International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

Comparative Study of Deep Learning Models for Network Intrusion Detection

Reeja Susan Reji[1], Anjitha Raj[2], Riya Raju[3], Manya[4], Sunandha Rajagopal[5]

[1] *PG – Computer Application, Kristu Jyoti College of Management and Technology, MG University, Kottayam, Kerala*

[2] *PG - Computer Application, Kristu Jyoti College of Management and Technology, MG University, Kottayam, Kerala*

[3] *PG - Computer Application, Kristu Jyoti College of Management and Technology, MG University, Kottayam, Kerala*

[4] *PG - Computer Application, Kristu Jyoti College of Management and Technology, MG University, Kottayam, Kerala*

[5]*Assistant Professor - Computer Application, Kristu Jyoti College of Management and technology, Changanacherry, Kottayam, Kerala*

ABSTRACT

In this paper, we present a relative assessment of profound learning ways to deal with network interruption recognition. An Organization Interruption Recognition Framework (NIDS) is a basic part of each and every Web associated framework due to likely goes after from both outer and inside sources. A NIDS is utilized to distinguish network conceived goes after like Forswearing of Administration (DoS) assaults, malware replication, and interlopers that are working inside the framework. Various profound learning approaches have been proposed for interruption identification frameworks. We assess three models, a vanilla profound brain net (DNN), self-trained learning (STL) approach, and Repetitive Brain Organization (RNN) based Long Present moment Memory (LSTM) on their exactness and accuracy. Their exhibition is assessed utilizing the organization interruption dataset given by Information Disclosure in Data sets (KDD). This dataset was utilized for the third global Information Revelation and Information Mining Devices contest held related with KDD Cup 1999. The outcomes were then contrasted with a standard shallow calculation that utilizes multinomial strategic relapse to assess if profound learning models perform better on this dataset.

1. Introduction

Over the past few decades, the Internet has penetrated all aspects of our lives. Experts predict that by 2020 there would be 50 billion connected devices [1]. As innovation turns out to be increasingly coordinated, the test to keep the frameworks protected and away from weakness assaults increments. Throughout the long term we have seen an expansion in hacks in financial frameworks, medical services frameworks and may Web of Things (IoT) gadgets. These assaults cause billions of dollars in misfortunes consistently and loss of frameworks at vital times. This has prompted higher significance in digital protection explicitly in the interruption discovery frameworks. A connected test with most current frame work is that information prerequisites relating to security are in many cases a bit of hindsight. It is expected that this effects the consequences of any AI calculation applied towards the issue; nonetheless, an examination differentiating the distinctions are yet to be seen. In addition to this, there is little research in the results of applying next level analysis using deep learning algorithms to determine if there is an improvement in accuracy versus its traditional machine learning counterparts [2].

An organization interruption identification framework (NIDS) is a product application that screens the organization traffic for malevolent movement. One well known system is to screen an

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

organization's action for oddities, or anything that veers off from typical organization conduct. Abnormality discovery makes models of typical way of behaving for networks and different gadgets and afterward searches for deviations from those examples of conduct at a much quicker pace. AI is utilized to fabricate oddity identification models and there are two methodologies shallow learning and profound learning. Shallow students for the most part depend on the elements utilized for making the expectation model. Then again, profound students can possibly remove better portrayals from the crude information to make much better models. Profound students can learn better since they are made out of the different secret layers. At each layer the model can separate a superior portrayal from the list of capabilities when contrasted with shallow students who don't have stowed away layers.
In this paper, we assess three profound learning models that utilization general brain net, self-educated learning, and constancy. The latter two models we build are based on Autoencoder and Long ShortTerm Memory (LSTM). For this research, we use the KDD Cup 1999 Dataset for our deep learning models and compare them to soft-max regression (SMR) results performed on the NSL-KDD dataset. Soft-max regression performed yielded an accuracy of 75.23%, recall of 63.73%, and an f-measure of 75.46% [3].

