*applied sciences*

*Article*

# An ML-Based Solution in the Transformation towards a Sustainable Smart City

**Izabela Rojek [1,*], Dariusz Mikołajewski [1], Janusz Dorożyński [1], Ewa Dostatni [2] and Aleksandra Mreła [1]**

[1] Faculty of Computer Science, Kazimierz Wielki University, Chodkiewicza 30, 85-064 Bydgoszcz, Poland; dariusz.mikolajewski@ukw.edu.pl (D.M.); janusz.dorozynski@ukw.edu.pl (J.D.); aleksadra.mrela@ukw.edu.pl (A.M.)
[2] Faculty of Mechanical Engineering, Poznań University of Technology, Marii Skłodowskiej-Curie 5, 60-965 Poznań, Poland; ewa.dostatni@put.poznan.pl
* Correspondence: izabela.rojek@ukw.edu.pl

**Featured Application: Potential applications of the work include novel ML-based systems for sustainable smart cities and smart territory control.**

**Abstract:** The rapid development of modern information technology (IT), power supply, communication and traffic information systems and so on is resulting in progress in the area of distributed and energy-efficient (if possible, powered by renewable energy sources) smart grid components securely connected to entire smart city management systems. This enables a wide range of applications such as distributed energy management, system health forecasting and cybersecurity based on huge volumes of data that automate and improve the performance of the smart grid, but also require analysis, inference and prediction using artificial intelligence. Data management strategies, but also the sharing of data by consumers, institutions, organisations and industries, can be supported by edge clouds, thus protecting privacy and improving performance. This article presents and develops the authors' own concept in this area, which is planned for research in the coming years. The paper aims to develop and initially test a conceptual framework that takes into account the aspects discussed above, emphasising the practical aspects and use cases of the Social Internet of Things (SIoT) and artificial intelligence (AI) in the everyday lives of smart sustainable city (SSC) residents. We present an approach consisting of seven algorithms for the integration of large data sets for machine learning processing to be applied in optimisation in the context of smart cities.

**Keywords:** artificial intelligence; machine learning; data processing; smart sustainable city; Social Internet of Things; 6G; Industry 5.0

check for updates

**Citation:** Rojek, I.; Mikołajewski, D.; Dorożyński, J.; Dostatni, E.; Mreła, A. An ML-Based Solution in the Transformation towards a Sustainable Smart City. *Appl. Sci.* **2024**, *14*, 8288. https://doi.org/10.3390/app14188288

Academic Editors: Paolo Renna, Pedro Couto, Maja Trstenjak, Hrvoje Cajner and Tihomir Opetuk

Received: 31 July 2024
Revised: 5 September 2024
Accepted: 12 September 2024
Published: 14 September 2024



## 1. Introduction

The advancement of artificial intelligence, eHealth, information and communication technologies (ICT) and the Internet of Things (IoT) has caused the mindset of consumers (of services, products, water, energy and other goods) to change from being completely passive towards active monitoring of and even participation in the market. This is creating new smart city communities focused on sustainability and lowering the cost of living. Key factors shaping these attitudes are identified, and the relationship between them and socio-economic and behavioural factors is shown [1]. There is a need for further research and tests to modify individual factors and groups of factors. The social networks of today's cities are strongly and increasingly influenced by new ICT. Its integration into urban operations has fostered the development of information cities, ease of communication and the creation of smart communities. IoT applications and especially the Social IoT (SIoT) have resulted in the emergence of smart cities (SCs) supporting urban operations with minimal human intervention and often even minimal human interaction. This means that data generated by SC residents (also travellers, etc.) and collected by the SIoT can be used to create new smart services within the SC paradigm [2]. We are thus channelling dispersed people into virtual communities centred around work, interests or lifestyles. This can take forms distinct from the previous behaviour within a place of residence (street, quarter, neighbourhood) or workplace. The place of residence or work may become less about social class or even ease of commuting (nursery or children's schools, shops) and more about a conscious choice based on the financial situation and the ability to afford a particular one. Learning, working and providing services (banking, commercial, health and other) online is slowly becoming the norm. The concept of local and global trust in the provision of services in SC based on online social networks (OSNs) has emerged. The local trust model allows you to determine the so-called central user within the local cluster. The global trust model allows you to indicate the so-called opinion leaders. Filtration of the above-mentioned central users and opinion leaders results in the dispersion and significant limitation of networks such as Facebook or X. However, data posted by citizens on OSNs can enable the creation and promotion of new services, thus mitigating the impact of untrusted users [2].

Industry 5.0 is focused on people and the environment; therefore, it seems natural to integrate low-emission or zero-emission industry (production, services, etc.) into the human environment, so it can be included not only in smart factories, sustainable smart cities and smart territories, but also in the 15-min city paradigm. This places enormous demands on modern industrial solutions, service infrastructure and management systems, including the optimisation of operation, energy consumption and carbon footprint. The goal of the SC, since its appearance at the beginning of the 21st century, has been to develop optimal care for its inhabitants through the implementation of breakthrough technologies. The meaning of SC expands and becomes more precise as concepts, implementations, new technologies and trends develop, up to the concept of smart territories [1]. These trends have been changing: they started with the transformation of vast metropolitan areas, but their vastness and the dynamics of the surroundings result in the implementation of intelligent technologies only in a small proportion of the processes, and the progress generated by this may be unnoticeable to a large proportion of the inhabitants. The current stage involves focusing on smart micro-territories and satellite cities. The current global population is 8 billion, and estimates for 2050 say it will reach 9.7 billion, concentrated in and around cities, with their overpopulation threatening the quality of life of residents. This forces the optimisation of solutions in the SC, including towards sustainable use of resources that offset the effects of economic crises, uneven access to goods, contamination, diseases, and ineffective services and processes [3]. Aggregation, anonymisation, pre-processing and cascaded learning in machine learning systems remain major challenges in smart home environments, especially in areas such as health monitoring, energy consumption management and mobility/transportation. These tasks need to be performed efficiently and cost-effectively while ensuring sustainability, but current methods often fail to meet these goals. The complexity of integrating different data sources and ensuring privacy while maintaining system performance highlights a critical research gap. Filling this gap is critical to developing smart home capabilities to be both smart and practical. As the demand for smart home technology increases, solving these challenges becomes more urgent. The aim of this article is to develop and initially test the concept of an AI solution taking into account the research gap and research problem discussed above, emphasising practical aspects and cases of using the SIoT and AI in the everyday life of SSC residents (health, sustainable energy use, sustainable transport and mobility of residents/SSC users). The challenge of developing simple, transparent, low-cost and fast solutions for smart home data cascading remains unsolved. Current approaches are either based on simple algorithms tailored to specific types of data, limiting their generalisability, or on complex deep learning models that lack transparency and struggle to meet evolving legal standards, especially for sensitive data such as emotions. These deep learning methods, while powerful, are often opaque and cannot always accommodate the diverse data streams in the smart home. As a result, they do not offer a comprehensive solution that balances efficiency, transparency and compliance. This ongoing issue highlights an important gap in the development of effective smart home technologies. This will constitute the basis for further development of the proposed solution, both in the areas of horizontal scaling (adding more sensors, effectors and analysis segments) and vertical scaling (greater emphasis on edge computing, as well as delegating authority to lower decision-making levels despite maintaining a reliable overall image at cloud level).

The novelty of the work lies in defining from scratch the structure for analysing real data (from data sets regarding health, renewable energy sources and energy management, as well as mobility and transport of residents/users) and the algorithms used, as well as creating the basis for further development of this group of algorithms and solutions for SSC.

The contribution of our solution is not only the expansion of application possibilities in the area of SSC, but also fundamental research on new families of simple algorithms capable of quickly learning based on initial data sets, learning based on new data, and adding new segments without losing reliability and redundancy. This seems to be the strength of our solution: scaling horizontally instead of vertically, whenever possible, while avoiding deep learning that is difficult for humans to understand. With the help of ML, it is possible to reduce computational complexity, energy intensity, environmental impact and carbon footprint by choosing between cascading processing with simple ML algorithms and deep learning. The chosen solution is determined by the complexity of the running system and the data set being analysed, as well as the selection priorities, which are often conflicting (e.g., speed of operation vs. low energy consumption). Difficult analyses (controlling temperature, monitoring safety, performing supervisory or maintenance activities) can be simplified and performed completely independently by ML under human supervision. The ultimate criterion still seems to be to improve/maintain people's quality of life; otherwise, they will not want to live/stay in an SSC.

Such a defined contribution and research gap should be highlighted in relation to the current state of knowledge. AI is the cornerstone of the sustainable smart city concept, enabling more efficient resource management, improved public services and a higher quality of life while addressing environmental challenges. The continued development and integration of AI technologies will be crucial to the future of sustainable urban development. The development of artificial intelligence (AI) in the context of sustainable smart cities (SSCs) is progressing rapidly, playing a key role in optimising city life, reducing environmental impact and improving quality of life for residents (Figure 1).

We are looking for AI systems that interact with humans in their immediate environment, not replace them. Hence, the solution to the research problem will be a cascade of a group of simple AI algorithms: understandable and friendly, also in terms of service and functionality. The excessive complexity of AI systems creates a sense of threat and resistance, and the number of people who do not understand how basic devices work should not increase.

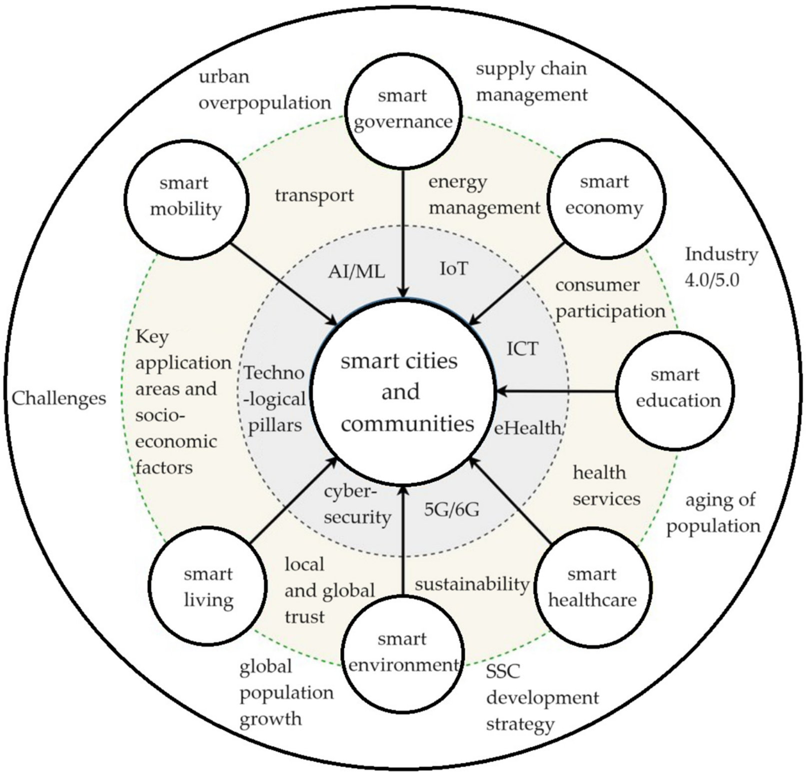


**Figure 1.** Framework of smart city evolution and AI integration (own version based on [4,5]).

## 2. AI in SSC

AI and the IoT enable the optimisation of citizen-centric SCs by

- Performing collection, analysis, inference and prediction from large amounts of data describing residents, services and facilities of the SC;
- Process automation;
- Improving efficiency and supporting the economy;
- Securing and creating opportunities for residents;
- Respecting the environment and its resources [3].

