Big Data Analysis and Deep Learning in Smart Environments: Uncovering Key Technologies for an Intelligent Future

Jiahao Huang[a]
Emergency Technology and Management, Nanjing Tech University, Puzhu South Road, Nanjing, China

Keywords: Smart Environment, Deep Learning, Big Data.

Abstract: Smart City (SC) is a hot issue in the economic and social development of various countries today. In the developed digital era, big data analysis (BDA) plays an indelible role in all aspects of SC. This study pays special attention to the environmental aspect in the development of SC. This article summarizes previous research, analyzes and summarizes the application of multiple digital technologies such as BDA and deep learning (DL) in smart environment (SE) in air pollution index monitoring and urban Solid Waste Management Applications. The prediction model processed by BDA realizes the prediction of PM2.5 and other air pollutants and data visualization processing to achieve real-time monitoring of urban air pollutants. City managers can understand the air quality status through real-time monitoring data; through DL The architecture's data model enables intelligent management of urban waste, further enhancing waste processing and energy recycling. Research points out that BDA and DL still have great potential in the development of smart environments. The development of smart environments can be studied through a deep neural network model for a comprehensive and systematic model. In the era of information management, city managers use data-driven decisions to help improve the efficiency and quality of urban governance.

1 INTRODUCTION

In recent years, smart cities (SC) (Gracias et al., 2023) have been receiving widespread attention around the world. With the advancement of the times and technology, the growth of data is tending to a new peak. Statistics show that by 2050, nearly 70% of the world's population will live in cities (Mirza et al., 2024). The concept of smart cities is proposed to provide people with a more convenient and intelligent life. Governments around the world are strengthening the construction of smart cities. With the explosion of big data, people's interest in smart cities and big data analysis (BDA) (Olaniyi et al., 2023) is increasing steadily. Big data and smart cities have become two unavoidable topics. Indeed, a smart city is a comprehensive and complex system. The problems faced in smart cities from education, health, environment, transportation, management, and all aspects of life cannot be separated from the support of data. As a data-driven tool, BDA has great potential. The BDA framework supports the ability of data-driven decision-making in SC (Olaniyi, 2023).

At present, economic, social and other sustainable development issues are attracting widespread attention around the world, among which the environment is one of the most important factors (Ullo, S. L., and Sinha, G. R., 2020). And the growth of population migration to cities has led to excessive density of residents and increased urban economic burdens. Automobilization has caused traffic problems. Technological development has led to changing requirements of residents and enterprises for the quality and capabilities of the urban environment. All of this ultimately leads to increased environmental pressure and the emergence of environmental problems (Turgel et al., 2019). Therefore, it is particularly important for the construction of smart environment in smart cities. At this stage, there are many studies using big data processing platforms, machine learning methods and Artificial Intelligence (AI) intelligent technology to analyze and process smart environments including air pollution index prediction, urban sewage system

[a] https://orcid.org/0009-0006-6895-1023

Huang, J.
Big Data Analysis and Deep Learning in Smart Environments: Uncovering Key Technologies for an Intelligent Future.
DOI: 10.5220/0012909100004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 80-84
ISBN: 978-989-758-713-9

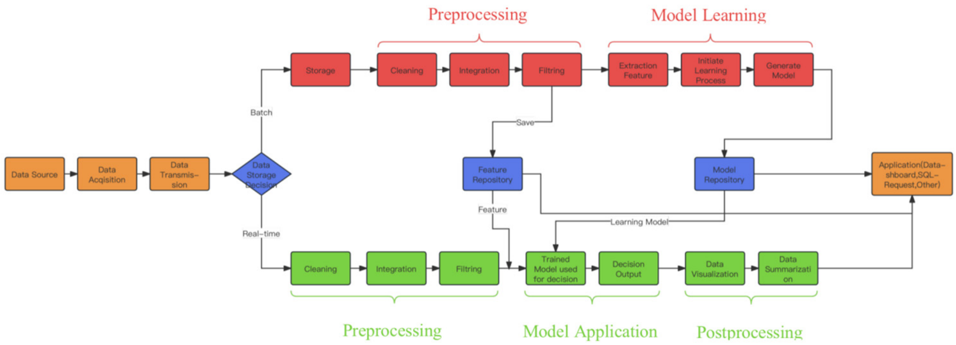

Figure 1: BDA in smart environment (Photo/Picture credit: Original).

