People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University Larbi Tébessi - Tébessa
Faculty of Exact Sciences and Natural and Life Sciences
Department: Mathematics & Computer science

Submitted in partial fulfillment of the requirements for MASTER Degree
Domain: Computer Science
Field: Networks and IT security

Theme
Presented by:
Guenez Yamina

Cyberattack detection in mobile cloud computing: a deep learning approach

In front of the jury:

Mr.Sahraoui abdelatif	MCB	Univérsity Larbi Tébessi	President
Mr.Gasmi Mohamed	MCB	Univérsity Larbi Tébessi	Examiner
Mr.Amroune Mohamed	MCA	Univérsity Larbi Tébessi	Supervisor
Mrs.Salima Bourougaa-tria	MCB	Univérsity Larbi Tébessi	C.Supervisor

Date: 15 /09/ 2020


بسم الله الرحمن الرحيم

Acknowledgement

During my final thesis, I received moral and technical support from several people which made my work environment very pleasant.

Foremost, I would like to express my sincere gratitude to my advisor Mr. Amroune & Mrs.Bourougaa for the continuous support of my research, for their patience, motivation, enthusiasm, and immense knowledge.

I would like to express my gratitude and high consideration to Professor sahraoui abdelatif, professor at the University of Larbi Tébessade, for having honored me by accepting to chair the jury.

Finally, I would like to thank the member of the jury professor of the University of Larbi Tébessa: Mr.Gasmi Mohamed for the honor he did me by agreeing to examine my work.

Dedication

I dedicate my dissertation work to my family and many friends. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears. My Brother Nassim, Ahmed and Abderahmane, and my sister Wafa and my best friend zaineb who have never left my side and are very special.

Table of Contents

Abstract / Résumé / الملخص.....viii

General Introduction.

1. Scientific Framework ***.....1***

2. Problematic ***.....2***

3. The purpose of our work ***.....2***

4. The structure of the dissertation ***.....3***

5. Introduction ***.....5***

6. A Brief History ***.....5***

7. Cloud Computing ***.....6***

7.1 Definition 6

7.2 Objectives 8

7.3 Features of Cloud Computing..... 8

7.4 Elements of the Cloud Computing 9

7.5 Forms of Cloud Computing deployment 11

7.6 Cloud Computing services 12

7.7 Main Cloud applications 13

7.8 Cloud Computing Actors.....14

7.9 Cloud Computing advantages.....15

7.10 Cloud Computing Limits.....15

8. Security in Cloud Computing ***.....16***

8.1 Objectives and main security services 16
8.2 Security issues in Cloud Computing 17
8.3 Classification of attackers 19
8.4 Classification of attacks 20

9. Data Security in Cloud Computing 26
9.1 Data Life Cycle in Cloud Computing 26
9.2 Protective Measures 27
9.3 Different solutions proposed 28

10. Conclusion 34

Chapter 01: Cyberattacks in mobile cloud computing.

11. Introduction 36

12. Machine Learning 37
12.1 Classification 37
12.1.1 Types of classification 38
12.1.2 Bayesian Optimization methods 39
12.2 Stochastic gradient descent 40

Chapter 02: Deep Learning.

13. Deep Learning 40
13.1 History of Deep Learning 40
13.2 Definition of Deep Learning 45
13.3 How Deep learning works 46

14. Fields of applications of deep learning 47

15. Architectures of deep neural networks 48
15.1 Convolutional neural networks (CNN) 48
15.1.1 The convolution operation 49
15.1.2 CNN layers 51
15.1.3 CNN parameters 54
15.2 Recurrent neural network (RNN) 57
15.2.1 What is RNN ? 58
15.2.2 RNN application 59

15.3 Reinforcement learning 59

16. Conclusion 60

Chapter 03: Cyberattack detection based on deep learning.

17. Introduction: 62

18. Related works: 62

19. Conclusion: 73

Chapter 04: Cyberattack detection in mobile cloud computing based on deep learning(Contribution).

20. Introduction 75

21. Execution environment 75

22. System Architecture: 77

23. Datasets and Features Analysis: 77

23.1 KDD-Cup99: 78

23.2 NSL-KDD: 78

23.3 UNSW-NB15: 81

24. Attack detection module architecture: 83

25. Neural network architecture: 84

26. Results and Discussions 86

26.1 Learning and testing the proposed models 86

26.2 Our models comparison results: 90

26.3 Evaluation Methods : 90

26.4 The basic models used for the comparison: 93

26.5 The performance of the classification model 94

27. Conclusion 95

28. Bibliography 99

List of Figures

Figure 7.1: the different components of cloud computing. 7
Figure 7.2: Mobile cloud computing architecture. 7
Figure 7.3: a Datacenter example. 10
Figure 7.4: Cloud computing deployment models 11
Figure 7.5: the different layers of the Cloud Computing. 12
Figure 8.1: Types of attacks in Cloud Computing. 20
Figure 8.2: Man in The Middle attack. 25
Figure 9.1: security model in the Cloud 31
Figure 9.2: Keyless Encryption principale 32
Figure11.1: Relationship between AI & ML & Deep Learning. 36
Figure 13.1: On the left the diagram of a biological neuron and on the right the diagram of the formal neuron from. 41
Figure 13.2: Deep Learning schema. 42
Figure 13.3: Example on how deep learning works. 47
Figure 15.1: Standard Architecture of a deep neural network. 48
Figure 15.2: Diagram of the route of the filter window on the image. 50
Figure 15.3: Example of a convolution whose configuration is: Operation = Maximum argument, horizontal step = 1 pixel, vertical step = 1 pixel. 50
Figure 15.4: Pooling with a 2x2 filter & a step of 2. 52
Figure 15.5: Importance of context in the recognition of handwriting. 58
Figure 15.6: example of an agent using reinforcement learning. 59
Figure 18.1: A typical three-layer network, with five inputs, eight hidden units, and two outputs. 63
Figure 18.2: Precision of different algorithms used to recognize normal traffic. 64
Figure 18.3: Precision of different algorithms used to recognize DOS traffic. 64
Figure 18.4: The unfolded Recurrent Neural Network. 65
Figure 18.5: Block diagram of proposed RNN-IDS. 66
Figure 18.6: The Accuracy on the KDDTest+ and KDDTest−21datasets in the Binary Classification. 66
Figure 18.7:Visualizations of three datasets: (a) NSL-KDD, (b) UNSW-NB15, and (c) KDDcup 1999,

by using PCA with 3 most important features. Grey circlesrepresent normal packets, while circles with other colors than grey express the different types of attacks.....67
Figure 18.8: Steps of creating attack images. 69
Figure 18.9: Design of CNN model. 70
Figure 18.10: Cont.....71
Figure 21.1: The growth in popularity of TensorFlow. 76
Figure 22.1 : System model architecture for cyberattack detection in mobile cloud computing systems.....77
Figure 23.1 : Attacks Categories with its types of the KDD-CUP99 Dataset 78
Figure 23.2 : Attacks distribution of the NSL-KDD & KDD-CUP99 Datasets. 80
Figure 23.3 : NSL-KDD dataset features. 80
Figure 24.1 : Attack detection module architecture for cyberattack detection in mobile cloud computing system. 84
Figure 25.1 :The first CNN Architecture of NSL_KDD 85
Figure 25.2 : The second CNN Architecture of NSL_KDD 85
Figure 25.3 : CNN Architecture model of UNSW-NB15 85
Figure 25.4 : CNN Architecture model of KDD_CUP99.....85
Figure 26.1 : CNN first model architecture training & test accuracy epochs for NSL_KDD 87
Figure 26.2 : CNN first model architecture training & test loss epochs for NSL_KDD 87
Figure 26.3 : CNN first model architecture training & test accuracy epochs for NSL_KDD with different feature analysis method. 87
Figure 26.4 : CNN first model architecture training & test accuracy epochs for NSL_KDD with different feature analysis method. 87
Figure 26.5 : CNN second model architecture training & test accuracy epochs for NSL_KDD. 87
Figure 26.6 : CNN second model architecture training & test loss epochs for NSL_KDD. 87
Figure 26.7 : CNN first model architecture training & test accuracy epochs for KDD_CUP99. 88
Figure 26.8 : CNN first model architecture training & test loss epochs for KDD_CUP99. 88
Figure 26.9 : CNN first model architecture training & test accuracy epochs for KDD_CUP99 with different feature analysis method. 88
Figure 26.10 : CNN first model architecture training & test loss epochs for KDD_CUP99 with different feature analysis method. 88
Figure 26.11 : CNN first model architecture training & test accuracy epochs for UNSW_NB15. 88
Figure 26.12 : CNN first model architecture training & test loss epochs for UNSW_NB15. 88
Figure 26.13 : CNN second model architecture training & test accuracy epochs for UNSW_NB15. 89
Figure 26.14 : CNN second model architecture training & test loss epochs for UNSW_NB15. 89

Figure 26.15: TP, FN, FP, TN of DOS Class. 92
Figure 26.16: TP, FN, FP, TN of Probe Class. 92
Figure 26.17: TP, FN, FP, TN of R2L Class. 92
Figure 26.18: TP, FN, FP, TN of Normal Class. 92
Figure 26.19: Confusion matrix without normalization for attack detection first model. 94
Figure 26.20: Normalized Confusion matrix for attack detection first model. 94
Figure 26.21: Confusion matrix without normalization for attack detection second model. 94
Figure 26.22: Normalized Confusion matrix for attack detection second model. 94

List of Tables

Table 18.1: Selected studies results. 72
Table 23.1: Features of UNSW-NB15 dataset 81
Table 26.1: accuracy results of NSL_KDD. 90
Table 26.2: accuracy results of KDD_CUP99. 90
Table 26.3: accuracy results of UNSW_NB15. 90
Table26.4: Precision and Recall comparison. 91
Table 26.5: Accuracy comparison with other models. 93

Abstract / Résumé / الملخص

Abstract: The security of our data and systems was and will always be the main subject that we're trying to tackle, especially with the fast growth of technology in different fields such as mobile cloud computing. This fast growth is always accompanied by serious security issues that threaten our data privacy and integrity. That's why we need an effective approach of detection in order to prevent those cyber threats and protect our data in an efficient way. In this work, we used a promising Deep Learning approach based on CNN 1D (Convolutional Neural Networks) with different architectures to tackle these kinds of security issues.

Keywords: Cybersecurity; cyberattack; mobile cloud computing; deep learning; CNN 1D.

Résumé: La sécurité de nos données et de nos systèmes a été et sera toujours le principal sujet que nous essayons d'aborder, en particulier avec la croissance rapide de la technologie dans différents domaines tels que le cloud computing mobile. Cette croissance rapide s'accompagne toujours de graves problèmes de sécurité qui menacent la confidentialité et l'intégrité de nos données. C'est pourquoi nous avons besoin d'une approche efficace de détection afin de prévenir ces cyber menaces et de protéger nos données de manière efficace. Dans ce travail, nous avons utilisé une approche prometteuse de l'apprentissage profond basée sur CNN 1D (Convolutional Neural Networks) avec différentes architectures pour résoudre ce type de problèmes de sécurité.

Mots clés: cyber-sécurité; cyber-attaque; cloud computing mobile; l'apprentissage profond; CNN 1D.

الملخص:أمن البيانات والأنظمة كان وسيظل دائمًا الموضوع الرئيسي الذي نحاول معالجته ، خاصة مع النمو السريع للتكنولوجيا في مجالات مختلفة مثل مجال الحوسبة السحابية المتنقلة.

دائمًا ما كان هذا النمو السريع مصحوبًا بمشاكل أمنية خطيرة تهدد خصوصية بياناتنا وسلامتها. لهذا السبب نحتاج إلى نهج فعال للكشف عن مثل هذه التهديدات من أجل منع تلك الهجمات الإلكترونية وحماية بياناتنا بطريقة فعالة.

في هذا العمل، استخدمنا نهجًا واعدًا للتعلم العميق يعتمد على CNN 1D لمعالجة هذه الأنواع من المشاكل الأمنية.

الكلمات المفتاحية: الأمن السيبراني ؛ الهجمات الألكترونية ؛ الحوسبة السحابية المتنقلة ؛ التعلم العميق؛ CNN 1D

General

Introduction

1. Scientific Framework

In recent years, the number of network attacks has increased dramatically; the interest in spotting cyber attacks has increased among researchers because of the growing need for information security, as computing resources are more vulnerable and more dependent on them than before, mobile Cloud Computing in particular is more vulnerable than any other resource or host, because all user's private and important informations are stored online and users has no control to isolate or keep them safe.

Cyberattack detection is considered a necessary security mechanism to deal with network attacks and identify malicious activity in Mobile cloud computing environment. It plays a vital role in information security technology and helps to discover, determine and identify the unauthorized access, use, duplication, modification and destruction of information. This important research achievement is generally equivalent to a classification problem, such as a binary problem or a multi-class classification problem that identifies whether the behavior of the network is normal or abnormal.

In short, Cyberattack detection is primarily driven by improving the accuracy of classifiers with respect to effectively identify attacks.

Artificial Intelligence (AI) is one of the newest and most popular fields. The serious use of Artificial Intelligence began in earnest after World War II, and the name itself was coined in 1956.

Its popularity is due to recent developments in deep learning area which is one of the most promising data mining methods and the successor to ML algorithms.

ML algorithms work well for a wide variety of problems. However, they failed to solve a few major AI problems such as speech recognition and object recognition. The development of deep learning was driven in part by the failure of traditional algorithms in such AI tasks. But it was only after larger amounts of data became available, thanks in particular to Big Data and

connected objects, and because of that computing machines became more powerful, which helped us understand the real potential of deep learning.

Deep Learning is a new area of ML research, which was introduced with the aim of bringing ML closer to its main goal: artificial intelligence.
Deep learning uses algorithms inspired by the structure and function of the brain. They can learn several levels of representation in order to model complex relationships between data.
In this study, we create several Deep Learning models and apply them to the fields of cyberattack detection.

2. Problematic

User's Private data is always a priority that needs to be protected from malicious users, and cyberattack detection plays a vital role in information security technology and helps discover, determine and identify unauthorized users. Therefore, improving its performance and efficiency in detecting multiple cyberattacks simultaneously and with a high accuracy is a priority.

3. The purpose of our work

The aim of this work is to study and compare and propose a new cyberattack detection system based on Deep Learning.
Our main goals consist of:

- Propose a cyberattack detection system architecture that detects malicious users early in the mobile cloud computing environment and helps detect various and multiple attacks simultaneously.

- Develop different features extraction methods and Create different deep learning model architectures and test them on several datasets.

- Finally, compare the results to see which one is the optimal to use in order to ensure the security of the cloud computing network.

4. The structure of the dissertation

- Chapter I: First of all, we started our thesis with the first chapter which introduced the basic concepts of the cloud computing environment and security in the mobile cloud computing.

- Chapter II: We then presented the field of Deep Learning in the second chapter.

- Chapter III: Here, we presented cyberattacks detection models based on deep learning.

- Chapter IV: In the last chapter, we presented our different architectures and implementation of deep learning for cyberattack detection.

CHAPTRE I

Cyberattacks in mobile cloud computing

5. Introduction

Information and communication technologies (ICT) are developing faster and progressively. In recent years there is a new destination, its goal is to improve services in the ICT field, it's called "Cloud Computing".

Cloud security is a sub-domain of Cloud Computing in relation to IT security. It involves concepts such as network, hardware and control security that are deployed to protect the data, applications and infrastructure associated with Cloud Computing. An important aspect of the Cloud is the notion of interconnection with various hardware which makes it difficult and necessary to secure these environments.

In this chapter we give some general information on Cloud Computing, namely its definition, its different types, the services it offers, its advantages and disadvantages, as well as the security of the Cloud. We will start with the security objectives and services, and then we will present the security problems in the cloud as well as the different attackers. Finally, we will finish with the different types of attacks on Cloud Computing and their solutions.

6. A Brief History

The first statement of the concept of Cloud Computing dates from 1960, when John McCarthy affirmed that the computer resource would be accessible and consumed by the public in the same way as the distribution of water and energy [1].

The concept of Cloud Computing was adopted for the first time in 2002 by Amazon, a leader in e-business, who had invested in a huge fleet of machines, sized to absorb the large load of orders placed on their site during the holidays Christmas, but relatively untapped the rest of the year. Under sizing their fleet would have caused their site to be unavailable during peak periods, thus jeopardizing their business during the holidays (a large part of their turnover). Their idea was therefore to open all these unused resources to companies, so that they could rent them on demand. Since then, Amazon has invested heavily in this area and continues to expand its fleet and services.

7. Cloud Computing

7.1 Definition

The term Cloud Computing is a concept which represents access on demand, to information and services located on a remote server.

Mobile Cloud Computing (MCC) is the combination of cloud computing and mobile computing to bring rich computational resources to mobile users, network operators, as well as cloud computing providers.

The main idea to remember is that the Cloud is not a set of technologies, but a model for the provision, management and consumption of IT services and resources. The concept of Cloud Computing is broad and it cannot be reduced to a simple definition. Indeed, many definitions exist, here are the main ones:

NIST (National Institute of Standards and Technology) defines Cloud Computing as follows: "Cloud Computing is the set of disciplines, practices, technologies and business models used to deliver IT capabilities as a service on demand and over the network (software, platforms, hardware) " [2].

For the CIGREF working group (Large enterprise networks) Cloud Computing is defined by the following four points [3]:

- A Cloud is always a virtual space.
- Containers of information that are fragmented.
- The fragments are duplicated and distributed in this virtual / physical space.

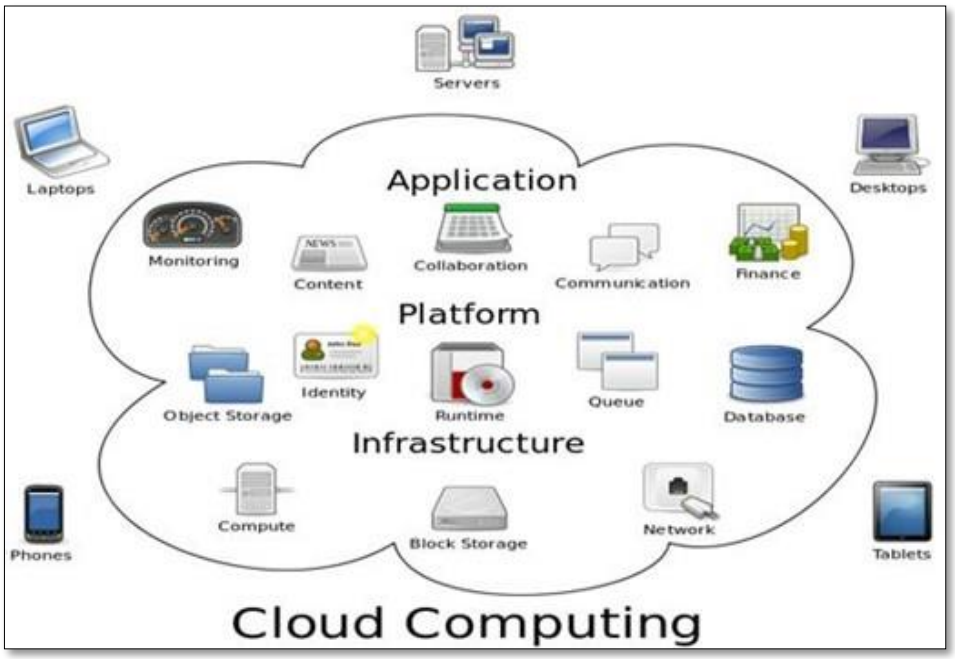


Figure 7.1: the different components of cloud computing.

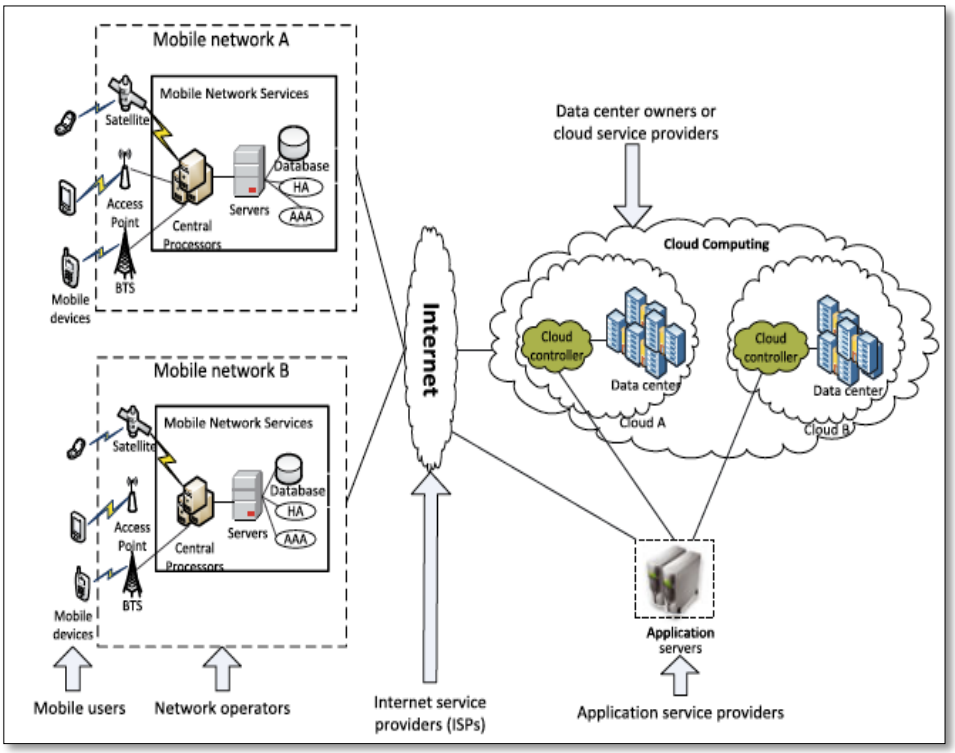


Figure 7.2: Mobile cloud computing architecture.

7.2 Objectives

The main objective of Cloud Computing is to create a community of practice on the concept of Cloud Computing. The specific objectives pursued are:

- Provide objective data on the field to build a dynamic information base;
- Generate interest and provoke the reflection of stakeholders on the field ;
- Stimulate participation, collaboration and promote the enrichment of the information base of the field;
- Ensure the permanence and visibility of the community of practice on Cloud Computing.

7.3 Features of Cloud Computing

The main features of Cloud Computing are:

- Elasticity: It defines the capacity of a given infrastructure to adapt dynamically to change [4].
- Pay-as-you-use: The cost is proportional to the usage, so the user pays for exactly what he uses [5].
- Self-service: Resources are available when and where the client wants [5].
- Quality of the services measurement: assess and guarantee a level of performance and availability adapted to specific customer needs.
- Universal Network Access: Access to resources is very fast and using a network (Internet), by standard protocols in a very elastic way [6].
- pooling: In a Cloud Computing environment, we do not think of the number of servers, disk size, number of processors, etc. but in terms of computing power, total storage capacity, bandwidth available thanks to virtualization [2].
- Multi-tenancy: In the Cloud, the same application can be used by several customers at the same time, preserving the security and private data of each customer. This is possible by using virtualization tools which allow sharing a server among several users [7].

- Availability: The high availability of the platform is obtained thanks to the implementation of redundancy and / or replication techniques. So the Cloud provides a reliable service that is not susceptible to failure [4].
- Self-healing: Any cloud system must contain one or more copies of each deployed application, so that in case of malfunction of the current application, the application in copy will replace it. Copy applications must be maintained and updated each time the current application is modified [8].
- SLA (Service Level Management): With cloud services, a customer can negotiate the level of the service they want and they must pay for it. In the event that the Cloud resources are overloaded, the system creates other Cloud application entities using the virtualization tools available in order to comply with the terms of the SLA contract [9].

7.4 Elements of the Cloud Computing

The elements that can constitute the Cloud system are as follows:

- **Virtualization**

Virtualization is the main technology in the Cloud, it allows optimized management of hardware resources by having several virtual machines on a physical machine. It is a technology that allows greater modularity in load distribution and reconfiguration of servers in the event of evolution or momentary failure.

The principle of virtualization makes it possible to integrate different servers in more flexible ways for easier use.

The goal of virtualization is to make the use transparent and the use of resources efficient, to ensure the operation of different services and the separation between multiple tenants (users) involved in physical hardware [10].

- **Infrastructure**

Cloud computing infrastructure is an assembly of servers, storage and network components organized to allow incremental growth greater than that achieved with conventional infrastructure. These components must be selected for their ability to meet the requirements for extensibility, efficiency, robustness and security. Conventional enterprise servers are unaware

of the network capabilities, reliability, or other qualities necessary to efficiently and securely secure Service Level Agreements (SLAs). In addition, cloud servers have lower operating costs and can be more reliable if they are not all equipped with internal disks [6].

- **Datacenter**

A datacenter is a physical site on which are grouped equipment constituting the company's information system (mainframes, servers, storage racks, network and telecommunications equipment, etc.). It can be internal or external to the company, operated or not with the support of service providers. It generally includes environmental control (air conditioning, fire prevention system, etc.), emergency and redundant power, as well as high physical security. Individuals or companies can come and store their data there according to well-defined methods [11]. We distinguish four forms of Community Cloud. In the following, we describe each of them in detail:

Figure 7.3: a Datacenter example.

- **Collaborative platform**

A collaborative platform is a virtual workspace. It is a tool, sometimes in the form of a website that centralizes all the tools related to the conduct of a project and makes them available to stakeholders (customers). The objective of collaborative work is to facilitate and optimize communication between individuals within the framework of work or a task [11].

