Improving language modelling with overlapping

This repos implements the overlapping mechanism on state-of-the-art model.

Requirements

Python 3.6, PyTorch 0.4.1

Installation via virtualenv

Create a virtualenv:

```
python3 -m virtualenv venv
```

Activate the virtualenv:

```
source venv/bin/activate
```

Install torch and numpy:

```
pip3 install torch==0.4.1
pip3 install numpy
```

Download the data

Note: `wget` is required.

```
./get_data.sh
```

Train the models (to reproduce our results)

Penn Treebank

First, train the model

```
python main.py --data data/penn --dropouti 0.4 --dropoutl 0.29 --dropouth 0.225
--seed 28 --batch_size 12 --lr 20.0 --epoch 1000 --nhid 960 --nhidlast 620 --
emsize 280 --n_experts 15 --save PTB --single_gpu
```

Second, finetune the model

```
python finetune.py --data data/penn --dropouti 0.4 --dropoutl 0.29 --dropouth
0.225 --seed 28 --batch_size 12 --lr 25.0 --epoch 1000 --nhid 960 --emsize 280 -
-n_experts 15 --save PATH_TO_FOLDER --single_gpu
```

where `PATH_TO_FOLDER` is the folder created by the first step (concatenation of PTB with a timestamp).

Third, run dynamic evaluation

```
python dynamiceval.py --model PATH_TO_FOLDER/finetune_model.pt --lamb 0.075
```

WikiText-2

First, train the model

```
python main.py --epochs 1000 --data data/wikitext-2 --save WT2 --dropouth 0.2 --seed 1882 --n_experts 15 --nhid 1150 --nhidlast 650 --emsize 300 --batch_size 15 --lr 15.0 --dropoutl 0.29 --small_batch_size 5 --max_seq_len_delta 20 --dropouti 0.55 --single_gpu
```

Second, finetune the model

```
python finetune.py --epochs 1000 --data data/wikitext-2 --save PATH_TO_FOLDER --dropouth 0.2 --seed 1882 --n_experts 15 --nhid 1150 --emsize 300 --batch_size 15 --lr 20.0 --dropoutl 0.29 --small_batch_size 5 --max_seq_len_delta 20 --dropouti 0.55 --single_gpu
```

Third, run dynamic evaluation

```
python dynamiceval.py --data data/wikitext-2 --model PATH_TO_FOLDER/finetune_model.pt --epsilon 0.002
```

Acknowledgements

A large portion of this repo is borrowed from the following repo: https://github.com/zihangdai/mos