Qubahan Academic Journal
گوڤارا قبهان یا ئەکادیمی
مجلة قبهان الأكادمية
Doi: 10.48161/issn.2709-8206

A Survey of Data Mining Implementation in Smart City Applications

1st Zainab Salih Ageed
Translation Dept.
Nawroz University
Duhok, Iraq
zainab.ageed@nawroz.edu.krd

2nd Subhi R. M. Zeebaree
Culture Center
Duhok Polytechnic University
Duhok, Iraq
subhi.rafeeq@dpu.edu.krd

3rd Mohammed A. M.Sadeeq
Quality Assurance
Duhok Polytechnic University
Duhok, Iraq
mohammed.abdulrazaq@dpu.edu.krd

4th Shakir Fattah Kak
Information Technology Dept.
Duhok Polytechnic University
Duhok -Iraq
shakir.fattah@dpu.edu.krd

5th Zryan Najat Rashid
Computer Network Dept.
Sulaimani Polytechnic University
Sulaimani-Iraq
zryan.rashid@spu.edu.iq

6th Azar Abid Salih
Information Technology Management Dept.
Duhok Polytechnic University
Duhok -Iraq
azar.abid@dpu.edu.krd

7th Wafaa M. Abdullah
Computer Science Dept.
Nawroz University
Duhok, Iraq
wafaa.mustafa@nawroz.edu.krd

https://doi.org/10.48161/qaj.v1n2a52

***Abstract*— Many policymakers envisage using a community model and Big Data technology to achieve the sustainability demanded by intelligent city components and raise living standards. Smart cities use different technology to make their residents more successful in their health, housing, electricity, learning, and water supplies. This involves reducing prices and the utilization of resources and communicating more effectively and creatively for our employees. Extensive data analysis is a comparatively modern technology that is capable of expanding intelligent urban facilities. Digital extraction has resulted in the processing of large volumes of data that can be used in several valuable areas since digitalization is an essential part of daily life. In many businesses and utility domains, including the intelligent urban domain, successful exploitation and multiple data use is critical. This paper examines how big data can be used for more innovative societies. It explores the possibilities, challenges, and benefits of applying big data systems in intelligent cities and compares and contrasts different intelligent cities and big data ideas. It also seeks to define criteria for the creation of big data applications for innovative city services.**

Keywords*— *Data Computing, Internet of Things, Data Mining, Smart City, Cloud Computing.

I. Introduction

The Internet has experienced immense growth in recent years, and its content is continually growing and extending. The Internet of Things is viewed as an Internet application creation [1]. This means that customers, customers, paper, objects, and paper are all linked through contact and exchange of information, but the Internet is central to IoT [2].

The International Telecommunications Union Internet Report provides the following definition of IoT: The connection between the objects and the Internet and the interaction and communication of information through different protocols is achieved through various kinds of sensor systems to achieve an intelligent network identity, location, management, and control [3]. It has three different features: Intelligent Twitter sharing in real-time [4].

There have also been many studies in this area on innovative and wired companies' needs. There have been many facts, such as fixed and mobile sensors, internet data, and social data, from several outlets [5, 6].

Just a few of the data collection fields include agriculture, civic infrastructure, catastrophe management, education and apprehension, electricity, the efficiency of the environment, health and wellbeing, including medical, resilience, welfare, social services, telecommunications, transport, and mobility [7]. In several ways, a Smart City should be able (including Big-Data, traditional data sources, and personal information for users) to retrieve historical and real-time data from a wide range of sources [8].

The paper structure: have the Background Theory in Section II, Literature Review In Section III, discussion in Section IV, the conclusion in section V.

II. Background Theory

A. Data Mining

Significant figures are explosive, with an expected annual increase in global data production of 40% compared to just 5 percent in global IT expenditure. About 90% of digitized data worldwide have been registered over the past two years [9]. As a result, many municipalities worldwide have started using big data to help intelligent cities develop and sustain [10]. By recognizing their key smart city features, the cities have retained standards, values, and specifications for innovative city applications. Sustainability, long service cycle, governance, improved quality of life, and smart use of natural and urban resources comprise these qualities [11]. Smart City's well-defined components are mobility, governance, climate, people, and applications and services such as healthcare, transport, smart education, and electricity [12]. Sustainability, longevity, management, more outstanding living standards, and intelligent utilization of natural and urban resources are all virtues. The smart city is well developed in mobility, governance, atmosphere, inhabitants, wellbeing, mobility, smart education, and energy facilities [13, 14]. Big data has developed into a strategic weapon of immense potential importance that promotes industry's upgrade and growth as a critical driver for advancement [15]. It also affects science and methodology. Big data offers many benefits, including a large pool of capital and specialized training measurement technology [16]. As a result of large, complex, and volatile data, the storage and computational bottlenecks impede conventional data processing systems [17]. Working environments improved with time and resolved a range of measurement problems, including administrative functions at a high level, program upgrades, and the use of other computer series [18]. Big data mining is a service that collects the essential information and expertise from and provides the customer with a large, complex, competitive, high-volume, and low-density data set [19]. It helps to find valuable knowledge and expertise instead of traditional data mining. There are, however, technological, historical shortcomings, data environments, and mining scope [20]. The diagram below shows the layout of big data based on data mining techniques. The three levels of the architecture are networks, operating layers, and facility supports [21]:

1) Layer platform: The integration of big mixed data with a range of support technologies dependent on cloud infrastructure can support big data mining. The integration of big mixed data with a range of cloud computing support technologies can support them. Big data mining is also supported [22]. This cloud environment can not only provide the rest of the world with information, hardware, and software. However, it can also quantify moving data to allow more efficient preprocessing, analysis, and mining of complex data in several sources [23].

2) Functional layer: this layer can interpret and dig out data depending on users' requirements. The high efficiency of storages and computers made available to users as visuals, data sources, and other high-scalabilité and expandable technologies is essential for scientific, mining, and other tools [24].

3) Application Level Layers: Big data mining communicates automatically with customers, service providers, and users. The mining results lead to preprocessing, analysis, and extraction from various dynamic data sources [25].

