```
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.1, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 5.00 | ms/batch  7.27 | loss  7.05
| epoch   1 |   400/  686 batches | lr 5.00 | ms/batch  6.13 | loss  6.49
| epoch   1 |   600/  686 batches | lr 5.00 | ms/batch  6.01 | loss  6.04
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.49s | valid loss  6.11 | valid acc   0.0183
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 5.00 | ms/batch  6.08 | loss  5.87
| epoch   2 |   400/  686 batches | lr 5.00 | ms/batch  6.10 | loss  5.76
| epoch   2 |   600/  686 batches | lr 5.00 | ms/batch  6.02 | loss  5.64
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.23s | valid loss  5.84 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 5.00 | ms/batch  6.03 | loss  5.52
| epoch   3 |   400/  686 batches | lr 5.00 | ms/batch  6.03 | loss  5.42
| epoch   3 |   600/  686 batches | lr 5.00 | ms/batch  6.08 | loss  5.31
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.25s | valid loss  5.69 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 5.00 | ms/batch  6.25 | loss  5.22
| epoch   4 |   400/  686 batches | lr 5.00 | ms/batch  6.18 | loss  5.13
| epoch   4 |   600/  686 batches | lr 5.00 | ms/batch  6.09 | loss  5.04
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.34s | valid loss  5.62 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 5.00 | ms/batch  6.13 | loss  4.96
| epoch   5 |   400/  686 batches | lr 5.00 | ms/batch  6.13 | loss  4.87
| epoch   5 |   600/  686 batches | lr 5.00 | ms/batch  6.15 | loss  4.79
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.23s | valid loss  5.61 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 5.00 | ms/batch  6.15 | loss  4.72
| epoch   6 |   400/  686 batches | lr 5.00 | ms/batch  6.03 | loss  4.64
| epoch   6 |   600/  686 batches | lr 5.00 | ms/batch  6.15 | loss  4.57
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.26s | valid loss  5.62 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 5.00 | ms/batch  5.94 | loss  4.47
| epoch   7 |   400/  686 batches | lr 5.00 | ms/batch  6.23 | loss  4.40
| epoch   7 |   600/  686 batches | lr 5.00 | ms/batch  6.20 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.28s | valid loss  5.61 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 5.00 | ms/batch  6.17 | loss  4.38
| epoch   8 |   400/  686 batches | lr 5.00 | ms/batch  6.14 | loss  4.32
| epoch   8 |   600/  686 batches | lr 5.00 | ms/batch  6.12 | loss  4.29
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.28s | valid loss  5.62 | valid acc   0.0442
-----------------------------------------------------------------------------------------
```

```
| epoch   9 |   200/  686 batches | lr 5.00 | ms/batch  6.41 | loss  4.32
| epoch   9 |   400/  686 batches | lr 5.00 | ms/batch  6.39 | loss  4.26
| epoch   9 |   600/  686 batches | lr 5.00 | ms/batch  6.20 | loss  4.24
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.39s | valid loss  5.62 | valid acc   0.0453
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 5.00 | ms/batch  6.37 | loss  4.29
| epoch  10 |   400/  686 batches | lr 5.00 | ms/batch  6.13 | loss  4.25
| epoch  10 |   600/  686 batches | lr 5.00 | ms/batch  6.17 | loss  4.22
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.33s | valid loss  5.62 | valid acc   0.0442
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.70 | test acc   0.0105
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.1, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 5.00 | ms/batch 10.08 | loss  7.03
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.34s | valid loss  6.66 | valid acc   0.0097
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 5.00 | ms/batch  8.93 | loss  6.16
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.13s | valid loss  6.06 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 5.00 | ms/batch  8.95 | loss  5.85
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.12s | valid loss  5.96 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 5.00 | ms/batch  8.47 | loss  5.66
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.02s | valid loss  5.83 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 5.00 | ms/batch  8.94 | loss  5.47
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.12s | valid loss  5.74 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 5.00 | ms/batch  8.95 | loss  5.29
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.11s | valid loss  5.67 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 5.00 | ms/batch  8.97 | loss  5.11
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.12s | valid loss  5.64 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 5.00 | ms/batch  8.69 | loss  4.95
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.08s | valid loss  5.61 | valid acc   0.0399
```

```
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 5.00 | ms/batch  8.99 | loss  4.80
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.08s | valid loss  5.62 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 5.00 | ms/batch  8.80 | loss  4.62
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.10s | valid loss  5.59 | valid acc   0.0453
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.68 | test acc   0.0116
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.1, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 5.00 | ms/batch 12.35 | loss  7.05
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.85s | valid loss  6.82 | valid acc   0.0065
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 5.00 | ms/batch 11.33 | loss  6.49
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.64s | valid loss  6.34 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 5.00 | ms/batch 11.30 | loss  6.01
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.64s | valid loss  6.07 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 5.00 | ms/batch 11.37 | loss  5.85
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.65s | valid loss  6.02 | valid acc   0.0259
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 5.00 | ms/batch 11.22 | loss  5.72
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.62s | valid loss  5.88 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 5.00 | ms/batch 11.47 | loss  5.58
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.68s | valid loss  5.77 | valid acc   0.0248
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 5.00 | ms/batch 11.22 | loss  5.45
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.62s | valid loss  5.73 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 5.00 | ms/batch 11.26 | loss  5.32
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.63s | valid loss  5.66 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 5.00 | ms/batch 11.32 | loss  5.20
-----------------------------------------------------------------------------------------
```

```
| end of epoch   9 | time:  2.66s | valid loss  5.65 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 5.00 | ms/batch 11.36 | loss  5.08
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.65s | valid loss  5.63 | valid acc   0.0388
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.69 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.2, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 5.00 | ms/batch  7.11 | loss  7.05
| epoch   1 |   400/  686 batches | lr 5.00 | ms/batch  6.06 | loss  6.52
| epoch   1 |   600/  686 batches | lr 5.00 | ms/batch  6.06 | loss  6.05
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.43s | valid loss  6.11 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 5.00 | ms/batch  6.07 | loss  5.89
| epoch   2 |   400/  686 batches | lr 5.00 | ms/batch  6.04 | loss  5.77
| epoch   2 |   600/  686 batches | lr 5.00 | ms/batch  6.10 | loss  5.66
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.23s | valid loss  5.85 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 5.00 | ms/batch  6.11 | loss  5.56
| epoch   3 |   400/  686 batches | lr 5.00 | ms/batch  6.16 | loss  5.46
| epoch   3 |   600/  686 batches | lr 5.00 | ms/batch  6.11 | loss  5.35
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.28s | valid loss  5.68 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 5.00 | ms/batch  6.19 | loss  5.27
