Benha Journal of Applied Sciences (BJAS) print: ISSN 2356–9751
Vol. (8) Issue (5) (2023), (89-104) online: ISSN 2356–976x
http://bjas.journals.ekb.eg

Recent Studies and A Review about Malware detection and classification by using Artificial Intelligence Techniques

Abdelrhman Samy, Ahmed A. El-Sawy, and Fatma Sakr
Faculty of Computers And Artificial intelligence, Department of Computer science, Benha University
E-mail:- abdelrahman.mohamed21@fci.bu.edu.eg

1. Abstract

Due to the harmful and unsafe broad utilization of Malware emergency as a result of various sorts of malware, perilous programs, and scripts that are accessible on the tremendous virtual world known as the Web. This study centers on learning around most later different sorts of malware and strategies to induce freed of them by finding them and kicking them out of the framework, which isn't simple since these little pieces of script or code can be found all over within the client framework. In this paper, we highlight malware collection, conglomeration, and dispersal challenges in client framework environment and show a comprehensive dialog on the later ponders that utilized different AI strategies to meet particular destinations of most malware location frameworks, from 2017 to 2022. We compare and differentiate diverse calculations based on optimization criteria, recreation, genuine sending, malware sorts, and execution parameters. We conclude with conceivable future inquire about headings. This would direct the peruser towards an understanding of up-to-date applications of ML methods concerning malware acknowledgment, accumulation, and spread challenges. At that point, we offer a common assessment and comparison of diverse ML strategies used, which is able be a direct for the investigate community in recognizing the foremost adjusted strategies and the benefits of utilizing different AI and machine learning strategies for tackling the challenges related to getting freed of these destructive malware. At long last, we conclude the paper by expressing the open issues of investigate and unused conceivable outcomes for future ponders.

Keywords: deep learning, machine learning, malware and software defined networking.

2. Introduction

Each day, the number of unused mal- product variations develops. Agreeing to G-research Data's on Malware Patterns in 2021, there was a 45 percent in- wrinkle in modern malware variations in 2020 [1]. In expansion to recognizing perilous computer program utilizing heuristic and signature-based strategies that battle to keep up with malware development, certain novel machine learning-based strategies have as of late been explored. Machine learning calculations such as Bolster Vector Machines (SVM), Irregular Woodlands (RF), Gullible Bayes (NB), or Neural Systems (NN) are utilized to analyze malware to illuminate the issue of recognizing and classifying perilous computer program. To extricate highlights from malware, both energetic and inactive investigation are utilized; a few of the ways incorporate: (i) Utilizing n-gram investigation to clarify a grouping of malware's hex values, op-codes, and other characteristics. (ii) API and work calls have long been utilized to recognize unsafe computer program. (iii) Malware is treated as an picture: the vector of bi- nary zeros and ones recovered from the malware record was bended into a framework so that the malware record might be seen as a gray-scale picture, and after that classified utilizing SVM, K-Nearest Neighbor (KNN), or Convectional Neural Net- work (CNN).

2.1. Research methodology

The strategy utilized in this investigate is part into two fundamental parts, Articles choice and Articles classifications, the articles choice stage (Stage 1) is comprised of two steps: database source determination and article determination and sifting. The classification of articles is secured in Stage 2.

2.1.1. Articles selection phase

Database sources choice way were made firstly by measure of information at that point in moment part by number of utilize in papers and within the third part by sort of information (gray-scale picture or http re-journeys). Articles determination and sifting way product understanding of other inquire about related words, brief. We carried out look by selecting one of the AI watchwords with Malware discovery watchword. The look comes about are restricted to diary articles and the extend of a long time (2017 to 2022) figure 1. At that point all the look comes about are combined and sifted.

2.1.2. Articles classification

We outline a malware discovery scientific classification based on machine learning approaches figure 2. Agreeing to this figure, the API calls highlights, gathering highlights, and double highlights are existing approaches for malware discovery strategy. These highlights utilize ma- chine learning strategies for foreseeing and recognizing noxious records.

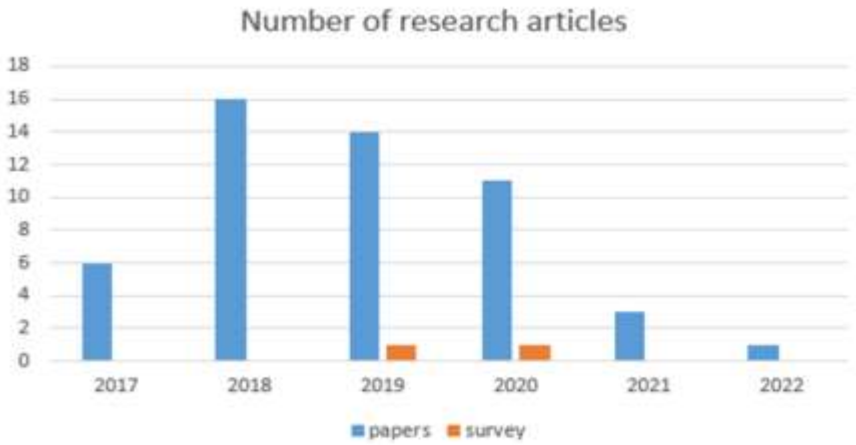

Fig. (1) Number of research articles selected per year from 2017 to 2022

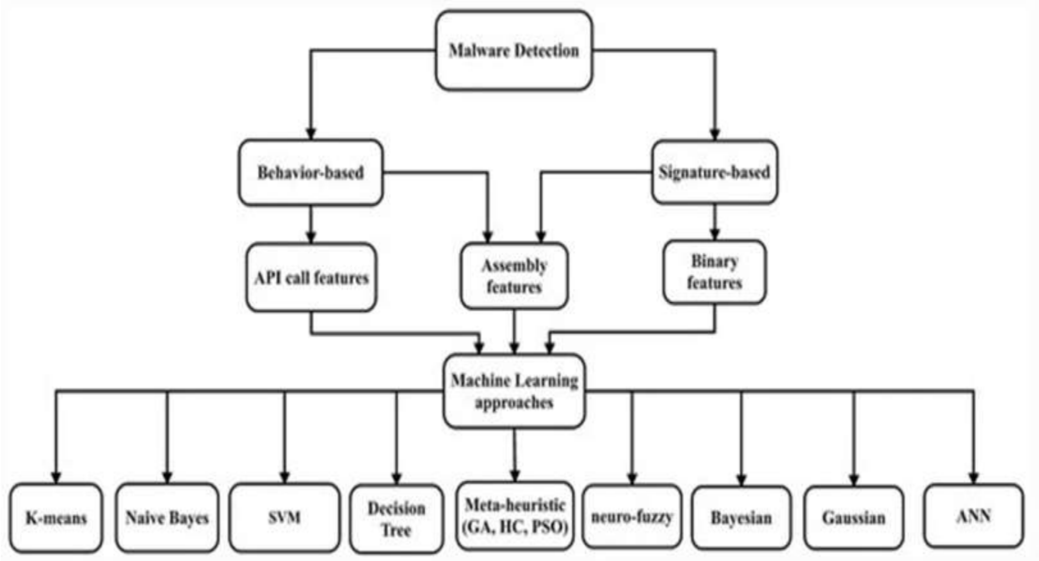

Fig. (2) papers classification algorithms

2.2. Paper abbreviations

Table (1) Table of Main Used Abbreviations

Abbreviation	Description
DPN	Dual Path Network
SVM	Support-Vector Machine
CNN	Convolutional Neural Network
KNN	K-Nearest Neighbor
MLP	Multilayer Perceptron
LBP	Local Binary Patterns
ASVM	Active Learning Support Vector Machine
FCM	Fuzzy C-Means
LGBM	Light Gradient Boosting Machine
RF	Random Forest
LSTM	Long Short-Term Memory
TF-IDF	Term frequency–inverse document frequency
SDN	Software-Defined Networking