B. Smart City and Cloud

City and metropolitan areas are abstruse social ecosystems, like local government, residents, and organizations [26]. The ICT is becoming progressively facilitating and enabling the ICTs to meet particular criteria relating to key themes, including enterprise and job growth, economic development, energy and water, public security, the atmosphere, health care, education, and public services. Simultaneously, urban spending is gradually being pressured by the new tumultuous global economic crisis, which is causing disastrous impacts not only on maintaining and upgrading existing ICT infrastructures and facilities but also on future innovation policies [27]. However, it has been identified, as an exemplary example of an answer to current and future complex challenges of the resource efficiencies, reducing emissions, sustainable health care services for older people, strengthening young people, and integral cities, which can be used for intelligent cities, information cities, digital towns, e-cities, and virtual towns. Smart cities, with their cameras, built-in computers, vast data

Fig. 1. Smart city G-Cloud platform [28].

Fig. 2. Pilot smart city G-Cloud project [28].

Sets and knowledge, and responses have been characterized to generate a specific spatial intelligence and creativity [29]. It proposes a comprehensive concept of a smart city that embodies a city that fosters sustainable economic development and high quality of life through investment in human and social capital, transportation and modern ICT infrastructure, and sound management of natural resources by participatory governments [30]. The requirement for investing in modern ICT and participatory government requirements include, in particular, the concept of empowering cities [31].

A form of democratic innovation that first becomes popular with the increasing capacity of businesses and consumers, using software products and services, t Two of the main issues highlighted in the definition mentioned above [32]. Propose a typology that defines the traditional roles inherent in the intelligent City that include smart market (competitiveness), intelligent citizens (social and human capital); intelligent administration (intervention); smart mobility (transport and ICTs) (quality of life) [33]. It is possible to use creative ICT to build sustainable strategies that cut expense, concentrate attention on top public challenges and align groups with common agendas by providing visual and innovative leadership [34]. The role of innovation facilitators in vital sectors such as industry, health, the environment, and ICT is played by cities worldwide. In their administration of daily public utilities, city councils look for bright, cost-efficient ICT strategies [35]. City and urban policymakers will use advanced low-cost technology tools such as cloud analytics to analyze data and market indicators for effective decision-making and predict problems to address them proactively and organize city capital to function efficiently [36].

C. Internet of Things

The three critical innovations used in the practical application of the Internet of Things are sensor technology, radio frequency identification technology, and embedded computer technology [37]. Sensors are components or tools that can obtain information on a sensing object and convert it into a relevant information type using rules or methods. Identification of the radio frequency merges radio frequency technologies with embedded technology to define particular objectives through radio signals and read and write specific target data [38]. The technology comprises computer software, electronic application technology, integrated circuit technology, and sensor technology as one unified technology. The embedded system consists of CPU, storage, I/O and applications, lightweight and high stability, controls computing capability, and fulfills the application device's different functions [39]. The increasing demand for intelligent and integrated communities has led to various research in this field, including the flagship, mobile devices, web-related data, and social data [40]. This abundance of knowledge comes from the various fields in which it has been used, including agriculture, civilian infrastructure, disaster relief, and response, preparation, learning, electricity, the efficiency of the atmosphere, health and wellbeing including, health and human services [41]. The use of technical means to identify individuals is the for persons with less physical contact. Passwords from bank ATM to the Internet are essential [42]. Although it is essential to understand, interpret and process data, politicians, municipal operators, residents and others can significantly benefit from the fact that data is viewed in real-time. This highlighted the need to display data meaningfully on a graphical user interface [43].

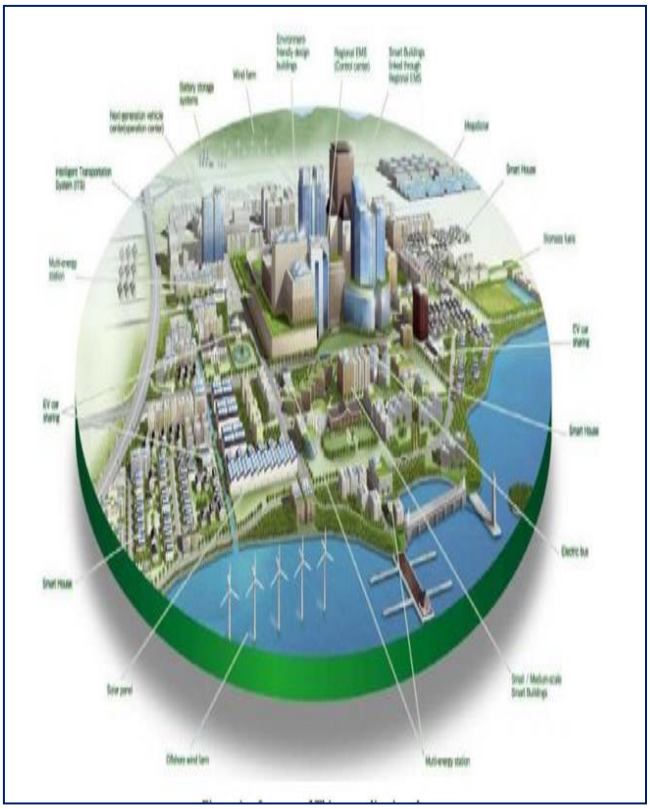

Fig. 3. Smart IOT application chart [44].

"smart city" means using applied research and technology, and information technology to refine and upgrade the infrastructure and quality of services in the City to increase municipal resource use efficiencies. [45]. It can considerably improve and maximize urban management, services, and people's happiness with urban life. The Internet of Things has been interwoven with various fields and has pervaded every aspect of our lives due to its exponential progress [46]. Smart houses, clever college, intelligent medical treatment, intelligent transportation, intellectual enterprise, and intelligent agriculture are one of the significant Internet-related fields of affairs, as shown in Figure 3. [47].

