A Secure Mechanism for Safeguarding Cloud Infrastructure

Khare Pratyush, Vivek Kumar Prasad(✉), Rachana Mehta, and Madhuri Bhavsar

Department of CSE, Nirma University, Ahmedabad, India
{20bce519,vivek.prasad,rachana.mehta,
madhuri.bhavsar}@nirmauni.ac.in

Abstract. Security is frequently viewed as the largest impediment to a cloud-based approach, but in actuality, it can be the majorenabler. Cloud security guarantees that your information and apps are easily accessible to authorized users. In this paper, we shall be putting forward, the cloud ecosystem's security concerns. The most crucial concerns for the popularity of cloud computing services are privacy and security. Here we try to depict a study of data that is hosted on the cloud and the issues in its security. The study will examine the particular data protection practices used globally to offer optimum data security while reducing threats and risks. Although many apps benefit by having access to data on the cloud, but doing so poses concerns since it makes data accessible to apps that could already contain security flaws. Analog to this, data may be at risk if a guest OS operates on top of the hypervisor, without consideration for dependability of the guest OS, resulting in a flaw in security. The paper ends with a case study where the request has been classified as safe or malicious. If the malicious request is identified, then these requests are to be discarded so that the cloud remains safe. The classification has been conducted using Machine Learning and Deep Learning concepts and an accuracy of 85% has been achieved.

Keywords: Threats · Risks · Data security · Data protection

1 Introduction

Cloud computing has fundamentally changed how end users access computers as well as other IT resources in the contemporary technological era. But yet computing based on cloud system is a new term which has not yet been widely accepted. Among the various definitions available, "A network solution for providing inexpensive, reliable, easy and simple access to IT resources" is one of the most basic definition. Cloud Computing provides IT access that is low-cost, dependable, as well as straightforward. Cloud is service focused rather than application oriented [8].

Cloud Computing's nature not only reduces the timeframe needed to execute a task, but it also minimizes the amount of infrastructure and ownership costs, and also provides end-user flexibility and enhanced efficiency. The latest web-based computer network, cloud computing, provides users with simple and adaptable resources to access or use

S. Rajagopal et al. (Eds.): ASCIS 2022, CCIS 1760, pp. 144–158, 2022.
https://doi.org/10.1007/978-3-031-23095-0_11

various cloud apps. Without explicit user's active control, cloud computing is the availability of computer network services, primarily for storing data and computing power [9].

Research on cloud computing is now being extensively used in both academia and business. Cloud computing benefits users as well as cloud service providers (CSPs). The difficulties with security that come with cloud computing have been extensively researched and we try to propose solutions to those limitations. In a cloud environment information may be shared throughout different businesses, which is an advantage of Cloud Computing. However with greater conveniences, comes great responsibility of safeguarding that shared data so that it is not misused by any illegitimate user or an attacker [10].

As and when one decides to use the services of cloud to store data, a crucial decision has to be formed as to employ a 3rd- party cloud service provider or develop one's own personal business cloud. Data pertaining to national security or highly secretive upcoming product information, for instance, is often too delicate to be kept on a public cloud. Data pertaining such information can be exceedingly sensitive, and moreover exposing it to the public cloud can have catastrophic effects. On such cases, it is strongly advised to store data in an organization's personal internal cloud. This approach can promote data security by enforcing on-premises data usage regulations. Furthermore, many businesses lack the expertise to apply all necessary layers of security to critical data, therefore total data security and privacy are still not guaranteed [11].

In order to safeguard and secure information stored in the cloud, this study examines data security strategies that are employed worldwide. It examines potential hazards to cloud data as well as the precautions taken by different service providers in order to ensure its security.

The remaining work is structured as hereunder. A survey of the literature in Sect. 2 provides an overview of earlier studies in this area. We try to look at the several different types of cloud computing security in Sect. 3. The dangers to data in the cloud are discussed in Sect. 4. Section 5 looks at some of the most effective security techniques used for data security around the globe. This is followed by Sect. 6, which presents the major security challenges that are posed upon cloud infrastructure. Section 7 shows how we can use Encryption in order to provide Security to the data. A Case Study on Cloud Security Model is presented in Sect. 8. The paper's last section, the conclusion, offers an overview of the whole body of research.

