

# IFT 6266: Algorithmes d'apprentissage

**Yoshua Bengio**, `Yoshua.Bengio@umontreal.ca`

**James Bergstra**, `bergstrj@iro.umontreal.ca`

Quiz Surprise #1 (1 point bonus dans le bulletin)

15 septembre 2008

Nom de l'étudiant: \_\_\_\_\_

---

Complétez les trous dans les phrases ci-bas, en utilisant les termes suivants: exemples, optimiste, pessimiste, algorithme d'apprentissage, prédiction, ensemble d'apprentissage, erreur d'apprentissage, hyper-paramètre, régularisation, critère d'apprentissage, cible, norme des paramètres, augmente, diminue, erreur de généralisation, l'erreur quadratique, ensemble de test, paramètres, erreur de test.

1. On fournit des exemples regroupés dans un \_\_\_\_\_ à un \_\_\_\_\_ et on obtient une fonction apprise. Cette fonction résultante peut être appliquée sur de nouveaux \_\_\_\_\_. Les exemples qui ont servi à faire l'entraînement constituent \_\_\_\_\_. L'erreur mesurée sur ces exemples est appelée \_\_\_\_\_. Les nouveaux exemples qui servent à estimer la généralisation de la fonction apprise constituent \_\_\_\_\_. L'erreur mesurée sur ces exemples est appelée \_\_\_\_\_.
2. Certains algorithmes d'apprentissage fonctionnent par la minimisation d'un \_\_\_\_\_, qui est une fonction des données d'entraînement et de \_\_\_\_\_ que l'on veut ajuster. Pour la régression linéaire on s'intéresse en particulier à minimiser \_\_\_\_\_, qui est le carré de la différence entre \_\_\_\_\_ (sortie de la fonction apprise) et une valeur \_\_\_\_\_ observée.
3. L'erreur d'apprentissage est généralement un reflet \_\_\_\_\_ de la véritable \_\_\_\_\_, qui est l'espérance de l'erreur mesurée sur de nouveaux tirages d'exemples. On ne peut donc pas utiliser l'\_\_\_\_\_ (seule) pour évaluer la qualité d'une fonction apprise: on va estimer l'\_\_\_\_\_ en mesurant la moyenne des erreurs sur un \_\_\_\_\_. Cette \_\_\_\_\_ est plus ou moins fiable selon la taille de l'ensemble de test.
4. Les \_\_\_\_\_ sont ajustés en fonction des exemples de l'\_\_\_\_\_. Mais certains algorithmes, comme la régression linéaire avec régularisation L2, ont un \_\_\_\_\_ qui peut être choisi pour contrôler le degré de complexité de la solution. On force ainsi la régression linéaire à donner une solution plus simple en ajoutant un terme de régularisation, qui pénalise la \_\_\_\_\_. Le coefficient devant ce terme de régularisation est appelé \_\_\_\_\_, et c'est lui qui nous permet d'ajuster la complexité du modèle. Plus on pénalise (coefficient élevé) et plus l'erreur d'apprentissage \_\_\_\_\_. Plus on pénalise et plus la différence entre l'erreur d'apprentissage et l'erreur de généralisation \_\_\_\_\_.