The idea of a smart city differs from a people's perspective in terms of technology[48]. This is clear as the countries embark on smart cities projects because they provide a wide range of insights into the problem. While intelligent cities worldwide are becoming increasingly common, their significance is still a mystery [49]. "The smart cities market continues in the 'I know it when I see it' process without a generally recognized definition." The notion of an intelligent city is not widely known, and identifying one has proved difficult [50]. On the other hand, several definitions stress standard features, characteristics, and elements that can help define intelligent city perspectives [51]. The use of hardware, applications, and networks for information technology, data on different urban areas and services, for instance, would enhance the quality of life for a particular population group. [52]. Figure 4 depicts how big data technologies are used in smart cities.

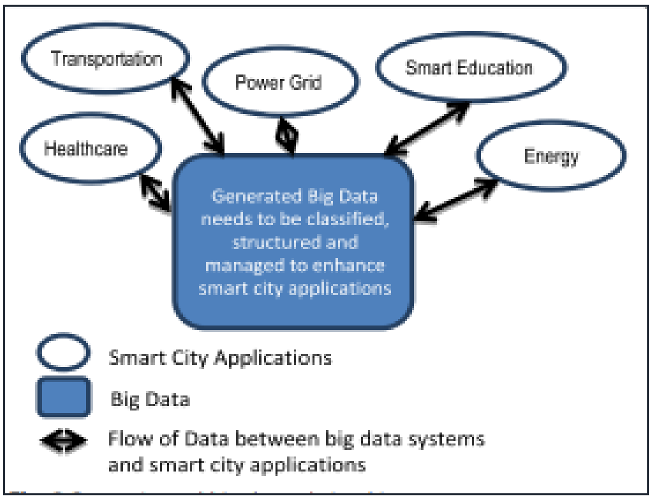

Fig. 4. Smart City and big data relationship [53].

Smart city applications Produce vast numbers of data that are then used by big data networks for knowledge to enhance intelligent urban applications[54]. Big data platforms can archive, process, and obtain information effectively from intelligent city applications and collect data to enhance various intelligent city services [55].

III. Literature Review

In recent work and research projects in this field, big data promoting intelligent city apps and services was emphasized. Furthermore, some study has examined some of the problems of using big data in intelligent cities.

Kolozali et al. [56] Proposed a novel architecture that allows for real-time tracking of the City's pulse with a fast memantine data pipeline. The proposed architecture allows for the robust semantic convergence of data sources and dynamic event handling in addition to real-time data aggregation and quality monitoring in a Semantic Web framework. The author tests our framework on the open-source Open Data Aarhus with real-time sensor observations published by the City of Aarhus.

Kotevska et al. [57] Submitted a dynamic network model to improve data failure services' resilience. In the absence of data, the network model detects statistically significant mutual time patterns through multivariant streams and uses these trends to improve prediction accuracy. The network model uses these patterns. The system's complexities often cause it to respond to data stream changes, including the loss of integration of new data flows. Montgomery County, Maryland, USA, city-based crime patterns represent the network model.

M. V. Moreno et al. [58] The standardized Internet proposed to be used in various smart urban implementations of Things-based architecture passed along two actual examples of data research. One of them displays some of the smart campus projects implemented by the University of Murcia. The second is a tram operation case with thousands of transactions via transit pass by discussing two major city applications of computer technology: smart campus buildings and urban mass transit, energy storage, and comfort control schemes. They showed that in the initial smart city scenario, the author could achieve median energy savings of 23% per month by applying efficient big-data techniques to problems such as indoor location, power usage modeling, and optimization.

H. Zhang et al. [59] The distinctive architecture showcasing the City's biodiversity, including cameras, sensors, and other physical equipment, was unmistakable. Safe City's three-tier architecture consists of a data security layer, a data computing layer, and a decision-making layer. The first level of payload-dependent symmetrical encryption protects the data against intrusions by exchanging accurate data among physical devices. The second layer is used to calculate secure data. Visions from data are extracted in the third layer. The use of Raspberry Pi boards guarantees a stable exchange of data, while the Hadoop framework is used for testing accurate data computations.

A. Gyrard [60] demonstrated how ontology catalogs could plan and grow smart city apps more efficiently. Four ontological catalogs are critical for IoT and intelligent cities: READY4SmartCities, LOV, Open Sensing City, and LOV4IoT. They also suggested a technique to enrich ontology catalogs with some ontologies to allow semantic interoperability with the reuse of smart city apps based on ontology. The approach is sufficiently general to be applied to any other area, shown by its adoption of the OSC and LOV4IoT Ontology Catalog methodology. In this appraisal, the importance of the ontology catalogs has led to their continuous evolution and maintenance.

S. Fiore [61] A public transport research platform based on the EUBra-BIGSEA platform and used by stakeholders in the City of Curitiba to deal with the urban traffic data processing and planning issues, focusing on the City administration dashboard. A large and fast data analytics platform, a scalable, diverse cloud architecture, data quality modules, protection and privacy techniques, a rich programmable system layer, and a web interface with various views were specified for the proposed solution.

M. Shen [62] Proposed Stable SVM, a blockchain-based cryptographic IoT data privacy-preserving SVM training scheme. Using blockchain technologies, a secure, reliable data-sharing network is developed for multi-data providers to encrypt IoT data and then archive it on a distributed directory. Their use of a homomorphic cryptosystem, Paillier, is designed to construct robust building blocks such as safe multiplication of polynemes, secure comparison, and a secure SVM training algorithm that only requires two connections one iteration and does not need a trustworthy third party.

J. Qiu [63] Presented a multi-phase association search tool to derive non-taxonomic associations from domain records. Various types of semantic data are used to improve the consistency of the method. They initially suggested an approach based on the semantic diagram, which incorporates structural knowledge from the semantic diagram and meaning data from words to define nontaxonomic relations. This has been influenced by network representation work. Secondly, various memantine verb sets depend on syntactic details about the dependence and are classified as non-taxonomic names. Extensive tests demonstrate the efficiency of the proposed structure.

