决策树是机器学习的常用方法。 树学习是如今最能满足于数据挖掘的方法，因为它在特征值的缩放和其他各种转换下保持不变，对无关特征是稳健的，而且能生成可被检查的模型。然而，它通常并不准确。

特别的，生长很深的树容易学习到高度不规则的模式，即过学习，在训练集上具有低偏差和高变异数的特点。随机森林是平均多个深决策树以降低变异数的一种方法，其中，决策树是在一个数据集上的不同部分进行训练的。这是以偏差的小幅增加和一些可解释性的丧失为代价的，但是在最终的模型中通常会大大提高性能。

随机森林训练算法把bagging的一般技术应用到树学习中,在训练结束之后，对未知样本x的预测可以通过对x上所有单个回归树的预测求平均来实现或者在分类任务中选择多数投票的类别。这种bagging方法在不增加偏置的情况下降低了方差，从而带来了更好的性能。这意味着，即使单个树模型的预测对训练集的噪声非常敏感，但对于多个树模型，只要这些树并不相关，这种情况就不会出现。