2 Literature Review

We utilized a variety of sources in order to gain an insight over the principles of cloud computing and also on how we can securely store data over there. A literature review is included in this part to set the stage for examining various data security challenges.

J. Shrinivas talked about how consumers are concerned about shifting their data to the cloud. He asserts that security concerns are the main causes why big businesses are still hesitant to migrate their data on to the cloud. The authors did an excellent job of analyzing security issues in data and privacy preservation challenges in the cloud. Additionally, he also discusses some viable remedies for these problems [1].

M. A. Vouk, on the other hand, designed a standard for protecting cloud data while it is in motion. A starting point of encryption has been considered for data security during transfer. Additional encryption is necessary for trustworthy security, but it greatly lengthens the latency. The benchmark they utilized in their investigation finds an equilibrium between security and encryption overhead [2].

By giving the end user power over their data, P. S. Wooley investigates the privacy concern in an effort to inspire confidence. He examines numerous attacks based on cloud, and also proposes ways to combat them [3].

V. Kavitha and S. Subashini offer a cloud computing data protection paradigm which is based on architecture of the cloud. They have also developed tools to add to the effort in developing a prototype for security of data in Cloud Computing [5].

An efficient method for implementing security policies in web services was put out by Ranchal and Bhargava. This method can preserve user privacy, provide data owners control over data disclosure decisions, and lower the danger of unauthorized access [6].

According to a study by Martin, consumers pay greater concern to privacy, but it doesn't provide them the capacity to regulate their online-experience. That evaluated how important privacy expectations violations are in relation to consumer website trust [7].

Several dangers in the cloud environment have been thoroughly analyzed and studied in [22, 23]. According to current trends, machine learning techniques are being used to advance cloud and network security, as mentioned in [24].

To improve cloud security, Support Vector Machine (SVM)-based classifiers are used in [25, 26]. Cloud abnormalities were found using a different machine learning classifier that is frequently employed to increase cloud security in [27, 28].

3 Types of Security in Cloud Computing

A set of security measures known as cloud security, often referred to as Cloud Security, is intended to safeguard the architecture, applications, and data stored in the cloud. The different sorts of Security Models have been depicted in Fig. 1.

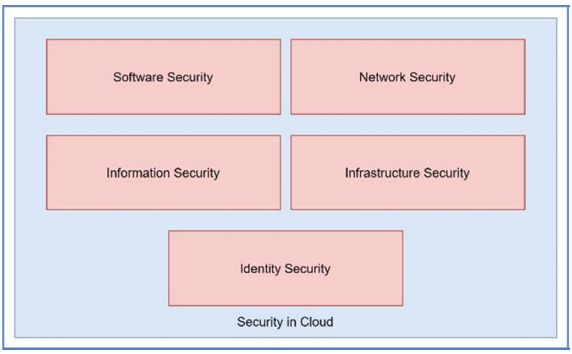

Fig. 1. Different types of security in cloud

3.1 Information Security

No matter if the data is encoded, moved, processed, or deposited, maintaining a set of business policies that will secure data resources is necessary for information protection [13].

3.2 Identity Security

It is referred to be the privacy and professional approach that "allows the authenticating users to obtain the resources at the proper moment and for the excellent aims" [14]. While preserving the confidentiality and security of data and programs, it enhances access to certified users.

3.3 Network Security

An important computing necessity is network security. It entails taking protective hardware and software measures to guard against unauthorized users, breakups, revisions, violations, degradation, or inappropriate dissemination of the existing networking infrastructure, thereby providing a stable platform for machines, clients, and services to carry out their essential tasks in a secure setting [15]. Web systems in particular may be influenced by network level issues, which fundamentally reduce capacity and lengthen device delay.

3.4 Software Security

Security concerns for software should be established as a procedure of security analysis, starting with the concept of the program and continuing through the design and execution procedures. To provide the highest level of software security, each of these procedures depends on the others [16]. Although the complexity of software development efforts varies greatly, they always require security assurance.