L. Qi [64] Evaluated a series of experiments using the classic Locality-Sensitive Hashing (LSH) technique to guarantee that consumer anonymity is hidden in the data before data is combined from smart cities additional mining, analysis, and forecast. However, it is also a daunting challenge to maintain user confidentiality while providing accurate data processing and forecast results following data fusion, as data protection and data accessibility come together in a mutual agreement. To answer this challenge, the authors offer a new solution based on the traditional LSH methodology for a brilliant industrial setting concerning data fusion confidentiality.

V. Dattana [65] Rrepresented Biography of data object distribution across different agents. Identifying the individual or entities responsible for the data breach is aimed at protecting confidential data. The Scheme uses robust encryption mechanisms to ensure the safety of the protocol. Biography is used for the distribution of data between different agents. By combining the probability of data leaks with data secrecy, the data leaker's identity is determined.

M. Zhou and X. Chen [66] The Dendrobium preparation antioxidant process has been investigated from a compositional point of view, and a study has been conducted on the potential nutrient advantage of Dendrobium compound preparation. After preparation for the Dendrobium compound, the human body's physiological changes were studied in an intelligent big data city's health environment. Dendrobium compound preparation extract (DCPE) has evaluated the antioxidant activity on the cell level and decide its overall composition and sugar content. DCPE has effectively reduced the reactive oxygen species (ROS) level and reduced glutathione in human cloned colon cancer cells (Caco-2) in the data-sized smart city health setting with dose-dependent effects showing that Caco-2 cells will substantially improve their function at the cell level with antioxidants.

C. Xu et al. [67] created Sanya tourism inspiration in a Smart Tourism area, combining current Sanya tourism, All-for-One Tourism, and Smart Cities. Analyzes the benefits of building a Smart Tourism City in Sanya and the drawbacks and proposes solutions to these issues. By constructing a Wisdom Travel Framework using a cell phone app, how to use the Internet and Big Data would become a Smart Cities of Sanya building strategy.

L. Wei and S. Yang [68] analyzed the study of big data urban digital management needs, referred to the critical challenges of smart city development in the early stages, and proposes a model and countermeasures for big data city's construction and activity. Establish a sustainable environmental resource system and urban development course by beginning with environmental protection and energy. Amsterdam, for example, has sustainable transportation and public space.

S. Xiong and B. Ye [69] introduced smart cities' implementation in China and conducted a theoretical study of their construction. The paper then examines the case of Singapore's smart national planning, summarizes Singapore's experience, and offers some recommendations for China's smart city growth. China can make big data a new driving force for technological progress and economic transformation by expanding its creative applications in cities. The author can complete the entire phase of urban planning, design, and management using big data as our guide.

B. N. Silva [70] proposed an intelligent city architecture with Big Data Analytics (BDA) built in. The proposed Scheme's main aim is to improve the efficiency of real-time decision-making by implementing adequate Big Data (BD) processing. The review results provide valuable information for community growth while ensuring that the proposed framework's processing time and throughput are improved.

IV. Discussion

Many experiments have been used to clarify the thesis and examine current trends, as shown in Table 1. It goes through the methods used, the approaches that were taken, and the flaws found. This approach exploration explains the theoretical and explorative approaches and the disadvantages of the described problems. [56] Examine the system by reducing data streams with Symbolic Aggregate Approximation and performing quality analysis on multiple data streams. Based on the proposed stream quality analysis and data aggregation techniques, the author also optimizes semantic data discovery and integration. Nevertheless, [57] Compared to single city-based auto-regression, built and shared temporal patterns between cities provided improved crime rate prediction and robustness to data loss. Silver Spring has the highest increase in performance at 7.8%, with an average improvement of 5.6 % among cities with high crime rates. [58] The proposed trends piqued the service provider's interest, who wanted to learn more about how people use the transportation system. This was extremely beneficial in developing improved preparation protocols and more appealing advertising strategies. The current work focuses on incorporating people's behavior into the operational loop of these smart city systems. As a result, in smart building applications, consumers will be motivated to participate in energy conservation by their participation. On the other hand, data from crowdsensing projects would be incorporated into the public tram service to enhance the estimation of urban mobility trends. [59] Presents useful insights into a cleaner and clever City in an IoT world built on sensors. This paper imagined the essential role of safety and protection in IoT-enabled data calculation and communication in ensuring safe and stable decisions. By leveraging the interaction between the different data features, data generated by IoT sensors allow the sense of a healthy urban environment. [60] evaluated the quality of IoT and smart city ontologies to increase the consistency of the ontology catalog, they also share what they have learned about ontology best practices and provide software resources to help with ontology improvements. They created a system for enriching ontology catalogs, which is used in the LOV4IoT ontology catalog. The Open Sensing City and LOV4IoT ontology catalogs are a massive undertaking. However, they are a necessary first step toward encouraging the reuse of existing ontologies and fostering semantic interoperability among ontology-based smart city applications. [61] The author suggested a solution that combines a broad and fast, scalable data analytics platform, flexible and dynamic cloud infrastructure, data quality and algorithms matching entities, and security and privacy practices into a single solution. The interoperable programming platform based on Python Application Program Interface (API) enables the easy, quick, and transparent creation of smart city applications. [62] Presented Safe SVM, a new SVM privacy-preserving training scheme that solves privacy and integrity issues using blockchain techniques to generate an encrypted multi-part scenario SVM training algorithm, collecting IoT information from various data sources. The homomorphic Paillier cryptosystem creates a fast and accurate SVM training algorithm that respects privacy. Author have shown safe productivity and safety of SVM. [63] The extraction, as the basis for knowledge representation, of non-taxonomic relations has been studied. A term used to describe the large quantities of media data generated by different intelligent city devices. Multi-elements and multi-information should be considered in order to improve system performance in the future. Combining knowledge about semantic graphic structure with background details will help to identify better, non-taxonomic relationships. [64] Used a sequence of actual data sets-based tests to evaluate them. Experiments reveal that our system performs in estimation precision against other rivals. A novel data fusion and forecasting methodology is proposed for intelligent cities. Finally, the feasibility of our proposal confirms several real-world experiments. The suggested solution to privacy-conscious data fusion in this paper may be useful for big data and smart city applications, such as medical data fusion [65] Conducted C/C++ model experiments and a simulated data leakage environment. By comparing its probability with a function known as the weight factor WF, the effectiveness of a system is evaluated. The weight factor consists of a proportion of the total number of data objects available for allotment of all data objects assigned. The Scheme uses robust encryption mechanisms to ensure the safety of the protocol. [66] It was shown to contain a total of 28 member compounds in Dendrobium fruit compounds, with eight first-time Dendrobium amine derivatives. In vitro tests showed that DCPE supplied, respectively, the total ingredient and the total sugar of GAE 100g-1DW 3,320 mg and DCPE 100g-1 DW 570 mg. [67]. He examined the advantages of developing a city for intelligent tourism and found how the Internet and Big Data can develop a plan for Smart Cities. These problems can be solved by creating a Wisdom Travel Framework using the smartphone app across the Internet and Big Data as smart cities in Sanya. [68] To understand how big data should be used, analysis at the application level needs to be accelerated. A standard set of service standards and recommendations for large data applications developed. Wisdom is pushing the City for the safe use of big data.[69] The City was reexamined and understood through data, intelligent plan collection, science-driven urban building, and new urban wisdom development environment were studied, urban planning problems resolved, and urban management operations improved continuity and efficiency. [70] A three-fold architecture has been proposed for data storage and data analyzes and for data application. They used the planned BDA embedded Smart City to assess authentic data on water use, pollution, parking and air quality control. The architecture suggested analyzes data and decides intelligently to provide passengers with traffic level warnings. It shows the effects of the study of vehicle density between two lane sites listed. Congestion alerts are issued if the density exceeds the threshold.

