A Data Management System for Smart Cities Leveraging Artificial Intelligence Modeling Techniques to Enhance Privacy and Security

Dr.V. Jyothi[1*], Dr. Tammineni Sreelatha[2], Dr.T.M. Thiyagu[3], R. Sowndharya[4] and N. Arvinth[5]

[1*]Assistant Professor, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad, Telangana, India. jyothinaikv@gmail.com, https://orcid.org/0000-0002-5899-2183

[2]Assistant Professor, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur District, Andhra Pradesh, India. sreelatha457@gmail.com, https://orcid.org/0000-0002-0951-2796

[3]Assistant Professor, Division of Artificial Intelligence and Machine Learning, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore, India. t.m.thiyagu@gmail.com, https://orcid.org/0000-0002-4902-3153

[4]Assistant Professor, Department of Computer Science and Engineering, Sona College of Technology, Salem, India. rajendransowndharya@gmail.com, https://orcid.org/0009-0006-6209-0357

[5]Research Associate, National Institute of STEM Research, India. arvinthwork@gmail.com, https://orcid.org/0009-0000-9798-9828

Received: October 04, 2023; Accepted: December 12, 2023; Published: February 29, 2024

Abstract

Smart cities are metropolitan areas that use sophisticated technology to increase efficiency, sustainability, and overall quality of life. The potential for transformation is tremendous, with applications ranging from Internet of Things (IoT)-driven infrastructure to data-driven governance. Effectively handling the abundant data produced in smart cities requires stringent security and privacy protocols. This research aims to tackle these difficulties by introducing the suggested Artificial Intelligence-based Data Management System (AI-DMS) for Smart Cities. AI-DMS seeks to optimize the data processing pipeline, guaranteeing effectiveness throughout the process, from data extraction to publication. Implementing a Multi-Level Sensitive Model is a notable addition, as it classifies data into three categories: sensitive, quasi-sensitive, and public. This allows for more nuanced sharing of data. Privacy preservation is accomplished using Principal Component Analysis (PCA), a comprehensive technique encompassing feature mapping, selection, normalization, and transformation. The simulation results demonstrate that AI-DMS outperforms other methods. It achieves a Data Quality Score of 95.12% (training) and 93.76% (testing), a Privacy Preservation Rate of 85.23% (training) and 82.76% (testing), a Processing Efficiency of 90.54% (training) and 88.76% (testing), a Sensitivity Model Accuracy of 80.12% (training) and 78.45% (testing), and a Data Access Time of 22.76 ms (training) and 21.32 ms (testing). The results highlight AI-DMS as a

Journal of Internet Services and Information Security (JISIS), volume: 14, number: 1 (February), pp. 37-51.
DOI: *10.58346/JISIS.2024.I1.003*

*Corresponding author: Assistant Professor, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad, Telangana, India.

reliable and effective system, guaranteeing superior smart city data management that is secure and precise. This contribution aligns with the changing urban scene, offering improvements in decision-making based on data while still ensuring privacy and security.

Keywords: Smart City, Data Management, Security, Privacy, Artificial Intelligence.

1 Introduction to Smart City and Data Management Issues

Smart city technology is expected to enhance the efficiency and productivity of communities in the following decades in response to the steady growth of urban populations globally (Zhao et al., 2021). The Smart City idea encompasses the integration of Information and Communication Technology (ICT) and a multitude of physical devices connected to the Internet of Things (IoT) network Bhattacharya et al., 2022; Raj et al., 2022). This integration aims to improve the efficiency of the city and its interaction with public data management systems (Ahmad et al., 2022). A smart city primarily consists of ICT, the framework used to establish, implement, and promote sustainable development principles to tackle the growing issues of urbanization. The network of interconnected items and machines that communicate data wirelessly and use cloud technology is an integral component of this ICT system. System IoT applications enable municipalities, enterprises, and individuals to acquire, analyze, and manage data in real-time, leading to improved decision-making and an enhanced quality of life (Liu et al., 2021; Gaber et al., 2022). Artificial Intelligence (AI) and big data have been combined in several study fields to create intelligent data management systems that improve safety in smart city structures (Herath & Mittal, 2022; Talebkhah et al., 2021).

Smart cities are urban areas that extensively use ICT. The emergence of IoT methods, defined by integrating technology, physical and digital communication, and sophisticated algorithms, has reinforced the technological aspect of determining smart societies. Computerized smart cities generate several forms of data, which are often converted to facilitate the process of discovery and access. This procedure is referred to as data publication. The data that has been published is then made available using query and publish/subscribe mechanisms. The data is incorporated into higher-level activities or applications via accessible interfaces, resulting in improved quality and effectiveness of smart city services (Kaginalkar et al., 2021).

Smart city data is publicly released to enable apps and municipal services to provide new possibilities for decision-making based on information. The data is gathered from several heterogeneous and complex sources, such as distinct kinds, forms, and quantities. Smart city data often needs several quality challenges, such as lack of uniformity, fullness, accuracy, and reliability. Data quality is essential for accurate decision-making. Low-quality data result in inaccurate analytical findings and misguided decision-making, possibly leading to catastrophic outcomes (Braga et al., 2021). While numerous platforms share information without proper processing processes, quality Key Performance Indications (KPIs), and quality monitors, a reliable data structure should offer information with a certain level of quality assurance. For instance, it must release well-prepared data sets or show indications highlighting data quality concerns (Rahman et al., 2018).

The primary contributions are listed below:

- **Streamlined Data Processing:** The suggested architecture guarantees effective and well-organized data handling by streamlining every step of the data management procedure for smart cities—from extraction to publishing.

