

# A Survey of Explainable Artificial Intelligence for Smart Cities

Abdul Rehman Javed <sup>1,2,\*</sup> , Waqas Ahmed <sup>1</sup>, Sharnil Pandya <sup>3,4</sup> , Praveen Kumar Reddy Maddikunta <sup>5</sup> , Mamoun Alazab <sup>6</sup>  and Thippa Reddy Gadekallu <sup>2,5</sup> 

<sup>1</sup> Department of Cyber Security, Air University, Islamabad 44000, Pakistan

<sup>2</sup> Department of Electrical and Computer Engineering, Lebanese American University, Byblos P.O. Box 36, Lebanon

<sup>3</sup> Department of Computer Science and Media Technology, Faculty of Technology, Linnaeus University, 351 95 Vaxjo, Sweden

<sup>4</sup> Faculty of Technology, Symbiosis Institute of Technology, Symbiosis International Deemed University, Maharashtra 412115, India

<sup>5</sup> School of Information Technology and Engineering, Vellore Institute of Technology, Tamil Nadu 632014, India

<sup>6</sup> College of Engineering, IT and Environment, Charles Darwin University, Casuarina, NT 0810, Australia

\* Correspondence: [abdulrehman.cs@au.edu.pk](mailto:abdulrehman.cs@au.edu.pk)

**Abstract:** The emergence of Explainable Artificial Intelligence (XAI) has enhanced the lives of humans and envisioned the concept of smart cities using informed actions, enhanced user interpretations and explanations, and firm decision-making processes. The XAI systems can unbox the potential of black-box AI models and describe them explicitly. The study comprehensively surveys the current and future developments in XAI technologies for smart cities. It also highlights the societal, industrial, and technological trends that initiate the drive towards XAI for smart cities. It presents the key to enabling XAI technologies for smart cities in detail. The paper also discusses the concept of XAI for smart cities, various XAI technology use cases, challenges, applications, possible alternative solutions, and current and future research enhancements. Research projects and activities, including standardization efforts toward developing XAI for smart cities, are outlined in detail. The lessons learned from state-of-the-art research are summarized, and various technical challenges are discussed to shed new light on future research possibilities. The presented study on XAI for smart cities is a first-of-its-kind, rigorous, and detailed study to assist future researchers in implementing XAI-driven systems, architectures, and applications for smart cities.

**Keywords:** Explainable Artificial Intelligence (XAI); smart cities; artificial intelligence; machine learning; future cities; Internet of Things (IoT); data mining; smart health; real-time systems



**Citation:** Javed, A.R.; Ahmed, W.; Pandya, S.; Maddikunta, P.K.R.; Alazab, M.; Gadekallu, T.R. A Survey of Explainable Artificial Intelligence for Smart Cities. *Electronics* **2023**, *12*, 1020. <https://doi.org/10.3390/electronics12041020>

Academic Editor: Gwanggil Jeon

Received: 14 January 2023

Revised: 9 February 2023

Accepted: 17 February 2023

Published: 18 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Information and communication technologies (ICT) and the Internet of Things (IoT) are essential components of smart cities, which may boost operational efficiency and improve services while helping residents lead sustainable lives [1,2]. Viably, the administration of accessible resources benefits from information and communication advances. There is a critical purpose for developing smart technologies to enable people to experience new and better things in their everyday lives [3,4]. In addition to making these technologies more efficient, increasing their efficiency may make them eco-friendly, more productive, and more flexible. While the digital revolution was important in many ways, it was also crucial for businesses to think about how easy and efficient it would be to run their businesses.

ICT is essential to the smart cities idea. It is not only crucial in policy formulation, decision-making, implementation and the provision of valuable services, but it is also essential in all other stages of the strategy. Artificial intelligence (AI) can help make cities more efficient in many areas of life, including their use in energy management, temperature management, education, health and human services, water management, air quality

management, traffic management, payments and finance, smart parking, and trash management [5–7]. A smart, AI-powered city will use energy and resources more efficiently, protect the environment, improve its citizens' lives, and enable them to adopt current ICT more quickly. Specifically, (i) technology and data availability and reliability, the dependency on third parties and the lack of skills are limiting factors; (ii) ethical issues when using AI are complicated; and (iii) regulatory issues when attempting to interconnect infrastructures and data are complex.

Explainable AI (XAI) in smart city development plays a crucial part. Recent applications based on deep learning, big data, and IoT architectures need intensive use of complicated computational solutions. Since these systems are closed to users, they are called “black boxes.” People will fear that their tools may be untrustworthy if this is true. In the last few years, attempts have been made to solve this problem using XAI methodologies to make things more transparent.

### *1.1. Clinical Decision Support Systems*

Clinical decision support systems (CDSS) are computerized systems that help healthcare providers use information more intelligently, helping to improve both patient health and the healthcare process. In order to accomplish many goals, CDSSs are developed with the following capabilities: diagnosis, prediction of treatment response, the suggestion of treatments (personalization), prognosis and the prioritizing of patient care based on risk [8]. In addition, CDSS can be beneficial in areas with limited resources, such as the number of healthcare facilities, equipment and physicians. CDSS may be categorized as knowledge-based or non-knowledge-based. While non-knowledge-based CDSS are usually based on AI, knowledge-based CDSS rely on medical guidelines and knowledge ([https://www.limswiki.org/index.php/Clinical\\_decision\\_support\\_system](https://www.limswiki.org/index.php/Clinical_decision_support_system) accessed date: 14 December 2022)). The AI-based CDSS examines past clinical data to produce prediction models to assess new input variables. When these results are used as guidance for doctors, these recommendations can aid them in their practices. There is tremendous promise for AI-based CDSS in clinical practice. Using a CDSS increases clinical choices while reducing medical mistakes since it is objective and relies only on input data and decision-making logic.

Nonetheless, their use of data depends on the amount and quality. Bias in training data results in skewed or inaccurate predictions for AI models. Biased or erroneous human decisions are likely to occur if this practice is widespread [9–11].

### *1.2. The Need for XAI: Fair and Ethical Decision-Making*

Understanding the mathematical underpinnings of existing machine learning architectures may only provide insights into how and why a result was obtained, not into the inner workings of the models. One must use explicit modeling and reasoning techniques to answer questions like “How did that happen?” One also knows that contextual adaptation, e.g., systems that aid in developing explanatory models for tackling real-world issues, will be a crucial difficulty for future AI. Human expertise should not be excluded, but AI should supplement it. When classification findings may lead to dangerous incidents for people, it is necessary to comprehend the mechanism that is at work behind such outcomes. Complex machine learning models are an essential focus of XAI research. Machine learning models may be classified according to their interpretability or opacity.

### *1.3. Motivation*

XAI has already risen to the top of research and development in the field of smart cities [1]. Large-scale decision-making will be required to carry out daily activities in smart cities. Since these decisions are made in black boxes, people need to know how a particular decision is made, as well as there is a developing situation amongst policymakers in smart cities with this loss of explainability of AI-based black-box models, which is a primary problem in the acceptability of AI-based decisions. Here, XAI comes into play to explain

these decisions. XAI converts black-box to white box to make understandable, interpretable, transparent, and explainable decisions [12].

There are a few surveys on XAI for smart cities; however, only a few go into more depth, look at the benefits of XAI for smart cities, key technologies of XAI and open research problems, and ongoing projects in XAI for smart cities. Authors in [13] provided a survey on explainable deep learning models for smart city solutions such as flood detection and drainage monitoring. The authors in [14] have broadly described the rapidly-growing field of XAI-related research. Authors in [15] provided thorough research on XAI related to medical applications. Authors in [16] suggested a taxonomy and classifying the XAI approaches based on their breadth of explanations, methodology used behind the algorithms, and explanation level of utilization.

Table 1 concisely presents a comparison of the important survey papers and also highlights the gap in existing surveys, which is the need for a comprehensive analysis of XAI for smart city decision-making that is trustworthy, responsible, and transparent. Keeping in mind the limitations and gaps of existing surveys, there is a high need for research focusing on XAI for smart cities.

**Table 1.** Summary of important surveys on XAI.

Ref.	XAI Driving Trends	Applications/Use Cases	Requirements/Vision	Technical Challenges	Enabling Technologies	Ongoing Activities	Research Directions	Remarks
[13]	L	H	L	M	M	L	M	This study provides a survey on explainable deep learning for smart cities solutions such as flood detection and drainage monitoring.
[14]	L	M	L	H	M	L	M	This study serves as a stepping stone for academics and practitioners who want to grasp the details of the rapidly-growing field of XAI-related research.
[15]	M	M	L	M	H	M	M	A thorough research on XAI related to medical usage and its branches.
[16]	M	L	M	H	M	H	H	This study suggested a taxonomy and classifying the XAI approaches based on their breadth of explanations, methodology used behind the algorithms and explanation level of utilization.
[17]	H	L	M	H	H	M	M	By exploring different application fields, such as medical informatics, law, the semantic web, security, robotics, and certain general-purpose systems, the researchers elucidate how XAI Argumentation aids in the construction of explainable systems.
[18]	L	L	M	H	H	M	G	This article shares XAI design and evaluation framework experiences from many fields. Supporting different design goals and assessment techniques is another purpose of this research project.
[19]	L	M	L	H	H	L	L	A survey of AI explainability research and techniques is presented in the literature. Initially, we surveyed the CS research community to locate the most relevant studies on AI explainability, or “XAI.” Then, we researched HCI.
[20]	H	M	H	H	H	L	M	In this review, they examine and overview several XAI techniques developed utilizing approaches derived from computational argumentation.
[21]	L	M	L	H	H	L	L	This article goes into great detail on current XAI approaches of possible importance for pathological imaging deep learning algorithms and organizes them based on three key facets. By including uncertainty analysis methodologies as an inherent component of the XAI environment, they provide their clients with more assurance in predictions.
[22]	L	M	L	H	H	L	H	They provide a cutting-edge, browser-based system, called XNLP, that showcases new and relevant research in XAI through a dynamic survey of recently published state-of-the-art work in the field of natural language processing (NLP).
This paper	H	H	H	H	H	H	H	This study addresses current, projected and future developments and use cases in XAI for smart cities.

L Low Coverage

M Medium Coverage

H High Coverage

#### 1.4. Contributions

As the XAI concept in smart cities is transforming with the advances in complementary technologies, a paradigm shift is needed to drive and transform AI-driven projects into XAI-

driven mode. To assist researchers in initiating their research work in the domain of XAI for smart cities, a rigorous and informative study is required to provide current, projected, and future insights. The presented study provides in-depth insights and analysis of various XAI technologies for Smart cities. The presented survey discusses use cases, standard practices, challenges, the latest trends, and future research directions. The contributions of the presented study are:

- The presented study on XAI for smart cities is a first-of-its-kind, rigorous, and detailed study to assist future researchers in implementing XAI-driven systems, architectures and applications.
- To assist readers, this paper discusses concise definitions and key technologies of XAI and its explanation using various smart city-driven use cases (see Section 2).
- In previous studies, more focus has been given to XAI applications for smart cities. The undertaken study discussed the concept of XAI for smart cities, various XAI technology use cases, challenges, applications, possible alternative solutions and current and future research enhancements (see Sections 2–7).
- The core contribution of the conducted study is to discuss open research problems and challenges (see Section 5).
- Furthermore, It also discusses ongoing smart city-driven XAI projects (see Section 6).
- Finally, security, privacy, ethical issues and regular compliance for implementing XAI in smart cities are discussed and future directions are provided (see Section 7).

### 1.5. Organization

Figure 1 depicts the taxonomy of the survey paper. Table 2 provides the list of abbreviations. The paper is further organized as follows: Section 2 defines key aspects and technologies of XAI. Section 3 explains the XAI technologies and their implementation in current and emerging smart cities. Section 4 identifies the applications of XAI in smart cities. Section 5 discusses the open research issues and challenges. Section 7 presents XAI's future development and research direction in smart cities. Finally, Section 8 concludes this paper. Section 6 presents the ongoing XAI projects for smart cities.

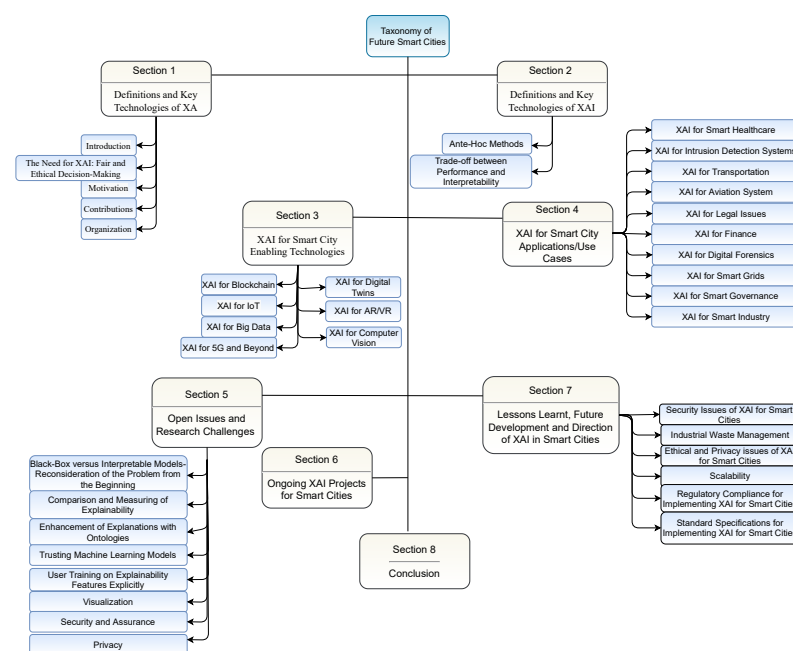


Figure 1. Taxonomy diagram of the explainable artificial intelligence for smart cities.

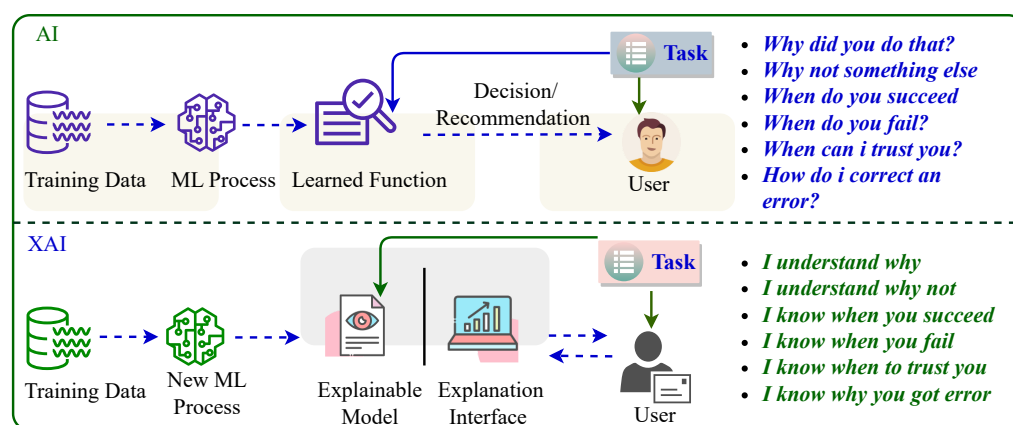
**Table 2.** List of Abbreviations.

Abbreviation	Description
5G	Fifth-Generation
6G	Sixth-Generation
AI	Artificial Intelligence
ADLs	Activities of Daily Living
ANN	Artificial Neural Networks
AR	Augmented Reality
AQI	Air Quality Index
AIoMT	Artificial Intelligence of Medical Things
BRB	Belief-Rule-Base
BRL	Bayesian Rule learning
CDSS	Clinical Decision Support Systems
CCTV	Closed-Circuit Television
CC	Cognitive Computing
DL	Deep Learning
DNN	Deep Neural Network
DT	Digital Twin
DT	Decision Tree
Grad-CAM	gradient-weighted class activation mapping
ICT	Information and Communication Technologies
IDS	Intrusion Detection System
IoT	Internet of Things
IoMT	Internet of Medical Things
KNN	K-Nearest Neighbours
LIME	Local Interpretable Model-Agnostic Explanations
ML	Machine Learning
NLP	Natural Language Processing
PDPs	Partial Dependence Plots
PDPs	Partial Dependence Plots
QoS	Quality of Service
R-CNN	Region-Based Convolutional Neural Networks
SCM	Supply Chain Management
SHAP	SHapley Additive exPlanations
SP	Service Provider
VR	Virtual Reality
VANET	Vehicular Ad-Hoc Networks
XAI	Explainable Artificial Intelligence
XML	Explainable Machine Learning
YOLO	You Only Look Once

## 2. Definitions and Key Technologies of XAI

XAI is a method that helps humans understand how the output is created by the machine/deep learning algorithm. It contributes to quantifying model correctness, fairness, and transparency and results in AI-assisted decision-making. XAI is critical for organizations' trust and confidence when using AI models. Additionally, AI explainability enables organizations to take a responsible approach to AI development [23].

Figure 2 shows the difference between the working methodology of AI and XAI. As AI becomes more sophisticated, humans face difficulty comprehending and retracing the algorithm's steps. The entire calculating process is transformed into what is often referred to as a "black box" that is impenetrable. These black-box models are constructed using the data directly, and not even the algorithm's developers or data scientists understand or can describe what is occurring inside them or how the AI algorithm arrived at a particular conclusion [24]. As machine learning models improve performance, explainable and interpretable predictions become increasingly hard to create. Black-box models may be defined as deep learning [25] or ensembles [26–28]. On the other hand, white-box or glass-box models are known as open-source models because it is straightforward to create understandable results using explainable examples, like the reference [14] and decision tree models. It is possible to make the new models more understandable and interpretable. However, these models still need state-of-the-art performance compared to the earlier models. It boils down to their frugal design; when they perform poorly or are well-interpreted and quickly explained, it is due to them having a frugal design [29]. Interpretability, as mentors remark, is beneficial for accomplishing other model goals, which may include establishing user confidence, recognizing the effect of certain factors, comprehending how a model would behave given inputs and ensuring that models are fair and impartial [30].



**Figure 2.** Difference between the working methodology of AI and XAI.

1. **Understandability :** It denotes that the goal of a model is to allow a human to comprehend how the model works, and a model is considered to be excellent if the model allows a human to gain understanding without being dependent on understanding the model's internal structure or algorithmic techniques by which the model processes data [31]. Descriptions should be composed of little pieces of information, which humans and computer algorithms can interpret and should connect quantitative and qualitative ideas [32].
2. **Comprehensibility:** The capacity of a learning system to express its learned information in a human-comprehensible manner is known as comprehensibility [33]. Computability is typically related to model complexity when comprehensibility is considered [34].
3. **Explainability:** As an interface between people and a decision-maker, an explainability is understandable to humans while also providing a close approximation of the decision-maker [34].



