Exercice 1 Soit un espace probabilisé comportant un couple aléatoire $(X, Y) \in \mathbb{R}^n \times \{-1, 1\}$. Soit $g : \mathbb{R}^n \mapsto \{-1, 1\}$ un classifieur quelconque.

1. Exprimer l'erreur de Bayes de ce classifieur.
2. Donner le classifieur optimal en minimisant son erreur de Bayes.
3. Si l'on pénalise les deux types d'erreurs différemment, donner le classifieur optimal.
4. Soit un triplet (X, Y, Z) dont les coordonnées sont i.i.d. d'une loi exponentielle et $T = 1_{X+Y+Z<1}$. Quel est le classifier de Bayes si on n'observe que le couple (X, Y)?
5. Generaliser: que se passe-t-il si on détruit de l'information? ($T(X)$ au lieu de X)
6. Observer le lien entre la classification et la régression. Plug-In Rule. Comment interpréter Perceptron?

Exercice 2 Dans le même cadre du problème précédent, au lieu d'avoir accès directement à la distribution, on ne possède qu'un nombre fini d'échantillons i.i.d. Supposons de plus qu'ils sont linéairement séparables.

1. Si l'on choisit de manière uniforme un entier m entre 1 and N pour construire un classifier à l'aide de l'algorithme de perceptron avec m observations, montrer que l'erreur de Bayes de ce perceptron randomisé est bornée. Observe l'impact de la marge sur le risque.
2. Formuler le problème de maximization de marge, les conditions KKT and les vecteurs supports.

Exercice 3 L'entropie de Shannon. Soit un signal aléatoire composé d'un alphabet de n lettres. Supposons que chaque lettre du signal est générée de manière indépendante suivant une probabilité $(p)_{1\leq k\leq n}$ satisfaisant $\min_k p_k > 0$.

1. Si l'on code la lettre k 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 d'optimisation pour retrouver l'expression de l'entropie de Shannon.
4. Rappel de la définition de l'entropie conditionelle $H(X|Y)$. Montrer $H(X|Y) \leq H(X)$ et $H(X, Y) \leq H(X) + H(Y)$.