TABLE I. COMPARISON OF PREVIOUS WORKS ABOUT DATA MINING IMPLEMENTATION IN SMART CITY APPLICATIONS

Ref.	Algorithms	Tools	Satisfied Objectives and Significant Results
[58], 2016	Mobile data sources and build a BMS system	Transit-card Mining with Predictive Models	The results from both scenarios demonstrate the potential for these types of techniques to provide good smart city services, such as energy use and comfort management in smart buildings and the identification of travel profiles in smart transportation.
[57], 2017	The Adaptive algorithms	geospatial analysis with using predictive models	It can be used as an input for quickly emerging recommendation motors as used in an inter-city network. It can be used in other networks of data streams, for instance, dispersed local clouds in a Smart City environment. This network system can also be scaled on common platforms by using traditional statistical techniques.
[66], 2018	Mobile edge computing framework	Dendrobium Compound Preparation Extract (DCPE)	Dendrobium compound preparation includes many component compounds, many of which have relatively high antioxidant activity when applied transdermal and can be found in various foods, pharmaceuticals, cosmetics, and natural products.
[67], 2018	Mobile Phone app. Mobile Terminal	Wisdom Travel Framework	analyzes the benefits of creating a Smart Tourism city, and determines how to use the Internet and Big Data to create a Smart Cities building strategy.
[68], 2018	Cloud service-based support construction model	Construction Operation Mod	accelerating research at the application level and establishing uniform requirements and service guidelines for big data applications. Wisdom propels the City forward to encourage the healthy development of big data.
[70], 2018	The Hadoop distributed file system (HDFS) MapReduce is used to perform data analysis	Big Data Analytics (BDA)	The findings showed that a Hadoop cluster with a dual node exceeded a Hadoop cluster with one node and dependent on jQuery. The processing time was halved compared to jQuery-based processing. The incorporation of Big Data Analytics facilitates comprehensive data analytics. By evaluating the results, the threshold value was computed.
[60], 2018	Ontology Catalogs	Semantic Web Technologies	IoT and smart cities, the author, looked at four ontology catalogs (Ready4SmartCities, LOV, LOV4IoT, and Open Sensing City). Studied the most up-to-date ontologies in order to build smart cities using IoT technologies.
[56], 2018	The Genetic Algorithms	C-SPARQL	- investigated large-scale stream processing in the context of innovative city applications, including data aggregation, quality analysis, semantic data discovery, convergence, and complex event processing issues. The author demonstrated how to minimize data size using techniques such as SAX and various techniques for accessing and manipulating semantically annotated data streams.
[63], 2018	Priority algorithm, Latent Dirichlet Allocation (LDA) and Word2Vec	Semantic Web Technologies With frequent itemsets for mining association rules.	Using big multimedia data generated by various smart city devices explained the massive amounts of multimedia data. To boost the device output in the future, multi-elements and multi-information should be considered.
[61], 2019	Spark and COMPSs run into a cloud	interoperable programming framework	The City administration has demonstrated its Dashboard, a public analysis program designed to address the Curitiba Municipality of Brazil's needs on the EUBra-BIGSEA

	environment	based on Python Application Programming Interface (API)	Website. –defines the approach proposed in terms of a broad and rapid data analytical network, a cloud server, flexible and complex data quality-aware plugins, data security and privacy strategies, and a web-based multi-viewed GUI.
[62], 2019	Blockchain-based IoT platform, and IoT data analysts.	Machine Learning - Blockchain, Homomorphic Cryptosystem	Safe SVM presented a new SVM data protection training scheme that addresses data protection and integrity challenges through blockchain technology, creating a secured SVM training algorithm in multiple scenarios that collect IoT data from a multifunctional data provider. The homomorphic cryptosystem from Paillier is used to create a compelling and reliable privacy-conserved SVM training algorithm. They demonstrated how stable and reliable SVM operates.
[65], 2019	Mobile MODEL through Biograph. Information leaker	Guilty Agent Identification Model	-The probability of an agent being guilty is calculated by using Biography. -The measured risk of leaked data and the secrecy of data is compared into an information leaking device. - In order to improve the confidentiality of the most sensitive documents, consider the threshold value
[69], 2019	Google Trends	Combination graph theory algorithm with spanning tree-based greed algorithm In the Urban environment	ensure data protection, rethink and understand the City through data, investigate the compilation of intelligent planning, scientifically straightforward urban design, and open up a new world of urban wisdom creation.
[64], 2020	WS-DREAM with cloud architecture	Python, LSH technique 64 pieces of QoS temporal information (at 15-minute interval) is reported in the dataset	A series of real-world experiments confirm the viability of our plan. This paper's proposed privacy-aware data fusion approach may help big data analyses and smart city applications, such as medical records fusion.
[59], 2020	Hadoop platform with Symmetric Encryption	Data Analytics using Hadoop platform	According to the evaluations, Safe City provides valuable insights into a stable, smart city in the sense of a sensor-based IoT world.