- **Multi-Level Sensitive Model:** Provides a nuanced approach to data sharing depending on its nature by introducing a three-level sensitive framework (sensitive, quasi-sensitive, and public) for data arrangement and dissemination.
- **Privacy Preservation:** Principal Component Analysis (PCA) is used to change raw data while protecting privacy by offering a reliable technique for feature mapping, selection, normalization, and transformation.

The following sections are organized in the given manner: Section 2 examines the current research and discoveries in the literature about data management in smart cities. Section 3 introduces an Artificial Intelligence-based Data Management System (AI-DMS) designed explicitly for Smart Cities. Section 4 explores the simulation analysis and results of implementing the suggested AI-DMS. The study is concluded in Section 5, where significant results are summarized, and prospective future advancements and research paths are outlined.

2 Literature Survey and Analysis

This section performs an extensive literature review, examining current research and knowledge in smart city data management. This investigation serves as a basis for comprehending the present difficulties and solutions. The study covers various viewpoints and methods used by researchers in this field.

Babar et al. introduce an Energy-Aware Smart City Management System (EASCM) that utilizes data analytics and IoT (Babar et al., 2021). The suggested strategy enhances energy efficiency by synergizing data analytics with IoT technology. The data indicates a noteworthy decrease in energy use by 25%, a rise in operational effectiveness by 30%, and enhanced sustainability measurements. Cha et al. propose a Blockchain-Empowered Cloud Architecture (BECA) that utilizes secret sharing to enhance smart city applications (Cha et al., 2021). BECA uses blockchain and private sharing methodologies to improve privacy in cloud-based smart city systems. The research presents a reduction of 20% in security breaches, an enhancement of 15% in data integrity, and a boost of 25% in total system resilience.

Rajawat et al. concentrate on augmenting healthcare data security in smart cities via artificial intelligence and blockchain technology (Rajawat et al., 2022). Their suggested approach combines artificial intelligence algorithms and blockchain technology to ensure healthcare information security. The results indicate a significant decrease of 35% in data breaches, a notable enhancement of 40% in patient privacy, and a substantial rise of 20% in the overall effectiveness of healthcare data administration. Esposito et al. provide a Blockchain-Based Authentication and Authorization (BBAA) system designed specifically for smart city applications (Esposito et al., 2021). BBAA utilizes blockchain technology to provide safe access and permission in smart city systems. The findings demonstrate a 30% decline in illegal access attempts, a 25% enhancement in authentication speed, and a 20% drop in security-related occurrences.

Qureshi et al. provide the Nature-Inspired Algorithm-Based Secure Data Dissemination Framework (NIASDDF) for smart city networks (Qureshi et al., 2021). This solution uses nature-inspired algorithms to distribute data securely in smart city contexts. The results demonstrate a 15% enhancement in the speed of data dissemination, a 20% decrease in data loss, and a 25% advancement in the network's overall security. Abd El-Latif et al. provide a Quantum-Inspired Blockchain-Based Cybersecurity (QIBBC) method to protect smart edge utilities in IoT-based smart cities (Abd El-Latif et al., 2021). The approach incorporates quantum-inspired ideas into blockchain technology, therefore bolstering cybersecurity. The findings indicate a significant enhancement of 30% in security strength, a notable decrease of 25% in cyber assaults, and a commendable advancement of 20% in the system's overall resilience.

Izonin et al. propose a method for managing missing data called Missing Data Management (MDM) utilizing an Enhanced Generalized Regression Neural Network-Self-Generating Takagi-Sugeno-Kang Fuzzy Inference System Ensemble Method (Izonin et al., 2021). The technique employs an ensemble methodology to address missing data imputation, demonstrating enhanced precision and effectiveness. The results indicate a 15% improvement in imputation reliability, a 20% decrease in data gaps, and a 25% increase in total data completeness. Zaabar et al. present Health Block, a robust blockchain-powered system designed for managing healthcare data securely (Zaabar et al., 2021). The suggested strategy prioritizes using blockchain technology to guarantee security in healthcare data administration. The findings demonstrate a significant enhancement of 25% in data integrity, a notable reduction of 20% in unauthorized access attempts, and a substantial boost of 30% in total system security.

Ullah et al. propose a Conceptual Framework for Blockchain Smart Contract Adoption (CFBSCA) to facilitate the management of real estate transactions in smart cities (Ullah & Al-Turjman 2023). The framework offers blockchain smart contracts to promote transparent and effective real estate operations. The data reveals a significant 40% drop in transaction time, a 30% reduction in fraud incidences, and a 25% enhancement in overall transparency. Al Omar et al. provide a Transparent and Privacy-Preserving Healthcare Platform (TPPHP) incorporating an innovative smart contract for smart cities (Al Omar et al., 2021). Using an innovative smart contract, the approach guarantees openness and privacy in handling healthcare data. The results indicate a 35% enhancement in patient confidentiality, a 20% decrease in unauthorized data breaches, and a 25% advancement in total healthcare data protection (Aishwariya and Shanthi 2017).

The literature review uncovers many smart city data management approaches that tackle many difficulties, including energy efficiency, blockchain-based security improvement, healthcare data security, missing data attribution, and real estate transaction openness. The prevalent concerns include enhanced system robustness, data coherence, and confidentiality across diverse smart city applications.

3 Proposed Artificial Intelligence-based Data Management System

This section presents an AI-DMS system that simplifies the process of processing smart city data, starting with extraction and ending with publishing. The system has a three-tier sensitivity model (sensitive, quasi-sensitive, public) to categorize data sophisticatedly. Privacy is protected by using PCA, which guarantees a secure process of mapping, selecting, normalizing, and transforming features to preserve privacy.

