

Perceptron, descente de gradient, évaluation.

Exercice 1 *Perceptron (Examen 2023 - S2)*

Soit le perceptron dont le vecteur de poids est $(w_0, w_1, w_2) = (2, 1, 1)$ et qui calcule le score :

$$y = \text{signe}(w_0 + w_1 x_1 + w_2 x_2).$$

Question 1. Dessiner dans \mathbb{R}^2 la frontière de décision de ce perceptron.

Question 2. Parmi les 4 modèles suivants, quel perceptron donne la même décision que le perceptron précédent ?

	$(w_0, w_1, w_2)^\top$
(I)	$(1, 0.5, 0.5)^\top$
(II)	$(200, 100, 100)^\top$
(III)	$(\sqrt{2}, \sqrt{1}, \sqrt{1})^\top$
(IV)	$(-2, -1, -1)^\top$

Question 3. Soit les données ci-dessous où \mathcal{P} correspond aux exemples positifs et \mathcal{N} aux exemples négatifs :

$$(1, 1)^\top \in \mathcal{P}, \quad (1, 0)^\top \in \mathcal{N}, \quad (0, 0)^\top \in \mathcal{P}, \quad (0, 1)^\top \in \mathcal{N}$$

Soit le vecteur initial $w = (1, 0, 0)$. En prenant les exemples dans l'ordre, de manière cyclique, appliquer l'algorithme du perceptron (en considérant un pas d'apprentissage = 1). Comment peut-on voir qu'il ne sera pas possible d'apprendre un perceptron dans ce cas ?

Exercice 2 *Perceptron (Examen 2024 - S1)*

On considère un perceptron simple avec deux entrées notées x_1 et x_2 et une sortie y telle que :

$$y = \begin{cases} 1 & \text{si } w_1 x_1 + w_2 x_2 - w_0 > 0 \\ 0 & \text{sinon} \end{cases}$$

On considère que 1 représente la valeur vrai et 0 la valeur faux.

Question 1. Trouver des poids pour que le perceptron calcule la fonction ET logique.

Question 2. Même question avec la fonction OU logique.

Question 3. Essayer de trouver des poids pour la fonction XOR et commenter.

Exercice 3 *Descente de gradient, fonctions de coût et évaluation (Examen 2024 - S1)*

Avec les notations du cours, une donnée décrite vectoriellement est notée $\mathbf{x} \in \mathbb{R}^d$ et associée à une étiquette y . Une base de données entière est notée : $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)\}$. On se limite dans cet exercice à l'étude des modèles linéaires : l'objectif est donc d'apprendre un vecteur \mathbf{w} .

Question 1. Donner la dimension du vecteur \mathbf{w} .

Question 2. L'idée générale d'une fonction de coût est de mesurer les erreurs d'un système d'apprentissage automatique. Par exemple, dans le cas de classification, on pourrait simplement compter le nombre d'erreurs de classification. Cependant, cette dernière proposition n'est pas utilisée en pratique : pourquoi ? Citer deux fonctions de coût classiques (noms des fonctions et expressions mathématiques). Expliquer les points forts et faibles de ces fonctions de coût.

Question 3. Rappeler le but et le principe général de l'algorithme de descente de gradient et décliner cet algorithme sur l'une des fonctions de coût précédentes. Expliquer le principe d'une descente de gradient stochastique ET le principe de la descente de gradient classique. Expliquer clairement ce qui est fourni en entrée et ce qui est obtenu en sortie de l'algorithme.

Question 4. Comment évaluer le modèle appris ? Expliquer la procédure et rappeler brièvement le principal piège à éviter.

Question 5. Discuter la complexité algorithmique et les performances des algorithmes de descente de gradient par rapport à l'algorithme des k -plus proches voisins.

Exercice 4 *Apprentissage du perceptron (examen 2024 - S2)*

N.B. Les notations sont les mêmes que dans l'exercice précédent.

On dispose d'une base de N exemples (observations), $\{\mathbf{x}_i\}_{i=1,\dots,N}$, dont les classes sont connues ; la classe de \mathbf{x}_i est notée $d(\mathbf{x}_i)$. On utilise l'algorithme donné plus loin (une des variantes de l'algorithme du perceptron) pour apprendre automatiquement la valeur des paramètres, c'est à dire du vecteur w : (ϵ est un nombre positif ; \mathbf{x}^T est le transposé de \mathbf{x})

Le critère d'arrêt peut être, par exemple, qu'il n'y a pas eu d'erreur de classification pendant un certain nombre d'itérations successives.

Question 1. Que signifie la condition $d(\mathbf{x}_i)\mathbf{x}_i w(t) \geq 0$? Expliquer le principe de l'algorithme.

Question 2. Faites tourner l'algorithme du perceptron sur le problème du OU logique en itérant sur la base d'apprentissage constituée des 4 exemples distincts possibles, avec successivement pour valeur initiale $w(0)$:

1. $w(0) = (0; 0; 0)$;
2. $w(0) = (1; 1; 1)$;
3. $w(0) = (1; -1; 1)$.

(Prendre $\epsilon = 1$) Représenter graphiquement l'évolution de la frontière de décision d'itération en itération.