

Network intrusion detection system: machine learning approach

Ameera S. Jaradat, Malek M. Barhoush, Rawan Bani Easa

Department of Computer Science, Faculty of Information Technology, Yarmouk University, Irbid, Jordan

Article Info

Article history:

Received May 1, 2021

Revised Dec 1, 2021

Accepted Dec 9, 2021

Keywords:

Classifier

Cyber security

Data analysis

Intrusion detection

KNIME

Machine learning

ABSTRACT

The main goal of intrusion detection system (IDS) is to monitor the network performance and to investigate any signs of any abnormalities over the network. Recently, intrusion detection systems employ machine learning techniques, due to the fact that machine learning techniques proved to have the ability of learning and adapting in addition to allowing a prompt response. This work proposes a model for intrusion detection and classification using machine learning techniques. The model first acquires the data set and transforms it in the proper format, then performs feature selection to pick out a subset of attributes that worth being considered. After that, the refined data set was processed by the Konstanz information miner (KNIME). To gain better performance and a decent comparative analysis, three different classifiers were applied. The anticipated classifiers have been executed and assessed utilizing the KNIME analytics platform using (CICIDS2017) datasets. The experimental results showed an accuracy rate ranging between (98.6) as the highest obtained while the average was (90.59%), which was satisfying compared to other approaches. The gained statistics of this research inspires the researchers of this field to use machine learning in cyber security and data analysis and build intrusion detection systems with higher accuracy.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ameera S. Jaradat

Department of Computer Science, Faculty of Information Technology, Yarmouk University

St. Shafiq Irshidat, Irbid, Jordan

Email: ameera@yu.edu.jo

1. INTRODUCTION

Over the last few years, networks have played big roles in modern style of life. Cybersecurity has therefore become a rich area of research that created a new anticipation in data innovation [1]. information security in networks is becoming a primary need of modern society to protect private information flowing over the networks [2]. An intrusion detection system (IDS) is concerned with controlling the events appearing in the different organizations and networks, and thus correlating them for marking of potential hazards, which is crucial for insuring the security over the network [3]. IDS eavesdrops the activities in a given environment and identify these activities as either "malicious ("intrusive") or "legitimate ("normal") given features gained from the network traffic data [4].

Due to the lack of trustworthy tests and validation knowledge, existing IDSs still have problems in improving accuracy of detection, detection of unknown attacks, reduction of false alarm rates. On the other hand, anomaly-based intrusion detection methods are still suffering constant and truthful performance development. A number of studies have focused on developing IDSs that benefit by different techniques known as classical machine learning (ML) and deep learning methods [5], [6]. Interest in intrusion detection was pioneered by J. P Anderson in 1980 where he offered the first technique reflecting intrusion detection system [7]. The world witnessed a massive evolution of the network related technologies since then. The

enormous growth in the network size, and the huge spread of the data generated and shared through the network generates security threats, which encourages the researches of this field to investigate methods for eliminating attacks and maintaining the security of the network.

Many of the evolved researches suffer high rates of false alarm and generate many alerts for low non-threatening acts. In this case, the increasing number of false alerts forced the security analysts to only detect the seriously harmful attacks on the network. As the network environments update fast, this may allow different kinds of attacks to arise continuously, thus creating new challenges in network security with existing IDSs that could not detect some unfamiliar attacks [8]. For this, many investigators in the field have concentrated their attention on developing IDSs that would ensure higher accuracy of detection rates of true attacks and reduce false alarm rates [3], [9].

