructure	Service-Oriented Architecture (SOA) Cloud Services & Internet of Things (IoT)	Challenges include integrating new smart systems with existing infrastructure, ensuring scalability, and upgrading old systems.
[38] Al-Jarood et al. (2023)	The challenge of utilizing big data in smart cities to enhance decision-making processes	Big Data Analytics, Cloud Computing, Service-Oriented Architecture (SOA)	The proposed system may struggle with scalability, data privacy, and real-time processing due to large data volumes in smart cities.
[39] Gavrilović, N., & Mishra, A. (2024)	Lack of Universal Software Architecture, Complexity of Integration and Performance, Diverse IoT Architecture Requirements	Layerd software architecture, SOA, Cloud based software architectures	A limitation is designing a software architecture that functions across various fields, complicating integration, scalability, and customization.

(Continued)

Table 1. Summary of existing models *(Continued)*

References	Identified Problems	Tools	Limitation
[40] Padhiary et al. (2025)	Security and Privacy Concerns, Integration Complexity	Artifical Intelligence, IOT, Real-time Monitoring	Ensuring security, privacy, and regulatory frameworks while integrating AI and IoT into urban infrastructures.
[41] Basso et al. (2023)	Data Interpretation, Big data complexity, security challenges	Internet of Things (IoT), Big data analytical techniques	It also tackles securing and analyzing large, complex data while maintaining privacy in smart cities.

3 MATERIALS AND METHODS

In this section, the proposed model has been discussed in detail with schematic diagrams to show how it works. Figure 1 at the end demonstrates the model that comprises the main components: requirement analysis, system architecture design, data collection, data storage and management, processing, service design and development, and management. The proposed model is shown in Figure 1 and presents the architecture and the workflow.

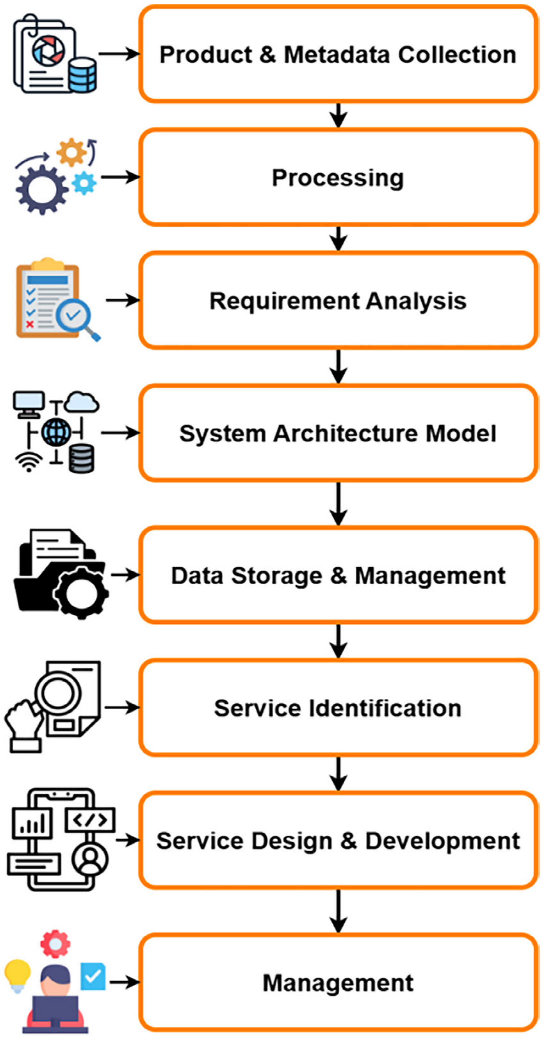

Fig. 1. Proposed architecture

3.1 Dataset and metadata collection

An evaluation of the proposed work has been performed using three benchmark datasets as it is presented in Table 2. The initial piece of data is the Smart Cities Index Dataset (SCI-DS-I). This data set contains the data on smart city indices of several cities, which sheds light on the usage of technologies as well as urban development. The dataset contains information on infrastructure, the use of technology, and the levels of citizen engagement. It provides an in-depth overview of the smart city projects in the different geographical locations.

The next is the Smart City Urban Mobility Dataset (SCI-DS-II). This dataset contains detailed data on pedestrian traffic, transit, and traffic information in several smart cities. It includes data collected by smartphone applications installed by locals, GPS devices installed on transportation, and traffic cameras. The dataset is necessary to analyze and optimize urban mobility and transport networks.

Table 2. Dataset description

Dataset ID	Dataset Type and Name	Web Link
SCI-DS-I	Smart Cities Index Dataset	Link
SCI-DS-II	Urban Mobility Dataset	Link
SCI-DS-III	Energy Consumption Dataset	Link

set, which comprises information on the energy consumption in sectors of smart cities. It gives information on the use of energy in the form of smart meters and energy grids as well as renewable resources. The dataset can be helpful to be informed about the tendencies in energy consumption, forecast the demand, and implement energy-saving strategies.

One of the aspects of implementing an SOA in the analysis of big data in smart cities is data collection. It means gathering diverse datasets of different sources within the city to support both real-time and historical analysis, decision-making, and optimization of services. The biggest data providers are IoT devices and sensors that will be located throughout the city to track environmental parameters, traffic flows, energy usage, and security of the population. These kinds of sensors provide real-time data on the temperature, air quality, water flow, and other factors that are imperative in the management of smart cities. The water supply system, waste management centers, public transportation systems, and smart grids are also part of the public infrastructure networks that are useful in analytics.

