

According to the StatQuest video, random forests are a powerful technique that builds upon decision trees. They offer several advantages including increased flexibility and improved accuracy.

The video explains how to create a random forest. First, it involves generating a bootstrap dataset by randomly sampling the original data. Then, decision trees are built using a random subset of features at each step. This process is repeated multiple times, resulting in a collection of diverse trees. Finally, when classifying new data points, the random forest considers the votes from each tree to make the final prediction.

Another key concept introduced in the video is out-of-bag data. This data is left out of the bootstrap process and can be used to assess the performance of the random forest.

Overall, the video highlights that random forests offer a simpler and more accurate approach compared to individual decision trees.