### *2.1. General Picture of AI in SSCs*

AI is playing an increasingly important role in the development of SSCs, supporting resource management, energy efficiency improvements and environmental monitoring. AI enables the optimisation of urban processes such as public transport, waste management or air quality control, contributing to the sustainability of cities. AI-based systems also support real-time data analysis, enabling faster responses to problems such as traffic jams or infrastructure failures. The developing AI technologies in SSCs are helping to integrate different city systems, leading to a more holistic approach to city management. At the same time, AI is becoming crucial in areas such as public health, where it can monitor and predict the spread of disease or support the care of seniors. As cities become smarter, AI is helping to drive data-driven decision-making, leading to more efficient use of resources. However, the development of AI in SSCs brings challenges, such as the need to protect privacy, transparency of algorithms and regulatory compliance. It is also important that AI technologies are accessible and understandable to all residents to avoid digital exclusion. As technology advances, AI has the potential to become a cornerstone of sustainable urban development, but this requires further research and innovation to maximise benefits and minimise risks. Currently, SC goals are generally focused on striving for

- Better education and work/career opportunities and achieving work–life balance (as part of the so-called individual success);

- Higher ecological performance (including natural diversity);
- Support of the health of residents, including support based on preventive medicine, faster and more accurate diagnoses and timely healthcare services, and the well-being of residents;
- Support of the social activity of residents as a community.

This means that SC 'in the background' supports economic, social and health activities while balancing budget, risk and technological upgrades. By monitoring, analysing and responding (both to correct operations and to anomalies such as failures) in real or near real time, smart city processes can be streamlined and optimised. This will become increasingly easier as platforms develop, standardise and the price of sensors, effectors and other IoT components (including cloud services and edge computing) falls [6]. An increase in available computing power at a much lower cost is expected, which is very beneficial for the broader use of AI [7]. Detailed visualisations of predictions, such as paraconsistent analysis, are needed, creating a framework for the use of other systems, both collecting data and carrying out the resulting analysis in SC [7]. The challenge here is the technological heterogeneity of SCs, which use very different approaches, protocols, methods and communication technologies (also depending on their manufacturers and SC implementers) (Table 1) [8,9].

**Table 1.** Developing status of AI in SSC.

| Area | Subarea | Tasks |
|---|---|---|
| Urban planning and development | Simulation and modelling | AI helps simulate urban development scenarios, enabling planners to assess the sustainability of different projects and optimise the use of space, resources and energy. |
| | Building smart infrastructure | AI helps in the design and management of smart buildings that optimise energy consumption, increase user comfort and reduce environmental impact. |
| Energy management and efficiency | Smart grids | AI is used to manage smart grids that optimise energy distribution and consumption in real time. This helps to reduce energy waste and integrate renewable energy sources more effectively. |
| | Predictive maintenance | AI algorithms predict when infrastructure, such as power lines or renewable energy equipment, will require maintenance, thus preventing downtime and reducing operational costs. |
| Transport and mobility | Traffic management | Traffic management systems based on artificial intelligence optimise traffic flow, reducing congestion and emissions. These systems can adjust traffic signals in real time based on traffic patterns. |
| | Deployment of autonomous vehicles | Self-driving cars and public transport systems based on artificial intelligence are being tested and deployed to reduce congestion and carbon emissions. |
| Water management | Intelligent water systems | AI monitors water quality and consumption, ensuring efficient distribution and reducing water waste. It also helps detect leaks and manage water resources during droughts. |
| | Flood prediction | AI models are used to predict flooding, enabling cities to take proactive measures to minimise damage and protect citizens. |
| Air quality monitoring | Real-time air quality monitoring | AI systems continuously monitor air quality, providing real-time data to help identify sources of pollution and take immediate action to mitigate them. |
| | Pollution prediction | AI can predict pollution levels based on weather patterns and human activity, enabling cities to implement preventative measures. |

**Table 1.** *Cont.*

| Area | Subarea | Tasks |
| --- | --- | --- |
| Waste management | Intelligent waste collection | AI is used in waste management to optimise collection routes, reducing fuel consumption and operational costs. AI can also help sort waste more efficiently, increasing recycling rates. |
| | Waste prediction models | AI models predict trends in waste generation, helping cities to better plan waste treatment and recycling. |
| Climate resilience | Climate modelling | AI models help cities understand the long-term impacts of climate change and develop mitigation strategies. |
| | Disaster response | AI is used in disaster management systems to predict natural disasters, such as earthquakes or hurricanes, and effectively coordinate response actions. |
| Other areas | Smart governance | AI is used to analyse data from citizen feedback and social media to improve public services and respond more effectively to citizens' needs. |
| | Public safety | AI enhances public safety with surveillance systems that can detect and respond to criminal activity or accidents in real time. |
| | Resource allocation | Artificial intelligence helps optimise the allocation of urban resources, ensuring that sustainability goals are achieved without compromising efficiency. |
| | Big Data analysis | AI processes vast amounts of data from a variety of city sensors and IoT devices, providing insights that help city managers make informed decisions about sustainability initiatives. |

Basic requirements for the current SC include

- Careful planning, taking into account the objectives of the SC and its inhabitants;
- Differentiation according to geographical, demographic, economic, social, etc., factors;
- Collection and analysis of data to extract valuable knowledge associated with location and navigation strategies;
- Optimal response of the SC according to social rules (cyberdemocracy);
- Sustainable management of resources;
- Strong and comprehensive ICT platform;
- Cybersecurity;
- Dynamic modernisation at different levels and time horizons of operational improvements;
- The search for new models, solutions and technologies (self-developing SC) [10–12].

Redundancy (and, indirectly, heterogeneity) of systems is already stardom. SC in its daily work should not depend on the correct functioning of a single platform, but rather on the management systems of a group of such platforms. Therefore, it must be possible to combine smart city platforms with other types of platforms and management systems for such scalable integrated group platforms [6]. Such platforms must be efficient, scalable, flexible and integrable. The challenge is to create smart systems to simultaneously manage smart territories, i.e., urban and rural areas for services targeting users with different needs. Residents do not need to be knowledgeable (or even aware) of the smart systems they use but must be able to rely on the reliability of the technological infrastructure [7,13].

### *2.2. AI in Selected Areas of SSCs*

Mobility within SCs is a problem where poor management will lead not only to traffic congestion and protests by residents, but also to a waste of fuel, loss of residents' time and their employers' time (e.g., due to lateness and absence from work) and delays in the supply chain (and, consequently, a shortage in products and services). So far, electric cars, car sharing or public transport as integrated transport solutions have seemed to be the right solution that could reduce traffic on the roads and reduce the emission of harmful gases into the atmosphere (i.e., reducing smog), often while maintaining the privacy of users [11,12]. With an ageing population, supporting elderly people's mobility and day-to-day activities, and preventing falls, require monitoring of residents without disrupting their daily activities or making them feel like they are being watched [6]. It seems that a responsible and well-thought-out combination of urban innovations will lead to the creation of effective SSCs that shape the comfort of, and have a positive impact on, the lives of their residents [13]. An underestimated aspect is the durability of SSCs despite their ongoing, required evolution in areas as important as transport and preventive medicine [3]. With the development of ICT (new generation, sixth generation communication system [6G]) and modifications in ways of life and methods of communication, the continuous evolution of urban environments has led to the emergence of the concept of a smart sustainable city (SSC). This will further connect city infrastructure (including vehicles, robots and sensors) in order to communicate, respond and act in real or near real time with high efficiency of city operations and services. 6G candidate technologies primarily include terrestrial networks, advanced mobile edge computing, vision-enhanced wireless communications, artificial intelligence (AI)-based wireless communications, and integrated sensing and communications [14]. This makes it possible for the SSC to constantly adapt the concept to the conditions [15]. Participation indicators, open innovations (not only technological, but also social) should be promoted in the SSC within social networks [16]. Currently, data aggregation within SSCs focuses on specific application areas, classifications or predictions, but the goal is to create conditions for a multi-layered "common operational picture" (military analogy) for imaging, analysis and management of data:

1. data from various IoT sources:
    - Scattered;
    - Heterogeneous;
    - Non-linear;
    - Monitoring and tracking objects;
2. data analysed using various methods:
    - Mathematical (including probabilistic);
    - Computational (including artificial intelligence) [17].

This requires energy-efficient communication along with quality of service (QoS) optimisation [17]. The transformation towards sustainable AI-based management of renewable energy sources is becoming a reality, including for energy cooperatives, which is of great importance for SSCs. AI enables faster responses, optimised operations, improved efficiency, lower costs and a shift to cleaner and more sustainable energy sources [18,19]. Production technologies in the areas of technology and materials take into account sustainable development through the selection of processing technologies (e.g., casting, machining, 3D printing, etc.), materials, energy savings, emitted particles and waste, which is supported by AI-based management, thus enabling the assessment of equipment consumption (as part of predictive maintenance), materials, the amount of pollution and the waste generated by production lines, packaging and transport systems. Self-learning programs adapt to changes in near real time, complementing previously used metrics and software [20,21]. Such systems collect data and analysis results, and draw conclusions to solve current and future problems, also allowing the support and education of people and specialists in creative design, and solving SSC issues.

SSCs will have a better ability to respond quickly and with fewer errors to health threats, including mass threats such as poisonings and pandemics. Policies, strategies and actions based on AI technology enable crisis management, improve the health of residents and increase the resilience of SSCs. In particular, coping and recovery are possible through:

- Improving and equalising access to digital SSC services;
- Improving the digital skills of residents (including children, elderly people and disabled people);

- Improving physical and mental health;
- Increasing social participation and connections;
- Maintaining the functionality of educational and economic systems [22–24].

E-health, smart manufacturing and logistics, traffic monitoring and car sharing (and, in the future, autonomous driving) are all data-driven. This requires the implementation of three main functionalities:

- Real-time intelligence;
- Distributed intelligence;
- Law enforcement with privacy;

Using technologies of the third decade of the 21st century:

- 6G mobile networks (bit-pipe connectivity);
- Smart edges (realising the burden of intelligent computing as close as possible to the consumers of services);
- AI/ML

Provides the Internet of Medical Things (IoMT, Figure 2) and electronic health records (EHR). In a sustainable, AI-powered smart city, the IoMT revolutionises healthcare by seamlessly integrating AI-powered insights with connected medical devices. Smart sensors and wearables continuously monitor the health of residents, transmitting real-time data to AI systems that analyse and predict potential health issues, thus enabling proactive interventions. These AI-powered insights optimise resource allocation, ensuring efficient healthcare delivery while minimising waste. Smart city infrastructure supports the IoMT network with renewable energy, reducing the environmental impact of healthcare operations.AI also enhances telemedicine by providing accurate diagnostics and personalised treatment plans, reducing the need for physical visits and lowering transportation emissions. Using machine learning, AI systems detect patterns in health data, therefore identifying public health trends and enabling a rapid response to potential disease outbreaks. Integrating the IoMT with AI ensures that healthcare is not only personalised but also scalable, meeting the needs of a growing urban population. This collaboration between the IoMT and AI contributes to the city's sustainability goals by improving health outcomes while reducing resource consumption. Additionally, robust cybersecurity measures have been implemented to protect sensitive medical data, and so ensure trust and privacy in the IoMT ecosystem. Ultimately, the combination of the IoMT and AI in a sustainable smart city creates a future where healthcare is efficient, environmentally friendly and highly responsive to the needs of its residents. More effective legacy data acquisition and centralised ML models have less data security and privacy, and mass sensing, data provisioning and SSC service delivery may be more difficult as the population and sensing population grow [23–29]. Sensors can measure air quality, traffic volume and other features of the urban environment, and in the "drive-by" paradigm, sensors can be installed in the vehicles of various companies, thus increasing range and reducing costs. Until now, the main problems have lain in reconciling the massiveness and privacy of data with the required short learning time of AI. It provides a distributed training approach capable of solving the aforementioned problems.