Many tools have been designed to stop internet-based attacks such as firewall, intrusion prevention system and IDS. ML is one of the artificial intelligences (AI) branches that acquire knowledge from training data depending on established facts. ML is recognized as a technique that makes computers gain knowledge automatically after being programmed [9]. ML is basically categorized into three broad categories: "supervised learning, unsupervised learning, and reinforcement learning" [10]. In "supervised learning (classification)", the instances (features) are classified in the training phase. There are varied "supervised learning" algorithms, including: "artificial neural network (ANN), Bayesian statistics, Gaussian process regression, lazy learning, nearest neighbor algorithm, support vector machine (SVM), hidden Markov model (HMM), and Bayesian networks" [10]. In "unsupervised learning", the data samples are unclassified. A notable way for the learning method using the unsupervised learning depends on the clustering technique. Some of the familiar unsupervised learners are cluster analysis "K-means clustering, fuzzy clustering, hierarchical clustering, self-organizing map, apriority algorithm, Eclat algorithm, and outlier detection" (Local outlier factor). In "reinforcement" learning, the computer interacts with a setting to realize a confirmed objective. Mohamed *et al.* [11] presented a classification model that was applied on network security layer-knowledge discovery in database (NSL-KDD) dataset. The model first implements correlation feature selection to select the dataset features, then applied three different machine learning algorithms to differentiate normal traffic from anomalous ones. The classifiers are random forest (RF), multi-layer perceptron (MLP), and library for support vector machine (LIBSVM).

Hamdi *et al.* [12] presented a dynamic model for the purpose of intrusion detection that chains machine learning approaches and preceding statistics to generate a trusted cloud environment. Kumar *et al.* [13] utilize deep learning approach through employing deep neural network (DNN) to dynamically detect cyberattacks. Sharafaldin *et al.* [14] examined the IDS dataset generation and evaluation through the analysis some available datasets since 1998. Then they utilized the study results to generate a new IDS dataset, which includes set of attacks that mimic real world criteria. Panwar *et al.* [15] utilized Waikato environment for knowledge analysis (WEKA) machine learning tool to apply and evaluate variety of classifiers including decision tree (DT) and J48 to experiment on CICIDS-2017 dataset. Yulianto *et al.* [16] focused on testing the intrusion detection system performance on CICIDS-2017 dataset through utilizing synthetic minority oversampling technique (SMOTE).

This work is proposing a machine learning approach for developing intrusion detection system. ML methods can automatically detect the abnormal data and show the essential differences in normal data with high accuracy rates. Also, ML methods have strong generalizability, in detecting unknown attacks [6]. The presented model uses supervised classification algorithm build a classifier that can screen the data and recognize intrusions. In the prospective approach, our goal is to build a model using selected supervised classification algorithm and test the attacks in terms of both number and accuracy. To accomplish this purpose, the study adopted "CICIDS2017" dataset, and utilized the Konstanz information miner (KNIME) analytics platform to build classifiers that can filter the data and detect intrusions.

The major contribution of this research is developing a scheme for intrusion detection by utilizing machine learning approaches. The model employs KNIME analytics platform to apply the classifiers on the refined Data set. For the aim of providing comparative analysis the system involves applying three different classifiers. The rest of the paper is organized as follows: section 2 describes the machine learning approach for intrusion detection. Section 3 illustrates the experimental results, which is followed by the conclusion.

2. RESEARCH METHOD

This section describes the process of building the intrusion detection (IDS) model, in order to gain a better insight into the capability of the classical machine learning approaches for intrusion detection. The steps of the intrusion detection model involve the acquirement of the raw dataset (CICIDS2017), which is followed by performing data preprocessing to reduce the complexity of the data by removing some of the non-descriptive, messed values. Then, the feature extraction step extracts selected representative set of

attributes from original dataset. Here, MATLAB environment is utilized to try a variety of algorithms for this purpose. The next step is building the classifiers using supervised machine learning to categorize unseen patterns in suitable classes using the KNIME analytics platform [17], [18]. Having the classifier model build, the accuracy results are generated and tested. Consequently, measure the ability of the classical machine learning approach of detecting intrusions with the appropriate features.

2.1. Dataset

The experiments were conducted on the CICIDS2017 dataset, which comprises categorized network inflows, including full packet payloads. The CICIDS2017 dataset included eight varied files containing the traffic data for attacks and five normal days of the Canadian Institute of Cybersecurity. Table 1 provides short description of dataset files for the CICIDS2017 dataset [19], [20].

