

EL流程

1. 根据股票名称库，去百科爬取信息，建立知识库。【产出kb_data.json】 【产出id_embedding.pkl】

a. 包含以下内容：

i. kb-data.json

```
1 {"entity_name":"","entity_id":"","ali":"","abstract":"","else}
```

ii. id_embedding.pkl

```
1 {"entity_id":[list,1*768]}
```

2. 新闻库，【已知new.json】

1 新闻库包含：

```
2 {"text_id":"","title":"","text":"","stocks_name":""}
```

将新闻库的text输入到NER中，找到ORG(机构名)，会得到

1 n个stocks_name: setSN: {A B C D}

2 m个ORG结果: setORG: {a b c d e f g 等}

3 分别计算A B C D 与SetORG的分数:

4 得到一个表格:

5

	A	B	C	D
--	---	---	---	---

a	1	2	4	3
---	---	---	---	---

b	4	3	2	1
---	---	---	---	---

c	2	1	3	4
---	---	---	---	---

d	3	4	1	0
---	---	---	---	---

e	*	*	*	*
---	---	---	---	---

f	0	*	*	*
---	---	---	---	---

g	*	*	*	*
---	---	---	---	---

14 等

15

