ML hw5

> b05902127 劉俊緯

RNN

Structure

```
x = Input(shape=(40,128))
output = LSTM(256 , recurrent_dropout =0.3, return_sequences =True)(output )
lstm1 = BatchNormalization (momentum=0.5)(output )
output = lstm1
output = LSTM(256, recurrent_dropout =0.3, return_sequences =True)(output )
lstm2 = BatchNormalization (momentum=0.5)(output )
lstm2 =  add([lstm2 , lstm1])
output = lstm2
output = LSTM(256, recurrent_dropout =0.3, return_sequences =True)(output )
lstm3 = BatchNormalization (momentum=0.5)(output )
lstm3 =  add([lstm3 , lstm2])
output = lstm3
output = LSTM(256, recurrent_dropout =0.3, return_sequences =False)(output )
output = BatchNormalization (momentum=0.5)(output )
output = BatchNormalization (momentum=0.5)(Dense(512,activation='relu' )(output ))
output = Dropout (0.3)(output )
output = BatchNormalization (momentum=0.5)(Dense(256,activation='relu' )(output ))
output = Dropout (0.3)(output )
output = BatchNormalization (momentum=0.5)(Dense(64,activation='relu' )(output ))
output = Dropout (0.3)(output )
output = BatchNormalization (momentum=0.5)(Dense(16,activation='relu' )(output ))
y = Dense(1, activation='sigmoid' , kernel_initializer ='glorot_normal' )(output )
```

training details

- epoch: 5
- optimizer & learning rate: keras default Adam
- loss function: binary crossentropy

accuracy

- private : 0.821980 , public : 0.82166.

BOW

Structure

```
x = Input(shape=(40,256,1))
output = Reshape((40,256))(x)
output = Lambda(lambda x : K.sum(x,1))(output)
output = BatchNormalization()(output)
output = Dense(32,activation='relu')(output)
output = Dropout(0.3)(output)
output = BatchNormalization()(output)
output = Dense(8,activation='relu')(output)
y = Dense(1, activation='sigmoid', kernel_initializer='glorot_normal')(output)
```

training details

- epoch: 5
- optimizer & learning rate: keras default Adam
- loss function: binary crossentropy

accuracy

- private: 0.78611, private : 0.78633

Compare senetence by RNN

- RNN(LSTM)
  - "today is a good day, but it is hot" : 0.43633416295051575
  - "today is hot, but it is a good day" : 0.8663930296897888
- BOW
  - "today is a good day, but it is hot" : 0.7951212478234873
  - "today is hot, but it is a good day" : 0.7951212478234873
- 因為RNN會判斷前後順序的語意，因此做出來的分數會是好的，而bow不管順序，加起來一樣，所以出來的結果會是一樣。

Token Comparism

- with token:
  - private : 0.82347 , public : 0.82364.
- without token:
  - private : 0.82198 , public : 0.82166.
- 我認為因為有些標點符號例如！／？其實有幫助了整段語意更加明確。

Semi-supervised

- supervised:
  - private : 0.82347 , public : 0.82364
- semi-supervised:

    - 每跑完兩個epoch，就取20萬信心度最高的data加入training data裡面，並做三輪。
        - 因此是對信心度做排序。
    - private : 0.79960 , public : 0.80124
- semi效果較差我認為應該是semi-supervised因為加入太多self-training的結果導致更容易overfitting，使validation acc上不去。