

aliceyangxi1987的博客

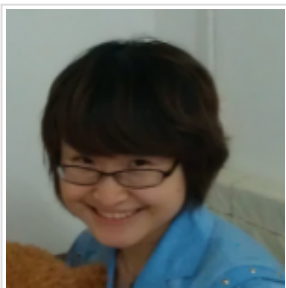
智能时代，做个终身学习者，持续迭代；公众号 - 极客X养成计划；简书

目录视图

摘要视图

RSS 订阅

个人资料



Alice喜爱学习



访问：365100次

积分：6007

等级：BLOC 6

排名：第4375名

原创：252篇 转载：12篇

译文：2篇 评论：159条

异步赠书：Kotlin领衔10本好书 SDCC 2017之区块链技术实战线上峰会 程序员9月书讯 每周荐书：Java Web、Python极客编程（评论送书）

用 TensorFlow 创建自己的 Speech Recognizer

标签：语音识别 人工智能

2017-04-26 11:06

7429人阅读

评论(1)

分类：TensorFlow (19)

版权声明：本文为博主原创文章，未经博主允许不得转载。

目录(?)

[+]

参考资料

源码请点:<https://github.com/llSourcecell/tensorf...>

语音识别无处不在，siri，google，讯飞输入法，讯飞语记，智能家居，车，etc。

每天都在用的，很好奇它是怎么实现的，今天来看看这么便利的东东到底是什么样子呢。

我的小程序

欢迎关注公众号：极客X养成计划

人工智能时代，学点机器学习，一起持续迭代，Run With AI！



文章搜索

博客专栏



机器学习

文章：16篇

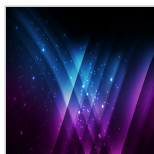
阅读：49243



自然语言处理

