A Robust Data Security Model Development to Maintain the Integrity of Information using Artificial Intelligence and Cyber Laws

G.Muthu Kumar,
Department of Artificial Intelligence and Data Science,
S.A. Engineering College(Autonomous),
Chennai, Tamil Nadu, India
muthukumarg79.cse@gmail.com

D. Hemanand,
Department of Computer Science and Engineering,
S.A. Engineering College (Autonomous),
Chennai, Tamil Nadu, India
hemanand1982@gmail.com

2024 International Conference on Intelligent Systems for Cybersecurity (ISCS) | 979-8-3503-7523-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCS61804.2024.10581024

Abstract—The field of artificial intelligence (AI) has seen significant advancements in recent years. These days, artificial intelligence (AI) tools are being utilized by organizations in both the public and commercial sectors all over the world. Individuals, organizations, and society as a whole will reap broad and significant advantages as a result of the capabilities of artificial intelligence (AI) both today and in the near future. Nevertheless, these very same technical advancements give rise to significant concerns, such as the question of how to ensure that artificial intelligence technology is built and implemented in a manner that is in accordance with the applicable data privacy laws and standards. The fast development of artificial intelligence presents substantial hurdles in terms of protecting customers' privacy and the confidentiality of their data. The purpose of this essay is to suggest an all-encompassing strategy for the development of a framework to solve these concerns. First, an overview of prior research on security and privacy in artificial intelligence is presented, with an emphasis on both the progress that has been made and the limits that still remain. In the same vein, open research topics and gaps that need to be addressed in order to improve existing frameworks are recognized. Regarding the development of the framework, the topic of data protection in artificial intelligence is discussed. This includes elaborating on the significance of protecting the data that is utilized in artificial intelligence models, as well as elaborating on the policies and practices that are in place to ensure the data's safety and the methods that are utilized to maintain the data's integrity. Additionally, the security of artificial intelligence is investigated, which includes an analysis of the vulnerabilities and dangers that are present in artificial intelligence systems, as well as the presentation of instances of potential assaults and malevolent manipulations, as well as security frameworks that are designed to prevent these risks. The model that has been suggested is known as the Artificial Intelligence based Secured Data Handler (AISDH), and in order to test the effectiveness of the proposed scheme, it is cross-validated with the traditional Blowfish Algorithm (BA).

***Index Terms*—Data Security, Data Integrity, Artificial Intelligence, AI, Cyber Laws, Secured Data Handler, AISDH, BA, Blowfish Algorithm**

I. INTRODUCTION

In the past few years, there has been a rapid development in artificial intelligence (AI) [1]. These days, artificial intelligence (AI) tools are being utilized by organizations in both the public and private sectors all over the world. Individuals, organizations, and society as a whole will reap widespread and significant benefits as a result of the capabilities of artificial intelligence (AI) both today as well as in the near future [2, 3].

On the other hand, these very same technical advancements bring up significant concerns, such as those regarding the conflict that exists between artificial intelligence and data protection legislation [4]. Furthermore, in light of the technical realities of the 21st century, we are in a position to review the efficacy of the data protection rules that are now in place, which are both an opportunity as well as an obligation. We require data protection rules and policies that not only successfully protect privacy in an era of artificial intelligence and the vast data on which it frequently depends, but also do not impose additional hurdles for the development of these breakthrough technologies in the future. We cannot make a decision among the already routine advantages of artificial intelligence and the safety of personal data; rather, we must develop practical solutions to ensure both [5]. This is something that has been emphasized in multiple reports by the government and regulators. There has been a significant increase in the analytical capacity of modern computers, which is posing an increasing challenge to the laws and conventions that govern data protection [6]. The term "artificial intelligence" is frequently used to refer to these advancements. This word is used to define the overarching objective of enabling "computer systems to perform tasks which normally require human intelligence, including visual perception, speech recognition, decision-making, and translation among languages."1. This one term incorporates a vast range of technical advancements, each of which may provide a unique set of difficulties to the data protection regulations that are already in place [7].

