

python 有趣

python项

python 鼠标手写输入

python环

dnn神经网络

短信发送接口

手写板

原油直播网

基于tensorflow的MNIST手写数字识别（二）

蓝莓对冲基金 图书馆

★★★★★

11483 馆藏 33861

2016-04-09 蓝莓对冲... 摘自 简书 阅 18

分享： 微信 转藏到我的图书馆

- 基于tensorflow的MNIST手写字识别（一）--白话卷积神经网络模型
- 基于tensorflow的MNIST手写数字识别（二）--入门篇

一、本文的意义

因为谷歌官方其实已经写了MNIST入门和深入两篇教程了，那我写这些文章又是为什么呢，只是抄袭？那倒并不是，更准确的说应该是笔记本吧，然后用更通俗的语言来解释，并且补充更多，官方文章中没有详细展开的一些知识点，不过建议与官方文章结合着阅读。

另外是代码部分的改动，官方的demo只提供了验证精确度，我将它改造成了能输入并预测输出结果的代码也就是说是一个从准备待测图片到最终是别的一个完整demo

中文版本：MNIST机器学习入门

http://wiki.jikexueyuan.com/project/tensorflow-zh/tutorials/mnist_beginners.html

demo截图如下，会将放进去的图片预测，然后输出结果，代码说明请看github的readme（最底下）



二、MNIST简介

官网：<http://yann.lecun.com/exdb/mnist/>

这个MNIST数据库是一个手写数字的数据库，它提供了六万的训练集和一万测试集。

它的图片是被规范处理过的，是一张被放在中间部位的28px*28px的灰度图

总共4个文件：

train-images-idx3-ubyte: training set images

train-labels-idx1-ubyte: training set labels

t10k-images-idx3-ubyte: test set images

t10k-labels-idx1-ubyte: test set labels

图片都被转成二进制放到了文件里面，

所以，每一个文件头部几个字节都记录着这些图片的信息，然后才是储存的图片信息

TA的推荐

TA的最新馆藏

永远成功的秘密，就是每天淘汰自己
我们将永生还是灭绝？人工智能很...
我们将永生还是灭绝？人工智能很...
[转] 赞美的大能
他们还不信我要到几时呢？
基督徒的委身【】



推荐阅读

更多

- BetaCat 的前生后世
- 揪出bug！解析调试神经网络的技巧
- 深度学习计算模型中“门函数（Ga...
- 简易的深度学习框架Keras代码解析...
- 国外公司开发新型移动无线网pCell...
- enum的用法
- 再谈：义和团史实（转）
- 是还没有受洗，还没有正式参加某...
- 帧缓存



- | | |
|-----------|-----------|
| 1 美亚保险官网 | 7 企业邮箱申请 |
| 2 美亚保险 | 8 企业邮箱注册 |
| 3 公司邮箱 | 9 led亮化照明 |
| 4 用英语介绍美国 | 10 企业邮箱 |
| 5 钱爸爸理财 | 11 中老年妈妈装 |
| 6 英语学习 | 12 北京口腔医院 |

TRAINING SET LABEL FILE (train-labels-idx1-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	60000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

The labels values are 0 to 9.

TRAINING SET IMAGE FILE (train-images-idx3-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	60000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

每个像素被转成了0-255,0代表着白色, 255代表着黑色。

TEST SET LABEL FILE (t10k-labels-idx1-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	10000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

The labels values are 0 to 9.

TEST SET IMAGE FILE (t10k-images-idx3-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	10000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

每个像素被转成了0-255,0代表着白色，255代表着黑色。

三、tensorflow手写数字识别的大致步骤

- 1、将要识别的图片转为灰度图，并且转化为28*28矩阵（单通道，每个像素范围0-255，0为黑色，255为白色，这一点与MNIST中的正好相反）
- 2、将28*28的矩阵转换成1维矩阵（也就是把第2,3,4,5....行矩阵纷纷接入到第一行的后面）
- 3、用一个1*10的向量代表标签，也就是这个数字到底是几，举个例子e数字1对应的矩阵就是[0,1,0,0,0,0,0,0,0,0]
- 4、softmax回归预测图片是哪个数字的概率
- 5、用交叉熵和梯度下降法训练参数