7.5 Forms of Cloud Computing deployment

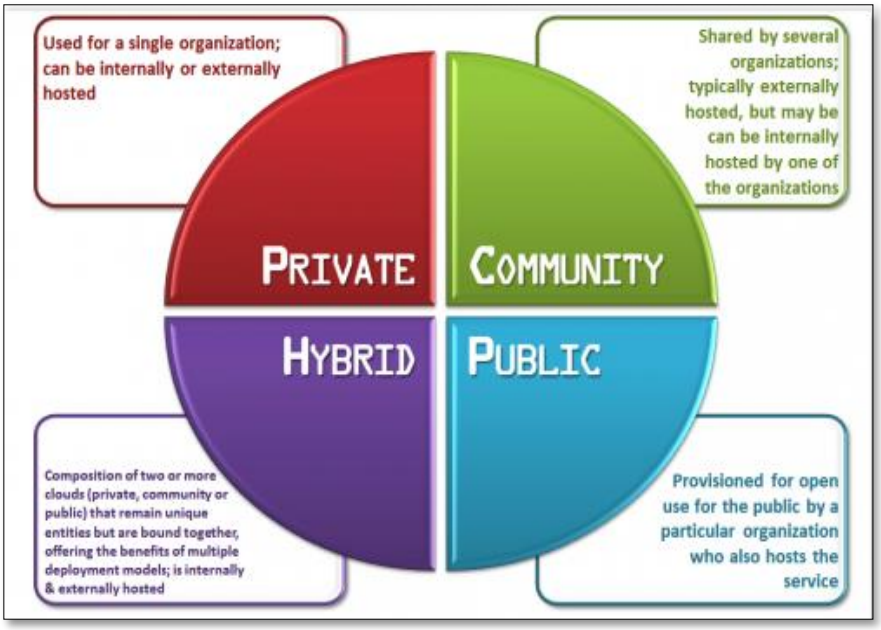


Figure 7.4: Cloud computing deployment models [5].

- **Public Cloud**

The public cloud also appeared first, its principle being to host applications, generally Web applications, on a shared environment with an unlimited number of users. It is a set of services and resources accessible via the Internet and managed by an external service provider (supplier), these resources and services are shared between several customers, used on demand and at any time without knowing where they exist, also these services can be free or paid. In the case of paid services, there are SLA contracts (Service Level Agreement) between customers and suppliers, SLA is a document that defines the quality of service required between the two [12].

- **Private Cloud**

A private cloud is a set of services and resources available to a single customer, for example a company, it can be managed by the company itself, or with its branches, in this case it is called "The private cloud Internal ", in other ways it can be managed by an external provider hired by the company, in this case called "The External Private Cloud", it is accessible via secure networks such as VPN (Virtual Private Network). The advantage of this type of cloud compared to the public cloud lies in the aspect of security and data protection [9].

- **Hybrid Cloud**

Hybrid cloud is the coexistence and communication between a private cloud and a public cloud in an organization sharing data and applications. That is to say, we can deport our applications to a public cloud which will consume data stored and exposed in a private cloud, or communicate two applications hosted in two separate private clouds, or even consume several services hosted in different clouds [13].

- **Community Cloud**

A community cloud is used by several organizations that have the same interests. In such architecture, the administration of the system can be carried out by one or more of the organizations sharing the resources of the Cloud. This can therefore involve hosting a very specialized business application, but common to very numerous companies, who decide to unite their efforts [14].

7.6 Cloud Computing services

Cloud Computing can be subdivided into 3 layers (see Figure 3.5): The infrastructure layer (IaaS) is managed by network architects, the platform layer (PaaS) intended for application developers and finally the application layer (SaaS) which is the end product for users.

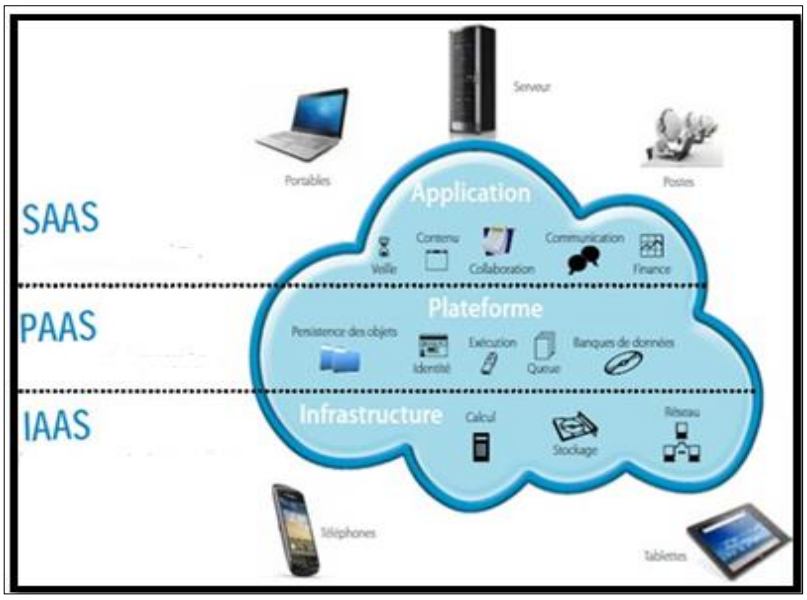


Figure 7.5: the different layers of the Cloud Computing.

IaaS (Infrastructure as a Service)

This is a model where the company has an IT infrastructure (computing capacity, storage and sufficient bandwidth) that is actually located at the supplier. This infrastructure is made available to automatically manage the workload required by the applications. However, the company has unrestricted access to it, as if the equipment were on its premises. This allows it to completely free itself from purchasing and managing equipment. This layer allows the company to focus first on its business processes without worrying about the hardware [15].

PaaS (Platform as a Service)

It is a platform for running, deploying and developing applications. PaaS brings together the developer (client) and system (supplier) part of Cloud Computing. It offers functions that deprive the developer of user management or availability issues, for example. The developer only needs to host his application so that it is available in SaaS [13].

SaaS (Software as a Service)

It is the provision via the Internet of computer applications (software) as a part of the subscription service, the data is also stored on a server of the operator SaaS. There is therefore no prerequisite on the client computer other than having network access to the Cloud (usually the Internet). Deployment, maintenance, supervision of the proper functioning of the application and data backup are then the responsibility of the service provider.
This is sort of the visible part of Cloud Computing for the end user, who no longer needs to install the application on his computer, and who accesses his account via the Web, on a secure environment [6].

7.7 Main Cloud applications

Many cloud applications have been proposed. The main applications found on the Clouds are as follows [15]:

- Messaging ;
- Collaborative and web-conferencing tools;
- Development and test environments;
- Sales force automation and Business Intelligence;
- Customer relationship management (CRM);
- Office utilities;
- Archiving and saving data;
- Mathematical engineering applications (3D modeling, simulation, CAD, etc.);
- Financial applications (stock market analysis, long-term analyzes, etc.);
- Accounting (cash management, invoicing, etc.);
- Human resources (recruitment management, payroll, etc.).

7.8 Cloud Computing Actors

Cloud technology has been adopted by the largest companies around the world, given the benefits it offers. The main cloud computing actors that are positioned are:

Amazon

Amazon has become the best-known cloud service provider in the world, thanks to all of the cloud services it offers and the contributions it makes [6].

Google

In 2008, Google launched its public cloud oriented for web services offering a platform (PaaS) called "Google App Engine" providing a large processing power and a large storage capacity, and allowing the hosting of Python applications or Java, as well as SaaS applications grouped

in the "Google App" range [16].

Microsoft

Microsoft provides a cloud platform called Windows Azure. It is an offer for hosting applications, data and storage services, data synchronization, message bus, contact, etc. [17] Windows azure also offers a set of services.

7.9 Cloud Computing advantages

Cloud Computing has many advantages, here are the main ones [10]:

- Quick start: Cloud Computing allows you to test the business plan quickly, at low cost and with ease.
- Agility for the company: Solving IT management problems simply without having to make a long-term commitment.
- Faster product development: Reduction of research time for developers on the configuration of applications.
- No capital expenditure: No need for premises to expand IT infrastructure.
- Cost reduction: Users only pay for what they consume. Strong savings in cost and energy, especially in cases of non-constant or linear needs.
- Mobility: Users can at any time and from any device (Computers, Smartphones, Lap top, etc.) have access to data, applications, servers or platform independently of the terminal [18].

7.10 Cloud Computing Limits

Cloud Computing has the following disadvantages (limits):

- Internet connection required: Access to Cloud Computing services is via the Internet bailiff. The breakdown of the Internet connection implies the loss of access to applications and data [18].

- Data security: In the case of Cloud Computing, the company will have to connect its workstations to the Internet, and expose them to the risk of attack and intrusion, and of data theft by hackers [18].

- Bandwidth can blow your budget: The bandwidth that would be required to store data in the Cloud is gigantic, and the costs would be so high that it is more advantageous to buy storage rather than paying someone else to do it [18].

- Data storage: Physical storage of data in the Cloud is carried out by service providers, which limits handling of the latter by customers [17].

- Identifying clients: With the increasing use of the Cloud and the multi-location use of these resources, it is becoming increasingly difficult to identify by whom and from where the data was modified [17].

- Company size: If the company is large then the resources are large, which includes a large consumption of the Cloud. You may find it more beneficial to develop your own cloud rather than using an outsourced one. The gains are much greater when we go from a small consumption of resources to a higher consumption [19].

8. Security in Cloud Computing

8.1 Objectives and main security services

Computer security is the set of resources implemented to reduce the vulnerability of a system against accidental or intentional threats.

In order to ensure secure data transfer over communication networks, a number of security services are required. IT security generally has five main objectives [20]:

- **Authentication:** Consists of ensuring the identity of a user that is to say to guarantee to each correspondent that his partner is who he thinks he is. Access control can allow (for example by means of a password which must be encrypted) access to resources only to authorized persons. Without authentication, an attacker can impersonate another user to conduct his attack on the network.

- **Confidentiality:** a service which ensures that only authorized persons have access to the resources exchanged.

- **Integrity:** This is to ensure that the data is what it is believed to be. Checking the integrity of the data consists of determining whether the data has not been altered or modified without prior authorization during communication.

- **Availability:** Ensures that the requested services or resources are available. It helps maintain the proper functioning of the information system. The goal of availability is to guarantee access to a service or resources.

- **Non-repudiation:** It is the possibility of verifying that the sender and the recipient are the parts who claim to have respectively sent or received the message. In other words, the non-repudiation of the origin proves that the data was sent, and the non-repudiation of the arrival proves that the data was received.

These security objectives require the use of certain security mechanisms and services to be implemented. A security mechanism can be defined as a process or device, which aims to detect, or prevent or recover from a security attack. Security mechanisms such as shorthand, encryption, hashing, etc. are commonly used to ensure the security of a system. A security service can be identified as a processing or communication service aimed at improving data security and information transfer on an entity. These services help in the fight against security attacks. Security services usually use one or more mechanisms to achieve their objectives.

8.2 Security issues in Cloud Computing

As far as the cloud is concerned, security is one of the first concerns of companies transferring elements of their infrastructure to the cloud. There is also a contradiction, some companies fear the loss of vital information (patent, customer data), damage to their reputation, malicious acts

and interruptions of services. However, companies recognize that Cloud Computing can improve their security because providers must meet security requirements and regulatory obligations, while guaranteeing performance and compliance with service level agreements (SLAs).

Frequent use of cloud computing brings to light several security risks and problems [21]:

- **Access:** In the field of Cloud Computing, security needs also to relate to the access control and identity management. It's about putting in place precise controls to effectively filter who is authorized to use the company's applications brought to the Cloud.

- **Availability:** Availability ensures fast and reliable access to cloud data or Cloud Computing resources by the right person. Availability ensures that systems work properly when needed.

- **Network load:** In a Cloud strategy, we must be attentive to the availability of the network and therefore of the services. It is therefore important for a CIO (Information Systems Department) to validate with its supplier service commitments (SLA) and its mastery of physical and logical security issues, its performance and the speed of the proposed network.

- **Integrity:** The concept of cloud information integrity requires the following three principles:

  . Data changes are not made by unauthorized person or processes.
  . Unauthorized changes to data are not made by authorized person or processes.
  . Internal and external data consistency.

- **Data security:** The security of the data becomes an issue for the IT department. The main areas of cloud data security can be distinguished as follows:

  . Data protection against breaches, attacks, losses, malicious acts.
  . Data localization.
  . Access to data.

- **Location of data:** The geographic location of data storage is an important point which is also topical with what is called "the Sovereign Cloud" and "the Cloud in France". A service contract will determine the traditional legal and legal aspects, and will have to specify the country of execution of the contract because each country has legislation.

- **Data separation:** Cloud Computing often works on the principle of several leases, so that its servers are shared by different parties with the intention of hosting data and applications. Separating customer data often becomes a challenge with data from one party that is accessible from the other party. This poses serious challenges to maintaining customer privacy and often leads to the violation of sensitive data which may be specific to customers.

8.3 Classification of attackers

With the advent of the Internet, cybercrime has become increasingly widespread. Crimes are committed by ordinary adolescents or by real professionals:

Scripts kiddies

The most common hacker category included teenagers. This is often done by playing with scripts and other programs downloaded from the Internet and in a random way these scripts are used against random targets. Script kiddies are luckily easier to detect. Despite their low or no qualification level, kiddie scripts are sometimes a real threat to the security of systems. On the one hand, kiddie scripts are very numerous, and on the other hand, they are often obstinate to the point of sometimes spending several days to try all possible combinations of a password.

Real Hackers

Beyond the kiddie scripts, this is the category of real hackers, who are above all "network enthusiasts". They want to understand how computer systems work and to test both the capabilities of the tools and their knowledge. Generally they have a relatively high level of knowledge and creativity. Real hackers like to explore and exploit the weaknesses of any type of database system, application server, web server, etc. Most hackers claim to break into systems out of a passion for computing, not to destroy or steal data.

The internal threat

The third category is represented by the indoor pirate. In general, he is an employee or a former employee of a company who act out of personal revenge or in the context of economic espionage. Many large companies tend to overlook this threat. Most intrusion detection

systems have been primarily geared towards Internet-only intrusion detection.

Organized Structures

The fourth and last category concerns pirates of governments and terrorists or criminals of organizations. Their motivations are economic or ideological.

8.4 Classification of attacks

More and more the world is moving towards Cloud Computing, it becomes more sophisticated and thus increases the interest of attackers in finding new vulnerabilities. Cloud Computing faces a certain types of attacks. An attack is any action that compromises the security of information held by an organization or an individual.

There are different attacks on Cloud Computing, the most potential of which are discussed below:

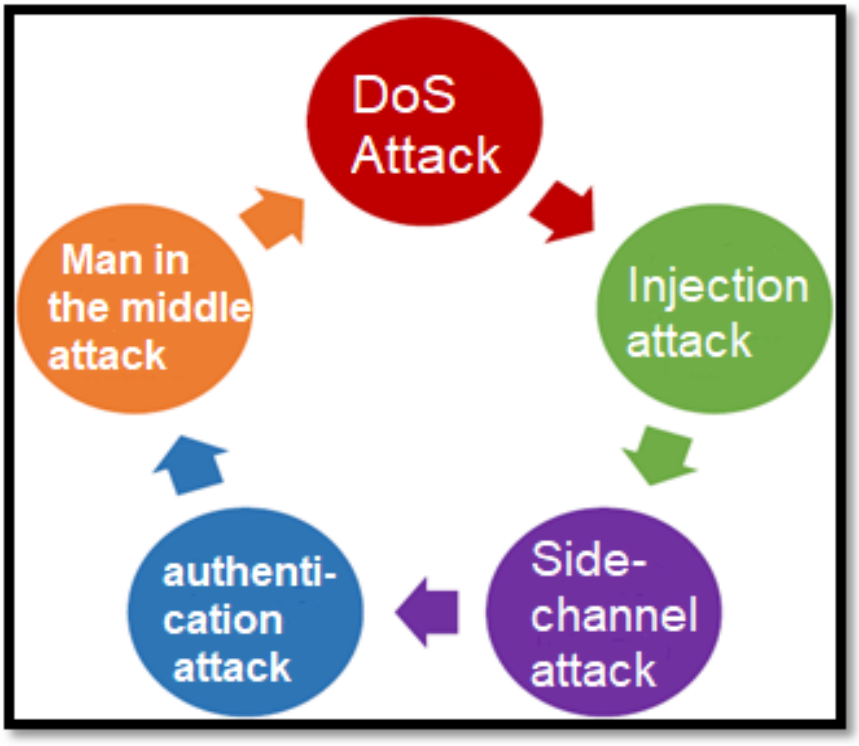


Figure 8.1: Types of attacks in Cloud Computing.

DoS Attacks

Principle

It is a type of attack aimed at making an organization's services or resources unavailable for an indefinite period of time. Most of these attacks are against the servers of an enterprise, so that they cannot be used and viewed. Denial of service attacks is a scourge that can affect any corporate server or anyone connected to the internet. The purpose of such an attack is not to recover or alter data, but to harm the reputation of companies with an Internet presence and possibly harm their operation if their activity is based on an information system. . From a technical point of view, these attacks are not very complicated, but they are no less effective against any type of machine with an operating system (Windows, Linux, Unix commercial) or any other system. Some Cloud Security Alliance has identified that the cloud is more vulnerable to denial of service attacks because it is used by many users which makes it much more damaging. The principle of denial of service attacks consists in sending IP packets or data of unusual size or constitution, in order to cause saturation or an unstable state of victim machines and thus prevent them from providing the network services they offer [22], [23].

There are usually two types of denial of service:

• Denial of service by saturation, consisting of submerging a request machine, so that it is no longer able to respond to real requests;

• Denials of service by exploiting vulnerabilities, consisting in exploiting a flaw in the remote system in order to make it unusable

Solution

To limit the DoS attack we can classify traffic on the basis of authorization, so that we can block traffic that identifies as unauthorized and allow traffic that is to be identified as authorized. For this firewall can be used to allow or deny traffic based on access protocols, ports or IP addresses. Today, most switches have a speed limit capability based on the access control list which can provide automatic limit rate form traffic, false IP filtering, binding and can deeply inspect the packets. As for switches and routers also have some capacity such as access control lists (ACLs), and speed limits which can be set manually to create rules. Request

endpoint equipment can be used on networks in colligation with routers and switches that can analyze data packets as they enter the network system to verify their authority and priority so that traffic flow can be controlled. The DoS attacker can send all attacked packet traffic to a null interface or to a non-existent interface, which helps reduce the effect of DoS attacks [24], [25].

Malware Injection Attacks

Principle

In Malware Attack, an attacker attempts to inject a malicious service or a virtual machine into the cloud. In this type of attack, the attacker creates his own malicious services implementation module (SaaS or PaaS) or a virtual machine instance (IaaS), and tries to add it to the Cloud system. Then, the attacker must behave in such a way as to make it a valid service to the Cloud system, which is a new instance of service implementation among the valid instances. If the attacker succeeds, the Cloud automatically redirects valid user requests to the implementation of malicious services, and the attacker's code begins to execute. The main scenario behind the Cloud injection attack is that an attacker transfers a malicious service instance to the cloud so that he can gain access to the victim's service requests. To do this, the attacker must gain control over the victim's data in the cloud.

According to the classification, this attack is the main representative of the exploit of the attack surface of the cloud service. The purpose of this attack may be something in which an attacker is interested; it can include data changes, changes to full functionality or reverse blockages [24], [26].

Solution

In Cloud Computing the system application managed by the customer is considered with great efficiency and integrity. So, to avoid the cloud injection attack, we can combine integrity with hardware or we can use hardware for integrity purposes because for an attacker, it is difficult to encroach on the IaaS level. For this we can use a file allocation table system (FAT). By using it, we can determine the validity and integrity of a new instance by comparing the current and previous instance. For this purpose, we need to deploy a hypervisor on the provider side. In the Cloud, the system hypervisor is considered to be the most secure and sophisticated part, the

security of which cannot be broken by any means. The hyperserver is responsible for planning all cases and the services we can do to check the file allocation table, to validate and integrate a client instance. Another approach is that we can maintain the platform version information of a client user to access the Cloud in the first phase when a client opens an account and can use this information to check the validity of the new instance of the client [27].

Side Channel Attacks

Principle

An attacker attempts to compromise the Cloud system by placing a malicious virtual machine near a target server system in the Cloud, and then launching a side-channel attack. The Side Channel Attack has emerged as a kind of effective security threat targeting the implementation of the cryptographic algorithm system. Evaluation of cryptographic systems resistant to side channel attacks is therefore important for the design of secure systems.
Side channel attacks use two stages:

- **VM CO- Residence & placement:** An attacker can often place his instance on the same physical machine as a target instance.

- **VM Extraction:** The ability of a malicious example to use side channels to learn information about co-resident cases.

It can be very easy to get secret information from a device so security against side channel attacks in Cloud Computing should be provided [24].

Solution

Virtual Firewall

The firewall is a collection of related programs that protect the resources of users on other networks and imposters. In this approach, the virtual firewall runs on the cloud server. It is possible to detect new malicious virtual machines in the Cloud Computing environment. With the help of the virtual firewall server these types of attacks can be prevented in cloud environments. The attacker attempts to place virtual machines in cloud environments. This

virtual firewall system blocks these types of new locations from malicious virtual machines [28].

Encryption & Decryption

Side channel attacks can be avoided in cloud computing environments by means of virtual firewalls. This can prevent side channel attacks in these environments. In order to provide greater security for the cloud, confidential data and information uses encryption and decryption. Client-side data is randomly encrypted using the concept of confusion and dissemination. Different security keys and different encryption algorithms are used to encrypt the data on the client side. Even if side channel attacks can occur it is difficult to decrypt client data. Which provides greater security in Cloud Computing environments [28], [29].

Authentication Attacks

Principle

These types of attacks can be easily produced in cloud environments. Attackers easily target servers with these types of authentication attacks. Attackers target the mechanism that is followed by the user. The mechanism used for authentication is captured and the attackers try to access confidential information. They use the various decryption mechanisms to transfer more confidential data. Service provider stores users' key value that must be authorized before accessing a service [30], [31].

Solution

This problem arises when using a simple authentication mechanism such as a simple username and password. Several authentication mechanisms must be established in environments to avoid these types of attacks [22], [31].

Man-In-The-Middle (Cryptographic Attacks)

Principle

In this attack the attacker intercepts messages while exchanging public keys, then retransmits them, substituting his own public key so that the two original parts still seem to be communicating with each other. The sender of the message does not recognize that the receiver is an unknown attacker by trying to access or modify the message before retransmitting to the

receiver. Thus, the attacker controls the entire communication [32].

In a man-in-the-middle attack (MITM) see figure (4.2), communication between two computers (here a personal computer on the left and a server on the right) is intercepted by a third part, here MITM. The computer and the server seem to be chatting together, but all messages actually go through MITM, which can read them and pretend to be one of them.

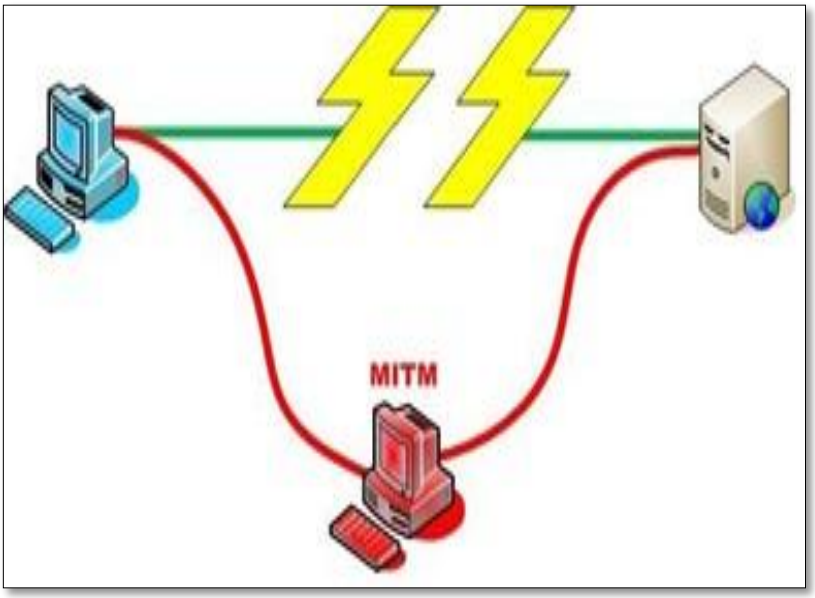


Figure 8.2: Man in The Middle attack.

Solution

This type of attack is avoided by the appropriate authentication mechanism. Encryption is used for the sender side and decryption is used for the receiver side. This device must be used. The attacker cannot modify the encrypted data. Several encryption and decryption algorithms are used such as AES, DES, triple DES, etc. [22].

Several solutions are used against MITM which are:

- Using a password because a one-time password is safe from MITM attacks.
- Mutual authentication, with many client and server implementations, initial trust is confirmed only by a path check between the client and the server. With mutual authentication, the server checks the client and the client checks the server to ensure

that legitimate communications are exchanged. Verification can be performed using public and private keys.

9. Data Security in Cloud Computing

9.1 Data Life Cycle in Cloud Computing

The life cycle of data in the Cloud can be broken down into five (5) main steps: Data transfer, data storage, use, recovery and destruction of data [33].

Data Transfer Phase

The phases related to sending data from a company's internal systems to the Cloud or repatriating it are the most mature. Where the data can be encrypted internally by the company and then sent, or else we use a transport layer integrating this encryption function. In this second category, the standard protocols which are IPSEC (Internet Protocol Security) and SSL (Secure Socket Layer) are very widespread. In connection with an authentication based on asymmetric keys (certificates with public key for example), these protocols make it possible to transmit data securely to or from the Cloud, thus the systems become reliable and easy to use.

Data Storage Phase

Once the data has arrived in the Cloud, it is stored. In the absence of recognized standards, the implementation of the encryption functions is dependent on the service provider. Some of them will offer systems whose operation may not always be very clear. In the event that data is stored in the cloud to ensure its availability, it is best to encrypt the data before sending it. This task will be performed by the Cloud client. Obviously, in the case of a SaaS (Software as a Service) type of encryption, it can only be carried out by the supplier, the end customer has an almost non-existent role in this encryption step.

Use of Data in the Cloud

A virtual machine (VM) deployed on an IaaS (Infrastructure as a Service) cloud. This VM uses a file system to store the operating system, applications and application data. Even if the file system is encrypted, the decryption keys must be present in the VM for it to work. An attacker could therefore, if he manages to recover these keys, access the data present on the VM disk. In this specific case, data security will be based on the access control measures put in place to access the data, both for external access and for access by administrators and cloud operators. Trusting your supplier is therefore more important when you give them VMs, same for an application positioned in the Cloud (web mail, CRM application, document management, etc.)

