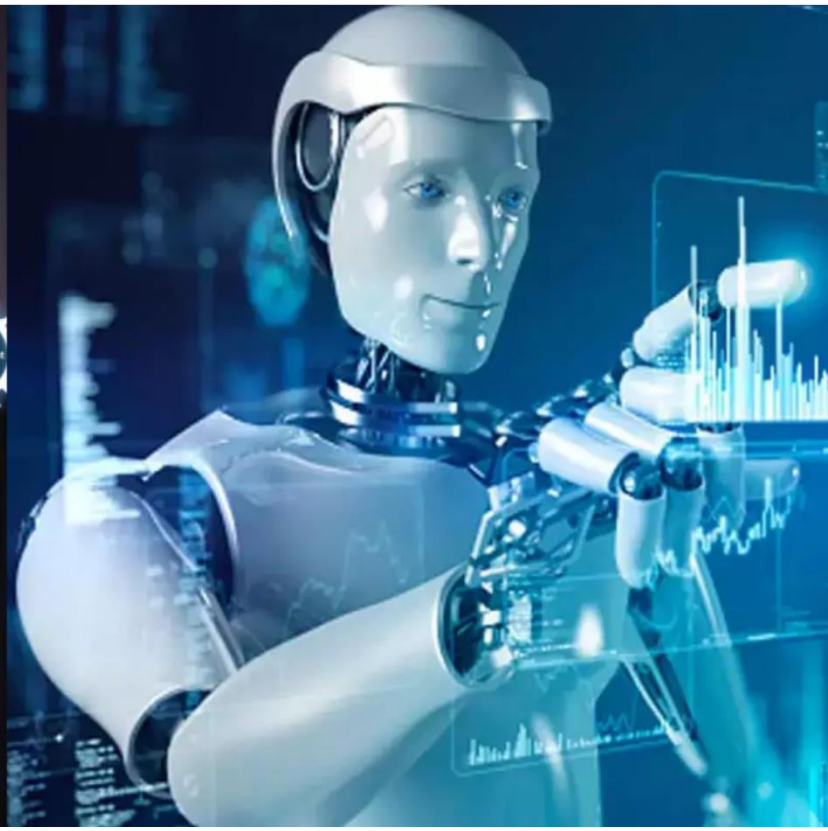
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*

**Impact Factor: 9.935**

**Volume 15, Issue 6, June 2026**

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 6, June 2026

|DOI: 10.15680/IJIRSET.2026.1506052|

# AI-Powered Android Malware Detection using Machine Learning

**Peram Venkata Lakshmi Narayana Reddy[1], K. Raman[2]**

PG Student, Dept. of CSE, Siddartha Educational Academy Group of Institutions, Tirupati, India[1]

Associate Professor, Dept. of CSE (AI&ML), Siddartha Educational Academy Group of Institutions, Tirupati, India[2]

**ABSTRACT:** The widespread adoption of Android smartphones has significantly increased the risk of malware attacks targeting mobile devices. Traditional signature-based malware detection methods often fail to identify newly emerging and obfuscated malicious applications. Artificial Intelligence (AI) and Machine Learning (ML) techniques offer a promising solution by automatically learning patterns from application behavior and detecting previously unseen malware. This research presents an AI-powered Android malware detection framework that utilizes machine learning algorithms to classify Android applications as benign or malicious. Features such as permissions, API calls, intents, and network activities are extracted from Android application packages (APKs). Several machine learning models, including Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Networks, are evaluated to determine their effectiveness. Experimental results demonstrate that AI-based approaches achieve high detection accuracy and significantly improve the identification of zero-day malware threats.

**KEYWORDS:** Android Malware, Artificial Intelligence, Machine Learning, Cybersecurity, Malware Detection, Deep Learning, APK Analysis.

## I. INTRODUCTION

Android is the most widely used mobile operating system globally due to its open-source nature and extensive application ecosystem. However, this popularity has made Android devices a prime target for cybercriminals. Malware applications can steal sensitive information, monitor user activities, send unauthorized messages, and cause financial losses.

Traditional malware detection systems rely on signature databases that require continuous updates and are ineffective against newly developed malware variants. AI-powered malware detection systems can analyze application characteristics and identify malicious behavior without requiring predefined signatures.

The objective of this research is to develop an intelligent malware detection framework that leverages machine learning techniques to improve Android security and detect both known and unknown malware threats.

## II. LITERATURE REVIEW

Several researchers have explored machine learning techniques for Android malware detection.

