Global Transitions Proceedings 3 (2022) 243–250

Contents lists available at ScienceDirect

Global Transitions Proceedings

journal homepage: http://www.keaipublishing.com/en/journals/global-transitions-proceedings/

Modelling of smart risk assessment approach for cloud computing environment using AI & supervised machine learning algorithms

Abhishek Sharma*, Umesh Kumar Singh

Institute of Computer Science,Vikram University, Ujjain,456001, INDIA

ARTICLE INFO

Keywords:
Artificial intelligence
Cloud computing (CC)
Cloud security
Machine learning
Risk assessment

ABSTRACT

Major backbone of today's competitive and upcoming market is definitely becoming Cloud computing & hence corporate utilize capabilities of cloud computing services. To improve security initiatives by cloud computing service or CRPs, novel types of tools and protocols finds themselves always in demand. In order to build comprehensive risk assessment methodology, extensive literature review was conducted to identify risk factors that may affect cloud computing adoption. In this context various risk factors were identified. After feature selection and identification of risk factors, utilized to select most effective features using linear regression algorithms. Then AI-ML techniques like Decision Tree (DTC), Randomizable Filter Classifier, k-star with RMSE method is used to analyse threats within CC environment. Experimental outcomes depicted that division of dataset to (95%-5%) provided best result out of every remaining partitioning and moreover put forth that DTC algorithm provided best outcomes out of entire data set used in experimental setups.

1. Introduction

Cloud [1] being one of most versatile and dynamic inventions has brought focus of technologists on the global stage. Although Cloud Computing comes with huge rewards like measured services, rapid elasticity, scalability, and major significant being low cost to companies, more over it is embedded with good proportion of security risks that is not comprised by any other enterprise that can be ignored. The security danger that emanate from the wide range of the vulnerabilities inherent in any type of Cloud computing system and in the absence of reliable security directives, there is an apparent reluctance on the part of organizations to adopt an otherwise a powerful environment called cloud computing [2].

Without bothering regarding physical and technical maintenance or management challenges of original resources, Cloud Computing makes users capable of controlling and reaching their resources online with help from internet at any time irrespective of their locations [3]. Moreover, its resources are scalable and dynamic. It is an autonomous computing platform that is entirely different than utility and grid computing. Cloud Computing best example can be seen as Google Apps; it helps in accessing services through browser and can be easily installed on extensive number of computers over Internet links [4]. Using internet, Resources are easily approachable from cloud environment from any place and at any time across globally. Comparatively to other computing models it is less as far as cost is concern. As service provider is accountable for accessibility and availability of various services, cost of maintenance entangled with it is negligible and clients remain free from management and maintenance issues of resources at provider's end. Because of these characteristics, Cloud Computing came to be called as simply IT on demand or utility computing. A main significant feature of Cloud Computing is its Scalability and is attained via virtualization of their servers [5,6]. It gives software, computation, storage services and data access that doesn't need any end-user knowledge of system configuration that delivers services or physical location.

Cloud Computing Technology works on three different SPI (Software Platform Infrastructure) models and four deployments (public, private, hybrid, and community) models [7]. As per the usage or requirement consumer can use the service(s) of the cloud and deploy the cloud. Presently, it has three types of service models also called SPI (Software Platform Infrastructure) models which are given below [8]:

Ø SaaS (Software as a Service) [9]:-

It is a Software distribution and deployment schemes where applications are delivered to clients in form of services. For accessing and utilization of service or an application which is embedded in cloud, costumers receive this facility with ease. Such application can execute on service provider's Web servers or client's computing machines. SaaS delivers services for patch management efficiently and also promotes collaboration.

Least extensibility and greatest amount of securing responsibility taken on by the cloud provider. Basically, it uses provider's applica-

* Corresponding author.
E-mail address: abhiujn9@gmail.com (A. Sharma).

https://doi.org/10.1016/j.gltp.2022.03.030

Available online 2 April 2022

tions over a network. “Salesforce.com” is an example, where necessary information for the interaction between the consumer and the service is hosted as part of the service in the cloud and deliverance of basic business services that comprises of word processing and email by Google Apps [10].

Ø PaaS (Service Platform) [11]:-

Client gains accessibility to platform by empowering them in organizing their own application and software and applications in cloud domain. IT lies in the middle at somewhere, with extensibility and security features that must be leveraged by the customer. Main responsibility of PaaS is deploying customer related application to a cloud.

Ø IaaS (Infrastructure as a Service) [12]:-

This model is utilized in delivering storage, lease processing, and various other important resources of computing for their client. These clients doesn’t manage or control basic cloud framework but has accessing capabilities on basic storage, operating system, computing skills as standardized services on given network and deployed applications. It requires greater extensibility and least amount of security responsibility taken on by cloud provider. Network capacity (Servers, Switches, Routers, and other Systems) Storage, Rent processing & other fundamental computing resources comes under this model.

Cloud Computing is quite rich in features but main futures of Cloud Computing are as follows [13]:

Ø Use of internet-based services to support business process: -

Platform of Cloud Computing is not constrained to particular networks but can be access via huge network of them all that is Internet [13].

Ø Utility-like based Renting IT-services: -

Cloud Computing comprises in renting of computing resources like network bandwidth, software and hardware, on demand basis or as-required, like utility computing which was a precursor to Cloud Computing; without worrying about installation and maintenance cost of Cloud Computing [13].

It is a combination of prevailing technologies and methods, embedded in a new infrastructure paradigm which provides improved elasticity, scalability, faster start up time, reduced management costs, just-in-time availability of resources and business agility.