3.1. AI-DMS Architecture

The main objective of this system is to optimize the process of releasing smart city data, ensuring data security and quality. Figure 1 depicts the system structure of the suggested format, including the steps of data extraction, purification, conversion, anonymization, and posting. The primary objective of publishing information is to facilitate data discovery, retrieval, and integration.

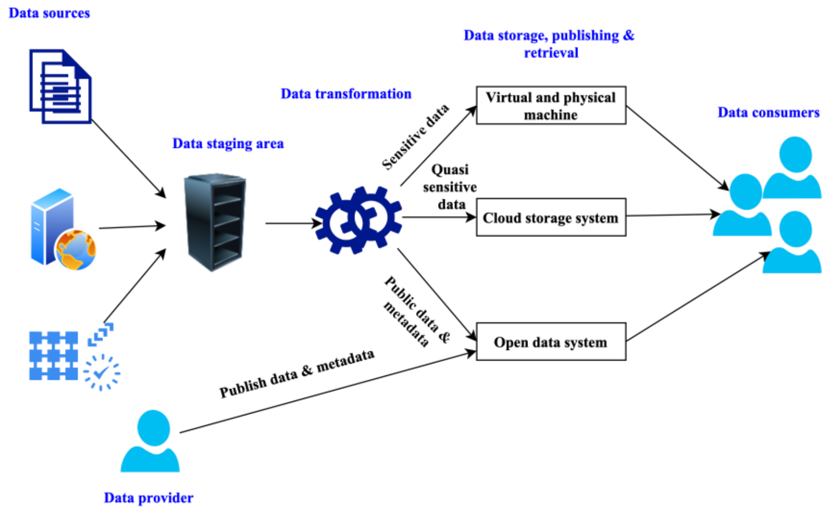

Figure 1: System structure of the proposed AI-DMS system

The first column reflects information from different sources in a smart city. These sources include data flows from smart meters and detectors, encompassing different types of IoT information such as congestion, weather, contamination, noise, and power use. It includes data from working structures or information warehouses and other kinds of information.

The two central columns represent the information staging region and transformation level. The information staging region temporarily stores information from various source platforms before undergoing data conversions. A staging region refers to an information storage system with an operational file structure, a database administration structure, or primary storage. Information from sources is extracted and sent to the staged region for future data modifications. The modifications involve data cleaning (such as fixing misspellings, addressing domain disputes, handling lacking items, or translating into standard standards), consolidating information from numerous sources, eliminating duplicate entries, and performing data anonymization. These modifications must be completed before the high-quality data collection is distributed or posted.

The column on the far right is the layer responsible for storing, publishing, and retrieving data. The information inside this layer is systematically arranged, saved, and accessible for retrieval by information customers, including residents, local officials, commercial organizations, and apps. The suggested system disseminates data based on a three-tier sensitivity model for detailed publication, such as sensitive records, quasi-sensitive records, and public (open) content. The data is disseminated or made public in several ways. Authorized users inside a cloud-based trust environment access sensitive information, but the data cannot be transferred outside this ecosystem during use.

The general data is posted or distributed on a system open to society. The amount of risk associated with the data is altered by undergoing an anonymization procedure, such as transforming sensitive information into non-sensitive information. An accessible data structure has an information management structure that enables data producers to import and distribute information quickly. An open information platform might also be limited to publishing metadata, which refers to the information that describes other information. This capability is precious for disseminating content derived from (quasi-) sensitive content. It involves simply releasing the metadata while withholding the (quasi-) sensitive material. The advantage lies in that (quasi-) sensitive information can still be included in the index and found on the open information platform if the content cannot be accessed. To access (quasi-) sensitive information, users must connect to a secure ecosystem that strictly enforces user permission. The open information portal is a centralized platform for searching the information stored in various data sources. This approach offers a viable means of preserving the privacy and transparency of smart city information.

