universität freiburg

Assignment 08

Machine Learning WS23/24

Assignment 08
Solution

Bagging

Bagging

Bootstrapping is random sampling with replacement. Discuss the advantages and disadvantages of bootstrapping.

- **Advantages**: Simple, Increases the sample size, introduces variance into the dataset.
- **Disadvantages**: It can lead to skewed samples, i.e. the instances are not i.i.d. anymore.

Bagging

What is bagging and why do we use it?

- Bagging is learning a set of individual models on bootstrapped samples and averaging their predictions. We use it to build ensembles which greatly profit from variances in the training data of the individual learned models.

Bagging

What is the difference between Bagging and Boosting? Given a scenario where your base model is performing well on the training dataset but not on the validation dataset, which of the two ensemble techniques you learned so far would be suitable to apply here? Comment on how the complexity of the model would matter when choosing an ensemble method.

- Bagging and Boosting are both ensemble methods that combines the prediction of many learners to give the final prediction. Bagging averages the prediction from a set of models that are trained in parallel on the bootstrapped samples, while boosting aims to penalize weak base learners sequentially by weighting the dominant learners higher, in turn to get a strong ensemble model.

Bagging

What is the difference between Bagging and Boosting? Given a scenario where your base model is performing well on the training dataset but not on the validation dataset, which of the two ensemble techniques you learned so far would be suitable to apply here? Comment on how the complexity of the model would matter when choosing an ensemble method.

- Since the base model is performing well on the training dataset and bad on the validation dataset, it indicates that the model has low bias and high variance. Using bagging in this case would be sufficient to reduce variance as it averages over the prediction of N such individual models, thus increasing the performance on the validation dataset.
- A high complexity model usually does not generalize well and is prone to overfitting, while a low complexity model can fail to even learn from the training dataset or can underfit. One might try bagging for the first case which can help to reduce variance, whereas Boosting can help to make a strong ensemble even from low complexity models.

Random Forest

Boosting

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}$$

Initialize the weights

$$w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 0.955$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$
- $L(\cdot)|x_2 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 0.955$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$ $\qquad$ $L(\cdot)|x_2 > 0 = 0.955$
- $L(\cdot)|x_3 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} = 2.131$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$ $\quad L(\cdot)|x_2 > 0 = 0.955$ $\quad L(\cdot)|x_3 > 0 = 2.131$
- $L(\cdot)|x_1 < 0 = \sum_{i=1}^{4} w_i e^{-y_i \hat{y}_i} = w_1 e^{-y_1 \hat{y}_1} + w_2 e^{-y_2 \hat{y}_2} + w_3 e^{-y_3 \hat{y}_3} + w_4 e^{-y_4 \hat{y}_4} = 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot -1)} = 2.131$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y,\hat{y},w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$ $\quad L(\cdot)|x_2 > 0 = 0.955$ $\quad L(\cdot)|x_3 > 0 = 2.131$
- $L(\cdot)|x_1 < 0 = 2.131$
- $L(\cdot)|x_2 < 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} = 2.131$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$ $\quad$ $L(\cdot)|x_2 > 0 = 0.955$ $\quad$ $L(\cdot)|x_3 > 0 = 2.131$
- $L(\cdot)|x_1 < 0 = 2.131$ $\quad$ $L(\cdot)|x_2 < 0 = 2.131$
- $L(\cdot)|x_3 < 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 0.955$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 0.955$ $\quad$ $L(\cdot)|x_2 > 0 = 0.955$ $\quad$ $L(\cdot)|x_3 > 0 = 2.131$
- $L(\cdot)|x_1 < 0 = 2.131$ $\quad$ $L(\cdot)|x_2 < 0 = 2.131$ $\quad$ $L(\cdot)|x_3 < 0 = 0.955$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$

Calculate the error and α:

- $err_1 = \frac{0.25(0+0+1+0)}{4 \cdot 0.25} = 0.25$
- $\alpha_1 = ln\left(\frac{1-err_1}{err_1}\right) = \ln\left(\frac{0.75}{0.25}\right) = ln3 = 1.099$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