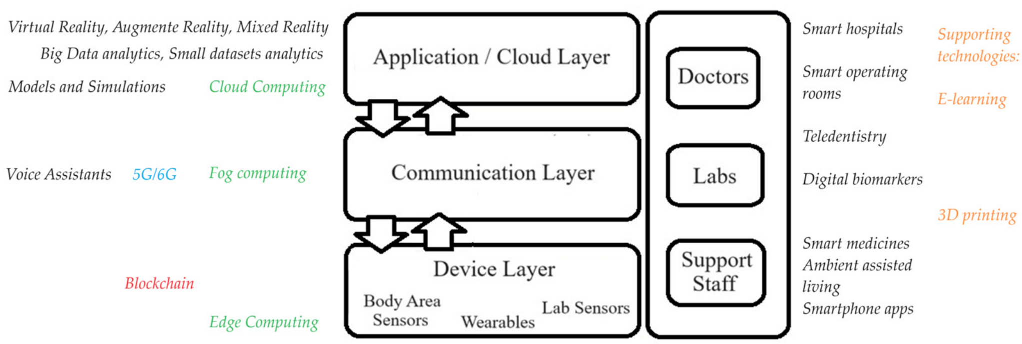


**Figure 2.** IoMT main concept (own version based on [23–29]).

Local data are used to train local models and, in turn, local models are used to update the global model. This aggregated global model is returned to the local models for further training and this procedure is repeated until the global model converges and works to further maintain this convergence, i.e., a common operational picture at all levels [30–38]. Large volumes of IoMT data can be harnessed without sharing, complex dynamics and complex data-sharing agreements. However, handling data from multiple locations presents another challenge [35–38]. Deep reinforcement learning (DRL), digital twins and generative adversarial networks (GANs) may be particularly useful here [38–42].

There is not much research on AI-based data and energy management within energy cooperatives (this is how SSCs should be considered); this area is only just being transformed [43]. In doing so, it is clear that SSCs will face the need for lighting and heating(especially in the global south), e.g., warming as a result of climate change and as a side effect of urbanisation, which will increase the need for cooling. GIS-based urban heat island estimation, energy modelling and rooftop solar potential have been used for residential energy stress (REST) buildings (Amaravati, India). This allows for the application of decentralized optimisation solutions for energy control and peer-to-peer energy sharing on a neighbourhood scale. In this solution, decision tree algorithms based on energy justice variables classify energy grid data in a sustainability framework to alleviate energy stress at the SSC level. This allows an up to 80% reduction in energy consumption in SSCs, including based on the optimisation of planning variables (such as floor area ratio and building density) and optimisation of current energy consumption. Therefore, these indicators must be taken into account when planning the construction of an SSC, and the mere use of AI for SSC management may produce lesser results [44–46].

The rapid development of modern IT, power, communication and road information systems, etc., develops distributed and energy-saving (if possible, powered by renewable energy sources) elements of the smart grid safely connected with entire smart city management systems. This enables a wide range of applications such as distributed energy management, system health forecasting and cybersecurity based on huge volumes of data that automate and improve the performance of the smart grid, but also require analysis, inference and prediction using artificial intelligence. Data management strategies, but also the sharing of data by consumers, institutions, organisations and industries, can be supported by edge clouds, thus protecting privacy and improving performance. Many solutions (e.g., mobility systems) usually operate based on the agents'/vehicles' self-interest, as part of integration with solutions for autonomous drivers [3]. Traffic simulation tools can be used to determine the duration of transport use. Pedestrian traffic routes pose a challenge; they can also be used for autonomous vehicles delivering parcels to shops, service premises and houses or apartments. Found in areas intended for pedestrian traffic stakeholders in such a logistics process are primarily courier service providers and their customers, but also SSC traffic and logistics managers—fragmentation of the "last mile" of transport services in this way will shorten the time in which the shipment reaches the recipient (also, regardless of the time, in 24 h) [3]. Automated diagnostics and support in apartments, nursing homes, rehabilitation centres, hospitals and sanatoriums require the use of edge computing, fog and clouds for rapid processing and response. It is also the only alternative in the absence of caregivers, physiotherapists or doctors, and perhaps also psychologists and family. The complexity of such a problem requires hybrid solutions combining mathematical models with ML [3,12].

SSCs can promote so-called active transport, such as walking or cycling. In turn, such appropriately measured activity may contribute to increasing the overall level of physical activity and changing activity habits, and, over time, dietary habits, work–private life balance or preferred forms of tourism. Apps for smartphones, tablets, smart TVs and vehicle infotainment can also promote this behavioural change. The aggregate benefits from such seemingly minor changes within the SSC may be significant and, in the long run, inevitable, and should therefore be monitored and stimulated to avoid going in the wrong direction [47–55]. It is worth looking for cheap solutions so that the application has the potential to affect many copies on a mass scale. The impact of the time it takes to get to work and school (on foot or by various means of transport) on the quality of life is large: the longer employees' commute time, the lower their satisfaction with work, life and health. It is worth knowing how many vehicles, appropriately equipped with sensors, would be needed to properly scan the SSC for a specific account. Publicly available data on the movement of individual travellers is contained both in their smartphones and in traffic lights and city monitoring. This allows, for example, the simplification of location based on movement trajectories. Even controlling pedestrian and vehicle traffic and effectively managing congestion and delays help reduce travel times and save many valuable resources. Basing such a system on three basic elements: vehicle, infrastructure and events, enables taking all scenarios and possible problems of the transport system into account. (including the machine learning-based DBSCAN clustering method for anomaly detection).

As people migrate between SSCs or from rural to urban environments, cities must optimise their services (services, logistics and transport) to best respond to changing demand [3]. ML can solve some of the problems associated with decentralised computing, enhancing cyber security and data privacy, but it needs further development. This is implemented, among other things, as part of a health impact assessment (HIA), health economics, the impact of health recommendations on urban policies, the burden of disease/injury at the city level and the involvement of stakeholders (residents, city authorities, representatives of associations of medical specialists and patients) [45–48]. This is the result of a simple observation that if the aspirations, goals (and sometimes behaviour patterns) of the upper and middle classes differ, such a distinction (provided with appropriate data) can be made in more detail, sometimes even to the level of personalisation of interactions within preventive medicine (healthy people) [45–48]. However, this imposes significant organisational and technical requirements, such as

- Large amount of high-quality data;
- Agreement at the level of policies, their coherence and the participation of all stakeholder groups;
- Health equity (measured cross-sectionally);
- Creative removal of barriers across various levels and areas.

Factors such as the use of energy in wastewater treatment plants (WWTPs) and the need to ensure the energy balance of wastewater treatment plants become more important [49,50].

AI plays a leading role in the transformation towards SSC by promoting and managing smart grids, energy demand monitoring, service products and load/supply management in line with energy efficiency, maintenance and asset management, carbon reduction, risk management and even compliance and cyber security [18,19]. The current state of knowledge indicates that attempts to develop AI solutions for SSC have focused on large systems, which, with a large number of sensors, effectors and analysed data, tend to transform over time and towards development into a "black box", not susceptible to the analysis of decision rules and causes in making a decision. Therefore, specific decisions are made at a level that is not fully understandable to a human operator, even an engineer or scientist (experts in their field). There is a need for simplification of the above analytical and decision-making processes so that they become understandable to humans while maintaining their effectiveness. There is a research gap in the development of cascading systems using simple algorithms that are both easy to analyse and understand by humans. These algorithms, when combined, should generate a sufficiently complex and accurate picture to support decision-making in various aspects of smart sustainable cities (SSCs). However, current research has not yet fully explored or achieved the balance between simplicity, transparency and the ability to deal with the complexity of real-world data. The challenge is to create a system where each algorithm makes a meaningful contribution to the overall decision-making process without sacrificing clarity or interpretability. Addressing this gap is critical to the development of practical and reliable decision support in SSC.

## 3. Materials and Methods

The primary directive of the SSC is to ensure the well-being and comfort of residents and maintain the continuity of city functions, including crisis management. We chose three key areas of SSCs for modelling: health, energy management and mobility/transport. This choice is supported by the fact that three areas of sustainable community management turned out to be critical during the last two major crises: the pandemic and the war in a neighbouring country (Ukraine). We assumed that additional areas may be added in subsequent studies, as the number of sensors and effectors within each area/segment may increase. This means system scalability.

### *3.1. Computational Methods*

The first computational approach (named cascade of simple algorithms) includes several levels of analysis shown in Figure 3: from edge computing (close to the source, possibly AI-based, STAGE 1, algorithms 1–3 selected to suit the data) through data aggregation from particular segments (STAGE 2, algorithms 4–6 selected to suit the data) to central AI-based management at the level of the entire SSC (STAGE 3, algorithm 7 selected to suit the data). ML with a distributed, local approach to data collection and pre-analysis (edge computing) may provide a potential solution to various challenges regarding processing speed, data confidentiality and cybersecurity (Figure 3). Here, ML provides conditions in which multiple clients (data sources) jointly learn a model based on decentralised data sets.

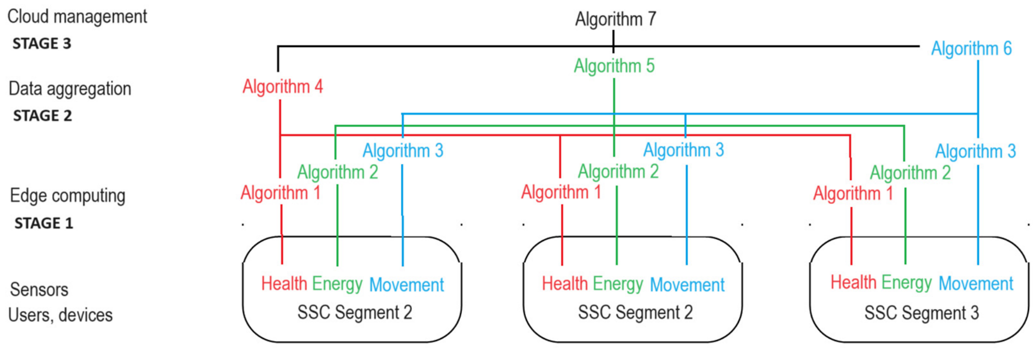


**Figure 3.** General architecture of AI-based management system (cascade of simple algorithms).

The algorithm numbers make orientation easier. Algorithms 1–7 are divided into three stages:

- Stage 1 (algorithms 1–3): algorithms for edge processing;
- Stage 2 (algorithms 4–6): algorithms for data aggregation;
- Stage 3 (algorithm 7): global model algorithm.