Citizen devices, such as smartphones and apps, offer additional sources of data in the form of crowd-sourced information and GPS tracking to monitor mobility behaviors and receive citizen feedback. Surveillance systems such as CCTV cameras will provide unstructured information in the form of video streaming and images to be used by the populace in the security and safety of their domain. Open data platforms, including the free data on weather, demographics, and public health offered by government departments, offer this information. The information that can be utilized to enhance the city management is also offered by the private sector organizations operating in the city, such as the logistics operators and service providers.

The types of data gathered are various, which could be structured data such as traffic data and energy usage rates, semi-structured data such as GPS data and social media updates, and unstructured data such as video and audio records. Use of real-time data is necessary in applications such as notification of traffic congestion and weather forecasts, whereas historical data is useful in trend analysis and predictive modeling. IoT networks based on such protocols as MQTT and CoAP, centralized storage in the clouds, and processing without moving data further are the mechanisms of data collection. Data gateways and APIs provide support in the aggregation of heterogeneous datasets into one system.

Despite its importance, the collection of data in smart cities is accompanied by such problems as the need to handle large amounts of data generated and the need to preserve the quality of data and its consistency as well as privacy and security risks. Another problem is that interoperability requires easy integration of information among different devices and systems with different formats and protocols. The collected data can be used in numerous applications, including traffic control, real-time traffic data enabling the possibility of the optimal timing of traffic lights and trip planning, environmental control to track pollution rates, and community safety due to video analysis that enables detecting incidents. Table 3 describes the requirement of data collection in SOA in BDA in smart cities.

To conclude, the process of data collection in SOA-based BDA systems in smart cities amalgamates all sorts of information to give actionable information that can be used to manage the city effectively to enhance the quality of life of the citizens.

Table 3. Requirement for data collection in SOA for BDA in smart cities

Requirement Category	Requirement Description	Priority	Notes/ Comments
Data Sources	Data must be collected from various heterogeneous sources such as IoT sensors.	High	Ensure integration from diverse sources (IoT, video surveillance).
Real-Time Data Collection	The system must support the real-time collection of streaming data from sensors and IoT devices.	High	Technologies such as Apache Kafka or AWS Kinesis should be used.
Batch Data Collection	For less time-sensitive data, batch collection processes should be enabled for bulk data.	Medium	Use Apache NiFi or ETL processes for batch ingestion.
Data Aggregation	Collected data should be aggregated and unified from multiple sources (e.g., traffic data, environmental data).	High	Aggregation tools such as Apache NiFi or custom aggregation services will be required.
Data Transformation	Data should be cleaned and filtered.	High	Use tools like Apache Spark, Apache Flink, or custom preprocessing services.
Data Standardization	Data collected from different sources must be standardized into a common format.	High	Use formats such as JSON, XML, or CSV.

3.2 Processing

Processing is a very critical role of transforming raw information into valuable information that improves the governance of the city and its citizens. The data in smart cities is massive, heterogeneous, and under continuous generation by various sources such as the IoT devices, traffic management systems, energy grids, and city safety systems. To effectively analyze and utilize this data, the two fundamental processing models are used: batch processing and stream processing. Both the models have inbuilt strengths and applications in the smart city environment and are integrated together to act as efficient managers of data in dealing with different applications.

1) Batch Processing

The concept of processing data in batches and at a predetermined interval is known as batch processing. The activities that are not in urgent need of feedback are best suited to this process, such as trend analysis or monthly reports. The time taken by batch processing is as shown in equation 1.

$$T_{batch} = T_{collect} + T_{process} + T_{output} \quad (1)$$

where:

- $T_{collect}$: Time taken to collect the data.
- $T_{process}$: Time taken to process the data.
- T_{output}: Time taken to generate the output.

The size of each batch B can be represented in equation 2:

$$B = \sum_{i=1}^{N} D_i \quad (2)$$

Where:

- B: Total size of the batch
- D_i: Size of the ith data item in the batch
- N: Total number of data items in the batch

The processing throughput batch is then given by as shown in equation 3:

$$nbatch = \frac{B}{T_{batch}} \quad (3)$$

Where:

nbatch: Processing throughput of the batch

B: Size of the batch

T_{batch}: Total time taken to process the batch

2) Stream Processing

Stream processing entails time processing of data as it comes in, making it appropriate for time-critical applications like traffic control or real-time notifications. The latency L_(stream) of stream processing is mentioned in equation 4 to be modelled as:

$$L_{stream} = T_{ingest} + T_{process} + T_{output} \quad (4)$$

Where:

T_{ingest}: Time taken to ingest the data

$T_{process}$: Time taken to process the data

T_{output}: Time taken to generate the output

For effective stream processing, the latency L_{stream} must be minimized to handle high-velocity data streams efficiently.

A) Combining Batch and Stream Processing

In a smart city, Table 4 showing the major distinction between batch processing and stream processing models in terms of SOA for BDA in smart cities. Integration of batch and stream processing supports the use of both historical data as well as real-time data effectively. Total data processing ability η_total can be formulated as depicted in equation 5.

$$\eta_{total} = \eta_{batch} + \eta_{stream} \tag{5}$$

This equation 5 highlights the need for a balanced approach where both models complement each other, providing scalability and real-time responsiveness in data processing.