2. Intrusion Detection

The development of the Web and information traffic displayed various issues in respects to security the board. As the Web was not created with security in mind, the expansion in the quantity of clients all over the planet had acquainted the need with integrate access controls. Gate crashers have gotten imaginative in their strategies to penetrate or upset network traffic. They keep on adjusting to counteraction systems set up and keep on tracking down ways of taking advantage of the frameworks that are set up to forestall these interruptions from happening. First designed as a rule based system in 1987, Dorothy E. Denning and Peter Neumann where the first to pioneer the Intrusion Detection Expert System (IDES) using statistical models to achieve detection of anomalies [4]. Since then, methods of attack and prevention have adapted to utilizing different mediums as those innovations continue to release new methods of connection, thus opening the window to increased vulnerabilities that have yet to be discovered [5].
Various designs, or blends thereof, can be utilized to recognize an assortment of known assaults. The genuine benefit of NIDS is the framework will order examined network traffic to decide whether traffic or action is typical end-client movement or pernicious movement. Normal strategies for assault are perceptible by a NIDS. Overall NIDS can be arranged to two unique models on the host or organization. The first is to identify irregularities. The location of irregularities are accomplished by laying out a gauge of ordinary ways of behaving and hailing ways of behaving that go astray from that gauge. The second configuration relies on the comparison of known unwanted behaviors or misuse detection [6].
Assaults can come in many structures. A few conduct models that would set off a banner for an oddity can be port sweeps coming from one host on an organization across an whole subnet, download document count/size in a common organization envelope, various USB record moves, and so forth. NIDS can be designed to represent numerous conduct models. More designated setups can represent known marks for malware moved across the organization contrasted with a data set containing hashes for the malware. NIDS in this structure can be dealt with on a host-based arrangement. All the more strikingly, furthermore, frequently revealed in the news, DDOS assaults can utilize comparable arrangements to hinder the mind-boggling association demands. This is accomplished through arrangement if the framework can analyse against realized IP addresses; notwithstanding, can likewise be designed to identify obscure solicitations that show a similar example as a DDOS assault. Altogether occasions, a reaction is fundamental whether latent or dynamic. In case of social triggers, a uninvolved reaction, or a banner can inform a security chairman of potential split the difference to scholarly resources in the event that a representative was to download all documents from a record share. The case of a DDOS join; be that as it may, would require obstructing approaching traffic demands from the identified IP address to keep the solicitations from affecting the accessibility of a

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

framework. In all accounts, NIDS can be a powerful solution to mitigating for policing the massive amounts of data that can travel across networks, but does not replace the need for human intervention when further analysis is required to identify new threats or false positive detections [7].

Table 1. Definitions of attack types in the KDD Cup 1999 Dataset [8].

Attack Type	Description
DoS	A DoS attack is a type of attack in which the hacker makes a computing or memory resources too busy or too full serve legitimate networking requests and hence denying user access to a machine.
Probe	Probing is an attack in which the hacker scans a machine or a networking device in order to determine weaknesses or vulnerabilities that may later be exploited so as to compromise the system.
R2L	A remote to user attack is an attack in which a user sends packets to a machine over the internet, which s/he does not have access to in order to expose the machines vulnerabilities and exploit privileges which a local user would have on the computer.
U2R	User to root attacks are exploitations in which the hacker starts off on the system with a normal user account and attempts to abuse vulnerabilities in the system in order to gain super user privileges.

3. KDD and NSL-KDD Dataset

For our work, we utilize both the KDD and NSL-KDD dataset to see the distinction in execution. The KDD Cup dataset was arranged utilizing the organization traffic caught by 1998 DARPA IDS assessment program. The organization traffic incorporates typical and various types of assault traffic, like DoS, Testing, client to-root (U2R), and root-to-neighborhood (R2L). The organization traffic for preparing was gathered for quite some time followed by about fourteen days of traffic assortment for testing in crude tcpdump design. The test information contains many assaults that were not infused during the preparation information assortment stage to make the interruption discovery task practical. It is accepted that the greater part of the original assaults can be gotten from the known assaults. At long last, the preparation and test information were handled into the datasets of five million and two million TCP/IP association records, individually.

The KDD Cup dataset has been broadly utilized as a benchmark dataset for a long time in the assessment of NIDS. One of the significant disadvantage with the dataset is that it contains a huge measure of excess records both in the preparation and test information. It was observed that almost 78% and 75% records are redundant in the training and test dataset, respectively [9]. This overt repetitiveness makes the learning calculations one-sided towards the continuous assault records and

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

prompts unfortunate arrangement results for the rare, however destructive records. The preparation and test information were grouped with the base precision of 98% and 86% individually utilizing an extremely straightforward AI calculation. It made the examination task hard for different IDSs in view of various learning calculations.

