

Project Milestone-1

Overview:

This project detects malicious Android applications using system call fingerprints. The initial model is a decision tree; future work will enhance data preprocessing and incorporate additional models.

Current Model & Results:

Model: Decision Tree

Hyperparameters:

- Maximum Depth = 15
- Minimum Samples Split = 2

Evaluation (Test Set): Precision = 0.823, Recall = 0.837, F1-score = 0.820

Kaggle Score: 0.84

The tuned decision tree effectively balances precision and recall.

Next Steps:

Data Preprocessing Enhancements: Apply normalization (zero mean, unit variance) to balance feature contributions and use feature engineering (e.g., PCA) to reduce noise and redundancy.

Model Expansion:

- **Perceptron Models:** Test standard, averaged, and margin variants.
- **Ensemble Methods:** Combine multiple models (e.g., perceptrons with decision trees) for enhanced robustness.
- **Other Models:** Incorporate neural networks to further boost classification accuracy.

Conclusion:

The decision tree (depth 15, min split 2) achieved promising results (F1 = 0.82 on test set; Kaggle score = 0.84). Future efforts will focus on improved preprocessing and expanding the model portfolio to enhance detection performance.