Each one of algorithms 1–7 was tested and selected for the data it was to analyse:

- Algorithms 1–3: for segmented sensor data;
- Algorithms 4–6: for Stage 1 analysis data (to aggregate the segmented data);
- Algorithm 7: for Stage 2 (aggregated) analysis data.

The whole procedure for constructing the model includes, in turn

- selection of the overall structure of the cascade model (e.g., Figure 3: for 3 types of input sensors/data sets),
- selection of data sets (Table 2),
- selection of the best algorithms for STAGE 1 for the given data sets (Tables 3–5),
- selection of the best STAGE 2 algorithms for the data sets (Tables 6–8),
- selection of algorithms for STAGE 3 best for given data sets,
- completion and tuning of the overall model (e.g., Figure 4: for 3 types of input sensors/data sets).

This approach provides maximum flexibility for a cascade model based on simple, transparent algorithms, easy to increase the number and types of input data, and easy to learn and upgrade, e.g., as new, more efficient algorithms emerge.

For greater readability of the models, the data used in the article are described in Section 3.2 Data sets and the algorithms used in the Results section. These data and models will allow cities and local managers to better understand the burden and geographic distribution of health, energy use and mobility/transportation variables in their jurisdictions and help them plan interventions in the above areas. This enables in-depth analysis and modelling of energy consumption behaviour and the discovery of complex factors influencing changes in the studied areas.

The paper provides a framework for future work on integrating large data sets for analysis using machine learning techniques to integrate data from different sources occurring in SSCs. For the above-mentioned reasons, the study uses real data with naturally occurring distributions. The integration procedure includes preprocessing the data by checking the correctness and completeness of the data and their repeatability, possibly removing outliers or unreliable data (e.g., out of range), if necessary, also normalising the data, balancing them in classes and formatting them in the form of standard anonymised feature vectors/matrices (usually significantly shortened in relation to the raw data). Such a form of data is accepted by communication and data analysis systems, while improving network throughput and data security. Section 3 describes the use of seven algorithms selected in terms of criteria such as data prediction/classification accuracy and duration. This selection consists in applying a set of algorithms and searching for those that give the best results at each stage. Of course, in a real system, this process will take much less time due to previously established, repeatable data preprocessing procedures and the repeatable characteristics of device data.

For the purpose of reproducibility and replication of the study, we have included a formal description of each of the stages, including the establishment of possible algorithms to be used and their key hyperparameters. This will allow us to reproduce the work performed, as well as to look for other approaches to improve the obtained results. Figure 3 shows a schematic general architecture of the AI-based management system. The algorithm numbers in Figure 3 facilitate orientation. The selection of the number of algorithms was made experimentally. The algorithms marked with numbers 1–7 are divided into three stages: algorithms 1–3 for edge processing, algorithms 4–6 for data aggregation, and algorithm 7 for the global model.The details of the algorithms (during the selection process) are presented in Tables 2–8, and the final version with details of the selected algorithms is presented in Figure 4.

In the process of selecting optimal algorithms, different data sets were searched and machine learning algorithms were used, and then those that gave the best results were selected. This approach results from the fact that despite the usual repeatability (similarity) of data during normal operation, the data may also be subject to unexpected changes that must be detected and properly classified, and a decision must be developed and an appropriate response must be made on them. In the context of the decision-making process, such selection and application of the above-mentioned algorithms contribute to

- reducing errors (i.e., incorrect response to input vectors/matrices);
- shortening response time;
- training the system to better adapt to the specificity of the data.

The process of integrating information from other areas important for smart cities, such as data on citizen safety, education, etc., is analogous, but for the purposes of this study, it was limited by the availability of databases. Expanding to a full model in all 10 areas of SSC activity supported by AI is our goal in subsequent studies.

It seems that the data in the SSC can be so volatile that the flexibility of the development environments will be key when, for example, the sensor software is substantially replaced by a new generation, and this cannot be achieved manually. The openness of the environment to the automated use of large data sets is also important. An important assumption may be that the implementation of connecting new sensors and effectors, and even whole clusters of them, will take place in stages, with gradual learning by the algorithms.

We would like to point out that, in the study, we did not use complex systems, but simple algorithms in order to exclude the "black box" phenomenon even in such a large-scale solution as the handling of SSC functionality. This will preserve people's understanding of the rules governing SSC, although as advanced eHealth or sustainable energy solutions are implemented, there is a risk of algorithms evolving into systems that are more complex than they are today. Detailed data and descriptions of the algorithms used are described in the Results section. The second computational approach is derived from an article by Wu et al. concerning deep learning (DL) in smart cities [56]. We have used three deep neural networks (Figure 5) [56]. Their results were aggregated as mean. Their results, in order to obtain a single accuracy in calculations, were aggregated as average values. The comparison was performed using the Bayesian Regularization of Artificial Neural Network (BRANN) models dedicated to smart city data, presented in Figure 5, based on the inverse of the Hessian data calculated in the evidence maximization loop, described in detail in [56].

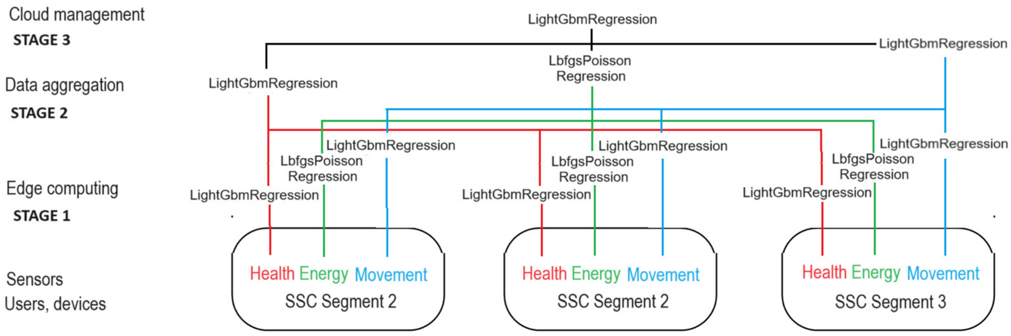


**Figure 4.** AI algorithms used within AI-based management system.

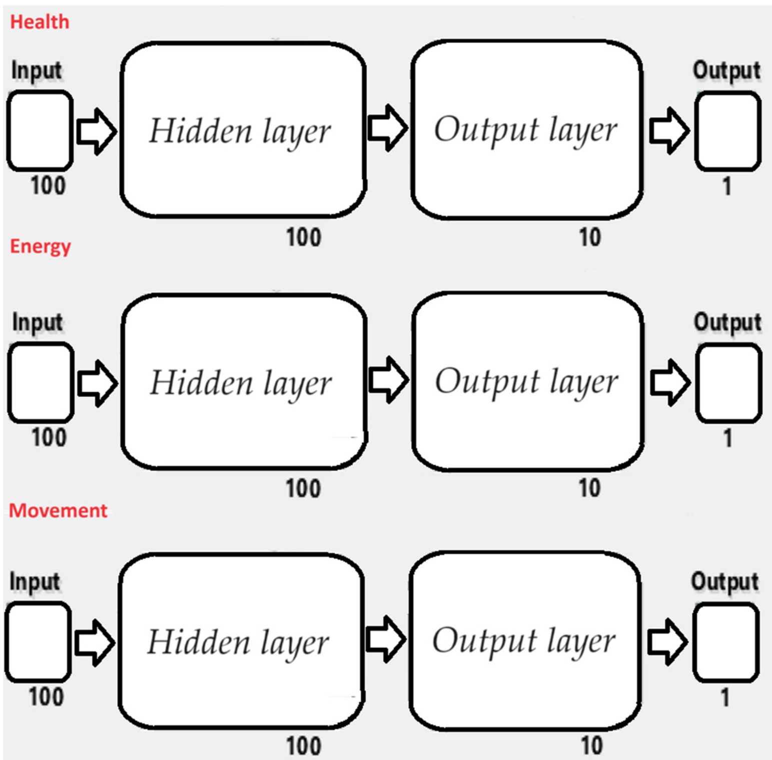


**Figure 5.** General architecture of AI-based management system (deep learning) (own version based on [56–60].

### 3.2. Data Sets

In order to ensure the feasibility of the analyses, aggregated real data from the Kaggle database covering three types of IoT sources were used for modelling:

- Medical data;
- Energy consumption data;
- Data on the movement of people and vehicles.

To develop and test our solution, we used three data sets from Kaggle dedicated to three areas of SSC sensors: health management, energy consumption and human and vehicle movement (Table 2). This approach stems from the difficulty of collecting a consistent real-world data set from three such different areas. Table 2 describes the three data sets used in the study.

**Table 2.** Data sets used in the study.

| Data Set Type | Data Set Name | Source |
|---|---|---|
| Medical data | 500 cities local data for better health, 2018 [61] | https://www.kaggle.com/datasets/jaimeblasco/500-cities-local-data-for-better-health-2019 City- and census region-level estimates for small area chronic disease risk factors, health outcomes and use of clinical preventive services for the largest 500 cities in the United States. |
| | United States of America Health Indicators [62] | https://www.kaggle.com/datasets/mahdiehhajian/united-states-of-america-health-indicators The database contains data from the World Health Organisation's data portal covering basic healthcare categories. |
| Energy production and consumption data | Hourly Energy Consumption [63] | https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption Database of the Eastern Interconnection network operating the electric transmission system serving parts of the US containing hourly energy consumption data. |
| | Energy consumption prediction [64] | https://www.kaggle.com/datasets/mrsimple07/energy-consumption-prediction The data set includes temperature, humidity, occupancy, HVAC and lighting use, renewable energy contribution, time stamp. |
| Movement of people and vehicles | Smart city traffic patterns [65] | https://www.kaggle.com/datasets/utathya/smart-city-traffic-patterns The data set form Mckinsey Analytics Online Hackathon. |
| | Traffic Prediction Dataset [66] | https://www.kaggle.com/datasets/fedesoriano/traffic-prediction-dataset This dataset 48120 observations of the number of vehicles each hour in four different junctions. |

A comparison was made for the proposed cascade solution with deep learning and the results are shown in Figure 6 in the Results section. Data sets were divided into three sample groups:

- Training (70%): presented to the network during training as the network is adjusted according to this error;
- Validation (20%): is used to measure the generalisation of the network and to stop training when the generalisation stops improving;
- Testing (10%): provides an independent measurement of network performance during and after training [56].

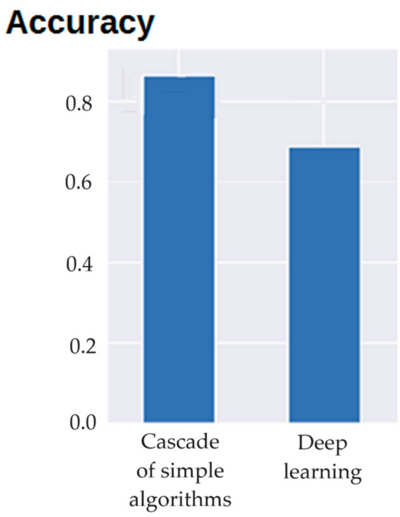


**Figure 6.** Cross-validation results between cascade of simple algorithms and deep learning.

The above data sets were used in various configurations in the search for the best solution.

