electronics

Review

A Review of Machine Learning Algorithms for Cloud Computing Security

Umer Ahmed Butt [1], Muhammad Mehmood [1], Syed Bilal Hussain Shah [2], Rashid Amin [1], M. Waqas Shaukat [1], Syed Mohsan Raza [3], Doug Young Suh [4,*] and Md. Jalil Piran [5,*]

[1] Department of Computer Science, University of Engineering and Technology, Taxila 47080, Pakistan; umerahmed.butt@yahoo.com (U.A.B.); muhammad.mehmood@students.uettaxila.edu.pk (M.M.); rashid.amin@uettaxila.edu.pk (R.A.); waqaxch@gmail.com (M.W.S.)
[2] School of Software, Dalian University of Technology, Dalian 116000, China; bilalshah@dlut.edu.cn
[3] Department of Computer Science, Abasyn University, Peshawar 25000, Pakistan; smohsanraza@gmail.com
[4] Department of Electronics Engineering, Kyung Hee University, Yong-in 17104, Korea
[5] Department of Computer Science and Engineering, Sejong University, Seoul 05006, Korea
* Correspondence: suh@khu.ac.kr (D.Y.S.); piran@sejong.ac.kr (M.J.P.)

Received: 19 July 2020; Accepted: 9 August 2020; Published: 26 August 2020

Abstract: Cloud computing (CC) is on-demand accessibility of network resources, especially data storage and processing power, without special and direct management by the users. CC recently has emerged as a set of public and private datacenters that offers the client a single platform across the Internet. Edge computing is an evolving computing paradigm that brings computation and information storage nearer to the end-users to improve response times and spare transmission capacity. Mobile CC (MCC) uses distributed computing to convey applications to cell phones. However, CC and edge computing have security challenges, including vulnerability for clients and association acknowledgment, that delay the rapid adoption of computing models. Machine learning (ML) is the investigation of computer algorithms that improve naturally through experience. In this review paper, we present an analysis of CC security threats, issues, and solutions that utilized one or several ML algorithms. We review different ML algorithms that are used to overcome the cloud security issues including supervised, unsupervised, semi-supervised, and reinforcement learning. Then, we compare the performance of each technique based on their features, advantages, and disadvantages. Moreover, we enlist future research directions to secure CC models.

Keywords: cloud computing; cloud security; security threats; cybersecurity; machine learning; network-based attacks; VM-based attacks; storage-based attacks; application-based attacks

1. Introduction

Cloud Computing (CC) has recently arisen as a new framework for facilitating and delivering services over the Internet [1]. The common financial restrictions and growing computational cost require storage, analysis, and presentation of data that have imposed critical modifications for the present day cloud model [2,3]. CC is the on-demand accessibility of end-users' resources, especially information storage and processing power, without a direct special organization by the client. Distributed computing is a popular articulation that implies different things to different people. Distributed computing offers public and private data to the client on a single platform across the Internet [4]. However, CC has several security challenges that delay the rapid adoption of the computing model, such as vulnerability for client and associations [5,6].

Edge computing, a version of CC used to process time-sensitive data, offers application designers and service providers distributed computing ability at the edge of a system [7]. Current edge processing extends this methodology through virtualization innovation to simplify sending and operating a more extensive scope of use on edge servers. The distributed concept of this paradigm presents a shift in security plans used in distributed computing. In addition encrypting information unique encryption systems should be embraced because information may travel between various distributed hubs associated with the Web before eventually arriving at the cloud. Edge hubs may likewise be asset-obliged devices, restricting the decision concerning security strategies. By maintaining information at the edges, it is conceivable to move the responsibility for information from service providers to end-users.

The center concept is to permit computers to modify without human mediation or help and change activities as prerequisites. CC has service models such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), deployment models such as public, private, community, and hybrid cloud are also discussed. The major security concerns in CC are classified under integrity, availability, and confidentiality threats. Cloud services range from information storage to managing software services, with requirements for unlimited availability. CC is usually designed as an impervious surrounding that can supply architecture, offerings, and processing power and strength on request [8]. The cloud model prefers and supports large applications of hardware assets (for providing the supported service), and infrastructure [9]. As a new model for computing, CC has challenges despite its advantages. Not all cloud deployment styles are appropriate for every service, every provider customer, or all involved parties [10]. This paper describes the security issues and challenges in CC and associated solutions using Machine learning (ML) algorithms.

ML algorithms are used to solve security issues and manage data more efficiently [11]. ML is the use of man-made consciousness that enables frameworks to normally take in and improve truly without being expressly customized [12]. ML focuses on the advancement of computer programs that can find a suitable pace use it to learn for themselves [13]. The approach toward learning starts with perceptions or data, such as models, direct understanding, or heading, to channel for structures in information and pick better decisions later on the subject to the models that are given.

The purpose of this paper is to present an analysis of the legal issues and security threats in distributed computing using ML algorithms. The growing recognition of distributed computing offerings is an integral driver behind changing it into a possible enterprise suggestion for decreasing the price of both architecture and operations, Accordingly, manipulating safety and privacy risks is critical in a distributed environment as properly tool fault troubles associated with it [14]. Safety problems and issues associated with distributed computing using ML algorithms are analyzed and discussed with valid steps to justifying such issues. The major inspiration was developed as a request for distributed services that are efficient, value impressive, and secure. Currently, users aware of choosing their cloud carrier contributor usually cannot manage to pay for any agreement on security risk and privacy, which is an extensive assignment for cloud system service suppliers. Using different algorithms it is critical to evaluate the security troubles faced with the aid of cloud provider vendors and the associated criminal implications [10]. The primary problem that we study is security threats in distributed computing. We explain the algorithms used to solve issues and improve performance. Moreover, we discuss the use of various ML algorithms to solve security threats in CC. Furthermore, we present research directions that should be studied in the future.

The rest of the paper is organized as follows. Section 2 discusses the related survey articles and compares them with our review paper. Section 3 presents the background study of CC models, threats, and attacks and an introduction to ML algorithms. Section 4 discusses ML techniques for CC security and the objective, techniques, advantages, and disadvantages of all papers discussed in this section. Furthermore, we present research directions that should be studied in future in Section 5. Finally, Section 6 concludes the paper.

2. Related Work

In this section, we study related papers that considered the issue of cloud security using ML algorithms. Then, we discuss the comparison of the related papers with our paper.

Khan et al. [15] discussed an algorithm to solve security issues to improve the performance of the cloud system. The lack of interest of information still exists because of information instability by the outsider, which stores, manages, and forms the information. The authors used artificial neural networks (ANNs) over the scrambled information. Khilar et al. [16] surveyed trust-based security issues and challenges in the cloud models. They defined CC as an appropriated processing condition that hosts devoted registering resources to any time from any place. This creates flexibility of information, information inescapability, and versatility. The authors proposed a trust-based access control model as an efficient method for security in distributed computing frameworks. The primary motivation behind their model is to offer access to an approved client in the cloud and choose an asset for the calculation. Both the client and cloud assets are assessed based on their trust estimation.

The authors in [17], discussed cloud security issues and models. The authors examined the distinctive security issues of distributed computing resulting organization movement models. However, the crucial advancement of the cloud without any other person produces the possibility of significant security. The methodology of a moved model should not wrangle with the necessary functionalities and capacities present in the current model. Another model concentrating on improving the features of a presentation model must not risk or undermine other critical features of the current model.

Bhamare et al. in [18], discussed ML models to improve data security. The concept of distributed computing was discussed because it expands critical traction and virtualized server farms becoming well-known as a practical framework and an answer to enormous business applications [19]. The researchers used an assistance model to solve security threats and protection challenges in distributed computing [20]. They examined basic threats and protection challenges in distributed computing, various existing arrangements, and analyzed their strengths and limitations.

The authors in [21], discussed the CC threat classification model based on the feasibility of ML algorithms to detect and resolve security issues. Furthermore, they proposed the CC risk grouping model based on the feasibility of ML algorithms to distinguish them. ML algorithms and defensive techniques were used to solve security threats and issues in CC [22,23]. Furthermore, they identified five notable trends introduced in the search for security threats and defensive strategies of ML that merit examination.

Selamat et al. in [24], studied ML algorithms used to solve malware security threats and security in CC. The researchers proposed a barrier framework that uses three ML algorithm examination and selected them based high-accuracy malware discovery. Shamshirband et al. in [25] introduced a complete review of interruption recognition frameworks that use a computational insight strategy in a mobile cloud condition. Their study presented provided a diagram of CC and mobile CC (MCC) standards and administration models, likewise auditing security dangers in these unique circumstances.

The previous literature studied security issues and threats using either one or two ML techniques to solve the problem of cloud security. This paper discusses multiple types of ML algorithms to solve cloud security issues. Table 1 presents a comparison of our review paper with the related papers. We also compare different algorithms and verify which techniques are optimal for solving the issue. For the main comparison with other surveys and papers we use a different supervised and unsupervised algorithm to analyze the problem and solve the legal issues.

Table 1. Comparison of related research.

Reference	Year	Areas Focused	ML Techniques	Security Issues	Impact in Cloud
[21]	2019	Protection preserved encrypted data	Supervised and unsupervised learning	Limited	Minor or Intermediate Issues
[24]	2019	Trust-based access control	Unsupervised learning	No	A few solutions accessible
[25]	2020	Security issues	Supervised and unsupervised learning	Limited	Minor issues
[26]	2011	Security and threat issues	Supervised learning	Yes	Long term issues
[27]	2016	Security issues and datasets	Supervised learning	Limited	Minor or intermediate issues
[28]	2018	Cloud Security	Supervised and unsupervised learning	Limited	Minor or intermediate issues
[29]	2017	Cloud threats classification	Supervised and unsupervised learning	No	A few solutions accessible
[30]	2019	Malware security threats and protection	Supervised learning	Yes	Long term issues
[31]	2020	Security and threat Issues	Supervised learning	Limited	Minor or intermediate issues

3. Background Study

3.1. CC

CC is an on-demand accessibility of end users' resources, particularly information storage and processing power, without an immediate one-of-a-kind association with the client [26]. The term is normally used to describe server farms open to different clients over the Internet [27]. Huge clouds, common today, have limits ignored in various areas from central servers [32]. If the relationship with the customer is respectably close, it might be assigned to an edge server [33].

CC contains three explicit types of organizations preparation and prodices these remotely using the Web. Clients regularly pay monthly or yearly to cover supplier costs, to find a workable pace convey to IaaS, PaaS, and SaaS to endorsers. Longer periods of huge capital interests in programming and information technology (IT) foundations are now obsolete for any effort to manage using the circulated processing model for procurement organizations [34]. The capacity to identify functional IT resources on consistent explanation is making everything sensible for small and medium-sized affiliations, providing them the key instruments and development to fight in general business arena, without the need for on-premise IT assets [26]. Customers who become tied up with enrolling organizations who have avoided cloud can significantly decrease the IT organization uses for their affiliations, and access sensibly sorted out and flexible endeavor level figuring organizations, all the method [35].

