iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52704

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Malware Detection Using Machine Learning and Deep Learning

Siddheshwar Midgule[1] Nikesh Konde[2] Suraj Kale[3] Prof. Sayli Haldavanekar[4]

Abstract: ***Android application security is based on permission-based mechanisms that restrict third-party Android applications' access to critical resources on an Android device. The user must accept a set of permissions required by the application before proceeding with the installation. This process is intended to inform users about the risks of installing and using applications on their devices. However, most of the time, even with a well-understood permission system, users are not fully aware of endangered threats, relying on application stores or the popularity of applications and relying on developers to You are accepting the install without trying to analyze your intent.***

Keywords: Android applications, malware detection, permission-related APIs, random forests, software security.

I. INTRODUCTION

Permissions are the foundation of the Android security concept. Permission is a security feature that restricts access to a portion of the code or data on the device.

The restriction is in place to safeguard sensitive data and code from being abused to distort or harm the user experience. Permissions are used to provide or deny access to APIs and resourcesthat are restricted. The Android INTERNET permission, for example, is needed for applications to execute network communications; hence, the INTERNET permission restricts the establishing of a network connection.

Malware Detection It is generally a program that is installed outside the user's will and can cause damage to both the operating system and the hardware (physical) elements of a computer. 1)Effects generated by the virus: 2)destruction of files. changing the file size. 3) Delete all information on the disc, including formatting it. destruction of the file allocation table, which makes it impossible to read the information on the disk.

II. LITERATURE SURVEY

1) *Malware Detection Using Machine Learning (2021) : Authors: Omar Sh. Ahmed Aboosh With over 852. Android Malware Detection based on Useful API Calls and Machine Learning (2018)*

Authors:Jaemin Jung .

Among security researchers and engineers, there has been anincrease in interest in and attempts to use machine learning techniques to detect malicious Android apps (also knownas Android malwares) [1-6]. Feature selection, which is the process of choosing a subset of pertinent characteristics tobe used in model creation, is crucial in machine learning generally [2,5,6]. The list of Android APIs called inside apps, as well as permissions and app intents, have been suggested as features in a number of recent research projects that used machine learning techniques to detect Android malware [1-4]. Using API call information as the feature, we providea machine learning-based methodology in this study. On the official website for Android developers, we select Android 7.0 Nougat (API level 24) as the particular operating system.

2) *Dynamic Malware Detection using API Similarity*

Authors: Ehab M. Alkhateeb.

Malware is a program intended to harm, damage orto perform illegal activities. According to AV-test [1], there are over 390,000 new malicious programs detectedevery day, therefore, malware detection and analysis is considered to be one of the most important field of Information Security.

Unfortunately, the currently used malware-detection techniques are insufficient, because the development of new malware and the modularization of attack tools increases the simplicity and effectiveness of these tools – unlike classic tools that implement one-type attack which gives cause for quickly evolving, and at maximum results in metamorphic and polymorphic tools [2]. Basically, signature-based detection remains today the main technique used for malware detection by Anti-Virus programs

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

III. PROPOSED SYSTEM

A. Login Module

User login

B. MasterPage Module

Master page module you can upload the dataset and checkthe analysis of that dataset.

C. Classification Module

Classification module we can used the naive bayes algo- rithm and support vector machine algorithm to classify the module.

1) Take the exam and submit paper.
2) Verifying answer using Natural Language process-ing(NLP).
3) Result Generation.
4) Display Evaluation Report.

IV. ALGORITHM

A. Naive Bayes

Naive Bayes is a fantastic illustration of how the most straightforward answers are frequently the most effective. Despite recent developments in machine learning, it has shown to be not only quick, accurate, and dependable but also simple. It has been used successfullyfor many things, but it excels at solving natural language processing (NLP) issues. The Bayes Theorem is the founda- tion of the probabilistic machine learning method known as Naive Bayes, which is utilised for a variety of classification problems. In order to eliminate any potential for misunder- standing, we will thoroughly explain the Naive Bayes methodin this post.

