

目前我已完成的工作：

- 1, 写了两个软件（只要修改配置文件就可以运行），分别对应于有控制组的查找和无控制组的查找。
- 2, 多进程优化，根据实验测得的结果，至少有10倍的提升，如果机器的核数更多，则优化的结果更好，为大数据集的计算提供了方便。
- 3, 根据软件测得的数据显示，DNB子网的存在很明显，在两个数据集中都显示出较大的突变。
- 4, 人种的数据，正在进行，在数据解析上遇到困难。

下面是具体结果：

1.对于 liver_case_data.txt 数据集

具体的数据集描述：

10729 个特征
5个 sample
5个 采样时间点

采用单进程 运行时间为： 2210.847s

运行多进程 运行时间为： 244.355s

找到的 DNB 分子在五个不同的采样时间点的 CI 值为：

0.35883081250218973, 1.7131192380251437, 1.5796586827518977,
95.860459889094926, 9.8820995875867403

2 对于 GSE64538_case_data.txt 数据集

具体的数据集描述

30246 个特征
3个 sample
4个 采样时间点

多进程运行时间 using time: 2458.0296s

找到的 DNB 在五个不同的采样时间点的 CI 值为：

0.75463542729576127, 116.21874716665505, 0.61741229946992371, 1.2609393741583383