
TRAVAUX DIRIGÉS N° 1 : Classifieur de Bayes

Stéphan CLÉMENÇON <stephan.clemencon@telecom-paristech.fr>
Joseph SALMON <joseph.salmon@telecom-paristech.fr>

EXERCICE 1. On considère un modèle de classification où le couple aléatoire (X, Y) est de loi P décrite par :

$$\mathcal{L}(X \mid Y = 0) = \mathcal{U}([0, \theta])$$
$$\mathcal{L}(X \mid Y = 1) = \mathcal{U}([0, 1])$$
$$p = \mathbb{P}\{Y = 1\}$$

où $p, \theta \in]0, 1[$ sont fixés. On note $\eta(x) = \mathbb{P}(Y = 1|X = x)$ la fonction de régression correspondante. Donner sa valeur en fonction de p et θ. Donner l'application numérique pour $\theta = 1/2$.

EXERCICE 2. On considère un modèle de classification où le couple aléatoire (X, Y) est de loi P décrite par :

— la loi de X est une loi de probabilité P_X sur $\mathbb{R}$

— la fonction de régression $\eta(x) = \dfrac{x}{x + \theta}$ où $\theta > 0$ fixé.

On note h^* le classifieur de Bayes. Expliciter le classifieur de Bayes dans ce modèle. Montrer ensuite que son risque "0-1" vaut

$$R(h^*) = \int \min(\eta(x), 1 - \eta(x)) dP_X(x).$$

Calculer le risque de Bayes lorsque $P_X = \mathcal{U}([0, \alpha\theta])$ où $\alpha > 1$.

EXERCICE 3. On considère un modèle de classification général. On considère le risque de classification pondéré :

$$L_\omega(g) = \mathbb{E}\big(2\omega(Y) \cdot \mathbb{1}_{\{Y \neq g(X)\}}\big)$$

où $\omega(0) \geq 0, \omega(1) \geq 0$ et $\omega(0) + \omega(1) = 1$. Donner le classifieur de Bayes et le risque de Bayes pour ce critère. Quel est l'intérêt de considérer un tel critère ?

EXERCICE 4. Soit $X = (T, U, V)^T$ où T, U, V sont des variables aléatoires réelles i.i.d. de loi exponentielle $\mathcal{E}(1)$. Soit $Y = \mathbb{1}_{\{T+U+V<\theta\}}$ où $\theta \in \mathbb{R}_+$ fixé.

1) Calculer le classifieur de Bayes $g^*(T, U)$, lorsque V n'est pas observée. Calculer le risque de Bayes associé à ce classifieur. Donner l'application numérique lorsque $\theta = 9$.
2) On suppose à présent que, seul T est observée. Reprendre les calculs précédents et comparer les risques bayésiens obtenus lorsque $\theta = 9$.
3) Proposer un classifieur lorsque X n'a aucune composante qui soit observée. Calculer son erreur de classification.

Conseils bibliographiques

Vous trouverez ci-dessous quelques points d'entrée utiles pour l'apprentissage automatique :

— Théorique et porté sur les aspects probabilistes : [DGL96]

— Utilitaire et porté sur les aspects pratiques : [HTF09]

— Livre récent porté essentiellement sur l'aspect optimisation : [SSBD14] (et du même auteur sur l'apprentissage en ligne [Sha11])

— Méthodes Bayésiennes et modèles graphiques : [Mur12]

Références

[DGL96] L. Devroye, L. Györfi, and G. Lugosi. *A probabilistic theory of pattern recognition*, volume 31 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1996. 2

[HTF09] T. J. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York, second edition, 2009. 2

[Mur12] K. P. Murphy. *Machine learning : a probabilistic perspective*. MIT press, 2012. 2

[Sha11] S. Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2) :107–194, 2011. 2

[SSBD14] S. Shalev-Shwartz and S. Ben-David. *Understanding machine learning : From theory to algorithms*. Cambridge University Press, 2014. 2