**Gradient boosting算法（GBM）和随机森林都是基于树的算法，它们有什么区别？**

//参考http://www.jianshu.com/p/18a400ad47cc

相同点：

(1)都是由多棵树组成的，都是集成学习算法

(2)最终的结果都是由多颗树一起决定

不同点：

(1)组成随机森林的树可以是分类树，也可以是回归树，但是GBDT只能由回归树组成。

(2)组成随机森林的树可以并行生成，但是组成GBDT的树只能串行生成。

(3)对于最终的输出结果，随机森林采用多数投票；而GBDT是将所有的结果累加起来，或者加权起来

(4)随机森林对异常值不敏感，而GBDT对异常值非常敏感

(5)随机森林通过减小方差来提高性能，GBDT通过减小偏差来提高性能

**gbdt在调参时为什么树的深度很少就能达到很高的精度？**

对于Bagging算法来说，由于我们会并行地训练很多不同的分类器的目的就是降低这个方差(variance) ,因为采用了相互独立的基分类器多了以后，h的值自然就会靠近.所以对于每个基分类器来说，目标就是如何降低这个偏差（bias),所以我们会采用深度很深甚至不剪枝的决策树。

  对于Boosting来说，每一步我们都会在上一轮的基础上更加拟合原数据，所以可以保证偏差（bias）,所以对于每个基分类器来说，问题就在于如何选择variance更小的分类器，即更简单的分类器，所以我们选择了深度很浅的决策树。