Slot Tagging of Natural Language Utterances

CSE 244 Machine Learning for NLP, Homework 2

Instructor: Dilek Hakkani-Tur
TA: Jialu Wang

Due: 02/23/2020 11:59PM

1 Introduction

The goal of this homework is to tag slots in natural language utterances of users addressing a virtual personal assistant. The tagged slots and the associated values are traditionally used to accomplish the request of the users. For example, if a user asks to learn about movies of a specific director, we would need to identify the name of the director (as well as the user's intention to find out about movies), and issue a query to the backend knowledge graph to get information for formulating a response back to the user.

You will be given a training set of utterances, each paired with a sequence of IOB (inside-outside-begin) tags. You can use these to sequence tagging models. Here is an example utterance from the data set:

`Show me movies directed by Woody Allen recently`

There are two slots in this utterance:

director = Woody Allen
release_year = recently

As we learned in class, the sequence of IOB tags corresponding to these slots would be the following tag sequence:

`O O O O O B_director I_director B_release year`

.

You are expected to train deep neural networks to tag sequential data using the training examples. Here are some example models you could try:

- vanilla RNN/Elman RNN we covered in class
- RNN with LSTM or GRU cells

For each of these, you can change the depth of layers, numbers of neurons in each layer, etc. You can also try different loss functions and learning rates

to train your models. You may try any technique and trick to improve your models' performance. You need to describe your methods in details and how they improve the performance in your report.

2 Dataset Format

Please note that the data was collected through crowd sourcing, by priming crowd workers to imagine a hypothetical assistant. The annotations for this dataset is generated based on a simplified version of film schema of Freebase knowledge graph. There are five files.

hw2_train.txt The `train` set has two columns:

- *utterances*: the natural language text you need to label the slots.
- *IOB Slot tags*: the targeted sequence of IOB tags corresponding to the utterance.

hw2_utterance_dev.txt The `dev` data. It has only one column:

- *utterances*: the natural language text you need to label the slots.

hw2_tags_dev.txt The `dev` label. It has only one column:

- *IOB Slot tags*: the targeted sequence of IOB tags corresponding to the utterance in the same row in *hw2_utterance_dev.csv*.

hw2_utterance_test.txt The `test` data. It has only one column:

- *utterances*: the natural language text you need to label the slots.

The ratio of numbers of data examples in `dev` set and `test` set is approximately 30 : 70.

3 Submission Files

You need to submit a zip file on canvas including:

1. Your prediction on test data. The name must be *prediction.txt* so that we can run the evaluation script directly.
2. A report in pdf.
3. Your source code.

4 Evaluation

Submissions will be evaluated based on their mean F-1 score considering the context. We provide our evaluation script on Canvas. To evaluate your prediction's performance on the dev set, simply run

```
python3 evaluation.py
```

and evaluate on the test set

```
python3 evaluation.py --test
```