3.1.1. Cloud Service Models

With the developing appropriation of cloud infrastructures to convey several IT administrations, observing organization execution has become pivotal [9]. Cloud suppliers likely disclose subjective data about system execution, which hinders productive cloud selection, and causes performance issues, vulnerabilities concerning the conduct of facilitated administrations, and imperfect deployment decisions [36]. CC service models have three types: IaaS, SaaS, and PaaS. These cloud models serve various businesses and have different benefits. Figure 1 illustrates the CC service models.

- IaaS; has many benefits but also some issues. IaaS provides the infrastructure through the virtual machine (VM), but VMs are gradually becoming obsolete. This is due to mismatching the cloud to provide security and VM security. Data deletion and issues can be solved by deciding the time frame for data deletion by both the client and the cloud provider. Compatibility issue occurs in IaaS as client-only run legacy software, which may increase the cost [10]. The security of the hypervisor is important splitting physical resources between the VMs.
- PaaS; is a web-based software creation and delivery platform offered as a server for programmers, enabling the application to be developed and deployed [10]. The security issues of PaaS are inter-operation, host vulnerability, privacy-aware authentication, continuity of service, and fault tolerance.

- SaaS; has no practical need for indirect deployment because it is not geographically dispersed and is delivered nearly immediately. Security issues in the SaaS are authentication, approval, data privacy, availability, and network security [28].

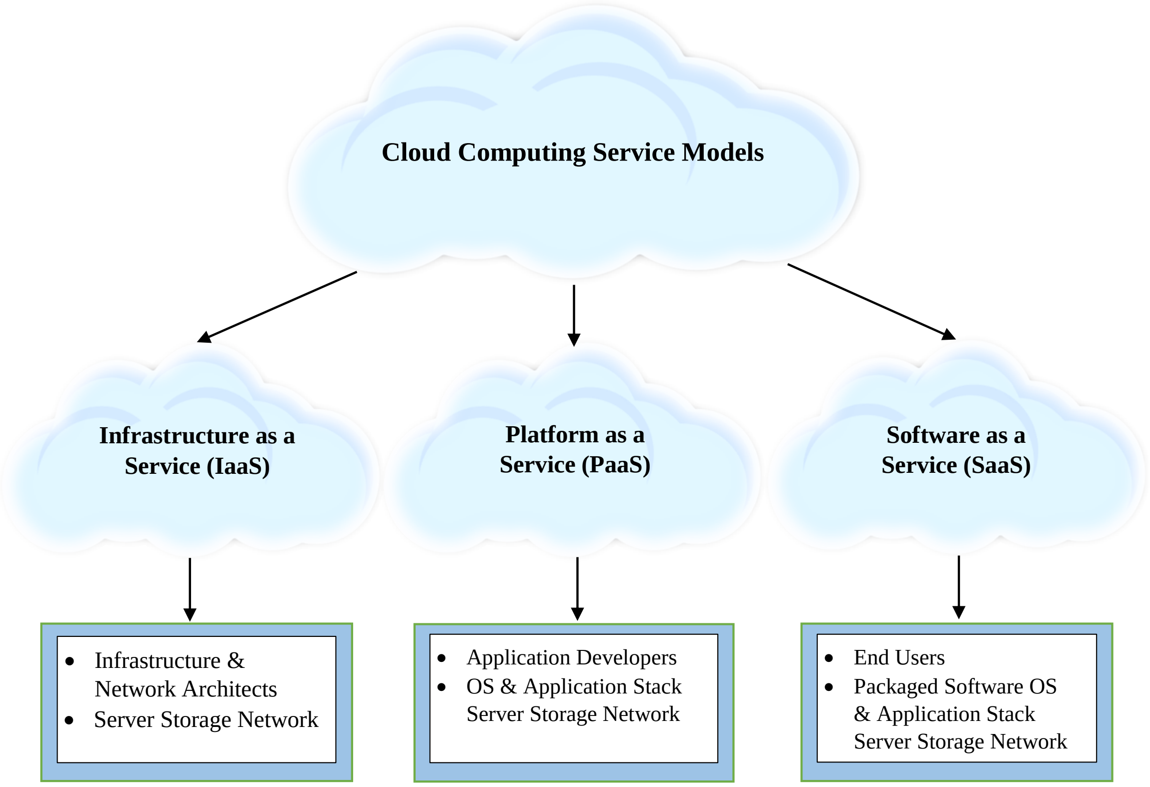
Cloud Computing Service Models
Infrastructure as a Service (IaaS)
Platform as a Service (PaaS)
Software as a Service (SaaS)
• Infrastructure & Network Architects
• Server Storage Network
• Application Developers
• OS & Application Stack Server Storage Network
• End Users
• Packaged Software OS & Application Stack Server Storage Network

Figure 1. Cloud Computing (CC) service models.

3.1.2. Design of the Cloud

As indicated by CC architecture, the five most significant elements affect and are affected by CC, alongside its security suggestions. Figure 2 illustrates the design of CC that includes a start-to-finish reference design that represents the layers of the Open Systems Interconnection (OSI) Model. CC is a complicated design with multiple zones of vulnerability [37]. The components of CC are as follows:

- Cloud Consumer: An individual or association that maintains career, relationship, and utilization administrations from the cloud providers [29].
- Cloud Provider: An individual or organization for manufacturing, or administration, available to invested individuals.
- Cloud Auditor: A gathering that can direct the self-sufficient examination of cloud organizations, information system activities, implementation, and security of cloud users.
- Cloud Broker: A substance that manages the usage, implementation, and conveyance of cloud benefits and arranges links between cloud purchasers and cloud suppliers [29].
- Cloud Carrier: A medium that offers a system of cloud administrations from cloud suppliers to the cloud consumers.

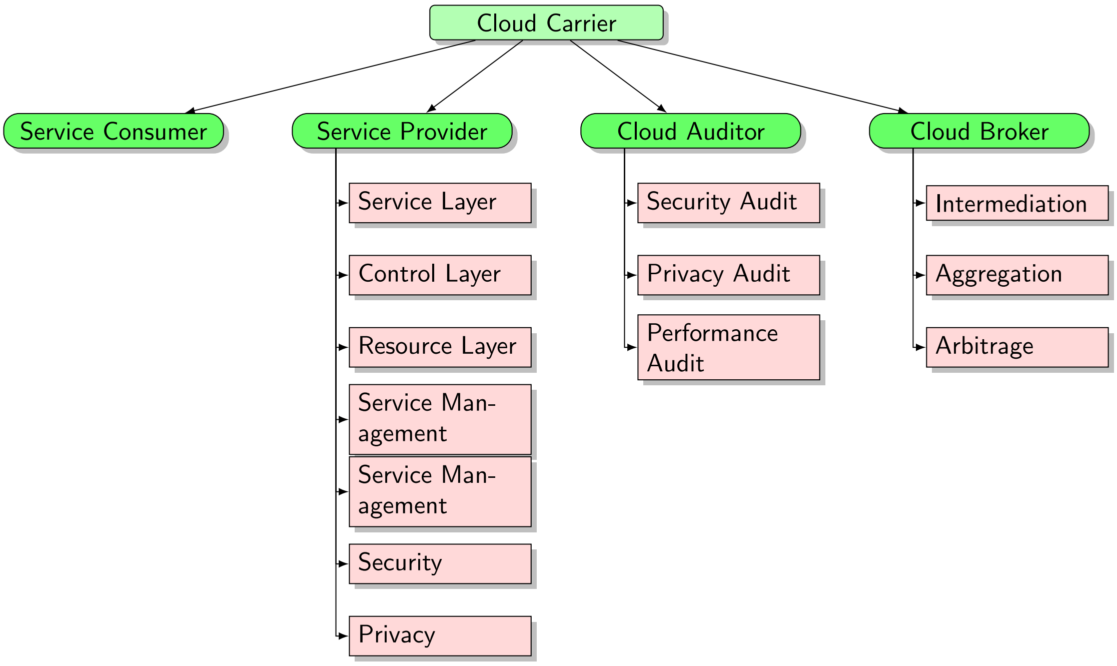

Figure 2. Cloud computing architecture.

3.1.3. Cloud Deployment Models

CC has four deployment models: private, public, hybrid, and community cloud [28]. Each deployment model has different costs and value propositions. Therefore, deciding the deployment model is a difficult and critical decision. Figure 3 illustrates the cloud deployment models. Table 2 presents a comparative analysis of the benefits and issues of cloud models.

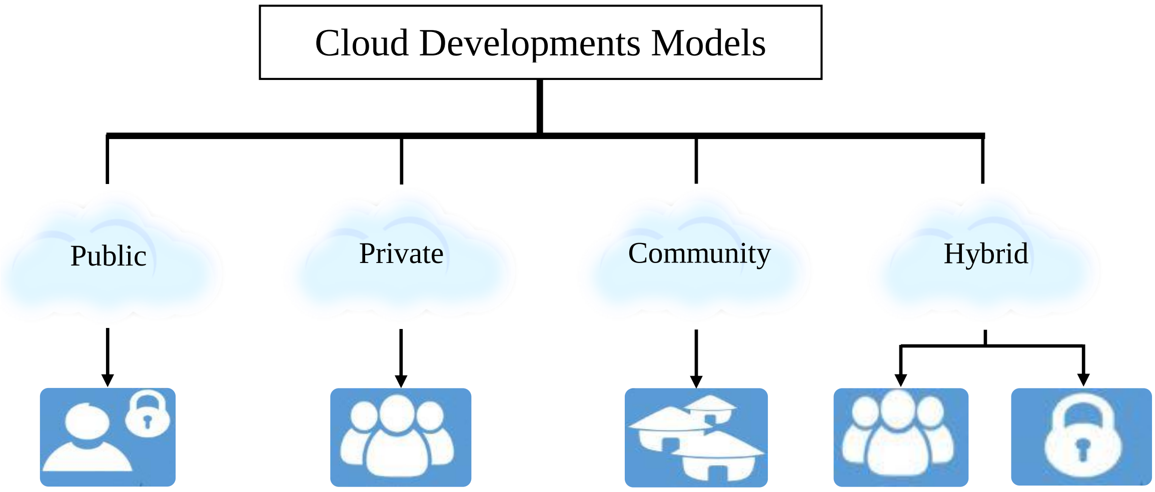

Figure 3. Cloud deployment models.

Table 2. Comparative analysis of the cloud deployment models.

Cloud Models	Pros	Cons
Public	• High scalability • Flexibility • Cost-effective • Reliability • Location independence	• Less secure • Less customizability
Private	• More reliable • More control • High security and privacy • Cost and energy efficient	• Lack of visibility • Scalability • Limited services • Security breaches • Data loss
Community	• More secure than public Cloud • Low cost than private Cloud • More flexible and Scalable	• Data segregation • Responsibilities allocation within the organization
Hybrid	• High scalability • Low cost • More flexible • More secure	• Security compliance • Infrastructure dependent

3.2. Cloud Threats

CC is a developing model that has significant potential to grow and is becoming widely popular. However, even with its unique characteristics, it has various security threats and protection challenges, as discussed in this section [30]. The categorization is performed based on the CIA Triad and attacks on cloud components [31,38].

3.2.1. Cloud Security Threats

The major security threats in CC are classified under confidentiality, integrity and availability. These issues are discussed briefly here.