NSL-KDD was proposed to beat the impediment of KDD Cup dataset. The dataset is gotten from the KDD Cup dataset. It worked on the past dataset in two ways. To start with, it wiped out every one of the excess records from the preparation and test data. Second, it divided every one of the records in the KDD Cup dataset into different difficulty levels in light of the quantity of learning calculations that can accurately characterize the records. Further, it chose the records by irregular examining of the unmistakable records from various trouble levels in a part that is contrarily corresponding to their portions in the particular records. Each record in the NSL-KDD dataset comprises of 41features and is named with one or the other typical or a sort of assault. These highlights incorporate essential elements got straightforwardly from a TCP/IP association, traffic highlights collected in a window stretch, either time, for example two seconds, or numerous associations, and content elements separated from the application layer information of associations. While contrasting the precision of our model against the KDD and NSL-KDD dataset, KDD fared improved yielding a higher exactness. However the NSL-KDD dataset has been cleaned and upgraded for AI purposes, we find that decrease in dimensionality in our profound learning models considerably affect the exactness while executing against the test set of the first KDD dataset. Moreover, there are tremendous contrasts in the measures of the preparation sets for the two datasets. KDD dataset contains 370,515 records while the NSL-KDD dataset contains 125,974. Successfully, the profound learning model has lightened the necessity of a manual information step.

4. Deep Learning Models

Profound learning was motivated by the design and profundity of human cerebrum. Due to the various degrees of reflection, the organization figures out how to plan the info elements to the result. The most common way of learning doesn't rely upon human-made highlights. Given a bunch of conditions, the machine can utilize a progression of numerical strategies to decide whether a characterization is exact in light of the probability of blunder. Inside the domain of profound learning, we center around profound organizations where the characterization preparing is directed via preparing with many layers in various leveled networks with solo learning. Profound organization interruption location frameworks can be characterized in view of how the models and procedures are being utilized.

In this segment, we notice the models utilized for examination. The primary model is a vanilla profound brain net classifier, which can be considered stacked calculated regressors. The second is the self-trained learning model utilizing autoencoder and the third is Intermittent Brain Organization will utilize Long Momentary Memory. To gauge the presentation of these models we utilize the measurements referenced in table 2

Table 2. Model Evaluation Metrics

Attack Type	Description
Accuracy	Characterized as the level of accurately ordered records over the complete number of records

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

Precision (P)	Characterized as the % proportion of the quantity of valid up-sides (TP) records isolated by the quantity of valid up-sides (TP) and misleading up-sides (FP) characterized records. P =TP/(TP + FP) ×100%
Recall (R)	Characterized as the % proportion of number of genuine up-sides records partitioned by the quantity of genuine up-sides and misleading negatives (FN) characterized records. R = TP/(TP+ FN) ×100%
F-Measure (F)	Characterized as the symphonious mean of accuracy and review and addresses a harmony between them. F =2.P.R/(P+R)

4.1 Deep Neural Net

A Profound Brain Organization is basically a multi-facet perceptron, which was at first evolved by stacking straight classifiers. This is the most essential kind of Profound Brain Organization that exists. The model is taken care of information sources, inputs get increased by loads and the passed into an enactment capability. In a Profound Brain Organization, this cycle happens over different layers. The model purposes backpropagation to change loads and increment exactness. Any model than contains at least 3 layers is viewed as a profound organization.

4.1.1 Model Setup

Before preparing the model the information were ready by changing unmitigated highlights over completely to numeric qualities. The information were standardized to lessen preparing time and increment execution. The last component of the dataset were 41 distinct highlights with 5 different anticipated classes.

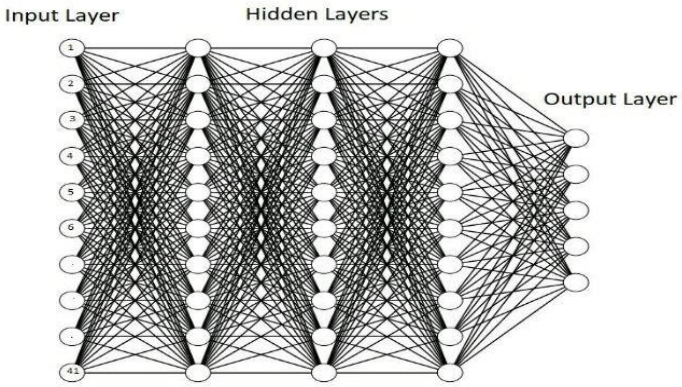

Fig. 1. A Deep Neural Network with 3 hidden layers.

