
Machine Learning Final Project

Listen and Translate
Kaggle

Task Description

Given a Taiwanese audio signal, select the most possible Chinese translations from the given options.



(A) 廠廠

(B) 來來來來來你來

(C) 你好帥

(D) 放棄final寒假就到了

Data Preprocessing

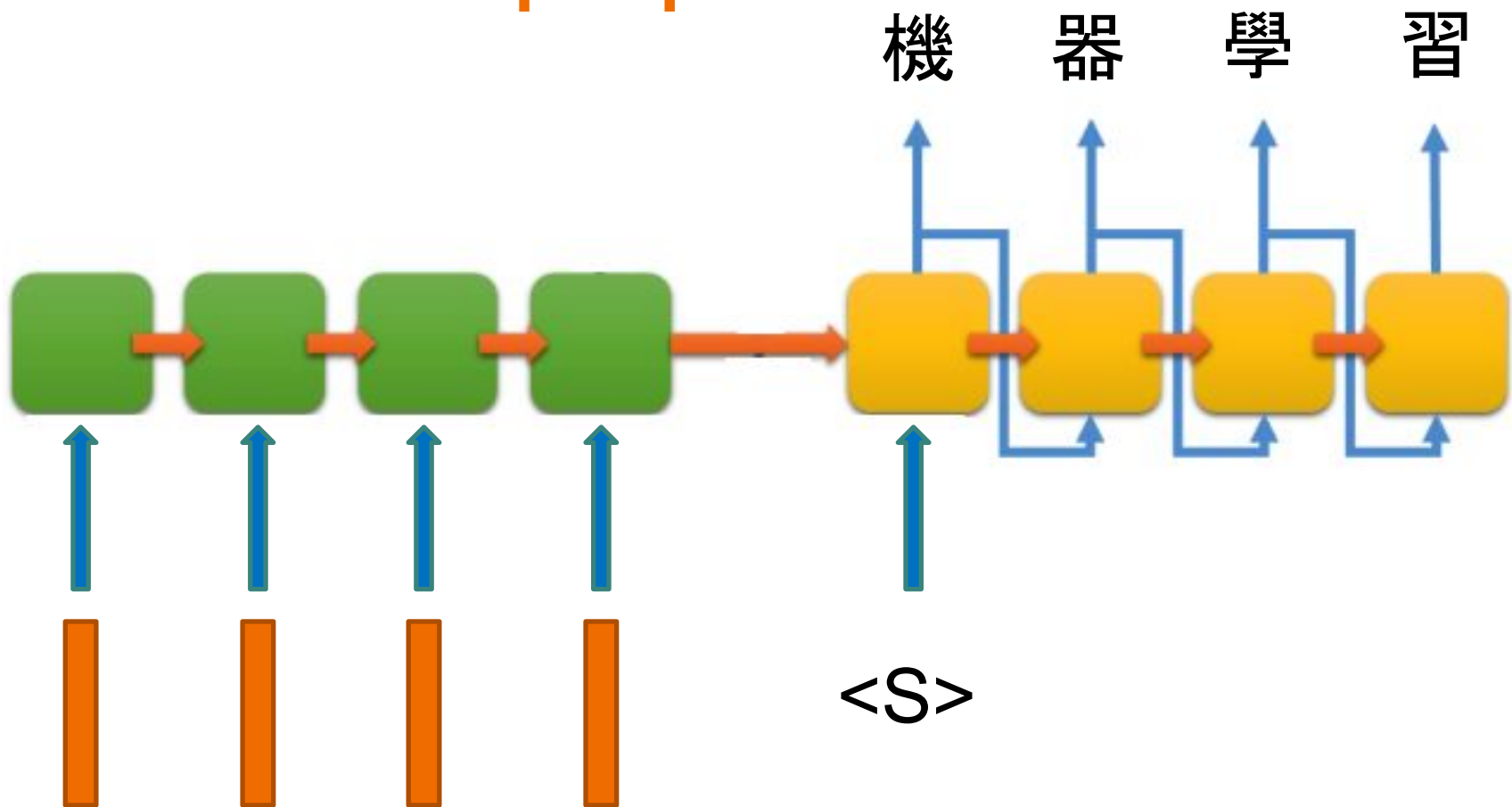
Use Kaldi to extract a series of MFCC feature vectors(39 dim) from audio signal.