1. Confidentiality threats involves an insider threat to client information, risk of external attack, and data issues [39]. First, insider risk to client information is related to unapproved or illegal access to customer information from an insider of a cloud service provider is a significant security challenge [31]. Second, the risk of outside attack is increasingly relevant for cloud applications in unsecured area. This risk includes remote software or hardware hits on cloud clients and applications [40]. Third, information leakage is an unlimited risk to cloud bargain data because of human mistake, lack of instruments, secured access failures, after which anything is possible.
2. Integrity threats involve the threats of information separation, poor client access control, and risk to information quality. First is the risk of information isolation, which inaccurately joins the meanings of security parameters, ill-advised design of VMs, and off base client-side hypervisors. This is complicated issue inside the cloud, which offers assets connecting the clients; if assets change, that could affect information trustworthiness [41,42]. Next is poor client access control, which because of inefficient access and character control has various issues and threats that enable assailants harm information assets [43,44].
3. Availability threats include the effect of progress on the board, organization non-accessibility, physical interruption of assets, and inefficient recovery strategies. First is the effect of progress on the board that incorporates the effect of the testing client entrance for different clients, and the effect of foundation changes [31]. Both equipment and application change inside the cloud condition negatively affect the accessibility of cloud organizations [45]. Next is the non-accessibility of services that incorporate the non-accessibility of system data transfer capacity, domain name system (DNS) organization registering software, and assets. It is an external risk

that affects all cloud models [46]. The third is its physical disturbance IT administrations of the service providers, cloud customers, and wide area network (WAN) specialist organization. The fourth are weak recuperation techniques, such as deficient failure recovery which impacts recovery time and effectiveness if there should develop an occasion of a scene.

3.2.2. Attacks on the Cloud

Four relative analysis attacks are classified by their segments: network-based, VM-based, storage-based, and application-based as depicted in Figure 4 [46].

1. Network-based attacks: Three types of system attacks discussed here are port checking, botnets, and spoofing attacks. A port scan is useful and of considerable interest to hackers in assessing the attacker to collect relevant information to launch a successful attack [46]. Based on whether a network's defense routinely searches ports, the defenders usually do not hide their identity, whereas the attackers do so during port scanning [47]. A botnet is a progression of malware-contaminated web associated devices that can be penetrated by hackers [48,49]. A spoofing assault is when a hacker or malicious software effectively operates on behalf of another user (or system) by impersonating data [46]. It occurs when the intruder pretends to be someone else (or another machine, such as a phone) on a network to manipulate other machines, devices, or people into real activities or giving up sensitive data.
2. VM-based attacks: Different VMs facilitated on a frameworks cause multiple security issues. A side-channel assault is any intrusion based on computer process implementation data rather than flaws in the code itself [25]. Malicious code that is placed inside the VM image will be replicated during the creation of the VM [46]. VMs picture the executive's framework offers separating and filtering for recognizing and recovering from the security threats.
3. Storage-based attacks: A strict monitoring mechanism is not considered then the attackers steal the important data stored on some storage devices. Data scavenging refers to the inability to completely remove data from storage devices, in which the attacker may access or recover this data. Data de-duplication refers to duplicate copies of the repeating data [50]. This attack is mitigated by ensuring the duplication occurs when the precise number of file copies is specified.
4. Application-based attacks: The application running on the cloud may face many attacks that affect its performance and cause information leakage for malicious purposes. The three primary applications-based attacks are malware infusion and stenography attacks, shared designs, web services, and convention-based attacks [46].

3.3. ML and Cloud Security

ML is the logical examination of calculations and measurable models that computer systems use to implement a specific endeavor without using express headings, contingent upon models, and acceptance. It is a subset of computerized reasoning [51]. ML is so significant in the cloud that each cloud will use ML in near future [52]. In this section, discussions about asset designation and focus, rather than on how ML can help with the security of distributed computing will be presented [53].

With the expansion of general information in the cloud, there has likewise been an expansion of delicate information in the cloud, motivating the requirement for higher security in CC. This section describes methodologies proposed to improve cloud security using more accurate risk identification [54]. We begin by describing a general methodology to decide threats and dangers through the summation of hazard levels. Then, we describe the approaches for addressing dangers that use signature recognition and anomaly detection to create a half-and-half model for threat detection [55].

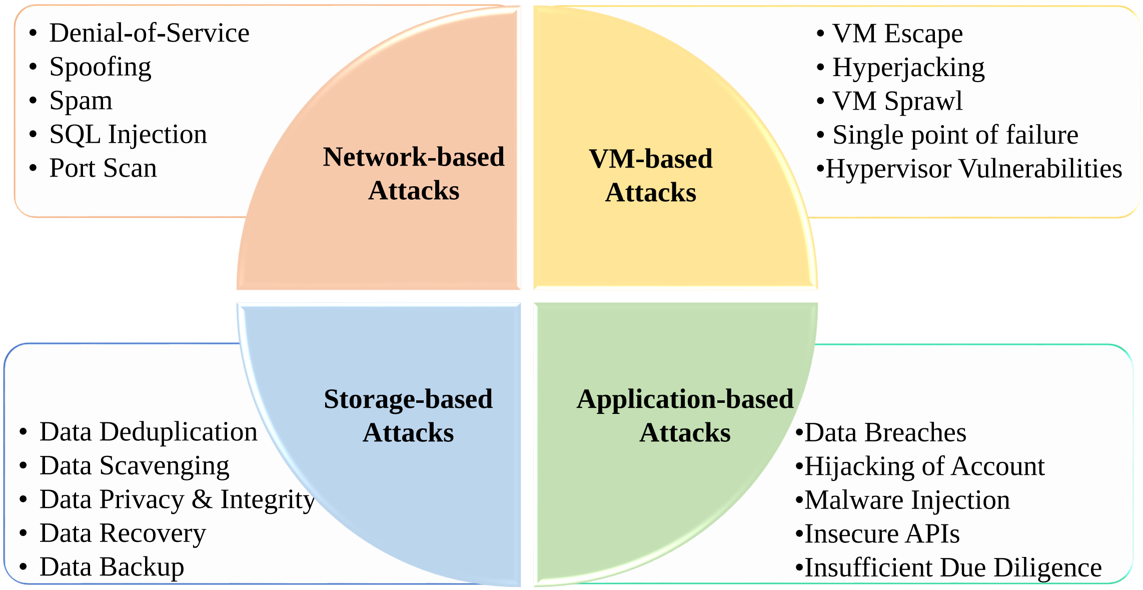

Figure 4. Attacks on cloud components.

3.3.1. Types of ML Algorithms

1. Supervised learning is an ML task of learning a function that maps a contribution to the yield subject to procedure data yield sets. It prompts a capacity for naming data involving many of the preparation models. Managed learning is a significant part of the data science [56]. Administered learning is the ML assignment of initiating a limit from named getting ready data, preparing data involves many getting ready models.

    (a) Supervised Neural Network: In a supervised neural network, the yield of the information is known. The predicted yield of the neural system is compared with the real yield. Given the mistake, the parameters are changed and afterward addressed the neural system once more. The administered neural system is used in a feed-forward neural system [57].
    (b) K-Nearest Neighbor (K-NN): A basic, simple to-execute administered ML calculation that can be used to solve both characterization and regression issues. A regression issue has a genuine number (a number with a decimal point) as its yield. For instance, it uses the information in the table below to appraise somebody's weight given their height.
    (c) Support Vector Machine (SVM): A regulated ML algorithm used for both gathering and relapse challenges. It is generally used in characterization issues. The SVM classifier is a frontier that separates the two classes (hyper-plane).
    (d) Naïve Bayes: A regulated ML algorithm that uses Bayes' theorem, which accepts that highlights are factually free. Despite this assumption, it has demonstrated itself to be a classifier with effective outcomes.

2. Unsupervised learning is a type of ML algorithm used to draw deductions from datasets consisting of information without marked reactions. The most widely recognized unsupervised learning strategy is cluster analysis, which is used for exploratory information analysis to discover hidden examples or grouping in the information [58].

    (a) Unsupervised Neural Network: The neural system has no earlier intimation about the yield of the information. The primary occupation of the system is to classify the information based on several similarities. The neural system verfies the connection between diverse source of information and gatherings.
    (b) K-Means: One of the easiest and renowned unsupervised ML algorithms. The K-means algorithm perceives *k* number of centroids, and a short time later generates each data point

to the closest gathering, while simultaneously maintaining the centroids as little as could be typical considering the present circumstance.

(c) Singular Value Decomposition (SVD): One of the most broadly used unsupervised learning algorithms, at the center of numerous proposals and dimensionality reduction frameworks that are essential to worldwide organizations, such as Google, Netflix, and others.

3. Semi-Supervised Learning is an ML method that combines a small quantity of named information with abundant unlabeled information during training. Semi-supervised learning falls between unsupervised and supervised learning. The objective of semi-supervised learning is to observe how combining labeled and unlabeled information may change the learning conduct and to structure calculations that exploit such a combination.
4. Reinforcement Learning (RL) is a territory of ML that emphasizes programming administrators should use activities in a scenario to enlarge some idea of the total prize. RL is one of three major ML perfect models, followed closely by supervised learning and unsupervised learning. One of the challenges that emerges in RL, and not in other types of learning, is the exchange of the examination and abuse. Of the extensive approaches to ML, RL is the nearest to humans and animals.

4. ML Algorithms for the Cloud Security

In this section, we study different ML algorithms that have been used to overcome the security issues in CC.

4.1. Supervised Learning

Supervised learning is the ML task of learning a limit that maps a commitment to a yield based on model data yield sets. It infers a limit from data involving many planning models. The supervised ML algorithms are those algorithms that require outside help.

4.1.1. Supervised ANNs

ANNs are the bits of a computing framework intended to recreate how the human mind analyzes and processes data. They are the establishments of ML that solve issues otherwise impossible or troublesome for humans or statistical principles.

Hussin et al. [59] predicted basic distributed computing security issues using ANN algorithms. An ANN algorithm was used to determine security issues in a banking organization. ANNs were used for improving the execution and learning neural capacities. Levenberg-Marquardt (LMBP) algorithms were used to predict the presentation for the cloud security level. LMBP is a nonlinear improvement model used to measure the exactness of the forecasts present and decrease the error between genuine yields and focus for the preparation procedure; the mean square error (MSE) is estimated to decide the presentation. The cloud Delphi procedure was used for informal social events and investigation. The Delphi strategy was used to collect information as qualified sources. The ANN algorithm was used as the measurable information model to forecast distributed computing issues. The LMBP algorithm was utilized for predicting cloud security issues. In the CC security issue with banking organizations, LMBP algorithms have been confirmed to be extremely productive for testing and prepartion systems.

It is challenging to perceive advanced attacks in cloud conditions because of the composite and circulated structures of cloud frameworks. Furthemore, various portable registering and putting away devices are connected to cloud structures to support a clients' entrance, which raises the multifaceted nature and difficulty of identifying digital assaults. Sayantan et al. [60] identified digital assaults in cloud conditions that are significant for protecting cloud arrangements from digital assaults. A proficient methodology was proposed for the discovery of digital assaults in cloud foundations just as for the remote processing devices. The proposed strategy was associated with using an ANN. The ANN was prepared by using the system traffic information on the joining connections of the cloud stops. Becuase ANN is computationally thorough, a technique that uses a hereditary calculated to diminish the number of structures mined from the system-traffic information is built up and joined in this methodology. This technique was exhibited by methods for two major informational collections of system-traffic, demonstrating improved results compared to those of present methodologies for identifying digital assaults in cloud arrangements. In the strategy used, the ordered informational collections of system-traffic are used for preparing and testing of directed ML in the ANN. The classes of assaults can be identified based on the names accessible in the preparation informational collection.

