

B06902125 黃柏瑋 SDML 作業2-0 & 2-1

2-0 training

運用簡單的char-level seq2seq model：

在encoder中將input過一層onehot embedding之後，送進512層的雙層GRU；

在decoder中用標準答案當作input(teacher forcing)，過一層onehot embedding後，送進以雙向GRU final state作為initial state的1024層GRU。

在訓練的過程中，loss = categorical cross entropy、batch = 128、learning rate = $1e-3$ 、random seed = 871024。

然而，這樣char-level的cross-entropy accuracy大約只能0.95，在整句答對的evaluation下很難拿到高分，於是我之後又將encoder output接在每個decoder input embedding之後，其餘不變，在大約25個epoch時便能達到training = 0.998、valid = 0.991的categorical cross entropy accuracy。

2-0 testing and evaluating

由於太晚發現助教將大小寫改為小寫，來不及再重新訓練模型，原先大小寫能在evaluation.py中拿到0.96的成績，一切都變成浮雲。

為了在課堂上testing的時候不要太難看，我將testing data的第一個字母都轉為大寫丟入model中，輸出再全轉為小寫，雖然成績可能還是會垮掉，但至少會讓模型適應開頭為大寫的input。

2-1-1 training

大致上與2-0幾乎一樣，而control signal也當一般字元投入encoder input中，只是將模型改成encoder gru = 128層、decoder gru = 256層。此外，為了強調control signal，在decoder中我將decoder input embedding + encoder output + control signal embedding投入真正的decoder gru中。

2-1-1 testing and evaluating

以char-level的categorical cross-entropy而言，這樣的設計能達到83%的accuracy，evaluation也有大約96%的成績。

2-1-2 training

和2-1-1相比沒有太多差別，只是當面對control signal只有一個時，為了再次強調control signal的重要性，與其利用padding補足長度，我將其增為兩倍處理(例如: 1 我 -> 1 我 1 我)。

2-1-2 testing and evaluating

以char-level的categorical cross-entropy而言，這樣的設計能達到84%的accuracy，evaluation也有大約90%的成績。

Command line for both

```
python train.py [train_dataset] [dict_name(exist or not both ok)]  
[model(not exist)]
```

```
python test.py [test_dataset] [dict_name(exist)] [model(exist)] >  
[result_file]
```

2-1 Remind

Before 2-1 training or testing, we must restrict the length of data in dataset first. And use the new_data as training/testing data.

```
python strict_data.py [original_data] [new_data]
```

Also, after 2-1 testing, we must transform it into the required format.

```
python transform_result.py [original_data] [new_data]
```

Packages

1. keras==2.3.0
2. tensorflow==2.0.0
3. numpy==1.17.2