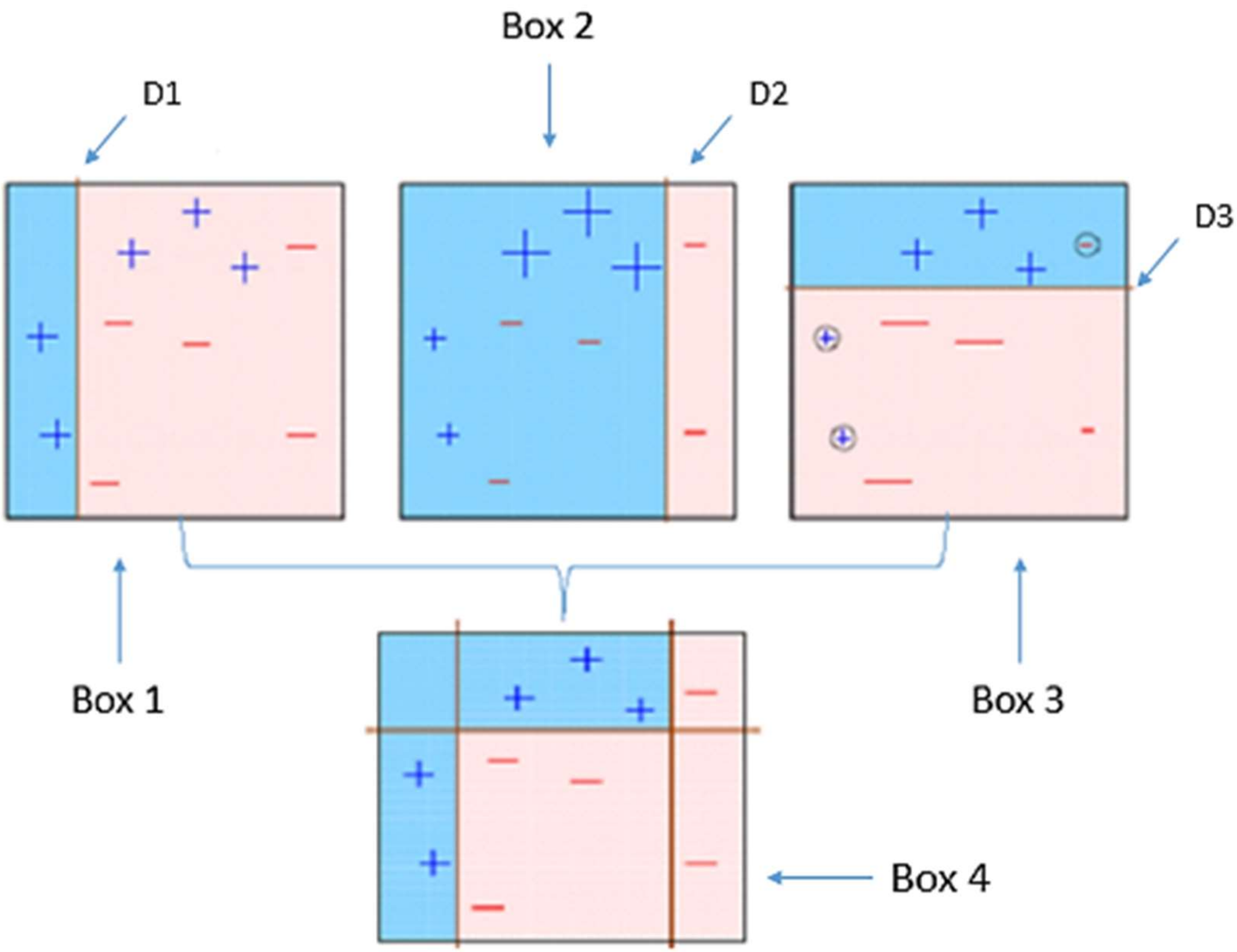
AdaBoost Regressor

A Journey through Pseudocode

SmartScience


Box 2
D1
D2
D3
Box 1
Box 3
Box 4

Ensembling

Ensembling is **a process where multiple diverse models are created, either by using many different modeling algorithms or using different training data sets**.

The ensemble model then aggregates the prediction of base models and results in once final prediction for the unseen data.