| epoch   4 |   400/  686 batches | lr 5.00 | ms/batch  6.27 | loss  5.18
| epoch   4 |   600/  686 batches | lr 5.00 | ms/batch  6.07 | loss  5.09
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.31s | valid loss  5.61 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 5.00 | ms/batch  6.22 | loss  5.02
| epoch   5 |   400/  686 batches | lr 5.00 | ms/batch  6.12 | loss  4.94
| epoch   5 |   600/  686 batches | lr 5.00 | ms/batch  6.27 | loss  4.87
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.31s | valid loss  5.59 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 5.00 | ms/batch  6.16 | loss  4.81
| epoch   6 |   400/  686 batches | lr 5.00 | ms/batch  6.04 | loss  4.73
| epoch   6 |   600/  686 batches | lr 5.00 | ms/batch  6.00 | loss  4.67
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.23s | valid loss  5.60 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 5.00 | ms/batch  6.16 | loss  4.59
```

```
| epoch   7 |   400/  686 batches | lr 5.00 | ms/batch  6.18 | loss  4.52
| epoch   7 |   600/  686 batches | lr 5.00 | ms/batch  6.25 | loss  4.48
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.32s | valid loss  5.58 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 5.00 | ms/batch  6.31 | loss  4.51
| epoch   8 |   400/  686 batches | lr 5.00 | ms/batch  6.13 | loss  4.46
| epoch   8 |   600/  686 batches | lr 5.00 | ms/batch  6.07 | loss  4.43
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.30s | valid loss  5.58 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 5.00 | ms/batch  6.12 | loss  4.45
| epoch   9 |   400/  686 batches | lr 5.00 | ms/batch  6.08 | loss  4.41
| epoch   9 |   600/  686 batches | lr 5.00 | ms/batch  6.22 | loss  4.38
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.27s | valid loss  5.59 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 5.00 | ms/batch  6.10 | loss  4.43
| epoch  10 |   400/  686 batches | lr 5.00 | ms/batch  6.05 | loss  4.39
| epoch  10 |   600/  686 batches | lr 5.00 | ms/batch  6.07 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.24s | valid loss  5.59 | valid acc   0.0431
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.65 | test acc   0.0105
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.2, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 5.00 | ms/batch  9.98 | loss  7.04
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.33s | valid loss  6.67 | valid acc   0.0097
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 5.00 | ms/batch  9.04 | loss  6.16
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.14s | valid loss  6.19 | valid acc   0.0237
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 5.00 | ms/batch  8.94 | loss  5.86
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.12s | valid loss  5.93 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 5.00 | ms/batch  8.95 | loss  5.68
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.11s | valid loss  5.84 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 5.00 | ms/batch  9.00 | loss  5.49
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.12s | valid loss  5.74 | valid acc   0.0356
-----------------------------------------------------------------------------------------
```

```
| epoch   6 |   200/  343 batches | lr 5.00 | ms/batch  8.91 | loss  5.33
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.05s | valid loss  5.67 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 5.00 | ms/batch  8.90 | loss  5.17
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.07s | valid loss  5.62 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 5.00 | ms/batch  8.32 | loss  5.02
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.90s | valid loss  5.59 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 5.00 | ms/batch  8.43 | loss  4.89
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.92s | valid loss  5.61 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 5.00 | ms/batch  8.45 | loss  4.73
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.02s | valid loss  5.56 | valid acc   0.0431
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.64 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.2, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 5.00 | ms/batch 12.60 | loss  7.04
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.91s | valid loss  6.79 | valid acc   0.0065
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 5.00 | ms/batch 11.40 | loss  6.50
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.66s | valid loss  6.25 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 5.00 | ms/batch 11.35 | loss  6.02
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.65s | valid loss  6.08 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 5.00 | ms/batch 11.40 | loss  5.86
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.66s | valid loss  5.96 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 5.00 | ms/batch 11.48 | loss  5.74
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.66s | valid loss  5.92 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 5.00 | ms/batch 11.42 | loss  5.61
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.67s | valid loss  5.82 | valid acc   0.0313
```

```
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 5.00 | ms/batch 11.41 | loss  5.48
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.66s | valid loss  5.72 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 5.00 | ms/batch 11.37 | loss  5.37
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.66s | valid loss  5.69 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 5.00 | ms/batch 11.29 | loss  5.25
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.64s | valid loss  5.63 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 5.00 | ms/batch 11.42 | loss  5.15
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.67s | valid loss  5.59 | valid acc   0.0367
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.66 | test acc   0.0116
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.5, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 5.00 | ms/batch  7.20 | loss  7.05
| epoch   1 |   400/  686 batches | lr 5.00 | ms/batch  6.09 | loss  6.55
| epoch   1 |   600/  686 batches | lr 5.00 | ms/batch  6.18 | loss  6.08
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.49s | valid loss  6.11 | valid acc   0.0194
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 5.00 | ms/batch  6.30 | loss  5.93
| epoch   2 |   400/  686 batches | lr 5.00 | ms/batch  6.17 | loss  5.83
| epoch   2 |   600/  686 batches | lr 5.00 | ms/batch  6.16 | loss  5.74
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.33s | valid loss  5.89 | valid acc   0.0280
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 5.00 | ms/batch  6.19 | loss  5.67
| epoch   3 |   400/  686 batches | lr 5.00 | ms/batch  6.14 | loss  5.59
| epoch   3 |   600/  686 batches | lr 5.00 | ms/batch  6.19 | loss  5.50
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.30s | valid loss  5.72 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 5.00 | ms/batch  6.29 | loss  5.45
| epoch   4 |   400/  686 batches | lr 5.00 | ms/batch  6.13 | loss  5.38
| epoch   4 |   600/  686 batches | lr 5.00 | ms/batch  6.14 | loss  5.31
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.31s | valid loss  5.63 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 5.00 | ms/batch  6.23 | loss  5.27
| epoch   5 |   400/  686 batches | lr 5.00 | ms/batch  6.31 | loss  5.21
```

```
| epoch   5 |   600/  686 batches | lr 5.00 | ms/batch  6.30 | loss  5.15
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.39s | valid loss  5.59 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 5.00 | ms/batch  6.20 | loss  5.12
| epoch   6 |   400/  686 batches | lr 5.00 | ms/batch  6.19 | loss  5.07
| epoch   6 |   600/  686 batches | lr 5.00 | ms/batch  6.10 | loss  5.02
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.30s | valid loss  5.56 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 5.00 | ms/batch  6.29 | loss  5.00
| epoch   7 |   400/  686 batches | lr 5.00 | ms/batch  6.06 | loss  4.94
| epoch   7 |   600/  686 batches | lr 5.00 | ms/batch  6.15 | loss  4.90