文章：14篇

阅读：12296



TensorFlow

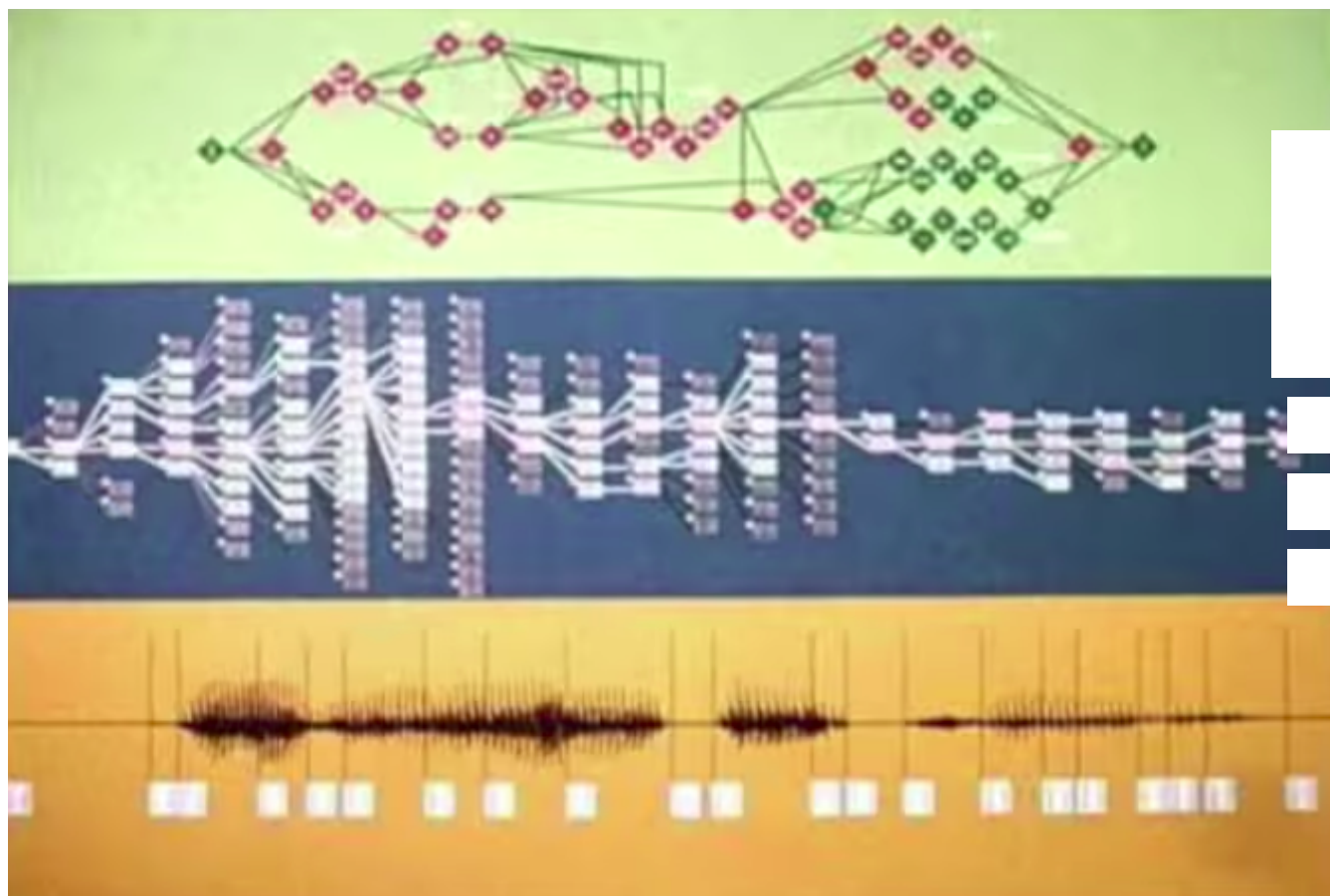
文章：19篇

阅读：116665

进化史

最开始的 speech recognizer 只能识别 0-9 这几个数字，说别的单词是识别不了滴。

后来有一个叫做 DARPA 的梦想家 team 孜孜不倦地研究。



他们用 15000 个节点代表可能的发音，然后用暴力搜索 brute force search 算法来找到节点对应的文字。

后来 IBM 用 Hidden Markov Model 来预测每个点最大概率可能表示的文字。

文章分类

技术博文链接汇总 (1)

- - - 数据结构 - - - (0)

LEETCODE (138)

CTCI (2)

算法 (2)

数据结构 - Tree (23)

数据结构 (4)

- - - Python - - - (0)

coursera-python (1)

Python学习笔记 (4)

- - - 机器学习 - - - (0)

MachineLearning (34)

DeepLearning (14)

Theano (1)

TensorFlow (20)

- - - 人工智能 - - - (0)

神经网络 (8)

计算机视觉 (2)

自然语言处理 (18)

聊天机器人 (3)

- - - R 统计 - - - (0)

R学习笔记 (3)

- - - 移动开发 - - - (0)

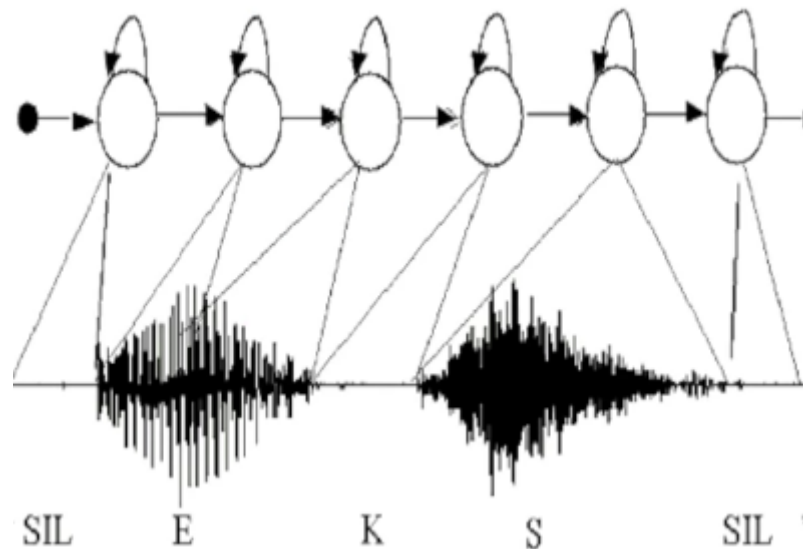
iOS学习笔记 (5)

