```
-------------------------------------------------------------------------------
learning_rate: 0.1, dropout: 0.1, batch_size: 32
-------------------------------------------------------------------------------
| epoch   1 |    200/ 2104 batches | lr 0.100 | ms/batch 8.21 | loss  0.69 |
| epoch   1 |    400/ 2104 batches | lr 0.100 | ms/batch 7.05 | loss  0.69 |
| epoch   1 |    600/ 2104 batches | lr 0.100 | ms/batch 7.08 | loss  0.69 |
| epoch   1 |    800/ 2104 batches | lr 0.100 | ms/batch 7.07 | loss  0.69 |
| epoch   1 |   1000/ 2104 batches | lr 0.100 | ms/batch 7.11 | loss  0.68 |
| epoch   1 |   1200/ 2104 batches | lr 0.100 | ms/batch 7.05 | loss  0.68 |
| epoch   1 |   1400/ 2104 batches | lr 0.100 | ms/batch 6.98 | loss  0.69 |
| epoch   1 |   1600/ 2104 batches | lr 0.100 | ms/batch 7.02 | loss  0.69 |
| epoch   1 |   1800/ 2104 batches | lr 0.100 | ms/batch 7.03 | loss  0.69 |
| epoch   1 |   2000/ 2104 batches | lr 0.100 | ms/batch 7.32 | loss  0.69 |
-------------------------------------------------------------------------------
| end of epoch   1 | time: 15.20s | valid loss  0.71 | valid acc    0.540 |
-------------------------------------------------------------------------------
| epoch   2 |    200/ 2104 batches | lr 0.100 | ms/batch 7.77 | loss  0.69 |
| epoch   2 |    400/ 2104 batches | lr 0.100 | ms/batch 7.30 | loss  0.69 |
| epoch   2 |    600/ 2104 batches | lr 0.100 | ms/batch 7.35 | loss  0.69 |
| epoch   2 |    800/ 2104 batches | lr 0.100 | ms/batch 7.32 | loss  0.69 |
| epoch   2 |   1000/ 2104 batches | lr 0.100 | ms/batch 7.33 | loss  0.68 |
| epoch   2 |   1200/ 2104 batches | lr 0.100 | ms/batch 7.34 | loss  0.68 |
| epoch   2 |   1400/ 2104 batches | lr 0.100 | ms/batch 7.41 | loss  0.68 |
| epoch   2 |   1600/ 2104 batches | lr 0.100 | ms/batch 7.35 | loss  0.68 |
| epoch   2 |   1800/ 2104 batches | lr 0.100 | ms/batch 7.30 | loss  0.68 |
| epoch   2 |   2000/ 2104 batches | lr 0.100 | ms/batch 7.33 | loss  0.68 |
-------------------------------------------------------------------------------
| end of epoch   2 | time: 15.59s | valid loss  0.71 | valid acc    0.540 |
-------------------------------------------------------------------------------
| epoch   3 |    200/ 2104 batches | lr 0.100 | ms/batch 7.33 | loss  0.69 |
| epoch   3 |    400/ 2104 batches | lr 0.100 | ms/batch 7.32 | loss  0.68 |
| epoch   3 |    600/ 2104 batches | lr 0.100 | ms/batch 7.33 | loss  0.68 |
| epoch   3 |    800/ 2104 batches | lr 0.100 | ms/batch 7.29 | loss  0.68 |
| epoch   3 |   1000/ 2104 batches | lr 0.100 | ms/batch 7.31 | loss  0.68 |
| epoch   3 |   1200/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.68 |
| epoch   3 |   1400/ 2104 batches | lr 0.100 | ms/batch 7.29 | loss  0.68 |
| epoch   3 |   1600/ 2104 batches | lr 0.100 | ms/batch 7.29 | loss  0.68 |
| epoch   3 |   1800/ 2104 batches | lr 0.100 | ms/batch 7.26 | loss  0.67 |
| epoch   3 |   2000/ 2104 batches | lr 0.100 | ms/batch 7.29 | loss  0.66 |
-------------------------------------------------------------------------------
| end of epoch   3 | time: 15.39s | valid loss  0.64 | valid acc    0.664 |
-------------------------------------------------------------------------------
| epoch   4 |    200/ 2104 batches | lr 0.100 | ms/batch 6.77 | loss  0.65 |
| epoch   4 |    400/ 2104 batches | lr 0.100 | ms/batch 7.25 | loss  0.64 |
| epoch   4 |    600/ 2104 batches | lr 0.100 | ms/batch 7.09 | loss  0.63 |
| epoch   4 |    800/ 2104 batches | lr 0.100 | ms/batch 7.12 | loss  0.63 |
| epoch   4 |   1000/ 2104 batches | lr 0.100 | ms/batch 7.11 | loss  0.61 |
| epoch   4 |   1200/ 2104 batches | lr 0.100 | ms/batch 7.14 | loss  0.61 |
| epoch   4 |   1400/ 2104 batches | lr 0.100 | ms/batch 7.15 | loss  0.60 |
| epoch   4 |   1600/ 2104 batches | lr 0.100 | ms/batch 7.14 | loss  0.59 |
| epoch   4 |   1800/ 2104 batches | lr 0.100 | ms/batch 7.13 | loss  0.59 |
```

```
| epoch   4 |   2000/ 2104 batches | lr 0.100 | ms/batch  7.12 | loss  0.57 |
-----------------------------------------------------------------------------
| end of epoch   4 | time: 15.01s | valid loss  0.62 | valid acc    0.740 |
-----------------------------------------------------------------------------
| epoch   5 |    200/ 2104 batches | lr 0.100 | ms/batch  7.22 | loss  0.58 |
| epoch   5 |    400/ 2104 batches | lr 0.100 | ms/batch  7.11 | loss  0.57 |
| epoch   5 |    600/ 2104 batches | lr 0.100 | ms/batch  7.30 | loss  0.57 |
| epoch   5 |    800/ 2104 batches | lr 0.100 | ms/batch  7.31 | loss  0.55 |
| epoch   5 |   1000/ 2104 batches | lr 0.100 | ms/batch  7.35 | loss  0.55 |
| epoch   5 |   1200/ 2104 batches | lr 0.100 | ms/batch  7.34 | loss  0.55 |
| epoch   5 |   1400/ 2104 batches | lr 0.100 | ms/batch  7.31 | loss  0.53 |
| epoch   5 |   1600/ 2104 batches | lr 0.100 | ms/batch  7.34 | loss  0.53 |
| epoch   5 |   1800/ 2104 batches | lr 0.100 | ms/batch  7.24 | loss  0.53 |
| epoch   5 |   2000/ 2104 batches | lr 0.100 | ms/batch  7.07 | loss  0.51 |
-----------------------------------------------------------------------------
| end of epoch   5 | time: 15.31s | valid loss  0.55 | valid acc    0.774 |
-----------------------------------------------------------------------------
| epoch   6 |    200/ 2104 batches | lr 0.100 | ms/batch  7.07 | loss  0.52 |
| epoch   6 |    400/ 2104 batches | lr 0.100 | ms/batch  7.03 | loss  0.51 |
| epoch   6 |    600/ 2104 batches | lr 0.100 | ms/batch  7.00 | loss  0.52 |
| epoch   6 |    800/ 2104 batches | lr 0.100 | ms/batch  7.22 | loss  0.50 |
| epoch   6 |   1000/ 2104 batches | lr 0.100 | ms/batch  7.22 | loss  0.50 |
| epoch   6 |   1200/ 2104 batches | lr 0.100 | ms/batch  7.21 | loss  0.50 |
| epoch   6 |   1400/ 2104 batches | lr 0.100 | ms/batch  7.29 | loss  0.48 |
| epoch   6 |   1600/ 2104 batches | lr 0.100 | ms/batch  7.32 | loss  0.48 |
| epoch   6 |   1800/ 2104 batches | lr 0.100 | ms/batch  7.25 | loss  0.49 |
| epoch   6 |   2000/ 2104 batches | lr 0.100 | ms/batch  7.28 | loss  0.46 |
-----------------------------------------------------------------------------
| end of epoch   6 | time: 15.19s | valid loss  0.55 | valid acc    0.760 |
-----------------------------------------------------------------------------
| epoch   7 |    200/ 2104 batches | lr 0.025 | ms/batch  7.26 | loss  0.46 |
| epoch   7 |    400/ 2104 batches | lr 0.025 | ms/batch  7.21 | loss  0.45 |
| epoch   7 |    600/ 2104 batches | lr 0.025 | ms/batch  7.18 | loss  0.46 |
| epoch   7 |    800/ 2104 batches | lr 0.025 | ms/batch  7.18 | loss  0.45 |
| epoch   7 |   1000/ 2104 batches | lr 0.025 | ms/batch  7.17 | loss  0.45 |
| epoch   7 |   1200/ 2104 batches | lr 0.025 | ms/batch  7.18 | loss  0.46 |
| epoch   7 |   1400/ 2104 batches | lr 0.025 | ms/batch  7.33 | loss  0.44 |
| epoch   7 |   1600/ 2104 batches | lr 0.025 | ms/batch  7.04 | loss  0.44 |
| epoch   7 |   1800/ 2104 batches | lr 0.025 | ms/batch  7.28 | loss  0.44 |
| epoch   7 |   2000/ 2104 batches | lr 0.025 | ms/batch  7.48 | loss  0.43 |
-----------------------------------------------------------------------------
| end of epoch   7 | time: 15.28s | valid loss  0.55 | valid acc    0.793 |
-----------------------------------------------------------------------------
| epoch   8 |    200/ 2104 batches | lr 0.006 | ms/batch  7.30 | loss  0.44 |
| epoch   8 |    400/ 2104 batches | lr 0.006 | ms/batch  7.12 | loss  0.44 |
| epoch   8 |    600/ 2104 batches | lr 0.006 | ms/batch  7.10 | loss  0.45 |
| epoch   8 |    800/ 2104 batches | lr 0.006 | ms/batch  7.31 | loss  0.44 |
| epoch   8 |   1000/ 2104 batches | lr 0.006 | ms/batch  7.42 | loss  0.44 |
| epoch   8 |   1200/ 2104 batches | lr 0.006 | ms/batch  7.41 | loss  0.44 |
| epoch   8 |   1400/ 2104 batches | lr 0.006 | ms/batch  7.25 | loss  0.44 |
| epoch   8 |   1600/ 2104 batches | lr 0.006 | ms/batch  7.38 | loss  0.43 |
```

```
| epoch   8 |  1800/ 2104 batches | lr 0.006 | ms/batch  7.43 | loss  0.44 |
| epoch   8 |  2000/ 2104 batches | lr 0.006 | ms/batch  7.40 | loss  0.42 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time: 15.46s | valid loss  0.55 | valid acc    0.793 |
-----------------------------------------------------------------------------------------
| epoch   9 |   200/ 2104 batches | lr 0.002 | ms/batch  7.42 | loss  0.44 |
| epoch   9 |   400/ 2104 batches | lr 0.002 | ms/batch  7.32 | loss  0.44 |
| epoch   9 |   600/ 2104 batches | lr 0.002 | ms/batch  7.07 | loss  0.44 |
| epoch   9 |   800/ 2104 batches | lr 0.002 | ms/batch  7.17 | loss  0.43 |
| epoch   9 |  1000/ 2104 batches | lr 0.002 | ms/batch  7.13 | loss  0.44 |
| epoch   9 |  1200/ 2104 batches | lr 0.002 | ms/batch  7.06 | loss  0.44 |
| epoch   9 |  1400/ 2104 batches | lr 0.002 | ms/batch  7.04 | loss  0.43 |
| epoch   9 |  1600/ 2104 batches | lr 0.002 | ms/batch  7.08 | loss  0.43 |
| epoch   9 |  1800/ 2104 batches | lr 0.002 | ms/batch  7.14 | loss  0.44 |
| epoch   9 |  2000/ 2104 batches | lr 0.002 | ms/batch  7.12 | loss  0.41 |
-----------------------------------------------------------------------------------------
| end of epoch   9 | time: 15.13s | valid loss  0.55 | valid acc    0.793 |
-----------------------------------------------------------------------------------------
| epoch  10 |   200/ 2104 batches | lr 0.000 | ms/batch  7.62 | loss  0.44 |
| epoch  10 |   400/ 2104 batches | lr 0.000 | ms/batch  7.38 | loss  0.43 |
| epoch  10 |   600/ 2104 batches | lr 0.000 | ms/batch  7.33 | loss  0.44 |
| epoch  10 |   800/ 2104 batches | lr 0.000 | ms/batch  7.21 | loss  0.43 |
| epoch  10 |  1000/ 2104 batches | lr 0.000 | ms/batch  7.28 | loss  0.44 |
| epoch  10 |  1200/ 2104 batches | lr 0.000 | ms/batch  7.35 | loss  0.44 |
| epoch  10 |  1400/ 2104 batches | lr 0.000 | ms/batch  7.28 | loss  0.43 |
| epoch  10 |  1600/ 2104 batches | lr 0.000 | ms/batch  7.44 | loss  0.43 |
| epoch  10 |  1800/ 2104 batches | lr 0.000 | ms/batch  7.40 | loss  0.44 |
| epoch  10 |  2000/ 2104 batches | lr 0.000 | ms/batch  7.36 | loss  0.42 |
-----------------------------------------------------------------------------------------
| end of epoch  10 | time: 15.57s | valid loss  0.55 | valid acc    0.781 |
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  0.56 | test acc    0.767 |
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 0.1, dropout: 0.1, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |   200/ 1052 batches | lr 0.100 | ms/batch  8.82 | loss  0.69 |
| epoch   1 |   400/ 1052 batches | lr 0.100 | ms/batch  7.50 | loss  0.69 |
| epoch   1 |   600/ 1052 batches | lr 0.100 | ms/batch  7.52 | loss  0.68 |
| epoch   1 |   800/ 1052 batches | lr 0.100 | ms/batch  7.48 | loss  0.69 |
| epoch   1 |  1000/ 1052 batches | lr 0.100 | ms/batch  7.55 | loss  0.69 |
-----------------------------------------------------------------------------------------
| end of epoch   1 | time: 8.22s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   2 |   200/ 1052 batches | lr 0.100 | ms/batch  7.46 | loss  0.69 |
| epoch   2 |   400/ 1052 batches | lr 0.100 | ms/batch  7.34 | loss  0.69 |
| epoch   2 |   600/ 1052 batches | lr 0.100 | ms/batch  7.34 | loss  0.68 |
| epoch   2 |   800/ 1052 batches | lr 0.100 | ms/batch  7.38 | loss  0.68 |
```

```
| epoch   2 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.45 | loss  0.69 |
-----------------------------------------------------------------------------------------
| end of epoch   2 | time: 7.85s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   3 |    200/ 1052 batches | lr 0.100 | ms/batch  7.69 | loss  0.69 |
| epoch   3 |    400/ 1052 batches | lr 0.100 | ms/batch  7.43 | loss  0.69 |
| epoch   3 |    600/ 1052 batches | lr 0.100 | ms/batch  7.52 | loss  0.68 |
| epoch   3 |    800/ 1052 batches | lr 0.100 | ms/batch  7.44 | loss  0.68 |
| epoch   3 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.41 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time: 7.94s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   4 |    200/ 1052 batches | lr 0.100 | ms/batch  7.35 | loss  0.69 |
| epoch   4 |    400/ 1052 batches | lr 0.100 | ms/batch  7.32 | loss  0.68 |
| epoch   4 |    600/ 1052 batches | lr 0.100 | ms/batch  7.18 | loss  0.68 |
| epoch   4 |    800/ 1052 batches | lr 0.100 | ms/batch  7.19 | loss  0.68 |
