

Examen semestriel

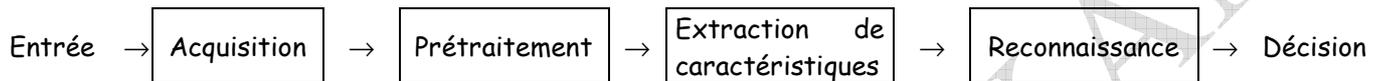
Module optionnel : Reconnaissance des
Formes

Corrigé

Exercice :

Question 1 : Donnez le schéma général d'un processus de reconnaissance.

Réponse :



(02 points)

Question 2 : Expliquez le rôle de l'étape « pré-traitement » d'un système de reconnaissance de formes.

Réponse :

L'objectif des pré-traitements est de faciliter la caractérisation de la forme (caractère, chiffre, mot) ou de l'entité à reconnaître soit en nettoyant la forme (élimination du bruit) ou en réduisant la quantité d'information à traiter pour ne garder que les informations les plus significatives.

(02 points)

Question 3 : Certains systèmes de reconnaissances de formes ont une étape de traitement supplémentaire appelée « post-traitement ». Quel peut être son rôle ?

Réponse :

L'étape de « post-traitement » vient après la sortie du processus de reconnaissance. Elle sert pour le test et la validation. C'est une étape importante car elle peut mettre en cause le choix des primitives ou le choix de la méthode d'apprentissage.

(02 points)

Question 4 : Expliquez la différence entre les méthodes d'apprentissage « supervisée » et « non supervisée ».

Réponse :

Dans le cas de l'apprentissage supervisé, un échantillon représentatif de l'ensemble des formes à reconnaître est fourni au module d'apprentissage. Chaque forme est étiquetée par un opérateur appelé professeur, cette étiquette permet d'indiquer au module d'apprentissage la classe dans laquelle le professeur souhaite que la forme soit rangée.

Dans le cas de l'apprentissage non supervisé, aucun exemple n'est fourni à l'apprenant. Il doit lui même deviner les différentes caractéristiques communes à chaque classe de formes.

(02 points)

Question 5 : Un algorithme de classification inspiré de la règle de Bayes est appelé « naïf ». Pourquoi ?

Réponse :

L'algorithme de classification de Bayes est qualifié de « naïf » parce qu'il suppose l'indépendance des attributs (descriptions), et par conséquent exprime la probabilité d'appartenance à une classe comme un produit de probabilités, mais en général ce n'est pas toujours le cas.

(02 points)

Question 6 : Qu'est ce qui est caché dans un HMM ?.

Réponse :

Ce qui est caché dans un HMM c'est la succession d'états suivie pour générer une séquence de symboles.

Exercice 2 :

On considère le modèle de Markov caché (HMM) suivant:

Les probabilités de transitions sont :

	Article	Adjectif	Nom	Verbe	End
Start	1				
Article		X	Y		
Adjectif			0.4		0.6
Nom	0.2	0.2		0.3	0.3
Verbe	1.0				

Les probabilités de génération sont :

	Le	La	Les	pecheur	Grand	Petit	gibier	chasse	Attire	trompe
Article	0.1	0.5	0.4							
Adjectif				0.2	0.4	0.4				
Nom				0.2			0.4	0.4		
Verbe								0.1	0.6	0.3

Soit la phrase P : «Le pecheur chasse le gibier ».

1/Trouver les différents chemins possibles permettant de générer la phrase P.

Réponse :

1er chemin : Start - Article - Adjectif - Nom - Article - Nom - End

2nd chemin : Start - Article - Nom - Verbe - Article - Nom - End

(01 points)

2/Calculer, en fonction de X, la probabilité de génération selon chacun des chemins en utilisant la méthode directe.

Réponse :

en constatant que $Y = 1 - X$, les probabilités recherchées sont :

$$P(\text{chemin1}) = x(1 - x) \cdot 0.768 \cdot 10^{-4}$$

$$P(\text{chemin2}) = (1 - x) \times 2 \cdot 0.072 \cdot 10^{-4}$$

(02 points)

Réponse :

En fonction de la valeur de X, il existe deux possibilités d'étiquetages :

Première possibilité :

Le	Pêcheur	Chasse	Le	Gibier	Si $X > 0.09375$
Article	Adjectif	Nom	Article	Nom	

Seconde possibilité :

Le	Pêcheur	Chasse	Le	Gibier	Si $X \leq 0.09375$
Article	Nom	Verbe	Article	Nom	

(03 points)

4/ Que faut-il faire si on veut que la probabilité de génération par le HMM de la phrase «Le pêcheur chasse le gibier » soit égale à 1 ? Justifiez.

Réponse :

Pour avoir une probabilité de 1 avec la phrase évoquée, il faut ré-entraîner le modèle en utilisant un algorithme d'entraînement avec la phrase en question.

(02 points)