



1

Module 3: Boosted Algorithms

Data Science Immersive

S

2

Lesson Overview

Aim: SWBAT differentiate between bagging and boosting ensemble methods, as well as explain how two specific boosting methods (AdaBoost and Gradient Boosting) differ.

Agenda:

- Review Ensemble Methods
- Learn about the AdaBoost method
- Learn about the Gradient Boosting

3

Boosting

Boosting (originally called hypothesis boosting) refers to any ensemble method that train predictors sequentially, each trying to correct its predecessor in some fashion.

Model 1, ..., N are individual models (e.g. decision trees)

```
graph LR; M1[Model 1] --> E[Ensemble]; M2[Model 2] --> E; MN[Model N] --> E; E --> O[Output];
```

There are many boosting methods available, but by far the most popular are AdaBoost(short for Adaptive Boosting) and Gradient Boosting.

4

Bagging vs. Boosting

```
graph LR; subgraph Bagging; B1[Model 1] --> B2[Model 2] --> B3[Model 3]; end; subgraph Boosting; B1[Model 1] --> B2[Model 2] --> B3[Model 3]; end;
```

5

Boosting Models

- Boosting is a **generic algorithm rather than a specific model**. Boosting needs you to specify a weak model (e.g. regression, shallow decision trees, etc) and then improves it.
- This allowed different loss functions to be used, expanding the technique beyond binary classification problems to support regression, multi-class classification and more.
- Boosting is a statistical framework where the objective is to minimize the loss of the model by adding weak learners using a **gradient descent** like procedure.
- This class of algorithms were described as a **stage-wise additive** model. This is because one new weak learner is added at a time and existing weak learners in the model are frozen and left unchanged.

6

AdaBoost - Adaptive Boosting

AdaBoost is a specific Boosting algorithm developed for classification problems (also called discrete AdaBoost). The weakness is identified by the weak estimator's error rate.

Module 3: Boosted Algorithms

Data Science Immersive