Data Recovery

It is essential to have the guarantee of having the means for recovering data in the event of problems other than cases of unavailability. The recovery must be able to be carried out under conditions of deadlines respecting the constraints expressed and the business needs. However, the dissemination of data must be carried out in a manner transparent to the user of the Cloud.

Data Destruction

Once data has been recovered from a cloud, it is important to ensure that it is destroyed. We should ask what are the commitments, means and procedures implemented by a supplier to erase all traces of our data. In this case, encryption can be used. In fact, without the key to decrypt them, previously encrypted data is completely unusable. So to destroy data, just destroy the encryption key. It is this concept which makes it possible to ensure that data is indeed inaccessible even in the extreme case where a cloud provider has just destroyed the key without warning.

9.2 Protective Measures

At each stage of the data life cycle, different measures can be implemented to ensure data security. There are two types of protection: Access control and encryption.

Access Control

Controlling access to data relies on authentication mechanisms. A person, a system or a program must be reliable in order to be able to access the data. All techniques, systems and means of access control are grouped under the acronym IAM (Identity and Access Management) [33].

Encryption

Thanks to specific software, an individual can "encrypt" their own documents. Their access is therefore limited since you must have the decryption key to be able to access and read them. Consequently, only the person who holds the key can access the documents [33].

9.3 Different solutions proposed

Cloud computing is a new model of IT service delivery using many existing technologies. However, we are seeing clouds appearing today specific to functional needs such as the archiving cloud or the workstation cloud, so for each of them security adapted to their functions is applied. We must also think of a security approach adapted to the Cloud service such as IaaS with strong infrastructure and identity security management or PaaS centered on the application and its data.

The considerations to be taken into account to adopt a security solution are therefore numerous and complex. What is certain is that the Cloud has made it possible to develop security in multiple areas such as the evolution of log management, identity federation and multiple login.

Security is often considered the main obstacle to the adoption of Cloud Computing services. This is how a lot of work has been devoted to finding solutions to remedy this problem.

In this section, we present the main research works that have proposed solutions to ensure data security in the clouds.

Security of access and storage of data in the Cloud

Jensen & al have listed the different techniques used in Cloud Computing to secure access to data and they have identified the shortcomings of these techniques to implement their solution

which is based on TLS (Transport Layer Security) protocol and XML cryptography (Extensible Markup Language) [34].

This solution responds to the problem of web browser which has security gaps. The idea proposed consists of using the TLS protocol and adapting the browser by integrating XML cryptography.

However Wang & al have proposed a solution based on the erasure-correcting code in order to allow redundancy and guarantee the reliability of the data [35]. They used the homomorphic token for storage accuracy and to locate errors. The proposed solution is capable of detecting data corruption during storage, it can guarantee the location of erroneous data and identify the server that has bad behavior [35].

Danwei and Yanjun have proposed a security algorithm that restores data if certain servers fail. It is a data separation algorithm.

This algorithm is only an extension of the fundamental theorem of K equation in algebra, Shamir's secret key sharing algorithm which is a cryptography algorithm based on sharing the secret (random number), the algorithm of Abhishek's online data storage [36]and number theory. The idea is to divide the data d into k parts of d = d1, d2, d3; up to dk. This sharing is done using the data separation algorithm to store it later on randomly chosen servers, denoted S = s1, s2, s3 until, S_m with m> k. The process of storing data in the clouds is therefore done in two stages, the first consists in dividing and storing the data on an arbitrarily chosen server and the second consists in being able to restore this data. Through these processes, data is ready to be transferred, stored, processed, securely since it is encrypted. The researchers deduced that the time complexity of the algorithm is the same for generating k data blocks and for restoring data. They have shown that even if an attacker invades a storage node, steals a block of data and tries to restore the data series, the time complexity necessary to perform the processing cannot be supported by today's IT environments. Among the other advantages that distinguish this proposal is the ability to restore data even if one or more storage nodes are not available which cannot be the case with a traditional cryptography solution.

Logical security of Cloud Computing

In IAAS (Infrastructure as a service) clouds, users have access to virtual machines (VMs) on which they can install and run their software. These virtual machines are created and managed by a virtual machine monitor (VMM) which is a software layer between the physical machine and the operating system. The VMM controls the resources of the physical machine and creates several virtual machines that share these resources.

Virtual machines have independent operating systems running independent applications and are isolated from each other by the VMM. This type of device has caused a lot of vulnerability problems in the virtual machine, which has prompted the authors to work in this area to find effective solutions. Zhou & al. have proposed a solution to eliminate the vulnerability of the virtual machine [37]. Discovering the limits of the XEN hypervisor used by AMAZON was their starting point. They proposed four approaches to improve the performance of this hypervisor which are based on Poisson's law, Bernoulli's law, the uniform law and finally the exact law. After a comparison between the four new models, they deduced that the Poisson law strategy is the best in practice to prevent cycle theft. Wei & al. [38] proposed a virtual machine image management system which controls the access and provenance of these images through filters and scanners which make it possible to detect and repair violations using data mining techniques, this system is called Mirage. S. Berger & al. have also developed a technology that responds to the problems encountered by the virtual machine. This technology is called Trusted Virtual Data Center (TVDc) [39], it ensures that workloads can only be billed to the customer who has benefited from the service. It also ensures in the case of certain malicious programs such as viruses that they cannot spread and it also helps to prevent problems of misconfiguration. TVDc uses the isolation policy which is based on the separation of material resources used by customers.

It manages the data center, access to virtual machines and the passage from one virtual machine to another.

Data security model offered in the Cloud

The layered approach is presented in Figure (5.1) where the first layer is responsible for user authentication. The second layer is responsible for the anonymization of data and the protection of user privacy and the third layer is responsible for data recovery and decryption [40], [41].

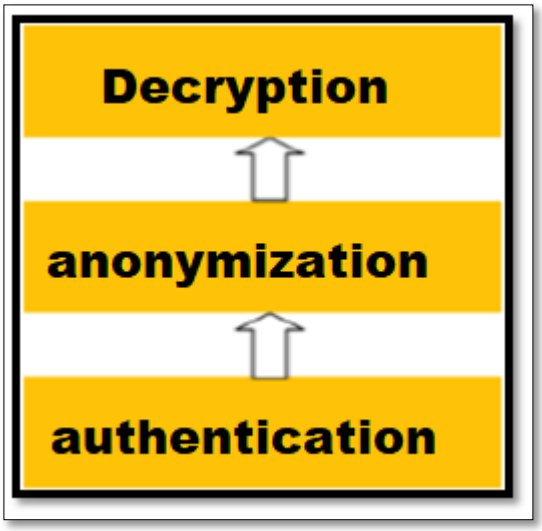


Figure 9.1: security model in the Cloud [40].

OTP Mechanism

One time password (OTP) mechanism or the one-time password is a password that is only valid for a session or a transaction. The use of multi-factor authentication with OTP reduces the risks associated with logging into the system from an unsecured workstation [42]. OTP is like a validation system that provides an extra layer of security for sensitive data and information by asking for a password that is only valid for a single connection. In addition, this password is no longer chosen by the user, but automatically generated by a pre-calculated method, which will eliminate certain gaps associated with static passwords such as gaps in password longevity, simplicity of password and brute force attack. OTPs are generated on the server side and sent to the user using a telecommunication channel. They are not susceptible to malicious users finding the username and password to access the resource. There's nothing you can do to get in the cloud without the right combination of username, password, and one-time password. In order to secure the system more efficiently, the WBS generated must be difficult to estimate, find, or trace by hackers. Therefore, it is very important to develop secure OTP generation algorithms [43]. Several elements can be used to generate a single-use password that is difficult

to guess [44], namely, International Mobile Equipment Identity (IMEI), International Mobile Subscriber Identity (IMSI), user name, PIN, minute, hour , etc.

Encryption Techniques

Current research is likely to use encrypted data with applications in the cloud. Indeed, today, the data must be in clear to be used by applications on the Cloud. If the data is encrypted, the result will also be encrypted (example of an SQL query which will return an encrypted result) [45]. The use of a key is not possible here, because otherwise it will be published on the Internet (Figure 5.2).

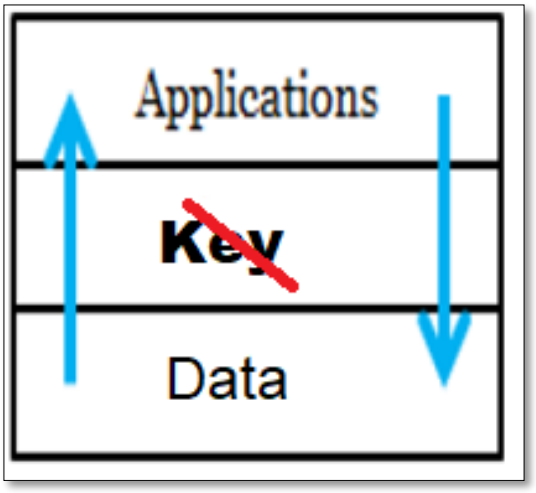


Figure 9.2: Keyless Encryption principale [45] .

You can store or archive encrypted data, but a cloud application cannot work on encrypted data. Research to remedy this problem has led to an attempt to find a solution to use encrypted data anyway (research by Craig Gentry). The method consists in implementing a fully homomorphic algorithm [35]. The latter works for addition operations. The most advanced in this research are the researchers from ENS Lyon, and part of their work will also consist in improving the complexity of the algorithm which is a real problem when the key becomes very large. This research could be the future of cloud security.

Approaches to data security in the Cloud

Privacy and data security are paramount when using cloud services. There are several works done in this area. Models, approaches and techniques are proposed to protect data.

In [46], Singh and Singh proposed a multi-level authentication system aimed at enhancing security in financial transactions. In [47], Satish and Anita proposed a false screen method to provide two-level authentication in Cloud Computing. In [48], Tandis et al. have proposed a method using the message authentication code in which the cryptographic key, the message and the hash function are concatenated together to provide authentication. In [49], Parsi and Sudha proposed a method using the RSA algorithm (Rivest Shamir and Adman) for authentication and secure data transfer. This method involves a key generation phase, encryption and decryption. In [50], the author proposed a technique for data security in the Cloud by the combination of different mechanisms, namely, multi-factor authentication by a single-use password and the authentication code of a message cryptographic fingerprint with a key. In [51], the concept of digital signature with the RSA algorithm was proposed to encrypt the data before transmitting it over the network. This technique solves the problem of authentication and security using anonymization techniques. In [52], Balasaraswathi and Manikandan proposed a multiple cloud architecture based on the partitioning of encrypted data with a dynamic approach in order to secure information in transit or stored.

We have analyzed several approaches to secure data transfer, these approaches mainly focus on authentication parameters. Indeed, data in transit to the Cloud can be attacked by different unauthorized interceptors. A particular method is not sufficient to deal with all questions of data security and confidentiality. Therefore, different techniques and integrated mechanisms should be used.

Other approaches

Today we see the appearance of other approaches that seeking to guarantee the security of the cloud resources more.

The considerations to be taken into account to adopt a security solution are therefore numerous and complex. What is certain is that the Cloud has made it possible to develop security in multiple areas such as the evolution of log management, "identity federation" and "multiple login".

10. Conclusion

Cloud computing is a very promising technology allowing its customers to reduce operating costs, administration etc. While increasing efficiency, however, the adoption of this technology remains low, and this comes back to security issues, in particular the security of data exchanged over the Internet.

The attacks affect the cloud computing environment. They lead to data loss and also financial loss for cloud owners, providers and users. Attacks should be avoided before they happen. For effective use of Cloud Computing, we must reduce and avoid attacks on vulnerabilities and improve security by applying the various possible solutions and appropriate mitigation techniques.

That's why in the next chapters we are going to propose an effective detection system based on deep learning approach with which we can secure our Cloud resources and maintain users privacy.

CHAPTRE II

Deep Learning

11. Introduction

Many of us are unable to distinguish between artificial intelligence (AI), machine learning (ML) and deep learning (DL), but if the three terms are often used interchangeably, they do not refer to the same things.

Here is an image that tries to visualize the distinction or the relationship between them:

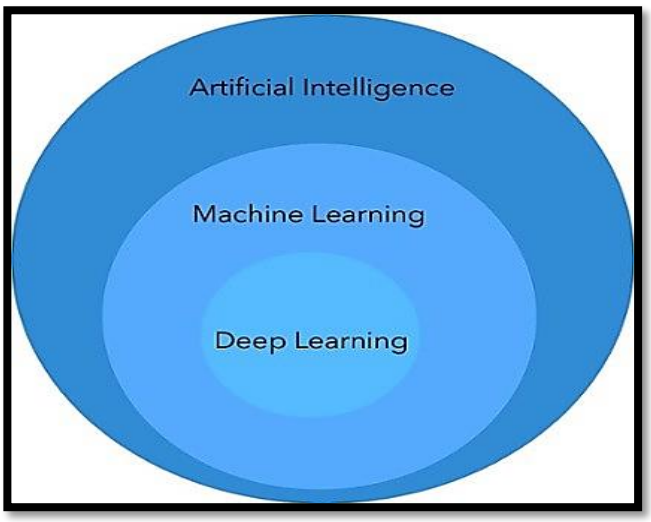


Figure11.0.1: Relationship between AI & ML & Deep Learning.

Machine learning only focuses on solving concrete problems, it also takes some ideas of artificial intelligence. Machine learning goes through neural networks designs to mimic human decision-making capabilities. ML tools and techniques are the two narrow subsets focused only on deep learning. We must apply it to solve any problem that requires thinking.

Any deep neural network will have three types of layers:

- The input layer
- The hidden layer
- The output layer

We can say that deep learning is the most recent term in the field of machine learning. It's a way to implement Machine Learning.

We use a machine algorithm to analyze data, learn from it, and make informed decisions based on what we have learned. Basically, deep learning is used in layers to create an "artificial neural network" capable of learning and making intelligent decisions on its own.

We can say that deep learning is a sub-domain of machine learning.

Different neural network architectures are used in DL, each with its own advantages and disadvantages. The convolutional neural networks or CNN for "Convolutional Neural Network" are an extension of MLP allowing to respond effectively to the main defects of MLP. They are designed to automatically extract the characteristics of images and also input text (1D), are invariant to slight distortions of the image, and implement the concept of weight sharing making it possible to considerably reduce the number of network parameters.

12. Machine Learning

"Machine Learning is the Field of study that gives computers the ability to learn without being explicitly programmed" L. Samuel, 1959(IBM)

Machine learning is a way to analyze data. It is also seen as a branch of artificial intelligence built on the idea that systems can learn from data, determine patterns and make decisions autonomously, with minimal human intervention.

Because of new computer technologies, machine learning is not comparable to machine learning of the past. It was born out of pattern recognition and the theory that computers can learn without being programmed to perform specific tasks. Researchers interested in artificial intelligence wanted to know if computers could learn from data. The iterative aspect of machine learning is important because, as models are exposed to new data, they can adapt independently. They learn from previous calculations to produce reliable and reproducible decisions and results. This is a science that is not new, but has gained new momentum.

12.1 Classification

Classification is a method of dividing a large group of data into a number of desired and distinct categories, in which a classification can be assigned to each category.

In other words, the classification belongs to the category of supervised learning, which has the advantage of providing the objectives of the input data. Classification applications are

presented in many areas, including credit accreditation, medical diagnosis, targeted marketing, etc.

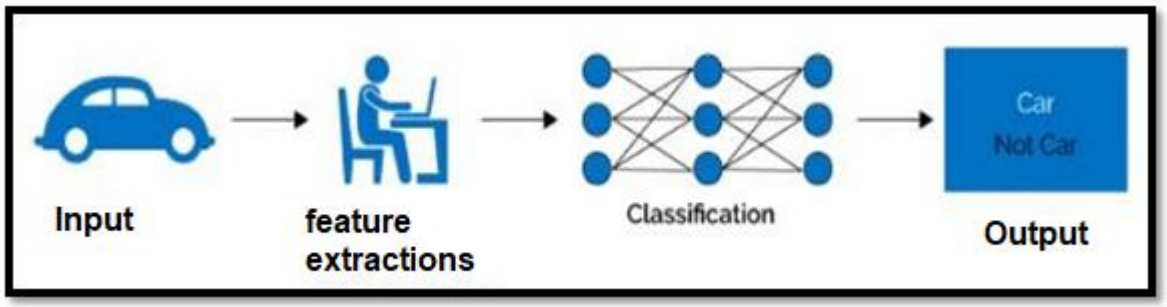


Figure12.1: classification with machine Learning Example.

As the name suggests, classification is obtained by classifying objects into subcategories from a general group, but can the machine distinguish people or animals using images, or does it distinguish and classify texts according to their content? Yes, the machine can perform all classifications using the appropriate data and algorithms, and give accurate results in no time. So it is an example of pattern recognition.

But in machine learning, classification gives a problem of identifying a new observation to which group of categories (sub-populations) belongs, based on a set of training data that contains observations and whose categories are known.

12.1.1 Types of classification

Two types of classifications :

Binary Classification: This classification is followed if we have to divide and classify the available data into two different categories (identifying the group to which each belongs) and the basis of the classification rule. Includes contexts that require deciding whether the item has a specific property or a typical binary classification, for example, if a person is sick and one wishes to diagnose the type of disease, "x" if is infected or not.

Multi-Class Classification: In this case, the number of classes should be more than two, depending on the wide range of data available. In machine learning, classification of multiple or multi-border classes is a classification problem in one or more of the three categories, despite the existence of classification algorithms allowing the use of more than two classes .

Error function and optimization of hyperparameters

Hyperparameters are essential in machine learning because they are often generated models with very different performance.

Tuning machine learning hyperparameters is a tedious task, but crucial, because the performance of an algorithm can strongly depend on the choice of hyperparameters, it is performed using automated methods which aim to find less optimal hyperparameters time, using an informed search, without any manual effort beyond the initial configuration [53].

12.1.2 Bayesian Optimization methods

In short, Bayesian optimization finds the value which minimizes an objective function by constructing a substitution function (probability model) on the basis of the results of previous evaluation of the objective. Optimizing substitution is less expensive than optimizing the goal. The concept is as follows: limit costly evaluations of the objective function by choosing the following input values based on those that have worked well in the past.

In the case of optimizing hyperparameters, the objective function is the validation error of a machine learning model using a set of hyperparameters. The goal is to find the hyperparameters that generate the lowest error on the validation set, in the hope that these results generalize to the test set. The evaluation of the objective function is expensive, because it requires learning the machine learning model with a specific set of hyperparameters. The Bayesian hyperparameter setting uses a constantly updated probability model to "focus" on promising hyperparameters based on previous results [53].

12.2 Stochastic gradient descent

The stochastic gradient descent (often abbreviated SGD), also called incremental gradient descent, is known as an iterative method to optimize an evolutionary function, a stochastic approximation of the optimization of the gradient descent. It is called stochastic due to the random selection of samples [54].

Since an iteration of the gradient descent algorithm requires prediction for each instance of the training dataset, managing millions of instances can be time consuming. In situations where you have large amounts of data, you can use a gradient descent variation called stochastic gradient descent. In this variant, the gradient descent procedure described above is executed, but the updating of the coefficients is carried out for each learning instance, rather than at the end of the batch of instances. Learning can be much faster with stochastic gradient descent for very large training data sets, and often you only need a small number of passes through the data set to reach a set satisfactory or sufficient coefficient [55].

13. Deep Learning

13.1 History of Deep Learning

- 1943 – The first mathematical model of a neural network

  In early 1943, Walter Pitts and Warren McCullochont created a computer model based on the neural networks of the human brain, using a set of mathematical-based algorithms to mimic human thought. This model was called "threshold logic". Work on "deep learning" continued by relating to a fluctuating phase of artificial intelligence

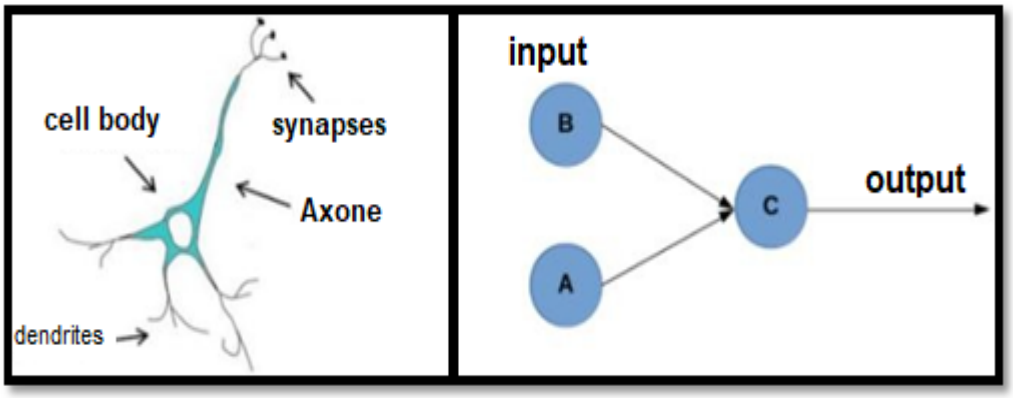


Figure 13.1: On the left the diagram of a biological neuron and on the right the diagram of the formal neuron from.

- 1950 - The prediction of machine learning.

Turing, a mathematician who broke the codes during World War II, predicted the development of machine learning in 1947. In 1950, he invented a similar machine and always spoke about "computers and intelligence" and developed a "Turing test" to determine if the computer could "think".

- 1952 – First machine Learning programs.

Arthur Samuel continued to create the first computer learning programs. And started to show the learning features through the game of checkers, which allowed the computer to correct its errors and find better ways to win. It was one of the first examples of machine learning.

- 1957 – Laying the foundation for deep neural networks

In 1957, at Cornell Aviation Laboratory, Rosenblatt wrote an article entitled "Perspective perception and recognition automaton" in which he declared that he would create a system capable of learning to recognize similarities or identities between visual information models, electrical or tonal in the same way to cognitive processes in the brain. This idea laid the groundwork for bottom-up learning and has been widely disseminated as the basis of deep neural networks (DNN).

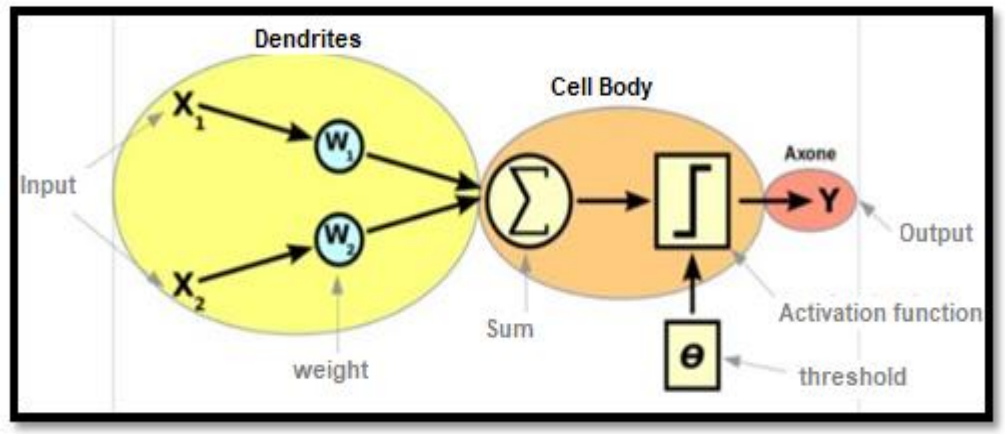


Figure 13.2: Deep Learning schema.

- 1959 – Discovery of simple & complex cells

In 1959, David Hubble and Torsten Wiesel discovered simple and complex cells. These biological activities have been a positive factor in the area of deep learning, with many artificial neural networks (ANN) being inspired by biological observations.

- 1960 – Control Theory

In 1960 Henry J. Kelley was recognized for developing the foundations of a continuous backward propagation model. Two years later, more precisely in 1962, Stuart Dreyfus developed a simple version based only on the chain rule, but the concept of back propagation, supposed to extend beyond errors for training purposes, remained ineffective until 1985.

- 1965 –The first deep learning networks.

Efforts continued, notably those of Alexey Grigoryevich Ivakhnenko, who worked on deep learning algorithms and developed a data processing method. Valentin Grigor'evich lapa developed cybernetics and forensic techniques, all dating from 1965. These models were used with multidisciplinary activation functions (complex equations), then analyzed statistically.

During the 1970s, a group of people continued their research in this area, but without any funding.

- 1979-80 - ANN learns to recognize visual patterns

Thanks to Fukushima, who created a Neocognitron neural cell network, capable of learning and identifying visual patterns and of being used in the recognition of handwriting and other patterns, and even in the processing of natural languages. His work led him to develop the first transformational neural networks, based on the regulation of the visual cortex in animals.

- 1982 - Creation of Hopfield networks

Hopfield Networks is a frequent neural network that functions as a recallable memory system and remains a common implementation tool for deep learning in this century.

- 1985 – Program learns to pronounce English words

In 1985, Terry Segnovsky created the program "NETtalk" which allows the pronunciation of English words and manages to improve with time.

- 1986 - Improved pattern recognition and word prediction

In 1986 Romlhart, Hinton and Williams published an article entitled "Representations of Retrograde Error Learning", which showed how existing neural networks could be greatly improved for many tasks such as pattern recognition, word prediction, etc. He was considered the spiritual father of deep learning.

- 1989 - Q-Learning and machines read handwritten numbers

In 1970, it was at this time that Seppo Linnainmaa wrote a master's thesis including the FORTRAN code for posterior propagation. This concept was not implemented until 1985. Rumelhart, Williams and Hinton demonstrated that repression of the neural network could provide "interesting" representations of the distribution, which raised the question of whether human understanding was based on symbolic logic or distributed representation.

In 1989, Yann LeCun identified the first practical presentation of reverse depression, combining neurotransmitters and the spread of “handwritten” figures. This system made it possible to read many handwritten checks 1993 - a very deep learning task is solved.

In 1993, the German computer scientist Schmidopfer solved the task of "deep learning" which required more than 1000 layers in the repeated neural network.

- 1995 – Support vector machines

SVM is essentially a system for identifying and planning similar data. It can be used to classify texts, recognize handwritten characters and classify images for further learning.

- 1997 – Long-term memory proposal.
- 1998 – Gradient Learning.
- 2009 – Launch of Image Net

ImageNet was launched in 2009 by "Fei Fei Lee", professor and director of the artificial intelligence laboratory at Stanford University.