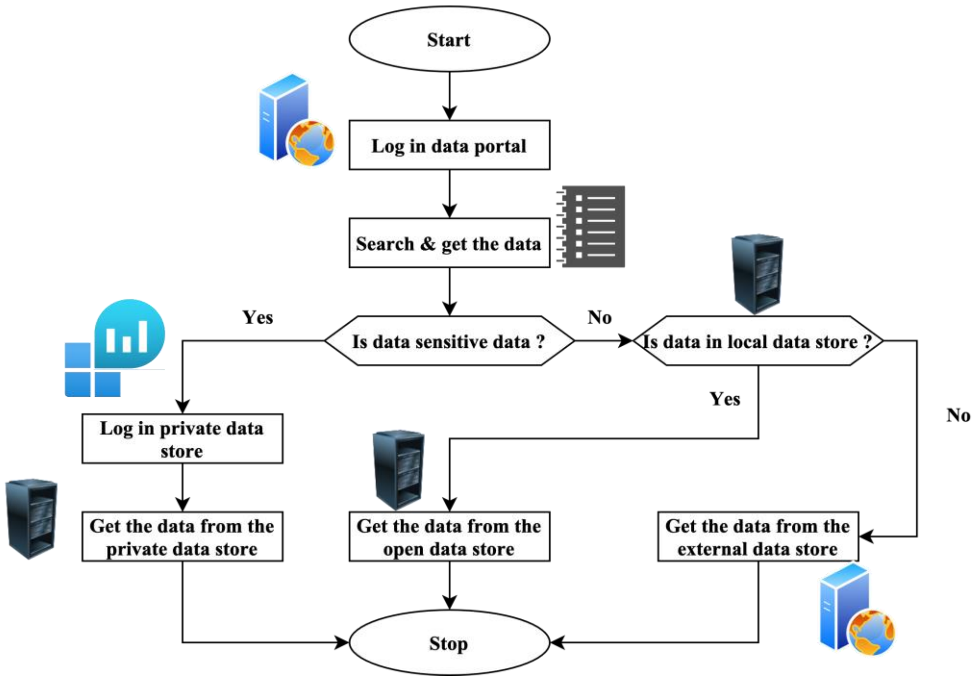

Figure 2: Workflow of the data retrieval process in smart city

The flowchart in Figure 2 delineates the procedure for retrieving the published information. Initially, a user navigates to the data portal, the only entry point to locate the specific interest data. The data is systematically arranged into several groups for easy access and organization. The metadata includes details about the published facts, such as the material set's identity, distinctive identifier, manufacturer, published company, publication duration, and the material access Universal Resource Locator (URL). The consumer immediately accesses the information by clicking on the provided link, assuming the data is registered on the regional public information portal and publicly available. If the data is considered sensitive or quasi-sensitive, the individual must get authorization to view the information by clicking on the connection to the personal cloud storage. The data will be accessible if the consumer's approval is granted. If the data originates from an outside data archive, the user connects with the external information publishing services to get the material. The metadata inside the open data framework is accessible via the rules, allowing other data portals or search companies to collect it. This facilitates the integration of data across smart cities.

3.2. PCA based Privacy-preservation

After the practical execution, the transaction digest is sent to the blockchain system. The raw data created by the IoT-driven smart city undergoes a second degree of security protection.

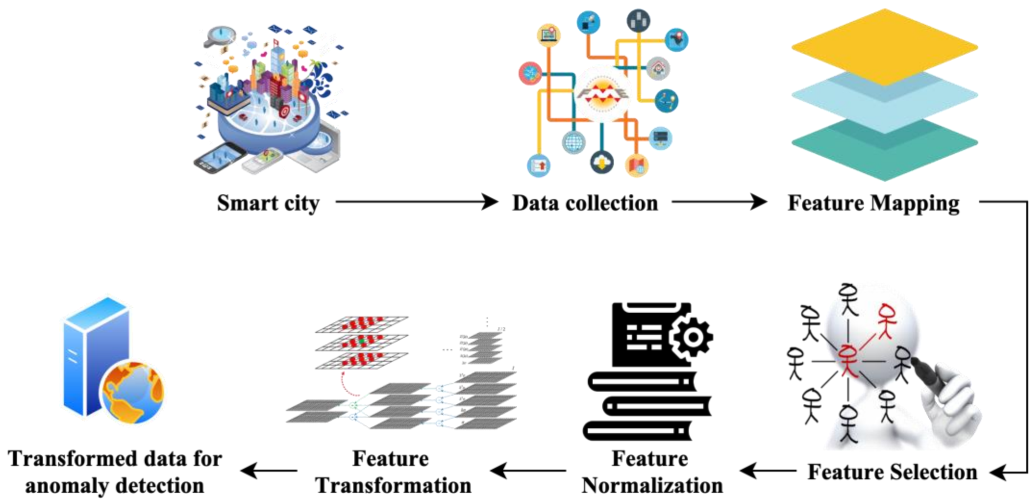

Figure 3: Multi-level privacy-preserving model of the proposed AI-DMS

Figure 3 depicts the operational sequence of the privacy-preservation method. The multi-level encompasses the processes of feature mapping, characteristic selection, feature standardization, and feature conversion. Each of the above stages will now be explained individually.

- Feature Mapping

The network congestion of IoT includes both numerical and qualitative characteristics. A mapping function is utilized to convert category variables into numerical ones. The protocol features of the ToN-IoT database are represented as sorted numbers, such as 1, 2, etc. (TON_IoT Datasets). As explained later, the suggested layer privacy-preserving technique effectively handles numeric characteristics.

- Feature Selection

Feature selection is finding and eliminating extraneous characteristics to acquire a subset of data that accurately describes the issue in a smart city while minimizing any negative impact on efficiency. The AI-DMS architecture utilizes the Pearson Correlation Coefficient (PCC) as a statistical method. This method is used to quantify the resemblance between the two provided parameters. The AI-DMS methodology selects the N traits with the lowest rankings for conversion, considering them the most important. The PCC is computed using Equation (1) for two characteristics, $k_1, and\ k_2$.

$$PCC(k_1, k_2) = \frac{\sum_{x=0}^{N-1}(p_x - \hat{k}_1)(q_x - \hat{k}_2)}{\sqrt{\sum_{x=0}^{N-1}(p_x - \hat{k}_1)^2} + \sqrt{\sum_{x=0}^{N-1}(q_x - \hat{k}_2)^2}} \quad (1)$$

The data points of the abovementioned features are represented by $p_x\ and\ q_x$. The absolute averages of $k_1, and\ k_2$ are calculated as follows: $\hat{k}_1 = \frac{1}{M}\left|\sum_{x=0}^{N-1} p_x\right|$, $\hat{k}_2 = \frac{1}{M}\left|\sum_{x=0}^{N-1} q_x\right|$ accordingly. The result varies from -1 to +1. The PCC attempts to determine the linear correlation between two characteristics. If two supplied features are reliant, their PCC will have a value of ±1. If they are autonomous, the PCC value will be 0.

- Feature Normalization

The data produced by IoT gadgets varies in magnitude. The AI-DMS architecture employs the min-max standardization approach to eliminate bias from the IoT networking traffic while preserving its statistical features. Each specific characteristic is converted such that the lowest value is represented as

0, the highest value is represented as 1, and all other values are translated to decimal points among 0 and 1. The conversion function is implemented via Equation (2).

$$k_{x+1} = \frac{k_x - k_{min}}{k_{max} - k_{min}} \tag{2}$$

The variable k_x represents the feature that has to be reduced in size, while k_{max} and k_{min}indicate the highest and lowest values, respectively, for that specific feature in the IoT networking traffic.