四、过程讲解

4.1 准备要识别的图片

这个部分其实是比较重要的，因为如果处理不当可能并不一定会有很好的结果，所以按照mnist的标准规范需要将待测图片转为28×28且文字居中的灰度图（其实彩色的也可以，不过就是最后代码需要改一下），目前介绍两种获得待测图片的方法：

- 1、自己用ps或者真的手写一些数字
- 2、将MNIST数据库中的二进制转化成图片，然后用来做测试

ps:图片解析 点击进入

4.2 将待测图片转换为矩阵

如图所示，根据黑色部分的浓淡将其转化成微一个浮点数的数组，（白色0,黑色1）



看到这里，如果你跟我一样不熟悉python，是不是开始方了，没事，其实python很厉害，自带的PIL图片库一句话就可以搞定

```
img=array(Image.open(filename))    //打开然后就被numpy转化了
```

如果是彩色的图片，则需要先将它这样子转换一下（我当初并不知道可以转化，傻不垃圾地自己写了一个转化，所以python还是好好学习啊）

```
im=Image.open("test_num3/8_3.png")
Lim = img=array(im.convert("L"))
```

4.3将矩阵转化为一维矩阵,以及标签的介绍

转化为一维的矩阵其实并不难，用python的reshape就能搞定，还是要讲一下标签的表示方法，这个曾经令队友疑惑不久，直到我把这个数组打印出来

4.3.1标签的来历--有监督学习 和 无监督学习

监督学习：利用一组已知类别的样本调整分类器的参数，使其达到所要求性能的过程，也称为监督训练或有教师学习举个例子，MNIST自带了训练图片和训练标签，每张图片都有一个对应的标签，比如这张图片是1，标签也就是1,用他们训练程序，之后程序也就能识别测试集中的图片了，比如给定一张2的图片，它能预测出他是2

无监督学习：其中很重要的一类叫聚类举个例子，如果MNIST中只有训练图片，没有标签，我们的程序能够根据图片的不同特征，将他们分类，但是并不知道他们具体是几，这个其实就是“聚类”

4.3.2 标签的表示

在这里标签的表示方式有些特殊，它也是使用了一个一维数组，而不是单纯的数字，上面也说了，他是一个一位数组，0表示方法[1,0,0,0,0,0,0,0,0,0],1表示[0,1,0,0,0,0,0,0,0,0],.....，

主要原因其实是这样的，因为softmax回归处理后会生成一个1*10的数组，数组[0,0]的数字表示预测的这张图片是0的概率，[0,1]则表示这张图片表示是1的概率.....以此类推，这个数组表示的就是这张图片是哪个数字的概率（已经归一化），

因此，实际上，概率最大的那个数字就是我们所预测的值。两者对应来看，标准的标签就是表示图片对应数字的概率为100%，而表示其它数字的概率为0，举个例子，0表示[1,0,0,0,0,0,0,0,0,0]，可以理解为它表示0的概率为100%，而表示别的数字的概率为0。

4.4 softmax回归

这是一个分类器，可以认为是Logistic回归的扩展，Logistic大家应该都听说过，就是生物学上的S型曲线，它只能分两类，用0和1表示，这个用来表示答题对错之类只有两种状态的问题时足够了，但是像这里的MNIST要把它分成10类，就必须用softmax来进行分类了。

$P(y=0)=p_0, P(y=1)=p_1, P(y=2)=p_2, \dots, P(y=9)=p_9$ 。这些表示预测为数字i的概率，（跟上面标签的格式正好对应起来了），它们的和为1，即 $\sum(p_i)=1$ 。

tensorflow实现了这个函数，我们直接调用这个softmax函数即可，对于原理，可以参考下面的引文，这里只说一下我们这个MNIST demo要用softmax做什么。

（注：每一个神经元都可以接收来自网络中其他神经元的的一个或多个输入信号，神经元与神经元之间都对应着连接权值，所有的输入加权和决定该神经元是处于激活还是抑制状态。感知器网络的输出只能取值0或1，不具备可导性。而基于敏感度的训练算法要求其输出函数必须处处可导，于是引入了常见的S型可导函数，即在每个神经元的输出之前先经过S型激活函数的处理。）