- Arp et al. proposed DREBIN, a lightweight malware detection framework using static analysis and machine learning.
- Yuan et al. developed DroidDetector, combining static and dynamic features for malware classification.
- Hou et al. applied deep learning techniques to classify Android applications based on behavioral characteristics.
- Recent studies have demonstrated that ensemble learning models such as Random Forest and Gradient Boosting achieve superior detection performance compared to conventional classifiers.

Despite significant advancements, challenges such as feature selection, computational overhead, and adversarial attacks remain open research problems.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 6, June 2026

|DOI: 10.15680/IJIRSET.2026.1506052|

## III. PROBLEM STATEMENT

Current Android malware detection systems face several limitations:

- Inability to detect zero-day malware.
- Dependence on manually updated signature databases.
- Poor performance against obfuscated malware.
- High false-positive rates in some detection systems.
- Limited scalability for large application repositories.

Therefore, there is a need for an AI-driven solution capable of accurately identifying malicious applications in real-time.

## IV. PROPOSED METHODOLOGY

**System Architecture**

The proposed framework consists of the following stages:

- APK Collection
- Feature Extraction
- Data Preprocessing
- Feature Selection
- Machine Learning Model Training
- Malware Classification
- Performance Evaluation

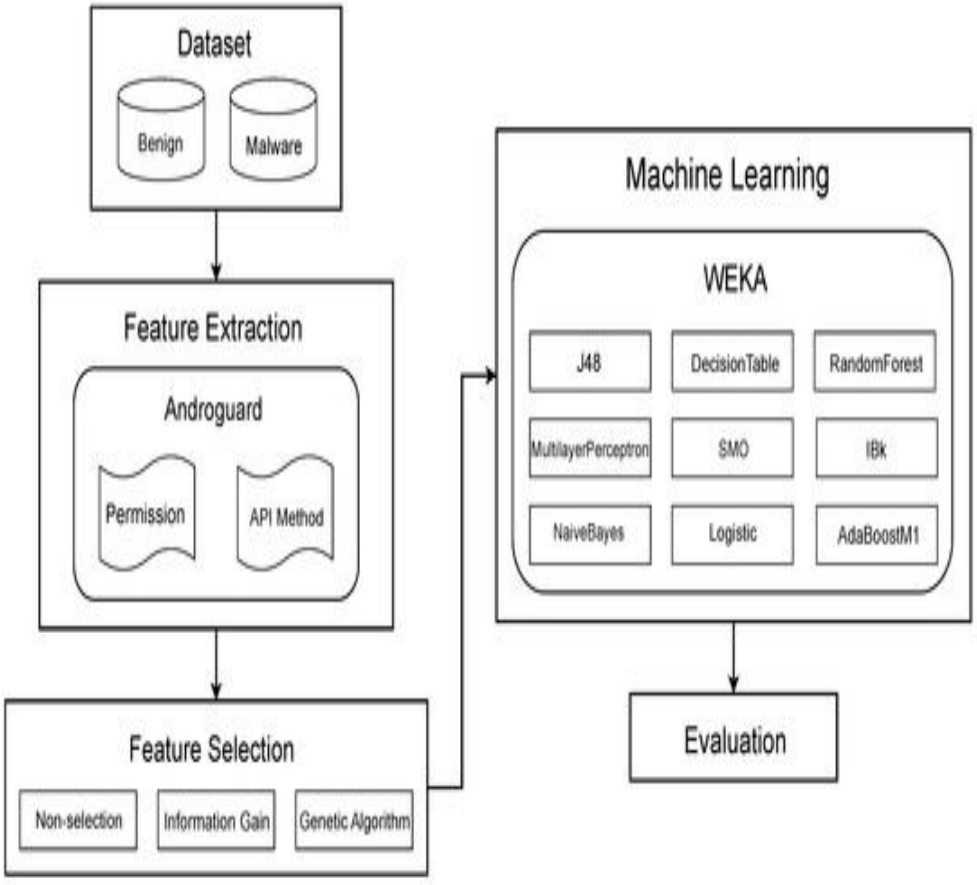



Fig.1: experiment progression.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 6, June 2026

|DOI: 10.15680/IJIRSET.2026.1506052|

**Dataset Collection**

- Android applications are collected from:
- Google Play Store (Benign Apps)
- Android Malware Genome Project
- CICMalDroid Dataset
- AndroZoo Repository

The dataset contains both benign and malicious applications representing different malware families.

**Feature Extraction**

- Static analysis features include:
- Requested permissions
- API calls
- Intents
- Hardware components
- Manifest file information

**Dynamic analysis features include:**

- CPU usage
- Network traffic
- System calls
- Battery consumption
- Runtime behavior

**Data Preprocessing**

- The extracted features are processed using:
- Missing value handling
- Feature normalization
- One-hot encoding
- Dimensionality reduction

**Machine Learning Algorithms**

The following algorithms are implemented:

**Decision Tree**

Provides interpretable classification rules.

