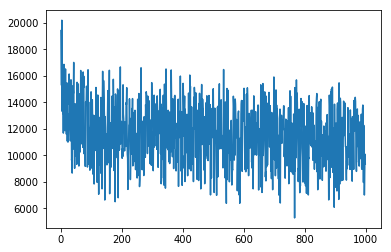
FINAL PROJECT TESTS

1. batch\_size = 256, hidden\_dim = 256, lr = 0.01, reg\_param = 0.003, train\_iter = 10000. 3:2 partitioning of train:test

Train loss= 1.78 (*not a typo, actually this low*)

Test loss = 16000 +

2. batch\_size = 512, hidden\_dim = 128, lr = 0.001, reg\_param = 0.01, train\_iter = 1000. 3:2 partitioning of train:test



Train loss (final 3): 8708.805

10089.638

9358.905

Test mse is: 14988.09806766329

3. Increased train\_iter to 2000. Changed partition of data to 4:1

batch\_size = 512, hidden\_dim = 128, lr = 0.001, reg\_param = 0.01, train\_iter = 2000. 3:2 partitioning of train:test

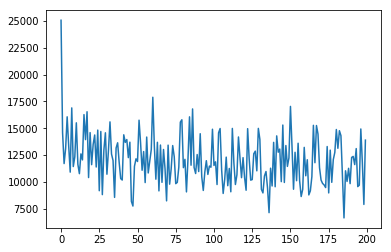
Training iteration: 1997, Loss: 10523.8349609375

Training iteration: 1998, Loss: 9022.9638671875

Training iteration: 1999, Loss: 9883.408203125

Training iteration: 2000, Loss: 9826.5576171875

Test mse is: 13372.60358642749



*Loss graph prints 1 in every 10 iters*

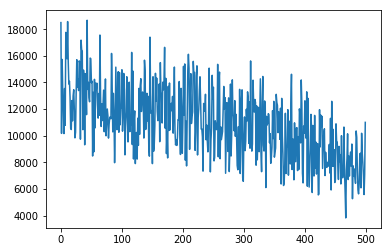
4. Increased num\_iters to 5000.

batch\_size = 512, hidden\_dim = 128, lr = 0.001, reg\_param = 0.01, train\_iter = 5000. 3:2 partitioning of train:test

Training iteration: 4998, Loss: 10030.92578125

Training iteration: 4999, Loss: 8605.8251953125

Training iteration: 5000, Loss: 7190.03466796875



Test mse is: 14028.862189816384

5. Same trained model as (4), but changed input seq len to **3** at test time.

batch\_size = 512, hidden\_dim = 128, lr = 0.001, reg\_param = 0.01, train\_iter = 1000. 3:2 partitioning of train:test

Test mse is: 13098.331856319302

**6. Decreased hidden\_dim to 64**

batch\_size = 512, hidden\_dim = 64, lr = 0.001, reg\_param = 0.01, train\_iter = 5000. 3:2 partitioning of train:test

Training iteration: 1, Loss: 17707.6796875

(512, 10, 10)

Training iteration: 2, Loss: 21311.0

(512, 10, 10)

Training iteration: 3, Loss: 21780.34375

……….

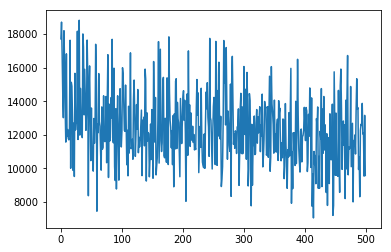
Training iteration: 4998, Loss: 8060.33642578125

(512, 10, 10)

Training iteration: 4999, Loss: 12742.0712890625

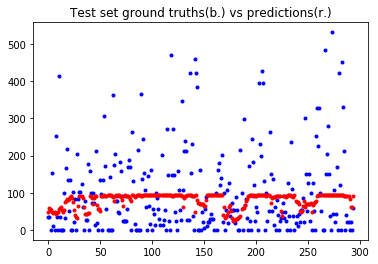
(512, 10, 10)

Training iteration: 5000, Loss: 10084.2607421875



Test mse is: 12383.044540755754 (**input\_seq\_len = 7**)

*Remarks: Test mse varies with the input sequence len at test time, even though input length at train time is const at 10*

**

**7. Increased hidden\_dim to 512 to reduce bias in test-time. Trained for 2400 iters.**

batch\_size = 512, hidden\_dim = 512, lr = 0.001, reg\_param = 0.01, train\_iter = 2400. 3:2 partitioning of train:test