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.31s | valid loss  5.58 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 5.00 | ms/batch  6.28 | loss  4.86
| epoch   8 |   400/  686 batches | lr 5.00 | ms/batch  6.16 | loss  4.81
| epoch   8 |   600/  686 batches | lr 5.00 | ms/batch  6.19 | loss  4.78
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.32s | valid loss  5.55 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 5.00 | ms/batch  6.17 | loss  4.81
| epoch   9 |   400/  686 batches | lr 5.00 | ms/batch  6.16 | loss  4.77
| epoch   9 |   600/  686 batches | lr 5.00 | ms/batch  6.32 | loss  4.74
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.35s | valid loss  5.55 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 5.00 | ms/batch  6.34 | loss  4.77
| epoch  10 |   400/  686 batches | lr 5.00 | ms/batch  6.14 | loss  4.73
| epoch  10 |   600/  686 batches | lr 5.00 | ms/batch  6.27 | loss  4.71
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.36s | valid loss  5.55 | valid acc   0.0399
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.62 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.5, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 5.00 | ms/batch 10.13 | loss  7.03
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.36s | valid loss  6.48 | valid acc   0.0194
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 5.00 | ms/batch  9.00 | loss  6.15
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.14s | valid loss  6.08 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 5.00 | ms/batch  8.80 | loss  5.90
```

```
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.05s | valid loss  5.93 | valid acc   0.0280
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 5.00 | ms/batch  8.97 | loss  5.75
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.13s | valid loss  5.87 | valid acc   0.0280
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 5.00 | ms/batch  9.00 | loss  5.61
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.14s | valid loss  5.77 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 5.00 | ms/batch  8.90 | loss  5.49
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.12s | valid loss  5.70 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 5.00 | ms/batch  9.02 | loss  5.37
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.14s | valid loss  5.65 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 5.00 | ms/batch  9.00 | loss  5.27
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.13s | valid loss  5.62 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 5.00 | ms/batch  9.00 | loss  5.16
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.15s | valid loss  5.60 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 5.00 | ms/batch  8.86 | loss  5.07
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.11s | valid loss  5.58 | valid acc   0.0399
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.65 | test acc   0.0105
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 5, dropout: 0.5, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 5.00 | ms/batch 12.45 | loss  7.05
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.87s | valid loss  6.77 | valid acc   0.0097
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 5.00 | ms/batch 11.35 | loss  6.48
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.65s | valid loss  6.24 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 5.00 | ms/batch 11.28 | loss  6.04
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.64s | valid loss  6.09 | valid acc   0.0259
-----------------------------------------------------------------------------------------
```

```
| epoch   4 |   200/  228 batches | lr 5.00 | ms/batch 11.32 | loss  5.90
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.64s | valid loss  6.00 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 5.00 | ms/batch 10.22 | loss  5.80
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.39s | valid loss  5.91 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 5.00 | ms/batch 10.72 | loss  5.70
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.53s | valid loss  5.85 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 5.00 | ms/batch 11.29 | loss  5.60
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.64s | valid loss  5.76 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 5.00 | ms/batch 11.37 | loss  5.51
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.65s | valid loss  5.71 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 5.00 | ms/batch 11.35 | loss  5.42
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.65s | valid loss  5.66 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 5.00 | ms/batch 10.91 | loss  5.34
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.54s | valid loss  5.61 | valid acc   0.0378
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.68 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.1, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 10.00 | ms/batch  7.11 | loss  7.02
| epoch   1 |   400/  686 batches | lr 10.00 | ms/batch  6.07 | loss  6.07
| epoch   1 |   600/  686 batches | lr 10.00 | ms/batch  6.02 | loss  5.81
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.42s | valid loss  5.94 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 10.00 | ms/batch  6.08 | loss  5.57
| epoch   2 |   400/  686 batches | lr 10.00 | ms/batch  6.00 | loss  5.38
| epoch   2 |   600/  686 batches | lr 10.00 | ms/batch  6.05 | loss  5.22
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.22s | valid loss  5.63 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 10.00 | ms/batch  6.07 | loss  5.05
| epoch   3 |   400/  686 batches | lr 10.00 | ms/batch  6.17 | loss  4.90
| epoch   3 |   600/  686 batches | lr 10.00 | ms/batch  6.04 | loss  4.76
```

```
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.25s | valid loss  5.63 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 10.00 | ms/batch  6.08 | loss  4.61
| epoch   4 |   400/  686 batches | lr 10.00 | ms/batch  6.04 | loss  4.47
| epoch   4 |   600/  686 batches | lr 10.00 | ms/batch  6.06 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.23s | valid loss  5.69 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 10.00 | ms/batch  6.09 | loss  4.18
| epoch   5 |   400/  686 batches | lr 10.00 | ms/batch  6.04 | loss  4.06
| epoch   5 |   600/  686 batches | lr 10.00 | ms/batch  6.03 | loss  3.99
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.22s | valid loss  5.67 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 10.00 | ms/batch  6.05 | loss  4.02
| epoch   6 |   400/  686 batches | lr 10.00 | ms/batch  5.99 | loss  3.93
| epoch   6 |   600/  686 batches | lr 10.00 | ms/batch  6.10 | loss  3.89
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.21s | valid loss  5.68 | valid acc   0.0442
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 10.00 | ms/batch  6.04 | loss  3.97
| epoch   7 |   400/  686 batches | lr 10.00 | ms/batch  6.14 | loss  3.90
| epoch   7 |   600/  686 batches | lr 10.00 | ms/batch  6.03 | loss  3.86
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.23s | valid loss  5.68 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 10.00 | ms/batch  6.10 | loss  3.96
| epoch   8 |   400/  686 batches | lr 10.00 | ms/batch  6.04 | loss  3.89
| epoch   8 |   600/  686 batches | lr 10.00 | ms/batch  6.02 | loss  3.85
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.22s | valid loss  5.68 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 10.00 | ms/batch  6.08 | loss  3.96
| epoch   9 |   400/  686 batches | lr 10.00 | ms/batch  6.02 | loss  3.89
| epoch   9 |   600/  686 batches | lr 10.00 | ms/batch  6.02 | loss  3.85
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.21s | valid loss  5.68 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 10.00 | ms/batch  6.08 | loss  3.96
| epoch  10 |   400/  686 batches | lr 10.00 | ms/batch  6.20 | loss  3.89
| epoch  10 |   600/  686 batches | lr 10.00 | ms/batch  6.31 | loss  3.85
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.31s | valid loss  5.68 | valid acc   0.0431
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.70 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
```