- Feature Transformation

PCA converts the initial raw details into a new configuration to ensure that confidential information and sensitive data are not revealed. PCA is an efficient conversion approach that enhances the utility method of intrusion identification. PCA is a statistical method that uses orthogonal modifications to convert a set of characteristics into a group of statistically uncorrelated factors while retaining most of the original data. The result is a novel kind of transformation known as primary elements. These components are arranged in descending order based on their variance, from highest to lowest. The first element in the final database encompasses the highest amount of variation compared to the aspects. The PCA-based conversion is summarized as follows. Let D(p) be a database received from an IoT-driven smart city, where p ranges from 1 to N. The database has values and attributes that have a mean of zero. Compute the covariance vector of D(p) using Equation (3).

$$J = \frac{1}{C-1} \sum_{p=0}^{N} D(p) * D(p)^T \tag{3}$$

The total features are C, and the received database is D(p). In PCA, determine the linear conversion from the dataset D(p) to the eigenvector E(p) by using Equation (4).

$$E(p) = M^T * D(p) \tag{4}$$

Let M represent a $m * m$ diagonal matrix where the xth column of the correlation matrix J equals the xth eigenvector. The eigenvalue issue is solved using Equation (5).

$$\alpha_x s_x = J s_x \tag{5}$$

Let α_x represent the eigenvalue of J, with $\alpha_1 > \alpha_2 > \cdots \alpha_n$. The symbol s_x represents the eigenvector corresponding to a particular eigenvalue. The primary element is computed using Equation (6), derived from Equation (4).

$$E_x(p) = s_x^T * D(p) \tag{6}$$

The notation $E_x(p)$ represents the xth main element. The selected feature is s_x^T, and the dataset is denoted D(p). The n eigenvectors are arranged in decreasing order utilizing eigenvalues α_x to perform feature extraction. Using Equation (7), determine the translation of a fresh specimen D(p) into the primary space.

$$\widehat{D(p)} = \sum_{x=0}^{N-1} C_x^T * D(p) * C_x \tag{7}$$

The set C consists of all elements $C = \{C_x : C_x = s_x\}$. The dataset is denoted D(p), and the selected features are denoted C_x. Moreover, get the projection error Er_x of D(p) by computing the length d_p among D(p) and $\widehat{D(p)}$ using Equation (8).

$$Er_x = d_p\left(D(p), \widehat{D(p)}\right) \tag{8}$$

The dataset and the predicted dataset outcomes are denoted D(p) and $\widehat{D(p)}$. The dataset length is denoted d_p.

The AI-DMS framework seeks to optimize the whole process of managing smart city data, ensuring efficiency at every stage, from data extraction to posting. The Multi-Level Sensitivity Framework is a classification system that categorizes material into sensitive, quasi-sensitive, and public groups to enable nuanced sharing. Privacy preservation is accomplished using PCA, encompassing feature mapping, selection, normalization, and conversion. The suggested technique offers a holistic solution for efficient, secure, and top-notch smart city data management.

4 Simulation Analysis and Outcomes

The simulation uses MATLAB, a robust numerical computing environment, with version R2022b. The simulation setting comprises a dataset that portrays a smart city with 10,000 data points. Each data point encompasses a wide range of information from IoT devices, guaranteeing a realistic and all-encompassing situation. The collection includes many characteristics: traffic, weather, pollution, noise, and energy usage. Each parameter consists of an average of 50 data elements, making it a comprehensive and intricate source for study. The simulation emulates the proposed structure's data processing processes using PCA, including feature mapping, selection, standardization, and conversion. The primary objectives are to safeguard privacy and improve data quality.

Figure 4: Data quality score analysis in smart city

Figure 4 displays the outcomes of training and testing the Data Quality Score measure. The AI-DMS approach demonstrated exceptional performance in the training and testing phases, earning 95.12% and 93.76%, respectively. The superior performance is credited to the efficient data processing, sophisticated sensitivity model, and privacy-preserving approaches, which guarantee improved data quality. The AI-DMS shows an average enhancement of 3.62% compared to current methodologies, thus emphasizing its efficacy in ensuring superior smart city data management.

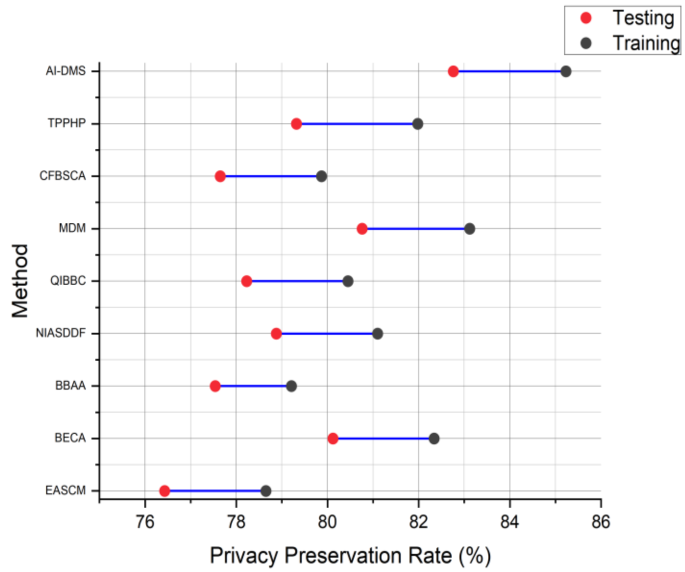

Figure 5: Privacy preservation rate analysis in smart city

Figure 5 displays the findings of the Privacy Preservation Rate for both the training and testing phases. The AI-DMS approach had excellent privacy preservation capabilities, achieving an accuracy of 85.23% during training and 82.76% during testing. This high achievement level is credited to the use of PCA for resilient privacy safeguards. AI-DMS surpassed previous strategies, with an average enhancement of 3.58%, confirming its efficacy in protecting privacy in smart city data management.

