

WEB SECURITY IN IoT NETWORKS USING DEEP LEARNING MODEL

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF DEGREE
OF

MASTER OF TECHNOLOGY
IN
SOFTWARE ENGINEERING

Submitted by:

INDERPREET SINGH BAINS

2K18/SWE/19

Under the supervision of

Mr. Sanjay Patidar
(Assistant Professor)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JUNE, 2020

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi - 110042

CANDIDATE'S DECLARATION

I, Inderpreet Singh Bains, Roll No. 2K18/SWE/19 student of M.Tech (Software Engineering), hereby declare that the project Dissertation titled “**WEB SECURITY IN IoT NETWORKS USING DEEP LEARNING MODEL**” which is submitted by me to the Department of Computer Science and Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any degree, Diploma Associateship, Fellowship or other similar title or recognition.

A rectangular box containing a handwritten signature in black ink that reads "I. P. S. Bains". The signature is written in a cursive style and is underlined with a single horizontal line.

Place: Delhi

INDERPREET SINGH BAINS

Date: 29-6-2020

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi - 110042

CERTIFICATE

I hereby certify that the Project Dissertation titled “**WEB SECURITY IN IoT NETWORKS USING DEEP LEARNING MODEL**” which is submitted by Inderpreet Singh Bains, 2K18/SWE/19 Department of Computer Science Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has never been submitted in part or full for any Degree or Diploma to this University.



Place: Delhi

Mr. Sanjay Patidar

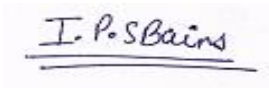
Date: 29-6-2020

SUPERVISOR

Assistant Professor

ACKNOWLEDGMENT

I express my gratitude to my major project guide Mr. Sanjay Patidar, Assistant Professor, Department of Computer Science Engineering, Delhi Technological University, for the valuable support and guidance he provided in making this major project. It is my pleasure to record my sincere thanks to my respected guide for his constructive criticism and insight without which the project would not have shaped as it has. I humbly extend my words of gratitude to other faculty members of this department for providing their valuable help and time whenever it was required.

A handwritten signature in black ink on a light-colored background. The signature reads "I. P. S Bains" and is underlined with a double horizontal line.

Inderpreet Singh Bains

Roll No. 2K18/SWE/19

M.Tech (Software Engineering)

E-mail: inder.rockstar07@gmail.com

ABSTRACT

The vision of IoT is to interface day by day utilized items (which have the capacity of detecting and activation) to the Internet. This may or might possibly include human. IoT field is as yet developing and has many open issues. We develop on the digital security issues. The Web of things (IoT) is still in its beginning phases and has pulled in much enthusiasm for some mechanical parts including clinical fields, coordination's following, savvy urban communities and autos. Anyway, as a paradigm, it is defenseless to a scope of significant intrusion threats. In IoT whenever there is a web attack then we need to remove the attack by installing software so by using these models we can remove the attack from the system. It presents a threat investigation of the IoT and uses an Artificial Neural Network (ANN) to battle these threats. In this, profound learning method for digital security and prevention of attacks is used in which a convolution 1d with multiple convolutions is used to increase the accuracy of the user. We have proposed profound models of learning and assessed those utilizing most recent CICIDS2017 datasets for DDoS assault recognition which has given most noteworthy precision as 99.38%. It is essential to create an effective intrusion discovery framework which uses deep learning mechanism to overcome attack issues in IOT framework. In this, a CNN i.e convolutional neural system is developed with various convolution layers and accuracy of attack detection is increased.

CONTENTS

Candidate's Declaration	i
Certificate	ii
Acknowledgment	iii
Abstract	iv
Contents	v
List of Figures	vii
List of Tables	ix
List of Abbreviations	x
CHAPTER 1 INTRODUCTION	1
1.1 Context	1
1.2 Security Problems with the Internet of Things	2
1.3 Challenges of IoT	2
1.4 IoT Hazards and Risks	3
1.5 Deep Learning	4
1.6 Contributions	6
CHAPTER 2 BACKGROUND WORK	7

CHAPTER 3 CYBERSECURITY	10
3.1 Significance of Cybersecurity	10
3.2 Challenges to Cybersecurity	10
3.3 Managing Cybersecurity	11
CHAPTER 4 DEEP LEARNING MODELS	12
CHAPTER 5 LITERATURE REVIEW	16
CHAPTER 6 PROPOSED WORK	18
CHAPTER 7 EVALUATION AND CONCLUSION	25
7.1 Dataset & Environment	25
7.2 Performance Metrics	26
7.3 Results	26
7.4 Conclusion & Future Works	31
REFERENCES	32
LIST OF PUBLICATIONS OF THE CANDIDATE’S WORK	36

LIST OF FIGURES

Figure 1. Fog-to-Node engineering for IoT system	2
Figure 2. IoT architecture	3
Figure 3. The design of profound learning model	4
Figure 4. CNN methodology	5
Figure 5. Biological Neuron vs. Artificial Neuron	7
Figure 6. Step Function	8
Figure 7. Sigmoid Function	8
Figure 8. Tanh Function	8
Figure 9. Relu Function	9
Figure 10. Architecture of CNN	12
Figure 11. Recurring Neural System Architecture	13
Figure 12. Structure of LSTM memory cell	14
Figure 13. Structure of an Autoencoder network	14
Figure 14. Multilayer Perceptron	15
Figure 15. Proposed Methodology	19

Figure 16. Proposed CNN3 Layer model	21
Figure 17. Proposed Multiheaded CNN Layer model	22
Figure 18. 1D CNN architecture	23
Figure 19. Confusion matrix of CNN model	27
Figure 20. Confusion matrix of Proposed CNN 3 Layer model	27
Figure 21. Confusion matrix of Proposed Multiheaded CNN layer model	28
Figure 22. Epochs vs Accuracy and Epochs vs Loss curve of CNN model	29
Figure 23. Epochs vs Accuracy and Epochs vs Loss curve of Proposed CNN 3 layer model	29
Figure 24. Epochs vs Accuracy and Epochs vs Loss curve of Proposed Multiheaded CNN layer model	30
Figure 25. Comparison of Evaluation parameters of all models	30

List of Tables

Table 1: Comparing ANN with BNN	7
Table 2: Performance Metrics Evaluation Table	28

List of Abbreviations

1. IoT: Internet of Things
2. CPS: Cyber-Physical System
3. DDoS: Distributed Denial of Service
4. RFID: Radio Frequency Identification
5. M2M: Machine to Machine
6. M2H: Machine to Human
7. IDS: Intrusion Detection Systems
8. NN: Neural Networks
9. CNN: Convolutional Neural Network
10. RNN: Recurrent Neural Network
11. MLP: Multilayer Perceptron
12. LSTM: Long Short Term Memory
13. SGD: Stochastic Gradient Descent
14. MSE: Mean Square Error
15. FP: False Positive
16. TP: True Positive
17. FN: False Negative
18. TN: True Negative

CHAPTER 1: INTRODUCTION

1.1 Context

We live during a time that is fuelled by data. Data has found every single point of our being. Many people believe this phenomenon is the product of an industry that, like all other technological revolutions, makes our lives quicker than before [1]. Industry allowed computers and physical systems to cooperate. This relationship is known as cyber-physical frameworks (CPS). Through embedding sensors, controls, and actuators, the physical device then generated the web of things (IoT). The result of this transition is the huge information that is created, and that must be handled, called Big Data.

To research these technologies more easily, there are several instruments to recreate the IoT climate, and ultimately Big Data. Cooja [2], GNS-3 [3], Iotify [4], and MATLAB [5] are the most common simulators to use. As the amount of data produced has increased, the term data protection has become a critical term explicitly for the protection of delicate information as per the three guidelines of information security (Confidentiality, Transparency, and Accessibility). There are several threats to IoT (car hacking, DDoS, or physical attacks), due to the absence of directing conventions. Insights say that in 2017 DDoS assaults expanded by 91 percent due to the abuse of IoT contraptions [6]. IoT, which is within all life structures, is frail against different forms of assault. Yet there's additionally no viable way to protect our lives from these attacks. Nowadays, AI is the most common recognized subject for distinguishing IoT security digital assaults since ML-based approaches can provide a vigorous mechanism for concealed assaults.

The Internet of Things is the most exciting technology emerging that links all around the globe via the web. IoT innovation guarantees that our personal, professional, and societal lives will be enhanced and helped [7]. IoT consists of a system of savvy objects across the world over the web with no human intervention, which is awesome, yet like some other system, it is vulnerable to digital assaults. DDoS is one of the biggest cyberattacks in the recent past that has plagued the IoT network and brought about significant losses. In DDoS assault, a programmer utilizes a variety of hosts to overload the objective server, resulting in a total network crash, thereby preventing legitimate users from accessing the server network service. The denial of service attack is expected to hit 17 million by 2020, according to the information given in [8].

The list of IoT areas is a smart house, and its apps, remote sensors, brilliant locks, shrewd meters, wearable gadgets, surveillance cameras, savvy plugs, Radio Frequency Identification (RFID), Machine to Machine (M2M), and Machine to Human (M2H) gadgets, etc. Consumption principles and territories have shown that the IoT brush (brushes) every single point in human life.

An important technique for detecting cyberattacks in any network is the intrusion detection system (IDS). A considerable lot of the new IDS depend on network-based AI algorithms to train and detect cyberattacks. Fog computing is an improved augmentation of brought together distributed computing wherein disseminated mist hubs are nearer to IoT arrange articles and address adaptability congestion, high data transmission usage nature of administration humbling, and minimal cloud computing high inertness. Fog-to-node registering is suitable for IoT

networks being deployed in operation and being efficient. The figure below illustrates the haze to-hub model architecture with appropriated equal calculation giving insight to the disseminated hazes by giving IDS closer to the IoT organize objects computation, control, and storage. Compared with the cloud, IDS detect cyber threats easily and rapidly at fog nodes. IoT network consists of links between various types of savvy objects going from supercomputers to minuscule gadgets, which may have low processing capacity, so it is difficult to access these types of networks. Therefore, cybersecurity is a significant shortcoming in IoT network implementation [9].

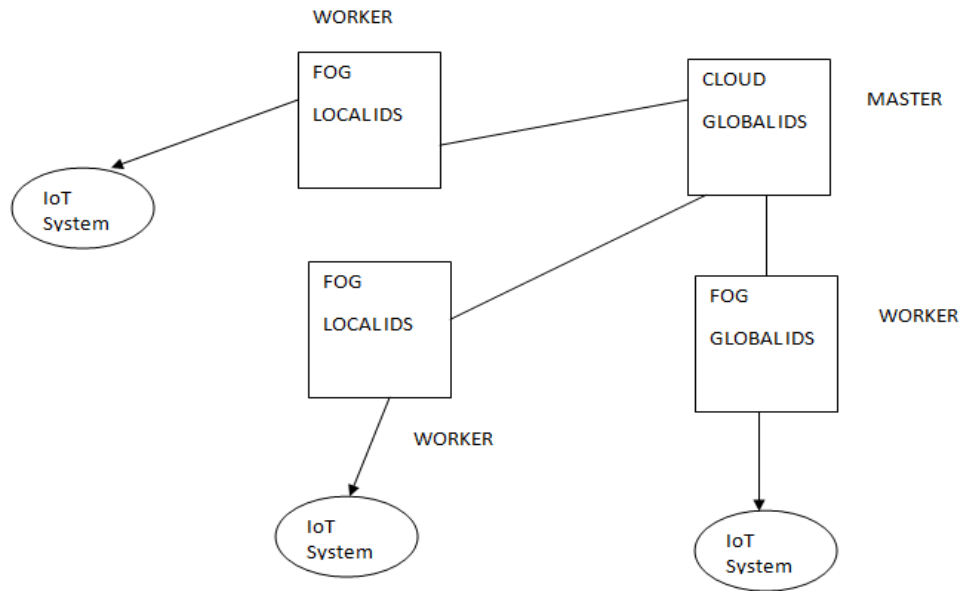


Fig.1. Fog-to-Node engineering for IoT system

1.2 Security Problems with the Internet of Things

IoT is at danger on account of its heterogeneous structure, which allows cyber domain and physical domain cooperation. The list includes deficient approval, inconsistent framework infrastructure, and nonappearance of transport encryption and check of integrity, security issues, unsafe programming or firmware, poor physical protection. Additionally, the IoT vulnerabilities may be inadequate routing protocols [10].