## 4. Results

For the purpose of this study, we tested a number of machine learning automation environments. For the implementation, we chose ML in Visual Studio 2022 because of its large selection of algorithms, fast semi-automatic learning and testing, and the ability to download a file or API for further use within the SSC. Here is a brief description of the algorithms used:

- LightGBM regression: A gradient boosting framework that uses tree-based learning algorithms. It is designed to be fast and efficient, especially for large data sets, and is known for its high performance and scalability.
- FastTreeTweedie regression: A variant of the FastTree algorithm that models the Tweedie distribution, suitable for tasks such as insurance claims, where the target variable has both zero and positive continuous values. It combines decision tree learning with the Tweedie distribution to handle complex data distributions.
- LbfgsPoisson regression: A regression algorithm that uses the limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) optimization method to fit a Poisson distribution model. It is useful for numerical data where the output is a non-negative integer, such as the number of occurrences of an event.
- FastTree Regression: A gradient-boosting algorithm that builds an ensemble of decision trees for regression tasks. It is designed for fast execution and can efficiently handle large data sets, making it suitable for high-dimensional data.
- FastForest Regression: An ensemble method that combines multiple decision trees to increase predictive accuracy. It is a type of random forest algorithm that is particularly good at reducing overfitting and improving generalization in regression tasks.
- SdcaRegression: Stochastic Dual Coordinate Ascent (SDCA) regression is an optimization-based method for linear regression tasks. It is well-suited for large-scale problems and achieves efficient, fast convergence by solving a dual original optimisation problem.

For each data set 1–7, we tested 41–78 algorithms, selecting the best ones (algorithms 1–7) against the set criteria (high accuracy, low RMSE and short execution time). The best selected algorithms are shown in Figure 4.

Algorithms 1–3 (Stage 1), considered best for analysing individual data, are shown in Figure 5. In the individual areas, the best algorithms were for

- Health data: LightGbm Regression (0.8510) (Table 3);
- Energy consumption data: LbfgsPoissonRegression (0.7688) (Table 4);
- Traffic data: LightGbm Regression (0.8687) (Table 5).

Important note: the LightGbm Regression algorithms for health data analysis and traffic data analysis are algorithms trained on different data; thus, they result in different models. Similarly, the LightGbm Regression algorithms used in Stage 1, Stage 2 and Stage 3 are different algorithms because we used different data sets for learning.

**Table 3.** Best algorithms to select the best algorithm for data sets from the health area (Stage 1).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| LightGbmRegression | 0.8510 | 0.37 | 0.24 | 0.49 |
| LightGbmRegression | 0.8383 | 0.38 | 0.25 | 0.50 |
| LightGbmRegression | 0.8357 | 0.39 | 0.25 | 0.50 |
| FastTreeTweedieRegression | 0.8258 | 0.38 | 0.26 | 0.51 |
| FastTreeTweedieRegression | 0.8179 | 0.39 | 0.27 | 0.52 |

**Table 4.** Best algorithms to select the best algorithm for data sets from the energy consumption area (Stage 1).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| LbfgsPoissonRegression | 0.7688 | 975.90 | 1539806.29 | 1240.89 |
| LbfgsPoissonRegression | 0.7088 | 1046.44 | 1939766.85 | 1392.76 |
| LbfgsPoissonRegression | 0.6942 | 1128.64 | 2037087.58 | 1427.27 |
| LbfgsPoissonRegression | 0.5871 | 1277.95 | 2750434.96 | 1658.44 |
| LbfgsPoissonRegression | 0.5585 | 1317.30 | 2940633.20 | 1714.83 |

**Table 5.** Best algorithms to select the best algorithm for data sets from the movement area (Stage 1).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| LightGbmRegression | 0.8687 | 4.96 | 55.89 | 7.48 |
| FastTreeRegression | 0.8639 | 5.11 | 57.91 | 7.61 |
| LightGbmRegression | 0.8506 | 5.30 | 63.56 | 7.97 |
| LightGbmRegression | 0.8388 | 5.53 | 68.61 | 8.28 |
| FastTreeRegression | 0.8370 | 5.54 | 69.35 | 8.33 |

Algorithms 4–6 (Stage 2), considered best for data aggregation, were from

- Health data (algorithm 4): LightGbm Regression (0.8501);
- Energy consumption data (algorithm 5): LbfgsPoissonRegression (0.7597);
- Traffic data (algorithm 6): LightGbm Regression (0.8567).

The algorithm considered best for cloud management (central/general view: algorithm 7, Stage 3) was

- LightGbm Regression (0.8322).

The selection of SSC areas to be modelled (health, energy and mobility) and the levels and modelling approaches seem correct, touching on the main current and partly future SSC problem areas and related ESG policies and reports.

We considered all types of algorithms and networks, including traditional multilayer perceptron (MLP) when comparing algorithms in each area (health, energy management, traffic and transport). However, assuming a minimum accuracy threshold of 0.60, none of them achieved the required results in a reasonable time (they took too long to learn). The likely reason for this was that the data sets were relatively simple, but there were many of them; hence, approaches that were effective for smaller data sets or complex data sets were not optimal for their analysis. We have presented example results of comparisons with other algorithms at a threshold of 0.60 in Table 6 (health), Table 7 (energy management) and Table 8 (traffic and transport).We note that for other data sets (with different characteristics of features, abundance, etc.) the optimal algorithms could be different.

**Table 6.** Comparison of different health assessment algorithms (53 algorithms tested, only best shown) to select the best algorithm for aggregation of data sets from the health area (Stage 2).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| FastForestRegression | 0.6944 | 7.65 | 145.13 | 11.92 |
| LightGbmRegression | 0.7623 | 7.19 | 111.22 | 11.01 |
| FastForestRegression | 0.7521 | 6.78 | 114.43 | 11.17 |
| FastTreeTweedieRegression | 0.8367 | 5.55 | 66.22 | 7.97 |
| SdcaRegression | 0.7790 | 7.11 | 101.77 | 10.17 |
| FastTreeRegression | 0.8283 | 6.56 | 82.83 | 9.43 |
| LbfgsPoissonRegressionRegression | 0.7912 | 6.78 | 97.99 | 9.84 |
| LightGbmRegression | 0.8439 | 5.61 | 72.55 | 8.97 |
| FastForestRegression | 0.8407 | 5.19 | 67.43 | 8.43 |
| FastTreeTweedieRegression | 0.7979 | 6.78 | 101.11 | 10.01 |
| SdcaRegression | 0.8320 | 6.43 | 85.65 | 9.44 |
| FastTreeRegression | 0.7789 | 7.13 | 108.32 | 10.22 |
| LbfgsPoissonRegressionRegression | 0.7645 | 7.08 | 104.41 | 10.13 |

**Table 7.** Comparison of different energy consumption algorithms (48 algorithms tested, only best shown) to select the best algorithm for aggregation of data sets from the energy management area (Stage 2).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| LbfgsPoissonRegressionRegression | 0.6534 | 1153.40 | 2311041.62 | 1520.21 |
| LbfgsPoissonRegressionRegression | 0.6542 | 1139.34 | 2305941.48 | 1518.53 |
| FastTreeRegression | 0.6283 | 1221.52 | 2478228.27 | 1574,24 |
| LbfgsPoissonRegressionRegression | 0.6972 | 1111.92 | 2018649.36 | 1420.79 |
| FastTreeRegression | 0.7427 | 1016.11 | 1715846.96 | 1309.90 |
| LbfgsPoissonRegressionRegression | 0.6895 | 1083.60 | 2070330.99 | 1438.86 |
| FastTreeRegression | 0.6145 | 1247.94 | 2570239.36 | 1603.20 |
| LbfgsPoissonRegressionRegression | 0.7056 | 1099.95 | 1963138.93 | 1401.2 |
| FastTreeTweedieRegression | 0.6013 | 1267.53 | 2658605.93 | 1630.52 |
| FastTreeRegression | 0.6049 | 1260.06 | 2634609.08 | 1623.15 |
| LbfgsPoissonRegressionRegression | 0.6112 | 1267.24 | 2592538.55 | 1610.14 |
| FastTreeRegression | 0.7604 | 974.84 | 1597501.02 | 1263.92 |
| LbfgsPoissonRegressionRegression | 0.6404 | 1219.63 | 2397823.37 | 1548.49 |

**Table 8.** Comparison of different movement and transport assessment algorithms (53 algorithms tested, only best shown) to select the best algorithm for aggregation of data sets from the health area (Stage 2).

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| FastForestRegression | 0.7759 | 7.05 | 102.82 | 10.14 |
| LightGbmRegression | 0.7677 | 7.12 | 106.56 | 10.32 |
| FastForestRegression | 0.7542 | 7.44 | 112.75 | 10.62 |
| LbfgsPoissonRegressionRegression | 0.6620 | 7.42 | 155.06 | 12.45 |
| FastTreeRegression | 0.8010 | 6.65 | 91.32 | 9.56 |
| LightGbmRegression | 0.7808 | 6.69 | 100.56 | 10.03 |
| FastTreeTweedieRegression | 0.6788 | 7.82 | 147.36 | 12.14 |
| FastForestRegression | 0.7773 | 7.03 | 102.17 | 10.11 |
| LbfgsPoissonRegressionRegression | 0.7540 | 6.61 | 112.85 | 10.62 |
| FastTreeRegression | 0.8577 | 5.51 | 65.30 | 8.08 |
| FastForestRegression | 0.7801 | 7.00 | 100.86 | 10.04 |
| LbfgsPoissonRegressionRegression | 0.8173 | 6.28 | 83.83 | 9.16 |
| LightGbmRegression | 0.7845 | 6.89 | 98.86 | 9.94 |
| FastTreeRegression | 0.8442 | 5.74 | 71.48 | 8.45 |
| FastTreeTweedieRegression | 0.8507 | 5.20 | 68.50 | 8.28 |

**Table 8.** *Cont.*

| Agorithm | RSsquared | Absolute Loss | Squared Loss | RMS Loss |
|---|---|---|---|---|
| FastForestRegression | 0.7820 | 6.98 | 100.00 | 10.00 |
| LightGbmRegression | 0.8150 | 6.06 | 84.85 | 9.21 |
| FastTreeRegression | 0.7627 | 7.02 | 108.89 | 10.44 |
| FastForestRegression | 0.7759 | 7.05 | 102.82 | 10.14 |

We have developed and tested a conceptual framework based on the data sets, but it is the resulting model that needs further refinement and development, especially based on real data. In the individual areas, the best deep learning results were for

- Health data: 0.8322;
- Energy consumption data: 0.6513;
- Traffic data: 0.7523.
- Mean aggregated value: 0.7453.