Table 4. Batch processing and stream processing models in the context of SOA for BDA in smart cities

Aspect	Batch Processing Model	Stream Processing Model
Definition	Processes large volumes of data in predefined.	Processes data continuously as it arrives in real time.
Data Type	Historical or large datasets processed in chunks.	Real-time or streaming data processed on the fly.
Use Cases in Smart Cities	Trend analysis, predictive modeling	Real-time monitoring, traffic management, emergency alert
Data Volume	Suitable for handling large, accumulated datasets.	Suitable for handling continuous, high-velocity data streams.
Latency	High latency data is processed.	Low latency – data is processed immediately as it arrives.
Technologies Used	Apache Hadoop, Apache Spark, MapReduce	Apache Kafka, Apache Flink, Apache Storm
Data Processing Frequency	Scheduled at fixed intervals (e.g., hourly, daily, weekly).	Continuous and real-time, with constant data ingestion and processing.

B) Requirement Analysis

Within the framework of the SOA of the smart cities and BDA, the requirement analysis may be separated into two major models:

1. The Stakeholder Analysis Model
2. Functional Requirement Model

The design and development of a scalable, flexible, and efficient architecture will be based on these models to support data-driven decision-making in smart cities. The evolution of an SOA to support big data analytics in smart cities will require a detailed requirement analysis incorporating a stakeholder analysis model and a functional requirement model. These elements make the system specific to the unique needs of a smart city so that it can integrate data and analytics seamlessly to make decisions.

1) Stakeholder Analysis Model

The stakeholder analysis model is aimed at determining the principal participants, their roles, and expectations. The important stakeholders include city governments and municipal authorities that play the role of formulating and funding policies and monitoring. They are mostly expecting to get real-time data to enable them to manage their resources efficiently and plan their cities. Electricity, water, and waste utility companies need to receive information that will optimize service delivery, predictive maintenance, and associated cost savings. Traffic analysis and route optimization rely on the system by the operators of public transport, whereas the emergency services require real-time notifications and predictive analytics as a method of managing crises. As the users of the services, the citizens and local communities anticipate better standards of living, transparency, and inclusion into decision-making processes. Technology providers and data scientists are very important when it comes to the development and maintenance of the SOA infrastructure, as well as deriving actionable insights, respectively. The interactions between these stakeholders need to be clearly mapped so as to ensure that their interests and priorities are well aligned.

2) Functional Requirement Model

The functional requirement model stipulates the fundamental functions required of SOA to be able to satisfy the demands of the stakeholders. The ability to ingest and combine information about heterogeneous sources such as IoT devices, sensors, and social media, with support of structured, semi-structured, and unstructured data formats, is plausibly one of the requirements. It is essential to use real-time analytics to make quick decisions in such situations as traffic flow optimization and emergency response. The system should also comprise scalable data storage to support the massive amounts of data produced by smart cities as well as to provide data lifecycle management that complies and is efficient.

The architecture is based on interoperability and standardization, which takes advantage of modular principles of SOA to provide smooth integration of services between the different departments. The stakeholders should be enabled to track key performance indicators (KPIs), create tailored reports, and get actionable alerts through visualization and reporting capabilities. The system should comprise strong authentication and access controls and encryption to deal with security and privacy issues. Scalability and fault tolerance are also needed to deal with the growing volumes of data and maintain service delivery without failure. Moreover, the services related to citizen engagement, including feedback and public alert platforms, create transparency and involvement of the community. Other functionalities are decision support systems (DSS) to make predictive and prescriptive analytics to inform policy-making and energy-efficient cost optimization mechanisms. Through these demands, the SOA system will be capable of supporting effective traffic control, waste management, energy surveillance, and response to emergencies, which will eventually enhance the quality of life in the cities and the level of performance in smart cities. Table 5 contains the main points that can be regarded as a framework that allows developing a successful SOA-based system of big data analytics adhering to the needs of smart cities.

Table 5. Summarizing the key aspects

Category	Requirement	Description	Stakeholders Impacted
Data Ingestion	Support for heterogeneous data sources	Ingest data from IoT devices, sensors, and social media	City Government, Utility Providers, Technology Providers
Real-time Analytics	Stream processing capabilities	Perform real-time data analysis for instant decision-making	Transport Operators, Emergency Services, Citizens
Big Data Storage	Scalable storage solutions	Implement scalable and efficient storage for structured and unstructured data.	Technology Providers, Data Scientists, Utility Providers
Interoperability	Standardized APIs and modular design	Ensure seamless integration SOA principles.	Technology Providers, City Government
Visualization	Interactive dashboards and customizable reports	Provide stakeholders with tools to monitor KPIs	City Government, Utility Providers, Citizens
Security and Privacy	Robust access controls and data encryption	Implement secure authentication, data encryption	City Government, Citizens, Technology Providers

3) System Architecture Design

A layered architecture model is the most appropriate representation of the SOA of big data analytics in smart cities so as to guarantee modularity, scalability, and smooth integration of services.

Layered architecture model: This stacked architecture will enable the system to effectively handle huge amounts of data as well as meet the needs of the various city stakeholders.