- - - 积淀 - - - (0)

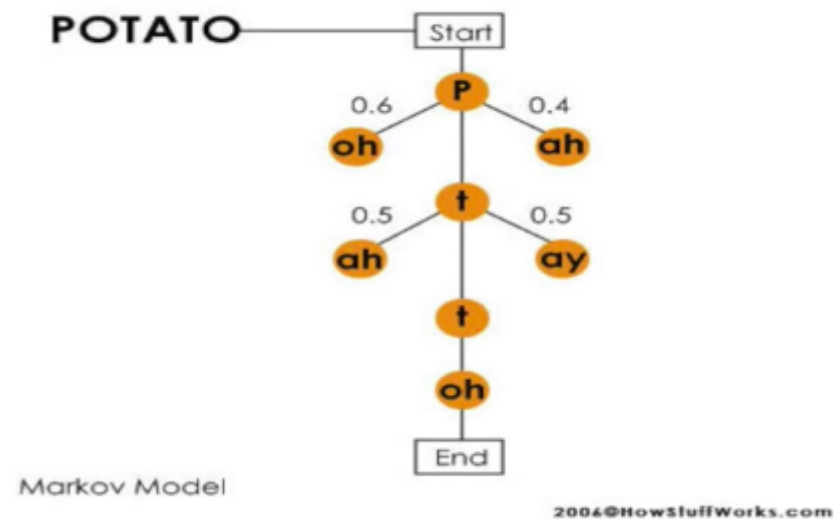
spark (1)

文章存档

2017年08月 (2)



How Speech Recognition Works



再后来人们尝试用 NN 神经网络来做这个任务，但是很长时间没太大进展，直到 深度学习之父 Geoffrey Hinton 研究出个 Deep Learning 模型，语音识别的效果显著提高。

[2017年07月](#) (12)[2017年06月](#) (23)[2017年05月](#) (42)[2017年04月](#) (13)[展开](#)

阅读排行

- [用 TensorFlow 做个聊天](#) (24965)
- [简单粗暴地入门机器学习](#) (16705)
- [如何选择优化器 optimize](#) (14337)
- [一个框架解决几乎所有机](#) (13383)
- [用 TensorFlow 让你的机](#) (13041)
- [自己动手写个聊天机器人](#) (8564)
- [自己写个 Prisma](#) (8042)
- [用 TensorFlow 创建自己](#) (7423)
- [轻松看懂机器学习十大常](#) (6496)
- [【Programming for Ever](#) (6273)

评论排行

- [轻松看懂机器学习十大常](#) (20)
- [从 0 到 1 走进 Kaggle](#) (14)
- [机器学习 人工智能 博文](#) (13)
- [一个框架解决几乎所有机](#) (12)
- [使聊天机器人的对话更有](#) (12)
- [简单粗暴地入门机器学习](#) (9)



Yours ~~

像 Siri , Google 一样 , 现在我们来看看怎样用 TensorFlow 创建自己的 Speech Recognizer , 来识别数字吧。

Steps:

- 导入库
- 定义参数

- [用 TensorFlow 做个聊天机器人](#) (9)
- [推荐系统](#) (9)
- [TensorFlow-7-TensorBoard](#) (4)
- [用 TensorFlow 让你的机器学习](#) (4)

推荐文章

- * [CSDN日报20170828——《4个方法快速打造你的阅读清单》](#)
- * [Android检查更新下载安装](#)
- * [动手打造史上最简单的Recycleview 侧滑菜单](#)
- * [TCP网络通讯如何解决分包粘包问题](#)
- * [SDCC 2017之区块链技术实战线上峰会](#)
- * [快速集成一个视频直播功能](#)

最新评论

[TensorFlow-7-TensorBoard Embedding](#)
kinbo88: "" fetching sprite images parsing metadata ""这个原因其...

推荐系统

[Jack床长](#): 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

用ARIMA模型做需求预测

[Jack床长](#): 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

了解 Sklearn 的数据集

[Jack床长](#): 我最近在写一系列的人工智能教程，通俗易懂，无需...