2.3. Paper contributions

About the papers surveyed, different AI methods in SDNs have been checked and grouped. An outline of these procedures is introduced at first. This characterization of AI procedures is then utilized during the conversation of each test in SDNs to show how AI methods dealt with each test. A few papers might cover numerous view- points and will be overviewed for every classification. The paper examinations the exploration appropriation and patterns that portray the utilization of AI in SDN. Moreover, the paper distinguishes difficulties, promising examination headings in applying AI-based answers for different SDN challenges, with the plan to advance and work with additional exploration. In summary, the following are the survey's major contributions:

1. We review the existing AI techniques and their applications in SDNs to overcome the challenge issues of SDNs.
2. We present an overview of the major challenges in SDNs and the various AI techniques to Malware classification and Malware detection challenges.
3. A comprehensive discussion on the recent studies that utilized various AI methods to meet specific objectives of SDN during the span of 2017 to 2021 is given.

3. Artificial Intelligence Techniques

The field of machine learning (ML) is committed to creating frameworks that can consequently learn from information and recognize covered up designs without genuine programming. ML calculations are classified based on the learning fashion they utilize and the useful similitude in how they work. Here's an diagram of machine learning approaches based on your learning fashion. Machine learning strategies are considered viable ways to extend location rates, decrease untrue cautions, and decrease computational and communication costs. Machine learning strategies can be partitioned into administered learning, unsupervised learning and semi-supervised learning.

In directed learning, an calculation learns representations from settled input information to anticipate questionable circumstances. Illustrations of directed machine learning calculations incorporate back vector machines (SVM) for classification issues and arbitrary woodlands for classification and relapse issues. The Bolster Vector Machine (SVM) calculation is broadly utilized in NIDS inquire about due to its capable classification capabilities and computational utility. In spite of the fact that it is reasonable for high-dimensional information, it is critical to select a sensible bit work. It is asset seriously and requires number-crunching preparing units and memory.

The Random Forest Calculation could be a capable ensemble-based learning approach planned to work productively with different information, but it goes past. In an unsupervised learning plot, an calculation learns the structure and representation of unlabeled input information. The reason of unsupervised learning calculations is to show the fundamental structure or dissemination of information to foresee obscure information. Illustrations of unsupervised learning calculations incorporate highlight diminishment procedures such as central component examination (PCA) and clustering methods such as self-organizing maps (SOM). Foremost Component Examination (PCA) is an calculation utilized to greatly speed up unsupervised include learning. Numerous analysts utilize PCA for highlight choice some time recently applying classification. Clustering calculations like K-Means and other remove learning calculations are utilized to distinguish inconsistencies.

Self-organizing outline (SOM) is an counterfeit neural arrange utilized to decrease the burden of NIDS. A impediment of employing a clustering calculation for peculiarity discovery is that the clustering calculation depends on starting conditions such as the center of gravity, which can create a tall untrue positive rate. Semi-supervised learning may be a sort of directed learning that employments unlabeled information for preparing. The preparing information comprises of little sums of labeled data and large sums of unlabeled information. Reasonable for circumstances where expansive sums of labeled information are not accessible, such as photo files where as it were a few of the pictures (eg individuals) are labeled and most are unlabeled.

To progress the precision of NIDS, a semi-supervised back vector machine was utilized. Two semi-supervised classification strategies utilized to identify obscure assaults, a ghastly chart converter and a semi-supervised clustering strategy from the MPCK asset utilized to progress the execution of the Gaussian field guess and discovery framework. Profound learning calculations are the most up to date upgrade to manufactured neural systems that use large-scale and reasonable computing. Profound learning permits calculations to memorize how to speak to information with distinctive levels of generalization. These methods have been connected to visual protest acknowledgment, question location, arrange interruption discovery, and numerous other areas. Profound learning calculations can be prepared using both directed and unsupervised strategies. Administered profound learning calculations, such as convolutional neural systems (CNNs), are ordinarily prepared on directed information. CNNs are presently benchmarks in computer vision. CNN design is utilized to construct

2D images and one of the most highlights of CNN is confront acknowledgment.

Unsupervised Deep Learning Calculations: Autoencoders are utilized to memorize representations (encodings) of datasets for dimensionality lessening. Profound Belief Networks (DBNs) can learn to recreate their inputs when prepared unsupervised employing a set of illustrations. The layer at that point acts as a include finder on the input. After this preparing stage, DBNs experience extra administered preparing to perform the classification. DBNs, such as Confined Boltzmann Machines (RBMs) or Autoencoders, can be utilized for dimensionality lessening, relapse, collaborative sifting, include learning, and point modeling, among others. Utilizing directed or unsupervised calculations is considered a directed or unsupervised learning strategy such as Repeat. - Leasing Neural Systems (RNNs). RNNs can process subjective input arrangements utilizing inner memory. Discourse recognition is a common application of RNNs. RNNs are great at anticipating characters in content and can learn long conditions and particular contentions.

4. Previous Work

Programming Software Defined Networking Technology (SDN) gives a possibility to actually identify and screen network security issues attributing to the rise of the programmable highlights. As of late, Machine Learning (ML) approaches have been executed in the SDN-based Network Intrusion Detection Systems (NIDS) to safeguard PC organizations and to survive network security issues. A surge of cutting edge AI draws near - the profound learning innovation (DL) begins to arise in the SDN setting. In this study, they audited different late chips away at AI (ML) techniques that influence SDN to carry out NIDS. All the more explicitly, they assessed the methods of profound learning in creating SDN-based NIDS. In [2], they covered devices that can be utilized to foster NIDS models in SDN climate. This review is finished up with a conversation of continuous difficulties in executing NIDS utilizing ML/DL and future works.

