

机器学习

第1次实验：单变量线性回归

实验目的

01

根据数据创建单变量线性回归模型

02

使用线性回归模型进行预测

03

对预测结果进行评价



实验内容

假设某披萨店的披萨价格和披萨直径之间有下列数据关系：
训练数据：

训练样本	直径 (英寸)	价格 (美元)
1	6	7
2	8	9
3	10	13
4	14	17.5
5	18	18

根据上面的训练数据，预测12英寸的披萨的可能售价。

实验内容

题目1：直径为自变量 x ，价格为因变量 y ，画出二者的散点图，并给出结论。

题目2：根据现有的训练数据求线性回归模型，并画出拟合直线。：

可以使用sklearn库中的sklearn.linear_model.LinearRegression对象来进行线性拟合，给出拟合直线方程

步骤：准备训练数据

创建模型对象

拟合

求线性方程的截距和斜率

画拟合直线

实验内容

题目3：预测12英寸披萨的价格。（使用predict函数）

题目4：评价模型的准确率，分析模型预测结果。

测试数据：

测试数据	直径（英寸）	价格（美元）
1	8	8.5
2	9	11
3	11	12
4	12	15
5	16	18

模型准确率计算方法

A. 手动计算方法:

假设hpyTrain代表针对训练数据的预测y值, hpyTest代表针对测试数据的预测y值

- ✓ 训练数据残差平方和: $ssResTrain = \sum((hpyTrain - yTrain) ** 2)$
- ✓ 测试数据残差平方和: $ssResTest = \sum((hpyTest - yTest) ** 2)$
- ✓ 测试数据偏差平方和: $ssTotTest = \sum((yTest - np.mean(yTest)) ** 2)$
- ✓ R方: $Rsquare = 1 - ssResTest / ssTotTest$

B. Python的LinearRegression对象提供的方法:

- ✓ 训练数据残差平方和: `model._residues`
- ✓ R方: `model.score(xTest, yTest)`

结果图