This is owing to advancements in processing power as well as the massive amounts of digital data that are available for study [8]. The foundation of other tools, a few of which are described below, is machine learning, which is widely used today to perform a wide variety of tasks. These tasks include the detection of fraud, the filtering of email, the detection of cyber-threats including network intruders or malicious insiders, the recommendation of books or films, and the provision of other services based on previous or unusual behavior. Deep learning frequently makes use of larger data sets in order to generate larger models and train those models in the most effective manner [9].

II. RELATED STUDY

With the rise of new technologies like distributed computing, utility computing, and grid computing, the cloud computing movement is picking up speed [11]. One of the most important trends in modern computing, cloud computing has the ability to give accessible, elastic, and powerful resources over the internet at a reasonable price. Virtual computers access data through a network resource and the cloud provides the ability to store and access data remotely. In addition, the fourth industrial revolution is spearheaded by cloud computing. When using Dropbox, other Google services, or Microsoft Office 365, everyone is always interacting with the cloud. Although there are numerous benefits to such an environment, there are also security concerns, including data availability, privacy, and security; access control; cyber-attacks; and performance and reliability challenges. Providers of cloud services have an obligation to their customers to take reasonable precautions to protect their data in a way that is both private and accessible at all times. However, consumers have not been well served by cloud service providers in terms of dependable and secure services. Cloud computing is being enhanced by a technology called blockchain. In order to address security issues, this groundbreaking technique provides data integrity features that are convincing. The difficulties associated with cloud privacy and security is examined in depth in this study. In this case study, we focus on smart campus security to highlight the significance of security challenges; hopefully, this will inspire future studies to investigate cloud computing security [11].

Everyone can benefit from enhanced physical health by getting enough quality sleep [12]. In order to maintain a healthy body and mind, it is essential to evaluate the sleep cycle, since irregular sleep patterns may be a sign of the illness leading to persistent sadness. The provision of high-quality health care facilities through the use of cost-effective instruments and technology has been one of many issues investigated in the context of globalization, alongside the expansion of facilities. Creating a reliable and affordable system to track patients' sleep quality is one of the stated goals of IoT technology development. There are a number of alternative systems that provide this function, but they are prohibitively expensive and complex to set up. An innovative approach to track and analyze sleep patterns based on environmental factors is proposed as a solution to this problem in this study. The suggested method is robust enough to accurately track patients' sleep using COS sensors and make predictions based on the data provided by the random forest model. The study's findings demonstrate that the recommended method has a 95% success rate. Through the validation of manual results, which yield the lowest error rate, this method is tested using the patient's sleep data. The use of the Internet of Things (IoT) to create a smart system to monitor sleep quality on a variable number of people at a low cost is highlighted in this work [12].

The term "cloud computing" refers to an emerging paradigm in the IT industry that uses the Internet to provide computer resources to users on demand [13]. Small and medium-sized enterprises (SMEs) often choose cloud computing over building their own IT infrastructure due to the significantly reduced cost of the services offered. Computing and data storage are outsourced to third-party companies under the cloud computing model. Apps and customer data are not entirely under the control of the customer. Customers' principal concern when contemplating cloud services is the security of their data, which carries with it the additional burden of security concerns. This study delves into different cloud computing service deployment patterns and the Internet of Things (IoT), identifies the data security challenges that come along with each, and proposes a metric-based approach to evaluate the security of cloud services [13].

A critical issue for the network community is managing the massive IoT infrastructure [14]. Many believe that Software Defined Networking (SDN) is the best way to oversee the Internet of Things (IoT) because of its centralized network management features. In order to localize cloud demands, edge computing provides cloud resources close to the IoT. Thus, a framework that combines SDN, the Internet of Things (IoT), and edge computing may be built to establish a powerful SDIoT-Edge architecture. This architecture can effectively orchestrate cloud services and make flexible use of IoT devices with constrained resources. In addition to the widespread use of IoT, attackers can exploit the OpenFlow channel through Distributed Denial of Service (DDoS) attacks due to the vulnerabilities in this less secure architecture. A major challenge with SDIoT-Edge is providing security for the OpenFlow channel, since DDoS attacks on this channel might impair the entire network. In order to address the security issues with this design, we offer a framework that we term SDIoT-Edge Security (SIESec). For effective defence against DDoS attacks, the SIESec prototype uses a categorization technique based on machine learning, integrates blacklists, and filters contextual network flows. Using the Mininet network simulator and the Floodlight controller, we run comprehensive simulations. Our findings demonstrate that SIESec offers robust protection against distributed denial of service (DDoS) attacks over OpenFlow channels while imposing minimal network burden [14].