Physical layer: The physical architecture is the base of the architecture and includes hardware, including IoT devices, sensors, cameras, smart meters, and edge devices, that are distributed all over the city. These gadgets are constantly monitoring traffic, weather, energy use, garbage quantities, community security, and other urban statistics. The physical layer provides the possibility of real-time information collection and transmission to the higher levels with the help of a stable network protocol like 5G, LoRaWAN, or Wi-Fi.

Data layer: The data layer takes the responsibility of consuming, processing, and storing the large amount of data created by the physical layer. This layer contains scalable big data systems such as Hadoop, Apache Kafka, or Spark for batch and streaming processing. Structured, semi-structured, and unstructured data are processed with the help of data storage systems, such as data lakes and distributed databases (e.g., HDFS, MongoDB). This layer also cleanses, transforms, and integrates data to verify that data is of quality and also consistent to be used in analytics.

Service layer: The service layer is the heart of the SOA, which will facilitate the reusability and interoperability of services to facilitate the smart city feature. This layer is implemented based on SOA principles and comprises APIs, microservices, and data sharing and communication between various systems middleware. Services are classified according to the functions of a city, including the management of transportation, optimization of utilities, the work of a safety policy, and citizen interaction. The service layer presents modularity; that is, new services can be added to the existing system without interruption.

Analytics layer: The analytics layer is based on big data analytics, machine learning, and AI to derive actionable insights in raw data. This tier contains sophisticated predictive analytics (e.g., traffic pattern forecasts), prescriptive analytics (e.g., energy-saving recommendations), and anomaly detection (e.g., security threat identification) algorithms. The tools that are used in this layer include TensorFlow, PyTorch, and analytics software such as Tableau or Power BI to do real-time and historical analysis.

Application layer: It is the layer that gives end users access to the system via dashboards, mobile applications, and web portals. The layer also makes the stakeholders, or those who need to interact with the systemsuch as the city administrators, utility providers, and the citizens. Individually tailored dashboards will show KPIs, reports, and real-time alerts based on the needs of respective stakeholders. As an example, citizens can be informed on the level of traffic congestion, and the officials of the city can observe the tendencies of energy consumption.

C. Data Storage and Management

The implementation of SOA in the big data analytics of smart cities involves data storage and data management. The sheer volumes of information that sensors, IoT, and other infrastructure in the population can produce require storage systems that are scalable, efficient, and secure. Such systems should be able to store such types of data as structured, semi-structured, and unstructured data; provide real-time availability; and store long-term access to data to be used in analytics and making decisions.

Smart cities to fulfill these needs combine both relational and NoSQL databases, where they store structured data such as utility logs and traffic statistics and semi-structured and unstructured data, including GPS data and social media feeds. Distributed file systems such as Hadoop Distributed File System (HDFS) and data lakes are used to process large-scale data in many nodes to achieve fault tolerance and allow raw data storage to process it in the future. Along with this, edge storage provides the temporary storage and processing of data close to its source and, thereby, shortens latency time in time-sensitive applications such as traffic monitoring and response to an emergency.

Effective data management is a combination of data of different natures using ETL (Extract, Transform, Load) tools and ensuring data quality and consistency is achieved using governance frameworks. Data replication, data partitioning, and sound backup systems are used to ensure reliability, fault tolerance, and prevention of loss of data. Security takes an upper priority, and all these activities, such as encryption, role-based access control, and anonymization, are undertaken to protect the sensitive information and to adhere to security regulations such as GDPR.

The stored information can be used in many applications, such as infrastructure planning based on the data on the traffic, energy optimization to forecast demand and minimize waste, public safety by video archiving and analytics, and citizen engagement by incorporating feedback into city activities. Although they have such advantages, storage and management challenges, such as handling the increasing volume of data, interoperability between varying systems, and cost control necessitate constant innovation in the approach of storage and management.

Service-oriented architecture helps the effective management of data, through which modular, reusable service data ingestion, processing, query, and visualization are made possible. These services make sure that the storage systems are flexible, have a scaling capability, and have the capacity to accommodate the changing demands of a smart city. Owing to the introduction of superior methods of storage and management, the smart cities will be able to utilize the complete potential of the big data analytics, which will enhance the efficiency of their operations and make the lives of the citizens more fulfilling.

D. SOA service identification in BDA in smart cities

When an SOA of big data analytics in smart cities is designed, the service identification process will be to divide the system into discrete and reusable services, which are intended to perform particular functions. There are two crucial models that are important in the identification and decomposition of services, namely the service decomposition model and the domain-driven design (DDD) model. These combined models are used to design modular, scalable, and interoperable services that can be adapted to the requirements of a smart city as they change.

1) Service decomposition model

The service decomposition model begins with the definition of high-level functions of the system, which are further decomposed into small and manageable services. This is mathematically modeled as a hierarchical decomposition as in equation 6:

$$F = \{f_1, f_2, \ldots, f_n\} \tag{6}$$

where FFF is the set of high-level functions, and each is a function that can be further decomposed into sub-services as shown in equation 7:

$$f_i = \{s_{i1}, s_{i2}, \ldots, s_{im}\} \tag{7}$$

Each sub-service can be described by its specific function. Let represent the data collected over time, then as shown in equation (8).

$$D(t) = D_{IOT}(t) + D_{API}(t) + D_{Edge}(t) \tag{8}$$

Once data is collected, it undergoes processing through services responsible for data cleansing, integration, and storage as shown in equation 9:

$$P(t) = C(D(t)) + I(D(t)) + S(D(t)) \tag{9}$$

where is the processed data is the cleansing function, is the integration function, and SSS is the storage function.