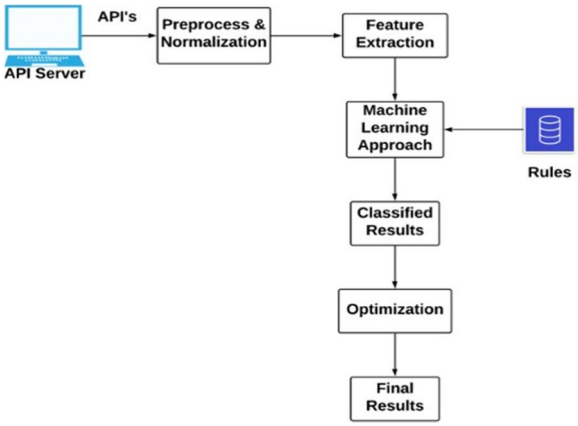


Fig.1. System Architecture

B. Support Vector Machine

The algorithm is implementedas below given steps. Input: Selected feature of all test instances D[i. . . .n], Training policies T[1]. ..... T[n]
Output: No. of probable classified with label. Step 1: Read (D into D[i]) V ←Extract features (D) Step 2: N ← Count Features(D) Step 3: for each (c into Train DB) Step 4: Nc[i]
> Ext Features(c) Step 5: select relevant features of w=Nc[i], N Step 6: Statement (w>t) Step 9: Return label

C. RNN

RNN Training Process for classification
Input: Training dataset TrainData[], Various activation functions[], Threshold Th
Output: Extracted Features Feature_set[] for completed trained module.

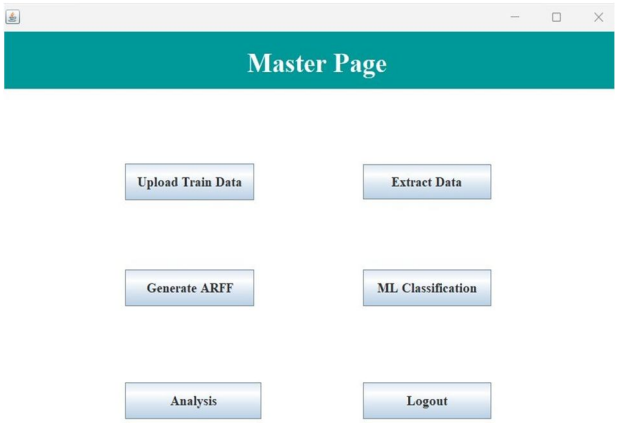
- Step 1: Set input block of data d[], activation function, epoch size,
- Step 2 : Features.pkl ExtractFeatures(d[])
- Step 3 : Feature_set[] optimized(Features.pkl)
- Step 4 : Return Feature_set[]

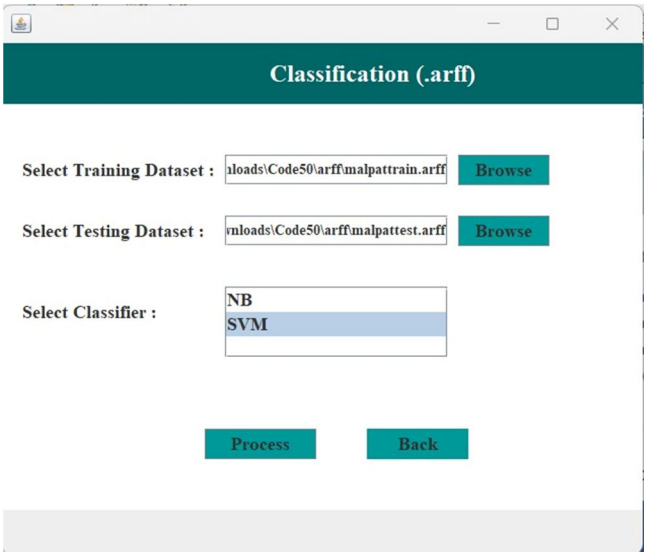
V. RESULTS

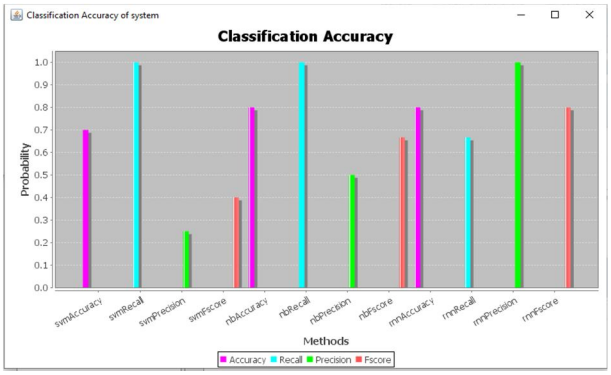
In our system we can detect the malware files in Android api using web base application

VI. CONCLUSIONS

In this paper, we explore the various different types of the modals which are used by various researchers in the malwaredetection and highlight the accuracy of these models. Asper the accuracy of the results we found that the Malware detection based on the Naive Bayes are more effective and accurate as compared to the other approache