4. **Transparency:** For a model to be transparent, it must be comprehensible by itself. Simulatable, decomposable and algorithmically transparent models are all types of transparent models [35].

### 2.1. Ante Hoc Methods

Predictive models and explainability techniques have two fundamental connections between them. An illustration of ante hoc is using the same model to predict and explain, such as using feature weights to predict a linear regression. It is critical to point out that there are assumptions associated with these approaches and they must be met if the explanation is to operate as expected [36,37]. Clear indications about the underlying prediction algorithm and data, whether transparent or opaque, should be used to have an explainable approach. Can someone describe the workings of a predictive model without knowing how it works? When explaining data or forecasts, is it necessary to make data characteristics intelligible to humans first? For example, while describing a forecast, think of how it may be stated in terms of the temperature in a room compared to calculating a squared sum of the room temperature and the height. Sometimes the system designer wants to provide a list of data transformations or an example-based explanation in instances when the input domain is not intelligible to humans [35,38].

### 2.2. Trade-Off between Performance and Interpretability

As with any other bold statement, the issue of interpretability versus performance becomes bogged down in myths and misconceptions. It might be true that the more complicated the model is, the better it is for predictions, but it is only sometimes true [39]. In such circumstances, the prediction is wrong. This phenomenon is quite prevalent in several industries. For example, because features must be tested in confined physical settings where all of the characteristics are closely linked and no wide range of potential values is represented in the data, many issues have just a limited subset of features [40]. More predictive models allow for more complex functions to make a prediction. Predictability involves having specific complexity and the data is widely distributed across the world of suitable values for each variable. There is also sufficient data available to use a complex model. Therefore, “more complex models are more accurate” appears correct. The trade-off between performance and interpretability may be seen in this circumstance. Also, it is essential to remember that, while attempting to resolve problems that do not follow the abovementioned principles, an organization will risk fixing too simplistic (variance) problems. A key difference between a complicated model and a sophisticated one is that a complicated one makes the prediction process more complicated, whereas a sophisticated one keeps the model simple but accurate. The performance brings complexity with it and as a result, interpretability confronts itself on a downhill slope. However, when these newer techniques of explainability arise, this would flip or nullify the increase in explainability [16].

## 3. XAI for Smart City Enabling Technologies

In general, smart cities are defined as technology-enabled, socially intensive and environmentally friendly urban areas. This section discusses key enabling technologies of smart cities such as blockchain, IoT, big data, 5G and beyond technologies, digital twins, AR/VR, and computer vision. Researchers have recently attempted to address smart city problems using predictive and advanced analytics, smart environmental monitoring and smart mobility. Still, the design and development of smart cities have remained an open research problem [41]. With the advancements in technologies such as computer vision, IoT, and big data and AI, it has become feasible to address smart cities-related challenges [42–44]. Recently, numerous machine learning and deep learning methodologies have been applied to applications related to smart cities. For instance, fine k-nearest neighbors (KNN) [45], decision tree [46], medium KNN [47], You only look once (YOLO) v4 and YOLOv5 [48] and mass region-based convolutional neural networks (R-CNN) methodologies are ap-

plied for image classification, traffic congestion, and autonomous driving vehicle-related problems [49,50]. However, most machine learning models depend highly on training data sets, pre-processing data methodologies, and fine-tuning non-transparent machine learning models. The current data-driven methodologies cannot understand, interpret, and explain complex decision-making-related tasks [51]. Such black box-like methodologies cannot identify and explain decision-making in smart cities. XAI algorithms can be applied to computer vision problems, traffic congestion, self-driving cars, intrusion detection systems (IDS) [52–54], etc., to address the above issues. It is also essential to monitor and validate the behavior of the smart city application before deploying advanced machine and deep learning models.

### 3.1. XAI for Blockchain

Blockchain is an immutable ledger that can process various transactions and carry out asset management and tracking. Blockchain is widely known for tracking and trading physical assets, such as houses, cars, and real estate properties and virtual assets such as patents, copyrights, and many more [51]. Furthermore, blockchain technology is suitable for providing real-time information in a completely transparent and secure environment. Blockchain technology can be implemented using smart contracts. Smart contracts are predetermined constraint-based programs that are executed on a blockchain network. It enables the automation of agreements without any interruptions.

Blockchain technology is one of the essential solutions to drive cloud-based data centers. Moreover, blockchain technology brings reliability and trustworthiness aspects in designing and developing secure engineering solutions [55]. However, the merits of blockchain technology also bring two critical challenges: (i) cross-layer implementation of blockchain technology in cloud computing environments and (ii) the need for a control mechanism due to the automation of tasks in most blockchain-based solutions. Blockchain technology systematically organizes data to prevent intruders from manipulating or hacking the system. The increasing complexity of data-driven AI methodologies and their non-transparent behavior presents numerous security challenges to smart city problems [56]. XAI methodologies have extended and improved the explainability of AI models. Especially in decentralized AI applications, integrating blockchain with AI models guarantees data privacy and security. It also provides data accessibility and traceability. Integrating AI and blockchain architecture can implement a secure decentralized framework for storing and retrieving data generated with AI models.

As every coin has two sides, combining blockchain with XAI methodologies will also bring numerous future challenges which need to be tackled smartly. Firstly, the validation of explanations given by XAI methodologies is a huge challenge due to the minimization of humans in the loop. Secondly, achieving the real-timeliness of applications is a big concern, which requires an immediate resolution [57].

### Applications of XAI for Blockchain

Integrating XAI with blockchain can support diverse trusted, non-transparent, secure, decentralized and undisputed systems and application domains.

- a. **Customer Profile Assessment:** Integrating XAI with blockchain will massively affect banking and finance operations. The banking and finance applications are jointly integrated with blockchain technology (BCT) and XAI-based multi-agent systems (MAS). The XAI and blockchain-based multi-agent systems comprise various intelligent expert agents. The expert agents analyze bank customers' profiles, credit history, demographic information and health history. The blockchain-based multi-agent systems also enable effective decision-making and minimize risk probabilities in investments by fostering trust and transparency [57]. Furthermore, the integration of XAI with blockchain also assists in identifying creditworthy customers and decision-making about loan allotment, providing business finance, or empowering start-ups.



- b. **Medical Imaging:** Integrating XAI with blockchain technology can implement a secure decentralized medical diagnosis framework for medical image-based diagnosis. The XAI and blockchain-based medical diagnosis framework uses block-wise encryption and histogram shifting methodology to ensure secure transmission of historical patient data and provide trustable information about patient history, such as how, who, when, and where the patient profile data is created. This methodology can also assist radiologists in making decisions by explaining critical patient conditions. The combination of medical imaging technology, XAI and blockchain can assist doctors in making transparent, trustable, unbiased and effective decisions for critical patients [58].
- c. **Auditing:** The XAI with blockchain applications can protect organizations from money laundering, bank transaction fraud, and income and sales tax fraud. The integration of BCT and MAS uses a widely known consensus algorithm with the SHA-256 secure hashing algorithm to ensure the safety of banking and financial transactions. In addition, a joint integration of BCT and XAI-based multi-agent systems can analyze biased and disputed decisions, explain them, and assist government officials, judges, and lawyers in identifying and detecting various frauds. Furthermore, the integration of XAI with blockchain can also detect election-related frauds, such as manipulating votes and election results. [59].
- d. **Real-time Decisions:** Based on the intelligent sensing units connected to smart vehicles, XAI with blockchain solutions can assist in making real-time decisions such as fatal accidents and traffic congestions [60]. The XAI and blockchain-integrated systems contain deterministic and non-deterministic predictors. The deterministic predictors assist in making accurate decisions based on the inputs and data of the XAI methodology. For inexact decisions, the non-deterministic predictors produce inexact decisions. The deterministic or probabilistic predictors can make all types of AI-related decisions, such as frequent pattern mining, optimization-related decisions, clustering, classification-related decisions, and many more [57]. In non-real-time situations, XAI with blockchain solutions can be a blessing in identifying fatal accidents and handling traffic congestion solutions by diverting vehicles in the right direction. Figure 3 demonstrates a use case of XAI with blockchain for smart cities. As shown in Figure 3, XAI-based multi-agent systems integrated with blockchain methodology in applications, such as banking and finance, medical imaging and accounting, assist in identifying credit-worthy customers, and making decisions about their loans and finance. The blockchain methodology contains block-wise encryption, histogram shifting methodology, and SHA-256 cryptography. The XAI-based multi-agent systems integrated with blockchain methodology can explain the reasons for customers' credit and health history, banking, and finance transactions to make critical real-time decisions.

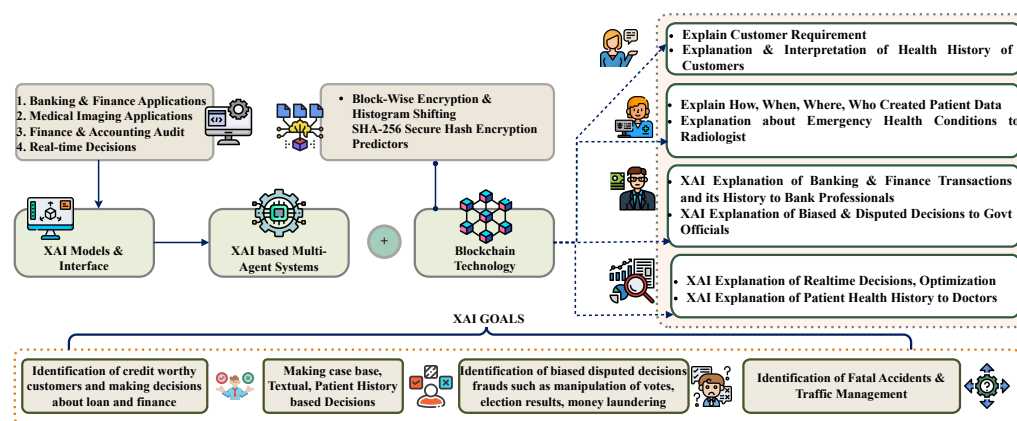


Figure 3. A Usecase of XAI with Blockchain for Smart Cities.

### 3.2. XAI for IoT

In the development of innovative smart city solutions, Internet of things (IoT) technologies has played a vital role in shaping the life of citizens by addressing issues such as traffic congestion, theft detection, geospatial farming, telemedicine, remote healthcare monitoring, and many more [61,62]. In recent times, smart city problems have been addressed using various data sensed via smart sensing devices. Furthermore, the introduction of intelligent edge computing and edge-AI-enabled devices and their integration with IoT technologies have envisioned the future directions of smart city concepts [63,64]. Edge-AI-enabled devices have been integrated with powerful machine and deep learning methodologies to analyze and predict smart city-related issues, such as the air quality index (AQI), weather prediction, accident prevention, and traffic monitoring.

However, Edge-AI-enabled devices and robust machine learning methodologies can not assist in real-time or near-real-time decision-making. To achieve this objective, XAI methodologies can be integrated with intelligent IoT, the Internet of medical things (IoMT), the AI of medical things (AIoMT), and edge AI-enabled smart devices to identify, interpret, and explain a particular situation which helps in making critical decisions. Figure 4 represents a use case of integrated XAI and IoT-enabled architecture for smart cities. As shown in Figure 4, edge-level XAI is responsible for collecting sensing information from Edge-AI-enabled smart devices and sending it to the application servers for processing via a cloud gateway [65,66]. Furthermore, the application server also notifies individuals of emergency alerts via network operators. The XAI-enabled IoT systems use methodologies such as QARMA, CDSS, LIME, etc. Thus, for example, XAI and IoT-integrated devices can notify the family members and explain the whole situation to them after a fatal accident, make critical decisions based on the critical health condition of patients, provide real-time notification and explain traffic congestion situations.

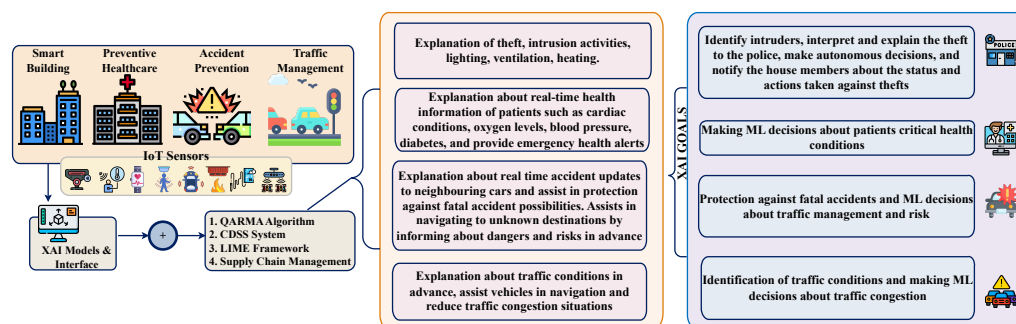
Combining XAI with blockchain and IoT infrastructure brings economic and scalability challenges. The scalability issue comes from the overall size of each block and its adaptability with the increase in the number of transactions. Along with the increase in the number of transactions, the handling and maintenance costs will also increase due to increased traffic. Furthermore, the increase in the number of users and transactions will also increase the latency time for processing. Still, the challenges of combining XAI with blockchain and IoT have remained open research issues [67]. Researchers have introduced methodologies such as Segwit, Sharding, Plasma, etc.

#### Applications of XAI for IoT

Integrating XAI with IoT can support various critical, trusted, decentralized and undisputed systems and application domains.

- a. Preventive Healthcare: The XAI integrated clinical decision support systems (CDSS) explain the relevance of XAI methodologies from various perspectives: (i) medical, (ii) technological, (iii) legal, and (iv) end-user (patient) perspectives. The XAI-integrated CDSS system uses the analysis findings to conduct a detailed ethical assessment of patients' profiles with appropriate explanations [68]. Integrating XAI, CDSS and edge, and edge AI-enabled smart devices can provide real-time information about patients, such as cardiac conditions, oxygen levels, blood pressure, diabetes, and many more; it also assists caretakers and healthcare experts and family members in making critical healthcare decisions. In addition, XAI and IoT-enabled frameworks can also be applied to advanced analytics on patients' vital health information and can predict health diseases in advance [69].
- b. Smart Building Management: XAI and IoT-enabled smart building/home architectures can autonomously control building operations [70,71]. The XAI systems integrated with QARMA algorithms and models monitor smart building operations. The QARMA methodologies can formulate quantitative rules for creating, updating, and managing smart building operations such as protection against thefts and intrusion activities, lighting, ventilation, heating, etc. Furthermore, XAI-integrated

- QARMA methodologies can also identify intruders, interpret, and explain the theft to the police, make autonomous decisions, and notify the house members about the status and actions taken against thefts [72].
- c. Accident Prevention: The XAI and IoT integrated frameworks, such as the local interpretable model-agnostic explanations (LIME) framework, can easily be integrated with LoRA and the LIME framework can explain the classification results generated by XAI algorithms. For example, The LIME integrated XAI and IoT-based systems can provide real-time accident updates to neighboring cars and protect against fatal accident possibilities. It also assists in navigating to unknown destinations by informing about dangers and risks in advance [73].
  - d. Traffic Management: Based on the intelligent sensing units connected to smart vehicles, XAI with IoT solutions can assist in smart vehicle management [74]. The XAI, supply chain management (SCM) and blockchain-integrated heuristic search methodology can assist in avoiding traffic congestion situations and identifying traffic conditions in advance. The XAI-enabled SCM system stores the information and time of every service provider (SP). The XAI-enabled SCM system is connected with smart networks such as vehicular ad-hoc networks (VANET). Every service provider has an open location key against the traffic information gathered by the XAI-enabled SCM system; finally, the stored information is integrated to identify traffic conditions in advance, assist vehicles in navigation, and reduce traffic congestion situations [75].



**Figure 4.** A use case of XAI with IoT for Smart Cities.

### 3.3. XAI for Big Data

Big data is a collection of structured, unstructured and semi-structured data, information and knowledge [76–79]. Integrating AI and big data-enabled systems collect, interpret, process and store large amounts of data. It also applies advanced analytics that can be used for predictive modeling and analytics [80]. However, due to the black box-like behavior of most AI methodologies, machine, and deep learning methodologies fail to interpret and explain large volumes of data resulting in poor decision-making. Integrating XAI algorithms with big data can assist big data systems in understanding, interpreting, and processing diverse and large volumes of data [81]. It can also assist in identifying complex patterns, categorization, dimensionality adjustments, and maintaining transparency and accountability of data [15]. The big-data-based XAI integrated decision support systems, multi-agent systems, and healthcare systems can assist organizations in explaining and making decisions about customer's geographical segmentation, health and personal history, selection and prediction of stocks, identification of anomalies in lifestyles of elderlies and planning and formulating supply chain strategies. Figure 5 represents a use case of integrated XAI and big data applications for smart cities. Furthermore, the knowledge layer is responsible for the timely generation of processed information to assist various big data systems and tools in decision-making. For example, healthcare experts and clinicians can apply medical approaches with caution based on the inputs received from XAI-integrated big data systems, which will assist them in selecting appropriate medical practices, medical

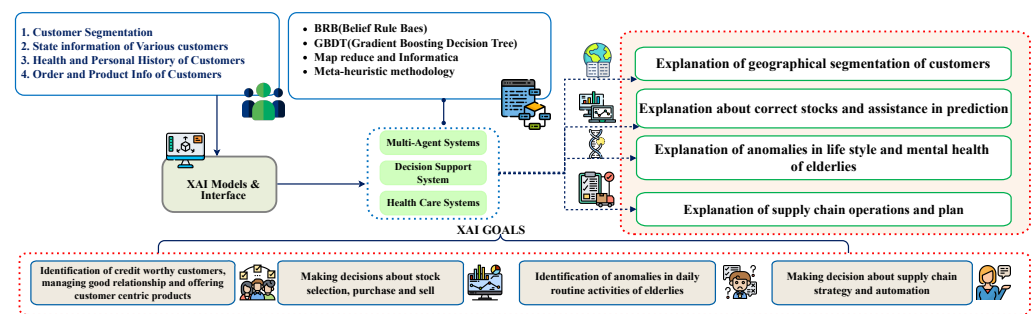
operations, and critical healthcare decisions. Finally, the service layer is responsible for notifying the real-time updates and explaining a particular event.

