#### 数据预处理面试问题汇总

1. 数值特征的预处理办法有哪些? 为什么要进行这种预处理呢?
2. 数值特征的预处理办法有归一化和标准化。
3. 假设有两种特征，一个特征的数量级在(0,10)，另一个特征数量级在(100,10000)，第二个特征对结果的影响是比较大的。所以在把特征处理成相似的数量级。
4. 类别特征的处理办法有哪些?

有LabelEncoder和OneHotEncoder，还有一个我不熟的BinaryEncoder。