  Énoncés des exercices

#### Les données et la problématique

Étant donné un ensemble de données de m exemples d'apprentissage, chacun contenant des informations sous la forme de diverses caractéristiques et d'une étiquette. Chaque étiquette correspond à une classe, à laquelle appartient l'exemple d'apprentissage. Dans la classification multi classe, nous avons un ensemble fini de classes. Chaque exemple de formation a également n fonctionnalités.

Par exemple, dans le cas de l'identification de différents types de fruits, « Forme », « Couleur », « Rayon » peuvent figurer et « Pomme », « Orange », « Banane » peuvent être des étiquettes de classe différentes.

Dans une classification multi classe, nous formons un classificateur à l'aide de nos données d'apprentissage et utilisons ce classificateur pour classer de nouveaux exemples.

Objectif de ce TP

– Nous utiliserons différentes méthodes de classification multiclasses telles que KNN, Arbres de décision, SVM, etc. Nous comparerons leur précision sur les données de test. Nous allons effectuer tout cela avec sci-kit learn (Python).

*Les données*

Les jeux de données utilisés par scikit contiennent un certain nombre d'éléments, parmi lesquels :

***.data*** : est un tableau de dimensions nm ou n est le nombre d'instances, et m le nombre d'attributs

.***target :*** stocke les classes (etiquettes) de chaque instance (dans le cas supervise) .***target\_names***: donne le nom des classes .feature\_names donne le nom des attributs

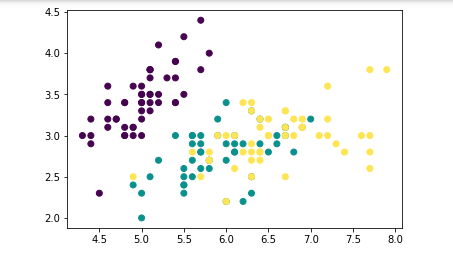
.***DESCR*** : est une description complete du jeu de donnees au format texte.

# charge tous les jeux de donnees

from sklearn import datasets

irisData=datasets.load\_iris()

*Approcher* -

1. ****Une fois notre dataset chargé, comme nous l'avons vu dans le TP précédent, explorer les donner et afficher les.
2. Divisez l'ensemble de données en données « d'entraînement » et « de test ».
3. Entraînez les classificateurs Arbre de décision, SVM et KNN sur les données d'entraînement.
4. Utilisez les classificateurs ci-dessus pour prédire les étiquettes des données de test.
5. Mesurez la précision et visualisez la classification pour chaque classifieur.