**附录：**

**附录一：stroke.csv数据集标签的详细说明**

1） id：唯一标识符

2） gender：“男”、“女”或“其他”

3） age：患者年龄

4） hypertension：如果患者没有高血压，则为0；如果患者患有高血压，则为1

5） heart\_disease：如果患者没有任何心脏病，则为0；如果患者患有心脏病，则为1

6） ever\_married：“否”或“是”

7） work\_type：“children”，“Govt\u jov”，“Never\u worked”，“Private”或“自营职业者”

8） Residence\_type：住宅类型“农村”或“城市”

9） avg\_glucose\_level：平均葡萄糖水平，血液中的平均葡萄糖水平

10） bmi：体重指数

11） smoking\_status：“以前吸烟”，“从未吸烟”，“吸烟”或“未知”

12） stroke：如果患者中风，则为1；如果没有中风，则为0

注意：smoking\_status中的“未知”表示该患者无法获得该信息。

**附录二：heart.csv数据集标签的详细说明**

* 1. Age：患者年龄[年]
  2. Sex：患者性别[男：男，女：女]
  3. ChestPainType：胸痛型[TA：典型心绞痛，ATA：非典型心绞痛，NAP：非心绞痛，ASY：无症状]
  4. RestingBP：静息血压[毫米汞柱]
  5. Cholesterol：血清胆固醇
  6. FastingBS：空腹血糖[1：如果空腹血糖>120 mg/dl，0：否则]
  7. RestingECG：静息心电图结果[正常：正常，ST：ST-T波异常（T波倒置和/或ST抬高或压低>0.05 mV），LVH：根据Estes标准显示可能或明确的左心室肥大]
  8. MaxHR：达到的最大心率[数值介于60和202之间]
  9. ExerciseAngina：运动诱发的心绞痛
  10. Oldpeak：Oldpeak=ST[抑郁症测量的数值]
  11. ST\_Slope：峰值运动ST段的斜率[向上：向上倾斜，平坦：平坦，向下：向下倾斜]
  12. HeartDisease：是否有心脏病 [1：心脏病，0：正常]

**附录三：cirrhosis.csv数据集标签的详细说明**

1） ID：唯一标识符

2） N\_Days：从登记到死亡、移植或研究分析时间（以较早者为准）之间的天数

3） Status：患者的状态C（被审查），CL（肝脏被审查）或D（死亡）

4） Drug：药物类型D-青霉胺或安慰剂

5） Age：年龄（天）

6） Sex：M（男）或F（女）

7） Ascites：存在腹水N（否）或Y（是）

8） Hepatomegaly：存在肝肿大N（否）或Y（是）

9） Spiders：Spiders N（否）或Y（是）的存在

10）Edema：存在水肿N（无水肿，无水肿利尿剂治疗），S（无利尿剂存在的水肿，或利尿剂解决的水肿）或Y（尽管有利尿剂治疗的水肿）

11） Bilirubin：血清胆红素[毫克/分升]

12） Cholesterol：血清胆固醇[毫克/分升]

13） Albumin：白蛋白[克/分升]

14） Copper：尿铜[微克/天]

15） Alk\_Phos：碱性磷酸酶，单位为[单位/升]

16） SGOT：SGOT单位[单位/毫升]

17） Triglycerides：甘油三酯（mg/dl）

18） Platelets：每立方血小板[毫升/千]

19） Prothrombin：凝血酶原时间（秒）

20） Stage：疾病的组织学阶段（1、2、3或4）