However, integrating XAI with big data technologies also challenges future directions. For instance, in applications such as healthcare, XAI can visualize the AI model and assist healthcare experts in decision-making. However, generating accurate medical diagnosis reports, deriving conclusions from them, and validating these reports is a major concern for healthcare experts [82].

#### Applications of XAI for Big Data

Integrating XAI with big data can support diverse, decentralized, risk-oriented, unbiased systems and application domains.

- a. **Customer Segmentation and Management:** The big data-based XAI integrated decision support systems can assist organizations in geographical segmentation, understanding, and interpretation of various types of customers. Such systems consist of belief-rule-base (BRB) and factual and heuristic rules. Furthermore, it can understand, analyze, extract, and interpret historical data using supervised machine learning and deep learning methodologies. The big data-based XAI integrated decision support systems also assist in managing customer relationships and offering them various customer-centric products, their explanation and benefits [83]. Furthermore, XAI methodologies can also assist in understanding the needs and requirements of regular customers in offering suitable products to customers.
- b. **Stock Prediction and Management:** The XAI integrated multi-agent stock prediction systems can assist stakeholders in picking the right stocks concerning the current know-how. Such systems are integrated with gradient boosting decision trees methodology, which can interpret and predict stock price inclines and declines [84]. It can also assist finance portfolio managers in making the right decisions about stock selection, purchase, and selling for their esteemed customers [85]. Furthermore, the XAI integrated Big data systems can better understand customer finance history, current needs, and aspirations to gain more profits and make future stock decisions.
- c. **Health Analytic of the Elderly:** The big data-enabled XAI integrated healthcare systems can analyze activities of daily living (ADLs) of the elderly, their daily lifestyle monitoring status and anomalies in routine activities; such systems use the concept of data-driven AI to identify the cognitive decline of activities of daily living in a smart home. It analyses cognitive decline and explains why a sudden decline has happened in routine ADLs. The data-driven big data-enabled XAI systems also identify the possibility of health diseases such as Dementia and Parkinson's [86]. Furthermore, such systems can assist healthcare experts in understanding the elderly's mental health conditions, cardiac conditions, blood pressure, and oxygen levels to predict the probability of critical health conditions and diseases in advance [87].
- d. **Business Analytics for Supply Chain Management:** The big data-integrated-XAI supply chain systems can plan and organize supply chain operations, gain real-time insights into customer feedback and build cost-effective business strategies for finance and marketing. Such systems use a robust meta-heuristics base that can analyze customers' and vendors' histories and formulate a cost-effective business strategy [88]. Furthermore, big data-integrated-XAI supply chain systems can also provide further recommendations for improving supply chain strategies based on customer necessities and feedback using meta-heuristics [89].



**Figure 5.** A Use case of XAI with Big Data for Smart Cities.

### 3.4. XAI for 5G and Beyond

Recently, 5G technologies have faced challenges such as connection density, network access issues in basements, underwater, and space, 24 × 7 availability, delay in communication, quality of service (QoS), and many more. Several research works have discussed 6G technologies, benefits, and standards [90]. However, their practical usage and applications still need to be completed. Standard machine and deep learning methodologies integrated with 5G and beyond technologies can implement technologies such as remote healthcare monitoring, telemedicine, Industry 4.0 and digital twin, but fail to understand, interpret, and assist in real-time decision-making [91–93].

However, by integrating XAI methodologies with 5G and beyond technologies, edge-AI-enabled smart devices can facilitate humans with various intelligent applications such as connected robots, collaborative autonomous driving, smart and interpretable health, remote surgery, connected restaurants, connected cars, connected assembly, and many more [94]. However, XAI for 5G and beyond technologies also brings challenges and research directions. In applications such as precision manufacturing and autonomous vehicles, a human-computer interface-driven interface is required for smooth functioning and communication for smart city applications [95,96].

#### Applications of XAI for 5G and Beyond

The integration of XAI with 5G and beyond technologies can change accessing real-time updates, bringing humans and machines together on a trusted platform and automating information delivery. As a result, it can play a vital role in smart city design and development and global societal development.

- a. **Precision Manufacturing:** The 5G and XAI-enabled manufacturing systems with human and machine participation in smart factories can increase productivity, organize production processes, automate factory operations, and increase flexibility [97,98]. Such systems consist of a gradient-boosting decision tree methodology to identify machine errors or any errors made by tools/soft-wares. Furthermore, the 5G and XAI-enabled manufacturing systems can perform focused analytics using a reliable AI-integrated prediction model for maintenance tasks. It can develop a flexible production environment for future product trends and customer needs [99].
- b. **Connected Robotics:** The 5G-enabled explainable agents and robots can automate restaurant operations (connected restaurants), drive and control autonomous vehicles, automate and control cooking tasks and connect assembly systems [100]. The explainable agents and robots can explain their behavior to humans and carry out intra-agent explanations for a particular task [101]. Furthermore, the 5G enabled explainable agents and robots can understand, interpret and explain the allotted tasks to robots and provide real-time status to the humans [102].
- c. **Collaborative Autonomous Driving:** The vision-based autonomous driving systems, along with 5G and beyond technologies, can enable collaborative autonomous driving, inter-vehicle, vehicle-to-infrastructure and vehicle-to-vehicle communications in near real-time using technologies such as LoRa. The combination of behavior cloning and reinforcement learning methodology can carry out imitation-based learning from



- human driving lessons and make safety-critical decisions [103]. Furthermore, vision-based autonomous driving systems can also provide background insights about road accidents, geographical conditions, alternative routes, smart car parking, platooning, assistance for changing and merging lanes, and managing intersections [104].
- d. Targeted Healthcare: The 5G enabled EMR systems connected with clinicians and medical experts can perform feature interpretability analysis of health patients and provide essential information, quantitative and qualitative assessment of patient health history using methodologies such as local explanations, global explanations, local interpretable model-agnostic explanations (LIME), SHapley additive exPlanations (SHAP), Example-based Techniques, and Feature-based Techniques [105]. Furthermore, such intelligent methodologies, integrated with 5G-enabled smart and connected health systems, can also assist healthcare experts in combating fatal diseases such as COVID-19 [106].

### 3.5. XAI for Digital Twins

A virtual replica spans a physical system or object's lifecycle called a digital twin (DT). A DT is updated frequently with real-time data and uses ML simulation and reasoning for decision-making. DT is a highly complex virtual model and replica of its physical counterpart, which can be anything ranging from a car to a person to a building to a city to a bridge. The data from the physical assets are collected from the sensors connected to them and are mapped to the virtual model. The behavior of the real-world object/thing can be understood or visualized by looking at the DT. Using DT, we can understand how the objects are behaving presently and predict how they will behave in the future by analyzing the data from the sensors [107]. Even though DTs were originally designed to improve the manufacturing process using simulations, they are now being used in several application domains, such as healthcare and smart cities, due to the increase in big data generated from several IoT-based applications [108]. The potential applications of DT in building/designing effective smart city services are increasing every year due to the increased connectivity by IoT devices [109]. The potential applications of DTs in the smart city include planning and developing smart cities and energy saving in smart cities. The data from the utilities in smart buildings gives insights into the usage patterns and distribution of the utilities, through which decisions can be taken based on the predictions made by ML algorithms and big data analytics [110]. DTs can facilitate the growth of a smart city by creating testbeds inside a virtual twin. A smart city DT can achieve two objectives; the first thing is that DTs can act as a testbed for testing the scenarios and the other one is that DTs can analyze the changes in the collected data and learn from the environment that can be used for monitoring and data analytics [111]. For instance, DT-integrated clinical decision support systems can set a threshold for doctors using patient history and meta-heuristics. The decision thresholds can assist doctors in recommending medicines and deciding tests and treatments based on patient conditions.

A smart city DT can be created by integrating building information models with the big data generated by sensors from IoT devices in a smart city. The public can walk around the city's accurate 3D model created by the DT. It can observe the proposed changes in the policies and urban planning that will pave the way for public opinion before the decisions come into practice [112,113]. AI-enabled smart city DTs can be used effectively to plan for preparation, mitigation, and response during natural disasters and calamities during floods and earthquakes [114,115]. Even though AI can help DTs simulate smart cities, through which the authorities can take necessary actions, due to the black-box nature of AI/ML algorithms, it is challenging for the authorities to understand the reason for the predictions/classifications. In mission-critical applications such as traffic control and disaster management, wrong decisions taken by authorities can affect millions of lives and properties in urban areas. Through the transparency and justification of the predictions, XAI can alleviate these problems faced by the authorities in making decisions based on the predictions given by AI algorithms in the smart city DTs. Moreover, combining



XAI methodologies with digital twins brings interface and optimization challenges [95]. The XAI-based digital twin systems are at an initial stage and cannot handle massive data processing and self-optimization in digital twin systems [116].

### 3.6. XAI for AR/VR

Augmented reality (AR) can use sounds, digital visual elements, sound, or other sensors to enhance the physical world. The main aim of AR is to highlight some essential features of the real world, understand those features better, and come up with smart insights that can be applied to the applications in real-world [117,118]. Virtual reality (VR) generates a virtual environment with objects and scenes, making users feel as if they are physically immersed in the virtual environment. The virtual environment created by VR can be perceived using a VR headset [119]. AR/VR coupled with AI/ML are key enablers of the smart city through the urban planners and the general public can view the virtual urban planning and simulations of events in a smart city. The AI/ML algorithms' lack of reasoning/justification by the AI/ML algorithms on predictions of some of the events, such as disasters/accidents, makes it difficult for the concerned to make decisions solely based on the predictions from the AI/ML algorithms in AR/VR applications in smart cities [118]. XAI can address these issues by providing interpretability and justification for the prediction results of the AI/ML algorithms. For instance, 5G and AR/VR-enabled recommender systems use explanation-enabled recommendation methodology (XARSSA). The XARSSA methodology can address the influence of customer demographics, extracting customer demands and choices using AR/VR-based shopping assistant applications. The XARSSA methodology uses a design science approach to attract a large customer base towards AR/VR-enabled shopping assistant applications and gain more insights about their needs [120].

Furthermore, combining XAI methodologies with AR/VR techniques will assist in smooth functioning and interfacing; however, achieving complete transparency and trust between the human brain and machine interface has remained an open research problem [19]. In addition, XAI-based AR/VR systems have to deal with massive data processing and high-performance computing challenges.

### 3.7. XAI for Computer Vision

Computer vision is an application of AI that enables the systems to understand meaningful information from videos, digital images and closed-circuit television (CCTV) footage and make recommendations based on the information extracted from these sources [121,122]. Computer vision has many applications in smart cities, such as object detection in autonomous vehicles that can avoid collisions/accidents, reduction in traffic, monitoring of suspected criminals [123] that will, in turn, reduce the crimes, structural monitoring, combating disasters [124,125]. The 5G and computer vision integrated multi-agent systems use reinforcement learning methodology to analyze traffic congestion situations. Such systems are integrated with interpretations techniques such as SHAP, LIME and gradient-weighted class activation mapping (Grad-CAM) for explaining various traffic situations to a driver and assisting in making driving-related decisions in dense traffic situations [126]. The applications mentioned above are very sensitive as the lives of the citizens are at stake if wrong decisions are taken. Traditional computer vision applications do not explain or justify the classification of images/videos. Hence, making real-time decisions based on the classifications given by computer vision-based applications in scenarios in smart cities, such as collision avoidance, traffic monitoring and crime prevention, may incur severe costs, such as loss of lives and ethical issues. XAI can solve the issues related to computer vision-based applications in smart cities through explainability, interpretability, and justification of classification results.

Combining XAI with computer vision methods also brings a few challenges, such as data exploration and measuring the complexity of black-box models. XAI with computer vision can easily work with text, image, and audio data [127]. However, it cannot interpret

spatiotemporal quantities, matrices and vectors. Furthermore, the complexity of interpretability highly depends on black-box models, such as the depth of trees, the presence of non-zero weights in neural network models, etc.

#### 4. XAI for Smart City Applications/Use Cases

Currently, AI is playing the leading role in several domains, such as healthcare, entertainment, and business, regarding progress and innovations. From the current solutions point of view, the influence of deep learning (DL) is based on machine learning (ML) and artificial neural networks (ANN) have shown good robustness and excellent performance. Still, from the decisions and human-understandable perspectives, they lack the abilities. Therefore, the need for more abilities in the more sensitive domain is crucial. Some of the applications of smart cities that can benefit from XAI integration are discussed below.

##### 4.1. XAI for Smart Healthcare

Advanced technologies like AI, the IoT and cloud computing refer to smart healthcare systems in the existing health system. Figure 6 presents smart healthcare systems in smart cities. The advanced technologies enable the healthcare system to be more personalized, convenient, and efficient [128]. Furthermore, advanced technologies help in health monitoring in real-time through available applications of healthcare on wearable or smartphone devices. Thus, it motivates individuals to take care of and control their health issues. Furthermore, the collected health data at the individual level may be shared with doctors and clinicians through cloud computing for other diagnoses [128].

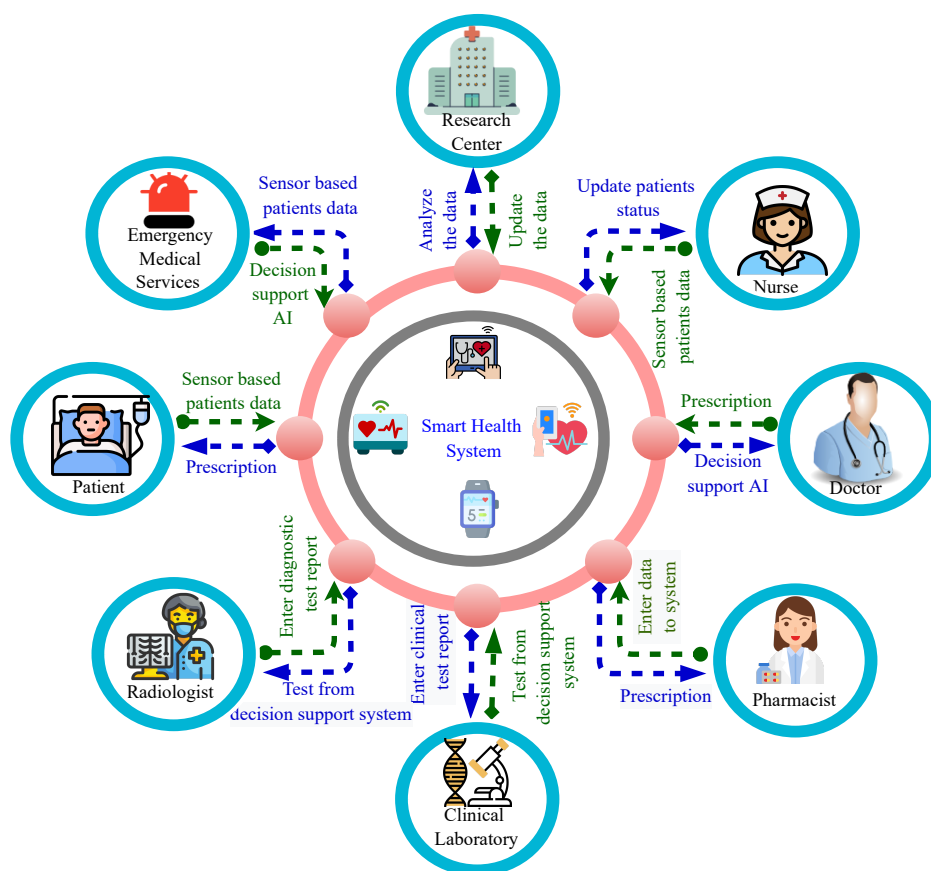


Figure 6. Smart Health Care System in Smart Cities.

Moreover, in health screening, AI may be used for the disease's early diagnosis and the selection of the treatment plan [129]. Ref. [130] presented the ethical issues of trust in the operation of block boxes in AI systems and transparency related to AI frameworks and

applications. The techniques of AI used in explaining the AI framework and the predictions of AI are known as XAI techniques [129]. Ref. [129] proposed the XAI techniques in the existing AI techniques to increase predictions in AI-based models in healthcare. Following are some of the benefits of the XAI in healthcare:

- **Result Tracing:** The factors that may affect the AI-based system's outcome can be traced with the help of explanations generated by XAI techniques.
- **Increased Transparency:** XAI can be used to improve trust levels and help increase AI system transparency; it also explains the arrival of the AI system at the specific decision.
- **Model Improvement:** For making an outcome AI framework learned from the provided dataset. The erroneous predictions may be because of the faulty rules of AI. To improve the model accuracy and identify the errors in the learning process, explanations generated by XAI are useful.

#### 4.2. XAI for Intrusion Detection Systems

Deep neural networks (DNNs) are becoming popular because of their exact predictions [131]. These types of models are hard to interpret but useful. For instance, through the model of DNN, a self-driving car control requires several thousands of tuning parameters [132]. From the perspective of IDS, it is tough for a network administrator if such models of DNN are implemented to comprehend the cognition provided using ML methods. Therefore, the DNN method is also considered a black box method [133]. The black-box method retains its decision-making process stimulating to construe until finding the best solution. DNN mainly modifies several parameters using a trial-and-error method [134]. To increase the known attack classification accuracy, several studies used ML techniques in the IDS, automated model construction, and recognized anomalous network packets [135]. However, some of these frameworks emphasize output interpretation as an understanding of predicting attacks and how ML techniques conclude.

Attacks can be interpreted easily if clarified in rule forms, such as simple construed rules in the if, Then statement. To impose security policies for recognized attacks, some of the challenges include: linking the IDS with the network, analyzing the network traffic, explaining network traffic, and deliver explanations to the network administrators. The explanation is required to understand (i) the parts of violated security policies, (ii) the target parts of the attack, and (iii) the network feature parts. For the logically construed outputs, crucial pillar models for the processes are required. A decision tree (DT) is the best solution for such tasks. DTs are the most suitable models where the scenario requires explaining the output. DTs do not need any assumptions on data distribution like other supervised ML techniques. In the context of IDS, the interpretative approach is necessary.