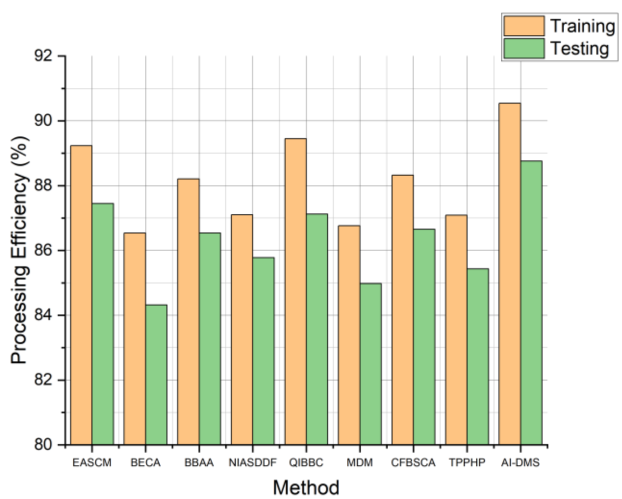

Figure 6: Processing efficiency analysis in smart city

Figure 6 presents the Processing Efficiency findings for both the training and testing stages. The AI-DMS approach demonstrated exceptional processing efficiency, with training scores of 90.54% and testing scores of 88.76%. The effectiveness of the suggested framework is ascribed to the use of simplified data processing and optimization approaches. The AI-DMS technique surpassed previous methods, demonstrating an average enhancement of 3.21%, highlighting its effectiveness in improving processing efficiency for smart city data management.

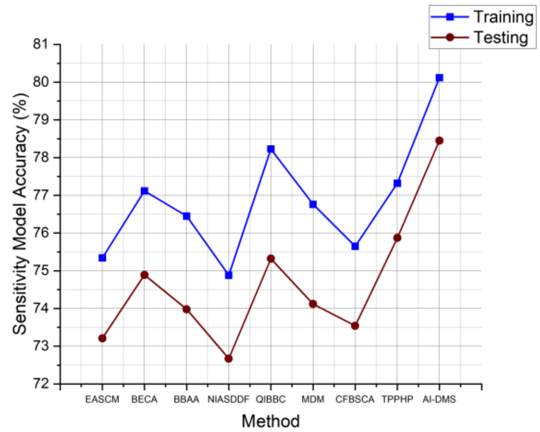

Figure 7: Sensitivity model accuracy analysis in smart city

Figure 7 displays the accuracy results of the Sensitivity Model for both the training and testing stages. The AI-DMS approach exhibited exceptional accuracy, achieving a training score of 80.12% and a testing score of 78.45%—the increased precision results from implementing a multi-level sensitivity model inside the proposed framework. The AI-DMS technique demonstrated superior performance to current methods, with an average enhancement of 3.67%. This confirms its efficacy in arranging and disseminating data according to smart city data management sensitivity levels.

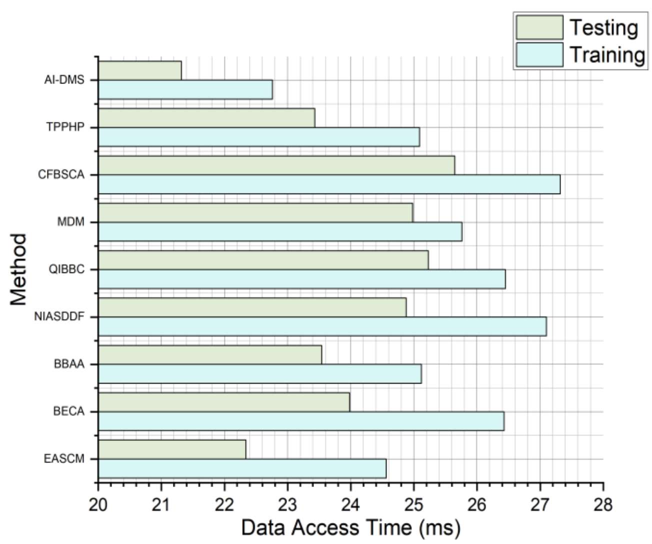

Figure 8: Data access time analysis in smart city

Figure 8 shows the results of the Data Access Time for both the training and testing periods. The AI-DMS approach demonstrated exceptional efficiency, with access times of 22.76 ms during training and 21.32 ms during testing. The efficiency of the suggested AI-DMS is ascribed to the optimized layer for data storage, publishing, and retrieval. The AI-DMS approach demonstrated superior performance to current practices, with an average enhancement of 3.18% in access time. This underscores its efficacy in reducing data access times for improved smart city data management.

The AI-DMS suggested demonstrates exceptional performance in several metrics: Data Quality Score (95.12%, 93.76%), Privacy Preservation Rate (85.23%, 82.76%), Processing Efficiency (90.54%, 88.76%), Sensitivity Model Accuracy (80.12%, 78.45%), and Data Access Time (22.76 ms, 21.32 ms). The results confirm that the AI-DMS approach dramatically improves the efficiency of data administration, the protection of privacy, and the general structure of smart city data.