4.1.2 Results

The profound brain network accomplished an exactness of 66%, the order of each assault type is displayed underneath in Figure 2. The model had the option to order DoS and test goes after well however had little outcome in accurately characterizing ordinary harmless solicitations and U2R assaults. The exactness is a ton lower than anticipated from a profound organization.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

	Precision	Recall	f 1-score	Support
DoS	0.83	0.62	0.71	4342
Probe	0.80	0.68	0.74	2402
R2L	0.67	0.00	0.00	2753
U2R	0.00	0.00	0.00	200
Normal	0.23	0.69	0.34	2152
Average/Total	0.66	0.49	0.47	11849

Fig. 2. Deep neural net results.

4.2 Self-Taught Learning Approach

Self-educated inclining (STL) is a profound learning approach that comprises of two phases for grouping. The primary stage is Unaided Component discovering that comprises of gaining a decent element portrayal from an enormous assortment of unlabeled information. This stage is carried out utilizing a meager Autoencoder. A meager autoencoder is a brain network that comprises of information, stowed away and yield layers. The information and result layers contain equivalent N hubs, while the secret layer contains K hubs. The result from the autoencoder is then gone through a delicate max relapse (SMR) for the characterization task.

4.2.1 Model Setup

Before using the training dataset, we first convert the categorical features to numeric values. We then perform a min-max normalization on this feature vector. The labels are one hot encoded. Therefore, the input dimension is 41 and output dimension is 5 (4 attacks and 1 normal). We pass the feature vector through a two-layer stacked autoencoder, the first autoencoder has a hidden layer of 20 and the second layer has a hidden layer of 10. The output from the encoder of the second layer is then passed through a soft max regressor to classify the input to one of the 5 labels

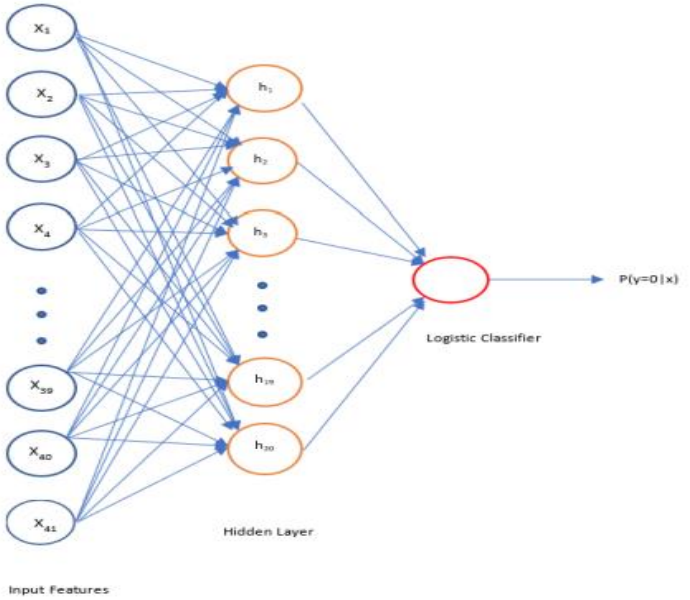

Fig. 3. Autoencoder dimensionality reduction and features input to a logistic classifier.

