Homework Assignment

Bookmark this page

Homework Assignment

This homework assignment is based on Lab6_TextClassification_with_LSTM.ipynb notebook.

Before doing this assignment, you should first complete the following tutorials:

- Understanding the ATIS Dataset
- Training and Evaluating Text Classification using Recurrent Network

Run the notebook and observe the test error. Right now, your test error hovers around 2.1x% to 2.2x%. Let's take this as our baseline and start experimenting.

Question 1

1/1 point (graded)

Currently, the number of embedding dimensions is set to 150. Experiment with the following embedding dimensions: 50, 150, and 300, and observe the test error. Which of the setting results in the **highest** test error?

- (•) emb_dim = 50 ✔
- ( ) emb_dim = 150
- ( ) emb_dim = 300

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

Once you've finished experimenting with the embedding dimensions, revert back to the original setting, that is, set the embedding dimensions back to 150.

Question 2

0/1 point (graded)

Another variable of importance is the number of hidden dimensions in the LSTM layer. Currently, the number of hidden dimensions is set to 300. Experiment with the following hidden dimensions: 100, 300, and 500, and observe the test error. Which of the setting results in the **lowest** test error?

- ○ 100
- ◉ 300
- ○ 500

Submit You have used 1 of 1 attempt

✖ Incorrect (0/1 point)

Once you've finished experimenting with the hidden dimensions, revert back to the original setting, that is, set the hidden dimensions back to 300.

Question 3

1/1 point (graded)

Up to now, you have seen several learners. With the MNIST data, we used SGD. With the Solar Panel data, we used fsadagrad. And with the ATIS data, we currently used adam as our learner. Experiment with SGD and fsadagrad as the learner for the ATIS data. Use the same learning rate and momentum as the ones used with adam. Which of the following learner results in the **highest** test error?

- ◉ SGD ✔
- ○ fsadagrad
- ○ adam

Submit You have used 1 of 1 attempt

✔ Correct (1/1 point)

Once you've finished experimenting with different learners, revert back to the original setting, that is, set adam as the learner for your model, and run the notebook.

Question 4

1/1 point (graded)
Currently, we are classifying a sequence: `BOS flights from new york to seattle EOS`

Let's try to classify another sequence:
`BOS i want round trip flights from new york to toronto EOS` . What is the label of the word **toronto**?

- B-fromloc.city_name
- I-fromloc.city_name
- B-toloc.city_name ✔
- I-toloc.city_name

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

Question 5

1/1 point (graded)
What is the index of the word **toronto** in the ATIS vocabulary?

- 855
- 856 ✔

- 857
- None of the above

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

Question 6

1/1 point (graded)

With the model you have, now try to classify the following sequence: `BOS flights from new york to paris EOS`. What do you observe?

- paris is classified as O
- paris is classified as B-fromloc.city_name
- paris is classified as B-toloc.city_name
- the code run into an error ✔

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

Question 7

1/1 point (graded)

The word **paris** is not in your vocabulary. One popular option to tackle this problem is to classify words that are not in the vocabulary as 'Unknown'. Currently, you have a vocabulary of 943 words with 129 labels. Equipped with updated `query_wl` and `slots_wl`, which three of the following options are needed to handle out of vocabulary word not present in training set?

- ☑ Add a new word such as 'Not known' to the vocabulary

- [x] Add a label 'Unknown' to the slots.wl for the new 'Not known' word
- [ ] Retrain the model (with some code change in `create_model`)
- [x] Retrain the model (no code change in `create_model`)
- [ ] Use the original model (no retraining needed)

✔

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

English ▾