**ILPD (Indian Liver Patient Dataset) Data Set** : ILPD.csv

该数据集共包括416名肝病患者和167名非肝病患者。11个属性如下：

1. Age：Age of the patient   
2. Gender：Gender of the patient   
3. TB：Total Bilirubin **（总胆红素）**  
4. DB：Direct Bilirubin **（直接胆红素）**  
5. Alkphos：Alkaline Phosphotase**（碱性磷酸酶）**   
6. Sgpt：Alamine Aminotransferase **（丙氨酸转氨酶）**  
7. Sgot：Aspartate Aminotransferase **（天冬氨酸转氨酶）**  
8. TP：Total Protiens**（总蛋白）**   
9. ALB：Albumin **（白蛋白）**  
10. A/G：Ratio Albumin and Globulin Ratio **（白蛋白及球蛋白比例）**  
11. Class：liver patient**(1)** / non-liver patient**(0)**

1. **分类**

**数据文件ILPD.csv中的记录已经预先被随机打乱顺序**

**（1）利用样本节点,选取数据集前466条记录作为训练集，剩余的数据作为测试集，并保存为相应的csv文件。(20分)**

**（2）使用C5.0算法，利用训练集训练出一个分类器模型（15分）**

**（3）将分类器模型应用到测试集上，使用表节点查看预测结果。（15）分）**

**（4）使用输出->分析节点,查看分类器的评价指标（15分）**

**（5）建第2个流,使用字段->分区节点，分为训练集与测试集（20分）,并完成C5.0分类建模和分析节点的分析**

**（6）使用图形->评估节点，模型的ROC曲线（15分）**