# Titanic总结

1. 数据预处理部分
   1. 加载数据
   2. 处理缺失值：数值型缺失值可以填0也可以填中位数，pandas提供了相关的库。非数值型缺失可以使用get\_dummies中设置dummy\_na=True把所有非数值特征转换成为独热编码
   3. 主要有三个类型需要考虑：DataFrame、Tensor、numpy，在进行类型转换的时候要注意这三个数据结构，不要转出错来。
2. 定义网络结构 – 这一块需要摸索几层硬藏层合适、每层几个节点合适。
3. 数据分批次加载，注意这里要把feature和label分别加载到DataLoader中，然后在后续处理的时候方便进行
4. 网络正向传播，三件套：
   1. L = loss(model(train\_loader), labels) # 计算损失
   2. Optimizer.zero\_grad() # 梯度清零
   3. Optimizer.setp() # 更新参数
5. 然后就是测试，用model传进test\_feature生成预测结果。