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

Cloud Data Security by Hybrid Machine Learning and Cryptosystem Approach

Aditya Kumar Shukla[1], Ashish Sharma[2]

Submitted: 15/07/2023 **Revised**: 01/09/2023 **Accepted**: 22/09/2023

Abstract: Cloud computing is a key platform for achieving secure data exchange and offering increased data storage with no additional burden. Cloud environments provide these benefits. In order to achieve greater levels of security inside a cloud environment, one of the most important aspects of cloud service provisioning is the prevention of data breaches caused by unauthorized access to personal information. There are still many security measures that have not been refined to an acceptable level. To address this issue, we provide an unique machine learning programme that use cryptographic techniques to protect data both while it is being stored in the cloud and while it is being accessed. According to the results of the study, the suggested approach has attained a sensitivity of 87.5%, an F1-score of 93.5%, and a specificity of 97.5%.

Keywords: *Cloud computing, ML in cloud, Cryptography, Security in cloud, Data privacy*

1. Introduction

The term "cloud computing" refers to a delivery model in which resources including servers, storage, networks, development platforms, and applications are made available to users on-demand and charged on a per-use basis [1]. Cloud computing is expected to become an integral part of the enterprise IT infrastructure of the future. Storage and transmission of data between server and client occurs in data centres, much as in the cloud. Users may keep their data in the cloud and take use of high-quality cloud apps without having to worry about maintaining their own servers or installing new programmes [2][3]. The field of cloud computing is expanding quickly in both the business and academic worlds. By facilitating the access and sharing of computer resources as virtual resources, cloud computing provides a significant benefit in corporate data centres by allowing the data centres to function like the internet. The key features that make cloud computing more effective include reduced storage costs, virtualization, grid computing technologies, service-oriented software, power efficiency, and the management of massive facilities [4].

To simplify and safeguard everything in this industry, machine learning plays a crucial role using a variety of current algorithms. Currently, all cloud tiers may benefit from using cloud security. The cloud provider takes complete responsibility for the client's data and makes it possible to store massive amounts of data in a virtual place. Cloud databases were developed in response to the ever-increasing need for more and better data storage space as well as improved internet infrastructure. As cloud computing becomes more widely used, so does the need for cloud-based database services. Multiple agreed computers are used in this type of attack to target network servers and resources, resulting in packet flooding, corrupted messages, connection requests, and denial of service to the legitimate consumer [27]. Protecting customers' private information is a top priority for every cloud service. Safeguarding user data is the first consideration while storing files in the cloud. The most difficult aspect of cloud computing's data management is ensuring the privacy, confidentiality, integrity, and security of client data. Figure 1 demonstrates the many data security layers made available by the cloud.

[1,2]*Department of Computer Engineering & Applications, GLA University, NH#2, Delhi Mathura Highway, Post Ajhai, Mathura (UP) India*
1. uraditya@gmail.com, 2.ashishs.sharma@gla.ac.in

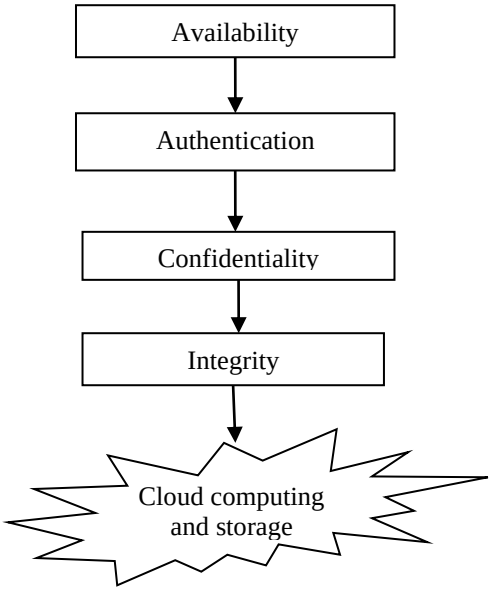

Fig 1. Different layers of data security [5].

1.1 Security on clouds

The security of the data stored in the cloud is an important concern for the whole corporation. There are a variety of strategies available to protect the information stored on various servers. Data integration, confidentiality, dependability, scalability, and authentication are among the most important main points regarding the security of data stored in the cloud. Another important big point is the privacy of data stored in the cloud [6]. The protection of data, information systems, and networks against infiltration and intrusions necessitates the implementation of defensive measures [28].

1.1.1 Cloud security and challenges

The integrity of sensitive data is a key factor that has to be evaluated. A lack of confidence prevents many businesses from buying a vendor's guarantee of data protection. They worry that the data order of purchasers may be disrupted in the case of a dispute. One of the hazards is a hacker or cracker who is responsible for defending a network against intrusions. IDS uncovered three distinct types of intruders[26]. Many models don't show the actual amassing area, which raises safety problems for businesses [7]. In contemporary models, this private data is protected by firewalls across server farms (as declared by attempts). There is a risk that cloud service providers won't be vigilant about data security, which puts the onus on the businesses using them [8].

1.1.2 Administrative and Compliance Restrictions

Private information and other sensitive information cannot be disseminated outside of the state or nation in various European countries due to government limitations. If they want to operate legally, cloud providers will need to set up shop in the form of a server ranch or other restricted location inside the nation. It is very challenging for cloud service providers to maintain such a structure, and it may not be possible in the long run [9].

1.1.3 Data Recovery and availability

It's common knowledge that any enterprise software has a set of Service level guidelines that must be strictly adhered to. Utilitarian groups often play a central role in the development of strategic plans and the management of day-to-day operations. Utilitarian assemblies are helpful in developing environments .

- Recovery of data
- Replication of data
- Monitoring of data
- Management of data

1.2 Symmetric view to secure the cloud-Machine learning.

The terms "plan affirmation" and "computational learning" are sometimes used interchangeably when referring to "machine learning intelligence,"

which refers to both a system and a group of advances that use AI thinking. It's an old idea, first articulated in 1959 as giving computers the ability to learn without having to reproduce. Once upon a time, artificial intelligence was out of reach for the majority of businesses' budgets [29]. However, nowadays, public cloud providers are able to provide an AI service, which makes this advancement financially feasible [11]. I should bring you up to speed on artificial intelligence and its significance to the ongoing demands for IT improvement and delivery, particularly for those who operate in a system that is based on cloud computing. In the context of industry 4.0, additional applications and methodologies may be useful [12][13]. Through access to managed services offering a broad range of hardware heterogeneity and great horizontal scalability, MLaaS may assist minimise the entry barrier to the application of machine learning and deep learning. MLaaS is now being offered by a number of the world's most prominent companies, including Google, Microsoft, and Amazon. Customers may reduce their computing cost by using MLaaS services, which let them upload their model and data for training on the cloud. Transfer learning is another option. In this method, a model that has already been trained is modified so that it may be used for a different task .