| epoch   4 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.28 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 7.71s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   5 |    200/ 1052 batches | lr 0.100 | ms/batch  7.38 | loss  0.69 |
| epoch   5 |    400/ 1052 batches | lr 0.100 | ms/batch  7.36 | loss  0.68 |
| epoch   5 |    600/ 1052 batches | lr 0.100 | ms/batch  7.37 | loss  0.68 |
| epoch   5 |    800/ 1052 batches | lr 0.100 | ms/batch  7.30 | loss  0.68 |
| epoch   5 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.18 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 7.75s | valid loss  0.69 | valid acc    0.612 |
-----------------------------------------------------------------------------------------
| epoch   6 |    200/ 1052 batches | lr 0.100 | ms/batch  7.47 | loss  0.67 |
| epoch   6 |    400/ 1052 batches | lr 0.100 | ms/batch  7.36 | loss  0.66 |
| epoch   6 |    600/ 1052 batches | lr 0.100 | ms/batch  7.44 | loss  0.65 |
| epoch   6 |    800/ 1052 batches | lr 0.100 | ms/batch  7.37 | loss  0.64 |
| epoch   6 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.18 | loss  0.63 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time: 7.80s | valid loss  0.60 | valid acc    0.702 |
-----------------------------------------------------------------------------------------
| epoch   7 |    200/ 1052 batches | lr 0.100 | ms/batch  7.21 | loss  0.62 |
| epoch   7 |    400/ 1052 batches | lr 0.100 | ms/batch  7.16 | loss  0.61 |
| epoch   7 |    600/ 1052 batches | lr 0.100 | ms/batch  7.53 | loss  0.60 |
| epoch   7 |    800/ 1052 batches | lr 0.100 | ms/batch  7.19 | loss  0.59 |
| epoch   7 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.41 | loss  0.58 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time: 7.75s | valid loss  0.63 | valid acc    0.719 |
-----------------------------------------------------------------------------------------
| epoch   8 |    200/ 1052 batches | lr 0.025 | ms/batch  7.35 | loss  0.57 |
| epoch   8 |    400/ 1052 batches | lr 0.025 | ms/batch  7.41 | loss  0.56 |
| epoch   8 |    600/ 1052 batches | lr 0.025 | ms/batch  7.34 | loss  0.56 |
| epoch   8 |    800/ 1052 batches | lr 0.025 | ms/batch  7.13 | loss  0.55 |
| epoch   8 |   1000/ 1052 batches | lr 0.025 | ms/batch  7.46 | loss  0.55 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time: 7.78s | valid loss  0.54 | valid acc    0.748 |
```

```
-------------------------------------------------------------------------------
| epoch   9 |   200/ 1052 batches | lr 0.025 | ms/batch  7.50 | loss  0.55 |
| epoch   9 |   400/ 1052 batches | lr 0.025 | ms/batch  7.30 | loss  0.55 |
| epoch   9 |   600/ 1052 batches | lr 0.025 | ms/batch  7.47 | loss  0.55 |
| epoch   9 |   800/ 1052 batches | lr 0.025 | ms/batch  7.40 | loss  0.54 |
| epoch   9 |  1000/ 1052 batches | lr 0.025 | ms/batch  7.46 | loss  0.54 |
-------------------------------------------------------------------------------
| end of epoch   9 | time: 7.88s | valid loss  0.53 | valid acc    0.745 |
-------------------------------------------------------------------------------
| epoch  10 |   200/ 1052 batches | lr 0.025 | ms/batch  7.76 | loss  0.54 |
| epoch  10 |   400/ 1052 batches | lr 0.025 | ms/batch  7.56 | loss  0.54 |
| epoch  10 |   600/ 1052 batches | lr 0.025 | ms/batch  7.43 | loss  0.54 |
| epoch  10 |   800/ 1052 batches | lr 0.025 | ms/batch  7.44 | loss  0.53 |
| epoch  10 |  1000/ 1052 batches | lr 0.025 | ms/batch  7.35 | loss  0.52 |
-------------------------------------------------------------------------------
| end of epoch  10 | time: 7.97s | valid loss  0.53 | valid acc    0.752 |
-------------------------------------------------------------------------------
=================================================================================
==
| End of training | test loss  0.58 | test acc    0.743 |
=================================================================================
==
-------------------------------------------------------------------------------
learning_rate: 0.1, dropout: 0.2, batch_size: 32
-------------------------------------------------------------------------------
| epoch   1 |   200/ 2104 batches | lr 0.100 | ms/batch  8.66 | loss  0.69 |
| epoch   1 |   400/ 2104 batches | lr 0.100 | ms/batch  7.15 | loss  0.69 |
| epoch   1 |   600/ 2104 batches | lr 0.100 | ms/batch  7.02 | loss  0.69 |
| epoch   1 |   800/ 2104 batches | lr 0.100 | ms/batch  6.91 | loss  0.69 |
| epoch   1 |  1000/ 2104 batches | lr 0.100 | ms/batch  6.41 | loss  0.68 |
| epoch   1 |  1200/ 2104 batches | lr 0.100 | ms/batch  5.99 | loss  0.68 |
| epoch   1 |  1400/ 2104 batches | lr 0.100 | ms/batch  6.94 | loss  0.69 |
| epoch   1 |  1600/ 2104 batches | lr 0.100 | ms/batch  5.73 | loss  0.69 |
| epoch   1 |  1800/ 2104 batches | lr 0.100 | ms/batch  6.81 | loss  0.69 |
| epoch   1 |  2000/ 2104 batches | lr 0.100 | ms/batch  6.94 | loss  0.69 |
-------------------------------------------------------------------------------
| end of epoch   1 | time: 14.49s | valid loss 0.71 | valid acc    0.540 |
-------------------------------------------------------------------------------
| epoch   2 |   200/ 2104 batches | lr 0.100 | ms/batch  6.92 | loss  0.69 |
| epoch   2 |   400/ 2104 batches | lr 0.100 | ms/batch  6.87 | loss  0.69 |
| epoch   2 |   600/ 2104 batches | lr 0.100 | ms/batch  6.93 | loss  0.69 |
| epoch   2 |   800/ 2104 batches | lr 0.100 | ms/batch  6.96 | loss  0.69 |
| epoch   2 |  1000/ 2104 batches | lr 0.100 | ms/batch  6.95 | loss  0.68 |
| epoch   2 |  1200/ 2104 batches | lr 0.100 | ms/batch  7.02 | loss  0.68 |
| epoch   2 |  1400/ 2104 batches | lr 0.100 | ms/batch  7.01 | loss  0.68 |
| epoch   2 |  1600/ 2104 batches | lr 0.100 | ms/batch  7.02 | loss  0.68 |
| epoch   2 |  1800/ 2104 batches | lr 0.100 | ms/batch  7.08 | loss  0.68 |
| epoch   2 |  2000/ 2104 batches | lr 0.100 | ms/batch  7.22 | loss  0.68 |
-------------------------------------------------------------------------------
| end of epoch   2 | time: 14.80s | valid loss 0.71 | valid acc    0.540 |
-------------------------------------------------------------------------------
```

```
| epoch   3 |   200/ 2104 batches | lr 0.100 | ms/batch 5.82 | loss  0.69 |
| epoch   3 |   400/ 2104 batches | lr 0.100 | ms/batch 5.92 | loss  0.68 |
| epoch   3 |   600/ 2104 batches | lr 0.100 | ms/batch 5.79 | loss  0.68 |
| epoch   3 |   800/ 2104 batches | lr 0.100 | ms/batch 7.20 | loss  0.68 |
| epoch   3 |  1000/ 2104 batches | lr 0.100 | ms/batch 7.22 | loss  0.68 |
| epoch   3 |  1200/ 2104 batches | lr 0.100 | ms/batch 7.27 | loss  0.68 |
| epoch   3 |  1400/ 2104 batches | lr 0.100 | ms/batch 7.16 | loss  0.68 |
| epoch   3 |  1600/ 2104 batches | lr 0.100 | ms/batch 7.19 | loss  0.68 |
| epoch   3 |  1800/ 2104 batches | lr 0.100 | ms/batch 7.51 | loss  0.67 |
| epoch   3 |  2000/ 2104 batches | lr 0.100 | ms/batch 7.25 | loss  0.66 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time: 14.47s | valid loss 0.65 | valid acc   0.674 |
-----------------------------------------------------------------------------------------
| epoch   4 |   200/ 2104 batches | lr 0.100 | ms/batch 7.78 | loss  0.65 |
| epoch   4 |   400/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.64 |
| epoch   4 |   600/ 2104 batches | lr 0.100 | ms/batch 7.17 | loss  0.64 |
| epoch   4 |   800/ 2104 batches | lr 0.100 | ms/batch 7.25 | loss  0.63 |
| epoch   4 |  1000/ 2104 batches | lr 0.100 | ms/batch 7.25 | loss  0.62 |
| epoch   4 |  1200/ 2104 batches | lr 0.100 | ms/batch 7.33 | loss  0.62 |
| epoch   4 |  1400/ 2104 batches | lr 0.100 | ms/batch 7.17 | loss  0.61 |
| epoch   4 |  1600/ 2104 batches | lr 0.100 | ms/batch 7.07 | loss  0.60 |
| epoch   4 |  1800/ 2104 batches | lr 0.100 | ms/batch 6.93 | loss  0.60 |
| epoch   4 |  2000/ 2104 batches | lr 0.100 | ms/batch 6.91 | loss  0.58 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 15.21s | valid loss 0.60 | valid acc   0.745 |
-----------------------------------------------------------------------------------------
| epoch   5 |   200/ 2104 batches | lr 0.100 | ms/batch 6.89 | loss  0.59 |
| epoch   5 |   400/ 2104 batches | lr 0.100 | ms/batch 6.84 | loss  0.58 |
| epoch   5 |   600/ 2104 batches | lr 0.100 | ms/batch 6.83 | loss  0.58 |
| epoch   5 |   800/ 2104 batches | lr 0.100 | ms/batch 7.02 | loss  0.56 |
| epoch   5 |  1000/ 2104 batches | lr 0.100 | ms/batch 7.26 | loss  0.56 |
| epoch   5 |  1200/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.56 |
| epoch   5 |  1400/ 2104 batches | lr 0.100 | ms/batch 7.22 | loss  0.55 |
| epoch   5 |  1600/ 2104 batches | lr 0.100 | ms/batch 7.28 | loss  0.54 |
| epoch   5 |  1800/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.54 |
| epoch   5 |  2000/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.53 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 15.02s | valid loss 0.53 | valid acc   0.764 |
-----------------------------------------------------------------------------------------
| epoch   6 |   200/ 2104 batches | lr 0.100 | ms/batch 7.27 | loss  0.53 |
| epoch   6 |   400/ 2104 batches | lr 0.100 | ms/batch 7.15 | loss  0.52 |
| epoch   6 |   600/ 2104 batches | lr 0.100 | ms/batch 7.10 | loss  0.53 |
| epoch   6 |   800/ 2104 batches | lr 0.100 | ms/batch 7.08 | loss  0.51 |
| epoch   6 |  1000/ 2104 batches | lr 0.100 | ms/batch 7.12 | loss  0.51 |
| epoch   6 |  1200/ 2104 batches | lr 0.100 | ms/batch 7.12 | loss  0.51 |
| epoch   6 |  1400/ 2104 batches | lr 0.100 | ms/batch 7.12 | loss  0.49 |
| epoch   6 |  1600/ 2104 batches | lr 0.100 | ms/batch 7.12 | loss  0.50 |
| epoch   6 |  1800/ 2104 batches | lr 0.100 | ms/batch 7.24 | loss  0.50 |
| epoch   6 |  2000/ 2104 batches | lr 0.100 | ms/batch 7.05 | loss  0.47 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time: 15.08s | valid loss 0.52 | valid acc   0.786 |
```

```
--------------------------------------------------------------------------------------
| epoch   7 |    200/ 2104 batches | lr 0.100 | ms/batch 7.39 | loss  0.49 |
| epoch   7 |    400/ 2104 batches | lr 0.100 | ms/batch 7.37 | loss  0.48 |
| epoch   7 |    600/ 2104 batches | lr 0.100 | ms/batch 7.41 | loss  0.48 |
| epoch   7 |    800/ 2104 batches | lr 0.100 | ms/batch 7.31 | loss  0.47 |
| epoch   7 |   1000/ 2104 batches | lr 0.100 | ms/batch 7.28 | loss  0.47 |
| epoch   7 |   1200/ 2104 batches | lr 0.100 | ms/batch 7.25 | loss  0.47 |
| epoch   7 |   1400/ 2104 batches | lr 0.100 | ms/batch 7.22 | loss  0.46 |
| epoch   7 |   1600/ 2104 batches | lr 0.100 | ms/batch 7.22 | loss  0.45 |
| epoch   7 |   1800/ 2104 batches | lr 0.100 | ms/batch 7.06 | loss  0.46 |
| epoch   7 |   2000/ 2104 batches | lr 0.100 | ms/batch 6.98 | loss  0.44 |
--------------------------------------------------------------------------------------
| end of epoch   7 | time: 15.27s | valid loss  0.53 | valid acc    0.781 |
--------------------------------------------------------------------------------------
| epoch   8 |    200/ 2104 batches | lr 0.025 | ms/batch 7.01 | loss  0.44 |
| epoch   8 |    400/ 2104 batches | lr 0.025 | ms/batch 7.12 | loss  0.43 |
| epoch   8 |    600/ 2104 batches | lr 0.025 | ms/batch 7.08 | loss  0.44 |
| epoch   8 |    800/ 2104 batches | lr 0.025 | ms/batch 7.12 | loss  0.43 |
| epoch   8 |   1000/ 2104 batches | lr 0.025 | ms/batch 7.08 | loss  0.43 |
| epoch   8 |   1200/ 2104 batches | lr 0.025 | ms/batch 7.09 | loss  0.43 |
| epoch   8 |   1400/ 2104 batches | lr 0.025 | ms/batch 7.07 | loss  0.43 |
| epoch   8 |   1600/ 2104 batches | lr 0.025 | ms/batch 7.06 | loss  0.42 |
| epoch   8 |   1800/ 2104 batches | lr 0.025 | ms/batch 6.97 | loss  0.43 |
| epoch   8 |   2000/ 2104 batches | lr 0.025 | ms/batch 7.07 | loss  0.41 |
--------------------------------------------------------------------------------------
| end of epoch   8 | time: 14.93s | valid loss  0.56 | valid acc    0.779 |
--------------------------------------------------------------------------------------
| epoch   9 |    200/ 2104 batches | lr 0.006 | ms/batch 7.11 | loss  0.43 |
| epoch   9 |    400/ 2104 batches | lr 0.006 | ms/batch 7.10 | loss  0.42 |
| epoch   9 |    600/ 2104 batches | lr 0.006 | ms/batch 7.08 | loss  0.43 |
| epoch   9 |    800/ 2104 batches | lr 0.006 | ms/batch 7.07 | loss  0.42 |
| epoch   9 |   1000/ 2104 batches | lr 0.006 | ms/batch 7.11 | loss  0.42 |