Today, the models for proficient and secure system framework structures, such as the Internet of Things (IoT)and big data analytic, are developing at a quicker pace than at any time in recent memory. Edge processing of an IoT structure is a data arrangement that is performed at or approaches the identifiers of data in an IoT system [61]. Al-Janabi et al. [62] explained secure edge computing in the IoT to secure data and improve performance. Their paper intended to rapidly review the thoughts, features, security, and use of IoT-empowered edge preparation similar to its security perspectives in our data-driven world [63]. The authors aimed to quickly audit the ideas, highlights, security, and use of IoT enabled edge processing just as its security aspects in our information-driven world. The authors explained the considerations designing a scalable, reliable, secure, and circulated computing framework. Moreover, the authors summarized the fundamental thoughts concerning security chance moderation strategies. They also explored the difficulties and opportunities in the field of edge computing. Finally, the authors surveyed two contextual investigations, keen stopping and substance conveyance arrangement (CDN), and examined various techniques in which IoT frameworks can be used to complete day-to-day tasks.

DeepRM and DeepRM2 were modified to address the cloud scheduling problem. Resource scheduling in CC is one of the most challenging jobs in which resources must be allocated to the required tasks or jobs, based on the required Quality of Services (QoS) of the cloud applications. Due to the cloud environment, uncertainty, and heterogeneity, resource allocation cannot be addressed with existing policies. El-Baghdadi et al. [64] used one of the ML emergent algorithms, deep RL (DRL), referred to DRL for Cloud Scheduling (DRLCS), to solve the problem of resource scheduling in CC. In cloud scheduling, there are different parameters to be considered, such as the required CPU, memory, job deadline, and VM load balancing, following the same approach used by DeepRM and DeepRM2. However, DeepRM and DeepRM2 only use CPU and memory parameters.

The authors in [65] proposed and executed a conventional model of the arrangement. The versatile MCC condition has been achieved using cross-stage innovation in designing Internet applications. This innovation empowered hybrid applications to be designed with code move that sudden spike in demand for various working frameworks, (for example, Android or Windows), which decreases the amount of work required from engineers, as a similar code is executed on a cell phone and in the cloud. The authors in [15] investigated both supervised and unsupervised ML capabilities through ANNs over encrypted data from a semantically secure cryptosystem based on homomorphic properties. The lack of interest of information owners still exists because of information instability by the outsider, which stores, manages, and forms the information. Information owners in the cloud need information protection and effective information the board over cloud. Example coordinating is one of the basic tools in various fields.

By employing ANN, the authors in [66] tried to detect cyberattacks in MCC. They have been able to demonstrate that their proposed framework improves the accuracy up to 97.11% in detecting the attacks. In Reference [67], ANN was used to detect intrusion and attacks. The authors tested their model for NSL-KDD and KDD-CUP datasets. They claimed that their proposed model was able to detect attacks and intrusions by unauthorized users.

MEC plans to overcome the restricted terminal battery and handling abilities related to running applications in the portable terminal and the high latency acquainted by offloading these applications with the cloud. It expands the processing resource of the cloud at the edge of the cell closer to the mobile client [68]. Asset the board the board in mobile edge registering is one of the principal issues recently, according to numerous analysts. It consists of resource distribution and computation offloading. The allocation of assets includes managing and booking the assets to achieve the requests of the clients. It relies upon the accessibility and the limit of the assets. Based on the cutoff time of each task, the service provider will allocate adequate assets to every client the adequate assets to every client. Computation offloading is executed at an outside stage (edge or then again cloud server) and depends upon the handling ability and the capacity limit of the device. It is challenging to provide an ideal answer for an asset to the executives in a unique framework due to the random varieties of undertakings required by the clients and the portability of these clients, therefore, artificial intelligence (AI) strategies are proposed to take care of this optimization issue. Zamzam et al. [69] discussed resource management using ML in mobile edge computing (MEC) to solve the issue and improve performance. The authors cutting-edge AI to advance resources in portable edge processing. They separated the research into four classifications: limiting cost, minimizing energy consumption, limiting inactivity, and limiting both latency and energy consumption. The authors classified the system model, the imperatives, and the type of AI procedures that are used in every optimization issue.

4.1.2. K-NN

K-NN is likely the simplest algorithm among the ML algorithms for relapse and classification issues. The K-NN algorithm uses information and characterizes new information based on similarity measures (e.g., distance). Classification is finished by a larger part vote to its neighbors.

The security of information in the cloud remains challenge. Different frameworks are being used to enhance cloud data security, such as data encryption. The methodologies of information security cannot be applied. The comprehension of the necessities of security is fundamental to the legitimate use of these measures. Zardari et al. [70] proposed a data classification approach based on data confidentiality. The authors described a methodology of information grouping that depends on the security and protection of information. The K-NN method of information arrangement was executed in the cloud administrations and virtual conditions. The target of using K-NN incorporates the grouping of information based on their security prerequisites. The information was grouped into two classes: touchy and non-delicate (or open) information. The order of the information helped in the recognizable proof of the information that is intended to be ensured. Only the touchy and non-open information groups were required to be ensured. The order ot security and privacy-based information was proposed using a model for distributed computing. An examination was performed on the arrangement of the information based on security needs. The commitment of this investigation is the information privacy order procedure using a K-NN classifier method of ML. The delicate and private information requires greater security and encryption by using the RSA calculation. The proposed model of the information order for the security of cloud information has been achieved in the cloud simulation test system.

Cybersecurity controls are not sufficient to protect the networks from highly-skilled cyber criminals [71]. Intrusion detection and prevention services (IDPS) are not adequate for manageing the threats; fortunately, AI and ML provide great support to IDPS and increase the rate of detection and prevention. AI provides effective results in data mining and intrusion detection, but introduces many other risks; cybersecurity experts must find a balance between risks and benefits. Cybersecurity faces several new issues, which have existed for a long time, but cybersecurity experts must to discover superior approaches to protect systems from existing issues. The two key issues are botnets (used to dispatch DDoS) and IDPSs (creates bogus alerts that divert from discovering genuine dangers). Botnets assume a significant role in DDoS assaults: the larger the botnet, the more successful the DDoS assault. Furthermore, botnets are used for wholesale fraud and stealing information. An IDPS is an innovation that system and framework managers use to recognize interruptions. After the IDPS recognizes interruption, the approved heads may receive email cautions. ML IDPS enhances defenses. ANNs and ML genetic algorithms are two extraordinary ML methods for cybersecurity. GA uses the past examples to settle on choices on new examples that the framework cannot perceive.

Cybercriminals are progressively proficient designing new devices that use AI to abuse vulnerabilities. This permits cybercriminals to hide their expectations when testing systems, and sending malware [72].

The authors in [70], proposed a data classification approach based on data confidentiality. They employed K-NN in order to classify the data based on their security requirements. In the proposed method, the data is classified into two classes, e.g., sensitive and nonsensitive data. Then, in order to take care of secrity, the authors utilized RSA algorithm to encrypt the sensitive data. Such kind of methodology help to easily decide the level of security that different data requires. The authors in [73], studied the issue of privacy preserving in e-health cloud. They proposed a system that achieves the privacy of medical dataset, symptoms, and diagnosis results and hide the data access pattern. They designed a novel privacy preserving protocol for finding k data with the highest similarity.

4.1.3. Naive Bayes

In ML, Navies Bayes classifiers are a group of basic "probabilistic classifiers" that apply Bayes' hypothesis with solid (naive) freedom suppositions between the highlights. They are among the least complex Bayesian system models.

Zekri et al. [74] designed a distributed denial-of-service (DDoS) detection system based on the C4.5 algorithm to mitigate the DDoS threat. The hidden innovations and legacy conventions contain bugs and vulnerabilities that can enable interruption by the attackers. Assaults, such as DDoS, cause serious harm and influence the performance of the cloud [75]. DDoS assaults have become one of the fundamental dangers to security. A DDoS attack executes an assault by permitting an interloper to interact with a computerized PC organization. Infected with malware, PCs and different machines (e.g., IoT devices) transform into bots (or zombie). Then, the assailant has remote control over the bots, which is known as a botnet. The traditional intrusion detection techniques have limitations such as large false alarms, noise that reduces the capabilities of the IDS by generating the rate of a false alarm, and constant updating of software to track the new threats. ML methods are acquainted with call attention to the dangers more productively than traditional IDS. Distinctive ML algorithms are used to identify the threat in a DDoS. C4.5 performed a calculation that attempts to locate the smallest decision tree. The decision tree created by C4.5 can be used for the order. After investigations, the C4.5 is being the optimal technique for grouping. From the results of the detection of DDoS using C4.5, it was found that the detection rate is more than 98%; moreover, the greater the DDoS attack duration, the higher the detection rate using this algorithm. The C4.5 algorithm produces a more accurate result than other detecting techniques and addresses the issue of incomplete data effectively; it also functions with both discrete and continuous data.

Hanna et al. [76] discussed and analyzed how to achieve mitigation for CC security risks as a basic step toward obtaining a secure and safe environment for CC. This primary point was to discuss and how to achieve moderation for distributed computing security risks as an essential advance toward acquiring a safe condition for distributed computing [77]. The outcomes indicated that using a basic decision tree model Chaid algorithm security rating for the ordering approach is a powerful strategy that empowers the leader to quantify the degree of cloud, ensuring offered types of assistance. The Naive Bayes, multilayer perceptron, SVM, choice tree (C4.5), and Partial Tree (PART) algorithm were used for secure information. Distributed computing exhibits risks that may affect administrations and data bolstered utilizing this innovation. The outcomes indicated that utilizing a basic decision tree model, (the Chaid algorithm security rating), for the grouping approach is a robust system that empowers the chief to quantify the degree of cloud making and the offered types of assistance. ML procedures can be utilized effectively for the grouping of any movement relying upon predefined classes. ML procedures are accessible from the computational intelligence network. From the access list of algorithms in ML, in [76], the authors employed Naive Bayes, a multi-layer perceptron, SVM, decision tree (C4.5), and PART for grouping the information. These algorithms were helpful in solving security threats and risks.

The authors in [78] discussed web pre-fetching schemes using ML algorithms in mobile computing to resolve the issue. Pre-fetching is one of the innovations utilized in reducing the latency of organizing traffic on the Internet. The paper proposed this innovation to use MCC conditions to deal with inactivity issues in the setting of information management. Overaggressive utilization of the pre-bringing method causes overhead and hinders the framework execution becuase of the pre-fetching of inappropriate item information storage and the capacity limit of a cell phone. An enhanced security framework that was used for intrusion detection in CC was proposed in [79]. The authors combined signature and anomaly based technique to detect the attacks. They employed Navie Bayes and some other algorithms in order to increase the efficiency of the proposed method.

