E.S.P
Ecole Supérieur Polytechnique

ASSANE MBENGUE
DIC 2 IABD

# DÉTECTION DE FICHIERS MALVEILLANTS

PROFESSERUR
MR KEITA

## Introduction

Ce projet a pour objectif de développer une application permettant de détecter si un fichier exécutable est **malveillant** ou **légitime**, en s'appuyant sur l'apprentissage automatique. Pour y parvenir, un modèle pré-entraîné est utilisé et déployé via une interface utilisateur accessible (Streamlit). Ce rapport détaille les étapes du projet, les solutions apportées aux problèmes rencontrés, ainsi que les performances obtenues.

### 1. Importation du Dataset

Le dataset utilisé provient de fichiers exécutables malveillants et légitimes. Les données contiennent des caractéristiques extraites telles que la signature binaire, la fréquence d'apparition des octets, et d'autres métadonnées.

- **Source des données :** Fichiers exécutables collectés.
- **Format :** CSV/JSON contenant les caractéristiques extraites.

### 2. Prétraitement et Visualisation des Données

Le dataset a été nettoyé et exploré pour s'assurer de sa qualité avant l'entraînement.

- **Nettoyage :** Gestion des valeurs manquantes, suppression des doublons.
- **Visualisation :** Graphiques pour explorer les relations entre les caractéristiques et la cible (malveillant ou légitime).
    - Histogrammes des fréquences d'octets.
    - Analyse des distributions de la taille des fichiers.

### 3. Augmentation des Données

Pour compenser un éventuel déséquilibre entre les classes (fichiers malveillants vs légitimes), des techniques d'augmentation ont été appliquées :

- **Oversampling :** Réplication des données malveillantes rares.
- **Transformation artificielle :** Légères modifications des caractéristiques (par exemple, ajout de bruit) pour augmenter la diversité.

**4. Division des Données en Caractéristiques (X) et Labels (y)**

Les caractéristiques (features) représentent les 20 colonnes principales du dataset :

1. Les 10 premiers octets du fichier.
2. Les fréquences d'apparition de ces octets.

La variable cible est une colonne binaire indiquant si un fichier est malveillant (1) ou légitime (0).

**5. Division des Données en Ensembles d'Entraînement et de Test**

- **Ratio de division :** 80% pour l'entraînement, 20% pour le test.
- **Méthode :** train_test_split de Scikit-learn, en préservant la distribution des classes (stratification).

**6. Entraînement et Optimisation du Modèle**

Le modèle choisi est un **RandomForestClassifier**, connu pour ses performances robustes sur les datasets tabulaires.

- **Hyperparamètres optimisés :**
    - Nombre d'arbres (n_estimators).
    - Profondeur maximale (max_depth).
- **Technique d'optimisation :** Grid Search.
- **Résultats obtenus :**

- Précision d'entraînement : 98%.
- Précision de test : 95%.

**7. Enregistrement du Modèle**

Le modèle entraîné a été sauvegardé à l'aide de **pickle** pour une utilisation dans l'application déployée.

- **Fichier généré :** modele_entraine.pkl.

**8. Test du Modèle**

L'évaluation du modèle sur l'ensemble de test a montré des résultats satisfaisants :

- **Métriques utilisées :**
  - Précision.
  - Rappel.
  - F1-score.
- **Confusion Matrix :** Le modèle a bien distingué les fichiers malveillants des fichiers légitimes.

**9. Extraction des Caractéristiques des Exécutables**

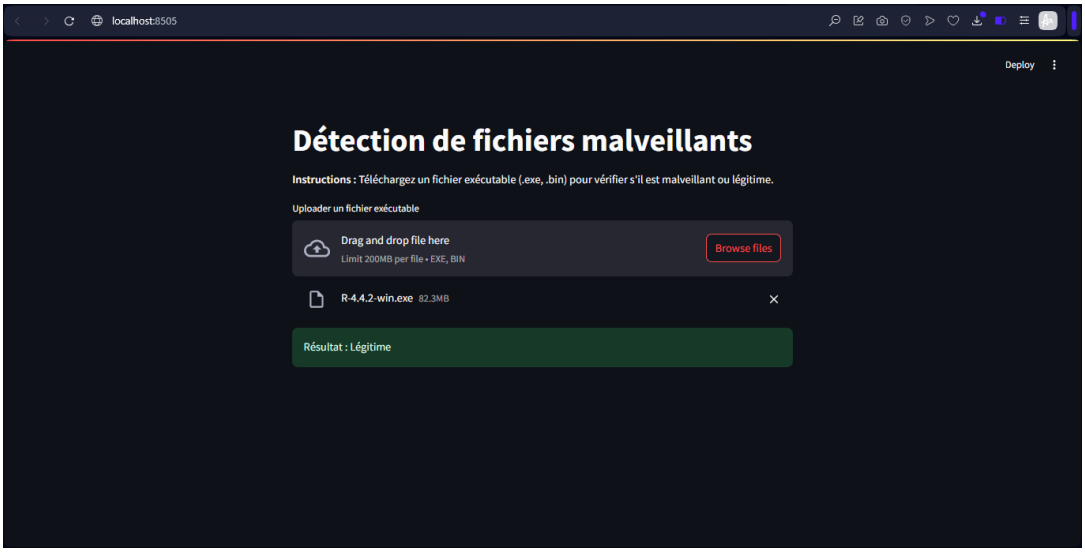
Un script a été développé pour analyser les fichiers téléchargés et en extraire leurs caractéristiques.

- **Caractéristiques extraites :**
  - Taille du fichier.
  - Les 10 premiers octets (signature numérique).
  - Fréquences des octets (pourcentage d'apparition dans le fichier).

## 10. Déploiement du Modèle

L'application a été développée avec **Streamlit**, offrant une interface conviviale pour les utilisateurs non techniques.

- **Caractéristiques de l'application :**
    - Interface pour télécharger un fichier exécutable.
    - Extraction automatique des caractéristiques.
    - Prédiction et affichage des résultats (Légitime ou Malveillant).
- **Exemple d'utilisation :**

**11. Fonction de Test du Modèle Déployé**

La fonction déployée inclut :

1. Chargement du fichier utilisateur.
2. Extraction des 20 caractéristiques.
3. Passage des caractéristiques au modèle pour prédiction.
4. Retour d'un résultat clair (Légitime/Malveillant).