

Pour  $N_C$ , il existe un  $N_C$  suffisant pour un projet d'apprentissage. Si  $N_C$  est trop petit, le  $J_T$  diminue très lent, et le  $\tau_g$  et  $\tau_{app}$  sont petits. Dans cet exemple, le  $N_C=6$  est suffisant, pour  $N_C$  plus grand, le temps de traitement est très long et  $\tau_g$  et  $\tau_{app}$  change pas évidemment. Pour  $\eta$ , si  $\eta$  est très petit, sur le graphe de  $J_T$  il y a trop de bruit, et le  $\tau_g$  et  $\tau_{app}$  sont dans les bas niveaux. Si  $\eta$  très grand, le  $J_T$  diminue rapidement mais  $\tau_g$  et  $\tau_{app}$  sont encore dans les bas niveaux, donc on doit choisir un  $\eta$  adaptable pour l'exemple. Pour le nombre de passés, pour un exemple spécifique, il existe un nombre de passés adaptable pour différents  $P_{app}$ . En fait on doit maintenir la balance entre  $P_{app}$  et nombre de passés. C'est pourquoi on utilise nombre de passés\* $P_{app}$  dans la figure 6. Si  $P_{app}$  assez grand, on a besoin moins du nombre de passés, si le nombre de passés est très grand, on perd trop de temps. Je trouve que pour certain de  $N_C$ , il existe plusieurs étapes sur la figure 7, comme un escalier, et je ne sais pas pourquoi.

Sur les 2 et 3, on trouve que FA gaussien, le  $\tau_g$  diminue avec plus de nombre de chiffres mais les changements sont petits. Mais pour le RN,  $\tau_g$  est très petit pour un grand nombre de chiffres. Donc on pense RN n'est pas adaptable pour les grands nombres.

Après on utilise l'amplitude connue, on n'a trouvé pas la différence entre FA et RN,  $\tau_g$  sont très grands, mais s'on utilise l'amplitude inconnue,  $\tau_g$  de FA diminue évidemment, mais  $\tau_g$  de RN ne change pas. Donc on pense la connaissance de  $\mu_c$  n'influence pas la performance de RN. Donc le RN est plus fort pour les données inconnues.

Partie 5, on apprend, si l'ensemble de app et l'ensemble de gen sont les mêmes types, la performance de RN est très bien, mais si l'ensemble de app et gen sont différents,  $\tau_g$  de RN est petit et la performance est mauvais. Ça veut dire que les apprentissages et les généralisations doivent être similaires. On trouve aussi que pour le grand nombre de chiffres, la performance de RN n'est pas bien, qui correspond au résultat de partie 3, mais je ne sais pas pourquoi.

Samuel

(travaillé avec Etienne)