- 2011 – Creation of AlexNet

Its success began to revive the neural network in the deep learning community.

- 2012 – Cat experience

"The cat experience" was a big step forward. A neural network spread over thousands of computers was used, and 1,000,000 non-pending images - randomly captured on YouTube - were detected and analyzed. They found that the program had learned to recognize and recognize cats and that the proportion of learning attempts had improved considerably.

- 2014 - DeepFace

It was developed in 2014, the deep social media system - DeepFace - uses neural networks to identify faces with an accuracy of 97.35%.

- 2014 – Generative opponent networks

It was introduced by a team of researchers led by Ian Goodfellow in 2014. The Generative opponent network uses two competing networks: the first takes the data and attempts to create non-characteristic samples, the second receives the data and the samples generated and must determine whether each data point is correct or created.

- 2016 – Powerful Machine Learning products

Some companies, such as Cray Inc, have been able to provide powerful machine and deep learning products and solutions. An example is the use of Microsoft's neural network software on supercomputers equipped with graphics processors to perform deep data learning tasks in a very short time.

In 2016, the AlphaGo program, developed by Google DeepMind, caused a sensation by beating the game of go by a score of 4 to 1 the South Korean Lee Sedol, who is considered the best player in the world.

13.2 Definition of Deep Learning

With deep learning, computational models with multiple processing layers can recognize most representations of data at multiple levels of abstraction.

Deep Learning is based on:

- Deep neural networks
- Specific training algorithm
- Learning more abstract representations
- Larger datasets

Deep Learning is a type of machine learning that aims to train and teach the computer to perform human functions such as distinguishing visual objects and identifying sounds and pictures. Instead of organizing data, deep learning sets its own basic parameters, which allow the machine to learn on its own, and the current interest in deep learning is partly due to buzz surrounding artificial intelligence. Deep learning techniques have improved the ability to classify, recognize, detect, etc.

Deep Learning discovers the complex structure in large data sets using the back propagmentation algorithm to indicate how a machine should modify its internal parameters used to calculate the representation of each layer from the representation of the previous layer.

Deep convolutional networks have enabled advances in image, video, speech and audio processing, while recurrent networks have illuminated sequential data such as text and speech [56].

13.3 How Deep learning works

Deep learning changes the way we think about the problems we solve with analysis. It's going to tell the computer how to solve a problem and then train it to solve the problem itself.

A classic approach to analysis consists in using the available data to design entities in order to deduce new variables, then select an analytical model and finally estimate the parameters (or the unknowns) of this model. These techniques can give predictive systems that do not generalize well, since their completeness and accuracy depend on the quality of the model and its characteristics. Adding more data means starting over. The new approach based on deep learning consists in replacing the formulation and the specification of the model by hierarchical characterizations (or layers) which learn to recognize the latent characteristics of the data starting from the regularities of the layers. The paradigm shift with deep learning goes from functional engineering to functional representation.

To better understand how to do this, let's take a simple example: we want to identify a fish among a large group of fish that are filmed, regardless of how they are photographed.

To this end, a large collection of different images of fish and images of other objects must be assembled in order to train the artificial neural network for the practice of deep learning. These images are converted into data and transmitted to the network after the weight is assigned to different elements by artificial nerve cells. The information is then collected by the last layer of neurons to determine whether it is a fish or not.
Then, the artificial neural network will compare this result to the response provided. If there is a match, the network will record this success because it will use it in other parts. If the result is negative, the network will recognize the imbalance and adjust the weight of the different neurons to correct its error. The process is done thousands of times until the image is recognized, this type of learning is called "supervised learning".

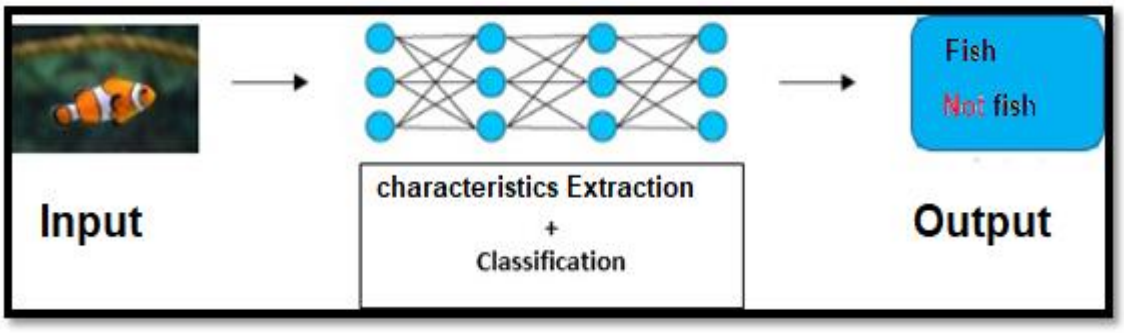


Figure 13.3: Example on how deep learning works.

The second method is unsupervised learning, which relies on unclassified data, the network must identify patterns in the datasets to learn for themselves.

14. Fields of applications of deep learning

Deep learning has been developed to a large extent, sweeping across different ICT sectors, for example, visual recognition of the traffic sign by a robot or an independent vehicle and also image comprehension with deep convolutional networks : Since the early 2000s <<ConvNet>> has successfully applied to the detection, segmentation and recognition of objects and regions in images. New classes of neural networks have been developed and are ideal for applications such as text translation and image classification.

Today, companies use many practical applications in the field of deep learning, including:

- Speech recognition: Take for example Xbox, Skype, Google Now, and Apple's Siri, which uses deep learning techniques in their systems to recognize human speech patterns and sounds.
- Image recognition: The practical application of image recognition is machine translation of images and description of scenes. This application is very effective in the investigation of criminal activity among thousands of images provided and cars can benefit from it through image identification using 360-degree camera technology.
- Natural language processing: Neural networks have for many years been an essential

component of deep learning in the processing and analysis of written texts. It can also be used to discover and analyze doctor's notes or reports.

- Recommendation systems: Both Amazon and Netflix have successfully integrated the concept of the recommendation system with the ability to recognize your level of interest next time based on past behavior. It is also used in other fields such as music and other fields.

At the same time, human-machine interfaces (HMI) have evolved a lot. Including the mouse and keyboard where they have been replaced by gestures…, sparking renewed interest in AI and deep learning.

15. Architectures of deep neural networks

There are a large number of deep architectural variables. But we will mention the most important / standard.

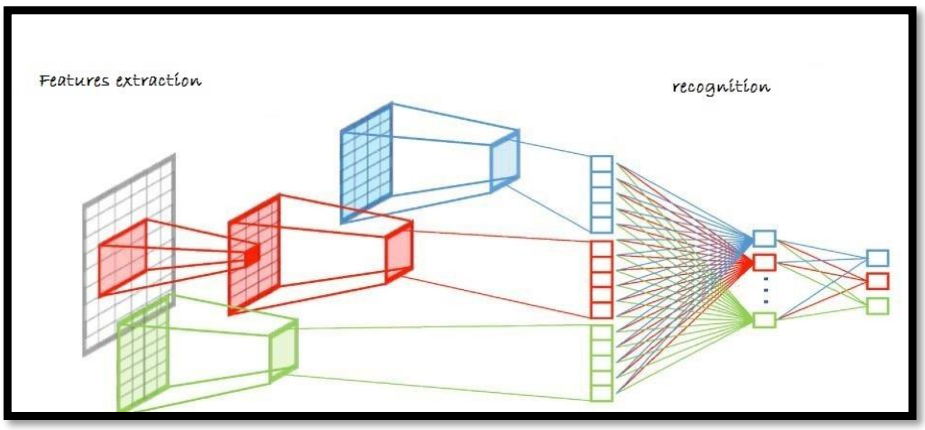


Figure 15.1: Standard Architecture of a deep neural network. [53] [54]

15.1 Convolutional neural networks (CNN)

The convolutional neural network (CNN) is a type of artificial neural network, it is widely used in graphics applications such as image processing and video recognition, as well as in

recommendation systems. In particular, it is widely used in the field of image analysis. CNN has the advantage of using kernel filters and implementing pooling.

15.1.1 The convolution operation

In mathematics, a convergence product is originally an alternative product and a binary operator mixed at the same time and whose symbol is "*"

The mathematical representation of the concept of linear filter is the generalization of the idea of average regression by the product of convolution, which makes it valid for all temporal data such as signal processing and spatial data in the field of data processing.

The convolution process: So what exactly is the convolutional neural network? According to Chris Olah, research scientist at Google Brain:

«Basically, convolutional neural networks can be thought of as a type of neural network using multiple identical copies of the same neuron. This allows the network to have a large number of neurons and to express large computer models while keeping the number of real parameters (the values describing the behavior of the neurons) that need to be learned relatively small." [57]

As mentioned earlier, convolution is a very useful mathematical tool in the field of image processing, which explains the use of convolutional neural networks in the field.

Convolution acts like filtering. We define a window size that will wrap around the entire image (remember that an image can be seen as a table). At the very beginning of the convolution, the window will be positioned at the top left of the image then it will shift by a certain number of boxes (this is called the step) to the right and when it reaches the end of the frame, it will shift down one step, and so on until the filter has gone through the entire frame: [58]

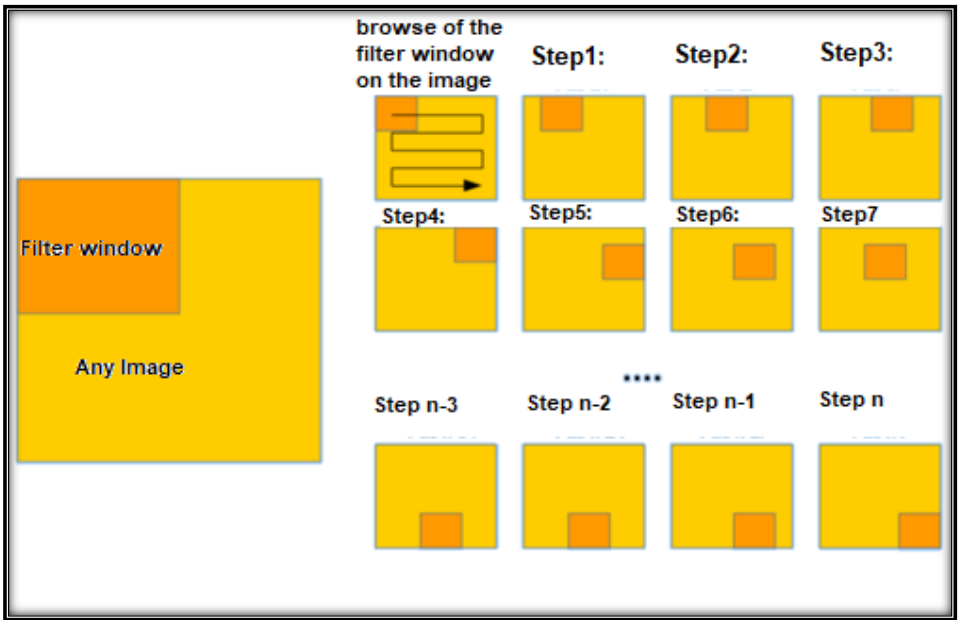


Figure 15.2: Diagram of the route of the filter window on the image. [58]

The goal is to use the values present in the filter at each step. For example if we define a window 3 by 3, this will represent 9 boxes of the table (that is to say 9 pixels). The convolution will perform an operation with these 9 pixels. It can be any operation, for example we extract the largest value (i.e. the pixel with the largest value).

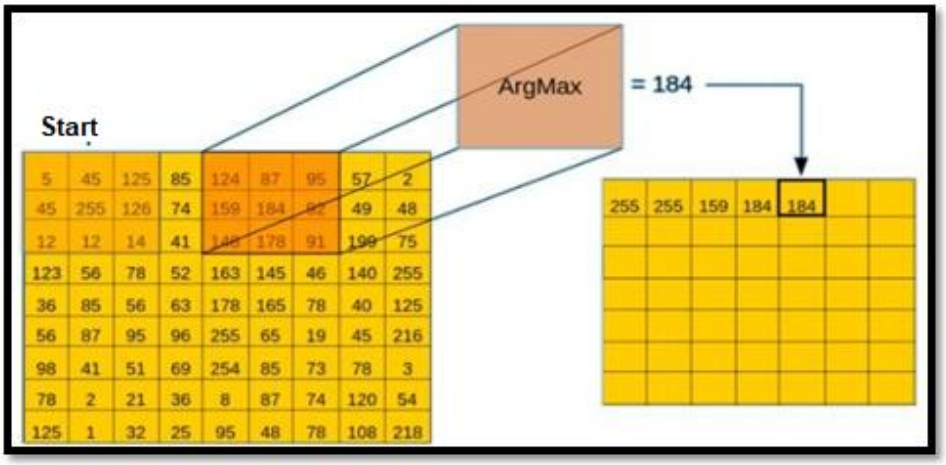


Figure 15.3: Example of a convolution whose configuration is: Operation = Maximum argument, horizontal step = 1 pixel, vertical step = 1 pixel. [58]

We drag the window in orange and at each step we recover the largest value among the 9 pixel values. [58] The output of this operation (convolution) which is called "feature map" at dimensions is smaller than that of the input image.

15.1.2 CNN layers

The structure of CNN consists of a set of independent processing layers:

- The convolution layer (CONV) competent to process data from the receiver field.
- The pooling layer (POOL), which has the characteristic of reducing its size, allowing information to be compressed, sometimes by sampling.
- The correction layer (ReLU) is defined by its name.
- The "fully connected" check mark (FC), which is a perception type layer.
- The loss layer (LOSS).

Convolution layer

This layer is the main component of convolutional neural networks and constitutes their first layer. Its function is to locate a set of features in the images that have been given as inputs, which is why we perform a process called meta-filtering, in which we drag the image function into a form of window and calculate the convolution product between the feature and each portion of the scanned image. This process produces an image function visualized as a filter.

To measure the volume of the convolutional layer, we need the following three hyperparameters:

– **Depth of the layer**: The number of convolution nuclei or the number of neurons which have a direct correlation with the same field of the receptor.

– **The step**: The step controls the overlap of the receptive fields. The smaller the pitch, the more the receptive fields overlap and the greater the output volume will be.

– **The margin (at 0) or zero padding**: sometimes it is convenient to put zeros at the border of the input volume. The size of this zero-padding is the third hyperparameter. This margin is used to control the spatial dimension of the output volume [59].

Pooling layer (POOL)

The most common thing about this layer is that it is placed between two convolution layers, which allows it to receive many feature maps as input, as it applies to every pooling process.

The pooling (or sub sampling) operation consists of minimizing the size of the image and preserving its important properties.

That's why we divide the image into regular cells, while keeping the maximum value in each cell. The cells used are square and small to avoid losing a lot of information. The most common choices are adjacent cells of size 2 × 2 pixels which do not overlap, or cells of size 3 × 3 pixels, spaced from each other by a step of 2 pixels (which therefore overlap). The output is the same number of feature maps as the input, but they are much smaller. The pooling layer reduces the number of parameters and calculations in the grid (network). [59] This increases the efficiency of the network and limits over-learning. And therefore to also control the over-adjustment. The pooling layer operates independently on each input depth slice and resizes it spatially, using the MAX operation.

So the pooling layer makes the network less sensitive to the position of features, which is slightly higher or lower and may have a slightly different trend, but should not cause any change in the classification of images.

Among the known characteristics of pooling is to give great computing power, and despite the very low representation (and therefore the loss of associated information), all current trends are to use small filters (type 2 × 2). But the risk of over-learning posed many problems which made it possible to avoid the pooling layer.

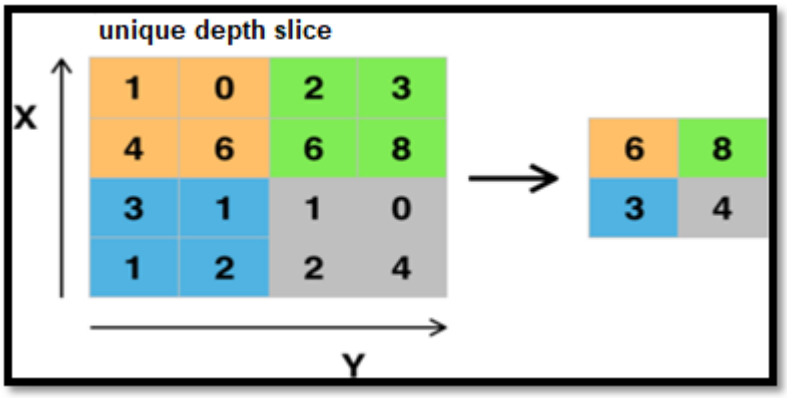


Figure 15.4: Pooling with a 2x2 filter & a step of 2. [54]

Correction layer (ReLU)

To improve the efficiency of the processing by inserting between the processing layers a layer which will operate a mathematical function (activation function) on the output signals. In this framework we find:

ReLU (Rectified Linear Units) designates the real non-linear function, this function also called "non-saturating activation function", increases the non-linear properties of the decision function and of the whole network without affecting the receptive fields of the convolution layer defined by ReLU (x) = max (0, x).

The ReLU correction layer therefore replaces all the negative values received as inputs with zeros. It plays the role of activation function.

- Often the proofread correction is preferable, but there are other forms: $f(x) = \tanh(x)$,
- Correction by hyperbolic tangent :
- Correction by the saturating hyperbolic tangent : $f(x) = |\tanh(x)|$,
- Correction by the sigmoid function:$(1 + e^{-x})^{-1}$. [14]

Fully connected layer

After several layers of convolution and max-pooling, it is necessary to think high level in the neural network which is formed of layers, the latter being fully connected. The fully connected layer contains the neurons which in turn contain links to all of the outputs of the previous layer and are managed in

regular neural networks, their activation function can be easily calculated using a multiplication matrix followed by an increase in polarization.

This type of layer receives an input vector and produces a new output vector. To do this, it applies a linear combination and possibly an activation function to the values received as input.

Loss layer

The loss layer specifies how network dragging penalizes the gap between the expected and actual signal. It is normally the last layer in the network. Various loss functions suitable for different tasks can be used there. The Softmax is used to predict a single class among K mutually exclusive classes [54].

15.1.3 CNN parameters

Hyperparameters are the most common use of neural networks, when compared to standard MLP. It is necessary to take into account the concepts notions of number of filters, their shape and the shape of max pooling, although the usual rules for learning rates and regularization constants are still valid.

5.1.3.1 Number of filtres

Due to the reduction in the size of the intermediate images as a function of the processing depth, the number of filters in the layers close to the input is reduced, while the reverse is found in the layers closest to the output where they are larger. This layer is selected to be almost constant across all layers. It is very important to keep the number of intermediate outputs (i.e., multiply the number of intermediate images by the number of pixel positions) in order to preserve the input information. In order to increase (in the broad sense) from one layer to another. The number of intermediate images directly controls the power of the system, depends on the number of examples available and the complexity of the processing.

5.1.3.2 Shape of filters

Forms of filtering are often chosen based on the dataset, due to the great diversity of their forms in the literature. The best results are those obtained on MNIST (28 × 28) in a range of 5 × 5 in the first layer, while natural images (often with hundreds of pixels in each dimension) tend to use first-class filters 12 × 12 or even 15 × 15. As for creating abstractions in the appropriate range and adapting them to each case, we have to find the appropriate level of granularity for that

5.1.3.3 Shape of Max Pooling

Typical values are 2 × 2. Very large input volumes may justify pooling 4 × 4 in the lower layers. However, choosing larger shapes will greatly reduce the size of the signal and may result in excessive loss of information. Often times, non-overlapping pooling windows give the best results.

5.1.1 Regularization method

One of the central problems in machine learning is how to create an algorithm that will work well not only on training data, but also on new entries. Many strategies used in machine learning are explicitly designed to reduce the test error, possibly at the expense of an over-learning error. These strategies are known collectively as regularization. A large number of forms of regulation are available to the discouraging practitioner.

Regularization has been used for decades before deep learning.

5.1.4.1 Empirical

5.1.4.1.1 Dropout

The dropout method consists in “deactivating” the outputs of neurons randomly (with a predefined probability, for example 0.5 for the hidden layers and 0.8 for the input layer) during the learning phase. This is equivalent to simulating a set of different models (bagging) and learning them jointly (although none is learned from start to finish). During the learning iteration, it is possible that all the neurons are inactive, so that each unit "learns well" independently of the others and avoids “Co-adaptation”, because dropout can lead to accelerated learning.

The dropout technique is used in particular in image recognition systems, voice recognition, document classification and on calculation problems in biology [60].

5.1.4.1.2 DropConnect

DropConnect is an alternative to dropout consisting in inhibiting a connection (the equivalent of the synapse), and this always in a random manner. The results are similar (speed, ability to generalize learning) to dropout. A layer “completely connected” with a DropConnect can be compared to a layer with a “diffuse” connection [61].
When training with Dropout, a randomly selected subset of activations are zeroed in each layer. Instead, DropConnect sets a randomly selected subset of weights in the network to zero. Each unit thus receives an input from a random subset of units from the previous layer.

5.1.4.1.3 Stochastic pooling

The same principle that we find in Max-pooling is also taken by stochastic pooling, but the outputs will be chosen randomly according to the multinomial distribution which is also determined according to the activity of the zone that the pool addressed. In fact, this system is similar to the creation of Max-pooling with a large number of similar images, which vary only by localized deformations. This method can also be considered as an adaptation to the elastic deformations of the image.

This is why this method is very effective on MNIST images (database of images representing handwritten numbers). The strength of stochastic pooling is to see its performance increase exponentially with the number of layers in the network.

5.1.4.2 Explicit

5.1.4.2.1 Network size

The convergence of processing is one of the most important problems of what is called overfitting, and the simplest solution is to reduce the number of network layers and the free parameters (connections) of the network, which at least directly reduces the energy and predictive potential of the network. It is equivalent to having a "zero standard".

5.1.4.2.2 Weight degradation

A simple form of added regularizer is the weight decrease, which simply adds an additional error, proportional to the sum of the weights (standard L1) or the magnitude squared (standard L2) of the weight vector, to the error at each node. The level of acceptable model complexity can be reduced by increasing the proportionality constant, thereby increasing the penalty for large vectors.

- **Regularization of L1**: The L1 regularization leads to the scarcity of weight vectors during optimization. In other words, neurons with L1 regularization only use a small subset of their most important inputs and become almost invariant with respect to noisy inputs.
- **Regularization of L2**: The L2 regularization intuitively interprets the very disadvantageous peak weight vectors and favors the diffuse weight vectors. Due to the multiplicative interactions between weights and inputs, this has the useful property of encouraging the network to use all of its inputs somewhat, rather than some of its inputs.

15.2 Recurrent neural network (RNN)

During the 90s, many researchers focused on recurrent neural networks, as they were an important part of neural network research during this period, they were applied to various problems including chronological sequence events and ordered data.

Recurrent neural networks are based on the work of David Rumelhart in 1986. The simplest version of RNNs proposed by Jeff Elman in 1990 and in which links are added to an MLP to give as input a layer of it's own network output at the previous time step in addition to the current output of the previous layer. In 1993, a system of neural history compression solved a "deep learning" task that required more than 1,000 subsequent layers in a time unfolded RNN.

15.2.1 What is RNN ?

A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a sequence. This allows it to exhibit temporal dynamic behavior for a temporal sequence. Unlike anticipatory neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as handwriting recognition or connected, non-segmented speech recognition.

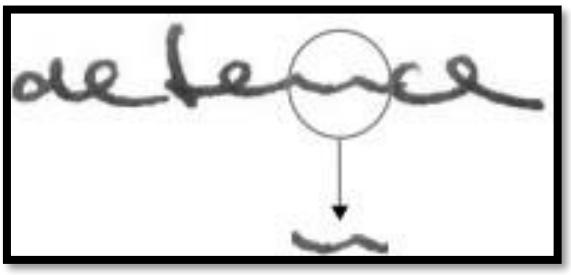


Figure 15.5: Importance of context in the recognition of handwriting. [62]

The term "recurrent neural network" is used interchangeably to denote two major classes of networks having a similar general structure, where one is a finite momentum and the other infinite momentum. Both classes of networks exhibit dynamic temporal behavior [63].

A finite-pulse recurrent network is a directed acyclic graph which can be unwound and replaced by a strictly anticipatory neural network, while an infinite-pulse recurrent network is a directed cyclic graph which cannot be unwound.

Finite pulse networks and infinite pulse recurrent networks can have an additional stored state and the storage can be under the direct control of the neural network. The storage can also be replaced with another network or graphics, if it incorporates delays or feedback loops. Such controlled states are called stuck state or stuck memory and are part of short term memory arrays (LSTMs) and triggered recurring units [64].

RNNs have had great success in labeling and sequence prediction tasks, such as handwriting recognition and language modeling. In acoustic modeling for speech recognition, however, where deep neural networks (DNN) are the established state of the art, RNNs have recently received little attention beyond small telephone recognition tasks scale, the notable exceptions being the work of Robinson.

A recurrent neural network and a ConvNet network work together to recognize an image and give a description of it if there is no name. This combination of neural network works beautifully and produces fascinating results. Here is a visual description of the procedure. The combined model even aligns the generated words with the characteristics of the images [65].

15.2.2 RNN application

Recurrent neural networks play an active role in the field of speech recognition, as used by the Google Translation Network. It has already been applied in the field of machine translation and robot control, as well as in predicting time series. It has been successful in rhythm learning, musical composition and grammar learning, detecting protein parity, and predicting cellular sub-cellularization and predictability in the healthcare industry, as well as numerous forecasting tasks in the field of business process management, among others.

15.3 Reinforcement learning

Reinforcement learning is an area of machine learning Enhancement. It is about taking appropriate steps to maximize the benefits in a given situation. It is used by various software and machines to find the best possible behavior or the best path to follow in a given situation. Reinforcement learning differs from supervised learning in that, in supervised learning, training data has the key to the answer; the model is therefore formed with the correct answer while in reinforcement learning there is no answer, but the reinforcement agent decides what to do to perform the given task. In the absence of a training dataset, he is required to learn from his experience.

