



基於Word2Vec的情緒分析規則探勘

Low-Power 6-Gbit/s Spread Spectrum Clock Generator

國立臺北大學 電機工程學系

指導老師：吳柏翰 老師 林嘉淦 老師 專題生：黃子育

研究目標

- 利用Jieba進行斷詞
- 利用word2vec轉換詞彙成詞向量
- 利用mobile01的十萬則評論進行情緒分析
- 利用台大開放的情緒字典等等資源進行情緒分析

實作結果

```
請問suv第三排的舒適度及空間如何?
為什麼擁有40年經驗的車廠，確輸給不到10年的Tesla?
Luxgen7 suv豪華型和尊爵型各有哪些配備?
Luxgen7 suv豪華型和尊爵型各有哪些配備?
Luxgen7 suv豪華型和尊爵型各有哪些配備?
比原廠更靚的suv運動版大包開箱
比原廠更靚的suv運動版大包開箱
Luxgen7 suv豪華型和尊爵型各有哪些配備?
為什麼擁有40年經驗的車廠，確輸給不到10年的Tesla?
LUXGEN 5 RF-Design Sport Styling Package 空力套件版
luxgen 5 sedan 新年繼續改
luxgen 5 sedan 新年繼續改
luxgen 5 sedan 新年繼續改
luxgen 5 sedan 新年繼續改
luxgen 5 sedan 新年繼續改
luxgen 5 sedan 新年繼續改
```

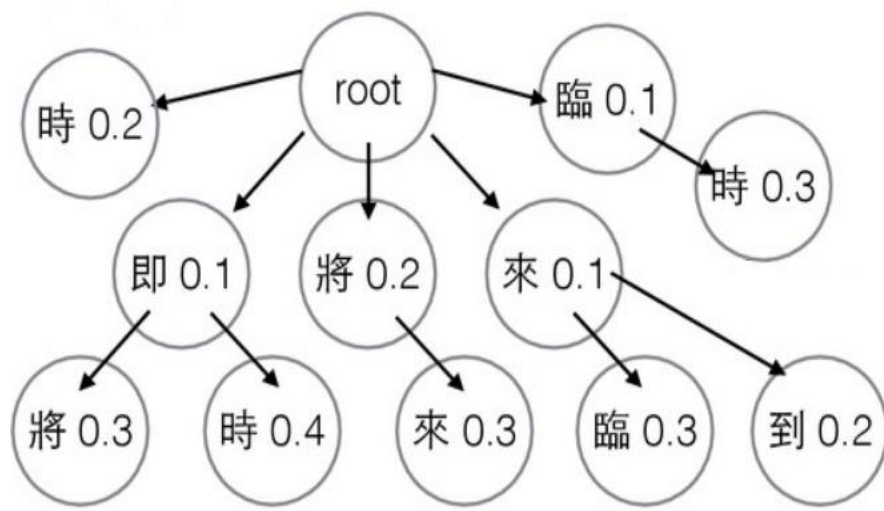
```
INFO: PROGRESS: at 3.32% examples, 75793 words/s, in_qs
INFO: PROGRESS: at 6.37% examples, 76183 words/s, in_qs
INFO: PROGRESS: at 9.84% examples, 74121 words/s, in_qs
INFO: PROGRESS: at 12.74% examples, 69242 words/s, in_q
INFO: PROGRESS: at 15.97% examples, 69716 words/s, in_q
INFO: PROGRESS: at 19.71% examples, 71789 words/s, in_q
INFO: PROGRESS: at 23.79% examples, 74526 words/s, in_q
INFO: PROGRESS: at 28.08% examples, 77781 words/s, in_q
INFO: PROGRESS: at 32.90% examples, 80504 words/s, in_q
INFO: PROGRESS: at 37.61% examples, 82968 words/s, in_q
INFO: PROGRESS: at 42.49% examples, 84927 words/s, in_q
```

```
>>> result=model.most_similar("保固")
>>> for e in result:
... print e[0], e[1]
...
快過 0.924601256847
零件 0.733630597591
不會 0.691097795963
單單 0.689484536648
不二價 0.688739895821
趕快 0.686941325665
很爛 0.68683642149
只是 0.683317780495
而是 0.680120646954
說好 0.675670862198
```