Physical characteristics of wounds, such as their width, form, color, etc., have been the primary focus of clinical wound assessment studies [15]. However, the most important component influencing the healing process is the wound's appearance. But there are other aspects outside the wound's appearance that play a role in the healing process. An example of a factor that might have an effect on healing is the wound's internal and exterior environment. The widespread use of the Internet of Things (IoT) in nearly every industry and household over the past decade has contributed significantly to its meteoric rise in popularity. Thus, in this study, we present an Internet of Things (IoT) based intelligent wound assessment system that uses decision tree entropy and information gain statistics to classify assessment results as good, satisfactory, or alarming, reflecting the status of the wound assessment. Using MATLAB, we built a decision tree and used the ID3 algorithm to choose the optimum feature to split the tree based on entropy and information gain. Decision tree

performance and training accuracy were both enhanced by an effective feature split [15].

III. METHODOLOGY

When it comes to physical protection, the application of artificial intelligence in information defense creates new obstacles. Even if it is essential to make use of artificial intelligence technologies in order to identify and combat malware threats, it is also possible for cyber attackers to use AI tools in order to carry out progressive behavior attacks. In part, this is due to the fact that access to advanced artificial intelligence technology, in addition to machine learning methodologies, is increasing at a time when the costs of creating and deploying these breakthroughs are decreasing. As a result, malicious software can be developed by computer attackers in a more expedient and cost-effective manner, allowing them to create increasingly sophisticated and effective applications. A definition of hostile artificial intelligence provided by Accenture describes it as a situation that "causes machine learning systems to misunderstand inputs into the structure and react in a way that is beneficial to the intruder." In essence, this happens when neural networks in an artificial intelligence programme are tricked into incorrectly identifying or erroneously depicting artifacts as a result of modified inputs that have been purposefully altered. When it comes to building privacy remedies for big data, one of the challenges that arise is the sheer volume of data sources. When there are several data sources, it is possible for there to be a variety of settings in which different stakeholders can have differing degrees of possession of the processed data.

A single data owner has the ability to encrypt their data using their own keying material and to apply data analytics to the encrypted data neither locally or offloaded to an external platform. This means that the data owner has the ability to encrypt their data. On the other hand, in today's world, data are gathered by a wide variety of apps and services, which are utilized by a variety of businesses. The data in question are frequently subjected to in-depth analysis in order to derive knowledge that is beneficial to the companies in question. The following figure Fig.1 shows the flow diagram for the construction of a structure that assures the privacy and security of AI.

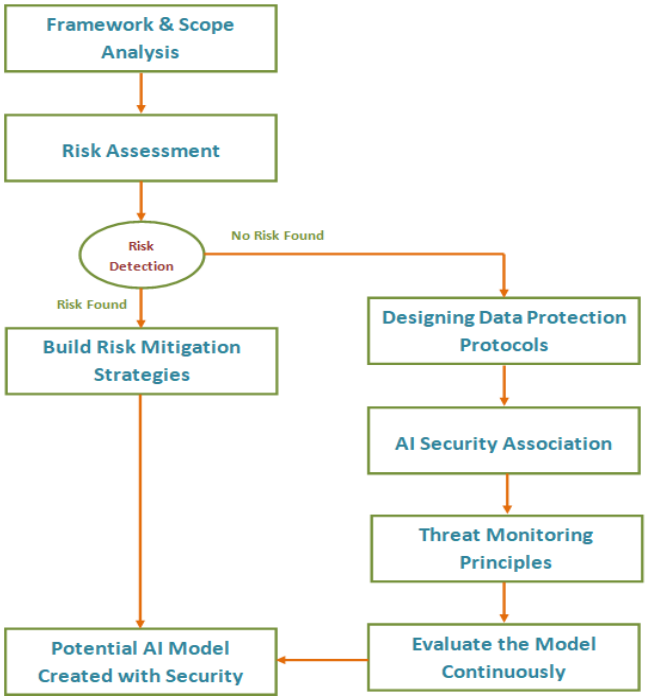

Fig.1 Flow Diagram for the Construction of a Structure that Assures the Privacy and Security of AI

