

### compte rendu du TP 3 16241081 Paul

Je trouve que la frontière est tangente aux droites qui présentent les neurones cachés, et je trouve que les performances des discriminateurs diffèrent. Quand le nombre de neurones est grand, la performance de RN augmente, et quand on choisit le base d'apprentissage grande, je trouve que la performance augmente aussi. (Je sais pas comment on peut écrire la fonction de J et comment on peut choisir la fonction de produit scalaire, la fonction sigmoïde ou les autres fonctions). (Et quand  $N_c$  est trop grand la performance diminue, pourquoi ?) Et je trouve que si le nombre de passage (époque) augmente alors le taux d'apprentissage et le taux de généralisation et le loss diminuent.

En changeant le nombre de classes (c'est le nombre de chiffres à distinguer ?), je trouve que si le nombre de classes est petit, je trouve que la performance est mieux que la performance dont le nombre de classes est grand. Je pense si le problème est difficile à résoudre il faut que plus de neurones pour terminer le modèle mais les paramètres augmentent aussi donc c'est difficile d'avoir la bonne performance ?

Je trouve que si  $\eta$  est petit les paramètres  $\omega$  vont converger simplement au minimum local donc la performance n'est pas bonne. Est-ce que c'est correct ou pas ? Et pourquoi quand  $\eta$  est grand on ne peut pas avoir une bonne performance et le taux de généralisation est environ 50% ?

Base d'apprentissage et base de généralisation basées toutes sur (51). Je ne comprends pas pourquoi si l'on prend un nombre de chiffres plus grand que 2, le taux  $\tau_g$  du réseau est toujours environ 50%.