4.1.4. SVM

SVM is an ML algorithm that investigates information for grouping and regression analysis. SVM is a supervised learning technique that analyzes at information and sorts it into one of two classes. An SVM outputs a guide of the arranged information with the edges between the two as far separated as could reasonably be expected.

Grusho et al. [80] discussed AI methods and models to solve information security problems. The most significant security threats to cloud-computing environments were abusive and malicious use of cloud services, architectural limitations to cloud-infrastructure access and the dynamic nature of CC environments the control-ability of all CCE-deployed VMs, remote access to computing resources, dynamic changes in the current landscape of involved VMs, the vulnerability of idle VMs, unsafe interfaces (unauthorized entry or hacking of interfaces and API4), malicious insiders (intruders–insiders), and problems with the distributed use of technologies (including those with the synchronized use of cloud infrastructure resources by various users). The data leaks, account or service breaches (vulnerabilities), unknown risk profiles (insufficient awareness of users of CC features and incipient risks and threats when transferring to a CC environment), compromising profiles and bypassing standard authentication procedures, unauthorized use (theft) of profiles, targeted cyber attacks on cloud infrastructure elements and malicious impacts on the architecture and control mechanisms of cloud-based services, DDoS5 attacks, intrusions or unauthorized agents, rogue devices, rogues into the cloud infrastructure that are closed to them, and wireless access vulnerabilities (including threats and attacks specifically for wireless access, including network configuration intelligence) are most important. The solution propose to use "instrumental" software systems that identify incidents, accumulate detailed incident information, and try to intercept them and form incident reports for IS administrators as reference models for IDPSs. IDPSs are used to identify problems with security policies, document existing threats, and maintain the participants in information exchange from violating security policies. The typology of IDPSs depends on the specific events that they should track, and the means ("channels") to implement these events. The types of such systems include decisions oriented at controlling the configuration correctness of computer networks, at wireless technologies that analyze the behavior of computer networks, and, finally, at the operational analysis of host computers. The weakness is that the system requires significant development and an infrastructure of tools and technologies of cloud security IS alternate expert tools.

Hanna et al. [76] discussed how to achieve moderation for distributed computing security risk as an essential advance toward acquiring a safe condition of distributed computing [77]. The outcomes indicated that utilizing a basic decision tree model Chaid algorithm security rating for ordering approach is a powerful strategy that empowers the leader to quantify the degree of the cloud making sure about the offered types of assistance. The Naive Bayes, multi-layer perceptron, SVM, choice tree (C4.5), and Partial Tree algorithm are utilized for secure information. Distributed computing is surrounded by numerous dangers that may effectively affect administrations and data bolstered using this innovation. The outcomes indicated that utilizing a basic decision tree model Chaid algorithm security rating for the grouping approach is a robust system that empowers the chief to quantify the degree of cloud making sure about and the offered types of assistance. From the access list of algorithms in ML, the authors chose Naive Bayes, a multi-layer perceptron, SVM, decision tree (C4.5), and PART for grouping our information. These algorithms, are helpful in solving security threats and risks.

Hou et al. [54] explained the detection of the network security of edge computing systems using ML to solve the problem. Their investigation include the manufactured simulation of a smart home framework by the Alibaba ECS. The equipment architect was designed with an edge computing innovation. The entire strategy would structure a reasonable classifier to discover the limit between regular and transformation codes. It could be applied in the identification of the change code of the system. The task utilized a dataset vector to partition them into positive and negative type, and the results demonstrate that the RBF-work SVM strategy performs most effectively in this mission. This research has suitable system security recognition in IoT frameworks and expanded the uses of ML. DDoS attacks are considered as sophisticated attacks and detection of such attacks have become a challenging task. The authors in [81] studied the a cloud environment using Tor Hammer as an attacking toop and then they created a dataset in order to detect the intrusions. They have utilized different ML algorithms including SVM, Navie Bayes, Random forest for classification and they demonstrated that SVM had the highest accuracy e.g., 99.7%.

4.1.5. Discussion and Lessons Learned

In this subsection we summarize the methodologies and advantages and the disadvantages of the above previous studies that discussed the use of supervised algorithms for cloud security as shown in Table 3.

Hussin et al. [59] used an ANN algorithm to resolve security issues. The ANN algorithm was used to resolve security threats in a banking organization. ANNs are used for improving performance and learning, neural-functions [59]. Sayantan et al. in [60] described a proficient methodology for the discovery of digital assaults in cloud foundations similar to the remote processing devices. This proposed strategy was associated with using an ANN [60]. In Reference [59], the advantage is to improve the data analysis, and the disadvantage is how to influence the performance of the network. In Reference [60], the advantage is to provide parallel processing capability and the disadvantage is an increase in computational costs.

Al et al. [62] explained secure edge computing in the IoT to secure data and improve performance. The advantages are to provide secure and reliable data, while the disadvantages are security and storage issues. Edge Computing of an IoT structure is a data arrangement that is performed at or approaches the finders of data in an IoT system. [65] reviewed explicit parts of the edge computing design and its relationship to industrial applications as a feature of an exacting revision, performed to confirm supporting the use of edge arrangements in testing conditions that emerge in the industry, including smart factories [82]. The advantages of their study are to provide secure and accurate data using ML and the disadvantages are security and accuracy issues.

Zardari et al. [70] discussed the K-NN method of information arrangement in the cloud administrations and virtual conditions. The target of utilizing K-NN incorporates the grouping of information based on their security prerequisites [70]. The authors in [83] discussed the possibilities of using ML in cybersecurity by demonstrating both the benefits and the risks. Cybersecurity controls are not adequate to protect networks from highly skilled cybercriminals. IDPSs are not adequate for managing the threats; fortunately, AI and ML provide significant support to IDPSs and increase the rate of detection and prevention. In Reference [70], the advantage of K-NN is simple, and intuitive while the disadvantages are difficulties in finding the optimal k value. In Reference [83], the advantage is a superior classification over large data sets using IDPSs and the disadvantage is time-consuming with high memory utilization.

Zekri et al. [74] designed a DDoS detection system based on the C.4.5 algorithm to mitigate a DDoS threat. CC empowers consumers and organizations to utilize applications without establishment and access their records on any PC with Internet access. Based on the investigation results, C4.5 is the optimal technique for grouping. Based on the results of the detection of DDoS through C4.5, the detection rate is more than 98%, and the DDoS attack duration is higher than detection rate using this algorithm. Different ML algorithms are used to detect the threat in DDoS. The C4.5 attempts to find the smallest decision tree [74]. Hussein et al. in [78] discussed web pre-fetching schemes using ML for MCC. Pre-fetching is one of the innovations utilized in reducing the latency of organizing traffic on the Internet. Their paper proposed this innovation to use MCC conditions to deal with inactivity issues in the setting of information management. The advantage of [78] is efficient data handling, and the disadvantages are time and storage issues.

Hanna et al. [76] proposed a decision tree (C4.5) algorithm used for classification and securing data. The study discussed and break down of how to achieve moderation for distributed computing security risks as an essential advance toward acquiring a safe condition for distributed computing. The outcomes indicated that using a basic decision tree model Chaid algorithm security rating for the ordering approach is a powerful strategy that empowers the leader to quantify the degree of cloud and the offered types of assistance [76]. In Reference [74], the advantage of C4.5 is that accepts both continuous and discrete values; the disadvantage is a small variety of data may produce different decision trees. In Reference [76], the advantage is efficient data handling, and the disadvantages are time and space issues.

El-Baghdadi et al. [64] used one of the ML emergent algorithms, DRLCS, to solve the problem of resource scheduling in CC. Ahmad et al. [15] discussed the lack of interest of the data owner that still exists due to data insecurity by a third party that stores, manage, and processes the data. In Reference [64], the advantage is to improve the data analysis and the disadvantage is the challenge of selecting the appropriate resource scheduling algorithm for a specific workload. In Reference [15], the advantage is cloud workload protection and the disadvantage is to dedicate and specialized client-server applications for proper functionality. Zamzam et al. [69] discussed resource management using ML in MEC to solve the issue and improve performance. In their reseach, the advantages are improved performance and secure resource management, and the disadvantages are network and optimization error.

Grusho et al. [80] discussed AI methods and models to solve information security problems. The most significant security threats to CC environments are the abusive and malicious use of cloud services computing, architectural limitations to cloud infrastructure access, and the dynamic nature of CC environments. IDPs are used to identify problems with security policies, document existing threats, and maintain the information exchange participants from violating security policies. The typology of IDPS systems depends on the specific events that they should track, and the means ("channels") to the implement these events. Hanna et al. [76] described that a decision tree (C4.5) Algorithm used for classification and secure data. They discussed how to achieve moderation for distributed computing security risks as an essential advance toward acquiring a safe condition for distributed computing. In Reference [80], the advantage is to provide parallel processing capability and the disadvantage is increase computational cost. In Reference [76], the advantage is efficient data handling and the disadvantages are time and space issues. Hou et al. [54] proposed a security detecting network of edge computing systems using ML to solve the security problem. The advantages are providing secure data and improving security issues, and the disadvantages are network error and storage issues.

Table 3. Comparison of supervised learning techniques for cloud security.

Reference	Objective	Technique	Advantages	Disadvantages
[21]	Public Cloud and private Cloud authorities	ANN	Ensure high data privacy; Cloud workload protection	Dedicated and specialized client-server applications for proper functionality
[64]	Supervised and unsupervised for secure cryptosystems	SVM	Secure Data; Improve Security Issues	Storage Issues; Network Error; Security Issues
[66]	Attack detection MCC	ANNs	High accuracy	Time and Storage
[67]	Attack and intrusion detection	ANNs	Tested on different dataset	Accuracy was not reported.
[70]	Reliable resource provisioning in joint edge Cloud environments	K-NN and Data Mining Techniques	K-NN is very simple and intuitive; Better classification over large data sets	Difficulties in finding optimal *k* value; Time Consuming; High memory utilization
[73]	Privacy Preserving	K-NN	Time efficiency	Accuracy was not reported.
[74]	ML for Cloud Security & C4.5 Algorithm for better protection in the Cloud	C4.5 Algorithm and signature detection Techniques	C4.5 algorithm deals with noise; C4.5 accepted both continues and discrete values	The small variation of data may produce different decision trees; Over-fitting
[78]	Web pre-fetching scheme in MCC	Naive Bayes	Efficient data handling	Time and Storage issues
[79]	Intrusion detection	Navie Bayes	Compatability	Accuracy was not reported.
[80]	Security and privacy issues identification & clarifies the information transfer using ML	ANN	Cloud workload protection and transfer data easily	Dedicate and specialized client-server application for proper functionality; Security issues
[81]	Intrusion detection	SVM and Navie Bayes	High Accuracy	Limited test environments.
[68]	Pros and cons of different authentication strategies for Cloud authentication	ANN & Cloud Delphi techniques	Improved data analysis; ANN gets lower detection precision	Unexplained behavior of ANN; Influence the performance of the network
[84]	Attacks launched on different level of Cloud	ANN & NN Techniques	Provide parallel processing capability	Computational cost increases

4.2. Unsupervised Learning

Unsupervised learning is a type of ML algorithm used to draw inductions from datasets, including data without naming responses. The most broadly perceived unsupervised learning strategy is the grouping assessment, which is used for data examination to find covered models or gathering in the data. The unsupervised learning algorithms receive several features from the information [85]. At the time that new data are presented, the algorithm uses the recently learned features to observes the class of the data. It is generally used for clustering and feature decline.

