

**HackMD** 

ADL 2022 Fall - HW1 Rep...

Q1: Data processing (2%)
Q2: Describe your intent...
Q3: Describe your slot ta...

Q4: Sequence Tagging E... Q5: Compare with differ...

全部展開

回到頂部

移至底部

Prime Full-text search on all paid plans

ADL 2022 Fall - HW1 Report

Describe how do you use the data for intent\_cls.sh , slot\_tag.sh :

str.split() 針對空白字串 '' 將句子內的單字切成 token。

對 intent classification、 slot tagging 兩個 task,都先對每個句子以 python 內建的

接著以[GloVe 的 Common Crawl glove.840B.300d] 這個 pre-trained 的 300 維 word

embedding 將上一步切好的 token 向量化。然而某些不被 GloVe 所涵蓋的 token,其

輸入為 300 維的 word embedding,接著以 bidirectional GRU 作為特徵抽取層,其中  $h_t$  作為

ightarrow bidirectional GRU 的輸出會將  $h_t \mathrel{^{ee}} h_t$  concat 在一起,再經過 batch normalization 成為

故整體會變成 MLP(BatchNorm( $[h_t, h_t]$ )) 或是 MLP(BatchNorm( $[GRU(h_t, h_{t-1}), GRU(h_t, h_{t+1})]$ ))

bidirectional GRU 在 t 時間的輸出。 $\rightarrow$  代表 forward layer 的方向; $\leftarrow$  代表 backward layer 的方

後面則為 2 層的 MLP (multi-layer perceptron),每層的 activation function 為 LeakyReLU,並

AccumulateGrad AccumulateGrad

♡ 讚賞

口 收藏

● 0 篇留言

♣ 已訂閱 >

🏙 變更於 20 分鐘前

tags: 課程報告

Describe

1. your model

 $egin{aligned} \stackrel{
ightarrow}{\overrightarrow{h_t}} &= \operatorname{GRU}(\stackrel{
ightarrow}{h_t}, \stackrel{
ightarrow}{h_{t-1}}) \ \stackrel{
ightarrow}{\leftarrow} & \stackrel{
ightarrow}{\leftarrow} \ h_t &= \operatorname{GRU}(\stackrel{
ightarrow}{h_t}, \stackrel{
ightarrow}{h_{t+1}}) \end{aligned}$ 

 $\operatorname{BatchNorm}([\overset{
ightarrow}{h_t},\overset{\leftarrow}{h_t}])$ 

且接著 Dropout。

Yang, Yu Chun (D11922022)

Q1: Data processing (2%)

1. How do you tokenize the data.

2. The pre-trained embedding you used.

embedding 則用 300 維的隨機值作為其向量。

Q2: Describe your intent classification model. (2%)