

周报

2020 年 4 月 13 日

1 上周任务

- 对非隐喻词进行统计；
- 看论文；

1.1 任务详细

- 预测词组成（示例如下表）：
 - 取预测列表前 50 词；
 - 取隐喻词预测位置前 50 词；
 - 在 1-10000，10001-20000，20001-30000 每段中随机抽取 50 词；
- 预测词处理：
 - 通过词表过滤掉不正确的词；
- 实验结果：
 - 从当前的抽样结果中，看不出来有什么规律。。。

2 本周计划

- 改变抽样方法，再找一下规律；
- 看论文；

表 1: 预测词示例

| Position | Predictive word |
|-----------------|---|
| TOP50 | handed,deaf,gen,mi,lane,grave,adverse,notorious,me,divers, skater,overcame,converge,fin,pile,bearing,daring |
| MTOP50 | muscle,finalist,westphalia,phillips,brook,compute,engineering,mouse, paid,strategic,courtroom,danish,cypress,diver,integration,flaw, tri,doris,intent,biographer,antiquity,arctic |
| RAND1_50 | gliding,tactical,disqualification,creeping,postage,negligence,carrying, howie,regain,scorer,bleed,zeta,hiding,choreography,feeder,xp,loyal, spanned,counterattack,werner,alternating,resort,swollen,elise,smashing,direct |
| RAND2_50 | papal,cautiously,trustee,reservoir,fitting,competence,sunken,creamy, executive,despite,chico,romanian,hostile,roam,recover,reckon,ham,elder, airlift,panties,stallion |
| RAND3_50 | ina,startup,euclidean,geek,discuss,exception,earning,authentication, archipelago,austin,interference,casablanca,finest,warlock,harbour,cannes,theme |

句子: It seemed to me an enormously powerful tool if you could [MASK] out a gene responsible for a virus or cancer.

隐喻词: knock; 列表位置: 4509;

TOP50: 预测列表前50; MTOP50: 隐喻词预测位置之前50个预测词;

RAND1_50: 在 1-10000 中随机获取50个预测词;

RAND2_50: 在 10001-20000 中随机获取50个预测词;

RAND3_50: 在 20001-30000 中随机获取50个预测词;