learning_rate: 10, dropout: 0.1, batch_size: 64

```
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 10.00 | ms/batch 10.20 | loss  7.03
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.37s | valid loss  6.15 | valid acc   0.0291
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 10.00 | ms/batch  9.00 | loss  5.90
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.13s | valid loss  5.88 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 10.00 | ms/batch  9.06 | loss  5.49
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.13s | valid loss  5.71 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 10.00 | ms/batch  8.97 | loss  5.18
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.13s | valid loss  5.64 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 10.00 | ms/batch  8.99 | loss  4.92
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.11s | valid loss  5.63 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 10.00 | ms/batch  8.74 | loss  4.66
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.03s | valid loss  5.67 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 10.00 | ms/batch  8.28 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.97s | valid loss  5.64 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 10.00 | ms/batch  8.74 | loss  4.27
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.07s | valid loss  5.64 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 10.00 | ms/batch  8.56 | loss  4.25
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.02s | valid loss  5.64 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 10.00 | ms/batch  8.79 | loss  4.24
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.09s | valid loss  5.64 | valid acc   0.0345
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.69 | test acc   0.0116
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.1, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 10.00 | ms/batch 11.07 | loss  6.99
```

```
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.56s | valid loss  6.37 | valid acc   0.0162
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 10.00 | ms/batch 10.00 | loss  6.08
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.34s | valid loss  6.04 | valid acc   0.0205
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 10.00 | ms/batch 11.45 | loss  5.74
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.67s | valid loss  5.84 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 10.00 | ms/batch 11.36 | loss  5.47
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.65s | valid loss  5.72 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 10.00 | ms/batch 11.39 | loss  5.24
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.67s | valid loss  5.68 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 10.00 | ms/batch 11.38 | loss  5.05
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.66s | valid loss  5.64 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 10.00 | ms/batch 11.39 | loss  4.86
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.67s | valid loss  5.63 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 10.00 | ms/batch 11.40 | loss  4.68
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.66s | valid loss  5.64 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 10.00 | ms/batch 11.33 | loss  4.44
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.65s | valid loss  5.62 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 10.00 | ms/batch 11.37 | loss  4.37
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.66s | valid loss  5.63 | valid acc   0.0378
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.69 | test acc   0.0105
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.2, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 10.00 | ms/batch  7.28 | loss  7.04
| epoch   1 |   400/  686 batches | lr 10.00 | ms/batch  6.17 | loss  6.13
| epoch   1 |   600/  686 batches | lr 10.00 | ms/batch  6.19 | loss  5.85
-----------------------------------------------------------------------------------------
```

```
| end of epoch   1 | time:  4.52s | valid loss  5.89 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 10.00 | ms/batch  6.22 | loss  5.60
| epoch   2 |   400/  686 batches | lr 10.00 | ms/batch  6.17 | loss  5.43
| epoch   2 |   600/  686 batches | lr 10.00 | ms/batch  6.17 | loss  5.27
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.32s | valid loss  5.62 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 10.00 | ms/batch  6.18 | loss  5.13
| epoch   3 |   400/  686 batches | lr 10.00 | ms/batch  6.15 | loss  4.99
| epoch   3 |   600/  686 batches | lr 10.00 | ms/batch  6.22 | loss  4.86
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.32s | valid loss  5.62 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 10.00 | ms/batch  6.15 | loss  4.69
| epoch   4 |   400/  686 batches | lr 10.00 | ms/batch  6.19 | loss  4.59
| epoch   4 |   600/  686 batches | lr 10.00 | ms/batch  6.22 | loss  4.52
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.30s | valid loss  5.57 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 10.00 | ms/batch  6.16 | loss  4.54
| epoch   5 |   400/  686 batches | lr 10.00 | ms/batch  6.12 | loss  4.46
| epoch   5 |   600/  686 batches | lr 10.00 | ms/batch  6.33 | loss  4.41
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.33s | valid loss  5.58 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 10.00 | ms/batch  6.18 | loss  4.43
| epoch   6 |   400/  686 batches | lr 10.00 | ms/batch  6.11 | loss  4.36
| epoch   6 |   600/  686 batches | lr 10.00 | ms/batch  6.34 | loss  4.32
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.35s | valid loss  5.58 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 10.00 | ms/batch  6.38 | loss  4.39
| epoch   7 |   400/  686 batches | lr 10.00 | ms/batch  6.18 | loss  4.33
| epoch   7 |   600/  686 batches | lr 10.00 | ms/batch  6.18 | loss  4.30
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.35s | valid loss  5.58 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 10.00 | ms/batch  6.37 | loss  4.38
| epoch   8 |   400/  686 batches | lr 10.00 | ms/batch  6.20 | loss  4.33
| epoch   8 |   600/  686 batches | lr 10.00 | ms/batch  6.17 | loss  4.29
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.34s | valid loss  5.58 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 10.00 | ms/batch  6.25 | loss  4.38
| epoch   9 |   400/  686 batches | lr 10.00 | ms/batch  6.29 | loss  4.32
| epoch   9 |   600/  686 batches | lr 10.00 | ms/batch  6.28 | loss  4.29
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.37s | valid loss  5.58 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 10.00 | ms/batch  6.31 | loss  4.38
```

```
| epoch  10 |   400/  686 batches | lr 10.00 | ms/batch  6.31 | loss  4.32
| epoch  10 |   600/  686 batches | lr 10.00 | ms/batch  6.17 | loss  4.29
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.35s | valid loss  5.58 | valid acc   0.0421
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.64 | test acc   0.0084
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.2, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 10.00 | ms/batch  9.94 | loss  6.97
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.32s | valid loss  6.19 | valid acc   0.0259
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 10.00 | ms/batch  8.96 | loss  5.89
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.12s | valid loss  5.89 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 10.00 | ms/batch  8.99 | loss  5.54
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.13s | valid loss  5.70 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 10.00 | ms/batch  8.94 | loss  5.25
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.12s | valid loss  5.62 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 10.00 | ms/batch  8.93 | loss  5.00
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.12s | valid loss  5.59 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 10.00 | ms/batch  8.98 | loss  4.76
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.12s | valid loss  5.59 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 10.00 | ms/batch  8.97 | loss  4.48
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.13s | valid loss  5.56 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 10.00 | ms/batch  8.96 | loss  4.40
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.13s | valid loss  5.57 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 10.00 | ms/batch  8.79 | loss  4.33
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.08s | valid loss  5.58 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 10.00 | ms/batch  8.96 | loss  4.30
-----------------------------------------------------------------------------------------
```

```
| end of epoch  10 | time:  3.13s | valid loss  5.58 | valid acc   0.0410
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.64 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.2, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 10.00 | ms/batch 12.41 | loss  7.01
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.86s | valid loss  6.57 | valid acc   0.0162
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 10.00 | ms/batch 11.24 | loss  6.09
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.63s | valid loss  6.10 | valid acc   0.0248
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 10.00 | ms/batch 11.35 | loss  5.76
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.65s | valid loss  5.89 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 10.00 | ms/batch 11.28 | loss  5.50
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.63s | valid loss  5.73 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 10.00 | ms/batch 11.35 | loss  5.30
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.65s | valid loss  5.61 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 10.00 | ms/batch 11.36 | loss  5.11
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.65s | valid loss  5.63 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 10.00 | ms/batch 11.26 | loss  4.90
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.63s | valid loss  5.55 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 10.00 | ms/batch 11.23 | loss  4.84
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.63s | valid loss  5.56 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 10.00 | ms/batch 11.25 | loss  4.80
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.63s | valid loss  5.56 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 10.00 | ms/batch 11.43 | loss  4.78
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.66s | valid loss  5.56 | valid acc   0.0367
-----------------------------------------------------------------------------------------
=========================================================================================
```

