桂林电子科技大学2021-2022学年 第2学期

**机器学习 实验报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 实验名称 | **实验一、线性回归与决策树** | | |  |
| 学 号 | **1901510216** | 姓名 | **李赏** |

**一、实验目的**

（1）掌握线性回归算法和决策树算法ID3的原理

（2）学会线性回归算法和决策树算法 ID3 的实现和使用方法。

**二、实验内容**

1、线性回归模型实验

（1）假设 line-ext.csv 是对变量 y 随变量 x 的变化情况的统计数据集。请根据教 材的公式（3.7，3.8），使用 Python 语言编程计算线性回归模型的系数，建立一个 线性回归模型。

1）计算出回归系数，输出模型表达式，绘制散点图和回归 直线，输出均方误差。

2）给出当自变量 x =0.8452 时，因变量 y 的预测值。

（2）对于上面的数据集 line-ext.csv，分别使用第三方模块 sklearn 模块中的类

LinearRegression、第三方模块 statsmodels 中的函数 ols()来计算线性回归模型的系 数，建立线性回归模型，验证上面计算结果及预测结果。

2. 决策树算法实验

（1）隐形眼镜数据集 glass-lenses.txt 是著名的数据集。它包含了很多患者眼部 状况的观察条件以及医生推荐的隐形眼镜类型。请使用 Python 语言建立决策树模 型 ID3，划分 25%的数据集作为测试数据集，并尝试进行后剪枝操作。要求

1）使用 Graphviz 工具，将此决策树绘制出来。

1. 输出混淆矩阵和准确率。

（2）对于数据集 glass-lenses.txt，使用 sklearn 模块中的类DecisionTreeClassifier

（其中，criterion=’entropy’）来建立决策树模型 ID3，重复 2（1）中的操作，验证 上面计算结果。

**三、实验代码和过程**

1. 线性回归模型实验

代码如下：

2、决策树

代码如下：

## ****四、实验结果截图及结果分析****

1、线性回归模型实验

### 2、 决策树算法实验

## ****五、实验中遇到的问题和解决问题的方法****