In machine learning algorithms, several key hyperparameters have been tuned to optimise model performance:

- LightGbm Regression: The number of boosting iterations (num_iterations), learning rate (learning_rate) and maximum tree depth (max_depth).These control the number of trees built, the step size of model updates and the complexity of each tree;
- FastTreeTweedie Regression: the number of leaves (num_leaves), the minimum number of examples per leaf (min_leaf_count) and learning rate (learning_rate).These affect the granularity of the tree, the minimum data needed in a leaf for further splitting, and the speed of convergence;
- LbfgsPoissonRegression: the number of iterations (num_iterations), the strength of regularisation (l2_regularization) and the tolerance (convergence_tolerance).These govern the maximum number of optimisation steps, the penalisation of large coefficients and the convergence criteria of the algorithm;
- FastTreeRegression: number of trees (num_trees), minimum split gain (min_split_gain) and learning rate (learning_rate).These control the ensemble size, the threshold for splitting nodes and the rate of model adaptation;
- FastForestRegression: number of trees (num_trees), the number of features to consider per split (num_features_per_split) and the minimum number of samples per leaf (min_samples_per_leaf).These control the diversity and depth of the trees, as well as the minimum data required to create a leaf;
- SdcaRegression: L1 and L2 regularisation terms (l1_regularization and l2_regularization) and convergence tolerance (convergence_tolerance). These parameters control the sparsity of feature selection, the model complexity penalty and the stopping criterion of the optimisation process.

Cross-validation showed the superiority of the cascade of simple algorithms over the concurrent solution described in [56] (Figure 6). It is worth considering that for a small sample of smart city transportation data, this deep learning approach achieved 0.99978 [56].

## 5. Discussion

SSCs that are socially inclusive, safe (also from an energy and cybersecurity point of view), resilient, and take into account the health and well-being of residents are a key element of the Sustainable Development Goals (SDG) and the New Urban Agenda [45–48]. The operation of such a system is predictable in the short term, allowing the ongoing management of SSCs based on current status, historical data and forecasts made 24 hours in advance. Predictions for longer periods depend more strongly on the accuracy of other predictions (e.g., weather forecasts) and can be shaped by long-term factors (e.g., seasonal variation, seasonal employment, etc.); hence, they require greater management flexibility.

The entire algorithm selection process is novel. We have provided a framework that not only sets the stage for future work on integrating large data sets but also uses machine learning techniques to synthesise data from different sources in SSCs. In particular, we have described the use of seven carefully selected algorithms, selected based on criteria such as prediction and classification accuracy, as well as computational efficiency. The selection process involved systematically applying different algorithms and identifying those that consistently delivered optimal results at each stage of the data processing process. These algorithms, numbered 1 to 7, are divided into three distinct stages to facilitate a structured approach to handling data:

- Algorithms 1–3 are dedicated to edge computing, where initial data refinement and feature extraction occur;
- Algorithms 4–6 focus on data aggregation, where information from multiple sources is consolidated;
- Algorithm 7 is used in the final stage to develop a global model that synthesises the aggregated data into actionable insights.

This methodical approach enables the most effective algorithms to be applied at each stage, which maximises overall system performance and reliability when handling large-scale data integration tasks in SSC environments.

### 5.1. Comparison with Results of Previous Studies

In this article, we present an approach consisting of seven algorithms for the integration of large data sets for machine learning processing to be applied in optimisation in the context of smart cities. The results of the bibliographic search on technical aspects (including AI-based ones) in the transformation towards SSC showed 78 publications between 2012 and 2024, including 36 journal articles, 4 books or chapters and 37 conference/workshop papers. As many as 64 (82.1%) of these were published in the last 5 years. Only 19 of these were published as open access. However, the range of topics covered in them was very broad: from the cognitive aspects of SSCs and their friendliness for residents/users through the logistics and ecology of SSCs to meta-analyses of previous experiences and publications. In the area addressed by this paper, a number of metaheuristic solutions are proposed for the effective management of the big data collections generated by SSCs [51,52].

The problem of using simple AI/ML algorithms for simultaneous multi-level analysis of different areas of SSC operations, including such key areas as health, energy consumption, traffic and transport has not been explored in this way before. Single areas or groups of areas can be modelled using deep learning, but in this case, they are resource-intensive and large, multi-year projects are required to explore them, which are not cost-effective with limited funding, nor are they time and energy-efficient. For the aforementioned reasons, the proposed approach raises hopes for further systematic development by other researchers.

While deep neural networks (DNNs) are increasingly being used for IoT device identification due to their high learning capability, they are vulnerable to attacks that can significantly reduce device identification accuracy. DNN-based design of reliable device identification schemes and investigation of the impact of untargeted and targeted adversary attacks on device identification require enrichment of evaluation criteria. Effective evaluation metrics should effectively show differences in the signal of individual devices, even though the accuracy of IoT device identification decreases with increasing attack/disturbance level and iteration step size [53].

Even more challenging are the formation flights of multiple unmanned aerial vehicles (UAVs) with dynamic spectrum interaction for ordered communication of multiple UAVs with limited bandwidth and in a countermeasure and jamming environment. Deep learning algorithms with reinforcement (DRL) and networks with long- and short-term memory (LSTM) provide the optimal strategy here, and also under conditions of environment interaction, UAV sharing and M/G/1 queueing are used for UAV prioritisation and packet loss assessment. Therefore, faster convergence and better performance with limited bandwidth are achieved [54].

### 5.2. Opportunities for Further Exploration

AI-based solutions offer promising advances in SSC development, but there are several limitations and challenges that need to be addressed. These include scalability, high initial costs and infrastructure requirements, integration challenges, controlled energy consumption, limited predictive accuracy and technology dependency. Scalability is a significant constraint in sustainable AI-supported smart cities due to the huge number of sensors, which can reach up to 50 million. This enormous scale creates challenges in managing and processing the huge amount of data generated, leading to potential bottlenecks in real-time decision-making. The complexity of coordinating thousands of segments and ensuring seamless communication between them can place a strain on infrastructure and computing resources. Additionally, the cost and energy consumption required to maintain such a large system can be prohibitive. Finally, as the system grows, it becomes increasingly difficult to ensure data security and privacy as more nodes create more points of vulnerability.

High initial costs and infrastructure requirements are significant constraints for the development of sustainable AI-supported smart cities, especially in developing countries. These regions often lack the financial resources and existing infrastructure needed to implement advanced technologies, making it difficult to adopt sustainable solutions. The costs of installing sensors, upgrading networks and providing robust data management systems can be prohibitively high, therefore delaying or preventing smart city initiatives. Furthermore, the ongoing costs of maintenance, training and technology upgrades add to the financial burden. As a result, the digital divide between developed and developing regions may widen, exacerbating global inequalities in access to sustainable technology.

Integration challenges are a major constraint in sustainable AI-supported smart cities, especially when integrating AI into existing urban infrastructure. Older systems, which are often outdated and incompatible with new technologies, require significant modification or replacement, making the integration process complex and costly. The lack of standardisation between different technologies further complicates matters, as different systems may not communicate effectively or share data seamlessly. These challenges can lead to inefficiencies, delays and increased costs in implementing smart city initiatives. Additionally, the need for expertise to manage these integrations can be a barrier, especially in regions with limited technical resources.

Controlled energy consumption is an important constraint in sustainable AI-supported smart cities as ML technologies are often energy-intensive. The computing power required to process large data sets, run complex algorithms and support real-time decision-making can lead to significant energy consumption. This increased energy demand can counteract smart cities' sustainability goals, potentially contributing to an increased carbon footprint. As cities seek to balance the benefits of artificial intelligence with the need for energy efficiency, this paradox challenges overall environmental goals. Furthermore, reliance on energy-intensive technologies may require additional infrastructure to support clean energy sources, thus adding complexity and cost.

Limited predictive accuracy is a significant limitation in SSCs supported by artificial intelligence (AI) as machine learning (ML) models can struggle to accurately predict complex urban dynamics. The challenge for these models is often the uncertainty of human behaviour, which can be unpredictable and highly variable in different contexts. Additionally, factors such as changing weather conditions, economic changes and unexpected events further complicate the accuracy of predictions. The inherent limitations of machine learning algorithms in capturing and responding to these dynamic variables can lead to suboptimal decision-making and planning. Consequently, reliance on AI forecasts can sometimes result in ineffective or inefficient urban management strategies.

Dependence on technology is a significant limitation in sustainable AI-enabled smart cities as over-reliance on AI systems can introduce vulnerabilities. In the event of technical failures, cyber-attacks or natural disasters, critical systems can be compromised, leading to disruptions in essential services such as transport, healthcare and utilities. The lack of human oversight and backup mechanisms in highly automated systems can exacerbate this risk, potentially leaving cities unprepared for unexpected events. This dependency highlights the need for robust contingency plans to manage disruptions and ensure continuity of services during emergencies. Without these safeguards, the resilience and sustainability of smart cities could be compromised, undermining their intended benefits.

The framework is considered to be at an early stage of development, with challenges primarily related to data origination, data aggregation, scalability and data security, especially in target production versions on SSCs of various sizes (number of inhabitants, data, devices, employees, etc.) and their use cases. The next stage of the research will be to create large data sets and evaluate them in difficult conditions. Processing based on AI integrated with the cloud of data from many personalized IoMT devices (bedside, wearable) must be characterised by high performance to provide the basis for building next-generation intelligent healthcare. The SSC also needs a holistic approach to early detection of pandemics such as COVID-19 based on molecular diagnostics and computer-aided detection (AI-based models) [55–59]. There is a certain challenge still in collecting and integrating data (e.g., on health) from a wide variety of devices, not only from smart watches, scales and smart shoes but also from thermographic sensors and sensors that examine people's movements as part of screening. Other data are easier to integrate if they are linked in terms of geolocation.

Addressing these limitations requires a multidisciplinary approach involving urban planners, policymakers, technologists (including AI specialists) and the public to develop SSC solutions that are inclusive, secure and ethically sound [67–70].

On the other hand, automated machine learning (AutoML) can lead to high energy consumption due to a large number of calculations and thus can significantly worsen its carbon footprint. For this reason, there is a need to apply energy efficiency metrics to advanced optimisation algorithms within AutoML so that the costs do not exceed the results. These metrics allow the evaluation and optimisation of the algorithm's energy consumption, taking into account accuracy, sustainability and reduced environmental impact with a minimal decrease in validation accuracy (e.g., by 0.5%). ML can be made more sustainable by carefully considering the energy efficiency of computational processes [71,72]. As biotechnological approaches evolve within smart cities, smart factories and smart territories, these technologies can be integrated with IoT, especially in rain gardens, urban vertical farming systems and urban photobioreactors. Biofuel cells can also be used to power low-power sensor networks or self-powered biosensors (synthetic biology: cell-based biosensors, bioactuators with synthetic genetic circuits as a development direction for cyber-physical system—CPS) [73]. Technological advancement and innovation path reviews provide a methodology for technological improvements and directions for their search. This can be guided by building on R&D strengths and business value, and structured technology exploration can be extended to explore and stimulate other forms of development [74]. It is necessary to carefully consider whether and which industrial and service optimisation processes (including the optimisation of machine resources as part of preventive maintenance) are consistent with sustainable development [75–77].

## 6. Conclusions

The SSC policy emphasises the environmental aspects of urban areas, and their social and economic sectors, enabling the development of practical plans for the sustainable development of cities towards SSC. The pandemic, high energy prices, lack of drinking water and security threats have accelerated the processes of transition to intelligent city management (including ML-based management of urban services), and thus the transition towards SSC.