| epoch   9 |   1200/ 2104 batches | lr 0.006 | ms/batch 7.18 | loss  0.43 |
| epoch   9 |   1400/ 2104 batches | lr 0.006 | ms/batch 7.22 | loss  0.42 |
| epoch   9 |   1600/ 2104 batches | lr 0.006 | ms/batch 7.07 | loss  0.41 |
| epoch   9 |   1800/ 2104 batches | lr 0.006 | ms/batch 7.10 | loss  0.42 |
| epoch   9 |   2000/ 2104 batches | lr 0.006 | ms/batch 7.13 | loss  0.40 |
--------------------------------------------------------------------------------------
| end of epoch   9 | time: 15.03s | valid loss  0.55 | valid acc    0.790 |
--------------------------------------------------------------------------------------
| epoch  10 |    200/ 2104 batches | lr 0.002 | ms/batch 6.66 | loss  0.42 |
| epoch  10 |    400/ 2104 batches | lr 0.002 | ms/batch 7.18 | loss  0.42 |
| epoch  10 |    600/ 2104 batches | lr 0.002 | ms/batch 7.23 | loss  0.43 |
| epoch  10 |    800/ 2104 batches | lr 0.002 | ms/batch 7.26 | loss  0.41 |
| epoch  10 |   1000/ 2104 batches | lr 0.002 | ms/batch 7.16 | loss  0.42 |
| epoch  10 |   1200/ 2104 batches | lr 0.002 | ms/batch 7.12 | loss  0.43 |
| epoch  10 |   1400/ 2104 batches | lr 0.002 | ms/batch 7.12 | loss  0.42 |
| epoch  10 |   1600/ 2104 batches | lr 0.002 | ms/batch 7.25 | loss  0.41 |
| epoch  10 |   1800/ 2104 batches | lr 0.002 | ms/batch 7.26 | loss  0.42 |
| epoch  10 |   2000/ 2104 batches | lr 0.002 | ms/batch 7.25 | loss  0.40 |
--------------------------------------------------------------------------------------
```

```
| end of epoch  10 | time: 15.11s | valid loss  0.55 | valid acc    0.798 |
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  0.56 | test acc    0.776 |
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 0.1, dropout: 0.2, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |    200/ 1052 batches | lr 0.100 | ms/batch  8.69 | loss  0.69 |
| epoch   1 |    400/ 1052 batches | lr 0.100 | ms/batch  7.52 | loss  0.69 |
| epoch   1 |    600/ 1052 batches | lr 0.100 | ms/batch  7.51 | loss  0.68 |
| epoch   1 |    800/ 1052 batches | lr 0.100 | ms/batch  7.52 | loss  0.69 |
| epoch   1 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.62 | loss  0.69 |
-----------------------------------------------------------------------------------------
| end of epoch   1 | time: 8.23s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   2 |    200/ 1052 batches | lr 0.100 | ms/batch  7.35 | loss  0.69 |
| epoch   2 |    400/ 1052 batches | lr 0.100 | ms/batch  7.29 | loss  0.69 |
| epoch   2 |    600/ 1052 batches | lr 0.100 | ms/batch  7.41 | loss  0.68 |
| epoch   2 |    800/ 1052 batches | lr 0.100 | ms/batch  7.32 | loss  0.68 |
| epoch   2 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.23 | loss  0.69 |
-----------------------------------------------------------------------------------------
| end of epoch   2 | time: 7.76s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   3 |    200/ 1052 batches | lr 0.100 | ms/batch  7.18 | loss  0.69 |
| epoch   3 |    400/ 1052 batches | lr 0.100 | ms/batch  7.20 | loss  0.69 |
| epoch   3 |    600/ 1052 batches | lr 0.100 | ms/batch  7.20 | loss  0.68 |
| epoch   3 |    800/ 1052 batches | lr 0.100 | ms/batch  7.24 | loss  0.68 |
| epoch   3 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.36 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time: 7.68s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   4 |    200/ 1052 batches | lr 0.100 | ms/batch  7.38 | loss  0.69 |
| epoch   4 |    400/ 1052 batches | lr 0.100 | ms/batch  7.33 | loss  0.68 |
| epoch   4 |    600/ 1052 batches | lr 0.100 | ms/batch  7.27 | loss  0.68 |
| epoch   4 |    800/ 1052 batches | lr 0.100 | ms/batch  7.37 | loss  0.68 |
| epoch   4 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.53 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 7.83s | valid loss  0.71 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   5 |    200/ 1052 batches | lr 0.100 | ms/batch  7.37 | loss  0.69 |
| epoch   5 |    400/ 1052 batches | lr 0.100 | ms/batch  7.44 | loss  0.68 |
| epoch   5 |    600/ 1052 batches | lr 0.100 | ms/batch  7.47 | loss  0.68 |
| epoch   5 |    800/ 1052 batches | lr 0.100 | ms/batch  7.48 | loss  0.68 |
| epoch   5 |   1000/ 1052 batches | lr 0.100 | ms/batch  7.46 | loss  0.68 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 7.89s | valid loss  0.69 | valid acc    0.614 |
-----------------------------------------------------------------------------------------
| epoch   6 |    200/ 1052 batches | lr 0.100 | ms/batch  7.43 | loss  0.68 |
```

```
| epoch   6 |    400/ 1052 batches | lr 0.100 | ms/batch 7.44 | loss  0.67 |
| epoch   6 |    600/ 1052 batches | lr 0.100 | ms/batch 7.32 | loss  0.65 |
| epoch   6 |    800/ 1052 batches | lr 0.100 | ms/batch 7.51 | loss  0.65 |
| epoch   6 |   1000/ 1052 batches | lr 0.100 | ms/batch 7.55 | loss  0.63 |
-----------------------------------------------------------------------------------
| end of epoch   6 | time: 7.90s | valid loss  0.60 | valid acc    0.702 |
-----------------------------------------------------------------------------------
| epoch   7 |    200/ 1052 batches | lr 0.100 | ms/batch 7.65 | loss  0.63 |
| epoch   7 |    400/ 1052 batches | lr 0.100 | ms/batch 7.41 | loss  0.62 |
| epoch   7 |    600/ 1052 batches | lr 0.100 | ms/batch 7.54 | loss  0.61 |
| epoch   7 |    800/ 1052 batches | lr 0.100 | ms/batch 7.55 | loss  0.60 |
| epoch   7 |   1000/ 1052 batches | lr 0.100 | ms/batch 7.41 | loss  0.59 |
-----------------------------------------------------------------------------------
| end of epoch   7 | time: 7.97s | valid loss  0.62 | valid acc    0.719 |
-----------------------------------------------------------------------------------
| epoch   8 |    200/ 1052 batches | lr 0.025 | ms/batch 7.49 | loss  0.58 |
| epoch   8 |    400/ 1052 batches | lr 0.025 | ms/batch 7.46 | loss  0.57 |
| epoch   8 |    600/ 1052 batches | lr 0.025 | ms/batch 7.37 | loss  0.57 |
| epoch   8 |    800/ 1052 batches | lr 0.025 | ms/batch 7.31 | loss  0.56 |
| epoch   8 |   1000/ 1052 batches | lr 0.025 | ms/batch 7.27 | loss  0.56 |
-----------------------------------------------------------------------------------
| end of epoch   8 | time: 7.82s | valid loss  0.54 | valid acc    0.740 |
-----------------------------------------------------------------------------------
| epoch   9 |    200/ 1052 batches | lr 0.025 | ms/batch 7.31 | loss  0.56 |
| epoch   9 |    400/ 1052 batches | lr 0.025 | ms/batch 6.59 | loss  0.56 |
| epoch   9 |    600/ 1052 batches | lr 0.025 | ms/batch 5.33 | loss  0.56 |
| epoch   9 |    800/ 1052 batches | lr 0.025 | ms/batch 6.25 | loss  0.55 |
| epoch   9 |   1000/ 1052 batches | lr 0.025 | ms/batch 7.65 | loss  0.55 |
-----------------------------------------------------------------------------------
| end of epoch   9 | time: 7.09s | valid loss  0.53 | valid acc    0.755 |
-----------------------------------------------------------------------------------
| epoch  10 |    200/ 1052 batches | lr 0.025 | ms/batch 7.64 | loss  0.55 |
| epoch  10 |    400/ 1052 batches | lr 0.025 | ms/batch 7.68 | loss  0.55 |
| epoch  10 |    600/ 1052 batches | lr 0.025 | ms/batch 7.51 | loss  0.55 |
| epoch  10 |    800/ 1052 batches | lr 0.025 | ms/batch 7.51 | loss  0.54 |
| epoch  10 |   1000/ 1052 batches | lr 0.025 | ms/batch 7.50 | loss  0.53 |
-----------------------------------------------------------------------------------
| end of epoch  10 | time: 8.02s | valid loss  0.53 | valid acc    0.745 |
-----------------------------------------------------------------------------------
===================================================================================
==
| End of training | test loss  0.56 | test acc    0.738 |
===================================================================================
==
-----------------------------------------------------------------------------------
learning_rate: 1, dropout: 0.1, batch_size: 32
-----------------------------------------------------------------------------------
| epoch   1 |    200/ 2104 batches | lr 1.000 | ms/batch 7.99 | loss  0.70 |
| epoch   1 |    400/ 2104 batches | lr 1.000 | ms/batch 6.92 | loss  0.69 |
| epoch   1 |    600/ 2104 batches | lr 1.000 | ms/batch 6.94 | loss  0.69 |
| epoch   1 |    800/ 2104 batches | lr 1.000 | ms/batch 6.94 | loss  0.69 |
```

```
| epoch   1 |  1000/ 2104 batches | lr 1.000 | ms/batch 7.01 | loss 0.69 |
| epoch   1 |  1200/ 2104 batches | lr 1.000 | ms/batch 6.96 | loss 0.69 |
| epoch   1 |  1400/ 2104 batches | lr 1.000 | ms/batch 7.07 | loss 0.69 |
| epoch   1 |  1600/ 2104 batches | lr 1.000 | ms/batch 7.10 | loss 0.68 |
| epoch   1 |  1800/ 2104 batches | lr 1.000 | ms/batch 6.94 | loss 0.65 |
| epoch   1 |  2000/ 2104 batches | lr 1.000 | ms/batch 6.99 | loss 0.60 |
-----------------------------------------------------------------------------------
| end of epoch   1 | time: 14.96s | valid loss 0.67 | valid acc   0.662 |
-----------------------------------------------------------------------------------
| epoch   2 |   200/ 2104 batches | lr 1.000 | ms/batch 7.12 | loss 0.55 |
| epoch   2 |   400/ 2104 batches | lr 1.000 | ms/batch 7.07 | loss 0.51 |
| epoch   2 |   600/ 2104 batches | lr 1.000 | ms/batch 6.97 | loss 0.49 |
| epoch   2 |   800/ 2104 batches | lr 1.000 | ms/batch 7.08 | loss 0.47 |
| epoch   2 |  1000/ 2104 batches | lr 1.000 | ms/batch 7.02 | loss 0.45 |
| epoch   2 |  1200/ 2104 batches | lr 1.000 | ms/batch 7.10 | loss 0.44 |
| epoch   2 |  1400/ 2104 batches | lr 1.000 | ms/batch 7.29 | loss 0.41 |
| epoch   2 |  1600/ 2104 batches | lr 1.000 | ms/batch 7.33 | loss 0.39 |
| epoch   2 |  1800/ 2104 batches | lr 1.000 | ms/batch 7.30 | loss 0.38 |
| epoch   2 |  2000/ 2104 batches | lr 1.000 | ms/batch 7.30 | loss 0.36 |
-----------------------------------------------------------------------------------
| end of epoch   2 | time: 15.13s | valid loss 0.84 | valid acc   0.764 |
-----------------------------------------------------------------------------------
| epoch   3 |   200/ 2104 batches | lr 0.250 | ms/batch 7.33 | loss 0.32 |
| epoch   3 |   400/ 2104 batches | lr 0.250 | ms/batch 7.39 | loss 0.30 |
| epoch   3 |   600/ 2104 batches | lr 0.250 | ms/batch 7.20 | loss 0.31 |
| epoch   3 |   800/ 2104 batches | lr 0.250 | ms/batch 7.06 | loss 0.29 |
| epoch   3 |  1000/ 2104 batches | lr 0.250 | ms/batch 7.12 | loss 0.28 |
| epoch   3 |  1200/ 2104 batches | lr 0.250 | ms/batch 7.14 | loss 0.29 |
| epoch   3 |  1400/ 2104 batches | lr 0.250 | ms/batch 7.22 | loss 0.27 |
| epoch   3 |  1600/ 2104 batches | lr 0.250 | ms/batch 7.24 | loss 0.26 |
| epoch   3 |  1800/ 2104 batches | lr 0.250 | ms/batch 7.27 | loss 0.26 |
| epoch   3 |  2000/ 2104 batches | lr 0.250 | ms/batch 7.23 | loss 0.24 |
-----------------------------------------------------------------------------------
| end of epoch   3 | time: 15.25s | valid loss 0.65 | valid acc   0.812 |
-----------------------------------------------------------------------------------
| epoch   4 |   200/ 2104 batches | lr 0.250 | ms/batch 7.30 | loss 0.28 |
| epoch   4 |   400/ 2104 batches | lr 0.250 | ms/batch 7.19 | loss 0.27 |
| epoch   4 |   600/ 2104 batches | lr 0.250 | ms/batch 7.08 | loss 0.27 |
| epoch   4 |   800/ 2104 batches | lr 0.250 | ms/batch 7.23 | loss 0.26 |
| epoch   4 |  1000/ 2104 batches | lr 0.250 | ms/batch 7.28 | loss 0.26 |
| epoch   4 |  1200/ 2104 batches | lr 0.250 | ms/batch 7.30 | loss 0.27 |
| epoch   4 |  1400/ 2104 batches | lr 0.250 | ms/batch 7.35 | loss 0.24 |
| epoch   4 |  1600/ 2104 batches | lr 0.250 | ms/batch 7.28 | loss 0.24 |
| epoch   4 |  1800/ 2104 batches | lr 0.250 | ms/batch 7.26 | loss 0.24 |
| epoch   4 |  2000/ 2104 batches | lr 0.250 | ms/batch 7.32 | loss 0.22 |
-----------------------------------------------------------------------------------
| end of epoch   4 | time: 15.34s | valid loss 0.64 | valid acc   0.817 |
-----------------------------------------------------------------------------------
| epoch   5 |   200/ 2104 batches | lr 0.250 | ms/batch 7.33 | loss 0.26 |
| epoch   5 |   400/ 2104 batches | lr 0.250 | ms/batch 7.17 | loss 0.24 |
| epoch   5 |   600/ 2104 batches | lr 0.250 | ms/batch 7.11 | loss 0.25 |
```

```
| epoch   5 |   800/ 2104 batches | lr 0.250 | ms/batch 7.11 | loss  0.24 |
| epoch   5 |  1000/ 2104 batches | lr 0.250 | ms/batch 7.22 | loss  0.23 |