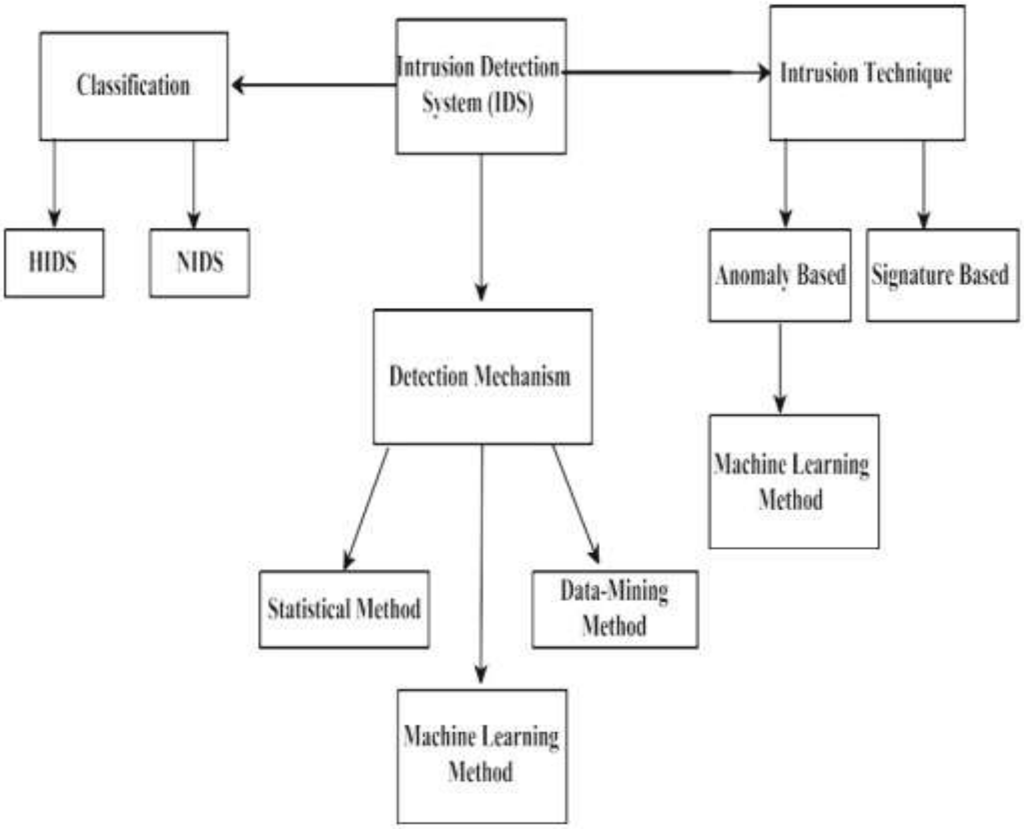

Fig. (3) Overview of intrusion detection system

However, in [3] they discussed a number of the current approaches and difficulties in the categorization of malware, including data gathering, feature extraction, model design, and evaluation. They discussed the limitations that must be taken into account for machine learning-based solutions in this space as well as unresolved issues for which machine learning may also offer a solution.

5. Intrusion Detection System

5.1. Malware classification

Malware records are open and dynamic in the real world, and new malware samples from old and new classes are continually being added. They need a malware classification approach that allows for incremental learning to learn new information quickly. However, so far, the effort has mostly concentrated on feature engineering, which use machine learning as a tool. They will use an incremental malware classification framework named ”IMC” to address this challenge, which comprises of op-code sequence extraction, selection, and incremental learning approaches. Based on a thorough support vector machine (SVM), they devised an incremental learning approach [4]. A machine learning honeypot is another technique to identify malware. Machine learning can detect malware by classifying it into classes, and honey- pots operate as traps for suspicious packets [5].

System administrators should re- view the requirements before putting them on the server to ensure security. A rogue link based on TF-IDF technology to evaluate the ”importance” of a term or symbol in a user request link from a user’s browser using machine learning techniques is one of the most frequent approaches. What is a web application firewall with a detection algorithm? Improves the precision with which illicit links are identified [6].

In [7] paper, they show Cardi-Wall, a novel location and avoidance framework outlined to secure ICDs from cyber-attacks pointed at the software engineer gadget. Their framework has six distinctive layers of assurance, leveraging restorative experts' information, factual strategies, and machine learning calculations. They assessed the Cardi-Wall framework broadly in two comprehensive tests. For the assessment, they accumulated information for a period of four a long time and utilized 775 kind clinical commands that are related to hundreds of diverse patients (gotten from distinctive software engineer gadgets found at Barzilai College Therapeutic center) and 28 pernicious clinical commands (made by two cardiology specialists from distinctive healing centers). The assessment comes about appear that as it were two out of the six layers proposed in CardiWall framework given a tall location capability related with tall rates of genuine positive, and moo rates of untrue positive.

Other publications identify relevant patterns for identifying as many harmful computer attacks as possible and categorizing them as methods for predicting future criminal activities using intrusion detection systems. The goal is to collect evidence of suspicious activities and improve network security. To that purpose, this research looks into computer crimes that occur over time through the transportation network of a petrochemical subsidiary as a result of attack modeling utilizing the SVM algorithm. [8] The topic of viruses and Trojan horses attacking network protocols in the application layer of industrial control systems is investigated using Modbus / TCP communication protocol rules. Sup- port and Clustering they suggested a vector machine-based approach for detecting intrusion attempts. To deter- mine the communication data of an industrial control network and the distance between cluster centers, this approach blends unsupervised fuzzy c- means clustering (FCM) with a super- vised support vector (SVM) machine [9].

Characterization of the problem by identifying several scenarios based on the availability of valid and attack data for training. They also offer two alter- natives. The first is a multiclass approach to valid and attack data in cases when both are present. Second when only valid data is provided, a one class solution [10].

Another option is to create a three- phase procedure that confirms the identification of all sorts of malware on the host. It also protects you against zero-day assaults. Zero-day attacks are notoriously difficult to defend against and can be extremely harmful [11].

Another paper was used as Associate in nursing aggressive system for net applications and varied operative systems. The projected methodology achieved Associate in nursing accuracy of 96.111 for network intrusion detection, however due to this document variety [12]

Paper [13] anticipated CNN show is hybridized with a bolster vector machine. Instead of mis-treatment Softmax as enactment perform. SVM performs the assignment of classifying the malware upheld alternatives extricated by the CNN demonstrate. The anticipated fine-tuned demonstrate of CNN produces a well chosen choices vector of 256 Neurons with the FC layer that's input to SVM. Straight SVC bit changes the parallel SVM classifier into multi-lesson SVM , that classifies the malware tests abuse the one-against- one technique Relate conveys an exactness of 99.59%.

In [14], [15] they absolutely were found that artificial neural network (ANN)-based machine learning to pick out wrapper options outperforms sup- port vector machine (SVM) ways for classifying network traffic. To gauge the performance, the NSLKDD dataset was employed to classify net- work traffic mistreatment machine learning techniques below the supervising of SVMs and ANNs.

Paper [16] reports the use of as a data mining method for handling custom malware using 4gram API calls. For comparison, a set of 552 Windows Portable Executable (PE) files were used for API calls, and PE was run in a Windows 7 virtual environment using the cuckoo sandbox. The highlight of that API 4gram call is excluded using the term frequency in- verse document frequency (ofidf). In- formation was prepared and tested using Naive Gauss Bayes, SVM, random forest, and decision trees. They show that this method is effective with an accuracy of 92 to 96.4. At exactly there is inherent variability, SVMs and decision trees work best, and Gaussian naive Bayesian methods give very horrendous results.

In [17] too utilized progresses in common dialect handling and a few strategies rain. Utilize ngram, doc2vec (or section vectors), TFIDF to change over these API bunches to numeric vectors some time recently utilizing the classifier. The proposed approach can be partitioned into three unmistakable ways to characterize official malware. That's, a passage vector with words disseminated and a section vector with dispersed memory. But moreover can accomplish tall location rates in genuine time when conveyed on the well-known cloud stage OpenStack as appeared in [18]. In a study based on Spark MLlib (Machine Learning Library), efficient and intelligent electronic identifiers are extracted using modern DLs such as Convolutional Automatic Encoder (ConvAE) for misuse attacks and powerful classic machine learning classifiers for anomaly detection, as cleared in [46]. A system that detects and classifies unpredictable malicious attacks was developed. To measure the effectiveness of the pro- posed identification system, they used several important performance metrics such as FAR, DR, and accuracy, and experiments were performed using publicly available datasets, especially the latest heterogeneous dataset CSE- CICIDS2018.

