Ensemble Methods in a Nutshell

Junio de 2018

Contenido

Introduction

- These are general methodologies for combining a collection of simpler models.
- We give a very brief overview of:
  1. Bagging and Sub-bagging.
  2. Boosting.
  3. Stacking.

Contenido

In a nutshell

- Bagging is a general methodology for avergaging models and reducing variance.
- Bagging is a bootstrap of the prediction.
- Sub-bagging is a special case, it is also a boostrap of the prediction and balances clases when dispropotiantly unbalanced.

In a nutshell

- Bagging is a general methodology for avergaging models and reducing variance.
- Bagging is a bootstrap of the prediction.
- Sub-bagging is a special case, it is also a boostrap of the prediction and balances clases when dispropotiantly unbalanced.

In a nutshell

- Bagging is a general methodology for avergaging models and reducing variance.
- Bagging is a bootstrap of the prediction.
- Sub-bagging is a special case, it is also a boostrap of the prediction and balances clases when dispropotiantly unbalanced.

Contenido

Setup

- Viene de Adaptive Boosting.
- Supongamos que tenemos una muestra $\tau_n = \{(x_1, y_1), ..., (x_m, y_m)\}$ donde $y_i \in \{-1, 1\}$.
- Defina unos pesos iniciales para cada observación i: $D_1(i) = \frac{1}{m}$. D_t siempre será una distribución sobre las m observaciones.

AdaBoost

- Para cada $t = 1, ..., T$
- Construir un clasificador (puede ser débil) h_t que minimice la función de pérdida:
  1. Defina el error e_t como:

$$e_t = \sum_{i=1}^{m} D_t(i) I(y_i \neq h_t(x_i)) \tag{1}$$

  2. Calcular $\alpha_t = \frac{1}{2} \log(\frac{1-e_t}{e_t})$
  3. Modificar los pesos:

$$D_{t+1}(i) \rightarrow \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

  donde $Z_t = \sum_{i=1}^{m} D_t(i)$
- $H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x))$

Contenido

Optimal combination of models

- Take the prediction of many models as features in a regression problem.
- A simple example is to use regularization techniques to make combined model (e.g., Ridge, Lasso, etc)