2. Related work

Computing in cloud is a natural development for computation centres and information/data enabled with virtualization technologies, workload balancing and automated systems management [14]. With all these evolutions of technologies, there arise complexities that contribute towards reluctance to adopt cloud by the market. One the main issue is security threats for migrating the data and application in the cloud because of obvious reasons. Further, security threats also have various categories like data security threats, network security threats, server security threats, application security threats, authentication & authorization security threats, web security threats, virtualization security threats. Each security danger constitutes of their own issues or concerns. One of prominent problems in network security and data security is Cyber-attacks. Therefore, this whole literature review is mainly concerned with different security threats or issues in cloud environment.

In [15], the authors give a review of a study that looked into CC security problems, challenges, and solutions that included one or more machine learning techniques. This includes a look at several machine learning methods, such as supervised, unsupervised, semi-supervised, and reinforcement learning, that are used to address cloud security concerns. Then, based on their theoretical qualities, benefits, and downsides, it compares the performance of each approach. Furthermore, it identifies future research possibilities for CC models to ensure their security.

With an aim in promoting application of best practices to deliver security and protection guarantee within Cloud Computing domain, Cloud Security Alliances (CSA) is a not-for profit organization. CSA also provides knowledge on how to use Cloud Computing in assisting of safeguarding other forms of computing entirely. Hence It is being identified as top seven dangers to cloud platform through an article “Top Threats to Cloud Computing V1.0” that is as followed [16,17]:

Ø Nefarious and Abusive application of Cloud Computing
Ø Application Programming Interfaces Insecurely
Ø Malicious programs Inside machines
Ø Vulnerabilities associated with Shared Technology
Ø Loss/Leakage of Information
Ø Services and Account
Ø Hijacking of Traffic

Cloud computing is prone to manifold security threats varying form network level threats to application level threats [18–20]. The reason for conducting literature review is to understand what exactly cloud computing means, working of cloud computing, and what are the problems in the cloud computing. It also focused on how it can mitigate the risks and encourage the companies/consumers to adopt the cloud computing environment.

The dangers, weaknesses, and risk mitigation, as well as the norms and legislation, are highlighted in [21]. The three technologies were then compared against international standards (OWASP, NIST, ISO, and GDPR), with the findings revealing that the majority of AI and IoT standards and regulations are still under development, while cloud computing has a sufficient basis of supporting standards. To counter the DDoS threat, the author created a DDoS detection system based on the C.4.5 algorithm in [22]. This method, when combined with signature detection approaches, provides a decision tree that can identify signature assaults for DDoS flooding attacks automatically and effectively. It chose various machine learning approaches and compared the results to validate the system. The authors show in [23] that their suggested system not only recognizes a wide range of cyber-attacks, but also detects them with great accuracy (up to 97.11%). It also presents comparisons with current machine learning-based methodologies in order to illustrate the efficacy of its suggested solution. The processing and analysis of massive data acquired from 3S methods and sensors, as described by the author in [24], give intriguing views for developing an integrated technological system for excavation.

The application of machine learning algorithms for risk assessment is clearly a developing topic of study, according to [25], as seen by the increasing trend in annual publications. Machine learning approaches may improve traditional risk assessment by offering data-driven insights as more data is collected on various socio-technical systems. The author of [26] provides a mathematical technique for automating anomaly detection by merging principles for cognitive engine design, edge computing, Artificial Intelligence, and Machine Learning. By embedding Artificial Intelligence and Machine Learning at the edge of IoT networks, this engine creates a step change in delivering secure and functional real-time intelligence for predictive cyber risk analytics.

Since the early days of computer design, network-based intrusion detection systems (NIDS) that employ statistical metrics or computer thresholds have been linked to security research [27]. However, because they have a high rate of false negatives (failure to detect) and false positives, they are useless for contemporary cyber risk analytics of networked and extremely complex ICT systems (false alerts). In the case of IoT, distributed threat detection at the fog level has been shown to be more scalable than centralized cloud [28]. If the attack vectors are known, a form of attack using bidirectional long short-term memory (LSTM) units introduced to a recurrent neural network (RNN) can achieve 99.999% accuracy [29]. Similarly, when compared to other algorithms, a Siamese Network Classification Framework (SNCF) may re-

lieve risk prediction imbalance and provide more trustworthy findings [30].

Cloud network security brings out the taxonomy of various types of cloud attacks that have occurred in recent past and it also lists out the successfully implemented solutions to mitigate the risks [31,32]. The review of literature reveals many different types of threats and attacks on cloud networks includes Denial of Service (Distributed DoS, HDoS, XDoS Attack, Man-in-the-middle attack, IP spoofing Attack, Sniffer Attack, Replay Attack, Dictionary Attack, Injection Attack, Hypervisor, Wrapping Attack and so on [33]. Prominent ones are Do's with the categories of DDoS and Man-in-the-Middle attacks. Under cloud network security threats, it has chosen to analyze two categories of the attacks, which are more prominent on cloud networks. These two categories are DoS (DDoS, XDoS, HDoS) attacks, and Man-in-the-middle attack. DoS attacks are very strong attack which interrupts services for a long duration of time. Instead if an attacks gains access maliciously in communication link so as to control and monitor communication and tampers messages for malevolent intentions, then Man-in-the-middle attack takes place. With assistance from authentication and identification procedures by validating identity of users such kinds of damages can be avoided. Few of networked attacks are describe below:

i Man-in-the-Middle Attack
ii Denial of Service (DoS) Attack
iii Distributed DoS Attack (DDoS
iv Repudiation
v Privileges Elevation
vi Worms and Viruses Attack.
vii Spoofing attack.
viii Reused IP Addresses.
ix Cookie Poisoning.
x CAPTCHA Braking.
xi Google Hacking
xii Dictionary Attack.
xiii Malware Injection Attacks.
xiv Sniffer Attack.
xv Tampering.
xvi Eavesdropping/Information Disclosure.
xvii Replay Attack
xviii Wrapping Attack.