The paper [20] proposed a convolutional neural network-based multi- component classification engineering show for anticipating progressed abilities in resumes, indeed in case not expressly said in resumes. In tests con- ducted with mysterious IT resumes collected from the Web, the viability of our strategy was found to reach 98.79 memory and 91.34 exactness. In expansion, capacities (terms) recognized by the convolution channel are anticipated onto the input rundown, giving the client the terms that impacted the show choice. Numerous analysts have proposed machine learning-based IDS for interruption discovery, and presently they are moving to profound learning for more prominent precision. Profound learning can give a noteworthy level of conceptual data by dynamically coordination essential capacities into complex capacities layer by layer.

Another paper [21] proposes a CNN-based architecture for classifying malware samples. It transforms the malware binary into a grayscale image and then trains the CNN to classify it. Experiments with Malimg and Microsoft malware, two complex datasets for malware classification, show that our method outperforms the current one. The proposed method provides
98.52 and 99.97 accuracy for Malimg and Microsoft data sets, respectively. Classify images with multi-class net- work attack classifiers that can be deployed on routers. Genetic Algorithms (GAs) are used to find high-quality solutions by changing the layout of in- put features and reducing the number of other features as needed.

In [22] presents a novel hyper- chart based peculiarity disclosure strategy with upgraded central component investigation and convolution neural arrange (EPCAHGCNN) to recognize crack hones of such systems. The proposed EPCAHGCNN strategy incorporates two stages: (1) Dimensionality diminish utilizing overhauled PCA and (2) inconsistency revelation with hypergraph based convolution neural net- work. Encourage, the show of that strategy is surveyed with SUTD's Se- remedy Water Treatment Framework (SWaT) and the test comes about appear that the proposed EPCAHGCNN has recognized unusual conduct of the data with tall revelation rate, moo sham up- sides, and way better gathering precision. Significant learning methods can fulfill tall precision with a moo fake caution rate to recognize arrange intrusions.

In [23] the approach utilized a mixture calculation of convolutional neural network (CNN) and long short term memory (LSTM). The approach is related to further developed interruption identification. This bidirectional calculation showed the highest known ac- curacy of 99.70 of the standard data set known as the NSL KDD. The performance of this calculation is assessed using an outline that looks promising for the organization based on its accuracy, false positives, F1 scores, and real tissue.

The structure used in paper [27] has improved performance with a sophisticated one with 94.4 curacy of one-dimensional convolutional neural network (CNN) and long short-term memory (LSTM) networks. They got 84.9. Transboundary tuning is accomplished through deep engineering learning to set boundaries. A learning rate of 0.01 was taken for all exams. A 70-30 training test split was completed during the course of the experiment. It works in conjunction with observing how well the model performs with disproportionate information gathering. Common strategies such as dis- assembly, de-compilation, reverse congestion or pairing do not require this proposed method. Paper [25] used a character-level convolutional neural network (CLCNN) with exceptionally massive world-wide pooling to segment HTTP request elements and recognize them as benign or malicious requests. They evaluated the structure of the http dataset CSIC 2010 dataset and achieved an accuracy of 98.8 in 10x cross-validation, with a typical processing time per request of 2.35 ms In paper [26], they have proposed a clever interruption location framework model dependent on model based interpretability, called Interpretable Intrusion Detection System (I2DS). First and foremost, typical and assault tests are recruited via AutoEncoder (AE) with preparing tests to feature the ordinary and assault highlights, so the classifier has an exquisite impact. Added substance tree (addtree) is utilized as a double classifier, which can give fantastic prescient execution in the joined dataset while keeping up with great model based interoperability. In the test, UNSWNB15 dataset is utilized to assess their proposed model. For identification execution, I2DS accomplishes a location exact- ness of 99.95, which is better compared to the greater part of best in class interruption recognition techniques. Additionally, I2DS keeps up with higher simulation and catches the choice principles without any problem.

In paper [27] proposes a significant learning structure utilizing long transitory memory (LSTM) plan for a estimate of the region names that are created utilizing the DGAs. Twofold gathering had safe and DGA zone names and multiclass arrange were per- shaped utilizing 20 particular DGAs. For the twofold characterization, LSTM demonstrate gave precision of 98.7 and 71.3 on two differing test instructive col- lections and for the multiclass arrange, it gave accuracy of 68.3 and 67.0 independently. Two broadened educational records were utilized to break down the quality of the LSTM plan.

In [28] proposed a way to further reduce the component aspect by selecting factor based on the encoding, the study is divided into three publicly available Windows programming datasets. Primer test results show that the proposed procedure can successfully discriminate between malware families.

In [29], [30], and [31] there work demonstrates a hybrid semi- supervised machine learning technique that uses Active learning Support Vector Machine (ASVM), Fuzzy C- Means (FCM), Decision Trees, Multi- layer Perceptron and Multi SVM clustering in the design of an efficient IDS. This algorithm is tested on NSL KDD bench mark IDS data set and found to be promising.

5.2. Malware detection

As feature learning approaches, the other proposed method employs Stacked Auto encoder and Deep Belief Network. The training phase is classified using only regular data, using a single class of SVMs, isolated forests, and elliptical envelopes as classifiers [1]. They will employ these SDN capabilities to demonstrate the architecture of a hierarchical and lightweight intrusion detection system (IDS) for software-defined networking using the notion of SDN flow in the upcoming study [33].

Multiclass Support Vector Machine (SVM) classifier is used to classify logs generated by your firewall de- vice. SVM classification uses linear, polymorphic, sigmoid, and radial ba- sis function (RBF) functions as activation functions, the user has devised a method to make things easier. Other articles have proposed a support vector machine (SVM), which is a problem in terms of image processing for multiclass malware picture classification. Gabor wavelets, GIST, discrete wavelet transforms, and other features are utilized to generate effective texture feature vectors employing several resolutions and wavelets [34].

In [35] sort Suricata IDS/IPS is sent with a NN demonstrate for detached metaheuristic location of noxious activity on the target organize. Amid this ponder, numerical rationale was utilized for metaheuristic-based highlight choice, neural arrange, and peculiarity- based location, and thus the most recent steady Kali UNIX form 2020.3

In [36] it used a mixture of super- vised and unsupervised machine learning to higher classify malware whereas maintaining a suitable level of accuracy and reducing coaching time. The answer with reference variety needs an easy pure mathematics real number to classify malware supported representative digital pictures. They have a tendency to measure performance by reviewing publically obtainable image information sets Malimg, Ember, and large 2015. Their execution investigation demonstrates that their classifier outflanks state-of-the-art models and achieves classification exactness's of 0.998, 0.911 and 0.997 abuse Malimg, coal and expansive 2015 malware datasets, severally.

But in paper [37] it take manner adept grouping strategy is significant to beat the problem. Known AI ways particularly SVM, multi-class SVM, KNN, binary classification (BC) square measure connected. These methods square measure understood within the lightweight of their classification capabilities. The NSL dataset is employed, that could be a learning revelation and data mining that's thought-about a benchmark for evaluating interrupt-aware elements.