V. Conclusions

This paper's main contribution is to research extensive data in intelligent cities and explore the advantages and challenges of using extensive data in intelligent cities. The paper also looks at the general specification and deployment criteria for Big-Data architectures for smart city apps and services. Efficient use and protection of sensitive data are needed in order to implement the smart city definition effectively.[57] The model correctly calculates all optimal network links according to prediction error minimization. Weather proves to be a good predictor of crime, while the city-to-city gap is a predictor of shared trends in crime. [58] The general 4-layer smart city systems architecture that takes into consideration big data was proposed. There is a distinction between static and mobile data sources, with methods suggested for each extraction of relevant data. Two large data applications will be described for innovative city services. In particular, energy management and comfort regulation systems are in smart campus buildings and urban mass transit. On the other hand [59], Based on the data protected and the data calculation layer was processed. The decision-maker collects facts to assist you in making smart decisions. Using Raspberry Pi boards, omnipresent data security is analyzed, and Hadoop is used to calculate all kinds of omnipresent data. Compared to existing approaches, safe city handshake duration, answer time, and typical memory use are all unusual. The processing time is short, the performance is good and the data intake is more reliable. [60] LOV4IoT's ontology catalog was reviewed with users after the standards and tools for improving the quality of ontologies were created. These contributions help developers to repurpose current intelligent city and IoT ontologies to build new applications.

References

[1] M. A. Sadeeq, S. R. Zeebaree, R. Qashi, S. H. Ahmed, and K. Jacksi, "Internet of Things security: a survey," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 2018, pp. 162-166.

[2] A. A. Yazdeen, S. R. Zeebaree, M. M. Sadeeq, S. F. Kak, O. M. Ahmed, and R. R. Zebari, "FPGA Implementations for Data Encryption and Decryption via Concurrent and Parallel Computation: A Review," *Qubahan Academic Journal,* vol. 1, pp. 8-16, 2021.

[3] O. Alzakholi, H. Shukur, R. Zebari, S. Abas, and M. Sadeeq, "Comparison among cloud technologies and cloud performance," *Journal of Applied Science and Technology Trends,* vol. 1, pp. 40-47, 2020.

[4] A. Xin, L. Yiheng, T. Yawen, and B. Zhengde, "The typical application of internet of things in China in the future: The smart home," in *2018 International Conference on Electronics Technology (ICET)*, 2018, pp. 361-364.

[5] A. B. Sallow, H. I. Dino, Z. S. Ageed, M. R. Mahmood, and M. B. Abdulrazaq, "Client/Server Remote Control Administration System: Design and Implementation," *Int. J. Multidiscip. Res. Publ,* vol. 3, p. 7, 2020.

[6] A. I. Abdulla, A. S. Abdulraheem, A. A. Salih, M. A. Sadeeq, A. J. Ahmed, B. M. Ferzor, *et al.*, "Internet of Things and Smart Home Security," *Technol. Rep. Kansai Univ,* vol. 62, pp. 2465-2476, 2020.

[7] Q. Han, P. Nesi, G. Pantaleo, and I. Paoli, "Smart City Dashboards: Design, Development, and Evaluation," in *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, 2020, pp. 1-4.

[8] P. Bellini, M. Marazzini, N. Mitolo, M. Paolucci, D. Cenni, and P. Nesi, "Smart City Control Room Dashboards Exploiting Big Data Infrastructure (S)," in *DMSVIVA*, 2018, pp. 44-50.

[9] A. S. Abdulraheem, A. A. Salih, A. I. Abdulla, M. A. Sadeeq, N. O. Salim, H. Abdullah, *et al.*, "Home automation system based on IoT," 2020.

[10] D. H. Maulud, S. R. Zeebaree, K. Jacksi, M. A. M. Sadeeq, and K. H. Sharif, "State of art for semantic analysis of natural language processing," *Qubahan Academic Journal,* vol. 1, pp. 21-28, 2021.

[11] D. A. Hasan, B. K. Hussan, S. R. Zeebaree, D. M. Ahmed, O. S. Kareem, and M. A. Sadeeq, "The impact of test case generation methods on the software performance: A review," *International Journal of Science and Business,* vol. 5, pp. 33-44, 2021.

[12] F. Q. Kareem, S. R. Zeebaree, H. I. Dino, M. A. Sadeeq, Z. N. Rashid, D. A. Hasan, *et al.*, "A Survey of Optical Fiber Communications: Challenges and Processing Time Influences," *Asian Journal of Research in Computer Science,* pp. 48-58, 2021.

[13] E. Al Nuaimi, H. Al Neyadi, N. Mohamed, and J. Al-Jaroodi, "Applications of big data to smart cities," *Journal of Internet Services and Applications,* vol. 6, pp. 1-15, 2015.

[14] K. Pantelis and L. Aija, "Understanding the value of (big) data," in *2013 IEEE International Conference on Big Data*, 2013, pp. 38-42.

[15] S. Zeebaree, "Energy Management for Internet of Things via Distributed Systems," *Journal of Applied Science and Technology Trends,* vol. 2, pp. 59-71, 2021.

[16] Z. Ageed, M. R. Mahmood, M. Sadeeq, M. B. Abdulrazzaq, and H. Dino, "Cloud computing resources impacts on heavy-load parallel processing approaches," *IOSR Journal of Computer Engineering (IOSR-JCE),* vol. 22, pp. 30-41, 2020.

[17] Z. H. Deng, B. Fan, Y. J. Lu, and Z. F. Li, "Discussion about big data mining based on hadoop," in *Applied Mechanics and Materials*, 2013, pp. 2063-2066.

