ISSN 2278-3091

Volume 9, No.1.1, 2020

International Journal of Advanced Trends in Computer Science and Engineering

Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse0891.12020.pdf

https://doi.org/10.30534/ijatcse/2020/0891.12020

Machine Learning for Cyber Threat Detection

Pournima More[1], Mr. Pragnyaban Mishra[2]

[1]Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India, pournima.more1@gmail.com

[2]Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India, pragnyaban@kluniversity.in

ABSTRACT

In recent years, anomaly detection has more importance in networking domains. Machine learning is very effective in anomaly detection to improve accuracy in classification. To ensure automated and effective cyber threat detection, analysis of security logs from the dataset is required. Usage of the internet increases cyber-attacks and at present, the cyber security situation is pessimistic. Use of social media and networking has increased in daily life, nowadays all are learning and working by using the internet but on the other hand, it becomes serious security threats problem. Thus the development of the Intrusion Detection System (IDS) is essential to provide an extra level of security. Cyber threat is an important issue faced by all organizations. However, it has difficult to use machine learning algorithms for threat detection analysis, due to huge number of negative threats detection, especially in the case of large scale environments. In this paper, we surveyed clustering and classification of machine learning algorithms. Using machine learning algorithms cyber threats and false detection rates are reduced which increases the performance of the system.

Key words : Multi-Layer Perceptron (MLP), Intrusion Detection Systems (IDS), Principal Component Analysis (PCA), Fast Clustering based Feature Selection (FAST), Support Vector Machine (SVM)

1. INTRODUCTION

To detecting threats, researcher's has developed systems. The attacks are over the networks and therefore, consider as a malicious activity. Hence, IDS is used to secure the network from such activity and attacks. The working of IDS is to gather the information from network and used for threat detection analysis. There are many techniques to detect intrusions efficiently [12].

IDS are developed to identify unauthorized access or manipulate computer systems. IDS collect network data to identify different kinds of malicious services. IDS have two major classifications, signature-based detection and anomaly-based detection. In signature-based IDS, attack pattern of intruders are matched with existing database pattern then the system will notify once the match is identified. To identify the new attack pattern Signature databases have to be updated frequently. However, anomaly based detection system stores the behavior of the normal users over the network and find the any action different from normal behavior [7].

Cyber-security attacks are increasing in computer network. Therefore, robust and reliable security framework requires in society. An IDS is an important to prevent the threats in computer network system [1,21]. Unsupervised IDSs can differentiate between normal behavior and malicious behavior[17]. Machine Learning (ML) techniques has played important role in IDS. ML-IDS is analyzing and classifying the information [12, 21]. Data mining technologies such as naïve Bayes networks, fuzzy logic, neural networks, support vector machine and genetic algorithms are used for pattern recognition and classification as they have enhanced the performance of the models that use by such algorithms.

IDS can be used for big data over the network, which contains redundant and irrelevant attributes. Irrelevant attributes can make it difficult to identify malicious activity, causing slow testing process and low detection rate. Attribute selection is important in IDS as it gives high classification performance using training data. Some of the attributes may be duplicate and irregular [14]. Removing these attributes is important as they may give the low performance of classifiers. Attributes selection has feature to find the subset of attributes for accuracy [6, 8].

In this paper, a survey has been made based on IDS machine learning algorithms. To achieve this, different machine learning algorithms are considered to find out the best suitable among them. Machine learning algorithms can apply to calculate the accuracy of the detection rate.

The remaining section has organized as follows: In Section 2, we describe the literature survey. Finally, in Section 3, we summarize the present study and draw some conclusions.

2. LITERATURE SURVEY

In IDS and network security, Machine Learning classification algorithms has used to reduce the false detection rate in detection systems and differentiate between normal and abnormal behavior of network traffic.

2.1 Data Mining Techniques

Data mining algorithms have been recently used for development of intrusion detection models to reduce data overloading. These models extract useful knowledge by searching for patterns and relationships from the data collected thereby improving decision making. In classification, the features of the new object are observed and are assigned to one of the existing set of classes. Classifier models gain knowledge from the training data and detect the class for the new instances. Many supervised learning algorithms are used to solve classification problems [28].

SVM is one of the efficient techniques when the sample training data is small [22]. In recent years, many hybrids intelligent systems have been proposed to improve accuracy.

Data Mining has widely used in intrusion detection. It is used for removing unwanted information over the network and keeps the important information in the network. Data mining is used in computer network for security. The below algorithms of data mining techniques are used in the IDS.

