

**这部分是关于 DISTSDET 的 4.3 的代码，介绍如下：**

**第一步：**

**query\_process.py:** 查询 arena 数据服务进程的代码；

**query\_process\_o.py:** 查询本地主机服务进程的代码（可供测试用，无太大用处）。

**第二步：**

**process2vector.py:** 将第一步查询到的服务进程名生成对应的词向量，保存在 process\_embedding 目录下的相应文件中。

所用的词向量文件放置在 word2vec\_model 目录下，文件为 cc.en.300.bin。

**第三步、第四步：**

**process\_cluster.py:** process\_embedding 目录下的所有文件中的进程名词向量聚类，然后通过计算生成全局词向量，保存在 global\_process\_embedding 目录下的 global\_process\_embedding.csv 文件中。

**第五步：**

**distribute.py:** 全局词向量文件下发更新各主机的词向量文件。