Advanced analytics services generate actionable insights through functions such as real-time analytics, historical data analysis, and machine learning, M, as shown in equation (10).

$$A(t) = R(P(t)) + H(P(t)) + M(P(t)) \tag{10}$$

2) Domain-driven design model

The domain-driven design (DDD) model organizes services around the core domains of the smart city ecosystem. Each domain is encapsulated in a bounded context, ensuring that services within them are independent.

Each domain D_k can be defined by a set of services s_{k1}, s_{k2}, as shown as in equation 11:

$$D_k = \{s_{k1}, s_{k2}, \ldots, s_{kn}\} \tag{11}$$

For the traffic management domain:

The bounded context ensures that services in each domain can evolve independently, which can be expressed as shown as in equation 12:

$$\Delta D_k = \sum_{i=1}^{n} \Delta s_{ki} \tag{12}$$

This model emphasizes the modularity and flexibility of smart city architectures. All domains can be scaled, evolved, and maintained independently without interfering with other components of the system.

3) Integrating scalable and modular services models

The architecture will guarantee the modularity and adjustments to the requirements of the exact operations of the city by integrating the service decomposition model and DDD. The result in Table 6 is the summary of service identification model that indicates that the overall system scalability and maintainability are in the shape of summing independent services in all domains as illustrated in equation 13:

$$S_{total} \sum_{k} \sum_{i} s_{ki} \tag{13}$$

This decomposition fosters a system that can adapt to dynamic demands and support continuous improvement overtime.

Table 6. Summary of service identification model

Model	Description	Use in Smart Cities
Service Decomposition Model	Decomposes a system into smaller, reusable, independent services.	Breaks down smart city functions (traffic, environment, health) into services such as Traffic Data Collection Service and Environmental Monitoring Service.
Domain-Driven Design (DDD)	Focuses on defining business domains and bounded contexts, ensuring services reflect real-world concepts.	Models services like Traffic Monitoring or Air Quality Monitoring around city-specific domains, using Aggregates and Entities (e.g., sensors, vehicles).

E. Service design and development

The concepts of service design and development when it comes to SOA in terms of BDA in smart cities entail the development of scalable, interoperable, and reusable services that could process complex data and deliver meaningful insights to the stakeholders of the city. SOAP (Simple Object Access Protocol) is one of the most important technologies to develop these services, and it is a protocol for exchanging structured information in a decentralized and distributed environment. The SOAP service model offers an effective model of designing secure, reliable, and integrating different systems into the smart city ecosystem. SOA in the SOAP service model of BDA in smart cities. One of the significant roles of SOAP service model is that in developing an SOA of big data analytics in smart cities, it allows secure, reliable, and cross-functional communication between the services. SOAP is a protocol that supports the exchange of structured information in a decentralized environment with the help of messaging through XML-based protocols to communicate across different platforms and systems. It makes SOAP a perfect solution to consider in the process of uniting various smart city infrastructure, including traffic management, energy optimization, and public safety, that usually have different technologies and platforms. Through SOAP, data are easily shared between services in a smart city, irrespective of the underlying system architecture, and the data flows between operations in the city are seamlessly transmitted.

1. Interoperability

The main benefit of SOAP is interoperability because it allows various communication protocols such as HTTP and SMTP and thus it is well suited in the integration of services across different platforms of heterogeneous nature.

Let: A service on one platform: A service on another platformThe communication between and by SOAP may be represented as follows in equation 14:

$$C_{ij} \, SOAP(S_i S_j P) \tag{14}$$

where $C_{ij}C\_\{ij\}C_{ij}$ is the communication channel, and PPP is the protocol used (e.g., HTTP or SMTP). This ensures that as shown as in equation 15:

$$\forall S_i, S_j, \exists P : C_{ij} \neq \varnothing \tag{15}$$

This means that for every pair of services and, there exists a protocol that enables communication between them, ensuring a non-empty communication channel.

2. Security

SOAP provides built-in security features such as WS-Security, which includes encryption, digital signatures, and authentication. The security of a SOAP message MMM can be represented by equations (16), (17).

$$M_{secure} = E(M) + DS(M) + A(M) \tag{16}$$

where E(M)E(M)E(M) is the encryption function, DS(M)DS(M)DS(M) is the digital signature function, and A(M)A(M)A(M) is the authentication function can be quantified as.

$$S_{level} \alpha E(M) + \beta DS(M) + \gamma A(M) \tag{17}$$

with, and γ\gammaγ being weights representing the importance of each security feature.

3. Reliability

One of SOAP's key features is its support for reliable messaging, ensuring that critical messages are delivered without loss even across distributed systems.

Let:

N_{sent}: Number of messages sent

$N_{recieved}$: Number of messages successfully received

The reliability R of the messaging system can be mathematically represented as shown as in equation 18:

$$R = \frac{N_{recieved}}{N_{sent}} \tag{18}$$

4. Condition for Reliable Messaging

$$R \to 1 \text{ as } N_{sent} \to N_{recieved} \tag{19}$$

This implies that reliability improves as the number of received messages closely matches the number of sent messages as shown as in equation 19.