2.1.1 Machine Learning Algorithms

Machine learning is important in data mining algorithms as they are pointing the issues related to information.

Machine learning techniques are:

•Multi-Layer Perceptron (MLP): MLP is an algorithm that works on the concept of feed-forward a neural network with multiple hidden layers. MLP is common in the fields of pattern prediction, classification, approximation and recognition as it is addressing the issues that are not linearly separable.

•Support Vector Machine (SVM): SVM reduces the error without concerning the classification error and it can be used the duality theory of mathematical programming for efficient computational methods [26, 30,35].

•Bayes Net: Bayes Net gives the information about each cycle and relation between to each other cycle probabilistically. Bayes Net acquires Bayesian networks under predictions like nominal attributes and correct values. A Bayes Netis depends on nodes which saves the large number of computation. It is not required to save all the configurations of cycles.

•Random Forest: Random Forest algorithm can categorize each new object based on the separate choice analysis done by each tree.

2.2 Anomaly Detection

ANOMALY detection has different patterns to searching information for expected behaviors. Number of techniques has been implemented to solve this problem, in that one-class support vector machine (OCSVM) is the much useful method.

2.2.1 Support Vector Machine (Svm)

SVM is the most useful technique of Machine learning [3, 13]. An SVM plays effective role as a classification technique in a problem, such as signal processing, often providing a huge enhancement over challenging methods [4, 18, 29]. In cyber-security, the SVM can decrease false detection rate and increase the accuracy of IDSs. The classifier is useful to analysis the network traffic and differentiate it two possible outcomes malicious and non-malicious. Machine learning algorithms helps and contributes to detect a classification technique in anomaly-based IDS. It can measure the execution of IDS against different SVM methods one class and two class SVMs. Even though a two-class SVM gives much accurate performance, if one class SVM has robust, then it could reduce the necessity of offline dataset labeling process. A one-class SVM needs normal traffic of trained dataset for classification. Attribute subset selection is an actual procedure of separate not pertinent data, increasing quality, and refining result comprehensibility [11,27, 32].

It can be achieved regression, outlier detection, and classification using Support Vector Machines (SVMs)

Advantages of SVM are:

1. Efficient results in big dimensional spaces
2. Less usage of memory as SVM uses set of training data in the decision-making function
3. Various kernel functions can be used for the support vectors. [8,34]
4. It gives good performance when small trained data with more attributes.

Algorithm for SVM-

1. Importing Dataset
2. Dividing data into training data & testing data
3. Feature Scaling
4. Fitting SVM to the training data set
5. Predicting the test data
6. Creating the Confusion Matrix
7. Display the training data results
8. Display the test data results

2.3 Clustering Algorithms

Analyzing the internet traffic is very essential for the initial detection of cyber-attacks and network security. In large corporations, network traffic is high due to high level of internet usage, statistical tools are unable to identify malicious activities. Arithmetical Clustering can be used as an effective way to separate normal and malicious information. It could analyze and compare clustering algorithms like DBSCAN, K-Means and BIRCH for network traffic survey because of the arithmetical nature of the data-set, and due to K Means is giving good performance compare than other Hierarchical methods (Agglomeration, Divisive etc.) [33]. To detect the IDS K-means algorithm is good than other algorithms.

2.3.1 K-Means

K-means algorithm specifies the data point to a class. The data point is divided into clusters; the cluster center is the shortest distance from the data point. The clustering centers are optimized and a new clustering center is calculated for the data points allocated in the cluster [19,23]. The algorithm ends until the cluster center does not change. The main advantages of the algorithm are that the complexity of the algorithm is low, fast calculation, consumption of small resources and high efficiency. K-means algorithm is used for large data sets and better clustering. The K-means algorithm mainly includes the following aspects: K-means algorithm is used to select different initial clustering center which leads to the final clustering results. The clustering results depend on the selection of initial cluster centers which may also randomly select the initial clustering center, resulting in the K-means algorithm find a suboptimal solution[20,31,33].

Algorithm for K-Means-

1. Randomly choose n cluster centers

2. Assign each point to the group that has the closest centroid.

3. When all data points have been allotted, recalculate the locations of the n centroids.

4. Repeat Steps 2 and 3 until the centroids no longer change [10].

2.3.2density-Based Spatial Clustering (Dbscan)

DBSCAN generates cluster regions with higher spatial density of data points, whereas low density regions are marked as clusters of noise. DBSCAN requires two essential input variables, a radius of nearest cluster and minimum points of cluster neighborhood.

