

Overlap with language modelling

Hold experiments on 2 models using overlap:

- **simple-lstm**, a very basic lstm
- **awd-lstm**, [AWD](#) ASGD Weight-Dropped LSTM
- **mos-lstm**, [MOS](#) Mixture of Softmaxes

To specify which model to run, use `--main-model {simple-lstm | awd-lstm | mos-lstm}` argument. There are additional common parameters as well as specific parameters for each model. Those can be found in `main_run.py`.

Set-up

Download the data:

```
chmod +x get_data.sh
./get_data.sh
```

Install dependencies:

```
python3.6 -m virtualenv venv
source venv/bin/activate
pip3 install -r requirements.txt
```

Simple lstm

```
python3 main_run.py --main-model simple-lstm --epochs 1000
```

MOS lstm

TODO

AWD lstm

Fast run:

```
python3 main_run.py --batch-size 20 --data data/penn --dropouti 0.4 \
    --dropout 0.25 --seed 141 --epoch 1 \
    --nhid 5 --emsize 5 --nlayers 1 --bptt 5 \
    --main-model awd-lstm
```

Acknowledgements

Code is heavily borrowed from the following sources:

- simple-lstm: <https://github.com/deeplearningathome/pytorch-language-model>
- awd-lstm: <https://github.com/salesforce/awd-lstm-lm>
- mos-lstm: <https://github.com/zihangdai/mos>