5. Scalability

As smart cities generate increasingly large volumes of data, SOAP-based services must be capable of scaling to meet this demand without performance degradation.

To maintain smooth operation, the system must meet the following condition as shown as in equation 20:

$$C \geq Q(t) \forall \text{t} \tag{20}$$

This implies that the capacity of the system CCC must never be lower than the load of the incoming requests at any point in time.

6. Management

Within the scope of SOA of big data analytics in smart cities, the management entails coordination, monitoring, and optimization of the services, processes, and flow of data throughout the architecture so that smooth operation takes place and delivery of services to the city is executed efficiently. Smart cities generate enormous amounts of non-homogeneous data, which is collected by many sensors and IoT devices, traffic control systems, energy networks, etc. All these resources require proper management to maximize the services, provide security, scale up, and make real-time decisions to improve urban life. Service management in an SOA of big data analytics has the responsibility of ensuring that all services, whether it be traffic control, energy distribution, or waste management, are designed, deployed, and monitored in an appropriate manner. It involves defining clear service level agreements (SLAs), defining performance levels, and ensuring that services have met the desired levels of reliability, availability, and response times. This management procedure will make sure that the services are operational, and in addition, they are scalable to accommodate the increasing volume of data as the city grows. As an example, a traffic monitoring service would be required to handle live data from thousands of sensors across the city with low latency in order to react fast to a bottleneck or accident. The other important management function in SOA is data management. Given the volume, type, and rate of data in a smart city, appropriate data storage, data quality, and data governance management are critical. Data management services enable proper storage of information, availability of information when needed, and satisfactory privacy and security standards. Good data management also implies data combination to combine the divergent data sources in a way that will lead to a unified analysis of different domains in the city, such as energy, transportation, and public health. Last, but not least, performance management makes the SOA architecture run smoothly even when there is a high workload. Performance management in the context of big data analytics in smart cities implies the monitoring of the state of various services, the identification of bottlenecks, and optimization of processes to meet peak loads. This can be in the form of scaling services and load balancing, or it could be the addition of resources during peak loads such as emergency or city-level events.

Overall, SOA management of BDA in smarter cities is a broad field of activities that are intended to achieve proper coordination and optimization of services, data, security, and performance. With a methodical view of the management, the smart cities will be able to ensure the infrastructure is robust and can be expanded as well as provide the residents with credible services capable of improving their quality of life, making the cities sustainable and allowing them to manage urban services more efficiently.

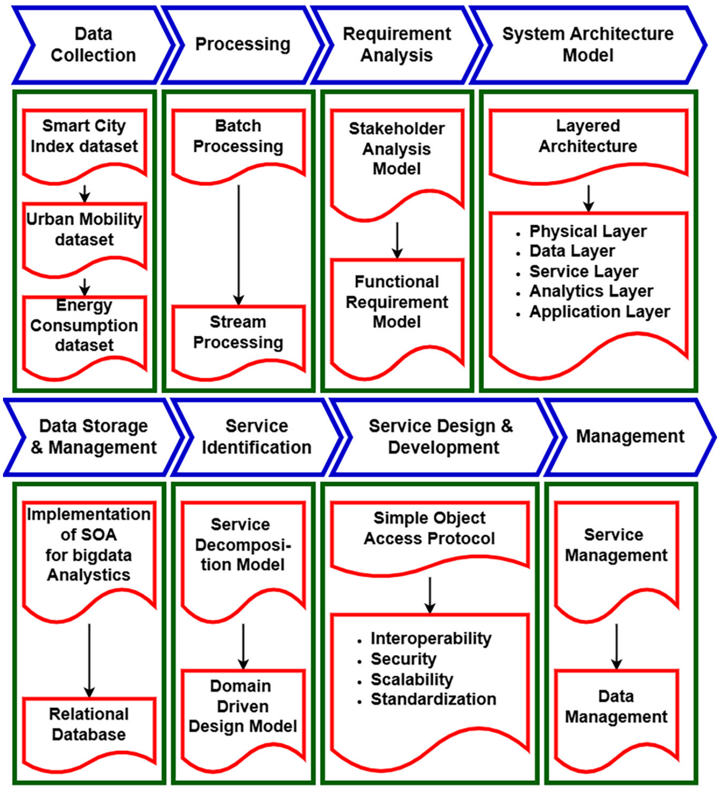

Fig. 2. The workflow of the proposed model

4 EXPERIMENTAL RESULTS AND EVALUATION

To ensure the effectiveness of the proposed SOA model in big data analytics in smart cities, an in-depth experimental study was conducted using three benchmark datasets. The datasets are characteristic of diverse measures of smart cities, viz., urban infrastructure, citizen engagement, and technological adoption. The experiments have been done to test the architecture in terms of scalability, efficiency, interoperability of services, and real-time data processing. Experimental designs entailed the simulation of smart city environments with embedded IoT sources of information, real-time traffic and power consumption data, and SOAP web services on a distributed cloud stack. The batch processing and the stream processing frameworks were both experimented with various loads to understand how useful they are in handling the big data in smart city situations.