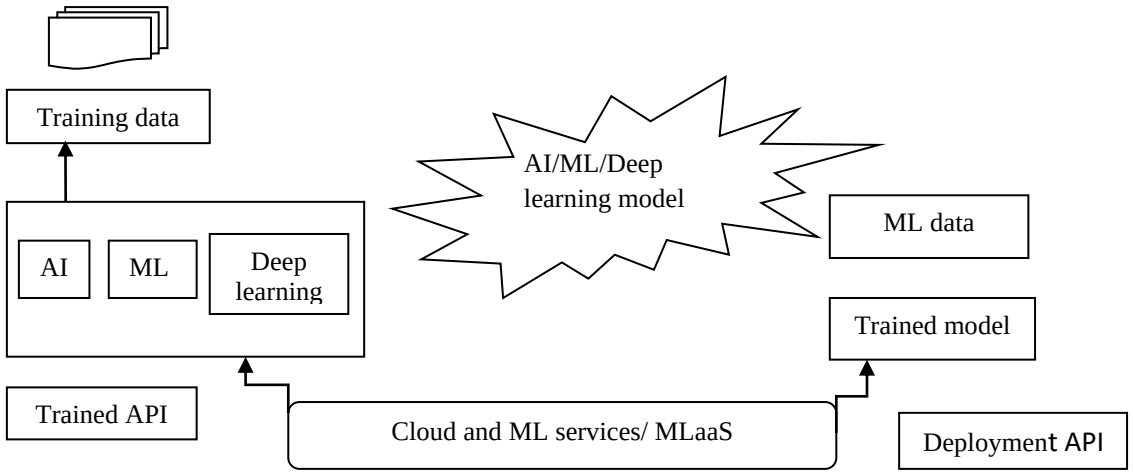

Fig 2. Role of ML in cloud computing security.

1.3 Cryptography in cloud computing

The process of encrypting data and messages by converting them into formats that are not easily decipherable is known as cryptography. The most important contribution that crypto cloud computing can make is to make it possible to share digital resources while maintaining a high level of data confidentiality and privacy. The approval of the digital signatures of each component is another method for providing security management for cloud computing. It is essential to pay extra attention to the protection of sensitive data in the event that the data are corrupted or changed as a result of unauthorised access. For this reason, cryptographic methods are being developed for the purpose of providing safe cloud storage. A cryptographic technique encrypts the data before sending it to the cloud so that it cannot be read by unauthorised parties. The structure of a cryptography-based secure cloud storage solution is shown in Figure 3.

A data owner uses cryptographic techniques to protect sensitive data from unauthorised access, as seen in Figure 3. These methods help to ensure that the data remain private. After the data has been encrypted and uploaded to the cloud environment by the owner of the data, an authorised user may access the data, decrypt it, and then retrieve the necessary file. Encryption of data provides an additional layer of protection for the data before it is uploaded to the cloud. When an unauthorised user attempts to access sensitive data, the decryption of such data provides a great deal of complication for the illegal user, making it more difficult for the illegal user to execute improper inception.

2. Review of literature

Rodríguez et al., (2023)[14] studied that the fast growth of the Internet of Things may be attributed to the recent proliferation of smart and networked gadgets made possible by advancements in

hardware and information technology (IoT). Smart cities, smart industries, autonomous cars, and eHealth are just a few examples of settings that have embraced IoT applications and services. As a result, IoT gadgets are permanently linked, automatically exchanging private information. Thus, protecting confidentiality of information is of paramount importance. This article provides a detailed overview of current approaches to privacy in the IoT that are based on ML and DL. First, we provide a comprehensive review of contemporary privacy-related concerns and assaults. Finally, we provide the implementations, details, and published results for each suggested ML architecture. At the end, we single out the best countermeasures to each type of assault or danger. According to the results, the SVM model has been extensively implemented, yielding precisions between 86.08% and 94.08%.

Ma et al., (2022)[15] stated that cloud-assisted multimedia data environments may benefit greatly from the privacy-preserving Content-Based Image Retrieval (CBIR) approach. However, the performance fell short of expectations because to inadequate ciphertext descriptors and ineffective feature-preserving techniques. As a result, there are still obstacles to privacy-preserving CBIR, such as designing secure encryption methods and extracting useful information from ciphertext pictures. To do this, we suggest using deep convolutional network features to perform private picture retrieval. First, photos are encrypted using an unique hybrid encryption approach, and then the encrypted images are used to fine-tune a DenseNet model and create a feature extractor. The next step is to upload the encrypted photos and fine-tune the feature extractor to a cloud server. Additionally, the CBIR service runs securely on the cloud. We investigate on two publicly available benchmark datasets to gauge performance in terms of mean absolute precision (mAP) and accuracy. The investigational findings show that the suggested strategy outperforms the state-of-the-art approaches, increasing performance on the two measures by 1.9% and 10%, respectively. The enhanced DenseNet model also uses depth wise divisible convolution, which reduces the computational cost and parameter by 8-9 times compared to the original DenseNet's conventional convolutions while maintaining a high level of accuracy.

Kulshrestha et al., (2022)16] analyzed that statistics are stored, managed, and backed up on the cloud, making it a popular computing paradigm with adaptable features. Over a network, businesses and individuals alike may access the services on demand, expanding the capabilities of the underlying hardware via efficient and collaborative use of those resources. However, data privacy and security has emerged as a major obstacle to the widespread use of CC. First and foremost, storing data on the cloud increases the possibility of data leaks and unauthorized access. Second, attacks and disruptions are increasingly being targeted at cloud a server, which poses a threat to cloud security. Third, the owners of data stored in the cloud could not fully trust the security of jobs such data storage, reinforcement, migration, removal, update, exploration, inquiry, and admission. Thus, a probabilistic hybrid security strategy has been developed to improve CC's security. Assuring cloud users of the safety of their data is a priority, and our architecture helps to make that a reality. Our explanation is based on the "hybrid probabilistic triple (RSA and AES) encryption" method, which is used to encode-decode the data before it is uploaded to the cloud along with a hash value for each piece of information. This method ensures the data's integrity, confidentiality, and user authenticity while also addressing a variety of privacy and security concerns.