4.2.2 Results

The STL approach brings about an exactness of 98.9% with the accompanying separation by each assault type. We see that since the portrayal of R2L and U2R kind of assaults are low the accuracy and review of these assault types are lesser when thought about to the next assault types. We see that the STL has taken in a decent portrayal of the include set to have the option to foresee with a serious level of exactness.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

	Precision	Recall	f 1-score	Support
DoS	1.00	0.99	0.99	97865
Probe	0.78	0.59	0.67	1027
R2L	0.51	0.11	0.19	281
U2R	0.00	0.00	0.00	13
Normal	0.95	0.98	0.96	24320
Average/Total	0.98	0.98	0.98	123506

Fig.4. Self-taught learning (autoencoder) results

4.3 Recurrent Neural Network

Intermittent brain networks are a class of Counterfeit Brain organization. They take as their input the ongoing info case as well as what they have seen already in time. This implies that they likewise have an extra memory input. The choice a RNN takes at time t-1 impacts the choice it requires at investment t. In this way, the repetitive brain networks have two wellsprings of info - the present and the new past, which join to decide how the RNN will answer the new information. This input circle is principal distinction among RNNs and the feed forward brain organization. One of the weaknesses of a RNN was the evaporating inclination issue. This happens when the inclination is tiny, and subsequently the loads can't be changed. This would forestall the brain net from preparing further. The Long Momentary Memory organizations (LSTM) are a unique sort of RNN, which dispenses with the evaporating inclination issue, as they can learn long haul conditions without any problem. Typical RNNs take in their past stowed away state and the ongoing info state to yield another secret state. The LSTM does the same, aside from it likewise takes an old cell state

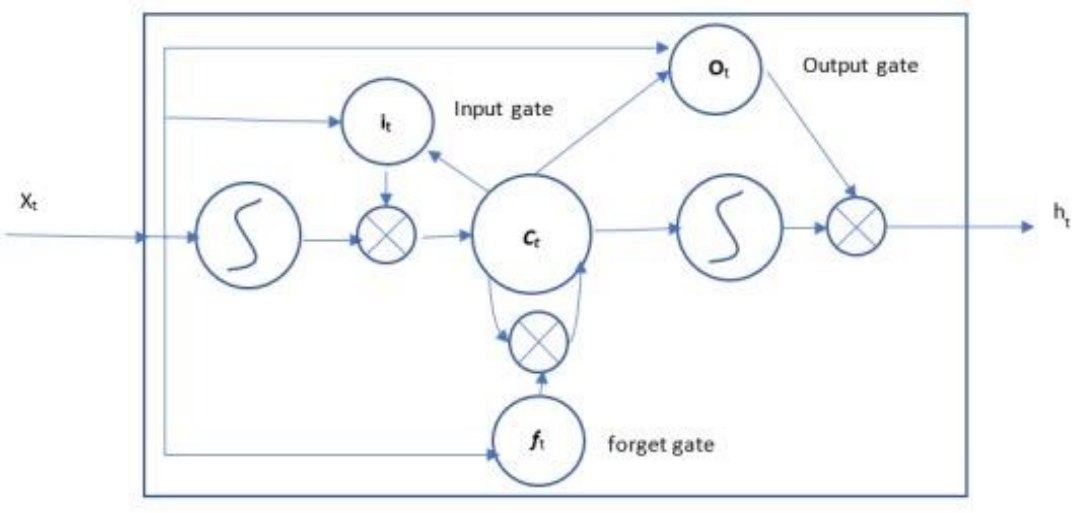

Fig.5. Long Short Term Memory (LSTM) cell.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

4.3.1 Model Setup

Like the past model arrangement we first believer the clear cut elements to numeric qualities. We then, at that point, play out a min-max standardization on this component vector. The marks are one hot encoded. Consequently, the information aspect is 41 and yield aspect is 5 (4 assaults and 1 ordinary). What's more, we apply LSTM engineering to the secret layer. The time step size, clump size, and ages are 100, 50, 5 individually. We use soft max for the result layer and stochastic angle plunge (SGD) for an optimizer. We use a learning rate of 0.01 and hidden layer of 80.

4.3.2 Results

The LSTM model outcomes in a precision of 79.2% with the accompanying separation by each assault type. We see that this model can't anticipate goes after other than DoS. This might be because of the preparation information having a higher circulation of DoS cases and may require further tuning of our model.

	Precision	Recall	f 1-score	Support
DoS	0.79	1.00	0.88	97780
Probe	0.00	0.00	0.00	1027
R2L	0.00	0.00	0.00	281
U2R	0.00	0.00	0.00	13
Normal	0.00	0.00	0.00	24304
Average/Total	0.63	0.79	0.70	123405

Fig. 6. Long Short Term Memory (LSTM) model result.