The comes about appear that Multiclass SVM bypasses elective techniques. The execution of such CNNs as include extractor's abuse back vector machines (SVMs) and closest neighbors (KNNs) for classification capacities inside the paper [38]. It also gives combining strategies to extra move forward execution. This ponder employments a publically reachable data given by the Microsoft Malware Classification Challenge (Huge 2015). Our add up to outturn sets a substitution standard at 99.4 on a bunch of two, 174 take a see at tests in nine grades.

Comparative ponder appears that the anticipated show is conservative than elective existing models with significance interruption location victory rate. In elective discovery framework is utilized in analyzing endless activity information; hence, Relate in nursing conservative classification method is vital to defeat the issue. This drawback is taken into consideration in dad- per [39] Well-known machine learning techniques, namely, SVM, irregular timberland, and extreme learning machine unit of measurement connected. These strategies area unit well-known inferable to their capability classification. The NSL discovery and mining dataset is utilized and is taken into consideration a benchmark for assessing interruption discovery instruments. The comes about appear that ELM is predominant to elective approaches. Comparative ponder appears that the anticipated show is conservative than elective existing models with significance interruption location victory rate. In elective discovery framework is utilized in analyzing endless activity information; hence, Relate in nursing conservative classification method is vital to defeat the issue. This drawback is taken into consideration in dad- per [39] Well-known machine learning techniques, namely, SVM, irregular timberland, and extreme learning machine unit of measurement connected. These strategies area unit well-known inferable to their capability classification. The NSL discovery and mining dataset is utilized and is taken into consideration a benchmark for assessing interruption discovery instruments. The comes about appear that ELM is predominant to elective approaches

In [40] confirming the model presented by was generated using a lean auto-encoder. This is a learning computation that can be performed to reconstruct representations of other components without assistance. Later in, new key points are taken into ac- count in the SVM calculation to further develop the ability to detect interruptions during the preparation phase and improve the accuracy of orders. They also examine the effectiveness of the methodology in parallel with multi-class ordering and compare it with planar feature estimation strategies such as J48, Creduras Bayesian, Irregular Forest, and SVM. The results show that our methodology reduces the time to prepare and test SVMs and outperforms most previous measurement approaches to characterize two and multiple classes.

The proposed STLIDS approach further develops network interrupt identification and provides another discovery method for interrupt detection. Some AI classifiers with alternative requirements and test rates on paper [41] thought about accuracy.

The results show that applications in the same category detail more accurate execution when detecting malicious applications as opposed to non-class applications. Backwoods' random classifier, which received 10x mutual approval using highlight positioning and data highlight definition, achieved high true positive scores (Book and Reference and Personalization), they also reached low false- positive rates ratio.

Each of them usually gives a good accuracy of paper [42] uses the smg-scnn (Static Convolutional Neural Net- work of Malware Gene Sequences) module to validate groups of malware quality and create a neural network that checks for malware sequences. Test results show a significant improvement in performance accuracy, up to 98 for the MCSMGS model. CNN models are more practical than traditional SVM models.

A Convolutional Neural Arrange (CNN)-based approach for real-time location of pernicious behavior of cloud clients. The proposed approach in [43], naturally learned a few arrangement designs from the grouping of framework calls and created a CNN demonstrate to proficiently distinguish noxious occupant behaviors. It moreover employments Start spilling innovation to handle expansive volumes of client information in genuine time within the cloud. Exploratory comes about appear that the proposed approach not as it were beats the three existing strategies, be that as it may, [44] proposed an anomaly-based arrange interruption discovery framework that employments a profound learning approach to SDN to identify different and obscure sorts of assaults.

In paper [45], they propose a Malscore classification framework based on likelihood scoring and machine learning that sets a likelihood sift- ancient for combining inactive examination (alluded to as step 1) and energetic examination (alluded to as step 2). In step 1, grayscale pictures (static functions) were analyzed employing a convolutional neural organize combined with spatial pyramids, and in step 2, groupings of fundamental API calls (energetic capacities) were analyzed utilizing ngram factors and machine learning. In [46], they change NetFlow information into NetFlow pictures utilizing highlight relationship examination and minimal relationship network (SC) to extricate and encode highlights from NetFlow information distributed in Endless 2013. The created NetFlow pictures were sent to the CNN demonstrate. The comes about appear that the proposed approach permits interruption location with 95.86 precision.

In [47] Testing was performed utilizing two freely accessible information sets with diverse assault rates:

UNSW (10 classes) and NSLKDD (4 classes). Both classifiers accurately separate between attacker and ordinary activity. Be that as it may, to accurately classify an assault, the last mentioned works way better since it can be relative between diverse classes, coming about in a cross-validated multi-class classifier with a K of 0.95.

In paper [48], a sequence structure of equipment including convolutional neural organization and discontinuous neural organization was proposed for malicious code recognition. This sur- vey used the Microsoft malware classification challenge (big 2015) dataset, which contains nine error-free classes.

This dataset contains outline reports and sorted sections for each malware. Organized records are imagined as pictures and are characterized utilizing convolutional neural net- works (CNNs). Get together records include of machine dialect opcodes that are recognized among classes utilizing long shorter memory (LSTM) systems consequent to changing them in progression. Additionally, highlights are expelled from these models (CNNs and LSTM) and are requested utilizing a offer assistance vector machine or key backslide. An precision of 97.2 is finished utilizing LSTM organize for isolating get together records, 99.4 utilizing CNN designing for gathering accumulated records and a common exactness of 99.8 utilizing the proposed equip approach in this way setting another benchmark. An independent and computerized characterization system for gathering or possibly requested archives gives the benefit to mal- product industry pros to choose the sort of system depending upon their available computational resources.

In paper [49] the system chair- man contributes a extraordinary bargain of vitality in keeping the system running, however they might fall flat to keep in mind the code implantation attacks. Along these lines, the essential task for system chiefs is to recognize organize as- saults at the application level utilizing a web application firewall and apply compelling calculations in this fire- divider to plan web application fire- dividers normally for extending their capability. The article presents a definition of the errand for extending the precision of address arrange by the subjective forest procedure, in this way, making the reason for recognizing as- saults at the application level. In paper [50] utilizing the same calculation, they proposed the utilization of a get together classifier, moreover called irregular timberland, that works within the presentation of other striking calculations by conglomerating person course desires to connect into a last figure.

In [51], the approach is to render the document as grayscale images and use DBNs to organize these images into the two previously mentioned classes. A vast information gathering of 10,000 documents is used to prepare the DBN. Approval is done using 4,000 documents that have not yet been submitted to the organization. The final result of whether a document is accepted or not is obtained by appearing in a proof-of-work contract on the block- chain network.

In [52] they find another type of malware in Portable document format (PDF) files, they presented detection technique that can examine PDF files to separate clean PDF files from malicious PDF files. The AdaBoost decision tree with optimal hyper parameters, which is trained and assessed on the contemporary inclusive data set known as Evasive-PDFMal2022, is used by the suggested system. With a prediction accuracy of 98.84 and a short prediction interval of 2.174 seconds, the investigative assessment reveals a simple and effective PDF detecting method.

We will now present a summary of Datasets in detail about the type of data, its number and year of issuance (table 2).