Artificial intelligence (AI) comes with both advantages and disadvantages. Despite the fact that technology is now helping workers in a variety of professions, it is conceivable that AI will lessen the necessity for workers in other professions. If the data used in the development of artificial intelligence only represents some segments of the population or reflects current societal bias, it may add bias and new kinds of discrimination. This is especially true if the data is used to develop AI. The conventional ideas of urban and residential design, which often involve substantial areas that are devoted to parking lots and garages, are expected to be challenged by artificial intelligence. Additionally, artificial intelligence may give rise to significant antitrust concerns, particularly in the event that the data required for its development is centralized in the hands of a small number of firms. When compared to the way that is now in use, which is known as Blowfish Algorithm (BA), the proposed method in question, which is dubbed Artificial Intelligence based Secured Data Handler (AISDH), considerably improves the capabilities of systems in terms of cyber-security. The following figure Fig.2 describes the process of implementing the framework for AI security and privacy.

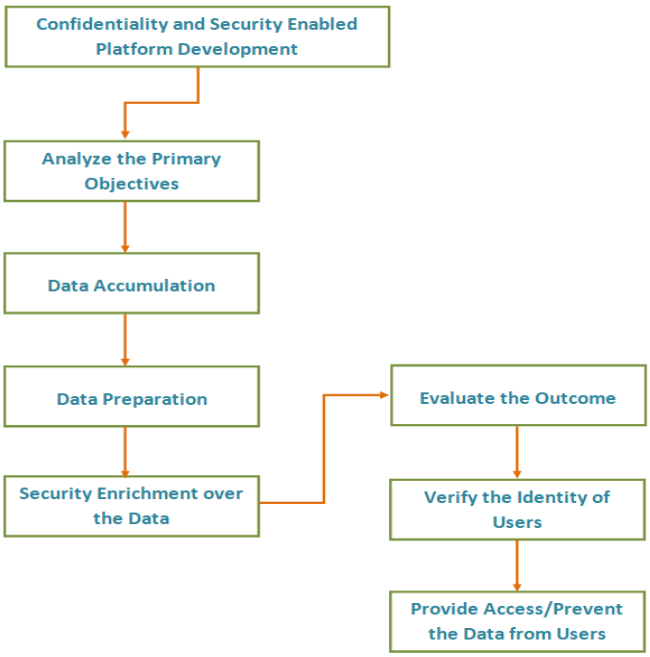

Fig.2 Process of Implementing the Framework for AI Security and Privacy

Deep learning has become increasingly important in computer security due to the rapid mobility of both its implementation in hardware and its use in graphical chipsets. New neural net innovations include third-generation cognitive nets, which speed up machine learning by better simulating artificial neurons and opening up more potential uses.

Agents with intelligence are programmes that run on computers. They could be capable of planning, organizing, and assessing. In fact, there is a concept of software agents within the software development community. These are objects that actively employ the agent-like language used for networking. Subjects, in contrast to agents, can be non-active and unable to communicate in order to be understood. Effective protection of cooperating agents from DDoS attacks was demonstrated in simulations using intelligent agents. A "cyber police force" made up of mobile smart officers could be a reality if all contractual and legal hurdles have been cleared. Included in this would be gear that allows cyber professionals to be mobile and connected, but which adversaries cannot access. Works well with Internet service providers (ISPs). Nevertheless, the search can be somewhat more effective if additional experience can be utilized to guide it. A smart system's total success is frequently dependent on the quality of its quest, which is present in nearly all of them. In order to address specific search challenges, a wide range of search strategies are developed. There are a lot of known AI search strategies that are utilized in a lot of applications, but they aren't employed as AI. One issue is that search isn't considered an AI function because it's part of the application stack. This is the main way that optimal security issues are addressed by dynamic analysis programming. Learning fortifies the data structure by extending, rearranging, or improving the existing body of knowledge. There has been a lot of research into it because it is one of the most crucial areas of AI. With the use of computation, we may expand our horizons, develop our skills, and find better ways to organize what we already know. This calls for computer learning. From the most fundamental kind of learning-parametric learning, or understanding the meaning of such quantities to the most advanced forms of abstract teaching concept, grammar, usability, and behavioral learning, there is a vast spectrum of difficulties.

