# **MusicBot Milestone 1**

Members: 王耀賢 陳品融 陳聖曄 李俊鈺 謝明勳

## 1. Support Function

#### a. Slot:

i. [s]: singer name

ii. [t]: track name

iii. [a]: album name

iv. [g]: genre

#### b. Intent:

i. search(artist, album, track):
 return the search result given the artist, album or track
 \*但在training data當中被標成given了。下文以及demo系統顯示的given其實就是search。

ii. Recommend(artist,track,genre):return the recommended tracks given artist or track or genre

iii. Play(track): play the track

iv. Next:

play next song

v. Previous:

play previous song

vi. Pause:

pause the song

# 2. How we build our system:

# a. Training data Generation:

i. 首先先生出許多句子的模板 ex: 我想聽[s]唱的[t]

ii. 再去爬singer, track ,album, genre的資料

iii. 最後再把這些slot填入句子的template中產生training data

iv. Statistics:

Input vocabulary size: 11984 Max sequence length: 130

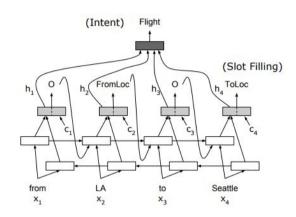
Intent distribution:

given recommend	askThis	popular	repeat	pause	next	previous	play
-----------------	---------	---------	--------	-------	------	----------	------

621259	572	15	12	11	9	9	7	4

## b. Language Understanding:

我們language understanding的部分是參考"Attention- Based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling"這篇論文,train一個可以同時做intent detection以及slot filling的 seq2seq model。我們的seq2seq model有採用attention的機制,而其中 encoder的部分爲一個bidirectional的RNN。其架構如下圖所示,foward 與backward的hidden states及其attention的weights可組合成attention states,而decoder便會根據這些attention states predict出intent以及每個step對應的slot。



由於每個input的長度都不盡相同,因此我們也有使用padding的方法,將每句話都補到最長句子的長度(130)。

以下爲我們的model predict出來的一些例子。其中如果intent是given,也就是給定歌手、歌名、專輯去找歌的狀況,我們可發現效果還不錯,然而如果是其他的intent如詢問歌曲、暫停播放等等,結果就不盡理想。原因是由於我們data絕大部分的intent都是given,其他intent的training example都太少,因此我們的model對於沒看過的input幾乎都會把intent predict成given。

```
>心情好差,聽聽opera歌
(['recommend'], ['0', '0', '0', '0', '0', '0', '0', 'g', '0'])
>我想聽五月天的乾杯
(['given'], ['0', '0', '0', 's', 's', 's', '0', 't', 't'])
>這是啥歌?
(['given'], ['t', 't', 't', 't'])
>播一下舊物
(['given'], ['0', '0', '0', 't', 't'])
```

此外,我們有發現training到最後perplexity可以降到1,也就是幾乎可以完美地predict training examples。我們認為這是由於句子都是由 template產生的,因此其結構大致上大同小異且沒有很複雜。而在 testing的時候我們有觀察到,只要輸入的input與template有些許差異便 會造成預測準確度的落差,所以我們認為在這個task上或許可以嘗試一 些比較簡單的model,比如單純的RNN等等,應該就可以有不錯的 performance,且同時較能避免overfit的問題。

#### c. Ontology:

我們不需要再建立database,直接使用spotify的API就可以完成所有我們所support的function,他可以做歌曲的搜尋;根據歌手、聽過的歌曲、曲風做出推薦;列出目前流行的歌曲;以及幫你建立播放清單下圖是API回傳的資訊

Track	Artist	Album	genre	popularity	
I See Fire	Ed Sheehan	x (Deluxe Edition)	рор	80	
麋途	麋先生Mixer	沒名字的人類	deep taiwanese pop	27	
殺死那石家莊人	萬念青年旅店	萬年青年旅店	deep taiwanese pop	30	
Two Moons	toe	For Long Tomorrow	post rock	40	

#### d. Language Generation:

用rule-based來generate回答