International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Explainable AI for Trustworthy Decision-making in IoT Environments

Dr. Huma Khan[1], Sheetal Pradip Patil[2], Arpit Namdev[3], Dr. Gitanjali Shrivastava[4,*], Nagarjuna Karyemsetty[5], Elangovan Muniyandy[6], Ankur Gupta[7]

Submitted: 03/02/2024 **Revised**: 11/03/2024 **Accepted**: 17/03/2024

Abstract: Artificial Intelligence (AI) is being widely incorporated into Internet of Things (IoT) contexts, leading to improved automation and efficiency. Given that AI algorithms have a substantial impact on decision-making processes in these intricate ecosystems, it becomes crucial to prioritize their reliability. Explainable AI (XAI) is becoming more important for promoting openness and accountability in decision-making inside the Internet of Things (IoT). Through the provision of explanations that humans can comprehend, explainable artificial intelligence (XAI) improves stakeholders' understanding of the underlying logic and allows for the detection and reduction of any biases or mistakes. This abstract explores the importance of Explainable Artificial Intelligence (XAI) in facilitating reliable decision-making in Internet of Things (IoT) contexts. It highlights the function of XAI in improving transparency, reducing risks, and building trust among stakeholders. This abstract emphasizes the crucial need to include explainability into AI-driven decision-making processes in order to guarantee their dependability and ethical soundness. It does so by thoroughly examining XAI concepts and techniques specifically designed for the problems posed by IoT ecosystems.

Keywords: *Machine learning, Blockchain, IoT, Predictive analytics*

1. Introduction

A new age of connectedness has dawned with the advent of Internet of Things (IoT) devices, which have allowed for the automation of many jobs across varied areas and the smooth interchange of data. Still, reliable methods of making decisions are critical as the Internet of Things (IoT) grows in both scope and connectivity. When it comes to optimizing resource allocation, boosting operational efficiency, and improving user experiences, many IoT applications rely heavily on judgments made by artificial intelligence (AI) algorithms. Concerns about AI models' dependability, accountability, and ethical consequences arise from the fact that they are sometimes difficult to comprehend and trust because to their complexity and opacity.The idea of Explainable AI (XAI) has grown in popularity as a solution to these problems; it aims to make AI decision-making processes more transparent and interpretable. The goal of XAI approaches is to make AI judgments more human-like so that people can understand why they happened and whether or not the results are reliable. There is an immediate and pressing demand for reliable decision-making in Internet of Things (IoT) settings since judgments in these settings affect several important applications, including smart cities, healthcare monitoring, industrial automation, and autonomous cars."Explainable AI for Trustworthy Decision-making in IoT Environments" aims to investigate and create XAI approaches that are particularly designed for IoT applications in order to tackle this urgent demand. The research's overarching goal is to pave the way for the broad application of AI technologies in IoT systems by making AI algorithms running in IoT settings more transparent and interpretable; this will increase trust, accountability, and user acceptability. When it comes to trustworthy decision-making in IoT contexts, there are both advantages and disadvantages, and we cover them all in this article. Our goal is to fill in the gaps and highlight the limits in the present literature on XAI approaches and how they are used in IoT scenarios.

1.1. Artificial Intelligence

At the vanguard of technical advancement, artificial intelligence (AI) is reshaping human-machine interaction, data processing, and problem-solving. Artificial intelligence (AI) is essentially the study and creation of computer systems that can learn, reason, perceive, and make decisions, all of which are traditionally associated with human intellect. With its revolutionary impact on businesses and cultural norms, artificial intelligence (AI)

[1]*Associate Professor, Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Bhilai, Chhattisgarh, India Email: huma@rungta.ac.in*
[2]*Assistant Professor, Department of Management Studies Bharati Vidyapeeth Deemed to be University, Navi Mumbai, Maharashtra, India Email: sheetal.patil3@bharatividyapeeth.edu*
[3]*Assistant Professor, Dept. of Information Technology, University Institute of Technology RGPV, Bhopal, Madhya Pradesh, India Email: namdev.arpit@gmail.com*
[4]*Assistant Professor, Department of Law, Symbiosis Law School, Pune, India Email: dr.gitanjali10@gmail.com*
[5]*Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India Email: nagarjunaatwork@gmail.com*
[6]*Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India Email: muniyandy.e@gmail.com*
[7]*Assistant Professor, Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India Email: ankurdujana@gmail.com*
[*]*Corresponding Author: Dr. Gitanjali Shrivastava (dr.gitanjali10@gmail.com)*

has invaded almost every facet of contemporary life, from financial analysis and medical diagnosis to autonomous cars and virtual assistants. Midway through the twentieth century, with seminal breakthroughs like the Turing Test and the advent of neural networks, artificial intelligence (AI) emerged as a distinct academic discipline. The availability of massive volumes of data, innovations in algorithms, and improvements in computing power have all contributed to AI's meteoric rise in the last few decades. These developments have taken artificial intelligence out of the lab and into the real world, where it is changing businesses, communities, and people's daily lives. Learning from data and improving over time via iterative methods is one of the fundamental aspects of current AI. Machine Learning (ML) is a branch of artificial intelligence that enables computers to learn correlations and patterns in data automatically, allowing them to spot patterns, make predictions, and extract insights from complicated datasets