| epoch   5 |  1200/ 2104 batches | lr 0.250 | ms/batch 7.49 | loss  0.25 |
| epoch   5 |  1400/ 2104 batches | lr 0.250 | ms/batch 7.22 | loss  0.22 |
| epoch   5 |  1600/ 2104 batches | lr 0.250 | ms/batch 7.31 | loss  0.22 |
| epoch   5 |  1800/ 2104 batches | lr 0.250 | ms/batch 7.26 | loss  0.22 |
| epoch   5 |  2000/ 2104 batches | lr 0.250 | ms/batch 7.39 | loss  0.21 |
-----------------------------------------------------------------------------------
| end of epoch   5 | time: 15.30s | valid loss  0.66 | valid acc    0.829 |
-----------------------------------------------------------------------------------
| epoch   6 |   200/ 2104 batches | lr 0.062 | ms/batch 7.02 | loss  0.24 |
| epoch   6 |   400/ 2104 batches | lr 0.062 | ms/batch 7.15 | loss  0.22 |
| epoch   6 |   600/ 2104 batches | lr 0.062 | ms/batch 7.22 | loss  0.23 |
| epoch   6 |   800/ 2104 batches | lr 0.062 | ms/batch 7.26 | loss  0.22 |
| epoch   6 |  1000/ 2104 batches | lr 0.062 | ms/batch 7.31 | loss  0.20 |
| epoch   6 |  1200/ 2104 batches | lr 0.062 | ms/batch 7.45 | loss  0.22 |
| epoch   6 |  1400/ 2104 batches | lr 0.062 | ms/batch 7.15 | loss  0.20 |
| epoch   6 |  1600/ 2104 batches | lr 0.062 | ms/batch 7.07 | loss  0.19 |
| epoch   6 |  1800/ 2104 batches | lr 0.062 | ms/batch 7.07 | loss  0.19 |
| epoch   6 |  2000/ 2104 batches | lr 0.062 | ms/batch 7.23 | loss  0.18 |
-----------------------------------------------------------------------------------
| end of epoch   6 | time: 15.21s | valid loss  0.67 | valid acc    0.821 |
-----------------------------------------------------------------------------------
| epoch   7 |   200/ 2104 batches | lr 0.016 | ms/batch 7.27 | loss  0.23 |
| epoch   7 |   400/ 2104 batches | lr 0.016 | ms/batch 7.23 | loss  0.21 |
| epoch   7 |   600/ 2104 batches | lr 0.016 | ms/batch 7.21 | loss  0.22 |
| epoch   7 |   800/ 2104 batches | lr 0.016 | ms/batch 7.23 | loss  0.21 |
| epoch   7 |  1000/ 2104 batches | lr 0.016 | ms/batch 7.19 | loss  0.19 |
| epoch   7 |  1200/ 2104 batches | lr 0.016 | ms/batch 7.21 | loss  0.22 |
| epoch   7 |  1400/ 2104 batches | lr 0.016 | ms/batch 7.33 | loss  0.19 |
| epoch   7 |  1600/ 2104 batches | lr 0.016 | ms/batch 6.94 | loss  0.18 |
| epoch   7 |  1800/ 2104 batches | lr 0.016 | ms/batch 6.75 | loss  0.18 |
| epoch   7 |  2000/ 2104 batches | lr 0.016 | ms/batch 5.84 | loss  0.17 |
-----------------------------------------------------------------------------------
| end of epoch   7 | time: 14.84s | valid loss  0.67 | valid acc    0.821 |
-----------------------------------------------------------------------------------
| epoch   8 |   200/ 2104 batches | lr 0.004 | ms/batch 7.32 | loss  0.23 |
| epoch   8 |   400/ 2104 batches | lr 0.004 | ms/batch 7.30 | loss  0.20 |
| epoch   8 |   600/ 2104 batches | lr 0.004 | ms/batch 7.28 | loss  0.22 |
| epoch   8 |   800/ 2104 batches | lr 0.004 | ms/batch 7.32 | loss  0.21 |
| epoch   8 |  1000/ 2104 batches | lr 0.004 | ms/batch 7.37 | loss  0.19 |
| epoch   8 |  1200/ 2104 batches | lr 0.004 | ms/batch 7.29 | loss  0.21 |
| epoch   8 |  1400/ 2104 batches | lr 0.004 | ms/batch 7.16 | loss  0.19 |
| epoch   8 |  1600/ 2104 batches | lr 0.004 | ms/batch 7.11 | loss  0.18 |
| epoch   8 |  1800/ 2104 batches | lr 0.004 | ms/batch 7.18 | loss  0.18 |
| epoch   8 |  2000/ 2104 batches | lr 0.004 | ms/batch 7.21 | loss  0.17 |
-----------------------------------------------------------------------------------
| end of epoch   8 | time: 15.32s | valid loss  0.67 | valid acc    0.829 |
-----------------------------------------------------------------------------------
| epoch   9 |   200/ 2104 batches | lr 0.001 | ms/batch 7.31 | loss  0.23 |
| epoch   9 |   400/ 2104 batches | lr 0.001 | ms/batch 7.25 | loss  0.21 |
```

```
| epoch   9 |    600/ 2104 batches | lr 0.001 | ms/batch  7.27 | loss  0.22 |
| epoch   9 |    800/ 2104 batches | lr 0.001 | ms/batch  7.19 | loss  0.21 |
| epoch   9 |   1000/ 2104 batches | lr 0.001 | ms/batch  7.23 | loss  0.19 |
| epoch   9 |   1200/ 2104 batches | lr 0.001 | ms/batch  7.33 | loss  0.21 |
| epoch   9 |   1400/ 2104 batches | lr 0.001 | ms/batch  7.26 | loss  0.19 |
| epoch   9 |   1600/ 2104 batches | lr 0.001 | ms/batch  7.28 | loss  0.18 |
| epoch   9 |   1800/ 2104 batches | lr 0.001 | ms/batch  7.10 | loss  0.18 |
| epoch   9 |   2000/ 2104 batches | lr 0.001 | ms/batch  7.23 | loss  0.17 |
-----------------------------------------------------------------------------------------
| end of epoch   9 | time: 15.28s | valid loss  0.67 | valid acc    0.831 |
-----------------------------------------------------------------------------------------
| epoch  10 |    200/ 2104 batches | lr 0.000 | ms/batch  7.01 | loss  0.23 |
| epoch  10 |    400/ 2104 batches | lr 0.000 | ms/batch  6.83 | loss  0.21 |
| epoch  10 |    600/ 2104 batches | lr 0.000 | ms/batch  7.39 | loss  0.22 |
| epoch  10 |    800/ 2104 batches | lr 0.000 | ms/batch  7.07 | loss  0.21 |
| epoch  10 |   1000/ 2104 batches | lr 0.000 | ms/batch  7.07 | loss  0.19 |
| epoch  10 |   1200/ 2104 batches | lr 0.000 | ms/batch  7.28 | loss  0.21 |
| epoch  10 |   1400/ 2104 batches | lr 0.000 | ms/batch  7.29 | loss  0.19 |
| epoch  10 |   1600/ 2104 batches | lr 0.000 | ms/batch  7.04 | loss  0.18 |
| epoch  10 |   1800/ 2104 batches | lr 0.000 | ms/batch  6.99 | loss  0.18 |
| epoch  10 |   2000/ 2104 batches | lr 0.000 | ms/batch  7.07 | loss  0.17 |
-----------------------------------------------------------------------------------------
| end of epoch  10 | time: 15.03s | valid loss  0.67 | valid acc    0.824 |
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  0.56 | test acc    0.855 |
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 1, dropout: 0.1, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch   1 |    200/ 1052 batches | lr 1.000 | ms/batch  8.68 | loss  0.69 |
| epoch   1 |    400/ 1052 batches | lr 1.000 | ms/batch  7.51 | loss  0.69 |
| epoch   1 |    600/ 1052 batches | lr 1.000 | ms/batch  7.44 | loss  0.69 |
| epoch   1 |    800/ 1052 batches | lr 1.000 | ms/batch  7.50 | loss  0.69 |
| epoch   1 |   1000/ 1052 batches | lr 1.000 | ms/batch  7.50 | loss  0.69 |
-----------------------------------------------------------------------------------------
| end of epoch   1 | time: 8.18s | valid loss  0.70 | valid acc    0.555 |
-----------------------------------------------------------------------------------------
| epoch   2 |    200/ 1052 batches | lr 1.000 | ms/batch  7.42 | loss  0.67 |
| epoch   2 |    400/ 1052 batches | lr 1.000 | ms/batch  7.44 | loss  0.63 |
| epoch   2 |    600/ 1052 batches | lr 1.000 | ms/batch  7.21 | loss  0.57 |
| epoch   2 |    800/ 1052 batches | lr 1.000 | ms/batch  7.47 | loss  0.53 |
| epoch   2 |   1000/ 1052 batches | lr 1.000 | ms/batch  7.40 | loss  0.49 |
-----------------------------------------------------------------------------------------
| end of epoch   2 | time: 7.83s | valid loss  0.57 | valid acc    0.736 |
-----------------------------------------------------------------------------------------
| epoch   3 |    200/ 1052 batches | lr 1.000 | ms/batch  7.49 | loss  0.45 |
| epoch   3 |    400/ 1052 batches | lr 1.000 | ms/batch  7.36 | loss  0.43 |
| epoch   3 |    600/ 1052 batches | lr 1.000 | ms/batch  7.36 | loss  0.41 |
```

```
| epoch   3 |    800/ 1052 batches | lr 1.000 | ms/batch  7.40 | loss  0.38 |
| epoch   3 |   1000/ 1052 batches | lr 1.000 | ms/batch  7.39 | loss  0.36 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  7.84s | valid loss  0.58 | valid acc    0.786 |
-----------------------------------------------------------------------------------------
| epoch   4 |    200/ 1052 batches | lr 0.250 | ms/batch  7.43 | loss  0.31 |
| epoch   4 |    400/ 1052 batches | lr 0.250 | ms/batch  7.49 | loss  0.30 |
| epoch   4 |    600/ 1052 batches | lr 0.250 | ms/batch  6.48 | loss  0.28 |
| epoch   4 |    800/ 1052 batches | lr 0.250 | ms/batch  7.51 | loss  0.27 |
| epoch   4 |   1000/ 1052 batches | lr 0.250 | ms/batch  7.44 | loss  0.26 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  7.72s | valid loss  0.61 | valid acc    0.812 |
-----------------------------------------------------------------------------------------
| epoch   5 |    200/ 1052 batches | lr 0.062 | ms/batch  7.52 | loss  0.27 |
| epoch   5 |    400/ 1052 batches | lr 0.062 | ms/batch  7.49 | loss  0.27 |
| epoch   5 |    600/ 1052 batches | lr 0.062 | ms/batch  7.46 | loss  0.26 |
| epoch   5 |    800/ 1052 batches | lr 0.062 | ms/batch  7.43 | loss  0.24 |
| epoch   5 |   1000/ 1052 batches | lr 0.062 | ms/batch  7.45 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  7.92s | valid loss  0.62 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch   6 |    200/ 1052 batches | lr 0.016 | ms/batch  7.55 | loss  0.26 |
| epoch   6 |    400/ 1052 batches | lr 0.016 | ms/batch  7.39 | loss  0.26 |
| epoch   6 |    600/ 1052 batches | lr 0.016 | ms/batch  7.15 | loss  0.25 |
| epoch   6 |    800/ 1052 batches | lr 0.016 | ms/batch  7.32 | loss  0.23 |
| epoch   6 |   1000/ 1052 batches | lr 0.016 | ms/batch  7.50 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  7.83s | valid loss  0.61 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch   7 |    200/ 1052 batches | lr 0.004 | ms/batch  7.62 | loss  0.26 |
| epoch   7 |    400/ 1052 batches | lr 0.004 | ms/batch  7.43 | loss  0.26 |
| epoch   7 |    600/ 1052 batches | lr 0.004 | ms/batch  7.47 | loss  0.24 |
| epoch   7 |    800/ 1052 batches | lr 0.004 | ms/batch  7.50 | loss  0.23 |
| epoch   7 |   1000/ 1052 batches | lr 0.004 | ms/batch  7.54 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  7.95s | valid loss  0.61 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch   8 |    200/ 1052 batches | lr 0.001 | ms/batch  7.43 | loss  0.26 |
| epoch   8 |    400/ 1052 batches | lr 0.001 | ms/batch  7.51 | loss  0.25 |
| epoch   8 |    600/ 1052 batches | lr 0.001 | ms/batch  7.51 | loss  0.25 |
| epoch   8 |    800/ 1052 batches | lr 0.001 | ms/batch  7.46 | loss  0.23 |
| epoch   8 |   1000/ 1052 batches | lr 0.001 | ms/batch  7.29 | loss  0.22 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time:  7.89s | valid loss  0.61 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch   9 |    200/ 1052 batches | lr 0.000 | ms/batch  7.41 | loss  0.26 |
| epoch   9 |    400/ 1052 batches | lr 0.000 | ms/batch  7.46 | loss  0.25 |
| epoch   9 |    600/ 1052 batches | lr 0.000 | ms/batch  7.40 | loss  0.24 |
| epoch   9 |    800/ 1052 batches | lr 0.000 | ms/batch  7.44 | loss  0.23 |
| epoch   9 |   1000/ 1052 batches | lr 0.000 | ms/batch  7.41 | loss  0.22 |
-----------------------------------------------------------------------------------------
```

```
| end of epoch   9 | time: 7.87s | valid loss  0.61 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch  10 |    200/ 1052 batches | lr 0.000 | ms/batch  7.45 | loss  0.26 |
| epoch  10 |    400/ 1052 batches | lr 0.000 | ms/batch  7.26 | loss  0.25 |
| epoch  10 |    600/ 1052 batches | lr 0.000 | ms/batch  7.32 | loss  0.25 |
| epoch  10 |    800/ 1052 batches | lr 0.000 | ms/batch  7.41 | loss  0.23 |
| epoch  10 |   1000/ 1052 batches | lr 0.000 | ms/batch  7.48 | loss  0.22 |
-----------------------------------------------------------------------------------------
| end of epoch  10 | time: 7.83s | valid loss  0.61 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  0.55 | test acc    0.779 |
=========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 1, dropout: 0.2, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |    200/ 2104 batches | lr 1.000 | ms/batch  8.10 | loss  0.70 |
| epoch   1 |    400/ 2104 batches | lr 1.000 | ms/batch  6.16 | loss  0.69 |
| epoch   1 |    600/ 2104 batches | lr 1.000 | ms/batch  6.76 | loss  0.69 |
| epoch   1 |    800/ 2104 batches | lr 1.000 | ms/batch  7.30 | loss  0.69 |
| epoch   1 |   1000/ 2104 batches | lr 1.000 | ms/batch  7.33 | loss  0.69 |
| epoch   1 |   1200/ 2104 batches | lr 1.000 | ms/batch  7.33 | loss  0.69 |
| epoch   1 |   1400/ 2104 batches | lr 1.000 | ms/batch  7.39 | loss  0.69 |
| epoch   1 |   1600/ 2104 batches | lr 1.000 | ms/batch  7.27 | loss  0.68 |
| epoch   1 |   1800/ 2104 batches | lr 1.000 | ms/batch  6.96 | loss  0.65 |
| epoch   1 |   2000/ 2104 batches | lr 1.000 | ms/batch  7.01 | loss  0.60 |