References: (DSP 2017)

<http://speech.ee.ntu.edu.tw/DSP2017Autumn/Slides/7.0.pptx>

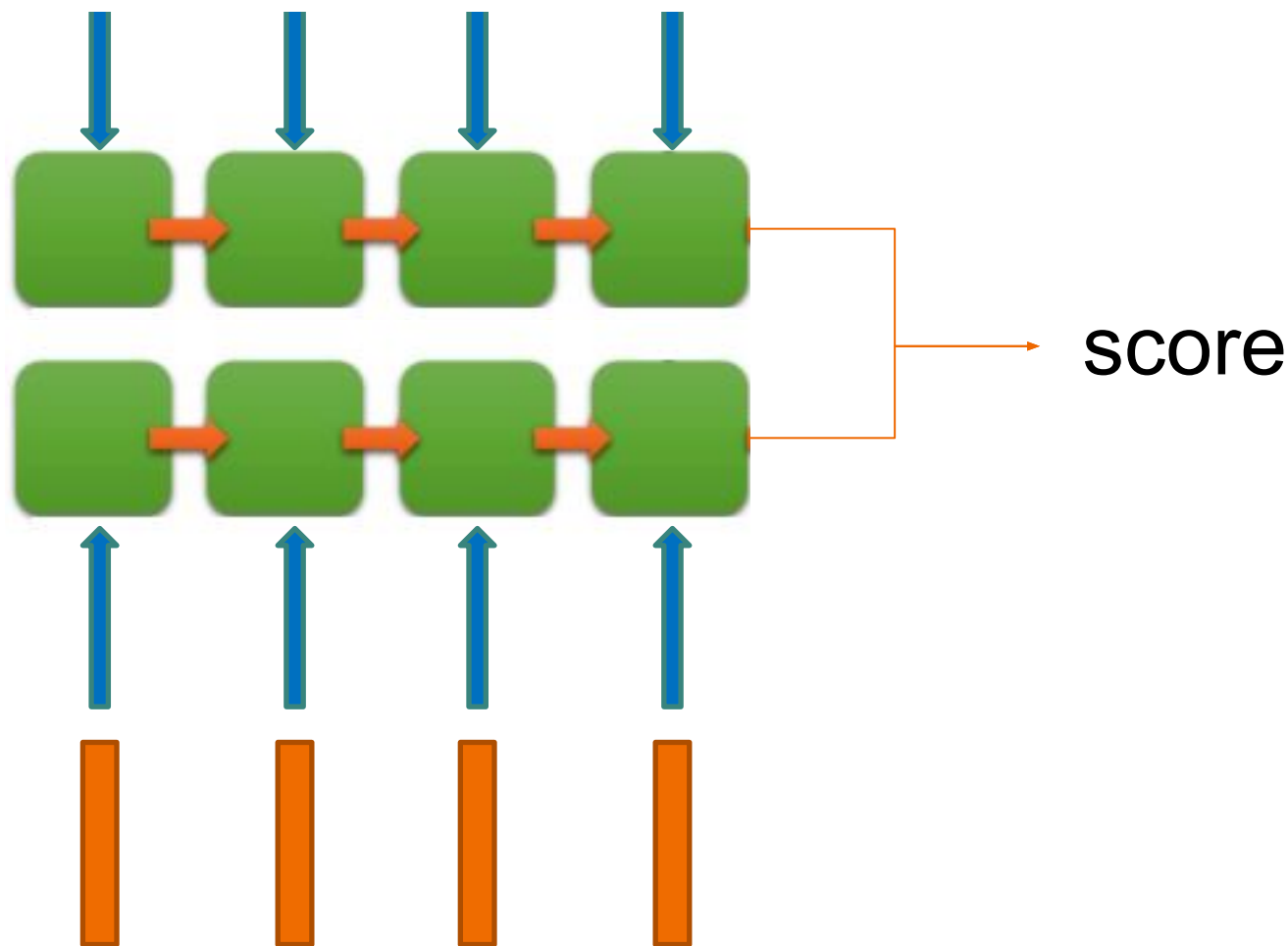
Basic model – seq2seq



Basic model – retrieval model

Hinge loss !!!

$$l(w, w_+, w_-) = \max(0, 1 - w^T \cdot w_+ + w^T \cdot w_-)$$



references

seq2seq

- <http://blog.csdn.net/jerry/article/details/53749693>
- (ML 2017)
[http://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2016/Lecture/RNN%20\(v2\).pdf](http://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2016/Lecture/RNN%20(v2).pdf)
- <http://cyruschiu.github.io/2017/02/24/learning-Tensorflow-Seq2Seq-for-translation/>

Hinge loss

- <http://blog.csdn.net/luo123n/article/details/48878759>

Training tips

- Word Embedding
 - fastText
- Handling of OOV

Data Files

Given Files:

- data.zip [Data]

- |__ train.data

- |__ train.caption

- |__ test.data

- |__ test.csv

- |__ example_wav.zip

Training Data

- train.data : a pickle file of a list of numpy tensors(39 x series_length)
- train.caption :

file format example: <word1><blank ><word2 ><blank>.....



你	要	把	握	這	個	機	會		
別	在	這	裡	丟	人	現	眼		
當	然	要	知	道	他	的	想	法	吧
鐵	的	紀	律	的	崇	尚	者	對	
好	不	好		我	身	邊	的	人	
不	要	牽	扯	她	被	你	欺	負	
