- 1. 读取 breast-cancer-wisconsin.csv 文件
- 2. 输出前五行查看数据内容
- 3. 将数据中的缺失值用该列特征的中位数填充(提示:将?替换为 nan,将 nan 修改为该列特征的中位数)
- 4. 划分数据集为训练集和测试集(其中第 1 列为 id 号, 第 2-10 列为特征, 11 列为标签(2 为良性、4 为恶性))
- 5. 使用三种朴素贝叶斯算法对训练集进行拟合,分别输出测试集的 accuracy_score
- 6. 分别输出三种朴素贝叶斯算法的十折交叉验证的平均准确度