DBSCAN makes more significant clusters in case of operational and behavioral studies being unaffected to noise which can handle clusters of different categories[33]. However, it’s dependent on its defined value of parameters. Generally, DBSCAN is not giving correct results in security, due to network heavy network traffic.

2.3.3 Balanced Iterative Reducing & Clustering Using Hierarchical (BIRCH)

BIRCH is capable of a multilevel hierarchical based clustering algorithm that produces hierarchical clustering dendrogram [33], known as a clustering feature (CF) tree. Each cluster in the CF tree is entitled three parameters (N: number of data points, LS: linear sum, SS: square sum). In case of larger datasets BIRCH clustering is not giving high performance as its working very slow since it has number of clustering cycles depending upon the data resolution and splitting level. Hence, it’s role in cyber security is only limited to small scale clustering on a subset of data, in combining with a large-scale algorithm (DBSCAN, K-Means etc.) clustering data at a higher level.

2.4 Feature Extraction

2.4.1 Principal Component Analysis (Pca)

PCA used statistical techniques in the field of data mining to reduce the dimensionality and to identify data points with the highest possible variance [8,25]. The PCA is used to differentiate network traffic data into normal and anomalous sub-regions. In this method, the focus is on the detection of anomalies in origin-destination flow aggregated in backbone networks and it is a necessary component within the IDS system. The PCA method identifies anomalous traffic on a particular link by comparing it with past values. Thus, PCA separates link traffic measurements into sub-regions representing normal and abnormal traffic. PCA is an effective technique in data compression and feature extraction. In Feature selection process information space is converted into a feature space, which has a lower dimension [11,15,16].

PCA minimizes the number of dimensions needed to classify new data and creates a set of principal components. This minimizes computational costs and the inaccuracy of parameter evaluation.

Algorithm for PCA-

1. Determine the covariance matrix of the normalized n-dimensional dataset.

2. Determine the eigenvectors and eigen values of the covariance matrix.

3. Sort the eigen values in descending order.

4. Select the k eigenvectors that correspond to the k largest eigen values where k is the number of dimensions of the new feature subspace.

5. Create the projection matrix from the k selected eigenvectors.

6. Convert the original dataset to build a new k-dimensional feature space [24].

2.4.2Fast Clustering Based Feature Selection Algorithm (FAST)

Attributes set selection is a way of detecting and eliminating irregular and duplicate attributes. Irrelevant attributes and duplicate attributes [9] are not giving correct enhanced predictor. Which solves using already present information of other attributes. FAST algorithm can remove the different attributes while maintained the duplicate attributes. In a network to identify anomalies, a FAST algorithm is a necessary component within the IDS system.

The FAST algorithm is working based on minimum spanning tree (MST) method to cluster attributes. It is not assuming the data cluster centers by a regular geometric curve. The FAST algorithm does not have any limit for any types of information [9].

Different attributes and duplicate attributes cause the accuracy of machine learning [9]. Attribute set selection can identify and eliminate irrelevant and redundant information. Good feature subsets contain features that are highly correlated with predictive class, yet uncorrelated with not predictive class [2, 9].

The FAST algorithm effectively and efficiently deals with the different and duplicate attributes. To obtain features through a new feature selection this repressed of two connected segments of different and duplicate attributes removal. Selecting relevant features from the object by removing different ones and then removes duplicate attributes from relevant ones by selecting from different attribute clusters and creates the concluding set [9].

The FAST algorithm involves three steps: first, the construction of the MST from a weighted complete graph; second, the partitioning of the MST into a forest with each tree is representing a cluster; third, the collection of representative attributes from the clusters [9].

```
Algorithm for FAST-
1. Collect a training data set from a particular domain.
2. Shuffle the data set.
3. Break it into m partitions
4. For each partition (k = 0, 1, ..., m-1 )
    a. Let Outer Train set(k) = all partitions except k.
    b. Let Outer Test set(k) = the k’th partition
  c. Let Inner Train(k) = randomly chosen 70% of the Outer Train set(k).
  d. Let Inner Test(k) = the remaining 30% of the Outer Train set(k).
    e. For l = 0, 1, ..., n
  Search for the best feature set with l components,. using leave-one-out on Inner Train(k) Let Inner Test Score(kl) = RMS score of on   Inner Test(k).
  End loop of (l).
  f. Select the best inner test score.
  g. Let Outer Score(k) = RMS score of the selected feature set on Outer Test set(k)
  End of loop of (k).
5. Return the mean Outer Score.
```

2.4.3 Intrusion Detection System (IDS)

An intrusion detection system observes the activity in a system and chooses whether these activities are malicious or not. Network-based IDS analysis for all activities in network traffic and set up an alarm whenever abnormal activity is observed.