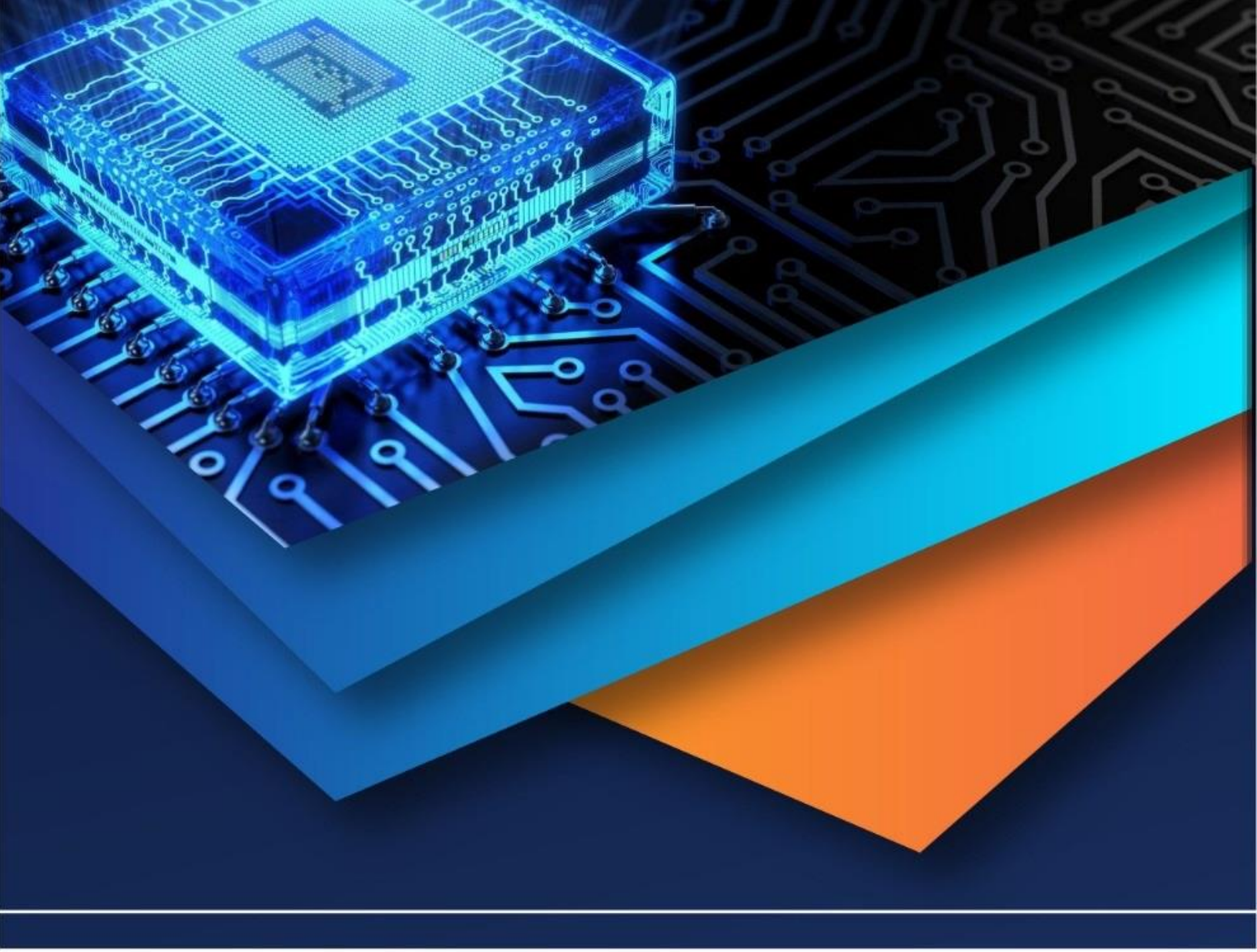






REFRENCES

[1] D. Arp, M. Spreitzenbarth, M. Hubner, H. Gascon, K. Rieck, and C.
Siemens, “DREBIN
[2] S. R. Sato, D. Chiba, S. Goto, “Detecting Android Malware by Manifest File Parsing,” Proceedings Asia-Pacific Advanced Network vol. 36, p. 23-31, 2
[3] Han, Zhuobing, et al. "Learning to predict severity of software vulner-ability using only vulnerability description." 2017 IEEE International conference on software maintenance and evolution (ICSME). IEEE, 2017.
[4] D. Arp, M. Spreitzenbarth, M. Hubner, H. Gascon, K. Rieck, and C. Siemens, “DREBIN: Effective ¨ and Explainable Detection of Android Malware in Your Pocket,” 2014M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
[5] Siddheshwar Midgule, Nikesh Konde, Suraj Kale, "Malware Detection Using machine Learning and Deep learning,” 2023



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY