

# 机器学习&tensorflow

Xuyadan - 2017年7月24日

## 机器学习

- ❖ 一个我觉得讲得非常清楚的机器学习入门指南，适合对机器学习一无所知的人

<http://blog.jobbole.com/67616/>

- ❖ Andrew ng教授的《machine learning》（网易公开课或者coursera）  
我只看了前两节，看到第二节后面对算法的证明就看不下去了。看完他的视频印象最深刻的是gradient descent，“I am going to take a small step in this direction of steepest descent”。等有点文化了再继续完成后面的视频。
- ❖ 跟着 莫烦python 学习了下tensorflow的基本Session，Variables，placeholder，Activation Function和def add\_player()

## Tensorflow

### 1. 安装tensorflow

- ❖ upgrade pip

pip is already installed if you're using Python 2 >=2.7.9 or Python 3 >=3.4 binaries downloaded from [python.org](https://python.org), but you'll need to [upgrade pip](#).

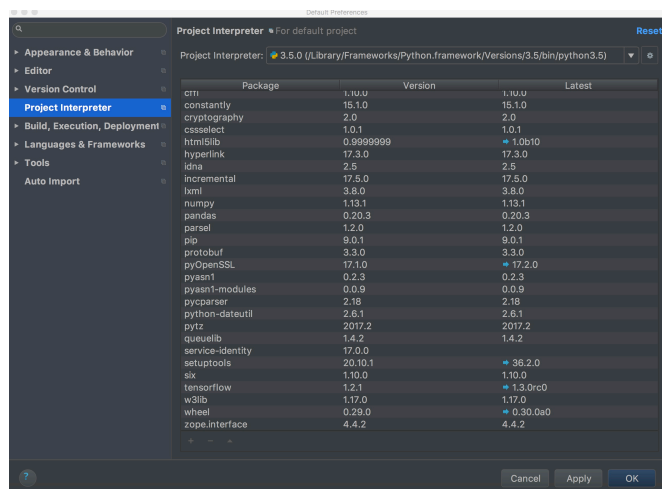
终端中输入：

```
$ pip3 install --upgrade pip
```

- ❖ 安装tensorflow(最简单的pip安装)

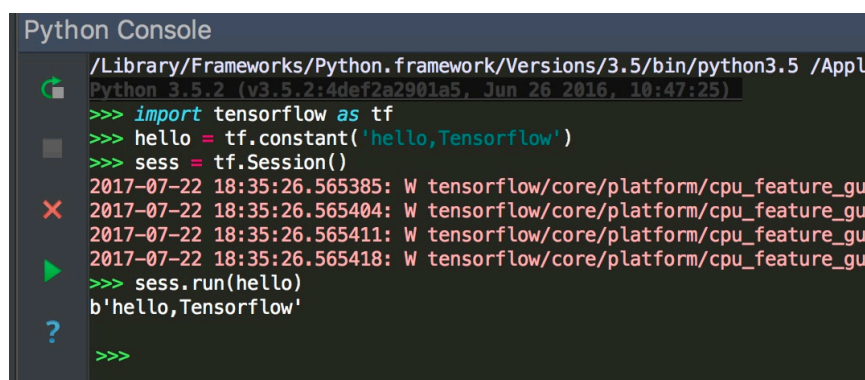
```
zhouyangdeMacBook-Pro:~ zhouyang$ pip3 install tensorflow
Collecting tensorflow
  Downloading tensorflow-1.2.1-cp35-cp35m-macosx_10_11_x86_64.whl (33.6MB)
    100% |#####| 33.6MB 38kB/s
Collecting wheel>=0.26 (from tensorflow)
  Downloading wheel-0.29.0-py2.py3-none-any.whl (66kB)
    100% |#####| 71kB 1.0MB/s
Collecting protobuf>=3.2.0 (from tensorflow)
  Downloading protobuf-3.3.0.tar.gz (271kB)
    100% |#####| 276kB 515kB/s
Collecting html5lib==0.9999999 (from tensorflow)
  Downloading html5lib-0.9999999.tar.gz (889kB)
```

- ❖ python3.6不支持tensorflow!在pycharm始终显示不出来
- ❖ 在pycharm中import tensorflow as tf，显示no module?



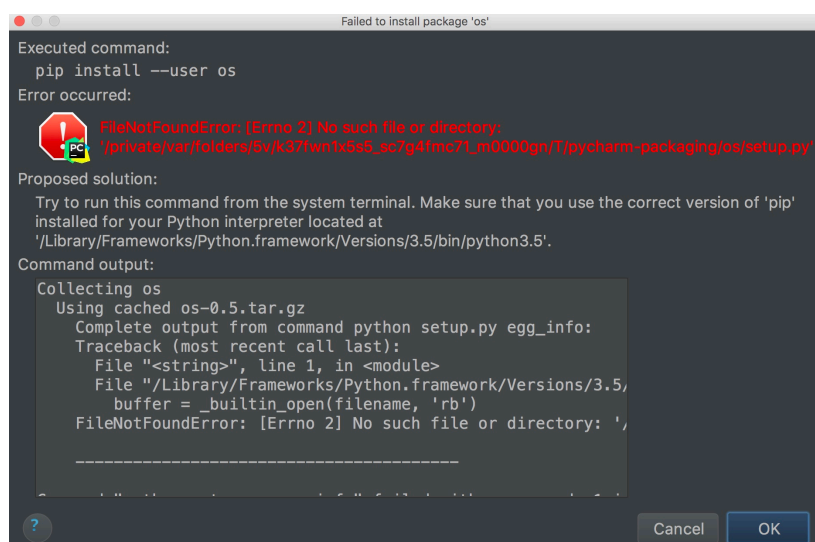
那么需要检查pycharm中选择的default setting中的interpreter是不是选的对  
对应下载的python3.5的版本