Training iteration: 1, Loss: 18207.099609375

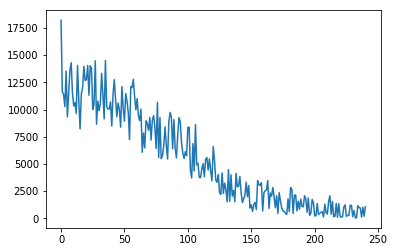
Training iteration: 2, Loss: 14612.6728515625

…

Training iteration: 2402, Loss: 1006.8379516601562

Training iteration: 2403, Loss: 151.23841857910156

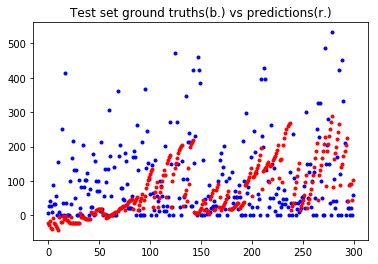
Training iteration: 2404, Loss: 1042.6748046875



Test time:

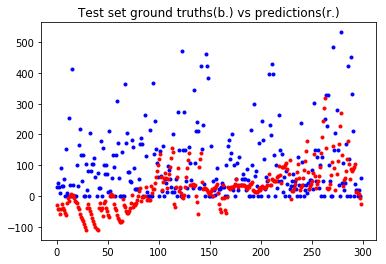
1

Test mse is: 16105.85651547784



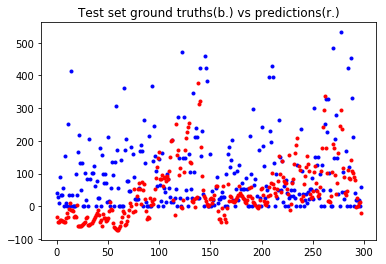
2

Test mse is: 19630.599959061317



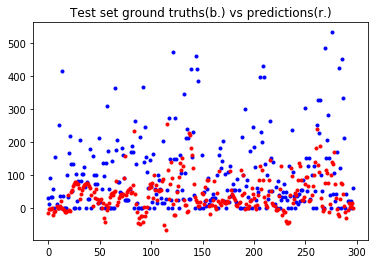
3

Test mse is: 19479.915241393774



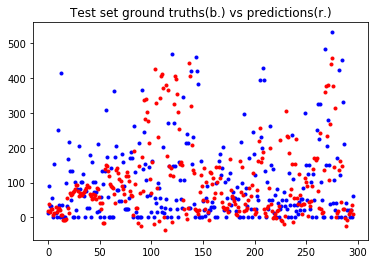
4

Test mse is: 17364.10705239845



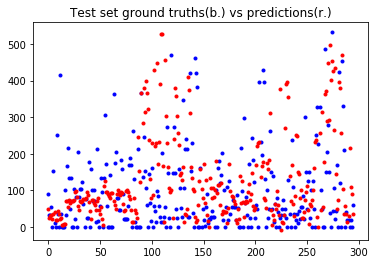
5

Test mse is: 19609.85366341085



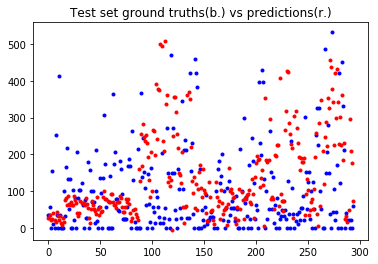
6

Test mse is: 22896.75983137835



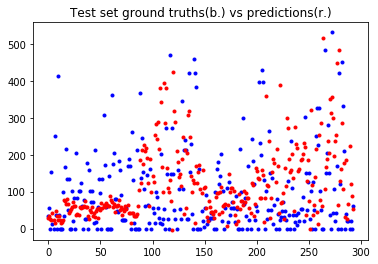
7

Test mse is: 24673.843902506233



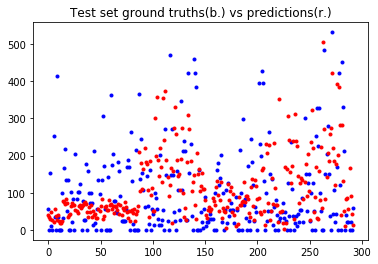
8

Test mse is: 18771.93061891684



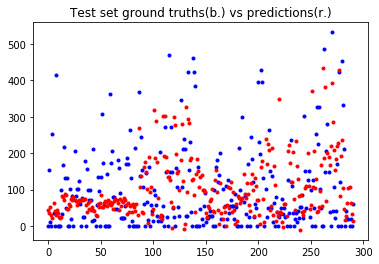
9

Test mse is: 16201.410963627657



