Recap

In this lesson, you learned about a number of techniques used in ensemble met
the techniques, you saw that there are two variables with tradeoffs **Bias** and **Var**

- **High Bias, Low Variance** models tend to underfit data, as they are not flexi
into this category of models.
- **High Variance, Low Bias** models tend to overfit data, as they are too flexib
into this category of models.

Ensemble Models

In order to find a way to optimize for both variance and bias, we have ensemble
methods have become some of the most popular methods used to compete in c
and used in industry across applications.

There were two randomization techniques you saw to combat overfitting:

1. **Bootstrap the data** - that is, sampling the data with replacement and fitting
fitting your algorithm to the sampled data.
2. **Subset the features** - in each split of a decision tree or with each algorithm
only a subset of the total possible features are used.

Techniques

You saw a number of ensemble methods in this lesson including:

- BaggingClassifier
- RandomForestClassifier
- AdaBoostClassifier

Another really useful guide for ensemble methods can be found in the documer
methods can also all be extended to regression problems, not just classification.

Additional Resources

Additionally, here are some great resources on AdaBoost if you'd like to learn so

- Here is the original paper from Freund and Schapire.
- A follow-up paper from the same authors regarding several experiments w
- A great tutorial by Schapire.