construction and real-time monitoring of traffic conditions. Smart cities must provide a unified framework to manage and control these services in an intelligent manner and, where possible, facilitate integration between these services. One service can have a significant impact on the quality of other services. Shih et al has established six integrated prediction models by collecting PM2.5 sensor data from the open-source community LASS to predict Taiwan's PM2.5 concentration in the next 30 to 180 minutes at a 30-minute frequency (Shih et al., 2021). In terms of taking the prediction of PM2.5 as an example, some research uses Gaussian SVM (Bhuvaneshwari et al., 2022) to predict air pollutant concentrations, classifying polluted areas into heavy traffic and not heavy traffic, this method only divides the areas with heavy traffic or not, and does not integrate the real-time traffic data of an area into the air pollutant pre-detection model. There is no intuitive and usable model for comprehensive preprocessing of multi-faceted data. Judging from current research, there are few studies that can provide a unified overview of such a complex and large system.

This article aims to review the views of latest studies on smart environment, analyzes the current development issues of smart environment, and comprehensively introduce the multiple applications of digital technologies such as BDA and machine learning methods in the field of smart environment, and propose innovative solutions to provide reference.

2 METHOD

2.1 The Framework of Big Data Technologies for the Environment in Smart Cities

The application of big data analysis in smart environments refers to the use of big data technology and analysis methods to process large amounts of data generated in smart environments, thereby achieving intelligent management and optimization of smart environments and helping people better understand the operation of the environment. Figure1 Demonstrates how static and real-time data flow in information systems, including how data flows through information systems, how external processes or entities create or use data, and how data is stored. The information system first collects data from the smart city environment through the data transmission mechanism or directly calls the data from the database for data preprocessing. The BDAFC architecture adopts a hybrid approach to process historical data and real-time data, while utilizing multiple learning algorithms to implement decision-making tasks such as classification, association rules, anomaly detection, and prediction. Data preprocessing methods are divided into 3 categories: data cleaning, data integration, and data filtering. Data preprocessing improves data quality by performing multiple tasks and activities such as data conversion, merging, and standardization, and then visualizes and summarizes the data to the application.

2.2 Air Pollution Index Monitoring

Air pollution problems are becoming increasingly serious, air quality is deteriorating, and air pollution levels in urban areas are rising, leading to many serious environmental problems in many cities. Adequate air quality prediction has become a serious challenge to population health, and the impact of air pollutants on Real-time monitoring is of great significance to the management of air environment.In recent years, big data science and technology has developed vigorously. People have built various models of big data in various aspects to find the patterns of problems and methods to solve problems. At this time, big data analysis is used to monitor air pollutants in real time and improve the air environment. The governance of the country plays an important role that cannot be ignored.

Shih et al. (Shih, et al., 2021) used PM2.5 sensor data from the open-source community LASS as input variables, and used the Spark computing framework and integrated machine learning methods to train the PM2.5 concentration prediction model, then evaluated the model through the evaluation indicators RMSE and R2. For prediction performance, the data visualization tool PlotDB was finally used to generate a map of PM2.5 concentration value distribution. Its research also found that using ensemble models with different time periods and other variables can improve forecasting performance. However, the input variables of the prediction model only include PM2.5, PM10, PM1, temperature and relative humidity. It has not been possible to design a unified model for various variables related to PM2.5 such as rainfall, wind direction and traffic data. predict. This is nothing more than a huge test for the data processing capabilities, which will require the update of more powerful algorithms and more powerful computing power to process the huge data in the context of smart environments. Myeong et al. (Myeong and Shahzad et al., 2021) used machine learning and deep learning methods to solve problems related to air pollution monitoring. They used Gaussian SVM to predict air pollutant concentrations and classify polluted areas into heavy and light traffic congestion areas. Huang et al developed a deep neural network model, the CNN-LSTM model (Huang et al., 2018), to learn and predict the PM2.5 concentration in the next hour through historical data (such as cumulative hourly rainfall, cumulative wind speed and PM2.5 concentration).

Different studies have proposed different monitoring models for various variables related to PM2.5 prediction. Taking into account the same differences in the models, a comprehensive BDA model is studied for real-time prediction of PM2.5.

2.3 Waste Management in Smart Cities

As the air pollution index continues to rise, waste management in smart cities has become a burning problem.In recent years, as urbanization, income and consumption have increased, so has the generation of waste.It is estimated that the amount of global waste is expected to increase to 2.8 billion tons by 2050 (Szpilko et al., 2023). The implementation of new waste monitoring and management systems has become an important area for the sustainable development of smart environments.