#### 4.3. XAI for Transportation

Self-driving cars are expected to reduce deaths and provide greater mobility, but they also question the explainability of AI decisions. Self-driving cars must make immediate decisions to categorize objects in real-time. If a car suddenly behaves abnormally due to misclassification issues, it may result in casualties. The significance of misclassification may be unsafe. It is not a prospect; it has previously happened. Recently, an Uber killed a female in Arizona. This is the first identified accident relating to an entirely autonomous vehicle. Now, the question is, how does this misclassification occur? XAI can answer this question. An XAI-based model can be trained to get an insight into the decision made by the model. It was reported that the car's software searched objects in front of the car but interpreted them as plastic bags or windblown weeds [136]. An explainable system can only simplify the uncertainty of this situation and prevent it from happening. In XAI, transportation is an important application area. The explanation of car behavior has even now begun, [137,138], but there is still a long way to go.

#### 4.4. XAI for Aviation Systems

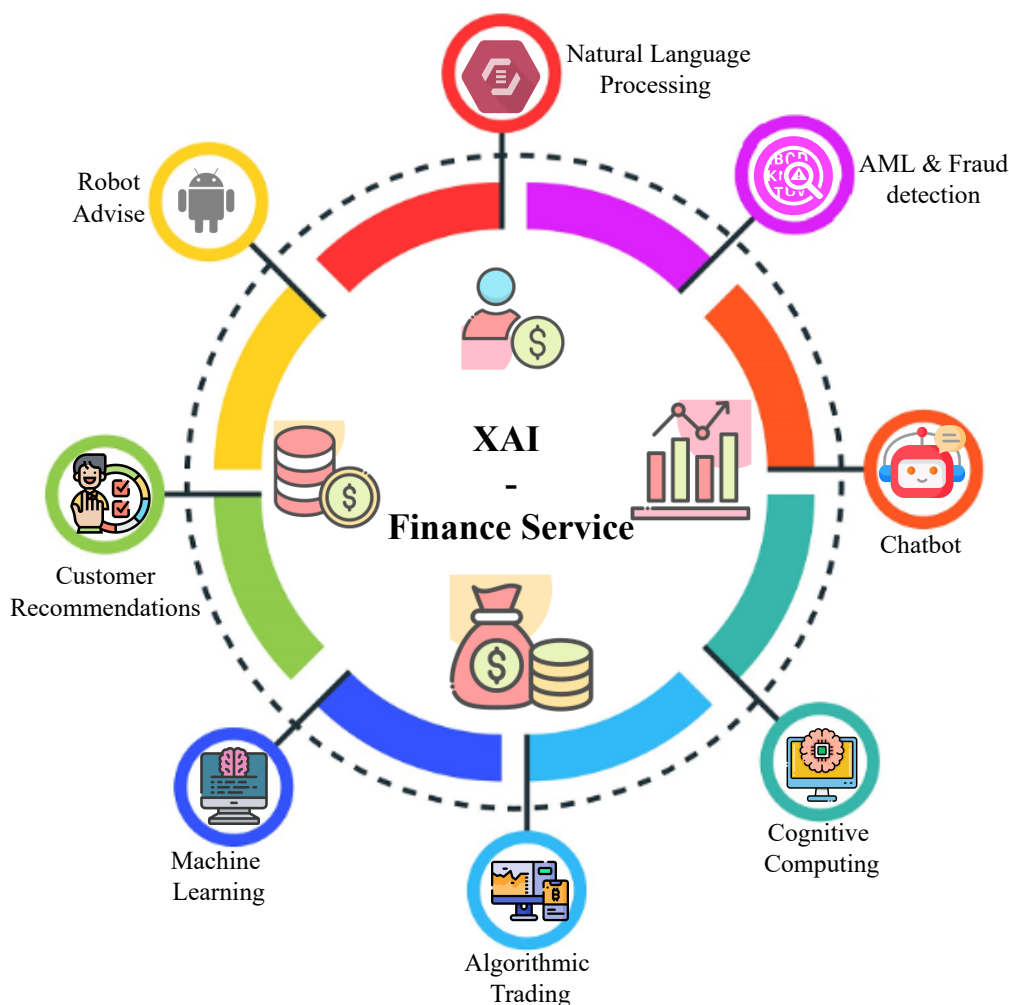
A repair card is used in the aviation maintenance management system to report anomalies or a failure in aviation maintenance. The incorrect maintenance may result from an incomplete or incorrect repair card, making the maintenance data hard to analyze. This incomplete reporting has several different reasons. Firstly, when the crew fills out maintenance of the card, some valuable information is frequently unknown. The findings mentioned on the card are usually written in the form of free-form text, making it hard to interpret the card findings automatically [139]. The failure assessed automatically will describe the repair cards as more consistent and complete. The automatically assessed failure would increase troubleshooting efficiency and add more diagnosis information that may otherwise be unavailable to the manual maintenance crew. Ref. [140] proposed an automatically diagnosed model for aviation failure. The proposed model combines usage and maintenance data and utilizes a data-driven approach. XAI plays an essential role in the aviation industry. The inclusion of XAI will provide interpretability and transparency of the identified and assessed diagnosis. With the help of XAI, the diagnosis is explained by comparing failure with the expected new values of different diagnoses.

#### 4.5. XAI for Legal Issues

In the criminal justice scenario, AI can better evaluate the risk of reoffending and decrease the expenses related to incarceration and crime. However, to predict the risk of reoffending using a criminal decision-making framework in court, it must be ensured that the framework acts reasonably, non-discriminatorily, and honestly. This critical area requires transparency in decision-making methods, but there needs to be more work that can explain automatic decision-making in the legal system [141–143]. For example, in [144], the case questioned the use of proprietary risk assessment software in the decision. The case supposed that the software “created an alternative case to deal with sanctions for inmates: COMPASS” [141] despoiled the right to judge based on race and gender. However, the Judge is still determining the causal verification process and the program used measured trade secrets.

#### 4.6. XAI for Finance

Using AI tools includes asset management activities, improving customer service, and obtaining investment advice in financial services. However, AI tools also raise questions about fair lending and data security. The financial sector is strictly regulated, and the law requires credit institutions to make the right decisions. Therefore, when AI systems are used in rating and loan models, the main challenge is making it more challenging to provide the valid code lenders need. For instance, an explanation of why the borrower was denied credit. Especially when the basis of the negation is the result of the opaque ML algorithm, some credit organizations, such as Experian and Equifax, have carried out valuable research projects to produce unconscious incentive codes to make it easier for auditors to interpret and use XAI-based credit rating decisions [145]. Figure 7 presents the role of XAI in financial services, including natural language processing (NLP), cognitive computing (CC), chatbot, anti-money laundering (AML) & Fraud detection, robot, customer recommendations, algorithmic trading, robot advice and machine learning [146]. XAI is useful in understanding the decisions made by chatbots [147] since chatbots are now being used in every field of life to interact with customers, entertain requests and respond to decisions. Authors in [148] used XAI to explain the role of AI in fraud detection. The authors state that using AI-based systems in finance is risky due to its nature (Black box). Therefore, XAI can be used to explain on what grounds did learning model made this prediction.



**Figure 7.** XAI in Financial Service.

Decision made by the AI model is challenging to understand due to the lack of logical reasoning behind any decision made against a customer recommendation. Authors in [149] used XAI and presented a case study to explain the model decision on customer recommendations using random forest. Furthermore, XAI can be used in cognitive computing and natural language processing. Authors in [150] use AI for cognitive computing for human behavior analysis. Since AI makes decisions in a black box, explaining how ML predicts human behavior is necessary. It is critical to understand what decision has been made and how it was made. Natural language processing uses AI-based models for various purposes, such as opinion mining, sentiment analysis, etc. XAI is useful here to explain these analyses.

#### 4.7. XAI for Digital Forensics

Authors in [151] emphasized the limitations of IT expertise in AI and automation, data detection/retrieval, device classification, network traffic analysis, forensics, encrypted data and programming in specific tools. Describing the applicability of AI in the reconstruction of forensic and multimedia events for different DF examiners (commercial or personal), each field has different practicalities and applicability. For example, AI/ML can help detect child pornography and other suspicious images. DF inspectors can use XAI to classify images, protect their origin and avoid seeing too many objectionable images in these frames. Alternatively, one can view the image/score similarity without displaying the image or reading the default image description (generated by the XAI/ML algorithm).



From the perspective of law enforcement, the use of XAI to support preliminary investigations may be restricted. Traditional ways of criminal forensics are changing dimensions into more technology-oriented investigations [152]. DF police officers emphasized that they were theoretically unjudgmental and did not provide objective reports, but most of DF's analysis showed that the suspect had been arrested and executed. Ref. [152] proposed the forensic analysis method to identify the communicating parties for suspecting communication for later generation and evaluation of the evidence. In this case, the investigator wants the DF expert to recover the data from the targeted devices and submit it to the court in a reportable format so that XAI can perform this operation automatically. Integrating XAI into DF space is a real challenge, but XAI offers the prospect of solving older issues that are increasingly challenging to solve. The increased amount of data essential for analysis, the difficulty of IT crimes and the diversity of evidence slow the transfer of valuable DF information, rely on data processing, and consume many resources.

#### 4.8. XAI for Smart Grids

Smart grids combine advanced measurement, control, and communication technologies to collect large amounts of multidimensional data related to network performance. Due to the rapid changes in advanced power systems, more scattered smart grid components such as distributed energy, electric vehicles, telecommunications infrastructure and smart metering infrastructure are tightly cohesive into advanced power systems [153]. These components generate a large quantity of data to improve and automate smart grid performance and support various applications such as network security [154], system prediction [155], distributed generation control [156] and FD [157]. XAI technologies are gaining widespread attention because traditional computing technologies cannot handle the large amounts of data flowing into smart grid systems. However, the application of AI technology in smart grids is becoming more evident due to the many limitations of traditional technologies for optimization, modeling, and control in data processing [158]. As these XAI technologies use large amounts of data to further improve the performance of smart grids, much research is being done to study these AI technologies and solve these problems.

#### 4.9. XAI for Smart Governance

The continuous development exacerbates the government's need to control the scale and speed of AI and the increasing application of XAI in autonomous, robotics, and deadly weapons systems. There are many research works on changing aspects of XAI, but the management of XAI is underdeveloped. While new AI programs offer opportunities to improve profitability and quality of life, they can also have unintended consequences and new forms of risk that need to be considered [159]. To maximize the benefits of XAI while mitigating risks, governments worldwide need to understand the depth and extent of risk, the regulatory and governance processes to address these challenges, and the need to design the structure. The autonomy of XAI solutions decreases human control over these solutions, creating new problems for other people and human operators in terms of liability and legal damages caused by AI. However, XAI depends heavily on ML to adapt and learn its own rules, leaving humans out of control and only sometimes responsible for intelligent behavior.

#### 4.10. XAI for Smart Industry

Smart manufacturing and intelligent machines in today's industry are being developed with the help of AI, ML and big data exchange technologies [160,161]. Figure 8 presents the role of XAI in Industry 4.0. In the smart Industry, manufacturing comprises data-driven decision creation to improve current manufacturing processes and the latest robots are used to solve complex problems increase [162–164]. With the IoT and other assistive technologies, Industry 4.0 and 5.0 have a long way to go to support these latest trends, but smart factories and smart manufacturing are picking up performance. Following Industry 4.0 and 5.0, smart enterprises offer smarter solutions for predicting machine



failures, safety, machine, and product failures, and suggestions for saving time and cost for the smart work machines need to communicate quickly with other machines in the manufacturing process. By introducing XAI technology into Industry 4.0 and 5.0, our manufacturing industry can be strengthened and form a modern industrial cluster with a resilient and dynamic ecosystem [165].

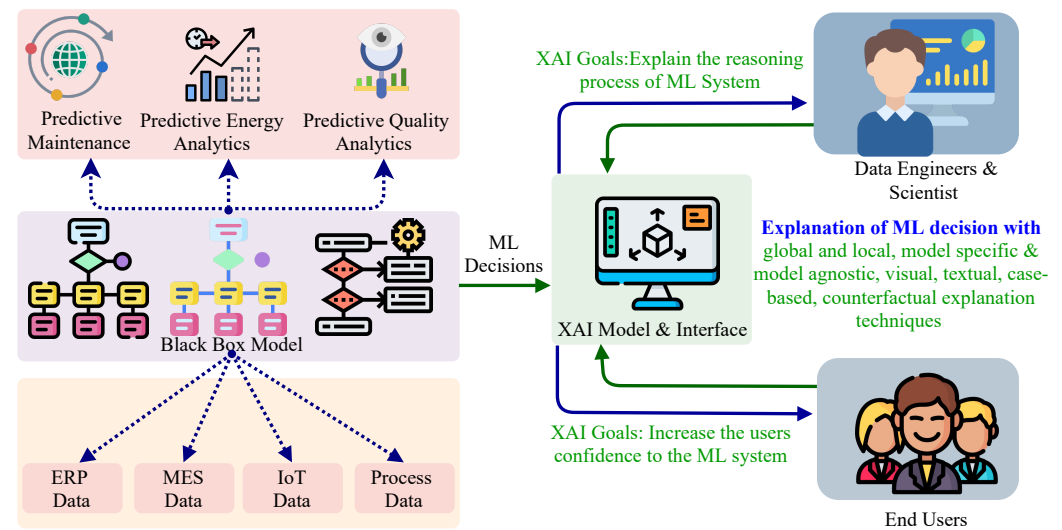


Figure 8. Role of XAI in Industry 4.0.

Table 3 presented the literature review of the XAI applications in smart cities in a tabular form.

Table 3. XAI Application Summary of related works.

Ref	Description	Application										Technical Aspect						
		Smart Transportation	Smart Industry	Smart Healthcare	Smart Governance	Aviation System	Intrusion Detection	Smart Grids	UAVs and drones	IoT	Scalability	Dynamacity	Recursion	Adaptivity	Security	Privacy	Communication	Resource Management
[102]	Proposed network slices for identified V2X use cases: autonomous driving, teleoperated driving, vehicle infotainment, remote diagnostic and management.	✓																✓
[128]	The author described the status of the smart healthcare system in different fields and identified the key technologies that can be incorporated into smart healthcare. The author also identified different research problems in smart healthcare and, in the end, proposed solutions for these problems.			✓				✓						✓		✓		

Table 3. Cont.

Ref	Description	Application									Technical Aspect						
		Smart Transportation	Smart Industry	Smart Healthcare	Smart Governance	Aviation System	Intrusion Detection	Smart Grids	UAVs and drones	IoT	Scalability	Dynamacity	Recursion	Adaptivity	Security	Privacy	Communication
[130]	Using XAI, the author proposed a novel approach to achieve model improvement, accountability, result tracking and transparency in the healthcare domain.			✓						✓	✓	✓					
[129]	The author describes interpretability and the need for interpretability in smart health care.				✓		✓									✓	
[137]	For controlling self-driving vehicles, the author proposed an attention branch network. The proposed network visually presented and analyzed self-driving decision-making through an attention map.	✓									✓	✓					
[138]	The author introduced and defined the importance of timing and explanations in AVs to promote trust.	✓				✓					✓						
[165]	The author provided a detailed overview of several aspects of Big Data and AI in Industry 4.0, particularly focused on specific technologies, applications, concepts, techniques, research perspectives and challenges in deploying Industry 5.0.		✓							✓	✓						
[159]	The author summarizes AI concepts and presents why AI Governance is gaining attention to solve research challenges in different fields.				✓						✓						
[139]	The author developed a prototype tool, Logic and Explained Process of AI Decisions, to validate and verify aviation systems based on AI.	✓				✓		✓					✓				✓

Table 3. Cont.

Ref	Description	Application								Technical Aspect							
		Smart Transportation	Smart Industry	Smart Healthcare	Smart Governance	Aviation System	Intrusion Detection	Smart Grids	UAV's and drones	IoT	Scalability	Dynamacity	Recursion	Adaptivity	Security	Privacy	Communication
[140]	The author developed a novel model to automatically diagnose an aviation system’s failure. The developed model used the data-driven technique with usage data and maintenance data.					✓		✓	✓			✓	✓				
[158]	The research survey reviewed common AI approaches used in security problems, load forecasting, faults detection, power grid stability assessment in the power systems and smart grid.		✓					✓					✓	✓			✓
[154]	The study aims to design and develop a scalable IDS for smart grids. The developed model will differentiate between a disturbance and an actual cyber attack.		✓										✓	✓	✓		✓
[1]	The study surveys future smart cities, their requirements, emerging technologies, Applications, challenges and future aspects.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓
[166]	The study focused on developing human-centered smart cities. They also focused on critical analysis of smart city security, data management, and ethical challenges.	✓	✓	✓	✓	✓			✓	✓	✓	✓			✓	✓	✓
[167]	The study focused on AI and XAI-based methods adopted in the Industry 4.0 scenarios.	✓	✓	✓	✓					✓				✓			✓

## 5. Open Issues and Research Challenges

Machine learning's decision explainability has been one of the important research areas in recent years in smart cities. According to some XAI researchers, the understandability of models for reliability is essential. Meanwhile, some believe that explainability is not required in reality for every purpose because human reasoning is considered a black box [168]. For example, in the daily life activities of smart cities, someone would be more interested in why a loan was approved than a doctor who prefers a specific medical treatment. A recent real-life application where explainability is a crucial need is as detection of COVID patients [169]. Either the individual is COVID-positive or negative. Since it is a crucial decision to be made by the model, an explanation is required on which parameter this model predicts the patient COVID positive or negative. Despite the disagreement, both sides of XAI agree that understandable models are critical in regulated areas. It is

possible to achieve global model behavior using Bayesian rule learning (BRL) or a simple decision tree. There are no variations in approach for the rule or the tree itself. Decision trees must have a high level of understanding to make lay users understand the reason for a prediction. Decision trees identify relationships when the model learns only a few significant features. A decision tree maintains a preferred situation requiring little training. As a result, both rule and decision trees have reproducibility properties and can be referred to as human-simulatable [170].

A wide selection of training models can allow the users to guess correctly. For “lay users,” the explanation generation methods such as Partial Dependence Plots (PDPs) do not possess significant importance. Meanwhile, model evaluation based on features can be roughly approximated when knowledge of relevant techniques is absent among the users. Therefore, LIME and SHAP were applied regarding local surrogate explanations. Both of them follow a similar approach to explanation, and there is no real difference among users except for the actual graphical representation. This results in marginal differences among the values for single features.