4.2.1. Unsupervised ANNs

Jiafu et al. [82] presented the four-layer cloud-assisted smart factory (CaSF) design to build up the after-effects of flow inquires the manufacturing field. ML techniques using four-layer techniques to solve this problem. The combination of ML methods and techniques demonstrate greater trust, flexibility, and efficiency for a manufacturing organization, but still present several issues and technical challenges in this domain. Current engineering designs consist of an intelligent device, system, Cloud, and application layer. This technique is helpful for solving CaSF issues. The four-layer CaSF architecture technique is used to improve performance.

The fast improvements of the IoT and smart cell phones recently have drastically incentivized the advancement of edge registering. Edge computing has been extraordinarily helpful for lightweight devices to achieve complicated tasks effectively; however, its rushed improvement prompts the disregard of security dangers to an enormous degree in edge computing stages and their empowered applications. Xiao et al. [68] discussed the security issues and attacks and resolved problems using ML techniques. Their paper extensively reviews the most compelling and fundamental attacks the compared guard systems that have edge computing explicit attributes can be applied to real-world edge computing frameworks. The authors specifically foucsed on the accompanying four types of attacks: distribute denial of service attacks, side-channel attacks, malware injection attacks, and authentication and authorization attacks. [68] analyzed the underlying causes of these attacks, presented the typical operation and challenges in edge processing security, and proposed future examination directions.

The spread of handheld devices has prompted the remarkable development of traffic volume crossing both nearby systems and the Internet, designating mobile traffic order as a key instrument for social occasion high-valuable, important profiling data, other than traffic building and administering the board. Mobile traffic is a challenge for state-of-art ML approaches, because the rapidly advancing and extending set of applications creating traffic ruins ML-based approaches requiring a domain master structure. Deep learning (DL) may address this issue but results in higher completion times; we thus recommended the use of big data. In Reference [56] DL was used to classify encrypted mobile traffic. The authors investigated for the first time BD-empowered arrangement of scrambled mobile traffic-utilizing DL from an overall outlook, characterizing general plan rules, using an open cloud stage, based on a practical exploratory setup. The researchers found that, while big data represents a straightforward quickening agent for certain tasks, this is not the situation for the preparation period of DL models for traffic characterization, requiring a particular BD-informed structure. The trial arrangement is based upon a three-dimensional examination in the Big Data selection relevant to non-trivial trade-off, completion time, sending costs, and classification performance.

4.2.2. K-Means

K-means clustering is one of the most straightforward and well-known unsupervised ML algorithms. It accomplishes this goal by searching for a fixed number, *k*, of bunches in a dataset. A group refers to an assortment of accumulated information based in specific similarities. The authors in [86] studied the issue of attack disruption in order to overcome the security threats. The authors explained the use of K-means and some other Ml algorithm for improving CC security.

Ahmad et al. [15] performed a survey of ML techniques for a secure Cloud. The lack of interest of data owners still exists due to data insecurity by third-parties that store, manage, and processes the data. The opportunity for miuse remains. SMC has been used to securely process the data. Data owners in the cloud require data privacy and efficient data management over the Cloud. Pattern matching is one of the basic tools in different fields. Both supervised and unsupervised learning methods have been applied over the encrypted information to explore learning execution. Similar information supervised learning performs well in contrast with unsupervised learning outperforms unsupervised learning. The authors in [87], proposed a new intrusion detection technique by utilizing K-means. The main objective was to find the characteristics and security needs of CC. They were able to find out known and anomaly attacks in CC. The authors also claimed that the proposed method shown that it is able to decrease the flase positive and false negative rate as well as accelerating the speed of intrusion detection.

4.2.3. Singular Value Decomposition (SVD)

SVD provides another approach to factorize a grid into singular vectors and qualities. SVD is used broadly both in the count of other matrix activities, such as framework conversion, and as an information reduction strategy in ML.

Kumar et al. [84] described the framework, challenges, and open questions surrounding the successful operation of ML-based security detection in a cloud environment. Regular irregularity recognition does not create acceptable outcomes for investigators who examine security episodes in the Cloud. Model assessment introduces issues caused by the absence of benchmark datasets. While sending these discoveries, manage model consistency, confinement, and information storehouse issues, among numerous others. In their research, the authors represented the issue of "assault interruption" a posssible approach in the security information science space. They described the structure, difficulties, open inquiries, and strategy encompassing the fruitful of ML-based security locations in a cloud domain. A half-and-half approach of rules and ML yielded superior results and demonstrated how they can be consolidated as channels, or even as one single ML unit. Because there is no benchmark for assessing cloud interruption discovery frameworks, systems for consolidating high-quality assessment information with other security items or red teams (recommended) and developing the information collection with SMOTE or GANs are utilized. The involvement with the model clarifies capacity and exhibits how it is a higher priority than at any other time.

Khilar et al. in [16] considered the trust-based access control approach for the user. The authors sorted clients and cloud assets into different classes. The distinctive application is also based on an alternate degree of trust. The proposed ML approach was equipped to quickly manage an immense number of movement logs inside almost no time which significantly improves the speed. The proposed model outperforms the related models. The authors in [88], designed a novel scheme named as SHoSVD in the two-cloud model. They claimed that their proposed method could make sure that the outsourced data privacy of the users. Moreover, they claimed that the proposed method can support off-line users. Finally, they proved the superiority of the proposed method in terms of accuracy. Feng et al. in [89] proposed a novel orthogonal tensor SVD method utilizing data science methods for dimensionality reduction in big data, which is applicable for cyber security and cyber forensics. They presented a high-order lanczos-based orthogonal tensor SVD algorithm that was used for dimensionality reduction. Moreover, they developed a secure orthogonal tensor SVD method to outsource the computation task of the orthogonal SVD algorithm to cloud.

4.2.4. Discussion and Lessons Learned

In this subsection we analysis the above-mentioned papers that considered unsupervised learning algorithm to overcome the cloud security as summarized in Table 4.

Table 4. Comparison of unsupervised learning techniques for cloud security.

Reference	Objective	Technique	Advantages	Disadvantages
[21]	ML capability for secure cryptosystems K-Means	ANN Techniques	Ensure high data privacy; Cloud workload protection	Dedicated and specialized client-server applications or proper functionality
[24]	A trust evaluation strategy based on the ML approach predicting the trust values of user and resources	SVD Techniques	A trust-based access control model is an efficient method for security in CC; Privacy protection	Influence the performance of the network; Security Issues
[56]	The encrypted mobile traffic using deep learning	CNN & Deep learning	Secure data; Fast data transfer	Runtime error
[86]	Challenges and successful operationalization of ML based security detections	K-Means & Intrusion Detection Techniques	Ensure high data privacy consistency, restriction, and information	Difficulties to manage information
[87]	Intrusion detection	K-mean	High accuracy and consistency	Comparability
[88]	User privacy	SVD	High Accuracy	Tested on a single model
[89]	Dimensionality reduction	SVD	High accuracy	Comparability

Jiafu et al. [82] presented the four-layer CaSF design to build up the after-effects of flow inquires the manufacturing field. The combination of ML strategies and methods demonstrate greater trust, adaptability, and effectiveness in an assembling association. However, several issues and specialized difficulties persist in this area. Xiao et al. [68] discussed the security issues and attacks and resolved problems using ML techniques. Edge registering has enabled lightweight devices to accomplish unprecedented assignments adequately; however, its hurried improvement leads to the negligence of security threats in edge processing stages and associated applications. This paper is based on the evaluation of ML with neural networks to resolve security issues. In Reference [82], the advantage is to ensure high data privacy, and the disadvantages are issues and technical challenges in this domain. In Reference [68], the advantage is cloud-based security protection, and the disadvantage are security issues and challenges that need further investigations. Ahmad et al. [15] used unsupervised learning to infer functions from unlabeled training data. Clustering is regarded as unsupervised learning. In these studies, the advantage is cloud workload protection, and the disadvantage is the necessity of dedicated and specialized client-server applications for proper functionality.

Kumar et al. [84] described the framework, challenges, and open questions surrounding the successful operation of ML-based security detections in a cloud environment. Regular recognition does not create acceptable outcomes for investigators who examine security episodes in the Cloud. Khilar et al. [16] described CC that hosts devoted registering resources that got to whenever from anyplace. This creates flexibility in information retrieval, information inescapability, and versatility, and is used in ML methods. The primary motivation behind the model was to offer access to an approved client in the cloud and choose the confided in the asset for his calculation [16]. In Reference [84], the advantages are ensuring high data privacy consistency, restriction, and information and disadvantage are difficulties to manage information. In Reference [16], the advantage is that a trust-based access control model is an efficient method for security in CC, and the disadvantage is security issues.

5. Future Research Directions

In the future, cloud security will use different ML models [90]. The followings are several dataset and directions that required further studies:

- An appropriate investigation of overhead should be performed before including new progressions, for example, virtualization could be used to produce the preferred position concerning essential capabilities.
- ML datasets: a collection of AI datasets across numerous fields, for which there exist security-applicable datasets associated with themes, such as spam, phishing, and so on [91].
- HTTP dataset CSIC: The HTTP dataset CSIC contains a substantial number of automatically-produced Web demands and could be used for the testing of Web assault protection frameworks.
- Expose deep neural system: This is an open-source deep neural system venture that endeavors to distinguish malicious URLs, document ways, and registry keys with legitimate preparation. Datasets can be found in the information or model's registry in the sample scores.json documents.
- Although the exploration of ML with crowdsourcing has advanced significantly in the recent years, there are still some basic issues that remain to be studied [92].
- Potential directions exist to of positioning innovation by coordinating heterogeneous LBS frameworks and consistently indoor and outdoor situations [93]. There remain numerous challenges that can be explored in the future.

6. Conclusions

In this study, security threats and attacks as the most challenging issues in CC were analyzed. Different types of ML algorithms e.g., ANNs, K-NN, Naïve Bayes, SVM, K-Means, and SVD were investigated as solutions to address the security issues in CC. We reviewed several proposed techniques that used ML algorithms for cloud security. We presented an analytical review and analysis of the proposed techniques and highlighted their advantages and disadvantages. We also introduced several research directions that need more investigations in future.