4.4 Results Analysis

Figure 7 shows a comparison of the performance measures. The outputs of our three models—DNN, RNN, and Autoencoder deep learning algorithms—are compared. Overall, Autoencoder had the best performance when it came to separating DoS type assaults from regular network data, scoring the highest in precision, recall, and f1-score. Due to a lack of data to properly classify attacks, all models failed to recognize U2R-type attacks.

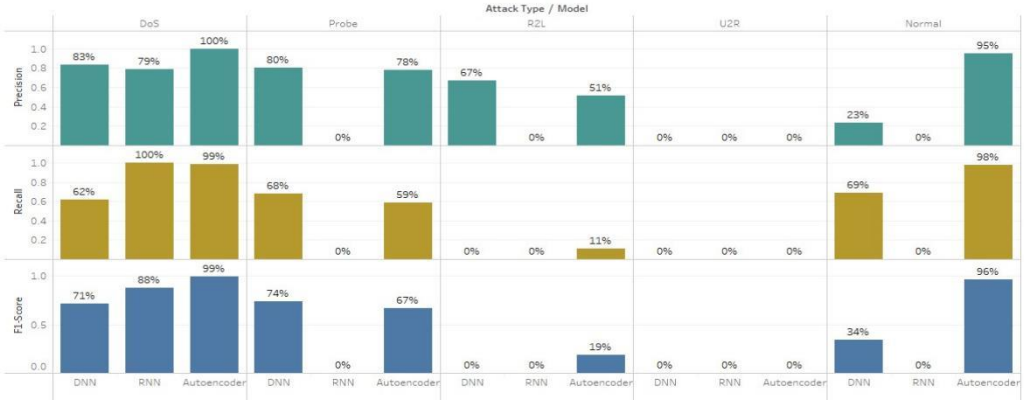

Fig. 7. Performance metric result comparison between DNN, RNN, and Autoencoder.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

5 Ethical Considerations

Given the Internet's environment, machine learning can be used to manage the vast volumes of traffic and distinguish between dangerous and benign content. It can help optimize the use of resources, whether human or machine; however, it cannot be the sole solution in the attempt to mitigate the risk of intrusion [4]. The answer to the larger problem of cyber security is that several solutions must be present at each network tier in order to meaningfully refer to a network as "secure." Though a very significant role, machine learning is only one aspect of a larger picture. Consider how many workers it would take to manually evaluate a million records to get an estimate of how much. Deep learning-based NIDS implementation would significantly reduce the strain now imposed on resources during the detection phase. There are inherent hazards with this. Maintaining the highest level of integrity is essential in the continuous use of these huge volumes of data since it is possible for one person or a group of people to draw conclusions or inferences from them. This alludes us to the ACM Code of Ethics and Professional Conduct [9]. From a broad perspective, the professional has an obligation to "contribute to society and human well-being." However, there are several occasions where study findings might unintentionally injure a person or a group of people. Even worse, there are situations where these actions might be planned.

Examples of the mishandling of data can be found in across different industries from the accidental termination of employees to the intentional manipulation of earnings reports [10][11]. A scientist with access to the data will ethically handle any nuances that are attributed to the dataset throughout their engagement in the research and report in order to align these incidents with the theme of NIDS. For instance, a training set contains false positives for IP addresses that are actually from satellite offices that are part of the company's known network, despite the scientist's knowledge that the addresses belong to a dubious location. However, it would be unethical for a scientist to describe the findings of their research on behalf of the organization for whom they are doing the study while failing to disclose any biases that may have been present. Even more unethical would be if the scientist were to alter the training data for this particular satellite office out of personal animus toward a certain person and then later claim that it was an oversight.

Using a dataset with known biases would be considered unethical in the context of the earlier example if these results were to be utilized to make choices. The ethical conclusion in this situation would be that any findings of the study would be deemed useless because of the biases that are known to exist. These factors explain why ethics continues to be essential to the advancement of deep learning algorithms.

6 Conclusion

We note that the autoencoder has a classification accuracy of 98.9% for the different attack types. The Long Short Term Memory (LSTM) model, in comparison, produced a score of 79.2%. To increase the accuracy of the LSTM model, additional hyperparameter adjustment is probably necessary. Because these models' predictions are dependent on the training data, a class imbalance may be the root of their decreased accuracy. The self-taught learning model reduces the number of features in the autoencoder to 10 as a result of dimensionality reduction. Compared to SMR results on the cleaned NSL-KDD dataset, which produced a much lower 75.23% when accounting for all 41 features in the original dataset, the outcome is more accurate. We can get the conclusion that a good model for NIDS is the autoencoder deep learning algorithm.