In this study, we were able to analyse and manage three important areas of SSCs (health, energy consumption, and traffic and transport) with accuracies of over 0.8322 at the global level and 0.7688–0.8687 at the local level (the lowest values for energy management) using a cascade of a group of fairly simple algorithms. This shows that it is effective, and further research work can contribute to higher values.

The SSC conceptual framework shown in the article emphasises the priorities, practicalities and use cases of SIoT and AI in the daily lives of SSC residents and users. The experiences to date can be further extended to include new data and functional areas of SSCs (e.g., energy cooperatives or the sales and marketing network). The proposed AI/ML-based computational model in the transformation of SCs to SSCs with energy markets in mind can lay the foundations for the further development of AI-based end-to-end management of SSCs, with a focus on critical infrastructure, constraints and priority development directions. It is worth combining various management strategies based on both simpler methods (decision trees) and deep learning; even simple artificial intelligence solutions may prove effective in reasoning and forecasting and making SSC management more predictable and efficient for the benefit of residents and the environment.

Researchers and practitioners can contribute to the development of SSCs by prioritising ethical issues, community engagement, and long-term environmental and social benefits, in particular through research on

- Interdisciplinary approaches: interdisciplinary collaboration between computer scientists, urban planners, environmental scientists and social scientists can lead to more holistic and effective AI-based solutions for SSCs;
- Privacy-preserving AI: effective data analysis is possible while protecting the privacy of SSC residents;
- Community engagement and citizen empowerment: AI can be used to increase citizen engagement, participation and empowerment in sustainable urban development projects;
- Resilience and disaster management: AI can contribute to building resilient smart cities by developing robust systems for disaster prediction, response and recovery, integrated with urban planning.

**Author Contributions:** Conceptualization, I.R., D.M. and J.D.; Methodology, I.R., D.M. and J.D.; Software, I.R., D.M. and J.D.; Validation, I.R., D.M. and J.D.; Formal analysis, I.R. and D.M.; Investigation, I.R., D.M., J.D., E.D. and A.M.; and Resources, I.R., D.M. and J.D.; Data curation, I.R., D.M. and J.D.; Writing—original draft preparation, I.R., D.M. and J.D.; Writing—review and editing, I.R., D.M., J.D., E.D. and A.M.; Visualisation, I.R., D.M., J.D., E.D. and A.M.; Supervision, I.R., D.M. and J.D.; Project administration, I.R. and D.M.; Funding acquisition, I.R. and D.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work presented in this paper has been financed under a grant to maintain the research potential of Kazimierz Wielki University and grant No. 0613/SBAD/4888.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available as Kaggle datasets.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Oprea, S.-V.; Bâra, A.; Ciurea, C.-E.; Stoica, L.F. Smart Cities and Awareness of Sustainable Communities Related to Demand Response Programs: Data Processing with First-Order and Hierarchical Confirmatory Factor Analyses. *Electronics* **2022**, *11*, 1157. [CrossRef]
2. Rehman, A.U.; Naqvi, R.A.; Rehman, A.; Paul, A.; Sadiq, M.T.; Hussain, D. A Trustworthy SIoT Aware Mechanism as an Enabler for Citizen Services in Smart Cities. *Electronics* **2020**, *9*, 918. [CrossRef]
3. Corchado, J.M.; Trabelsi, S. Advances in Sustainable Smart Cities and Territories. *Electronics* **2022**, *11*, 1280. [CrossRef]
4. Herath, H.M.K.K.M.B.; Mittal, M. Adoption of artificial intelligence in smart cities: A comprehensive review. *Int. J. Inf. Manag. Data Insights* **2022**, *2*, 100076. [CrossRef]
5. Adenekan, O.A.; Ezeigweneme, C.; Chukwurah, E.G. The evolution of smart cities: Integrating technology, governance, and sustainable development. *Int. J. Appl. Res. Soc. Sci.* **2024**, *6*, 891–902. [CrossRef]

6. Yigitcanlar, T.; Corchado, J.M.; Mehmood, R.; Li, R.Y.M.; Mossberger, K.; Desouza, K. Responsible Urban Innovation with Local Government Artificial Intelligence (AI): A Conceptual Framework and Research Agenda. *J. Open Innov. Technol. Mark. Complex.* **2021**, *7*, 71. [CrossRef]
7. Mondal, M.; Srivastava, D. A Genetic Algorithm-Based Approach to Solve a New Time-Limited Travelling Salesman Problem. *Int. J. Distrib. Syst. Technol.* **2023**, *14*, 1–14. [CrossRef]
8. Namazi, E.; Li, J.; Lu, C. Intelligent Intersection Management Systems Considering Autonomous Vehicles: A Systematic Literature Review. *IEEE Access* **2019**, *7*, 91946–91965. [CrossRef]
9. Lin, D.; Jabari, S.E. Pay for Intersection Priority: A Free Market Mechanism for Connected Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 5138–5149. [CrossRef]
10. Kocsis, M.; Susmann, N.; Buyer, J.; Zollner, R. Safety concept for autonomous vehicles that operate in pedestrian areas. In Proceedings of the IEEE/SICE International Symposium on System Integration (SII), Taipei, Taiwan, 11–14 December 2017; pp. 841–846.
11. Artuñedo, A.; del Toro, R.M.; Haber, R.E. Consensus-based cooperative control based on pollution sensing and traffic information for urban traffic networks. *Sensors* **2017**, *17*, 953. [CrossRef]
12. Valença, G.; Moura, F.; de Sá, A.M. Main challenges and opportunities to dynamic road space allocation: From static to dynamic urban designs. *J. Urban Mobil.* **2021**, *1*, 100008. [CrossRef]
13. Alam, F.; Mehmood, R.; Katib, I.; Albogami, N.; Albeshri, A. Data fusion and IoT for smart ubiquitous environments: A survey. *IEEE Access* **2017**, *5*, 9533–9554. [CrossRef]
14. Kim, N.; Kim, G.; Shim, S.; Jang, S.; Song, J.; Lee, B. Key Technologies for 6G-Enabled Smart Sustainable City. *Electronics* **2024**, *13*, 268. [CrossRef]
15. Yun, J.J.; Zhao, X.; Jung, K.; Yigitcanlar, T. The culture for open innovation dynamics. *Sustainability* **2020**, *12*, 5076. [CrossRef]
16. Flores, P.; Carrillo, F.J.; Robles, J.G.; Leal, M.A. Applying open innovation to promote the development of a knowledge city: The Culiacan experience. *Int. J. Knowl. Based Dev.* **2018**, *9*, 312–335. [CrossRef]
17. Sodhro, A.; Gurtov, A.; Zahid, N.; Pirbhulal, S.; Wang, L.; Rahman, M.; Abbasi, Q. Towards convergence of AI and IoT for energy efficient communication in smart homes. *IEEE Internet Things J.* **2020**, *8*, 9664–9671. [CrossRef]
18. Rojek, I.; Dostatni, E.; Mikołajewski, D.; Pawłowski, L.; Wegrzyn-Wolska, K. Modern approach to sustainable production in the context of Industry 4.0. *Bull. Pol. Acad. Sci. Tech. Sci.* **2022**, *70*, e143828. [CrossRef]
19. Rojek, I.; Mikołajewski, D.; Mroziński, A.; Macko, M. Machine Learning- and Artificial Intelligence-Derived Prediction for Home Smart Energy Systems with PV Installation and Battery Energy Storage. *Energies* **2023**, *16*, 6613. [CrossRef]
20. Rojek, I.; Mikołajewski, D.; Macko, M.; Szczepański, Z.; Dostatni, E. Optimization of extrusion-based 3D Printing process using neural networks for sustainable development. *Materials* **2021**, *14*, 2737. [CrossRef]
21. Rojek, I.; Mikołajewski, D.; Kotlarz, P.; Macko, M.; Kopowski, J. Intelligent system supporting technological process planning for machining and 3D printing. *Bull. Pol. Acad. Sci. Tech. Sci.* **2021**, *69*, e136722. [CrossRef]
22. Hassankhani, M.; Alidadi, M.; Sharifi, A.; Azhdari, A. Smart City and crisis management: Lessons for the COVID-19 pandemic. *Int. J. Environ. Res. Public. Health* **2021**, *18*, 7736. [CrossRef] [PubMed]
23. Khan, M.A.; Alkaabi, N. Rebirth of distributed AI—A review of eHealth research. *Sensors* **2021**, *21*, 4999. [CrossRef] [PubMed]
24. Al-Rawashdeh, M.; Keikhosrokiani, P.; Belaton, B.; Alawida, M.; Zwiri, A. IoT Adoption and Application for Smart Healthcare: A Systematic Review. *Sensors* **2022**, *22*, 5377. [CrossRef]
25. Ali, H.; Alam, T.; Househ, M.; Shah, Z. Federated Learning and Internet of Medical Things—Opportunities and Challenges. *Stud. Health Technol. Inform.* **2022**, *295*, 201–204. [CrossRef]
26. Jiang, J.C.; Kantarci, B.; Oktug, S.; Soyata, T. Federated Learning in Smart City Sensing: Challenges and Opportunities. *Sensors* **2020**, *20*, 6230. [CrossRef] [PubMed]
27. Ismail, L.; Buyya, R. Artificial Intelligence Applications and Self-Learning 6G Networks for Smart Cities Digital Ecosystems: Taxonomy, Challenges, and Future Directions. *Sensors* **2022**, *22*, 5750. [CrossRef]
28. Mohsen, F.; Ali, H.; El Hajj, N.; Shah, Z. Artificial intelligence-based methods for fusion of electronic health records and imaging data. *Sci. Rep.* **2022**, *12*, 17981. [CrossRef]
29. Ali, H.; Shah, Z.; Alam, T.; Wijayatunga, P.; Elyan, E. Editorial: Recent advances in multimodal artificial intelligence for disease diagnosis, prognosis, and prevention. *Front. Radiol.* **2024**, *3*, 1349830. [CrossRef]
30. Khan, M.J.; Khan, M.A.; Malik, S.; Kulkarni, P.; Alkaabi, N.; Ullah, O.; El-Sayed, H.; Ahmed, A.; Turaev, S. Advancing C-V2X for Level 5 Autonomous Driving from the Perspective of 3GPP Standards. *Sensors* **2023**, *23*, 2261. [CrossRef]
31. Wang, W.; Liang, H.; Zhang, Z.; Xu, C.; Wei, D.; Li, W.; Qian, Y.; Zhang, L.; Liu, J.; Lei, D. Comparing three-dimensional and two-dimensional deep-learning, radiomics, and fusion models for predicting occult lymph node metastasis in laryngeals quamous cell carcinoma based on CT imaging: A multicentre, retrospective, diagnostic study. *eClinicalMedicine* **2024**, *67*, 102385. [CrossRef]
32. Gonzalez, R.; Saha, A.; Campbell, C.J.V.; Nejat, P.; Lokker, C.; Norgan, A.P. Seeing the random forest through the decision trees. Supporting learning health systems from histopathology with machine learning models: Challenges and opportunities. *J. Pathol. Inform.* **2023**, *15*, 100347. [CrossRef] [PubMed]
33. Ali, H.; Qureshi, R.; Shah, Z. Artificial intelligence-based methods for integrating local and global features for brain cancer imaging: Scoping review. *JMIR Med. Inform.* **2023**, *11*, e47445. [CrossRef] [PubMed]

