

epoch 1, step: 0, tr_loss: 1369.63, vl_loss: 1644.89, tr_err: 6.31, vl_err: 7.77
epoch 2, step: 0, tr_loss: 1507.26, vl_loss: 1606.90, tr_err: 7.77, vl_err: 7.46
epoch 3, step: 0, tr_loss: 1527.81, vl_loss: 1551.59, tr_err: 7.46, vl_err: 6.80
epoch 4, step: 0, tr_loss: 1506.10, vl_loss: 1481.24, tr_err: 6.80, vl_err: 5.98
epoch 5, step: 0, tr_loss: 1458.91, vl_loss: 1406.71, tr_err: 5.98, vl_err: 5.26
epoch 6, step: 0, tr_loss: 1400.53, vl_loss: 1333.47, tr_err: 5.26, vl_err: 4.48
epoch 7, step: 0, tr_loss: 1338.64, vl_loss: 1264.62, tr_err: 4.48, vl_err: 3.56
epoch 8, step: 0, tr_loss: 1277.75, vl_loss: 1200.58, tr_err: 3.56, vl_err: 2.69
epoch 9, step: 0, tr_loss: 1219.36, vl_loss: 1141.08, tr_err: 2.69, vl_err: 2.04
epoch 10, step: 0, tr_loss: 1163.93, vl_loss: 1086.84, tr_err: 2.04, vl_err: 1.58
time_elapsed for 200 steps: 1.117,

('Sample mini-batch results: \npredicted string: ', array(['', 'T', '', 'R', '', 'Y', 'S', 'A', '', 'J', 'B',
'A', 'E',

'A', 'N', 'Y', 'B', 'K', 'A', 'F', 'O', '', 'A', 'M', 'A', 'R',
'A', 'M', 'R', 'A', 'I', 'P', 'R', '', 'R', 'Y', '', 'A', '',
'A', 'S', 'E', '', 'H', 'N', 'Q', 'L', 'E', 'D', '', 'D',
'W', 'D', 'U', '', 'A', '', 'L', 'U', 'Q', 'C', '', 'I', 'E',
'P', 'U', 'H', 'U', 'M', 'D', 'P', '', 'A', '', 'P', '', 'A',
'', 'H', 'T', '', 'Z', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',
'', '', '', '', '',],
dtype='|S1'))

('actual string1: ', array(['I', '', 'W', 'O', 'U', 'L', 'D', '', 'L', 'I', 'K', 'E', '',
'T', 'O', '', 'F', 'I', 'N', 'D', '', 'A', 'N', '', 'I', 'N',
'D', 'I', 'A', 'N', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A',
'N', 'T', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H', 'O',
'U', 'L', 'D', '', 'B', 'E', '', 'I', 'N', '', 'T', 'H', 'E',
'', 'G', 'I', 'R', 'T', 'O', 'N', '', 'A', 'R', 'E', 'A'],
dtype='|S1'))

epoch 11, step: 0, tr_loss: 1112.55, vl_loss: 1037.14, tr_err: 1.58, vl_err: 1.22
epoch 12, step: 0, tr_loss: 1064.85, vl_loss: 992.15, tr_err: 1.22, vl_err: 1.07
epoch 13, step: 0, tr_loss: 1021.18, vl_loss: 950.26, tr_err: 1.07, vl_err: 0.95
epoch 14, step: 0, tr_loss: 980.22, vl_loss: 911.08, tr_err: 0.95, vl_err: 0.89
epoch 15, step: 0, tr_loss: 941.65, vl_loss: 874.63, tr_err: 0.89, vl_err: 0.84
epoch 16, step: 0, tr_loss: 905.56, vl_loss: 841.51, tr_err: 0.84, vl_err: 0.82
epoch 17, step: 0, tr_loss: 872.58, vl_loss: 810.78, tr_err: 0.82, vl_err: 0.83
epoch 18, step: 0, tr_loss: 841.83, vl_loss: 781.39, tr_err: 0.83, vl_err: 0.80
epoch 19, step: 0, tr_loss: 812.35, vl_loss: 753.43, tr_err: 0.80, vl_err: 0.83
epoch 20, step: 0, tr_loss: 784.24, vl_loss: 727.06, tr_err: 0.83, vl_err: 0.81
time_elapsed for 200 steps: 1.136,

('Sample mini-batch results: \npredicted string: ', array(['', 'N', '', 'A', '', 'A', '', 'E', '', 'E', 'K',
'U', '',

```

    'N', '', 'A', '', 'T', '', 'L', '', 'A', '', 'A', '', 'A',
    '', 'A', '', 'A', '' ],
    dtype='|S1'))
('actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',
    'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',
    'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',
    '', 'A', 'R', 'E', 'A'],
    dtype='|S1'))

```

epoch 21, step: 0, tr_loss: 757.66, vl_loss: 702.44, tr_err: 0.81, vl_err: 0.78
epoch 22, step: 0, tr_loss: 732.77, vl_loss: 679.32, tr_err: 0.78, vl_err: 0.72
epoch 23, step: 0, tr_loss: 709.34, vl_loss: 657.66, tr_err: 0.72, vl_err: 0.75
epoch 24, step: 0, tr_loss: 687.33, vl_loss: 637.28, tr_err: 0.75, vl_err: 0.74
epoch 25, step: 0, tr_loss: 666.57, vl_loss: 618.10, tr_err: 0.74, vl_err: 0.69
epoch 26, step: 0, tr_loss: 647.01, vl_loss: 599.96, tr_err: 0.69, vl_err: 0.68
epoch 27, step: 0, tr_loss: 628.47, vl_loss: 582.84, tr_err: 0.68, vl_err: 0.65
epoch 28, step: 0, tr_loss: 610.94, vl_loss: 566.64, tr_err: 0.65, vl_err: 0.57
epoch 29, step: 0, tr_loss: 594.33, vl_loss: 551.33, tr_err: 0.57, vl_err: 0.56
epoch 30, step: 0, tr_loss: 578.60, vl_loss: 536.83, tr_err: 0.56, vl_err: 0.53
time_elapsed for 200 steps: 1.126,

```

('Sample mini-batch results: \npredicted string: ', array([' ', 'O', '', 'O', 'I', 'N', '', 'I', '', 'N', 'T',
'E', 'R',
    '', 'E', 'T', '', 'O', '', 'E', 'T', 'I', '', 'N', '', 'L',
    'O', 'L', '', 'N', '', 'C', '', 'R', 'L', 'R', 'N', '' ],
    dtype='|S1'))
('actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',
    'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',
    'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',
    '', 'A', 'R', 'E', 'A'],
    dtype='|S1'))

```

epoch 31, step: 0, tr_loss: 563.69, vl_loss: 522.97, tr_err: 0.53, vl_err: 0.49
epoch 32, step: 0, tr_loss: 549.43, vl_loss: 509.72, tr_err: 0.49, vl_err: 0.40
epoch 33, step: 0, tr_loss: 535.77, vl_loss: 496.99, tr_err: 0.40, vl_err: 0.34
epoch 34, step: 0, tr_loss: 522.65, vl_loss: 484.79, tr_err: 0.34, vl_err: 0.26
epoch 35, step: 0, tr_loss: 510.07, vl_loss: 473.16, tr_err: 0.26, vl_err: 0.21
epoch 36, step: 0, tr_loss: 498.06, vl_loss: 462.08, tr_err: 0.21, vl_err: 0.21
epoch 37, step: 0, tr_loss: 486.61, vl_loss: 451.53, tr_err: 0.21, vl_err: 0.19
epoch 38, step: 0, tr_loss: 475.69, vl_loss: 441.44, tr_err: 0.19, vl_err: 0.20
epoch 39, step: 0, tr_loss: 465.24, vl_loss: 431.82, tr_err: 0.20, vl_err: 0.18
epoch 40, step: 0, tr_loss: 455.26, vl_loss: 422.57, tr_err: 0.18, vl_err: 0.16
time_elapsed for 200 steps: 1.126,

```

('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'O', 'K', 'I', 'N', 'G', '',
'F', 'O', '',
    'A', 'N', '', 'E', 'X', 'P', 'E', 'N', 'S', 'I', '', 'R', 'E',
    'S', 'T', 'A', 'U', 'R', 'A', 'N', 'T', '', '', '', '', '',
    '', '', '', '', '', '', '', '', '', '', '', '', '',

```


epoch 60, step: 0, tr_loss: 318.04, vl_loss: 295.71, tr_err: 0.09, vl_err: 0.09
time_elapsed for 200 steps: 1.130,
('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O',
'R', '', 'A', 'N', '', 'E', 'X', 'P', 'E', 'N', 'S', 'I', 'V',
'', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A', 'N', 'T', '', '',
'', '', '', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '',],
dtype='|S1'))
('actual string1: ', array(['I', '', 'W', 'O', 'U', 'L', 'D', '', 'L', 'I', 'K', 'E', '',
'T', 'O', '', 'F', 'I', 'N', 'D', '', 'A', 'N', '', 'I', 'N',
'D', 'I', 'A', 'N', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A',
'N', 'T', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H', 'O',
'U', 'L', 'D', '', 'B', 'E', '', 'I', 'N', '', 'T', 'H', 'E',
'', 'G', 'I', 'R', 'T', 'O', 'N', '', 'A', 'R', 'E', 'A'],
dtype='|S1'))

epoch 61, step: 0, tr_loss: 313.32, vl_loss: 291.34, tr_err: 0.09, vl_err: 0.09
epoch 62, step: 0, tr_loss: 308.73, vl_loss: 287.11, tr_err: 0.09, vl_err: 0.08
epoch 63, step: 0, tr_loss: 304.30, vl_loss: 283.01, tr_err: 0.08, vl_err: 0.09
epoch 64, step: 0, tr_loss: 299.99, vl_loss: 279.02, tr_err: 0.09, vl_err: 0.09
epoch 65, step: 0, tr_loss: 295.80, vl_loss: 275.14, tr_err: 0.09, vl_err: 0.09
epoch 66, step: 0, tr_loss: 291.72, vl_loss: 271.36, tr_err: 0.09, vl_err: 0.09
epoch 67, step: 0, tr_loss: 287.76, vl_loss: 267.69, tr_err: 0.09, vl_err: 0.09
epoch 68, step: 0, tr_loss: 283.90, vl_loss: 264.11, tr_err: 0.09, vl_err: 0.09
epoch 69, step: 0, tr_loss: 280.14, vl_loss: 260.63, tr_err: 0.09, vl_err: 0.09
epoch 70, step: 0, tr_loss: 276.47, vl_loss: 257.24, tr_err: 0.09, vl_err: 0.09
time_elapsed for 200 steps: 1.155,
('Sample mini-batch results: \npredicted string: ', array(['L', 'O', 'K', 'I', 'N', 'G', '', 'F', 'O', 'R', ' ', 'A', '',
'P', 'U', 'B', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H',
'O', 'U', '', 'H', '', 'E', '', 'I', 'N', 'T', 'E', 'R', 'N',
'E', 'T', '', 'C', 'O', 'N', 'E', 'C', 'T', 'I', 'O', 'N', '',
'A', 'N', 'D', '', 'A', 'L', 'O', 'W', '', 'C', 'H', 'I', 'L',
'D', 'R', 'E', 'N', '', ''],
dtype='|S1'))
('actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',
'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',
'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',
'', 'A', 'R', 'E', 'A'],
dtype='|S1'))

epoch 71, step: 0, tr_loss: 272.91, vl_loss: 253.94, tr_err: 0.09, vl_err: 0.09
epoch 72, step: 0, tr_loss: 269.43, vl_loss: 250.71, tr_err: 0.09, vl_err: 0.08
epoch 73, step: 0, tr_loss: 266.04, vl_loss: 247.56, tr_err: 0.08, vl_err: 0.08
epoch 74, step: 0, tr_loss: 262.72, vl_loss: 244.49, tr_err: 0.08, vl_err: 0.08

epoch 75, step: 0, tr_loss: 259.49, vl_loss: 241.50, tr_err: 0.08, vl_err: 0.08
epoch 76, step: 0, tr_loss: 256.34, vl_loss: 238.59, tr_err: 0.08, vl_err: 0.08
epoch 77, step: 0, tr_loss: 253.28, vl_loss: 235.74, tr_err: 0.08, vl_err: 0.08
epoch 78, step: 0, tr_loss: 250.27, vl_loss: 232.95, tr_err: 0.08, vl_err: 0.08
epoch 79, step: 0, tr_loss: 247.34, vl_loss: 230.23, tr_err: 0.08, vl_err: 0.08
epoch 80, step: 0, tr_loss: 244.47, vl_loss: 227.56, tr_err: 0.08, vl_err: 0.08
time_elapsed for 200 steps: 1.096,
('Sample mini-batch results: \npredicted string: ', array(['L', 'O', 'K', 'I', 'N', 'G', '', 'F', 'O', 'R', ' ',
', 'A', ''],
 'P', 'U', 'B', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H',
 'O', 'U', 'I', 'D', '', 'H', '', 'E', '', 'I', 'N', 'T', 'E',
 'R', 'N', 'E', 'T', '', 'C', 'O', 'N', 'E', 'C', 'T', 'I', 'O',
 'N', '', 'A', 'N', 'D', '', 'A', 'L', 'O', 'W', '', 'C', 'H',
 'I', 'L', 'D', 'R', 'E', 'N', '', ''],
 dtype='|S1'))
('actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',
 'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',
 'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',
 '', 'A', 'R', 'E', 'A'],
 dtype='|S1'))

epoch 81, step: 0, tr_loss: 241.66, vl_loss: 224.96, tr_err: 0.08, vl_err: 0.08
epoch 82, step: 0, tr_loss: 238.92, vl_loss: 222.40, tr_err: 0.08, vl_err: 0.08
epoch 83, step: 0, tr_loss: 236.22, vl_loss: 219.90, tr_err: 0.08, vl_err: 0.08
epoch 84, step: 0, tr_loss: 233.58, vl_loss: 217.45, tr_err: 0.08, vl_err: 0.08
epoch 85, step: 0, tr_loss: 231.00, vl_loss: 215.05, tr_err: 0.08, vl_err: 0.07
epoch 86, step: 0, tr_loss: 228.48, vl_loss: 212.71, tr_err: 0.07, vl_err: 0.07
epoch 87, step: 0, tr_loss: 226.01, vl_loss: 210.41, tr_err: 0.07, vl_err: 0.07
epoch 88, step: 0, tr_loss: 223.58, vl_loss: 208.16, tr_err: 0.07, vl_err: 0.07
epoch 89, step: 0, tr_loss: 221.21, vl_loss: 205.96, tr_err: 0.07, vl_err: 0.08
epoch 90, step: 0, tr_loss: 218.89, vl_loss: 203.80, tr_err: 0.08, vl_err: 0.08
time_elapsed for 200 steps: 1.115,
('Sample mini-batch results: \npredicted string: ', array(['L', 'O', 'K', 'I', 'N', 'G', '', 'F', 'O', 'R', ' ',
', 'A', ''],
 'P', 'U', 'B', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H',
 'O', 'C', 'I', 'T', '', 'H', 'E', '', 'I', 'N', 'T', 'E', 'R',
 'N', 'E', 'T', '', 'C', 'O', 'N', 'E', 'C', 'T', 'I', 'O', 'N',
 '', 'A', 'N', 'D', '', 'A', 'L', 'O', 'W', '', 'C', 'H', 'I',
 'L', 'D', 'R', 'E', 'N', '', ' ', ' ', ' ', ''],
 dtype='|S1'))
('actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',
 'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',
 'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',
 '', 'A', 'R', 'E', 'A'],
 dtype='|S1'))

epoch 91, step: 0, tr_loss: 216.62, vl_loss: 201.69, tr_err: 0.08, vl_err: 0.07

epoch 92, step: 0, tr_loss: 214.39, vl_loss: 199.63, tr_err: 0.07, vl_err: 0.07
epoch 93, step: 0, tr_loss: 212.21, vl_loss: 197.60, tr_err: 0.07, vl_err: 0.07
epoch 94, step: 0, tr_loss: 210.07, vl_loss: 195.61, tr_err: 0.07, vl_err: 0.07
epoch 95, step: 0, tr_loss: 207.97, vl_loss: 193.67, tr_err: 0.07, vl_err: 0.07
epoch 96, step: 0, tr_loss: 205.92, vl_loss: 191.76, tr_err: 0.07, vl_err: 0.07
epoch 97, step: 0, tr_loss: 203.90, vl_loss: 189.88, tr_err: 0.07, vl_err: 0.07
epoch 98, step: 0, tr_loss: 201.92, vl_loss: 188.04, tr_err: 0.07, vl_err: 0.07
epoch 99, step: 0, tr_loss: 199.97, vl_loss: 186.23, tr_err: 0.07, vl_err: 0.07
epoch 100, step: 0, tr_loss: 198.06, vl_loss: 184.46, tr_err: 0.07, vl_err: 0.07
time_elapsed for 200 steps: 1.137,

('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O',

'R', '', 'P', 'U', 'B', '', 'F', 'O', 'D', '', 'R', 'E', 'S',
'T', 'A', 'U', 'R', 'A', 'N', 'T', '', 'A', 'N', 'D', '', 'I',
'T', '', 'S', 'H', 'O', 'U', 'L', 'D', '', 'B', 'E', '', 'I',
'N', '', 'T', 'H', 'E', '', 'U', 'I', 'D', 'Y', '', 'C', 'E',
'N', 'T', 'E', 'R', '', 'A', 'R', 'E', 'A'],

dtype='|S1'))

('actual string1: ', array(['I', '', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', ' ', 'F',
'O', 'R', '', 'A', '', 'M', 'E', 'D', 'I', 'T', 'E', 'R', 'R',
'A', 'N', 'E', 'A', 'N', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R',
'A', 'N', 'T'],

dtype='|S1'))

epoch 101, step: 0, tr_loss: 196.19, vl_loss: 182.72, tr_err: 0.07, vl_err: 0.07
epoch 102, step: 0, tr_loss: 194.35, vl_loss: 181.01, tr_err: 0.07, vl_err: 0.07
epoch 103, step: 0, tr_loss: 192.55, vl_loss: 179.34, tr_err: 0.07, vl_err: 0.07
epoch 104, step: 0, tr_loss: 190.78, vl_loss: 177.69, tr_err: 0.07, vl_err: 0.07
epoch 105, step: 0, tr_loss: 189.04, vl_loss: 176.08, tr_err: 0.07, vl_err: 0.07
epoch 106, step: 0, tr_loss: 187.34, vl_loss: 174.49, tr_err: 0.07, vl_err: 0.07
epoch 107, step: 0, tr_loss: 185.66, vl_loss: 172.93, tr_err: 0.07, vl_err: 0.07
epoch 108, step: 0, tr_loss: 184.01, vl_loss: 171.39, tr_err: 0.07, vl_err: 0.07
epoch 109, step: 0, tr_loss: 182.39, vl_loss: 169.89, tr_err: 0.07, vl_err: 0.08
epoch 110, step: 0, tr_loss: 180.79, vl_loss: 168.40, tr_err: 0.08, vl_err: 0.07
time_elapsed for 200 steps: 1.222,

('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O',

'R', '', 'P', 'U', 'B', '', 'F', 'O', 'D', '', 'R', 'E', 'S',
'T', 'A', 'U', 'R', 'A', 'N', 'T', '', 'A', 'N', 'D', '', 'I',
'T', '', 'S', 'H', 'O', 'U', 'L', 'D', '', 'B', 'E', '', 'I',
'N', '', 'T', 'H', 'E', '', 'U', 'I', 'D', 'Y', '', 'C', 'E',
'N', 'T', 'E', 'R', '', 'A', 'R', 'E', 'A'],

dtype='|S1'))

('actual string1: ', array(['I', '', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', ' ', 'F',
'O', 'R', '', 'A', '', 'M', 'E', 'D', 'I', 'T', 'E', 'R', 'R',
'A', 'N', 'E', 'A', 'N', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R',
'A', 'N', 'T'],

dtype='|S1'))

epoch 111, step: 0, tr_loss: 179.23, vl_loss: 166.95, tr_err: 0.07, vl_err: 0.07
epoch 112, step: 0, tr_loss: 177.69, vl_loss: 165.51, tr_err: 0.07, vl_err: 0.07
epoch 113, step: 0, tr_loss: 176.17, vl_loss: 164.11, tr_err: 0.07, vl_err: 0.07
epoch 114, step: 0, tr_loss: 174.68, vl_loss: 162.72, tr_err: 0.07, vl_err: 0.07
epoch 115, step: 0, tr_loss: 173.22, vl_loss: 161.36, tr_err: 0.07, vl_err: 0.07
epoch 116, step: 0, tr_loss: 171.78, vl_loss: 160.03, tr_err: 0.07, vl_err: 0.07
epoch 117, step: 0, tr_loss: 170.36, vl_loss: 158.71, tr_err: 0.07, vl_err: 0.07
epoch 118, step: 0, tr_loss: 168.97, vl_loss: 157.41, tr_err: 0.07, vl_err: 0.07
epoch 119, step: 0, tr_loss: 167.60, vl_loss: 156.13, tr_err: 0.07, vl_err: 0.07
epoch 120, step: 0, tr_loss: 166.24, vl_loss: 154.86, tr_err: 0.07, vl_err: 0.07

time_elapsed for 200 steps: 1.233,

('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O',

'R', '', 'P', 'U', 'B', '', 'F', 'O', 'D', '', 'R', 'E', 'S',
'T', 'A', 'U', 'R', 'A', 'N', 'T', '', 'A', 'N', 'D', '', 'I',
'T', '', 'S', 'H', 'O', 'U', 'L', 'D', '', 'B', 'E', '', 'I',
'N', '', 'T', 'H', 'E', '', 'U', 'L', 'D', 'Y', '', 'C', 'E',
'N', 'T', 'E', 'R', '', 'A', 'R', 'E', 'A'],

dtype='|S1'))