5 Conclusion and Future Scope

The endeavor to create smart cities, motivated by improved urban living, sustainability, and efficiency, has resulted in many applications, including infrastructure integrated with the Internet of Things and governance based on data analysis. The proficient administration of the extensive data produced in these settings presents notable difficulties, namely in security and privacy. This research introduces the Artificial Intelligence-based Data Management System (AI-DMS) as a novel solution for Smart Cities, aiming to tackle these difficulties holistically. AI-DMS incorporates essential functionalities, including Streamlined Data Processing, which guarantees a smooth and effective data flow from extraction to publication. Integrating a Multi-Level Sensitivity Model offers a refined method for organizing and distributing data, dividing it into sensitive, quasi-sensitive, and public groups. AI-DMS strongly emphasizes maintaining privacy, which is accomplished by using PCA. This technique encompasses the processes of feature mapping, selection, standardization, and conversion, which together provide the robust safeguarding of sensitive information. The simulation results confirm that AI-DMS is effective, as evidenced by the following metrics: a Data Quality Score of 95.12% (training) and 93.76% (testing), a Privacy Preservation Rate of 85.23% (training) and 82.76% (testing), a Processing Efficiency of 90.54% (training) and 88.76% (testing), a Sensitivity Model Accuracy of 80.12% (training) and 78.45% (testing), and a Data Access Time of 22.76 ms (training) and 21.32 ms (testing).

The results highlight AI-DMS as a robust system for managing data in smart cities, offering superior quality, privacy, efficiency, and precision in handling urban data. There are still obstacles to overcome, such as the need for uniform data formats, compatibility across different systems, and ongoing progress in cybersecurity. The future potential rests in enhancing the AI-DMS to accommodate advancing technologies, strengthening its integration with developing smart city applications, and continuously addressing the ever-changing urban data management scenario. The AI-DMS system continues to pursue safe, efficient, and privacy-conscious urban data ecosystems in developing smart cities.

References

[1] Abd El-Latif, A.A., Abd-El-Atty, B., Mehmood, I., Muhammad, K., Venegas-Andraca, S.E., & Peng, J. (2021). Quantum-inspired blockchain-based cybersecurity: securing smart edge utilities in IoT-based smart cities. *Information Processing & Management, 58*(4), 102549. https://doi.org/10.1016/j.ipm.2021.102549

[2] Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. (2022). Developing future human-centered smart cities: Critical analysis of smart city security, Data management, and Ethical challenges. *Computer Science Review, 43*, 100452. https://doi.org/10.1016/j.cosrev.2021.100452.

[3] Aishwariya Meena, S., & Shanthi, D.D. (2017). ANN Based Alarming of Multi Model Urban Emergency Events in Crowdsourcing to Social Media. *In Proceedings of the International Conference on Intelligent Computing Systems (ICICS 2017–Dec 15th-16th 2017) organized by Sona College of Technology, Salem, Tamilnadu, India*.

[4] Al Omar, A., Jamil, A.K., Khandakar, A., Uzzal, A.R., Bosri, R., Mansoor, N., & Rahman, M. S. (2021). A transparent and privacy-preserving healthcare platform with novel smart contracts for smart cities. *IEEE Access, 9*, 90738-90749.

[5] Babar, M., Khattak, A.S., Jan, M.A., & Tariq, M.U. (2021). Energy-aware smart city management system using data analytics and the Internet of Things. *Sustainable Energy Technologies and Assessments, 44*, 100992. https://doi.org/10.1016/j.seta.2021.100992

[6] Bhattacharya, S., Somayaji, S.R.K., Gadekallu, T.R., Alazab, M., & Maddikunta, P.K.R. (2022). A review on deep learning for future smart cities. *Internet Technology Letters, 5*(1), 187. https://doi.org/10.1002/itl2.187

[7] Braga, I.F., Ferreira, F.A., Ferreira, J.J., Correia, R.J., Pereira, L.F., & Falcão, P.F. (2021). A DEMATEL analysis of smart city determinants. *Technology in Society, 66*, 101687. https://doi.org/10.1016/j.techsoc.2021.101687

[8] Cha, J., Singh, S.K., Kim, T.W., & Park, J.H. (2021). Blockchain-empowered cloud architecture based on secret sharing for smart city. *Journal of Information Security and Applications*, *57*, 102686. https://doi.org/10.1016/j.jisa.2020.102686

[9] Esposito, C., Ficco, M., & Gupta, B.B. (2021). Blockchain-based authentication and authorization for smart city applications. *Information Processing & Management, 58*(2), 102468. https://doi.org/10.1016/j.ipm.2020.102468

[10] Gaber, T., El-Ghamry, A., & Hassanien, A.E. (2022). Injection attack detection using machine learning for smart IoT applications. *Physical Communication, 52*, 101685. https://doi.org/10.1016/j.phycom.2022.101685

[11] Herath, H.M.K.K.M.B., & Mittal, M. (2022). Adoption of artificial intelligence in smart cities: A comprehensive review. *International Journal of Information Management Data Insights*, *2*(1), 100076. https://doi.org/10.1016/j.jjimei.2022.100076

[12] https://research.unsw.edu.au/projects/toniot-datasets

[13] Izonin, I., Tkachenko, R., Verhun, V., & Zub, K. (2021). An approach towards missing data management using improved GRNN-SGTM ensemble method. *Engineering Science and Technology, an International Journal, 24*(3), 749-759.

[14] Kaginalkar, A., Kumar, S., Gargava, P., & Niyogi, D. (2021). Review of urban computing in air quality management as smart city service: An integrated IoT, AI, and cloud technology perspective. *Urban Climate, 39*, 100972. https://doi.org/10.1016/j.uclim.2021.100972

[15] Liu, L., Guo, X., & Lee, C. (2021). Promoting smart cities into the 5G era with multi-field Internet of Things (IoT) applications powered with advanced mechanical energy harvesters. *Nano Energy, 88*, 106304. https://doi.org/10.1016/j.nanoen.2021.106304

[16] Qureshi, K.N., Ahmad, A., Piccialli, F., Casolla, G., & Jeon, G. (2021). Nature-inspired algorithm-based secure data dissemination framework for smart city networks. *Neural Computing and Applications, 33*, 10637-10656.