IV. Results and Discussions

Machine learning is an AI technique that eliminates the need for explicit programming and allows computers to learn from their own experiences and the data given to them. Many areas, including data security, can benefit from this rapidly developing technology. These days, data security is a hot subject because of all the dangers that come with dealing with massive amounts of data. When it comes to data security, machine learning technologies can help identify potential risks. This paper's objective, then, is to investigate how machine learning contributes to data protection. Machine learning will make it feasible to secure massive volumes of data. We are living in the data-rich era of the internet and rapidly expanding technology across all industries. The correct protection of data is of the utmost importance; hence, machine learning is one area that might be merged with data security to facilitate data management. Additionally, data security will be achieved by appropriate data handling. However, data security using machine learning is still in its infancy and is only beginning to develop in tandem with related research and best practices.

Data security can benefit from some of the machine learning methods that have recently been developed. Data security may be achieved by the application of machine learning techniques such as supervised learning, unsupervised learning, and reinforcement learning. It is also possible to utilize other algorithms. When it comes to protecting sensitive information, machine learning has several potential uses. Secure data storage will benefit from all of these. Data security is aided by machine learning, a prominent technology. Artificial Intelligence based Secured Data Handler (AISDH) is a machine learning algorithm and approach that can aid in data security; to evaluate the suggested scheme's efficacy, it is cross-validated with the standard Blowfish Algorithm (BA). Appropriate data classification is facilitated by the suggested classification methods. It aids data security by classifying fraudulent and accurate or secure data. As an example, we may learn about fraudulent emails since machine learning models categorize spam emails differently. Therefore, we shall take precautions to protect our data by avoiding providing any information in response to any unsolicited communication, especially if it appears to be spam. Classification of user behaviour patterns allows for user authentication, which in turn helps in identifying both authorized and unauthorized users of data or systems. Malware detection, questionable data access, and how to prevent it can all be taught to classification algorithms. Data and security issues, such as fraudulent users, may be better anticipated using this information. Face recognition is one of the most well-known uses of machine learning. To prevent unauthorized individuals or entities from gaining access to systems or

data, facial recognition technology may be utilized as a login credential for any system. System administrators can determine who should have access to sensitive information by implementing biometric authentication, which compares a user's face to a database of previously stored images. Using facial recognition technologies, it is possible to detect and track people who are unauthorized accessing data or systems.

This approach may be utilized to verify the identification of individuals in many settings, such as business enterprises, government offices, educational institutions, and more. It can authorize users to make transactions and prevent fraudulent activities in the financial industry. Data security is achieved by the widespread usage of facial recognition technology, which aids in authenticating users. One use of machine learning is human identity identification, which helps identify people by seeing patterns in their individual characteristics. Users can be granted access to data and systems using this method. Online marketplaces are becoming increasingly popular these days. In order to safeguard the system and its data, we must first identify the specific data security issue that needs fixing, and then think about the best machine learning approach to take. Overall, data security may benefit from machine learning if its many ideas are adequately understood.

Fig. 3 shows the results of a cross-validation test between the suggested model, AI-based Secured Data Handler (AISDH), and the traditional Blowfish Algorithm (BA), which was used to determine the model's classification accuracy. What follows is a descriptive table of the same information, Table-1.

TABLE-1: CLASSIFICATION ACCURACY

Data Size (bits)	BA (%)	AISDH (%)
1500	91.26	98.35
2000	91.32	98.39
2500	91.36	98.36
3000	91.39	98.64
3500	91.42	98.29
4000	91.47	98.64
4500	91.51	98.59
5000	91.55	98.63
5500	91.58	98.67
6000	91.62	98.71
6500	91.66	98.75
7000	91.70	98.79

Fig.3 Classification Accuracy

The prediction accuracy of the proposed model, AISDH, is shown in Fig-4. To evaluate the accuracy of the model, it is cross-validated with the traditional BA. What follows is a descriptive table of the same information, Table-2.