- 导入数据
- 建立模型
- 训练模型并预测

1. 导入库

需要用到 tflearn，这是建立在 TensorFlow 上的高级的库，可以很方便地建立网络。

还会用到辅助的类 speech_data，用来下载数据并且做一些预处理。

```
1 from __future__ import division, print_function, absolute_import
2 import tflearn
3 import speech_data
4 import tensorflow as tf
```

2. 定义参数

很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

Sklearn 快速入门

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

决策树的python实现

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

通过一个kaggle实例学习解决机器学习问题

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

机器学习算法应用中常用技巧-2

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

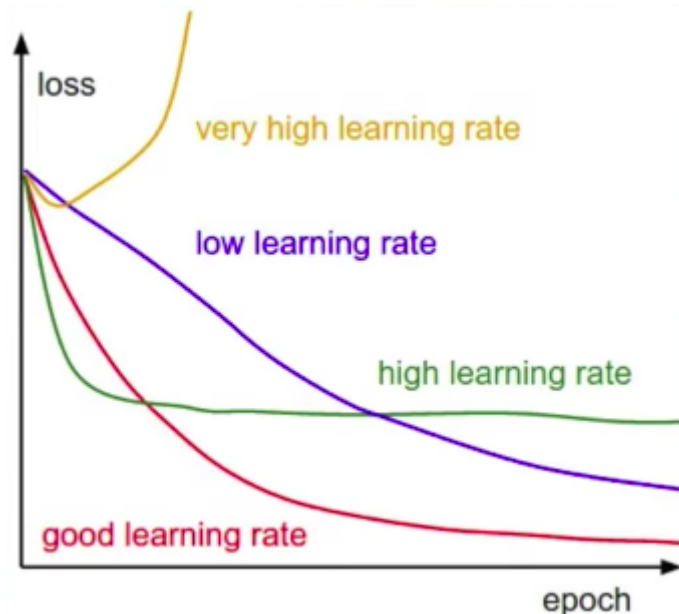
机器学习算法应用中常用技巧-1

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

特征工程怎么做

Jack床长: 我最近在写一系列的人工智能教程，通俗易懂，无需很高的数学基础，教程也力求风趣幽默，倡导快乐学习，欢迎...

learning rate 是在更新权重的时候用，太高可以很快，但是loss大，太低较准但是很慢。



```
1 learning_rate = 0.0001
2 training_iters = 300000 # steps
3 batch_size = 64
4
5 width = 20 # mfcc features
6 height = 80 # (max) length of utterance
7 classes = 10 # digits
```

