

EXAMEN

Matière: IML
 Enseignants: A.Dridi, Naïma Daghfous
 Classes :DSI3* , MDW3

Calculatrice Scientifique: **Autorisée**

Documents: **Non autorisés**
 Durée: **1h 30'**
 Date: JANVIER/2023

EXERCICE 1

3 Pts

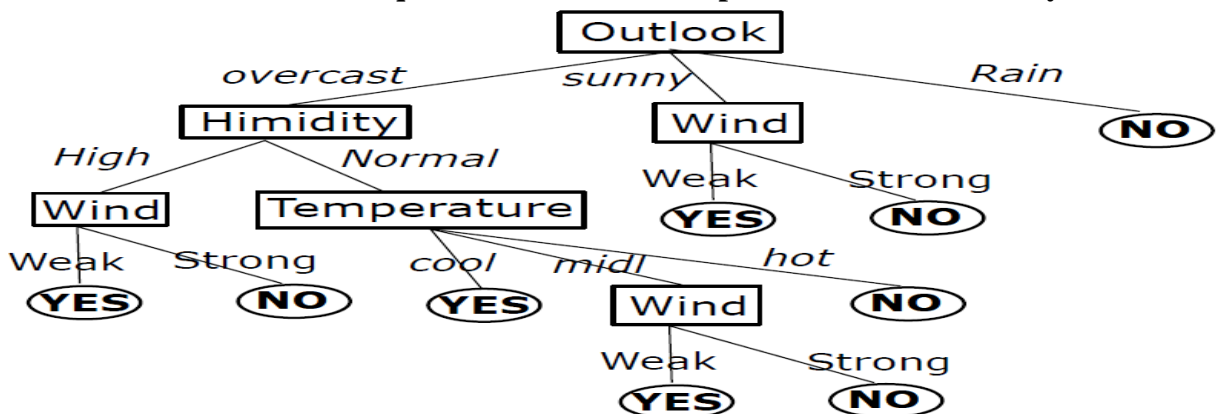
Donner les noms des algorithmes des modèles de machine Learning du tableau ci-dessous

Nom de l'algorithme	Principe de l'algorithme
.....	L'algorithme est un discriminant d'apprentissage supervisé non paramétrique, qui utilise la proximité pour effectuer des classifications ou des prédictions sur le regroupement d'un point de données individuel.
....	L'algorithme trouve un hyperplan qui sépare les données d'un jeu en classes spécifiques et maximise les marges (les distances entre les frontières de séparation et les échantillons les plus proches) de chaque classe.
.....	L'algorithme représente graphiquement les données en branches pour montrer les résultats possibles de diverses actions. Il classe et prédit les variables de réponse en fonction des décisions passées.

EXERCICE 2

6 Pts

A) Soit l'arbre de décision suivant qui prédit si nous pouvons jouer au ballon ou pas en fonction des conditions climatiques. En l'utilisant compléter le tableau TabPlayBall



Outlook	Temperature	Humidity	Wind	Playball ???
Sunny	Hot	High	Weak	...
Overcast	Mild	High	Strong
Overcast	Mild	Normal	Weak	...
Rain	Mild	High	Strong

Tableau : TabPlayBall

B) Nous utilisons sur un ensemble de 8 personnes les attributs (âge, mal de tête, fièvres) pour prédire si la personne a la grippe ou pas. Ces données sont résumées dans le tableau TabGrippe, construire un arbre de décision en utilisant l'algorithme CART (détailler les calculs)

âge	Mal de tête	fièvre	Grippe ???
adulte	doux	non	non
âgé	doux	oui	oui
jeune	fort	oui	oui
jeune	doux	non	non
adulte	néant	non	non
âgé	fort	non	oui
adulte	fort	non	oui
âgé	doux	non	non

Tableau : TabGrippe

EXERCICE 3

5 Pts

Soit un ensemble 'E' de 15 observations avec $Y_i \in \{\text{triangle, carré, rond}\}$ représenté sur le quadrillage de la figure QuadRéalité

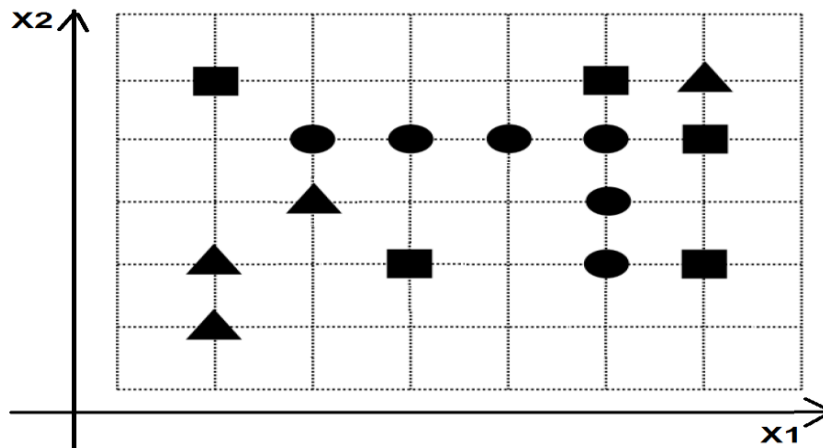


Figure : QuadRéalité

Nous avons construit un modèle dont les prévisions sur cet ensemble 'E' sont représentées sur le quadrillage de la figure QuadPrédiction

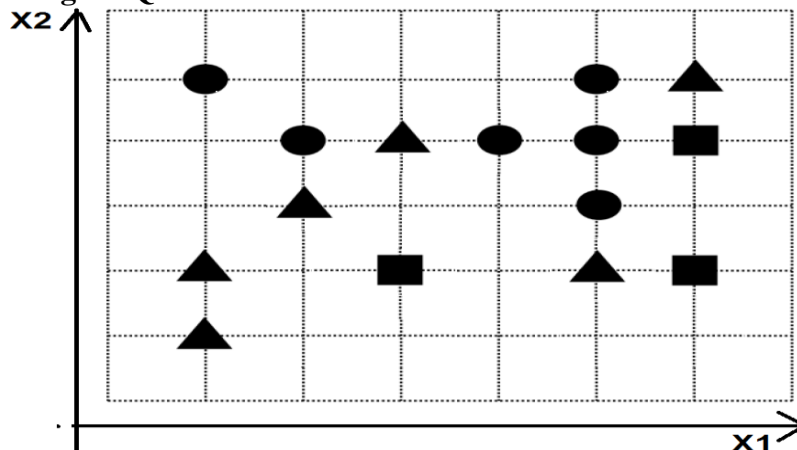


Figure : QuadPrédiction

- 1) Calculer les matrices de confusion du modèle
- 2) Calculer la précision du modèle en utilisant l'approche micro.

$$\text{Precision} = \frac{tp}{tp + fp}$$

EXERCICE 4**6PTS**

Le fichier « orientation.csv » contient un dataset avec
3 features : il s'agit de 3 notes pour 3 modules de bases (NoteMA, NoteMB, NotesMC) ,
2 classes : il s'agit des deux parcours MDW et DSI
de 500 étudiants.

Nous voulons créer un modèle pour prédire le parcours qui convient le mieux à un étudiant en se basant sur ses (NoteMA, NoteMB, NotesMC).

Nous hésitons entre deux modes : un K_NN (avec K=7) ou SVM (avec kernel='linear').

Donc nous voulons créer ces deux modèles et les comparer

Voici le début du code qu'il faut compléter

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix, classification_report

df=pd.read_csv('orientation.csv')
X=df.drop('Parcours',axis=1) #features
Y=df['Parcours']           #target
```

A compléter ...

Bon travail 😊