# Intelligence Artificielle - TD 7

Exercices partiellement tirés de `www.grappa.univ-lille3.fr/grappa/index.php3?info=apprentissage`

**Exercice 1** Quelques valeurs de $log_2$: $log_2(1/3) = -1.585$, $log_2(2/3) = -0.585$, $log_2(2/5) = -1.322$, $log_2(3/5) = -0.737$, $log_2(5/6) = -0.263$, $log_2(1/6) = -2.585$, $log_2(3/4) = -0.415$, $log_2(5/9) = -0.848$, $log_2(4/9) = -1.17$, $log_2(2/9) = -2.17$

On considère l'échantillon suivant :

| $N^o$ | $T_1$ | $T_2$ | $T_3$ | Classe |
|---|---|---|---|---|
| 1 | 0 | $V$ | $N$ | $A$ |
| 2 | 1 | $V$ | $I$ | $A$ |
| 3 | 0 | $F$ | $O$ | $B$ |
| 4 | 1 | $V$ | $N$ | $A$ |
| 5 | 1 | $V$ | $O$ | $A$ |
| 6 | 1 | $F$ | $N$ | $A$ |
| 7 | 0 | $F$ | $O$ | $B$ |
| 8 | 0 | $V$ | $I$ | $A$ |
| 9 | 0 | $F$ | $N$ | $B$ |
| 10 | 1 | $V$ | $I$ | $B$ |
| 11 | 1 | $F$ | $O$ | $A$ |
| 12 | 1 | $F$ | $I$ | $A$ |
| 13 | 0 | $V$ | $O$ | $B$ |

- Soit l'ensemble d'apprentissage constitué des exemples $\{1, \ldots 9\}$. Donnez le classificateur naïf de Bayes pour les trois descriptions des exemples 11, 12 et 13.
- Soit l'ensemble d'apprentissage constitué des exemples $\{1, \ldots 9\}$. Construire un arbre de décision $t_1$ en choisissant les attributs dans l'ordre $T_3$, $T_2$, $T_1$. Est-ce que $t_1$ est parfait ?
- Même question avec $t_2$ en utilisant l'ordre $T_1$, $T_2$, $T_3$.
- Calculer quel test réalise le gain maximal à la racine (en utilisant la mesure *gini*)
- Construire un arbre de décision parfait en choisissant, pour chaque nœud le meilleur test à lui appliquer (avec la mesure *entropie*).
- Peut-on trouver un arbre de décision parfait si on considère l'ensemble d'apprentissage constitué des exemples $\{1, \ldots 10\}$ ?
- Soit l'ensemble d'apprentissage $\mathcal{A}$ constitué des exemples $\{1, \ldots 9\}$ et l'ensemble test $\mathcal{T}$ constitué des exemples $\{11, 12, 13\}$. Soit les arbres de décision $t_3 = A$ et $t_4 = T_1(A, B)$. Calculer, pour chacun des arbres $t_1$, $t_2$, $t_3$ et $t_4$, l'erreur apparente sur l'ensemble d'apprentissage $\mathcal{A}$, sur l'ensemble test $\mathcal{T}$ et sur l'échantillon complet.

**Exercice 2** On reprend l'échantillon $\{1, \ldots, 9\}$ de l'exercice 1. On considère que $N^o$ est un attribut. Quel test est choisi en premier avec la mésure *Entropie* ? L'utilisation de *GainRatio* change-t-elle quelque chose ?

**Exercice 3** On considère un échantillon avec 200 individus qui peuvent être de classe 1 ou 2 et qui ont deux attributs binaires A et B. Il y a 100 individus de classe 1 et 100 de classe 2. Leur répartition est donnée comme suit :

| A | B | Classe 1 | Classe 2 |
|---|---|---|---|
| faux | faux | 0 | 50 |
| faux | vrai | 50 | 0 |
| vrai | faux | 50 | 0 |
| vrai | vrai | 0 | 50 |

La première ligne signifie qu'en tout il y a 50 individus dont les attributs A et B sont faux. Sur ces 50 individus 0 sont de classe 1 et 50 de classe 2.

- Donnez pour chaque valeur d'attribut le nombre d'individus qui sont dans chaque classe. Par exemple donnez le nombre d'individus de classe 1 avec attribut A faux, etc.
- En utilisant ces valeurs, appliquez l'algorithme de base pour construire un arbre de décision avec les fonctions Entropie et Gain pour cet échantillon. Quel problème rencontre-t-on pour choisir le premier test ? Doit-on s'arrêter ? Choisissez un attribut et poursuivez la construction de l'arbre. Que remarque-t-on ?
- Supposons maitenant que le langage de description contienne non seulement les attributs A et B mais aussi d'autres attributs C, ..., Z. Que va-t-on obtenir en appliquant l'algorithme de base ? Pourrait-on remédier à cette situation ?