Apart from these problems, it has found that industries are reluctant to adopting the cloud computing and different authors have different views for the reluctance. So there is a need for investigation into reasons for reluctance to adopt cloud services by the companies and consumers.

In [34], Secure Cloud Computing's Control Framework, Harshit Srivastava et al have opined that next big technology is cloud computing with its utility for different size of organization but the security &privacy issues are causing a serious concern for adoption which requires attention. In this paper authors have used many survey results to conclude that security and privacy of physical, environmental, and virtualization security is vendor's responsibility. This paper indicates, that organizations can exercise control on three main layers such as Physical, Logical and Methodology Layer for tackling threats at data centre, network security and insider security respectively. Authors have proposed a governing body with automated control framework which aims to computes threat index for solving security challenges by creating association within CSPs based on existing attacks.

In [35], author explains protocols of this emerging and new technology of cloud computing which delivers services and shared resources at reduced price of software and hardware along with few related security challenges during use of cloud's services. Additionally, author focuses on characteristic of multiple occupancy while also conversing information security problems on cloud exposed by CSA. Cloud computing Security issue can be reduced by designing or changing robust and strong multitenancy's architecture was concluded by Authors.

With regards to [36], author postulates that, for securing resources and the users data, the most important goal should be to maintain CIA (Confidentiality, Integrity and Availability) in order to continue the cloud service for business under emerging challenges of security which threaten this technology. Authors have used many survey results such as in 2013 DoS attack was declared fifth threat amongst top of infamous nine in cloud computing which might lead to damage of hardware or data and consequent loss of money if service is hijacked by CSA (Cloud Security Alliance). This paper also describes few DoS & DDoS attack and its effects on cloud with present resolutions; however, this paper focuses mainly on two category of DoS attack i.e. H-DoS and X-DoS and suggests Cloud protector and Cloud Trace Back (CTB) to remove these types of attacks and Cloud Defender System (CSQD) to alleviate XML vulnerabilities on web.

In [37], author discussed cloud computing as a technical shift to provide the services remotely by the third party called service providers. Authors have looked at insider threat with two outlooks, primarily with a view of cloud service provider and later with cloud outsourcer point of view and have suggested countermeasures accordingly. Against cloud outsourcer Countermeasures for user's perspective is host based IDS/IPS, log auditing, and with provider side is multi-factor authentication, anomaly detection and separation of duties. Researchers are more intense on 3 various kinds of attacks in current category like changing constituent of users file without their knowledge, obtaining the private keys of users 'encrypted files, web template poisoning and their mitigation techniques. Authors conclude by stating that organizations should be aware of vulnerabilities exposed by the utilization of cloud services and remain mindful of the availability of cloud services to employees with the organizations.

In the paper [38], author states that cloud security is an evolving sub-domain of information security, network security and computer security. Astounding security considerations of information security professionals need to be considered when evaluating the risks of cloud computing. The fundamental issues are application & data security and cloud user and provider both are responsible for this. However, providers must ensure that their infrastructure is secure and client's applications and data are protected and user should adopt measures to use strong passwords and authentication measures, the authors discussed. Researchers emphasis on privacy issues and cloud security, cloud security controls, its algorithms such as AES, MD-5, RSA, with their disadvantages and concluded that innovative decryption and encryption must be deployed in improving security across network.

In the paper [39], Pallavi Marathe et al. argue that, because cloud computing is outsourced through third party so inherently there is an added security risk which makes harder to maintain security, availability, confidentiality of data and it is also a hindrance for adoption of cloud computing. Cloud securities issues are broadly categorizing in two classes like security related problems as experienced by service providers with cloud and security challenges as sensed by end users. In this case, customer ensures that provider has taken proper security measures to protect their data and provider must ensure that their infrastructure is secure and customers' data and applications are protected. Cloud Computing services and storage are widely provided by Google and Amazon and VMware provide the software to create a privately owned cloud. Along with benefits from the cloud computing there are inherent security risks. As more data transits/moves from centrally located server-storage to another location in the cloud, the chances of compromise of private data also increase. In this paper authors are mainly focused on threats to information from confidentiality, integrity, availability and suggest the use of some tools, which are available in the market to reduce the risks of these threats such as Viivo, SkyCrypt, CipherCloud, Bitglass, Skyhigh explaining about their working, advantages, disadvantages. They conclude that Skyhigh and Bitglass tool is best in encryption of information and discover the cloud usage. Writers also propose that if corporate are capable of maintaining control and coordination of encryption keys, then they can guar-

antee agreement with external regulatory requirements and internal policies.

In [40], author Smita Parte et al. have discussed on how cloud computing offers attractive technological and financial advantages and the facilities of building, managing deploying and designing their autonomous applications remotely with no additional requirement of software and hardware. They also underline that the security considerations remain the major crucial features in cloud computing because of confidential information in the cloud. Authors have focuses on privacy, trust & security challenges (deficiency of user control, unapproved memory practice, data explosion, dynamic provisioning, access, trans border data flow, multi-tenancy, audit, availability etc.) issues, taxonomy of security aspects, security concerns (access management, encryption, key management, and other danger management), already present resolutions like firewall, IDS/IPS, antiviruses etc. Hence they concluded by stating that stakes holder, vendors, enterprises, organization have to take serious note about security concern of cloud computing prior to embracing cloud systems.