- ❖ 运行tensorflow



## 2.构建DNN对iris数据集分类

- ❖ 导入os urllib模块发生错误



---

原因：所导入的模块与当前的python3.5版本不符

解决方法：换一下模块的版本，urllib5,os0

### ❖ 无\_\_future\_\_模块

```
2 from __future__ import absolute_import
3 from __future__ import division
4 from __future__ import print_function
```

该模块是在旧版本中引用新版本的功能，故只有旧版python才有此模块，新版无需安装

### ❖ 包含iris数据集可以下载，讲解清楚的一个博客地址

<http://blog.csdn.net/mebiuw/article/details/53222000>

- 按照它上面的运行出现了错误 not found file iris\_training.csv

solution：将下载的iris\_training.csv&iris\_test.csv放到.py的目录下

- 无效命令load\_csv

solution:换成load\_csv\_with\_header,在target\_dtype后面加上features\_dtype变量（features\_dtype=np.float32）

### ❖ 基本步骤

- 1.从原始数据集的CSV里面读取数据，并且加载到Tensorflow当中

- 2.构建一个基于神经网络的分类器

- 3.使用训练数据进行模型训练

- 4.使用测试数据进行模型评估

- 5.使用训练好的模型对新的数据进行分类

```
import tensorflow as tf
```

```
import numpy as np
```

```
#设定数据集的位置
```

```
IRIS_TRAINING = "iris_training.csv"
```

```
IRIS_TEST = "iris_test.csv"
```

```
#使用tensorflow内置的方法进行数据加载，target_dtype是最终的label的类型，这里只有012三个取值，所以用int
```

```
training_set = tf.contrib.learn.datasets.base.load_csv_with_header(filename =  
IRIS_TRAINING,target_dtype = np.int,features_dtype=np.float32)
```

```
test_set = tf.contrib.learn.datasets.base.load_csv_with_header(filename =  
IRIS_TEST,target_dtype = np.int,features_dtype=np.float32)
```

```
# 每行数据4个特征，都是real_value的
feature_columns = [tf.contrib.layers.real_valued_column("",dimension = 4)]

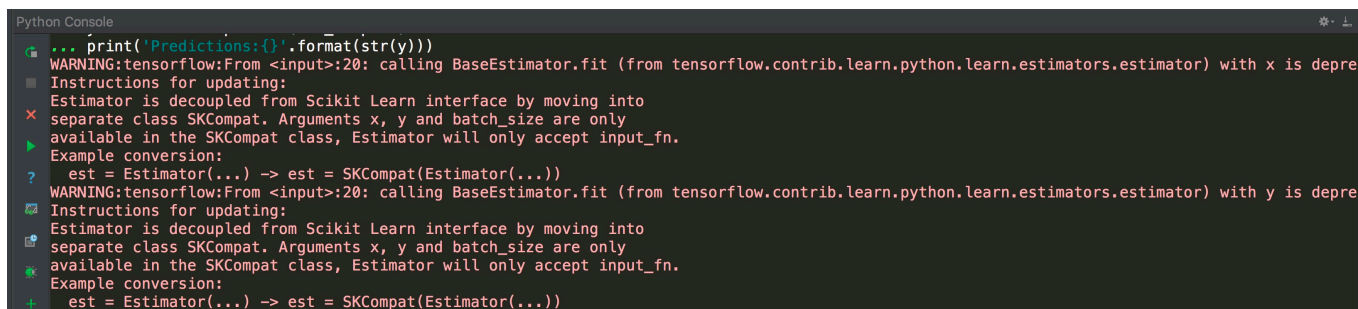
# 构建一个DNN分类器，3层，其中每个隐含层的节点数量分别为10，20，
# 10，目标的分类3个，并且指定了保存位置
classifier =
tf.contrib.learn.DNNClassifier(feature_columns=feature_columns,hidden_un
its=[10,20,10],n_classes=3,model_dir="/tmp/iris_model")

# 指定数据，以及训练的步数
classifier.fit(x=training_set.data,y=training_set.target,steps=2000)
accuracy_score = classifier.evaluate(x=test_set.data,y=test_set.target)
["accuracy"]
print('Accuracy:{0:f}'.format(accuracy_score))

new_samples = np.array([[6.4,3.2,4.5,1.5],[5.8,3.1,5.0,1.7]],dtype=np.float32)
y = classifier.predict(new_samples)
print('Predictions:{}'.format(str(y)))
```

## ❖ 结果

出现了warning，但是可以运行



```
Python Console
... print('Predictions:{}'.format(str(y)))
WARNING:tensorflow:From <input>:20: calling BaseEstimator.fit (from tensorflow.contrib.learn.python.learn.estimators.estimator) with x is depre
Instructions for updating:
Estimator is decoupled from Scikit Learn interface by moving into
separate class SKCompat. Arguments x, y and batch_size are only
available in the SKCompat class, Estimator will only accept input_fn.
Example conversion:
? est = Estimator(...) -> est = SKCompat(Estimator(...))
WARNING:tensorflow:From <input>:20: calling BaseEstimator.fit (from tensorflow.contrib.learn.python.learn.estimators.estimator) with y is depre
Instructions for updating:
Estimator is decoupled from Scikit Learn interface by moving into
separate class SKCompat. Arguments x, y and batch_size are only
available in the SKCompat class, Estimator will only accept input_fn.
Example conversion:
+ est = Estimator(...) -> est = SKCompat(Estimator(...))
```

warning还在找原因

## 下周计划

- 看机器学习的视频
- 通过阅读的论文了解视频预测的基本概念，运行的步骤
- 用tensorflow 模仿一些简单预测的案例