

暑期创造营Python作业

孙健老师团队

2021年7月30日

1. 使用 `numpy` 和循环，对于任意一个给定的 n ，构造下面这个 $n \times n$ 的矩阵。

$$\begin{bmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \\ 3 & 4 & 5 & \cdots & 1 & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ n-1 & n & 1 & \cdots & n-3 & n-2 \\ n & 1 & 2 & \cdots & n-2 & n-1 \end{bmatrix}$$

2. (1) 编写一个 `for` 循环，对于任意给定的正整数 n ，计算 n 的阶乘。
(2) 把这段代码变成一个函数 `factorial(n)`，使得这个函数对于任何正整数 n ，都能输出 n 的阶乘。

3. 函数优化求解。

- (1) 写一个 python 函数 $f(x)$ ，计算如下式子的值。

$$f(x) = x^4 + x^2$$

- (2) 调用 (1) 中的函数，计算 $f(2)$ 。
- (3) 调用 `scipy.optimize.minimize`，求解当初始值为 $-2, -0.5, 0, 1$ 等四种情况下， $f(x)$ 的极小值点。思考每种情况结果的原因。

4. 在 Python 环境下，使用

```
from sklearn.datasets import load_boston
boston = load_boston()
x = boston.data
y = boston.target
```

获取波士顿房价数据。随机取 100 个数据点作为测试集，其他数据点作为训练集。建立线性回归模型，查看模型的 `score`。

5. 分类模型对比。

- (1) 自行做出一个具有 10 个特征的数据集，分别是特征 X_1, \dots, X_{10} 。每个特征分别取 0, 1。随机产生 2000 条数据，其中每个特征取 0, 1 的概率分别是 $1/2$ 。令标签向量为 Y 。

$$Y = \text{mod}(X_1 * X_2 + X_3 * X_4 + X_5 * X_6 + X_7 * X_8 + X_9 * X_{10}, 2)$$

- (2) 取前面 1000 条数据作为训练集，取后面 1000 条数据作为测试集。

- (3) 使用训练集训练逻辑回归、决策树模型、KNN模型。比较三种方法在训练集和测试集上的优劣。

注：生成随机整数可以使用 `numpy.random.randint`，具体使用方式可查阅 numpy 文档。
公式中的 mod 是指取余，可使用 `numpy.mod` 计算。