Bal et al., (2022)[17] examined that the fast expansion of the cloud computing environment has made it difficult for cloud firms to manage the vast quantities of data and the many resources available in the cloud, as their clientele has expanded from individual users to major corporations. Cloud computing performance may be hindered by inefficient resource management. As a result, there has to be fair distribution of resources among all parties involved, without jeopardising either the company's bottom line or the happiness of its customers. Customer service cannot be put on hold forever because a once-available component is now unavailable. Resource allocation security with effective task scheduling in cloud computing is a hybrid machine learning technique proposed in this study to solve these concerns (RATS-HM). Some proposed RATS-HM strategies are as follows: To begin, we use a modified version of the cat swarm optimization–based short scheduler for task scheduling (ICSTS) to minimise make-span time and maximise throughput. Second, a group

optimization-based deep neural network (GO-DNN) considers a wide variety of constraints during design in order to allocate resources most effectively. Finally, to safeguard information storage, we offer a lightweight authentication mechanism called NSUPREME. Finally, a simulation setup is used to test the proposed RATSHM method, and the results are compared to those obtained using state-of-the-art methods. The outcomes in terms of resource usage, energy consumption, reaction time, etc. reveal that the suggested method is more efficient than the current one.

Sri et al., (2022)[18] proposed that medical computing systems that run in real time on the cloud have been bolstered by machine learning. However, owing to high processing cost and time, most computing servers are autonomous in terms of data security and recovery plan in many virtual machines. In addition, the security of cloud data used by these cloud-based medical applications necessitates the use of static security settings. Multiple servers are needed for cloud-based medical apps to store patient data or ML patterns for decision making. Due to high Expensive computing memory and time, these cloud platforms need effective data security architecture to enable robust data access control among the various users. A hybrid cloud data security framework is built in this study in order to better secure the data of huge ML patterns in a real-time cloud computing environment. There are two stages to this process: the first phase of data replication, and the subsequent stage of multi-user data access security. In the beginning of the unreliable data recovery process, the machine decision patterns are replicated across several servers. We deploy a hybrid multi-access key based data encryption and decryption model for huge machine learning medical patterns for data recovery and protection in the multi-access cloud data security framework. The experimental findings demonstrated the improved computing efficiency of the new two-phase data recovery and security framework compared to the standard techniques on big medical decision patterns.

Chen et al., (2021)[19] examined that data owners are increasingly interested in cloud storage solutions that include robust sharing tools. Cloud server providers (CSPs) claim to safeguard user data, but it's not clear whether that's really the case. This work offers a novel threshold hybrid encryption for integrity auditing approach without a trusted centre, which can be used to check data integrity while still protecting data and key privacy in the group. The suggested technique is based on ECC and AES, and it makes use of secret sharing based on the Shamir algorithm. This protects both the AES key and the users' private key by preventing a central authority from controlling the distribution and management of the key. Additionally, we address the cloud-user collusion problem by designing and deploying an unique integrity auditing and re-signature approach. The suggested approach achieves accuracy, security, and efficiency with minimal communication and computing cost, as shown by the security analysis and performance assessment.

Prabhakaran et al., (2021)[**20]** analyzed that consumers may pay as they go to access their data from anywhere over the internet due to cloud computing. Because of the prevalence of hacking events in the cloud, users are understandably worried about the safety of their data stored there. When it comes to the cloud, both CSPs and customers are targets of a wide range of assaults that might have disastrous consequences. As a result, it's important to set up reliable IDS to detect and prevent attacks in the cloud before they do significant damage. The current state of the art IDS system, based on shallow machine learning approaches, has a greater false alarm rate and slower reaction time and has poor accuracy. In this study, we propose a solution to this issue by combining the LSTM, CNN, and SVM architectures into a hybrid semantic deep learning (HSDL) framework. Through the use of a semantic layer called the Word2Vec embedding layer, the linguistic context of the data flowing through the network may be determined. The HSDL model assigns a category to the assault type that matches the kind of intrusion found in the text. To strengthen the safety of cloud storage, we encrypt non-modified, error-free text using the AES method and choose the AES key with the longest key breaking time using the crossover-based mine blast optimization technique (CMBA). In order to evaluate the proposed HSDL system, we use the NSL-KDD and UNSW-NB15 real-time intrusion detection benchmark datasets. The proposed model achieves an accuracy of 99.98% on the NSL-KDD dataset and 98.47% on the UNSW-NB15 dataset. The effectiveness and

reliability of the suggested approach have been shown via experimental and security studies.

Praveena et al., (2020)[21] proposed that the public, businesses, and new applications like healthcare and natural disaster relief may all benefit greatly from cloud computing. The two main types of cloud computing may be broken down into "private cloud" and "public cloud" in accordance with the services they provide and the details of their design. Both of these categories of cloud services are essential in this context if a company is to give adequate support to its constituents. So, a new cloud service, hybrid cloud service, combining elements of both private and public clouds, was launched. Today, cloud security is a complex topic, and the hybrid cloud poses unique challenges owing to its mix of public and private cloud services. Although there is several security procedures described in the literature, none of them have yet managed to provide a reliable safeguard. To this end, we present a novel machine learning programme to safeguard hybrid cloud networks during data storage and retrieval. Using a combination of the current Enhanced C4.5, the suggested deduplication processing method, and the proposed dynamic access control mechanism, this new technique is developed. Furthermore, we provide a novel deduplication processing technique for redundant-free, safe storage and retrieval. The proposed security architecture also takes use of the Dynamic Spatial Role Based Access Control Algorithm, a newly released dynamic access control mechanism. The security level of the hybrid cloud during data storage, retrieval, and access has been examined, and the recommended security architecture has been implemented.

