

Statistiques pour l'ingénieur

Validation croisée

Étienne Batise - Thibaud Dauce

19 mai 2014

1 Qu'est-ce que la validation croisée ?

- Le principe
- L'utilité
- 3 types de validation croisée

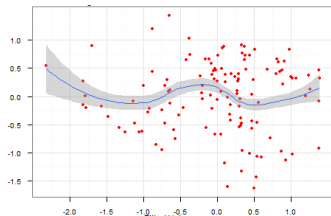
2 Démonstration Octave

3 Et maintenant, dans l'informatique

- Exemple des réseaux de neurones 1/2
- Exemple des réseaux de neurones 2/2

4 Questions

Le principe



- après avoir trouvé la fonction
- tester sa validité
- deux groupes de données :
 - groupe d'apprentissage
 - groupe de test



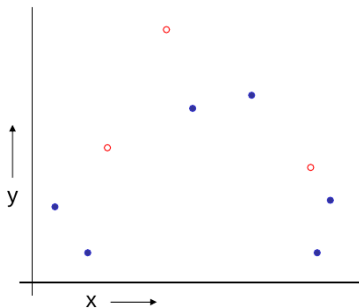
- tester les données...
- ...avec peu de points
- ...sans faire trop de calculs

3 types de validation croisée



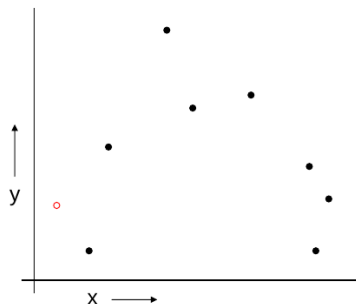
- Rapport données / temps de calcul
 - testset validation
 - leave-one-out cross-validation
 - k-fold cross-validation

testset validation



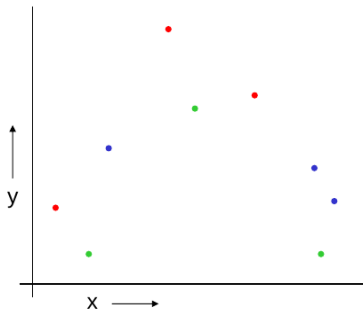
- choisir 30% des données pour le test
- Résultats :
 - très facile à mettre en place
 - perte de 30% de l'échantillon
 - peu précis en cas de petit échantillon

leave-one-out cross-validation



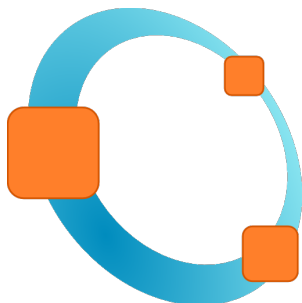
- choisir 1 donnée pour le test
- recommencer pour chaque donnée
- faire la moyenne des erreurs
- Résultats :
 - aucune perte de données
 - très couteux en temps

k-fold cross-validation



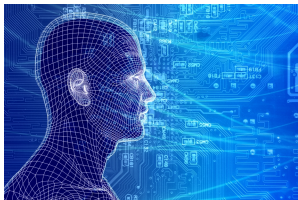
- découper les données en k parties
- Résultats :
 - perte de données relative à k
 - complexité relative à k
 - si $k = n$, équivalent à un leave-one-out

Présentation du code



- génération de points aléatoires
- calcul de la fonction par 2 méthodes :
 - régression linéaire
 - régression quadratique
- calcul de la validation croisée

Exemple des réseaux de neurones 1/2



- modèle de calcul basé sur les neurones biologiques
- une ou plusieurs entrées, une sortie
- fonctionne en couche
- permet de résoudre des problèmes statistiques
- basé sur l'apprentissage

Exemple des réseaux de neurones 2/2



- tester son réseau de neurones
- ajuster l'apprentissage
 - nombre de noeuds
 - paramètres d'apprentissage
- pour en savoir plus : cours "Machine Learning" sur Coursera

Questions

On vous écoute :)

- <http://www.autonlab.org/tutorials/overfit10.pdf>
- https://fr.wikipedia.org/wiki/Réseaux_de_neurones
- <https://jamesmccaffrey.wordpress.com/2013/10/25/k-fold-cross-validation-for-neural-networks>
- <http://visualstudiomagazine.com/articles/2013/10/01/understanding-and-using-kfold.aspx>
- https://fr.wikipedia.org/wiki/Validation_croisée
- <https://www.coursera.org/course/ml>