-----------------------------------------------------------------------------------------
| end of epoch   1 | time: 15.11s | valid loss  0.66 | valid acc    0.681 |
-----------------------------------------------------------------------------------------
| epoch   2 |    200/ 2104 batches | lr 1.000 | ms/batch  7.38 | loss  0.54 |
| epoch   2 |    400/ 2104 batches | lr 1.000 | ms/batch  7.02 | loss  0.51 |
| epoch   2 |    600/ 2104 batches | lr 1.000 | ms/batch  7.04 | loss  0.50 |
| epoch   2 |    800/ 2104 batches | lr 1.000 | ms/batch  7.12 | loss  0.48 |
| epoch   2 |   1000/ 2104 batches | lr 1.000 | ms/batch  7.02 | loss  0.46 |
| epoch   2 |   1200/ 2104 batches | lr 1.000 | ms/batch  7.28 | loss  0.44 |
| epoch   2 |   1400/ 2104 batches | lr 1.000 | ms/batch  7.28 | loss  0.41 |
| epoch   2 |   1600/ 2104 batches | lr 1.000 | ms/batch  7.27 | loss  0.39 |
| epoch   2 |   1800/ 2104 batches | lr 1.000 | ms/batch  7.29 | loss  0.38 |
| epoch   2 |   2000/ 2104 batches | lr 1.000 | ms/batch  7.27 | loss  0.37 |
-----------------------------------------------------------------------------------------
| end of epoch   2 | time: 15.21s | valid loss  0.80 | valid acc    0.769 |
-----------------------------------------------------------------------------------------
| epoch   3 |    200/ 2104 batches | lr 0.250 | ms/batch  7.32 | loss  0.33 |
| epoch   3 |    400/ 2104 batches | lr 0.250 | ms/batch  7.21 | loss  0.31 |
| epoch   3 |    600/ 2104 batches | lr 0.250 | ms/batch  7.22 | loss  0.32 |
| epoch   3 |    800/ 2104 batches | lr 0.250 | ms/batch  7.05 | loss  0.30 |
| epoch   3 |   1000/ 2104 batches | lr 0.250 | ms/batch  7.07 | loss  0.30 |
| epoch   3 |   1200/ 2104 batches | lr 0.250 | ms/batch  7.16 | loss  0.30 |
| epoch   3 |   1400/ 2104 batches | lr 0.250 | ms/batch  7.10 | loss  0.28 |
```

```
| epoch   3 |   1600/ 2104 batches | lr 0.250 | ms/batch  7.08 | loss  0.27 |
| epoch   3 |   1800/ 2104 batches | lr 0.250 | ms/batch  7.03 | loss  0.27 |
| epoch   3 |   2000/ 2104 batches | lr 0.250 | ms/batch  7.26 | loss  0.25 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time: 15.13s | valid loss  0.61 | valid acc    0.814 |
-----------------------------------------------------------------------------------------
| epoch   4 |    200/ 2104 batches | lr 0.250 | ms/batch  7.36 | loss  0.30 |
| epoch   4 |    400/ 2104 batches | lr 0.250 | ms/batch  7.31 | loss  0.27 |
| epoch   4 |    600/ 2104 batches | lr 0.250 | ms/batch  7.28 | loss  0.28 |
| epoch   4 |    800/ 2104 batches | lr 0.250 | ms/batch  7.22 | loss  0.27 |
| epoch   4 |   1000/ 2104 batches | lr 0.250 | ms/batch  7.05 | loss  0.27 |
| epoch   4 |   1200/ 2104 batches | lr 0.250 | ms/batch  7.05 | loss  0.28 |
| epoch   4 |   1400/ 2104 batches | lr 0.250 | ms/batch  6.98 | loss  0.25 |
| epoch   4 |   1600/ 2104 batches | lr 0.250 | ms/batch  6.99 | loss  0.25 |
| epoch   4 |   1800/ 2104 batches | lr 0.250 | ms/batch  7.04 | loss  0.25 |
| epoch   4 |   2000/ 2104 batches | lr 0.250 | ms/batch  7.01 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 15.05s | valid loss  0.65 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
| epoch   5 |    200/ 2104 batches | lr 0.062 | ms/batch  7.04 | loss  0.27 |
| epoch   5 |    400/ 2104 batches | lr 0.062 | ms/batch  7.05 | loss  0.25 |
| epoch   5 |    600/ 2104 batches | lr 0.062 | ms/batch  7.02 | loss  0.26 |
| epoch   5 |    800/ 2104 batches | lr 0.062 | ms/batch  7.02 | loss  0.25 |
| epoch   5 |   1000/ 2104 batches | lr 0.062 | ms/batch  7.00 | loss  0.24 |
| epoch   5 |   1200/ 2104 batches | lr 0.062 | ms/batch  6.97 | loss  0.25 |
| epoch   5 |   1400/ 2104 batches | lr 0.062 | ms/batch  6.99 | loss  0.23 |
| epoch   5 |   1600/ 2104 batches | lr 0.062 | ms/batch  7.02 | loss  0.22 |
| epoch   5 |   1800/ 2104 batches | lr 0.062 | ms/batch  7.02 | loss  0.22 |
| epoch   5 |   2000/ 2104 batches | lr 0.062 | ms/batch  7.35 | loss  0.21 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 14.91s | valid loss  0.66 | valid acc    0.819 |
-----------------------------------------------------------------------------------------
| epoch   6 |    200/ 2104 batches | lr 0.016 | ms/batch  7.33 | loss  0.26 |
| epoch   6 |    400/ 2104 batches | lr 0.016 | ms/batch  7.32 | loss  0.24 |
| epoch   6 |    600/ 2104 batches | lr 0.016 | ms/batch  7.26 | loss  0.25 |
| epoch   6 |    800/ 2104 batches | lr 0.016 | ms/batch  7.25 | loss  0.24 |
| epoch   6 |   1000/ 2104 batches | lr 0.016 | ms/batch  7.16 | loss  0.23 |
| epoch   6 |   1200/ 2104 batches | lr 0.016 | ms/batch  7.20 | loss  0.25 |
| epoch   6 |   1400/ 2104 batches | lr 0.016 | ms/batch  7.22 | loss  0.22 |
| epoch   6 |   1600/ 2104 batches | lr 0.016 | ms/batch  7.18 | loss  0.21 |
| epoch   6 |   1800/ 2104 batches | lr 0.016 | ms/batch  7.32 | loss  0.22 |
| epoch   6 |   2000/ 2104 batches | lr 0.016 | ms/batch  7.28 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time: 15.33s | valid loss  0.65 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
| epoch   7 |    200/ 2104 batches | lr 0.004 | ms/batch  7.33 | loss  0.26 |
| epoch   7 |    400/ 2104 batches | lr 0.004 | ms/batch  7.40 | loss  0.23 |
| epoch   7 |    600/ 2104 batches | lr 0.004 | ms/batch  7.32 | loss  0.25 |
| epoch   7 |    800/ 2104 batches | lr 0.004 | ms/batch  7.28 | loss  0.24 |
| epoch   7 |   1000/ 2104 batches | lr 0.004 | ms/batch  7.30 | loss  0.23 |
| epoch   7 |   1200/ 2104 batches | lr 0.004 | ms/batch  7.30 | loss  0.25 |
```

```
| epoch   7 |  1400/ 2104 batches | lr 0.004 | ms/batch  7.40 | loss  0.22 |
| epoch   7 |  1600/ 2104 batches | lr 0.004 | ms/batch  7.43 | loss  0.21 |
| epoch   7 |  1800/ 2104 batches | lr 0.004 | ms/batch  6.90 | loss  0.21 |
| epoch   7 |  2000/ 2104 batches | lr 0.004 | ms/batch  7.07 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time: 15.37s | valid loss  0.64 | valid acc    0.812 |
-----------------------------------------------------------------------------------------
| epoch   8 |   200/ 2104 batches | lr 0.001 | ms/batch  7.34 | loss  0.25 |
| epoch   8 |   400/ 2104 batches | lr 0.001 | ms/batch  7.28 | loss  0.23 |
| epoch   8 |   600/ 2104 batches | lr 0.001 | ms/batch  7.38 | loss  0.25 |
| epoch   8 |   800/ 2104 batches | lr 0.001 | ms/batch  7.45 | loss  0.23 |
| epoch   8 |  1000/ 2104 batches | lr 0.001 | ms/batch  7.11 | loss  0.23 |
| epoch   8 |  1200/ 2104 batches | lr 0.001 | ms/batch  7.17 | loss  0.24 |
| epoch   8 |  1400/ 2104 batches | lr 0.001 | ms/batch  7.25 | loss  0.22 |
| epoch   8 |  1600/ 2104 batches | lr 0.001 | ms/batch  7.47 | loss  0.21 |
| epoch   8 |  1800/ 2104 batches | lr 0.001 | ms/batch  7.38 | loss  0.21 |
| epoch   8 |  2000/ 2104 batches | lr 0.001 | ms/batch  7.31 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time: 15.45s | valid loss  0.65 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
| epoch   9 |   200/ 2104 batches | lr 0.000 | ms/batch  7.32 | loss  0.25 |
| epoch   9 |   400/ 2104 batches | lr 0.000 | ms/batch  7.30 | loss  0.23 |
| epoch   9 |   600/ 2104 batches | lr 0.000 | ms/batch  7.32 | loss  0.25 |
| epoch   9 |   800/ 2104 batches | lr 0.000 | ms/batch  7.26 | loss  0.23 |
| epoch   9 |  1000/ 2104 batches | lr 0.000 | ms/batch  7.25 | loss  0.22 |
| epoch   9 |  1200/ 2104 batches | lr 0.000 | ms/batch  7.29 | loss  0.24 |
| epoch   9 |  1400/ 2104 batches | lr 0.000 | ms/batch  7.12 | loss  0.22 |
| epoch   9 |  1600/ 2104 batches | lr 0.000 | ms/batch  7.05 | loss  0.21 |
| epoch   9 |  1800/ 2104 batches | lr 0.000 | ms/batch  7.18 | loss  0.21 |
| epoch   9 |  2000/ 2104 batches | lr 0.000 | ms/batch  7.10 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch   9 | time: 15.23s | valid loss  0.65 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
| epoch  10 |   200/ 2104 batches | lr 0.000 | ms/batch  7.01 | loss  0.25 |
| epoch  10 |   400/ 2104 batches | lr 0.000 | ms/batch  6.99 | loss  0.23 |
| epoch  10 |   600/ 2104 batches | lr 0.000 | ms/batch  7.05 | loss  0.25 |
| epoch  10 |   800/ 2104 batches | lr 0.000 | ms/batch  7.10 | loss  0.24 |
| epoch  10 |  1000/ 2104 batches | lr 0.000 | ms/batch  7.04 | loss  0.22 |
| epoch  10 |  1200/ 2104 batches | lr 0.000 | ms/batch  6.98 | loss  0.24 |
| epoch  10 |  1400/ 2104 batches | lr 0.000 | ms/batch  7.10 | loss  0.22 |
| epoch  10 |  1600/ 2104 batches | lr 0.000 | ms/batch  7.00 | loss  0.21 |
| epoch  10 |  1800/ 2104 batches | lr 0.000 | ms/batch  7.11 | loss  0.21 |
| epoch  10 |  2000/ 2104 batches | lr 0.000 | ms/batch  7.02 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch  10 | time: 14.87s | valid loss  0.65 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
=========================================================================================
==
| End of training | test loss  0.55 | test acc    0.850 |
=========================================================================================
==
```

```
--------------------------------------------------------------------------------------
learning_rate: 1, dropout: 0.2, batch_size: 64
--------------------------------------------------------------------------------------
| epoch   1 |    200/ 1052 batches | lr 1.000 | ms/batch 8.60 | loss  0.69 |
| epoch   1 |    400/ 1052 batches | lr 1.000 | ms/batch 7.26 | loss  0.69 |
| epoch   1 |    600/ 1052 batches | lr 1.000 | ms/batch 7.49 | loss  0.69 |
| epoch   1 |    800/ 1052 batches | lr 1.000 | ms/batch 7.49 | loss  0.69 |
| epoch   1 |   1000/ 1052 batches | lr 1.000 | ms/batch 7.49 | loss  0.69 |
--------------------------------------------------------------------------------------
| end of epoch   1 | time: 8.12s | valid loss 0.69 | valid acc   0.557 |
--------------------------------------------------------------------------------------
| epoch   2 |    200/ 1052 batches | lr 1.000 | ms/batch 7.57 | loss  0.67 |
| epoch   2 |    400/ 1052 batches | lr 1.000 | ms/batch 7.56 | loss  0.63 |
| epoch   2 |    600/ 1052 batches | lr 1.000 | ms/batch 7.60 | loss  0.58 |
| epoch   2 |    800/ 1052 batches | lr 1.000 | ms/batch 7.53 | loss  0.53 |
| epoch   2 |   1000/ 1052 batches | lr 1.000 | ms/batch 7.53 | loss  0.50 |
--------------------------------------------------------------------------------------
| end of epoch   2 | time: 8.01s | valid loss 0.57 | valid acc   0.750 |
--------------------------------------------------------------------------------------
| epoch   3 |    200/ 1052 batches | lr 1.000 | ms/batch 7.48 | loss  0.46 |
| epoch   3 |    400/ 1052 batches | lr 1.000 | ms/batch 7.22 | loss  0.44 |
| epoch   3 |    600/ 1052 batches | lr 1.000 | ms/batch 7.24 | loss  0.41 |
| epoch   3 |    800/ 1052 batches | lr 1.000 | ms/batch 7.31 | loss  0.39 |
| epoch   3 |   1000/ 1052 batches | lr 1.000 | ms/batch 7.22 | loss  0.37 |
--------------------------------------------------------------------------------------
| end of epoch   3 | time: 7.75s | valid loss 0.55 | valid acc   0.798 |
--------------------------------------------------------------------------------------
| epoch   4 |    200/ 1052 batches | lr 1.000 | ms/batch 7.52 | loss  0.36 |
| epoch   4 |    400/ 1052 batches | lr 1.000 | ms/batch 7.57 | loss  0.35 |
| epoch   4 |    600/ 1052 batches | lr 1.000 | ms/batch 7.47 | loss  0.34 |
| epoch   4 |    800/ 1052 batches | lr 1.000 | ms/batch 7.39 | loss  0.32 |
| epoch   4 |   1000/ 1052 batches | lr 1.000 | ms/batch 7.45 | loss  0.31 |
--------------------------------------------------------------------------------------
| end of epoch   4 | time: 7.92s | valid loss 0.60 | valid acc   0.819 |
--------------------------------------------------------------------------------------
| epoch   5 |    200/ 1052 batches | lr 0.250 | ms/batch 7.30 | loss  0.27 |
| epoch   5 |    400/ 1052 batches | lr 0.250 | ms/batch 7.45 | loss  0.27 |
| epoch   5 |    600/ 1052 batches | lr 0.250 | ms/batch 7.41 | loss  0.26 |
| epoch   5 |    800/ 1052 batches | lr 0.250 | ms/batch 7.41 | loss  0.24 |
| epoch   5 |   1000/ 1052 batches | lr 0.250 | ms/batch 7.42 | loss  0.23 |
--------------------------------------------------------------------------------------
| end of epoch   5 | time: 7.84s | valid loss 0.60 | valid acc   0.812 |
--------------------------------------------------------------------------------------
| epoch   6 |    200/ 1052 batches | lr 0.062 | ms/batch 7.47 | loss  0.25 |