Table (2) benchmark datasets comparison

Num	Dataset name	Dataset	Dataset type
1	VX HEAVENS	5647	grayscale images
2	MALIMG	9339	grayscale images
3	MALICIA PROJECT	23080	grayscale images
4	NSL-KDD	25252	grayscale images
5	MALICIOUS PE’S	552	pe files
6	GOOGLE PLAY APPS	3413	android app
7	UNIVERSITY	5195	grayscale images
8	GOOGLE DRIVE	6000	grayscale images
9	HTTP CSIC 2010	25000	network requests
10	FIRAT UNIVERSITY	65532	network requests
11	PKDD	15110	network requests
12	DRUPAL	2226	network requests
13	CLAMP-RW-5184	5184	portable executable
14	CSDMC-API-TRAIN	776	portable executable
15	MICROSOFT	10868	grayscale images
16	NETFLOW	70m	network requests
17	BIG 2015	2174	grayscale images
18	KDDCUP-99	65535	network requests
19	EMBER	900000	network requests
20	SWAT	946722	network requests
21	ADFA-LD	5971	network requests

6. Summary of Papers

Here in this section the table will give a review and summary about the malware classification and detection variant methods which used artificial intelligence techniques.

Table (3) Summary of malware classification and detection by using Artificial Intelligence Techniques.

Paper No.	Name	Field	Algorithm	Accuracy	Year
1	Decentralized firewall for malware detection	Malware Detection	DBN	89.28	2017
2	MCSMGS: Malware Classification Model Based on Deep Learning	Malware Classification	SVM-CNN	98	2017
3	Malware Class Recognition Using Image Processing Techniques	Malware Classification	SVM	98.88	2017
4	NLP-based Approaches for Malware Classification from API Sequences	Malware Classification	KNN-SUM- MLP	98.7	2017
5	Binary Malware image Classification using Machine Learning with Local Binary Pattern	Malware Classification	LBP-SVMLBP -KNNLBP	93.17	2017
6	A semi-supervised Intrusion Detection System using active learning SVM and fuzzy c-means clustering	Malware detection	ASVM-FCM	99.6	2017
7	Malware Classification using API System Calls	Malware Classification	GNB-SVM-RD-DT	96.4	2018

8	Android Malware Classification Base On Application Category Using Static Code Analysis	Malware Classification	NB-SUM- DT-RF	98.2	2018
9	A zero-day resistant malware detection method for securing Cloud using SVM and Sandboxing Techniques	Malware detection	SVM	not result	2018
10	Malware Classification Using Machine Learning Algorithms	Malware Classification	DT-MLP- MSVM	98	2018
11	Deep Learning Approach Combining Sparse Auto-encoder with SVM for Network Intrusion Detection	Malware detection	J48-NB- RF-SVM	99.3	2018
12	Visual Malware Classification Using Local and Global Malicious Pattern	Malware Classification	LGMP	94.08	2018
13	Web Application Firewall using Character-level Convolutional Neural Network	Malware detection	CLCNN	98.8	2018
14	Classification of Firewall Log Files with Multiclass Support Vector Machine	Malware Classification	SVM	76.4	2018
15	Web Application Attacks Detection Using Machine Learning Techniques	Malware detection	SVM	95.34	2018
16	Multiclass network attack classifier using CNN tuned with Genetic Algorithms	Malware Classification	CNN	97.76	2018
17	An Anomaly Detection Method to Detect Web Attacks Using Stacked Auto-Encoder	Malware detection	SAE	88.32	2018
18	A Hierarchical Intrusion Detection System using Support Vector Machine for SDN Network in Cloud Data Center	Malware detection	SVM	99.42	2018
19	Methodology for Malware Classification using a Random Forest Classifier	Malware Classification	RF	89.67 - 98.72	2018
20	Malware Classification with Deep Convolutional Neural Networks	Malware Classification	CNN	98.99 - 99.97	2018
21	Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection	Malware detection	SVM-RF- ELM	99.5	2018
22	Research on Industrial Control Anomaly Detection Based on FCM and SVM	Malware detection	SVM	96.85	2018
23	Predicting Network Attacks with CNN by Constructing Images from NetFlow Data	Malware detection	CNN	95.86	2019
24	Investigation and classification of cyber-crimes through IDS and SVM algorithm	Malware Classification	SVM	99	2019
25	Malware classification using probability scoring and machine learning	Malware Classification	CNN	98.82	2019
26	Network Intrusion Detection using Supervised Machine Learning Technique with Feature Selection	Malware detection	SVM- ANN	94.02	2019
27	Survey on SDN based network intrusion detection system using machine learning approaches	Malware detection	survey		2019
28	Deep Learning Framework And Visualization for Malware Classification	Malware Classification	CNN- LSTM	95.5	2019
29	Convolutional Neural Networks as Classification Tools and Feature Extractors for Distinguishing Malware Programs	Malware Classification	CNN- SVM- KNN	99.4	2019
30	Leveraging deep neural networks for anomaly-based web application firewall	Malware Classification	SVM	89.24	2019

31	Intrusion Detection System Using Deep Learning for Software Defined Networks (SDN)	Malware detection	CNN	95	2019
32	An Improved Intrusion Detection System Using Support Vector Machine	Malware detection	SVM- KNN	89.02	2019
33	Algorithm for detecting illegal links using the association rule for improving the web attack detection accuracy of web application firewall	Malware detection	SVM	98.9261	2019
34	Deep Learning Framework for Domain Generation Algorithms Prediction Using Long Short-term Memory	Malware Classification	LSTM	98.7	2019
35	Malware Detection Using Honeypot and Machine Learning	Malware detection	SVM	90	2019
36	Enhancing Firewall Filter Performance Using Neural Networks	Malware Classification	ANN	-	2019
37	Ensemble Malware Classification System Using Deep Neural Networks	Malware Classification	CNN- LSTM	99.8	2020
38	Convolutional Neural Networks with LSTM for Intrusion Detection	Malware detection	CNN- LSTM	97.36	2020
39	Skills prediction based on multi-label resume classification using CNN with model predictions explanation	Malware Classification	CNN	99	2020
40	Malware Classification with Improved Convolutional Neural Network Model	Malware Classification	CNN- SVM	99.59	2020
41	Malware classification using compact image features and multiclass support vector machines	Malware Classification	MSVM	99.97	2020
42	Improving Efficiency of Web Application Firewall to Detect Code Injection Attacks with Random Forest Method and Analysis Attributes HTTP Request	Malware Classification	RF	98.72	2020
43	Detection of Cyberattacks in Industrial Control systems using Enhanced Principal Component Analysis and Hypergraph based Convolution Neural Network (EPCA-HG-CNN)	Malware detection	EPCA- HG- CNN	98.02	2020
44	Toward Developing Ecient ConvA E-Based Intrusion Detection System Using Heterogeneous Dataset	Malware detection	CNN	98.2	2020
45	A Trusted Firewall for the Detection of Malicious Clinical Programming of Cardiac Implantable Electronic Devices	Malware detection	SVM	94.7	2020
46	Incremental Learning for Malware Classification in Small Datasets	Malware Classification	IMCSVM	98	2020
47	Real-time detection of cloud tenant malicious behavior based on CNN	Malware detection	CNN	98	2020
48	Interpretable Intrusion Detection System Using Autoencoder and Additive Tree	Malware detection	ADDTREE- AE	99.26	2021