3.5 Infrastructure Security

It is very vital for an enterprise to be able to verify that the infrastructure is secure in order to do business. It's important to keep things separate [17].

4 Security Concerns in Cloud

Cloud-computing and its data are related with a number of dangers and security issues. However, this research will focus on virtualization, multitenancy and public cloud storage, of which, all are connected to cloud-computing's data security. The areas pertaining to Cloud Security are showcased in Fig. 2.

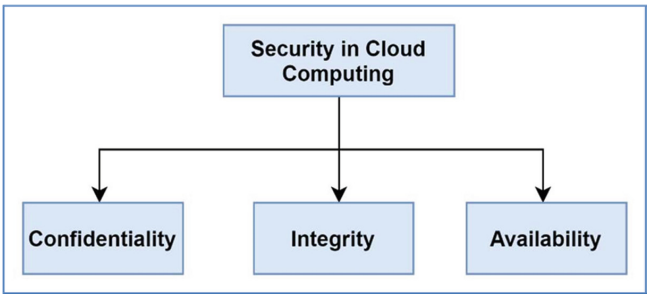

Fig. 2. CIA appendix in cloud

4.1 Virtualization

Virtualization is a technique for copying the image of a completely operational operating system to some another operating system. A specific component known as a hypervisor is essential in order to execute a guest operating-system as a virtual- machine in a host operating-system [18]. The architecture of Virtual Machines is shown in Fig. 3.

Virtualization is the primary technology that has altered how cloud computing data centers operate. Multiple copies of VMs may be created over the same physical infrastructure, which improves resource utilization and boosts return over Investment. Matter of concern in this case is the breach of the hypervisor itself. A weak hypervisor can end up becoming the primary target. The entire system and the data are at danger if a hypervisor is compromised [12].

The de-allocation and assignment of resources is another danger associated with virtualization. If the data of a particular VM is directed to the memory and not destructed before memory is allocated to some other VM, there are high chances that the data will

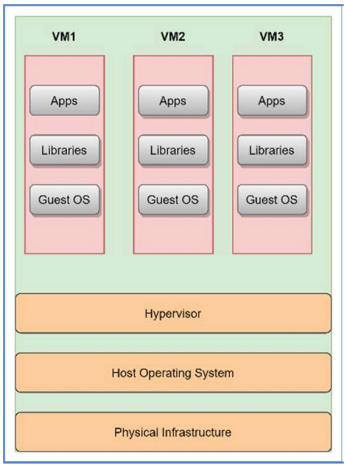

Fig. 3. Virtual machine infrastructure

be visible to new VM that has been allocated the same memory which is not ideal at all [19].

Improved and enhanced planning for virtualization's usage is a solution to the challenges listed above. Before de-allocating resources, care should be taken in their use and data must be properly validated.

A virtualized infrastructure or environment is protected by a combination of policies, practices, and procedures known as virtualization security. It examines security concerns that affect the elements of a virtualized system and solutions for mitigating or preventing them. A wide range of techniques are included in the broad notion of virtualization security, which may be used to assess, deploy, monitor, and manage security inside a virtualized infrastructure or environment. Virtualization security involves procedures like granular implementation of security processes and controls at each virtual machine, protecting virtual networks, virtual computers, and other virtual appliances from threats and flaws that the underlying physical equipment may have revealed and ensuring that each virtual machine is within your control and responsibility.

4.2 Public-Cloud Storage

Public clouds are affordable, incredibly adaptable, and infrastructure ready. However, there are significant problems when it comes to sensitive data. Normally, clouds use centralized storage, which makes them an easy victim [20].

Resources for storage are complex systems which combine software as well as hardware solutions, and in the public cloud, even a little compromise would result in data breach. For particularly sensitive data, it is usually advisable to have a private cloud if possible to eliminate such hazards.

