Classifieur de Bayes, Perceptron et Conditions de KKT

Exercice 1 Soit un couple aléatoire $(X, Y) \in \mathbb{R}^n \times \{-1, 1\}$.

1. Exprimer l'erreur de Bayes d'un classifieur quelconque $g : \mathbb{R}^n \mapsto \{-1, 1\}$.
2. Donner le classifieur optimal qui minimise l'erreur de Bayes.
3. Qu'en est-il si l'on pénalise les deux types d'erreurs différemment?
4. Soit un triplet (X, Y, Z) i.i.d. d'une loi exponentielle et $T = 1_{X+Y+Z<1}$. Quel est le classifier de Bayes de T sachant le couple (X, Y)?
5. Qu'en est-il si on détruit de l'information? (On observe $T(X)$ au lieu de X.)
6. Etablir le lien entre la classification et la régression. (Plug-In Rule)

Exercice 2 On se place dans le même cadre de classification binaire. Au lieu d'avoir accès à la loi du couple (X, Y) linéairement séparables, on ne possède qu'un nombre fini d'échantillons i.i.d.

1. Rappeler l'algorithme de perceptron. Observer l'impact de la marge sur le risque.
2. On construit un perceptron avec m observations où m suit une loi uniforme dans $\{1, \cdots, n\}$. Montrer que l'erreur de Bayes du classifieur résultant est bornée.
3. Formuler le problème de maximization de marge. Utiliser les conditions KKT pour déduire son problème dual (SVM).

Exercice 3 Soit un signal aléatoire composé d'un alphabet de n lettres. Supposons que les lettres du signal sont i.i.d. suivant une probabilité $(p)_{1 \leq k \leq n}$ satisfaisant $\min_k p_k > 0$.

1. (Inégalité de Kraft) Si l'on code toute lettre par une suite de 0 et 1 de longueur finie, quelle condition qu'il faudra imposer pour que ce signal soit décodable?
2. Formuler le problème de codage optimal.
3. Résoudre ce problème pour retrouver l'expression de l'entropie de Shannon.
4. Rappeler la définition de l'entropie conditionelle $H(X|Y)$. Montrer $H(X|Y) \leq H(X)$.