Sana Shahab et al (Shahab et al., 2022) found that compared with traditional machine learning and image processing methods, DL models provide more advanced technology in smart waste management (SWM) with significant effects and efficiency. Therefore, deep learning models have gained enough momentum in the SWM solid waste management research community to solve many problems. Patric Marques, et al. (Marques et al., 2019) implemented waste management in smart cities by implementing a smart waste management architecture based on indoor and outdoor waste management scenarios with three different protocols designed to provide secure communication - message queues Telemetry Transport Protocol (MQTT), Constrained Application Protocol (CoAP) and Hypertext Transfer Protocol Secure (HTTPS) are implemented together. Evaluation metrics such as energy consumption, latency, jitter, and throughput were considered when evaluating the system. Based on the results obtained, scalability was also analyzed, taking into account the impact on the system of increasing the number of concurrent bins in the system.

2.4 BDAs Application Analysis of Air Pollution Index and Waste Management

First, in response to the air pollution problem, researchers used big data technology and machine learning algorithms to build an air quality prediction model. Take the study by Shih et al. as an example. They used PM2.5 sensor data from the open source community LASS and trained a PM2.5 concentration prediction model through the Spark computing framework and integrated machine learning methods. The model evaluates its predictive performance by evaluating indicators such as RMSE and R2, and uses data visualization tools to generate maps of the distribution of PM2.5 concentration values, providing an important reference for decision-makers. On the other hand, in the field of waste management, big data analysis also shows great potential. Researchers have made significant progress in smart waste management through big data analysis technology, especially deep learning models. Compared with traditional machine learning and image processing methods, deep learning models show higher efficiency and effectiveness in smart waste management.

Although big data analysis has made some progress in air pollution index monitoring and waste management in smart cities, it still faces many challenges. First, the quality and completeness of data are critical to the accuracy and reliability of the model. Secondly, model complexity and computational resource requirements limit the possibility of large-scale applications. In addition, privacy and security issues also need to be fully considered. In the future, we can address these challenges by further improving algorithms and models, improving data quality and integrity, strengthening data sharing and cooperation, and strengthening privacy protection and security measures. At the same time, interdisciplinary cooperation is also the key to promoting smart city environmental monitoring and management, combining professional knowledge and technology in computer science, environmental science, sociology and other fields to jointly promote the sustainable development of smart cities.

3 DISCUSSIONS

Environmentally sustainable smart cities are a rapidly developing trend (Bibri et al., 2023). It can be seen from the above research that with the advent of the digital age, intelligent management of cities and intelligent monitoring of the environment are particularly important. With the advent of the digital age and the intelligent management of cities, intelligent monitoring of the environment has become particularly important, including air pollution index monitoring, urban waste management, etc. Relying on powerful computing algorithms, it can be analyzed real-time data of the environment. However, it can also be seen from a large number of studies that in the smart environment, Construction-wise, people still have a long way to go.

The PM2.5 prediction model studied by Shih et al (Shih et al., 2021) has only five input variables, namely PM2.5, PM10, PM1, temperature and humidity, excluding other variables related to PM2.5, such as rainfall, Wind direction and traffic data. K.S.Bhuvaneshwari et al. (Bhuvaneshwari et al., 2022) used Gaussian vector computer to roughly consider traffic data into PM2.5 prediction, but there are few studies on a comprehensive and systematic model. However, in the face of such a huge amount of data, a model for real-time prediction of environmental monitoring will undoubtedly require a very large algorithm to process data from many aspects in real time. The update of computing power may become a major limitation in the subsequent development of smart environments.

Data show that 700-900 million tons of waste are produced globally every year (Chen et al., 2020). Especially in developing countries, SWM has received more and more attention in recent years. Real-time monitoring and effective management of urban waste have become the key to smart environment construction. an important topic. Lin, Kunsen, et al. (Lin et al., 2022) reviewed the application of deep learning in solid waste management and used different algorithm models such as ANN, CNN, RNN/LSTM, Attention and GAN to recycle waste. The DL model has been used in the field of SWM. Four main applications include waste detection, identification, bin level detection and waste generation prediction (Shahab et al., 2022; Lin K et al., 2022). The use of deep learning algorithms to solve solid waste problems has great potential, but the application of deep learning in solid waste management is still at a relatively preliminary stage and requires further development. In the future development process, the use of deep learning algorithms to analyze historical and real-time data can be studied to predict the amount and type of solid waste generated in the future. This helps plan and optimize the construction and operation of waste treatment facilities, realize automated control and intelligent management of waste treatment equipment, improve processing efficiency and reduce energy consumption.

