

TD 2

Exercice 1 :

- 1- En apprentissage supervisé, quelle est la différence entre un ensemble de test et un ensemble de validation ?
- 2- Citez les méthodes d'évaluation des modèles d'apprentissage.
- 3- Illustrer chacune d'elles à travers l'exemple fourni en annexe.

Exercice 2 :

classe réelle \ classe prédite	A	B	C	D	Total
A	41	1	1	2	45
B	1	12	1	1	15
C	1	2	18	0	21
D	1	1	0	17	19
Total	44	16	20	20	100

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. La *sensibilité* (ou sélectivité) d'un test mesure sa capacité à donner un résultat positif lorsqu'une hypothèse est vérifiée. Elle s'oppose à la *spécificité*, qui mesure la capacité d'un test à donner un résultat négatif lorsque l'hypothèse n'est pas vérifiée.

Soit la matrice ci-dessus obtenue par application du modèle M1. Pour les questions 2) à 9), vous devez **indiquer aussi la formule de calcul**.

- 1- Comment appelle-t-on la matrice ci-dessus ?
- 2- Calculer le taux d'erreur en généralisation ;
- 3- Calculer le taux de généralisation (accuracy rate) ;
- 4- Calculer le taux de faux positifs (FP rate) pour la classe A ;
- 5- Calculer le taux de vrais positifs (TP rate) pour la classe B ;
- 6- Calculer le taux de faux négatifs (FN rate) pour la classe A ;
- 7- Calculer le taux de vrais négatifs (TN rate) pour la classe B ;
- 8- Calculer la précision pour la classe C ;
- 9- Calculer le rappel pour la classe D ;
- 10- Calculer la sensibilité pour la classe A ;
- 11- Calculer la spécificité pour la classe B ;
- 12- Calculer la précision et le rappel du modèle M1 en tenant compte de toutes les classes.
- 13- Ecrire un algorithme (pseudo-code) qui, à partir de 2 matrices T1 et T2 identiques à celle ci-dessus, obtenues sur un même jeu de données (15324 exemples, 21 attributs) par application respective des modèles M1 et M2, donne comme résultat le modèle à choisir par l'utilisateur pour la résolution de son problème.
- 14- Comment peut-on s'assurer que ce mode de comparaison de modèles est robuste ?

Exercice 3 : Matrice de coût

Soit la matrice de coût suivante, où c_{ij} est le coût associé à la classification d'un exemple de la classe i (réelle) dans la classe j (prédite).

Matrice	Classe Prédite		
Classe réelle	c_{ij}	+	-
	+	-1	100
	-	1	0

Modèle M ₁	Classe Prédite		
Classe réelle		+	-
	+	300	80
	-	120	500

Modèle M ₂	Classe Prédite		
Classe réelle		+	-
	+	500	90
	-	10	400

- 1- Etant donné les matrices de confusion des 2 modèles M₁ et M₂, calculer l'erreur apparente sans tenir compte de la matrice de coût. Quel est le meilleur modèle ?
- 2- Calculer le coût de classification. Quel est le meilleur modèle ?

Exercice 4 :

Etant donné un problème d'apprentissage supervisé à deux classes A et B. L'échantillon de données comprend 10000 exemples répartis comme suit : 9990 pour la classe A, et 10 pour la classe B.

- 1- Quel problème pose cet échantillon ?
- 2- Etant donné le modèle d'apprentissage M₁ suivant : « Tout objet est classé dans la classe A ». Quelle est l'erreur apparente de ce modèle d'apprentissage ?
- 3- Etant donné le modèle d'apprentissage M₂ suivant : « Tout objet est classé dans la classe B ». Quelle est l'erreur apparente de ce modèle d'apprentissage ?
- 4- Peut-on espérer qu'un utilisateur accepte d'utiliser le modèle M₁ ?
- 5- Comment peut-on procéder pour construire un modèle qui soit moins biaisé par l'échantillon de données ?

Annexes :

```
@relation contact-lenses
```

```
@attribute age {young, pre-presbyopic, presbyopic}
@attribute spectacle-prescrip {myope, hypermetrope}
@attribute astigmatism {no, yes}
@attribute tear-prod-rate {reduced, normal}
@attribute contact-lenses {soft, hard, none}
```

```
@data
```

```
%
```

```
% 24 instances
```

```
%
```

```
young,myope,no,reduced,none
pre-presbyopic,myope,no,normal,soft
young,myope,yes,reduced,none
young,myope,yes,normal,hard
young,hypermetrope,no,normal,soft
presbyopic,myope,yes,normal,hard
young,hypermetrope,no,reduced,none
young,hypermetrope,yes,reduced,none
young,hypermetrope,yes,normal,hard
young,myope,no,normal,soft
pre-presbyopic,myope,no,reduced,none
pre-presbyopic,myope,yes,reduced,none
pre-presbyopic,myope,yes,normal,hard
pre-presbyopic,hypermetrope,no,reduced,none
pre-presbyopic,hypermetrope,no,normal,soft
pre-presbyopic,hypermetrope,yes,reduced,none
pre-presbyopic,hypermetrope,yes,normal,none
presbyopic,myope,no,reduced,none
presbyopic,myope,no,normal,none
presbyopic,myope,yes,reduced,none
presbyopic,hypermetrope,no,reduced,none
presbyopic,hypermetrope,no,normal,soft
presbyopic,hypermetrope,yes,reduced,none
presbyopic,hypermetrope,yes,normal,none
```