Update the weights $w^{(1)}$

- $w^{(1)} = (0.25 \cdot e^{1.099 \cdot 0} \quad 0.25 \cdot e^{1.099 \cdot 0} \quad 0.25 \cdot e^{1.099 \cdot 1} \quad 0.25 \cdot e^{1.099 \cdot 0}) = (0.25 \quad 0.25 \quad 0.75 \quad 0.25)$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.75 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 2.131$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 =$*2.131*
- $L(\cdot)|x_2 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \hat{y}_i} = w_1 e^{-y_1 \hat{y}_1} + w_2 e^{-y_2 \hat{y}_2} + w_3 e^{-y_3 \hat{y}_3} + w_4 e^{-y_4 \hat{y}_4} = 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.75 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 1.139$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y,\hat{y},w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)} = (0.25 \quad 0.25 \quad 0.75 \quad 0.25)}$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $\mathrm{L}(\cdot)|\mathrm{x}_1 > 0 = 2.131$ $\qquad L(\cdot)|x_2 > 0 = 1.139$
- $L(\cdot)|x_3 > 0 = \sum_{i=1}^{4} w_i e^{-y_i\widehat{y_i}} = w_1 e^{-y_1\widehat{y_1}} + w_2 e^{-y_2\widehat{y_2}} + w_3 e^{-y_3\widehat{y_3}} + w_4 e^{-y_4\widehat{y_4}} = 0.25 \cdot e^{-(-1\cdot 1)} + 0.25 \cdot e^{-(1\cdot 1)} + 0.75 \cdot e^{-(1\cdot -1)} + 0.25 \cdot e^{-(1\cdot -1)} = 3.490$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$

Calculate the errors for the splits:

- $\mathrm{L}(\cdot)|\mathrm{x}_1 > 0 = 2.131$ $\qquad L(\cdot)|x_2 > 0 = 1.139$ $\qquad L(\cdot)|x_3 > 0 = 3.49$
- $L(\cdot)|x_1 < 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.75 \cdot e^{-(1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot -1)} = 2.315$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$

Calculate the errors for the splits:

- $\mathrm{L}(\cdot)|\mathrm{x}_1 > 0 = 2.131$ $\quad L(\cdot)|x_2 > 0 = 1.139$ $\quad L(\cdot)|x_3 > 0 = 3.49$
- $L(\cdot)|x_1 < 0 = 2.315$
- $L(\cdot)|x_2 < 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 0.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} + 0.75 \cdot e^{-(1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot -1)} = 3.49$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y,\hat{y},w) = \sum_{i=1}^{N} w_i e^{-y_i\hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25 \quad 0.25 \quad 0.25 \quad 0.25})$ split 1: $x_1 > 0$ $err_1 = 0.25$ $\alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25 \quad 0.25 \quad 0.75 \quad 0.25})$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 2.131$ $L(\cdot)|x_2 > 0 = 1.139$ $L(\cdot)|x_3 > 0 = 3.49$
- $L(\cdot)|x_1 < 0 = 2.315$ $L(\cdot)|x_2 < 0 = 3.49$
- $L(\cdot)|x_3 < 0 = \sum_{i=1}^{4} w_i e^{-y_i\hat{y}_i} = w_1 e^{-y_1\hat{y}_1} + w_2 e^{-y_2\hat{y}_2} + w_3 e^{-y_3\hat{y}_3} + w_4 e^{-y_4\hat{y}_4} = 0.25 \cdot e^{-(-1\cdot -1)} + 0.25 \cdot e^{-(1\cdot -1)} + 0.75 \cdot e^{-(1\cdot 1)} + 0.25 \cdot e^{-(1\cdot 1)} = 1.139$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$L(y,\hat{y},w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}$$

$w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$w^{(1)} = (0.25 \quad 0.25 \quad 0.75 \quad 0.25)$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 2.131$ $\quad L(\cdot)|x_2 > 0 = 1.139$ $\quad L(\cdot)|x_3 > 0 = 3.49$
- $L(\cdot)|x_1 < 0 = 2.315$ $\quad L(\cdot)|x_2 < 0 = 3.49$ $\quad L(\cdot)|x_3 < 0 = 1.139$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

