

Examen – Apprentissage**1h30.****Aucun document autorisé.**

Prenez soin de lire tous les exercices avant de commencer. La notation est donnée à titre indicatif.

Exercice 1 : (5 pts)

1. Expliquer ce qu'est l'apprentissage artificiel (1 pt)
2. Donnez le principe des réseaux de neurones multicouches, et illustrez-le. (2 pts)
3. Comparez trois méthodes d'évaluation des modèles d'apprentissage. (2 pts)

Exercice 2 : (15 pts)

```
@relation weather.symbolic
@attribute outlook      {sunny, overcast, rainy}
@attribute temperature {hot, mild, cool}
@attribute humidity     {high, normal}
@attribute windy        {TRUE, FALSE}
@attribute play         {yes, no}

@data
overcast, hot,      high,      FALSE,    yes
rainy,    mild,     high,      FALSE,    yes
rainy,    cool,     normal,    FALSE,    yes
overcast, cool,     normal,    TRUE,     yes
sunny,    cool,     normal,    FALSE,    yes
rainy,    mild,     normal,    FALSE,    yes
sunny,    hot,      high,      FALSE,    no
sunny,    hot,      high,      TRUE,     no
rainy,    cool,     normal,    TRUE,     no
sunny,    mild,     normal,    TRUE,     yes
sunny,    mild,     high,      FALSE,    no
overcast, mild,     high,      TRUE,     yes
overcast, hot,      normal,    FALSE,    yes
rainy,    mild,     high,      TRUE,     no
```

On dispose du fichier ci-dessus possédant une variable de classe PLAY matérialisant les conditions d'entraînement d'un joueur de tennis. On découpe l'ensemble en 2 : D_1 et D_2 . **D_1 contient les 9 premiers objets**, et D_2 contient les **5 derniers**.

- 1- A quoi correspond le type de fichier ci-dessus ? Quel logiciel l'utilise ? (1 pt)
- 2- Construire le modèle bayésien naïf M_1 en utilisant D_1 et en appliquant la formule de Laplace (voir annexe) (4 pts)
- 3- Donner la matrice de confusion sur D_2 ; (2 pts)
- 4- On souhaite construire le modèle M_2 **d'arbre de décision** en utilisant *l'indice d'erreur en classification* (voir annexe)
 - a. Construire l'arbre de décision M_2 sur l'ensemble d'apprentissage D_1 ; (4 pts)
 - b. Donner sa matrice de confusion sur D_2 ; (2 pts)
- 5- Comparer les 2 modèles M_1 et M_2 ; (2 pts)

ANNEXES

A_i : une valeur de l'attribut A

N_{ic} : Nombre d'objets ayant la valeur A_i dans la classe c

N_c : Nombre d'objets de la classe c

k : nombre de valeurs de l'attribut A

p : probabilité a priori

m : paramètre

$$\text{Original: } P(A_i | C) = \frac{N_{ic}}{N_c}$$

$$\text{Laplace: } P(A_i | C) = \frac{N_{ic} + 1}{N_c + k}$$

$$\text{m-estimate: } P(A_i | C) = \frac{N_{ic} + mp}{N_c + m}$$

Arbres de décision

$p(j | t)$ est la fréquence relative de la classe j au nœud t.

$$GINI(t) = 1 - \sum_j [p(j | t)]^2$$

Indice de Gini pour le nœud t :

$$GINI_{split} = \sum_{i=1}^k \frac{n_i}{n} GINI(i)$$

Indice de Gini pour l'attribut *split* :

Gain d'information avec l'indice de Gini pour l'attribut *split* : $\text{Gain}_{split} = \text{Gini}(r) - \text{Gini}_{split}$

Le nœud parent **r** a n objets, et est divisé en k partitions. La partition i possède n_i objets.

$$Entropy(t) = -\sum_j p(j | t) \log p(j | t)$$

Entropie du nœud t:

$$GAIN_{split} = Entropy(p) - \left(\sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right)$$

Gain d'information avec l'entropie pour l'attribut *split*:

Le nœud parent **p** a n objets, et est partitionné en k partitions. La partition i possède n_i objets.

$$Error(t) = 1 - \max_i P(i | t)$$

Indice d'Erreur en classification au nœud t :

Gain d'information avec l'indice d'erreur en classification : $\text{Gain}_{split} = \text{Error}(r) - \text{Error}_{split}$

Le nœud parent **r** a n objets, et est partitionné en k partitions. La partition i possède n_i objets.

La **précision** pour une classe donnée mesure le taux d'exemples corrects parmi les exemples prédits dans cette classe.

Le **rappel** mesure le taux d'exemples corrects parmi les exemples de la classe.

Le taux de **faux positifs** d'une classe mesure le nombre d'objets positifs parmi ceux n'appartenant pas à la classe.

Le taux de **vrais positifs** d'une classe mesure le nombre d'objets positifs parmi les vrais objets de la classe.

Le taux de **faux négatifs** d'une classe mesure le nombre d'objets négatifs parmi ceux appartenant à la classe.

Le taux de **vrais négatifs** d'une classe mesure le nombre d'objets négatifs parmi ceux n'appartenant pas à la classe.

La **sensibilité** est la probabilité qu'un test soit positif si l'objet appartient à la classe.

La **spécificité** est la probabilité qu'un test soit négatif si l'objet n'appartient pas à la classe.