**Random Forest**

Combines multiple decision trees to improve accuracy.

**Support Vector Machine (SVM)**

Separates malware and benign applications using optimal hyperplanes.

**Artificial Neural Network (ANN)**

Learns complex feature relationships through multiple hidden layers.

**Deep Learning Model**

A deep neural network automatically extracts meaningful patterns from feature data.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 6, June 2026

|DOI: 10.15680/IJIRSET.2026.1506052|

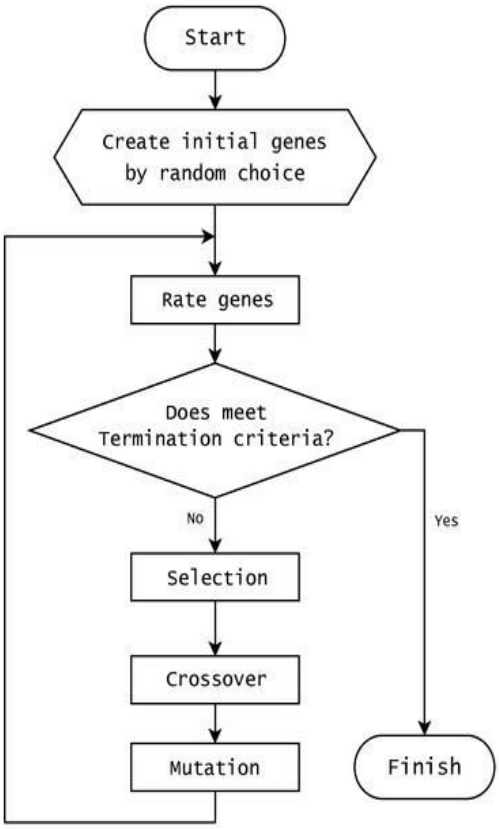



Fig.2: Simplified flowchart of the genetic algorithm's progression.

## V. EXPERIMENTAL RESULTS

**Performance Metrics**

The models are evaluated using:

1. Accuracy
2. Precision
3. Recall
4. F1-Score
5. ROC-AUC

| Algorithm | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| Decision Tree | 92.4 | 91.8 | 90.6 | 91.2 |
| SVM | 94.1 | 93.7 | 92.5 | 93.1 |
| Random Forest | 97.3 | 96.8 | 97.1 | 96.9 |
| Neural Network | 98.2 | 97.9 | 98.0 | 97.9 |

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 6, June 2026

|DOI: 10.15680/IJIRSET.2026.1506052|

The results indicate that AI-powered malware detection significantly outperforms traditional signature-based approaches. Machine learning models can identify previously unseen malware by learning behavioral and structural characteristics of malicious applications.

**Advantages include:**

- Detection of zero-day attacks
- Reduced reliance on signatures
- Improved scalability
- Faster classification

**Challenges include:**

- Dataset imbalance
- Adversarial machine learning attacks
- Resource constraints on mobile devices
- Privacy concerns during behavioral monitoring

## VI. CONCLUSION

This research presents an AI-powered Android malware detection framework using machine learning techniques. The proposed system effectively classifies Android applications based on extracted static and dynamic features. Experimental evaluation demonstrates that machine learning algorithms, particularly deep neural networks and Random Forest classifiers, achieve high detection accuracy and robustness against emerging malware threats. Future work will focus on federated learning, explainable AI, and lightweight edge-based malware detection models for real-time deployment on mobile devices.

## VII. FUTURE SCOPE

- Federated learning-based malware detection.
- Explainable AI for transparent security decisions.
- Real-time on-device malware analysis.
- Integration with cloud-based threat intelligence.
- Detection of advanced persistent threats (APTs).

## REFERENCES

1. Arp, D., et al. "DREBIN: Effective and Explainable Detection of Android Malware in Your Pocket." NDSS, 2014.
2. Yuan, Z., Lu, Y., and Xue, Y. "DroidDetector: Android Malware Characterization and Detection." IEEE TIFS, 2016.
3. Hou, S., et al. "Deep Neural Networks for Android Malware Detection." IEEE Access, 2017.
4. Sihag, V., et al. "Machine Learning Techniques for Android Malware Detection." Computers & Security, 2020.
5. Sharma, T., and Sahay, S. "Evolution and Detection of Android Malware." Procedia Computer Science, 2021.
6. Goodfellow, I., Bengio, Y., and Courville, A. Deep Learning. MIT Press, 2016.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details