EmuBidirLSTM((bidir): LSTM(1024, 1024, bidirectional=True) (layer1): LSTM(1024, 1024) (layer2): LSTM(1024, 1024)) [forward_compute_time=14.126,backward_compute_time=0.019,activation_size=26214400.0,parameter_size=67174400.0] $Dropout(p=0.2)\\ [forward_compute_time=0.186,backward_compute_time=0.285,activation_size=26214400.0,parameter_size=0.0]$ LSTM(2048, 1024)
[forward_compute_time=10.298,backward_compute_time=18.777,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] $Dropout(p=0.2) \\ [forward_compute_time=0.124,backward_compute_time=0.159,activation_size=13107200.0,parameter_size=0.0]$ LSTM(1024, 1024)
[forward_compute_time=6.887,backward_compute_time=13.549,activation_size=14155776.0,parameter_size=33587200.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] $Dropout(p=0.2) \\ [forward_compute_time=0.129,backward_compute_time=0.167,activation_size=13107200.0,parameter_size=0.0]$ LSTM(1024, 1024)
[forward_compute_time=6.903,backward_compute_time=13.518,activation_size=14155776.0,parameter_size=33587200.0] $_getitem_(0) \\ [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0]$ [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] $Dropout(p=0.2) \\ [forward_compute_time=0.122,backward_compute_time=0.160,activation_size=13107200.0,parameter_size=0.0]$ LSTM(1024, 1024) [forward_compute_time=6.889,backward_compute_time=13.594,activation_size=14155776.0,parameter_size=33587200.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Add [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.120,backward_compute_time=0.160,activation_size=13107200.0,parameter_size=0.0] LSTM(1024, 1024)
[forward_compute_time=7.019,backward_compute_time=13.628,activation_size=14155776.0,parameter_size=33587200.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.120,backward_compute_time=0.164,activation_size=13107200.0,parameter_size=0.0] LSTM(1024, 1024)
[forward_compute_time=6.991,backward_compute_time=13.614,activation_size=14155776.0,parameter_size=33587200.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.120,backward_compute_time=0.169,activation_size=13107200.0,parameter_size=0.0] LSTM(1024, 1024)
[forward_compute_time=7.055,backward_compute_time=14.112,activation_size=14155776.0,parameter_size=33587200.0] Input2 [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] hidden [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Embedding(32320, 1024, padding_idx=0) [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.128,backward_compute_time=0.460,activation_size=13107200.0,parameter_size=132382720.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] RecurrentAttention((rnn): LSTM(1024, 1024) (attn): BahdanauAttention((linear_q): Linear(in_features=1024, out_features=1024, __getitem__(7)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] [forward compute time=15.210,backward compute time=23.086,activation size=27582976.0,parameter size=41979904.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward compute time=0.000,backward compute time=0.000,activation size=0.0,parameter size=0.0] __getitem__(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=524288.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.104,backward_compute_time=0.262,activation_size=13107200.0,parameter_size=0.0] __getitem__(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] __getitem__(3)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] [forward_compute_time=10.517,backward_compute_time=18.744,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] __getitem__(5)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] $Dropout(p=0.2) \\ [forward_compute_time=0.109, backward_compute_time=0.277, activation_size=13107200.0, parameter_size=0.0]$ Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] __getitem__(6)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0] LSTM(2048, 1024)
[forward_compute_time=10.553,backward_compute_time=18.771,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0) [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.120,backward_compute_time=0.256,activation_size=13107200.0,parameter_size=0.0] Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] LSTM(2048, 1024)
[forward_compute_time=10.725,backward_compute_time=18.802,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Add [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] $Dropout(p=0.2)\\ [forward_compute_time=0.121,backward_compute_time=0.263,activation_size=13107200.0,parameter_size=0.0]$ Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] LSTM(2048, 1024)
[forward_compute_time=10.643,backward_compute_time=18.863,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.131,backward_compute_time=0.263,activation_size=13107200.0,parameter_size=0.0] Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] LSTM(2048, 1024)
[forward_compute_time=10.651,backward_compute_time=19.036,activation_size=14155776.0,parameter_size=50364416.0] $_getitem_(0) \\ [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0]$ [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.126,backward_compute_time=0.261,activation_size=13107200.0,parameter_size=0.0] Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=10.659,backward_compute_time=19.145,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] Dropout(p=0.2)
[forward_compute_time=0.144,backward_compute_time=0.201,activation_size=13107200.0,parameter_size=0.0] Concat(2)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=10.611,backward_compute_time=66.214,activation_size=14155776.0,parameter_size=50364416.0] __getitem__(0)
[forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0] [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=13107200.0,parameter_size=0.0]

Classifier((classifier): Linear(in_features=1024, out_features=32320, bias=True))
[forward_compute_time=24.782,backward_compute_time=5.373,activation_size=413696000.0,parameter_size=132512000.0]

Input1 [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0]

 $Input 0 \\ [forward_compute_time=0.000,backward_compute_time=0.000,activation_size=0.0,parameter_size=0.0]$

 $Embedding (32320,\ 1024,\ padding_idx=0)\\ [forward_compute_time=0.140,backward_compute_time=25.538, activation_size=13107200.0, parameter_size=132382720.0]$