



結合孿生神經網路之FAQ問答系統

指導教授：陳柏琳 學生：吳峻銘

National Taiwan Normal University, Taiwan

目的與動機

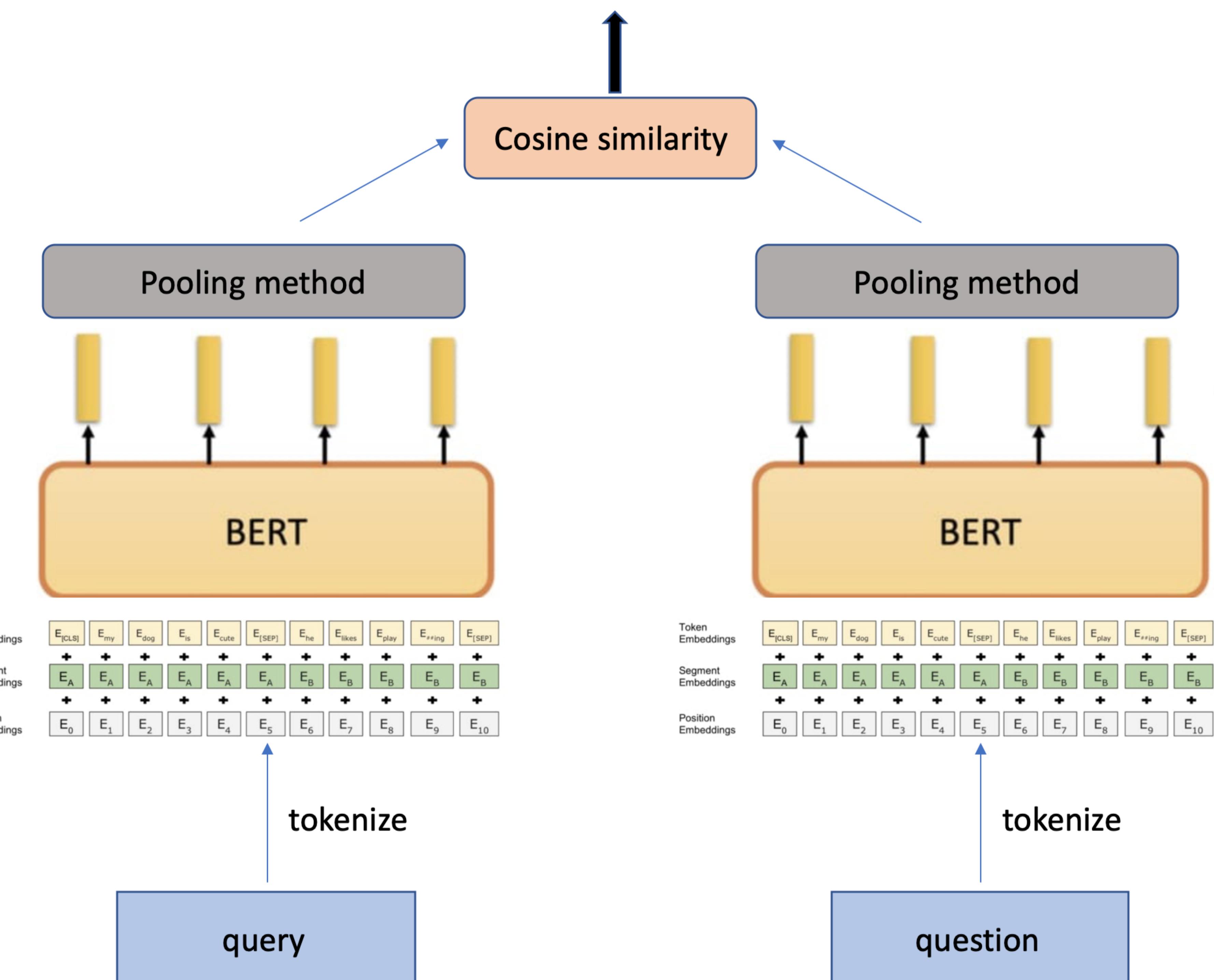
- FAQ-based QA是一種問答系統根據使用者的問句去FAQ pair中擷取語意最相近的一個，再把此FAQ pair的答案回傳給使用者，用以幫助使用者獲取特定領域的專業知識，FAQ-based QA的好壞將取決於常見問題檢索(FAQ Retrieval)的效能。
- 學生神經網路(Siamese Network)是一種神經網路結構常常用來處理兩個輸入資料之間相似度的問題，其架構為兩個神經網路共享權重。
- 本計畫建置一個結合以BERT為基礎的孿生神經網路的FAQ問答系統，希望能改善檢索不佳的問題。