The biggest DDoS attack was carried out with the use of IoT botnets in October 2016. In particular, PayPal, The Guardian, Netflix, Reddit, and CNN transformed into the objective. The botnets were developed using a malware called Mirai. This malware abused the security powerlessness of the login data provided by the IoT system. The devices used were guided to goals. The Mirai attack was triggered by the use of default username, secret key, non-interesting passwords, and the absence of device and firmware refreshes [11].

1.3 Challenges of IoT

IoT tools have information collection, storage, and handling capabilities in savvy applications

[12]. These types of data spread to many fields, including safety, transportation, military, etc. Protection of sensitive data, which poses the greatest danger to IoT, comes to the fore. That risk is rooted in two major vulnerabilities. First, heterogeneous devices and interoperable interfaces make it more difficult to handle IoT systems. Second, various devices have limited resources, nonappearance of computational capacity, low inertness. This additionally makes it difficult to detect probable and unknown attacks on IoT devices [13]. These problems can be in network security, as it is possible to hack IoT and put the security at stake. Private IoT details can be hacked, a risk that is another. The challenges can also be in networking because the inability to connect may pose a major challenge due to billions of devices on a centralized server. Another challenge in the field of IoT is the compatibility issues that require the deployment of extra hardware and software.

1.4 IoT Hazards and Risks

Getting perils and dangers of IoT is a prologue to understanding the assaults on IoT. For this, the engineering of IoT is studied.

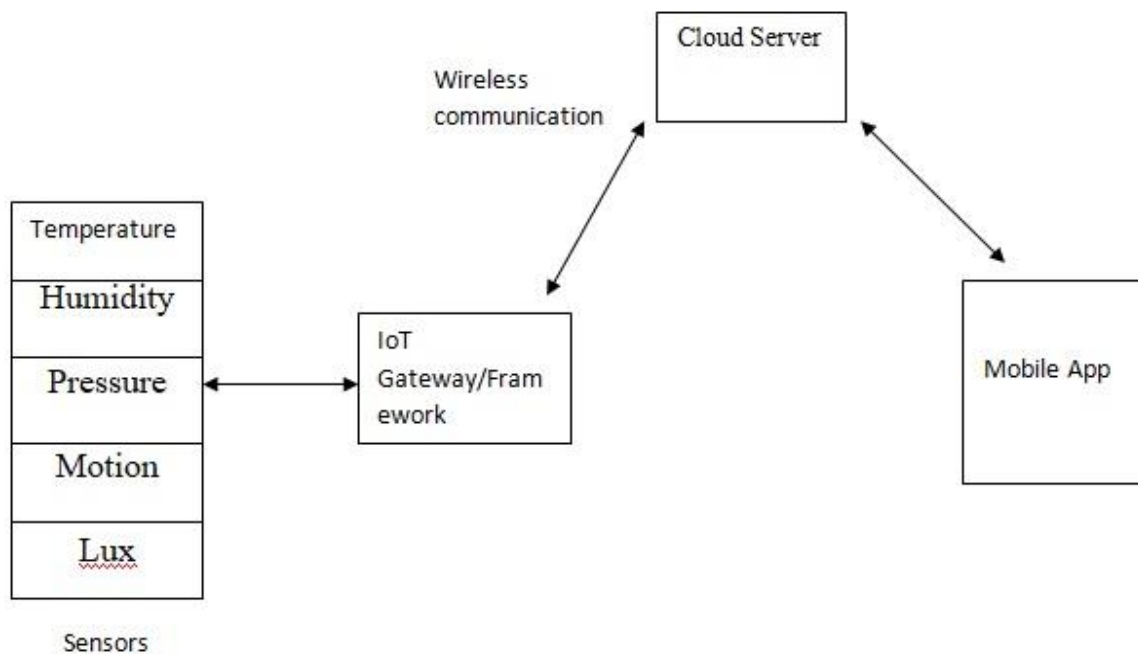


Fig.2. IoT Architecture

IPv6 as an enabler is an essential source of IoT, as IPv4 can't support the scale of IoT frameworks. IPv6 protection recommendations and contemplations form the premise of IoT security [14]. IoT experiences indistinguishable dangers from IPv4. In addition, IoT is the target of threats because of its position at an intersection point between the digital area and the physical space. There are several forms of IoT attacks that include physical attacks, intrusion attacks, DoS attacks, access attacks, privacy attacks, cybercrimes, disruptive attacks, and SCADA [15]. Routing attacks are carried out on the network layer, which is more important than other attacks; they can be updated to IoT for the rest of the attacks.

1.5 Deep Learning

Significant learning is such a training Neural Networks (NN) and has the plan of the neural network. The contrast between old school neural networks and profound learning is that there are many concealed layers of deep learning. Deep learning additionally learns the highlights themselves, making the learning procedure progressively exact and besides proving to be more effective and precise than shallow learning [16].

Profound learning is a more extensive AI sub-field, which is a bigger profound neural system that can be used for managed, unaided, and semi-administered learning. The idea of profound learning was first introduced in [17] Based on a system of profound beliefs and shown to be successful in fields, for example, picture handling, regular language processing, and self-driving vehicle, and so on. One downside of profound learning is the lengthy preparing period it takes, the bigger the preparation information, the more prominent the preparation time, however in order to perform well, profound learning techniques need enormous information for preparing.

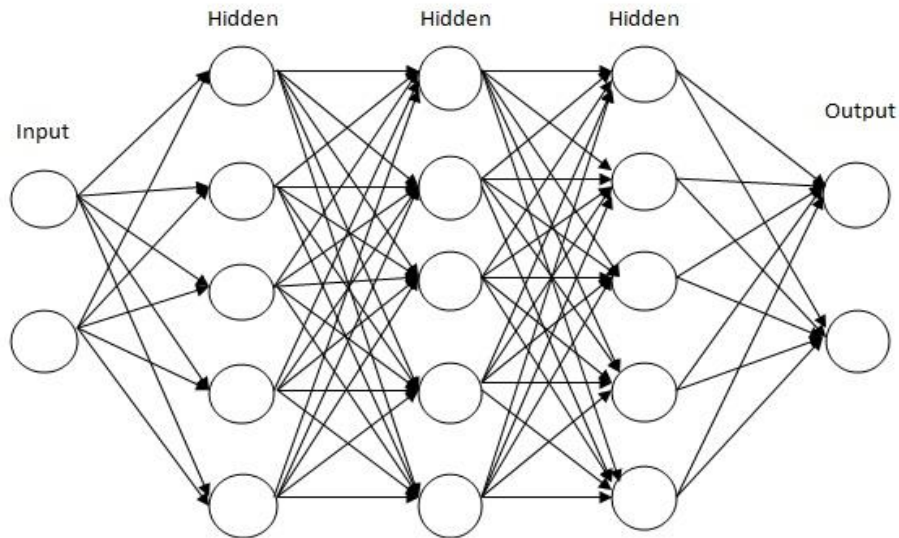


Fig.3.The design of profound learning model

Fig.3 demonstrates the fundamental design of the profound learning model. It consists of one info layer, trailed by various shrouded layers that took care of the yield layer further. CNN (Convolution Neural Network) is a type of profound learning used in PC image processing [18, 19] and language preparing [20]. Without pre-processing, a crude picture is taken care of directly to the CNN model; it then evaluates the features through convolution activities [21]. RNN (Recurrent Neural Network) is another form of profound learning model that has made promising ground in fields, for example, the processing of regular languages [22] and text preparing [23]. LSTM (Long Short Term Memory) organize is an RNN development capable of learning designs in successions; this can be utilized to identify information as an assault and as natural. One of the benefits of LSTM is that it very well may be used legitimately on crude information without using any form of a determination of features.

The main objectives are to implement deep learning method with higher accuracy in cyber security to compare the accuracy with existing methods. Using the convolutional operation on the input datasets with pre processing and special layers and filters are applied in CNN method. Fig.4 shows the base model representation of the CNN based model. The input data sets are fed using their features. Then convolution operation is applied. Maxpool, dropout and fully connected layer parts of CNN methodology is applied for the final output through dense layers.

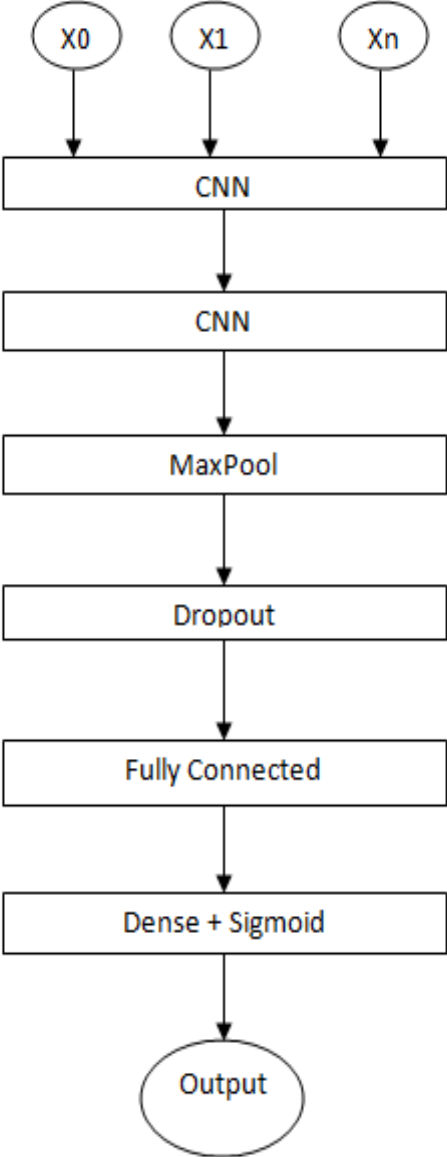


Fig.4 CNN methodology

1.6 Contributions

We used the CICIDS2017 datasets in this, which contain benign and most cutting-edge regular assaults that take after true information. It likewise includes the system traffic test result using CICFlowMeter with labeled streams subject to timestamp, source, & objective IPs, source & objective ports, shows, and ambush. The highest accuracy is 99.38 percent, which is likewise contrasted to machine learning calculations with the proposed models. In this we have proposed 2 models which give better accuracy as compared to the base model as we have used 1d CNN and added more layers to make it deeper by adding more convolutional layers, as well as maxpooling layers. Max pooling layer is included to dispose of highlights with low score and keep just highlights with most elevated score.