In [41], authors discuss the phenomenon of growth of cloud computing along with its challenges and issues are also growing as rapidly. This paper mainly covers overview, architecture, threats and existing countermeasures of threats to cloud computing. Various security attacks and threats at many strata (physical-IaaS levels, application-SaaS, virtual-PaaS,) as well as their influence like insider, flooding, user to root, port scanning, virtualization, backdoor channel(DDoS), storage allocation, authorization & authentication and data modification are discussed by the authors.

The following risk factors are identified for evaluation as:

i Authentication and Access Control (A&AC)
ii Insufficient due Diligence (IDD)
iii Data Loss (DL)
iv Insecure Application Programming (IAP)
v Data Transfer (DT)
vi Business Continuity and Service Availability (BC& SA)
vii Shared Environment (ShE)
viii Regulatory Compliance (RC)
ix Data Breaches (DB)
x Data location and Investigative Support (DL&IS)
xi Third Part Management (TPM)
xii Data Segregation (DS)
xiii Recovery (R)
xiv Data Integrity (DI)
xv Virtualization Vulnerabilities (VV)
xvi Resource Exhaustion (RE)
xvii Service Level Agreement (SLA)
xviii Interoperability and Portability (I& P)

3. Method

This stage compares the estimated risk levels against a risk acceptance criterion, which is a threshold established by business executives.

The goal of this work is to provide a methodological instrument for risk assessment in the cloud computing environment that is both trustworthy and effective. The proposed smart risk assessment modelling approach is implemented using ML model in three phases as shown in Fig. 1. The following precise research objectives were set in order to achieve this goal:

1- To evaluate and identify the body of information about risk issues related with cloud computing.
2- To run a simulation of the dataset using previously established risk variables.
3- To use machine learning techniques to create a feasible risk assessment model for cloud computing environments. Authors in [41] listed some issues should be put in considerable when selecting the most suitable assessment technique:

- Resource availability for analysis.
- Complexity and size of process that are analysed.
- The phase in which danger evaluation will be considered in process lifecycle.
- Information availability.

The Following predictive model is used to construct model for evaluation of performance through AI techniques & Linear regression algorithms (Machine learning) as represented in Fig. 2:-

Step 1: A literature evaluation of cloud computing was conducted, and the related risk factors were discovered. There were certain difficulties that were discovered. The same risk factor was defined by many studies, but they gave it different names. Other studies define risk variables, although they might be integrated and categorised under a different term. As a result, 18 risk factors for these disorders have been discovered. The goal of the project was to identify the most critical risk variables that can affect cloud computing adoption, as well as to assess which elements had a significant impact on the organization's objectives, so that they could be included and added to the risk factors already identified. All 18 risk factors are used as input variables in order to create a dataset with just one output, which is the estimated risk. Each variable is then divided into four categories: low, medium, high, and extremely high. Next, it uses one of the data measurement methods known as interval scale to assign numeric values to each variable; each variable has a numerical range value.

Step 2: After preparing the dataset, it is necessary to minimise the data's dimensionality, which allows the data analysis algorithm to run more quickly and efficiently. This job was completed using feature selection methods in this study. This work is done with the aid of the WEKA / orange tool, which is a proposed feature selection technique implementation. Best first, random search and ranker were always the feature selection methods employed.

Step 3: Three methods were employed as basis algorithms for assessing the risk variables connected with the cloud computing environment throughout this research. Extremely Randomized Decision Trees, K*, and randomizable Filter Classifier are the methods under question. These algorithms are well-known in the field of data analysis and have shown to be effective in practice. These algorithms are used with the appropriate customisation from WEKA / orange tools. According to the standard top-down technique, the Extremely Randomized Decision Trees or extra tree constructs an ensemble of un-pruned decision or regression trees. The data is split completely or partially randomly by the extra tree. It differs from existing decision tree induction techniques in two ways: it separates nodes by picking cut-points completely at random and it grows trees using the whole learning sample. The entropy distance metric is used in K*, an instance-based learning method, to quantify the distance between two instances. The entropy distance metric has various advantages, including a uniform method to addressing symbolic, real-valued, and missing-value properties. Typically used for running an arbitrary classifier on data that has been passed via an arbitrary filter in the case of randomizable Filter Classification. The structure of the filter, like that of the classifier, is solely dependent on the training data, and test instances will be handled by the filter without any changes to their structure.

Step 4: An ensemble is a group of learning machines whose judgments are pooled to improve the overall system's performance. It mix two datasets after applying machine learning methods to them to construct ensemble model. In trials, the ensemble model is built using the vote technique. A vote algorithm is a type of prediction algorithm that combines many predictors. For regression, several combinations of probability estimates are possible. Each prediction gets one vote in the voting procedure, and the

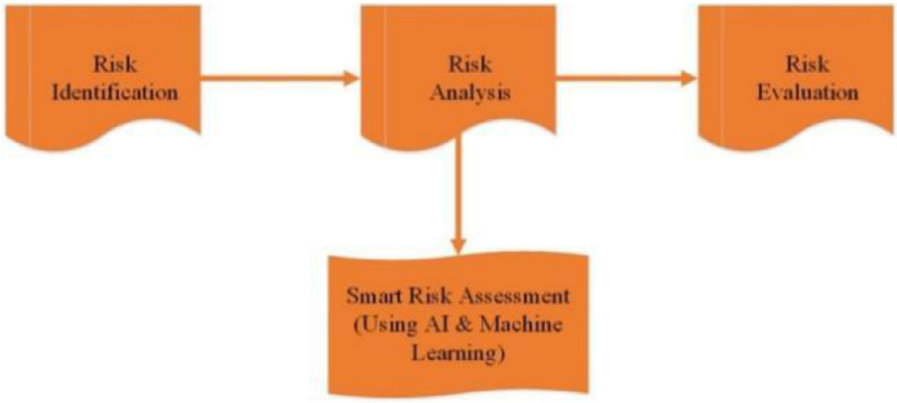

Figure 1. Smart risk assessment modelling approach.