4.3 Multitenancy

The greatest threats to cloud's data is shared access, commonly known as multitenancy. Multitenancy in cloud-computing refers to shared hosting, where server resources are distributed among various clients. It is the opposite of single tenancy, which occurs when a computer system or software instance has just one end user or set of users [21].

Multitenancy is posing problem to the clients as if there are issues with any of the client, it will impact the others in the same pool too. It is so because the other clients will be vulnerable to attacks as they are using shared server resources.

By thoroughly confirming individuals' identities before providing anyone access to the data, these issues may be avoided.

4.4 Identity and Access Management (IAM)

IAM relates to the user accounts' accessibility privileges. Managing user account authentication and authorization also applies here. Access controls are essential for preventing users, both good and bad, from accessing and jeopardizing systems and sensitive data. IAM encompasses techniques like password management and multi- factor authentication, among others.

5 Security of Data in Cloud-Computing

Data security is much more than only encryption of data in Cloud. Data security requirements are differing in all the three cloud service models. The Architecture of Cloud is depicted in Fig. 4.

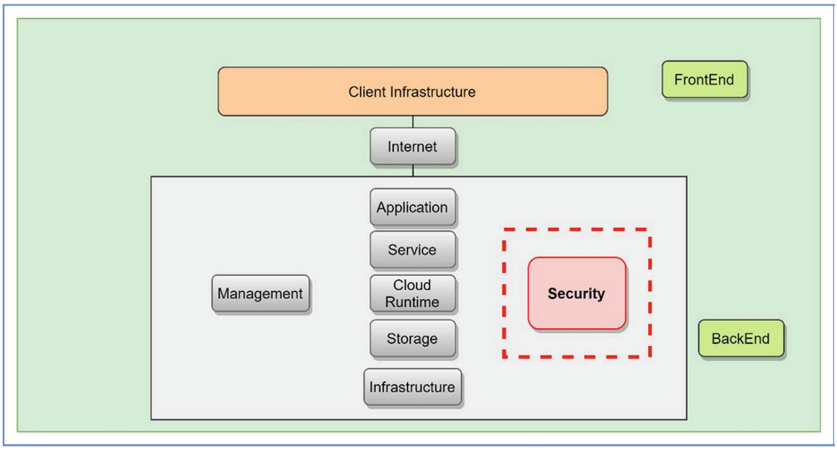

Fig. 4. Cloud system architecture

Cloud's data can be categorized in several ways. The type of data protection techniques, processes and procedures determines data confidentiality and integrity. The most important issue is data exposure in the two states stated below.

Types of Data in Cloud:
A. Data at Rest
Data at rest or data on cloud is the data which can be easily accessed with the help of internet. This shall include both – backed up data as well as the live-data. It contains everything, including the database for the program, log files, system configuration files, backups, and archives.

B. Data in Motion
In general, moving data pertains to information that enters and exits the cloud every now and then. This is information that the program is actively accessing and using. It could involve the transfer of data between two separate apps or services, or even between clients and servers or other parts of the same application.

Because the later must travel from one area to another, data-in-transit is often much vulnerable when compared to prior. There are several ways that intermediate software might monitor data and, occasionally, change it as it moves toward its target. One of the best methods for securing data in transit is encryption. Both kinds of data can be identified in Fig. 5.

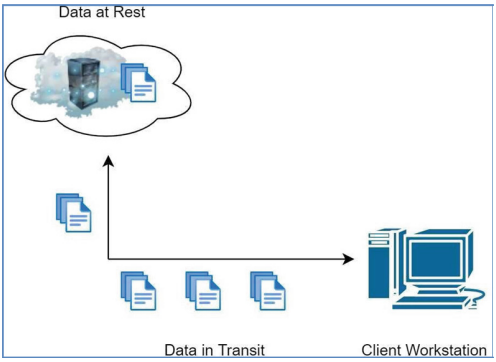

Fig. 5. Data-at-rest and data-in-transit

6 Major Security Challenges

There is no question that using the cloud instead of more established on premise systems has the ability to increase security. The term "potential" is crucial here. Businesses may not always benefit from increased protection when they go to the cloud, despite the fact that it might be more secure. The following are the primary obstacles to overcome:

6.1 Internal Attacks

The design of cloud computing networks might occasionally put customers' security and privacy at risk. Despite the fact that it occurs infrequently, this risk is extremely difficult to manage. For instance, admins and managers may occasionally act as nefarious agents, endangering the security of customers who utilize cloud computing services.. This type of attack is also known as Insider Attack in Security paradigm. Types of Internal Attacks are showcased in Fig. 6.

