



1

Should we trust an expert or a crowd?

You are at an auction trying to determine whether or not a specific painting is a fraud. You decide to ask around and get opinions from people at the auction site.

Option A:

An expert art inspector has an accuracy of 80%. There is only one on site.

Option B:

Other attendees, who are not quite as knowledgeable, accurately identify whether or not a piece of art is fake 55% of the time. There are at most 100 other attendees you can ask.

Assuming you can only have time for one option, which would you choose? Why?

Hint: Think about how you might be able to use a binomial distribution perhaps....

2

Bagging Methods with Random Forests

Data Science Immersive

// FLATIRON SCHOOL

3

Wisdom of the Crowd!

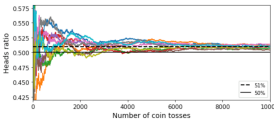
At a 1906 fair in Plymouth, England, statistician Francis Galton noticed how when 800 people guessed how much a "dressed" or weighed. It turns out that the actual guess had only a 1% error from the median guess



4

Why Ensemble Methods

Which would you rather use to solve a problem, a bunch of simple models that aren't that accurate or one complex model that is very accurate?



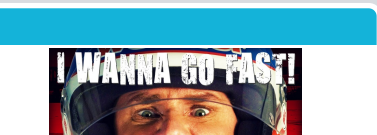
You have a biased coin that has a 51% chance of coming up heads and a 49% chance of coming up tails. If you toss it 1,000 times the probability of having a majority of heads is 75%. If you increase your tosses to 10,000, that probability rises to 97%.

5

Objectives

- Understanding Bootstrapping and Aggregating (Bagging) of models
- Understanding the Random Forests algorithm

6



31