49	Optimization of a Depiction Procedure for an Artificial Intelligence-Based Network Protection System Using a Genetic Algorithm	Malware Classification	CNN	99.84	2021
50	Malware classification using Km-SVM	Malware Classification	KM-SVM	86	2021
51	The Anomaly- and Signature-Based IDS for Network Security Using Hybrid Inference Systems	Malware Classification	SVM - DECISION TREE	96.11	2021
52	Malware Detection Based on Optimizable Decision Trees	Malware Detection	SVM - DECISION TREE	98.84	2022

7. Open Researches Issues and Challenges

Based on all that was mentioned in the previous review (discussion), we found that there are some problems and challenges faced by the various previous scientific researches, and they are divided into the following problems, the first of which is the problem of the huge number of viruses present around the world, which is constantly renewed day after day, which makes it difficult for the process of teaching a machine to face all this The second challenge is the problem of integrating the sys- tem that monitors the system's Inter- net network from unknown requests to enter the system or files of unknown source, which makes it difficult to protect our system and slows the entry and exit of data There is no doubt that these problems constitute difficult and a great challenge for science and last but not least the problem of implementing a system that has the ability to monitor the network at the same time and protect it based on what it learned from previous experiences from viruses that it learned in the system learning stage

8. Analysis Review

After all the previous researches were searched, it became clear that some researches use dataset grayscale images and some researches use dataset network requests and some researches use dataset PE files, and some researches use the dataset calls, simple to notice and difficult to stow away. In any case papers pro- poses a malware grouping philosophy utilizing various calculations was proposed, the errand was completed utilizing two unique datasets every one of them involve tests of malware and harmless applications from Windows OS, coming about highlights were pre- pared by the android app but it was the best researches that used Dataset grayscale images in terms of efficiency and accuracy.

Also, what was concluded from the research is that there are researches that used the following algorithms (CNN – ANN – LSTM – SVM – CLCNN – RF – KNN – ADDTREE –MLP) and the best results were when using CNN with Microsoft dataset and with MALIMG dataset.

9. Conclusion

These days, with the headway of science and innovation and the increment in the degree of mindfulness and information on people, as seen step by step, there are boundless assaults on different business, correspondence, data, and different organizations. A portion of these assaults target significant and delicate information and data contained in these organizations, and some different assaults attempt to produce additional traffic in the correspondence courses with the assets and servers which are offering various types of assistance for clients. A few papers proposes a CNN-based methodology for recognizing constant malignant conduct in the cloud climate, which has a lot of information and utilizations Spark for ongoing investigation. It utilizes the framework call grouping to distinguish the conduct of the occupant, contrasted and other decimal standard distinctive calculation and contrasted and other notable calculations broadly utilized in malware identification draws near.

Acknowledgements

This work was supported by Department of Computer science, Benha University, Egypt.

References

[1] Einy, Sajad and Oz, Cemil and Navaei, Yahya Dorostkar, “The anomaly- and signature-based IDS for network security using hybrid inference systems” Mathematical Problems in Engineering,Hindawi, 2021.

[2] Sultana, Nasrin and Chilamkurti, Naveen and Peng, Wei and Alhadad, Rabei, “Survey on SDN based network intrusion detection2019tem us- ing machine learning approaches” Peer-to-Peer Networking and Applica- tions,Springer, 2019.

[3] Raff, Edward and Nicholas, Charles, “A survey of machine learning meth- ods and challenges for windows malware classification” arXiv preprint arXiv:2006.09271,arXiv, 2022.

[4] Li, Jingmei and Xue, Di and Wu, Weifei and Wang, Jiaxiang, “Incremental learning for malware classification in small datasets” Security and Commu- nication Networks,Hindawi, 2020.

[5] Matin, Iik Muhamad Malik and Rahardjo, Budi, “Malware detection using honeypot and machine learning” 2019 7th international conference on cyber and IT service management (CITSM),IEEE, 2019.

[6] Manh, Thang Nguyen, “Algorithm for detecting illegal links using the asso- ciation rule for improving the web attack detection accuracy of web applica- tion firewall” International Journal of Open Information Technologies,2019.

[7] Kintzlinger, Matan and Cohen, Aviad and Nissim, Nir and Rav-Acha, Moshe and Khalameizer, Vladimir and Elovici, Yuval and Shahar, Yuval and Katz, Amos, “CardiWall: a trusted firewall for the detection of mali- cious clinical programming of cardiac implantable electronic devices” IEEE Access,IEEE, 2020.

[8] Zolfi, Hamid and Ghorbani, Hamidreza and Ahmadzadegan, M Hossein, “Investigation and classification of cyber-crimes through IDS and SVM algorithm” 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC,IEEE, 2019.

[9] Shang, Wenli and Cui, Junrong and Song, Chunhe and Zhao, Jian- ming and Zeng, Peng, “Research on industrial control anomaly detec- tion based on FCM and SVM” 2018 17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (Trust- Com/BigDataSE),IEEE, 2018.

[10] Betarte, Gustavo and Pardo, Alvaro and Martnez, Rodrigo, “Web appli- cation attacks detection using machine learning techniques” 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA),IEEE, 2018.

[11] Kumar, Saket and Singh, Chandra Bhim Bhan, “A zero-day resistant mal- ware detection method for securing cloud using SVM and sandboxing techniques” 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT),IEEE, 2018.

[12] Ghorpade, Ashish, “Malware Classification using Km-SVM” Dublin, Na- tional College of Ireland, National College of Ireland, 2020.

[13] Lad, Sumit S and Adamuthe, Amol C and others, “Malware classification with improved convolutional neural network model” Int. J. Comput. Netw. Inf. Secur,2020.

[14] Taher, Kazi Abu and Jisan, Billal Mohammed Yasin and Rahman, Md Mahbubur, “Network intrusion detection using supervised machine learning technique with feature selection” 2019 International conference on robotics, electrical and signal processing techniques (ICREST),IEEE, 2019.

[15] Saleous, Heba and Trabelsi, Zouheir, “Enhancing firewall filter performance using neural networks” 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC),IEEE, 2019.

[16] Ninyesiga, Allan and Ngubiri, John, “International Journal of Technology and Management” International Journal of Technology and Management, 2018.

[17] Tran, Trung Kien and Sato, Hiroshi, “NLP-based approaches for malware classification from API sequences” 2017 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES),IEEE, 2017.

[18] Dolezel, Petr and Holik, Filip and Merta, Jan and Stursa, Dominik, “Opti- mization of a depiction procedure for an artificial intelligence-based network protection system using a genetic algorithm” Applied Sciences,MDPI, 2021.

[19] Khan, Muhammad Ashfaq and Kim, Juntae, “Toward developing efficient Conv-AE-based intrusion detection system using heterogeneous dataset” Electronics,MDPI, 2020.

[20] Jiechieu, Kameni Florentin Flambeau and Tsopze, Norbert, “Skills predic- tion based on multi-label resume classification using CNN with model predictions explanation” Neural Computing and Applications,Springer, 2021.

[21] Kalash, Mahmoud and Rochan, Mrigank and Mohammed, Noman and Bruce, Neil DB and Wang, Yang and Iqbal, Farkhund, “Malware classifica- tion with deep convolutional neural networks” 2018 9th IFIP international conference on new technologies, mobility and security (NTMS),IEEE, 2018.