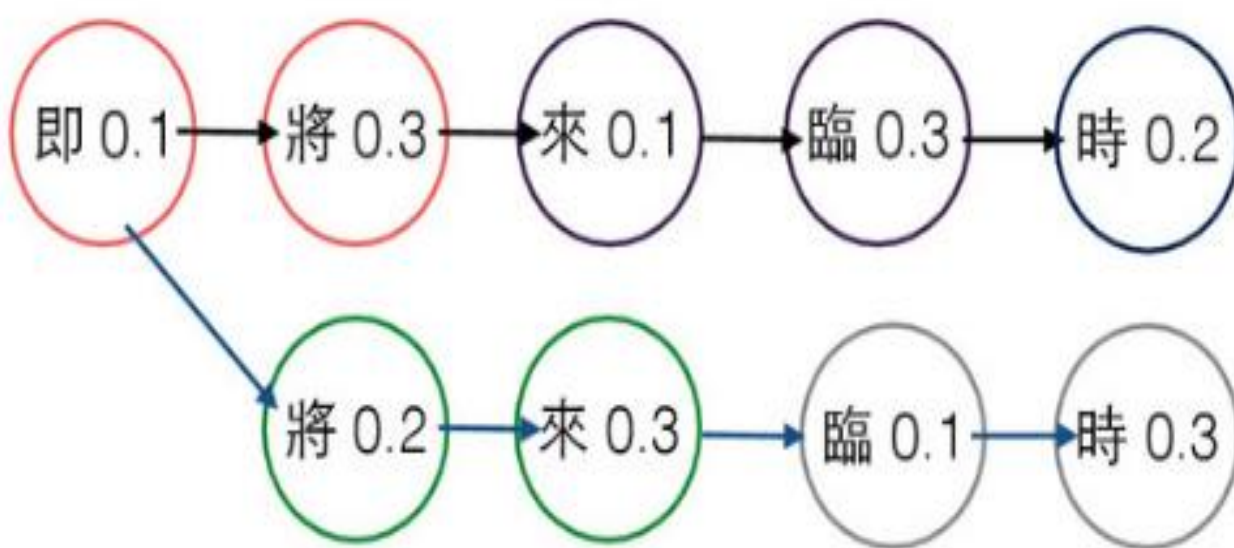
```
>>> result=model.most_similar("英俊")
>>> for e in result:
... print e[0], e[1]
...
嗆嚷的 0.191532313824
使碎裂 0.184567078948
刻薄的 0.179777443409
杼發 0.179080232978
打耳光 0.170517712831
不可愛 0.166952252388
最高 0.162761792541
失效 0.156733453274
喋喋不休地講話 0.15637640655
豪華的 0.156299352646
```

```
>>> result=model.most_similar("酒醉")
>>> for e in result:
... print e[0], e[1]
...
苛刻 0.186927571893
基礎薄弱 0.174567013979
蠢話 0.171918794513
有得瞧 0.170480147004
逆境 0.166295751929
無法預料 0.165853306651
未完全發展的 0.165104150772
興沖沖 0.159420341253
劈啪 0.159055486321
暗中 0.158660337329
```