The users likely focus on the most significant contributions and general direction, so the underlying methods are unimportant. The literature mentions feature reduction methods when explainability comes to light, in particular, [171]. However, in explainability, one needs to be careful with dimension reduction approaches as they can conceal explainability. The importance of preprocessing data can be accepted due to the more understandable model that may result. Whether the aim is to achieve accuracy or interpretability for the model, considering the quality of the input data for the model is always a good practice. As various explainability approaches are illustrated, it becomes evident that the intrinsic quality of explanations also tends to experience an influence from what was done by [172]. Numerous explanations consider feature importance, giving statisticians an intuitive sense, but a “lay user” would unlikely comprehend the process. This is one of many problems. Many other things need to be done in the future.

### *5.1. Black-Box versus Interpretable Models—Reconsideration of the Problem from the Beginning*

Before training models to address the smart cities problem, one should address what is being sought and whether there is a real need for a black-box model. According to [168], it is a problem that the field of XAI is exploding with research into prediction models for precise decisions. They argued that the accuracy of black-box models is worth considering in building interpretable models. Hence, there is a significant gap (i.e., lack of understandability, comprehensibility, explainability and transparency) in the literature regarding competing black boxes and an interpretable model. It is worth exploring if there would still be a need to use interpretable models when proven to be as reliable as a black box.

### *5.2. Comparison and Measuring of Explainability*

This research establishes that standard measurement, quantification, and comparison procedures for enhancing approaches to explainability need to be developed to make smart cities reliable. Such procedures can then enable researchers to compare different approaches. However, despite some research work, there remains a massive gap (i.e., trustworthiness) in the standardized procedure to be covered in future studies. Evaluation measures such as accuracy, recall and F1-score are required to evaluate the classifier’s performance. It is important to emphasize that there is a need for similar metrics to evaluate explainability. According to authors in [173], XAI can be evaluated based on the four measures. The first difference between the explanation’s logic and the agent’s actual performance is the number of rules resulting from the explanation, the number of features used to generate that explanation and the stability of the explanation [173].

### *5.3. Enhancement of Explanations with Ontologies*

Combining explanations with ontologies presents another area of exploration for future research. Enhancement of explanations with ontologies is a challenge for future

studies that can work on practical use cases to expand the current XAI research. Additionally, various case studies can be presented to investigate the advantages (i.e., reasoning) associated with the combination of ontologies [174].

#### 5.4. *Trusting Machine Learning Models*

Several issues and questions still need to be addressed despite extensive exploration. What if trust and explainability are missing while measuring both within a model? Are there potential smart city scenarios for raising trust in the model? Can more trust be offered just in time in a real-time smart city environment? There is a significant gap in this field, especially in user-centered experiments, that future researchers can fill. A quantitative measure exists in the literature for measuring trust in ML decisions [175]. Two methods were studied in the experiment of [176]. These methods are LIME, a black-box method, and COVAR, a glass-box method. According to their findings, COVAR gave more interpretable explanations than the others. The simple method's usefulness was highlighted in this regard. As per [35], to pursue meaningful progress in this area, exploring what explainability is, in reality, is vital for addressing the root of the problem.

#### 5.5. *User Training on Explainability Features Explicitly*

The literature on algorithmic solving of explainability problems is a challenge for stakeholders in smart cities. There remains a gap in giving attention to user-based studies due to their relevance to real-world problems. Though few studies exist on user experiments, there remains room for more experiments to give the topic holistic coverage addressing all aspects by determining the differences between the participant's prediction and prediction of the model; [177] measured trust in this regard. For example, in smart cities, housing prices can be predicted [178]. Similarly, the study was conducted for users trusting the prediction model in [179]. Authors in [180] evaluated three types of explanation approaches based on user trust using a within-subject design. Ref. [181] measured the frequency of revising predictions to match them with the self-reported trust levels and model predictions. Similarly, a user study can be found in [182] research where human-interpretable explanations were studied as properties of explanations were varied systematically.

The impact of these variations was seen in various tasks' performance. These tasks include simulation of the response from the system, verification of the suggested response and counterfactual queries showing low accuracy. The difference between persuasive and descriptive explanation generation tasks was studied by [183]. While persuasive explanations added user preferences, cognitive functions and expertise, the other types only described explanations within features created by an explainable approach.

#### 5.6. *Visualization*

After a complete XAI-based system overview, the subsequent issue is how to present the system's decision to the end-users and proper explainability, which is still a big challenge for researchers. Most end-users in smart cities are a layman. They need simple explanations of all smart city decisions. Therefore, visualization is one more problem for those considering implementing XAI systems. Research communities must balance the transparency and secrecy that target user groups need to maintain their users' trust and the security of their confidential data by using privacy preservation techniques such as federated learning.

#### 5.7. *Security and Assurance*

With the advancement of AI technology, the protection of smart cities-based autonomous systems has become a significant challenge since a cyberattack can disrupt them. An important application of smart cities is the need to drive self-driving cars worldwide, whether commercial or military, to ensure self-driving. The idea of using XAI to ensure vehicle operational autonomy while monitoring system security is an important area of research. An XAI-based system monitoring framework is required to ensure the natural

autonomy of autonomous systems and the security status of computer and sensor data. Extracting information from the systems used to explain navigation autonomy is also an important future area for research. The newly developed security monitor should assist autonomous navigation in decision-making during an attack. The new framework should enable an XAI-Monitor to explain its capability to find unproductive or false decisions generated by autonomous systems and find cybersecurity barriers that lead to unexpected navigation [184].

### 5.8. Privacy

Nowadays, privacy is a significant concern and a challenge for all stakeholders of smart cities. The following are some examples of how privacy can be violated. The authors of [1] show that smart homes in a smart city environment are a growing market that requires privacy sensors and reliable, interpretable, and explainable control systems. AI's recent advancement has been integrated into all aspects of human life, including the home, and is becoming increasingly resilient. This rise in AI has sparked mutual concern about the impact of confidentiality and resilience in adopting various sensors, such as home cameras. AI explainability and transparency are becoming more responsive as a result. Despite their high accuracy, current black-box AI models could be more trustworthy due to ambiguous solutions. XAI requires a short-range radar sensor with privacy protection and an intuitive AI system to solve privacy issues and organize gestures indoors in a smart home environment. Many traditional encryption concepts and products can be modified and reused to ensure the confidentiality and security of scenarios such as smart homes [185]. XAI is needed to ensure that the AI system adheres to privacy policies defined by organizations such as EU GDPR, etc.

## 6. Ongoing XAI Projects for Smart Cities

This section presents some of the key research projects in the context of XAI and their relevance to smart city applications and technologies.

### 6.1. European Union (EU) Projects

EU has several funding programs and projects in the context of XAI.

#### SPATIAL

Security and Privacy Accountable Technology Innovations, Algorithms and Machine Learning (SPATIAL) [186] is an EU-funded project under the Horizon 2020 (H2020) funding framework. The SPATIAL project focuses on developing accountable, resilient, and trustworthy AI-based security and privacy-preserving methods for future ICT systems. Thus, the SPATIAL project focuses on using XAI to ensure the security and privacy of the 5G and 6G networks. Several beyond 5G and 6G use cases are considered in this project.

#### XAI

The Explanation of AI for decision making (XAI) [187] project is an EU-funded project under the H2020 European Research Council (ERC) funding framework. XAI project aims to construct useful explanations of opaque AI/ML systems using a local-to-global framework to offer a black-box explanation. Moreover, the project will focus on developing an explanation infrastructure that can benchmark AI methods' explainability. The XAI project considers use cases such as health and fraud detection to introduce an explanation-by-design approach.

#### NL4XAI

Interactive Natural Language Technology for XAI (NL4XAI) [188] is an EU-funded project under the H2020 Marie Skłodowska-Curie funding European Training Networks (ETN) framework. NL4XAI makes AI self-explanatory by utilizing natural language generation and processing, argumentation technology, and interactive technology for XAI



systems. Moreover, the project focuses on training 11 Early-Stage Researchers (ESRs) in the domain of XAI.

### **XMANAI**

Explainable Manufacturing Artificial Intelligence (XMANAI) [189] is an EU-funded project under the H2020 funding framework. XMANAI applies XAI for manufacturing by promoting the concept that “our AI is only as good as we are.” To demonstrate the benefits of using xAI for manufacturing, XMANAI considers four real-life pilot use cases at CNH Industrial, Ford, UNIMETRIK, and Whirlpool plants.

### **DEEPCUBE**

XAI Pipelines for Big Copernicus Data (DEEPCUBE) [190] is a EU funded project under the H2020 funding framework. European Union’s Earth observation program manages Copernicus Space Programme [191], which observes the earth and its environment. It collects a great amount of data and DEEPCUBE aims to analyze the big Copernicus data by utilizing AI technologies efficiently.

### **AI4EU**

In 2019, the AI4EU [192] project was started to build the first European AI on-demand platform to share AI resources developed by EU-funded projects. AI4EU is also an EU-funded project under the H2020 funding framework. AI4EU focuses on five main interconnected AI domains, i.e., XAI, Collaborative AI, Physical AI, Integrative AI and Verifiable AI. Moreover, AI4EU is developing a comprehensive strategic AI research innovation agenda for Europe.

Research activities in five key interconnected AI scientific areas (XAI, Physical AI, Verifiable AI, Collaborative AI, Integrative AI), which arise from the application of AI in real-world scenarios;

### **STAR-AI**

Safe and Trusted Human Centric Artificial Intelligence in Future Manufacturing Lines (STAR) [193] project is an EU-funded project under the H2020 funding framework. It focuses on implementing secure, safe, reliable, and trusted human-centric AI systems in manufacturing environments. In the STAR project, XAI boosts the transparency of manufacturing-related AI systems to improve user trust in AI systems. In this regard, the STAR-AI project focuses on three main use cases, i.e., human–robot collaboration for robust quality inspection, human-centered AI for agile manufacturing 4.0 and human behavior prediction, and safe zone detection for routing.

### **FeatureCloud**

FeatureCloud project [194] aims to integrate security-by-design and privacy-by-architecture concepts on medical health systems to reduce the possibility of cybercrime and facilitate safe cross-border collaborative data-mining efforts. It is got funded project under the H2020 funding framework. FeatureCloud integrates blockchain and federated learning techniques to eliminate sharing sensitive mining data via any communication channels and centralized data storage. Moreover, the FeatureCloud project incorporates “supervised machine learning” with XAI to improve compatibility with legal considerations and international policies.

### **GECKO**

Building Greener and more sustainable societies by filling the Knowledge gap in social science and engineering to enable responsible artificial intelligence co-creation (GECKO) [195] is an EU-funded project under the H2020 Marie Skłodowska-Curie funding European Training Networks (ETN) framework. GECKO contributes to developing Accountable, Responsible, and Transparent AI (ART AI) designs considering technological,

ethical, and social science angles to support the European green deal ambition. The GECKO project focuses on training 15 Early-Stage Researchers (ESRs) who will explore interpretability and XAI models to mitigate unintentionally harmful and poorly designed AI models.

#### 6.2. Defense Advanced Research Projects Agency (DARPA) Projects

In 2016, Defense Advanced Research Projects Agency (DARPA) Information Innovation Office (I2O) launched the XAI program (Explainable Artificial Intelligence (XAI) <https://www.darpa.mil/program/explainable-artificial-intelligence>). Under this program, DARPA has funded several projects focusing on different aspects of XAI.

##### 6.2.1. Driving-X

This project is conducted by the University of California, Berkeley (UCB) and the University of Amsterdam. This project focuses on using XAI for self-driving vehicles to improve user acceptance.

##### 6.2.2. Retrieval: Saliency Driven Retrieval

This project is conducted by the University of California, Berkeley (UCB). This project focuses on using XAI to improve the accuracy of exemplar-based image retrieval.

##### 6.2.3. RISE: High-Fidelity Saliency for Black-Box Networks

This project is conducted by the University of California, Berkeley (UCB) and Boston University. This project focuses on high-fidelity saliency maps to indicate each pixel level's salient for the model's prediction, mainly for black-box models.

##### 6.2.4. Rollouts: Informative Rollouts and the Critical States

This project is conducted by the University of California, Berkeley (UCB). This project focuses on explaining a robot's policy and how that robot acts in different scenarios to offer safe and comfortable human–robot interaction.

##### 6.2.5. SNMN: Transparent Multi-Step Reasoning with Stack Neural Module Networks

This project is conducted by the University of California, Berkeley (UCB) and Boston University. This project focuses on reasoning procedures in the UCB-developed SNMN model-generated explanations. It studies how this SNMN model improves users' satisfaction with the explanation and improves truthful beliefs about the model's behavior.

##### 6.2.6. StarCraft: Heterogeneous Hierarchical Policies

This project is conducted by the University of California, Berkeley (UCB) and Boston University. This project focuses on developing an AI model to play the StarCraft II game. This AI model should be capable of automatically explaining its behavior.

##### 6.2.7. Textual: Textual Justification with Grounding

This project is run by the University of California, Berkeley (UCB) and the University of Amsterdam. This project combines explanation language generation and natural language grounding models to provide textual output to explain the image classification.

##### 6.2.8. CAMEL: Causal Models to Explain Learning

This project (CAMEL <https://cra.com/projects/camel/>) is run by Charles River Analytics, the University of Massachusetts and Brown University. And develops Causal Models to Explain Learning (CAMEL) which offer humans understandable and trusted explanations based on causality to manage machine learning (ML) techniques of AI systems.

#### 6.2.9. Learning and Communicating Explainable Representations for Analytics and Autonomy

This project is run by the University of California, Los Angeles (UCLA), Oregon State University and Michigan State University and develops an XAI framework for multi-model analytics and autonomous recognition, reasoning, and planning domains.

#### 6.2.10. xACT: Explanation-Informed Acceptance Testing of Deep Adaptive Programs

This project is run by Oregon State University (OSU). This project develops a new paradigm of explanation-informed acceptance testing (xACT) tools that can observe and evaluate machine-learned systems' behavior and explain the decisions leading to that behavior.

#### 6.2.11. Common Ground Learning and Explanation (COGLE)

This project (COGLE [https://www.markstefik.com/?page\\_id=2262](https://www.markstefik.com/?page_id=2262)) is run by Palo Alto Research Center, Inc. (PARC), Carnegie Mellon University, Florida Institute for Human and Machine Cognition and the United States Military Academy. This project develops an interactive sense-making system called COGLE (COMmon Ground Learning and EXplanation). The COGLE system can explain the learned performance capabilities of an autonomous system.

#### 6.2.12. XAI for Assisting Data Scientists

This project is run by Carnegie Mellon University (CMU) and experimentally analyzes the effectiveness of the XAI system in debugging standard ML models.

#### 6.2.13. Deep Attentional Representations for Explanations (DARE)

This project is run by SRI International, the University of Toronto and the University of California, San Diego, and focuses on augmenting several explainable deep learning models to enable multiple modes of explanation to improve accuracy.

#### 6.2.14. Explainable Question Answering System (EQUAS)

This project (<https://www.raytheonintelligenceandspace.com/news/feature/trust-machine>) is run by Raytheon BBN Technologies (Raytheon BBN), Georgia Tech Research Corporation, the University of Texas Austin and Massachusetts Institute of Technology. This project develops a new Explainable Question Answering System (EQUAS) based on pedagogical and argumentation theories. It can analyze the important explanation elements, the behavior of explanation space, and user expectations. The above two theories help define the foundation of EQUAS "explanation space," which provides analytics, visualization, cases and rejected alternatives.

#### 6.2.15. Tractable Probabilistic Logic Models: A New Deep Explainable Representation

This project (<https://www.eng.ufl.edu/ai-university/research/tractable-probabilistic-logic-models-a-new-deep-explainable-representation/>) is conducted by The University of Texas at Dallas (UTD), University of California, Los Angeles, the University of Florida and Indian Institute of Technology Delhi. It studies new explainable systems for fake news detection and videos' automatic action and object detection. These systems can explain the reasons for produced results.

#### 6.2.16. Transforming Deep Learning to Harness the Interpretability of Shallow Models: An Interactive End-to-End System

This project is conducted by Texas A&M University (TAMU) and the University of Florida and develops and analyzes an explainable system named XFake, on fake news detection. This is an XFake system with three different frameworks, i.e., MIMIC, ATTN, and PERT, which can analyze the news from different perspectives.

#### 6.2.17. Model Explanation by Optimal Selection of Teaching Examples

This project is conducted by Rutgers University and analyzes the Explanation-by-Examples system to improve the user understanding of the inference of black-box machine learning models.

#### 6.3. Other Projects

In South Korea, the XAI Center [196] is focusing on developing XAI models by developing new or modified ML techniques. This XAI center aims to support research projects that use XAI to offer sufficient human-understandable explanations about the results and decision-making ML model. In this regard, the XAI center primarily focuses on to Medical and Financial industry, where adaptation of AI is a high risk without explainability. Moreover, the Austrian Science Fund has funded a project titled “Reference Model of XAI for the Medical Domain” (<https://www.aholzinger.at/austrian-science-fund-project-on-explainability-granted/>).

The Christ-Era organization had a funding call on Explainable Machine Learning-based Artificial Intelligence (XAI) (<https://www.chistera.eu/projects-call-2019>). It has funded 12 projects in different domains, including digital medicine, robotics, and manufacturing.

### 7. Lessons Learned, Future Development, and Direction of XAI in Smart Cities

This section discusses the lessons learned from surveying the state-of-the-art in XAI for smart cities. Based on the detailed analysis, it synthesizes future research directions to enable XAI for smart cities.

#### 7.1. Security Issues of XAI for Smart Cities

From the above discussion, it can be found that XAI has the potential to revolutionize smart cities by rendering justification/explanation to the decision-makers/governments, through which better decisions can be made. However, due to the sensitive nature of the data generated from smart cities, the deployment of XAI-integrated smart cities applications will face critical security issues, such as authentication and authorization, integrity,  $24 \times 7$  availability, and monitoring and audit of explanatory and interpretable processes [197]. The security challenges faced by the applications of XAI-enabled smart cities, along with the possible solutions, are discussed below:

##### 7.1.1. Authentication and Authorization

Several stakeholders, such as intelligent agents, sensing nodes, IoT sensor nodes, and machines involved in smart cities, should be authenticated and authorized to ensure that they do not misuse or expose the AI model to malicious users. If the details of the training of AI models are exposed, malicious users can tamper with the training data to change the decision of AI models. To address these issues, blockchain can be used in smart city applications to ensure that the stakeholders participating in smart city applications are authentic [198].