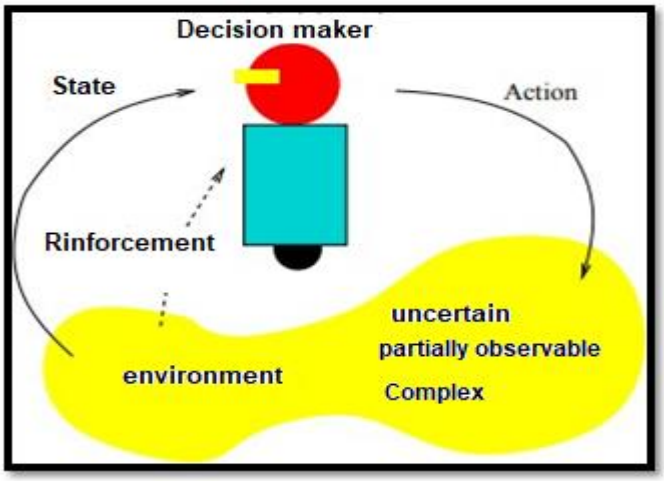


Figure 15.6: example of an agent using reinforcement learning.

Reinforcement learning is often used for robotics, games and navigation. With reinforcement learning, the algorithm discovers through trial and error which actions generate the greatest rewards. This type of learning has three main components: the agent (the learner or the decision maker), the environment (all interactions with the agent) and actions (what the agent can do). The goal is for the agent to choose actions that maximize the expected reward over a period of time. The agent will reach his goal much faster by following a good policy. Reinforcement learning therefore aims to learn the best policy.

16. Conclusion

In this chapter, we have talked about machine learning and tried to clarify the concepts that interest us in this area, including the classification and implications of the relationship between machine learning and deep learning. We have mentioned how it has evolved throughout history and what areas of application it is used for.

We didn't stop at this point because we needed to clarify as well the concept of the CNNs neural network, which is used in image classification. These networks are able to extract characteristics of images presented as input and to classify them. Convolutional neural networks, however, have a number of limitations; first, the hyper parameters of the network are difficult to assess a priority. Indeed, the number of layers and the number of neurons per layer or even the different connections between layers is crucial elements and essentially determined by good intuition or by a succession of tests / calculation of errors.
In the next chapter, (the related works chapter) we will see some cyberattack detection systems based on deep learning that was created by other researchers.

CHAPTRE III

Cyberattack Detection Based on Deep learning

17. Introduction

Mobile cloud computing is an emerging architecture which has been developed based on the power of cloud computing to serve mobile devices [66] .

However, the Mobile cloud computing is challenged by cybercrimes and in order to detect those cybercrimes and prevent them from harming cloud resources, in this context, we propose a new Deep Learning approach that is being used to early detect them.

But before presenting our contribution, in this chapter, we present some various Deep Learning models that have been applied to the areas of cyberattack detection.

18. Related works

There have been rich literatures dealing with cyberattacks in the cloud computing environment. Recently, many network attack detection techniques have been introduced to differentiate abnormal behavior from normal ones, in order to, detect unwanted or suspicious activity. Attacks detection methods can be classified into three distinct groups [67]:

- Supervised: In this class, a training set containing labeled instances for both the normal and anomalous class is available.
- Semi-supervised: The training here only contains instances for the normal class. Anything that cannot be characterized as normal is thus marked as anomalous.
- Unsupervised anomaly detection: No training set is available nor is it needed.

In [68], they have proposed a semi-supervised RBM-based detection system where the classifier is trained only with normal traffic data, so that the knowledge of abnormal behavior can dynamically evolve. The main advantage of this discriminatory model is its effectiveness in dealing with zero-day attacks, since they are not limited to prior knowledge. RBM is a generative classifier aimed at capturing as much as possible the variational potential of the inputs to describe the input data much better. In contrast, the discriminative RBM aims at combining the descriptive power with a sharp classification ability. In order to make an RBM operate in a supervised fashion, and they introduced an additional input containing the targets.

The datasets were thus structured as sequences of pairs (v, y) comprising an input vector v and a class y∈ {1, ..., C}. It is actually convenient to "vectorize" the class, introducing a vector y whose j-th component is $\delta_{j,C}$, using the Kronecker delta.

As it can be seen, DRBMs have many things in common with feed forward neural networks. The added value of an energy-based model, such as the one at the basis of RBMs, is that it introduces generative capabilities in an explicit, controllable, way [69]. Whereas the use of binary variables has been previously assumed, RBMs can easily accommodate different variables.

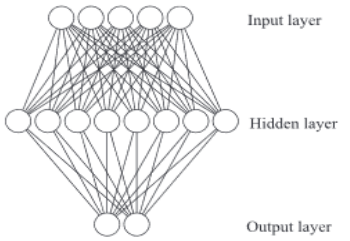


Figure 18.1: A typical three-layer network, with five inputs, eight hidden units, and two outputs. [68]

In their experiment, the objective was to test the accuracy of the DRBM in order to recognize abnormal traffic on real data. In the second experiment, the DRBM was trained with 10% learning with the KDD CUP dataset and tested against the actual data. In the experiments, 28 out of 41 characteristics related to network traffic were used. The KDD-CUP99 dataset containing 494,021 attack data records in the training set, 11,850 in the test set, and 41 attributes are used to evaluate the performance of the proposed model. The experimental results prove that there will be a dramatic reduction in classification performance, when the environment for the test data is very different from the network from which the training data was extracted. Therefore, more research should be done on the nature of abnormal traffic and how it differs from normal traffic.

Authors in [70] proposed detecting model consists of three phases. Firstly is the training phase, which focuses on constructing a profile of a normal network traffic behavior (baseline profile). It can be done by capturing the normal traffic than transfers it into the matching covariance matrix. Secondly, the resulted covariance matrix from the first phase is compared with a new covariance matrix of new traffic which can be normal or abnormal. The aim of this experiment

is to examine the effectiveness of the covariance matrix method to detect flooding based DoS attack.

Though this study showed a clear effectiveness to detect DoS attacks but its results was not that accurate as well as their methods.

By incorporating a different perspective, Dong in [71]have implemented different ML methods on the KDD'99 set to classify network traffic and distinguish normal traffic from attacks, DOS, Probe, U2R and R2L. They applied the Synthetic Minority over Sampling Technique (SMOTE) to solve the problem of imbalance in the dataset. They then compared classical ML methods such as NB, SVM and DT with SVM-RBM in terms of accuracy. The results show that SVM-RBM produces the best results compared to other traditional shallow learning architectures. Nevertheless, the maximum precision of SVM-RBM is acquired on 80% of the training samples, or 82% for DOS attacks, 58% for U2R, and 42% for R2L, which are remarkably low values compared to the advanced approaches in this domain. In addition, no other measurement, such as a false positive, a false negative, a recall or an F1 score is provided in the analysis part. Overall, the approach brings no value to the research community.

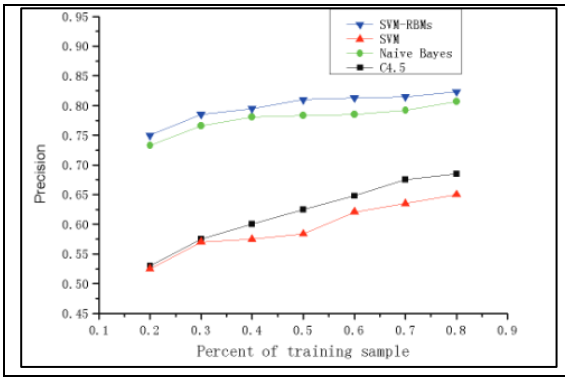


Figure 18.2: Precision of different algorithms used to recognize normal traffic. [71]

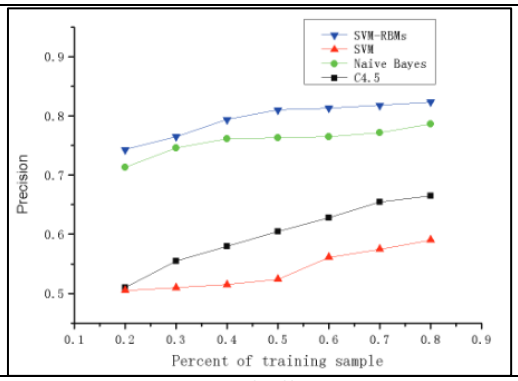


Figure 18.3: Precision of different algorithms used to recognize DOS traffic. [71]

Authors in paper [72] proposed a deep learning approach for attack detection using recurrent neural networks (RNN). Moreover, they studied the performance of the model in binary classification and multiclass classification, and the number of neurons and different learning rate impacts on the performance of their proposed model. In the data analysis phase authors in [72] used NSL_KDD dataset and they used it because all the researchers use it as the benchmark dataset, which not only effectively solves the inherent redundant records problems

of the KDD Cup1999 dataset but also makes the number of records reasonable in the training set and testing set [72]. For preprocessing, they converted non-numeric features to numeric values, a map of 41 dimensional feature vectors into 122 dimensional feature vectors, so the whole process went through two steps of NUMERICALIZATION and NORMALIZATION, after that in the training phase they used two parts Forward Propagation and Back Propagation. Forward Propagation is responsible for calculating the output values, and Back Propagation is responsible for passing the residuals that were accumulated to update the weights, which is not fundamentally different from the normal neural network training.

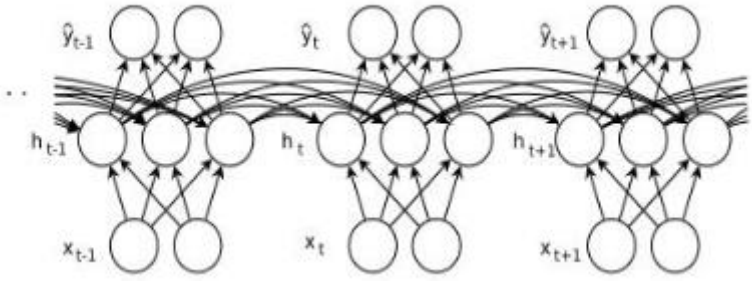


Figure 18.4: The unfolded Recurrent Neural Network. [72]

Recurrent neural networks have introduced a directional loop that can memorize the previous information and apply it to the current output, which is the essential difference from traditional Feed-forward Neural Networks (FNNs). The preceding output is also related to the current output of a sequence, and the nodes between the hidden layers are no longer connectionless; instead, they have connections. Not only the output of the input layer but also the output of the last hidden layer acts on the input of the hidden layer [72].

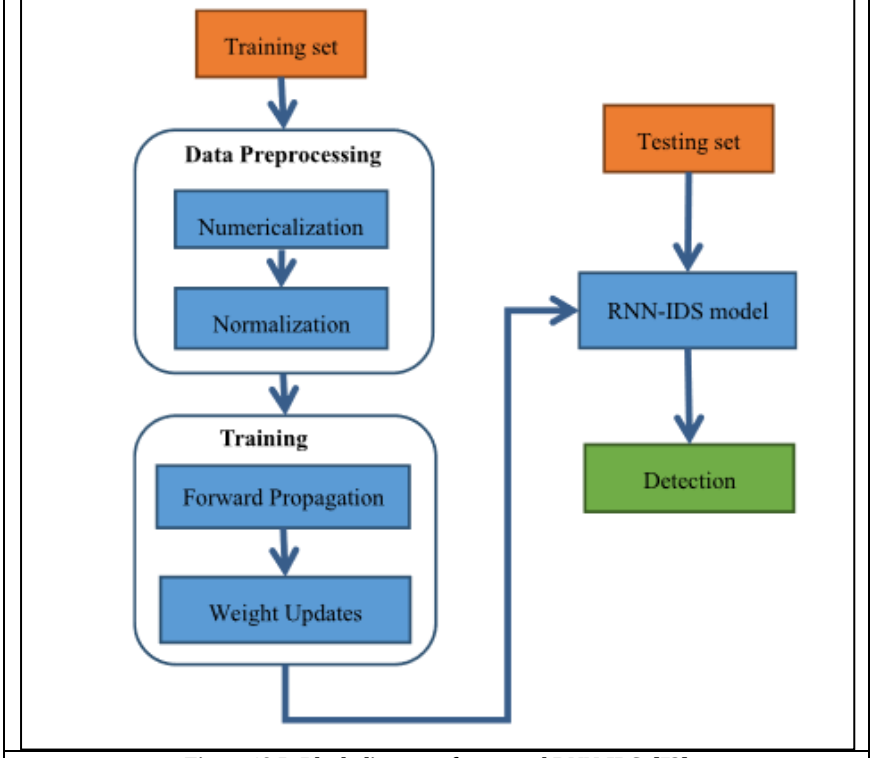


Figure 18.5: Block diagram of proposed RNN-IDS. [72]

The accuracy results for the binary classification of this model on the testing set of NSL_KDD dataset equals 83.28%, while the accuracy for the multiclass classification of it is 81.29%.

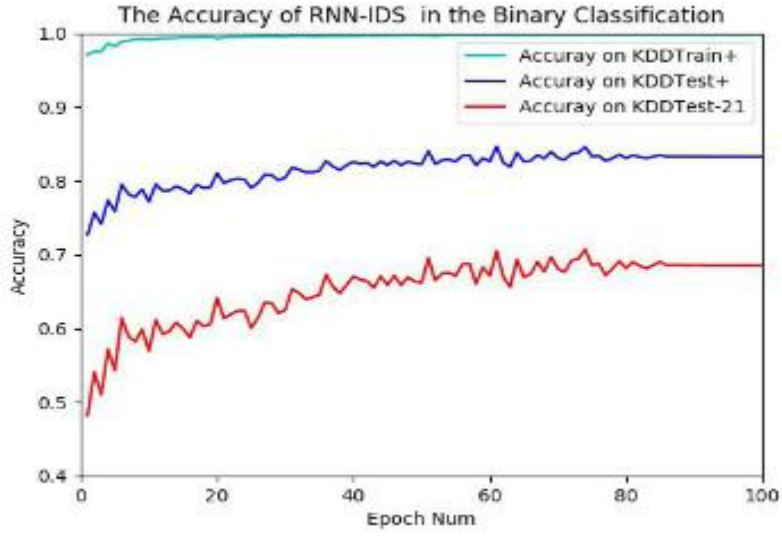


Figure 18.6: The Accuracy on the KDDTest+ and KDDTest−21datasets in the Binary Classification. [72]

However, this study has a high accuracy in both binary and multiclass classification. Compared with traditional classification methods, but the detection rate by attack category is very low at 24.69% and 11.50% for the R2L and U2R categories respectively and not acceptable for DoS and Probe (83.4%). In addition, they did not indicate the number of layers of RNN.

Furthermore, Authors in [73] proposed a novel framework that leverages a deep learning approach to detect cyberattacks in mobile cloud environment, through experimental results. They used an RBMs model with GRBM and PCA to create an effective detection system that can classify the attacks into 5 different classes using softmax regression as an output layer.

Principal Component Analysis (PCA) is used as an effective technique to emphasize variation and determine strong patterns in a dataset. The core idea of PCA is to reduce the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much as possible the variation presented in the data set [74] . Here they adopted the PCA to reduce dimensions for considered datasets.

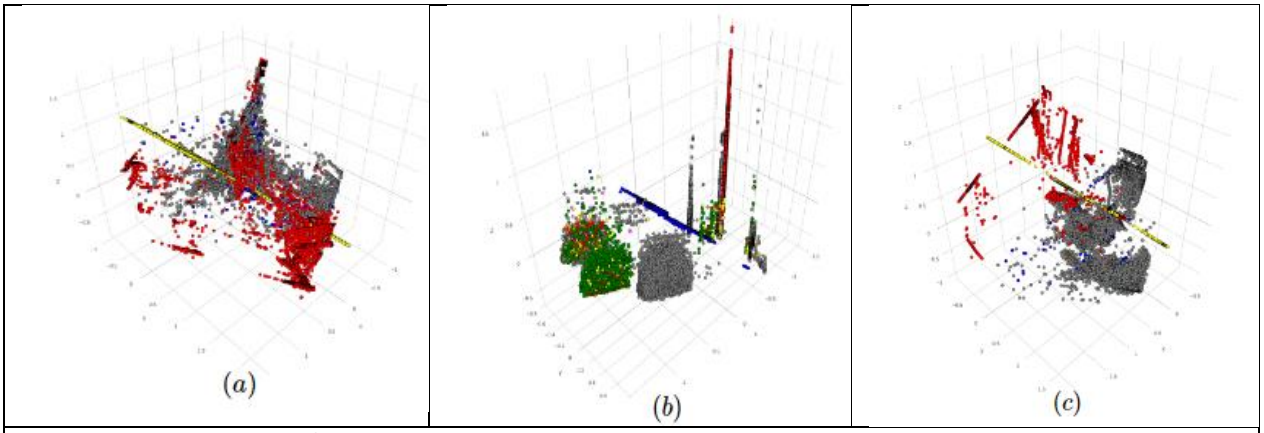


Figure 18.7:Visualizations of three datasets: (a) NSL-KDD, (b) UNSW-NB15, and (c) KDDcup 1999, by using PCA with 3 most important features. Grey circlesrepresent normal packets, while circles with other colors than grey express the different types of attacks [73]

In the pre-learning step, they used a Gaussian Binary Restricted Boltzmann Machine (GRBM) to transform real values into binary codes which will be used in the hidden layers.

The next step includes a series of learning processes which are performed in sequence to adjust weights of the neural network. Each learning process is performed between two successive

layers in the hidden layers through a Restricted Boltzmann Machine (RBM). The RBM is a particular type of Markov random field. It has a two-layer architecture in which the visible binary stochastic are connected to the hidden binary stochastic units. [73]

Finally, the last step was to obtain the output of the last hidden layer and pass it as an input to the softmax regression in order to classify it as normal or attack. The performance of the proposed RBMs model is tested on the NSL-KDD, UNSW_NB15 and KDD_CUP99 datasets and compared to K_means and SVM and some other algorithms. The accuracy of the proposed RBMs model has been shown to be 90% on the NSL_KDD and 95.84%, 97.11% on UNSW_NB15 and KDD_CUP99 datasets respectively.
However, the NSL_KDD accuracy was low compared to the other datasets also no serious comparison with other deep learning systems and no detailed plots were reported in this article, which shows a considerable weakness of the proposed system.

Authors in [75] used a KDD dataset and rescaled all its numerical features to be between 0 to 255 to convert the 117-dimensional vector into images with 13×9 pixels. Each color channel of the image should be represented with the value between 0 to 255. They then feed these images to their CNN model. The reason to convert the numerical samples into images is that CNN is a DL model for image training. Then they generated two types of image datasets. One is an RGB set which has 3 color channels (Red, Green, and Blue) and the other one is a grayscale set that has a single channel. An RGB image is an overlaid structure of the three types of color images and is converted into an array of M×N×3 pixels.
M and N are the number of columns and rows. After that they tried to observe how accuracy varies depending on the type of image.

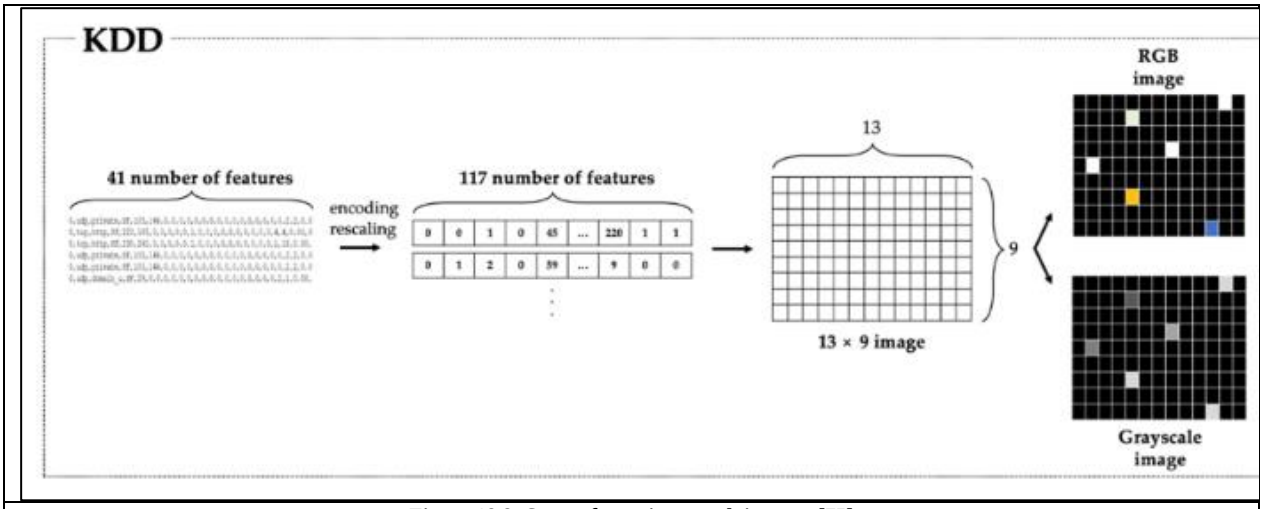


Figure 18.8: Steps of creating attack images. [75]

After the extractions of their features, the authors in [75] designed the CNN model in figure2.8, The proposed CNN model receives grayscale or RGB images as its input. The CNN model is also able to change two more parameters such as the number of convolutional layer and size of kernel

The CNN model consists of 1, 2, or 3 convolutional layers, and the number of kernels corresponding to the number of neurons per layer increases by a multiple of 2.the kernel size is usually set to 3×3. However, they set it to 3×3 as a median value and done experiment on sizes of 2×2 and 4×4 to find out the optimal size. The kernel generates a feature map by moving over the image as much as stride which is designated value. They set the stride to 1 to extract the feature densely. Figure 2.9 shows examples of their CNN design where they chose 6 scenarios (RGB-1, GS-1, RGB-5, GS-5, RGB-9 and GS-9) that can show various CNN designs with a different number of layers, kernels, and color channels.

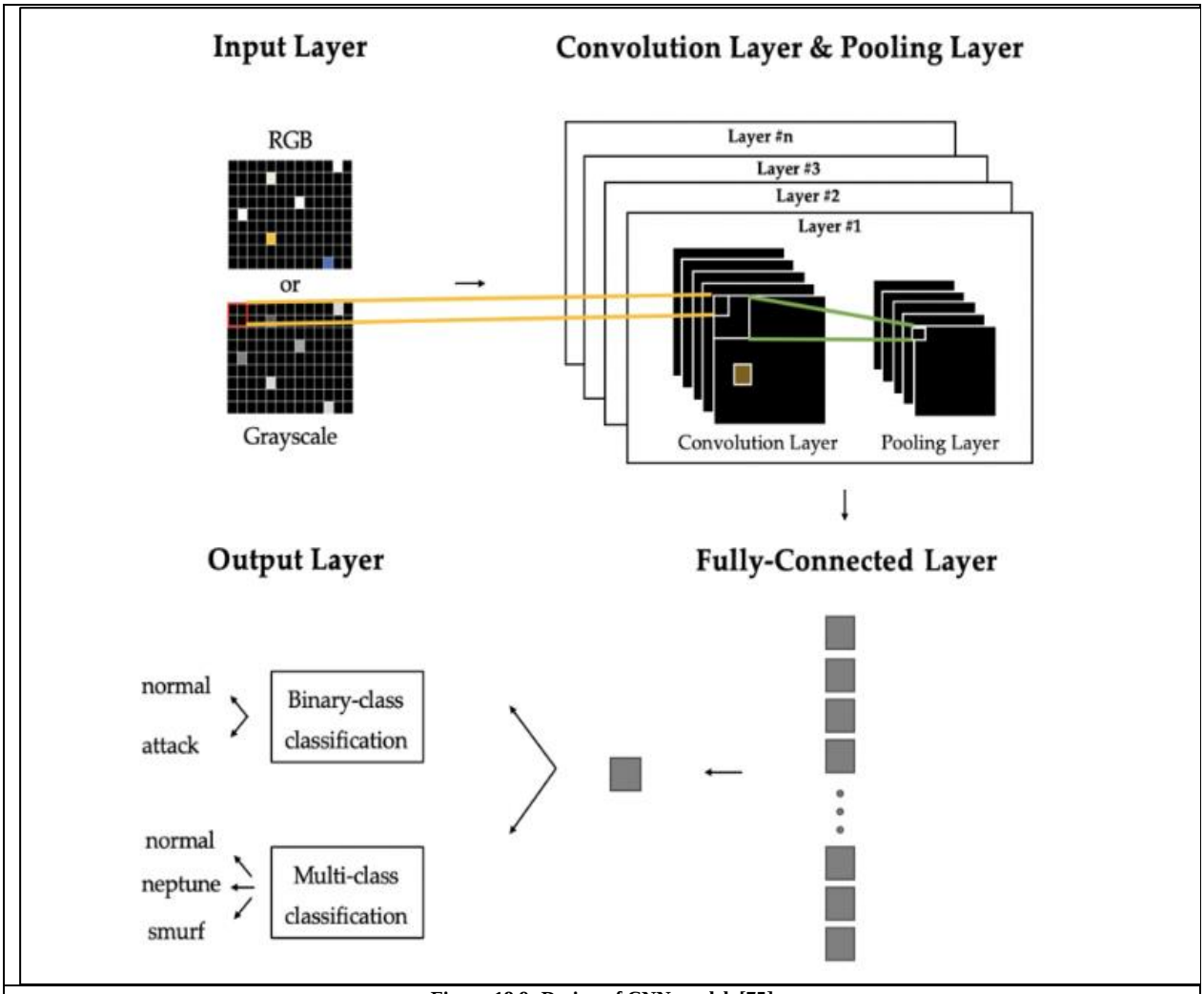


Figure 18.9: Design of CNN model. [75]

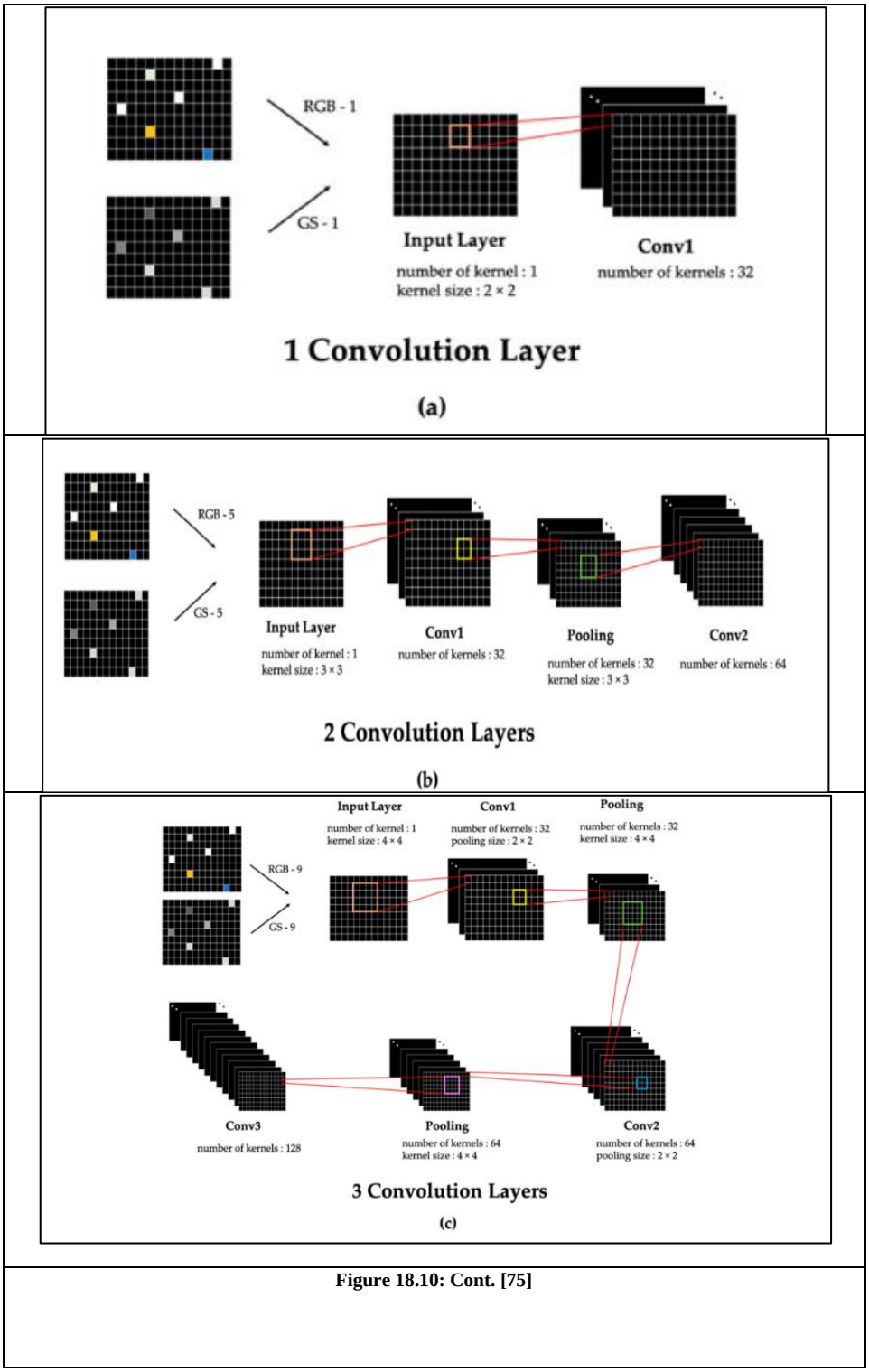


Figure 18.10: Cont. [75]

The experiments in [75] are performed with binary and multiclass classifications. However, their result of accuracy of the neptune is 80%, which show how weak is their CNN model in detecting DoS attacks.