CHAPTER 2: BACKGROUND WORK

Neurons are the principal actor in the segment of learning. With the relation weights, they take one or more inputs from the previous neurons, summarize them and put in the actuation work, and generate a yield that is or isn't essentially fired. The figure demonstrates the mathematical representation of the neurons. After this addition, Y is placed into the process by the activation function.

$$Y = \sum (\text{input}) * (\text{weight}) + \text{bias}$$

$$\text{Output} = f(Y)$$

'Fire' signifies activating; the name is enlivened by the brain's biochemical workings. The figure depicts the similarity between neurons and artificial neurons. For neurons, dendrites transmit electrical signs from different cells to the phone body and then forward feedback signs to other brain cells along the axon. There is a common cycle with artificial neurons.

Table 1: Comparing ANN with BNN

Biological Neural Network	Artificial Neural Network
Soma	Node
Dendrites	Input
Synapse	Weights or Interconnections
Axon	Output

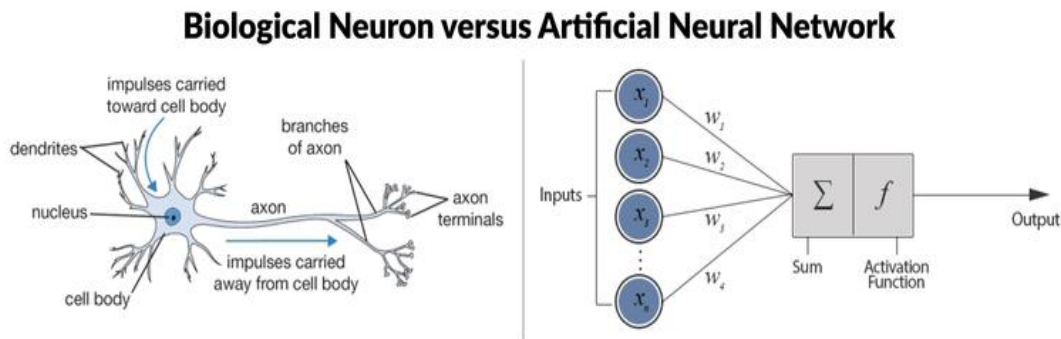


Fig.5. Biological Neuron vs. Artificial Neuron

Types of activation functions:

1. **Step Function-** The step function is activation, which is a neural network decision-making unit. This measures a neural node's net output. The progression work is an enactment work that takes Y; if Y reaches a given worth (or limit), the step function yield is initiated; something else, the yield isn't actuated. The graph for step work is shown in Figure 7. x is the threshold.

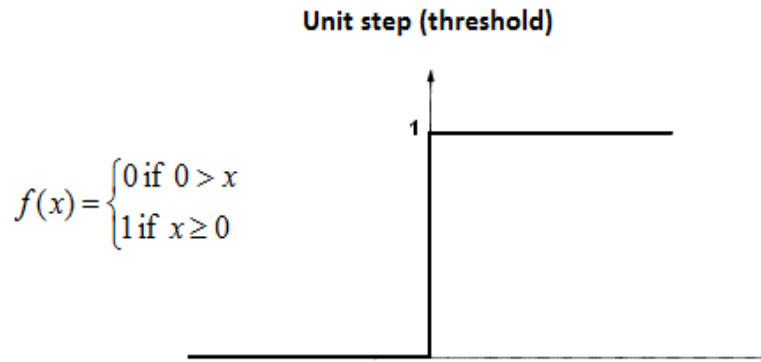


Fig.6. Step Function

2. **Sigmoid Function**- Sigmoid functions smoothly and is differentiable continuously. Its scale is from 0 to 1. It is a curve in the S-shaped form. The benefits are the non-linearity and the effects of better classification.

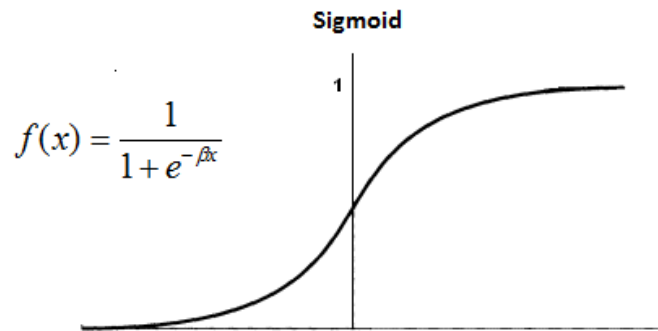


Fig.7. Sigmoid Function

3. **Hyperbolic Tangent function**- With sigmoid function, the Tanh function has a very similar structure. It has limits (-1, 1).

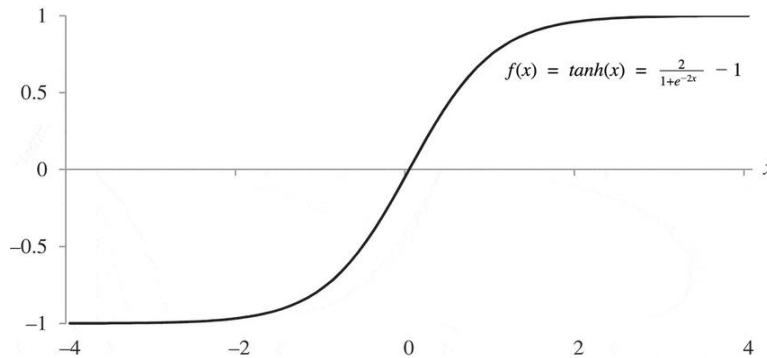


Fig.8. Tanh Function

4. **ReLU**- Rectified Linear Units are the function that is the most powerful solution to the problem of gradient loss. It has been very much favored in the last couple of years. It has a six times increase in Tanh function convergence. These days virtually all deep learning models are using ReLU. Its limitation is that a Neural Network Model can only be used inside hidden sheets.

$$y = f(x) = \text{maximum}(0, x)$$

The yield is zero for the under zero information, while the yield is equal to the more than zero input. The ReLU work is better suitable for double characterization, and we use it as the initiation function in the hidden layers.

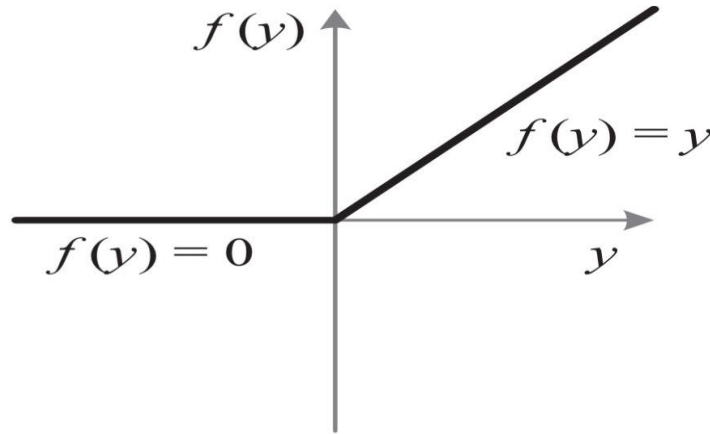


Fig.9.ReLu Function

5. **Softmax**- Softmax work is frequently portrayed as a blend of different sigmoids. We realize that sigmoid returns esteems somewhere in the range of 0 and 1, which can be treated as probabilities of information direct having a place toward a specific class. Along these lines sigmoid is broadly utilized for twofold characterization issues. The softmax capacity can be utilized for multiclass order issues. This capacity restores the likelihood for a datapoint having a place with every individual class. Here is the numerical articulation of the equivalent.

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K.$$

Softmax capacities can deal with various classes when contrasted with other actuation capacities. They are additionally utilized distinctly for the yield layer, for neural systems that need to arrange contributions to various classifications.

CHAPTER 3: CYBER SECURITY

Cybersecurity [24] insinuates the collection of advancements, methods, and practices planned to secure frameworks, computers, services, and information from attacks, lawsuits, or uncertified access. Cybersecurity is otherwise called security in the field of information technology. Application protection, information security, disaster recovery, and network security are the main areas addressed in cybersecurity. Application security defends applications from threats that can result from application design, development, implementation, update, or maintenance flaws. The protection of information protects the data from uncertified access to identify the threat and protect privacy. Disaster recovery arrangement is a mechanism that involves conducting risk evaluation, setting goals, establishing plans for recovery in the event of a calamity. Network security requires activities designed to protect network accessibility, reliability, honesty, and protection.

3.1 Significance of Cyber Security

Cybersecurity is noteworthy as government, military, business, monetary, and clinical firms assemble, strategy, and store extraordinary measures of information on PCs and various gadgets. Enormous amounts of information might be confidential data, including licensed innovation, monetary information, personal data, or other information that may have adverse consequences for non-official access or submission. When doing business, an organization transmits confidential information through frameworks and to various devices, and cybersecurity characterizes the way to protect the information and frameworks used to process or store this information. As the recurrence and multifaceted nature of computerized ambushes assembles, organizations and firms need to make a transition to ensure their secret business and individual information, particularly those blamed for protecting data identified with national security, prosperity, or financial records. Computerized ambushes and propelled observation are the highest threats to national security, which also surpass psychological oppression.

3.2 Challenges to Cyber Security

For convincing cybersecurity, an affiliation needs to harmonize its undertakings every single through its entire information structure. Parts of cybersecurity are:

- Network security
- Application security
- Endpoint security
- Data security
- Identity Management
- Database and establishment security
- Cloud security
- Mobile security
- Disaster recovery/ business movement arranging

- End-customer guidance

In cybersecurity, the most daunting test is the ever-developing existence of security dangers. Companies and governments have historically based much of their cybersecurity assets on external protection to secure even their most basic network parts and to ensure against known alerts. This strategy is inadequate today, as the dangers are evolving and shifting faster than companies can keep up. Consequently, advisory bodies advocate more constructive and adaptive approaches to cybersecurity.

3.3 Managing Cyber Security

Cybersecurity management can be portrayed as doing everything an association can do to ensure its data frameworks and PC systems from digital threats, interruptions, malware, and different forms of reality penetrates. There are five stages to oversee cybersecurity threats:

- Identify and assess dangers
- Assess capacity to bear the hazard
- Develop and actualize chance decrease measures
- Implement, screen and update
- Disclose dangers and methodologies

CHAPTER 4: DEEP LEARNING MODELS

1. Convolutionary Neural Networks (CNN)

The Convolutionary Neural System is a complex neurological feed-forward system which makes use of perceptrons for supervised learning and knowledge review. This is often used for visual data, such as the categorization of images. A convolutionary neural system comprises of one information and one yield layer, just as different shrouded layers. A CNN's secret layers are usually a progression of convolutionary layers that lap with augmentation or other result of spots. The initiation feature is always a layer with RELU. It is later accompanied by extra convolutions, for instance, blend bed, totally related bed, and uniform bed, called hidden layers, as their wellsprings of info and yields are secured by the activation work and by the last convolution. The CNN algorithm is highly adaptable and effective at recognition. Training is also simple as there are fewer training conditions, and when combined with backpropagation, it is scalable. The CNN algorithm can be used for image processing, acknowledgment, classification, video recognition, pattern recognition, recommendation engines, and medical image analysis [25].