1.2. Internet of Things (IoT)

The Internet of Things (IoT) is revolutionizing our relationship with the physical world and is causing a paradigm change. Through the use of embedded software, sensors, and networking capabilities, the Internet of Things (IoT) enables a system of interconnected computer devices, physical objects, and physical sensors to collect, send, and analyze data autonomously, all without the need for any kind of human intervention. Embedded inside this interconnected environment is a vast network of "things" that can converse and act in tandem with ease. This encompasses everything from smart household appliances and wearable electronics to large machinery and urban infrastructure. Advancements in connectivity, shrinkage, and the cost of hardware components are driving the expansion of IoT technology, which is in turn being propelled by the rise of cloud computing and big data analytics. These advancements have paved the way for the merging of the digital and physical worlds, which in turn has unleashed a flood of new opportunities across several industries. The proliferation of Internet of Things (IoT) smart devices has revolutionized consumer electronics by making previously impossible tasks easier, more convenient, and more efficient.

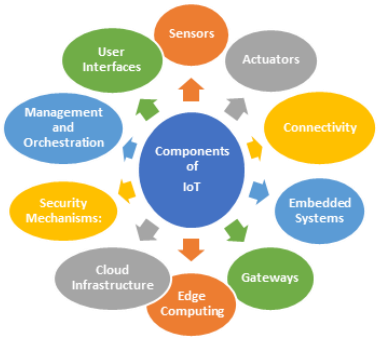

Fig. 1. Component of IoT

1.3. Decision-making

In IoT contexts, where networked devices create large volumes of data and enable automated decision-making, the incorporation of Explainable AI (XAI) is crucial for promoting confidence and dependability. This is a methodical approach that gives importance to openness, comprehensibility, and ethical concerns at each step of the decision-making process. Transparent data collecting techniques are essential for stakeholders to have a clear understanding of the sources and features of the data being used. Equally crucial is the choice of AI models that not only provide high accuracy but also allow for interpretation, either via naturally explainable algorithms or hybrid models that strike a compromise between complexity and transparency. Explainability approaches are crucial in clarifying the reasoning behind AI-driven judgments throughout the decision-making process. Offering both local and global explanations enables stakeholders to comprehend not only specific results, but also broader trends and patterns within the data. Visualizations and natural language explanations are effective methods for presenting complicated information in a clear and understandable manner, so improving understanding and building trust.

Fig. 2. Application of AI

1.4. Role of AI in IoT Environments

In the age of networked gadgets and digital transformation, the confluence of Artificial Intelligence (AI) with the Internet of Things (IoT) constitutes a critical nexus of technological innovation. The capabilities of Internet of Things settings are being revolutionized by artificial intelligence (AI), which has the potential to replicate human intellect and make sense of huge volumes of data. This enables IoT environments to become smarter, more efficient, and more responsive than they have ever been before. The essential role that artificial intelligence plays in Internet of Things settings is investigated in this introduction.

Fig. 3. Role of AI in IoT environment

It explains how AI technologies enable Internet of Things systems to analyze data, make intelligent choices, and uncover revolutionary value across a wide range of industries and applications respectively. As the number of Internet of Things devices continues to increase without stopping, the amount of data and the types of data that are produced by these networked devices have reached levels that have never been seen before. Artificial intelligence acts as the engine that drives insights and actionable intelligence from the flood of Internet of Things data in an environment that is rich in data.

1.5. Role of Decision-making in IoT Environments

A key component of operational efficiency, performance improvement, and value generation across linked ecosystems in the ever-changing IoT environment is decision-making. The term "Role of Decision-making in IoT Environments" highlights the crucial importance of decision-making processes in coordinating the actions and results of IoT installations in different fields and for different purposes. Understanding the complex nature of decision-making in IoT contexts and its far-reaching consequences for driving innovation and realizing the potential of linked systems is the basic focus of this introduction. With the proliferation of IoT technology comes a deluge of data, sometimes in real-time, from the network of linked sensors, devices, and systems. Organizations maximize operations, gain value, and extract useful insights from this data-rich environment via decision-making. Choices taken in IoT settings have an effect on performance, dependability, and efficiency in areas such as equipment failure prediction, resource allocation optimization, and adapting to changing circumstances.