Table 2.1 shows a selection of investigations of attack detection systems: using Deep Learning algorithms based on models, characteristics, datasets, and performance measures.

Year	Authors	Deep Learning Model	Features	Dataset	Accuracy
2013	U. Fiore, F. Palmieri, A. Castiglione, A. De Santis.	DRBM	28 out of 41 characteristics linked to network traffic	Real traffic including normal traffic on one host and traffic infected by a bot on another host & training with KDD dataset of 10% and test with real data	94%
2014	Mohd Nazri Ismail, Abdulaziz Aborujilah ,Shahrulniza Musa ,AAmir Shahzad	Covariance Matrix	IP flags behavior such as TCP FIN, RST and FIN and TCP retries	Normal and Abnormal Traffic	/
2016	Bo Dong, Xue Wang	SVM-RBM	41 features	10% KDD dataset	PR(DOS)=82% PR(U2R)=58% PR(R2L)=42%
2017	CHUANLONG YIN, YUEFEI ZHU,JINLONG FEI, AND XINZHENG HE	RNN	41 features	NSL-KDD	ACC=83.28%/81.29%
2018	Khoi Khac Nguyen, Dinh Thai Hoang, Dusit Niyato, Ping Wang, Diep Nguyen, and Eryk Dutkiewicz	RBMs	41 features	NSL-KDD KDD_CUP99 UNSW_NB15	ACC(NSL)=90.99% ACC(KDD99)=97.11% ACC(NB15)=95.84%
2020	Jiyeon Kim, Jiwon Kim, Hyunjung Kim, Minsun Shim and Eunjung Choi	CNN	41 features	KDD	ACC(neptune)=80%

Tableau 18.1: Selected studies results.

19. Conclusion

Cyberattack detection is considered as a necessary security mechanism to face network attacks and identify malicious activities.

It is highly recommended to use Deep Learning in various fields and especially in information security technologies to help discover, determine and identify various possible threats.

Deep learning has played a key role in complex problem solving, thanks to their various advantages over other traditional machine learning (ML) techniques, that's why in the next and last chapter we are going to present and discuss the work and results of our cyberattack detection system using deep learning.

However, in this chapter we have presented a survey of an important approaches presented in the area of Mobile Cloud Computing.

In the next chapter we present in details our contribution to take into account Cyber Attack detection for Mobile Cloud Computing.

CHAPTRE IV

Cyberattack Detection in Mobile Cloud Computing using A deep learning approach (Contribution)

20. Introduction

Cyberattack detection systems have been the subject of much research, but until now this domain requires more investigations.

That is why in this chapter we will create a system capable of detecting different types of attacks with a high accuracy using a Deep Learning technology; the proposed system is trained with the 3 datasets: NSL-KDD, KDD-Cup99 and UNSW_NB15 [76] [77].
But, first of all, we are going to focus on processing the datasets in the Features Analysis phase, with two different methods; the first one, by using old methods then by using new methods created by Google and after that, we are going to try two different model architectures to see at the end of this chapter whether we chose the right methods to process our data where accuracy is most affected or choosing the model architecture is really what we should be focusing on, to achieve higher accuracy.

21. Execution environment

- **Python:** Python is a high-level, object-oriented, interpreter programming language. Created by Guido van rossum and first released in 1991. Python supports modules and packages, which is why python is popular and used by a large community of developers and programmers.

- **Google Colaboratory**: or “Colab” for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary **python** code through the browser, and is especially well suited to machine learning, data analysis and education.
  This tool allows us to develop deep learning applications in Python in a flash. To get started, all we need to do is have a Gmail account.
  One of the great strengths of the Python language lies in the large number of software libraries available. Python 3.5 with Pandas, Numpy, H5py, Scikit-Learn, TensorFlow and Keras were chosen for this task.
- **Numpy and Pandas** are used to access and reorganize data.

- **H5py** is used to store the data in HDF5 format on the hard drive between different stages.
- **Scikit-Learn** which includes many predefined machine learning algorithms and which makes it possible to experiment with different techniques quickly and easily was used.
- **Keras** is used to create neural networks. Keras understands many types of neural layers and supports GPUs. It allows users to choose whether the models they build run on the Theano or TensorFlow framework. It is written and maintained by Francis Chollet, a member of the Google Brain team.

From the available frameworks, our choice fell on the Google TensorFlow framework created by the Google Brain team to conduct research on ML and Deep Learning. The main reason for this choice is the very large and also very active community of this library.

The other frameworks available are:

- CNTK: by Microsoft.
- Neon: by Nervana Systems. It was recently ranked as the fastest frameworks in several categories
- Deeplearning4j: It supports java language
- Caffe: by Berkeley Vision and Learning Center.

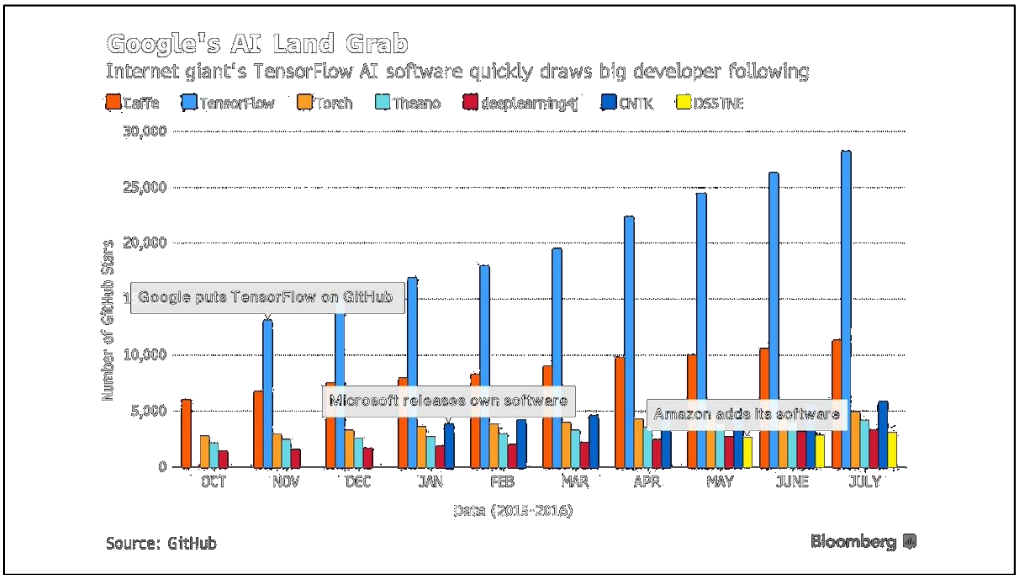


Figure 21.1: The growth in popularity of TensorFlow.

22. System Architecture

Users are trying, every day, to access their own data stored in the mobile cloud computing environment, but daily we find different types of users, the ones who try to access it and others who try to steal it. So, in order to, assure security for users we design a system presented in **figure22.1** to make sure that only confidential users will access their private data safely.

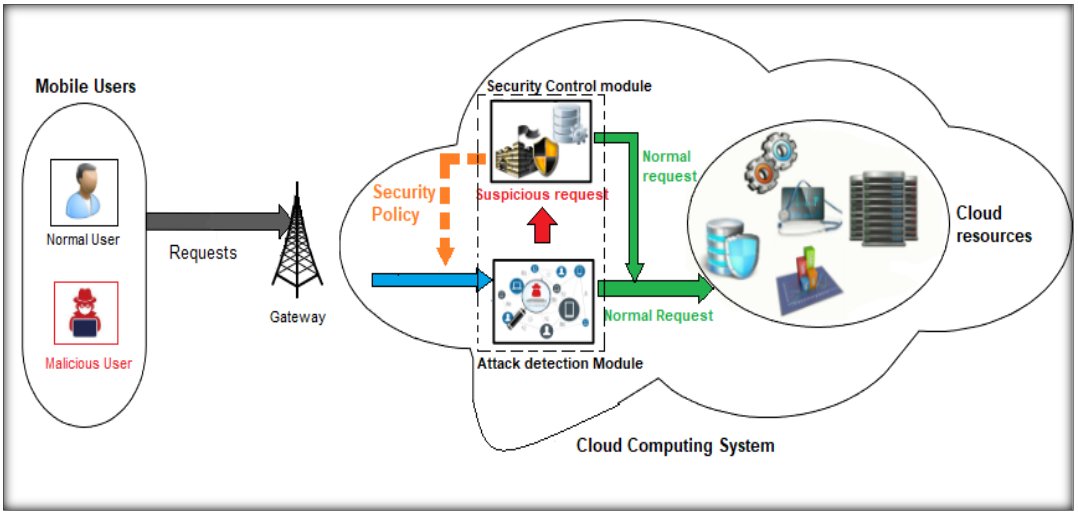


Figure 22.1 : System model architecture for cyberattack detection in mobile cloud computing systems. [73]

23. Datasets and Features Analysis

Feature analysis is what we first do in deep learning model, and we do this step to extract features and learn from them that's why in our case we did two different types of extractions and analysis using old and new methods in order to see how will that affect our accuracy results . Malicious packets features could be different from the normal ones, that's why by extracting and analyzing the abnormal attributes of those packets, we can decide whether the packet is malicious or not and we want to know by the end of this chapter how it affects our accuracy the way we process it.

Among the existing datasets we decided to work with the most popular and most used datasets in the field that contains different types of attacks and many different types of features. The three datasets are NSL-KDD, KDD-CUP99 and UNSW-NB15.

23.1 KDD-Cup99:

this dataset is the mostly widely and commonly used dataset and it consists of five million records, The KDD training dataset consist of 10% of original dataset that is approximately 494,020 single connection vectors each of which contains 41 features and is labeled with exact one specific attack type i.e., either normal or an attack. Each vector is labeled as either normal or an attack, with exactly one specific attack type. Deviations from 'normal behavior', everything that is not 'normal', are considered attacks. [78]Attacks labeled as normal are records with normal behavior. A smaller version 10% training dataset is also provided for memory constrained machine learning methods. The training dataset has 19.69% normal and 80.31% attack connections. KDD CUP 99 has been most widely used in attacks on network. The simulated attack falls in one of the following four categories [76] as shown in **figure23.1**.

Attacks Class	Attacks Types
DoS	Back, Neptune, Land, Pod, Smurf, Udpstorm, Teardrop, Apache2, Worm, Processtable
Probe	Nmap, Ipsweep, Portsweep, Saint, Mscan, Satan
R2L	Ftp_write, Warezmaster, Httptunnel, Guess_Password, Phf, Warezclient, Snmpguess, Multihop, Xlock, Spy, Xsnoop, Sendmail, Snmpgetattack, Imap, Named
U2R	Rootkit, Loadmodule, Xterm, Sqlattack, Perl, PS, Buffer_overflow

Figure 23.1 : Attacks Categories with its types of the KDD-CUP99 Dataset

23.2 NSL-KDD:

The NSL-KDD data set is the refined version of the KDD cup99 data set [79].

This dataset is designed by Tavallaee et al [76]; it is developed after the removal of redundant and duplicate records from training and test data of KDDCup. It contains only selected and necessary records from, and there are total of 37 attacks out of which 27 attacks are used by testing dataset and 23 attacks are used by training dataset for experiments [8].

The number of feature in NSL-KDD dataset has same as that of in KDDCup. This dataset contains 41 features and 5 attack classes as shown in **figure 23.2** and **23.3**. There is one normal class and other 4 are different types of attack. These different attacks are grouped into four categories: Probe attack, Denial of service attack (DoS), User to Root (U2R) and Remote to Local (R2L). The above dataset holds a binary class attributes as well as reasonable number of training and test instances [80].

There are total 21 different types of attacks in the NSL-KDD dataset which are present in training dataset, while test dataset contain 16 additional attacks. Major attacks are categorized as Probe, DoS, U2R and R2L [81]

Attack category	Attack type	KDD Cup 99		NSL-KDD	
		Training set kddcup.data10percent	Testing set corrected.gz	Training set KDDTrain+20Percent	Testing set KDDTest+
Denial of Service (DoS)	Neptune	107201	58001	8282	4657
	Smurf	164091	280790	529	665
	Pod	264	87	38	41
	Teardrop	979	12	188	12
	Land	21	9	1	7
	Back	2203	1098	196	359
	Apache2	-	794	-	737
	Udpstorm	-	2	-	2
	Process-table	-	759	-	685
	Mail-bomb	-	5000	-	293
User to Root (U2R)	Buffer-overflow	30	22	6	20
	Load-module	9	2	1	2
	Perl	3	2	0	2
	Rootkit	10	13	4	13
	Spy	2	-	1	-
	Xterm	-	13	-	13
	Ps	-	16	-	17
	Http-tunnel	-	158	-	133
	Sql-attack	-	2	-	2
	Worm	-	2	-	2
	Snmp-guess	-	2406	-	331
Remote to Local (R2L)	Guess-password	53	4367	10	1231
	Ftp-write	8	3	1	3
	Imap	12	1	5	1
	Phf	4	2	2	2
	Multihop	7	18	2	18
	Warezmaster	20	1602	7	944
	Warezclient	1020	-	181	-
	Snmpgetattack	-	7741	-	178
	Named	-	17	-	17
	Xlock	-	9	-	9
	Xsnoop	-	4	-	4
	Send-mail	-	17	-	14
Probe	Port-sweep	1040	354	587	157
	IP-sweep	1247	306	710	141
	Nmap	231	84	301	73
	Satan	1589	1633	691	735
	Saint	-	736	-	319
	Mscan	-	1053	-	996

Figure 23.2 : Attacks distribution of the NSL-KDD & KDD-CUP99 Datasets.

No.	Feature Name	No.	Feature Name
1	Duration	22	is_guest_login
2	protocol type	23	**count**
3	service	24	srv_count
4	**flag**	25	**serror_rate**
5	src_bytes	26	**srv_serror_rate**
6	**dst_bytes**	27	**rerror_rate**
7	land	28	srv_rerror_rate
8	wrong_fragment	29	**same_srv_rate**
9	urgent	30	diff_srv_rate
10	hot	31	srv_diff_host_rate
11	num_failed_logins	32	dst_host_count
12	logged_in	33	**dst_host_srv_count**
13	num_compromised	34	dst_host_same_srv_rate
14	root_shell	35	dst_host_diff_srv_rate
15	su_attempted	36	dst_host_same_src_port_rate
16	num_root	37	dst_host_srv_diff_host_rate
17	num_file_creations	38	**dst_host_serror_rate**
18	num_shells	39	**dst_host_srv_serror_rate**
19	num_access_files	40	dst_host_rerror_rate
20	num_outbound_cmds	41	dst_host_srv_rerror_rate
21	is_host_login	42	

Figure 23.3 : NSL-KDD dataset features.

23.3 UNSW-NB15:

UNSW-NB 15 data set was created in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) by the IXIA Perfect Storm tool. This dataset contains a hybrid of normal activities and attack behaviors. Tcp-dump tool was used to capture 100 GB of the raw traffic. Twelve algorithms and tools such as Argus, Bro-IDS were used to generate UNSW-NB15. It contains 49 features including a class label [82] as shown in **Table 23.1**.

The UNSW-NB15 dataset has 2.5 million records in total. The number of records in the learning set is 36,020 and in the test set 2,477,719, with 47 characteristics and two tags as target: is the recording an attack called "attack or not" and category. The category specifies the type of record as shown in the table down below.

Table 23.1: Features of UNSW-NB15 dataset [83]

SNo.	Name	Type	Description
36	Is_sm_ips_ports	Binary	”If source (1) and destination (3)IP addresses equal and port numbers (2)(4) equal then this variable takes value 1 else 0”
39	Is-ftp-login		If the ftp session is accessed by user and password then 1 else 0.
49	Label		0 for normal and 1 for attack records
7	Dur	Float	Record total duration
15	Sload		Source bits per second
16	Dload		Destination bits per second
27	Sjit		Source jitter (mSec)
28	Djit		Destination jitter (mSec)
31	Sintpkt		Source interpacket arrival time (mSec)
32	Dintpkt		Destination interpacket arrival time (mSec)
33	Tcprtt		”TCP connection setup round-trip time, the sum of synack and ackdat.”
34	Synack		”TCP connection setup time the time between the SYN and the SYN ACK packets.”

35	Ackdat		”TCP connection setup time the time between the SYN ACK and the ACK packets.”
2	Sport	Integer	Source port number
4	Dsport		Destination port number
8	Sbytes		Source to destination transaction bytes
9	Dbytes		Destination to source transaction bytes
10	Sttl		Source to destination time to live value
11	Dttl		Destination to source time to live value
12	Sloss		Source packets retransmitted or dropped
13	Dloss		Destination packets retransmitted or dropped
17	Spkts		Source to destination packet count
18	Dpkts		Destination to source packet count
19	Swin		Source TCP window advertisement value
20	Dwin		Destination TCP window advertisement value
21	Stcpb		Source TCP base sequence number
22	Dtcpb		Destination TCP base sequence number
23	Smeansz		Mean of the ?ow packet size transmitted by the src
24	Dmeansz		Mean of the ?ow packet size transmitted by the dst
25	Trans_depth		Represents the pipelined depth into the connection of http request/response transaction
26	res bdy_len		Actual uncompressed content size of the data transferred from the servers http service.
37	Ct_state_ttl		No. for each state (6) according to specific range of values for source/destination time to live (10) (11).
38	Ct-flw-http-mthd		No. of flows that has methods such as Get and Post in http service.
40	ct ftp-cmd		No of flows that has a command in ftp session.
41	Ct_srv_src		No. of connections that contain the same service (14) and source address (1) in 100 connections according to the last time (26).
42	Ct_srv_dst		No. of connections that contain the same service (14) and destination address (3) in 100 connections according to the last time (26).

43	Ct_dst_ltm		No. of connections of the same destination address (3) in 100 Connections according to the last time (26).
44	Ct_src_ltm		No. of connections of the same source address (1) in 100 connections according to the last time (26).
45	Ct_src_dport_ltm		No of connections of the same source address (1) and the destination port (4) in 100 connections according to the last time (26).
46	Ct_dst_sport_ltm		No of connections of the same destination address (3) and the source port (2) in 100 connections according to the last time (26).
47	Ct_dst_src_ltm		No of connections of the same source (1) and the destination (3) address in 100 connections according to the last time (26).
1	Srcip	Nominal	Source IP address
3	Dstip		Destination IP address
5	Proto		Transaction protocol
6	State		Indicates to the state and its dependent protocol
14	Service		”http ftp smtp ssh dns ftp-data
48	attack_cat		”The name of each attack category. In this data set nine categories *eg* Fuzzers Analysis Backdoors DOS exploits Generic Reconnaissance Shellcode and Worms”
29	Stime	Timestamp	record start time
30	Ltime		record last time

24. Attack detection module architecture:

Our main work here is to build the attack detection model from scratch as shown in **Figure 24.1** and train it with Convolutional Neural network (CNN 1D), in order to, help the detection of the highest number possible of attacks, which our network could be affected from.

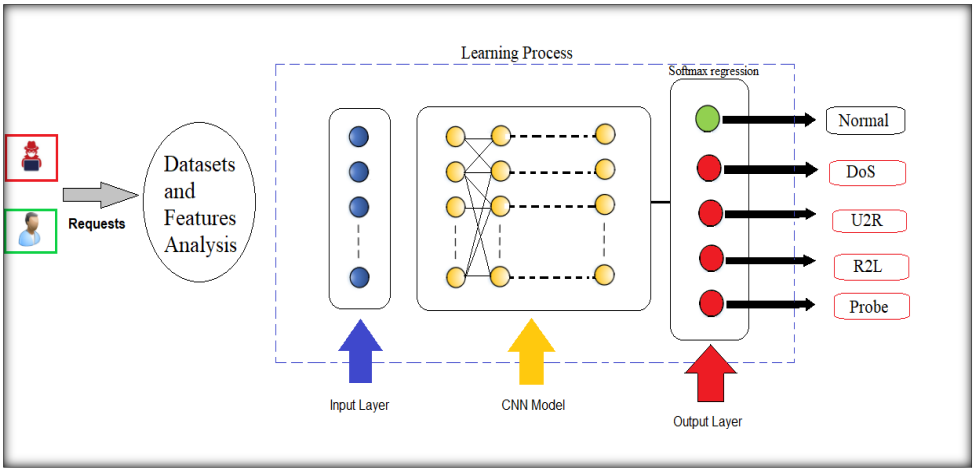


Figure 24.1 : Attack detection module architecture for cyberattack detection in mobile cloud computing system.

25. Neural network architecture:

The architecture of the neural network is different and depends on the model used.

In our case we used 2 different architectures based on CNN 1D to see how choosing 2 different architectures affects our accuracy, we used a ReLU activation function in both architectures because it always has the best results and The model can stack a chain of convolutional layers with pooling layers (MaxPooling1D) and then we used a ReLU in the output layer for the second architecture while we used a SoftMax activation function For each output layer in the first architecture because the sample labels are categorical. The dimensions are the same as the number of different classes. The output of the last CNN 1D model layer of our first model architecture will be used as the input of the softmax regression for the output layer to classify the packet. A packet can be classified into M= (K+ 1) classes, where K denotes all types of attacks. Mathematically, the probability that an output prediction Y is class i, is determined by:

$$p(Y = i|\mathbf{x}, \mathbf{W}, \mathbf{b}) = softmax_i(\mathbf{W}\mathbf{x} + \mathbf{b}) = \frac{e^{W_i \mathbf{x} + b_i}}{\sum_j e^{W_j \mathbf{x} + b_j}}, \quad [76]$$

Where W is a weight matrix between the last layer and the output layer, and b is a bias vector. Then, the model's prediction y_{pd} is the class whose probability is maximal, specifically:

$$y_{\text{pd}} = \arg\max_i \left[p(Y = i|\mathbf{x}, \mathbf{W}, \mathbf{b})\right], \forall i \in \{1, 2, \ldots, M\}. \quad [84]$$

The architecture of each neural network model is described in the figures down below.

