**徽标

描述已自动生成**

卡通画

描述已自动生成

**数智NLP组2022第一阶段考核报告**

**softmax回归模型的原理和实现**

学 院 计算机学院

专 业 计算机

年级班别 2021级（6）班

学 号 3121004890

学生姓名 赵继业

2022年4月1日

目录

1.实现原理**2**

1.1softmax回归.....................................................................................................................................2

1.2代价函数............................................................................................................................................2

1.2.1定义softmax回归的代价函数...................................................................................................2

1.2.2加入正则化惩罚项.........................................................................................................................2

1.3梯度下降...........................................................................................................................................3

1.3.1梯度下降的优化公式3

1.3.2对代价函数的求导3

2.实现过程**4**

2.1数据预处理5

2.2梯度下降...................................................................................................................................................5

2.3保存参数...................................................................................................................................................5

2.4预测结果..................................................................................................................................................5

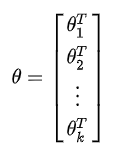
参考文献6

**1.实现原理**

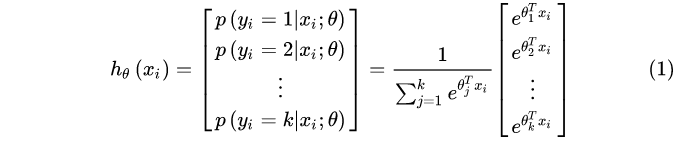
**1.1softmax回归**

softmax回归用于解决多分类问题，对于输入数据 有k个类别，用softmax 回归估算输入数据 属于每一类的概率。

将softmax 回归的参数矩阵记为

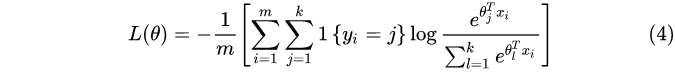
（1）

得到属于每一类的概率为

**** （2）

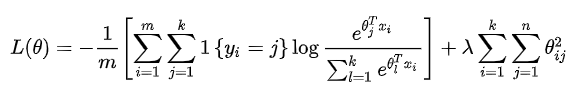
**1.2代价函数**

1.2.1定义 softmax 回归的代价函数

**** （3）

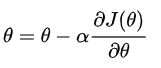
其中，1{}是示性函数，即1{值为真的表达式}=1，1{值为假的表达式}=0

1.2.2加入正则化惩罚项

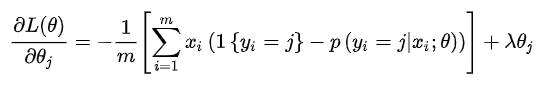
(4)

**1.3梯度下降**

1.3.1梯度下降的优化公式

 （5）

1.3.2对代价函数求导

（6）

注意这里的为常数项，不需要进行惩罚

**2.实现过程**

**2.1数据预处理**

读取训练集和测试集，并将训练集的特征和标签分离，得到标签的one hot矩阵。对训练集和测试集进行归一化，归一化方法为最大最小值归一化，

**2.2梯度下降**

这一步需要使用softmax激活函数输出一个矩阵，矩阵包含所有样本的softmax输出结果。根据公式（6）计算梯度，这里不单独计算的梯度，用one hot矩阵减去公式（2）softmax输出矩阵再与特征矩阵相乘，再加上正则化惩罚性，即可得到整个梯度矩阵。再根据公式（5），完成一次梯度下降的迭代。

**2.3保存参数**

保存参数矩阵，可以再下次计算时接着上次的计算，不需要再随机初始化参数矩阵。

**2.4预测结果**

根据softmax函数得到每个样本属于每个类别的概率，概率最大的类别即为输出的标签

**参考文献**

[1] [softmax回归原理与实现 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/98061179)

[2] [Python、Numpy 教程 | NumPy 中文](https://www.numpy.org.cn/article/basics/python_numpy_tutorial.html#python)

[3] [01 - Lesson · Pandas 官方教程 (gitbooks.io)](https://wizardforcel.gitbooks.io/pandas-official-tut-zh/content/learn-pandas/1.html)