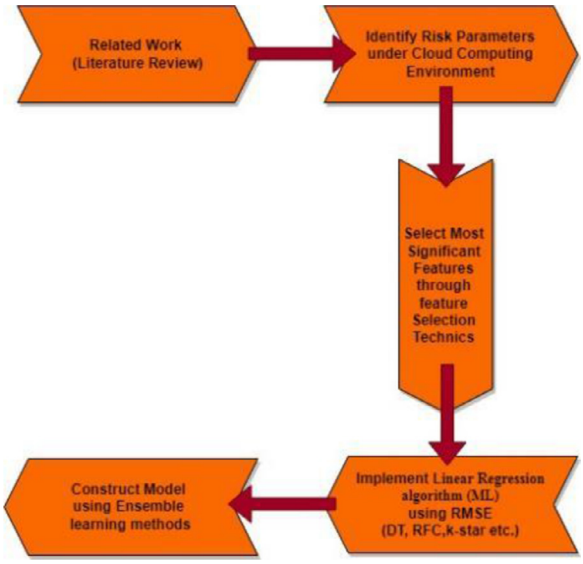

Figure 2. Methodology used to construct the proposed predictive Model.

majority wins. As a combination rule, the vote algorithm is implemented using the average of probability approach. Machine Learning Techniques used for Cloud Risk Assessment are as follows:

) Decision Trees

For supervised classifier, it is tree-based ensemble techniques. Decision-Trees is dependent on process of randomization, splitting rules are arbitrarily strained at every element of tree, and depends on selected one of such guidelines to be related with that node is best practice as per scoring computational evaluation. Such methods help in improving training speed, wearying correlation amongst induced decision-trees, and decreasing complication with induction procedures.

) Randomizable Filter Classifier

Utilised for executing a random classifier on information which has been transferred via an arbitrary filter. Similar to classifier, filter's structure is dependent exclusively on test instances and training information that will be executed by filter without altering their basic structure.

) k-Nearest Neighbours (k-NN or K*)

This algorithm, known as k-NN, is instance-dependent learning, in which the entire training sample is kept and no model is established until a novel instance is required to be grouped, and they use a few domain-specific distance functions to recover the single most identical instance from a set of training instances.

The integration of above machine learning methodologies is performed together with monotonous adaptive risk assessment system using AI technique. The outcome of this AI enabled proposed system with adaptive capabilities are continues evolving with the help of threat predictions and mitigation. Certain conventional performance indicators are employed in this research to assess the effectiveness of resultant approaches. It utilised two statistical metrics to solve this problem, first is Correlation Coefficient (R), and second is Root Mean Square Error (RSME) [46,47,48,49] as shown in Eqs. (1) and (2):

$$R = \sqrt{1 - \frac{\sum_1^n (Pi - Ai)^2}{\sum_1^n Ai^2}} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (Pi - Ai)^2} \tag{2}$$

Actual (desired) and fitted (predicted) output values are represented by Pi and Ai, respectively. As a consequence, it expects a value of one or close to one from the Correlation Coefficient (CC) metrics which is evaluated using Eq. (1), and low values from the Root Mean Square Error (RMSE) metrics as a final result which is evaluated using Eq. (2).

4. Result & discussion

For finalizing risk factors carried out a survey. In this, it questions volunteers to classify risk factors in three distinct strata as per possibilities of occurrence and their influence on CC. The following are the results of these classes: The terms "not important," "important," and "neutral" are used interchangeably. From various nations 35 international experts answered in this survey, and every one of them approved that formerly prescribed factors are significant, signifying that they have a lot of leverage in the cloud computing world. Later, it gave numeric range of values to each risk factor, and lastly, it established expert protocols and rules hence using some statistical techniques to produce information depending on those rules. 18 input attributes were contained by dataset and consists of around 1940 instances. Labelling of 18 attributes with their respective ranges for numeric values, Risk factors are provided in Table 1. It applied through splitting percentage to evaluate and test algorithm. Through splitting percentage, dataset is arbitrarily dividing testing and training information that follows:

- 50–50% (A)
- 65–35% (B)
- 85–15% (C)

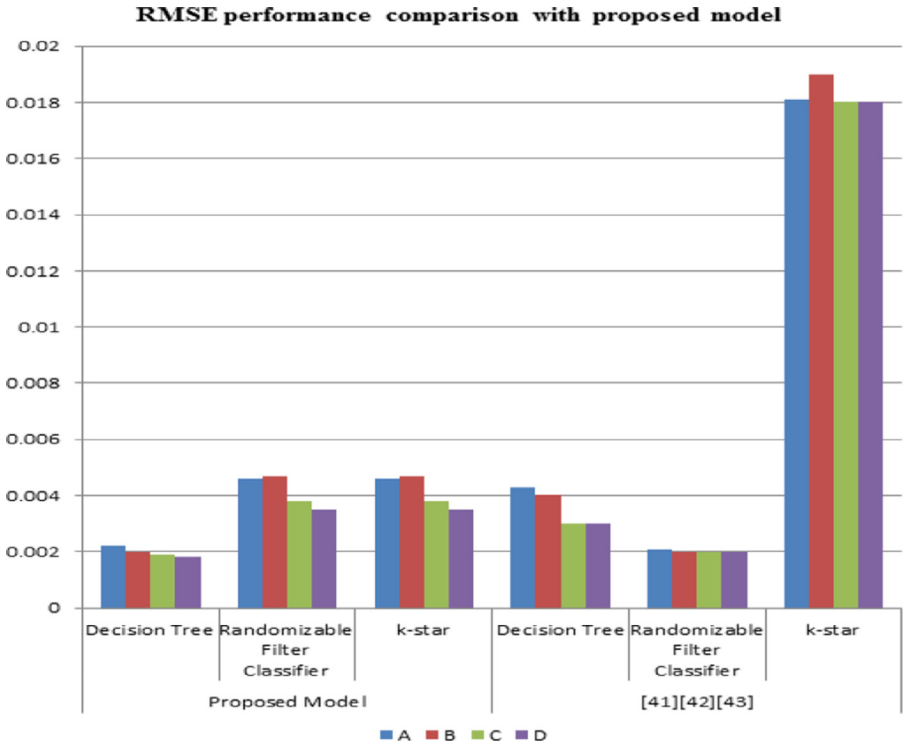

Figure 3. RMSE performance comparison for all test datasets with proposed model.

