Sentence Generator

Overview

Sentence Generator is a program to predict following words after entered words.
Sentence Generator uses a language model trained with PTB dataset.
Deep Recursive Neural Network (with 2 LSTM Layers) is applied to implement the language model.

Although the generated sentences are not quite perfect
You will see it mimics the structure of sentences.

Requirements

- Python > 3.5
- TensorFlow 1.4

Install all python packages required using pip

```
$ pip install -r requirements.txt
```

Using Virtualenv is recommended.

Dataset

To train the model, Penn Tree Bank(PTB) dataset is used.
Download 'ptb.train.txt' here and place in ./data.

Train Model

```
$ python word_sequence.py --mode train
```

Generate Sentence

```
$ python word_sequence.py --mode pretrained
Enter words(enter '!' to exit):
```

You can enter either a single word or multiple words.
In case you enter words out of the vocabulary, the program ends.

***N** is numbers.* ***<unk>** is a tag indicating unknown words.* *(They were already preprocessed in the PTB dataset)*

Generate Sentence using Pretrained Model

If you want to use the sentence generator without training the model,
Use '--model pretrained' option.
(You need to download './pretrained')

```
$ python sentence_generator.py --mode generate --model pretrained
```

Result

```
Enter words(enter '!' to exit): i
predicted_sentence:  i contract it 's a bill <unk>
Enter words(enter '!' to exit): it is
predicted_sentence:  it is n't given that the <unk> will be <unk>
Enter words(enter '!' to exit): it is n't
predicted_sentence:  it is n't a <unk>
Enter words(enter '!' to exit): the
predicted_sentence:  the company said it will be a N N union in the company
Enter words(enter '!' to exit): how
predicted_sentence:  how many companies are <unk> to be able to do the <unk>
Enter words(enter '!' to exit): the companies
predicted_sentence:  the companies were n't to be able to be a <unk> for the <unk>
Enter words(enter '!' to exit): !
Bye!
```