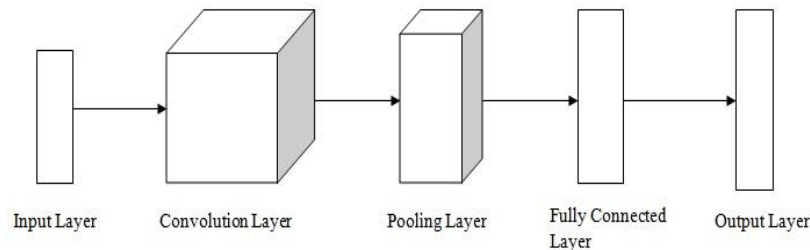


Fig.10. Architecture of CNN

2. Recurring Neural Network (RNN)

The purpose of the repetitive neural system is to recognize the successive component of an informational collection and to use patterns to predict the next probable scenario. It is an efficient approach to processing sequential information, for example, tone, time-arrangement data, and normal written language. The stochastic gradient descent (SGD), along with the backpropagation calculation, is utilized to train the system. A repetitive neurological framework is a kind of counterfeit neural framework where node-to-node connections form a directed graph along one chain. Often, RNNs is a feed-forward system, with repeated memory circles taking the contribution from the past layers or states. The secret layer in RNN retains sequential information from preceding phases. It means that the output from an earlier stage is fed in as the input to a current stage, often using the same weights and bias for prediction purposes. Instead, the layers are joined together to create a new recurrent layer. These feedback loops process sequential data, as in memory, allowing information to remain and inform the final output. RNNs are useful

for the classification of emotions, image captioning, speech recognition, processing of the natural language, machine translation, search prediction, video classification [26].

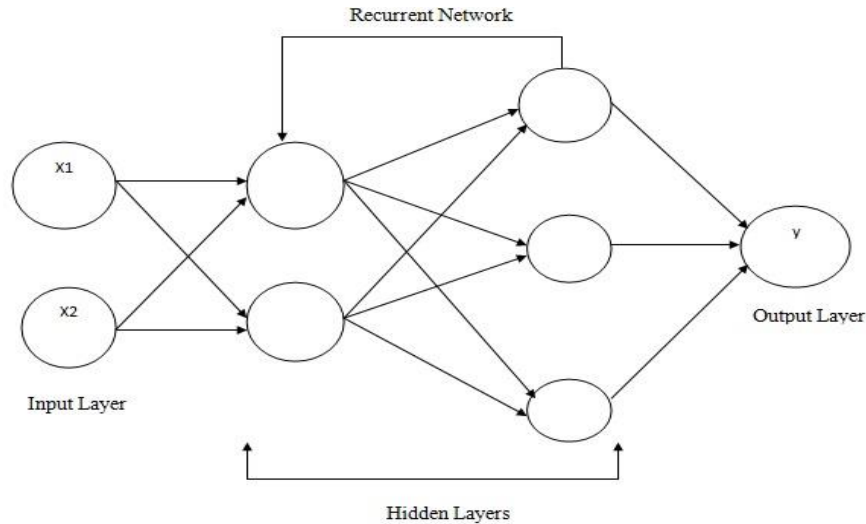


Fig.11. Recurring Neural System Architecture

3. Long Short-Term Memory (LSTM)

The algorithm for the long momentary memory is a type of RNN which takes into account the preparation of profound intermittent systems without making the slopes that update loads unsteady. Patterns may be retained for longer periods in memory, with the ability to selectively retrieve or erase the data. It uses backpropagation but is equipped to use memory blocks linked to layers instead of neurons to learn sequence information. Because the information is stored in layers, data can be added, removed, or changed by the architecture as needed. This calculation is unmistakably appropriate for classification and prediction dependent on information from time arrangement, providing sophisticated outcomes for various problems. These enable data scientists to build deep models using large stacked networks and deal more effectively with complex sequence problems in machine learning. A typical LSTM unit comprises of a cell, an info door, a yield entryway, and an overlook entryway. The battery recalls esteems over self-assertive timeframes, and the three doors control the progression of data inside and outside the cell. LSTM systems are used to identify, process, and make forecasts dependent on time arrangement data because there may be lags of uncertain time in a time series between important occurrences. LSTM applications include robot control, prediction of the time series, recognition of speech, rhythm learning, learning grammar, and recognition of human action [27].

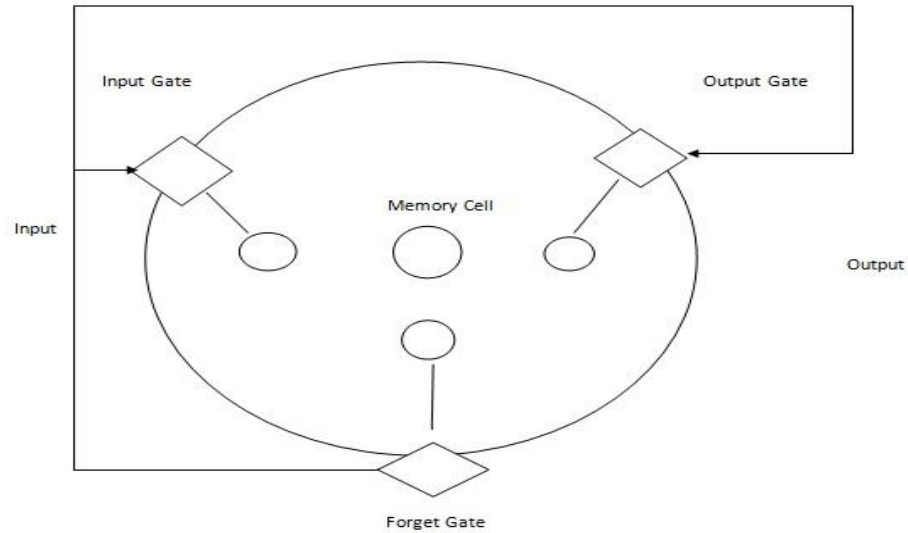


Fig.12. Structure of LSTM memory cell

4. Autoencoders

Autoencoders are often used as an unmonitored device, and their key use is the reduction of dimensions and compression. They are seeking to make output equal to the input. It is made out of an encipherer and a decipherer. The encipherer receives the information and encodes it in a lower-dimensional latent space. The decoder takes the vector back to the original input and decodes it back. Autoencoder applications include dimension reduction, information recovery, identification of anomalies, and image processing [28].

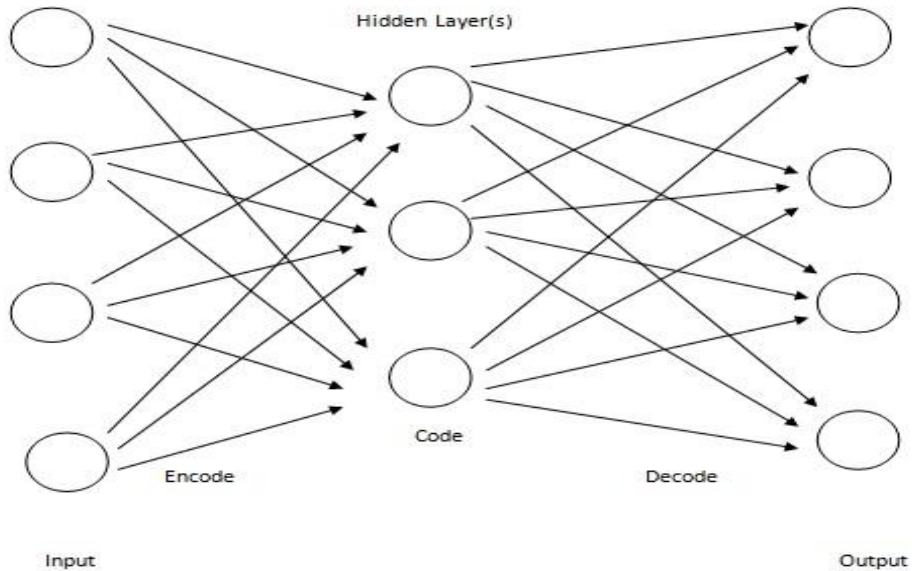


Fig.13. Structure of an Autoencoder network

5. Multi-layer Perceptron (MLP)

A multi-layer perceptron (MLP) is a fake neurological feed-forward system that produces a series of yields from an input series. It is characterized by different layers of connected info hubs as a coordinated chart between the information and yield layers. It employs backpropagation for arrange preparing. An MLP consists of an aggregate of three-hub beds: an information bed, a shrouded bed, and a yield bed. Each hub, barring the information hubs, is a neuron that utilizes a nonlinear initiation work. They by and large utilize some non-straight actuation functions like Relu or Tanh and calculate misfortunes like Mean Square Error (MSE), Logloss, etc. This failure is propagated in reverse to change the loads and preparing to reduce the damage or make the models increasingly precise [29].

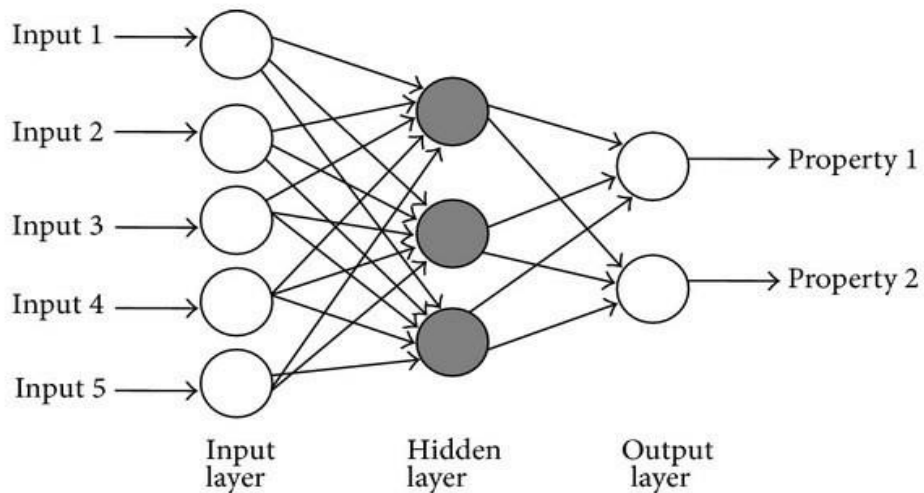


Fig.14. Multilayer Perceptron

CHAPTER 5: LITERATURE REVIEW

A. The Internet of Things

The Web of Things as an idea has its roots in the early 1990s. [30] was likely one of the first people to recognize the eventual rise of a system where on-demand computing was available to everyone through a combination of hardware and software, connected using wires or radio communication. In modern times, this idea has materialized as a viable and imminent future technology, characterized by a massively connected system of items or devices which can associate with each another over a network connection. [31]

I) Factors Influencing Rise of IoT

Today's widespread, worldwide telecommunications network lays the foundation for massive IoT in the near future. The rise of IoT as a technology platform is also partially attributable to the rapid downward scaling (miniaturization) of transistors; a trend that has been fairly consistent since the late 20th century and continues into the 21st century. Transistors form the bedrock of all silicon-based intelligence in today's time. This observed law of miniaturization is well documented [32] in Dennard's Scaling and Moore's Law; two well-known observations relating to computing in recent times. Current advances in miniaturization allow small sensors and computing modules to operate cheaply, efficiently, and be deployed at a large scale across a wide range of real-world applications. [33]

