IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 5, June 2022

Comparative Analysis of Ransomeware Using Deep Learning

Aruna[1], Vivekanadan S J[2], Reni Hena Helan R[3], Abirami G[4], Dhatchayani L[5], Surjitha R[6]

Assistant Professor, Department of Computer Science and Engineering[1,2,3,4]

Dhanalakshmi College of Engineering, Chennai, TamilNadu, India

Abstract: *Ransomware is a type of virus that encrypts a victim's data and demands payment in exchange for it. Critical data belonging to a person or organisation is encrypted, making it impossible for them to access files, databases, or apps. In order to gain access, a ransom is asked. An automated solution based on machine learning based classification algorithms is proposed in the research to prevent fraudulent job postings on the internet. For checking fraudulent posts on the internet, many classifiers are utilised, and the results of those classifiers are compared in order to determine the optimum employment scam detection model. For the detection of fake job postings, two types of classifiers are used: single classifiers and ensemble classifiers.*

Keywords: Ransomware

I. INTRODUCTION

Ransomware is becoming a major cyber threat to businesses and individuals all around the world. Ransomware hijacks the system and demands money to stop the attack, unlike ordinary malware. According to a recent study, the number of new ransomware strains and attacks has exploded, with little sign of slowing down. Ransomware samples can be differentiated from other samples within the same family using deep learning classification methods. The goal of this research is to use behaviour to identify new ransomware versions that have been updated. We looked at behavioural reports from different ransomware families and categorised modified ransomware versions depending on their behaviour in our study. It's plausible to assume that samples from the same family of ransomware are related based on the existing classification of malware.

II. SYSTEM ANALYSIS EXISTING SYSTEM

Malware has become far more common in recent years. In fact, ransomware has become one of the most wellknown forms of cybercrime. In order to avoid successful malware attacks, ransomware mitigation strategies must be devised. Malware detection and mitigation have been the subject of various research. These research, on the other hand, concentrate on ransomware detection via HTTP. If an API log file is smaller than 10 KB in size, it is erased because we believe it is not being executed properly.

While the core approach for distributing and launching ransomware is similar, the individual traffic shapes for different malware flavours are likely to differ slightly depending on the ransomware developer.

Disadvantages

- Less accuracy.
- Loss of Data and Information.
- High economical data is loss.

2.1 Proposed System

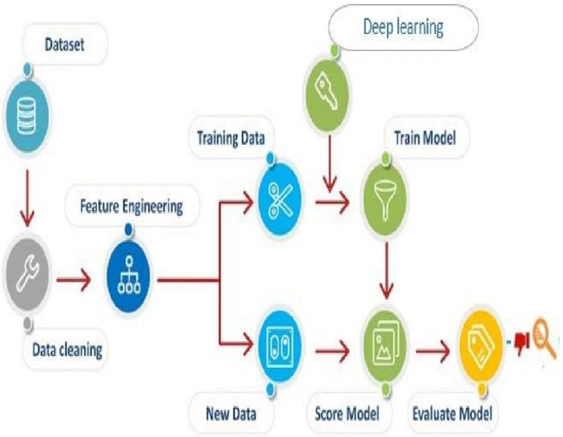
The Deep learning algorithm method is used to detect ransomware. The input files or data were divided into a training set and a testing set for fivefold crossvalidation, and deep learning techniques were applied. Each classifier performs classification after receiving a testing set as a hyper parameter. The classification findings were then given back into the Deep learning model to help it improve.

DOI: 10.48175/IJARSCT-4813

Advantages

- Different types of attack classify the security level increase.
- High accuracy

III. SYSTEM ARCHITECTURE



3.1 Modules Description

A. Data Collection and Pre-processing

The gathering of data is one of the most important tasks in the development of an AI model. It's a social gathering of errand-related data that's based on a few key elements to examine and produce a meaningful outcome. In any event, some of the facts may be shocking, such as qualities that are incorrect, insufficient, or wrong. As a result, it's necessary to manage the data before breaking it down and analysing the results. Cleaning, changing, and determining information should all be viable steps in the prehandling process.

B. Data Cleaning

Cleaning information entails filling in blanks, smoothing out tumultuous data, recognising and eliminating exceptions, and resolving inconsistencies. Smoothing, accumulation, conjecture, and change that improves the quality of the information are all examples of information change. Information selection refers to a set of approaches or abilities that enable us to select the most useful data for our system.

