

# TP2 IAA

## Exercice 1 :

fichier artificialData.py

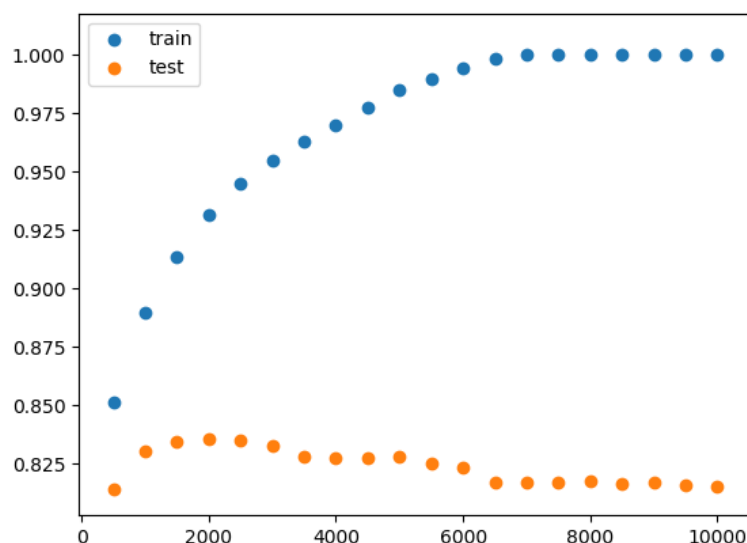
## Exercice 2 :

1//TP2prog1.py : Il y a 9 feuilles, cf iris.pdf.

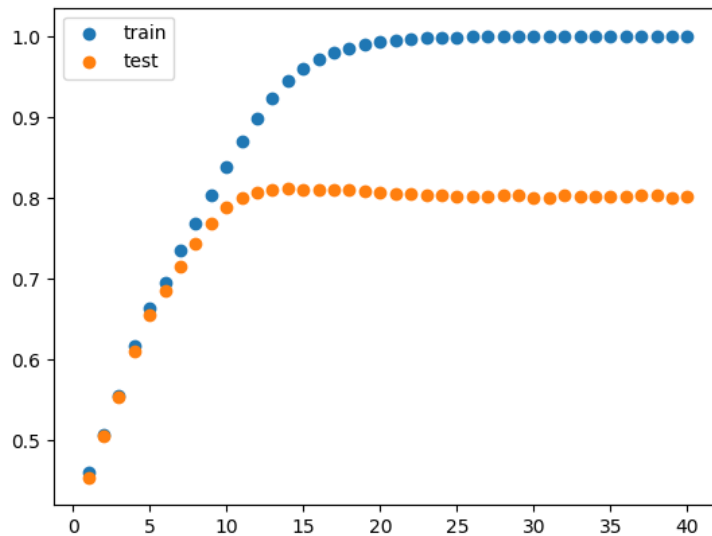
Pour les fichiers allant de iris9leaf.pdf jusqu'à iris3leaf.pdf , le nombre de feuilles diminue de 9 jusqu'à 3 .

2//TP2prog2.py :Les deux arbres dans les fichiers tree\_Entropie et tree\_Gini sont très similaires.Nous pensons que c'est dus à la petite taille des données.

3//TP2prog3.py : On constate que pour train il y a une augmentation du score jusqu'au maximum .Tandis que pour test il y a un pic du score autour de 2000 feuilles, puis le score décroît. On peut penser que pour train l'arbre de décision apprend par coeur tout les cas rencontré dans le jeu d'entrainement, de ce fait il ne fait quasiment plus d'erreur avec 10000 feuilles. L'arbre de décision passe par test et train de façon distincte , il a donc plus de chance de faire d'erreurs sur test .



4//TP2prog4.py : On remarque de train et test on un comportement similaire à la question précédente. Néanmoins avec une faible profondeur, ils sont sujet à beaucoup d'erreurs car ils ne peuvent pas explorer beaucoup de cas différents.



#### Exercice 4 :

TPprog5.py : On s'attend à trouver que 95 % des f soit dans I. Expérimentalement on trouve en effet aux alentours de 5 % de f en dehors de I.

#### Exercice 5 :

TPprog6.py :

-On peut en déduire que l'erreur du classifieur est très proche entre 0,15 pour Entropie et 0,16 pour Gini.

-D'après le test de McNemar nous ne concluons pas qu'un classifieur est meilleur que l'autre.