Fig. 6. Internal attacks

6.2 Partial/Incomplete Data Deletion

It is very critical to erase all the data from the cloud when there is no need of it or the client has asked to do so. If the data is not deleted precisely and the same resources are allocated to some other user, then there is a high probability of information leakage. This makes it more difficult for customers to sign up for cloud computing services.

6.3 Interception of Data

When compared to traditional computing, cloud computing divides and distributes data while it is in transit. Because of the weakness and fragility of computing technologies, attacks like reply assaults, 3rd party attacks and spoofing-snipping attacks offers agreater hazard.

6.4 Failure of Isolation

The pooling of resources that occurs as a result of cloud computing's multi-tenancy is a problematic quality in and of itself. For a business, not having separate storage might be deadly. The use and adoption of cloud-based services are considered as being significantly hampered by additional concerns regarding guest hop attacks and their effects.

7 Using Encryption for Data Protection

Encryption is the method of converting plain text into encrypted text using an algorithm in order to make sure that private information is unreadable by unauthorized users. Encrypted data often looks as a long list of random letters and numbers. Once information has been encoded, using the proper encryption key is the only method to decode it and regain access to it.

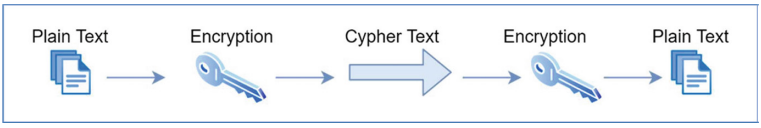

Fig. 7. Cryptography process

These days, several cryptographic algorithms are utilized to encrypt data. We use a key in order to generate cipher text and vice versa to get the plain text back in its actual form. The process of cryptography is shown in Fig. 7.

Cryptography typically has four basic applications:

7.1 Block-Cipher

Block cypher is an encryption technique that generates a cipher text of the same size as the input, say b bits, using a set input size of b bits. The dissemination of a block cypher is high. Because symbols cannot readily be inserted in the midst of a block, it is also quite challenging for an attacker to insert them without being noticed.

This technique ensures that similar blocks of text in a message are not encrypted in the same way. The cypher text from the preceding encrypted block is often used to encrypt the next block in the sequence. Block Cipher Algorithm is depicted in Fig. 8.

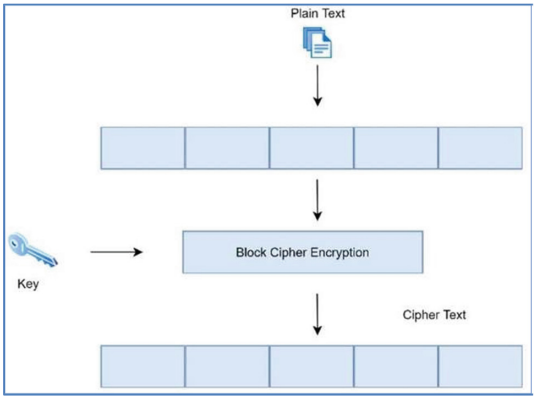

Fig. 8. Block cipher

7.2 Stream Ciphers

A stream cypher is an encryption method that converts plain text into codes that is unintelligible to anybody without the right key by working byte by byte. The same key is used to encrypt and decode messages with stream cyphers since they are linear.

Stream cyphers often operate more quickly than block cyphers due to their minimal hardware complexity. However, if not handled appropriately, this strategy might lead to major security issues. Figure 9 clearly shows the representation of Stream Cipher.

Stream cyphers are dependent on:

Plaintext: A message that you want to encrypt must already exist.