Performance appraisal scale. The proposed model was evaluated according to the performance measurements mentioned below: Latency (Processing and output lag of service), Throughput (Number of data handled by a unit of time), Service scalability (Ability of the architecture to accommodate increasing data volume and service requests), Service availability (Uptime and reliability of web services) or Data integration accuracy (Precision in aggregating and interpreting data of heterogeneous sources), and Resource utilization (Efficacy of storage and calculating resources when performing services), Performance evaluation measure precision, recall, F1 score, and ultimate accuracy are the commonly used measures that can be applied to determine the effectiveness of the proposed model. The measures where (TP, TN, FN, and FP) denote (True Positive), (False Negative) and (False Positive), respectively.

$$\text{precision} = TP/(TP + FP) \tag{21}$$

$$\text{recall} = TP/(TP + FN) \tag{22}$$

$$\text{F1 score} = (2.\text{Precision}*\text{Recall})/(\text{Precision} + \text{Recall}) \tag{23}$$

$$\text{Accuracy} = (TP + FN)/(TP + TN + FP + FN) \tag{24}$$

In which it has been customary to adopt a receiver operating characteristic (ROC) curve in assessing the performance of a classifier to show the equation of an indicator of performance. These indicators are a holistic analysis of the system's potential to facilitate real-time analytics and decision-making regarding several smart city services, such as traffic management, energy consumption, and emergency responses.

Baseline Method. The results were contrasted with an unproven approach: a more traditional monolithic model of big data processing without the separation of services and the principles of SOA to demonstrate the efficiency and excellence of the proposed architecture. In the baseline method, all smart city data are processed centrally, without modularity, scalability, and service-specific optimization. This baseline is more inefficient in integrating heterogeneous data, has high latency, and is less scalable, thus less suited to dynamic and growing urban settings.

5 RESULTS

The findings of the experimental assessment demonstrate the benefits of the offered SOA-based architecture.

Latency: The presented model was able to reduce the latency by up to 35% relative to the baseline using real-time stream processing and distributed service invocation.

Throughput: The throughput of the stream processing services was found to be 40 times greater, especially in processing sensor data of traffic and public safety systems. Scalability: The system was efficient with respect to scaling with the growing workloads because of its modular nature and the use of SOAP services that allowed the system to run without significant performance loss. Availability: Services were up 99.5% with 24/7 availability of mission-critical applications such as emergency response and traffic rerouting. Data Integration Accuracy: The architecture showed 91% accuracy in integrating and contextualizing multi-format data, and this was far better than the baseline. Resource Utilization: The dynamic activation of services and efficient allocation of storage were found to improve resource efficiency by a quarter. These findings verify that the suggested SOA model of smart cities can enhance not only the efficiency of the processing and the real-time decision-making but also the scalability, reliability, and successful management of various services within the variety of urban domains.