Author Contributions: All the authors had equal contribution. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program(IITP-2020-2015-0-00742) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
CaSF	Cloud-Assisted Smart Factory
CC	Cloud Computing
CCE	Contact Center Enterprise
CDN	Content Delivery Network
CIA	Confidentiality , Integrity , Availability
CNN	Convolutional Neural Network
DDoS	Distributed Denial of Service
DeepRM	Deep Reinforcement Learning
DRLCS	Deep Reinforcement Learning for Cloud Scheduling
ECS	Elastic Compute Service
GA	Genetic Algorithm
GAN	Generative Adversarial Network
IaaS	Infrastructure as a Service
IDPS	Intrusion Detection and Prevention Service
IDS	Intrusion Detection System
IoT	Internet of Things
K-NN	K-Nearest Neighbors
LAMB	Levenberg-Marquardt Back Propagation
MCC	Mobile Cloud Computing
MEC	Mobile Edge Computing
ML	Machine Learning
PaaS	Platform as a Service
PART	Partial Tree
RBF	Radial Basis Function
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SaaS	Software as a Service
SMOTE	Synthetic Minority Oversampling Technique
SMP	Secure Multi-party Computation
SVD	Singular Value Decomposition
SVM	Support Vector Machine
UNSW	University of New South Wales
VM	Virtual Machine

References

1. Lim, S.Y.; Kiah, M.M.; Ang, T.F. Security Issues and Future Challenges of Cloud Service Authentication. *Polytech. Hung.* **2017**, *14*, 69–89.
2. Borylo, P.; Tornatore, M.; Jaglarz, P.; Shahriar, N.; Cholda, P.; Boutaba, R. Latency and energy-aware provisioning of network slices in cloud networks. *Comput. Commun.* **2020**, *157*, 1–19. [CrossRef]
3. Carmo, M.; Dantas Silva, F.S.; Neto, A.V.; Corujo, D.; Aguiar, R. Network-Cloud Slicing Definitions for Wi-Fi Sharing Systems to Enhance 5G Ultra-Dense Network Capabilities. *Wirel. Commun. Mob. Comput.* **2019**, *2019*, 1–17. [CrossRef]
4. Dang, L.M.; Piran, M.; Han, D.; Min, K.; Moon, H. A Survey on Internet of Things and Cloud Computing for healthcare. *Electronics* **2019**, *8*, 768. [CrossRef]
5. Srinivasamurthy, S.; Liu, D. Survey on Cloud Computing Security. 2020. Available online: https://www.semanticscholar.org/ (accessed on 19 July 2020).
6. Mathkunti, N. Cloud Computing: Security Issues. *Int. J. Comput. Commun. Eng.* **2014**, *3*, 259–263
7. Stefan, H.; Liakat, M. Cloud Computing Security Threats And Solutions. *J. Cloud Comput.* **2015**, *4*, 1. [CrossRef]

8. Fauzi, C.; Azila, A.; Noraziah, A.; Tutut, H.; Noriyani, Z. On Cloud Computing Security Issues. *Intell. Inf. Database Syst. Lect. Notes Comput. Sci.* **2012**, *7197*, 560–569.
9. Palumbo, F.; Aceto, G.; Botta, A.; Ciuonzo, D.; Persico, V.; Pescapé, A. Characterizing Cloud-to-user Latency as perceived by AWS and Azure Users spread over the Globe. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Taipei, Taiwan, 7–11 December 2019; pp. 1–6.
10. Hussein, N.H.; Khalid, A. A survey of Cloud Computing Security challenges and solutions. *Int. J. Comput. Sci. Inf. Secur.* **2017**, *1*, 52–56.
11. Le Duc, T.; Leiva, R.G.; Casari, P.; Östberg, P.O. Machine Learning Methods for Reliable Resource Provisioning in Edge-Cloud Computing: A Survey. *ACM Comput. Surv.* **2019**, *52*, 1–39. [CrossRef]
12. Li, K.; Gibson, C.; Ho, D.; Zhou, Q.; Kim, J.; Buhisi, O.; Gerber, M. Assessment of machine learning algorithms in cloud computing frameworks. In Proceedings of the IEEE Systems and Information Engineering Design Symposium, Charlottesville, VA, USA, 26 April 2013; pp. 98–103.
13. Callara, M.; Wira, P. User Behavior Analysis with Machine Learning Techniques in Cloud Computing Architectures. In Proceedings of the 2018 International Conference on Applied Smart Systems, Médéa, Algeria, 24–25 November 2018; pp. 1–6.
14. Singh, S.; Jeong, Y.-S.; Park, J. A Survey on Cloud Computing Security: Issues, Threats, and Solutions. *J. Netw. Comput. Appl.* **2016**, *75*, 200–222.
15. Khan, A.N.; Fan, M.Y.; Malik, A.; Memon, R.A. Learning from Privacy Preserved Encrypted Data on Cloud Through Supervised and Unsupervised Machine Learning. In Proceedings of the International Conference on Computing, Mathematics and Engineering Technologies, Sindh, Pakistan, 29–30 January 2019; pp. 1–5.
16. Khilar, P.; Vijay, C.; Rakesh, S. Trust-Based Access Control in Cloud Computing Using Machine Learning. In *Cloud Computing for Geospatial Big Data Analytics*; Das, H., Barik, R., Dubey, H., Roy, D., Eds.; Springer: Cham, Switzerland, 2019; Volume 49, pp. 55–79.
17. Subashini, S.; Kavitha, V. A Survey on Security Issues in Service Delivery Models of Cloud Computing. *J. Netw. Comput. Appl.* **2011**, *35*, 1–11. [CrossRef]
18. Bhamare, D.; Salman, T.; Samaka, M.; Erbad, A.; Jain, R. Feasibility of Supervised Machine Learning for Cloud Security. In Proceedings of the International Conference on Information Science and Security, Jaipur, India, 16–20 December 2016; pp. 1–5.
19. Li, C.; Song, M.; Zhang, M.; Luo, Y. Effective replica management for improving reliability and availability in edge-cloud computing environment. *J. Parallel Distrib. Comput.* **2020**, *143*, 107–128. [CrossRef]
20. Purniemaa, P.; Kannan, R.; Jaisankar, N. Security Threat and Attack in Cloud Infrastructure: A Survey. *Int. J. Comput. Sci. Appl.* **2013**, *2*, 1–12.
21. Yuhong, L.; Yan, S.; Jungwoo, R.; Syed, R.; Athanasios, V. A Survey of Security and Privacy Challenges in Cloud Computing: Solutions and Future Directions. *J. Comput. Sci. Eng.* **2015**, *9*, 119–133.
22. Chirag, M.; Dhiren, P.; Bhavesh, B.; Avi, P.; Muttukrishnan, R. A survey on security issues and solutions at different layers of Cloud computing. *J. Supercomput.* **2013**, *63*, 561–592.
23. Behl, A.; Behl, K. An analysis of cloud computing security issues. In Proceeding of the World Congress on Information and Communication Technologies, Trivandrum, India, 30 October–2 November 2012; pp. 109–114.
24. Selamat, N.; Ali, F. Comparison of malware detection techniques using machine learning algorithm. *Indones. J. Electr. Eng. Comput. Sci.* **2019**, *16*, 435. [CrossRef]
25. Shamshirband, S.; Fathi, M.; Chronopoulos, A.T.; Montieri, A.; Palumbo, F.; Pescapè, A. Computational Intelligence Intrusion Detection Techniques in Mobile Cloud Computing Environments: Review, Taxonomy, and Open Research Issues. *J. Inf. Secur. Appl.* **2019**, 1–52.
26. Farhan, S.; Haider, S. Security threats in cloud computing. In Proceedings of the Internet Technology and Secured Transactions (ICITST), Abu Dhabi, UAE, 11–14 December 2011; pp. 214–219.
27. Shaikh, F.B.; Haider, S. Security issues in cloud computing. In Proceedings of the International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 15–16 May 2015; pp. 691–694.
28. Hourani, H.; Abdallah, M. Cloud Computing: Legal and Security Issues. In Proceedings of the International Conference on Computer Science and Information Technology (CSIT), Helsinki, Finland, 13–14 June 2018; pp. 13–16.
29. Alam, M.S.B. Cloud Computing-Architecture, Platform and Security Issues: A Survey. *World Sci. News* **2017**, *86*, 253–264.

30. Shukla, S.; Maheshwari, H. Discerning the Threats in Cloud Computing Security. *J. Comput. Theor. Nanosci.* **2019**, *16*, 4255–4261. [CrossRef]
31. Alsolami, E. Security threats and legal issues related to Cloud based solutions. *Int. J. Comput. Sci. Netw. Secur.* **2018**, *18*, 156–163.
32. Badshah, A.; Ghani, A.; Shamshirband, S.; Aceto, G.; Pescapè, A. Performance-based service-level agreement in cloud computing to optimise penalties and revenue. *IET Commun.* **2020**, *14*, 1102–1112. [CrossRef]
33. Tsuruoka, Y. Cloud Computing—Current Status and Future Directions. *J. Inf. Process.* **2016**, *24*, 183–194. [CrossRef]
34. Nagaraju, K.; Sridaran, R. A Survey on Security Threats for Cloud Computing. *Int. J. Eng. Res. Technol.* **2012**, *1*, 1–10.
35. Mozumder, D.P.; Mahi, J.N.; Whaiduzzaman, M.; Mahi, M.J.N. Cloud Computing Security Breaches and Threats Analysis. *Int. J. Sci. Eng. Res.* **2017**, *8*, 1287–1297.
36. Gessert, F.; Wingerath, W.; Ritter, N. Latency in Cloud-Based Applications. In *Fast and Scalable Cloud Data Management*; Springer: Cham, Switzerland, 2020.
37. De Donno, M.; Giaretta, A.; Dragoni, N.; Bucchiarone, A.; Mazzara, M. Cyber-Storms Come from Clouds: Security of Cloud Computing in the IoT Era. *Future Internet* **2019**, *11*, 127. [CrossRef]
38. Xue, M.; Yuan, C.; Wu, H.; Zhang, Y.; Liu, W. Machine Learning Security: Threats, Countermeasures, and Evaluations. *IEEE Access* **2020**, *8*, 74720–74742. [CrossRef]
39. Deshpande, P.; Sharma, S.C.; Peddoju, S.K. Security threats in cloud computing. In Proceedings of the International Conference on Computing, Communication and Automation, Greater Noida, India, 11–14 December 2011; pp. 632–636.
40. Varun, K.A.; Rajkumar, N.; Kumar, N.K. Survey on security threats in cloud computing. *Int. J. Appl. Eng. Res.* **2014**, *9*, 10495–10500.
41. Kazim, M.; Zhu, S.Y. A survey on top security threats in cloud computing. *Int. J. Adv. Comput. Sci. Appl.* **2015**, *6*. [CrossRef]
42. Barona, R.; Anita, M. A survey on data breach challenges in cloud computing security: Issues and threats. In Proceedings of the International Conference on Circuit, Power and Computing Technologies (ICCPCT), Paris, France, 17–18 September 2017; pp. 1–8.
43. Aawadallah, N. Security Threats of Cloud Computing. *Int. J. Recent Innov. Trends Comput. Commun.* **2015**, *3*, 2393–2397. [CrossRef]
44. Nadeem, M. Cloud Computing: Security Issues and Challenges. *J. Wirel. Commun*. **2016**, *1*, 10–15. [CrossRef]
45. Nicho, M.; Hendy, M. Dimensions Of Security Threats in Cloud Computing: A Case Study. *Rev. Bus. Inf. Syst.* **2013**, *17*, 159. [CrossRef]
46. Khan, M. A survey of security issues for cloud computing. *J. Netw. Comput. Appl.* **2016**, *71*, 11–29. [CrossRef]
47. Lin, C.; Lu, H. Response to Co-resident Threats in Cloud Computing Using Machine Learning. In Proceedings of the International Conference on Advanced Information Networking and Applications, Caserta, Italy, 15–17 April 2020; Volume 926, pp. 904–913.
48. Venkatraman, S.; Mamoun, A. Use of data visualisation for zero-day malware detection. *Secur. Commun. Netw.* **2018**, 1–13. [CrossRef]
49. Venkatraman, S.; Mamoun, A.; Vinayakumar, R. A hybrid deep learning image-based analysis for effective malware detection. *J. Inf. Secur. Appl.* **2019**, *47*, 377–389. [CrossRef]
50. Lee, K. Security threats in cloud computing environments. *Int. J. Secur. Its Appl.* **2012**, *6*, 25–32.
51. Liu, Q.; Li, P.; Zhao, W.; Cai, W.; Yu, S.; Leung, V.C. A Survey on Security Threats and Defensive Techniques of Machine Learning: A Data Driven View. *IEEE Access* **2018**, *6*, 12103–12117. [CrossRef]
52. Sarma, M.; Srinivas, Y.; Ramesh, N.; Abhiram, M. Improving the Performance of Secure Cloud Infrastructure with Machine Learning Techniques. In Proceedings of the International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, India, 19–21 October 2016; pp. 78–83.
53. Malomo, O.; Rawat, D.B.; Garuba, M. A Survey on Recent Advances in Cloud Computing Security. *J. Next Gener. Inf. Technol.* **2018**, *9*, 32–48.
54. Hou, S.; Xin, H. Use of machine learning in detecting network security of edge computing system. In Proceedings of the 4th International Conference on Big Data Analytics (ICBDA), Suzhou, China, 13–15 March 2019; pp. 252–256.