眼	睜	睜	看	快	打	電	話	回	去
文	進	你	趕						

Testing Data

- test.data : a pickle file of a list of numpy tensors(39 x series_length)
- test.csv :

file format example: <option0><comma><option1><comma><option2>
<comma><option3>

是長官,阿花早,初吻呢,快去啦
孩子,第二,你呢,我想
下他們一組就換文隆哥他們了,幸好我不用再來這裡上班,阿嬤說出事的是她不是我,不要叫我保護>
你乾脆改名叫白癡好了,只知道在外面與人火拚,以後別再提美純就對了,對了她在美國那邊的>
錢
結果我發現妳的個性跟我一樣,為什麼要把事情搞得這麼複雜,愛上黃正源那種每天無所事事,>
跟妳上次一起扯下水那件婚紗
叫妳老公才不都要自己說出來了,人家是一日不見如隔三秋,你的工作是玉婷幫忙找的,待會兒妳看到>
媽妳就知道付給妳們的賠償金,跟妳上次一起扯下水那件婚紗,但不要让阿嬤知道曉萍出事了,>
他不能相信這段時間妳跟我相處
能不是相信這
媽,就,但,爸

Submission Format

- Submission format: a **.csv** file with problem id and the ranking of options(lowercase)
- With header row: "id,answer"
- id + , + predicted answer

```
id,answer  
1,0  
2,3  
3,2  
4,1
```

evaluation

- Accuracy
- Example:
- **m** correct answers out of **N** questions:
- $\text{accuracy} = \mathbf{m} / \mathbf{N}$

Contact

- 有問題請利用TA hours、信箱或FB 社團, **請不要FB私訊助教！！**
- 寄信詢問請在**title**前面加上“<台語翻譯>”, 確保負責助教會讀到信
- If you have other questions,
 - please contact TAs via ntu.mlta@gmail.com
 - post your questions on [facebook group](#)
 - go to TA office hours
 - 劉達融 TA Hour: TBA
 - 陳品融 TA Hour: TBA

QAO

Question and Questions?