##### 7.1.2. Integrity

Integrity is a significant concern in XAI-integrated smart city applications due to real-time data monitoring and decision-making based on the explainability and interpretability of LIME, SHAP, and Grad-CAM methodologies. Furthermore, emergency notifications, such as making decisions about a patient’s health diseases, are sent over third-party networks that demand a secure and flawless communication channel [199]. Blockchain and federated learning can be used in smart city applications to preserve integrity. Even the combination of these two techniques can be used to preserve integrity [200].

### 7.1.3. 24 × 7 Availability

The availability and access of real-time and near-real-time data and XAI-based decision-making of smart city applications such as targeted healthcare, connected restaurants, connected cars and smart manufacturing are paramount. Furthermore, performing pervasive computing and providing robust real-time decisions in smart city applications is challenging and demands high-performance computing-intensive applications. To provide 24 × 7 availability, devices with long battery life and a lightweight machine learning model may be introduced. Furthermore, dead devices or sensor nodes can be identified by their readings using machine learning to keep the system up and running [1].

### 7.1.4. Monitoring and Audit

The XAI integrated smart cities applications demand regular monitoring and maintenance of all the smart city stakeholders using audit-log management and preventive maintenance mechanisms [197]. To address this issue, AI-based auditing can be performed where data is acquired from smart devices. Two types of auditing/forensics can be performed: software-based and hardware-based [201]. Data is acquired in two modes in software-based auditing: root and no root mode. In rooting mode, most of the data can be acquired. However, in non-rooting mode, data loss can occur due to required permissions. Various approaches, such as “JTGA” and “chip-off,” can be used in the hardware-based acquisition. However, it can result in permanent device damage. After data acquisition, AI is applied to classify or predict any output labels [201].

## 7.2. Ethical and Privacy Issues of XAI for Smart Cities

All the applications of smart cities depend on various AI algorithms and methodologies such as reinforcement learning, collaborative and networking nodes, confidential data and logs [202]. So it is essential to maintain the confidentiality of various smart city stakeholders. Therefore, an XAI-integrated blockchain framework must ensure stakeholders' data and controlled information exchange over a third-party network. The XAI-integrated smart cities applications overcome the disadvantages of black-box AI systems, such as lack of transparency and trust; however, the XAI-integrated smart cities applications need solutions for issues such as unreasonableness and injustice.

### Unreasonableness and Injustice

The XAI-integrated smart city applications can make decisions based on data availability. However, as the decisions are based on AI systems, it is challenging to confirm that they are fair and non-biased [203]. As a solution, XAI can ensure that the AI system provides enough explanation about how a model made a decision and can be evaluated using the difference between the logic and the models' actual performance, the number of rules resulting from the explanation, the number of features used to generate that explanation and the stability of the explanation. The decisions made using XAI integrated smart cities applications and systems should be evaluated based on ethical and moral behaviors [68].

### 7.3. Scalability

Scalability is a cardinal feature in assessing any system's performance and throughput [202]. The dynamic XAI integrated smart city system functions are based on machines, various AI algorithms and methodologies, sensing data, and third-party networks. It is important to ensure flexibility and responsiveness among various collaborative and networking nodes and AI methodologies. In the future, XAI architecture can be combined with Responsive AI to achieve scalability in smart city applications and systems [204].

### 7.4. Regulatory Compliance for Implementing XAI for Smart Cities

Regulatory compliance means formulating guidelines, regulations and laws [205]. Formulating appropriate rules and regulations is essential to avoid misuse of the latest technologies, such as XAI. Furthermore, smart city applications and systems blend the latest



technologies, such as Blockchain, computer vision, IoT, Big data, etc. Such diverse applications and systems need to formulate legal policies and compliance regulations to protect smart city applications and systems from fraud, misuse, malfunctioning, and manipulation.

#### 7.5. Standard Specifications for Implementing XAI for Smart Cities

Standards specifications are documents that describe a set of rules and conditions. Therefore, publishing standard specifications and documents for using technologies such as XAI is important. Setting high standards for the usage of XAI for smart cities is crucial in a manner that will lessen vagueness and increase robustness and could assist in enhancing living/working understanding, advanced transportation [206,207], and quicker availability of enough information for informed decision making [208,209]. Furthermore, it is also essential to make the stakeholders of smart cities aware of various standards and policies for implementing XAI for smart cities applications and systems.

### 8. Conclusions

This paper comprehensively surveyed recent and future developments in XAI for smart cities. This paper envisaged the societal, industrial, and technological trends considering the 2030 vision. Previous studies have given more focus to XAI applications for smart cities. This study discussed the concept of XAI for smart cities: various XAI technology use cases, challenges, applications, and possible alternative solutions. The paper discussed concise definitions and key technologies of XAI and its explanation using various smart city-driven use cases to assist readers. The paper also highlighted ongoing smart city-driven XAI projects, standardized practices, applications, alternative solutions, and future developments focusing on developing XAI for smart cities. In the end, the paper discussed security, privacy, and ethical issues and regular compliance for implementing XAI in smart cities and provided a roadmap for future research directions. In the future, It is intended to extend this research to the combination of XAI and recent security frameworks for future smart cities.

**Author Contributions:** Conceptualization, A.R.J.; methodology, A.R.J. and W.A.; validation, P.K.R.M. and S.P.; formal analysis, A.R.J. and T.R.G.; writing—original draft preparation, A.R.J., T.R.G., W.A.; writing—review and editing, M.A., S.P. and P.K.R.M.; supervision, W.A. and M.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** No external funding is required for this study.

**Conflicts of Interest:** The authors share no conflict of interest.

### References

1. Javed, A.R.; Shahzad, F.; ur Rehman, S.; Zikria, Y.B.; Razzak, I.; Jalil, Z.; Xu, G. Future smart cities requirements, emerging technologies, applications, challenges and future aspects. *Cities* **2022**, *129*, 103794. [\[CrossRef\]](#)
2. Pandya, S.; Srivastava, G.; Jhaveri, R.; Babu, S. P.; Bhattacharya, S.; Maddikunta, P.K.R.; Mastorakis, S.; Piran, M.J.; Gadekallu, T.R. Federated learning for smart cities: A comprehensive survey. *Sustain. Energy Technol. Assess.* **2023**, *55*, 102987. [\[CrossRef\]](#)
3. Javed, A.R.; Faheem, R.; Asim, M.; Baker, T.; Beg, M.O. A smartphone sensors-based personalized human activity recognition system for sustainable smart cities. *Sustain. Cities Soc.* **2021**, *71*, 102970. [\[CrossRef\]](#)
4. Javed, A.R.; Fahad, L.G.; Farhan, A.A.; Abbas, S.; Srivastava, G.; Parizi, R.M.; Khan, M.S. Automated cognitive health assessment in smart homes using machine learning. *Sustain. Cities Soc.* **2021**, *65*, 102572. [\[CrossRef\]](#)
5. Sajid, F.; Javed, A.R.; Basharat, A.; Kryvinska, N.; Afzal, A.; Rizwan, M. An Efficient Deep Learning Framework for Distracted Driver Detection. *IEEE Access* **2021**, *9*, 169270–169280. [\[CrossRef\]](#)
6. Shabbir, A.; Shabir, M.; Javed, A.R.; Chakraborty, C.; Rizwan, M. Suspicious transaction detection in banking cyber-physical systems. *Comput. Electr. Eng.* **2022**, *97*, 107596. [\[CrossRef\]](#)



7. Fayyaz, M.; Farhan, A.A.; Javed, A.R. Thermal Comfort Model for HVAC Buildings Using Machine Learning. *Arab. J. Sci. Eng.* **2022**, *47*, 2045–2060. [[CrossRef](#)]
8. Osherooff, J.A.; Teich, W. A.; Middleton, B.; Steen, E.B.; Wright, A.; Detmer, D.E. A roadmap for national action on clinical decision support. *J. Am. Med. Inform. Assoc.* **2007**, *14*, 141–145. [[CrossRef](#)]
9. Sutton, R.T.; Pincock, D.; Baumgart, D.C.; Sadowski, D.C.; Fedorak, R.N.; Kroeker, K.I. An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digit. Med.* **2020**, *3*, 1–10. [[CrossRef](#)]
10. Obermeyer, Z.; Emanuel, E.J. Predicting the future—Big data, machine learning and clinical medicine. *N. Engl. J. Med.* **2016**, *375*, 1216. [[CrossRef](#)]
11. Vucenovic, A.; Ali-Ozkan, O.; Ekwempe, C.; Eren, O. Explainable AI in Decision Support Systems: A Case Study: Predicting Hospital Readmission within 30 Days of Discharge. In Proceedings of the 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), London, ON, Canada, 30 August–2 September 2020; pp. 1–4.
12. Yenduri, G.; Gadekallu, T.R. XAI for Maintainability Prediction of Software-Defined Networks. In Proceedings of the 24th International Conference on Distributed Computing and Networking, Kharagpur, India, 4–7 January 2023; pp. 402–406.
13. Thakker, D.; Mishra, B.K.; Abdullatif, A.; Mazumdar, S.; Simpson, S. Explainable Artificial Intelligence for Developing Smart Cities Solutions. *Smart Cities* **2020**, *3*, 1353–1382. [[CrossRef](#)]
14. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
15. Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4793–4813. [[CrossRef](#)]
16. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Benetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]
17. Vassiliades, A.; Bassiliades, N.; Patkos, T. Argumentation and explainable artificial intelligence: A survey. *Knowl. Eng. Rev.* **2021**, *36*, e5. [[CrossRef](#)]
18. Mohseni, S.; Zarei, N.; Ragan, E.D. A multidisciplinary survey and framework for design and evaluation of explainable AI systems. *ACM Trans. Interact. Intell. Syst. (TiiS)* **2021**, *11*, 1–45. [[CrossRef](#)]
19. Ferreira, J.J.; Monteiro, M.S. What are people doing about XAI user experience? A survey on AI explainability research and practice. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 19–24 July 2020; pp. 56–73.
20. Čyras, K.; Rago, A.; Albini, E.; Baroni, P.; Toni, F. Argumentative XAI: A Survey. *arXiv* **2021**, preprint. arXiv:2105.11266.
21. Pocevičiūtė, M.; Eilertsen, G.; Lundström, C. Survey of XAI in digital pathology. In *Artificial Intelligence and Machine Learning for Digital Pathology*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 56–88.
22. Qian, K.; Danilevsky, M.; Katsis, Y.; Kawas, B.; Oduor, E.; Popa, L.; Li, Y. XNLP: A Living Survey for XAI Research in Natural Language Processing. In Proceedings of the 26th International Conference on Intelligent User Interfaces, College Station, TX, USA, 14–17 April 2021; pp. 78–80.
23. Ehsan, U.; Wintersberger, P.; Liao, Q.V.; Mara, M.; Streit, M.; Wachter, S.; Riener, A.; Riedl, M.O. Operationalizing Human-Centered Perspectives in Explainable AI. In Proceedings of the Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–6.
24. Rai, A. Explainable AI: From black box to glass box. *J. Acad. Mark. Sci.* **2020**, *48*, 137–141. [[CrossRef](#)]
25. Bengio, Y.; Lecun, Y.; Hinton, G. Deep learning for AI. *Commun. ACM* **2021**, *64*, 58–65. [[CrossRef](#)]
26. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News*. 2002. December 2002.
27. Chen, R.C.; Caraka, R.E.; Arnita, N.E.G.; Pomalingo, S.; Rachman, A.; Toharudin, T.; Tai, A.R.J.; Pardamean, B. An end to end of scalable tree boosting system. *Sylwan* **2020**, *165*, 1–11.
28. Livieris, I.E.; Pintelas, E.; Stavroyiannis, S.; Pintelas, P. Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms* **2020**, *13*, 121. [[CrossRef](#)]
29. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A review of machine learning interpretability methods. *Entropy* **2021**, *23*, 18. [[CrossRef](#)]
30. Hase, P.; Bansal, M. Evaluating explainable AI: Which algorithmic explanations help users predict model behavior? *arXiv* **2020**, preprint. arXiv:2005.01831.
31. Montavon, G.; Samek, W.; Müller, K.R. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing* **2018**, *73*, 1–15. [[CrossRef](#)]
32. Gleicher, M. A framework for considering comprehensibility in modeling. *Big Data* **2016**, *4*, 75–88. [[CrossRef](#)]
33. Fernandez, A.; Herrera, F.; Cordon, O.; del Jesus, M.J.; Marcelloni, F. Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for and where to? *IEEE Comput. Intell. Mag.* **2019**, *14*, 69–81. [[CrossRef](#)]
34. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–42. [[CrossRef](#)]
35. Lipton, Z.C. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* **2018**, *16*, 31–57. [[CrossRef](#)]

36. Islam, S.R.; Eberle, W.; Ghafoor, A.R.J.; Ahmed, M. Explainable Artificial Intelligence Approaches: A Survey. *arXiv* **2021**, preprint. arXiv:2101.09429.
37. Sokol, K.; Flach, P. Explainability fact sheets: A framework for systematic assessment of explainable approaches. In Proceedings of the 2020 Conference on Fairness, Accountability and Transparency, Barcelona, Spain, 27–30 January 2020; pp. 56–67.
38. Schneider, J.; Handali, J. Personalized explanation in machine learning: A conceptualization. *arXiv* **2019**, preprint. arXiv:1901.00770.
39. Rudin, C. Please stop explaining black box models for high stakes decisions. *STAT* **2018**, *1050*, 26.
40. Diez-Olivan, A.; Del Ser, J.; Galar, D.; Sierra, B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Inf. Fusion* **2019**, *50*, 92–111. [\[CrossRef\]](#)
41. Guo, J.; Ding, X.; Wu, W. A blockchain-enabled ecosystem for distributed electricity trading in smart city. *IEEE Internet Things J.* **2020**, *8*, 2040–2050. [\[CrossRef\]](#)
42. Li, M.; Tang, H.; Hussein, A.R.; Wang, X. A sidechain-based decentralized authentication scheme via optimized two-way peg protocol for smart community. *IEEE Open J. Commun. Soc.* **2020**, *1*, 282–292. [\[CrossRef\]](#)
43. Bhattacharya, S.; Somayaji, S.R.K.; Gadekallu, T.R.; Alazab, M.; Maddikunta, P.K.R. A review on deep learning for future smart cities. *Internet Technol. Lett.* **2022**, *5*, e187. [\[CrossRef\]](#)
44. Alazab, M.; Lakshmana, K.; Reddy, T.; Pham, Q.V.; Maddikunta, P.K.R. Multi-objective cluster head selection using fitness averaged rider optimization algorithm for IoT networks in smart cities. *Sustain. Energy Technol. Assess.* **2021**, *43*, 100973. [\[CrossRef\]](#)
45. Ghayvat, H.; Awais, M.; Gope, P.; Pandya, S.; Majumdar, S. Recognizing suspect and predicting the spread of contagion based on mobile phone location data (counteract): A system of identifying covid-19 infectious and hazardous sites, detecting disease outbreaks based on the internet of things, edge computing and artificial intelligence. *Sustain. Cities Soc.* **2021**, *69*, 102798.
46. Pandya, S.; Ghayvat, H. Ambient acoustic event assistive framework for identification, detection and recognition of unknown acoustic events of a residence. *Adv. Eng. Inform.* **2021**, *47*, 101238. [\[CrossRef\]](#)
47. Zhang, S.; Li, X.; Zong, M.; Zhu, X.; Wang, R. Efficient kNN classification with different numbers of nearest neighbors. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 1774–1785. [\[CrossRef\]](#)
48. Garg, D.; Goel, P.; Pandya, S.; Ganatra, A.; Kotecha, K. A deep learning approach for face detection using YOLO. In Proceedings of the 2018 IEEE Punecon, Pune, India, 30 November–2 December 2018; pp. 1–4.
49. Sharma, P.K.; Kumar, N.; Park, J.H. Blockchain-based distributed framework for automotive industry in a smart city. *IEEE Trans. Ind. Inform.* **2018**, *15*, 4197–4205. [\[CrossRef\]](#)
50. Liu, J.; Liu, Y.; Zhang, Q. A Weight Initialization Method Based on Neural Network with Asymmetric Activation Function. *Neurocomputing* **2022**, *438*, 171–182. [\[CrossRef\]](#)
51. Gai, K.; Guo, J.; Zhu, L.; Yu, S. Blockchain meets cloud computing: A survey. *IEEE Commun. Surv. Tutorials* **2020**, *22*, 2009–2030. [\[CrossRef\]](#)
52. Rahman, S.A.; Tout, H.; Talhi, C.; Mourad, A. Internet of things intrusion detection: Centralized, on-device, or federated learning? *IEEE Netw.* **2020**, *34*, 310–317. [\[CrossRef\]](#)
53. Khan, R.U.; Zhang, X.; Alazab, M.; Kumar, R. An improved convolutional neural network model for intrusion detection in networks. In Proceedings of the 2019 Cybersecurity and cyberforensics conference (CCC), Melbourne, Australia, 8–9 May 2019; pp. 74–77.
54. Gautam, S.; Henry, A.; Zuhair, M.; Rashid, M.; Javed, A.R.; Maddikunta, P.K.R. A Composite Approach of Intrusion Detection Systems: Hybrid RNN and Correlation-Based Feature Optimization. *Electronics* **2022**, *11*, 3529. [\[CrossRef\]](#)
55. Huang, C.; Wang, Z.; Chen, H.; Hu, Q.; Zhang, Q.; Wang, W.; Guan, X. RepChain: A Reputation-Based Secure, Fast and High Incentive Blockchain System via Sharding. *IEEE Internet Things J.* **2020**, *8*, 4291–4304.
56. Mumtaz, S.; Al-Dulaimi, A.; Gačanin, H.; Bo, A. Block Chain and Big Data-Enabled Intelligent Vehicular Communication. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 3904–3906. [\[CrossRef\]](#)
57. Nassar, M.; Salah, K.; ur Rehman, M.H.; Svetinovic, D. Blockchain for explainable and trustworthy artificial intelligence. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, e1340.
58. Kotter, E.; Marti-Bonmati, L.; Adrian P., Brady Nandita M., Desouza, E.S. ESR white paper: Blockchain and medical imaging. *Insights Into Imaging* **2021**, *12*, 1–7.
59. Calvaresi, D.; Dubovitskaya, A.; Calbimonte, J.P.; Taveter, K.; Schumacher, M. Multi-agent systems and blockchain: Results from a systematic literature review. In Proceedings of the International conference on practical applications of agents and multi-agent systems, Toledo, Spain, 20–22 June 2018; pp. 110–126.
60. Singh, A.R.J.; Rathore, S.; Park, J.H. Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence. *Future Gener. Comput. Syst.* **2020**, *110*, 721–743. [\[CrossRef\]](#)
61. Cirillo, F.; Gómez, D.; Diez, L.; Maestro, I.E.; Gilbert, T.B.J.; Akhavan, R. Smart city IoT services creation through large-scale collaboration. *IEEE Internet Things J.* **2020**, *7*, 5267–5275. [\[CrossRef\]](#)
62. Kimothi, S.; Thapliyal, A.; Singh, R.; Rashid, M.; Gehlot, A.; Akram, S.V.; Javed, A.R. Comprehensive Database Creation for Potential Fish Zones Using IoT and ML with Assimilation of Geospatial Techniques. *Sustainability* **2023**, *15*, 1062. [\[CrossRef\]](#)
63. Kirimtat, A.; Krejcar, O.; Kertesz, A.; Tasgetiren, M.F. Future trends and current state of smart city concepts: A survey. *IEEE Access* **2020**, *8*, 86448–86467. [\[CrossRef\]](#)

