

Malware detection using Machine learning methods

PROJECT REPORT

Submitted by

Anurag Chawla

20BIT0220

Arun V. S

20BIT0233

Divyanshu Kumar

20BIT0418

Atishay Jain

20BIT0015

Mukul Chavan

20BIT0238

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY



SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING

VIT UNIVERSITY: VELLORE 632014

AUGUST 2022.

INTRODUCTION

In the digital age, malware has impacted many computing devices. The term malware comes from malicious software which is designed to meet the harmful intent of the attacker. Malware can compromise computers/smart devices, steal confidential information, penetrate networks, Cripple critical infrastructures, etc. These programs include viruses, worms, trojans, spyware, bots, rootkits, ransomware, etc. According to Computer Economics, financial loss due to malware attacks has quadrupled from \$3.3 billion in 1997 to \$13.3 billion in 2006. Every few years the definition of the Year of Mega Breaches must be recalibrated based on attacks performed in that particular year. Recently in 2016, the WannaCry ransomware attack² crippled the computers of more than 150 countries, doing financial damage to different organizations. In 2016, Cybersecurity Ventures³ estimated the total damage due to malware attacks was \$3 trillion in 2015 and is expected to reach \$6 trillion by 2021. Antivirus software (such as Norton, McAfee, Avast, Kaspersky, AVG, Bitdefender, etc.) is a major line of defence against malware attacks. Traditional software used the signature-based method for malware detection. A signature is a short sequence of bytes which can be used to identify known malware. But the signature-based detection system cannot provide security against zero-day attacks. Also, malware generation toolkits like Zeus can generate thousands of variants of the same malware by using different obfuscation techniques. Signature generation is often a human driven process which is bound to become infeasible with the current malware growth.

ABSTRACT

Our project named “Malware Detection Using Machine Learning Algorithms” is aimed at detecting the malware infected program executable files using machine learning algorithms. Malware analysis and prevention methods are increasingly becoming necessary for computers systems connected to the Internet. This software exploits the system’s vulnerabilities to steal valuable information without the user’s knowledge, and stealthily sends it to remote servers controlled by attackers. This malware needs to be detected so that the user gets information that his system is getting corrupted, and he can therefore take appropriate measures to prevent further attacks of malware and he can take some steps to recover from the attacks. Machine learning has

been recently introduced into the field of Malware Detection. Many algorithms have been used which results in differing accuracies in predicting whether the input files are malware or not. Many different algorithms like Apache Spark and TuriGraph Lab have been used to predict the malware infected files, but the accuracy is less than 90%. Moreover, algorithms have become non-existent in today's world

PROBLEM STATEMENT:

To perform a comparative study on various ML algorithms used in Malware detection, and identify the best suited algorithm for the given dataset

OBJECTIVES LISTED AND COMPLETED:

- Train a Logistic regression model to detect malware
- Train a Decision tree model to detect malware
- Train a Random Forest model to detect malware
- Observe accuracy, precision, and other parameters for all models

LITERATURE SURVEY:

Title	Authors	Methodology/Algorithms	Metrics
Deep Learning in Cybersecurity: Challenges and Approaches (2020)	Yadigar Imamverdiyev, Fargana Abdullayeva	CNN, RBM, DBM, Hybrids	DBN achieved detection accuracy of 97.5% DL based Hybrid code attack detection method achieved 97.55% accuracy
Machine Learning in Cyber-Security threats (2020)	Alex Mathew	ANN + SVM, using Naïve Bayes and Random Forest algorithms	No metrics used

Determining Viability of Deep Learning on Cybersecurity Log analytics (2018)	Casey Lorenzen, Rajeev Agarwal, Jason King	Batch Normalization with ReLu activation function	http accuracy: 92.42% DNS accuracy: 96.88% conn accuracy: 84.42%
Machine Learning and Deep Learning methods for Cybersecurity (2018)	Yang Xin, Zhiu Liu, Lingshuang Kong, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, Chunhua Wang	SVM, K-nearest neighbour, Decision Tree, Deep Belief network, RNN	VPN traffic precision : 94.9% (6-class) and 92% (12-class) Non-VPN traffic precision: 85.5% (6-class) and 85.8% (12-class)
Analysis of Cybersecurity Threats in Cloud Applications Using Deep Learning Techniques (2019)	S.A. Sokolov, T.B. Iliev, I.S. Stoyanov	Combination of SVM and Neural networks	Accuracy: 85% (Pybull framework) Accuracy: 83% (in Weka)

