

公告

昵称: 冯瑾元  
园龄: 2年11个月  
粉丝: 1  
关注: 2  
+加关注

< 2018年8月 >						
日	一	二	三	四	五	六
29	30	31	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	1
2	3	4	5	6	7	8

搜索

找找看

谷歌搜索

常用链接

- 我的随笔
- 我的评论
- 我的参与
- 最新评论
- 我的标签

随笔分类(47)

- Deep Learning(2)
- LaTeX(1)
- Linear Algebra(8)
- Linux(1)
- Machine Learning(9)
- MIT\_Linear Algebra
- Opencv(1)
- Papers(1)
- Probability Statistics(1)
- Python(5)
- USACO(18)

随笔档案(47)

- 2016年6月 (12)
- 2016年5月 (8)
- 2016年4月 (5)
- 2015年8月 (4)
- 2014年8月 (7)
- 2014年7月 (5)
- 2014年6月 (6)

阅读排行榜

Logistic回归做多分类和Softmax回归

我们已经知道，普通的logistic回归只能针对二分类(Binary Classification)问题，要想实现多个类别的分类，我们必须改进logistic回归，让其适应多分类问题。

关于这种改进，有两种方式可以做到。

第一种方式是直接根据每个类别，都建立一个二分类器，带有这个类别的样本标记为1，带有其他类别的样本标记为0。假如我们有 $k$ 个类别，最后我们就得到了 $k$ 个针对不同标记的普通的logistic分类器。

第二种方式是修改logistic回归的损失函数，让其适应多分类问题。这个损失函数不再笼统地只考虑二分类非1就0的损失，而是具体考虑每个样本标记的损失。这种方法叫做softmax回归，即logistic回归的多分类版本。

我们首先简单介绍第一种方式。

对于二分类问题，我们只需要一个分类器即可，但是对于多分类问题，我们需要多个分类器才行。假如给定数据集 $\mathbf{X} \in \mathbb{R}^{m \times n}$ ，它们的标记 $\mathbf{Y} \in \mathbb{R}^k$ ，即这些样本有 $k$ 个不同的类别。

我们挑选出标记为 $c(c \leq k)$ 的样本，将挑选出来的带有标记 $c$ 的样本的标记置为1，将剩下的不带有标记 $c$ 的样本的标记置为0。然后用这些数据训练出一个分类器，我们得到 $h_c(x)$ （表示针对标记 $c$ 的logistic分类函数）。

按照上面的步骤，我们可以得到 $k$ 个不同的分类器。针对一个测试样本，我们需要找到这 $k$ 个分类函数输出值最大的那一个，即为测试样本的标记：

$$\arg \max_c h_c(x) \quad c = 1, 2, \dots, k$$

下面我们介绍softmax回归。

对于有 $k$ 个标记的分类问题，分类函数是下面这样：

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}, \theta) \\ p(y^{(i)} = 2|x^{(i)}, \theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)}, \theta) \end{bmatrix} = \frac{1}{\sum_{c=1}^k e^{\theta_c^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}$$

在这里，我们将上式的所有的 $\theta_1, \theta_2, \dots, \theta_k$ 组合起来，用矩阵 $\theta$ 来表示，即：

$$\theta = \begin{bmatrix} \theta_1^T \\ \theta_2^T \\ \vdots \\ \theta_k^T \end{bmatrix}$$

这时候，softmax回归算法的代价函数如下所示（其中 $\text{sign}(\text{expression is true}) = 1$ ）：

$$J(\theta) = - \sum_{i=1}^m \sum_{c=1}^k \text{sign}(y^{(i)} = c) \log p(y^{(i)} = c|x^{(i)}, \theta) = - \sum_{i=1}^m \sum_{c=1}^k \text{sign}(y^{(i)} = c) \log \frac{e^{\theta_c^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}}$$

很明显，上述公式是logistic回归损失函数的推广。

我们可以把logistic回归的损失函数改为如下形式：

1. Python中使用numpy创建的array之间的乘法(8213)
2. Logistic回归做多分类和Softmax回归(5783)
3. 对LeNet-5卷积神经网络的理解(3766)
4. 矩阵的迹运算 (trace operator) (2807)
5. 矩阵的伪逆 (pseudoinverse) (2791)

#### 推荐排行榜

1. Broken Necklace 坏掉的项链 (1)
2. Python中使用numpy创建的array之间的乘法(1)
3. 对LeNet-5卷积神经网络的理解 (1)

$$J(\theta) = - \sum_{i=1}^m y^{(i)} \log h_{\theta}(x^{(i)}) - (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)})) = - \sum_{i=1}^m \sum_{c=0}^1 \text{sign}(y^{(i)} = c) \log p(y^{(i)} = c | x^{(i)}, \theta)$$

但是，需要特别注意的是，对于 $p(y^{(i)} = c | x^{(i)}, \theta)$ ，softmax回归和logistic回归的计算方式是不同的。

对于选择softmax分类器还是 $k$ 个logistic分类器，取决于所有类别之间是否互斥。所有类别之间明显互斥用softmax分类器，所有类别之间不互斥有交叉的情况下最好用 $k$ 个logistic分类器。

参考资料：

[http://deeplearning.stanford.edu/wiki/index.php/Softmax\\_Regression](http://deeplearning.stanford.edu/wiki/index.php/Softmax_Regression)

<http://blog.csdn.net/u013239871/article/details/51291277>

分类: [Machine Learning](#)



冯瑾元  
关注 - 2  
粉丝 - 1

+加关注

0

0

« 上一篇: [二分类\(Binary Classification\)中用的衡量指标\(Precision, Recall, F1-Measure\)](#)

» 下一篇: [视频字幕识别的一般流程](#)

posted @ 2016-05-03 11:40 冯瑾元 阅读(5783) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问网站首页](#)。

#### 最新IT新闻：

- 苹果市值超万亿背后：苹果汽车月底“小批量送样”
  - 外媒：满帮集团寻求新一轮10亿美元融资 估值将达100亿美元
  - 高通发新品为Google Pixel Watch铺路？Wear OS真的准备好了吗
  - 微博高管电话会议实录：三年三线城市渗透率翻了三倍
  - 有了情商和智商的微软小冰，在商业化上有哪些想象力？
- » 更多新闻...

#### 最新知识库文章：

- 成为一个有目标的学习者
  - 历史转折中的“杭派工程师”
  - 如何提高代码质量？
  - 在腾讯的八年，我的职业思考
  - 为什么我离开了管理岗位
- » 更多知识库文章...

Copyright ©2018 冯瑾元