C. Data Transformation

The planning and transformation of information, beginning with one arrangement and progressing to the next, is known as information change. For example, XML data can be transformed from XML data valid for one XML Schema to a different XML record valid for a different XML Schema. Different approaches take into account the transformation of non-XML data to XML data.

D. Data Selection

The most frequent method of determining the proper information kind and source, as well as appropriate instruments to obtain information, is known as information determination. Information selection precedes the actual act of gathering information. The most popular method of selecting suitable data for an assessment project can compromise data integrity.

E. Data Input

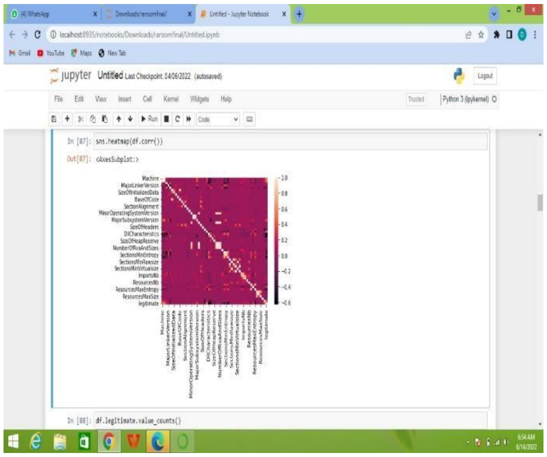
We used the formula to discover the spam site once we found the optimal calculation. We will then make a contribution to the calculation and determine the yield as a function of the yield.

DOI: 10.48175/IJARSCT-4813

F. Training and Testing

Train/Test is a technique for determining your model's accuracy. Because you split the data set into two sets: a training set and a testing set, it's termed Train/Test. Training accounts for 80% of the budget, while testing accounts for 20%. The training set is used to train the model.

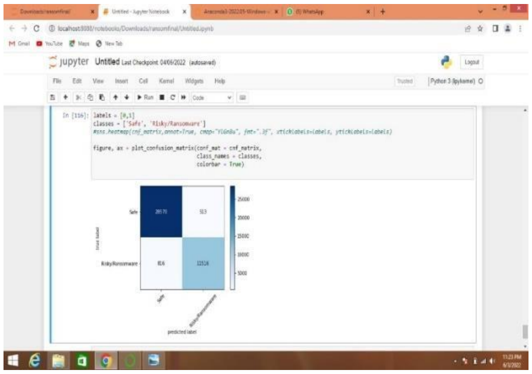
IV. OUTPUT RESULTS



```
In [87]: sns.heatmap(df.corr())
Out[87]: <AxesSubplot:>
```

```
In [88]: df.legitimate.value_counts()
```

Fig 1: Using CNN



```
In [116]: labels = [0,1]
          classes = ['Safe', 'Risky/Ransomware']
          #sns.heatmap(cnf_matrix,annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, yticklabels=labels)

          figure, ax = plot_confusion_matrix(conf_mat = cnf_matrix,
                                             class_names = classes,
                                             colorbar = True)
```

Fig 2: Accuracy from ANN

V. CONCLUSION

We investigated the use of deep learning techniques for malware classification based on malware sample behaviour in this article. To begin, we used Kaggle to collect behavioural analysis reports of ransomware samples and extract behavioural variables for categorization. We then conducted behavioural attribute selection trials in order to improve categorization outcomes. We discovered the collection of behavioural parameters that can be used for ransomware classification using an iterative technique to attain the best classification accuracy.

REFERENCES

[1]. AlexanderAdamov; Anders Accuracy depends on the size of the dataset. Here we Carlsson(2020),"Reinforcement Learning for have attained above 93% accuracy. Anti-Ransomware Testing" IEEE.

[2]. G Cusack, O Michel and E. Keller(2019), "API Call Based Ransomware Dynamic Detection Approach Using TextCNN". International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE).

[3]. Jack W. Stokes; KarthikSelvaraj; MadyMarinescu(2017), "Attention in Recurrent Neural Networks for

Ransomware

[4]. Detection" IEEE

[5]. Jagmeet Singh Aidan, Harsh Kumar Verma, Lalit Kumar Awasthi(2017), "Comprehensive Survey on Petya Ransomware Attack" International Conference on Next Generation Computing and Information Systems (ICNGCIS).

[6]. Jagmeet Singh Aidan; Harsh Kumar Verma; Lalit Kumar Awasthi(2017), "Comprehensive Survey on Petya Ransomware Attack", International Conference on Next Generation Computing and Information Systems (ICNGCIS).

DOI: 10.48175/IJARSCT-4813