Title	Author	Methodology	Metrics
Evaluating deep learning approaches to characterize and classify malicious URL's	R. Vinayakumara, K.P. Somana, Prabakaran Poornachandranb	Convolution Neural Network (CNN), Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM), deep learning mechanisms: Recurrent Neural Network (RNN), Identity-Recurrent Neural Network (I-RNN), Long Short-Term Memory (LSTM), Malicious uniform resource locator (URL) or malicious website	accuracy, precision, recall, F1-score
Machine Learning and Deep Learning methods for Cybersecurity	Prof. Nanda M B, Parinitha B S	Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Decision Tree	Accuracy, Performance, Precision

Deep Learning for Cybersecurity: A Review	Zhaolin Chen	Convolutional Neural Network (CNN), Recurrent neural network (RNN), Auto Encoders (AE), Deep Belief Networks (DBN), Supervised and Unsupervised learning, Reinforcement learning (RL)	Performance, Accuracy
DEEP LEARNING CLASSIFICATION METHODS APPLIED TO TABULAR CYBERSECURITY BENCHMARKS	David A. Noever, Samantha E. Miller Noever	MobileNetV2 architecture, TensorFlow, Convolutional Neural Network (CNN), XGBoost,	Accuracy, Spped, Performance,
Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study	Mohamed Amine Ferrag, Leandros Maglaras, Sotiris Moschoyiannis, Helge Janicke	deep neural network, feed forward deep neural network, recurrent neural network, convolutional neural network, restricted Boltzmann machine, deep belief network, deep auto-encoder, deep migration learning, self-	performance, accuracy, false alarm rate, detection rate, efficiency, f1-score, precision, recall

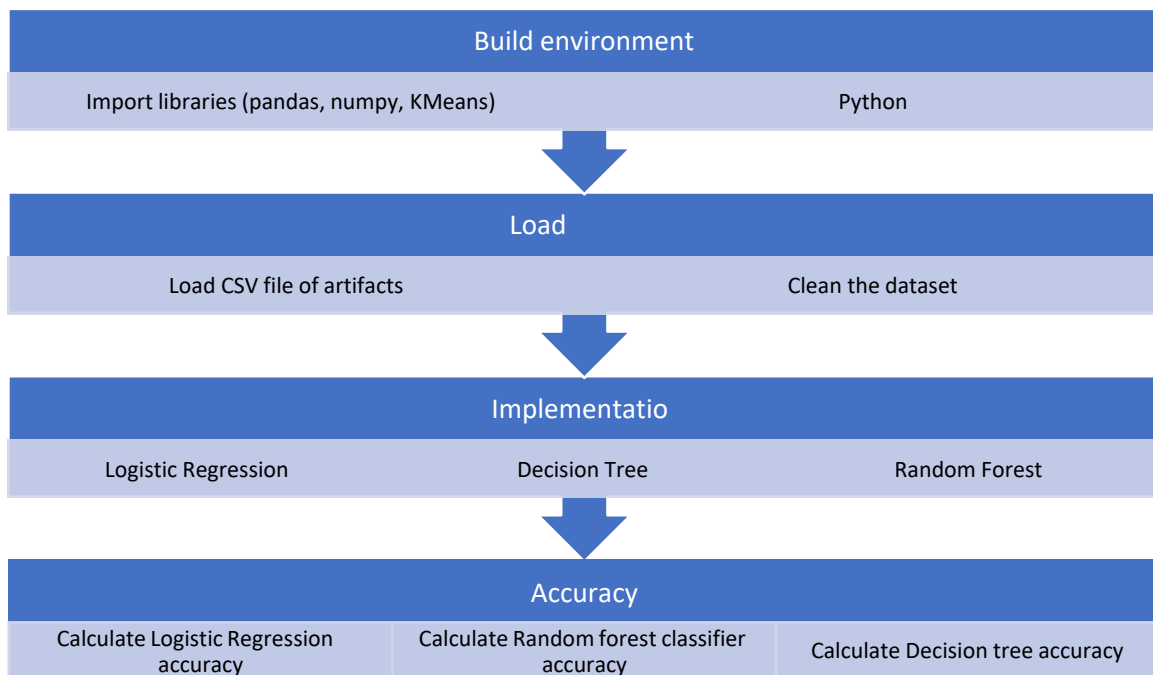
Title	Author	Methodology	Metrics
A Deep Learning Framework for Intelligent Malware Detection	William Hardy, Lingwei Chen, Shifu Hou, Yanfang Ye*, and Xin Li	Artificial Neural Network (ANNs), Support Vector Machine (SVM) , Naïve Bayes (NB) Decision Tree (DT)	PERFORMANCE MEASURES Accuracy F P Rate (FPR) T P Rate (TPR)
Apply Stacked Auto-Encoder to Spam Detection	Guyue Mi, Yang Gao, and Ying Tan	Naive Bayes Support Vector Machine Decision Tree Random Forest Boosting Artificial Neural Network	Performance Precision Recall Accuracy F1
Detecting Drive-by Download Attacks from Proxy Log Information using Convolutional Neural Network	Kohei Yamanishi	CNN EDCNN RandomForest	performance accuracy

