

1. 缺失值较多

直接将该特征舍弃掉，否则可能反倒会带入较大的nois，对结果造成不良影响

2. 缺失值较少

- (1) 将NaN直接作为一个特征，假设用0表示
- (2) 用均值填充
- (3) 用随机森林等算法预测填充