

Projet : Catégorisation de Fruits

Author : Badr TAJINI - Introduction to AI - ESIEE-IT 2024-2025

Cette première partie de notre projet se concentre sur des données tabulaires avec des caractéristiques simples et compréhensibles. Notre objectif, donc, est de comprendre la logique du machine learning en mode simplifié.

1. Collecte de Données (Manuelle ou Fournie) :

- **Scénario** : Imaginez que vous avez un panier de fruits et que vous voulez apprendre à un ordinateur à les identifier.
- **Caractéristiques** : Au lieu de pixels, nous utiliserons des caractéristiques simples :
 - Couleur (par exemple, Rouge, Jaune, Vert, Orange)
 - Taille (par exemple, Petite, Moyenne, Grande)
 - Forme (par exemple, Ronde, Allongée)
- **Étiquettes** : Le type de fruit (par exemple, Pomme, Banane, Orange, Raisin).
- **Format des Données** : Vous pouvez fournir aux étudiants un tableau simple (comme dans une feuille de calcul ou un simple fichier texte) qui ressemble à ceci :

```
Couleur,Taille,Forme,Fruit  
Rouge,Moyenne,Ronde,Pomme  
Jaune,Allongée,Courbée,Banane  
Orange,Moyenne,Ronde,Orange  
Vert,Petite,Ronde,Raisin  
Jaune,Allongée,Courbée,Banane  
Rouge,Petite,Ronde,Pomme  
... et ainsi de suite
```

2. Compréhension des Données :

- **Pseudo-code** :

```
# Regardez le tableau de données sur les fruits.  
# Quelles sont les différentes colonnes ? (Couleur, Taille, Forme, Fruit)  
# La colonne 'Fruit' est ce que nous voulons prédire (l'étiquette).  
# Les autres colonnes (Couleur, Taille, Forme) sont les caractéristiques que nous  
utilisons pour faire la prédiction (les attributs).
```

- **Activité** : Demandez aux étudiants d'examiner manuellement le tableau et de discuter :
 - Quelles sont les différentes couleurs de fruits ?
 - Quelles sont les différentes tailles ?
 - Quelles sont les différentes formes ?
 - Peuvent-ils voir des schémas ? Par exemple, les bananes sont-elles généralement jaunes et allongées ?

3. Construction d'un Classificateur Simple Basé sur des Règles (Approche Manuelle) :

- **Pseudo-code** :

```
# Essayons de créer des règles simples pour deviner le fruit en fonction de ses  
caractéristiques.  
  
# Règle 1 : SI la couleur est Jaune ET la forme est Allongée, ALORS le fruit est  
probablement une Banane.  
# Règle 2 : SI la couleur est Rouge ET la taille est Moyenne ET la forme est  
Ronde, ALORS le fruit est probablement une Pomme.  
# Règle 3 : SI la couleur est Orange ET la taille est Moyenne ET la forme est  
Ronde, ALORS le fruit est probablement une Orange.  
# ... (Les étudiants peuvent créer d'autres règles)
```

- **Activité** : Demandez aux étudiants de travailler en groupes pour créer leur propre ensemble de règles pour identifier les fruits. Cela introduit le concept d'un modèle sans aucune programmation.

4. Test des Règles (Évaluation Manuelle) :

- **Pseudo-code** :

```
# Voyons si nos règles fonctionnent bien.  
# Prenez un nouveau fruit (ou une ligne du tableau que nous n'avons pas utilisée  
pour créer les règles).  
# Vérifiez si nos règles prédisent correctement le type de fruit.  
  
# Exemple :  
# Fruit : Vert, Petite, Ronde  
# La règle 1 correspond-elle ? Non.  
# La règle 2 correspond-elle ? Non.  
# La règle 3 correspond-elle ? Non.  
# Nous avons besoin d'une nouvelle règle pour les fruits verts, petits et ronds  
(comme les Raisins).  
  
# Comptez combien de fois nos règles devinent correctement le fruit.  
# C'est comme mesurer la précision de notre "modèle".
```

- **Activité** : Fournissez aux étudiants quelques fruits de "test" (lignes des données) et demandez-leur d'appliquer manuellement leurs règles pour voir s'ils peuvent correctement identifier le fruit. Discutez des cas où les règles échouent et pourquoi.

5. Introduction à l'Apprentissage Automatique (Conceptuel) :

- **Expliquez** : "Ce que nous venons de faire en créant des règles est similaire à ce que fait un modèle d'apprentissage automatique. Au lieu que ce soit nous qui créons les règles, l'ordinateur apprend les règles à partir des données."
- **Analogie** : "Imaginez que vous apprenez à un enfant à reconnaître les fruits. Vous lui montrez des exemples de pommes, de bananes, etc., et il commence à apprendre les schémas (couleurs, formes) qui l'aident à identifier chaque fruit."

5.1. Passer à la Programmation de Base (Optionnel et Graduel)

:

Nous introduisons du Python très basique dans le but d'automatiser le système basé sur des règles ou d'utiliser un modèle pré-construit très simple.

- **Utilisation de Dictionnaires ou de Listes** : Au lieu de règles manuelles, représentez les règles dans le code :

```
# Exemple en pseudo-python (très basique)
regles = {
    ("Jaune", "Allongée"): "Banane",
    ("Rouge", "Moyenne", "Ronde"): "Pomme",
    ("Orange", "Moyenne", "Ronde"): "Orange"
}

caracteristiques_fruit = ("Jaune", "Allongée")
if caracteristiques_fruit in regles:
    fruit_predit = regles[caracteristiques_fruit]
    print("Le fruit prédit est :", fruit_predit)
else:
    print("Je ne connais pas encore ce fruit !")
```

5.2. Introduction d'un Modèle Scikit-learn Simple (Très Basique) :

- Avec la librairie scikit-learn installée, nous allons introduire un exemple très basique avec un arbre de décision :

```
# Pseudo-code très simplifié pour scikit-learn
from sklearn.tree import DecisionTreeClassifier

# Nos données de fruits (simplifiées)
attributs = [("Rouge", "Ronde"), ("Jaune", "Allongée"), ("Orange", "Ronde")] # Il
faudra encoder ces valeurs
etiquettes = ["Pomme", "Banane", "Orange"]

# Créer un modèle simple
modele = DecisionTreeClassifier()

# Apprendre au modèle (l'entraîner)
modele.fit(attributs, etiquettes)

# Faire une prédiction
nouveaux_attributs_fruit = ["Rouge", "Ronde"]
prediction = modele.predict([nouveaux_attributs_fruit])
print("Le fruit prédit est :", prediction)
```

Important : Si vous suivez cette voie, concentrez-vous sur les *étapes* (charger les données, créer le modèle, entraîner, prédire) plutôt que de vous enliser dans la syntaxe.

Une fois cette étape terminée, nous pourrions explorer la vraie partie du code (script `classification_fruits.py`), l'exécuter et observer les résultats.