

任务一-环境搭建

任务一-环境搭建

环境配置

基本库安装

安装TensorFlow（根据系统不同，命令略有区别）

安装 `git`

`python` 学习

`tensorflow` 学习

理论学习

环境配置

- 系统：ubuntu（推荐）/mac os/windows，另外如果是mac请不要强行安装三系统
- 环境：python 3.6（python已经更新到3.7，但TensorFlow还未支持到3.7，此外mac自带python2.7，不要卸载）
- 更换ubuntu apt源：系统设置 -> 软件和更新 -> 选择阿里云源（mirrors.aliyun）
- 根据是否有NVIDIA GPU 以及是否支持CUDA 安装相应驱动
`https://www.nvidia.cn/object/cuda_learn_products_cn_old.html`，注意版本选择最新版本不一定支持（显卡驱动 cuda cudnn）
可参考 `https://www.2cto.com/kf/201612/578337.html`，但先不要安装 `tensorflow`

基本库安装

`pip` 安装很慢可以更换国内源

- Windows用户（不推荐）：
 - 添加python和pip到环境变量Path
参考教程：`http://www.cnblogs.com/yuanzm/p/4089856.html`
 - Windows下可以使用CMD完成基本库安装
开始->系统->命令提示符->管理员权限打开，环境变量配置好后可以使用 `pip` 来完成基本库安装，先更新 `pip`

```
pip install --upgrade pip
```

使用 `pip` 安装 `numpy`、`pandas`、`matplotlib`、`jupyter` (IDE)

```
pip install ...
```

- Ubuntu用户：

- 安装并升级 pip：先使用 apt 更新 pip

```
sudo apt-get install python-pip python-dev build-essential
sudo pip install --upgrade pip
```

- 使用pip来安装 numpy 、 pandas 、 matplotlib 、 jupyter (IDE)

jupyter IDE的使用技巧：库函数参数、代码注释查看（shift+tab）

```
sudo pip install ...
```

- mac用户：

与Ubuntu基本相同，需要安装 brew 代替 apt，在 terminal 中输入：

```
ruby -e "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/master/install)"
```

- 安装相应库（同上），注意如果同时安装了Python2和Python3，则pip默认给Python2用，pip3指定给Python3用。

```
sudo pip3 install --upgrade pip3
sudo pip3 install ...
```

安装TensorFlow（根据系统不同，命令略有区别

- 不带GPU版本：

```
sudo pip install tensorflow
```

- 带GPU版本（要先安装好驱动）

```
sudo pip install tensorflow-gpu
```

安装git

根据不同系统，有不同的安装方式，具体参考：<https://git-scm.com/book/zh/v1/起步-安装-Git>

python 学习

- 学习资料可以看“廖雪峰 python3”

<https://www.liaoxuefeng.com/wiki/0014316089557264a6b348958f449949df42a6d3a2e542c000>

- 这期需要看的知识点：
 - 数据类型：整型、浮点数、字符串
 - 语句类型：条件语句、循环语句
 - 运算符：%、+、/、//
 - 字符串的索引、切片和编码
 - 列表和元组，列表和元组的索引和切片
 - 函数的定义、print 的格式化输出
 - 涉及的关键字：for、range、if else elif、while、is、and、or、True、False、enumerate、len、append、extend、in、print
- 练手题目：

现有一个字符串列表，自定义一个函数对该列表中的字符串元素按字符串长度进行排序，最后返回排序后的列表

tensorflow学习

- 阅读 tensorflow 白皮书（译），有能力的同学可以看一下官方的英文版 翻译版链接：<http://blog.csdn.net/xiaopihaierletian/article/details/63750807> 官方版：<http://download.tensorflow.org/paper/whitepaper2015.pdf>

也可以看群上的pdf TensorFlow 官方文档中文版 - v1.2

体会一些 变量、操作（op）、计算图、会话、张量（tensor）的概念。

（一维：线，二维：面，三维：体、N维：张量）
- 对比白皮书，学习 tf.placeholder、tf.Variable、tf.Session 的使用。

学习资料：http://www.tensorfly.cn/tfdoc/get_started/basic_usage.html
- 练手题目：使用 tensorflow 实现一个矩阵运算： $f(XW + b)$ ，其中 $f(x) = x^T x$ ，其中 W （矩阵）和 b （标量）为变量， X 为一个占位符。各个矩阵的大小为下： $X: 3 \times 8$ ； $W: 8 \times 3$ ， b 为标量。（提示：`tf.transpose`）

理论学习

- 数据集的划分方法：留出法，自助法，交叉验证法
- 特征与特征工程：
 - 两个思考：什么是样本的特征？为什么需要特征工程？
 - 参考链接：<http://www.360doc7.net/wxarticlenew/588303798.html>
- 机器学习中模型评估方法：精度，TP TN FP FN precision recall
- 线性回归模型：

学习资料：http://fangs.in/post/thinkstats/ml_0_linearregression/