Intelligence Artificielle - TD 6

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

Exercice 1 Soit une population Π d'individus qui consiste en un échantillon composé d'ouvrier, de médecins et d'employés des télécoms. On décrit les individus par un attribut logique *hautdebit* qui vaut vrai si l'individu possède l'internet haut débit et faux sinon. L'espace de description est donc égal à l'ensemble $\{hautdebit, \overline{hautdebit}\}$. On souhaite répartir les individus en trois classes *ouvrier*, *medecin* et *telecom*. On dispose des informations suivantes :

classe k	*telecom*	*medecin*	*ouvrier*
$P(k)$	0.2	0.3	0.5
$P(hautdebit/k)$	1	0.9	0.45

Une première règle possible pour le choix de la fonction de classement C pourraît être : "attribuer à chaque description la classe majoritaire", c'est-à-dire celle pour laquelle $P(k)$ est maximum; c'est la règle *majoritaire*.

Une seconde règle consiste à raisonner ainsi : "si j'observe d, je choisis la classe pour laquelle cette observation est la plus probable", c'est-à-dire $P(d/k)$ est maximum. C'est la règle dite du *maximum de vraisemblance*.

Une troisième règle (*règle de Bayes*) consiste à attribuer à une description d la classe k qui maximise la probabilité $P(k/d)$ qu'un élément ayant d pour description soit de classe k. La quantité $P(k/d)$ peut être estimée en utilisant la formule de Bayes, il suffit donc de choisir la classe k qui maximise le produit $P(d/k)P(k)$.

Décrire sur cet exemple les trois fonctions $C_{majoritaire}$, $C_{vraisemblance}$ et C_{Bayes}.

On peut définir la probabilité d'erreur d'une fonction de classement de la façon suivante : soit C une fonction de classement, l'erreur $E(d)$ (ou probabilité d'erreur) pour une description d est la probabilité qu'un élément de la population Π de description d soit mal classé par C, l'erreur $E(C)$ d'une fonction de classement est la moyenne pondérée des erreurs sur les descriptions d. Calculer les erreurs pour les trois procédures de classification trouvées précédemment.

Exercice 2 La population Π est un ensemble de champignons. Il y a deux classes $\{1, 2\}$ de champignons, où 1 est la classe des champignons vénéneux. Le langage de description est constitué de l'attribut binaire *volve* (c'est une membrane qui enveloppe certains champignons). On dispose des informations suivantes :

classe k	1 : vénéneux	2 : comestible
$P(k)$	0.05	0.95
$P(\text{volve}/k)$	0.9	0.2

- Je ramasse les champignons si le règle de Bayes les classifie dans la classe des comestibles. Donnez la fonction C_{Bayes} pour cet exemple. Est-ce-que je ramasse les champignons ayant une volve ? Appliqueriez-vous cette règle si vous alliez ramasser des champignons ?

- On définit un coût pour tout couple de classes (k, i) noté $\text{cout}(k, i)$. On définit alors le coût moyen de l'affectation à la classe k d'une description d de D par :

  $$\text{cout\_moyen}(k/d) = \sum_{i \in \{1, \dots c\}} \text{cout}(k, i) \times P(i/d).$$

  La règle de décision du coût minimum est : "Choisir $C_{\text{cout\_min}}$ qui à toute description d associe la classe k qui minimise $\text{cout\_moyen}(k/d)$".

  On définit sur notre exemple les coûts suivants :

  $$\text{cout}(1, 1) = \text{cout}(2, 2) = 0, \quad \text{cout}(1, 2) = 2, \quad \text{cout}(2, 1) = \infty.$$

  J'utilise la règle du coût minimal. Est-ce-que je ramasse les champignons ayant une volve ?

Exercice 3 On considère deux attributs pour déterminer la nationalité d'un individu. L'attribut taille qui peut prendre les valeurs grand ou petit, l'attribut couleur des cheveux qui peut prendre les valeurs brun ou blond. Les nationalités possibles sont français et allemand.

On suppose que les populations françaises et allemandes se répartissent selon le tableau suivant :

	petit,brun	petit,blond	grand,brun	grand,blond
français	25	25	25	25
allemand	10	20	30	40

- Est-ce que les attributs *taille* et *couleur* sont indépendants connaissant la classe (français ou allemand) ?
- Dans une assemblée comprenant 60% d'allemands et 40% de français, décrire
  1. la règle de décision majoritaire
  2. la règle du maximum de vraisemblance
  3. la règle de Bayes
- Calculer les probabilités d'erreur de chacunes des règles.
- On suppose maintenant que l'on ne connaît plus les proportions respectives des allemands et des français. On note p la proportion des allemands ($p \in [0, 1]$).
  - Décrire, selon les valeurs possibles de p, les règles de Bayes correspondantes.
  - Parmi les 5 règles que vous aurez detaillées à la question précédente, choisir celle qui **dans les pire des cas** possède la probabilité d'erreur minimale.