II) IoT in Industrial Applications

The Industrial applications of IoT (IoT) are of particular interest to this paper. IoT devices are primarily used to ingest and analyses data from industrial equipment, operational technology, physical locations, and human resources. Cyber security is an intensive industrial activity in the modern world; given its essential nature, IoT can make significant improvements to it. [34]

B. Deep Learning Methods for security in IoT

Subsequently, our exploration work advocates improving IoT security by utilizing Deep Learning calculations. Deep Learning [35] is in no way, shape or form an ongoing worldview. It is a subfield of Machine Learning that has its underlying foundations in Artificial Intelligence. Deep Learning helps essentially perform characterization assignments straightforwardly from writings, pictures and sounds. As of now, Deep Learning is to a great extent engaging the IT scene by tackling different issues. The brain cell of the ANN are utilized to shape complex theories; the more neurons, the progressively complex the speculations. Assessing the speculations is finished by setting the info hubs in a criticism procedure and the occasion streams are spread through the network to the yield where it is named typical or bargained. At this stage the inclination plunges are utilized to push the blunder in the yield hub back through the network by a back-proliferation process so as to gauge the mistake in the concealed hubs. The inclination of the expense – capacity would thus be able to be determined [36] – [37]. Neural network framework experiences preparing so as to gain proficiency with the example made in the framework. There are many techniques applied in cyber security with the help deep learning. Some of the techniques are LSTM which long term short memory. Recurrent neural network

known as RNN is used in some papers. Also, some researchers have used natural language processing techniques and bolster vector machine is also utilized.

A lot of work has been concluded in profound learning with the assistance of profound learning strategies. Chunyang Chen et al. [38] are exploring a profound learning way of assisting collaborative altering in Q&A pages. The main concept in this is to help inexperienced editors to alter posts with a wide variety of subjects, and to encourage the group to edit sentences. This exhibits the practicality of preparing a profound learning model with community post alters, and afterward utilizing the prepared model to help community post altering.

Subsequently, different systems like convolutional neural networks are likewise effectively investigated around there, which incorporates input surface, convolution surface, pooling surface, completely associated surface, and yield surface. Konstantinos P. Ferentinos [39] discusses plant disease detection and diagnosis models through a database that contains photographs of healthy and infected plant leaves. The future direction in this is to gather a wide scope of preparing information from various sources from various geographic areas, conditions of cultivation, and other factors.

In SenseBox architecture, H M Sajjad Hossain et al. [40] proposed a DeActive model. This algorithm is executed much more quickly than other algorithms. Dynamic learning can assist us with alleviating the manual effort expected to compile ground-level data about truth and decrease preparing time. With far less marked cases, DeActive can give better precision, which also ensures lower annotation effort.

Chris Xiaoxuan Lu et al. [41] examined security aspects by sniffing a deep learning smartwatch password that is a Snoopy method. Snoopy uses a uniform structure to separate movement information portions, albeit passwords are inserted, and utilizes new profound neurological systems directed toward surmise the real watchwords. This system can effectively spy information on moving out of sight while entering passwords. Without devouring noteworthy force/ computational assets, it can successfully extract password segments of motion data on smartwatches in real-time.

Parisa Pouladzadeh et al. [42] presented an app that utilizes the image of the nourishment, taken by the client's cell phone, to perceive different nourishment things in a similar food to evaluate the calorie and sustenance of the nourishment. In this, the client is challenged to rapidly recognize the broad territory of nourishment by an outline a bouncing ring on the nourishment image by contacting the canopy. The framework, at that point, utilizes picture handling and statistical insight for food item acknowledgment.

Nathan Shone et al. [43] presented a system that plays a crucial role in defending PC systems called Network Intrusion Detection frameworks. This paper gives another profound learning strategy for interference identification, addressing to the expanding levels of human cooperation required and diminishing degrees of exactness in detection. For Future work, the primary investigation road for development will be to evaluate and stretch out the model's ability to deal with zero-day assaults, and afterward, hope to develop existing assessments by using genuine backbone system freight to exhibit the benefits of the all-encompassing representation.

CHAPTER 6: PROPOSED WORK

In this section, detailed methodology and proposed models are explained. The main objectives are to implement deep learning method with higher accuracy in cyber security to compare the accuracy with existing methods.

In this model we have only considered convolutional layers; however, we will make it deeper by adding more convolutional layers, as well as maxpooling layers. Max pooling layer is included to dispose of highlights with low score and keep just highlights with most elevated score. The outcomes are down tested or pooled include maps that feature the most present component in the fix. This has been found to work preferred by and by over normal pooling for PC vision errands like picture arrangement. Dropout coating is added to spare framework from warming. Yield from the dropout sheet is taken care of to completely associated surface which than give contribution to the thick sheet with arched capacity.

A. Proposed Methodology

The flowchart of our framework is shown in figure 15; here the input is the csv files which were gathered with several diverse digital assaults alongside with usual data for five consecutive days. In this we first applied the convolution 1d to our system and then we divide the dataset into two categories: BENIGN and Web attack. Then in the next step we label the data by assigning benign as 0 and web attack as 1. Then we break the information into 70% preparing and 30% for testing. Then we set the batch size to 32 and epoch which are the number of rounds to 100. Then we use for loop for epoch=1, 2, 3 and so on. For this we set the count to 0 and repeat until count+batch size< no. of training samples else it goes back to the dataset division. Then we train the classifier on benign and web attack and then set count=batch size and then perform testing on benign and web attack. The last step is to calculate the precision, accuracy, review and f1-score from the disarray lattice to know the accuracy of model.

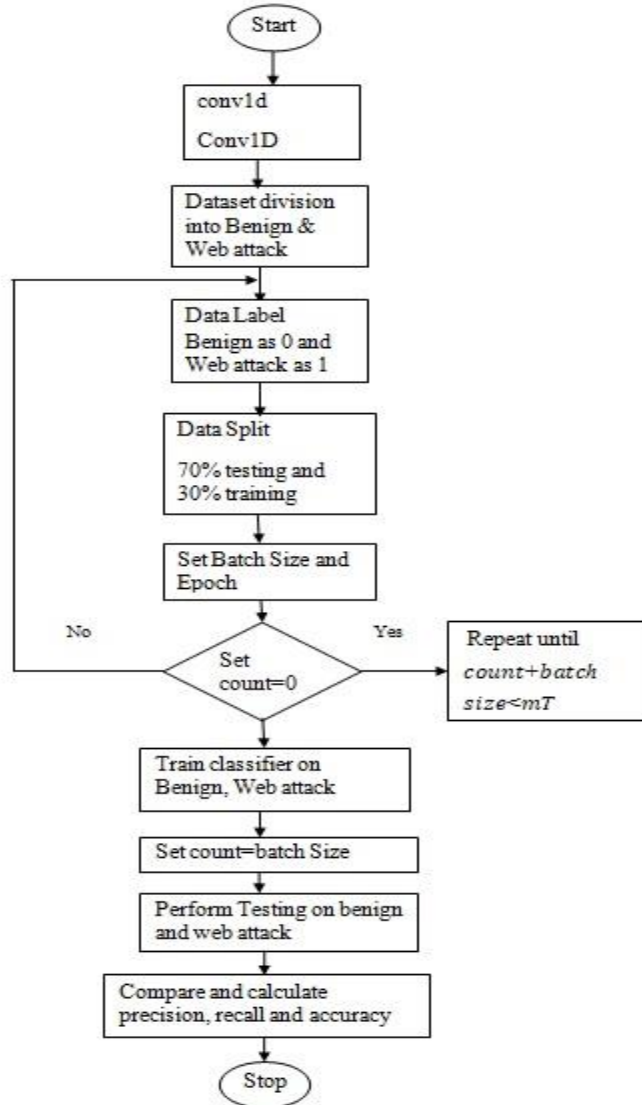


Fig. 15 Proposed Methodology

B. Algorithm

The final algorithm is given below:

Input: m - No. of samples
 n - No. of features
 0- Label for BENIGN data
 1- Label for Web attack
 mT - No. of training samples
 mt - No. of testing samples
 $XTmT \times n$ - Training samples
 $YTmT \times 1$ - Training labels
 $Xtmt \times n$ - Testing samples
 $Ytmt \times 1$ - Testing labels

$Y_{pred} \times 1$ - Predicted labels

Output: Calculation of accuracy, recall, and precision

1. Prepared dataset $X_{m \times n}$ and label $Y_{m \times 1}$ where $Y \in \{0, 1\}$
2. Split data $X_{m \times n}$ and label $Y_{m \times 1}$ into 70% training and 30% testing sets
3. $(X_{T \times mT \times n, \times 1})$ is the training set and $(X_{t \times mT \times n}, Y_{t \times mT \times 1})$ is the testing set
4. Set *batch size*=512 and *epochs*=100
5. for *epoch*=1,2,3,....., *epochs*
6. Set *count*=0
7. Repeat until *count*+*batch size*<*mT*. Train classifier on ($X_{T \times count+1 \text{ to } count+batch \text{ size}}$, $Y_{T \times count+1 \text{ to } count+batch \text{ size}}$)
6. Calculate accuracy
7. Set *count*=*batch size*
8. End for
9. Perform testing on $X_{t \times mT \times n}$ and find $Y_{pred} \times 1$
10. Compare $Y_{pred} \times 1$ and $Y_{t \times mT \times 1}$ and calculate precision, recall, F1-Score, Accuracy

First, divide the information into the preparation and testing part, where 70% of data have utilized for preparing and rest 30% part for testing. Since, the dataset is highly imbalanced, where attacks are in minimum quantity as compared to the BENIGN. Therefore, a different strategy such as multilevel classification can be adopted where the first decision will be whether the data is BENIGN or ATTACKED. If data comes in ATTACKED category then will predict the nature of the attack. Then set the batch size to 512 and epoch which are the number of rounds to 100. Then use for loop for epoch=1, 2, 3 and so on. For this set the count to 0 and repeat until *count*+*batch size*< no. of training samples else it goes back to the dataset division. Then train the classifier on benign and web attack and then set *count*=*batch size* and then perform testing on benign and web attack. The last step is to compute the exactness, accuracy, recall and f1-score from the confusion matrix to know the accuracy of model. We have tested 2 models on the dataset. The architectures of these models are shown below:

C. Proposed CNN 3 Layer Model

The modified CNN based deep learning algorithm is applied. Figure 17 shows multiple layers of deep learning-based convolution and max pooling is applied to improve the accuracy. In this there are convolution surface followed by max pooling coverings after which dropout surface is enforced. The dropout coatings are combined to spare the framework from warming. Yield from the dropout surface is taken care of to flatten sheet which at that point provide contribution for the dense sheet which then provide input to the second dropout layer. Yield from the dropout sheet is taken care of to the second dense sheet with sigmoid, relu, and softmax activation function. This CNN model is ideal when we require less computation as there is less parameter required in this model. Here the accuracy of the model is 99.10%.