Yousif et al., (2020)[22] stated that building a trustworthy Hadoop that can function as a cloud computing platform is a crucial task. The cloud computing security policy is adaptable enough to be utilised with a broad variety of cloud services, including Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS) (SaaS).This development underscores the need of a strategy to rein in these difficulties. Hadoop, which often employs the MapReduce architecture to organise massive volumes of data in the cloud system, is often suggested as a viable method to overcome this big data challenge. Hadoop's Distributed File System (HDFS) stores data without any kind of policy in place to guarantee the security or confidentiality of that data. Encryption techniques may play an essential role in ensuring the security of sensitive data stored in the cloud. A novel approach of encrypting HDFS data is proposed in this study, using two well-known asymmetric key cryptosystems (RSA and Rabin). Since this is the case, the suggested cryptosystem is used to encrypt data before it is placed in HDFS. To upload data to the cloud, users of the proposed system have the option of using either a secure or insecure method. The hybrid approach boasts superior computational complexity and lower latency when compared to only the RSA cryptosystem.

Sajay et al., (2019)[23] studied that the phrase "cloud computing" is used to refer to various computer architectures in which several computers are networked together in real time across a public network like the Internet. The potential for cloud computing to reduce computing costs is one of the main reasons why this emerging paradigm has attracted a large number of academics in recent years. It is remarkable how quickly cloud services have grown in popularity with the development of cloud computing methods. Maintaining secure information in the modern environment is a difficult task. Concerns about cloud security and how best to execute cloud services across networks are at the heart of the cloud computing debate. Protecting sensitive information in the cloud requires many layers of protection. The cloud services, deployment methodology, issues with and challenges to cloud security are all covered. Using the right encryption methods for storing data in the cloud has become more important in light of the recent rise in cybercrime. This research has proposed a hybrid method to improve encryption algorithm security for cloud data. When it comes to securing or archiving massive amounts of data, encryption technologies shine. To further strengthen cloud safety, this research combines homographic encryption with blowfish encryption. The future, it seems, would be solutions for cloud storage for both small and big businesses, provided that the current security concerns are addressed.

Saini et al., (2018)[24] examined that the safety and reliability of technological systems are of paramount importance in any region. Users are embracing every kind of technology, from the most fundamental aspects of computing to the most advanced forms of cloud computing. Without

confidence in the system's ability to keep data safe, no one will be willing to embrace it. Data security in a multicloud setup is discussed, along with the various algorithms and encryption methods that may be used. Regardless of the many other methods, such as RSA, AES, E2EE, and MD5, may be used to protect data in a hybrid cloud setting. The benefits and limitations of RSA in terms of time and storage space are discussed in this work. The benefits and drawbacks of the AES Algorithm and the MD5 technique are also examined in this work. In this paper, we'll look into End-to-End Encryption (E2EE) as a unique solution to data security for hybrid cloud services, discussing its advantages and drawbacks along the way. The data is encrypted using the Public key and decrypted using the Private key before being sent. Using this method, the server is prevented from ever gaining access to the unencrypted message and hence the stolen information. Based on the research, the authors believe that E2EE is a feasible alternative for hybrid security.

2.1 Comparison of reviewed technique

There is a diverse group of authors who have made use of their findings in their own study; a description of these authors can be found in table 1.

Table 1. Comparison of reviewed technique

Author [Ref.]	Technique	Outcome
Rodríguez et al., (2023)[14]	Machine learning and IoT	According to the results, the SVM model has been extensively implemented, yielding precisions between 86.08% and 94.08%.
Ma et al., (2022)[15]	CBIR	The investigational findings show that the suggested strategy outperforms the state-of-the-art approaches, increasing performance on the two measures by 1.9% and 10%, respectively.
Kulshrestha et al., (2022)16]	AES	This method ensures the data's integrity, confidentiality, and user authenticity while also addressing a variety of privacy and security concerns.
Bal et al., (2022)[17]	RATS-HM	The outcomes in terms of resource usage, energy consumption, reaction time, etc. reveal that the suggested method is more efficient than the current one.
Sri et al., (2022)[18]	ML	The experimental findings demonstrated the improved computing efficiency of the new two-phase data recovery and security framework compared to the standard techniques on big medical decision patterns.
Chen et al., (2021)[19]	ECC and AES	The suggested approach achieves accuracy, security, and efficiency with minimal communication and computing cost, as shown by the security analysis and performance assessment.
Prabhakaran et al., (2021)[20]	CMBA and AES	The proposed model achieves an accuracy of 99.98% on the NSL-KDD dataset and 98.47% on the UNSW-NB15 dataset.
Praveena et al., (2020)[21]	Dynamic Spatial Role Based Access Control Algorithm	The security level of the hybrid cloud during data storage, retrieval, and access has been examined, and the recommended security architecture has been implemented.
Yousif et al., (2020)[22]	HDFS	The hybrid approach boasts superior computational complexity and lower latency when compared to only the RSA cryptosystem.
Sajay et al., (2019)[23]	Encryption technique	This research has proposed a hybrid method to improve encryption algorithm security for cloud data.
Saini et al., (2018)[24]	E2EE	Using this method, the server is prevented from ever gaining access to the unencrypted message and hence the stolen information.

3. Novelty

Ensuring adequate security for a cloud environment is a difficult challenge in the modern day. There are a great deal of security procedures that have not yet reached an adequate degree of security. For this reason, we present a novel machine learning application that makes use of cryptographic methods to ensure data safety throughout the storage of data in the cloud, as well as during the process of retrieving or accessing data from the cloud. When compared to previously developed models, the time complexity is lower. Effectively manages the cloud data security risk that is

associated with the vast amount of cloud storage space. Provides a better result when it comes to protecting sensitive information, encrypting it, decrypting it, and analyzing how much time is needed to get it. An original classification strategy that integrates cryptographic techniques with modern-day machine learning models.

4. Problem formulation

Cloud data storage refers to the use of secure, remote hosting services for archival data. Files stored on external hard drives may also be uploaded through cloud. Data security is the process of keeping information private and secure either in transit or at rest. Outsourced data may be verified with the help of the storage security before being placed on untrustworthy cloud servers. Due to concerns about data security, it is important to take extra precautions when entrusting sensitive information to cloud-based storage services. Given the existing level of service, it seems clear that some kind of encryption (or other application-level data security) is necessary. This method enables (i) resolving trust concerns between the client and database, (ii) attaining the appropriate degree of flexibility in terms of allowing encryption at the finest level of characteristics, and (iii) performing out advanced computation over encrypted data.