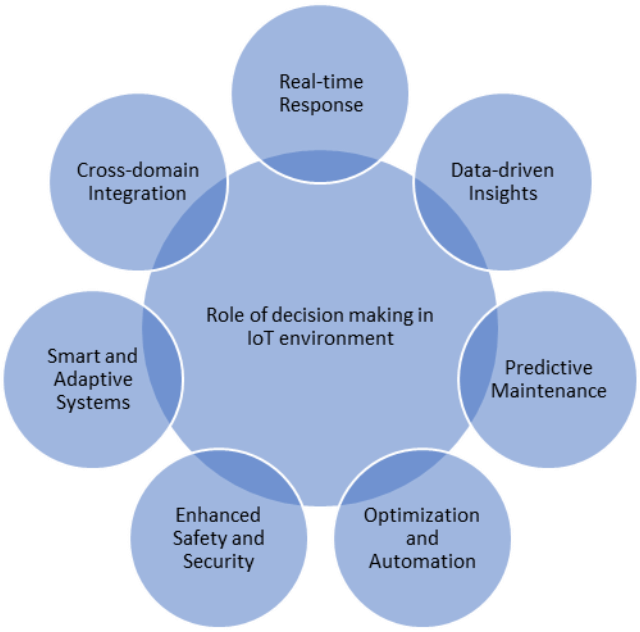

Fig. 4. Role of decision making in IoT Environment

1.6. Need of Research

There is a rising need for openness and responsibility in decision-making processes due to the increasing complexity and prevalence of AI algorithms in IoT contexts. In order for stakeholders to comprehend, analyze, and have faith in the results, explainable AI approaches strive to reveal how AI models arrive at their conclusions. This openness is critical for making sure that AI system judgments are in line with organization objectives, legal mandates, and moral principles.

- **Trust and Acceptance:** When it comes to using AI technology in IoT settings, trust is crucial. By giving consumers, decision-makers, and regulators transparent explanations of AI-driven actions, explainable AI technologies may contribute to trust building. Explainable AI reduces worries about algorithmic bias, unfairness, and dependability while simultaneously increasing adoption, cooperation, and confidence.
- **Regulatory Compliance:** In highly regulated sectors like healthcare, banking, and transportation, conformity with laws and regulations is of the utmost importance. By making decisions in an open and auditable manner, explainable AI helps businesses prove they are in compliance. Important regulations like HIPAA and the General Data Protection Regulation (GDPR) need this in order to guarantee accountability, reduce legal risks, and satisfy regulatory obligations.
- **Risk Management and Error Detection:** Artificial

intelligence systems may make mistakes, have biases, or act in unanticipated ways, no matter how smart they are. Through the use of explainable AI methodologies, stakeholders may recognize and comprehend any failures or abnormalities in decision-making processes, which aids in risk identification and mitigation. Error detection, debugging, and continual improvement in IoT installations are made easier by explainable AI, which provides insights into the performance and limits of AI models.

- **Human-Machine Interaction:** Explainable AI facilitates better human-machine communication and teamwork in Internet of Things (IoT) settings where AI systems communicate with people and other devices. To help people better comprehend AI suggestions, provide comments, and act when needed, it is important to offer interpretable explanations of AI judgments. User agency, participation, and empowerment are fostered by this people-first method of decision-making in IoT systems.
- **Ethical Considerations:** Talks regarding AI ethics are starting to center on questions of justice, accountability, and openness as they pertain to AI decision-making. With the use of explainable AI, businesses can deal with ethical concerns by learning how AI models handle various people and groups. Explainable AI helps with ethical decision-making in IoT settings by spotting and reducing biases, making sure everything is fair, and encouraging responsibility.

2. Literature Review

F. Wahab et al. (2022) conducted on big data and intrusion detection systems have inspired us to develop an approach that is adaptable and economical to protect critical environments from cyber attacks. The hybrid model described in our research consists of two components: Long short-term memory (LSTM) and gated recurrent unit (GRU). The recommended model underwent a thorough test using both the publicly available CICDDoS2019 dataset and conventional assessment techniques [1] F. Alenazi et al. (2021) used machine learning methods for traffic monitoring inside the SDN controller's Network Intrusion Detection System (NIDS). The objective was to detect malicious activities in the network. To demonstrate attack detection, a range of tree-based machine learning techniques such as Decision Tree, Random Forest, and XGBoost are used [2] Thakur, K., et al. (2021), produced domains in their first phases. This approach relies on simple and automated computable characteristics of actual domain name system (DNS) traffic. The system is meant to identify these domains without the need for reverse engineering or log monitoring, and without depending on external information such as WHOIS and DNS response. The IDGADS model is a supervised deep learning architecture that has undergone training using 17 million domains sourced from respectable entities [3]. Varghese et al. (2021) proposed a novel framework to address the performance challenges of intrusion detection systems (IDS) and the design problems of software-defined networking (SDN) related to distributed denial of service attacks. This framework involves incorporating intelligence into the data layer and utilizing the Data Plane Development Kit (DPDK) within the SDN architecture. The name of this cutting-edge framework is the DPDK-based DDoS Detection (D3) framework. [4] F. Wei et al. (2019) proposed using an anomaly-based intrusion detection system (IDS) to identify unusual network behavior and create software-defined networking (SDN) flow rules. This would enable the control of network access in a dynamic manner. By building an interpretable model to elucidate the outcomes of the anomaly-based intrusion detection system, we may get insights into network irregularities. Based on the above description, we may formulate protocols for access control. [5] J. Bhayo et al. (2023) proposed a machine learning-based approach for detecting distributed denial of service (DDoS) threats in an SDN-WISE Internet of Things (IoT) controller. A machine learning-based detection module has been included into the controller, and a testbed environment has been set up to simulate the generation of DDoS attack traffic. A supplementary logging method has been included in the SDN-WISE controller to capture the traffic. [6] P. Kumar e al. (2023) presented a vision for a future power transmission network that is intelligent and efficient, representing the next generation of technology. When functioning in an SG context, Smart Meters (SMs) generally engage in communication with Service Providers (SPs) across an unsecured public channel to exchange services and data. Consequently, the whole SG ecosystem is vulnerable to a diverse range of precarious security scenarios [7] M. Girdhar et al. (2023) proposed an architecture that combines an intrusion detection system (IDS) with software-defined networking (SDN) to detect and prevent the injection of malicious IEC 61850-based generic object-oriented substation event (GOOSE) messages in a digital substation [8] Kumar et al. (2023) explored the feasibility of developing security tools with the ability to analyze and detect malicious network traffic in a highly effective and efficient way, specifically for the goal of recognizing intrusions. Machine learning (ML) methods have recently attracted significant attention from both academia and industry for developing intrusion detection systems (IDSs) that incorporate logically centralized control and a comprehensive network perspective provided by software-defined networking (SDN). [9] S. Dahiya et al. (2021), conducted current improvements in network and communication have shown that attackers are presenting various difficulties to the