| epoch   6 |    400/ 1052 batches | lr 0.062 | ms/batch 7.51 | loss  0.24 |
| epoch   6 |    600/ 1052 batches | lr 0.062 | ms/batch 7.47 | loss  0.23 |
| epoch   6 |    800/ 1052 batches | lr 0.062 | ms/batch 7.46 | loss  0.22 |
| epoch   6 |   1000/ 1052 batches | lr 0.062 | ms/batch 7.51 | loss  0.21 |
--------------------------------------------------------------------------------------
| end of epoch   6 | time: 7.92s | valid loss 0.62 | valid acc   0.812 |
--------------------------------------------------------------------------------------
```

```
| epoch   7 |   200/ 1052 batches | lr 0.016 | ms/batch 7.56 | loss  0.24 |
| epoch   7 |   400/ 1052 batches | lr 0.016 | ms/batch 7.40 | loss  0.23 |
| epoch   7 |   600/ 1052 batches | lr 0.016 | ms/batch 7.46 | loss  0.23 |
| epoch   7 |   800/ 1052 batches | lr 0.016 | ms/batch 7.45 | loss  0.21 |
| epoch   7 |  1000/ 1052 batches | lr 0.016 | ms/batch 7.57 | loss  0.21 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time: 7.94s | valid loss  0.62 | valid acc    0.821 |
-----------------------------------------------------------------------------------------
| epoch   8 |   200/ 1052 batches | lr 0.004 | ms/batch 7.47 | loss  0.24 |
| epoch   8 |   400/ 1052 batches | lr 0.004 | ms/batch 7.51 | loss  0.24 |
| epoch   8 |   600/ 1052 batches | lr 0.004 | ms/batch 7.48 | loss  0.22 |
| epoch   8 |   800/ 1052 batches | lr 0.004 | ms/batch 7.25 | loss  0.21 |
| epoch   8 |  1000/ 1052 batches | lr 0.004 | ms/batch 7.27 | loss  0.21 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time: 7.84s | valid loss  0.62 | valid acc    0.819 |
-----------------------------------------------------------------------------------------
| epoch   9 |   200/ 1052 batches | lr 0.001 | ms/batch 7.54 | loss  0.24 |
| epoch   9 |   400/ 1052 batches | lr 0.001 | ms/batch 7.51 | loss  0.23 |
| epoch   9 |   600/ 1052 batches | lr 0.001 | ms/batch 7.55 | loss  0.22 |
| epoch   9 |   800/ 1052 batches | lr 0.001 | ms/batch 7.52 | loss  0.21 |
| epoch   9 |  1000/ 1052 batches | lr 0.001 | ms/batch 7.35 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch   9 | time: 7.94s | valid loss  0.62 | valid acc    0.819 |
-----------------------------------------------------------------------------------------
| epoch  10 |   200/ 1052 batches | lr 0.000 | ms/batch 7.30 | loss  0.24 |
| epoch  10 |   400/ 1052 batches | lr 0.000 | ms/batch 7.27 | loss  0.23 |
| epoch  10 |   600/ 1052 batches | lr 0.000 | ms/batch 7.31 | loss  0.23 |
| epoch  10 |   800/ 1052 batches | lr 0.000 | ms/batch 7.33 | loss  0.21 |
| epoch  10 |  1000/ 1052 batches | lr 0.000 | ms/batch 7.51 | loss  0.20 |
-----------------------------------------------------------------------------------------
| end of epoch  10 | time: 7.79s | valid loss  0.62 | valid acc    0.817 |
-----------------------------------------------------------------------------------------
=======================================================================================
==
| End of training | test loss  0.51 | test acc    0.810 |
=======================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.1, batch_size: 32
-----------------------------------------------------------------------------------------
| epoch   1 |   200/ 2104 batches | lr 10.000 | ms/batch 8.01 | loss  1.06 |
| epoch   1 |   400/ 2104 batches | lr 10.000 | ms/batch 7.38 | loss  1.04 |
| epoch   1 |   600/ 2104 batches | lr 10.000 | ms/batch 7.31 | loss  1.01 |
| epoch   1 |   800/ 2104 batches | lr 10.000 | ms/batch 7.40 | loss  1.02 |
| epoch   1 |  1000/ 2104 batches | lr 10.000 | ms/batch 7.32 | loss  1.02 |
| epoch   1 |  1200/ 2104 batches | lr 10.000 | ms/batch 7.29 | loss  1.03 |
| epoch   1 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.27 | loss  1.04 |
| epoch   1 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.42 | loss  1.01 |
| epoch   1 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.39 | loss  1.03 |
| epoch   1 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.43 | loss  1.05 |
-----------------------------------------------------------------------------------------
```

```
| end of epoch   1 | time: 15.65s | valid loss  1.10 | valid acc    0.483 |
-----------------------------------------------------------------------------
| epoch   2 |    200/ 2104 batches | lr 10.000 | ms/batch 7.37 | loss  1.04 |
| epoch   2 |    400/ 2104 batches | lr 10.000 | ms/batch 7.40 | loss  1.04 |
| epoch   2 |    600/ 2104 batches | lr 10.000 | ms/batch 7.36 | loss  1.01 |
| epoch   2 |    800/ 2104 batches | lr 10.000 | ms/batch 7.36 | loss  1.02 |
| epoch   2 |   1000/ 2104 batches | lr 10.000 | ms/batch 7.34 | loss  1.02 |
| epoch   2 |   1200/ 2104 batches | lr 10.000 | ms/batch 7.20 | loss  1.02 |
| epoch   2 |   1400/ 2104 batches | lr 10.000 | ms/batch 7.05 | loss  1.04 |
| epoch   2 |   1600/ 2104 batches | lr 10.000 | ms/batch 7.39 | loss  1.01 |
| epoch   2 |   1800/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.03 |
| epoch   2 |   2000/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.04 |
-----------------------------------------------------------------------------
| end of epoch   2 | time: 15.45s | valid loss  1.11 | valid acc    0.483 |
-----------------------------------------------------------------------------
| epoch   3 |    200/ 2104 batches | lr 2.500 | ms/batch 7.32 | loss  0.71 |
| epoch   3 |    400/ 2104 batches | lr 2.500 | ms/batch 7.01 | loss  0.70 |
| epoch   3 |    600/ 2104 batches | lr 2.500 | ms/batch 7.08 | loss  0.71 |
| epoch   3 |    800/ 2104 batches | lr 2.500 | ms/batch 7.10 | loss  0.70 |
| epoch   3 |   1000/ 2104 batches | lr 2.500 | ms/batch 7.13 | loss  0.70 |
| epoch   3 |   1200/ 2104 batches | lr 2.500 | ms/batch 7.05 | loss  0.70 |
| epoch   3 |   1400/ 2104 batches | lr 2.500 | ms/batch 7.05 | loss  0.70 |
| epoch   3 |   1600/ 2104 batches | lr 2.500 | ms/batch 7.17 | loss  0.70 |
| epoch   3 |   1800/ 2104 batches | lr 2.500 | ms/batch 7.29 | loss  0.70 |
| epoch   3 |   2000/ 2104 batches | lr 2.500 | ms/batch 7.37 | loss  0.71 |
-----------------------------------------------------------------------------
| end of epoch   3 | time: 15.14s | valid loss  0.69 | valid acc    0.605 |
-----------------------------------------------------------------------------
| epoch   4 |    200/ 2104 batches | lr 2.500 | ms/batch 7.32 | loss  0.67 |
| epoch   4 |    400/ 2104 batches | lr 2.500 | ms/batch 7.27 | loss  0.60 |
| epoch   4 |    600/ 2104 batches | lr 2.500 | ms/batch 7.26 | loss  0.55 |
| epoch   4 |    800/ 2104 batches | lr 2.500 | ms/batch 7.33 | loss  0.51 |
| epoch   4 |   1000/ 2104 batches | lr 2.500 | ms/batch 7.38 | loss  0.47 |
| epoch   4 |   1200/ 2104 batches | lr 2.500 | ms/batch 7.41 | loss  0.45 |
| epoch   4 |   1400/ 2104 batches | lr 2.500 | ms/batch 7.37 | loss  0.41 |
| epoch   4 |   1600/ 2104 batches | lr 2.500 | ms/batch 7.42 | loss  0.38 |
| epoch   4 |   1800/ 2104 batches | lr 2.500 | ms/batch 7.37 | loss  0.39 |
| epoch   4 |   2000/ 2104 batches | lr 2.500 | ms/batch 7.42 | loss  0.35 |
-----------------------------------------------------------------------------
| end of epoch   4 | time: 15.45s | valid loss  0.84 | valid acc    0.776 |
-----------------------------------------------------------------------------
| epoch   5 |    200/ 2104 batches | lr 0.625 | ms/batch 6.02 | loss  0.31 |
| epoch   5 |    400/ 2104 batches | lr 0.625 | ms/batch 7.32 | loss  0.28 |
| epoch   5 |    600/ 2104 batches | lr 0.625 | ms/batch 7.41 | loss  0.28 |
| epoch   5 |    800/ 2104 batches | lr 0.625 | ms/batch 7.42 | loss  0.26 |
| epoch   5 |   1000/ 2104 batches | lr 0.625 | ms/batch 7.38 | loss  0.25 |
| epoch   5 |   1200/ 2104 batches | lr 0.625 | ms/batch 7.11 | loss  0.26 |
| epoch   5 |   1400/ 2104 batches | lr 0.625 | ms/batch 6.34 | loss  0.23 |
| epoch   5 |   1600/ 2104 batches | lr 0.625 | ms/batch 7.22 | loss  0.22 |
| epoch   5 |   1800/ 2104 batches | lr 0.625 | ms/batch 7.38 | loss  0.23 |
| epoch   5 |   2000/ 2104 batches | lr 0.625 | ms/batch 7.38 | loss  0.20 |
```

```
-----------------------------------------------------------------------------------
| end of epoch   5 | time: 15.03s | valid loss  0.58 | valid acc    0.812 |
-----------------------------------------------------------------------------------
| epoch   6 |    200/ 2104 batches | lr 0.625 | ms/batch 7.49 | loss  0.25 |
| epoch   6 |    400/ 2104 batches | lr 0.625 | ms/batch 7.43 | loss  0.23 |
| epoch   6 |    600/ 2104 batches | lr 0.625 | ms/batch 7.35 | loss  0.23 |
| epoch   6 |    800/ 2104 batches | lr 0.625 | ms/batch 7.36 | loss  0.23 |
| epoch   6 |   1000/ 2104 batches | lr 0.625 | ms/batch 7.20 | loss  0.21 |
| epoch   6 |   1200/ 2104 batches | lr 0.625 | ms/batch 6.55 | loss  0.22 |
| epoch   6 |   1400/ 2104 batches | lr 0.625 | ms/batch 6.68 | loss  0.20 |
| epoch   6 |   1600/ 2104 batches | lr 0.625 | ms/batch 7.12 | loss  0.19 |
| epoch   6 |   1800/ 2104 batches | lr 0.625 | ms/batch 7.13 | loss  0.20 |
| epoch   6 |   2000/ 2104 batches | lr 0.625 | ms/batch 7.11 | loss  0.19 |
-----------------------------------------------------------------------------------
| end of epoch   6 | time: 15.09s | valid loss  0.57 | valid acc    0.810 |
-----------------------------------------------------------------------------------
| epoch   7 |    200/ 2104 batches | lr 0.625 | ms/batch 7.20 | loss  0.22 |
| epoch   7 |    400/ 2104 batches | lr 0.625 | ms/batch 7.24 | loss  0.21 |
| epoch   7 |    600/ 2104 batches | lr 0.625 | ms/batch 7.08 | loss  0.20 |
| epoch   7 |    800/ 2104 batches | lr 0.625 | ms/batch 7.07 | loss  0.20 |
| epoch   7 |   1000/ 2104 batches | lr 0.625 | ms/batch 7.31 | loss  0.18 |
| epoch   7 |   1200/ 2104 batches | lr 0.625 | ms/batch 7.25 | loss  0.20 |
| epoch   7 |   1400/ 2104 batches | lr 0.625 | ms/batch 7.15 | loss  0.18 |
| epoch   7 |   1600/ 2104 batches | lr 0.625 | ms/batch 7.32 | loss  0.17 |
| epoch   7 |   1800/ 2104 batches | lr 0.625 | ms/batch 7.18 | loss  0.19 |
| epoch   7 |   2000/ 2104 batches | lr 0.625 | ms/batch 7.18 | loss  0.17 |
-----------------------------------------------------------------------------------
| end of epoch   7 | time: 15.20s | valid loss  0.58 | valid acc    0.826 |
-----------------------------------------------------------------------------------
| epoch   8 |    200/ 2104 batches | lr 0.156 | ms/batch 7.46 | loss  0.20 |
| epoch   8 |    400/ 2104 batches | lr 0.156 | ms/batch 7.40 | loss  0.18 |
| epoch   8 |    600/ 2104 batches | lr 0.156 | ms/batch 7.21 | loss  0.18 |
| epoch   8 |    800/ 2104 batches | lr 0.156 | ms/batch 7.06 | loss  0.18 |
| epoch   8 |   1000/ 2104 batches | lr 0.156 | ms/batch 7.17 | loss  0.15 |
| epoch   8 |   1200/ 2104 batches | lr 0.156 | ms/batch 7.11 | loss  0.17 |
| epoch   8 |   1400/ 2104 batches | lr 0.156 | ms/batch 7.11 | loss  0.14 |
| epoch   8 |   1600/ 2104 batches | lr 0.156 | ms/batch 7.06 | loss  0.14 |
| epoch   8 |   1800/ 2104 batches | lr 0.156 | ms/batch 7.05 | loss  0.15 |
| epoch   8 |   2000/ 2104 batches | lr 0.156 | ms/batch 7.24 | loss  0.14 |
-----------------------------------------------------------------------------------
| end of epoch   8 | time: 15.18s | valid loss  0.58 | valid acc    0.845 |
-----------------------------------------------------------------------------------
| epoch   9 |    200/ 2104 batches | lr 0.039 | ms/batch 7.22 | loss  0.19 |
| epoch   9 |    400/ 2104 batches | lr 0.039 | ms/batch 7.20 | loss  0.17 |
| epoch   9 |    600/ 2104 batches | lr 0.039 | ms/batch 7.27 | loss  0.17 |
| epoch   9 |    800/ 2104 batches | lr 0.039 | ms/batch 7.27 | loss  0.17 |
| epoch   9 |   1000/ 2104 batches | lr 0.039 | ms/batch 7.03 | loss  0.14 |
| epoch   9 |   1200/ 2104 batches | lr 0.039 | ms/batch 7.27 | loss  0.16 |
| epoch   9 |   1400/ 2104 batches | lr 0.039 | ms/batch 7.16 | loss  0.14 |
| epoch   9 |   1600/ 2104 batches | lr 0.039 | ms/batch 7.22 | loss  0.13 |
| epoch   9 |   1800/ 2104 batches | lr 0.039 | ms/batch 7.18 | loss  0.14 |
```

```
| epoch   9 |  2000/ 2104 batches | lr 0.039 | ms/batch  7.18 | loss  0.13 |
-----------------------------------------------------------------------------
| end of epoch   9 | time: 15.21s | valid loss  0.58 | valid acc    0.845 |
-----------------------------------------------------------------------------
