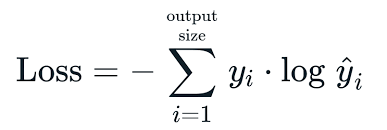
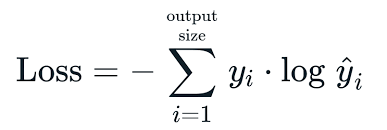
Applied Deep Learning HW1 report

B08310054 歐崇愷

1. Data Processing
   1. I use NLTK to tokenize 'text' field and didn’t remove stop word and didn’t set the max input length of text
   2. I use GLOVE as a fixed, pretrained word embedding. Embed each word into 300 dim vector
2. Describe your intent classification model
   1. I use LSTM model
      * *ht*, *ct* = LSTM(*wt*, *ht-1*, *ct-1*), where *wt* is the word embedding of the t-th token.
      * *ht* is a 2-dim vector, representing the hidden state of LSTM at the t-th timestamp
      * choose softmax as activation function
   2. performance of your model
      * **0.92222 acc**
   3. **cross entropy loss **
   4. I use torch.optim.Adam as the optimizer, with
      * Learning rate = 0.003
      * Batch size = 128
3. Describe your slot tagging model
   1. I use LSTM model
      * *ht*, *ct* = LSTM(*wt*, *ht-1*, *ct-1*), where *wt* is the word embedding of the t-th token.
      * *ht* is a 2-dim vector, representing the hidden state of LSTM at the t-th timestamp
      * choose softmax as activation function
   2. performance of your model
      * **0.92222 acc**
   3. **cross entropy loss **
   4. I use torch.optim.Adam as the optimizer, with
      * Learning rate = 0.003
      * Batch size = 128