Table 1
Associated range values with respective Risk factors.

Range Value	Risk factor
0-1	RC, I & P, IAS, DS
0-2	TPM, RE, DB, DI
0-3	DT,DL, DL & IS, SLA, A & AC
1-3	IDD, BC & SA, R, ShE, VV

Table 2
Comparison of RMSE performance of proposed model with others using all four classes of dataset.

Algorithm		A	B	C	D
Proposed Model	Decision Tree	**0.0022**	**0.002**	**0.0019**	**0.0018**
	Randomizable Filter Classifier	0.0046	0.0047	0.0038	0.0035
	k-star	0.0046	0.0047	0.0038	0.0035
[50]	Decision Tree	0.0043	0.004	0.003	0.003
[51]	Randomizable Filter	0.0021	**0.002**	0.002	0.002
[52]	Classifier				
	k-star	0.0181	0.019	0.018	**0.018**
[53]	K-mean	0.022			
[54]	SVM	0.1152			

- 95–5% (D)

Experiments were executed in WEKA which delivers a gathering of data pre-processing tools and machine learning algorithms in a GUI domain for algorithm evaluation and information exploration as shown in Table 1:

With every algorithm, Table 2 briefs best outcomes of test (RMSE) from all data % and comparison of proposed model RMSE performance with other model [50–54]:

The best performance in terms of RSME parameter is highlighted in each columns of Table 2. In [44,45] k-mean is used to evaluate the performance of the system whereas SVM is used to evaluate the system's performance. RMSE performance comparison between the proposed model with [41–43] is represented in Fig. 3 for entire dataset percentages and depicts that additional training data percentage of about (5% testing and 95% training) generates best outcomes, signifying better learning. From the results, it is clear that the RSME performance of proposed predictive model is better in case of decision tree and k* algorithm where in case of randomizable filter classifier previous was better.

The methodologies and learning approaches were explained first, followed by feature selection. Following that, several learning approaches such as decision trees, k-star, and others were considered. Finally, performance measurement measures are offered, which are utilised to assess the prediction models.

5. Conclusion

With the increase use of data as days goes by, systems of big data systems became one of important drives of innovation that delivers a path in managing information. Cloud domain extensively regulates big data resolutions by delivering modified domains to big data systems. While big data in cloud computing are robust and powerful systems enabling both, further research to develop and enhance and enterprises, but there arise few speculations regarding assessment of risk which is requited later and discussion and real investigation. Extra hard work must be used designing and developing risk assessment mechanism concerning to security in cloud computing domain for big data. Additional should be implemented in later but quickly to solve this risk assessment security issues. Major aim of this experimental hardship is to attain decrease of characteristics, best precision on data set of testing and finding out best existing schemes used for this dataset. It is inspected that the behaviour of various machine-learning algorithms for demonstrating risk factor involving cloud computing. Influence of subsets of testing and training of information is described in this paper by splitting sub dataset arbitrarily in four various classes. Experimental outcome depicts that splitting of dataset to (95% - 5%) provides best yield out of entire remaining partitioning and also illustrates that Decision Tree Classifier algorithm & k* delivers better results between all data sets in cloud computing

domain, whereas randomizable filter classifier was having fever performance that previous one.

Future Work: Because security is a top priority in the cloud computing environment, the security framework will inform both cloud customers and cloud service providers about individual boundaries and shared responsibilities at each level. Cloud actors can evaluate security parameters and compliance by simulating the Cloud Computing security framework in their in-house or external cloud environments. As a result, as a future project, it would like to conduct research on the comparative analysis of the performance of smart security assessment models or frameworks through simulation and integration of security standards and guidelines for service and delivery models, which will aid in the framework's benchmarking. Additional research is being done in the subject of integrating Cloud Service Level Agreements (SLAs) with adaptive and smart security frameworks, which will help a lot of CSPs to guarantee the service level that their customers need. In the future, the industrial need for real-time risk assessment may also fuel the adoption of machine learning techniques. Moving forward, procedures to validate the use of machine learning in risk assessment also need to be addressed by the various safety regulatory bodies.

References

[1] H. Guo, L. Wang, F. Chen, D. Liang, Scientific big data and digital earth, Chin. Sci. Bull. 59 (35) (2014) 5066–5073.

[2] M.G. Porcedda, Patching the patchwork: appraising the EU regulatory framework on cyber security breaches, Comput. Law Secur. Rev. 34 (5) (2018) 1077–1098.

[3] P. Subramani, P. BD, Prediction of muscular paralysis disease based on hybrid feature extraction with machine learning technique for COVID-19 and post-COVID-19 patients, Pers. Ubiquit. Comput. (2021) 1–14.

[4] Rajkumar Buyya, Rajiv Ranjan, Rodrigo N. Calheiros, Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities, in: 2009 international conference on high performance computing & simulation, IEEE, 2009, pp. 1–11.