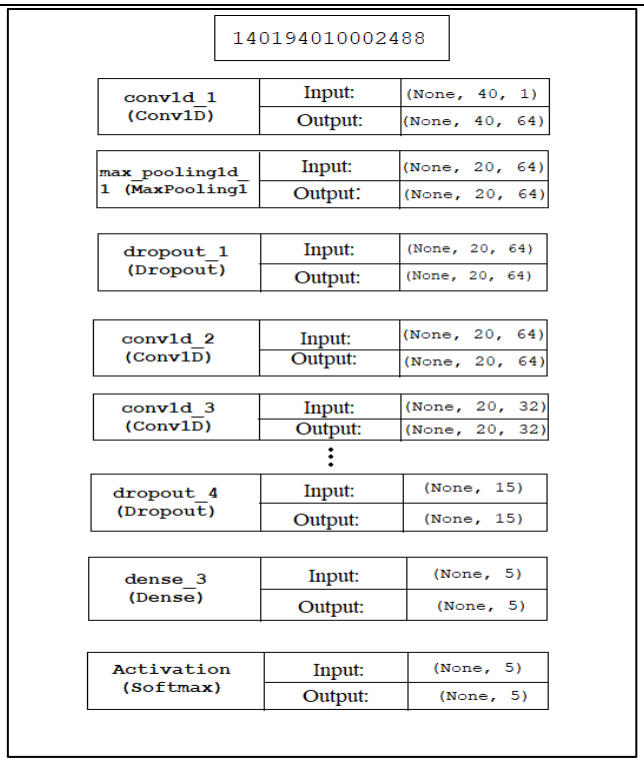


Figure 25.1 :The first CNN Architecture of NSL_KDD

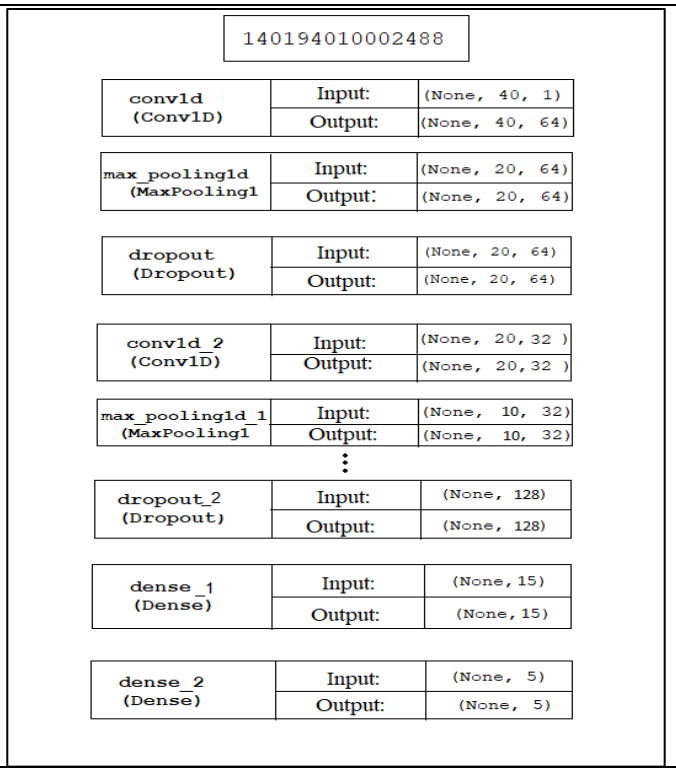


Figure 25.2 : The second CNN Architecture of NSL_KDD

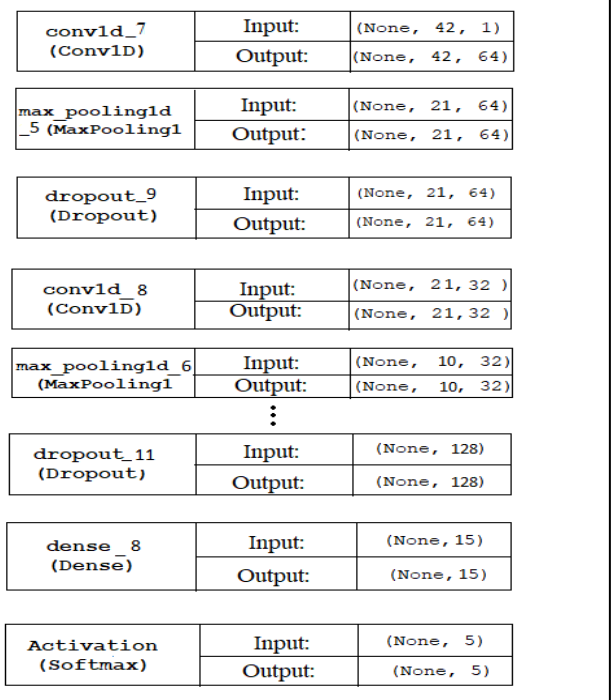


Figure 25.3 : CNN Architecture model of UNSW-NB15

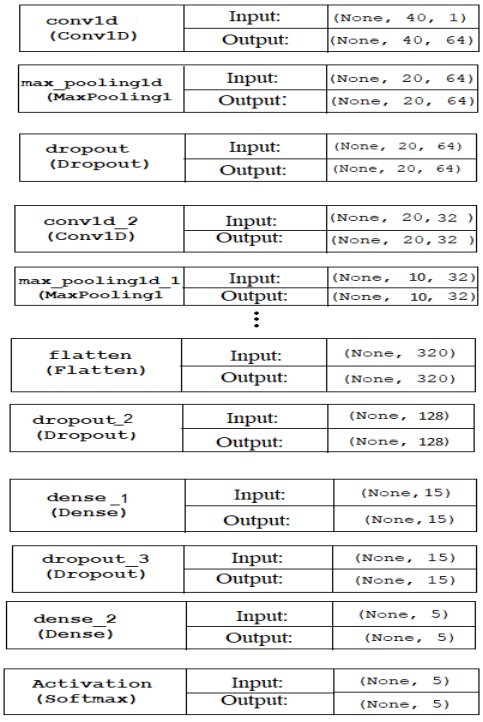
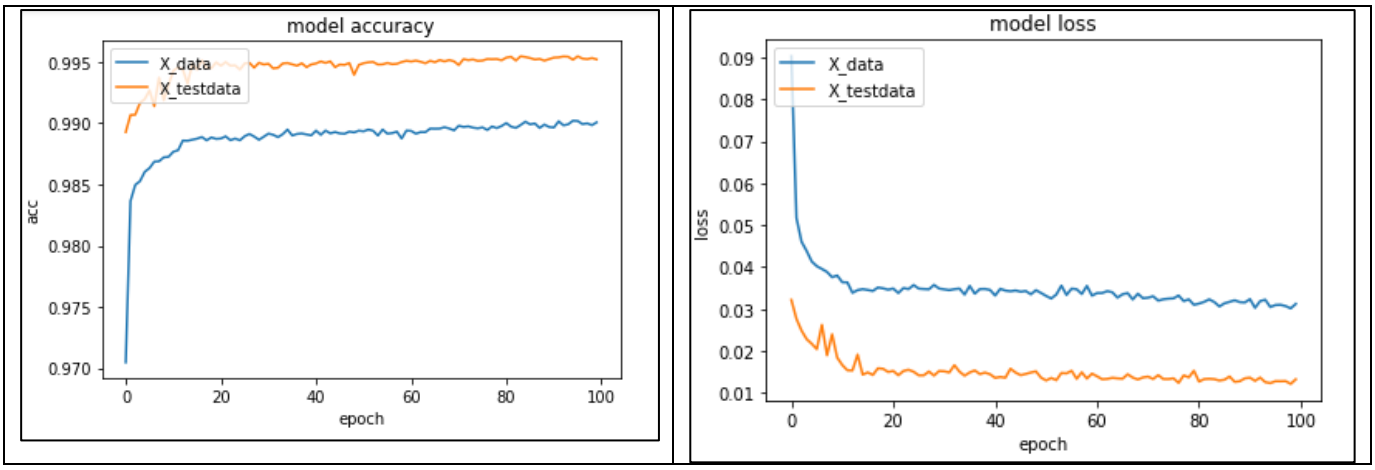


Figure 25.4 : CNN Architecture model of KDD_CUP99



26. Results and Discussions

26.1 Learning and testing the proposed models

During the learning and testing phase, our deep learning models were trained on 50 to 100 epochs, and our datasets were divided into 2 sets a training set and a validation set used as a pre-test dataset before we feed it into our actual testing set using the NSL_KDD, KDD_CUP99 and UNSW_NB15 datasets respectively .

On the NSL_KDD dataset we trained it using two different CNN 1D model architectures then we kept the same first architecture and we changed the methods we used to process our dataset. The Same things were done on the KDD_CUP99, where we used different methods to extract the features while on the UNSW_NB15 dataset we used two different CNN 1D model architectures. All the results obtained are shown in the figures bellow.

Figure 26.1 : CNN first model architecture training & test accuracy epochs for NSL_KDD	Figure 26.2 : CNN first model architecture training & test loss epochs for NSL_KDD

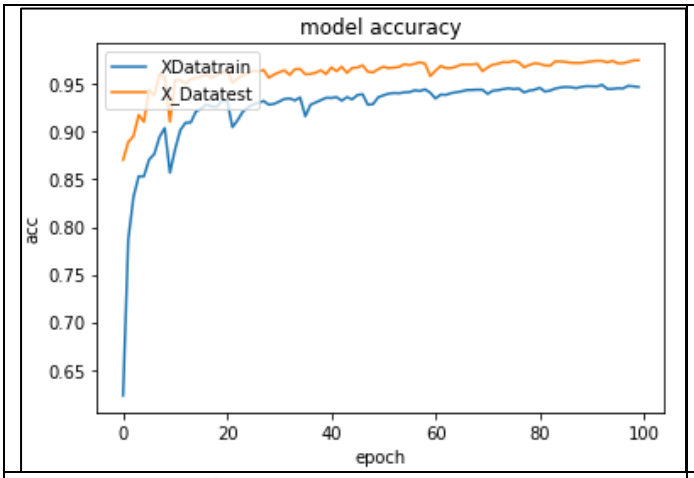


Figure 26.3 : CNN first model architecture training & test accuracy epochs for NSL_KDD with different feature analysis method.

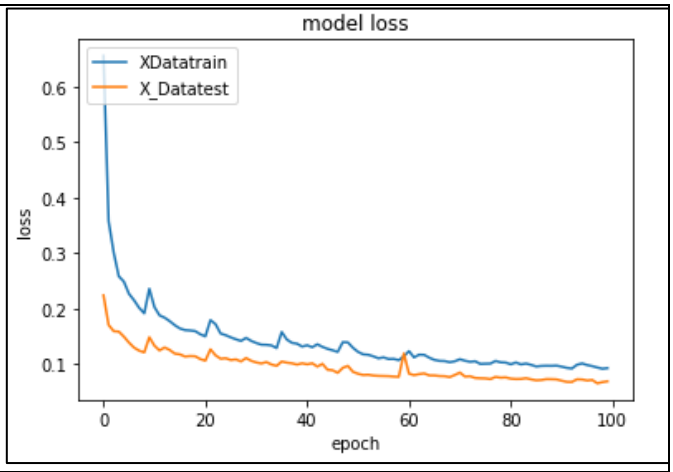


Figure 26.4 : CNN first model architecture training & test accuracy epochs for NSL_KDD with different feature analysis method.

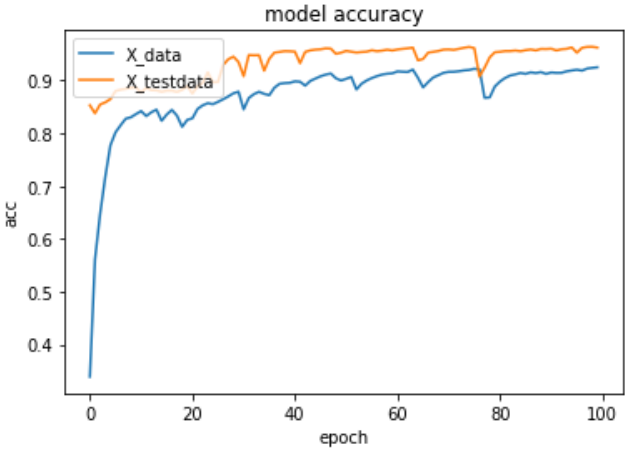


Figure 26.5 : CNN second model architecture training & test accuracy epochs for NSL_KDD.

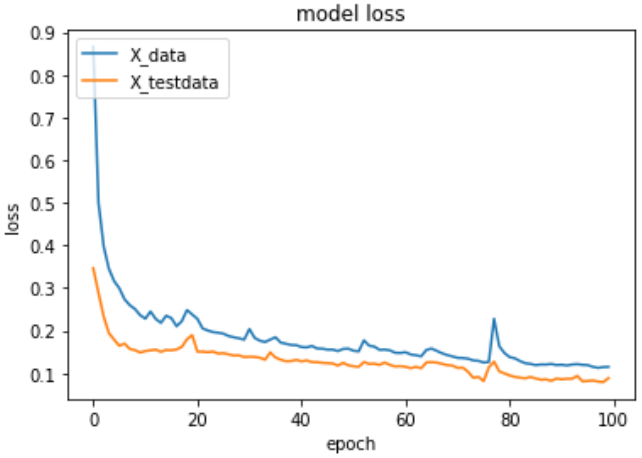


Figure 26.6 : CNN second model architecture training & test loss epochs for NSL_KDD.

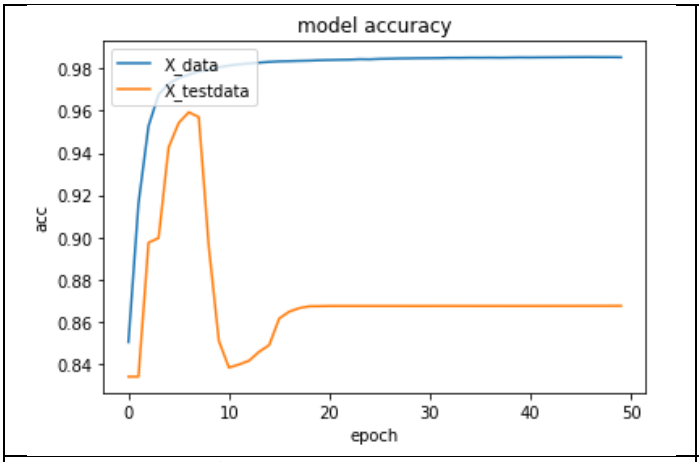


Figure 26.7 : CNN first model architecture training & test accuracy epochs for KDD_CUP99.

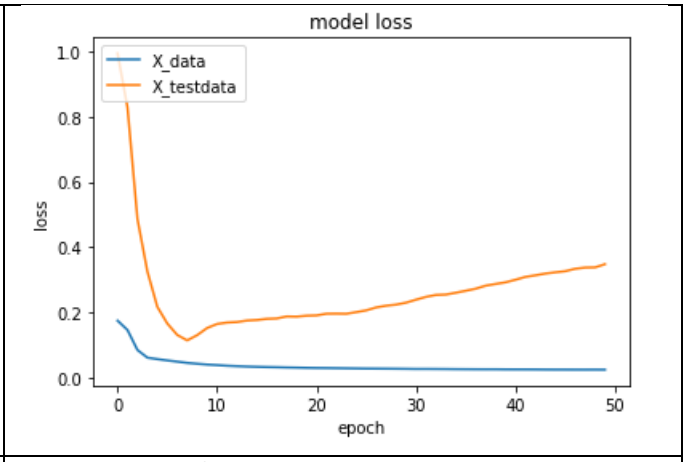


Figure 26.8 : CNN first model architecture training & test loss epochs for KDD_CUP99.

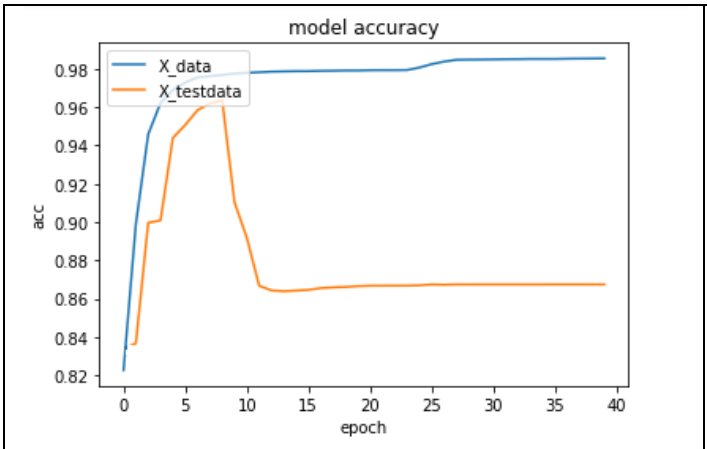


Figure 26.9 : CNN first model architecture training & test accuracy epochs for KDD_CUP99 with different feature analysis method.

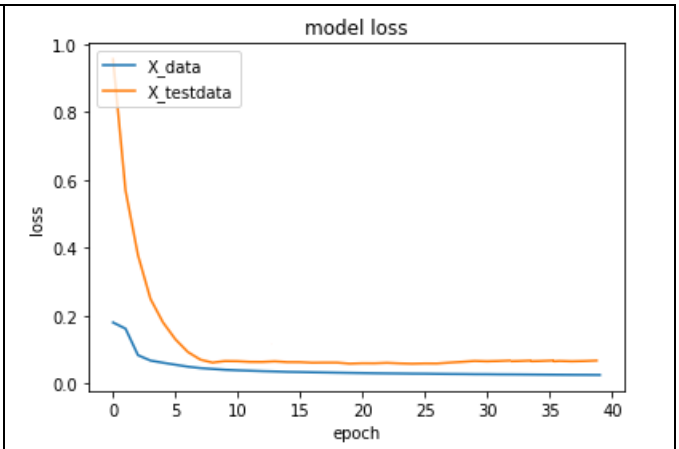


Figure 26.10 : CNN first model architecture training & test loss epochs for KDD_CUP99 with different feature analysis method.

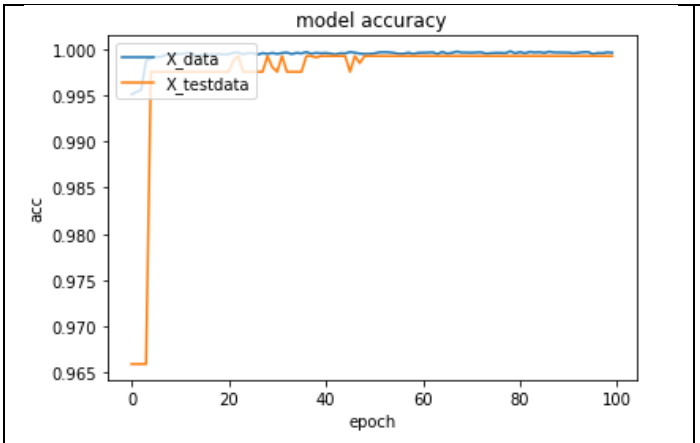


Figure 26.11 : CNN first model architecture training & test accuracy epochs for UNSW_NB15.

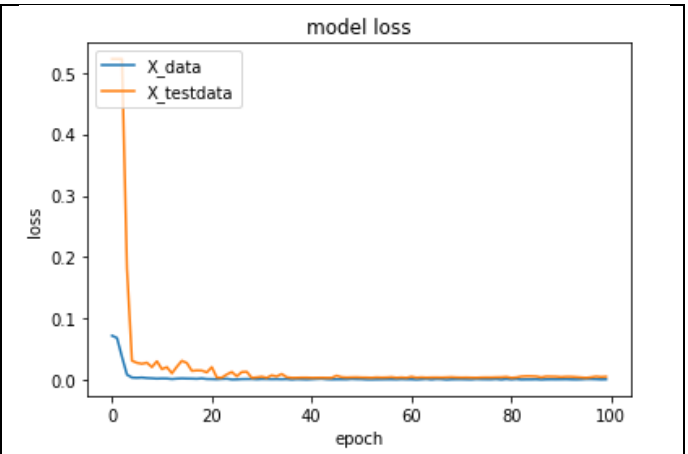


Figure 26.12 : CNN first model architecture training & test loss epochs for UNSW_NB15.

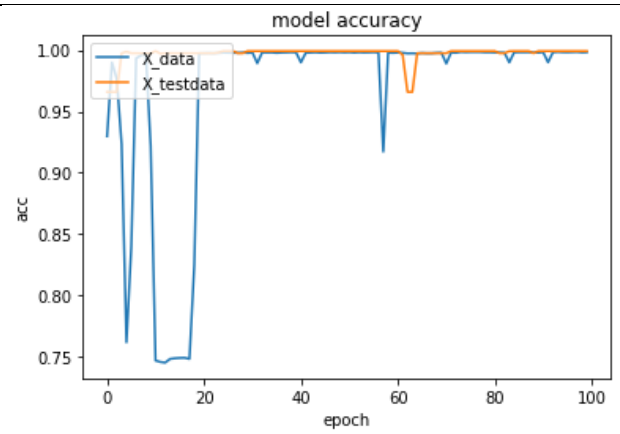


Figure 26.13 : CNN second model architecture training & test accuracy epochs for UNSW_NB15.

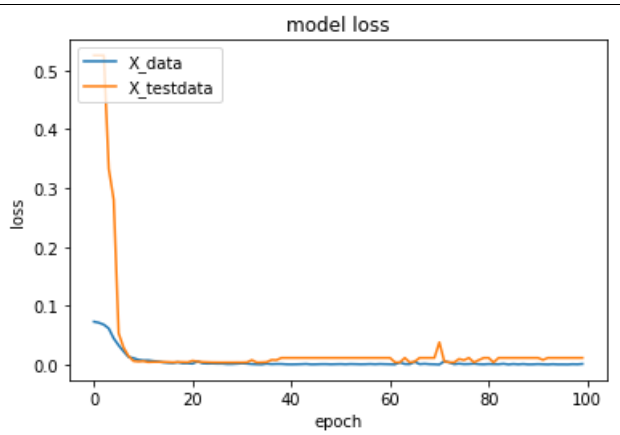


Figure 26.14 : CNN second model architecture training & test loss epochs for UNSW_NB15.

For the NSL-KDD dataset, we can notice that our loss values for the first model of the training set is in the range of 0.08-0.035 while the testing set is in the range of 0.03-0.015 **figure26.2**. On the other hand, the range of loss for the second model is between 0.9-0.15 for the training set and between 0.35-0.1 for the testing set **figure26.4**. And when we changed only the data processing methods while we kept the first model architecture the range of losses values we got were between 0.6-0.1 for the training set and between 0.24-0.08 for the testing set **figure26.6**.

For the same dataset, we find a validation accuracy of nearly 100% for each model, which indicates that our models have a good aptitude for generalization beyond the set of learning.

Then when we test our models on the test set. We see that the accuracy is unstable, but still remains in the range of 0.85 to 0.98 in **Figure26.1, 26.3, 26.5**.

Further, for the KDD_CUP99 dataset, we note that the accuracy rate for the first data processing methods is ranged between 0.85-0.98 for our training set and the accuracy rate for the testing set is between 0.82-0.96 **figure26.7**, whereas the accuracy rate for the second data processing methods is ranged between 0.79-0.97 for the training set and 0.83-0.974 for our testing set **figure26.9**.

On the other hand, for the UNSW_NB15 dataset the accuracy rate for the training set using the first model architecture is ranged between 0.995-0.998 and between 0.96-0.98 for the testing

set **figure26.11**, while the accuracy rate using the second model architecture is ranged between 0.93-0.98 for the training set and between 0.96-0.97 for the testing set **figure26.13**.

26.2 Our models comparison results:

The Comparison between our propose models Accuracies on three datasets in tables26.1, 26.2 and 26.3:

NSL_KDD		
First Model architecture accuracy	**Second model architecture accuracy**	**First Model arch accuracy with a different features analysis methods**
0.98	0.67	0.66

Table 26.1: accuracy results of NSL_KDD.

KDD_CUP99	
First Model architecture accuracy	**First Model arch accuracy with a different features analysis methods**
0.974	0.96

Table 26.2: accuracy results of KDD_CUP99.

UNSW_NB15	
First Model architecture accuracy	**Second model architecture accuracy**
0.98	0.97

Table 26.3: accuracy results of UNSW_NB15.

26.3 Evaluation Methods :

To evaluate our Deep-learning cyberattack detection model, we opted for the typical metrics used in deep learning as performance measures, namely: accuracy, precision, and recall.

- **Accuracy (ACC)** indicates the ratio of correct detection over total traffic trace:

$$ACC = \frac{1}{M}\sum_{i}^{M}\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}.$$

[85]

- **Precision (PPV)** shows how many attacks predicted are actual attacks. PPV is defined as the ratio of the number of TP records over the number of TP and FP records.

$$PPV = \frac{TP}{TP + FP}.$$

[86]

- **Recall (TPR)** shows the percentage of attacks that are correctly predicted versus all attacks happened. TPR is defined as the ratio of number of TP records divided by the number of TP and FN records.

$$TPR = \frac{TP}{TP + FN}.$$

[87]

After calculating the precision and recall of our model using the previous functions as shown in Table5.5, we present the TP,FN,FP,TN outcomes for each attack class DOS, Probe, R2L and Normal consecutively in Figures26.15, 26.16, 26.17 and 26.18 using our different two CNN architectures:

Methods	Precision	Recall
Random Forest	**71.21**	**70.99**
Linear SVM	**64.70**	**70.80**
K-means	**84.96**	**6.95**
Gaussian NB	**73.98**	**41.67**
Bernouli Naïve bayes	**87.47**	**36.49**
Logistic Regression	**62.04**	**73.79**
Decision Tree	**63.62**	**68.50**
RBMs	**81.95**	**77.48**
Our Proposed Model	**70**	**69.55**

Table26.4: Precision and Recall comparison.

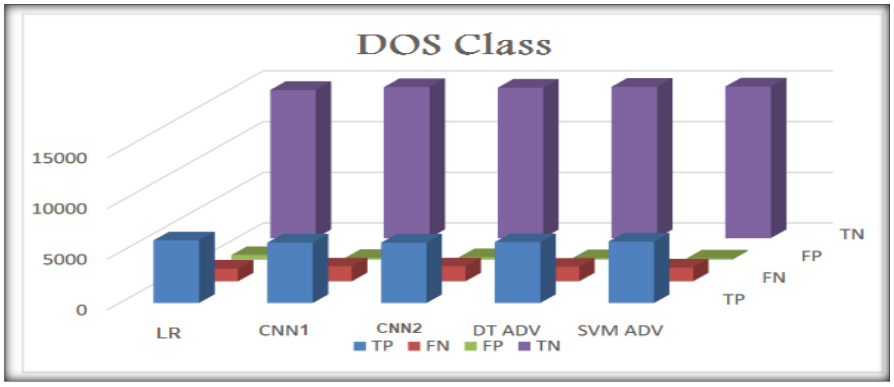


Figure 26.15: TP, FN, FP, TN of DOS Class.