Table 1. Network network flow analyzer for the CICIDS2017 dataset [13]

File	Day	Attack
File_1	Mon.	Benign (Normal human activities)
File_2	Tues.	Benign, FTP, SSH
File_3	Wed.	Benign, DoS (GoldenEye, Hulk, Slowhttptest, Sslwloris), Heartbleed
File_4	Thurs.	Benign, Web Attack (Brute Force, SQL Injection, XSS)
File_5	Thurs.	Benign, Infiltration
File_6	Fri.	Benign, Bot
File_7	Fri.	Benign, PortScan
File_8	Fri.	Benign, DDoS

2.2. Dataset preprocessing

Natural data are usually noisy, contain missing or redundant values, and incomplete. Data preprocessing phase is crucial in machine learning. This phase is necessary for assembling and fixing the raw data for the machine learning model. It comprises the implementation of techniques that aimed at reducing the complexity of the dataset by removing some of the non-descriptive, messed values, and non-necessary attributes from the original dataset [21].

2.3. Feature selection

The feature selection phase is geared to select all representatives and appropriate set of attributes from the set of raw attributes (raw dataset). The representative dataset keeps only relevant and critical attributes and cross any non-necessary attributes [22]–[25]. In this study, the approach used number of techniques and several algorithms for selecting relevant features from the raw dataset, thus gaining more facilitates for data visualization and data understanding. Table 2 provides a portrayal of the used techniques in the feature selection phase.

Table 1. The extraction techniques in dataset preprocessing

Technique	Description
Correlation-based feature selection method (CFS)	CFS aims to have new subsets of features highly correlated with a specific class (classes), and uncorrelated to each other (attributes).
Principal component analysis (PCA)	PCA aims to identify all uncorrelated features.
Information gain ratio-based feature selection (IGR)	IGR is used for splitting the attributes pattern distribution into classes, where a gain ratio of attribute decreases as the value of split information increases.
Minimum redundancy maximum relevance	The technique punishes a feature’s relevance based on its redundancy.

2.4. Classification

Recently, there has been a tendency to use deep learning techniques for most types of AI problems [25]–[31]. However, there are some benefits of using classical machine learning. In addition of being computationally inexpensive, machine learning algorithms perform better for smaller data sets.

The machine learning approach uses three classifiers through the KNIME analytics platform. Each classifier works separately for testing the accuracy of the generated files from the previous phase. Figure 1 shows the architecture of the intrusion detection model in KNIME analytics platform. Table 3 provides the description of the classifier model of supervised machine learning using the KNIME analytics platform. SVM classifier is a linear classification technique, which splits data through a hyperplane [32], [33]. DT classifier constantly splits the data conferring a certain parameter. Each node in the tree splits the data while the leaves represent the choices or the outcome decisions [34]. Resilient propagation (RPROP) is considered a robust machine learning scheme that adopts the local gradient information for the weight step [35], [36].

Table 2. The supervised machine learning model

Algorithm	KNIME Node	Description
SVM	SVM Learner SVM Predictor	This node trains a SVM on the input data and uses an SVM model generated by the SVM learner node to predict the output for given values.
RProp	RProp MLP Learner Multi-Layer Perceptron Predictor	Implementation of the RProp algorithm for multilayer feed forward networks. RPROP performs a local adaptation of the weight-updates according to the behavior of the error function.
Decision Tree	Decision Tree Learner Decision Tree Predictor	This node induces a classification decision tree in the main memory. The target attribute must be nominal (classes of attack).