3. 导入数据

用 `speech_data.mfcc_batch_generator` 获取语音数据并处理成批次，然后创建 training 和 testing 数据。

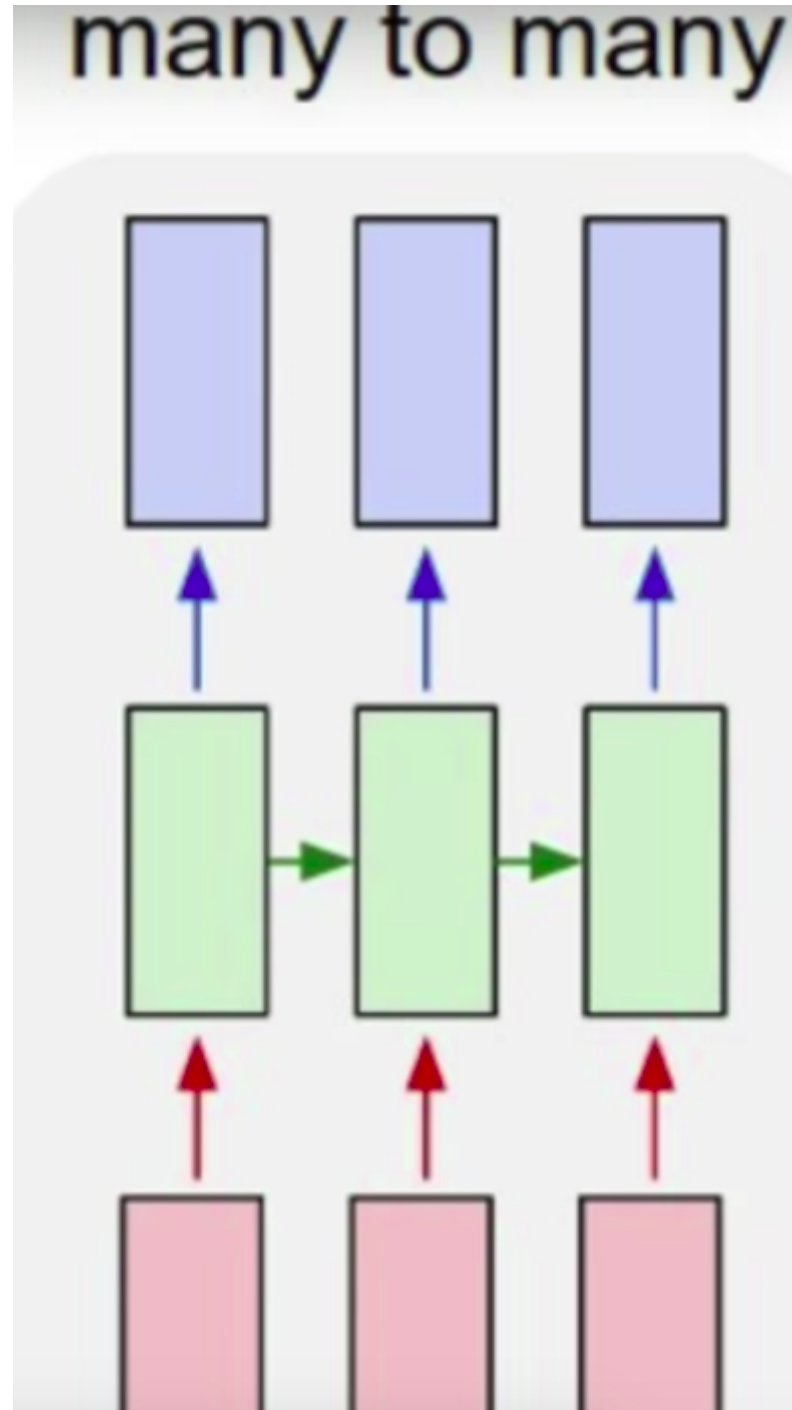
```
1 batch = word_batch = speech_data.mfcc_batch_generator(batch_size)
2 X, Y = next(batch)
3 trainX, trainY = X, Y
4 testX, testY = X, Y #overfit for now
```