Algorithm for IDS-

1. Gathering of data coming from network traffic and monitoring application log.

2. Feature extraction & Data selection- Select the required data for security analysis.

3. Differentiate between known and unknown packets- Partition the set into subsets using the different attribute

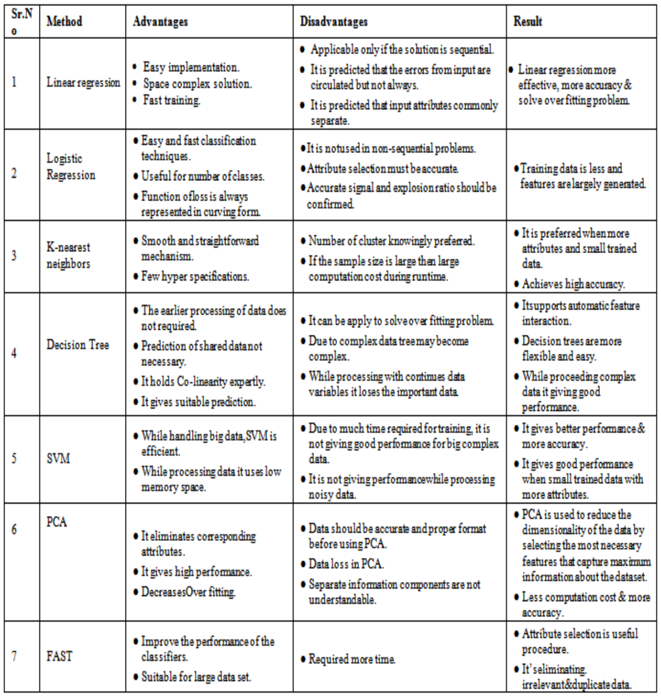
4. Data comparison- Matching selected traffic against stored rules

5. Build a decision tree node containing that attribute

6. Analysis of reports for anomaly detection

The table below shows analysis of different algorithm based on various features i.e. reduced noise and computational cost, enhanced accuracy and performance.

Table 1: Analysis of algorithms



Sr.No	Method	Advantages	Disadvantages	Result
1	Linear regression	• Easy implementation. • Space complex solution. • Fast training.	• Applicable only if the solution is sequential. • It is predicted that the errors from input are circulated but not always. • It is predicted that input attributes commonly separate.	• Linear regression more effective, more accuracy & solve over fitting problem.
2	Logistic Regression	• Easy and fast classification techniques. • Useful for number of classes. • Function of loss is always represented in curving form.	• It is not used in non-sequential problems. • Attribute selection must be accurate. • Accurate signal and explosion ratio should be confirmed.	• Training data is less and features are largely generated.
3	K-nearest neighbors	• Smooth and straightforward mechanism. • Few hyper specifications.	• Number of cluster knowingly preferred. • If the sample size is large then large computation cost during runtime.	• It is preferred when more attributes and small trained data. • Achieves high accuracy.
4	Decision Tree	• The earlier processing of data does not required. • Prediction of shared data not necessary. • It holds Co-linearity expertly. • It gives suitable prediction.	• It can be apply to solve over fitting problem. • Due to complex data tree may become complex. • While processing with continues data variables it loses the important data.	• Its supports automatic feature interaction. • Decision trees are more flexible and easy. • While proceeding complex data it giving good performance.
5	SVM	• While handling big data,SVM is efficient. • While processing data it uses low memory space.	• Due to much time required for training, it is not giving good performance for big complex data. • It is not giving performancewhile processing noisy data.	• It gives better performance & more accuracy. • It gives good performance when small trained data with more attributes.
6	PCA	• It eliminates corresponding attributes. • It gives high performance. • DecreasesOver fitting.	• Data should be accurate and proper format before using PCA. • Data loss in PCA. • Separate information components are not understandable.	• PCA is used to reduce the dimensionality of the data by selecting the most necessary features that capture maximum information about the dataset. • Less computation cost & more accuracy.
7	FAST	• Improve the performance of the classifiers. • Suitable for large data set.	• Required more time.	• Attribute selection is useful procedure. • It’s eliminating. irrelevant&duplicate data.