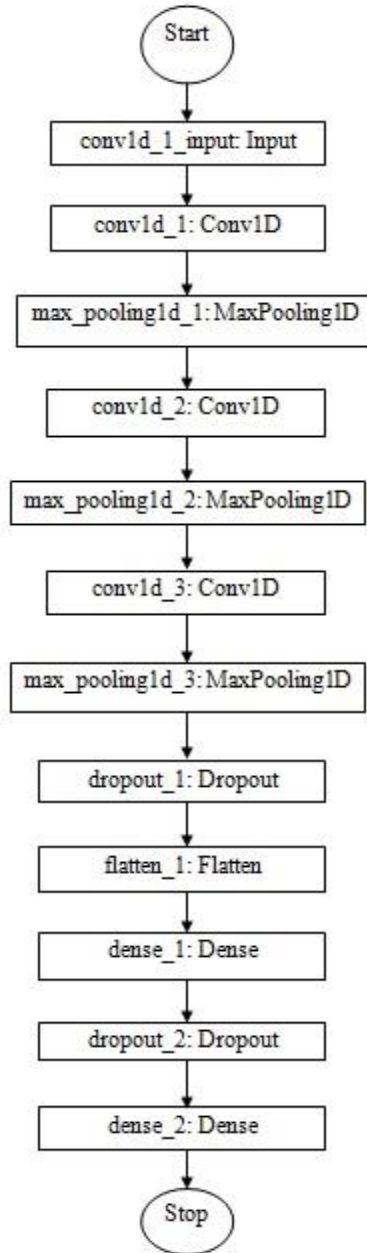


Fig.16 Proposed CNN 3 Layer Model

D. Proposed Multiheaded CNN Layer Model

The proposed model also known as multiheaded CNN layers is used, which basically concatenate three CNN processes as shown in the figure 16. They are followed by maxpooling layers, dropout layers and flatten layers and then it connects them with concatenate layer. Yield from the concatenate surface is connected with the dense coat which gives contribution to the dropout sheet. The dropout surface is added to spare framework from warming. Yield from the dropout bed is connected with the second dense surface with sigmoid, relu, and softmax activation function. In this model the layers are connected parallely which is why this model is also known as multiheaded CNN layers. It

gives significant improvement in the accuracy and other measures as it applies multiple CNN layers to reduce the error to minimum. This CNN model is ideal when we require high computation as it requires more parameters. Here the accuracy of the model is 99.38% which is the highest of all the models.

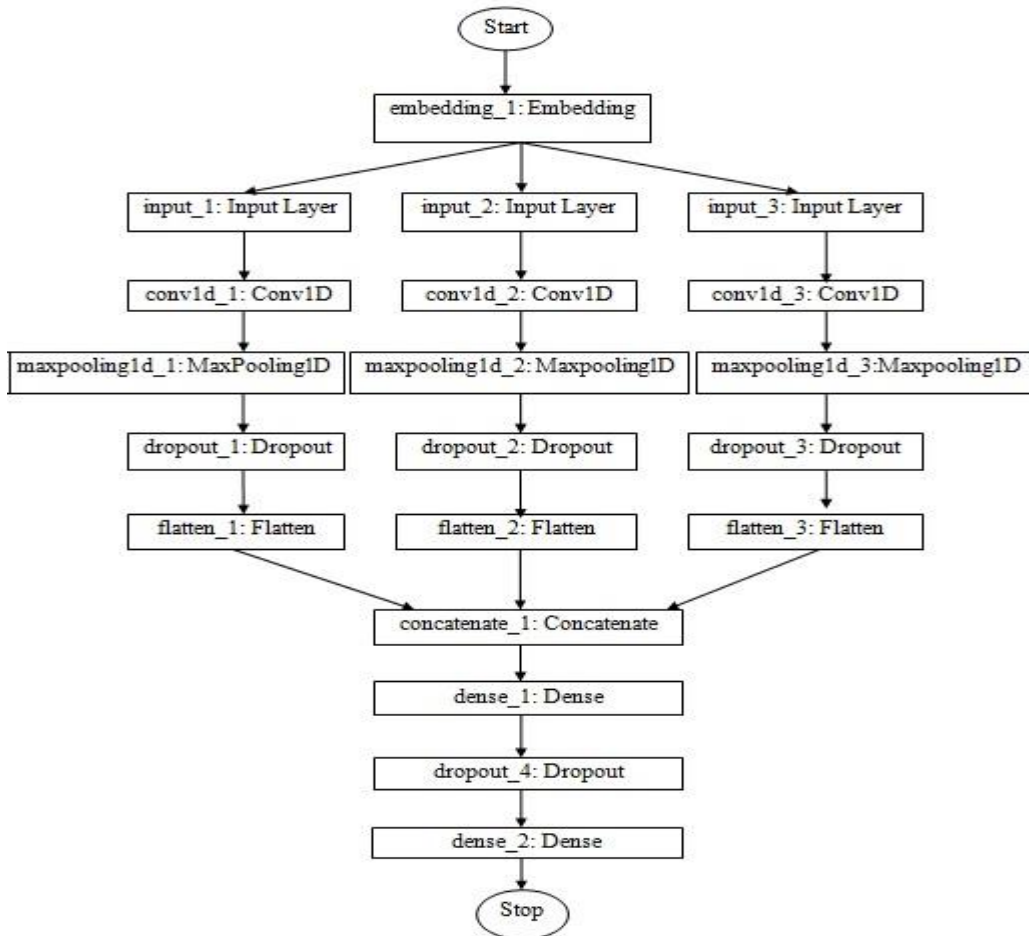


Fig.17 Proposed Multiheaded CNN Layer Model

The CNN based model is the base model. It consists of data input layer. Two convolution layers are applied before maxpooling of the features after which dropout is applied. The dropout sheet is combined to spare the framework from warming. Product from the dropout sheet is taken care of to flatten and sense layers are then applied for the final output. Here the accuracy of the model is 98.32%.

E. 1D CNN Architecture

The configuration of the 1D deep CNN model consists of an information blanket, a convolutional blanket, a max amalgamate sheet, a completely associated blanket, and a yield sheet as demonstrated in figure 18. We have applied a 1D Convolutional Neural Network on our data. A 1D CNN is exceptionally successful when you would like to get intriguing features from shorter (fixed-length) portions of the general enlightening file and where the region of the component inside the segment isn't of high significance. This applies well to the examination of time successions of sensor information. It additionally applies to the investigation of any sort of sign information over a fixed-length period, (for example, sound signs). Another application is NLP. First, we split the information into the preparation and testing part, where 70% of information have used for preparing and rest 30% part for examination. Since, the dataset is highly imbalanced, where attacks are in minimum quantity as compared to the BENIGN. Therefore, a different strategy such as multilevel classification can be adopted where the first decision will be whether the data is BENIGN or ATTACKED. If data comes in ATTACKED category then we will predict the nature of the attack.

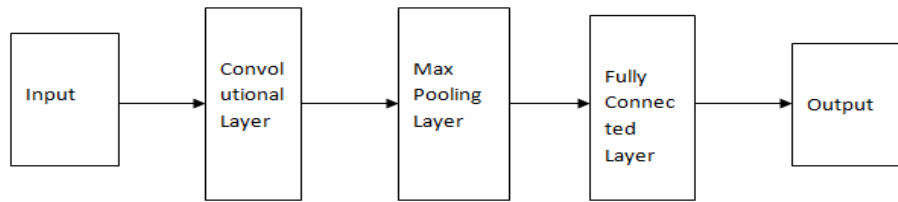


Fig.18 1D CNN architecture

Since the dataset is imbalanced, therefore the presentation of the models is assessed as far as Correctness, Recall, F1-Score, and Efficiency. The definition and formula for the performance metrics are shown below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

In this, TP is genuine positive, FN is bogus negative, TN is genuine negative and FP is bogus positive. F1 score depends on exactness. The performance metrics are then analyzed for each of the models and further explained in results section.

CHAPTER 7: EVALUATION AND CONCLUSION

We will clarify in this segment the informational set used and the examination environment. Instead, the measurements utilized will be talked about for the output assessment of the proposed models, and then the interpretation and proposals of future research will be debated.

7.1 Dataset and Environment

We used the new DDoS attack CICIDS2017 dataset for performing the work [44, 45]. Most DDoS assault datasets have several restrictions that are unreliable, for example, out of valid data, excess. CICIDS2017 databases state-of-the-art, data-like, genuine work networks. This dataset was gathered with several diverse digital assaults alongside with usual data for five consecutive days. This dataset contains the current modified system information with and without assault, which is fundamentally the same as the real system information of the activity. This dataset is uneven, so we have adjusted this dataset by copying the algorithm because it genuinely influences the deep learning system preparing and examination. This performance is utilized on Keras [46] on Tensorflow bundle for profound study.

For leading proposed endeavor we have utilized most recent DDoS assault CICIDS2017 dataset. CICIDS2017 datasets contain exceptional genuine work arrange taking after information. This dataset was accumulated for five sequential days with distinct cyberattacks alongside ordinary information. This dataset contains latest cutting-edge arrange information with and without assault which is near the genuine work organize data.

In this implementation, python language is used. There are 2830743 data samples and 79 features. After preprocessing like removing columns having all zeros and normalization, we have total 76 features.

In this implementation, further we will discuss about the dataset. It contains total 2830743 data samples from 15 different categories having 79 features. The categories and the percentage of data samples in the dataset are as follows: BENIGN (80.3%), Infiltration (0.0013%), DDoS (4.52%), Bot (0.069%), Web assault sql infusion (0.0007%), SSH-Patarator (0.21%), DoS slowloris (0.20%), DoS Hulk (8.16%), PortScan (5.61%), Heartbleed (0.0004%), DoS Slowhttpstest (0.19%), DoS GoldenEye (0.36%), FTP-Patarator (0.28%), Web assault Brute Force (0.053%), Web assault XSS (0.023%).

Out of 15 categories BENIGN is neural and rest are different web attacks. As we can see that percentage of individual attacks data in many cases are around 0%. BENIGN is the most used data sample in the dataset with the second being DoS Hulk. Therefore, we have divided data into two parts BENIGN and Web Attacks and then profound learning models are applied to arrange the information into either of the two categories. The dataset contains 79 features, some of highlights are: goal port, stream duration, aggregate forward parcels, complete in reverse bundles, absolute length of forward bundles, all out length of in reverse bundles, forward bundle length max, forward bundle length min, forward parcel length mean, forward parcel length standard, backward parcel length max, backward parcel length min, backward bundle length

mean, backward bundle length standard, stream byte/s, stream packet/s, stream IAT mean, stream IAT standard, stream IAT max, stream IAT min and so on.

7.2 Performance Metrics

High prediction accuracy rates and low error rates are the key targets in attack detection. If the prediction of the device is true, then the outcome is valid; else, it is considered bogus. When the forecast is about being targeted, this is considered a positive situation; else, it is negative. Therefore there are four prospects; expectation is valid and assault, valid and positive, bogus and assault, bogus and positive, respectively, Genuine Positive, Genuine Negative, Fake Positive and Fake Negative. The deep learning model's success in detecting DDoS attacks is calculated as accuracy, recall, and precision. The equations are given below for accuracy, recall, and precision:

$$\text{Accuracy} = (TP + TN) \div (TP + FP + TN + FN)$$

$$\text{Precision} = (TP) \div (TP + FP)$$

$$\text{Recall} = (TP) \div (TP + FN)$$

$$\text{F-Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

Where TP, TN, FP, FN stand for accurate positive, accurate negative, distorted positive, and distorted negative, respectively.