==
| End of training | test loss 5.63 | test acc   0.0095
==========
==
---
learning_rate: 10, dropout: 0.5, batch_size: 32

| epoch   1 | 200/ 686 batches | lr 10.00 | ms/batch 6.37 | loss 7.04
| epoch   1 | 400/ 686 batches | lr 10.00 | ms/batch 5.75 | loss 6.15
| epoch   1 | 600/ 686 batches | lr 10.00 | ms/batch 6.00 | loss 5.90

| end of epoch   1 | time: 4.16s | valid loss 5.93 | valid acc   0.0227

| epoch   2 | 200/ 686 batches | lr 10.00 | ms/batch 6.07 | loss 5.71
| epoch   2 | 400/ 686 batches | lr 10.00 | ms/batch 6.21 | loss 5.57
| epoch   2 | 600/ 686 batches | lr 10.00 | ms/batch 6.23 | loss 5.45

| end of epoch   2 | time: 4.30s | valid loss 5.66 | valid acc   0.0345

| epoch   3 | 200/ 686 batches | lr 10.00 | ms/batch 6.23 | loss 5.35
| epoch   3 | 400/ 686 batches | lr 10.00 | ms/batch 5.69 | loss 5.26
| epoch   3 | 600/ 686 batches | lr 10.00 | ms/batch 5.72 | loss 5.16

| end of epoch   3 | time: 4.12s | valid loss 5.58 | valid acc   0.0345

| epoch   4 | 200/ 686 batches | lr 10.00 | ms/batch 5.79 | loss 5.09
| epoch   4 | 400/ 686 batches | lr 10.00 | ms/batch 6.06 | loss 5.01
| epoch   4 | 600/ 686 batches | lr 10.00 | ms/batch 6.17 | loss 4.93

| end of epoch   4 | time: 4.20s | valid loss 5.55 | valid acc   0.0345

| epoch   5 | 200/ 686 batches | lr 10.00 | ms/batch 6.23 | loss 4.88
| epoch   5 | 400/ 686 batches | lr 10.00 | ms/batch 6.24 | loss 4.80
| epoch   5 | 600/ 686 batches | lr 10.00 | ms/batch 6.23 | loss 4.75

| end of epoch   5 | time: 4.34s | valid loss 5.57 | valid acc   0.0334

| epoch   6 | 200/ 686 batches | lr 10.00 | ms/batch 6.38 | loss 4.66
| epoch   6 | 400/ 686 batches | lr 10.00 | ms/batch 6.29 | loss 4.59
| epoch   6 | 600/ 686 batches | lr 10.00 | ms/batch 6.31 | loss 4.54

| end of epoch   6 | time: 4.39s | valid loss 5.55 | valid acc   0.0410

| epoch   7 | 200/ 686 batches | lr 10.00 | ms/batch 6.26 | loss 4.57
| epoch   7 | 400/ 686 batches | lr 10.00 | ms/batch 6.30 | loss 4.52
| epoch   7 | 600/ 686 batches | lr 10.00 | ms/batch 6.25 | loss 4.49