34. Siam, A.; Alsaify, A.R.; Mohammad, B.; Biswas, M.R.; Ali, H.; Shah, Z. Multimodal deep learning for liver cancer applications: A scoping review. *Front. Artif. Intell.* **2023**, *6*, 1247195. [CrossRef] [PubMed]
35. Manickam, P.; Mariappan, S.A.; Murugesan, S.M.; Hansda, S.; Kaushik, A.; Shinde, R.; Thipperudraswamy, S.P. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors* **2022**, *12*, 562. [CrossRef]
36. Rachakonda, S.; Moorthy, S.; Jain, A.; Bukharev, A.; Bucur, A.; Manni, F.; Quiterio, T.M.; Joosten, L.; Mendez, N.I. Privacy Enhancing and Scalable Federated Learning to Accelerate AI Implementation in Cross-Silo and IoMT Environments. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 744–755. [CrossRef]
37. Ahmed, J.; Nguyen, T.N.; Ali, B.; Javed, M.A.; Mirza, J. On the Physical Layer Security of Federated Learning Based IoMT Networks. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 691–697. [CrossRef]
38. Ali, M.; Naeem, F.; Tariq, M.; Kaddoum, G. Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 778–789. [CrossRef]
39. Irkham, I.; Ibrahim, A.U.; Nwekwo, C.W.; Al-Turjman, F.; Hartati, Y.W. Current Technologies for Detection of COVID-19: Biosensors, Artificial Intelligence and Internet of Medical Things (IoMT): Review. *Sensors* **2022**, *23*, 426. [CrossRef]
40. Irkham, I.; Ibrahim, A.U.; Pwavodi, P.C.; Al-Turjman, F.; Hartati, Y.W. Smart graphene-based electrochemical nanobiosensor for clinical diagnosis: Review. *Sensors* **2023**, *23*, 2240. [CrossRef]
41. Jain, S.; Nehra, M.; Kumar, R.; Dilbaghi, N.; Hu, T.; Kumar, S.; Kaushik, A.; Li, C.Z. Internet of medical things (IoMT)-integrated biosensors for point-of-care testing of infectious diseases. *Biosens. Bioelectron.* **2021**, *179*, 113074. [CrossRef]
42. Bi, S.; Wang, C.; Zhang, J.; Huang, W.; Wu, B.; Gong, Y.; Ni, W. A Survey on Artificial Intelligence Aided Internet-of-Things Technologies in Emerging Smart Libraries. *Sensors* **2022**, *22*, 2991. [CrossRef] [PubMed]
43. Rojek, I.; Mroziński, A.; Kotlarz, P.; Macko, M.; Mikołajewski, D. AI-Based Computational Model in Sustainable Transformation of Energy Markets. *Energies* **2023**, *16*, 8059. [CrossRef]
44. Beniwal, R.; Kalra, S.; Singh Beniwal, N.; Gupta, H.O. Smart photovoltaic system for Indian smart cities: A cost analysis. *Environ. Sci. Pollut. Res. Int.* **2023**, *30*, 45445–45454. [CrossRef] [PubMed]
45. Ait Abdelmoula, I.; Idrissi Kaitouni, S.; Lamrini, N.; Jbene, M.; Ghennioui, A.; Mehdary, A.; El Aroussi, M. Towards a sustainable edge computing framework for condition monitoring in decentralized photovoltaic systems. *Heliyon* **2023**, *9*, e21475. [CrossRef]
46. Bardhan, R.; Debnath, R.; Gama, J.; Vijay, U. REST framework: A modeling approach towards cooling energy stress mitigation plans for future cities in warming Global South. *Sustain. Cities Soc.* **2020**, *61*, 102315. [CrossRef]
47. Imam, A.U.; Banerjee, U.K. Urbanisation and greening of Indian cities: Problems, practices, and policies. *Ambio* **2016**, *45*, 442–457. [CrossRef] [PubMed]
48. Ramirez-Rubio, O.; Daher, C.; Fanjul, G.; Gascon, M.; Mueller, N.; Pajín, L.; Plasencia, A.; Rojas-Rueda, D.; Thondoo, M.; Nieuwenhuijsen, M.J. Urban health: An example of a "health in all policies" approach in the context of SDGs implementation. *Global Health* **2019**, *15*, 87. [CrossRef]
49. Camblong, H.; Curea, O.; Ugartemendia, J.; Boussaada, Z.; Lizarralde, I.; Etxegarai, G. Photovoltaic energy sharing: Implementation and tests on a real collective self-consumption system. *Heliyon* **2023**, *9*, e22252. [CrossRef]
50. Yang, X.; Wei, J.; Ye, G.; Zhao, Y.; Li, Z.; Qiu, G.; Li, F.; Wei, C. The correlations among waste water internal energy, energy consumption and energy recovery/production potentials in waste water treatment plant: An assessment of the energy balance. *Sci. Total Environ.* **2020**, *714*, 136655. [CrossRef]
51. Kwak, Y.H.; Lee, J. Toward Sustainable Smart City: Lessons From 20 Years of Korean Programs. *IEEE Trans. Eng. Manag.* **2023**, *70*, 740–754. [CrossRef]
52. Saraji, M.K.; Streimikiene, D. A Novel Extended Fermatean Fuzzy Framework for Evaluating the Challenges to Sustainable Smart City Development. In *Real Life Applications of Multiple Criteria Decision Making Techniques in Fuzzy Domain*; Sahoo, L., Senapati, T., Yager, R.R., Eds.; Springer: Singapore, 2022; pp. 37–58.
53. Bao, Z.; Lin, Y.; Zhang, S.; Li, Z.; Mao, S. Threat of Adversarial Attacks on DL-Based IoT Device Identification. *IEEE Internet Things J.* **2022**, *9*, 9012–9024. [CrossRef]
54. Lin, Y.; Wang, M.; Zhou, X.; Din, G.; Mao, S. Dynamic Spectrum Interaction of UAV Flight Formation Communication with Priority: A Deep Reinforcement Learning Approach. *IEEE Trans. Cogn. Commun. Netw.* **2020**, *6*, 892–903. [CrossRef]
55. Ek, A.; Alexandrou, C.; Delisle Nyström, C.; Direito, A.; Eriksson, U.; Hammar, U.; Henriksson, P.; Maddison, R.; Trolle Lagerros, Y.; Löf, M. The Smart City Active Mobile Phone Intervention (SCAMPI) study to promote physical activity through active transportation in healthy adults: A study protocol for a randomized controlled trial. *BMC Public Health* **2018**, *18*, 880. [CrossRef]
56. Wu, P.; Zhang, Z.; Peng, X.; Wang, R. Deep learning solutions for smart city challenges in urban development. *Sci. Rep.* **2024**, *14*, 5176. [CrossRef]
57. Han, L.; Peng, C.; Xu, Z. The Effect of Commuting Time on Quality of Life: Evidence from China. *Int. J. Environ. Res. Public Health* **2022**, *20*, 573. [CrossRef]
58. O'Keeffe, K.P.; Anjomshoaa, A.; Strogatz, S.H.; Santi, P.; Ratti, C. Quantifying the sensing power of vehicle fleets. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 12752–12757. [CrossRef] [PubMed]
59. Chen, L.; Lu, Y.; He, Z.; Chen, Y. Online Trajectory Estimation Based on a Network-Wide Cellular Fingerprint Map. *Sensors* **2022**, *22*, 1605. [CrossRef]

60. Lilhore, U.K.; Imoize, A.L.; Li, C.T.; Simaiya, S.; Pani, S.K.; Goyal, N.; Kumar, A.; Lee, C.C. Design and Implementation of an ML and IoT Based Adaptive Traffic-Management System for Smart Cities. *Sensors* **2022**, *22*, 2908. [CrossRef]
61. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/jaimeblasco/500-cities-local-data-for-better-health-2019 (accessed on 10 January 2024).
62. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/mahdiehhajian/united-states-of-america-health-indicators (accessed on 2 April 2024).
63. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption (accessed on 10 July 2024).
64. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/mrsimple07/energy-consumption-prediction (accessed on 10 July 2024).
65. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/utathya/smart-city-traffic-patterns (accessed on 10 July 2024).
66. Kaggle Dataset. Available online: https://www.kaggle.com/datasets/fedesoriano/traffic-prediction-dataset (accessed on 10 July 2024).
67. Alsuwian, T.; Shahid Butt, A.; Amin, A.A. Smart Grid Cyber Security Enhancement: Challenges and Solutions—A Review. *Sustainability* **2022**, *14*, 14226. [CrossRef]
68. Ding, J.; Qammar, A.; Zhang, Z.; Karim, A.; Ning, H. Cyber Threats to Smart Grids: Review, Taxonomy, Potential Solutions, and Future Directions. *Energies* **2022**, *15*, 6799. [CrossRef]
69. Choi, H.-S.; Song, S.-K. Direction for a Transition toward Smart Sustainable Cities based on the Diagnosis of Smart City Plans. *Smart Cities* **2023**, *6*, 156–178. [CrossRef]
70. Hashem, I.A.T.; Usmani, R.S.A.; Almutairi, M.S.; Ibrahim, A.O.; Zakari, A.; Alotaibi, F.; Alhashmi, S.M.; Chiroma, H. Urban Computing for Sustainable Smart Cities: Recent Advances, Taxonomy, and Open Research Challenges. *Sustainability* **2023**, *15*, 3916. [CrossRef]
71. Castellanos-Nieves, D.; García-Forte, L. Strategies of Automated Machine Learning for Energy Sustainability in Green Artificial Intelligence. *Appl. Sci.* **2024**, *14*, 6196. [CrossRef]
72. Castellanos-Nieves, D.; García-Forte, L. Improving Automated Machine-Learning Systems through Green AI. *Appl. Sci.* **2023**, *13*, 11583. [CrossRef]
73. Gotovtsev, P. How IoT Can Integrate Biotechnological Approaches for City Applications — Review of Recent Advancements, Issues, and Perspectives. *Appl. Sci.* **2020**, *10*, 3990. [CrossRef]
74. Trappey, A.J.C.; Chen, P.P.J.; Trappey, C.V.; Ma, L. A Machine Learning Approach for Solar Power Technology Review and Patent Evolution Analysis. *Appl. Sci.* **2019**, *9*, 1478. [CrossRef]
75. Dostatni, E.; Rojek, I.; Hamrol, A. The Use of Machine Learning Method in Concurrent Ecodesign of Products and Technological Processes. In *Advances in Manufacturing. Lecture Notes in Mechanical Engineering*; Hamrol, A., Ciszak, O., Legutko, S., Jurczyk, M., Eds.; Springer: Cham, Switzerland, 2018; pp. 321–330.
76. Rojek, I.; Jasiulewicz-Kaczmarek, M.; Piechowski, M.; Mikołajewski, D. An Artificial Intelligence Approach for Improving Maintenance to Supervise Machine Failures and Support Their Repair. *Appl. Sci.* **2023**, *13*, 4971. [CrossRef]
77. Narkhede, G.; Mahajan, S.; Narkhede, R.; Chaudhari, T. Significance of Industry 4.0 technologies in major work functions of manufacturing for sustainable development of small and medium-sized enterprises. *Bus. Strategy Dev.* **2023**, *7*, e325. [CrossRef]