$w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$w^{(1)} = (0.25 \quad 0.25 \quad 0.75 \quad 0.25)$ split 2: $x_2 > 0$

Calculate the error and α:

- $err_2 = \frac{0.25 \cdot 1 + 0.25 \cdot 0 + 0.75 \cdot 0 + 0.25 \cdot 0}{0.25 + 0.25 + 0.75 + 0.25} = \frac{0.25}{1.5} = 0.1667$
- $\alpha_2 = ln\left(\frac{1 - err_2}{err_2}\right) = 1.609$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y,\hat{y},w) = \sum_{i=1}^{N} w_i e^{-y_i\hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $\quad err_2 = 0.1667 \quad \alpha_2 = 1.609$

Update the weights $w^{(1)}$**:**

- $w^{(2)} = (0.25 \cdot e^{1.609 \cdot 1} \quad 0.25 \cdot e^{1.609 \cdot 0} \quad 0.75 \cdot e^{1.609 \cdot 0} \quad 0.25 \cdot e^{1.609 \cdot 0}) = (1.25 \quad 0.25 \quad 0.75 \quad 0.25)$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w^{(0)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)}} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $\quad err_2 = 0.1667 \quad \alpha_2 = 1.609$

$\boldsymbol{w^{(2)}} = (\mathbf{1.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = \sum_{i=1}^{4} w_i e^{-y_i \widehat{y_i}} = w_1 e^{-y_1 \widehat{y_1}} + w_2 e^{-y_2 \widehat{y_2}} + w_3 e^{-y_3 \widehat{y_3}} + w_4 e^{-y_4 \widehat{y_4}} = 1.25 \cdot e^{-(-1 \cdot -1)} + 0.25 \cdot e^{-(1 \cdot 1)} + 0.75 \cdot e^{-(-1 \cdot 1)} + 0.25 \cdot e^{-(1 \cdot 1)} = 2.682$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

$\boldsymbol{w}^{(0)} = (\boldsymbol{0.25} \quad \boldsymbol{0.25} \quad \boldsymbol{0.25} \quad \boldsymbol{0.25})$ split 1: $x_1 > 0$ $err_1 = 0.25$ $\alpha_1 = 1.099$

$\boldsymbol{w}^{(1)} = (\boldsymbol{0.25} \quad \boldsymbol{0.25} \quad \boldsymbol{0.75} \quad \boldsymbol{0.25})$ split 2: $x_2 > 0$ $err_2 = 0.1667$ $\alpha_2 = 1.609$

$\boldsymbol{w}^{(2)} = (\boldsymbol{1.25} \quad \boldsymbol{0.25} \quad \boldsymbol{0.75} \quad \boldsymbol{0.25})$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 2.682$ $\quad$ $L(\cdot)|x_2 > 0 = 3.858$ $\quad$ $L(\cdot)|x_3 > 0 = 6.208$
- $L(\cdot)|x_1 < 0 = 6.208$ $\quad$ $L(\cdot)|x_2 < 0 = 3.858$ $\quad$ $L(\cdot)|x_3 < 0 = 1.507$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w) = \sum_{i=1}^{N} w_i e^{-y_i \hat{y}_i}}$$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

$\boldsymbol{w^{(0)} = (0.25 \quad 0.25 \quad 0.25 \quad 0.25)}$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w^{(1)} = (0.25 \quad 0.25 \quad 0.75 \quad 0.25)}$ split 2: $x_2 > 0$ $\quad err_2 = 0.1667 \quad \alpha_2 = 1.609$

$\boldsymbol{w^{(2)} = (1.25 \quad 0.25 \quad 0.75 \quad 0.25)}$

Calculate the errors for the splits:

- $L(\cdot)|x_1 > 0 = 2.682$ $\qquad L(\cdot)|x_2 > 0 = 3.858$ $\qquad L(\cdot)|x_3 > 0 = 6.208$
- $L(\cdot)|x_1 < 0 = 6.208$ $\qquad L(\cdot)|x_2 < 0 = 3.858$ $\qquad L(\cdot)|x_3 < 0 = 1.507$