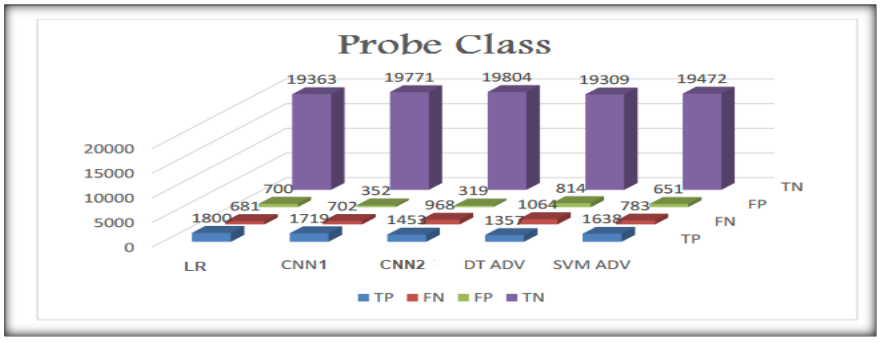


Figure 26.16: TP, FN, FP, TN of Probe Class.

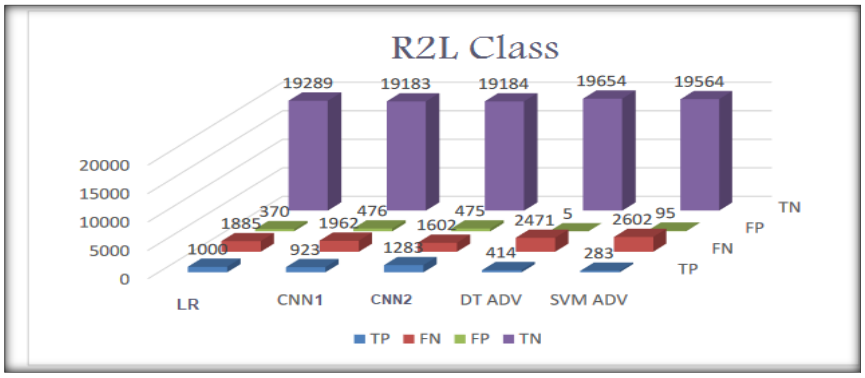


Figure 26.17: TP, FN, FP, TN of R2L Class.

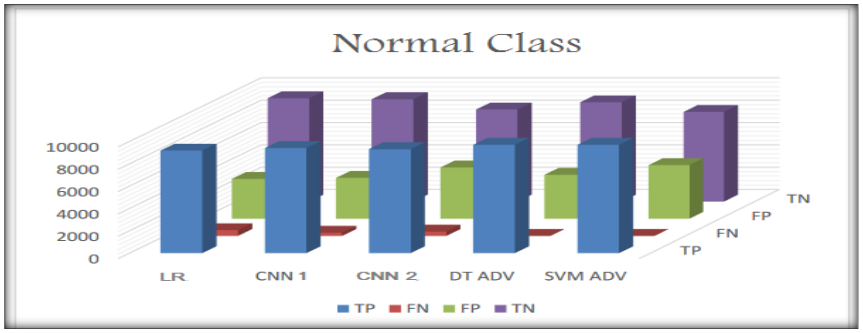


Figure 26.18: TP, FN, FP, TN of Normal Class.

26.4 The basic models used for the comparison:

A number of classifiers were trained and tested using the NSL-KDD, KDD_CUP99 and UNSW_NB15 datasets; we used those studies and some mentioned in the related works section as a baseline reference for our comparative study. The results are in Table 7.5.

	NSL_KDD	**KDD_Kup99**	**UNSW_NB15**
	Accuracy	Accuracy	Accuracy
Random Forest	88.39	97.02	94.44
Linear SVM	88.32	96.74	93.38
K-means	82.78	86.19	87.05
Gaussian NB	88.33	89.29	88.34
Bernouli Naïve bayes	74.60	90.94	91.31
Logistic Regression	89.52	96.2	92.52
Decision Tree	87.91	97.01	93.78
RNN	81.29	/	/
CNN	/	80	/
RBMs	90.99	97.11	93.38
Our Model with the First Architecture and Different Data analysis Method Architecture	66	96	/
Our Model with the First Architecture	98	97.4	98
Our Model with the Second Architecture	67	/	97

Tableau 26.5: Accuracy comparison with other models.

As we can see in the table above our model with the first architecture surpassed all the other models accuracies on the three datasets and proved that it's the best one that we can use to detect cyberattacks in the mobile cloud computing environment

26.5 The performance of the classification model

Our Classification results are presented in **Figures 26.19, 26.20, 26.21, 26.22.** For this we used a confusion matrix without normalization and a Normalized confusion matrix where each color signifies a different class (Normal, Probe, U2R, R2L and DOS).

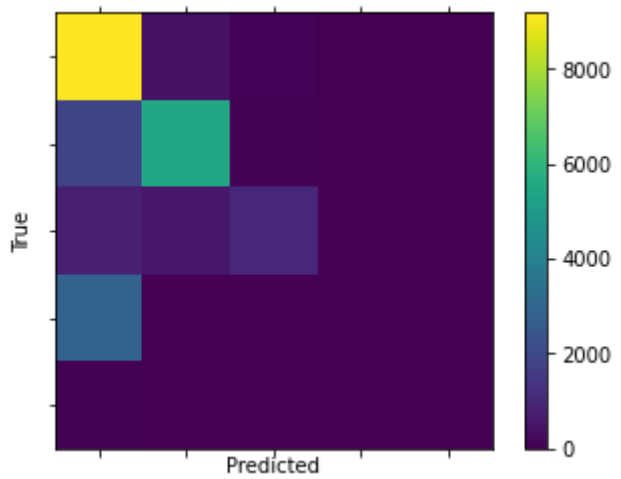


Figure 26.19: Confusion matrix without normalization for attack detection first model.

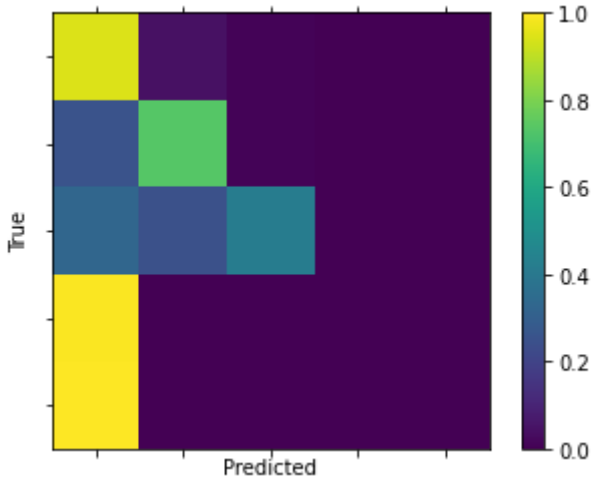


Figure 26.20: Normalized Confusion matrix for attack detection first model.

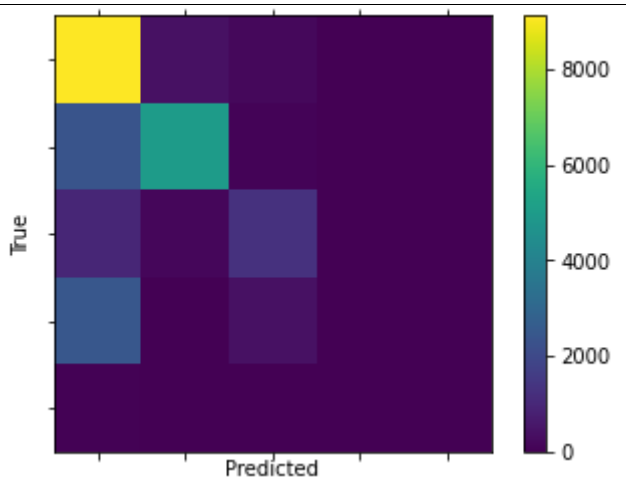


Figure 26.21: Confusion matrix without normalization for attack detection second model.

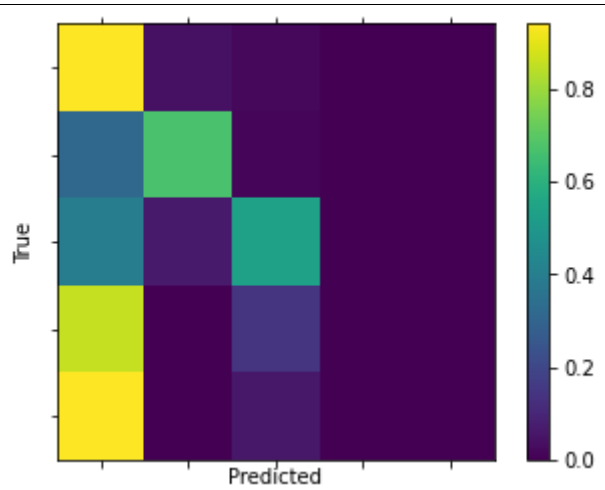


Figure 26.22: Normalized Confusion matrix for attack detection second model.

27. Conclusion

In this chapter we used two different methods to process our datasets NSL-KDD and KDD_CUP99 and we used as well two different neural networks architectures based CNN1D for both NSL_KDD and UNSW-NB15 datasets.

The datasets were divided into training, pre-testing and testing sets, these sets were then used to train our neural network to classify the data into different attack categories or into the normal category.

Development had many common issues. A large amount of data, lack of memory, and imbalance in class sizes were the most common issues, especially for the KDD_CUP99 dataset because of its big size and most of the development time has been devoted to solving these problems.

In the end, the final experimentation showed that our accuracies results of the various Deep Learning models developed were as good as or better and higher than those that could be obtained with the classic Machine Learning models and other deep learning models used with the same datasets and the same features.

General

Conclusion

General Conclusion

The future of artificial intelligence is promising thanks to deep learning.

In this work, we explored the field of deep learning used in the cyberattack detection systems and we chose to work with Keras library because it's easy to use and handle.

We used this library to train the three datasets NSL_KDD, KDD_CUP99 and UNSW_NB15, and for each dataset, we developed two different CNN 1D model architectures and two different methods to extract the datasets features, in order to, process them.
The trained models were used to classify the different types of attacks into five different categories (Normal, U2R, R2L, DoS and probe). Each model not only has a great generalization capacity for cyberattack detection, but also a high accuracy rate compared to other classification methods but one specific model architecture with which its datasets were processed with the new methods of tensorflow that google developed got good accuracy results on the NSL_KDD, KDD_CUP99 and UNSW_NB15 datasets respectively.

Finally, this work allowed us to put into practice our knowledge of neural networks.

Bibliography

28. Bibliography

[1] J. McCarthy, «utility-cloud computingflashback-to-1961-prof- john-mccarthy.,» *Informatique utilitaire : http ://computinginthe- cloud.wordpress.com/2008/09/25.,* Juin 2016.

[2] T. G. P. Mell, «The NIST Definition of Cloud Compu- ting, Recommendation of NIST. Special Publication 800-145,» *http ://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf,* 2011.

[3] S. B. A.-S. B. J. W. J-F Pépin, «Fondamentaux du Cloud computing- le point de vue des grandes entreprise. Réseau de Grandes Entreprises (CIGREF),» *http ://www.eurocloud.fr/doc/cigref2.pdf,* Mars 2013.

[4] N. Degroodt., « L'élasticité des bases de données sur le cloud computing. Mémoire de master en sciences informatiques,» *Université libre de bruxelles. Université d'Europe,* 2010.

[5] J. C. B. AZRIA, « L'impact du Cloud Computing dans les PME,» *mémoire, Ecole supérieure de génie informatique ESGI,* 2014.

[6] S.Lanani, «Une approche BPM (Business Process Managment) par composition d'ap- plications dans le cloud computing.,» *Mémoire de magister, Université Mohamed Khider de Biskra.,* 2015.

[7] «http ://whatiscloud.com/cloud_characteristics/multi_tenancy».

[8] X. Y. Z. D. Yuanshun, «Self-healing and Hybrid Diagnosis in Cloud Com- puting.,» Beijing, China, 2009.

[9] S. G. R. R. Buyya, «SLA-oriented resource provisioning for cloud computing : Challenges, architecture, and solutions,» chez *. Proceedings of the International Conference on Cloud and Service Computing (CSC)*, Hong Kong, China, 2011.

[10] I. Laribi, « La mise en place de la solution Openstack,» chez *Université de Tlemcen*, 2014.

[11] L. NOUMSI, «Etude et mise en place d'une solution "cloud computing " privée dans une entreprise moderne : cas de CAMTEL.,» chez *Ecole nationale supérieure des postes et télécommunications*, 2012.

[12] J. Baraban, «Private cloud, public cloud et hybrid cloud,http ://mysaas.fr/2010/10/04/private-Cloud-publique-Cloud-et-hybrid-Cloud/.,» chez *My saas*, 2010.

[13] A. Elwessabi, « Une approche basée agent mobile pour le cloud computing. Mé- moire de magister.,» chez *Université HADJ LAKHDAR - BATNA*, 2014.

[14] P. Codo., « Conception d'Une Solution de Cloud Computing Privé Basée sur un Algo- rithme de

Supervision Distribué : Application aux Services IAAS.,» chez *Ecole Polytechnique d'Abomey-Calavi (EPAC)*, 2012.

[15] J. T. J. M. R. B. J. M. L. B. a. D. F.Liu, « NIST SP 500-292, NIST Cloud Computing Reference Architecture ?,» chez *Leaf*, 2011.

[16] A.Prunier., « Le Cloud Computing : Réelle révolution ou simple évolution.,» 2011.

[17] H. Saouli., « Découverte de services web via le Cloud computing à base d'agents mobiles.,» chez *Thèse de doctorat. Université Mohamed Khider de Biskra.*, 2015.

[18] «http ://www.renaudvenet.com/cloud-computing-avantages-et-inconvenients-2011-01-26.html».

[19] A. M. K. Maioua, «Approche basée Agents Mobiles intelligents dans un environ- nement de cloud Computing.,» chez *Mémoire de master. Université Kasdi Merbah Ouargla*, 2014.

[20] «La sécurité informatique.Post BTS R2,» 2002.

[21] « Intrusion detection system in cloud computing environ- ment. Int. J. Cloud Computing,» 2012.

[22] B. D. T.K. Subramaniam, «Security attack issues and mitigation techniques in cloud computing environments.,» *International Journal of Ubicomp, http ://aircconline.com/iju/V7N1/7116iju01.pdf,* vol. 7, n° %11, Janvier 2016..

[23] D. E. P. H. T. A. M. Lonea, «Detecting ddos attacks in cloud computing environment.,» *International Journal of Computers,Communications & Control.,* vol. 8, n° %11, 2013.

[24] R. S. P. Chouhan, «Security attacks on cloud computing with possible solution. http://ijarcsse.com/docs/papers/Volume_6/01_January2016/V6I1-0140.pdf,» *International Journal of Advanced Research in Computer Science and Software Engineering.,* vol. 6, Janvier 2016 ISSN : 2277 128X..

[25] F. A.-H. K. S. M. H. Sqalli, « Edos-shield-a two-steps mitigation technique against edos attacks in cloud computing. in Utility and Cloud Computing (UCC).,» *2011 Fourth IEEE International Conference,* pp. 49-56, 2011.

[26] K. F. K. S. Qaisar, «Cloud computing : network/security threats and counter measures, http ://www.journal-archieves14.webs.com/1323-1329.pdf.,» *Interdisciplinary Journal of Contemporary Research in Business.,* vol. 3, n° %19, Janvier.

[27] S. V. K. Zunnurhain, «Security attacks and solutions in clouds.,» *In Proceedings of the 1st international conference on cloud computing. ,* pp. 145-156, 2010.

[28] B. Sevak., « Security against side channel attack in cloud computing.,» *International Jour- nal of Engineering and Advanced Technology (IJEAT),* vol. 2, n° %12, p. 183, 2013.

[29] Y. Q.Luo, «Algorithmic collision analysis for evaluating cryptographic systems and side channel attacks. In : Hardware-Oriented Security and Trust (HOST).,» *IEEE In- ternational Symposium .,* pp. 75-80, 2011.

[30] V. S.Subashini, «A survey on Security issues in service delivery models of Cloud Computing.,» *J Netw Comput Appl,* vol. 1, n° %11, p. 34, 2011.

[31] L. W.Li, «Trust model to enhance Security and interoperability of Cloud envi- ronment.,» *In : Proceedings of the 1st International conference on Cloud Computing. ,* pp. 69-79, 2009.

[32] R. K. M.K. Stine, «Guide for mapping types of information and information sys- tems to security categories.,» *National Institute of Standards and Technology.,* pp. NIST 800-60, 2008.

[33] J. Audenard., «Comprendre la protection des données dans le cloud.,» *Publication Orange Business Services,* 17 Mars 2011..

[34] L. L. M. J. N. G. J. Schwenk, «On technical security issues in cloud computing.,» *IEEE International Conference on Cloud Computing,* 2009.

[35] K. Q. W. C.Wang, «Ensuring data storage security in cloud compu- ting.,» *In the 17th IEEE InternationalWorkshop on Quality of Service (IWQoS'09).,* July.2009..

[36] S. K. A. Parakh, «Online data storage using implicit security.,» *Information Sciences.,* vol. 179, n° %13323-3331, 2009.

[37] M. G. P. D. R. S. F. Zhou, « Scheduler vulnerabilities and attacks in cloud computing.,» *College of Computer and Information Science Northeastern Uni- versity,* march.2011.

[38] X. Z. G. A. J. Wei, «Managing security of virtual machine images in a cloud environment.,» *Proceedings of the 2009 ACMworkshop on Cloud computing security.,* 2009.

[39] M. Q. J. L. T. G. D. T. S. M. L. B. R. H. F. Hu, « A review on cloud computing : design challenges in architecture and security,» *Journal of Computing and Information Technology - CIT 19 ,* p. 17, 2011.

[40] H. S. A. S. E.-E. E. M. Mohamed, «Enhanced data security model for cloud computing. In Informatics and Systems (INFOS).,» *8th International Conference ,* pp. CC-12, 2012.

[41] L. A. N. V. A. Irudayasamy, «Enhancing Data Security during Transit in Public Cloud.,» *International Journal of Engineering and Innovative Tech- nology (IJEIT). ,* vol. 3, July 2013.

[42] H. Z. D. Chen, « Data security and privacy protection issues in cloud computing,» pp. 647-651, 2012.

[43] G. S. S. S. S. K. S. Balakrishnan, «Introducing effective third party auditing (tpa) for data storage security in cloud.,» *International Journal of Com- puting and Technology,* pp. 397-400, June 2011..

[44] S. Z. W. E.-H. F. Aloul, « Two factor authentication using mobile phones.,» *In Computer Systems and Applications AICCSA,IEEE/ACS International Conference.,* pp. 641-644, 2009.

[45] F. S. J. D. L. R. M. D. T. G. D. Gelibert, «La sécurité et la Virtualisation.,» Mai 2012..

[46] S. S. S. Maninder, « Design and implementation of multi -tier authentication scheme in cloud,» *International Journal of Computer Science Issues. ,* vol. 9, n° %12.

[47] G. A. K. Satish, «Multi-authentication for cloud security : a framework.,» *Interna- tional Journal of Computer Computing and Engineering Technology (IJCSET).,* vol. 5, n° %14, Apr. 2014..

[48] B. S. A. S.Ezhil, *Privacy-preserving public Auditing in cloud using HMAC International Journal of Recent Technology and Engineering (IJRTE) ISSN : 2277-3878,,* vol. 2, March 2013..

[49] S. S. P. Kalpana, « Data security in cloud computing using RSA algorithm.,» *International Journal of Research in Computer and Communication technology (IJRCCT). ISSN 2278-5841,* vol. 1, pp. 143-146, September 2012.

[50] C. I. P. Pankaj, « A secure data transfer technique for cloud computing.,» *THA- PAR UNIVERSITY,* August 2014..

[51] K. L. a. M. M. U. Somani, «Implementing digital signature with RSA encryption algorithm to enhance the data security of cloud in cloud computing,» *In- ternational Conference on Parallel Distributed and Grid Computing (PDGC)..*

[52] S. M. V. R. Balasaraswathi, «Enhanced security for multi-cloud storage using cryptographic data splitting with dynamic approach. In Advanced Commu- nication Control and Computing Technologies(ICACCCT).,» *International Conference,* pp. 1190-1194.

[53] W. Koehrsen, «https://towardsdatascience.com/automated-machine-learning-hyperparameter-tuning-in- python-dfda59b72f8a,» Jul 3, 2018, /21/02/19.

[54] «https://R%C3%A9seau_neuronal_convolutif/16/02/19».

[55] «https://machinelearningmastery.com/gradient-descent-for-machine-learning/ 23/02/19».

[56] «REVIEW/ Deep learning/ Yann LeCun 1,2, Yoshua Bengio3 & Geoffrey Hinton4,5 / 28 MAY 2015 | VOL 521 | NATURE /doi:10.1038/nature14539/25/02/19».

[57] «https://towardsdatascience.com/convolutional-neural-networks-the-biologically- inspired-model-f2d23a301f71 28/02/19».

[58] «[12] : http://www.natural-solutions.eu/blog/la-reconnaissance-dimage-avec-les-rseaux-de-neurones-convolutifs /01/03/19».

[59] «[13] : https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches-dun-cnn /03/03/19».

[60] «J. Goodfellow, Yoshua Bengio et Aaron Courville, Deep Learning, MIT Press, 2016 (ISBN 0262035618), chapitre 7./06/03/19».

[61] «Regularization of Neural Networks using DropConnect | ICML 2013 | JMLR W&CP 07/03/19».

[62] A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke et J. (. Schmidhuber, «A Novel Connectionist System for Improved Unconstrained Handwriting Recognition,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* p. 855– 868.

[63] M. (. 2. Miljanovic, « "Comparative analysis of Recurrent and Finite Impulse Response Neural Networks in Time Series Prediction",» *Indian Journal of Computer and Engineering.,* /09/03/19.

[64] «https://towards/Recurrent_neural_network /11/03/19».

[65] «https://www.analyticsindiamag.com/overview-of-recurrent-neural-networks-and-their-applications/ 18/03/19».

[66] C. L. D. N. a. P. W. D. T. Hoang, «A survey of mobilecloud computing: Architecture, applications, and approaches,» *WirelessCommunications and Mobile Computing,* vol. 13, n° %118, pp. 1587-1611, 2013.

[67] A. B. V. K. [1]V. Chandola, «Anomaly detection: a survey,» *ACMComput. Surv,* 2009.

[68] F. P. A. C. A. D. S. U. Fiore, «Network anomaly detection with the restricted boltzmann machine,» *Neurocomputing,* n° %1122, pp. 13-23, 2013.

[69] Y. B. H. Larochelle, «Classification using Discriminative Restricted Boltz-mann Machines,» chez *Proceedings of the Twenty-Fifth International Conferenceon Machine Learning, ACM*, 2008.

[70] A. A. ,. M. ,. S. Mohd Nazri Ismail, «Detecting Flooding based DoS Attack in Cloud Computing Environment using Covariance Matrix Approach,» 2014.

[71] X. W. Bo Dong, «Comparison Deep Learning Method to Traditional Methods Using for Network Intrusion Detection,» chez *IEEE International Conference on Communication S oftw are and N etw ork s*, 2016.

[72] Y. Z. F. A. X. H. CHUANLONG YIN, «A Deep Learning Approach for IntrusionDetection Using Recurrent Neural Networks,» *IEEE ,* 2017.

[73] D. T. H. D. N. P. W. D. N. a. E. D. Khoi Khac Nguyen, «Cyberattack Detection in Mobile Cloud Computing:A Deep Learning Approach,» April 2018.

[74] I. T. Jolliffe, «Principal Component Analysis and Factor Analysis,» chez *Springer* , New York, 1986.

[75] J. K. H. K. M. S. a. E. C. Jiyeon Kim, «CNN-Based Network Intrusion Detectionagainst Denial-of-Service Attacks,» *MDPI,* 1 june 2020.

[76] E. B. W. L. a. A. G. M. Tavallaee, «A Detailed Analysis of the KDD'99 CUP Data Set,» chez *The 2nd IEEE Symposium on Computational Intelligence Conference for Security and Defense Applications (CISDA)*, 2009.

[77] W. F. W. L. A. P. a. P. K. C. S. J. Stolfo, «Cost-based modeling for fraud and intrusion detection: Results from thejam project,» *discex,* vol. 02, p. 1130, 2000.

[78] J. F. Nieves, «Data Clustering for Anomaly Detection in Network Intrusion Detection,» chez *Research Alliance in Math and Science*, 2009 August 14.

[79] E. B. W. L. a. A. A. G. Mahbod Tavallaee, «A Detailed Analysis of the KDD CUP 99Data Set,» chez *Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications*, 2009.

[80] A. Shaheen, «A comparative analysis of intelligent techniques for detecting anomalous internet traffic,» chez *MSc. Thesis, King Fahd University*, 2010.

[81] M. M. K. Dr.K.Arunesh, «A Comparative Study Of Classification Techniques For Intrusion Detection Using Nsl-Kdd Data Sets,» *in International Conference on Recent Trends in Engineering Science, Humanity and Management,* pp. 288-295, 2017.

[82] N. M. a. J. Slay, «The evaluation of networkanomaly detection systems: Statistical analysis ofthe UNSW-NB15 data set and the comparison withthe KDD99 data se,» *nformation Security Journal:A Global Perspective,* pp. 18-31, 2016.

[83] V. R. B. J. S. Yasir Hamid, «Benchmark Datasets for Network IntrusionDetection: A Review,» *International Journal of Network Security,* p. 7, 2018.

[84] S. O. a. Y.-W. T. G. E. Hinton, «A fast learning algorithmfor deep belief nets,» *Neural computation,* vol. 18, n° %17, pp. 1527-15545, 2006.

[85] G. E. H. a. R. R. Salakhutdinov, «Reducing the dimensionalityof data with neural networks,» *Science,* vol. 313, n° %15786, pp. 504-507, 2006.

[86] P. L. D. P. a. H. L. Y. Bengio, «Greedy layer-wise training of deep networks,» *Advances in neural information process-ing systems,* pp. 153-160, 2007.

[87] G. E. Hinton, «Training products of experts by minimizing contrastivedivergence,» *Neural computation,* vol. 14, n° %18, pp. 1771-1800, 2002.

[88] B. S. A. S.Ezhil, «Privacy-preserving public Auditing in cloud using HMAC International Journal of Recent Technology and Engineering (IJRTE) ISSN : 2277-3878.,» vol. 2, March 2013..