

SERIE TD N°7

Questions de cours :

1. La méthode bayésienne naïve est associée à quel modèle du data mining ?
2. Donner quelques avantages/inconvénients de la méthode bayésienne naïve.

Exercice ①

Objectif de cet exercice est de classer des phrases selon leur thème : la radio ou la télévision. L'échantillon disponible est le suivant :

- Classe télévision :
 - Le programme TV n'est pas intéressant.
 - La TV m'ennuie.
 - Les enfants aiment la TV.
 - On reçoit la TV par onde radio.
- Classe radio :
 - Il est intéressant d'écouter la radio.
 - Sur les ondes, les programmes pour enfants sont rares.
 - Les enfants vont écouter la radio ; c'est rare.

Le vocabulaire retenu est le suivant : $V = \{TV, \text{programme}, \text{intéressant}, \text{enfants}, \text{radio}, \text{onde}, \text{écouter}, \text{rare}\}$

En utilisant un classifieur de Bayes, à quel thème serait associée la phrase :

« Il a vu la radio de ses poumons à la TV »

Exercice ②

Soit le tableau suivant :

Age	Income	Student	Credit_rating	Class:Buys_comp
<=30	HIGH	N	FAIR	N
<=30	HIGH	N	EXCELLENT	N
31....40	HIGH	N	FAIR	Y
>40	MEDIUM	N	FAIR	Y
>40	LOW	Y	FAIR	Y
>40	LOW	Y	EXCELLENT	N
31....40	LOW	Y	EXCELLENT	Y
<=30	MEDIUM	N	FAIR	N
<=30	LOW	Y	FAIR	Y
>40	MEDIUM	Y	FAIR	Y
<=30	MEDIUM	Y	EXCELLENT	Y
31....40	MEDIUM	N	EXCELLENT	Y
31....40	HIGH	Y	FAIR	Y
>40	MEDIUM	N	EXCELLENT	N

Utiliser le principe du classifieur bayésien naïf pour classer l'élément X tel que :

X = (<=30, MEDIUM, Y, FAIR, ???)

Rappels

- On note :

$x = [x^1, x^2, \dots, x^i, \dots, x^p]$ un vecteur de descripteurs

$C = \{c_1, \dots, c_q\}$ l'ensemble des classes possibles

S : ensemble fini de couples de la forme (x, c_k)

n : nombre d'observations de S

n_k : nombre d'éléments de S appartenant à la classe c_k

- La règle de classification de Bayes recommande de classer le vecteur x dans la classe c_k pour laquelle $P(c_k/x)$ est maximal.

Ce qui revient à maximiser : $\frac{p(x/c_k)p(c_k)}{p(x)}$

soit encore $p(x/c_k)p(c_k)$, $p(x)$ ne dépendant pas de c_k

Pour pouvoir appliquer la règle de Bayes il faut donc pouvoir estimer : $p(x/c_k)$ et $p(c_k)$

- On estime $p(c_k)$ par : $\hat{p}(c_k) = \frac{n_k}{n}$
- et $p(x/c_k)$ par : $\hat{p}(x/c_k) = \prod_{i=1}^p p(x^i/c_k)$

Ce qui revient à considérer les attributs comme indépendants les uns des autres

- La règle de classification de Bayes devient alors : classer le vecteur x dans la classe c_k qui maximise :

$$\prod_{i=1}^p p(x^i/c_k)p(c_k)$$

Principe de la méthode de classification bayésienne naïve
--

- Une évidence X est donnée
- On calcule $P(C | X)$ pour toutes les valeurs de C
- Si $P(C = c | X)$ est maximum, alors on choisit: $C=c$