Keystreams: The plaintext characters are replaced with a set of random characters. There may be symbols, characters, or numbers among them.

Cipher text: The message that has been encrypted.

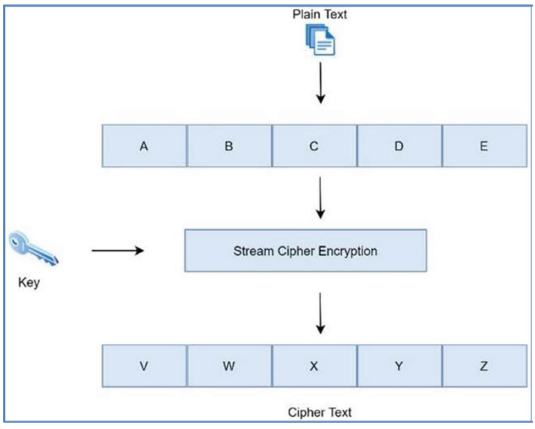

Fig. 9. Stream cipher

Stream cypher, as shown in Fig. 9, encrypts each bit rather than a block of text using an encryption key.

7.3 Hash Functions

Hash Function is a mathematical technique that involves a procedure that turns plaintext material of any size into a singular cipher text of a predetermined length. Generally, the size of the alphanumeric string output is not changed. This method makes guarantee that no two words will produce identical alphanumeric strings.

This hash function might be as basic as the one provided in Eq. (1) or also it might be quite sophisticated.

$$F(x) = x \bmod 20 \tag{1}$$

The method of hash function cryptography is depicted in Fig. 10.

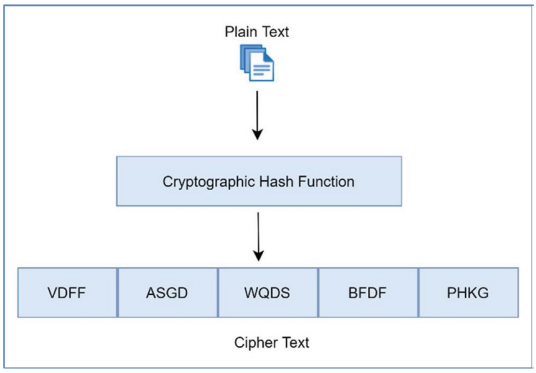

Fig. 10. Hash function mechanism

To protect data security, all of the methods and strategies covered above are frequently utilized to protect data in the cloud. Depending on the circumstance, various tactics are applied differently. Despite the type of technology used, it is highly recommended to safeguard information security both in private and public clouds.

8 Case Study on Cloud Based Cyber Security Model for Identification of Safe and Malicious Request

We try to depict a model in order to filter out the safe requests that might go in and out of the cloud every now and then. It is very crucial to check the requests as there is a high probability that it can contain malicious contents that can harm the integrity of the cloud data system.

The Table 1 shows the parameters of the data that has been used to implement the Case Study.

Table 1. Parameters of dataset

Parameter	Meaning
Base URL	URL of Application
Title of Person	Person's Title. Ex. Mr/Mrs
Name	Name of the Person
Body of Data	Message/Content
Host	Website Host
User-Agent	Request Header
Content-Type	Type - json/html etc
User's Session ID	Unique ID
Content Length	Length of Data
User Role	Role of Person
Protocol	Type of Protocol used. Ex. HTTP
IP Address	IP Address of User
isSafe	Added Parameter to mark a request as Safe/Unsafe

Models Used for filtration of data:

1. Exploratory Data Analysis (EDA) for Pre Processing of Data
It is a strategy for data analysis that uses visual methods. With the use of statistical summaries and graphical depiction, it is used to identify patterns, trends, and also to verify presumptions (Fig. 11).