64. Kherraf, N.; Alameddine, H.A.; Sharafeddine, S.; Assi, C.M.; Ghrayeb, A. Optimized provisioning of edge computing resources with heterogeneous workload in IoT networks. *IEEE Trans. Netw. Serv. Manag.* **2019**, *16*, 459–474. [\[CrossRef\]](#)
65. Shahzad, F.; Mannan, A.; Javed, A.R.; Almadhor, A.S.; Baker, T.; Al-Jumeily OBE, D. Cloud-based multiclass anomaly detection and categorization using ensemble learning. *J. Cloud Comput.* **2022**, *11*, 1–12. [\[CrossRef\]](#)
66. Kuppusamy, P.; Kumari, N.M.J.; Alghamdi, W.Y.; Alyami, H.; Ramalingam, R.; Javed, A.R.; Rashid, M. Job scheduling problem in fog-cloud-based environment using reinforced social spider optimization. *J. Cloud Comput.* **2022**, *11*, 99. [\[CrossRef\]](#)
67. Panarello, A.; Tapas, N.; Merlino, G.; Longo, F.; Puliafito, A. Blockchain and iot integration: A systematic survey. *Sensors* **2018**, *18*, 2575. [\[CrossRef\]](#)
68. Amann, J.; Blasimme, A.; Vayena, E.; Frey, D.; Madai, V.I. Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 1–9. [\[CrossRef\]](#)
69. García-Magariño, I.; Muttukrishnan, R.; Lloret, J. Human-centric AI for trustworthy IoT systems with explainable multilayer perceptrons. *IEEE Access* **2019**, *7*, 125562–125574. [\[CrossRef\]](#)
70. Hernandez, L.; Baladron, C.; Aguiar, W. A.; Carro, B.; Sanchez-Esguevillas, A.J.; Lloret, J.; Massana, J. A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings. *IEEE Commun. Surv. Tutorials* **2014**, *16*, 1460–1495. [\[CrossRef\]](#)
71. Rehman, S.U.; Javed, A.R.; Khan, M.U.; Nazar Awan, M.; Farukh, A.; Hussien, A. Personalised comfort: A personalised thermal comfort model to predict thermal sensation votes for smart building residents. *Enterp. Inf. Syst.* **2020**, *16*, 1–23. [\[CrossRef\]](#)
72. Soomro, S.; Miraz, M.H.; Prasanth, A.; Abdullah, M. Artificial intelligence enabled IoT: Traffic congestion reduction in smart cities. In Proceedings of the Smart Cities Symposium 2018, Bahrain, 22–23 April 2018.
73. Lavrenovs, A.; Graf, R. Explainable AI for Classifying Devices on the Internet. In Proceedings of the 2021 13th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 25–28 May 2021; pp. 291–308.
74. Celesti, A.; Galletta, A.; Carnevale, L.; Fazio, M.; Lay-Ekuakille, A.; Villari, M. An IoT cloud system for traffic monitoring and vehicular accidents prevention based on mobile sensor data processing. *IEEE Sens. J.* **2017**, *18*, 4795–4802. [\[CrossRef\]](#)
75. Ahamed, N.N.; Karthikeyan, P. A Reinforcement Learning Integrated in Heuristic search method for self-driving vehicle using blockchain in supply chain management. *Int. J. Intell. Netw.* **2020**, *1*, 92–101.
76. Lu, R.; Jin, X.; Zhang, S.; Qiu, M.; Wu, X. A study on big knowledge and its engineering issues. *IEEE Trans. Knowl. Data Eng.* **2018**, *31*, 1630–1644. [\[CrossRef\]](#)
77. Reddy, G.T.; Reddy, M.P.K.; Lakshmana, K.; Kaluri, R.; Rajput, D.S.; Srivastava, G.; Baker, T. Analysis of dimensionality reduction techniques on big data. *IEEE Access* **2020**, *8*, 54776–54788. [\[CrossRef\]](#)
78. Gadekallu, T.R.; Pham, Q.V.; Huynh-The, T.; Bhattacharya, S.; Maddikunta, P.K.R.; Liyanage, M. Federated Learning for Big Data: A Survey on Opportunities, Applications and Future Directions. *arXiv* **2021**, preprint. arXiv:2110.04160.
79. Sarkar, S.; Agrawal, S.; Baker, T.; Maddikunta, P.K.R.; Gadekallu, T.R. Catalysis of neural activation functions: Adaptive feed-forward training for big data applications. *Appl. Intell.* **2022**, *52*, 13364–13383. [\[CrossRef\]](#)
80. Cao, X.; Liu, L.; Cheng, Y.; Shen, X. Towards energy-efficient wireless networking in the big data era: A survey. *IEEE Commun. Surv. Tutorials* **2017**, *20*, 303–332. [\[CrossRef\]](#)
81. Li, X.H.; Cao, C.C.; Shi, Y.; Bai, W.; Gao, H.; Qiu, L.; Wang, C.; Gao, Y.; Zhang, S.; Xue, X.; et al. A survey of data-driven and knowledge-aware explainable AI. *IEEE Trans. Knowl. Data Eng.* **2020**, *34*, 29–49. [\[CrossRef\]](#)
82. Yoo, H.; Park, R.C.; Chung, K. IoT-Based Health Big-Data Process Technologies: A Survey. *KSII Trans. Internet Inf. Syst. (TIIS)* **2021**, *15*, 974–992.
83. Sachan, S.; Yang, J.B.; Xu, D.L.; Benavides, D.E.; Li, Y. An explainable AI decision-support-system to automate loan underwriting. *Expert Syst. Appl.* **2020**, *144*, 113100.
84. Ohana, J.J.; Ohana, S.; Benhamou, E.; Saltiel, D.; Guez, B. Explainable AI (XAI) Models Applied to the Multi-agent Environment of Financial Markets. In Proceedings of the International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems, Online, 3–7 May 2021; pp. 189–207.
85. Lee, R.S. Chaotic type-2 transient-fuzzy deep neuro-oscillatory network (CT2TFDNN) for worldwide financial prediction. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 731–745. [\[CrossRef\]](#)
86. Khodabandehloo, E.; Riboni, D.; Alimohammadi, A. HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline. *Future Gener. Comput. Syst.* **2021**, *116*, 168–189.
87. Brenas, J.H.; Shaban-Nejad, A. Health intervention evaluation using semantic explainability and causal reasoning. *IEEE Access* **2020**, *8*, 9942–9952. [\[CrossRef\]](#)
88. Mugurusi, G.; Oluka, P.N. Towards Explainable Artificial Intelligence (XAI) in Supply Chain Management: A Typology and Research Agenda. In Proceedings of the IFIP International Conference on Advances in Production Management Systems, Gyeongju, Republic of Korea, 25–29 May 2021; pp. 32–38.
89. Reyes, P.M.; Visich, J.K.; Jaska, P. Managing the dynamics of new technologies in the global supply chain. *IEEE Eng. Manag. Rev.* **2020**, *48*, 156–162.
90. Hūu, P.; Arfaoui, M.A.; Sharafeddine, S.; Assi, C.M.; Ghrayeb, A. A low-complexity framework for joint user pairing and power control for cooperative NOMA in 5G and beyond cellular networks. *IEEE Trans. Commun.* **2020**, *68*, 6737–6749. [\[CrossRef\]](#)
91. Zhang, C.; Ueng, Y.L.; Studer, C.; Burg, A. Artificial intelligence for 5G and beyond 5G: Implementations, algorithms and optimizations. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2020**, *10*, 149–163. [\[CrossRef\]](#)

92. Zhao, J.; Hu, T.; Zheng, R.; Ba, P.; Mei, C.; Zhang, Q. Defect recognition in concrete ultrasonic detection based on wavelet packet transform and stochastic configuration networks. *IEEE Access* **2021**, *9*, 9284–9295. [\[CrossRef\]](#)
93. Wang, S.; Qureshi, M.A.; Miralles-Pechuaán, L.; Huynh-The, T.; Gadekallu, T.R.; Liyanage, M. Explainable AI for B5G/6G: Technical Aspects, Use Cases and Research Challenges. *arXiv* **2021**, preprint. arXiv:2112.04698.
94. Li, C.; Guo, W.; Sun, S.C.; Al-Rubaye, S.; Tsourdos, A. Trustworthy deep learning in 6G-enabled mass autonomy: From concept to quality-of-trust key performance indicators. *IEEE Veh. Technol. Mag.* **2020**, *15*, 112–121. [\[CrossRef\]](#)
95. Guo, W. Explainable artificial intelligence for 6G: Improving trust between human and machine. *IEEE Commun. Mag.* **2020**, *58*, 39–45. [\[CrossRef\]](#)
96. Nassar, E.; El-Khalil, R. Assessing Agility Implementation in Manufacturing. In Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management, Detroit, MI, USA, 10–14 August 2020; pp. 10–14.
97. Lee, E.; Barthelmey, A.; Reckelkamm, T.; Kang, H.; Son, J. A Study on Human-Robot Collaboration based Hybrid Assembly System for Flexible Manufacturing. In Proceedings of the IECON 2019—45th Annual Conference of the IEEE Industrial Electronics Society, Lisbon, Portugal, 14–17 October 2019; Volume 1, pp. 4197–4202.
98. El-Khalil, R.; Nader, J. Impact of flexibility on operational performance: A case from us automotive manufacturing facilities. In Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management, Detroit, MI, USA, 10–14 August, 2020; pp. 10–14.
99. Hrnjica, B.; Softic, S. Explainable AI in Manufacturing: A Predictive Maintenance Case Study. In Proceedings of the IFIP International Conference on Advances in Production Management Systems, Novi Sad, Serbia, 30 August–3 September 2020; pp. 66–73.
100. Bailak, G.; Rubinger, B.; Jang, M.; Dawson, F. Advanced Robotics Mechatronics System: Emerging technologies for interplanetary robotics. In Proceedings of the Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No. 04CH37513), Niagara Falls, ON, Canada, 2–5 May 2004; Volume 4, pp. 2025–2028.
101. Anjomshoe, S.; Najjar, A.; Calvaresi, D.; Främling, K. Explainable agents and robots: Results from a systematic literature review. In Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, QC, Canada, 13–17 May 2019; pp. 1078–1088.
102. De Alwis, C.; Kalla, A.; Pham, Q.V.; Kumar, P.; Dev, K.; Hwang, W.J.; Liyanage, M. Survey on 6G frontiers: Trends, applications, requirements, technologies and future research. *IEEE Open J. Commun. Soc.* **2021**, *2*, 836–886. [\[CrossRef\]](#)
103. Zablocki, É.; Ben-Younes, H.; Pérez, P.; Cord, M. Explainability of vision-based autonomous driving systems: Review and challenges. *arXiv* **2021**, preprint. arXiv:2101.05307.
104. Malik, S.; Khan, M.A.; El-Sayed, H. Collaborative Autonomous Driving—A Survey of Solution Approaches and Future Challenges. *Sensors* **2021**, *21*, 3783. [\[CrossRef\]](#)
105. Dave, D.; Naik, H.; Singhal, S.; Patel, P. Explainable ai meets healthcare: A study on heart disease dataset. *arXiv* **2020**, preprint. arXiv:2011.03195.
106. Firouzi, F.; Farahani, B.; Daneshmand, M.; Grise, K.; Song, J.; Saracco, R.; Wang, L.L.; Lo, K.; Angelov, P.; Soares, E.; et al. Harnessing the Power of Smart and Connected Health to Tackle COVID-19: IoT, AI, Robotics and Blockchain for a Better World. *IEEE Internet Things J.* **2021**, *8*, 12826–12846. [\[CrossRef\]](#)
107. Rasheed, A.; San, O.; Kvamsdal, T. Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access* **2020**, *8*, 21980–22012. [\[CrossRef\]](#)
108. Alameddine, H.A.; Sharafeddine, S.; Sebbah, S.; Ayoubi, S.; Assi, C. Dynamic task offloading and scheduling for low-latency IoT services in multi-access edge computing. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 668–682. [\[CrossRef\]](#)
109. Ramu, S.P.; Boopalan, P.; Pham, Q.V.; Maddikunta, P.K.R.; The, T.H.; Alazab, M.; Nguyen, T.T.; Gadekallu, T.R. Federated Learning enabled Digital Twins for smart cities: Concepts, recent advances and future directions. *Sustain. Cities Soc.* **2022**, *7*, 103663. [\[CrossRef\]](#)
110. Shahat, E.; Hyun, C.T.; Yeom, C. City digital twin potentials: A review and research agenda. *Sustainability* **2021**, *13*, 3386. [\[CrossRef\]](#)
111. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital twin: Enabling technologies, challenges and open research. *IEEE Access* **2020**, *8*, 108952–108971. [\[CrossRef\]](#)
112. White, G.; Zink, A.; Codecá, L.; Clarke, S. A digital twin smart city for citizen feedback. *Cities* **2021**, *110*, 103064. [\[CrossRef\]](#)
113. Shirowzhan, S.; Tan, W.; Sepasgozar, S.M. Digital twin and CyberGIS for improving connectivity and measuring the impact of infrastructure construction planning in smart cities. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 240. [\[CrossRef\]](#)
114. Ford, D.N.; Wolf, C.M. Smart cities with digital twin systems for disaster management. *J. Manag. Eng.* **2020**, *36*, 04020027. [\[CrossRef\]](#)
115. Fan, C.; Zhang, C.; Yahja, A.; Mostafavi, A. Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. *Int. J. Inf. Manag.* **2021**, *56*, 102049. [\[CrossRef\]](#)
116. Lv, Z.; Xie, S. Artificial intelligence in the digital twins: State of the art, challenges and future research topics. *Digit. Twin* **2021**, *1*, 12. [\[CrossRef\]](#)
117. de Souza Cardoso, L.F.; Mariano, F.C.M.Q.; Zorzal, E.R. A survey of industrial augmented reality. *Comput. Ind. Eng.* **2020**, *139*, 106159. [\[CrossRef\]](#)



118. Gadekallu, T.R.; Huynh-The, T.; Wang, W.; Yenduri, G.; Ranaweera, P.; Pham, Q.V.; da Costa, D.B.; Liyanage, M. Blockchain for the Metaverse: A Review. *arXiv* **2022**, preprint. arXiv:2203.09738.
119. Zhan, T.; Yin, K.; Xiong, J.; He, Z.; Wu, S.T. Augmented reality and virtual reality displays: Perspectives and challenges. *Isience* **2020**, *23*, 101397. [\[CrossRef\]](#)
120. Mora, D.; Zimmermann, R.; Cirqueira, D.; Bezbradica, M.; Helfert, M.; Auinger, A.; Werth, D. Who Wants to Use an Augmented Reality Shopping Assistant Application? In Proceedings of the 4th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2020), Budapest, Hungary, 5–6 November 2020.
121. Hassaballah, M.; Awad, A.I. *Deep Learning in Computer Vision: Principles and Applications*; CRC Press: Boca Raton, FL, USA, 2020.
122. Chowdhary, C.L.; Alazab, M.; Chaudhary, A.; Hakak, S.; Gadekallu, T.R. *Computer Vision and Recognition Systems Using Machine and Deep Learning Approaches: Fundamentals, Technologies and Applications*; Institution of Engineering & Technology: London, UK, 2021.
123. Alazab, M.; Broadhurst, R. Spam and criminal activity. In *Trends and Issues in Crime and Criminal Justice*; Australian Institute of Criminology: Canberra, Australia, 2016; pp. 1–20.
124. Shorfuazzaman, M.; Hossain, M.S.; Alhamid, M.F. Towards the sustainable development of smart cities through mass video surveillance: A response to the COVID-19 pandemic. *Sustain. Cities Soc.* **2021**, *64*, 102582. [\[CrossRef\]](#)
125. Javed, A.R.; Jalil, Z.; Zehra, W.; Gadekallu, T.R.; Suh, D.Y.; Piran, M.J. A comprehensive survey on digital video forensics: Taxonomy, challenges and future directions. *Eng. Appl. Artif. Intell.* **2021**, *106*, 104456. [\[CrossRef\]](#)
126. Glomsrud, J.A.; Ødegårdstuen, A.; Clair, A.L.S.; Smogeli, Ø. Trustworthy versus explainable AI in autonomous vessels. In Proceedings of the International Seminar on Safety and Security of Autonomous Vessels (ISSAV) and European STAMP Workshop and Conference (ESWC), Helsinki, Finland, 17–19 September 2019 ; pp. 37–47.
127. Hamza, A.; Javed, A.R.R.; Iqbal, F.; Kryvinska, N.; Almadhor, A.S.; Jalil, Z.; Borghol, R. Deepfake Audio Detection via MFCC Features Using Machine Learning. *IEEE Access* **2022**, *10*, 134018–134028. [\[CrossRef\]](#)
128. Tian, S.; Yang, W.; Le Grange, W. A.; Wang, P.; Huang, W.; Ye, Z. Smart healthcare: Making medical care more intelligent. *Glob. Health J.* **2019**, *3*, 62–65. [\[CrossRef\]](#)
129. Pawar, U.; O'Shea, D.; Rea, S.; O'Reilly, R. Explainable ai in healthcare. In Proceedings of the 2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), Dublin, Ireland, 15–19 June 2020; pp. 1–2.
130. Doshi-Velez, F.; Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv* **2017**, arXiv:1702.08608.
131. Mourad, A.; Tout, H.; Wahab, O.A.; Otrok, H.; Dbouk, T. Ad hoc vehicular fog enabling cooperative low-latency intrusion detection. *IEEE Internet Things J.* **2020**, *8*, 829–843. [\[CrossRef\]](#)
132. Bojarski, M.; Del Testa, D.; Dworakowski, D.; Firner, B.; Flepp, B.; Goyal, P.; Jackel, L.D.; Monfort, M.; Muller, U.; Zhang, J.; et al. End to end learning for self-driving cars. *arXiv* **2016**, arXiv:1604.07316.
133. Zahavy, T.; Ben-Zrihem, N.; Mannor, S. Graying the black box: Understanding dqns. In Proceedings of the International Conference on Machine Learning, PMLR, New York, NY, USA, 20–22 June 2016; pp. 1899–1908.
134. Bengio, Y. *Learning Deep Architectures for AI*; Now Publishers Inc.: Norwell, MA, USA, 2009.
135. Kumar, G.; Kumar, K.; Sachdeva, M. The use of artificial intelligence based techniques for intrusion detection: A review. *Artif. Intell. Rev.* **2010**, *34*, 369–387. [\[CrossRef\]](#)
136. McFarland, M. Uber shuts down self-driving operations in Arizona. *CNNMoney Version* **2018**, *809*, 3.
137. Mori, K.; Fukui, H.; Murase, T.; Hirakawa, T.; Yamashita, T.; Fujiyoshi, H. Visual explanation by attention branch network for end-to-end learning-based self-driving. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 1577–1582.
138. Haspiel, J.; Du, N.; Meyerson, J.; Robert Jr, L.P.; Tilbury, D.; Yang, X.J.; Pradhan, A.K. Explanations and expectations: Trust building in automated vehicles. In Proceedings of the Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago IL, USA, 5–8 March, 2018; pp. 119–120.
139. Saraf, A.P.; Chan, K.; Popish, M.; Browder, J.; Schade, J. Explainable Artificial Intelligence for Aviation Safety Applications. In Proceedings of the AIAA Aviation 2020 Forum, Online, 15–19 June 2020; p. 2881.
140. Zeldam, S. Automated Failure Diagnosis in Aviation Maintenance Using Explainable Artificial Intelligence (XAI). Master's Thesis, University of Twente, Enschede, The Netherlands, 2018.
141. Tan, S.; Caruana, R.; Hooker, G.; Lou, Y. Detecting bias in black-box models using transparent model distillation. *arXiv* **2017**, preprint. arXiv:1710.06169.
142. Howell, C. A framework for addressing fairness in consequential machine learning. In Proceedings of the FAT Conference Tuts, New York, NY, USA, 23–24 February 2018; pp. 1–2.
143. Berk, R.A.; Bleich, J. Statistical procedures for forecasting criminal behavior: A comparative assessment. *Criminol. Pub. Pol'y* **2013**, *12*, 513. [\[CrossRef\]](#)
144. Lightbourne, J. Damned lies & criminal sentencing using evidence-based tools. *Duke Tech. Rev.* **2017**, *15*, 327.
145. Madi, J. *FinTech: Law and Regulation*; Edward Elgar Publishing: Cheltenham, UK, 2019.
146. Abbasi, A.; Javed, A.R.; Iqbal, F.; Kryvinska, N.; Jalil, Z. Deep learning for religious and continent-based toxic content detection and classification. *Sci. Rep.* **2022**, *12*, 17478. [\[CrossRef\]](#)
147. Gao, M.; Liu, X.; Xu, A.; Akkiraju, R. Chat-XAI: A New Chatbot to Explain Artificial Intelligence. In Proceedings of the SAI Intelligent Systems Conference, Amsterdam, The Netherlands, 2–3 September 2021; pp. 125–134.

