**随机森林**

(https://easyai.tech/ai-definition/random-forest/)

·随机森林的四个步骤



·优缺点

\* 优点

1. 它可以出来很高维度（特征很多）的数据，并且不用降维，无需做特征选择
2. 它可以判断特征的重要程度
3. 可以判断出不同特征之间的相互影响
4. 不容易过拟合
5. 训练速度比较快，容易做成并行方法
6. 实现起来比较简单
7. 对于不平衡的数据集来说，它可以平衡误差。
8. 如果有很大一部分的特征遗失，仍可以维持准确度。

\* 缺点

1. 在某些噪音较大的分类或回归问题上会过拟合
2. 对于有不同取值的属性的数据，取值划分较多的属性会对随机森林产生更大的影响，所以随机森林在这种数据上产出的属性权值是不可信的