4. 建立模型

接下来，用什么模型呢？

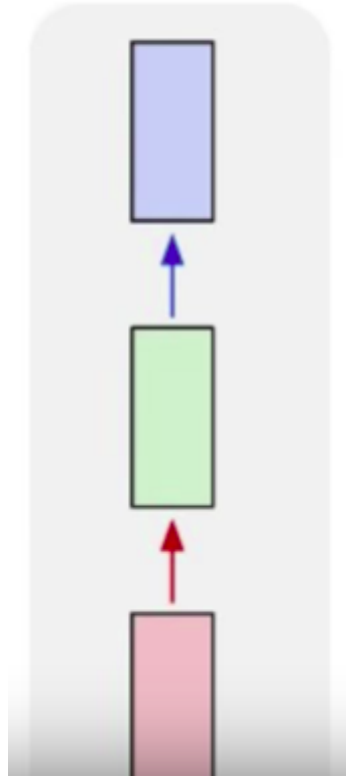
speech recognition 是个 many to many 的问题。

eg , speech recognition



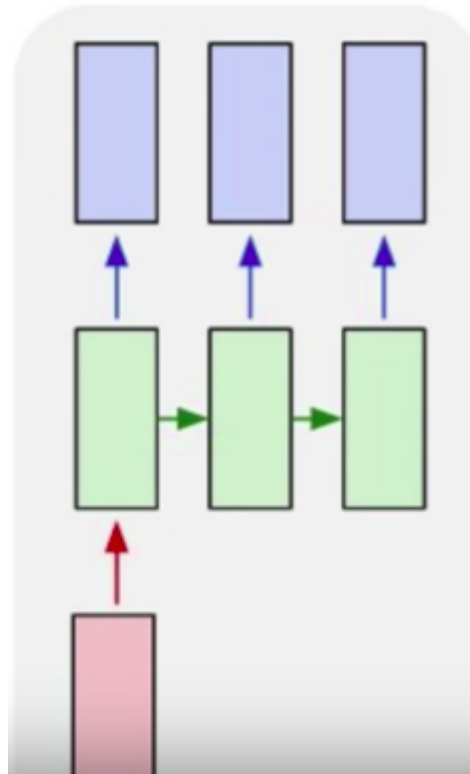
eg , image classification

one to one



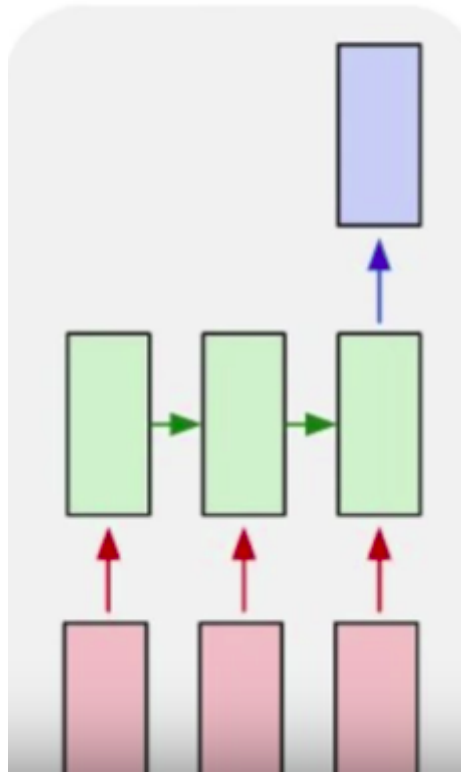
eg , image caption

one to many

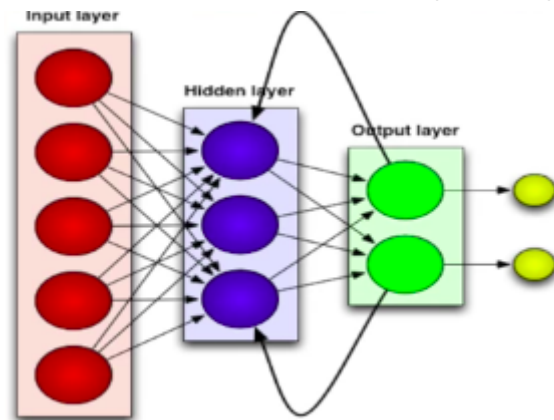


eg , sentiment analysis

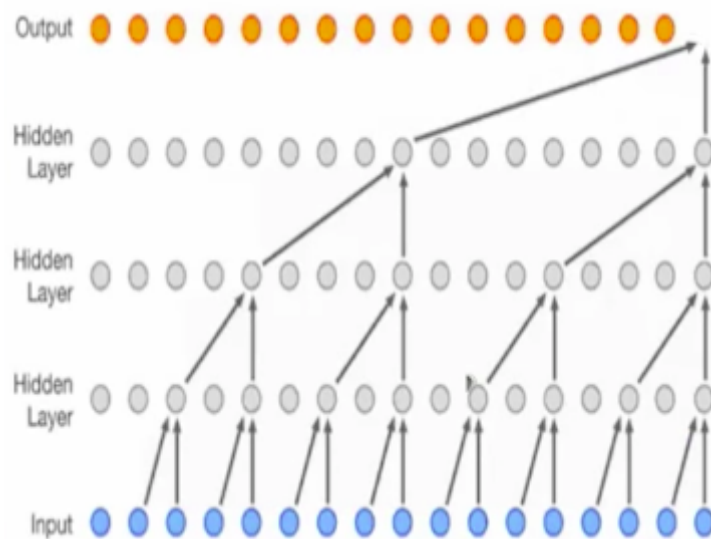
many to one



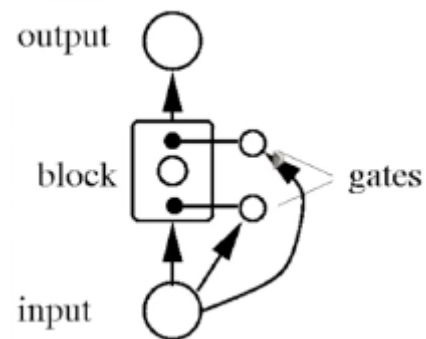
所以我们用 Recurrent NN 。



通常的 RNN ，它的输出结果是受整个网络的影响的。



而 LSTM 比 RNN 好的地方是，它能记住并且控制影响的点。所以这里我们用 LSTM。



每一层到底需要多少个神经元是没有规定的，太少了的话预测效果不好，太多了会 overfitting，这普遍的 128.

为了减轻过拟合的影响，我们用 dropout，它可以随机地关闭一些神经元，这样网络就被迫选择其进而生成想对 generalized 模型。

接下来建立一个 fully connected 的层，它可以使前一层的所有节点都连接过来，输出 10 类，因为数字是 0-9，激活函数用 softmax，它可以把数字变换成概率。

最后用个 regression 层来输出唯一的类别，用 adam 优化器来使 cross entropy 损失达到最小。

```
1 # Network building
2 net = tflearn.input_data([None, width, height])
3 net = tflearn.lstm(net, 128, dropout=0.8)
4 net = tflearn.fully_connected(net, classes, activation='softmax')
5 net = tflearn.regression(net, optimizer='adam', learning_rate=learning_rate, loss='categorical_crossentropy')
```

5. 训练模型并预测

然后用 tflearn.DNN 函数来初始化一下模型，接下来就可以训练并预测，最后再保存训练好的模型。

```
1 # Training
2 ### add this "fix" for tensorflow version errors
3 col = tf.get_collection(tf.GraphKeys.TRAINABLE_VARIABLES)
4 for x in col:
5     tf.add_to_collection(tf.GraphKeys.VARIABLES, x )
6
7 model = tflearn.DNN(net, tensorboard_verbose=0)
8
9 while 1: #training_iters
10     model.fit(trainX, trainY, n_epoch=10, validation_set=(testX, testY), show_metric=True,
11             batch_size=batch_size)
12     _y=model.predict(X)
13     model.save("tflearn.lstm.model")
14     print(_y)
15     print(y)
```

模型训练需要一段时间，一边碎觉一边等着模型出炉吧。

推荐阅读

[历史技术博文链接汇总](#)

也许可以找到你想要的

顶 踩
0 0

上一篇 [自己写个 Prisma](#)

下一篇 [用 TensorFlow 让你的机器人唱首原创给你听](#)

相关文章推荐

- TensorFlow-4: tf.contrib.learn 快速入门
- Android Studio如何格式化XML代码顺序。
- 【免费】深入理解Docker内部原理及网络配置--王...
- Android入门实战
- AndroidStudio-快捷键-格式化代码
- Android Studio设置代码风格
- SDCC 2017之区块链技术实战线上峰会--蔡栋
- 5天搞定深度学习框架Caffe
- mac android studio的代码格式化整理快捷键
- android studio 如何进行格式化代码 快捷键必备
- php零基础到项目实战
- Android Studio中格式化代码
- Android Studio格式化代码设置和代码风格设置、...
- 修改AndroidStudio虚拟机使用的内存大小
- C语言及程序设计入门指导
- Android studio 设置代码格式化不换行

查看评论

1楼 [Flody88](#) 2017-04-27 18:49发表



看了一遍，不错。。。

您还没有登录,请[登录](#)或[注册](#)

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

[公司简介](#) | [招贤纳士](#) | [广告服务](#) | [联系方式](#) | [版权声明](#) | [法律顾问](#) | [问题报告](#) | [合作伙伴](#) | [论坛反馈](#)

网站客服 杂志客服 微博客服 webmaster@csdn.net 400-660-0108 | 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 | 江苏乐知网络技术有限公司

京 ICP 证 09002463 号 | Copyright © 1999-2017, CSDN.NET, All Rights Reserved