[5] D.N. Tran, T.N. Nguyen, P.C.P. Khanh, D.T. Trana, An iot-based design using accelerometers in animal behavior recognition systems, IEEE Sensors J. (2021).

[6] Michael Miller, Cloud computing: Web-based applications that change the way you work and collaborate online, Que publishing, 2008.

[7] K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system, IEEE Trans. Intell. Transp. Syst. 22 (7) (2020) 4337–4347.

[8] Laxmana Rao Battula, Network security function virtualization (nsfv) towards cloud computing with nfv over openflow infrastructure: Challenges and novel approaches, in: 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2014, pp. 1622–1628.

[9] B.D. Parameshachari, Big data analytics on weather data: predictive analysis using multi node cluster architecture, Int. J. Comput. Appl. (2022) 0975–8887.

[10] Abhishek Sharma, Umesh Kumar Singh Dr., Cloud Computing Security Framework Based on Shared responsibility Model, "Cyber-Physical, IoT, and Autonomous Systems in Industry 4.0" (December) (2021) 39–56 ISBN: 9780367705152 3. In this issue, doi:10.1201/9781003146711-3.

[11] N.T. Le, J.W. Wang, D.H. Le, C.C. Wang, T.N. Nguyen, Fingerprint enhancement based on tensor of wavelet subbands for classification, IEEE Access 8 (2020) 6602–6615.

[12] Anurag Jain, Rajneesh Kumar, A taxonomy of cloud computing, Int. J. Scient. Res. Publ. 4 (7) (2014) 1–5.

[13] Z. Guo, Y. Shen, A.K. Bashir, M. Imran, N. Kumar, D. Zhang, K. Yu, Robust spammer detection using collaborative neural network in Internet-of-Things applications, IEEE Internet Things J. 8 (12) (2020) 9549–9558.

[14] Laura. Savu, Cloud computing: Deployment models, delivery models, risks and research challenges, in: 2011 International Conference on Computer and Management (CAMAN), IEEE, 2011, pp. 1–4.

[15] B. Rachana, T. Priyanka, K.N. Sahana, T.R. Supritha, B.D. Parameshachari, R. Sunitha, Detection of polycystic ovarian syndrome using follicle recognition technique, Global Trans. Proc. 2 (2) (2021) 304–308.

[16] Archana Mantri, Suman Nandi, Gaurav Kumar, Sandeep Kumar, High performance architecture and grid computing, International Conference, HPAGC, 2011.

[17] D.L. Vu, T.K. Nguyen, T.V. Nguyen, T.N. Nguyen, F. Massacci, P.H. Phung, HIT4Mal: Hybrid image transformation for malware classification, Trans. Emerg. Telecommun. Technol. 31 (11) (2020) e3789.

[18] edited by Mainak Adhikari, Aditi Das, Akash Mukherjee, Utility Computing and Its Utilization, in: Ganesh Chandra Deka, G.M. Siddesh, K.G. Srinivasa, L.M. Patnaik (Eds.), In Emerging Research Surrounding Power Consumption and Performance Issues in Utility Computing, IGI Global, Hershey, PA, 2016, pp. 1–21, doi:10.4018/978-1-4666-8853-7.ch001.

[19] L. Tan, K. Yu, F. Ming, X. Chen, G. Srivastava, Secure and resilient artificial intelligence of things: a HoneyNet approach for threat detection and situational awareness, IEEE Consumer Electronics Magazine, 2021.

[20] Daniel Beimborn, Thomas Miletzki, Stefan Wenzel, Platform as a service (PaaS), Bus. Inf. Syst. Eng. 3 (6) (2011) 381–384.

[21] Wesam Dawoud, Ibrahim Takouna, Christoph Meinel, Infrastructure as a service security: Challenges and solutions, in: 2010 the 7th International Conference on Informatics and Systems (INFOS), IEEE, 2010, pp. 1–8.

[22] Jong-Hei. Ra, Qualitative study on service features for cloud computing, J. Digit. Contents Soc. 12 (3) (2011) 319–327.

[23] Subashini Subashini, Veeraruna Kavitha, A survey on security issues in service delivery models of cloud computing, J. Netw. Comput. Appl. 34 (1) (2011) 1–11.

[24] Umer A. Butt, Muhammad Mehmood, Syed B.H. Shah, Rashid Amin, M.W. Shaukat, Syed M. Raza, Doug Y. Suh, Md.J. Piran, A review of machine learning algorithms for cloud computing security, Electronics 9 (9) (2020) 1379 9091379, doi:10.3390/electronics.

[25] Raj Samani, Jim Reavis, Brian Honan, CSA guide to cloud computing: Implementing cloud privacy and security, Syngress, 2014.

[26] Abhishek Sharma, & Dr, Umesh Kumar Singh, Deployment model of e-educational cloud for departmental academics automation using open source, HTL J. 27 (5) (2021) 36 ISSN 1006-6748, doi:10.37896/HTL27.5/3535.

[27] Dan Hubbard, Michael Sutton, Top threats to cloud computing v1. 0, Cloud Secur. Alliance (2010) 1–14.

[28] Abhishek Sharma, Umesh Kumar Singh, et al., An Investigation of Security Risk & Taxonomy of Cloud Computing Environment, IEEE 2ndInternational conference on Smart Electronics and Communication (ICOSEC 2021), 2022 ISBN: 978-1-6654-3368-6.

[29] Abhishek Sharma, Umesh Kumar Singh, et al., A Comparative analysis of security issues & vulnerabilities of leading Cloud Service Providers and in-house University Cloud platform for hosting E-Educational applications, IEEE Mysore Sub Section International Conference (MysuruCon), 2021 ISBN: 978-0-7381-4662-1.

