Encoder

file named fairseq/models/simple\_lstm.py:

**import** **torch.nn** **as** **nn**

**from** **fairseq** **import** utils

**from** **fairseq.models** **import** FairseqEncoder

**class** **SimpleLSTMEncoder**(FairseqEncoder):

**def** \_\_init\_\_(

self, args, dictionary, embed\_dim=128, hidden\_dim=128, dropout=0.1,

):

super().\_\_init\_\_(dictionary)

self.args = args

*# Our encoder will embed the inputs before feeding them to the LSTM.*

self.embed\_tokens = nn.Embedding(

num\_embeddings=len(dictionary),

embedding\_dim=embed\_dim,

padding\_idx=dictionary.pad(),

)

self.dropout = nn.Dropout(p=dropout)

*# We'll use a single-layer, unidirectional LSTM for simplicity.*

self.lstm = nn.LSTM(

input\_size=embed\_dim,

hidden\_size=hidden\_dim,

num\_layers=1,

bidirectional=**False**,

batch\_first=**True**,

)

**def** forward(self, src\_tokens, src\_lengths):

*# The inputs to the ``forward()`` function are determined by the*

*# Task, and in particular the ``'net\_input'`` key in each*

*# mini-batch. We discuss Tasks in the next tutorial, but for now just*

*# know that \*src\_tokens\* has shape `(batch, src\_len)` and \*src\_lengths\**

*# has shape `(batch)`.*

*# Note that the source is typically padded on the left. This can be*

*# configured by adding the `--left-pad-source "False"` command-line*

*# argument, but here we'll make the Encoder handle either kind of*

*# padding by converting everything to be right-padded.*

**if** self.args.left\_pad\_source:

*# Convert left-padding to right-padding.*

src\_tokens = utils.convert\_padding\_direction(

src\_tokens,

padding\_idx=self.dictionary.pad(),

left\_to\_right=**True**

)

*# Embed the source.*

x = self.embed\_tokens(src\_tokens)

*# Apply dropout.*

x = self.dropout(x)

*# Pack the sequence into a PackedSequence object to feed to the LSTM.*

x = nn.utils.rnn.pack\_padded\_sequence(x, src\_lengths, batch\_first=**True**)

*# Get the output from the LSTM.*

\_outputs, (final\_hidden, \_final\_cell) = self.lstm(x)

*# Return the Encoder's output. This can be any object and will be*

*# passed directly to the Decoder.*

**return** {

*# this will have shape `(bsz, hidden\_dim)`*

'final\_hidden': final\_hidden.squeeze(0),

}

*# Encoders are required to implement this method so that we can rearrange*

*# the order of the batch elements during inference (e.g., beam search).*

**def** reorder\_encoder\_out(self, encoder\_out, new\_order):

*"""*

*Reorder encoder output according to `new\_order`.*

*Args:*

*encoder\_out: output from the ``forward()`` method*

*new\_order (LongTensor): desired order*

*Returns:*

*`encoder\_out` rearranged according to `new\_order`*

*"""*

final\_hidden = encoder\_out['final\_hidden']

**return** {

'final\_hidden': final\_hidden.index\_select(0, new\_order),

}

DECODER

**import** **torch**

**from** **fairseq.models** **import** FairseqDecoder

**class** **SimpleLSTMDecoder**(FairseqDecoder):

**def** \_\_init\_\_(

self, dictionary, encoder\_hidden\_dim=128, embed\_dim=128, hidden\_dim=128,

dropout=0.1,

):

super().\_\_init\_\_(dictionary)

*# Our decoder will embed the inputs before feeding them to the LSTM.*

self.embed\_tokens = nn.Embedding(

num\_embeddings=len(dictionary),

embedding\_dim=embed\_dim,

padding\_idx=dictionary.pad(),

)

self.dropout = nn.Dropout(p=dropout)

*# We'll use a single-layer, unidirectional LSTM for simplicity.*

self.lstm = nn.LSTM(

*# For the first layer we'll concatenate the Encoder's final hidden*

*# state with the embedded target tokens.*

input\_size=encoder\_hidden\_dim + embed\_dim,

hidden\_size=hidden\_dim,

num\_layers=1,

bidirectional=**False**,

)

*# Define the output projection.*

self.output\_projection = nn.Linear(hidden\_dim, len(dictionary))

*# During training Decoders are expected to take the entire target sequence*

*# (shifted right by one position) and produce logits over the vocabulary.*

*# The \*prev\_output\_tokens\* tensor begins with the end-of-sentence symbol,*

*# ``dictionary.eos()``, followed by the target sequence.*

**def** forward(self, prev\_output\_tokens, encoder\_out):

*"""*

*Args:*

*prev\_output\_tokens (LongTensor): previous decoder outputs of shape*

*`(batch, tgt\_len)`, for teacher forcing*

*encoder\_out (Tensor, optional): output from the encoder, used for*

*encoder-side attention*

*Returns:*

*tuple:*

*- the last decoder layer's output of shape*

*`(batch, tgt\_len, vocab)`*

*- the last decoder layer's attention weights of shape*

*`(batch, tgt\_len, src\_len)`*

*"""*

bsz, tgt\_len = prev\_output\_tokens.size()

*# Extract the final hidden state from the Encoder.*

final\_encoder\_hidden = encoder\_out['final\_hidden']

*# Embed the target sequence, which has been shifted right by one*

*# position and now starts with the end-of-sentence symbol.*

x = self.embed\_tokens(prev\_output\_tokens)

*# Apply dropout.*

x = self.dropout(x)

*# Concatenate the Encoder's final hidden state to \*every\* embedded*

*# target token.*

x = torch.cat(

[x, final\_encoder\_hidden.unsqueeze(1).expand(bsz, tgt\_len, -1)],

dim=2,

)

*# Using PackedSequence objects in the Decoder is harder than in the*

*# Encoder, since the targets are not sorted in descending length order,*

*# which is a requirement of ``pack\_padded\_sequence()``. Instead we'll*

*# feed nn.LSTM directly.*

initial\_state = (

final\_encoder\_hidden.unsqueeze(0), *# hidden*

torch.zeros\_like(final\_encoder\_hidden).unsqueeze(0), *# cell*

)

output, \_ = self.lstm(

x.transpose(0, 1), *# convert to shape `(tgt\_len, bsz, dim)`*

initial\_state,

)

x = output.transpose(0, 1) *# convert to shape `(bsz, tgt\_len, hidden)`*

*# Project the outputs to the size of the vocabulary.*

x = self.output\_projection(x)

*# Return the logits and ``None`` for the attention weights*

**return** x, **None**