Since our data may be huge, in general, we don't want to train many models on the same data. This would be very expensive. Instead, we'll just take subsets of it and train a weak learner on each one of these subsets. Then we'll figure out how to combine these learners. A weak learner in this case is one that splits the data either vertically or horizontally.

Notice that we never partition the data. We are completely allowed to repeat points among our subsets and to even not consider some of the points at all. At every step, we pick a fully random subset of data.

How do we combine our weak learners?

By voting. We over impose each learner on the data. Because we have three learners in our example, if two or more of them predict blue, then that region is blue, and if two or more of them predict red, then that region is red. If we have an even number of models, we can pick any way we want to break ties. Although with lots of points and lots of models, it's hard to imagine that we would get a tie somewhere.

Quiz Question

What is the purpose of voting in Bagging?

a. Voting is used to set the priority of weak learners.
b. Voting is the last step after training and is used to combine the weak learner results.