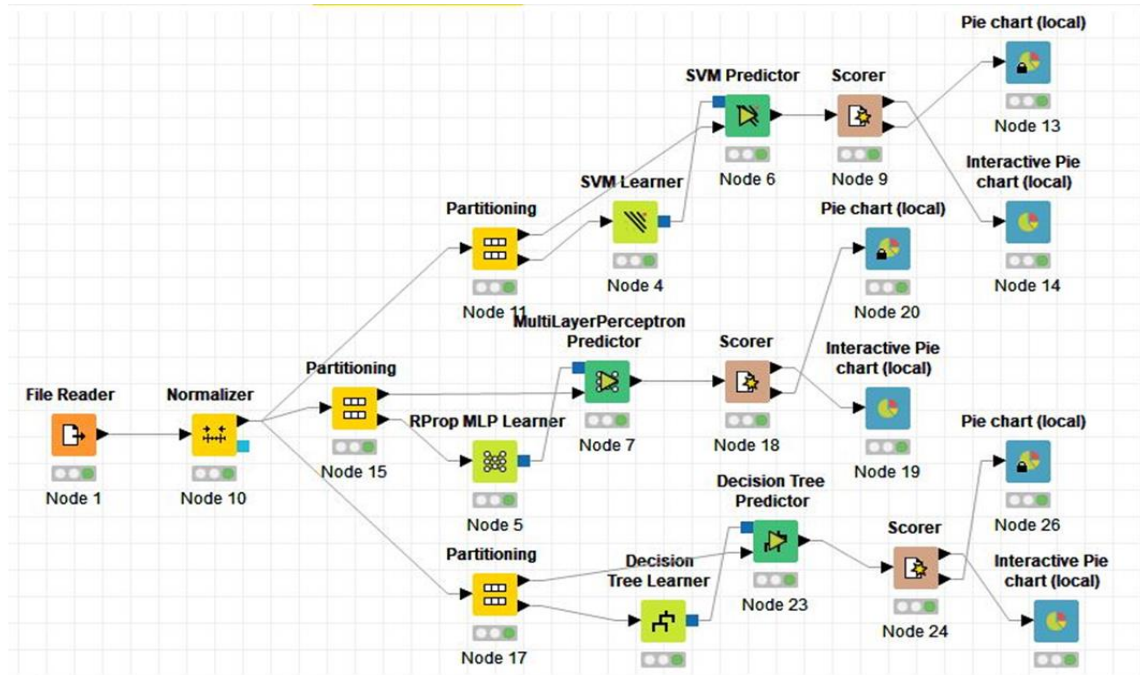


Figure 1. The classifier of KNIME model

3. RESULTS AND DISCUSSION

This section presents the experimental outcomes for the formulated methodology. Here, the first step after data preprocessing is carrying out feature extraction on the original dataset. Then, enter the refined dataset to KNIME analytics platform where classifiers (SVM, RProp, and decision tree) are build. Finally Generate the accuracy of each classifier. The accuracy results are then compared to other results from the related work. The research was conducted using MATLAB and KNIME. MATLAB tool was employed for feature extraction. The KNIME analytics platform was deployed for the purpose of building the classifier model and testing the results.

3.1. Data analysis and interpretation

The model utilized four common information retrieval evaluation metrics based on the confusion matrix, according to [31], [32]:

- The precision value (Pr) or known as positive predictive value (PPV), is the ratio of correctly classified attacks flows (TP), in front of all the classified flows (TP+CF).

$$Pr = \frac{TP}{TP+CF} \tag{1}$$

- Recall (Rc), is the ratio of correctly classified attack flows (TP), in front of all generated flows for the all experiments (TP+FN).

$$Rc = \frac{TP}{TP+Fn} \tag{2}$$

- F-measure (F1), is a hybrid combination of the Pr and Rc into a one measure.

$$F1 = \frac{2}{\frac{1}{Pr} + \frac{1}{Rc}} \tag{3}$$

- The accuracy or percentage of correct classification (PCC), can be calculated using (4):

$$PCC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP is the number of cases which is well classified as normal, TN is the number of cases which is well classified as Intrusion, FP is the number of cases that is classified as intrusion, but they were normal, and FN is the number of cases that is classified as normal, but they were attacks.

3.2. Experimental results

The CICIDS2017 dataset is adopted for the research. During the research process, the experiment starts by entering the raw dataset into MATLAB environment to perform data preprocessing and feature selection. Then, the generated file that contains the refined data was processed using KNIME platform. Figure 2 summarizes the building of the model for intrusion detection.