TABLE-2: PREDICTION ACCURACY

Data Size (bits)	BA (%)	AISDH (%)
1500	86.54	96.38
2000	86.39	96.37
2500	86.52	96.42
3000	86.47	96.49
3500	86.39	96.51
4000	86.40	96.48
4500	86.37	96.82
5000	86.35	96.73
5500	86.33	96.79
6000	86.31	96.85
6500	86.29	96.90
7000	86.26	96.96

Fig.4 Prediction Accuracy

Fig. 6 shows the data security ratio of the AISDH model, which is cross-validated with the standard BA to measure the degree to which the proposed system improves data security. Table 3 provides a descriptive representation of the same.

TABLE-3: DATA SECURITY ANALYSIS

Data Size (bits)	BA (%)	AISDH (%)
1500	91.83	98.76
2000	91.76	98.64
2500	91.82	98.52
3000	91.76	98.79
3500	91.83	98.67
4000	91.79	98.67
4500	90.64	98.66
5000	90.39	98.65
5500	90.52	98.73
6000	90.40	98.82
6500	90.20	98.78
7000	90.01	98.76

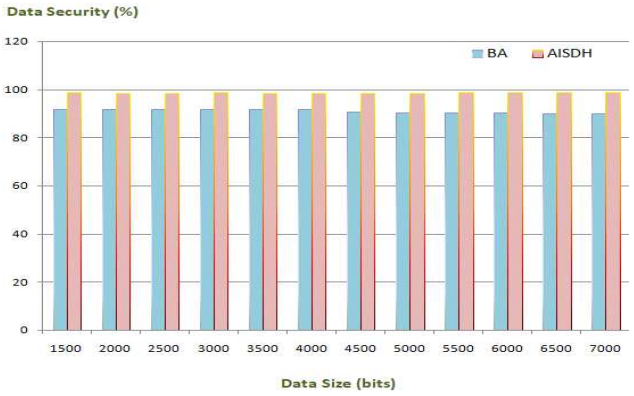

Fig.5 Data Security Analysis

Figure 6 shows the data integrity ratio of the AISDH model, which is cross-validated with the traditional BA to assess the data integrity ratio of the plan. What follows is a descriptive table of the same information, Table 4.

TABLE-4: DATA INTEGRITY ANALYSIS

Data Size (bits)	BA (%)	AISDH (%)
1500	90.26	95.64
2000	89.45	95.39
2500	90.01	95.48
3000	89.63	95.62
3500	89.51	95.54
4000	89.37	95.59
4500	89.24	95.58
5000	89.11	95.59
5500	88.97	95.60
6000	88.84	95.61
6500	88.71	95.62
7000	88.57	95.63

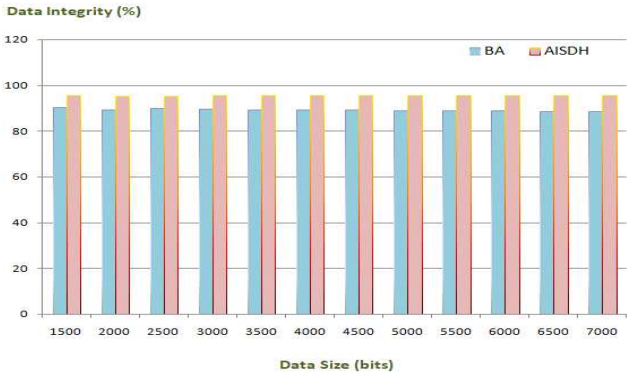

Fig.6 Data Integrity Analysis

V. CONCLUSION

For the purpose of ensuring the safety of data, we may make use of machine learning, which is one of the technologies that are experiencing rapid growth. Data may be protected through the use of machine learning by recognizing and anticipating potential dangers to the data. For the purpose of ensuring the safety of data in a variety of domains, a variety of machine learning algorithms and methods can be employed. Anomalies in data are discovered by machine learning through the analysis of patterns in data. Additionally, machine learning generates security warnings, which contributes to the prevention of data loss. Because machine learning models are able to process massive volumes of data and learn from their previous experiences and data, they are able to provide a variety of various ways to safeguard the data. Through these means, we are able to

utilize the ideas of machine learning in the context of data security if we have a thorough grasp of them.