7.3 Results

In this area, developments are conferred. All the models are assessed on a fair CICIDS2017 dataset. The testing of results is done in complete week bases with different attacks and methods on each day. The results are basically comparison of the performance metrics clarified in past segment. The exhibition of the classifier is conferred in Table 2.

The confusion matrix for all the models is presented in figure 19, 20, and 21. From this we can get the values of genuine positive, genuine negative, fake positive, fake negative and compute accuracy, recall, and precision.

A. Confusion matrix of CNN Model

The disarray network of the CNN model which is the base model is shown in figure 19. For example Heartbleed attack 100% incorrectly classifies the attack as BENIGN. However the second attack DoS slowloris 97% correctly classifies that the attack is DoS Slowloris but 2 % incorrectly the attack as BENIGN and DoS Slowhttptest. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

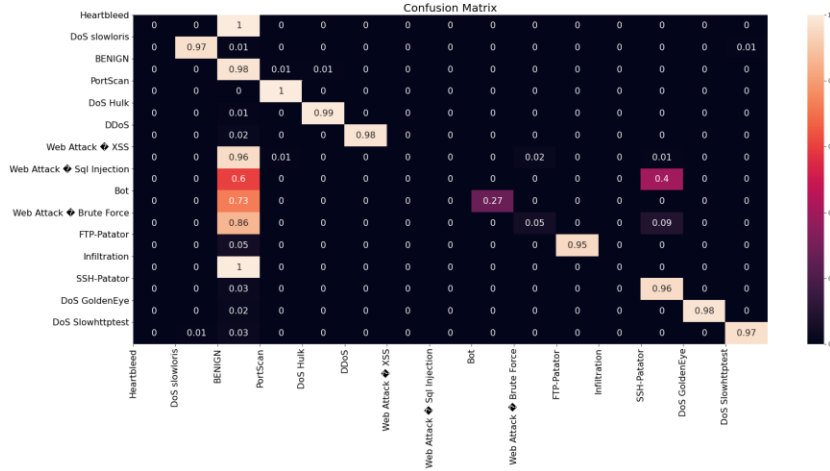


Fig. 19 Confusion matrix of CNN model

B. Confusion matrix of Proposed CNN 3 Layer Model

The disarray lattice of the proposed CNN 3 layer model is shown in figure 20. For example Web attack-brute force attack does not correctly classifies that it is the same attack. Instead it 92% incorrectly classifies the attack as BENIGN. Similarly the second attack Web attack- XSS 99% incorrectly classifies that the attack is BENIGN. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

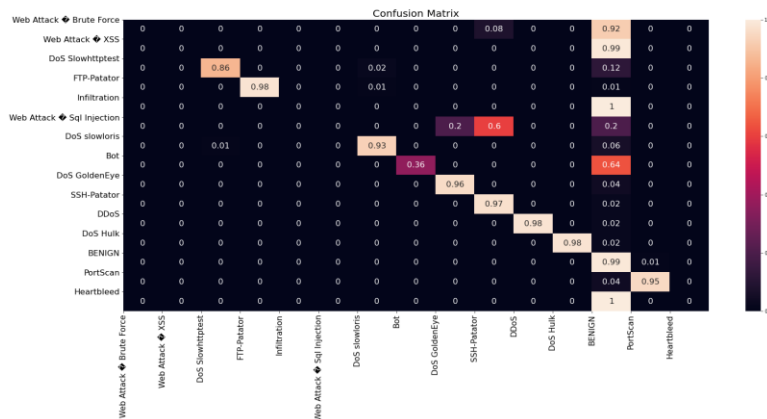


Fig. 20 Confusion matrix of Proposed CNN 3 layer model

C. Confusion matrix of Proposed Multiheaded CNN Layer Model

The disarray network of the proposed multiheaded CNN layer model is shown in figure 21. For example Heartbleed 100% incorrectly classifies the attack as BENIGN. The second attack SSH-patator 97% correctly classifies the attack as SSH-patator but 3% incorrectly classifies the attack as BENIGN. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

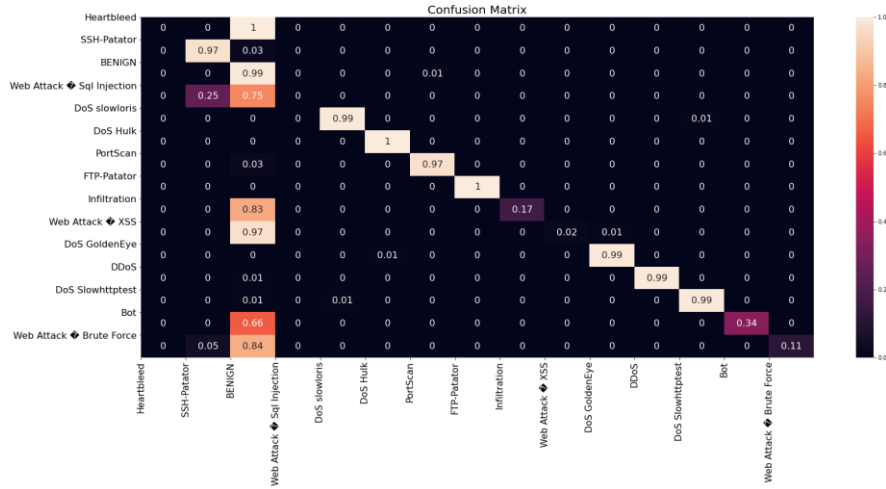


Fig. 21 Confusion matrix of proposed multiheaded CNN layer model

D. Performance Metrics Evaluation

The performance metrics evaluation is shown in table 2. In this the CNN model has precision of 94.33%, recall of 97.62%, f1-score of 96.41%, and accuracy of 98.32%. The second model which is the proposed CNN 3 layer model has precision of 96.54%, recall of 98.44%, f1-score of 97.48%, and accuracy of 99.10%. The third model which is the proposed multiheaded CNN layer model has precision of 98.70%, recall of 99.33%, f1-score of 99.01%, and has an accuracy of 99.38% which is the highest of all the models.

Table 2. Performance Metrics Evaluation Table

Model Name	Precision	Recall	F1-Score	Accuracy
Base Model	94.33 %	97.62 %	96.41 %	98.32 %
CNN 3 Layer	96.54 %	98.44 %	97.48 %	99.10 %
Proposed Multiheaded CNN Model	98.70 %	99.33 %	99.01 %	99.38 %

The time vs accuracy and time vs loss diagram for all models are presented in Figure 22, 23, and 24.

E. Epoch vs. Accuracy and Epoch vs. Loss curve of CNN Model

In this CNN model the testing accuracy versus epoch is significantly high in the below figure 22 for the first model. As we increase the number of epochs the testing accuracy also increases. Here epochs are the number of rounds in which we are running the model. The loss vs. epoch curve shows that loss is decreasing as the number of rounds increases.

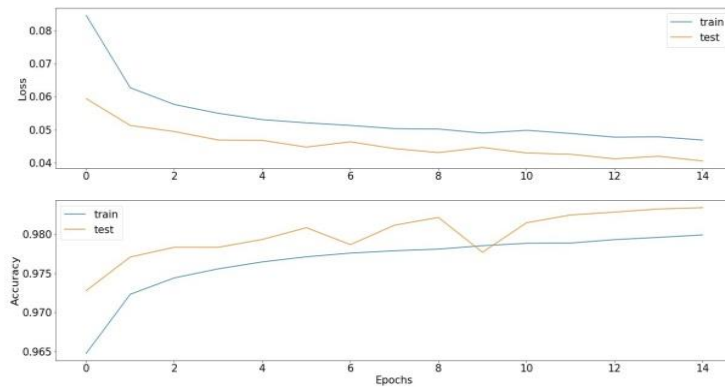


Fig. 22 Epoch vs Accuracy and Epoch vs Loss curve of CNN model

F. Epoch vs. Accuracy and Epoch vs. Loss curve of Proposed CNN 3 Layer Model

The CNN 3 layer model for the representation results is shown in figure 23; variation in accuracies is seen as per the increase in the epochs. As number of rounds increases there are variations in testing accuracy. The variation in accuracy shows that the accuracy is not constant it keeps on fluctuating. The loss vs. epoch curve shows variation in loss as per the increase in the epochs.

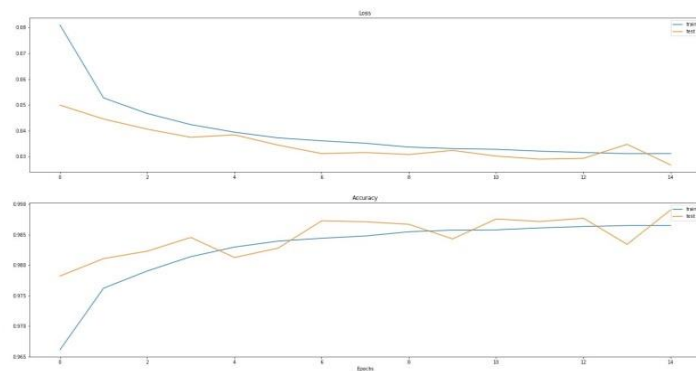


Fig.23 Epoch vs Accuracy and Epoch vs Loss curve of proposed CNN 3 layer model

G. Epoch vs. Accuracy and Epoch vs. Loss curve of Proposed Multiheaded CNN Layer Model

The final proposed multiheaded CNN model result for accuracy and loss curve is seen as shown in figure 24, which improved in terms of high accuracy and low loss techniques result. In this the testing accuracy is improved in terms of low loss in the model. The loss vs. epoch curve shows loss is decreasing and as a result testing accuracy of the model increases. This model has the highest accuracy as compared to the rest of the models

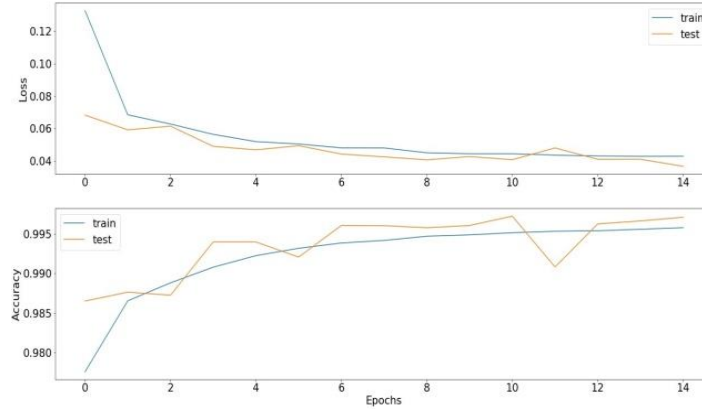


Fig.24. Epoch vs Accuracy and Epoch vs Loss curve of proposed Multiheaded CNN layer model

H. Comparison of Evaluation parameters for all models

Comparison of all three-model implementation on the basis of four parameters, accuracy, F1 score, recall and precision are shown in figure 25 for deep learning method in cyber security. Here model 1 is the cnn model which is the base model, model 2 is the proposed cnn 3 layer models, and model 3 is the proposed multiheaded cnn layer model. It is seen that proposed multiheaded cnn layer model is the best amongst the others and has a higher accuracy amongst all that means has the least error.