[30] R.A. Attar, J. Al-Nemri, A. Homsi, A. Qusef, Risk Assessment for Emerging Domains (IoT, Cloud Computing, and AI), in: 2021 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 2021, pp. 120–127, doi:10.1109/JEEIT53412.2021.9634156.

[31] M. Zekri, S.E. Kafhali, N. Aboutabit, Y. Saadi, DDoS attack detection using machine learning techniques in cloud computing environments, in: 2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech), 2017, pp. 1–7, doi:10.1109/CloudTech.2017.8284731.

[32] K.K. Nguyen, D.T. Hoang, D. Niyato, P. Wang, D. Nguyen, E. Dutkiewicz, Cyberattack detection in mobile cloud computing: A deep learning approach, in: 2018 IEEE Wireless Communications and Networking Conference (WCNC), 2018, pp. 1–6, doi:10.1109/WCNC.2018.8376973.

[33] Song-Shun Lin, Shui-Long Shen, Annan Zhou, Ye-Shuang Xu, Risk assessment and management of excavation system based on fuzzy set theory and machine learning methods, Autom. Constr. 122 (2021) 103490 ISSN 0926-5805, doi:10.1016/j.autcon.2020.103490.

[34] Jeevith Hegde, Børge Rokseth, Applications of machine learning methods for engineering risk assessment – A review, Saf. Sci. 122 (2020) 104492 ISSN 0925-7535, doi:10.1016/j.ssci.2019.09.015.

[35] P. Radanliev, D. De Roure, K. Page, et al., Design of a dynamic and self-adapting system, supported with artificial intelligence, machine learning and real-time intelligence for predictive cyber risk analytics in extreme environments – cyber risk in the colonisation of Mars, Saf. Extreme Environ. 2 (2020) 219–230, doi:10.1007/s42797-021-00025-1.

[36] R Vinayakumar, M Alazab, KP Soman, P Poornachandran, A Al-Nemrat, S Venkatraman, Deep learning approach for intelligent intrusion detection system, IEEE Access 7 (2019) 41525–41550, doi:10.1109/ACCESS.2019.2895334.

[37] AA Diro, N Chilamkurti, Distributed attack detection scheme using deep learning approach for Internet of Things, Futur. Gener. Comput. Syst. 82 (2018) 761–768, doi:10.1016/j.future.2017.08.043.

[38] D Berman, A Buczak, J Chavis, C Corbett, A survey of deep learning methods for cyber security, Information 10 (4) (2019) 122, doi:10.3390/info10040122.

[39] D. Sun, Z. Wu, Y. Wang, Q. Lv, B. Hu, Risk prediction for imbalanced data in cyber security: a Siamese network-based deep learning classification framework, in: Proceedings of the international joint conference on neural networks, 2019-July, 1–8, 2019, doi:10.1109/IJCNN.2019.8852030.

[40] Rohit Bhadauria, NabenduChaki RituparnaChaki, Sugata Sanyal, A survey on security issues in cloud computing, arXiv preprint arXiv:1109.5388 (2011) 1–15.

[41] Mohammad Masdari, Marzie Jalali, A survey and taxonomy of DoS attacks in cloud computing, Secur. Commun. Netw. 9 (16) (2016) 3724–3751.

[42] Harshit Srivastava, SathishAlampalayam Kumar, Control framework for secure cloud computing, J. Inf. Secur. 6 (01) (2014) 12.

[43] Amit Wadhwa Varsha, Swati Gupta, Study of security issues in cloud computing, Int. J. Comput. Sci. Mob. Comput. 4 (6) (2015) 230–234 ISSN 2320-088X, IJCSMCpg..

[44] Khalid H. Alotaibi, Threat in Cloud-denial of service (DoS) and distributed denial of service (DDoS) attack, and security measures, J. Emerg. Trends Comput. Inf. Sci. 6 (5) (2015) 241–244.

[45] Atulay Mahajan, Sangeeta Sharma, The malicious insiders threat in the cloud, Int. J. Eng. Res. Gen. Sci. 3 (2) (2015) 245–256.

[46] K. Vijayakumar, Security issues and algorithms in cloud computing, Global J. Adv. Res. 2 (3) (2022) 569–574.

[47] Pallavi Marathe, Cloud Computing Security threats and tools, 4, 2015 ISSN-2319—8354(E).

[48] Smita Parte, Noumita Dehariya, Cloud computing: issues regarding security, applications and mobile cloud computing, Int. J. Advanc. Res. Comp. Sci. Softw. Eng 5 (3) (2015) 403–406.

[49] ShaikKhaja Mohiddin, Suresh BabuYalavarthi, Research challenges in the emerging trends of cloud computing, Int. J. Adv. Comput. Sci. Technol. (IJACST) 4 (1) (2015) 4.

[50] Nada Ahmed Mohammednour Eisa, PhD diss, Sudan University of Science and Technology, 2016.
[51] M. Hall, et al., The WEKA data mining software: an update, ACM SIGKDD Explor. Newslett. 11 (1) (2009) 10–18.
[52] I.H. Witten, et al., Weka: Practical machine learning tools and techniques with Java implementations, 1999.
[53] Zhe Mi, Tiangang Wang, Zan Sun, Rajeev Kumar, Vibration signal diagnosis and analysis of rotating machine by utilizing cloud computing, Nonlinear Eng. 10 (1) (2021) 404–413, doi:10.1515/nleng-2021-0032.
[54] Priyanka Nehra, A. Nagaraju, Host utilization prediction using hybrid kernel based support vector regression in cloud data centers, J. King Saud Univ. - Comput. Inf. Sci. (2021) 1319–1578, doi:10.1016/j.jksuci.2021.04.011.