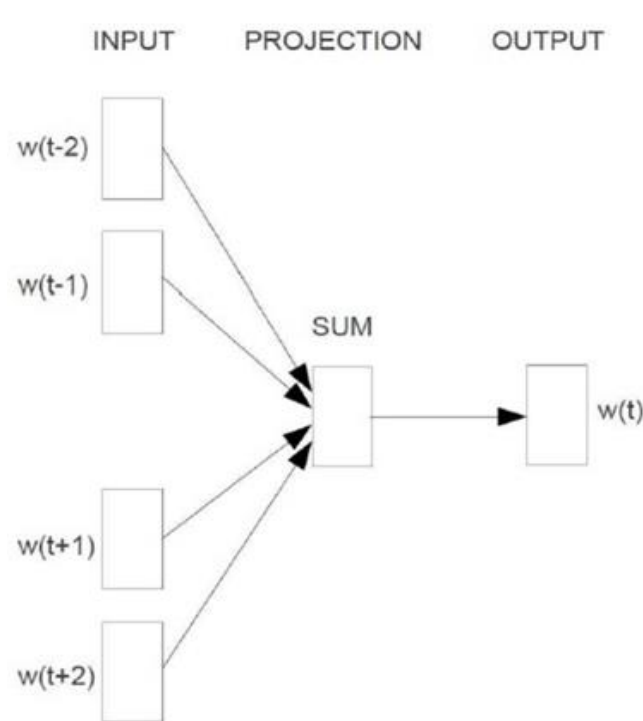
Jieba簡介



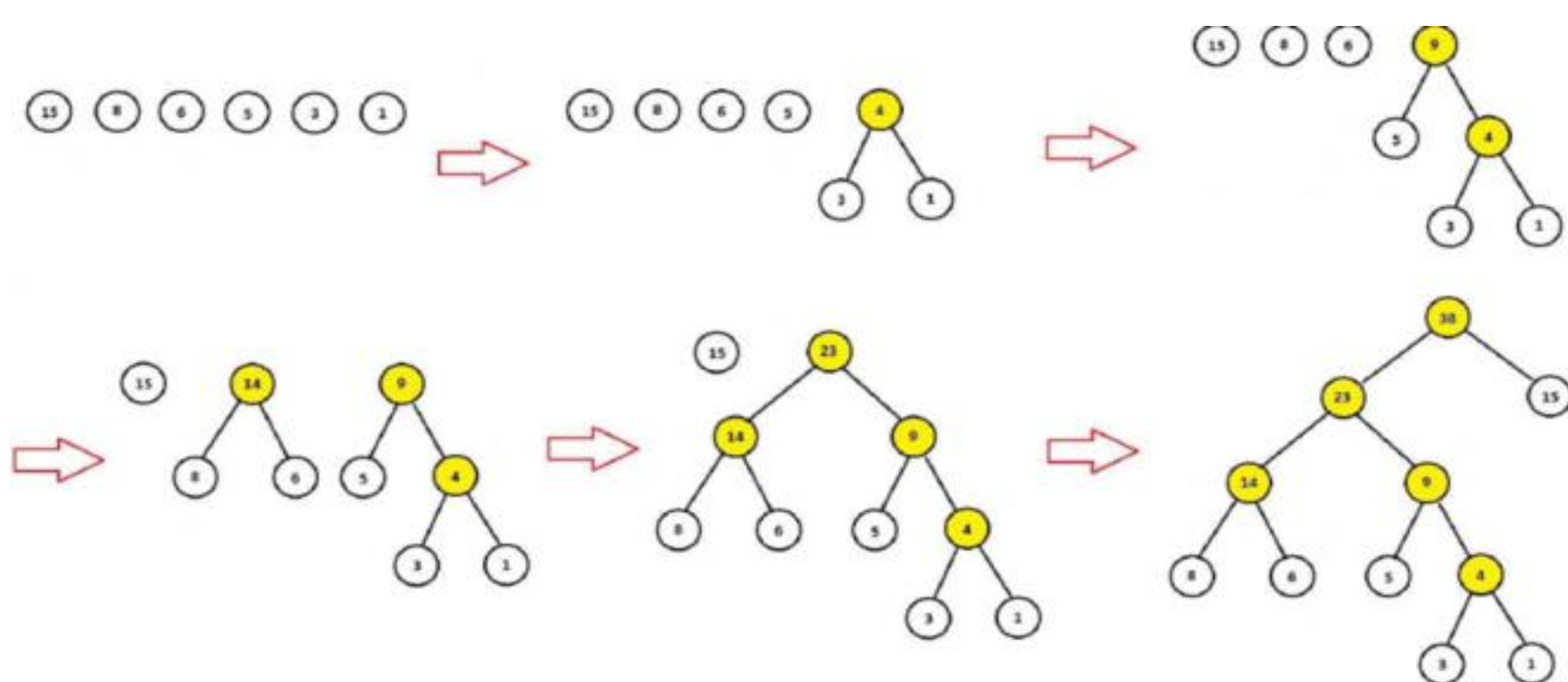
Trie樹



有向無環圖



CBOW 模型



Huffman樹