3. CONCLUSION

We have studied various clustering algorithms, classifier algorithms and features extraction algorithms. K-means[19,20,31], SVM [3, 22, 32, 34, 35] and FAST [2,5,9]algorithms has given better results. The survey shows that the system can speed the training process and tests the intrusions detection which is required in network security related applications.

REFERENCES

[1] K. G. Kyriakopoulos, F. J. Aparicio-Navarro, and D. J. Parish, Manual and automatic assigned thresholds in multi-layer data fusion intrusion detection system for 802.11 attacks, IET Information Security, Volume 8, Issue 1, 2014 https://doi.org/10.1049/iet-ifs.2012.0302

[2] Mr Avinash Godase, Mrs Poonam Gupta,A Survey on Clustering Based Feature Selection Technique Algorithm for High Dimensional Data, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 4, Issue 1, 2015

[3] Noreen Kausar, Brahim Belhaouari Samir, Azween Abdullah, Iftikhar Ahmad, Mohammad Hussain, Review of Classification Approaches Using Support Vector Machine in Intrusion Detection, Informatics Engineering and Information Science: International Conference, ICIEIS 2014

[4] T. S. Hai, and N. T. Thuy, Image classification using support vector machine and artificial neural network, International Journal of Information Technology and Computer Science (IJITCS), Volume 4, Issue 5, 2012

[5] Pawan Gupta, Susheel Jain, Anurag Jain A Review Of Fast Clustering-Based Feature Subset Selection Algorithm,

International Journal Of Scientific & Technology Research, Volume 3, Issue 11, 2014

[6] Molina L.C, Belanche L. and Nebot A, Feature selection algorithms: A survey and experimental evaluation, IEEE International Conference Data Mining, 2002

[7] Xiufeng Liu, Per Sieverts Nielsen, Regression-based Online Anomaly Detection for Smart Grid Data, International Journal of Energy Volume 16, Issue 6, 2016

[8] Sumaiya Thaseen Ikram and Aswani Kumar Cherukuri, Improving Accuracy of Intrusion Detection Model Using PCA and Optimized SVM, Journal of Computing and Information Technology, Volume 24, Issue 2, 2016

[9] Qinbao Song, Jingjie Ni, Guangtao Wang, A Fast Clustering-Based Feature Subset Selection Algorithm for High Dimensional Data ,IEEE Transactions On Knowledge And Data Engineering Volume 25, Issue 1,2013 https://doi.org/10.1109/TKDE.2011.181

[10] Trupti Kodinariya, Dr Prashant Makwana, Review on determining number of Cluster in K-Means Clustering, International Journal of Advance Research in Computer Science and Management Studies , Volume 1, Issue 6, 2013

[11] Goverdhan Reddy, SammulalPorika, Kernel Centric Machine Learning Classifiers for Anomaly Detection with Real Bank Datasets, IEEE International Conference on Innovations in Information Embedded and Communication Systems, 2015

[12] Inadyuti Dutt, Samarjeet Borah, Indrakanta Maitra, A Proposed Machine Learning based Scheme for Intrusion Detection, Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology, 2018

[13] Xuedan Miao, Ying Liu , Haiquan Zhao, Chunguang Li, Distributed online one class support vector machine for anomaly detection over networks, IEEE Transactions On Cybernetics, 2018

[14] Jun Yang , Chunjie Zhou , Shuanghua Yang, HaizhouXu ,Bowen Hu, Anomaly Detection Based on Zone Partition for Security Protection of Industrial Cyber-Physical Systems, IEEE Transactions On Industrial Electronics, Volume 65, Issue 5, 2018 https://doi.org/10.1109/TIE.2017.2772190

[15] Kun Xie , Xiaocan Li, Xin Wang, Jiannong Cao, GaogangXie, Jigang Wen, Dafang Zhang, and Zheng Qin, On-Line Anomaly Detection With High Accuracy, IEEE/ACM Transactions On Networking, 2018

[16] T. Morita, S. Yogo, M. Koike, T. Hamaguchi, S. Jung, I. Koshijima, Y. Hashimoto, Detection of Cyber-Attacks with Zone Dividing and PCA, Elsevier 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems, 2013

[17] Sujit Rokka Chhetri, SinaFaezi, Mohammad Abdullah Al Faruque, Information Leakage-Aware Computer Aided Cyber-Physical Manufacturing, IEEE Transactions On Information Forensics And Security, 2018

[18] Jianxin Wu, Hao Yang Linear Regression-Based Efficient SVM Learning for Large-Scale Classification, IEEE Transactions On Neural Networks And Learning Systems, Volume 26, Issue 10, 2015