| end of epoch   7 | time: 4.39s | valid loss 5.55 | valid acc   0.0399

| epoch   8 | 200/ 686 batches | lr 10.00 | ms/batch 6.39 | loss 4.55
| epoch   8 | 400/ 686 batches | lr 10.00 | ms/batch 6.34 | loss 4.50

```
| epoch   8 |   600/  686 batches | lr 10.00 | ms/batch  6.35 | loss  4.47
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.44s | valid loss  5.55 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 10.00 | ms/batch  6.35 | loss  4.54
| epoch   9 |   400/  686 batches | lr 10.00 | ms/batch  6.28 | loss  4.50
| epoch   9 |   600/  686 batches | lr 10.00 | ms/batch  6.42 | loss  4.47
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.42s | valid loss  5.55 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 10.00 | ms/batch  6.44 | loss  4.54
| epoch  10 |   400/  686 batches | lr 10.00 | ms/batch  6.29 | loss  4.49
| epoch  10 |   600/  686 batches | lr 10.00 | ms/batch  6.28 | loss  4.47
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.39s | valid loss  5.55 | valid acc   0.0410
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.63 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.5, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 10.00 | ms/batch 10.05 | loss  7.03
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.35s | valid loss  6.26 | valid acc   0.0248
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 10.00 | ms/batch  9.07 | loss  5.97
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.16s | valid loss  5.89 | valid acc   0.0259
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 10.00 | ms/batch  8.98 | loss  5.66
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.15s | valid loss  5.74 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 10.00 | ms/batch  8.93 | loss  5.44
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.12s | valid loss  5.67 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 10.00 | ms/batch  8.94 | loss  5.26
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.12s | valid loss  5.62 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 10.00 | ms/batch  8.57 | loss  5.10
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.06s | valid loss  5.58 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 10.00 | ms/batch  8.94 | loss  4.96
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.12s | valid loss  5.55 | valid acc   0.0378
```

```
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 10.00 | ms/batch  8.96 | loss  4.83
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.14s | valid loss  5.57 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 10.00 | ms/batch  8.30 | loss  4.66
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.90s | valid loss  5.53 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 10.00 | ms/batch  8.28 | loss  4.61
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.96s | valid loss  5.54 | valid acc   0.0378
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.61 | test acc   0.0074
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.5, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 10.00 | ms/batch 12.38 | loss  7.03
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.86s | valid loss  6.37 | valid acc   0.0194
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 10.00 | ms/batch 11.41 | loss  6.13
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.67s | valid loss  6.12 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 10.00 | ms/batch 11.48 | loss  5.85
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.67s | valid loss  5.85 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 10.00 | ms/batch 11.14 | loss  5.65
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.63s | valid loss  5.81 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 10.00 | ms/batch 11.34 | loss  5.49
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.64s | valid loss  5.66 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 10.00 | ms/batch 11.34 | loss  5.36
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.65s | valid loss  5.59 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 10.00 | ms/batch 11.29 | loss  5.24
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.64s | valid loss  5.56 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 10.00 | ms/batch 11.28 | loss  5.12
-----------------------------------------------------------------------------------------
```

```
| end of epoch   8 | time:  2.64s | valid loss  5.56 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 10.00 | ms/batch 11.30 | loss  4.98
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.64s | valid loss  5.53 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 10.00 | ms/batch 11.39 | loss  4.94
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.66s | valid loss  5.52 | valid acc   0.0356
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.60 | test acc   0.0116
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.1, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 20.00 | ms/batch  6.38 | loss  6.68
| epoch   1 |   400/  686 batches | lr 20.00 | ms/batch  5.20 | loss  5.92
| epoch   1 |   600/  686 batches | lr 20.00 | ms/batch  6.02 | loss  5.59
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.12s | valid loss  5.79 | valid acc   0.0227
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 20.00 | ms/batch  6.14 | loss  5.28
| epoch   2 |   400/  686 batches | lr 20.00 | ms/batch  6.08 | loss  5.07
| epoch   2 |   600/  686 batches | lr 20.00 | ms/batch  5.99 | loss  4.88
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.25s | valid loss  5.66 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 20.00 | ms/batch  6.27 | loss  4.66
| epoch   3 |   400/  686 batches | lr 20.00 | ms/batch  6.04 | loss  4.47
| epoch   3 |   600/  686 batches | lr 20.00 | ms/batch  6.07 | loss  4.30
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.27s | valid loss  5.77 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 20.00 | ms/batch  6.06 | loss  4.01
| epoch   4 |   400/  686 batches | lr 20.00 | ms/batch  6.06 | loss  3.84
| epoch   4 |   600/  686 batches | lr 20.00 | ms/batch  6.16 | loss  3.74
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.24s | valid loss  5.77 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 20.00 | ms/batch  6.04 | loss  3.75
| epoch   5 |   400/  686 batches | lr 20.00 | ms/batch  5.98 | loss  3.64
| epoch   5 |   600/  686 batches | lr 20.00 | ms/batch  5.98 | loss  3.57
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.19s | valid loss  5.79 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 20.00 | ms/batch  6.06 | loss  3.67
| epoch   6 |   400/  686 batches | lr 20.00 | ms/batch  6.08 | loss  3.58
| epoch   6 |   600/  686 batches | lr 20.00 | ms/batch  6.11 | loss  3.52
```

```
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.25s | valid loss  5.80 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 20.00 | ms/batch  6.11 | loss  3.66
| epoch   7 |   400/  686 batches | lr 20.00 | ms/batch  6.07 | loss  3.57
| epoch   7 |   600/  686 batches | lr 20.00 | ms/batch  6.04 | loss  3.51
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.24s | valid loss  5.80 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 20.00 | ms/batch  6.01 | loss  3.65
| epoch   8 |   400/  686 batches | lr 20.00 | ms/batch  6.02 | loss  3.56
| epoch   8 |   600/  686 batches | lr 20.00 | ms/batch  6.05 | loss  3.51
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.21s | valid loss  5.79 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 20.00 | ms/batch  6.24 | loss  3.65
| epoch   9 |   400/  686 batches | lr 20.00 | ms/batch  6.17 | loss  3.56
| epoch   9 |   600/  686 batches | lr 20.00 | ms/batch  6.16 | loss  3.50
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.30s | valid loss  5.79 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 20.00 | ms/batch  6.13 | loss  3.65
| epoch  10 |   400/  686 batches | lr 20.00 | ms/batch  6.05 | loss  3.56
| epoch  10 |   600/  686 batches | lr 20.00 | ms/batch  6.09 | loss  3.50
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.26s | valid loss  5.79 | valid acc   0.0399
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.73 | test acc   0.0063
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.1, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 20.00 | ms/batch 10.08 | loss  6.71
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.35s | valid loss  5.97 | valid acc   0.0280
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 20.00 | ms/batch  8.88 | loss  5.66
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.07s | valid loss  5.70 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 20.00 | ms/batch  8.84 | loss  5.14
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.11s | valid loss  5.62 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 20.00 | ms/batch  8.98 | loss  4.72
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.14s | valid loss  5.67 | valid acc   0.0345
-----------------------------------------------------------------------------------------
```

```
| epoch   5 |   200/  343 batches | lr 20.00 | ms/batch  8.79 | loss  4.26
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.09s | valid loss  5.62 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 20.00 | ms/batch  8.99 | loss  4.11
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.14s | valid loss  5.63 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 20.00 | ms/batch  9.03 | loss  4.06
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.16s | valid loss  5.64 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 20.00 | ms/batch  9.06 | loss  4.05
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.16s | valid loss  5.64 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 20.00 | ms/batch  8.96 | loss  4.05
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.13s | valid loss  5.64 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 20.00 | ms/batch  8.95 | loss  4.05
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.13s | valid loss  5.64 | valid acc   0.0399
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.68 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.1, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 20.00 | ms/batch 12.40 | loss  6.69
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.86s | valid loss  6.46 | valid acc   0.0183
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 20.00 | ms/batch 11.32 | loss  5.84
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.65s | valid loss  5.87 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 20.00 | ms/batch 11.31 | loss  5.44
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.64s | valid loss  5.66 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 20.00 | ms/batch 11.33 | loss  5.10
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.64s | valid loss  5.64 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 20.00 | ms/batch 11.32 | loss  4.80
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.64s | valid loss  5.67 | valid acc   0.0356
```

```
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 20.00 | ms/batch 11.35 | loss  4.44
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.65s | valid loss  5.60 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 20.00 | ms/batch 11.26 | loss  4.34
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.63s | valid loss  5.62 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 20.00 | ms/batch 11.41 | loss  4.25