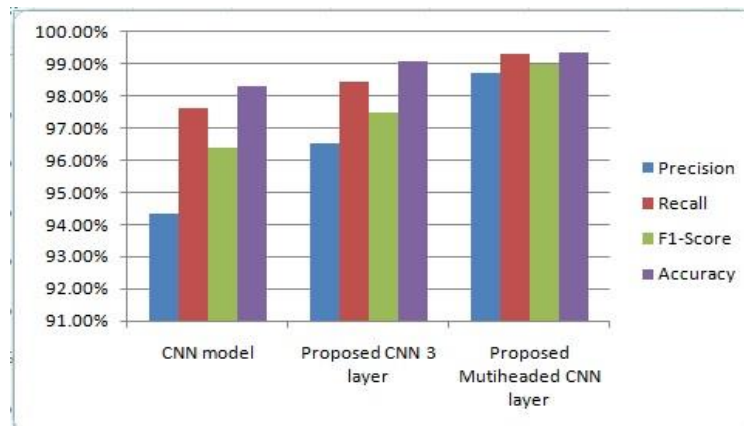


Fig.25. Comparison of evaluation parameters for all models

7.4 Conclusion & Future Works

Web of Thing is most recent rising bright development that interfaces all around the globe through web. IoT development helps to improve and bolster our own, proficient life and culture. The paper focused on deep learning method innovation using CNN and its different model variants. The datasets are used and tested with all possibilities to give better results and innovation in cyber security issues for internet of things. The loss and accuracy are main parameter for analysis in the ground pertaining to security improvement under the web of things. Hence, this paper provided an improved method to detect security issues in IOT using modified deep learning method. The proposed results indicate a higher accuracy in the CNN modified algorithm. In future it can be tested on edge servers and cloud assisted servers. As it is exceptionally unequal by copying information, we also balanced the data set for this study; this could be enhanced later on by building up a profound training standard that could run on the uneven dataset.

REFERENCES

1. W. Wahlster. From industry 1.0 to industry 4.0: Towards the 4th industrial revolution. In Forum Business meets Research, 2012.
2. F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt. Cross-level sensor network simulation with cooja. In Local computer networks, proceedings 2006 31st IEEE conference on, pages 641–648. IEEE, 2006.
3. Gns-3, 2017. URL <https://www.gns3.com/>.
4. Iotify, 2017. URL <https://iotify.io/iot-network-simulator/>.
5. Matlab, 2017. URL <https://www.mathworks.com/solutions/internet-of-things.html>.
6. DDoS attacks increased 91 URL <https://www.techrepublic.com/article/ddos-attacks-increased-91-in-2017-thanks-to-iot/>.
7. L. Coetzee and J. Eksteen, "The Internet of Things-promise for the future? An introduction," in *IST-Africa Conference Proceedings, 2011*, 2011, pp. 1-9.
8. X. Yuan, C. Li, and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," in *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2017, pp. 1-8.
9. A. Chadd, "DDoS attacks: past, present and future," *Network Security*, vol. 2018, pp. 13-15, 2018.
10. OWASP. Top IoT vulnerabilities. URL https://www.owasp.org/index.php/Top_IoT_Vulnerabilities.
11. Symantec. Internet security threat report. Technical report, Volume:22, April 2017.
12. J. A. M. M. Jazib Frahim, Carlos Pignataro. Securing the internet of things: A proposed structure. CISCO, 2016. URL <https://www.cisco.com/c/en/us/about/security-center/secure-iot-proposed-framework.html>.
13. A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng. Fog computing for the internet of things: Safety and privacy matters. *IEEE Internet Computing*, 21(2):34–42, 2017.
14. F. Gont. Results of a security evaluation of the internet protocol version 6 (ipv6).
15. M. Abomhara et al. Cybersecurity and the internet of things: vulnerabilities, threats, intruders, and attacks. *Journal of Cyber Security and Mobility*, 4(1):65–88, 2015.

16. A. A. Diro, and N. Chilamkurti. Distributed attack detection scheme using a deep learning approach for the internet of things: future Generation Computer Systems, 2017.
17. G. E. Hinton, "Deep belief networks. Scholarpedia, 4 (5), 5947," *Available electronically at [http://www. Scholarpedia.org/article/Deep_belief_networks](http://www.Scholarpedia.org/article/Deep_belief_networks) Hoppensteadt, FC*, pp. 129-35, 2009.
18. A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astonishing baseline for recognition," in *Proceedings of the IEEE seminar on computer vision and pattern identification workshops*, 2014, pp. 806-813.
19. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Picturenet categorization with deep convolutional neural systems," in *Advances in neural data processing systems*, 2012, pp. 1097-1105.
20. B. Hu, Z. Lu, H. Li, and Q. Chen, "Convolutional neural network architectures for matching natural language sentences," in *Advances in neural data processing systems*, 2014, pp. 2042-2050.
21. W. Hao, R. Bie, J. Guo, X. Meng, and S. Wang, "Optimized CNN Based Image Recognition Through Target Region Selection," *Optik- International Journal for Light and Electron Optics*, vol. 156, pp. 772-777, 2018.
22. T. Hori, Z. Chen, H. Erdogan, J. R. Hershey, J. Le Roux, V. Mitra, *et al.*, "Multi-microphone speech recognition integrating beamforming, robust feature extraction, and advanced DNN/RNN backend," *Computer Speech & Language*, vol. 46, pp. 401-418, 2017.
23. K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space quest," *IEEE transactions on neural networks and learning systems*, vol. 28, pp. 2222-2232, 2017.
24. Cybersecurity. URL <https://digitalguardian.com/blog/what-cyber-security>
25. Ciresan, Dan; Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber (2011). "Flexible, High-Performance Convolutional Neural Networks for Image Classification" (PDF). Proceedings of the Twenty-Second International Common Conference on Artificial Intelligence-Volume Volume Two. 2: 1237–1242.
26. Graves, A.; Liwicki, M.; Fernandez, S.; Bertolami, R.; Bunke, H.; Schmidhuber, J. (2009). "A Novel Connectionist System for Improved Unconstrained Handwriting Recognition"(PDF). *IEEE Proceedings on Pattern Examination and Machine Intelligence*. 31 (5): 855–868.
27. Sak, Hasim; Senior, Andrew; Beaufays, Francoise (2014). "Short-Term Long Memory recurrent neural network architectures for large scale acoustic modeling"(PDF).

28. Kramer, Mark A. (1991). "Nonlinear principal constituent analysis using auto-associative neural systems" (PDF). *AIChE Journal*. 37 (2): 233–243. DOI:10.1002/aic.690370209.
29. Multilayer Perceptron, 2016. URL <https://medium.com/pankajmathur/a-simple-multilayer-perceptron-with-tensorflow-3effe7bf3466>
30. M. Roopak, G. Yun Tian and J. Chambers, "Deep Learning Models for Cyber Security in IoT Networks," *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 2019, pp. 0452-0457, doi: 10.1109/CCWC.2019.8666588.
31. N. Y. Parotkin and V. V. Zolotarev, "Information Security of IoT Wireless Segment," *2018 Global Smart Industry Conference (GloSIC)*, Chelyabinsk, 2018, pp. 1-7, doi: 10.1109/GloSIC.2018.8570144.
32. Thamilarasu, Geethapriya & Chawla, Shiven. (2019). Towards Deep-Learning-Driven Intrusion Detection for the Internet of Things. *Sensors*. 19. 1977. 10.3390/s19091977.
33. R. Mahmoud, T. Yousuf, F. Aloul and I. Zualkernan, "Internet of things (IoT) security: Current status, challenges and prospective measures," *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, London, 2015, pp. 336-341, doi: 10.1109/ICITST.2015.7412116.
34. M. R. Schurgot, D. A. Shinberg and L. G. Greenwald, "Experiments with security and privacy in IoT networks," "2015 IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)", Boston, MA, 2015, pp. 1-6, doi: 10.1109/WoWMoM.2015.7158207".
35. W. ABBASS, Z. BAKRAOUY, A. BAINA and M. BELLAFKIH, "Classifying IoT security risks using Deep Learning algorithms," *2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Marrakesh, Morocco, 2018, pp. 1-6, doi: 10.1109/WINCOM.2018.8629709.
36. J. Lee, J. Kim, I. Kim and K. Han, "Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles," in *IEEE Access*, vol. 7, pp. 165607-165626, 2019, doi: 10.1109/ACCESS.2019.2953095.
37. D. Wilson, Y. Tang, J. Yan and Z. Lu, "Deep Learning-Aided Cyber-Attack Detection in Power Transmission Systems," *2018 IEEE Power & Energy Society General Meeting (PESGM)*, Portland, OR, 2018, pp. 1-5, doi: 10.1109/PESGM.2018.8586334.
38. Chunyang Chen and Zhenchang Xing. "Mining technology landscape from stack overflow," In *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. ACM, 14, 2016.

39. Dan, C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J. "Flexible, high execution convolutional neural systems for image classification," Events of the 22nd International Joint Conference on Artificial Intelligence, vol. 2, pp. 1237–1242, 2011.
40. Hande Alemdar, TLM van Kasteren, and Cem Ersoy, "Active learning with ambiguity sampling for broad-scale activity identification in a smart dormitory," *Journal of Ambient Intelligence and Smart Environments* 9, 2, 209-223, 2017.
41. Mariamn Harbach, Alexander De Luca, and Serge Egelman, "The anatomy of smartphone unlocking: A field study of android lockscreens" In *ACM Conference on Human Factors in Computing Systems, CHI*, 2016.
42. Parisa Pouladzadeh, Pallavi Kuhad, Sri Vijay Bharat Peddi, Abdulsalam Yassine, and Shervin shirmohammadi. "Calorie measurement and food classification using deep learning neural network," in *Proceedings of the IEEE International Conference on Instrumentation and Measurement Technology*, 2016.
43. B. Dong and X. Wang, "Comparison deep learning method to traditional methods using for network intrusion detection," in *Proc. 8Th IEEE Int. Conf. Commun. Softw. Network*, pp. 581-585, 2016.
44. I. Sharafaldin, A. H. Lashkari and A. A. Ghorbani, "Toward Creating a Modern Intrusion Detection Dataset and Intrusion Freight Identification," in *ICISSP*, 2018, pp. 108-116.
45. R. Vijayanand, D. Devaraj, and B. Kannapiran, "Intrusion detection arrangement for wireless mesh network using different support vector machine classifiers with hereditary calculation based element choice," *Computers & Security*, vol. 77, pp. 304-314, 2018.
46. Keras deep learning P.W.D. Charles Project Title Available at: <https://github.com/charlespwd/project-title>

LIST OF PUBLICATIONS OF THE CANDIDATE'S WORK

1. Sanjay Patidar and Inderpreet Singh Bains, "Web Security in IoT Networks using Deep Learning Model," 2020 3rd Scopus-Indexed IEEE International Conference on Smart Systems and Inventive Technology (ICSSIT), Tamil Nadu, India, 2020.
2. Sanjay Patidar and Inderpreet Singh Bains, "Intrusion Detection using Deep Learning," 2020 2nd Scopus-Indexed Springer International Conference on Inventive Computation and Information Technologies (ICICIT), Coimbatore, India, 2020.