

## Compte rendu TP3 Emma&Céline

1/

La surface discriminante est tangentielle à des droites.

Les performances : ddp connue > RN > quadratique

le nombre de neurones est plus grand, le critère  $J_T$  est plus petite; la performance est mieux.

La taille de base d'apprentissage est plus grande, la performance est mieux.

2/

Le nombre de chiffres est plus petit, la performance est mieux.

3/

Le nombre de chiffres est plus grand, la performance de RN est mieux. La raison est le nombre de chiffres qu'il faut traiter est plus grand, il est plus difficile de distinguer. Mais FA est plus mieux que RN.

4/

Dans le cas où les bases de généralisation sont générées suivant le modèle de l'équation (69).

Il n'y a pas de différence significative.

Quand la base de généralisation est générée selon le modèle de l'équation (51) avec

$A \leftarrow U[-5, 5]$ , il y a une différence significative.

Q : on ne sais pas la raison.

5/

Quand l'apprentissage est effectué avec une base qui vérifie le modèle de l'équation (69), la performance du réseau de neurones est faible.

Quand la base de généralisation est générée selon le modèle de l'équation (51) avec

$A \leftarrow U[-5, 5]$ , la performance du RN est mieux.

Quand le nombre de chiffres est plus haut, la performance est plus faible.

**Synthèse** : Pour le réseau de neurones, la taille d'apprentissage est plus grande, le nombre de neurones est plus grand, et avec le cas de l'amplitude de  $\mu_c$  inconnue, la performance est mieux. Pour distinguer  $\tau_g$  plus clairement, on peut augmenter  $P_{gen}$ .

Q : Pourquoi la performance est mieux avec la méthode de l'amplitude de  $\mu_c$  inconnue?

Q : Comment le nombre de passe,  $\eta$ , et le pas du gradient influence sur la performance?

Comment on peut expérimenter sur Matlab?