[22] Krithivasan, Kannan and Pravinraj, S and VS,

Shankar Sriram and oth- ers, “Detection of cyberattacks in industrial control systems using enhanced principal component analysis and hypergraph-based convolution neural network (EPCA-HG-CNN)” IEEE Transactions on Industry Appli- cations,IEEE, 2020.

[23] Ahsan, Mostofa and Nygard, Kendall E, “Convolutional Neural Networks with LSTM for Intrusion Detection.” CATA, 2020.

[24] Akarsh, S and Simran, K and Poornachandran, Prabaharan and Menon, Vijay Krishna and Soman, KP, “Deep learning framework and visualization for malware classification” 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS),IEEE, 2019.

[25] Ito, Michiaki and Iyatomi, Hitoshi, “Web application firewall using character-level convolutional neural network” 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA),IEEE, 2018.

[26] Xu, Wenfeng and Fan, Yongxian and Li, Changyong, “I2DS: interpretable intrusion detection system using autoencoder and additive tree” Security and Communication Networks,Hindawi, 2021.

[27] Akarsh, S and Sriram, S and Poornachandran, Prabaharan and Menon, Vi- jay Krishna and Soman, KP, “Deep learning framework for domain generation algorithms prediction using long short-term memory” 2019 5th Inter- national Conference on Advanced Computing & Communication Systems (ICACCS),IEEE, 2019.

[28] Naeem, Hamad and Guo, Bing and Naeem, Muhammad Rashid and Vasan, Danish, “Visual malware classification using local and global malicious pat- tern” Journal of Computers, 2019.

[29] Udayakumar, N and Saglani, Vatsal J and Cupta, Aayush V and Sub- bulakshmi, T, “Malware classification using machine learning algorithms” 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI),IEEE, 2018.

[30] Kumari, V Valli and Varma, P Ravi Kiran, “A semi-supervised intrusion detection system using active learning SVM and fuzzy c-means clustering” 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC),IEEE, 2017.

[31] Luo, Jhu-Sin and Lo, Dan Chia-Tien, “Binary malware image classification using machine learning with local binary pattern” 2017 IEEE International Conference on Big Data (Big Data),IEEE, 2017.

[32] Moradi Vartouni, Ali and Teshnehlab, Mohammad and Sedighian Kashi, Saeed, “Leveraging deep neural networks for anomaly-based web application firewall” IET Information Security,Wiley Online Library, 2019.

[33] Schueller, Quentin and Basu, Kashinath and Younas, Muhammad and Pa- tel, Mohit and Ball, Frank, “A hierarchical intrusion detection system using support vector machine for SDN network in cloud data center” 2018 28th International Telecommunication Networks and Applications Conference (ITNAC),IEEE, 2018.

[34] Ertam, Fatih and Kaya, Mustafa, “Classification of firewall log files with multiclass support vector machine” 2018 6th International symposium on digital forensic and security (ISDFS),IEEE, 2018.

[35] Makandar, Aziz and Patrot, Anita, “Malware class recognition using im- age processing techniques” 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI),IEEE, 2017.

[36] Ghouti, Lahouari and Imam, Muhammad, “Malware classification using compact image features and multiclass support vector machines” IET Information Security,Wiley Online Library, 2020.

[37] Mulay, Snehal A and Devale, PR and Garje, GV, “Intrusion detection system using support vector machine and decision tree” International journal of computer applications,Citeseer, 2010.

[38] Davuluru, Venkata Salini Priyamvada and Narayanan, Barath Narayanan and Balster, Eric J, “Convolutional neural networks as classification tools and feature extractors for distinguishing malware programs” 2019 IEEE National Aerospace and Electronics Conference (NAECON),IEEE, 2019.

[39] Ahmad, Iftikhar and Basheri, Mohammad and Iqbal, Muhammad Javed and Rahim, Aneel, “Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection” IEEE access,IEEE, 2018.

[40] Al-Qatf, Majjed and Lasheng, Yu and Al-Habib, Mohammed and Al- Sabahi, Kamal, “Deep learning approach combining sparse autoencoder with SVM for network intrusion detection” Ieee Access,IEEE, 2018.

[41] Aminordin, AZMI and MA, FAIZAL and Yusof, ROBIAH, “Android malware classification base on application category using static code analysis” J. Theor. Appl. Inf. Technol, 2018.

[42] Meng, Xi and Shan, Zhen and Liu, Fudong and Zhao, Bingling and Han, Jin and Wang, Hongyan and Wang, Jing, “MCSMGS: malware classification model based on deep learning” 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC),IEEE, 2017.

[43] Chen, Hao and Xiao, Ruizhi and Jin, Shuyuan, “Real-time detection of cloud tenant malicious behavior based on CNN” 2020 IEEE Intl Conf on

Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing& Networking(ISPA/BDCloud/SocialCom/SustainCom), IEEE,2020.

[44] Hande, Yogita and Muddana, Akkalakshmi, "Intrusion detection system us- ing deep learning for software defined networks (SDN)" 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT),IEEE, 2019.

[45] Xue, Di and Li, Jingmei and Lv, Tu and Wu, Weifei and Wang, Jiaxiang, "Malware classification using probability scoring and machine learning" IEEE Access,IEEE, 2019.

[46] Khan, Muhammad Ashfaq and Kim, Juntae, "Toward developing efficient Conv-AE-based intrusion detection system using heterogeneous dataset" Electronics,MDPI, 2020.

[47] Blanco, Roberto and Malago´n, Pedro and Cilla, Juan J and Moya, Jos´e M, "Multiclass network attack classifier using CNN tuned with genetic algorithms" 2018 28th International Symposium on Power and Timing Model- ing, Optimization and Simulation (PATMOS),IEEE, 2018.

[48] Narayanan, Barath Narayanan and Davuluru, Venkata Salini Priyamvada, "Ensemble malware classification system using deep neural networks" Elec- tronics,MDPI, 2020.

[49] Thang, Nguyen Manh, "Improving efficiency of web application firewall to detect code injection attacks with random forest method and analysis attributes HTTP request" Programming and Computer Software,Springer, 2020.

[50] Morales-Molina, Carlos Domenick and Santamaria-Guerrero, Diego and Sanchez-Perez, Gabriel and Perez-Meana, Hector and Hernandez-Suarez, Aldo, "Methodology for malware classification using a random forest classifier" 2018 IEEE International Autumn Meeting on Power, Elec- tronics and Computing (ROPEC),IEEvartouni2018anomalyE,Vartouni, Ali Moradi and Kashi, Saeed Sedighian and Teshnehlab, Mohammadvar- touni2018anomaly Vartouni, Ali Moradi and Kashi, Saeed Sedighian and Teshnehlab, Mohammad, "An anomaly detection method to detect web attacks using stacked auto-encoder" 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS),IEEE, 2018.

[51] Raje, Saurabh and Vaderia, Shyamal and Wilson, Neil and Panigrahi, Rudrakh, "Decentralised firewall for malware detection" 2017 Interna- tional Conference on Advances in Computing, Communication and Control (ICAC3),IEEE, 2017.

[52] Abu Al-Haija, Qasem and Odeh, Ammar and Qattous, Hazem, "PDF Mal- ware Detection Based on Optimizable Decision Trees" Electronics,MDPI, 2022.