| epoch  10 |   200/ 2104 batches | lr 0.010 | ms/batch  7.30 | loss  0.19 |
| epoch  10 |   400/ 2104 batches | lr 0.010 | ms/batch  7.08 | loss  0.17 |
| epoch  10 |   600/ 2104 batches | lr 0.010 | ms/batch  7.37 | loss  0.17 |
| epoch  10 |   800/ 2104 batches | lr 0.010 | ms/batch  7.16 | loss  0.17 |
| epoch  10 |  1000/ 2104 batches | lr 0.010 | ms/batch  7.17 | loss  0.14 |
| epoch  10 |  1200/ 2104 batches | lr 0.010 | ms/batch  7.06 | loss  0.16 |
| epoch  10 |  1400/ 2104 batches | lr 0.010 | ms/batch  7.26 | loss  0.13 |
| epoch  10 |  1600/ 2104 batches | lr 0.010 | ms/batch  7.23 | loss  0.12 |
| epoch  10 |  1800/ 2104 batches | lr 0.010 | ms/batch  7.09 | loss  0.14 |
| epoch  10 |  2000/ 2104 batches | lr 0.010 | ms/batch  7.09 | loss  0.13 |
-----------------------------------------------------------------------------
| end of epoch  10 | time: 15.17s | valid loss  0.57 | valid acc    0.845 |
-----------------------------------------------------------------------------
=============================================================================
==
| End of training | test loss  0.50 | test acc    0.850 |
=============================================================================
==
-----------------------------------------------------------------------------
learning_rate: 10, dropout: 0.1, batch_size: 64
-----------------------------------------------------------------------------
| epoch   1 |   200/ 1052 batches | lr 10.000 | ms/batch  8.75 | loss  1.07 |
| epoch   1 |   400/ 1052 batches | lr 10.000 | ms/batch  7.21 | loss  1.02 |
| epoch   1 |   600/ 1052 batches | lr 10.000 | ms/batch  7.51 | loss  1.03 |
| epoch   1 |   800/ 1052 batches | lr 10.000 | ms/batch  7.55 | loss  1.03 |
| epoch   1 |  1000/ 1052 batches | lr 10.000 | ms/batch  7.64 | loss  1.03 |
-----------------------------------------------------------------------------
| end of epoch   1 | time: 8.17s | valid loss  1.17 | valid acc    0.540 |
-----------------------------------------------------------------------------
| epoch   2 |   200/ 1052 batches | lr 10.000 | ms/batch  7.66 | loss  1.05 |
| epoch   2 |   400/ 1052 batches | lr 10.000 | ms/batch  7.64 | loss  1.03 |
| epoch   2 |   600/ 1052 batches | lr 10.000 | ms/batch  7.60 | loss  1.04 |
| epoch   2 |   800/ 1052 batches | lr 10.000 | ms/batch  7.54 | loss  1.03 |
| epoch   2 |  1000/ 1052 batches | lr 10.000 | ms/batch  7.58 | loss  1.03 |
-----------------------------------------------------------------------------
| end of epoch   2 | time: 8.05s | valid loss  1.29 | valid acc    0.540 |
-----------------------------------------------------------------------------
| epoch   3 |   200/ 1052 batches | lr 2.500 | ms/batch  7.61 | loss  0.71 |
| epoch   3 |   400/ 1052 batches | lr 2.500 | ms/batch  7.61 | loss  0.70 |
| epoch   3 |   600/ 1052 batches | lr 2.500 | ms/batch  7.56 | loss  0.69 |
| epoch   3 |   800/ 1052 batches | lr 2.500 | ms/batch  7.58 | loss  0.69 |
| epoch   3 |  1000/ 1052 batches | lr 2.500 | ms/batch  7.55 | loss  0.70 |
-----------------------------------------------------------------------------
| end of epoch   3 | time: 8.04s | valid loss  0.73 | valid acc    0.540 |
-----------------------------------------------------------------------------
| epoch   4 |   200/ 1052 batches | lr 2.500 | ms/batch  7.69 | loss  0.70 |
| epoch   4 |   400/ 1052 batches | lr 2.500 | ms/batch  7.53 | loss  0.70 |
```

```
| epoch   4 |   600/ 1052 batches | lr 2.500 | ms/batch  7.55 | loss  0.69 |
| epoch   4 |   800/ 1052 batches | lr 2.500 | ms/batch  7.52 | loss  0.68 |
| epoch   4 |  1000/ 1052 batches | lr 2.500 | ms/batch  7.53 | loss  0.63 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 8.01s | valid loss 0.83 | valid acc    0.650 |
-----------------------------------------------------------------------------------------
| epoch   5 |   200/ 1052 batches | lr 0.625 | ms/batch  6.44 | loss  0.47 |
| epoch   5 |   400/ 1052 batches | lr 0.625 | ms/batch  6.66 | loss  0.44 |
| epoch   5 |   600/ 1052 batches | lr 0.625 | ms/batch  7.48 | loss  0.41 |
| epoch   5 |   800/ 1052 batches | lr 0.625 | ms/batch  7.41 | loss  0.38 |
| epoch   5 |  1000/ 1052 batches | lr 0.625 | ms/batch  7.48 | loss  0.34 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 7.56s | valid loss 0.56 | valid acc    0.788 |
-----------------------------------------------------------------------------------------
| epoch   6 |   200/ 1052 batches | lr 0.625 | ms/batch  7.55 | loss  0.35 |
| epoch   6 |   400/ 1052 batches | lr 0.625 | ms/batch  7.53 | loss  0.35 |
| epoch   6 |   600/ 1052 batches | lr 0.625 | ms/batch  7.51 | loss  0.33 |
| epoch   6 |   800/ 1052 batches | lr 0.625 | ms/batch  7.56 | loss  0.31 |
| epoch   6 |  1000/ 1052 batches | lr 0.625 | ms/batch  6.53 | loss  0.29 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time: 7.79s | valid loss 0.56 | valid acc    0.798 |
-----------------------------------------------------------------------------------------
| epoch   7 |   200/ 1052 batches | lr 0.156 | ms/batch  7.44 | loss  0.28 |
| epoch   7 |   400/ 1052 batches | lr 0.156 | ms/batch  5.87 | loss  0.28 |
| epoch   7 |   600/ 1052 batches | lr 0.156 | ms/batch  6.97 | loss  0.26 |
| epoch   7 |   800/ 1052 batches | lr 0.156 | ms/batch  7.50 | loss  0.25 |
| epoch   7 |  1000/ 1052 batches | lr 0.156 | ms/batch  7.58 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time: 7.53s | valid loss 0.59 | valid acc    0.819 |
-----------------------------------------------------------------------------------------
| epoch   8 |   200/ 1052 batches | lr 0.039 | ms/batch  7.60 | loss  0.26 |
| epoch   8 |   400/ 1052 batches | lr 0.039 | ms/batch  7.61 | loss  0.26 |
| epoch   8 |   600/ 1052 batches | lr 0.039 | ms/batch  7.49 | loss  0.25 |
| epoch   8 |   800/ 1052 batches | lr 0.039 | ms/batch  7.56 | loss  0.23 |
| epoch   8 |  1000/ 1052 batches | lr 0.039 | ms/batch  7.62 | loss  0.22 |
-----------------------------------------------------------------------------------------
| end of epoch   8 | time: 8.03s | valid loss 0.59 | valid acc    0.812 |
-----------------------------------------------------------------------------------------
| epoch   9 |   200/ 1052 batches | lr 0.010 | ms/batch  7.61 | loss  0.25 |
| epoch   9 |   400/ 1052 batches | lr 0.010 | ms/batch  7.66 | loss  0.25 |
| epoch   9 |   600/ 1052 batches | lr 0.010 | ms/batch  7.68 | loss  0.24 |
| epoch   9 |   800/ 1052 batches | lr 0.010 | ms/batch  7.58 | loss  0.22 |
| epoch   9 |  1000/ 1052 batches | lr 0.010 | ms/batch  7.62 | loss  0.21 |
-----------------------------------------------------------------------------------------
| end of epoch   9 | time: 8.09s | valid loss 0.59 | valid acc    0.812 |
-----------------------------------------------------------------------------------------
| epoch  10 |   200/ 1052 batches | lr 0.002 | ms/batch  7.63 | loss  0.25 |
| epoch  10 |   400/ 1052 batches | lr 0.002 | ms/batch  7.30 | loss  0.25 |
| epoch  10 |   600/ 1052 batches | lr 0.002 | ms/batch  7.31 | loss  0.24 |
| epoch  10 |   800/ 1052 batches | lr 0.002 | ms/batch  7.31 | loss  0.22 |
| epoch  10 |  1000/ 1052 batches | lr 0.002 | ms/batch  7.44 | loss  0.21 |
```

```
--------------------------------------------------------------------------------
| end of epoch  10 | time:  7.86s | valid loss  0.59 | valid acc    0.812 |
--------------------------------------------------------------------------------
================================================================================
==
| End of training | test loss  0.53 | test acc    0.821 |
================================================================================
==
--------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.2, batch_size: 32
--------------------------------------------------------------------------------
| epoch   1 |   200/ 2104 batches | lr 10.000 | ms/batch 8.23 | loss  1.05 |
| epoch   1 |   400/ 2104 batches | lr 10.000 | ms/batch 7.06 | loss  1.03 |
| epoch   1 |   600/ 2104 batches | lr 10.000 | ms/batch 7.07 | loss  1.01 |
| epoch   1 |   800/ 2104 batches | lr 10.000 | ms/batch 7.23 | loss  1.02 |
| epoch   1 |  1000/ 2104 batches | lr 10.000 | ms/batch 5.73 | loss  1.02 |
| epoch   1 |  1200/ 2104 batches | lr 10.000 | ms/batch 6.51 | loss  1.03 |
| epoch   1 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.02 |
| epoch   1 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.24 | loss  1.02 |
| epoch   1 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.21 | loss  1.03 |
| epoch   1 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.04 |
--------------------------------------------------------------------------------
| end of epoch   1 | time: 15.00s | valid loss  1.13 | valid acc    0.483 |
--------------------------------------------------------------------------------
| epoch   2 |   200/ 2104 batches | lr 10.000 | ms/batch 5.76 | loss  1.02 |
| epoch   2 |   400/ 2104 batches | lr 10.000 | ms/batch 6.82 | loss  1.01 |
| epoch   2 |   600/ 2104 batches | lr 10.000 | ms/batch 7.29 | loss  1.02 |
| epoch   2 |   800/ 2104 batches | lr 10.000 | ms/batch 7.27 | loss  1.02 |
| epoch   2 |  1000/ 2104 batches | lr 10.000 | ms/batch 7.22 | loss  1.02 |
| epoch   2 |  1200/ 2104 batches | lr 10.000 | ms/batch 7.25 | loss  1.02 |
| epoch   2 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.02 |
| epoch   2 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.45 | loss  1.02 |
| epoch   2 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.03 |
| epoch   2 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.38 | loss  1.04 |
--------------------------------------------------------------------------------
| end of epoch   2 | time: 14.98s | valid loss  1.12 | valid acc    0.483 |
--------------------------------------------------------------------------------
| epoch   3 |   200/ 2104 batches | lr 10.000 | ms/batch 7.08 | loss  1.01 |
| epoch   3 |   400/ 2104 batches | lr 10.000 | ms/batch 7.11 | loss  1.03 |
| epoch   3 |   600/ 2104 batches | lr 10.000 | ms/batch 7.18 | loss  1.02 |
| epoch   3 |   800/ 2104 batches | lr 10.000 | ms/batch 7.22 | loss  1.02 |
| epoch   3 |  1000/ 2104 batches | lr 10.000 | ms/batch 7.20 | loss  1.02 |
| epoch   3 |  1200/ 2104 batches | lr 10.000 | ms/batch 7.20 | loss  1.03 |
| epoch   3 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.09 | loss  1.02 |
| epoch   3 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.16 | loss  1.01 |
| epoch   3 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.12 | loss  1.03 |
| epoch   3 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.15 | loss  1.04 |
--------------------------------------------------------------------------------
| end of epoch   3 | time: 15.11s | valid loss  1.09 | valid acc    0.483 |
--------------------------------------------------------------------------------
| epoch   4 |   200/ 2104 batches | lr 10.000 | ms/batch 7.29 | loss  1.04 |
```

```
| epoch   4 |   400/ 2104 batches | lr 10.000 | ms/batch 7.25 | loss  1.02 |
| epoch   4 |   600/ 2104 batches | lr 10.000 | ms/batch 7.19 | loss  1.03 |
| epoch   4 |   800/ 2104 batches | lr 10.000 | ms/batch 7.12 | loss  1.02 |
| epoch   4 |  1000/ 2104 batches | lr 10.000 | ms/batch 7.27 | loss  1.00 |
| epoch   4 |  1200/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.03 |
| epoch   4 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.38 | loss  1.02 |
| epoch   4 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.35 | loss  1.02 |
| epoch   4 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.35 | loss  1.03 |
| epoch   4 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.25 | loss  1.04 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time: 15.36s | valid loss  0.99 | valid acc    0.483 |
-----------------------------------------------------------------------------------------
| epoch   5 |   200/ 2104 batches | lr 10.000 | ms/batch 7.17 | loss  1.01 |
| epoch   5 |   400/ 2104 batches | lr 10.000 | ms/batch 7.06 | loss  1.02 |
| epoch   5 |   600/ 2104 batches | lr 10.000 | ms/batch 7.32 | loss  1.04 |
| epoch   5 |   800/ 2104 batches | lr 10.000 | ms/batch 7.35 | loss  1.02 |
| epoch   5 |  1000/ 2104 batches | lr 10.000 | ms/batch 7.47 | loss  1.02 |
| epoch   5 |  1200/ 2104 batches | lr 10.000 | ms/batch 7.41 | loss  1.02 |