5. Research methodology

The methodology that has been described is one that can be executed because, even if the data comes from the user, the database will be built if it is not already existent in a directory with tables that are known to the user. One table is for the data and IP address of the user which is sent and another table is about user information with credentials of user. IP address and email are the two methods that are used to validate the user. Logistic Regression is a mathematical modeling procedure which depicts the connection between numerous independent variables, X1... XK, and a dependant variable, D. The strategic model makes use of the estimated capacity as a numerical structure that can reach anywhere between 0 and 1 for certain random information. This range is determined by the strategic model. The system framework is shown in Figure 3.

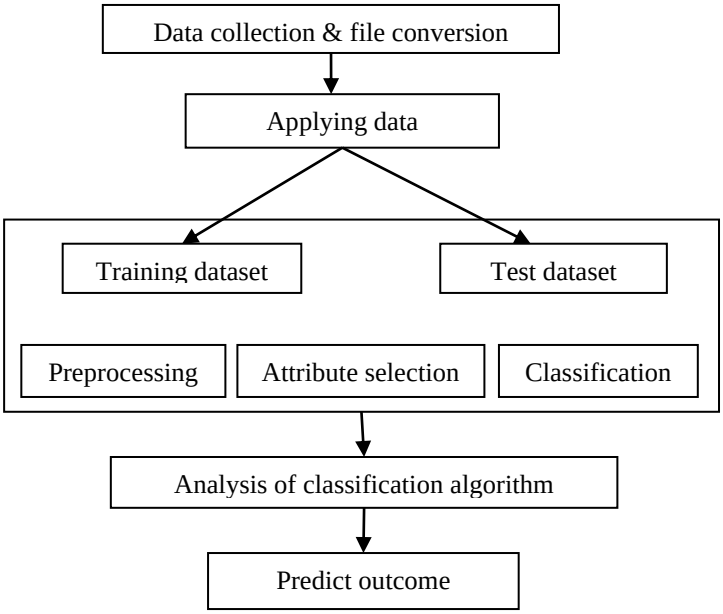

Fig 3. System framework

Decision Tree: A flowchart-like tree structure is represented by a decision tree model, which is also known as a decision tree. Each internal hub in this architecture addresses a test that is performed on a property of the dataset, while each tree limb addresses the outcome of the test.

Linear Regression Algorithm: An artificial intelligence computation that relies on controlled learning is called a linear regression algorithm. The regression analysis process includes the use of the linear regression method. The technique of predictive forecasting known as regression analysis is a tool that may aid you in determining the link that exists between the input variable and the objective variable.

Advanced Encryption Standard (AES): It is also called as Rijndael algorithm. The method is a symmetrical block cypher, and it turns plain text into cypher text by taking it in blocks of 128 bits and utilising keys that are 128, 192, or 256 bits long. It is byte-oriented. The structure is dependent on an organisation that manages replacement changes. The procedure for making this conclusion is top secret, despite the fact that it is accessible to public opinion.

Data Encryption Standard (DES): It is a computation for a block cypher algorithm that takes ordinary text in squares of 64 pieces and turns it to ciphertext using keys that are 48 bits long. This technique uses symmetric keys and has a bit-oriented structure. The cycle that is used to determine this is a closely guarded secret. Attacks against DES that have been documented in the past include Brute-power, Linear Grave Investigation, and Differential Grave Investigation.

Cloud computing would give a service in order to fulfil the requirements of its customers. Because of this, the cloud is becoming more important as more and more people depend on it. The detection of cloud-based attacks often makes use of machine learning in a variety of situations. It is able to identify the affected users when an attack takes place, and it is also able to prevent the attack from taking place by inspecting the security system for any potential flaws. Figure 4 depict the flow diagram as shown below.

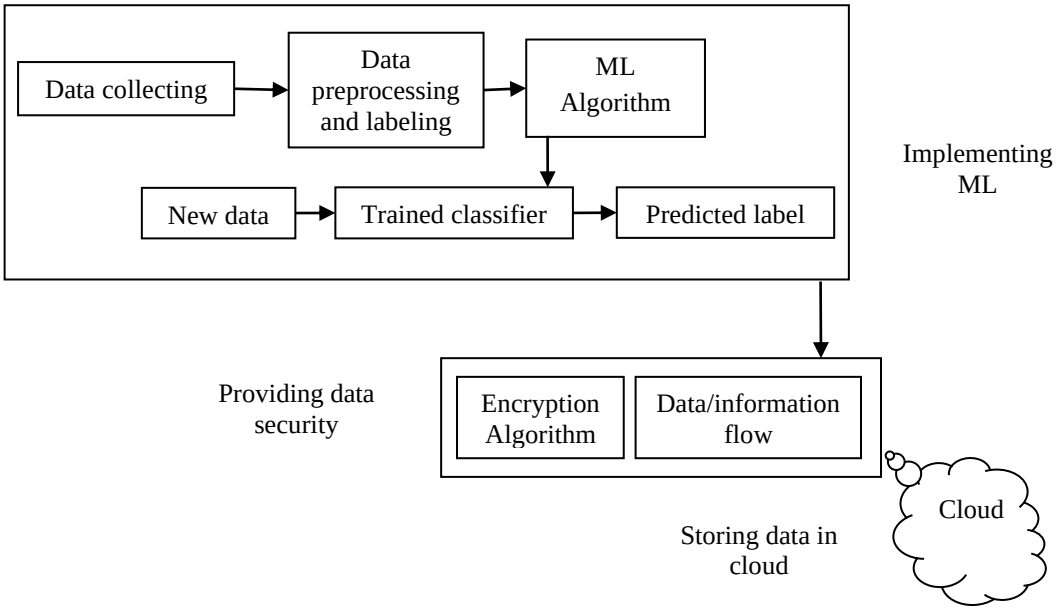

Fig 4. Flow diagram

The following is a description of each phase in the process.

Step1: Obtaining the data from the many sources of data.

Step2: Initially, the data are preprocessed in order to get a normalised dataset, and then the data rows are labeled.

Step3: The Machine Learning Algorithm is responsible for taking care of the aftereffect of the future advancement, which is the development and testing of the dataset.

Step4: The ML algorithm constructs a model with the help of the training data, and then it tests the model with the help of the test data.

Step5: The ML algorithm generates a trained model or trained classifier in order to predict the label of a new data row.

Step6: The data collected is guaranteed to be safe and secured due to the use of an AES encryption computation.

Step7: After that, data is encrypted and stored on a remote or local server.

