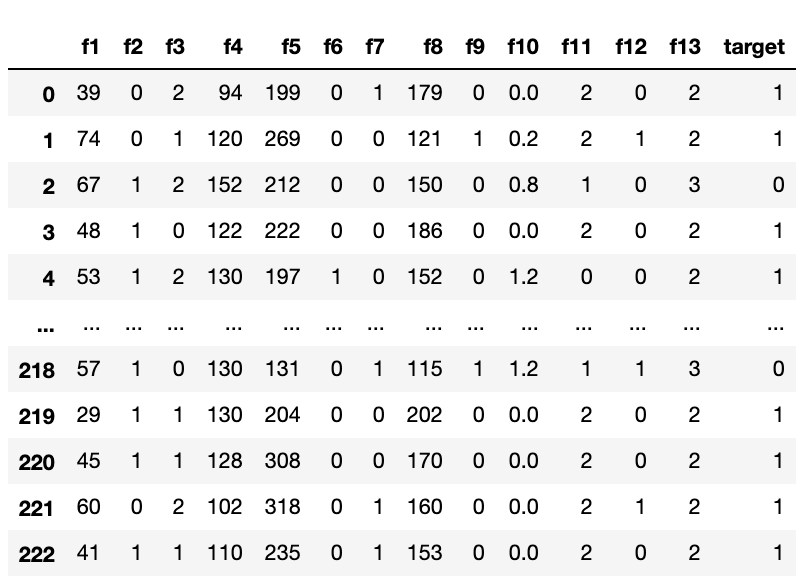
**实验6 数据挖掘分类预测实验I**

**1. 实验数据**

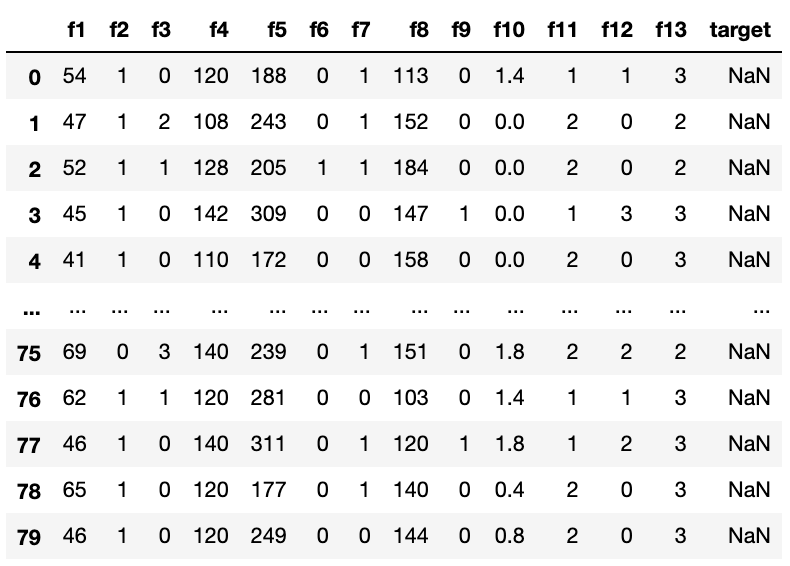
该数据集包含223条**训练数据**，存放于“data-train.csv”文件；

另有80条未知标签的**测试数据**，保存在“data-test.csv”文件中。

**训练集数**据共含与某种医学**疾病相关的13个（匿名）特征属性(f1~f13)**和**1个目标属性(target)**，具体字段如下:



**测试数据**的具体字段如下:



其中目标属性，即**target字段未知，待建模预测**。

## 2. 实验目的

本次实验的目的是利用机**器学习分类算法，**基于训练集构建分类器模型，进而预测测试集中全体样本的分类结果，即**测试样本的target 值：0 或1。**

## 3. 实验要求

## (1) 具体的机器学习算法不限，以预测效果最佳为目标。

### (2) 将预测结果保存在名为“predictions.txt”的文本文件中，内容为80行， 每一行只有0 或者1，代表你的算法对测试数据的预测结果。预测数据顺序须与测试集“data-test.csv”中的样本顺序保持一致。

(3) 将结果文件“predictions.txt”以附件形式提交至学习通。另外请注意：只需提交结果文件，无需提交本次实验报告文件。

(4) 本次实验成绩评定采用竞赛机制，即计算每位同学预测结果的F1 score，然后由高到低进行排序评定相应的实验成绩。