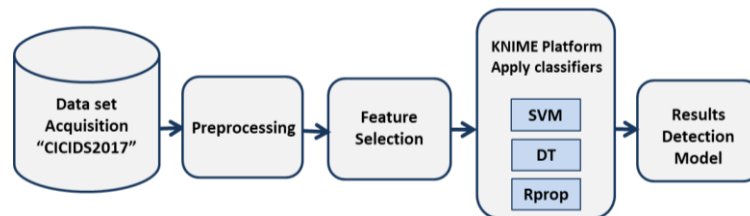


Figure 2. The main structure of the suggested IDS model

KNIME analytics platform contains several nodes that represent the stubs for data processing and building the classifier. The file reader node reads preprocessed data. The normalizer node normalizes the data by selecting the columns (attributes) to use in the experiment. The Partitioning node splits the dataset into two partitions (train and test data). This study used 60% of data for each Excel file as training data, while 40% as testing. The subsequent step is building the classifier. The experiment builds three classifiers (SVM, RProp, and decision tree). The scorer node generates a confusion matrix reflecting the number of attributes with their classification matches. It also, outputs accuracy statistics including precision, recall, F-measure, and accuracy. Table 4 lists the accuracy results. The results are also illustrated in Figure 3.

Table 3. The supervised machine learning model

Classifier	Precision	Recall	F-measure	Accuracy
SVM	97.35%	99.10%	84.48%	90.81%
RProp	95.35%	97.94%	90.16%	86.23%
DT	98.38%	99.33%	86.93%	94.72%

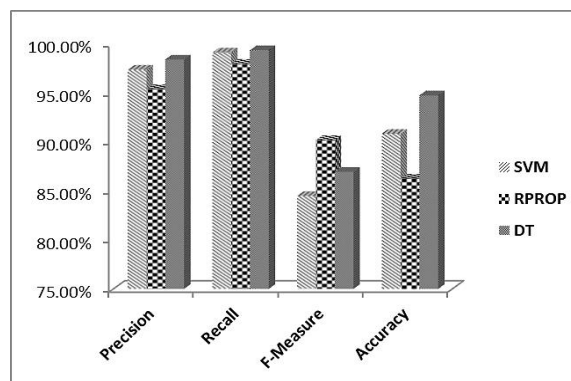


Figure 3. The accuracy result of the constructed approach

To evaluate the goodness of the suggested approach we compared the final results with other studies [11]–[16]. Table 5 lists the accuracy results along with other accuracy measures that are resulted of applying different methods from related works and the results of the anticipated method. The accuracy results seem satisfying when compared to other approaches in the same field. Figure 4 illustrates the average accuracy results of the other studies compared to the results of the proposed approach. The suggested approach gains the advantage of allowing control of traffic on the network and regulating the level of security. On the other hand, using classifiers has double the layer of security thereby making the network more protected.

Table 4. Comparing the results of the proposed approach against some related studies

Study	Algorithm	Precision	Recall	F-Measure	Accuracy
[11]	MLP	95.70%	97.10%	96.30%	93.80%
[11]	RF	99.60%	99.60%	99.60%	99.60%
[11]	LIBSVM	94.80%	95%	94.50%	97.20%
[12]	RF	96.48%	99.90%	94.03%	99.40%
[12]	BN	95.92	97.12%	94.03%	94.80%
[12]	RT	98.97%	99.80%	94.03%	99.70%
[12]	J48	98.88%	99.75%	95.04%	99.80%
[13]	KNN	78.10%	96.80%	86.50%	91.00%
[13]	DT	83.90%	96.50%	89.80%	94.00%
[13]	RF	84.90%	96.90%	90.50%	94.00%
[13]	SVM-rbf	99.30%	32.80%	49.30%	97.00%
[14]	Naive Bayes	99.24%	99.68%	94.03%	98.96%
[14]	J48	99.24%	99.68%	94.03%	98.96%
[14]	Decision Tree	99.24%	99.68%	94.03%	96.04%
[15]	AdaBoost	77%	88%	77%	77%
[16]	EFS	85.15%	94.92%	89.77%	81.47%
[16]	EFS + SMOTE	81.83%	100%	90.01%	81.83%
[16]	AdaBoost + PCA Feature	81.49%	99.93%	89.78%	81.47%
[16]	AdaBoost + PCA Feature + SMOTE	81.69%	95.76%	88.17%	81.47%
Proposed Approach	SVM	97.35%	99.10%	84.48%	90.81%
Proposed Approach	RPROP	95.35%	97.94%	90.16%	86.23%
Proposed Approach	DT	98.38%	99.33%	86.93%	94.72%

