

SORBONNE UNIVERSITÉ

RAPPORT DE ARF

TME 1 - 6

SAMU TAMMINEN

Table des matières

1	TME 1 : Arbres de décision, sélection de modèles	2
1.1	Q1.3 Interpretation de l'entropie	2
1.2	Q1.4 - 6 La profondeur d'arbre	2
1.3	Q1.7 - Sur et sous apprentissage	2
2	TME 2 : Estimation de densité	3
3	TME 3 : Descente de gradient	3
4	TME 4 et 5 : Perceptron	3
5	TME 5 : Support Vector Machine	3
5.1	Pistes de recherche	3

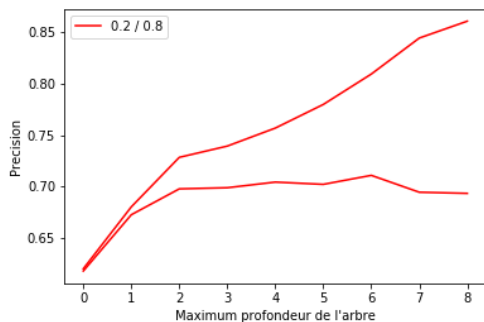


FIGURE 1 – Donnée apprentissage et test : 0.2 / 0.8

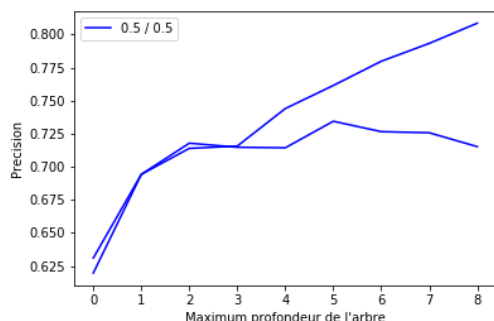


FIGURE 2 – Donnée apprentissage et test : 0.5 / 0.5

1 TME 1 : Arbres de décision, sélection de modèles

1.1 Q1.3 Interpretation de l'entropie

Calculer également la différence entre l'entropie et l'entropie conditionnelle pour chaque attribut. A quoi correspond une valeur de 0 ? une valeur de 1 ? Quel est le meilleur attribut pour la première partition ?

Quand la différence entre l'entropie et l'entropie conditionnelle est 0 quand la partition n'est pas du tout bon et un, quand la partition est le meilleur possible.

1.2 Q1.4 - 6 La profondeur d'arbre

En augmentant la profondeur maximale d'arbre, on obtient meilleur score. Mais ces scores ne sont pas un indicateur fiable, si nous ne faisons pas différence entre les données training et test.

1.3 Q1.7 - Sur et sous apprentissage

Le partition 0.8 / 0.2 1.3 semble d'être le plus robuste pour test et training.

Quand il y a que peu d'exemples d'apprentissage, l'erreur est plus élevée dans le test. Par contre quand on a beaucoup des d'exemples d'apprentissage, l'erreur dans le test est moins d'élévé.

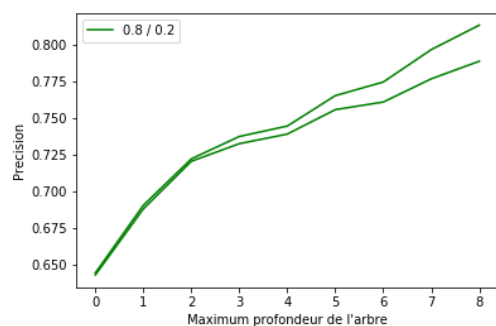


FIGURE 3 – Donnée apprentissage et test : 0.8 / 0.2

2 TME 2 : Estimation de densité

3 TME 3 : Descente de gradient

4 TME 4 et 5 : Perceptron

5 TME 5 : Support Vector Machine

5.1 Pistes de recherche