system via diverse approaches. Deploying an intrusion detection system (IDS) is essential to ensure the network's availability, integrity, and confidence. This system is specifically intended to proactively detect and block any possible unauthorized access to the network by continuously monitoring the flow of network data and identifying any malicious actions [10]. H. Zhang et al. (2019) introduced a software defined, artificial intelligence-based two-stage intrusion detection system in their publication. The network flows are collected comprehensively and adaptably, providing a worldwide outlook, and it detects attacks with advanced intelligence. To choose typical traits, we use the Bat technique, which involves the integration of swarm division and binary differential mutation. [11] A. Zaheer et al. (2021) provided a methodology for managing intrusions inside the Internet of Things (IoT) network after they have been detected by an Intrusion Detection System (IDS). By analyzing host logs and network traffic, the intrusion detection system (IDS) may detect intrusions [12].

3. Problem Statement

Explainable AI for Trustworthy Decision-making in IoT Environments" must overcome in order to be successful and to have practical relevance. When it comes to providing explanations that are sufficiently informative and understandable to end-users, traditional ways often fail. This is especially true in situations that are complex and include the Internet of Things (IoT). Although there have been attempts made to create strategies for interpretability, many of these methods are still at a point where they are unable to provide significant insights into the decision-making processes of artificial intelligence systems that are functioning inside ecosystems of the internet of things. Internet of Things settings, there are often a significant number of devices that are linked to one another and generate enormous volumes of data in real time. It is possible that conventional approaches are not able to handle the sheer amount and velocity of data that is created in such contexts. This might result in scalability concerns as well as significant performance reduction. The incorporation of explainability into AI models without compromising their predicted accuracy and efficiency is yet another significant obstacle that must be overcome. It is sometimes difficult for conventional techniques to achieve the appropriate balance between interpretability and performance, which may result in trade-offs that may impair the overall dependability and trustworthiness of decision-making systems in Internet of Things contexts. Additionally, the dynamic nature of surroundings that are connected to the internet of things presents hurdles for traditional research endeavors. Data patterns that are always changing, user needs that are constantly shifting, and unexpected occurrences are all characteristics of these settings.

4. Challenges

For the purpose of assuring trustworthy decision-making in Internet of Things contexts, where openness, interpretability, and accountability are of the utmost importance, explainable artificial intelligence (XAI) is essential. On the other hand, in order to develop successful XAI in IoT contexts, there are various problems that need to be addressed:

. Artificial intelligence models that are utilized in Internet of Things contexts, such as deep neural networks, may be very complicated and opaque, which makes it difficult to describe the decision-making processes that they employ. There is a substantial issue involved in simplifying and comprehending complicated models without the performance being compromised.

2. Interpretability vs Accuracy in Regards to When it comes to model interpretability, there is often a trade-off between accuracy and interpretability. It's possible that more interpretable models will compromise predictive performance, whereas models that are very accurate could not be transparent enough. The issue lies in striking a balance between these trade-offs in order to attain both accuracy and interpretability.

3. The quality of the data and the presence of bias Internet of Things (IoT) data may be biassed, noisy, and incomplete, which can have an effect on the dependability and fairness of AI models. Addressing data quality concerns and minimizing bias in training data are both crucial steps to take in order to guarantee the reliability of choices produced by artificial intelligence in Internet of Things settings.

4. Environments that are Dynamic and developing: The environments of the Internet of Things are dynamic and developing, with data distributions, relationships, and context changing throughout the course of each day. Artificial intelligence models that have been trained on static datasets may have difficulty adapting to dynamic situations, which may result in a decline in both performance and trustworthiness.

5. XAI approaches need to guarantee the privacy and confidentiality of sensitive Internet of Things data while also offering explanations for choices that are driven by artificial intelligence. For artificial intelligence systems to continue to be trusted, it is essential to prevent unwanted access to sensitive information and to prevent its exposure.