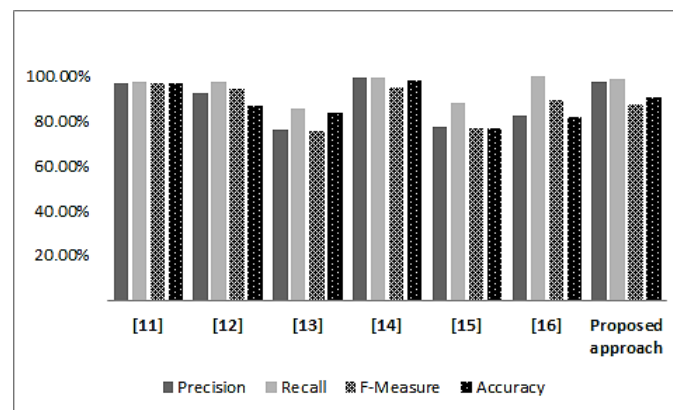


Figure 4. The average accuracy results of the proposed approach compared to related works result

4. CONCLUSION

Network security has been an obligatory subject in the system of distributed activities. Nonetheless, preserving the system security has becoming more demanding. Malicious attacks are constantly varying, which makes security sustaining a challenging practice. The main impact of this research is emerging a machine learning approach for intrusion detection. The model used the "CICIDS2017" dataset and employed Knime platform to apply the classifiers on the refined Data set. Three different classifiers were applied (SVM, RProp, and decision tree) for sorting out the filtered dataset. The research process includes data acquisition followed by preprocessing step to reduce the complexity of the data. Then, the feature selection step chooses relevant features from the raw dataset. After performing feature selection, the data was processed using KNIME analytics platform where the classifiers were applied to filter the data and detect intrusions. The main motivation for conducting this study is to build a robust system that can efficiently

detect malicious attacks and prevent unauthorized intrusions. Theoretically, the results may be helpful for research orientation; researchers and students in this field need to be in touch with the new research ideas, considering its strengths and weaknesses. Practically, the results of this research could be helpful to programmers and software designers how they can use ML in cyber security and data analysis. Researchers and experts in the field of computer sciences are therefore invited to build intrusion detection systems with higher accuracy and precision.