4 CONCLUSIONS

Good environmental quality is an important reflection of a city's image. This article mainly studies the application of BDA in smart environments. Through a review of previous research, we can understand that BDA penetrates into all aspects of smart cities in today's digital information age. This article specifically focuses on air The prediction of pollution index and the management and regulation of municipal solid waste are described. The BDA model for monitoring and prediction of multiple air pollutants such as PM2.5 uses a deep neural network model to combine CNN and LSTM architecture to predict PM2.5 concentration. Perform real-time monitoring; at the same time, deep learning models also play an important role in urban waste management. Research has found that DL provides more advanced technology in waste management, making waste identification and control more efficient. Through BDA's research, city managers can understand air quality conditions and identify solid waste through real-time monitoring data, make decisions based on scientific data, and take corresponding governance measures to improve air quality and protect citizens' quality of life. The above data-driven decision-making can help improve the efficiency and quality of urban governance, promote the development of circular economy, help reduce resource consumption and environmental load, promote the transformation of cities into a circular economy model, and achieve sustainable development goals.

REFERENCES

Bhuvaneshwari, K. S., Uma, J., Venkatachalam, K., Masud, M., Abouhawwash, M., & Logeswaran, T. 2022. Gaussian Support Vector Machine Algorithm Based Air Pollution Prediction. Computers, Materials & Continua, 71(1).

Bibri, S. E., Alexandre, A., Sharifi, A., & Krogstie, J. 2023. Environmentally sustainable smart cities and their converging AI, IoT, and big data technologies and solutions: an integrated approach to an extensive literature review. Energy Informatics, 6(1), 9.

Chen, D. M. C., Bodirsky, B. L., Krueger, T., Mishra, A., & Popp, A. 2020. The worlds growing municipal solid waste: trends and impacts. Environmental Research Letters, 15(7), 074021.

Gracias, J. S., Parnell, G. S., Specking, E., Pohl, E. A., & Buchanan, R. 2023. Smart Cities—A Structured Literature Review. Smart Cities, 6(4), 1719-1743.

Huang, C. J., & Kuo, P. H. 2018. A deep CNN-LSTM model for particulate matter (PM2. 5) forecasting in smart cities. Sensors, 18(7), 2220.

Lin, K., Zhao, Y., Kuo, J. H., Deng, H., Cui, F., Zhang, Z., ... & Wang, T. 2022. Toward smarter management and recovery of municipal solid waste: A critical review on deep learning approaches. Journal of Cleaner Production, 346, 130943.

Marques, P., Manfroi, D., Deitos, E., Cegoni, J., Castilhos, R., Rochol, J., ... & Kunst, R. 2019. An IoT-based smart cities infrastructure architecture applied to a waste management scenario. Ad Hoc Networks, 87, 200-208.

Mirza, N. M., Ali, A., & Ishak, M. K. 2024. The scheduling techniques in the Hadoop and Spark of smart cities environment: a systematic review. Bulletin of Electrical Engineering and Informatics, 13(1), 453-464.

Myeong, S., & Shahzad, K. 2021. Integrating data-based strategies and advanced technologies with efficient air pollution management in smart cities. Sustainability, 13(13), 7168.

Olaniyi, O., Okunleye, O. J., & Olabanji, S. O. 2023. Advancing data-driven decision-making in smart cities through big data analytics: A comprehensive review of existing literature. Current Journal of Applied Science and Technology, 42(25), 10-18.

Shahab, S., Anjum, M., & Umar, M. S. 2022. Deep learning applications in solid waste management: A deep literature review. International Journal of Advanced Computer Science and Applications, 13(3).

Shih, D. H., To, T. H., Nguyen, L. S. P., Wu, T. W., & You, W. T. 2021. Design of a spark big data framework for PM2. 5 air pollution forecasting. International journal of environmental research and public health, 18(13), 7087.

Szpilko, D., de la Torre Gallegos, A., Jimenez Naharro, F., Rzepka, A., & Remiszewska, A. 2023. Waste management in the Smart City: Current practices and future directions. Resources, 12(10), 115.

Turgel, I., Bozhko, L., Ulyanova, E., & Khabdullin, A. 2019. Implementation of the smart city technology for environmental protection management of cities: The experience of Russia and Kazakhstan. Environmental and Climate Technologies, 23(2), 148-165.

Ullo, S. L., & Sinha, G. R. 2020. Advances in smart environment monitoring systems using IoT and sensors. Sensors, 20(11), 3113.