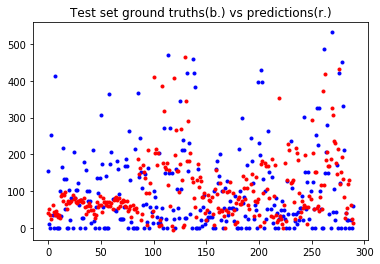
10

Test mse is: 15804.937109936342



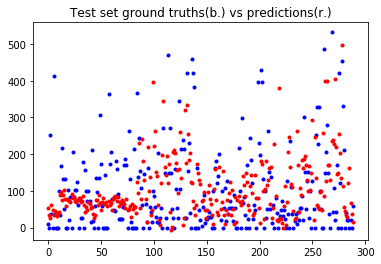
11

Test mse is: 16800.102402486977



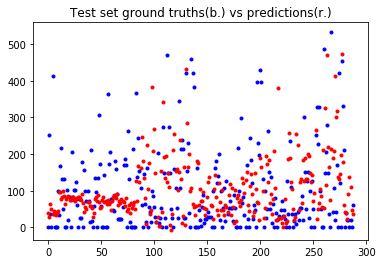
12

Test mse is: 15376.500265792367



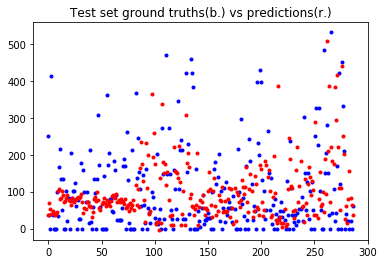
13

Test mse is: 14908.753613451845



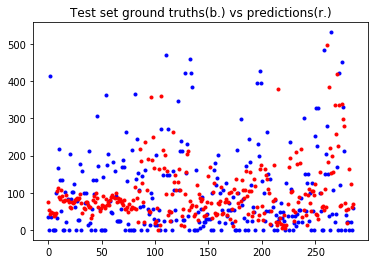
14

Test mse is: 14936.667901462853



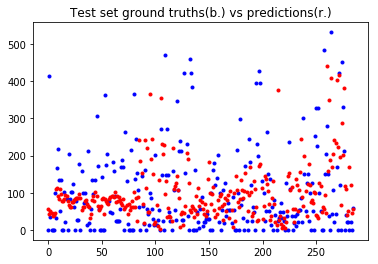
15

Test mse is: 15129.730018998962



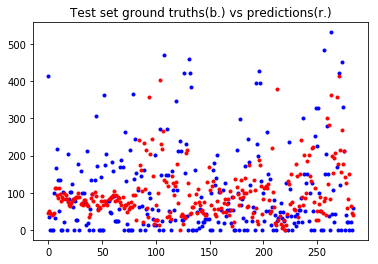
16.

Test mse is: 14633.555414118784



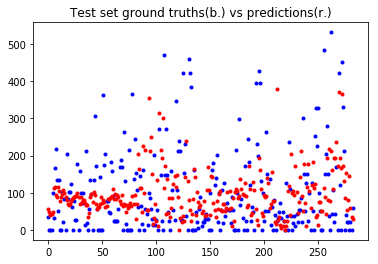
17

Test mse is: 14432.465322798016



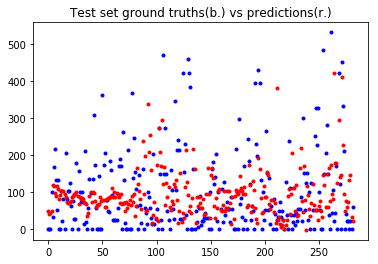
18.

Test mse is: 13977.601414407629



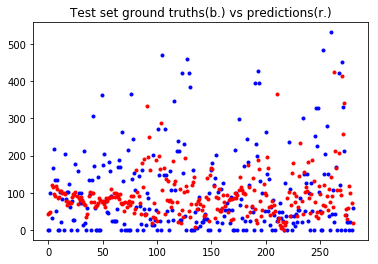
19.

Test mse is: 13791.237583072632



20.

Test mse is: 14421.850907507313

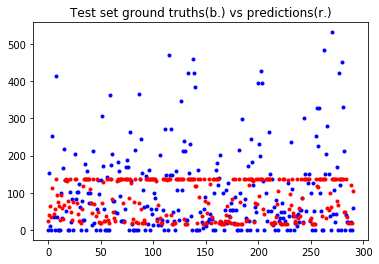


**8. Used one-hot representations of features. Same other features as last time.**

Number of features: 35

Started testing...

Test mse is: 11111.708348846361

****