The STL model might be applied in a situation when the data is not clean. However, it's vital to remember that the ideal method for using any model is to make sure the data is accurate. The analysis between the deep learning models shows that using deep learning in NIDS would be a good way to increase detection accuracy on dirty data; however, creating an environment that is specifically made for this use would greatly influence the choice of which model would work best in a particular environment

DOI:10.46647/ijetms.2023.v07i04.043 ISSN: 2581-4621

References

1. Evans, D.: The Internet of Things How the Next Evolution of the Internet Is Changing Everything (2011)
2. Gers, F.: Long Short-Term Memory in Recurrent Neural Networks (2001)
3. Niyaz, Q., Sun, W., Javaid, A.Y., Alam M..: A Deep Learning Approach for Network Intrusion Detection System (2015)
4. Amad, B., Jian, W., Hassan , B., Rehmatullah , S.: Hybrid Intrusion Detection Method to Increase Anomaly Detection by Using Data Mining Techniques. International Journal of Database Theory and Application (2017) 231-240
5. W. Lee and B. Rotoloni, Emerging Cyber Threats, Trends & Technologies Report. Georgia Tech Institue for Information Security & Privacy (2004)
6. Bruneau, G.: The History and Evolution of Intrusion Detection. https://www.sans.org/reading-room/whitepapers/detection/history-evolution-intrusiondetection-344 (2001)
7. Bunel, P.: An introduction to Intrusion Detection Systems. London, England: SANS Institute. https://www.giac.org/paper/gsec/4227/introduction-intrusion-detectionsystems/106775 (2017)
8. Paliwal, S., Gupta, R.: Denial-of-Service, Probing & Remote to User (R2L) Attack Detection using Genetic Algorithm, 60th ed. International Journal of Computer Applications (2017)
9. ACM Code of Ethics and Professional Conduct, Acm.org (2017)
10. Chew, J.: Marissa Mayer Just Fired Dozens of Yahoo Employees By Accident. Fortune (2016)
11. Trainer, D.: Four Reasons Executives Manipulate Earnings. Forbes (2017)
12. Dong B., Wang X.: Comparison Deep Learning Method to Traditional Methods Using for Network Intrusion Detection (2016)
13. Kruegel, C., Mutz, D., Robertson, W., Valeur. F.: Bayesian event classification for intrusion detection. Proc. 19th Annual Computer Security Applications Conference (2003) 14-23
14. Sinclair, C., Pierce, L., Matzner, S.: An application of machine learning to network intrusion detection. Proc. 15th Annual Computer Security Applications (1999) 371-377
15. Zhang, J., Zulkernine, M.: A hybrid network intrusion detection technique using random forests. Proc. First International Conference on Availability, Reliability and Security (ARES'06), April (2006) 8-16
16. Yang, J., Deng, J., Li, S., Hao, Y.: Improved traffic detection with support vector machine based on restricted Boltzmann machine. Soft Computing, vol. 19 (2015) 1-12
17. M. Lincoln Labs. DARPA Intrusion Detection Evaluation. http://www.ll.mit.edu/IST/ideval (1999)
18. Hinton, G.E.: The "wake-sleep" algorithm for unsupervised neural networks. Science 268.5214 (1995) 1158
19. Hinton, G.E., Osindero S., Teh Y.W.: A fast learning algorithm for deep belief nets. Neural computation 18.7 (2006) 1527-1554
20. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H.: Greedy layer-wise training of deep networks. Advances in neural information processing systems (2007)
21. Torres, P., Catania, C., Garcia, S., Garino, C.: An Analysis of Recurrent Neural Networks for Botnet Detection Behavior (2016)
22. Hayun, L.: Demystifying Machine Learning (‘Artificial Intelligence’) Use in Endpoint Security Products. Proc. Palo Alto Networks Ignite 2017 Conference (2017)
23. IT Industry Outlook 2017. CompTIA. https://www.comptia.org/resources/it-industrytrends-analysis-2017 (2017)
24. Building digital trust: The role of data ethics in the digital age. Accenture (2017)