系統展示

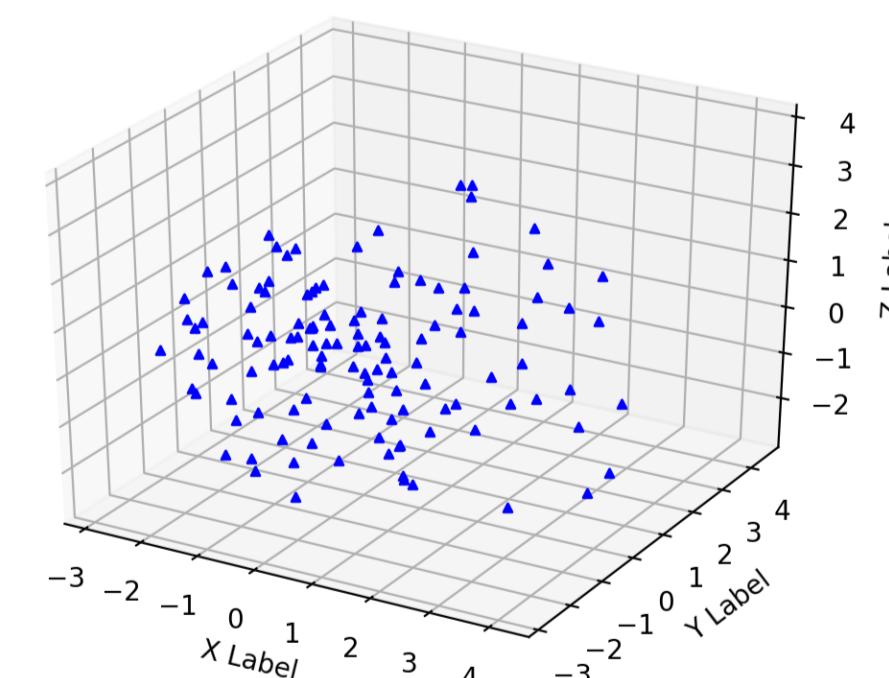
```
/lib/python2.7/site-packages/gmail/api.py:104: UserWarning: This method is deprecated and will be removed in a future version. It is recommended to use the new search method instead.
  warnings.warn("This method is deprecated and will be removed in a future version. It is recommended to use the new search method instead.", UserWarning)
[...]
How can I help you?(type by to exit):How can I filter my Gmail messages that aren't labeled?
  Updated: There's a blog today about updated Gmail search modifiers that allow you to do this with a simple search! has:userlabels Note: Because of Gmail's thread
ing, you will have labels on some of the messages in the derived list because some messages in a thread will have labeling while some won't. (From Gmail Help - Advanced
Search site.) Original answer: The Gmail advanced search help page says: label: Search for messages by label. There isn't a search operator for unlabeled messages
Example: from:david label:my-family Meaning: Messages from David that have t
he label "My Family"
[...]
How can I help you?(type by to exit):How can I embed a spreadsheet into my Google document on people docs?
  From today (2016-05-20) on, Google is starting to roll out the ability to embed a chart from a Google spreadsheet into a Google document. At this point, two of my 3
Google accounts make the option available in the Insert menu, below the Insert-Drawing item. When you insert such charts, they remain
linked to the original spreadsheet. If you want to edit your data, there's a table chart type that you can use to present your data as a simple table. There are some
limitations at this stage (e.g. regarding the size of the chart) and the chart must pre-exist in the spreadsheet before you can insert it, but this is going in the rig
ht direction... The chart is not updated live if you modify the source spreadsheet while your document is opened; however, the document will detect changes and offer you
to update your chart to reflect the latest changes in the source.
[...]
How can I help you?(type by to exit):Can google docs contain other objects such as spreadsheets?
  From today (2016-05-20) on, Google is starting to roll out the ability to embed a chart from a Google spreadsheet into a Google document. At this point, two of my 3
Google accounts make the option available in the Insert menu, below the Insert-Drawing item. When you insert such charts, they remain
linked to the original spreadsheet. If all you want is tabular data, there's a table chart type that you can use to present your data as a simple table. There are some
limitations at this stage (e.g. regarding the size of the chart) and the chart must pre-exist in the spreadsheet before you can insert it, but this is going in the rig
ht direction... The chart is not updated live if you modify the source spreadsheet while your document is opened; however, the document will detect changes and offer you
to update your chart to reflect the latest changes in the source.
[...]
How can I help you?(type by to exit):how can i remove my facebook account
  From here : Go here: http://www.facebook.com/help/delete_account Click "Submit" and follow the instructions. Your account will be deactivated for two weeks, and
if you DO NOT USE FACEBOOK IN ANY WAY during that period, your account is permanently deleted.
[...]
How can I help you?(type by to exit):I want to get rid of facebook, how can i do it?
  From here : Go here: http://www.facebook.com/help/delete_account Click "Submit" and follow the instructions. Your account will be deactivated for two weeks, and
if you DO NOT USE FACEBOOK IN ANY WAY during that period, your account is permanently deleted.
[...]
How can I help you?(type by to exit):bye
(base) bash-3.2$
```