Deep Packet: A Novel Approach For Encrypted Traffic Classification Using Deep Learning	Mohammad Lotfollahi · Mahdi Jafari Siavoshani · Ramin Shirali Hossein Zade · Mohammdsadegh Saberian	Convolutional Neural Network Autoencoder	accuracy performance
Real-Time Detection of False Data Injection Attacks in Smart Grid: A Deep Learning-Based Intelligent Mechanism	Youbiao He, Student Member, IEEE, Gihan J. Mendis, Student Member, IEEE, and Jin Wei, Member, IEEE	State Vector Estimator (SVE) Supervisory Control And Data Acquisition Artificial Neural Network	performance accuracy

Title	Authors	Methodology	Metrics
Machine Learning in Cyber-Security Threats	Alex Matthew	Cloud computing and machine learning algorithms	Accuracy, performance
Cybersecurity data science: an overview from machine learning perspective	Iqbal H. Sarker, A. S. M. Kayes, Shahriar Badsha, Hamed Alqahtani, Paul Watters & Alex Ng	machine learning based multi-layered framework	Accuracy, precision, performance
Deep Learning in Cybersecurity: Challenges and Approaches	Yadigar Imamverdiyev, Fargana Abdullayeva	Anomaly detection	Accuracy
Deep Learning for Cyber Security Applications: A Comprehensive Survey	Vinayakumar Ravi, Mamoun Alazab, Soman KP, Sriram Srinivasan, Sitalakshmi Venkatraman, Quoc-Viet Pham ,Simran K	Classical Machine learning algorithms and deep learning	Accuracy, precision, recall/true positive rate (TPR)/sensitivity, F1-measure/F1-score, false positive rate(FPR), true negative rate (TNR), and false negative rate(FNR)
A Hybrid Spectral Clustering and Deep Neural Network Ensemble Algorithm for Intrusion Detection in Sensor Networks	Tao Ma , Fen Wang , Jianjun Cheng , Yang Yu and Xiaoyun Chen	Sparse auto-encoder (SAE) for network intrusion detection	Accuracy, performance, precision

PROPOSED MODEL:

First, we build an environment for training the models, by installing the necessary tools and importing the necessary libraries. Then, we import the data, which is a CSVfile. For all three techniques: Logistic regression, Decision tree and Random Forest, we consider the first column (Address of Entry point) and the fifth column (Dll characteristics) as sample data and the last column (Legitimate = 0 or 1) as target class. We then fit each model into lr (Logistic regression), rfc (Random Forest classifier) and tree_classifier (Decision tree).



Each proposed model will be explained in detail...

