# Boosting a decision stump

**6/6** points earned (100%)

Quiz passed!

Continue Course (/learn/ml-classification/supplement/5PJ3W/slides-presented-in-this-module)

Back to Week 5 (/learn/ml-classification/home/week/5)

✓    1 / 1
points

1.

### Are you using GraphLab Create? Please make sure that

**1. You are using version 1.8.3 of GraphLab Create.** Verify the version of GraphLab Create by running

```
graphlab.version
```

inside the notebook. If your GraphLab version is incorrect, see this post (https://www.coursera.org/learn/ml-classification/supplement/LgZ3I/installing-correct-version-of-graphlab-create) to install version 1.8.3. **This assignment is not guaranteed to work with other versions of GraphLab Create.**

**2. You are using the IPython notebook** named module-8-boosting-assignment-2-blank.ipynb obtained from the associated reading.

This question is ungraded. Check one of the three options to confirm.

✔ 1 / 1
points

2.

Recall that the **classification error for unweighted data** is defined as follows:

$$\text{classification error} = \frac{\#\ \text{mistakes}}{\#\ \text{all data points}}$$

Meanwhile, the **weight of mistakes for weighted data** is given by

$$\text{WM}(\alpha, \hat{\mathbf{y}}) = \sum_{i=1}^{n} \alpha_i \times 1[y_i \neq \hat{y}_i].$$

If we set the weights **α=1** for all data points, how is the weight of mistakes **WM(α,ŷ)** related to the classification error?

---

1 / 1
points

3.

Refer to section **Example: Training a weighted decision tree**.

Will you get the same model as **small_data_decision_tree_subset_20** if you trained a decision tree with only 20 data points from the set of points in **subset_20**?

---

1 / 1
points

4.

Refer to the 10-component ensemble of tree stumps trained with Adaboost.

As each component is trained sequentially, are the component weights monotonically decreasing, monotonically increasing, or neither?

✔ 1 / 1
points

## 5.
Which of the following best describes a **general trend in accuracy** as we add more and more components? Answer based on the 30 components learned so far.

✔ 1 / 1
points

## 6.
From this plot (with 30 trees), is there massive overfitting as the # of iterations increases?

👍 👎 🚩