實驗方法與步驟

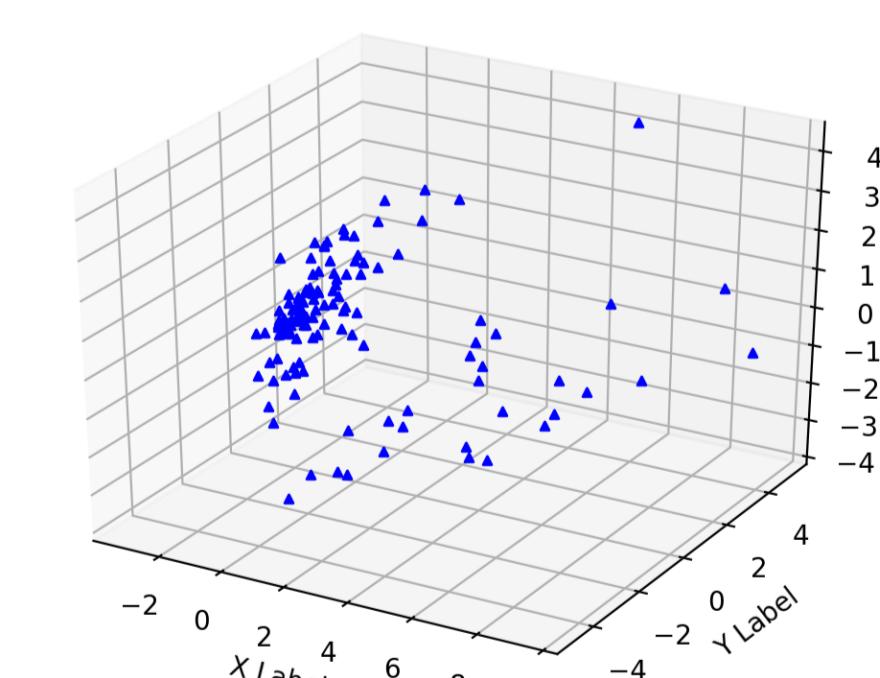
- Experiment dataset : StackExchange dataset
 1. 1250 query , 125 FAQ pair
 2. Web apps domain
- Training set(60%) , Development set(20%) , Test set(20%)
- 使用預訓練的BERT來產生詞嵌入(embedding)
- 實驗三種pooling method
 1. Average embedding
 2. [cls] token
 3. Max pool
- 擷取Cosine similarity值最高的FAQ pair



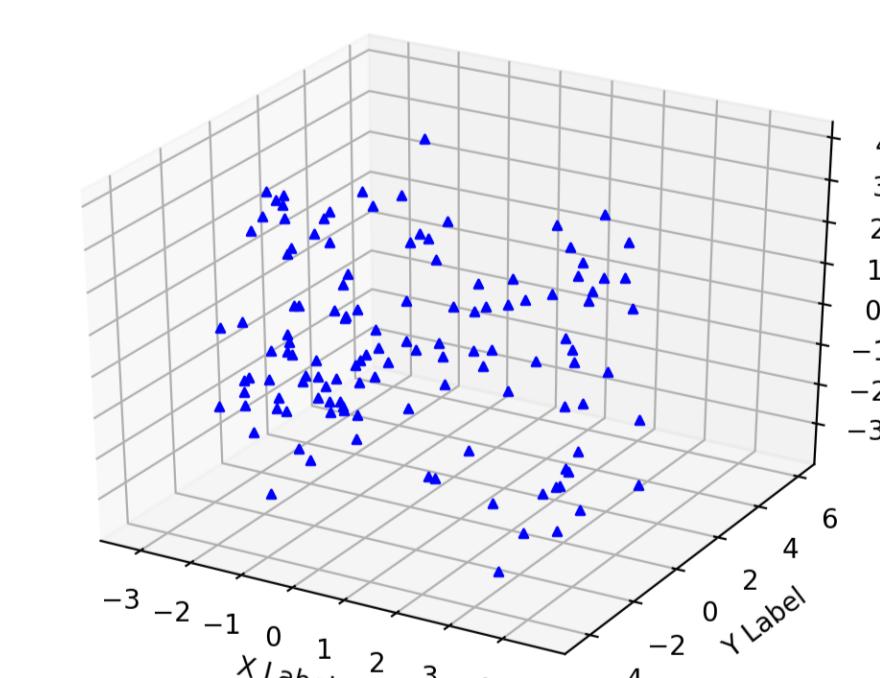
實驗結果



➤ BERT sentence embedding
use avg pooling method



➤ BERT sentence embedding
use cls pooling method



➤ BERT sentence embedding
use max pooling method

- 不同的pooling method在fine tune前後對系統效能比較(Map,Accuracy)

	avg without fine tune	avg	cls without fine tune	cls	max without fine tune	max
Map	0.5298	0.8324	0.2612	0.7722	0.4536	0.6141
Accuracy	0.4280	0.7640	0.1720	0.7000	0.3480	0.4960