6. Result and Analysis

Multiple results illustrate the discrepancy between the proposed model and previous research models. In addition, the findings are demonstrating the efficacy of the suggested model in terms of

specificity, sensitivity, F1 score, and confusion matrix.

Result 1

Figure 3 depict the confusion matrix as seen below. This matrix is described in terms of both its output class and its target value. Count values are used to provide a concise summary of the numbers of accurate and inaccurate predictions.

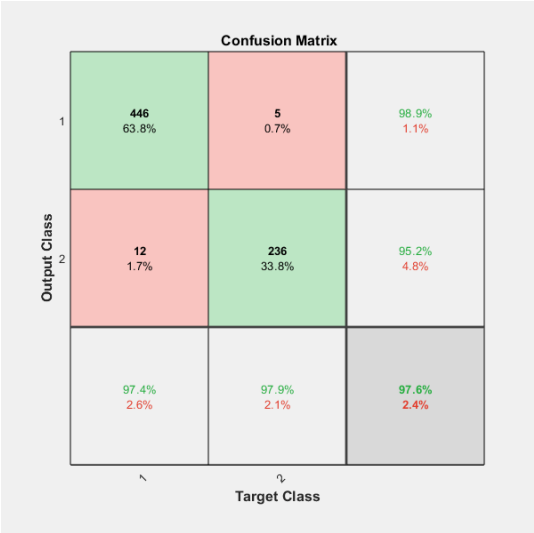

Fig 3. Confusion matrix

Result 2

Figure 4 illustrates the model accuracy and the loss that is shown below which is define in terms of train and test. The train is indicating by the blue line while test is signifying by the orange line as well as shown in model loss as shown below. Model accuracy was a representation of the model's performance, which was defined as a continuous improvement based on epochs in a manner that was easy to understand and computed as a percentage. The loss is stated in terms of between loss and epochs, and its interpretation is defined as being that models are continuously going down in decreasing form.

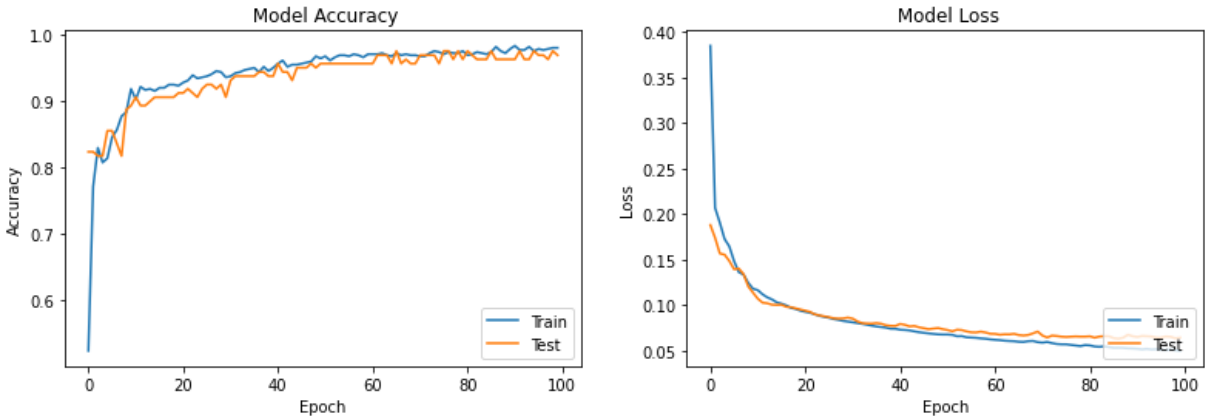

Fig 4. Model accuracy and loss.

Result 3

The area under the curve (ROC) is shown in Figure 5 and is defined as the ratio of the genuine positive rate to the false positive rate. The proportion of false positives is shown horizontally, while the proportion of correct positives is plotted vertically. The true positive rate is the percentage of observations that are expected to be positive when, in fact, they are positive. The percentage of observations that were assumed to be positive but turned out to be negative is known as the false positive rate. If the false positive rate and true positive rate are lower and higher values, respectively, it is desirable for classification. In this figure, No skill is indicating by the blue line while

logistic is signifying by the orange line as shown below. The ROC AUC curve of No skill is 0.500 and the ROC AUC curve of logistic is 0.903 as shown below.

```
No Skill: ROC AUC=0.500
Logistic: ROC AUC=0.903
```

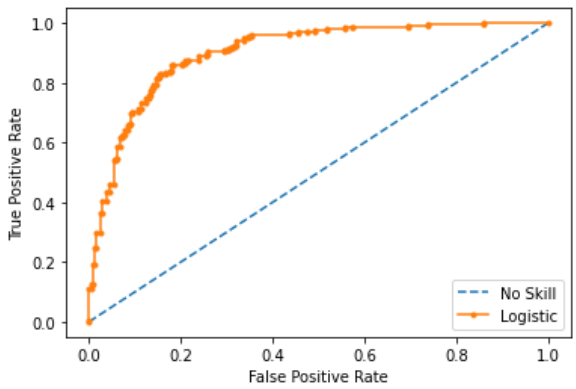

Fig 5. ROC curve

Result 4

The convergence curves of the Harris Hawks' optimization (HHO) are shown in Figure 6. The figure reports that the HHO has the best fitness curve, which is 0.130 with regard to the number of iterations, as is seen below.

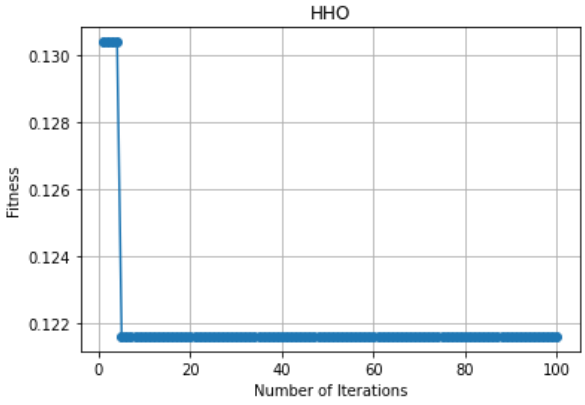

Fig 6. Convergence curve of HHO

Result 5

Figure 7 presents a demonstration of the suggested model's specificity, sensitivity, and F1-score value. This result suggests that the model that we have provided has the best value compared to other models. The suggested model's specificity is 97.5%, sensitivity, 87.5% and F1-score is 93.5% are all shown in Figure 7.

```
specificity is : 97.5
sensitivity is : 87.5
F1_score is : 93.5
```

Fig 7. Specificity, sensitivity, and F1 score of proposed method.