| epoch   5 |  1400/ 2104 batches | lr 10.000 | ms/batch 7.29 | loss  1.02 |
| epoch   5 |  1600/ 2104 batches | lr 10.000 | ms/batch 7.33 | loss  1.01 |
| epoch   5 |  1800/ 2104 batches | lr 10.000 | ms/batch 7.32 | loss  1.04 |
| epoch   5 |  2000/ 2104 batches | lr 10.000 | ms/batch 7.37 | loss  1.05 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time: 15.44s | valid loss  0.99 | valid acc    0.483 |
-----------------------------------------------------------------------------------------
| epoch   6 |   200/ 2104 batches | lr 2.500 | ms/batch 6.70 | loss  0.71 |
| epoch   6 |   400/ 2104 batches | lr 2.500 | ms/batch 7.27 | loss  0.70 |
| epoch   6 |   600/ 2104 batches | lr 2.500 | ms/batch 7.31 | loss  0.71 |
| epoch   6 |   800/ 2104 batches | lr 2.500 | ms/batch 7.10 | loss  0.70 |
| epoch   6 |  1000/ 2104 batches | lr 2.500 | ms/batch 7.14 | loss  0.70 |
| epoch   6 |  1200/ 2104 batches | lr 2.500 | ms/batch 7.05 | loss  0.70 |
| epoch   6 |  1400/ 2104 batches | lr 2.500 | ms/batch 7.05 | loss  0.70 |
| epoch   6 |  1600/ 2104 batches | lr 2.500 | ms/batch 7.08 | loss  0.70 |
| epoch   6 |  1800/ 2104 batches | lr 2.500 | ms/batch 7.07 | loss  0.68 |
| epoch   6 |  2000/ 2104 batches | lr 2.500 | ms/batch 7.04 | loss  0.63 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time: 14.96s | valid loss  0.67 | valid acc    0.683 |
-----------------------------------------------------------------------------------------
| epoch   7 |   200/ 2104 batches | lr 2.500 | ms/batch 6.81 | loss  0.54 |
| epoch   7 |   400/ 2104 batches | lr 2.500 | ms/batch 7.35 | loss  0.49 |
| epoch   7 |   600/ 2104 batches | lr 2.500 | ms/batch 7.33 | loss  0.48 |
| epoch   7 |   800/ 2104 batches | lr 2.500 | ms/batch 7.09 | loss  0.45 |
| epoch   7 |  1000/ 2104 batches | lr 2.500 | ms/batch 7.19 | loss  0.43 |
| epoch   7 |  1200/ 2104 batches | lr 2.500 | ms/batch 7.41 | loss  0.41 |
| epoch   7 |  1400/ 2104 batches | lr 2.500 | ms/batch 7.28 | loss  0.38 |
| epoch   7 |  1600/ 2104 batches | lr 2.500 | ms/batch 7.33 | loss  0.36 |
| epoch   7 |  1800/ 2104 batches | lr 2.500 | ms/batch 7.33 | loss  0.35 |
| epoch   7 |  2000/ 2104 batches | lr 2.500 | ms/batch 7.33 | loss  0.32 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time: 15.30s | valid loss  0.67 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
```

```
| epoch  8 |   200/ 2104 batches | lr 0.625 | ms/batch 7.16 | loss  0.29 |
| epoch  8 |   400/ 2104 batches | lr 0.625 | ms/batch 7.13 | loss  0.27 |
| epoch  8 |   600/ 2104 batches | lr 0.625 | ms/batch 7.35 | loss  0.27 |
| epoch  8 |   800/ 2104 batches | lr 0.625 | ms/batch 7.12 | loss  0.25 |
| epoch  8 |  1000/ 2104 batches | lr 0.625 | ms/batch 7.03 | loss  0.24 |
| epoch  8 |  1200/ 2104 batches | lr 0.625 | ms/batch 7.13 | loss  0.25 |
| epoch  8 |  1400/ 2104 batches | lr 0.625 | ms/batch 7.23 | loss  0.22 |
| epoch  8 |  1600/ 2104 batches | lr 0.625 | ms/batch 7.36 | loss  0.21 |
| epoch  8 |  1800/ 2104 batches | lr 0.625 | ms/batch 7.34 | loss  0.21 |
| epoch  8 |  2000/ 2104 batches | lr 0.625 | ms/batch 7.28 | loss  0.19 |
-----------------------------------------------------------------------------------------
| end of epoch  8 | time: 15.26s | valid loss 0.65 | valid acc   0.805 |
-----------------------------------------------------------------------------------------
| epoch  9 |   200/ 2104 batches | lr 0.625 | ms/batch 6.68 | loss  0.25 |
| epoch  9 |   400/ 2104 batches | lr 0.625 | ms/batch 7.34 | loss  0.23 |
| epoch  9 |   600/ 2104 batches | lr 0.625 | ms/batch 7.32 | loss  0.22 |
| epoch  9 |   800/ 2104 batches | lr 0.625 | ms/batch 6.71 | loss  0.22 |
| epoch  9 |  1000/ 2104 batches | lr 0.625 | ms/batch 7.19 | loss  0.20 |
| epoch  9 |  1200/ 2104 batches | lr 0.625 | ms/batch 7.11 | loss  0.22 |
| epoch  9 |  1400/ 2104 batches | lr 0.625 | ms/batch 7.33 | loss  0.19 |
| epoch  9 |  1600/ 2104 batches | lr 0.625 | ms/batch 7.29 | loss  0.19 |
| epoch  9 |  1800/ 2104 batches | lr 0.625 | ms/batch 7.13 | loss  0.19 |
| epoch  9 |  2000/ 2104 batches | lr 0.625 | ms/batch 7.21 | loss  0.17 |
-----------------------------------------------------------------------------------------
| end of epoch  9 | time: 15.07s | valid loss 0.61 | valid acc   0.821 |
-----------------------------------------------------------------------------------------
| epoch 10 |   200/ 2104 batches | lr 0.625 | ms/batch 7.26 | loss  0.22 |
| epoch 10 |   400/ 2104 batches | lr 0.625 | ms/batch 7.18 | loss  0.20 |
| epoch 10 |   600/ 2104 batches | lr 0.625 | ms/batch 7.17 | loss  0.20 |
| epoch 10 |   800/ 2104 batches | lr 0.625 | ms/batch 7.19 | loss  0.20 |
| epoch 10 |  1000/ 2104 batches | lr 0.625 | ms/batch 7.15 | loss  0.18 |
| epoch 10 |  1200/ 2104 batches | lr 0.625 | ms/batch 7.16 | loss  0.20 |
| epoch 10 |  1400/ 2104 batches | lr 0.625 | ms/batch 7.22 | loss  0.17 |
| epoch 10 |  1600/ 2104 batches | lr 0.625 | ms/batch 7.22 | loss  0.17 |
| epoch 10 |  1800/ 2104 batches | lr 0.625 | ms/batch 7.31 | loss  0.18 |
| epoch 10 |  2000/ 2104 batches | lr 0.625 | ms/batch 7.31 | loss  0.17 |
-----------------------------------------------------------------------------------------
| end of epoch 10 | time: 15.25s | valid loss 0.58 | valid acc   0.826 |
-----------------------------------------------------------------------------------------
===========================================================================================
==
| End of training | test loss 0.50 | test acc   0.819 |
===========================================================================================
==
-----------------------------------------------------------------------------------------
learning_rate: 10, dropout: 0.2, batch_size: 64
-----------------------------------------------------------------------------------------
| epoch  1 |   200/ 1052 batches | lr 10.000 | ms/batch 8.82 | loss  1.06 |
| epoch  1 |   400/ 1052 batches | lr 10.000 | ms/batch 7.40 | loss  1.02 |
| epoch  1 |   600/ 1052 batches | lr 10.000 | ms/batch 7.40 | loss  1.04 |
| epoch  1 |   800/ 1052 batches | lr 10.000 | ms/batch 7.17 | loss  1.03 |
```

```
| epoch   1 |   1000/ 1052 batches | lr 10.000 | ms/batch  7.28 | loss  1.03 |
-----------------------------------------------------------------------------------------
| end of epoch   1 | time:  8.05s | valid loss  1.18 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   2 |    200/ 1052 batches | lr 10.000 | ms/batch  7.45 | loss  1.05 |
| epoch   2 |    400/ 1052 batches | lr 10.000 | ms/batch  7.24 | loss  1.03 |
| epoch   2 |    600/ 1052 batches | lr 10.000 | ms/batch  7.23 | loss  1.04 |
| epoch   2 |    800/ 1052 batches | lr 10.000 | ms/batch  7.23 | loss  1.03 |
| epoch   2 |   1000/ 1052 batches | lr 10.000 | ms/batch  7.35 | loss  1.03 |
-----------------------------------------------------------------------------------------
| end of epoch   2 | time:  7.73s | valid loss  1.24 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   3 |    200/ 1052 batches | lr 2.500 | ms/batch  7.36 | loss  0.71 |
| epoch   3 |    400/ 1052 batches | lr 2.500 | ms/batch  7.52 | loss  0.70 |
| epoch   3 |    600/ 1052 batches | lr 2.500 | ms/batch  7.54 | loss  0.69 |
| epoch   3 |    800/ 1052 batches | lr 2.500 | ms/batch  7.41 | loss  0.69 |
| epoch   3 |   1000/ 1052 batches | lr 2.500 | ms/batch  7.46 | loss  0.70 |
-----------------------------------------------------------------------------------------
| end of epoch   3 | time:  7.90s | valid loss  0.73 | valid acc    0.540 |
-----------------------------------------------------------------------------------------
| epoch   4 |    200/ 1052 batches | lr 2.500 | ms/batch  7.49 | loss  0.70 |
| epoch   4 |    400/ 1052 batches | lr 2.500 | ms/batch  7.57 | loss  0.70 |
| epoch   4 |    600/ 1052 batches | lr 2.500 | ms/batch  7.51 | loss  0.69 |
| epoch   4 |    800/ 1052 batches | lr 2.500 | ms/batch  7.50 | loss  0.68 |
| epoch   4 |   1000/ 1052 batches | lr 2.500 | ms/batch  6.81 | loss  0.62 |
-----------------------------------------------------------------------------------------
| end of epoch   4 | time:  7.78s | valid loss  0.63 | valid acc    0.693 |
-----------------------------------------------------------------------------------------
| epoch   5 |    200/ 1052 batches | lr 2.500 | ms/batch  7.25 | loss  0.53 |
| epoch   5 |    400/ 1052 batches | lr 2.500 | ms/batch  7.40 | loss  0.49 |
| epoch   5 |    600/ 1052 batches | lr 2.500 | ms/batch  7.54 | loss  0.44 |
| epoch   5 |    800/ 1052 batches | lr 2.500 | ms/batch  7.60 | loss  0.40 |
| epoch   5 |   1000/ 1052 batches | lr 2.500 | ms/batch  7.43 | loss  0.38 |
-----------------------------------------------------------------------------------------
| end of epoch   5 | time:  7.87s | valid loss  0.62 | valid acc    0.790 |
-----------------------------------------------------------------------------------------
| epoch   6 |    200/ 1052 batches | lr 2.500 | ms/batch  7.74 | loss  0.35 |
| epoch   6 |    400/ 1052 batches | lr 2.500 | ms/batch  7.42 | loss  0.34 |
| epoch   6 |    600/ 1052 batches | lr 2.500 | ms/batch  7.44 | loss  0.31 |
| epoch   6 |    800/ 1052 batches | lr 2.500 | ms/batch  7.57 | loss  0.29 |
| epoch   6 |   1000/ 1052 batches | lr 2.500 | ms/batch  7.47 | loss  0.28 |
-----------------------------------------------------------------------------------------
| end of epoch   6 | time:  8.00s | valid loss  0.58 | valid acc    0.810 |
-----------------------------------------------------------------------------------------
| epoch   7 |    200/ 1052 batches | lr 2.500 | ms/batch  7.42 | loss  0.27 |
| epoch   7 |    400/ 1052 batches | lr 2.500 | ms/batch  7.59 | loss  0.26 |
| epoch   7 |    600/ 1052 batches | lr 2.500 | ms/batch  7.43 | loss  0.25 |
| epoch   7 |    800/ 1052 batches | lr 2.500 | ms/batch  7.38 | loss  0.22 |
| epoch   7 |   1000/ 1052 batches | lr 2.500 | ms/batch  7.23 | loss  0.23 |
-----------------------------------------------------------------------------------------
| end of epoch   7 | time:  7.85s | valid loss  0.57 | valid acc    0.831 |
```

```
-------------------------------------------------------------------------------------
| epoch   8 |    200/ 1052 batches | lr 2.500 | ms/batch  7.29 | loss  0.23 |
| epoch   8 |    400/ 1052 batches | lr 2.500 | ms/batch  7.28 | loss  0.21 |
| epoch   8 |    600/ 1052 batches | lr 2.500 | ms/batch  7.24 | loss  0.20 |
| epoch   8 |    800/ 1052 batches | lr 2.500 | ms/batch  7.29 | loss  0.19 |
| epoch   8 |   1000/ 1052 batches | lr 2.500 | ms/batch  7.24 | loss  0.19 |
-------------------------------------------------------------------------------------
| end of epoch   8 | time: 7.71s | valid loss  0.59 | valid acc    0.843 |
-------------------------------------------------------------------------------------
| epoch   9 |    200/ 1052 batches | lr 0.625 | ms/batch  7.33 | loss  0.18 |
| epoch   9 |    400/ 1052 batches | lr 0.625 | ms/batch  7.27 | loss  0.16 |
| epoch   9 |    600/ 1052 batches | lr 0.625 | ms/batch  7.34 | loss  0.15 |
| epoch   9 |    800/ 1052 batches | lr 0.625 | ms/batch  7.32 | loss  0.12 |
| epoch   9 |   1000/ 1052 batches | lr 0.625 | ms/batch  7.39 | loss  0.13 |
-------------------------------------------------------------------------------------
| end of epoch   9 | time: 7.78s | valid loss  0.61 | valid acc    0.845 |
-------------------------------------------------------------------------------------
| epoch  10 |    200/ 1052 batches | lr 0.156 | ms/batch  7.49 | loss  0.15 |
| epoch  10 |    400/ 1052 batches | lr 0.156 | ms/batch  7.50 | loss  0.14 |
| epoch  10 |    600/ 1052 batches | lr 0.156 | ms/batch  7.57 | loss  0.13 |
| epoch  10 |    800/ 1052 batches | lr 0.156 | ms/batch  7.61 | loss  0.10 |
| epoch  10 |   1000/ 1052 batches | lr 0.156 | ms/batch  7.53 | loss  0.10 |
-------------------------------------------------------------------------------------
| end of epoch  10 | time: 7.97s | valid loss  0.65 | valid acc    0.850 |
-------------------------------------------------------------------------------------
=====================================================================================
==
| End of training | test loss  0.51 | test acc    0.845 |
=====================================================================================
==
```