('actual string1: ', array(['I', '', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', ' ', 'F',

'O', 'R', '', 'A', '', 'M', 'E', 'D', 'I', 'T', 'E', 'R', 'R',
'A', 'N', 'E', 'A', 'N', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R',
'A', 'N', 'T'],

dtype='|S1'))

epoch 121, step: 0, tr_loss: 164.90, vl_loss: 153.62, tr_err: 0.07, vl_err: 0.07
epoch 122, step: 0, tr_loss: 163.59, vl_loss: 152.39, tr_err: 0.07, vl_err: 0.07
epoch 123, step: 0, tr_loss: 162.29, vl_loss: 151.19, tr_err: 0.07, vl_err: 0.07
epoch 124, step: 0, tr_loss: 161.01, vl_loss: 150.00, tr_err: 0.07, vl_err: 0.07
epoch 125, step: 0, tr_loss: 159.75, vl_loss: 148.83, tr_err: 0.07, vl_err: 0.07
epoch 126, step: 0, tr_loss: 158.52, vl_loss: 147.68, tr_err: 0.07, vl_err: 0.07
epoch 127, step: 0, tr_loss: 157.30, vl_loss: 146.55, tr_err: 0.07, vl_err: 0.07
epoch 128, step: 0, tr_loss: 156.11, vl_loss: 145.44, tr_err: 0.07, vl_err: 0.07
epoch 129, step: 0, tr_loss: 154.93, vl_loss: 144.34, tr_err: 0.07, vl_err: 0.07
epoch 130, step: 0, tr_loss: 153.76, vl_loss: 143.26, tr_err: 0.07, vl_err: 0.07

time_elapsed for 200 steps: 1.242,

('Sample mini-batch results: \npredicted string: ', array(['I', '', 'M', '', 'L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O',

'R', '', 'A', 'N', '', 'E', 'X', 'P', 'E', 'N', 'S', 'I', 'V',
'E', '', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A', 'N', 'T', '',
'', '', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '',],

dtype='|S1'))

('actual string1: ', array(['I', '', 'W', 'O', 'U', 'L', 'D', '', 'L', 'I', 'K', 'E', ' ',

```
'T', 'O', ' ', 'F', 'I', 'N', 'D', ' ', 'A', 'N', ' ', 'I', 'N',  
'D', 'I', 'A', 'N', ' ', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A',  
'N', 'T', ' ', 'A', 'N', 'D', ' ', 'I', 'T', ' ', 'S', 'H', 'O',  
'U', 'L', 'D', ' ', 'B', 'E', ' ', 'I', 'N', ' ', 'T', 'H', 'E',  
' ', 'G', 'I', 'R', 'T', 'O', 'N', ' ', 'A', 'R', 'E', 'A',  
dtype='|S1'))
```

epoch 131, step: 0, tr_loss: 152.62, vl_loss: 142.19, tr_err: 0.07, vl_err: 0.07
epoch 132, step: 0, tr_loss: 151.49, vl_loss: 141.14, tr_err: 0.07, vl_err: 0.07
epoch 133, step: 0, tr_loss: 150.37, vl_loss: 140.10, tr_err: 0.07, vl_err: 0.07
epoch 134, step: 0, tr_loss: 149.28, vl_loss: 139.08, tr_err: 0.07, vl_err: 0.07
epoch 135, step: 0, tr_loss: 148.20, vl_loss: 138.08, tr_err: 0.07, vl_err: 0.07
epoch 136, step: 0, tr_loss: 147.13, vl_loss: 137.09, tr_err: 0.07, vl_err: 0.07
epoch 137, step: 0, tr_loss: 146.08, vl_loss: 136.11, tr_err: 0.07, vl_err: 0.07
epoch 138, step: 0, tr_loss: 145.05, vl_loss: 135.15, tr_err: 0.07, vl_err: 0.07
epoch 139, step: 0, tr_loss: 144.03, vl_loss: 134.20, tr_err: 0.07, vl_err: 0.07
epoch 140, step: 0, tr_loss: 143.02, vl_loss: 133.26, tr_err: 0.07, vl_err: 0.07
time_elapsed for 200 steps: 1.199,

