bi-rnn

April 16, 2019

0.1 Bidirectional Recurrent Neural Networks

```
I am _____
I am _____ very hungry,
I am _____ very hungry, I could eat half a pig.
```

$$\overrightarrow{\mathbf{H}}_t = \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(f)} + \overrightarrow{\mathbf{H}}_{t-1} \mathbf{W}_{hh}^{(f)} + \mathbf{b}_h^{(f)}),$$
$$\overleftarrow{\mathbf{H}}_t = \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(b)} + \overleftarrow{\mathbf{H}}_{t+1} \mathbf{W}_{hh}^{(b)} + \mathbf{b}_h^{(b)}),$$

Output

$$\mathbf{O}_t = \mathbf{H}_t \mathbf{W}_{hq} + \mathbf{b}_q,$$

0.1.1 Doing it wrong

```python
In [ ]: import sys
        sys.path.insert(0, '..')

        import d2l
        from mxnet import nd
        from mxnet.gluon import rnn

        (corpus_indices, char_to_idx, idx_to_char, vocab_size) = d2l.load_data_time_machine()

        num_inputs, num_hiddens, num_layers, num_outputs = vocab_size, 256, 2, vocab_size
        ctx = d2l.try_gpu()
        num_epochs, num_steps, batch_size, lr, clipping_theta = 1000, 35, 32, 100, 1e-2
        pred_period, pred_len, prefixes = 200, 50, ['traveller', 'time traveller']

        lstm_layer = rnn.LSTM(hidden_size = num_hiddens, num_layers=num_layers,
                              bidirectional = True)
        model = d2l.RNNModel(lstm_layer, vocab_size)

In [1]: d2l.train_and_predict_rnn_gluon(model, num_hiddens, vocab_size, ctx,
                                        corpus_indices, idx_to_char, char_to_idx,
                                        num_epochs, num_steps, lr, clipping_theta,
                                        batch_size, pred_period, pred_len, prefixes)
```

```
epoch 200, perplexity 1.120558, time 0.10 sec
 - travellererererererererererererererererererererererer
 - time travellererererererererererererererererererererererer
epoch 400, perplexity 1.054240, time 0.10 sec
 - travellers that alosesesesesesesesesesesesesesesesese
 - time travellerly thickn--------------------------------------
epoch 600, perplexity 1.007709, time 0.10 sec
 - travellererer brececedededededededededededededededede
 - time travellererer brecedededededededededededededededede
epoch 800, perplexity 1.009584, time 0.10 sec
 - traveller (ricee thace absmondidiz getoninininininininini
 - time traveller (fffrfrfrf thee presesesesesesesesesesesesese
epoch 1000, perplexity 1.005260, time 0.10 sec
 - traveller hack, why cann of thace but, anomememe. the time!
 - time traveller hack, why car rare eximitigep pooveveve:e:e:e:e
```