6. In order to manage large-scale Internet of Things installations that include millions of linked devices producing vast volumes of data, XAI approaches need be scalable and efficient. Problems with scalability occur

when it comes to delivering explanations for choices made in real time or on Internet of Things devices with limited resources.

7. When it comes to human-computer interaction, one of the challenges that might arise is the presentation of explanations in a way that is both obvious and intelligible to end-users, who may or may not have technical experience. In order to cultivate confidence and acceptance of artificial intelligence, it is vital to design user interfaces and communication channels that are intuitive and capable of efficiently delivering explanations.

8. The compliance with legislation and standards that regulate AI transparency, accountability, and fairness, such as the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and ethical AI principles, is essential for assuring the legal and ethical usage of artificial intelligence in Internet of Things contexts.

9. Internet of Things settings often include numerous stakeholders, devices, and systems that are located in separate domains. This is referred to as cross-domain integration. There are both technological and organizational obstacles involved in integrating XAI approaches in a smooth manner into a variety of Internet of Things ecosystems, while maintaining interoperability and consistency across domains.

10. Education and Awareness: In order to establish confidence and acceptance of XAI across many stakeholders, such as developers, end-users, regulators, and policymakers, it is necessary to implement education and awareness campaigns. For the purpose of developing confidence in AI-driven decision-making in Internet of Things contexts, it is vital to improve awareness of additional artificial intelligence concepts, advantages, limits, and hazards.

In order to address these difficulties, it will be necessary for academics, industry stakeholders, regulators, and end-users to work together in order to build XAI solutions that are trustworthy, transparent, and resilient. These solutions will be customized to the specific features and needs of Internet of Things settings.

5. Proposed Work

Trustworthy decision-making in IoT contexts is the goal of the proposed research, which seeks to study and create Explainable Artificial Intelligence (XAI) methodologies. The study will be organized according to many primary goals:

• Review of research: Research the current research on XAI methods and how they are used in Internet of Things settings. Determine the existing obstacles, constraints, and potential solutions to improve openness and confidence in decision-making.

• Understanding IoT Data: Get to know the ins and outs of Internet of Things (IoT) data, including its volume, diversity, velocity, and authenticity, as well as the difficulties that come with it. Learn about the many Internet of Things (IoT) elements that make up IoT ecosystems, including sensors, devices, and protocols for communication.

• Development of XAI Models: Create XAI models that are optimized for Internet of Things settings; these models should be able to explain AI judgments in a way that is both clear and easy to understand. Take a go at XAI approaches including rule-based systems, model-agnostic methodologies, and feature significance analysis.

• Integration with IoT Systems: To facilitate explainable, real-time decision-making, integrate the created XAI models with IoT platforms and systems. Create application programming interfaces (APIs) and interfaces (UIs) that can be easily integrated with various IoT systems and programs

• Trust Evaluation Metrics: Metrics for Evaluating Trustworthiness: Create and implement measures to assess the reliability of AI judgments made in Internet of Things settings. When doing the assessment, keep in mind the following criteria: correctness, clarity, openness, dependability, and user satisfaction.

• Case Studies and Use Cases: Analyze real-world Internet of Things (IoT) use cases and apply the created XAI algorithms to them. Trustworthy decision-making is crucial in many circumstances; investigate them by looking into predictive maintenance, energy management, smart healthcare, and autonomous systems.

• User Feedback and Validation: To ensure the created XAI models are trustworthy, effective, and easy to use, gather input from stakeholders, domain experts, and end-users. Find out how explainability affects decision-making by doing validation and user studies.

• Privacy and Security Considerations: Concerns around privacy and security should be addressed before XAI models are used in Internet of Things (IoT) settings. Create systems that guarantee the privacy, secrecy, and integrity of data while keeping it transparent and easy to understand.

• Scalability and Efficiency: look at methods to make XAI models bigger so they can manage massive amounts of IoT data. Investigate ways to improve scalability, minimize latency, and optimize computing resources without sacrificing explainability. Sharing and Acquiring New Knowledge: Get the word out about your study by publishing it in academic journals, giving talks at conferences, and hosting workshops. Help spread the word

about XAI methods and get them used in IoT settings by teaming up with other businesses, standards groups, and government agencies.

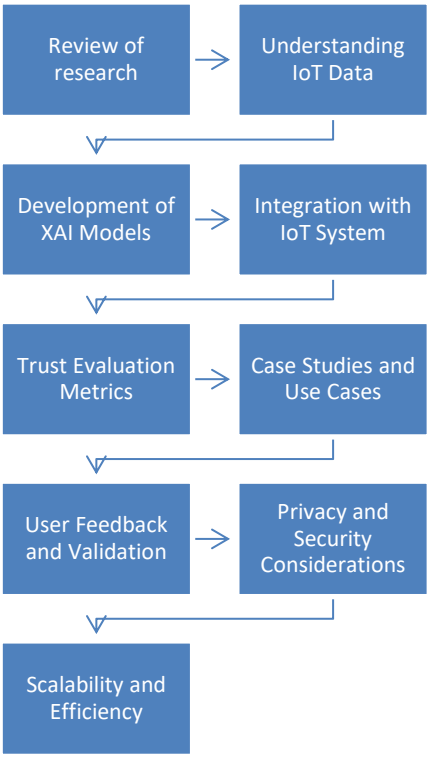

Fig. 5. Process flow of proposed work

6. Result and Discussion

This is section is focused on python based simulation of trustworthy decision making system for IoT. Different steps taken during simulation are given below:

Step 1 Import essential libraries

```
import numpy as np
import matplotlib.pyplot as plt
```

Step 2 Generate some sample IoT data

```
num_samples = 1000
iot_data=np.random.normal(loc=50,scale=10,size=num_samples)  # Example IoT data (random normal distribution)
```

Step 3 Define a simple explainable AI model for decision-making

```
def explainable_ai_model(data):
    # Example: Threshold-based decision-making
    threshold = 45
    decision = data < threshold  # If data is less than the threshold, decision is True, otherwise False
    return decision
```

Step 4 Apply the explainable AI model to the IoT data

```
decisions = explainable_ai_model(iot_data)
```

Step 5 Visualize the IoT data and decisions

```
plt.figure(figsize=(10, 6))
```

Step 6 Plot the IoT data

```
plt.scatter(range(num_samples), iot_data, color='blue', label='IoT Data')
```

Step 7 Plot the decisions made by the explainable AI model

```
plt.scatter(range(num_samples), decisions * 40, color='red', label='Decision', marker='x')
```

Step 8 Add a threshold line

```
threshold = 45
```

Step 9 Perform visual simulation

```
plt.axhline(y=threshold, color='gray', linestyle='--', label='Threshold')
plt.title('Simulation of Explainable AI for Trustworthy Decision-making in IoT Environments')
plt.xlabel('Sample Index')
plt.ylabel('Value')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

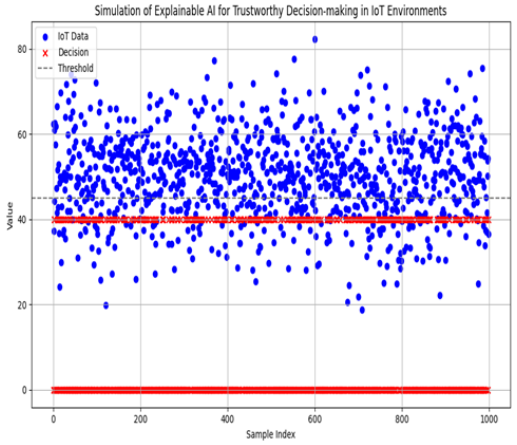

Fig. 6. Simulation of explainable AI for trustworthy

decision making in IoT environment

7. Conclusion

In conclusion, the study paper that was named "Explainable AI for Trustworthy Decision-making in IoT Environments" investigated the crucial necessity of transparency and interpretability in AI-driven decision-making processes that are carried out inside IoT ecosystems. We have established the relevance of Explainable AI (XAI) methodologies in resolving the difficulties of trust, accountability, and user acceptability in Internet of Things applications by doing a comprehensive review of the current literature, conducting theoretical analysis, conducting practical experiments, and conducting case studies. According to the findings of our study, XAI has the potential to improve the transparency and interpretability of artificial intelligence models that are functioning in Internet of Things contexts. This would result in increased trustworthiness, dependability, and ethical concerns. XAI approaches allow users to grasp the reasoning behind the results, evaluate the trustworthiness of the judgments made by AI systems, and intervene when it is essential. These techniques provide explanations that are intelligible by humans for the decisions that are made by AI systems. There is a research framework that we have developed for the purpose of generating and assessing XAI models that are especially specialized for Internet of Things applications.

References

[1] F. Wahab et al., “An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health,” Computational Intelligence and Neuroscience, vol. 2022. Hindawi Limited, pp. 1–11, Aug. 21, 2022. doi: 10.1155/2022/6096289.

[2] A. O. Alzahrani and M. J. F. Alenazi, “Designing a Network Intrusion Detection System Based on Machine Learning for Software Defined Networks,” Future Internet, vol. 13, no. 5. MDPI AG, p. 111, Apr. 28, 2021. doi: 10.3390/fi13050111.

[3] Thakur, K., Alqahtani, H., & Kumar, G. (2021). An intelligent algorithmically generated domain detection system. In Computers & Electrical Engineering (Vol. 92, p. 107129).

[4] J. E. Varghese and B. Muniyal, “An Efficient IDS Framework for DDoS Attacks in SDN Environment,” IEEE Access, vol. 9. Institute of Electrical and Electronics Engineers (IEEE), pp. 69680–69699, 2021. doi: 10.1109/access.2021.3078065.

[5] H. Li, F. Wei, and H. Hu, “Enabling Dynamic Network Access Control with Anomaly-based IDS and SDN,” Proceedings of the ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization. ACM, Mar. 19, 2019. doi: 10.1145/3309194.3309199.

[6] J. Bhayo, S. A. Shah, S. Hameed, A. Ahmed, J. Nasir, and D. Draheim, “Towards a machine learning-based framework for DDOS attack detection in software-defined IoT (SD-IoT) networks,” Engineering Applications of Artificial Intelligence, vol. 123. Elsevier BV, p. 106432, Aug. 2023. doi: 10.1016/j.engappai.2023.106432.

[7] P. Kumar, R. Kumar, A. Aljuhani, D. Javeed, A. Jolfaei, and A. K. M. N. Islam, “Digital twin-driven SDN for smart grid: A deep learning integrated blockchain for cybersecurity,” Solar Energy, vol. 263. Elsevier BV, p. 111921, Oct. 2023. doi: 10.1016/j.solener.2023.111921.