Ensembling

Bagging

- considers **homogeneous weak learners.**
- learns independently in **parallel.**
- **Combines them for overall prediction.**

Boosting

- considers **homogeneous weak learners.**
- learns **sequentially and adaptive manner.**.
- **Combines them for overall prediction.**

Stacking

- considers **heterogeneous weak learners.**
- Learns independently in **parallel.**
- **Combines them for overall prediction.**

AdaBoost Pseudocode walkthrough

1. Initialize the weights with $w_n^1 = \frac{1}{N}$ for $n = 1, 2, \dots, N$.
2. For $t = 1, 2, \dots, T$ or while $\bar{L}^t$, as defined below, is less than or equal to 0.5,
   - Draw a sample of size N from the training data with replacement and with probability w_n^t for $n = 1, 2, \dots, N$.
   - Fit weak learner t to the resampled data and calculate the fitted values on the original dataset. Denote these fitted values with $f^t(\mathbf{x}_n)$ for $n = 1, 2, \dots, N$.
   - Calculate the observation error L_n^t for $n = 1, 2, \dots, N$:

$$D^t = \max_n \{|y_n - f^t(\mathbf{x}_n)|\}$$

$$L_n^t = \frac{|y_n - f^t(\mathbf{x}_n)|}{D^t}$$

   - Calculate the model error $\bar{L}^t$:

$$\bar{L}^t = \sum_{n=1}^{N} L_n^t w_n^t$$

   If $\bar{L}^t \geq 0.5$, end iteration and set T equal to $t - 1$.
   - Let $\beta^t = \frac{\bar{L}^t}{1-\bar{L}^t}$. The lower β^t, the greater our confidence in the model.
   - Let $Z^t = \sum_{n=1}^{N} w_n^t (\beta^t)^{1-L_n}$ and update the model weights with

$$w_n^{t+1} = \frac{w_n^t (\beta^t)^{1-L_n}}{Z^t},$$

   which increases the weight for observations with a greater error L_n^t.
3. Set the overall fitted value for observation n equal to the weighted median of $f^t(\mathbf{x}_n)$ for $t = 1, 2, \dots, T$ using weights $\log(1/\beta^t)$ for model t.

https://dafriedman97.github.io/mlbook/content/c6/s2/boosting.html

1. Initialize the weights with $w_n^1 = \frac{1}{N}$ for $n = 1, 2, \ldots, N$.

Here;

N = number of observations

w_n^1 – initialize weight of each observation with 1/N

2. For $t = 1, 2, \ldots, T$ or while $\bar{L}^t$, as defined below, is less than or equal to 0.5,
   - Draw a sample of size N from the training data with replacement and with probability w_n^t for $n = 1, 2, \ldots, N$.
   - Fit weak learner t to the resampled data and calculate the fitted values on the original dataset. Denote these fitted values with $f^t(\mathbf{x}_n)$ for $n = 1, 2, \ldots, N$.
   - Calculate the observation error L_n^t for $n = 1, 2, \ldots, N$:

$$D^t = \max_n \{|y_n - f^t(\mathbf{x}_n)|\}$$

$$L_n^t = \frac{|y_n - f^t(\mathbf{x}_n)|}{D^t}$$

Here a loop is initiated(t-model index).

Draw bootstrap sample from the data with weights calculated

$f^t(x_n)$ – fitted weak learner

D_t – maximum of absolute error

L_n^t - Error ratio by row

- Calculate the model error $\bar{L}^t$:

$$\bar{L}^t = \sum_{n=1}^{N} L_n^t w_n^t$$

If $\bar{L}^t \geq 0.5$, end iteration and set T equal to $t-1$.

- Let $\beta^t = \frac{\bar{L}^t}{1-\bar{L}^t}$. The lower β^t, the greater our confidence in the model.
- Let $Z^t = \sum_{n=1}^{N} w_n^t (\beta^t)^{1-L_n}$ and update the model weights with

$$w_n^{t+1} = \frac{w_n^t (\beta^t)^{1-L_n}}{Z^t},$$

which increases the weight for observations with a greater error L_n^t.

$\bar{L}^t$ – model error

β^t – similar to 'amount of say' – takes large value when error is large.

w_n^{t+1} - updated weights which increases with error

3. Set the overall fitted value for observation n equal to the weighted median of $f^t(x_n)$ for $t=1,2,\ldots,T$ using weights $\log(1/\beta^t)$ for model t.

Thank You