

17 juin 2020

Question 1. Vous disposez d'un jeu de données contenant un millier d'observations. Les performances des modèles linéaires que vous avez essayés ne sont pas satisfaisantes. Vous décidez d'utiliser un réseau de neurones artificiels. Vaut-il mieux essayer

- ☐ un perceptron;
- ☐ un perceptron multi-couche avec une couche intermédiaire d'une dizaine de neurones;
- ☐ un perceptron multi-couche avec 4 couches intermédiaires d'une centaine de neurones chacune?

Question 2. La qualité du modèle appris par un réseau de neurones artificiel dépend

- ☐ de l'architecture de ce réseau;
- ☐ des fonctions d'activations;
- ☐ de la vitesse d'apprentissage (c'est-à-dire le pas de la descente de gradient);
- ☐ de la quantité de données utilisées.

Question 3. L'astuce du noyau s'applique

- ☐ à la régression ridge;
- ☐ à la régression lasso;
- ☐ aux arbres de décision.

Question 4. Considérons un arbre de décision appris sur un jeu de données contenant n observations décrites par p variables. Sa profondeur est au plus

- ☐ p .
- ☐ $\log_2(p)$.
- ☐ n .
- ☐ $\log_2(n)$.
- ☐ $\min(n, p)$.
- ☐ $\min(\log_2(n), \log_2(p))$.

Question 5. Le bagging permet de

- ☐ Combiner plusieurs jeux de données en un seul pour créer un meilleur modèle.
- ☐ Créer plusieurs jeux de données à partir d'un seul, en sélectionnant aléatoirement les observations.
- ☐ Créer plusieurs jeux de données à partir d'un seul, en sélectionnant aléatoirement les variables.
- ☐ Combiner plusieurs modèles simples en un meilleur modèle.

Solution

Question 1. Le perceptron est un modèle linéaire, il n'aura pas une meilleure performance. Le perceptron multi-couche avec 4 couches intermédiaires d'une centaine de neurones chacune contient beaucoup de paramètres pour 1 000 observations seulement et risque de surapprendre.

Question 2. Toutes les réponses sont valides.

Question 3. Ni le lasso (à cause de la norme ℓ_1) ni les arbres de décision (non paramétriques) ne peuvent s'écrire en faisant apparaître les observations uniquement au sein de produits scalaires entre observations. Seule la régression ridge est donc *kernelizable*.

Question 4. Dans le pire des cas, chaque niveau de l'arbre de décision met un seul échantillon dans la branche de gauche, et tous les autres dans la branche de droite. On obtient donc une profondeur de n .

Question 5. Le bagging consiste à créer plusieurs jeux de données à partir d'un seul, en sélectionnant aléatoirement les observations, afin de créer autant de modèles simples combinés en un modèle robuste.