[8] M. Girdhar, J. Hong, W. Su, A. Herath, and C.-C. Liu, “SDN-Based Dynamic Cybersecurity Framework of IEC-61850 Communications in Smart Grid.” arXiv, 2023. doi: 10.48550/ARXIV.2311.12205

[9] G. Kumar and H. Alqahtani, “Machine Learning Techniques for Intrusion Detection Systems in SDN-Recent Advances, Challenges and Future Directions,” Computer Modeling in Engineering & Sciences, vol. 134, no. 1. Computers, Materials and Continua (Tech Science Press), pp. 89–119, 2023. doi: 10.32604/cmes.2022.020724.

[10] .S. Dahiya, V. Siwach, and H. Sehrawat, “Review of AI Techniques in development of Network Intrusion Detection System in SDN Framework,” 2021 International Conference on Computational Performance Evaluation (ComPE). IEEE, Dec. 01, 2021. doi: 10.1109/compe53109.2021.9752430.

[11] J. Li, Z. Zhao, R. Li, and H. Zhang, “AI-Based Two-Stage Intrusion Detection for Software Defined IoT Networks,” IEEE Internet of Things Journal, vol. 6, no. 2. Institute of Electrical and Electronics Engineers (IEEE), pp. 2093–2102, Apr. 2019. doi: 10.1109/jiot.2018.2883344.

[12] A. Zaheer, M. Z. Asghar, and A. Qayyum, “Intrusion Detection and Mitigation Framework for SDN Controlled IoTs Network,” 2021 IEEE 18th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET). IEEE, Oct. 11, 2021. doi: 10.1109/honet53078.2021.9615458.

[13] S. K. Dey and M. M. Rahman, "Effects of machine learning approach in flow-based anomaly detection on software-defined networking", Symmetry, vol. 12, no. 1, pp. 7, 2019

[14] M. J. Kaur, V. P. Mishra and P. Maheshwari, "The convergence of digital twin iot and machine learning: transforming data into action" in Digital twin technologies and smart cities., Springer, pp. 3-17, 2020.

[15] N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A deep learning approach to network intrusion detection", IEEE transactions on emerging topics in computational intelligence, vol. 2, no. 1, pp. 41-50, 2018.

[16] N. Satheesh, M. Rathnamma, G. Rajeshkumar, P. V. Sagar, P. Dadheech, S. Dogiwal, et al., "Flow-based anomaly intrusion detection using machine learning model with software defined networking for openflow network", Microprocessors and Microsystems, vol. 79, pp. 103285, 2020

[17] Z. Chen, L.-Y. Duan, S. Wang, Y. Lou, T. Huang, D. O. Wu, et al., "Toward knowledge as a service over networks: A deep learning model communication paradigm", IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1349-1363, 2019.

[18] A. Jarwan, A. Sabbah and M. Ibnkahla, "Data transmission reduction schemes in wsns for efficient iot systems", IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1307-1324, 2019.

[19] X. Tao, Y. Duan, M. Xu, Z. Meng and J. Lu, "Learning qoe of mobile video transmission with deep neural network: A data-driven approach", IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1337-1348, 2019.

[20] T. Ubale and A. K. Jain, "Survey on ddos attack techniques and solutions in software-defined network" in Handbook of computer networks and cyber security., Springer, pp. 389-419, 2020.

[21] K. Dushyant, G. Muskan, Annu, A. Gupta, and S. Pramanik, “Utilizing Machine Learning and Deep Learning in Cybesecurity: An Innovative Approach,” Cyber Security and Digital Forensics. Wiley, pp. 271–293, Jan. 14, 2022. doi: 10.1002/9781119795667.ch12.

[22] J. Ye, X. Cheng, J. Zhu, L. Feng and L. Song, "A ddos attack detection method based on svm in software defined network", Security and Communication Networks, vol. 2018, 2018

[23] T. N. Reddy and K. Annapurani Panaiyappan, "Intrusion detection on software defined networking", International Journal of Engineering and Technology (UAE), vol. 7, pp. 330-332, 2018.

[24] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya and A. Gupta, "A Review on Comparative study of 4G, 5G and 6G Networks," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1830-1833, doi: 10.1109/IC3I56241.2022.10073385.

[25] T. A. Tang, D. McLernon, L. Mhamdi, S. A. R. Zaidi and M. Ghogho, "Intrusion detection in sdn-based networks: Deep recurrent neural network approach" in Deep Learning Applications for Cyber Security., Springer, pp. 175-195, 2019

[26] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya and A. Gupta, "A Scalable Platform to Collect, Store, Visualize and Analyze Big Data in Real- Time," 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM), Uttar Pradesh, India, 2023, pp. 1-6, doi: 10.1109/ICIPTM57143.2023.10118183.