('Sample mini-batch results: \npredicted string: ', array(['I', ' ', 'M', ' ', 'L', 'O', 'K', 'I', 'N', 'G', ' ',
'F', 'O',

```
'R', ' ', 'A', 'N', ' ', 'E', 'X', 'P', 'E', 'N', 'S', 'I', 'V',  
'E', ' ', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A', 'N', 'T', ' ',  
' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ',  
' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ',  
' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ', ' ',  
dtype='|S1'))
```

('actual string1: ', array(['I', ' ', 'W', 'O', 'U', 'L', 'D', ' ', 'L', 'I', 'K', 'E', ' ',
'T', 'O', ' ', 'F', 'I', 'N', 'D', ' ', 'A', 'N', ' ', 'I', 'N',
'D', 'I', 'A', 'N', ' ', 'R', 'E', 'S', 'T', 'A', 'U', 'R', 'A',
'N', 'T', ' ', 'A', 'N', 'D', ' ', 'I', 'T', ' ', 'S', 'H', 'O',
'U', 'L', 'D', ' ', 'B', 'E', ' ', 'I', 'N', ' ', 'T', 'H', 'E',
' ', 'G', 'I', 'R', 'T', 'O', 'N', ' ', 'A', 'R', 'E', 'A',
dtype='|S1'))

epoch 141, step: 0, tr_loss: 142.03, vl_loss: 132.34, tr_err: 0.07, vl_err: 0.07
epoch 142, step: 0, tr_loss: 141.05, vl_loss: 131.43, tr_err: 0.07, vl_err: 0.07
epoch 143, step: 0, tr_loss: 140.09, vl_loss: 130.53, tr_err: 0.07, vl_err: 0.07
epoch 144, step: 0, tr_loss: 139.13, vl_loss: 129.64, tr_err: 0.07, vl_err: 0.07
epoch 145, step: 0, tr_loss: 138.19, vl_loss: 128.77, tr_err: 0.07, vl_err: 0.07
epoch 146, step: 0, tr_loss: 137.27, vl_loss: 127.91, tr_err: 0.07, vl_err: 0.07
epoch 147, step: 0, tr_loss: 136.35, vl_loss: 127.05, tr_err: 0.07, vl_err: 0.07
epoch 148, step: 0, tr_loss: 135.45, vl_loss: 126.22, tr_err: 0.07, vl_err: 0.07
epoch 149, step: 0, tr_loss: 134.56, vl_loss: 125.39, tr_err: 0.07, vl_err: 0.07
epoch 150, step: 0, tr_loss: 133.68, vl_loss: 124.57, tr_err: 0.07, vl_err: 0.07
time_elapsed for 200 steps: 1.209,

('Sample mini-batch results: \npredicted string: ', array(['L', 'O', 'K', 'I', 'N', 'G', ' ', 'F', 'O', 'R', ' ',
'A', ' ',


```
'P', 'U', 'B', '', 'A', 'N', 'D', '', 'I', 'T', '', 'S', 'H',  
'O', 'C', 'L', 'D', '', 'H', 'A', 'E', '', 'I', 'N', 'T', 'E',  
'R', 'N', 'E', 'T', '', 'C', 'O', 'N', 'E', 'C', 'T', 'I', 'O',  
'N', '', 'A', 'N', 'D', '', 'A', 'L', 'O', 'W', '', 'C', 'H',  
'I', 'L', 'D', 'R', 'E', 'N', '', '', ''],  
dtype='|S1')  
(actual string1: ', array(['I', '', 'A', 'M', '', 'L', 'O', 'O', 'K', 'I', 'N', 'G', '',  
    'F', 'O', 'R', '', 'A', '', 'P', 'U', 'B', '', 'I', 'N', '',  
    'T', 'H', 'E', '', 'R', 'I', 'V', 'E', 'R', 'S', 'I', 'D', 'E',  
    '', 'A', 'R', 'E', 'A'],  
dtype='|S1'))
```

Finished

Process finished with exit code 0