148. Cirqueira, D.; Helfert, M.; Bezbradica, M. Towards Design Principles for User-Centric Explainable AI in Fraud Detection. In Proceedings of the International Conference on Human-Computer Interaction, Bari, Italy, 30 August–3 September 2021; pp. 21–40.
149. Souza, J.; Leung, C.K. Explainable Artificial Intelligence for Predictive Analytics on Customer Turnover: A User-Friendly Interface for Non-expert Users. In *Explainable AI Within the Digital Transformation and Cyber Physical Systems*; Springer: Cham, Switzerland, 2021; pp. 47–67.
150. Lv, Z.; Qiao, L.; Singh, A.K. Advanced machine learning on cognitive computing for human behavior analysis. *IEEE Trans. Comput. Soc. Syst.* **2020**, *8*, 1194–1202. [\[CrossRef\]](#)
151. Du, X.; Hargreaves, C.; Sheppard, J.; Anda, F.; Sayakkara, A.; Le-Khac, N.A.; Scanlon, M. SoK: Exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Online, 25–28 August 2020; pp. 1–10.
152. Ahmed, W.; Shahzad, F.; Javed, A.R.; Iqbal, F.; Ali, L. WhatsApp Network Forensics: Discovering the IP Addresses of Suspects. In Proceedings of the 2021 11th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 19–21 April 2021; pp. 1–7. [\[CrossRef\]](#)
153. Bhattacharya, S.; Chengoden, R.; Srivastava, G.; Alazab, M.; Javed, A.R.; Victor, N.; Maddikunta, P.K.R.; Gadekallu, T.R. Incentive Mechanisms for Smart Grid: State of the Art, Challenges, Open Issues, Future Directions. *Big Data Cogn. Comput.* **2022**, *6*, 47. [\[CrossRef\]](#)
154. Karimipour, H.; Dehghantanha, A.; Parizi, R.M.; Choo, K.K.R.; Leung, H. A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. *IEEE Access* **2019**, *7*, 80778–80788. [\[CrossRef\]](#)
155. Zhang, L.; Wang, G.; Giannakis, G.B. Real-time power system state estimation and forecasting via deep unrolled neural networks. *IEEE Trans. Signal Process.* **2019**, *67*, 4069–4077. [\[CrossRef\]](#)
156. Foruzan, E.; Soh, L.K.; Asgarpour, S. Reinforcement learning approach for optimal distributed energy management in a microgrid. *IEEE Trans. Power Syst.* **2018**, *33*, 5749–5758. [\[CrossRef\]](#)
157. Jiang, H.; Zhang, J.J.; Gao, W.; Wu, Z. Fault detection, identification and location in smart grid based on data-driven computational methods. *IEEE Trans. Smart Grid* **2014**, *5*, 2947–2956. [\[CrossRef\]](#)
158. Omitaomu, O.A.; Niu, H. Artificial Intelligence Techniques in Smart Grid: A Survey. *Smart Cities* **2021**, *4*, 548–568. [\[CrossRef\]](#)
159. Taeihagh, A. Governance of artificial intelligence. *Policy Soc.* **2021**, *40*, 137–157. [\[CrossRef\]](#)
160. El-Khalil, R. Classification, purpose, enablers of lean dimensions at automotive manufacturing industry: A case study. In Proceedings of the International Conference on Industrial Engineering and Operations Management, Paris, France, 26–27 July 2018; pp. 2743–2756.
161. Possik, J.; Zouggar-Amrani, A.; Vallespir, B.; Zacharewicz, G. Lean techniques impact evaluation methodology based on a co-simulation framework for manufacturing systems. *Int. J. Comput. Integr. Manuf.* **2022**, *35*, 91–111. [\[CrossRef\]](#)
162. Maddikunta, P.K.R.; Pham, Q.V.; Prabadevi, B.; Deepa, N.; Dev, K.; Gadekallu, T.R.; Ruby, R.; Liyanage, M. Industry 5.0: A survey on enabling technologies and potential applications. *J. Ind. Inf. Integr.* **2022**, *26*, 100257. [\[CrossRef\]](#)
163. Majid, M.; Habib, S.; Javed, A.R.; Rizwan, M.; Srivastava, G.; Gadekallu, T.R.; Lin, J.C.W. Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review. *Sensors* **2022**, *22*, 2087. [\[CrossRef\]](#)
164. Boopalan, P.; Ramu, S.P.; Pham, Q.V.; Dev, K.; Maddikunta, P.K.R.; Gadekallu, T.R.; Huynh-The, T.; et al. Fusion of Federated Learning and Industrial Internet of Things: A survey. *Comput. Netw.* **2022**, *212*, 109048. [\[CrossRef\]](#)
165. Jagatheesaperumal, A.R.J.; Rahouti, M.; Ahmad, K.; Al-Fuqaha, A.; Guizani, M. The Duo of Artificial Intelligence and Big Data for Industry 4.0: Review of Applications, Techniques, Challenges and Future Research Directions. *arXiv* **2021**, arXiv:2104.02425.
166. Ahmad, K.; Maabreh, M.; Ghaly, M.; Khan, K.; Qadir, J.; Al-Fuqaha, A. Developing future human-centered smart cities: Critical analysis of smart city security, Data management and Ethical challenges. *Comput. Sci. Rev.* **2022**, *43*, 100452. [\[CrossRef\]](#)
167. Ahmed, I.; Jeon, G.; Piccialli, F. From Artificial Intelligence to eXplainable Artificial Intelligence in Industry 4.0: A survey on What, How and Where. *IEEE Trans. Ind. Inform.* **2022**, *18*, 5031–5042. [\[CrossRef\]](#)
168. Shakur, A.H.; Qian, X.; Wang, Z.; Mortazavi, B.; Huang, S. GPSRL: Learning Semi-Parametric Bayesian Survival Rule Lists from Heterogeneous Patient Data. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 10608–10615.
169. Dhasarathan, C.; Hasan, M.K.; Islam, S.; Abdullah, S.; Mokhtar, U.A.; Javed, A.R.; Goundar, S. COVID-19 health data analysis and personal data preserving: A homomorphic privacy enforcement approach. *Comput. Commun.* **2023**, *199*, 87–97. [\[CrossRef\]](#)
170. Lipton, Z.C.; Kale, D.C.; Wetzel, R. Modeling missing data in clinical time series with RNNs. *Mach. Learn. Healthc.* **2016**, *56*.
171. Amin, N.; McGrath, A.; Chen, Y.P.P. FexRNA: Exploratory data analysis and feature selection of non-coding RNA. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2021**, *18*, 2795–2801. [\[CrossRef\]](#)
172. Patnaik, S.; Sidhardh, S.; Semperlotti, F. Towards a unified approach to nonlocal elasticity via fractional-order mechanics. *Int. J. Mech. Sci.* **2021**, *189*, 105992. [\[CrossRef\]](#)
173. Rosenfeld, A. Better metrics for evaluating explainable artificial intelligence. In Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems, Online, 3–7 May 2021; pp. 45–50.
174. Panigutti, C.; Perotti, A.; Pedreschi, D. Doctor XAI: An ontology-based approach to black-box sequential data classification explanations. In Proceedings of the 2020 Conference On Fairness, Accountability and Transparency, Barcelona, Spain, 27–30 January 2020; pp. 629–639.



175. Zhou, J.; Gandomi, A.H.; Chen, F.; Holzinger, A. Evaluating the quality of machine learning explanations: A survey on methods and metrics. *Electronics* **2021**, *10*, 593. [\[CrossRef\]](#)
176. Burkart, N.; Huber, M.F. A survey on the explainability of supervised machine learning. *J. Artif. Intell. Res.* **2021**, *70*, 245–317. [\[CrossRef\]](#)
177. Mencar, C.; Castiello, C.; Cannone, R.; Fanelli, A.M. Interpretability assessment of fuzzy knowledge bases: A cointension based approach. *Int. J. Approx. Reason.* **2011**, *52*, 501–518. [\[CrossRef\]](#)
178. Poursabzi-Sangdeh, F.; Goldstein, D.G.; Hofman, W. A.; Wortman Vaughan, J.W.; Wallach, H. Manipulating and measuring model interpretability. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–52.
179. Tulio Ribeiro, M.; Singh, S.; Guestrin, C. Model-Agnostic Interpretability of Machine Learning. *arXiv* **2016**, 1606. arXiv:1606.05386.
180. El Bekri, N.; Kling, J.; Huber, M.F. A study on trust in black box models and post-hoc explanations. In Proceedings of the International Workshop on Soft Computing Models in Industrial and Environmental Applications, Seville, Spain, 13–15 May 2019; pp. 35–46.
181. Yin, M.; Wortman Vaughan, J.; Wallach, H. Understanding the effect of accuracy on trust in machine learning models. In Proceedings of the 2019 Chi Conference on Human Factors in Computing Systems, Glasgow, Scotland, 4–9 May 2019; pp. 1–12.
182. Lage, I.; Chen, E.; He, J.; Narayanan, M.; Kim, B.; Gershman, S.; Doshi-Velez, F. An evaluation of the human-interpretability of explanation. *arXiv* **2019**, preprint. arXiv:1902.00006.
183. Herman, B. The promise and peril of human evaluation for model interpretability. *arXiv* **2017**, 8. preprint. arXiv:1711.07414.
184. Hamilton, D.; Kornegay, K.; Watkins, L. Autonomous Navigation Assurance with Explainable AI and Security Monitoring. In Proceedings of the 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 13–15 October 2020; pp. 1–7. [\[CrossRef\]](#)
185. Shahzad, F.; Iqbal, W.; Bokhari, F.S. On the use of CryptDB for securing Electronic Health data in the cloud: A performance study. In Proceedings of the 2015 17th International Conference on E-health Networking, Application Services (HealthCom), Boston, MA, USA, 14–17 October 2015; pp. 120–125. [\[CrossRef\]](#)
186. Security and Privacy Accountable Technology Innovations, Algorithms and machine Learning (SPATIAL). Available online: <https://cordis.europa.eu/project/id/101021808> (accessed on 1 September 2021).
187. Explanation of AI for Decision Making (XAI). Available online: <https://www.ibm.com/watson/explainable-ai> (accessed on 8 May 2012). (accessed on 1 September 2021).
188. Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI). Available online: <https://cordis.europa.eu/project/id/860621> (accessed on 1 September 2021).
189. Explainable Manufacturing Artificial Intelligence (XMANAI). Available online: <https://cordis.europa.eu/project/id/957362> (accessed on 1 September 2021).
190. Explainable AI Pipelines for Big Copernicus Data (DEEPCUBE). Available online: <https://ojs.mediageo.it/index.php/GEOmedia/article/view/1802> (accessed on 1 September 2021).
191. Copernicus:Europe’s eyes on Earth. Available online: <https://books.google.co.jp/books/about/Copernicus.html?id=LCvxjwEACAAJ> (accessed on 1 September 2021).
192. AI4EU. Available online: <https://www.ai4europa.eu/> (accessed on 1 September 2021).
193. Safe and Trusted Human Centric Artificial Intelligence in Future Manufacturing Lines (STAR). Available online: <https://cordis.europa.eu/project/id/956573/results> (accessed on 1 September 2021).
194. FeatureCloud. Available online: <https://featurecloud.eu/> (accessed on 1 September 2021).
195. Building Greener and More Sustainable Societies by Filling the Knowledge Gap in Social Science and Engineering to Enable Responsible Artificial Intelligence Co-Creation (GECKO). Available online: <https://vbn.aau.dk/en/projects/building-greener-and-more-sustainable-societies-by-filling-the-kn> (accessed on 1 September 2021).
196. Explainable Artificial Intelligence(XAI) Center. Available online: <https://www.imperial.ac.uk/explainable-artificial-intelligence/> (accessed on 1 September 2021).
197. Kabir, M.H.; Hasan, K.F.; Hasan, M.K.; Ansari, K. Explainable Artificial Intelligence for Smart City Application: A Secure and Trusted Platform. *arXiv* **2021**, arXiv:2111.00601.
198. Gadekallu, T.R.; Manoj, M.; Kumar, N.; Hakak, S.; Bhattacharya, S.; et al. Blockchain-Based Attack Detection on Machine Learning Algorithms for IoT-Based e-Health Applications. *IEEE Internet Things Mag.* **2021**, *4*, 30–33. [\[CrossRef\]](#)
199. Mahbooba, B.; Timilsina, M.; Sahal, R.; Serrano, M. Explainable artificial intelligence (xai) to enhance trust management in intrusion detection systems using decision tree model. *Complexity* **2021**, *2021*, 6634811. [\[CrossRef\]](#)
200. Javed, A.; Muhammad, A.H.; Faisal, S.; Waqas, A.; Saurabh, S.; Thar, B.; Gadekallu, T. Integration of Blockchain Technology and Federated Learning in Vehicular (IoT) Networks: A Comprehensive Survey. *Sensors* **2022**, *22*, 4394. [\[CrossRef\]](#)
201. Kim, S.; Jo, W.; Lee, J.; Shon, T. AI-enabled device digital forensics for smart cities. *J. Supercomput.* **2022**, *78*, 3029–3044. [\[CrossRef\]](#)
202. Das, A.; Rad, P. Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv* **2020**, preprint. arXiv:2006.11371.
203. Demollin, M.; Budzynska, K.; Sierra, C. Argumentation theoretical frameworks for explainable artificial intelligence. In Proceedings of the 2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence, Dublin, Ireland, November 2020 ; pp. 44–49.

204. Khamparia, A.; Gupta, D.; Khanna, A.; Balas, V.E. *Biomedical Data Analysis and Processing Using Explainable (XAI) and Responsive Artificial Intelligence (RAI)*; Springer: Berlin/Heidelberg, Germany, 2022.
205. Gerlings, J.; Shollo, A.; Constantiou, I. Reviewing the Need for Explainable Artificial Intelligence (xAI). *arXiv* **2020**, arXiv:2012.01007.
206. Al-Hilo, A.; Samir, M.; Assi, C.; Sharafeddine, S.; Ebrahimi, D. UAV-assisted content delivery in intelligent transportation systems-joint trajectory planning and cache management. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 5155–5167. [[CrossRef](#)]
207. Gupta, B.B.; Gaurav, A.; Marín, E.C.; Alhalabi, W. Novel graph-based machine learning technique to secure smart vehicles in intelligent transportation systems. *IEEE Trans. Intell. Transp. Syst.* **2022**. [[CrossRef](#)]
208. Samir, M.; Sharafeddine, S.; Assi, C.M.; Nguyen, T.M.; Ghayeb, A. UAV trajectory planning for data collection from time-constrained IoT devices. *IEEE Trans. Wirel. Commun.* **2019**, *19*, 34–46. [[CrossRef](#)]
209. Samir, M.; Assi, C.; Sharafeddine, S.; Ebrahimi, D.; Ghayeb, A. Age of information aware trajectory planning of UAVs in intelligent transportation systems: A deep learning approach. *IEEE Trans. Veh. Technol.* **2020**, *69*, 12382–12395. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.