REFERENCES

- [1] O. A. Raheem and E. Alomari, "An adaptive intrusion detection system by using decision tree," *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 10, no. 2, 2018, doi: 10.29304/jqcm.2018.10.2.387.
- [2] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, "Network intrusion detection," *IEEE Network*, vol. 8, no. 3, pp. 26–41, May 1994, doi: 10.1109/65.283931.
- [3] S. K. Biswas, "Intrusion detection using machine learning: a comparison study," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 19, pp. 110–114, Sep. 2018.
- [4] K. Jayakumar, T. Revathi, and S. Karpagam, "Intrusion detection using artificial neural networks with best set of features," *International Arab Journal of Information Technology*, vol. 12, no. 6A, pp. 728–734, 2015.
- [5] M. Almseidin, M. Alzubi, S. Kovacs, and M. Alkasasbeh, "Evaluation of machine learning algorithms for intrusion detection system," in *SISY 2017 - IEEE 15th International Symposium on Intelligent Systems and Informatics, Proceedings*, Sep. 2017, pp. 277–282, doi: 10.1109/SISY.2017.8080566.
- [6] H. Kaur, G. Singh, and J. Minhas, "A review of machine learning based anomaly detection techniques," *International Journal of Computer Applications Technology and Research*, vol. 2, no. 2, pp. 185–187, Jul. 2013, doi: 10.7753/ijcatr0202.1020.
- [7] J. P. Anderson, "Computer security threat monitoring and surveillance," *Technical Report James P Anderson Co Fort Washington Pa*, p. 56, 1980, doi: citeulike-article-id:592588.
- [8] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," *Applied Sciences (Switzerland)*, vol. 9, no. 20, p. 4396, Oct. 2019, doi: 10.3390/app9204396.
- [9] N. Farah, M. Avishek, F. Muhammad, A. Rahman, M. Rafni, and D. Md., "Application of machine learning approaches in intrusion detection system: a survey," *International Journal of Advanced Research in Artificial Intelligence*, vol. 4, no. 3, 2015, doi: 10.14569/ijarai.2015.040302.
- [10] A. KumarShrivias and A. Kumar Dewangan, "An ensemble model for classification of attacks with feature selection based on KDD99 and NSL-KDD data set," *International Journal of Computer Applications*, vol. 99, no. 15, pp. 8–13, Aug. 2014, doi: 10.5120/17447-5392.
- [11] H. Mohamed, H. Hefny, and A. Alsawy, "Network intrusion detection system: A machine learning approach," *Egyptian Computer Science Journal*, vol. 42, no. 3, pp. 44–56, Nov. 2018.
- [12] Z. Chkribene, A. Erbad, R. Hamila, A. Mohamed, M. Guizani, and M. Hamdi, "TIDCS: A dynamic intrusion detection and classification system based feature selection," *IEEE Access*, vol. 8, pp. 95864–95877, 2020, doi: 10.1109/ACCESS.2020.2994931.
- [13] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
- [14] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *ICISSP 2018 - Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 2018, vol. 2018-January, pp. 108–116, doi: 10.5220/0006639801080116.
- [15] S. Singh Panwar, Y. P. Raiwani, and L. S. Panwar, "Evaluation of Network Intrusion Detection with Features Selection and Machine Learning Algorithms on CICIDS-2017 Dataset," *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3394103.
- [16] A. Yulianto, P. Sukarno, and N. A. Suwastika, "Improving AdaBoost-based Intrusion Detection System (IDS) Performance on CIC IDS 2017 Dataset," *Journal of Physics: Conference Series*, vol. 1192, no. 1, p. 012018, Mar. 2019, doi: 10.1088/1742-6596/1192/1/012018.
- [17] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Systems with Applications*, vol. 36, no. 3 PART 1, pp. 5432–5435, Apr. 2009, doi: 10.1016/j.eswa.2008.06.054.
- [18] H. Chae, B. Jo, S. Choi, and T. Park, "Feature Selection for Intrusion Detection using NSL-KDD," *Recent Advances in Computer Science 20132*, vol. 20132, pp. 184–187, 2013.
- [19] R. Panigrahi and S. Borah, "A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems," in *International Journal of Engineering & Technology*, vol. 3, 2018, pp. 479–482.
- [20] A. Boukhamla and J. C. Gavio, "CICIDS2017 dataset: performance improvements and validation as a robust intrusion detection system testbed," *International Journal of Information and Computer Security*, vol. 16, no. 1/2, p. 20, 2021, doi: 10.1504/IJICS.2021.117392.
- [21] Y. Hamid, M. Sugumaran, and L. Journaux, "Machine Learning Techniques for Intrusion Detection," in *Proceedings of the International Conference on Informatics and Analytics*, Aug. 2016, pp. 1–6, doi: 10.1145/2980258.2980378.
- [22] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLOS ONE*, vol. 10, no. 3, p. e0118432, Mar. 2015, doi: 10.1371/journal.pone.0118432.
- [23] S. Juma, Z. Muda, M. A. Mohammed, and W. Yassin, "Machine Learning Techniques for Intrusion Detection: A Review," *Journal of Theoretical & Applied Information Technology*, vol. 72, no. 3, pp. 422–429, 2015.
- [24] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998, doi: 10.1023/A:1009715923555.
- [25] Yuanhang Su and C.-C. Jay Kuo, "Fast and robust camera's auto exposure control using convex or concave model," in *2015 IEEE International Conference on Consumer Electronics (ICCE)*, Jan. 2015, pp. 13–14, doi: 10.1109/ICCE.2015.7066300.
- [26] Y. Su, J. Y. Lin, and C.-C. J. Kuo, "A model-based approach to camera's auto exposure control," *Journal of Visual Communication and Image Representation*, vol. 36, pp. 122–129, Apr. 2016, doi: 10.1016/j.jvcir.2016.01.011.
- [27] Y. Su, Y. Huang, and C.-C. J. Kuo, "Efficient Text Classification Using Tree-structured Multi-linear Principal Component Analysis," in *2018 24th International Conference on Pattern Recognition (ICPR)*, Aug. 2018, pp. 585–590, doi: 10.1109/ICPR.2018.8545832.