[18] Z. S. Ageed, R. K. Ibrahim, and M. A. Sadeeq, "Unified Ontology Implementation of Cloud Computing for Distributed Systems," *Current Journal of Applied Science and Technology,* pp. 82-97, 2020.

[19] S. M. S. A. Abdullah, S. Y. A. Ameen, and S. Zeebaree, "Multimodal Emotion Recognition using Deep Learning," *Journal of Applied Science and Technology Trends,* vol. 2, pp. 52-58, 2021.

[20] C. Stergiou, K. E. Psannis, B.-G. Kim, and B. Gupta, "Secure integration of IoT and cloud computing," *Future Generation Computer Systems,* vol. 78, pp. 964-975, 2018.

[21] Y. Jiugen and X. Ruonan, "Cloud Computing-based Big Data mining Connotation and Solution," in *2020 15th International Conference on Computer Science & Education (ICCSE)*, 2020, pp. 245-248.

[22] A. B. Sallow, S. R. Zeebaree, R. R. Zebari, M. R. Mahmood, M. B. Abdulrazzaq, and M. A. Sadeeq, "Vaccine Tracker/SMS Reminder System: Design and Implementation."

[23] Z. S. Ageed, S. R. Zeebaree, M. M. Sadeeq, S. F. Kak, H. S. Yahia, M. R. Mahmood*, et al.*, "Comprehensive Survey of Big Data Mining Approaches in Cloud Systems," *Qubahan Academic Journal,* vol. 1, pp. 29-38, 2021.

[24] L. M. Abdulrahman, S. R. Zeebaree, S. F. Kak, M. A. Sadeeq, A.-Z. Adel, B. W. Salim*, et al.*, "A State of Art for Smart Gateways Issues and Modification," *Asian Journal of Research in Computer Science,* pp. 1-13, 2021.

[25] I. M. Ibrahim, "Task Scheduling Algorithms in Cloud Computing: A Review," *Turkish Journal of Computer and Mathematics Education (TURCOMAT),* vol. 12, pp. 1041-1053, 2021.

[26] M. Sadeeq, A. I. Abdulla, A. S. Abdulraheem, and Z. S. Ageed, "Impact of Electronic Commerce on Enterprise Business," *Technol. Rep. Kansai Univ,* vol. 62, pp. 2365-2378, 2020.

[27] H. Shukur, S. R. Zeebaree, A. J. Ahmed, R. R. Zebari, O. Ahmed, B. S. A. Tahir*, et al.*, "A State of Art Survey for Concurrent Computation and Clustering of Parallel Computing for Distributed Systems," *Journal of Applied Science and Technology Trends,* vol. 1, pp. 148-154, 2020.

[28] T. Clohessy, T. Acton, and L. Morgan, "Smart city as a service (SCaaS): A future roadmap for e-government smart city cloud computing initiatives," in *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing*, 2014, pp. 836-841.

[29] S. S. R. Zeebaree, S. Ameen, and M. Sadeeq, "Social Media Networks Security Threats, Risks and Recommendation: A Case Study in the Kurdistan Region," *International Journal of Innovation, Creativity and Change,* vol. 13, pp. 349-365, 2020.

[30] A. A. Salih, S. R. Zeebaree, A. S. Abdulraheem, R. R. Zebari, M. A. Sadeeq, and O. M. Ahmed, "Evolution of Mobile Wireless Communication to 5G Revolution," *Technology Reports of Kansai University,* vol. 62, pp. 2139-2151, 2020.

[31] A. A. Salih and M. B. Abdulrazaq, "Combining best features selection using three classifiers in intrusion detection system," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 94-99.

[32] A. M. Abdulazeez, S. R. Zeebaree, and M. A. Sadeeq, "Design and Implementation of Electronic Student Affairs System," *Academic Journal of Nawroz University,* vol. 7, pp. 66-73, 2018.

[33] M. A. Sulaiman, M. Sadeeq, A. S. Abdulraheem, and A. I. Abdulla, "Analyzation Study for Gamification Examination Fields," *Technol. Rep. Kansai Univ,* vol. 62, pp. 2319-2328, 2020.

[34] F. E. F. Samann, S. R. Zeebaree, and S. Askar, "IoT Provisioning QoS based on Cloud and Fog Computing," *Journal of Applied Science and Technology Trends,* vol. 2, pp. 29-40, 2021.

[35] Z. N. Rashid, S. R. Zeebaree, and A. Sengur, "Novel Remote Parallel Processing Code-Breaker System via Cloud Computing."

[36] S. R. Zeebaree, "Remote Controlling Distributed Parallel Computing System over the Cloud (RCDPCSC)," in *2020 3rd International Conference on Engineering Technology and its Applications (IICETA)*, 2020, pp. 258-258.

[37] M. M. Sadeeq, N. M. Abdulkareem, S. R. Zeebaree, D. M. Ahmed, A. S. Sami, and R. R. Zebari, "IoT and Cloud Computing Issues, Challenges and Opportunities: A Review," *Qubahan Academic Journal,* vol. 1, pp. 1-7, 2021.

[38] A. Iqbal, F. Ullah, H. Anwar, K. S. Kwak, M. Imran, W. Jamal*, et al.*, "Interoperable Internet-of-Things platform for smart home system using Web-of-Objects and cloud," *Sustainable Cities and Society,* vol. 38, pp. 636-646, 2018.

[39] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, "A review of smart home applications based on Internet of Things," *Journal of Network and Computer Applications,* vol. 97, pp. 48-65, 2017.

[40] H. Shukur, S. Zeebaree, R. Zebari, D. Zeebaree, O. Ahmed, and A. Salih, "Cloud computing virtualization of resources allocation for distributed systems," *Journal of Applied Science and Technology Trends,* vol. 1, pp. 98-105, 2020.

[41] J. Hale, "NSF Smart and Connected Communities 2021 Info Session," 2020.