Random Forest

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression

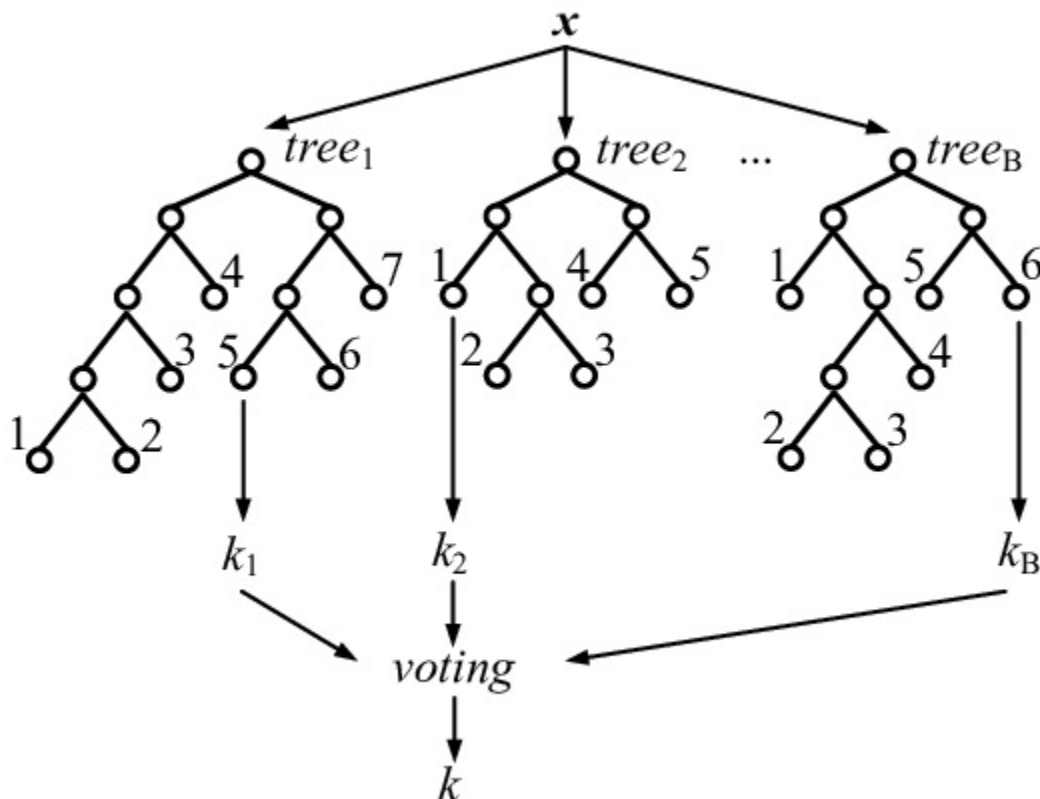
Select random samples from a given dataset.

Construct a decision tree for each sample and get a prediction result from each decision tree.

Perform a vote for each predicted result.

Select the prediction result with the most votes as the final prediction.

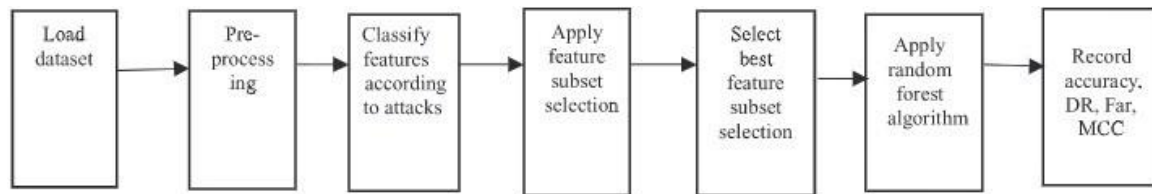
Random Forest grows multiple decision trees which are merged for a more accurate prediction. The logic behind the Random Forest model is that multiple uncorrelated models (the individual decision trees) perform much better as a group than they do alone.



Feature selection methods can be classified into three categories:

- Filter method

- Wrapper method
- Embedded method



Pros:

- The algorithm is robust as it consists of multiple Decision trees
- Missing values won't be a problem in Random Forest as median values can replace null values
- It can be used for both regression and classification problems

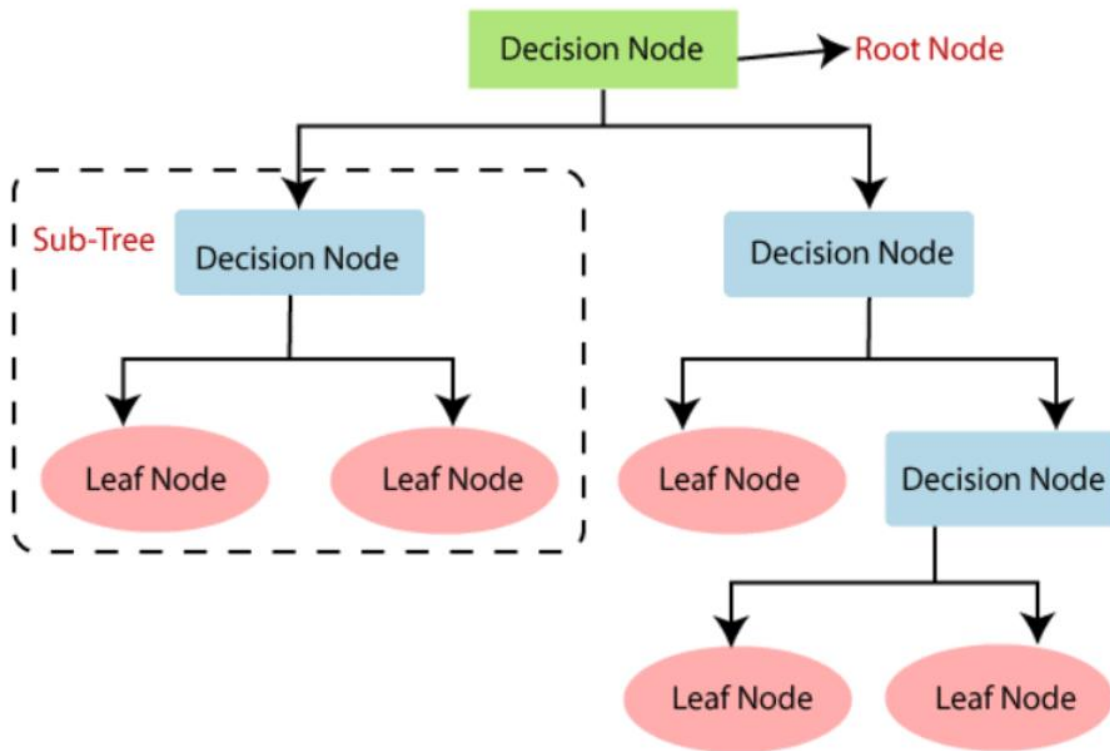
Cons:

- The algorithm is slow as it consists of multiple Decision trees
- It is unsuitable for real-time predictions
- Random forest algorithm computations are very complex

Decision Tree

Decision Tree is a kind of **supervised learning** technique that may be used for each classification and Regression issue, however principally it's most popular for determination Classification issues. it's a tree-structured classifier, wherever internal nodes represent the options of a dataset, branches represent the selection rules and each leaf node represents the result.

There are 2 nodes, which are called the Decision node and Leaf Node. The Decision node is used to make decisions and also the Leaf node is that the output.



Why use Decision Trees?

- Decision Trees typically mimic human thinking ability whereas creating a call, thus it's straightforward to grasp.
- The logic behind the Decision tree is often simply understood as a result of it shows a tree-like structure.

How The Decision Tree Algorithm Work?

- Begin the tree with the foundation node, says S, that contains the entire dataset.
- Notice the simplest attribute within the dataset victimization **Attribute choice measure (ASM)**.
- Divide the S into subsets that contain attainable values for the simplest attributes.
- Generate the choice tree node, that contains the simplest attribute.
- Recursively create new call trees victimization the subsets of the dataset created in step - 3. Continue this method till a stage is reached wherever you can't more classify the nodes and known as the ultimate node as a leaf node.

Pros:

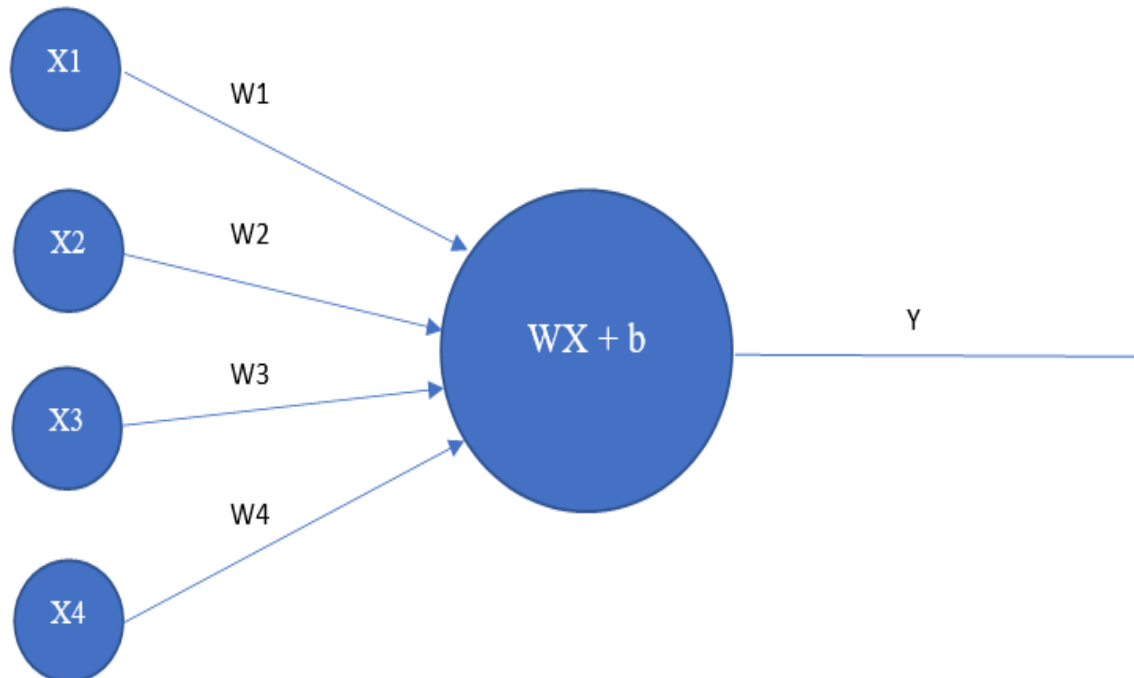
- Simple to understand and visualize
- Requires less effort from users in Data preparation phase
- Can handle both numerical and categorical data

