TP théorique réseau de neurones

Adrien CHAN-HON-TONG

Notation, le produit scalaire usuel est symbolisé par un point. On utilisera les 3 notations suivantes pour la fonction relu : $relu(x) = \max(0, x) = [x]_+$

Ce TP porte sur la classification binaire de point 2D : $x \in \mathbb{R}^2$ et $y(x) \in \{-1, 1\}$.

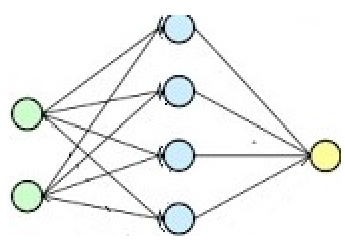
1 neurone

Q1.1 : Montrer qu'il est possible d'apprendre par coeur la base de données $((1,1),1)$, $((-1,-1),-1)$ avec 1 neurones sans biais et sans activation (puisqu'il est le dernier neurone). *Chercher des poids triviaux - inutile d'inverser des matrices.*

Q1.2 : Est-il possible d'apprendre par coeur la base de données $((0,1),1)$, $((0,-1),1)$, $((1,0),-1)$, $((-1,0),-1)$ avec 1 neurones sans biais ?

2 couches de neurones

Q2.1 : Montrer qu'il possible d'apprendre par coeur la base de données $((0,1),1)$, $((0,-1),1)$, $((1,0),-1)$, $((-1,0),-1)$ avec le réseau ci dessous (sans biais et avec activation relu) ? *Chercher des poids triviaux - inutile d'inverser des matrices.*



Le réseau prends en entrée un point 2D donc 2 valeurs (vert), a 1 couche cachée (bleu) de 4 neurones (avec relu) puis 1 dernière couche (jaune) avec 1 neurone (sans relu donc).

Q2.2 : Si oui, dessinez les zones classées comme 1 ($f(x) > 0$, f symbolisant le réseau de neurones) et celles classées comme -1 ($f(x) < 0$). *Chercher les frontières $f(x) = 0$, elle est linéaire par morçeau avec un réseau basé relu.*

Q2.3 : Est-il possible d'apprendre avec le même réseau (mais d'autres poids) la base $((0,2),1)$, $((0,-2),1)$, $((2,0),1)$, $((-2,0),1)$, $((0,0),-1)$?

2 couches de neurones avec biais

Q3 : Considérons encore même la base de données $((0,2),1)$, $((0,-2),1)$, $((2,0),1)$, $((-2,0),1)$, $((0,0),-1)$, ainsi que les 2 réseaux

— $f(x) = [(0,1).x]_+ + [(0,-1).x]_+ + [(1,0).x]_+ + [(1,0).x]_+ - 1$

— $f(x) = 2relu((-1,1).x - 1) + 2relu((-1,1).x - 1) - 1$

Dessinez les zones classées comme 1 ($f(x) > 0$, f symbolisant le réseau de neurones) et celles classées comme -1 ($f(x) < 0$).

Pour votre culture : *cette base est intéressant car il est possible de l'apprendre asymétriquement avec un réseau 2 neurones (2 relu) puis 1 neurone. Mais pour obtenir une solution symétrique, il faut 4 neurones (4 relu) puis 1 neurone. Ainsi, dans cet exemple (et cet exemple seulement), un peu plus de paramètres est nécessaire pour obtenir une solution plus régulière.*

Pourtant une ligne directrice très importante de l'apprentissage par ordinateur est que moins de paramètre conduit à plus de généralité.