REFERENCES

[1] Wang Shaohui, Pan Xiaoxiao, Wang Zhiwei, Xiao Fu and Wang Ruchuan, "Identity-based cloud storage data integrity verification scheme supporting forward security [J]", Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), vol. 39, no. 01, pp. 79-86, 2019.

[2] V. M et al, "Deep Reinforcement Learning for Energy Efficient Routing and Throughput Maximization in Various Networks," (I-SMAC), Dharan, Nepal, 2022, pp. 204-210, doi: 10.1109/I-SMAC55078.2022.9987395.

[3] Huang Xuerong and Guo Rongzuo, "Data integrity audit protocol based on identity proxy off-line signature [J]", COMPUTER ENGINEERING AND DESIGN, vol. 41, no. 06, pp. 1553-1561, 2020.

[4] T. M. et al, "Effective control of non linear parameters using artificial intelligence," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Dec. 2016, Published, doi: 10.1109/iccic.2016.7919659.

[5] Wang Yanling, "Cloud data integrity verification algorithm for accounting informationization in sharing mode [J]", Modern Electronics Technique, vol. 42, no. 05, pp. 87-89, 2019.

[6] S. S. Abdul-Jabbar, A. Aldujaili, S. G. Mohammed and H. S. Saeed, "Integrity and security in cloud computing environment: a review", Journal of Southwest Jiaotong University, vol. 55, no. 1, 2020.

[7] R. Umamaheswari, D. Lakshmi, V. S. Pandi, B. Geetha, S. Sumithra and R. P. Y, "An Advanced Deep Learning Approach for Primary Osteoporosis Prediction Using Radiographs with Clinical Covariates," *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, 2023, pp. 788-793, doi: 10.1109/ICECA58529.2023.10395285.

[8] Liu Xueyan, Lu Tingting and Yang Xiaotao, "Verifiable Attribute-based Keyword Search Scheme with Privacy Preservation [J]", Journal of Electronics & Infonnation Technology, vol. 43, no. 01, pp. 218-225, 2021.

[9] R. K. Sadavarte, G. D. Kurundkar and D. Parbhani, "Data security and integrity in cloud computing: Threats and Solutions", 2020.

[10] B. Sarwar, I. S. Bajwa, N. Jamil, S. Ramzan and N. Sarwar, "An intelligent fire warning application using IoT and an adaptive neurofuzzy inference system", Sensors, vol. 19, no. 14, pp. 3150, 2019.

[11] J. J. J, B. V. Prabha, B. Yasotha, J. J, C. Senthilkumar and V. S. Pandi, "Enhancing Residential Security with AI-Powered Intrusion Detection Systems," *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*, Theni, India, 2023, pp. 1510-1515, doi: 10.1109/ICSCNA58489.2023.10370042.

[12] K. Saleem, I. S. Bajwa, N. Sarwar, W. Anwar and A. Ashraf, "IoT healthcare: design of smart and cost-effective sleep quality monitoring system", Journal of Sensors, 2020.

[13] D. K. Saini, K. Kumar and P. Gupta, "Security Issues in IoT and CC Service Models with Suggested Solutions", Security and Communication Networks, vol. 2022, 2022.

[14] W. Rafique, M. Khan, N. Sarwar and W. Dou, "A security framework to protect edge-supported software-defined Internet of Things infrastructure", In International Conference on Collaborative Computing: Networking Applications and Worksharing, pp. 71-88, 2019.

[15] H. Sattar, I. S. Bajwa, R. U. Amin, N. Sarwar, N. Jamil, M. A. Malik, et al., "An IoT-based intelligent wound monitoring system", IEEE Access, vol. 7, pp. 144500-144515, 2019.