Model accuracy, model loss graph, and confusion matrix provide further evidence supporting the effectiveness of the suggested model. These graphs are generated by without optimization.

Result 6

Figure 8 illustrates the model accuracy and the loss that is shown below which is define in terms of train and test. The train is indicating by the blue line while test is signifying by the orange line as well as shown in model loss as shown below. Model accuracy was a representation of the model's performance, which was defined as a continuous vary with no. of epochs in a manner that was easy to understand and computed as a percentage. The loss is stated in terms of between loss and epochs,

and its interpretation is defined as being that models are continuously going down in decreasing form vary with epoch as shown below.

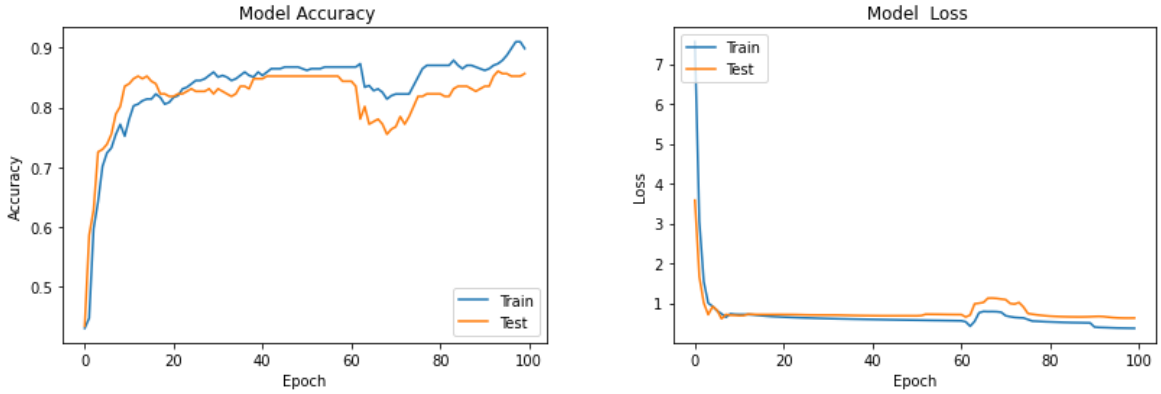

Fig 8. Model accuracy and loss (without optimization)

Result 7

The confusion matrix is shown in figure 9, which may also be seen below. The accuracy of the loss model is shown throughout a variety of time periods, or epochs. The performance of a classifier may be easily and effectively visualised with the use of confusion matrices. The fact that confusion matrices are easily interpretable is one of the most significant advantages they provide. While each row in the matrix displays the actual classes that were seen, each column in the matrix (which is sometimes represented graphically as a heatmap) displays expected classes. The ultimate result is that each cell represents a unique potential combination of the actual classes and those that were anticipated.

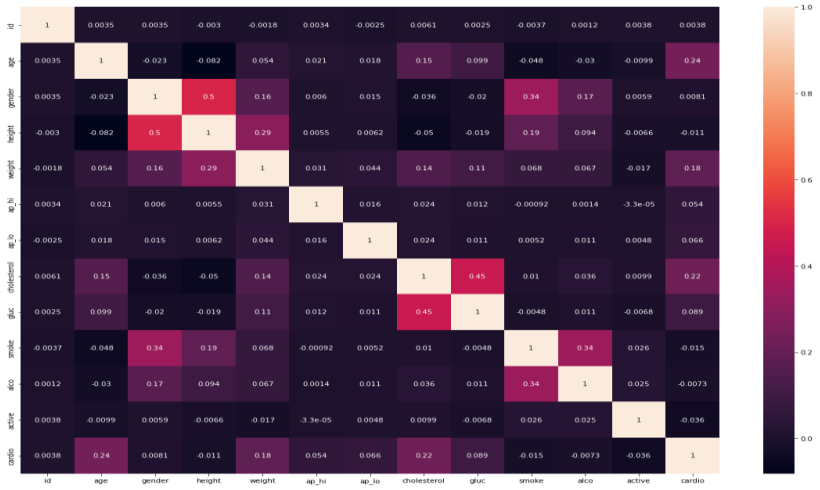

Fig 9. Confusion matrix of proposed model

7. Conclusion and Future Work

Cloud computing is a cutting-edge, future-proof service delivery innovation that can be used to any kind of data on any type of server. When it comes to safeguarding the Meta data utilizing various technologies, this presents a significant difficulty. Both from the perspective of the consumer and the cloud provider, information security are an extremely important factor. Information security was the primary focus of our investigation into many aspects of safety and security in this study. To ensure the security of cloud-based storage, we make use of encryption methods. Here, we apply efficient encryption techniques that boost speed both while encrypting and decrypting data. In comparison to other storage and retrieval methods, this one is much better both in terms of convenience and efficiency. A ROC AUC curve is constructed from the findings, and it indicates that logistic and NO skills are responsible. The value of NO talent is just 50.0%, whereas the value of logistics is 90.3%. The convergence curve of HHO

has been obtained. According to the results of the study, the suggested approach has attained a sensitivity of 87.5%, an F1-score of 93.5%, and a specificity of 97.5%. In future, further encryption methods will be developed for cloud service provisioning in order to increase the level of confidentiality achieved, in accordance with the data requirements.

References

[1] El-Attar, Noha E., Doaa S. El-Morshedy, and Wael A. Awad. "A New Hybrid Automated Security Framework to Cloud Storage System." *Cryptography* 5, no. 4 (2021): 37.

[2] Wang, Cong, Qian Wang, Kui Ren, Ning Cao, and Wenjing Lou. "Toward secure and dependable storage services in cloud computing." *IEEE transactions on Services Computing* 5, no. 2 (2011): 220-232.

[3] Gao, Jiechao, Haoyu Wang, and Haiying Shen. "Machine learning based workload prediction in cloud computing." In *2020 29th international conference on computer communications and networks (ICCCN)*, pp. 1-9. IEEE, 2020.

[4] Udendhran, R. "New Framework to Detect and Prevent Denial of Service Attack in Cloud Computing Environment." *Asian Journal of Computer Science and Information Technology* 4, no. 12 (2014): 87-91.

