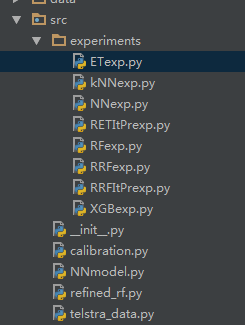
1. **Codes structure**



* 1. **Environment Preparation**

1. 安装数据挖掘需要的科学工具包numpy,scipy,
2. 安装scipy遇到的问题

numpy.distutils.system\_info.NotFoundError: no lapack/blas resources found

Reasons: 安装numpy后安装scipy失败（报错：

numpy.distutils.system\_info.NotFoundError）一般是缺少一些系统库，需要安装：libopenblas-dev、liblapack-dev、libatlas-dev、libblas-dev

<http://blog.csdn.net/lingfeng892/article/details/50425547>

1. 安装sacred及其使用方法

<http://sacred.readthedocs.io/en/latest/quickstart.html>

1. 运行sacred定义的实验时，遇到问题：

File "D:\Python27\lib\site-packages\pymongo\topology.py", line 210, in select\_server

address))

File "D:\Python27\lib\site-packages\pymongo\topology.py", line 186, in select\_servers

self.\_error\_message(selector))

pymongo.errors.ServerSelectionTimeoutError: localhost:27017: [Errno 10061]

应该是mongodb没有启动，需要先按照mongodb

1. 按默认路径安装mongodb，遇到链接失败错误如下：

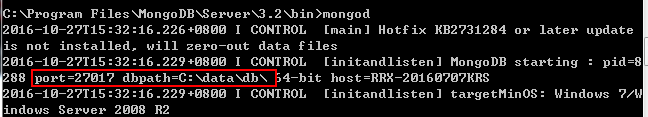


原因是没有指定data存放的文件夹路径

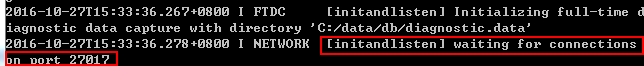
<https://segmentfault.com/a/1190000002744306>

mongod --dbpath {filedir}\data

但我通过指令指定出错，因此找到系统默认的路径，建立一样的文件夹路径



连接成功



1. Mongodb问题解决后出现新问题

File "pandas/parser.pyx", line 373, in pandas.parser.TextReader.\_\_cinit\_\_ (pandas\parser.c:4019)

File "pandas/parser.pyx", line 667, in pandas.parser.TextReader.\_setup\_parser\_source (pandas\parser.c:7989)

IOError: Initializing from file failed

原因：找不到要读入目标文件夹，修改文件路径即可解决

1. **Articles**
   1. **[Global refinement of random forest](https://github.com/gereleth/kaggle-telstra/blob/master/Global%20refinement%20of%20random%20forest.ipynb)**