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.67s | valid loss  5.62 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 20.00 | ms/batch 11.40 | loss  4.22
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.66s | valid loss  5.62 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 20.00 | ms/batch 11.15 | loss  4.21
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.61s | valid loss  5.62 | valid acc   0.0356
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.67 | test acc   0.0084
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.2, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 20.00 | ms/batch  7.44 | loss  6.71
| epoch   1 |   400/  686 batches | lr 20.00 | ms/batch  6.24 | loss  5.92
| epoch   1 |   600/  686 batches | lr 20.00 | ms/batch  6.29 | loss  5.61
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.60s | valid loss  5.76 | valid acc   0.0259
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 20.00 | ms/batch  6.30 | loss  5.32
| epoch   2 |   400/  686 batches | lr 20.00 | ms/batch  6.28 | loss  5.13
| epoch   2 |   600/  686 batches | lr 20.00 | ms/batch  6.16 | loss  4.96
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.36s | valid loss  5.62 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 20.00 | ms/batch  6.32 | loss  4.77
| epoch   3 |   400/  686 batches | lr 20.00 | ms/batch  6.23 | loss  4.60
| epoch   3 |   600/  686 batches | lr 20.00 | ms/batch  6.18 | loss  4.47
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.33s | valid loss  5.70 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 20.00 | ms/batch  6.27 | loss  4.22
| epoch   4 |   400/  686 batches | lr 20.00 | ms/batch  6.18 | loss  4.07
| epoch   4 |   600/  686 batches | lr 20.00 | ms/batch  6.19 | loss  3.98
-----------------------------------------------------------------------------------------
```

```
| end of epoch   4 | time:  4.33s | valid loss  5.69 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 20.00 | ms/batch  6.29 | loss  4.00
| epoch   5 |   400/  686 batches | lr 20.00 | ms/batch  6.13 | loss  3.90
| epoch   5 |   600/  686 batches | lr 20.00 | ms/batch  6.18 | loss  3.83
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.32s | valid loss  5.69 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 20.00 | ms/batch  6.16 | loss  3.93
| epoch   6 |   400/  686 batches | lr 20.00 | ms/batch  6.22 | loss  3.85
| epoch   6 |   600/  686 batches | lr 20.00 | ms/batch  6.22 | loss  3.80
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.33s | valid loss  5.69 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 20.00 | ms/batch  6.21 | loss  3.92
| epoch   7 |   400/  686 batches | lr 20.00 | ms/batch  6.11 | loss  3.84
| epoch   7 |   600/  686 batches | lr 20.00 | ms/batch  6.23 | loss  3.79
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.32s | valid loss  5.69 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 20.00 | ms/batch  6.23 | loss  3.91
| epoch   8 |   400/  686 batches | lr 20.00 | ms/batch  6.13 | loss  3.84
| epoch   8 |   600/  686 batches | lr 20.00 | ms/batch  5.44 | loss  3.78
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.09s | valid loss  5.69 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 20.00 | ms/batch  6.27 | loss  3.91
| epoch   9 |   400/  686 batches | lr 20.00 | ms/batch  6.15 | loss  3.83
| epoch   9 |   600/  686 batches | lr 20.00 | ms/batch  6.07 | loss  3.78
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.29s | valid loss  5.69 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 20.00 | ms/batch  6.05 | loss  3.91
| epoch  10 |   400/  686 batches | lr 20.00 | ms/batch  6.28 | loss  3.83
| epoch  10 |   600/  686 batches | lr 20.00 | ms/batch  6.20 | loss  3.78
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.30s | valid loss  5.69 | valid acc   0.0410
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.70 | test acc   0.0084
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.2, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  343 batches | lr 20.00 | ms/batch 10.05 | loss  6.68
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  3.36s | valid loss  5.95 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  343 batches | lr 20.00 | ms/batch  8.97 | loss  5.65
```

```
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  3.13s | valid loss  5.66 | valid acc   0.0334
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  343 batches | lr 20.00 | ms/batch  8.97 | loss  5.18
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  3.13s | valid loss  5.62 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  343 batches | lr 20.00 | ms/batch  9.06 | loss  4.82
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  3.15s | valid loss  5.68 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  343 batches | lr 20.00 | ms/batch  8.97 | loss  4.42
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  3.11s | valid loss  5.60 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  343 batches | lr 20.00 | ms/batch  9.07 | loss  4.29
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  3.15s | valid loss  5.62 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  343 batches | lr 20.00 | ms/batch  8.98 | loss  4.18
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  3.14s | valid loss  5.63 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  343 batches | lr 20.00 | ms/batch  9.06 | loss  4.15
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  3.16s | valid loss  5.63 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  343 batches | lr 20.00 | ms/batch  8.77 | loss  4.14
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  3.10s | valid loss  5.63 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  343 batches | lr 20.00 | ms/batch  8.87 | loss  4.14
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  3.14s | valid loss  5.63 | valid acc   0.0410
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.67 | test acc   0.0095
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.2, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 20.00 | ms/batch 12.50 | loss  6.67
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.88s | valid loss  6.24 | valid acc   0.0302
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 20.00 | ms/batch 11.41 | loss  5.84
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.66s | valid loss  5.82 | valid acc   0.0324
-----------------------------------------------------------------------------------------
```

```
| epoch   3 |   200/  228 batches | lr 20.00 | ms/batch 11.32 | loss  5.45
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.64s | valid loss  5.60 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 20.00 | ms/batch 11.26 | loss  5.13
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.63s | valid loss  5.58 | valid acc   0.0367
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 20.00 | ms/batch 11.23 | loss  4.87
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.62s | valid loss  5.66 | valid acc   0.0356
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 20.00 | ms/batch 11.51 | loss  4.55
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.69s | valid loss  5.57 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 20.00 | ms/batch 11.45 | loss  4.46
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.67s | valid loss  5.58 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 20.00 | ms/batch 11.42 | loss  4.38
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.67s | valid loss  5.58 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 20.00 | ms/batch 11.37 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.66s | valid loss  5.57 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 20.00 | ms/batch 11.37 | loss  4.36
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.66s | valid loss  5.57 | valid acc   0.0388
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.63 | test acc   0.0105
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.5, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  686 batches | lr 20.00 | ms/batch  7.24 | loss  6.74
| epoch   1 |   400/  686 batches | lr 20.00 | ms/batch  6.16 | loss  5.96
| epoch   1 |   600/  686 batches | lr 20.00 | ms/batch  6.18 | loss  5.70
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  4.52s | valid loss  5.77 | valid acc   0.0280
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  686 batches | lr 20.00 | ms/batch  6.16 | loss  5.48
| epoch   2 |   400/  686 batches | lr 20.00 | ms/batch  6.05 | loss  5.34
| epoch   2 |   600/  686 batches | lr 20.00 | ms/batch  6.03 | loss  5.21
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  4.23s | valid loss  5.57 | valid acc   0.0367
```

```
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  686 batches | lr 20.00 | ms/batch  6.14 | loss  5.11
| epoch   3 |   400/  686 batches | lr 20.00 | ms/batch  6.15 | loss  5.01
| epoch   3 |   600/  686 batches | lr 20.00 | ms/batch  6.37 | loss  4.91
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  4.32s | valid loss  5.55 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  686 batches | lr 20.00 | ms/batch  6.17 | loss  4.84
| epoch   4 |   400/  686 batches | lr 20.00 | ms/batch  6.03 | loss  4.76
| epoch   4 |   600/  686 batches | lr 20.00 | ms/batch  6.09 | loss  4.69
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  4.25s | valid loss  5.58 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  686 batches | lr 20.00 | ms/batch  6.04 | loss  4.56
| epoch   5 |   400/  686 batches | lr 20.00 | ms/batch  5.73 | loss  4.46
| epoch   5 |   600/  686 batches | lr 20.00 | ms/batch  6.09 | loss  4.40
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  4.17s | valid loss  5.54 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  686 batches | lr 20.00 | ms/batch  6.30 | loss  4.43
| epoch   6 |   400/  686 batches | lr 20.00 | ms/batch  6.56 | loss  4.36
| epoch   6 |   600/  686 batches | lr 20.00 | ms/batch  6.22 | loss  4.32
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  4.42s | valid loss  5.56 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  686 batches | lr 20.00 | ms/batch  6.11 | loss  4.34
| epoch   7 |   400/  686 batches | lr 20.00 | ms/batch  6.07 | loss  4.28
| epoch   7 |   600/  686 batches | lr 20.00 | ms/batch  6.03 | loss  4.25
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  4.24s | valid loss  5.56 | valid acc   0.0421
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  686 batches | lr 20.00 | ms/batch  6.24 | loss  4.31
| epoch   8 |   400/  686 batches | lr 20.00 | ms/batch  6.28 | loss  4.25
| epoch   8 |   600/  686 batches | lr 20.00 | ms/batch  6.22 | loss  4.23
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  4.35s | valid loss  5.56 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  686 batches | lr 20.00 | ms/batch  6.32 | loss  4.30
| epoch   9 |   400/  686 batches | lr 20.00 | ms/batch  6.47 | loss  4.24
| epoch   9 |   600/  686 batches | lr 20.00 | ms/batch  6.26 | loss  4.22
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  4.41s | valid loss  5.56 | valid acc   0.0431
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  686 batches | lr 20.00 | ms/batch  6.16 | loss  4.30
| epoch  10 |   400/  686 batches | lr 20.00 | ms/batch  6.11 | loss  4.25
| epoch  10 |   600/  686 batches | lr 20.00 | ms/batch  6.06 | loss  4.22
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  4.25s | valid loss  5.56 | valid acc   0.0431
-----------------------------------------------------------------------------------------
=========================================================================================
==
```