4.5 交叉熵

通俗一点就是，方差大家都知道吧，用它可以衡量预测值和实际值的相差程度，交叉熵其实也是一样的作用，那为什么不用方差呢，因为看sigmoid函数的图像就会发现，它的两侧几乎就是平的，导致它的方差在大部分情况下很小，这样在训练参数的时候收敛地就会很慢，交叉熵就是用来解决这个问题的，它的公式是

，其中，y是我们预测的概率分布,y'是实际的分布。

$$-\sum y' \log(y):$$

4.6 梯度下降

上面那步也说了，有个交叉熵，根据大伙对方差的理解，值越小，自然就越好，因此我们也要训练使得交叉熵最小的参数，这里梯度下降法就派上用场了，这个解释见上一篇系列文章吧，什么叫训练参数呢，可以想象一下，我们先用实际的值在二位坐标上画一条线，然后我们希望我们预测出来的那些值要尽可能地贴近这条线，我们假设生成我们这条线的公式 $ax+ax^2+bx^3+\dots$ ，我们需要生成这些系数，要求得这些系数，我们就需要各种点代入，然后才能求出，所以其实训练参数跟求参数是个类似的过程。

4.7 预测

训练结束以后我们就可以用这个模型去预测新的图片了，就像我们已经求出来了方程，以后只要随意输入一个x，就能求出对应的y。

5 代码

<https://github.com/wlmnzt/tensorflow-train/tree/master/mnist>

6 参考文章

<http://blog.csdn.net/acdreamers/article/details/44663305softmax回归>

http://wiki.jikexueyuan.com/project/tensorflow-zh/tutorials/mnist_beginners.html MNIST学习入门

<http://blog.csdn.net/u012162613/article/details/44239919交叉熵代价函数>

转藏到我的图书馆
 献花（0）
 分享：
 微信

来自：
 蓝莓对冲基金
 >
 《Tensorflow》
 以文找文 | 举报

上一篇：为什么要用activation functions？
 下一篇：基于tensorflow的MNIST手写数字识别（一）

猜你喜欢

360doc
个人图书馆
[首页](#)
[阅览室](#)
[馆友](#)
[我的图书馆](#)
[搜文章 找馆友](#)

霸业传奇

六房间直播

自我护理能力

角色扮演页游

原油分析师

现货白银

原油交易点差

原油模拟交易

现货贵金属

python趣味编程

类似文章

更多

精选文章

- java 中unsigned类型的转换
- 转帖-VHDL语法学习(1)-初步认识VHDL
- Java数据类型和MySQL数据类型对应一览
- C/C++中各种类型int、long、double、cha...
- CMPP2.0 协议SP端的·NET开发
- C/C 的内存模型
- PCM音频调节测试
- 单片机STM32的LCD12864显示(1)

- 谁能告诉他们有那么一片自由与幸福
- 1次面部刮痧=5次面部护理
- 全身经络排毒方法
- 中国五十六个民族风情
- 没事就用这些词练练你的嘴皮子
- 住房与风水
- 不老女神教你抗衰老秘籍
- 每个任性的女人背后都有一个宠她的男人

六房间直播

汉密尔顿手表

人有三个错误不能

lol竞猜的首页

调查问卷与量表的

《算命字典》举例

霍兰德职业兴趣测

中国居民膳食营养

苏果lol的首页

lol职业联赛2f的首

- 1 贵金属投资服务优选百利好..
- 2 看老股民如何从15万变300万
- 3 北京特价二手房急售,不限购..

- | | |
|----------|----------|
| 1 美亚保险官网 | 4 北京口腔医院 |
| 2 美亚保险 | 5 英语学习 |
| 3 公司邮箱 | 6 企业邮箱注册 |

发表评论：
 请
 登录
 或者
 注册
 后再进行评论
 社交帐号登录：

太阳公元二手房

关闭

台球桌生产厂家

文件管理系统

关闭