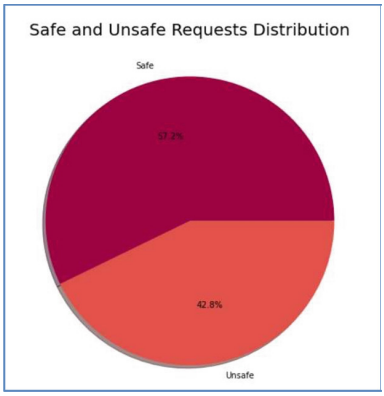

Fig. 11. Actual data distribution for safe and unsafe data

2. Bag-of-Words Model
Natural language processing employs the text modelling technique known as "bag of words." To explain it formally, it is a method for extraction of features from text data. This approach makes it simple and flexible to extract traits from documents.

Confusion Matrix for the above model is depicted in Table 2.

Table 2. Confusion matrix for bag of words model

49	32
0	119

Results of the model:

Table 3. Result obtained for Bag of Words Model

	Precision	Recall	F1 score	Support
0	1	0.6	0.75	81
1	0.79	1	0.88	119
Accuracy			0.84	200
Macro Average	0.89	0.8	0.82	200
Weighted Average	0.87	0.84	0.83	200

We could achieve an Accuracy of 84% for the above model and it is showcased in Table 3.

3. TF-IDF Technique
Term Frequency - Inverse Document Frequency scale is used to assess a word's uniqueness. Sentences are transformed into vectors (after tokenization, stemming, and lemmatization). The semantic significance of the term is not provided by the Bag of Words approach in this situation; instead, the TF-IDF is used.

We achieved an accuracy of 84% with this method too.

4. Deep Learning Model
Tokenization is the initial stage in text data modelling. To create tokens, the corpus is tokenized. The following step is to construct a glossary using the below-listed tokens. These are then classified as safe and malicious requests. To achieve this, Tensor Flow's tokenization has been used. An Accuracy of 84.5% could be achieved using this model which is depicted in Fig. 12.

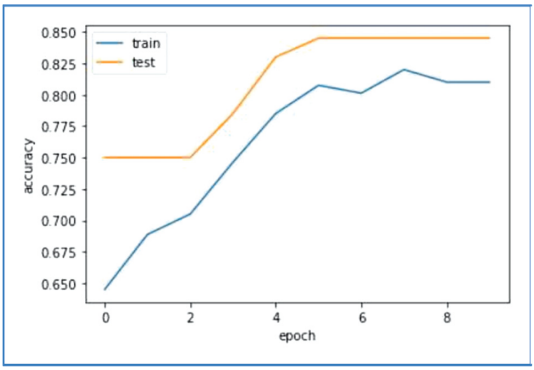

Fig. 12. Accuracy of deep learning tokenization model

9 Conclusion

The trend for improved cloud data storage techniques is surely intensifying as cloud computing is increasingly used for data storage. If data stored in the cloud is not properly safeguarded, it may be at danger. The hazards and security threats to data in the cloud were explored in this study. Also, an overview of three categories of security problems was discussed. The hazards that the hypervisor poses are examined when it comes to virtualization.

Multitenancy and risks associated with public clouds have also been addressed. One of the primary subjects of the article was data security, including its difficulties and solutions in cloud computing. It has been looked at how to encrypt data in the cloud using efficient methods and data at various stages.

The study examined block, stream, and hash ciphers—all of which are used to encrypt data in the cloud, whether it is at rest or in transit. An overview of Hash Functions is also provided discussing its functionalities and usage. Finally, a Case Study on Cloud based Cyber Security Model for Identification of Safe and Malicious Request has been presented which applies different models to filter the request and an accuracy of 85% has been achieved for the same.

References

1. Srinivas, J., Reddy, K., Qyser, A.: Cloud computing basics. Build Infrastruct. Cloud Secur. **1**(2011), 3–22 (2014)
2. Vouk, M.A.: Cloud computing - issues, research and implementations. In: Proceedings of the International Conference on Information Technologies Interfaces, ITI, pp. 31–40 (2008)
3. Wooley, P.S.: Identifying cloud computing security risks. Contin. Educ. **1277** (2011)
4. Alharthi, A., Yahya, F., Walters, R.J., Wills, G.B.: An Overview of Cloud Services Adoption Challenges in Higher Education Institutions (2015)
5. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. J. Netw. Comput. Appl. **34**(1), 1–11 (2011)
6. Ranchal, R., Bhargava, B.: Epics: A framework for enforcing security policies in composite Web services. IEEE Trans. Services Comput. **12**(3), 12–22(2019