[42] M. Abdulrazaq and A. Salih, "Combination of multi classification algorithms for intrusion detection system," *Int. J. Sci. Eng. Res.,* vol. 6, pp. 1364-1371, 2015.

[43] G. McArdle and R. Kitchin, "The Dublin Dashboard: Design and development of a real-time analytical urban dashboard," 2016.

[44] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, "Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios," *Ieee Access,* vol. 8, pp. 23022-23040, 2020.

[45] Z. Zhang, "Research on the Construction Mode of Intelligent Traffic Safety Integration Under the Background of Intelligent City," in *2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 2020, pp. 57-60.

[46] A. A. Salih and A. M. Abdulazeez, "Evaluation of Classification Algorithms for Intrusion Detection System: A Review," *Optimization (PSO),* vol. 42, p. 43, 2021.

[47] Q. Kun, "Design of Traffic Operation Situation Analysis System Based on Smart city Intelligent Transportation System [J]," *China Management Informatization,* vol. 19, pp. 208-210, 2016.

[48] H. I. Dino, S. R. Zeebaree, A. A. Salih, R. R. Zebari, Z. S. Ageed, H. M. Shukur*, et al.*, "Impact of Process Execution and Physical Memory-Spaces on OS Performance."

[49] A. AL-Zebari, S. Zeebaree, K. Jacksi, and A. Selamat, "ELMS–DPU ontology visualization with Protégé VOWL and Web VOWL," *Journal of Advanced Research in Dynamic and Control Systems,* vol. 11, pp. 478-85, 2019.

[50] A. Zeebaree, A. Adel, K. Jacksi, and A. Selamat, "Designing an ontology of E-learning system for duhok polytechnic university using protégé OWL tool," *J Adv Res Dyn Control Syst Vol,* vol. 11, pp. 24-37.

[51] S. A. A.-Z. A. Selamat, "Electronic Learning Management System Based on Semantic Web Technology: A Review," *Int. J. Adv. Electron. Comput. Sci,* vol. 4, pp. 1-6, 2017.

[52] W. Fan and A. Bifet, "Mining big data: current status, and forecast to the future," *ACM SIGKDD explorations newsletter,* vol. 14, pp. 1-5, 2013.

[53] P. Michalik, J. Štofa, and I. Zolotova, "Concept definition for Big Data architecture in the education system," in *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2014, pp. 331-334.

[54] A. Al-Zebari and A. Sengur, "Performance Comparison of Machine Learning Techniques on Diabetes Disease Detection," in *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, 2019, pp. 1-4.

[55] A.-Z. Adel, S. Zebari, and K. Jacksi, "Football Ontology Construction using Oriented Programming," *Journal of Applied Science and Technology Trends,* vol. 1, pp. 24-30, 2020.

[56] Ş. Kolozali, M. Bermudez-Edo, N. FarajiDavar, P. Barnaghi, F. Gao, M. I. Ali*, et al.*, "Observing the pulse of a city: A smart city framework for real-time discovery, federation, and aggregation of data streams," *IEEE Internet of Things Journal,* vol. 6, pp. 2651-2668, 2018.

[57] O. Kotevska, A. G. Kusne, D. V. Samarov, A. Lbath, and A. Battou, "Dynamic network model for smart city data-loss resilience case study: City-to-city network for crime analytics," *IEEE Access,* vol. 5, pp. 20524-20535, 2017.

[58] M. V. Moreno, F. Terroso-Sáenz, A. González-Vidal, M. Valdés-Vela, A. F. Skarmeta, M. A. Zamora*, et al.*, "Applicability of big data techniques to smart cities deployments," *IEEE Transactions on Industrial Informatics,* vol. 13, pp. 800-809, 2016.

[59] H. Zhang, M. Babar, M. U. Tariq, M. A. Jan, V. G. Menon, and X. Li, "SafeCity: Toward safe and secured data management design for IoT-enabled smart city planning," *IEEE Access,* vol. 8, pp. 145256-145267, 2020.

[60] A. Gyrard, A. Zimmermann, and A. Sheth, "Building IoT-based applications for smart cities: How can ontology catalogs help?," *IEEE Internet of Things Journal,* vol. 5, pp. 3978-3990, 2018.

[61] S. Fiore, D. Elia, C. E. Pires, D. G. Mestre, C. Cappiello, M. Vitali*, et al.*, "An integrated big and fast data analytics platform for smart urban transportation management," *IEEE Access,* vol. 7, pp. 117652-117677, 2019.

[62] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani, "Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities," *IEEE Internet of Things Journal,* vol. 6, pp. 7702-7712, 2019.

[63] J. Qiu, Y. Chai, Y. Liu, Z. Gu, S. Li, and Z. Tian, "Automatic non-taxonomic relation extraction from big data in smart city," *IEEE Access*, vol. 6, pp. 74854-74864, 2018.

[64] L. Qi, C. Hu, X. Zhang, M. R. Khosravi, S. Sharma, S. Pang*, et al.*, "Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment," *IEEE Transactions on Industrial Informatics*, 2020.

[65] V. Dattana, K. Gupta, and A. Kush, "A probability based model for big data security in smart city," in *2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)*, 2019, pp. 1-6.

[66] M. Zhou and X. Chen, "Application of Dendrobium Compound Preparation in the Health Environment of Big Data Smart City," in *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 2018, pp. 65-68.

[67] C. Xu, X. Huang, J. Zhu, and K. Zhang, "Research on the construction of sanya smart tourism city based on internet and big data," in *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 2018, pp. 125-128.

[68] L. Wei and S. Yang, "Based On Big Data Technology Analysis On The Mode And Countermeasures Of Smart City Construction Operation Management," in *2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE)*, 2018, pp. 184-186.

[69] S. Xiong and B. Ye, "Analysis on the development path of smart city in the era of big data," in *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*, 2019, pp. 177-181.

[70] B. N. Silva, M. Khan, J. Seo, D. Muhammad, Y. Yoon, J. Han*, et al.*, "Exploiting Big Data Analytics for Urban Planning and Smart City Performance Improvement," in *2018 12th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2018, pp. 1-4.