Mini-Quiz IFT6266 A00

Nom de l'étudiant: ______________________________

1. Décrivez l'objectif et le cadre formel à l'intérieur duquel se situent la plupart des algorithmes d'apprentissage.

2. Quelle est la différence fondamentale entre la problématique de *l'optimisation* et celle des *algorithmes d'apprentissage*?

3. Proposez un algorithme d'apprentissage pour la classification qui peut toujours obtenir zéro erreur sur les exemples d'apprentissage. Écrivez-le en *pseudo-code*. Pouvez-vous dire quelque chose sur ses chances de généraliser?

4. Considérez le problème de l'estimation de densité $f(z) \in \mathcal{R}$ (les données sont des points i.i.d. z_i, et la perte est $L(f, z) = -\log f(z)$ sous contrainte que $f(z) \geq 0$ et $\int f(z)dz = 1$). Quelle est la fonction (au sens large) qui minimise la perte empirique parmi TOUTES les fonctions possibles?

5. Considérons un classifieur Gaussien pour deux classes $C = 1$ et $C = 2$ et une entrée X dans $\mathcal{R}^n$. On a donc $P(X|C = 1)$ et $P(X|C = 2)$ deux distributions Gaussiennes de moyennes $\mu_1 \in \mathcal{R}^n$ et $\mu_2 \in \mathcal{R}^n$, et avec des matrices de covariance $S_1 \in \mathcal{R}^{n\times n}$ et $S_2 \in \mathcal{R}^{n\times n}$ respectivement. La surface de décision est donnée par l'équation

$$P(C = 1|X = x) = P(C = 2|X = x).$$

Utilisez le théorème de Bayes et remplacez les densités Gaussiennes par leurs formule ($P(X = x|C = 1) = \frac{\exp((x-\mu_1)' S_1^{-1}(x-\mu_1))}{(2\pi)^{n/2}|S_1|^{n/2}}$, où $|A|$ est le déterminant de A) pour montrer que quand $S_1 = S_2$, cette surface de décision est linéaire.