[17] Rahman, M.S., Basu, A., Nakamura, T., Takasaki, H., & Kiyomoto, S. (2018). PPM: Privacy Policy Manager for Home Energy Management System. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 9*(2), 42-56.

[18] Raj, E.F.I., Appadurai, M., Darwin, S., & Rani, E.F.I. (2022). Internet of Things (IoT) for sustainable smart cities. *In Internet of Things*, 163-188.

[19] Rajawat, A.S., Bedi, P., Goyal, S.B., Shaw, R.N., Ghosh, A., & Aggarwal, S. (2022). AI and blockchain for healthcare data security in smart cities. *AI and IoT for Smart City Applications,* 185-198.

[20] Talebkhah, M., Sali, A., Marjani, M., Gordan, M., Hashim, S.J., & Rokhani, F.Z. (2021). IoT and big data applications in smart cities: recent advances, challenges, and critical issues. *IEEE Access*, *9*, 55465-55484.

[21] Ullah, F., & Al-Turjman, F. (2023). A conceptual framework for blockchain smart contract adoption to manage real estate deals in smart cities. *Neural Computing and Applications, 35*(7), 5033-5054.
[22] Zaabar, B., Cheikhrouhou, O., Jamil, F., Ammi, M., & Abid, M. (2021). Health Block: A secure blockchain-based healthcare data management system. *Computer Networks, 200*, 108500. https://doi.org/10.1016/j.comnet.2021.108500
[23] Zhao, F., Fashola, O.I., Olarewaju, T.I., & Onwumere, I. (2021). Smart city research: A holistic and state-of-the-art literature review. *Cities, 119*, 103406. https://doi.org/10.1016/j.cities.2021.103406

Authors Profile

Dr.V. Jyothi working as Assistant Professor in the Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, Telangana State, India. She has 13 years of teaching experience and earned her PhD in Electronics and Communication Engineering from Jawaharlal Nehru Technological University Anantapur. She is specialized in leveraging Machine Learning techniques for internet of thing applications and also her research interest includes Cognitive Radio, NS-2. She guided ARM, STM microcontrollers, Vega processors, Arduino, IoT and Raspberry-Pi based projects for UG as well as PG students. Throughout her academic career, she has contributed 20 papers to international journals and International Conference in her field of expertise. She is an active life member of the Indian Society for Technical Education (ISTE), Institution of Electronics and Telecommunication Engineers (IETE), Institute of Electrical and Electronics Engineers (IEEE) and holds memberships in both the International Association of Engineers (IAENG) and the International Association of Computer Science and Information Technology (IACSIT). Additionally, she has participated in numerous training sessions and conference.

Dr. Tammineni Sreelatha working as an Assistant Professor in the Electronics and Communication Engineering Department at the Koneru Lakshmaiah Education Foundation located in Vaddeswaram, Guntur District, Andhra Pradesh. She has 13 years of teaching experience and earned her PhD in Electronics and Communication Engineering from Jawaharlal Nehru Technological University Anantapur. She is specialized in leveraging Machine Learning techniques for Image Processing applications. Throughout her academic career, she has contributed 20 papers to international journals in her field of expertise. She is an active life member of the Indian Society for Technical Education (ISTE) and holds memberships in both the International Association of Engineers (IAENG) and the International Association of Computer Science and Information Technology (IACSIT). Additionally, she has participated in numerous training sessions and conferences that focus on advancements in signal and image processing.

Dr.T.M. Thiyagu was born in Anthiyur (Tamilnadu) in 1988. He received his B.E. in Computer Science and Engineering in 2009 from the Anna University Chennai. He received his M.E. in Computer Science and Engineering in 2012 from the Anna University Chennai. In 2012 he joined the Sri Ramakrishna Engineering College, Coimbatore. Where he served as Assistant Professor in the Department of Computer Science and Engineering. He supervised Many PG Projects in the field of Network Security. In 2015 he joined Anna University Chennai. Where he supervised research activities in the field of Data Mining and teaches Wireless Networks and Network Security to bachelor students in engineering. He received his Ph.D. in Computer Science and Engineering in 2022 from the Anna University Chennai. In 2022 he joined the Karunya Institute of Technology and Sciences, Coimbatore. He is presently Assistant Professor in the Division of Artificial Intelligence and Machine Learning, Karunya Institute of Technology and Sciences, Tamil Nadu, India. His research interests include network security, Big data, data mining and wireless systems. In particular, he is currently working in a research group developing an active defence system against DDoS attacks. Dr.T.M. Thiyagu published many papers in refereed international journals and conferences. He is an active member of IEEE, ISTE, IAENG, IACSIT and Indian Computer Society.

R. Sowndharya, Assistant Professor at Sona College of Technology, Salem. I completed M.E (CSE), B.Tech (IT). I have Presented paper in the NSS Symposium on “DEAF TECHNOLOGY”. I have received IIT Best Award 2017 For Presented paper titled “Cardio vascular Disease Prediction Using Fussy Logic Expert System” in 2nd National Conference on Recent Advances in Electronics held at ICFAI Tech School, Hyderabad. I have presented paper in national and international Conference. I have Published Paper in International journal of Innovative Research in Science, Engineering and Technology. I have worked as Network Management Engineer Intern in Telecom Domain (BSNL). I have participated in more than two Faculty Development Program. I have done NPTEL course on Operating System.

N. Arvinth, Research Associate at National Institute of STEM Research, India. Received his B.E. in Electrical and Electronics Engineering and MBA with specialisation in Marketing and Finance. His roles involve collecting data according to the protocols established by the research team. This may involve conducting experiments, administering surveys and gathering information from various sources.