7. Martin, K.: The penalty for privacy violations: how privacy violations impact trust online. J. Bus. Res. **82**, 103–116 (2018)
8. Ali, T.: The next generation cloud technologies: a review on distributed cloud, fog and edge computing and their opportunities and challenges. Res. Berg Rev. Sci. Technol. **1**(1), 1–15 (2021)
9. Ramesh, N.P., Guruprasad, N., Dankan Gowda, V.: A high-availability and integrity layer for cloud storage, cloud computing security: from single to multi-clouds. J. Phys. Conf. Ser. **1921**(1) (2021). (IOP Publishing)
10. Abhishek, S., et al.: A comparative analysis of security issues & vulnerabilities of leading cloud service providers and in-house university cloud platform for hosting e-educational applications. In: 2021 IEEE Mysore Sub Section International Conference (MysuruCon). IEEE (2021)
11. Tawalbeh, L.A., Saldamli, G.: Reconsidering big data security and privacy in cloud and mobile cloud systems. J. King Saud Univ. Comput. Inform. Sci. **33.7**, 810–819 (2021)
12. Bhardwaj, A., Krishna, C.R.: Virtualization in cloud computing: moving from hypervisor to containerization—a survey. Arab. J. Sci. Eng. **46**, 8585–8601 (2021)
13. Ali, O.: Assessing information security risks in the cloud: a case study of Australian local government authorities. Gov. Inf. Q. **37**(1), 101419 (2020)
14. Identity, I., Anand, P.M.R., Bhaskar, V.: Identity and access management in cloud environment: mechanisms and challenges. Eng. Sci. Technol. Int. J. **21.4**, 574–588 (2018)
15. Wu, H., et al.: Network security for virtual machine in cloud omputing. In: 5th International Conference on Computer Sciences and Convergence Information Technology. IEEE (2010)
16. Gururaj, R., Iftikhar, M., Khan, F.A.: A comprehensive survey on security in cloud computing. Procedia Comput. Sci. **110**, 465–472 (2017)
17. Hassan, R.: Data and infrastructure security auditing in cloud computing environments. Int. J. Inform. Manage. **34.3**, 364–368 (2014)
18. Hanan, S., et al.: Cloud computing virtualization of resources allocation for distributed systems. J. Appl. Sci. Technol. Trends **1.3**, 98–105 (2020)
19. Theodor, B., et al.: Digital transformation of manufacturing through cloud services and resource virtualization. Comput. Indust. **108**,150–162 (2019)
20. Xue, Y., et al.: An attribute-based controlled collaborative access control scheme for public cloud storage. IEEE Trans. Inform. Forens. Secur. **14.11**, 2927–2942 (2019)
21. Ru, J., et al.: A systematic review of scheduling approaches on multi-tenancy cloud platforms. Inform. Softw. Technol. **132**, 106478 (2021)
22. Kandukuri, B., Paturi, V., Rakshit, A.: Cloud security issues. In: International Conference on Services Computing, pp. 517–520. IEEE (2009)
23. Almulla, S., Yeun, C.: Cloud computing security management. In: 2nd International Conference on ICESMA, pp 1–7. IEEE (2010)
24. Palivela, H., Chawande, N., Wani, A.: Development of server in cloud computing to solve issues related to security and backup. In: IEEE CCIS, pp 158–163 (2011)
25. Laura, A., Moro, R.: Support vector machines (SVM) as a technique for solvency analysis. DIW Berlin discussion paper (2008)
26. Zhang, X., Zhao, Y.: Application of support vector machine to reliability analysis of engine systems. Telkomnika **11**(7), 3352–3560 (2013)
27. Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice Hall, Englewood Cliffs, NJ (2009)
28. Michalski, R., Carbonell, J., Mitchell, T.: Machine Learning: An Artificial Intelligence Approach. Springer, Berlin (2013)