[5] Udendhran, R. "A hybrid approach to enhance data security in cloud storage." In *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*, pp. 1-6. 2017.

[6] Wani, Abdul Raoof, Q. P. Rana, U. Saxena, and Nitin Pandey. "Analysis and detection of DDoS attacks on cloud computing environment using machine learning techniques." In *2019 Amity International conference on artificial intelligence (AICAI)*, pp. 870-875. IEEE, 2019.

[7] Mohiuddin, Irfan, Ahmad Almogren, Majed Alrubaian, and Muhammad Al-Qurishi. "Analysis of network issues and their impact on Cloud Storage." In *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1-4. IEEE, 2019.

[8] Bhagawat, V., and A. Kumar. "Survey on data security issues in cloud environment." *International Journal of Innovative Research in Advanced Engineering* 2, no. 1 (2015): 31-35.

[9] Darwazeh, Nour S., Raad S. Al-Qassas, and Fahd AlDosari. "A secure cloud computing model based on data classification." *Procedia Computer Science* 52 (2015): 1153-1158.

[10] Cyril, B. Rex, and DS KUMAR. "Cloud computing data security issues, challenges, architecture and methods-A survey." *International Research Journal of Engineering and Technology* 2, no. 4 (2015): 848-857.

[11] Timothy, Divya Prathana, and Ajit Kumar Santra. "A hybrid cryptography algorithm for cloud computing security." In *2017 international conference on microelectronic devices, circuits and systems (ICMDCS)*, pp. 1-5. IEEE, 2017.

[12] Singh, Rajesh, Anita Gehlot, Mamoon Rashid, Ritika Saxena, Shaik Vaseem Akram, Sultan S. Alshamrani, and Ahmed Saeed AlGhamdi. "Cloud server and internet of things assisted system for stress monitoring." *Electronics* 10, no. 24 (2021): 3133.

[13] Goel, Gaurav, Rajeev Tiwari, Vinay Rishiwal, and Shuchi Upadhyay. "Data preservation by hash algorithm for matrix multiplication over venomous cloud." In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 210-214. IEEE, 2018.

[14] Akhtar, Naveed, and Ajmal Mian. "Threat of adversarial attacks on deep learning in computer vision: A survey." *Ieee Access* 6 (2018): 14410-14430.

[15] Rodríguez, Eva, Beatriz Otero, and Ramon Canal. "A Survey of Machine and Deep Learning Methods for Privacy Protection in the Internet of Things." *Sensors* 23, no. 3 (2023): 1252.

[16] Ma, Wentao, Tongqing Zhou, Jiaohua Qin, Xuyu Xiang, Yun Tan, and Zhiping Cai. "A privacy-preserving content-based image retrieval method based on deep learning in cloud computing." *Expert Systems with Applications* 203 (2022): 117508.

[17] Kulshrestha, Vartika, Seema Verma, and C. Rama Krishna. "Hybrid probabilistic triple encryption approach for data security in

cloud computing." *International Journal of Advanced Intelligence Paradigms* 21, no. 1-2 (2022): 158-173.

[18] Bal, Prasanta Kumar, Sudhir Kumar Mohapatra, Tapan Kumar Das, Kathiravan Srinivasan, and Yuh-Chung Hu. "A joint resource allocation, security with efficient task scheduling in cloud computing using hybrid machine learning techniques." *Sensors* 22, no. 3 (2022): 1242.

[19] Sri, V. Devi Satya, and Srikanth Vemuru. "A Framework for Uncertain Cloud Data Security and Recovery Based on Hybrid Multi-User Medical Decision Learning Patterns." *International Journal of Communication Networks and Information Security* 14, no. 1s (2022): 136-152.

[20] Chen, Yange, Hequn Liu, Baocang Wang, Baljinnyam Sonompil, Yuan Ping, and Zhili Zhang. "A threshold hybrid encryption method for integrity audit without trusted center." *Journal of Cloud Computing* 10 (2021): 1-14.

[21] Prabhakaran, Varun, and Ashokkumar Kulandasamy. "Hybrid semantic deep learning architecture and optimal advanced encryption standard key management scheme for secure cloud storage and intrusion detection." *Neural Computing and Applications* 33, no. 21 (2021): 14459-14479.

[22] Praveena, D., and P. Rangarajan. "A machine learning application for reducing the security risks in hybrid cloud networks." *Multimedia Tools and Applications* 79 (2020): 5161-5173.

[23] Yousif, Raghad Z., Shahab W. Kareem, and Shadan M. Abdalwahid. "Enhancing approach for information security in hadoop." *Polytechnic Journal* 10, no. 1 (2020): 81-87.

[24] Sajay, K. R., Suvanam Sasidhar Babu, and Yellepeddi Vijayalakshmi. "Enhancing the security of cloud data using hybrid encryption algorithm." *Journal of Ambient Intelligence and Humanized Computing* (2019): 1-10.

[25] Saini, Kavita, Vaibhav Agarwal, Arjun Varshney, and Anushka Gupta. "E2EE for data security for hybrid cloud services: a novel approach." In *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 340-347. IEEE, 2018.

[26] A. K. Shukla and A. Sharma, "Cloud Base Intrusion Detection System using Convolutional and Supervised Machine Learning," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112007.

[27] A. K. Shukla and A. Sharma, "Classification and Mitigation of DDOS attacks Based on Self-Organizing Map and Support Vector Machine," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10111988.

[28] A. K. Shukla and A. Sharma, "Distributed Attacks Classification Based on Radical Basis Function and Particle Swarm Optimization In Hypervisor Layer," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-4, doi: 10.1109/ISCON57294.2023.10112162.

[29] Shukla, Aditya Kumar, and Ashish Sharma. "Cryptography Algorithm Using A Lightweight For Enhancing Data Security In Cloud Computing." Journal of Optoelectronics Laser 41.5 (2022): 606-624.

[30] Manikandan, J. ., & Uppalapati, S. L. . (2023). Critical Analysis on Detection and Mitigation of Security Vulnerabilities in Virtualization Data Centers. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 238–246. https://doi.org/10.17762/ijritcc.v11i3s.6187

[31] Carlos Silva, David Cohen, Takashi Yamamoto, Maria Petrova, Ana Costa. Ethical Considerations in Machine Learning Applications for Education. Kuwait Journal of Machine Learning, 2(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/