

# Contrôle SI 221

Novembre 2014

1h30

autorisés : calculatrice + polycopié + notes de cours

## Exercice 1

On suppose que l'on dispose d'un test médical qui permette de détecter la présence (classe  $\omega_2$ ) ou l'absence (classe  $\omega_1$ ) d'une maladie à partir d'un taux sanguin noté  $x$ . Quand un patient subit ce test, on extrait la valeur de son taux et on le compare à une valeur seuil  $x_0$ .

Si  $x \leq x_0$  le patient est estimé en bonne santé,  
si  $x > x_0$  le patient est estimé malade.

On veut maintenant évaluer la pertinence de ce test du point de vue théorique et pratique.

**Partie I** On suppose que les densités de probabilités conditionnelles  $p(x | \omega_i), i = 1, 2$  sont connues et gaussiennes de moyenne  $\mu_1$  et  $\mu_2$ , et de variance  $\sigma_1^2, \sigma_2^2$ , respectivement.

**Q. 1.1** Donner les expressions des densités conditionnelles en fonction de  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ .

**Q. 1.2** Remplir le tableau I en indiquant où se trouvent les fausses alarmes et les non détections, ainsi que les décisions correctes. On note  $C$ , la variable aléatoire de classe.

TABLE 1 –

	C=1	C=2
décision : absence de maladie		
décision : présence de maladie		

**Q. 1.3** Donner les expressions des probabilités de fausse alarme et de non détection en fonction des probabilités a priori (supposées connues), des moyennes  $\mu_1, \mu_2$ , des variance  $\sigma_1^2, \sigma_2^2$ , et du seuil  $x_0$ . Quelle est la probabilité d'erreur totale du système ?

**Partie 2** Dans cette partie, on veut estimer les performances du système de détection mais on ne dispose pas de la connaissance des densités de probabilité comme dans la partie 1. On a maintenant un ensemble de patients dont on connaît l'état réel de santé (bonne santé ou malade), et on leur fait subir le test. 4 cas de figure se présentent :

- un patient malade est correctement détecté comme malade. Le nombre de ces patients est noté  $VP$  (*Vrais Positifs*).
- un patient en bonne santé est incorrectement détecté comme malade. Le nombre de ces patients est noté  $FP$  (*Faux Positifs*).

- un patient malade est incorrectement détecté comme en bonne santé. Le nombre de ces patients est noté  $FN$  (*Faux Négatifs*).
- un patient en bonne santé est détecté comme en bonne santé. Le nombre de ces patients est noté  $VN$  (*Vrais Négatifs*).

**Q. 2.1**

Remplir de nouveau le tableau I mais en indiquant où se trouvent  $VP$ ,  $FP$ ,  $FN$  et  $VN$ . On note toujours  $C$ , la variable aléatoire de classe.

**Q. 2.2** On note  $P_{FA}^c$  la probabilité conditionnelle de fausse alarme :  $P_{FA}^c = P(\text{décider } \omega_2 \mid \omega_1)$ , et  $P_{ND}^c$  la probabilité conditionnelle de non détection :  $P_{ND}^c = P(\text{décider } \omega_1 \mid \omega_2)$ . Comment peut-on estimer  $P_{FA}^c$  et  $P_{ND}^c$  à partir de  $VP$ ,  $FP$ ,  $FN$  et  $VN$ .

**Q. 2.3** Comment définiriez-vous la probabilité conditionnelle de détection correcte et comment l'estimeriez-vous en fonction de  $VP$ ,  $FP$ ,  $FN$  et  $VN$  ?

## Exercice 2 : Apprentissage du perceptron monocouche

Soit un perceptron monocouche ayant comme fonction de transition  $f$ , la fonction seuil suivante : si  $a > \text{seuil}$ ,  $f(a) = 1$ , sinon  $f(a) = -1$ . Nous fixons ici le seuil à 0.2.

Soit la base composée de 4 exemples d'apprentissage :

	e1	e2	d
(1)	1	1	1
(2)	-1	1	-1
(3)	-1	-1	-1
(4)	1	-1	-1

e1 et e2 sont les entrées du réseaux de neurones et d la sortie correspondante, telle qu'étiquetée dans l'ensemble d'apprentissage.

**Q. 1** En appliquant l'algorithme d'apprentissage du perceptron avec un pas de modification des poids  $\eta$  égal à 0.1, déterminez les poids du perceptron. Vous initialiserez l'algorithme avec les valeurs suivantes ,  $w1 = -0.2$ ,  $w2 = 0.1$ .

**Q. 2** Représentez les données d'apprentissage dans un espace à 2 dimensions et tracez la droite séparatrice des 2 classes. Donnez son équation.