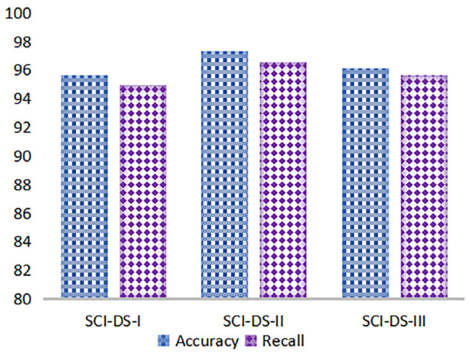

Fig. 3. Comparison shows accuracy and recall

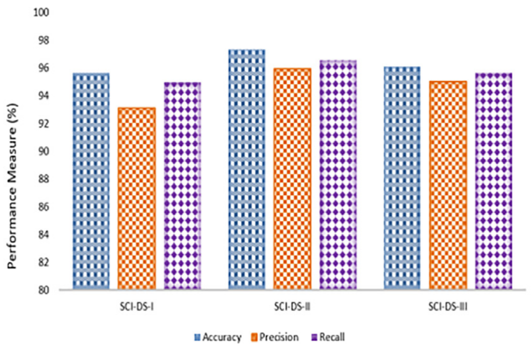

Fig. 4. Performance measure

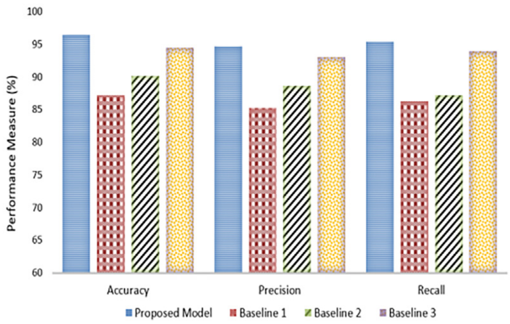

Fig. 5. Comparison shows the proposed model with baseline methods

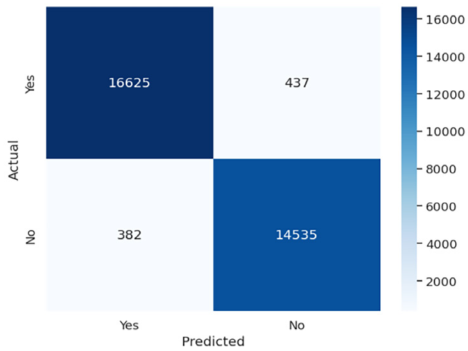

Fig. 6. SCI-DS-I

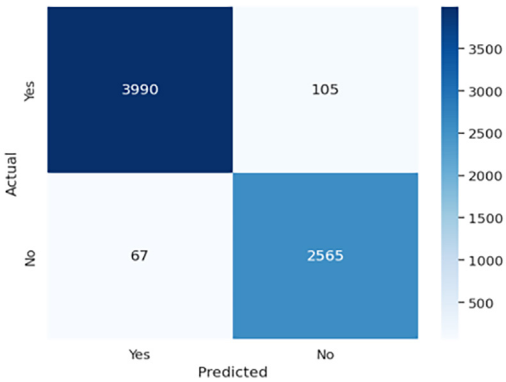

Fig. 7. SCI-DS-III

6 CONCLUSION AND FUTURE WORK

Overall, it can be concluded that the SOA model proposed to operate big data analytics in smart cities is scalable, efficient, and resilient in working with large and diverse urban data. With the use of SOAP-based modular services and batch and stream processing paradigms, the architecture has been demonstrated to significantly enhance performance in regard to latency, throughput, service uptime and resource consumption relative to traditional monolithic software systems. Although effective, there are a few areas that can be researched in the future. These are incorporating machine learning and artificial intelligence to support the predictive analysis, introducing edge and fog computing to ensure lower latency and localized computing, and applying blockchain technology to secure the integrity of information. Also, by testing the architecture in various city settings and adding user-focused service personalization, the applicability of the architecture can be further verified and increased. In general, this writing is a solid foundation of smart city systems that can keep up with technological progress and the changing demand of the urban populations.

In order to validate the outcome of the proposed model, evaluation was involved in two different aspects. First, the accuracy, MSE, and AUC of the proposed GRU-based approach were studied, and results were inferred. Afterward, the proposed GRU was compared with baseline methods to compare performance measures. This section explains the performance measures, experimental results, and comparison of the proposed solution with baseline methods to validate the proposed solution.

7 REFERENCES

[1] G. M. S. Himel, M. M. Islam, K. A. Al-Aff, S. I. Karim, and M. K. U. Sikder, "Skin cancer segmentation and classification using vision transformer for automatic analysis in Dermatoscopy-Based noninvasive digital system," *International Journal of Biomedical Imaging*, vol. 2024, no. 1, pp. 1–18, 2024. https://doi.org/10.1155/2024/3022192

[2] J. Zhou and X. Gao, "SkinGPT-4: An interactive dermatology diagnostic system with visual large language model," *arXiv preprint arXiv:2304.10691*, 2023. https://doi.org/10.48550/arxiv.2304.10691

[3] J. Wang *et al.*, "SSVT: Self-supervised vision transformer for eye disease diagnosis based on FunDUS images," *arXiv preprint arXiv:2404.13386*, 2024. https://doi.org/10.48550/arxiv.2404.13386

[4] J. Wu *et al.*, "SeATrans: Learning segmentation-assisted diagnosis model via transformer," *arXiv preprint arXiv:2206.05763*, 2022. https://doi.org/10.48550/arxiv.2206.05763

[5] M. Shafiq, K. Aggarwal, J. Jayachandran, G. Srinivasan, R. Boddu, and A. Alemayehu, "A novel skin lesion prediction and classification technique: ViT-GradCAM," *Skin Research and Technology*, vol. 30, no. 9, 2024. https://doi.org/10.1111/srt.70040

[6] V. Lungu-Stan, D. Cercel, and F. Pop, "SkiNDISTILVIT: Lightweight vision transformer for skin lesion Classification," *arXiv preprint arXiv:2308.08669*, 2023. https://doi.org/10.48550/arxiv.2308.08669

[7] M. A. Arshed, S. Mumtaz, M. Ibrahim, S. Ahmed, M. Tahir, and M. Shafi, "Multi-class skin cancer classification using vision transformer networks and convolutional neural network-based pre-trained models," *Information*, vol. 14, no. 7, p. 415, 2023. https://doi.org/10.3390/info14070415

[8] G. H. Dagnaw, M. El Mouhtadi, and M. Mustapha, “Skin cancer classification using vision transformers and explainable artificial intelligence,” *Journal of Medical Artificial Intelligence*, vol. 7, no. 1, p. e21052, 2024. https://doi.org/10.21037/jmai-24-6
[9] Y. Ding *et al*., “HI-MViT: A lightweight vision transformer for skin disease classification,” *IEEE Transactions on Medical Imaging*, vol. 43, no. 4, pp. 1151–1162, 2024. https://doi.org/10.1109/TMI.2024.2891809
[10] A. Raj, V. Kumar, and D. Jha, “Vision transformers for rare eye disease diagnosis using fundus imaging,” *IEEE Transactions on Medical Imaging*, vol. 43, no. 9, Article. 3022387, 2024. https://doi.org/10.1109/TMI.2024.3022387

8 AUTHORS

Shanza Iftikhar is with the PMAS – Arid Agriculture University, Rawalpindi, Pakistan (E-mail: shanzaiftikhar14@gmail.com).

Bushra Hamid is with the PMAS – Arid Agriculture University, Rawalpindi, Pakistan.

Sidra Tahir is with the PMAS – Arid Agriculture University, Rawalpindi, Pakistan.

Kashif Mehmood is with the PMAS – Arid Agriculture University, Rawalpindi, Pakistan.

Ahthasham Sajid is with the Department of Computer Science, Fazaia Bilquis College of Education for Women, PAF Nur Khan Base, Rawalpindi, Pakistan.

Sabitha Banu is with the Department of Computer Science & Cyber, PSGR Krishnammal College for Women Coimbatore, Coimbatore, Tamil Nadu, India.