- [28] Y. Su, R. Lin, and C.-C. Jay Kuo, "Tree-structured multi-stage principal component analysis (TMPCA): Theory and applications," *Expert Systems with Applications*, vol. 118, pp. 355–364, Mar. 2019, doi: 10.1016/j.eswa.2018.10.020.
- [29] Y. Su, K. Fan, N. Bach, C.-C. J. Kuo, and F. Huang, "Unsupervised Multi-modal Neural Machine Translation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nov. 2018, pp. 10482–10491, [Online]. Available: <http://arxiv.org/abs/1811.11365>.
- [30] Y. Su and C.-C. J. Kuo, "On extended long short-term memory and dependent bidirectional recurrent neural network," *Neurocomputing*, vol. 356, pp. 151–161, Sep. 2019, doi: 10.1016/j.neucom.2019.04.044.
- [31] F. Wang and C. Rudin, "Falling rule lists," *Journal of Machine Learning Research*, vol. 38, pp. 1013–1022, Nov. 2015, [Online]. Available: <http://arxiv.org/abs/1411.5899>.
- [32] R. AlAbdallah, A. Jaradat, I. Doush, and Y. Jaradat, "A Binary Classifier Based On Firefly Algorithm," *Jordanian Journal of Computers and Information Technology*, vol. 3, no. 3, p. 172, 2017, doi: 10.5455/jjcit.71-1501152301.
- [33] C. Igel and M. Hüsken, "Improving the Rprop learning algorithm," in *Proceedings of the second international ICSC symposium on neural computation (NC 2000)*, 2000, pp. 151–121.
- [34] H. F. Kareem, M. S. AL-Huseiny, F. Y. Mohsen, E. A. Khalil, and Z. S. Hassan, "Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, p. 1731, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1731-1738.
- [35] N. R. Shenoy and A. Jatti, "Ultrasound image segmentation through deep learning based improvised U-Net," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, p. 1424, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1424-1434.
- [36] K. M. O. Nahar, A. Jaradat, M. S. Atoum, and F. Ibrahim, "Sentiment analysis and classification of arab jordanian facebook comments for jordanian telecom companies using lexicon-based approach and machine learning," *Jordanian Journal of Computers and Information Technology*, vol. 6, no. 3, pp. 247–262, 2020, doi: 10.5455/jjcit.71-1586289399.

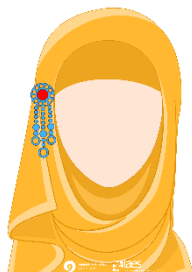
BIOGRAPHIES OF AUTHORS






Ameerah S. Jaradat    is an Associate professor at the Department of Computer Science, Yarmouk University-Jordan. She received her PhD in computer science at the University of Arkansas–USA in 2008. She received her MSc. in computer science at the University of Arkansas–USA in 2003. Her research interests are related to graph algorithms, structure and function of networks, specifically social and information networks. Other current research interests are in the fields of machine learning and data mining. She can be contacted at email: ameera@yu.edu.jo.



Malek M. Barhoush    is an assistant professor at computer science department at Yarmouk University (YU) since 2012. He received his Ph.D. in computer network security from Concordia University, Montreal-Canada in 2012. He was granted scholarship for MSc. and Ph.D. from YU in 2002 and 2005. During his work at YU, he was the dean assistant at Information Technology and computer science faculty for one year. He also worked as the chairman of the network security department during the years 2016-2018. He has several international journal and conference research publications in a number of research areas. His research interest focuses on Cloud Computing, Parallel and Distributed Systems, Computer & Network Security, Cyber security, Wireless Sensor Network, Mobile Computing, Image Processing and Natural Language Processing. He can be contacted at email: malek@yu.edu.jo.



Rawan Bani Easa    is a Computer Science graduate student at Yarmouk University. She received her master in computer science from Yarmouk University-Jordan, 2021. My research interest is specifically related to network security, Machine Learning, web application and web design. She can be contacted at email: rawan87.b.e@gmail.com.