Boosting

Given the following dataset:

1. **Perform three rounds of Adaboost with one-level decision trees (stumps). For calculating the splits use the loss:**

$$\boldsymbol{L(y, \hat{y}, w)} = \sum_{i=1}^{N} \boldsymbol{w_i e^{-y_i \hat{y}_i}}$$

$\boldsymbol{w}^{(0)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $\quad err_1 = 0.25 \quad \alpha_1 = 1.099$

$\boldsymbol{w}^{(1)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $\quad err_2 = 0.1667 \quad \alpha_2 = 1.609$

$\boldsymbol{w}^{(2)} = (\mathbf{1.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 3: $x_3 < 0$

Calculate the error and α:

- $err_3 = \frac{1.25 \cdot 0 + 0.25 \cdot 1 + 0.75 \cdot 0 + 0.25 \cdot 0}{1.25 + 0.25 + 0.75 + 0.25} = \frac{0.25}{2.5} = 0.1$
- $\alpha_3 = ln\left(\frac{1 - err_3}{err_3}\right) = 2.197$

x_1	x_2	x_3	y
-1	+1	+1	-1
+1	+1	+1	+1
-1	+1	-1	+1
+1	+1	-1	+1

Boosting

$\boldsymbol{w}^{(0)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $err_1 = 0.25$ $\alpha_1 = 1.099$

$\boldsymbol{w}^{(1)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $err_2 = 0.1667$ $\alpha_2 = 1.609$

$\boldsymbol{w}^{(2)} = (\mathbf{1.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 3: $x_3 < 0$ $err_3 = 0.1$ $\alpha_3 = 2.197$

Calculate the predictions for the following test set:

x_1	x_2	x_3
+1	-1	-1
-1	-1	-1
+1	-1	+1

Boosting

$\boldsymbol{w}^{(0)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $err_1 = 0.25$ $\alpha_1 = 1.099$

$\boldsymbol{w}^{(1)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $err_2 = 0.1667$ $\alpha_2 = 1.609$

$\boldsymbol{w}^{(2)} = (\mathbf{1.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 3: $x_3 < 0$ $err_3 = 0.1$ $\alpha_3 = 2.197$

Calculate the predictions for the following test set:

x_1	x_2	x_3	G_1	G_2	G_3
+1	-1	-1	+1	-1	+1
-1	-1	-1	-1	-1	+1
+1	-1	+1	+1	-1	-1

Boosting

$\boldsymbol{w}^{(0)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.25})$ split 1: $x_1 > 0$ $err_1 = 0.25$ $\alpha_1 = 1.099$

$\boldsymbol{w}^{(1)} = (\mathbf{0.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 2: $x_2 > 0$ $err_2 = 0.1667$ $\alpha_2 = 1.609$

$\boldsymbol{w}^{(2)} = (\mathbf{1.25} \quad \mathbf{0.25} \quad \mathbf{0.75} \quad \mathbf{0.25})$ split 3: $x_3 < 0$ $err_3 = 0.1$ $\alpha_3 = 2.197$

Calculate the predictions for the following test set:

$$\hat{y} = sign\left(\sum_{m=1}^{M} \alpha_m G_m(x)\right)$$

x_1	x_2	x_3	G_1	G_2	G_3
+1	-1	-1	+1	-1	+1
-1	-1	-1	-1	-1	+1
+1	-1	+1	+1	-1	-1

$$\hat{y} = sign\begin{pmatrix} 1.099 \cdot 1 + 1.609 \cdot (-1) + 2.197 \cdot 1 \\ 1.099 \cdot (-1) + 1.609 \cdot (-1) + 2.197 \cdot 1 \\ 1.099 \cdot 1 + 1.609 \cdot (-1) + 2.197 \cdot (-1) \end{pmatrix}$$

$$\hat{y} = sign\begin{pmatrix} 1.687 \\ -0.511 \\ -2.707 \end{pmatrix} = \begin{pmatrix} +1 \\ -1 \\ -1 \end{pmatrix}$$