Cons:

- Complicated trees can be created that do not generalize data well
- Small variations in data can lead to huge change in results
- The cost of computation is high

Logistic regression

X_1, X_2, X_3 and X_4 are the input features, which are connected to a single neuron. W_1, W_2, W_3 and W_4 are the weights of each input. First, the input features are multiplied with the weight vectors which results in WX . Then, bias (b) will be added to the product. In some cases, the activation function (σ) changes the output of the neuron, which is a probability between 0 and 1.



Pros

- Easy to implement, and train
- It is less inclined towards overfitting
- It has good scalability

Cons

- It cannot run on a non-linear model
- Data pre-processing is necessary as it can't handle missing values
- It is strictly a classification method and has lot of competition

EXPERIMENT SETUP:

Environment's hosted runtime was connected to Python 3 Google Compute Engine backend, on Google Collab, with a RAM utilization of 1.04 GB and Disk utilization of 39.53 GB

Dataset:

Training Dataset consist of 137445 files information which consist of both malware and original files information. Each column describes about the file properties and the last column legitimate can have only two possible values 0 or 1. If legitimate value equals to 1 then file is original and if legitimate value equals to 0 then file is malware.

original file 40919 (legit dataset)

malware file 96526 (malware dataset)

With the help of our training dataset and model we will try to identify that whether file is malware or not.

AutoSave MalwareArtifacts Search (Alt+Q) ATISHAY JAIN

File Home Insert Page Layout Formulas Data Review View Help

Clipboard Font Alignment Number Styles Cells Editing Analysis Sensitivity

POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma delimited (.csv) format. To preserve these features, save it in an Excel file format. Don't show again Save As...

	A	B	C	D	E	F	G	H	I	J	K	L	M
	AddressOfEntryPoint	MajorLinkerVersion	MajorImageVersion	MajorOperatingSystemVersion	DllCharacteristics	SizeOfStackReserve	NumberOfSections	ResourceSize	legitimate				
2	10407	9	6	6	33088	262144	4	952	1				
3	5354	9	6	6	33088	262144	4	952	1				
4	58807	9	6	6	33088	262144	4	136490	1				
5	25166	9	6	6	33088	262144	4	1940	1				
6	70387	9	6	6	33088	262144	4	83098	1				
7	5856	9	6	6	33088	262144	4	1064	1				
8	26798	9	6	6	33088	262144	4	1060	1				
9	28581	9	6	6	33088	262144	4	99567	1				
10	72841	9	6	6	33088	262144	4	1040	1				
11	6541	9	6	6	33088	262144	4	4264	1				
12	16320	9	6	6	33088	262144	4	24567	1				
13	35457	9	6	6	33088	262144	4	3138	1				
14	82426	9	6	6	33088	262144	4	99567	1				
15	56384	9	6	6	33088	262144	4	945	1				
16	5211	9	6	6	33088	262144	4	1030	1				
17	119081	9	6	6	33088	262144	4	904	1				
18	13410	9	6	6	33088	262144	4	956	1				
19	261025	9	6	6	33088	262144	4	904	1				
20	39564	9	6	6	33088	262144	4	1324	1				
21	107200	9	6	6	33088	262144	4	15992	1				
22	45623	9	6	6	33088	262144	4	916	1				
23	30579	9	6	6	33088	262144	4	52330	1				
24	29115	9	6	6	33088	262144	4	12392	1				
25	50674	9	6	6	33088	262144	3	0	1				
26	5641	9	6	6	32832	262144	4	896	1				

MalwareArtifacts

Ready Accessibility: Unavailable

24°C Rain to stop Search the web ENG IN 03:02 12-11-2022

IMPLEMENTATION:

Random Forest:

```
[1] #Import libraries
import pandas as pd
import numpy as np
from sklearn import *
from sklearn.metrics import accuracy_score

[2] #Read the CSV file
malware_dataset = pd.read_csv("/content/MalwareArtifacts.csv", delimiter=',')

[3] #Extracting artifacts samples fields AddressOfEntryPoint and DLLCharacteristics
samples = malware_dataset.iloc[:, [0, 4]].values
targets = malware_dataset.iloc[:, 8].values

[4] #Split dataset into training and test data
from sklearn.model_selection import train_test_split
training_samples, testing_samples, training_targets, testing_targets = train_test_split(samples, targets, test_size = 0.2, random_state = 0)

[5] #Calculate results and display accuracy
rfc = ensemble.RandomForestClassifier(n_estimators=50)
rfc.fit(training_samples, training_targets)
accuracy = rfc.score(testing_samples, testing_targets)
print("Random forest classifier accuracy : " + str(accuracy))

Random forest classifier accuracy : 0.9636581905489469
```

```

✓ [7] #Calculating TP, FP, TN, FN, Precision, Recall and F1 Score
10 res = rfc.predict(samples)
mt = confusion_matrix(targets, res)
print("True positive rate : %f %%" % ((mt[0][0] / float(sum(mt[0])))*100))
print("False positive rate : %f %%" % ((mt[0][1] / float(sum(mt[0])))*100))
print("True negative rate : %f %%" % ((mt[1][1] / float(sum(mt[1])))*100))
print("False negative rate : %f %%" % ((mt[1][0] / float(sum(mt[1])))*100))
precision = (mt[0][0]/(mt[0][0] + mt[0][1]))
print('Precision : ', precision)
recall = (mt[0][0]/(mt[0][0] + mt[1][0]))
print('Recall : ', recall)
f1score = 2 * (precision * recall) / (precision + recall)
print('F1 score : ', f1score)

```

```

True positive rate : 99.298635 %
False positive rate : 0.701365 %
True negative rate : 98.352803 %
False negative rate : 1.647197 %
Precision : 0.9929863456478047
Recall : 0.9930172083337651
F1 score : 0.9930017767509804

```

Decision Tree:

```

In [1]: #Import Libraries
import pandas as pd
import numpy as np
from sklearn import *
from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix

```

```

In [3]: #Load the CSV file
malware_dataset = pd.read_csv("MalwareArtifacts.csv", delimiter=',')

```

```

In [4]: #Extract artifacts samples files "AddressOfEntryPoint" and "DLLCharacteristics"
samples = malware_dataset.iloc[:, [0, 4]].values
targets = malware_dataset.iloc[:, 8].values

```

```

In [5]: #Split data into training and test data
from sklearn.model_selection import train_test_split
training_samples, testing_samples, training_targets, testing_targets = train_test_split(samples, targets, test_size = 0.2, random

```

```

In [6]: #Calculate results and display accuracy
from sklearn import tree
tree_classifier = tree.DecisionTreeClassifier()
tree_classifier.fit(training_samples, training_targets)
predictions = tree_classifier.predict(testing_samples)
accuracy = 100 * accuracy_score(testing_targets, predictions)
print("Decision Tree accuracy: " + str(accuracy))

```

Decision Tree accuracy: 96.27487358579796

```
In [7]: #Calculating TP, FP, TN, FN, Precision, Recall and F1 Score
res = tree_classifier.predict(samples)
mt = confusion_matrix(targets, res)
print("True positive rate : %f %%" % ((mt[0][0] / float(sum(mt[0])))*100))
print("False positive rate : %f %%" % ((mt[0][1] / float(sum(mt[0])))*100))
print("True negative rate : %f %%" % ((mt[1][1] / float(sum(mt[1])))*100))
print("False negative rate : %f %%" % ((mt[1][0] / float(sum(mt[1])))*100))
precision = (mt[0][0]/(mt[0][0] + mt[0][1]))
print('Precision : ', precision)
recall = (mt[0][0]/(mt[0][0] + mt[1][0]))
print('Recall : ', recall)
f1score = 2 * (precision * recall) / (precision + recall)
print('F1 score : ', f1score)

True positive rate : 99.355614 %
False positive rate : 0.644386 %
True negative rate : 98.228164 %
False negative rate : 1.771836 %
Precision : 0.9935561403145267
Recall : 0.9924970764470293
F1 score : 0.9930263260076105
```