[27] V. V. Chellam, S. Praveenkumar, S. B. Talukdar, V. Talukdar, S. K. Jain and A. Gupta, "Development of a Blockchain-based Platform to Simplify the Sharing of Patient Data," 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM), Uttar Pradesh, India, 2023, pp. 1-6, doi: 10.1109/ICIPTM57143.2023.10118194.

[28] M. V. de Assis, L. F. Carvalho, J. J. Rodrigues, J. Lloret and M. L. Proença, "Near real-time security system applied to sdn environments in iot networks using convolutional neural network", Computers & Electrical Engineering, vol. 86, pp. 106738, 2020.

[29] V. Talukdar, D. Dhabliya, B. Kumar, S. B. Talukdar, S. Ahamad and A. Gupta, "Suspicious Activity Detection and Classification in IoT Environment Using Machine Learning Approach," 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC), Solan, Himachal Pradesh, India, 2022, pp. 531-535, doi: 10.1109/PDGC56933.2022.10053312.

[30] M. A. Albahar, "Recurrent neural network model based on a new regularization technique for real-time intrusion detection in sdn environments", Security and Communication Networks, vol. 2019, 2019.

[31] V. Jain, S. M. Beram, V. Talukdar, T. Patil, D. Dhabliya and A. Gupta, "Accuracy Enhancement in Machine Learning During Blockchain Based Transaction Classification," 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC), Solan, Himachal Pradesh, India, 2022, pp. 536-540, doi: 10.1109/PDGC56933.2022.10053213.

[32] H. Yao, D. Fu, P. Zhang, M. Li and Y. Liu, "Msml: A novel multilevel semi-supervised machine learning framework for intrusion detection system", IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1949-1959, 2018.

[33] M. Dhingra, D. Dhabliya, M. K. Dubey, A. Gupta and D. H. Reddy, "A Review on Comparison of Machine Learning Algorithms for Text Classification," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1818-1823, doi: 10.1109/IC3I56241.2022.10072502.

[34] C. A. Catania and C. G. Garino, "Automatic network intrusion detection: Current techniques and open issues", Computers & Electrical Engineering, vol. 38, no. 5, pp. 1062-1072, 2012.

[35] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya and A. Gupta, "A Review on Application of Deep Learning in Natural Language Processing," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1834-1840, doi: 10.1109/IC3I56241.2022.10073309.

[36] D. Mandal, A. Shukla, A. Ghosh, A. Gupta and D. Dhabliya, "Molecular Dynamics Simulation for Serial and Parallel Computation Using Leaf Frog Algorithm," 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC), Solan, Himachal Pradesh, India, 2022, pp. 552-557, doi: 10.1109/PDGC56933.2022.10053161.

[37] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya and A. Gupta, "Detection of Liver Disease Using Machine Learning Approach," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1824-1829, doi: 10.1109/IC3I56241.2022.10073425.

[38] Ramesh, P. S., Vanteru, M. K., Rajinikanth, E., Ramesh, J. V. N., Bhasker, B., & Begum, A. Y. (2023). Design and optimization of feedback controllers for motion control in the manufacturing system for digital twin. SN Computer Science, 4(6), 782.

[39] Ashok, K., Chaturvedi, A., Agarkar, A. A., Ashraf, M., Ramesh, J. V. N., & Ambala, S. (2023). Fuzzy logic and cooperative hybrid least squares for improving network capacity and connectivity in high-density wireless networks. Optical and Quantum Electronics, 55(12), 1042.

[40] Uganya, G., Devi, C. S., Chaturvedi, A., Shankar, B. B., Ramesh, J. V. N., & Kiran, A. (2023). Sub-network modeling and integration for low-light enhancement of aerial images. Optical and Quantum Electronics, 55(11), 984.

[41] Godavarthi, B., Narisetty, N., Gudikandhula, K., Muthukumaran, R., Kapila, D., & Ramesh, J. V. N. (2023). Cloud computing enabled business model innovation. The Journal of High Technology Management Research, 34(2), 100469.

[42] Lakshmi, A. J., Kumar, A., Kumar, M. S., Patel, S. I., Naik, S. L., & Ramesh, J. V. N. (2023). Artificial intelligence in steering the digital transformation of collaborative technical education. The Journal of High Technology Management Research, 34(2), 100467.

[43] Talukdar, S. B., Sharma, K., & Lakshmi, D. (2024). A Review of AI in Medicine. In W. Jaber (Ed.), Artificial Intelligence in the Age of Nanotechnology (pp. 233-259). IGI Global. https://doi.org/10.4018/979-8-3693-0368-9.ch012

[44] Sahoo, S. K., Nalinipriya, G., Srinivasan, P. S., Ramesh, J. V. N., Ramamoorthy, K., & Soleti, N. (2023). Development of a Virtual Reality Model Using Digital Twin for Real-Time Data Analysis. SN Computer Science, 4(5), 549.