

TD PERCEPTRON

EXERCICE 1. ALGORITHME DE PERCEPTRON

Soit la base d'apprentissage de taille 4 suivante :

$$S = \left\{ \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, +1 \right); \left(\begin{pmatrix} -1 \\ 1 \end{pmatrix}, +1 \right); \left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}, -1 \right); \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, -1 \right) \right\}$$

1.1 Dessiner les points dans un repère orthonormé du plan.

1.2 On considère le modèle de perceptron pour séparer les points de classe +1 à ceux de la classe -1. On suppose que le vecteur des poids initial est le vecteur null, que le pas d'apprentissage est fixé à 1, et que le biais $w_0 = 0$. Quel est dans ce cas, le vecteur poids trouvé par l'algorithme du perceptron après 4 mises à jour si l'on considère que les mises à jour se font dans le sens trigonométrique en commençant par le point de coordonnées $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, c.à.d. $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ puis $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ puis $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ puis ... ?

1.3 Quelle est l'équation de la droite séparatrice trouvée après ces deux mises à jour?

1.4 À quelle distance se trouve les points de la base d'apprentissage à cette droite?

EXERCICE 2. THÉORÈME DE NOVIKOFF

Nous supposons qu'il existe un vecteur \mathbf{w}^* qui discrimine parfaitement tous les exemples de la base d'apprentissage $S = \{(\mathbf{x}_j, y_j), j \in \{1, \dots, m\}\}$, et nous définissons

$$\rho = \min_{j \in \{1, \dots, m\}} \left(y_j \left\langle \frac{\mathbf{w}^*}{\|\mathbf{w}^*\|}, \mathbf{x}_j \right\rangle \right).$$

2.1 À quoi correspond ρ ? Expliquer pourquoi elle serait une valeur strictement positive.

Nous supposons de plus que tous les exemples d'apprentissage sont contenus dans un hypersphère de rayon R (i.e. $\forall j, \|\mathbf{x}_j\| \leq R$). Sans restreindre la généralité, nous supposons de plus que le vecteur poids initial vaut $\mathbf{w}^{(0)} = 0$, et nous initialisons le pas d'apprentissage ϵ à 1.

2.2 Montrer qu'après l mises à jour, la norme du vecteur poids vérifie

$$\|\mathbf{w}^{(l)}\|^2 \leq l \times R^2 \quad (1)$$

2.3 Avec les mêmes conditions que précédemment, montrer qu'après l mises du vecteur poids nous avons

$$\left\langle \frac{\mathbf{w}^*}{\|\mathbf{w}^*\|}, \mathbf{w}^{(l)} \right\rangle \geq l \times \rho \quad (2)$$

2.4 Dédurre des questions précédentes que le nombre de mises à jour, l , de l'algorithme de Perceptron est borné par

$$l \leq \left\lceil \left(\frac{R}{\rho} \right)^2 \right\rceil$$

Où $\lfloor x \rfloor$ est la partie entière du réel x .

2.5 Ce résultat est-il vérifié pour le problème de l'exercice 1?