Logistic Regression:

```
✓ [11] #Import libraries
0s import pandas as pd
import numpy as np
from sklearn import *
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix

✓ [14] #Load the CSV file
0s malware_dataset = pd.read_csv("/content/MalwareArtifacts.csv", delimiter=',')

✓ [15] #Extracting artifacts samples fields AddressOfEntryPoint and DLLCharacteristics
0s samples = malware_dataset.iloc[:, [0, 4]].values
targets = malware_dataset.iloc[:, 8].values

✓ [16] #Split dataset into training and test data
0s from sklearn.model_selection import train_test_split
training_samples, testing_samples, training_targets, testing_targets = train_test_split(samples, targets, test_size = 0.2, random_state = 0)

✓ [20] #Calculate results and display accuracy
0s from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
lr.fit(training_samples, training_targets)
accuracy = lr.score(testing_samples, testing_targets)
print("Logistic regression accuracy : " + str(accuracy))

Logistic regression accuracy : 0.7058823529411765
```

```

[21] #Calculating TP, FP, TN, FN, Precision, Recall and F1 Score
res = lr.predict(samples)
mt = confusion_matrix(targets, res)
print("True positive rate : %f %% " % ((mt[0][0] / float(sum(mt[0])))*100))
print("False positive rate : %f %% " % ((mt[0][1] / float(sum(mt[0])))*100))
print("True negative rate : %f %% " % ((mt[1][1] / float(sum(mt[1])))*100))
print("False negative rate : %f %% " % ((mt[1][0] / float(sum(mt[1])))*100))
precision = (mt[0][0]/(mt[0][0] + mt[0][1]))
print('Precision : ', precision)
recall = (mt[0][0]/(mt[0][0] + mt[1][0]))
print('Recall : ', recall)
f1score = 2 * (precision * recall) / (precision + recall)
print('F1 score : ', f1score)

True positive rate : 99.990676 %
False positive rate : 0.009324 %
True negative rate : 1.036219 %
False negative rate : 98.963781 %
Precision : 0.9999067608727182
Recall : 0.7044470881899993
F1 score : 0.826567096434398

```

EXPERIMENTAL RESULTS:

On acquiring the results, we observe that the Random Forest has an accuracy of 96.36%, decision tree has an accuracy of 96.27% and logistic regression has an accuracy of 70.58%

PARAMETERS USED TO EVALUATE:

	Logistic Regression	Decision Tree	Random Forest
Accuracy	70.588%	96.271%	96.365%
Precision	0.999	0.993	0.992
F1 score	0.826	0.993	0.993
Recall	0.704	0.992	0.993
True positive	99.990676%	99.357686%	99.298635%
False positive	0.009324%	0.642314%	0.701365%
False negative	98.963781%	1.774280%	1.647197%
True negative	1.036219%	98.225720%	98.352803%

INFERENCES:

We infer that for the given dataset, Random Forest performs best and Logistic regression lags behind.

REFERENCES:

- [1] Yadigar Imamverdiyev, Fargana Abdullayeva “Deep Learning in Cybersecurity: Challenges and Approaches” (2020)
- [2] Alex Mathew “Machine Learning in Cyber-Security threats” (2020)
- [3] Casey Lorenzen, Rajeev Agarwal, Jason King “Determining Viability of Deep Learning on Cybersecurity Log analytics” (2018)
- [4] Yang Xin, Zhiu Liu, Lingshuang Kong, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, Chunhua Wang “Machine Learning and Deep Learning methods for Cybersecurity” (2018)
- [5] S.A. Sokolov, T.B. Iliev, I.S. Stoyanov “Analysis of Cybersecurity Threats in Cloud Applications Using Deep Learning Techniques” (2019)
- [6] Mathew, Alex, Machine Learning in Cyber-Security Threats (January 19, 2021). ICICNIS 2020, Available at SSRN: <https://ssrn.com/abstract=3769194> or
- [7] Sarker, I.H., Kayes, A.S.M., Badsha, S. et al. Cybersecurity data science: an overview from machine learning perspective. J Big Data 7, 41 (2020).
- [8] Imamverdiyev, Yadigar & Abdullayeva, Fargana. (2020). Deep Learning in Cybersecurity: Challenges and Approaches. International Journal of Cyber Warfare and Terrorism. 10. 82-105.
- [9] DOI:10.4018/IJCWT.2020040105
- [10] R, Vinayakumar; Alazab, Mamoun; KP, Soman; Srinivasan, Sriram; Venkatraman, Sitalakshmi; Pham, Quoc-Viet; et al. (2021): Deep Learning for Cyber Security Applications: A Comprehensive Survey. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.16748161.v1>
- [11] "School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China;
matlzu@163.com (T.M.); chengjianjun@lzu.edu.cn (J.C.);
yuyang_1102@163.com (Y.Y.)
- [12] 2 School of Mathematical and Computer Science, Ningxia Normal University, Guyuan 756000, China;
wangfen_2015@163.com

* Correspondence: chenxy@lzu.edu.cn; Tel.: +86-136-0931-9826"