[19] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, Angela Y. Wu, An Efficient k-Means Clustering Algorithm: Analysis and Implementation, IEEE Transactions On Pattern Analysis And Machine Intelligence, Volume 24, Issue 7, 2002 https://doi.org/10.1109/TPAMI.2002.1017616

[20] M. EmreCelebi a, Hassan A. Kingravi b, Patricio A. Vela, A comparative study of efficient initialization methods for the k-means clustering algorithm, Volume 40, Issue 1, Elsevier Expert Systems with Applications 2013

[21] Vasileios Mavroeidis, Siri Bromander, Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence, European Intelligence and Security Informatics Conference, 2017

[22] Peyman Asgharzadeh ,Shahram Jamali, A Survey On Intrusion Detection System Based Support Vector Machine Algorithm, International Journal Of Research In Computer Applications And Robotics, Volume 3, Issue 12, 2015

[23] Dasarisreelalitha , c visishta, Classification of attack types for intrusion detection system using machine learning algorithm: Random forest, International Journal of Advance Research, Ideas and Innovations in Technology Volume 5, Issue 2 , 2019

[24]Sumaiya Thaseen Ikram1, Aswani Kumar Cherukuri, Improving Accuracy of Intrusion Detection Model Using PCA and Optimized SVM, Journal of Computing and Information Technology, Volume 24, Issue 2, 2016 https://doi.org/10.20532/cit.2016.1002701

[25] Heba F. Eid, Ashraf Darwish, Aboul Ella Hassanien, Ajith Abraham, Principle Components Analysis and Support Vector Machine based Intrusion Detection System, IEEE 10th International Conference on Intelligent Systems Design and Applications, 2010

[26] Indu Kumar Kiran Dogra Chetna Utreja Premlata Yadav, A comparative study of supervised machine learning algorithms for stock market trend prediction, IEEE International Conference on Inventive Communication and Computational Technologies, 2018

[27] Xingzhi Zhang, Zhurong Zhou, Credit Scoring Model based on Kernel Density Estimation and Support Vector Machine for Group Feature Selection , IEEE International Conference on Advances in Computing, Communications and Informatics, 2018

[28] Vinayaka Nagendra, Harikishan Gude, Steven Corns, Suzanna Long, Evaluation of Support Vector Machines and Random Forest Classifiers in a Real-time Fetal Monitoring System Based on Cardiotocography Data, IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, 2017

[29] Ayumi Sada, Yuma Kinoshita, Sayaka Shiota, Hitoshi Kiya, Histogram-Based Image Pre-processing for Machine Learning, IEEE 7th Global Conference on Consumer Electronics, 2018

[30] ReemAlyami, Jinan Alhajjaj, Batool Alnajrani, Ilham Elaalami, Abdullah Alqahtani, Nahier Aldhafferi, Taoreed O. Owolabi b, Sunday O. Olatunji, Investigating the effect of Correlation based Feature Selection on breast cancer diagnosis using Artificial Neural Network and Support Vector Machines, IEEE International Conference on Informatics, Health & Technology, 2017
https://doi.org/10.1109/ICIHT.2017.7899011

[31] Gerhard M¨unz, Sa Li, Georg Carle, Traffic Anomaly Detection Using K-Means Clustering, International GI/ITG Conference, 2007

[32] P. R. Visali Lakshmi ; G. Shwetha ; N. Sri Madhava Raja, Preliminary Big Data Analytics of Hepatitis Disease by Random Forest and SVM Using R-Tool, Third International Conference on Biosignals, Images and Instrumentation, 2017

[33] K. Chitra , Dr. D. Maheswari, A Comparative Study of Various Clustering Algorithms in Data Mining, International Journal of Computer Science and Mobile Computing, Volume 6, Issue 8, 2017

[34] Reem Alyami, Jinan Alhajjaj, Batool Alnajrani, Ilham Elaalami, Abdullah Alqahtani, Nahier Aldhafferi, Taoreed O. Owolabi ,Sunday O. Olatunji, Investigating the effect of Correlation based Feature Selection on breast cancer diagnosis using Artificial Neural Network and Support Vector Machines, International Conference on Informatics, Health & Technology, 2017

[35] Chuan-Yu Chang, Man-Ju Cheng, Matthew Huei-Ming Ma, Application of Machine Learning for Facial Stroke Detection, IEEE 23rd International Conference on Digital Signal Processing, 2018
https://doi.org/10.1109/ICDSP.2018.8631568