55. Zhao, Y.; Chen, J.; Wu, D.; Teng, J.; Yu, S. Multi-Task Network Anomaly Detection using Federated Learning. In Proceedings of the Tenth International Symposium on Information and Communication Technology, Jeju Island, Korea, 16–18 October 2019; pp. 273–279.
56. Aceto, G.; Ciuonzo, D.; Montieri, A.; Persico, V.; Pescapé, A. Know your big data trade-offs when classifying encrypted mobile traffic with deep learning. In Proceedings of the Network Traffic Measurement and Analysis Conference (TMA), Paris, France, 19–21 June 2019; pp. 121–128.
57. Shamshirband, S.; Rabczuk, T.; Chau, K.W. A survey of deep learning techniques: Application in wind and solar energy resources. *IEEE Access* **2019**, *7*, 64650–164666. [CrossRef]
58. Usama, M.; Qadir, J.; Raza, A.; Arif, H.; Yau, K.L.A.; Elkhatib, Y.; Al-Fuqaha, A. Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges. *IEEE Access* **2017**, *7*, 65579–65615. [CrossRef]
59. Elzamly, A.; Hussin, B.; Basari, A.S. Classification of Critical Cloud Computing Security Issues for Banking Organizations: A Cloud Delphi Study. *Int. J. Grid Distrib. Comput.* **2016**, *9*, 137–158. [CrossRef]
60. Sayantan, G.; Stephen, Y.; Arun-Balaji, B. Attack Detection in Cloud Infrastructures Using Artificial Neural Network with Genetic Feature Selection. In Proceedings of the IEEE 14th International Conference on Dependable, Autonomic and Secure Computing, Athens, Greece, 12–15 August 2016; pp. 414–419.
61. Lee, Y.; Yongjoon, P.; Kim, D. Security Threats Analysis and Considerations for Internet of Things. In Proceedings of the International Conference on Security Technology (SecTech), Jeju Island, Korea, 25–28 Novemebr 2015; pp. 28–30.
62. Al-Janabi, S.; Shehab, A. Edge Computing: Review and Future Directions. *REVISTA AUS J.* **2019**, *26*, 368–380.
63. Pham, Q.V.; Fang, F.; Ha, V.N.; Piran, M.J.; Le, M.; Le, L.B.; Hwang, W.J.; Ding, Z. A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art. *IEEE Access* **2020**, *8*, 116974–117017. [CrossRef]
64. El-Boghdadi, H.; Rabie, A. Resource Scheduling for Offline Cloud Computing Using Deep Reinforcement Learning. *Int. J. Comput. Sci. Netw.* **2019**, *19*, 342–356.
65. Nawrocki, P.; Śnieżyński, B.; Słojewski, H. Adaptable mobile cloud computing environment with code transfer based on machine learning. *Pervasive Mob. Comput.* **2019**, *57*, 49–63. [CrossRef]
66. Nguyen, N.; Hoang, D.; Niyato, D.; Wang, P.; Nguyen, D.; Dutkiewicz, E. Cyberattack detection in mobile cloud computing: A deep learning approach, In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 15–18 April 2018; pp. 1–6.
67. Saljoughi, A.; Mehrdad, M.; Hamid, M. Attacks and intrusion detection in cloud computing using neural networks and particle swarm optimization algorithms. *Emerg. Sci. J.* **2017**, *1*, 179–191. [CrossRef]
68. Xiao, Y.; Jia, Y.; Liu, C.; Cheng, X.; Yu, J.; Lv, W. Edge Computing Security: State of the Art and Challenges. *Proc. IEEE* **2019**, *107*, 1608–1631. [CrossRef]
69. Zamzam, M.; Tallal, E.; Mohamed, A. Resource Management using Machine Learning in Mobile Edge Computing: A Survey. In Proceedings of the Ninth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 8–10 December 2019; pp. 112–117.
70. Zardari, M.A.; Jung, L.T.; Zakaria, N. K-NN classifier for data confidentiality in cloud computing. In Proceedings of the International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 3–5 June 2014; pp. 1–6.
71. Calderon, R. The Benefits of Artificial Intelligence in Cybersecurity. Available online: https://digitalcommons.lasalle.edu/ecf-capstones/36 (accessed on 19 July 2020).
72. Shamshirband, S.; Chronopoulos, A.T. A new malware detection system using a high performance-ELM method. In Proceedings of the 23rd International Database Applications & Engineering Symposium, Athens, Greece, 10–12 June 2019; pp. 1–10.
73. Park, J.; Lee, D. Privacy preserving K-nearest neighbor for medical diagnosis in e-health cloud. *J. Healthc. Eng.* **2018**, 1–11. [CrossRef] [PubMed]
74. Zekri, M.; El Kafhali, S.; Aboutabit, N.; Saadi, Y. DDoS attack detection using machine learning techniques in cloud computing environments. In Proceedings of the International Conference of Cloud Computing Technologies and Applications (CloudTech), Rabat, Morocco, 24–26 October 2017; pp. 1–7.
75. Kour, H.; Gondhi, N.K. Machine Learning Techniques: A Survey. In *Innnovative Data Communication Technologies and Application Lecture Notes on Data Engineering and Communications Technologies*; Springer: Cham, Switzerland, 2020; pp. 266–275.

76. Hanna, M.S.; Bader, A.A.; Ibrahim, E.E.; Adel A.A. Application of Intelligent Data Mining Approach in Securing the Cloud Computing. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 151–159.
77. Mishra, A.; Gupta, N.; Gupta, B.B. Security Threats and Recent Countermeasures in Cloud Computing. In *Modern Principles, Practices, and Algorithms for Cloud Security Advances in Information Security, Privacy, and Ethics*; IGI Global: Hershey, PA, USA, 2020; pp. 145–161.
78. Hussien, N.; Sulaiman, S. Web pre-fetching schemes using Machine Learning for Mobile Cloud Computing. *Int. J. Adv. Soft Comput. Appl.* **2017**, *9*, 154–187.
79. Arjunan, K.; Modi, C. An enhanced intrusion detection framework for securing network layer of cloud computing. In Proceeding of the ISEA Asia Security and Privacy (ISEASP), Surat, India, 29 January–1 February 2017; pp. 1–10.
80. Grusho, A.; Zabezhailo, M.; Zatsarinnyi, A.; Piskovskii, V. On some artificial intelligence methods and technologies for cloud-computing protection. *Autom. Doc. Math. Linguist.* **2017**, *51*, 62–74. [CrossRef]
81. Wani, A.; Rana, Q.; Saxena, U.; Pandey, N. Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques. In Proceedings of the Amity International Conference on Artificial Intelligence (AICAI), Dubai, UAE, 4–6 Febuary 2019; pp. 870–875.
82. Wan, J.; Yang, J.; Wang, Z.; Hua, Q. Artificial Intelligence for Cloud-Assisted Smart Factory. *IEEE Access* **2018**, *6*, 55419–55430. [CrossRef]
83. Abdurachman, E.; Gaol, F.L.; Soewito, B. Survey on Threats and Risks in the Cloud Computing Environment. *Procedia Comput. Sci.* **2019**, *161*, 1325–1332
84. Kumar, R.; Wicker, A.; Swann, M. Practical Machine Learning for Cloud Intrusion Detection: Challenges and the Way Forward. In Proceedings of the ACM Workshop on Artificial Intelligence and Security, Dallas, TX, USA, 3 November 2017; pp. 81–90.
85. Quitian, O.I.T.; Lis-Gutiérrez, J.P.; Viloria, A. Supervised and Unsupervised Learning Applied to Crowdfunding. In *Computational Vision and Bio-Inspired Computing. ICCVBIC 2019*; Springer: Cham, Switzerland, 2020.
86. Meryem, A.; Samira, D.; Bouabid, E.O. Enhancing Cloud Security using advanced Map Reduce k-means on log files. In Proceedings of the International Conference on Software Engineering and Information Management, New York, NY, USA, 4–6 January 2018; pp. 63–67. [CrossRef]
87. Zhao, X.; Zhang, W. An Anomaly Intrusion Detection Method Based on Improved K-Means of Cloud Computing. In Proceedings of the Sixth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC), Harbin, China, 21–23 July 2016; pp. 284–288.
88. Chen, J.; Liu, L.; Chen, R.; Peng, W. SHOSVD: Secure Outsourcing of High-Order Singular Value Decomposition. In Proceeding of the Australasian Conference on Information Security and Privacy, Perth, Australia, 30 November–2 December 2020; pp. 309–329.
89. Feng, J.; Yang, L.; Dai, G.; Wang, W.; Zou, D. A Secure High-Order Lanczos-Based Orthogonal Tensor SVD for Big Data Reduction in Cloud Environment. *IEEE Trans. Big Data* **2019**, *5*, 355–367. [CrossRef]
90. Subramanian, E.; Tamilselvan, L. A focus on future cloud: Machine learning-based cloud security. *Serv. Oriented Comput. Appl.* **2019**, *13*, 237–249. [CrossRef]
91. Alazab, M.; Layton, R.; Broadhurst, R.; Bouhours, B. Malicious spam emails developments and authorship attribution. In Proceedings of the Fourth Cybercrime and Trustworthy Computing Workshop, Sydney, Australia, 21–22 November 2013; pp. 58–68.
92. Sheng, V.; Zhang, J. Machine Learning with Crowdsourcing: A Brief Summary of the Past Research and Future Directions. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 9837–9843.
93. Li, Z.; Xu, K.; Wang, H.; Zhao, Y.; Wang, X.; Shen, M. Machine-Learning-based Positioning: A Survey and Future Directions. *IEEE Netw.* **2019**, *33*, 96–101 [CrossRef]