| End of training | test loss 5.64 | test acc   0.0116
==========
==
---
learning_rate: 20, dropout: 0.5, batch_size: 64

| epoch   1 | 200/ 343 batches | lr 20.00 | ms/batch 10.12 | loss 6.71

| end of epoch   1 | time: 3.36s | valid loss 5.98 | valid acc   0.0280

| epoch   2 | 200/ 343 batches | lr 20.00 | ms/batch 8.96 | loss 5.75

| end of epoch   2 | time: 3.13s | valid loss 5.71 | valid acc   0.0334

| epoch   3 | 200/ 343 batches | lr 20.00 | ms/batch 8.99 | loss 5.39

| end of epoch   3 | time: 3.13s | valid loss 5.61 | valid acc   0.0313

| epoch   4 | 200/ 343 batches | lr 20.00 | ms/batch 8.98 | loss 5.15

| end of epoch   4 | time: 3.13s | valid loss 5.60 | valid acc   0.0367

| epoch   5 | 200/ 343 batches | lr 20.00 | ms/batch 8.96 | loss 4.94

| end of epoch   5 | time: 3.12s | valid loss 5.54 | valid acc   0.0367

| epoch   6 | 200/ 343 batches | lr 20.00 | ms/batch 8.98 | loss 4.77

| end of epoch   6 | time: 3.14s | valid loss 5.56 | valid acc   0.0356

| epoch   7 | 200/ 343 batches | lr 20.00 | ms/batch 8.94 | loss 4.55

| end of epoch   7 | time: 3.13s | valid loss 5.54 | valid acc   0.0410

| epoch   8 | 200/ 343 batches | lr 20.00 | ms/batch 8.94 | loss 4.47

| end of epoch   8 | time: 3.12s | valid loss 5.55 | valid acc   0.0367

| epoch   9 | 200/ 343 batches | lr 20.00 | ms/batch 8.89 | loss 4.41

| end of epoch   9 | time: 3.13s | valid loss 5.55 | valid acc   0.0378

| epoch 10 | 200/ 343 batches | lr 20.00 | ms/batch 9.02 | loss 4.39

| end of epoch 10 | time: 3.15s | valid loss 5.55 | valid acc   0.0378

==========
==
| End of training | test loss 5.61 | test acc   0.0095
==========
==

```
-----------------------------------------------------------------------------------------
learning_rate: 20, dropout: 0.5, batch_size: 96
-----------------------------------------------------------------------------------------
| epoch   1 |   200/  228 batches | lr 20.00 | ms/batch 12.38 | loss  6.70
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  2.86s | valid loss  6.15 | valid acc   0.0259
-----------------------------------------------------------------------------------------
| epoch   2 |   200/  228 batches | lr 20.00 | ms/batch 11.29 | loss  5.88
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  2.64s | valid loss  5.85 | valid acc   0.0270
-----------------------------------------------------------------------------------------
| epoch   3 |   200/  228 batches | lr 20.00 | ms/batch 11.26 | loss  5.59
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  2.63s | valid loss  5.66 | valid acc   0.0378
-----------------------------------------------------------------------------------------
| epoch   4 |   200/  228 batches | lr 20.00 | ms/batch 11.24 | loss  5.36
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  2.63s | valid loss  5.59 | valid acc   0.0345
-----------------------------------------------------------------------------------------
| epoch   5 |   200/  228 batches | lr 20.00 | ms/batch 11.27 | loss  5.17
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  2.63s | valid loss  5.58 | valid acc   0.0313
-----------------------------------------------------------------------------------------
| epoch   6 |   200/  228 batches | lr 20.00 | ms/batch 11.23 | loss  5.02
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  2.62s | valid loss  5.53 | valid acc   0.0410
-----------------------------------------------------------------------------------------
| epoch   7 |   200/  228 batches | lr 20.00 | ms/batch 11.28 | loss  4.87
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  2.64s | valid loss  5.57 | valid acc   0.0324
-----------------------------------------------------------------------------------------
| epoch   8 |   200/  228 batches | lr 20.00 | ms/batch 11.25 | loss  4.68
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  2.63s | valid loss  5.51 | valid acc   0.0399
-----------------------------------------------------------------------------------------
| epoch   9 |   200/  228 batches | lr 20.00 | ms/batch 11.40 | loss  4.62
-----------------------------------------------------------------------------------------
| end of epoch   9 | time:  2.66s | valid loss  5.52 | valid acc   0.0388
-----------------------------------------------------------------------------------------
| epoch  10 |   200/  228 batches | lr 20.00 | ms/batch 11.26 | loss  4.58
-----------------------------------------------------------------------------------------
| end of epoch  10 | time:  2.63s | valid loss  5.52 | valid acc   0.0388
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  5.60 | test acc   0.0105
=========================================================================================
==
```