

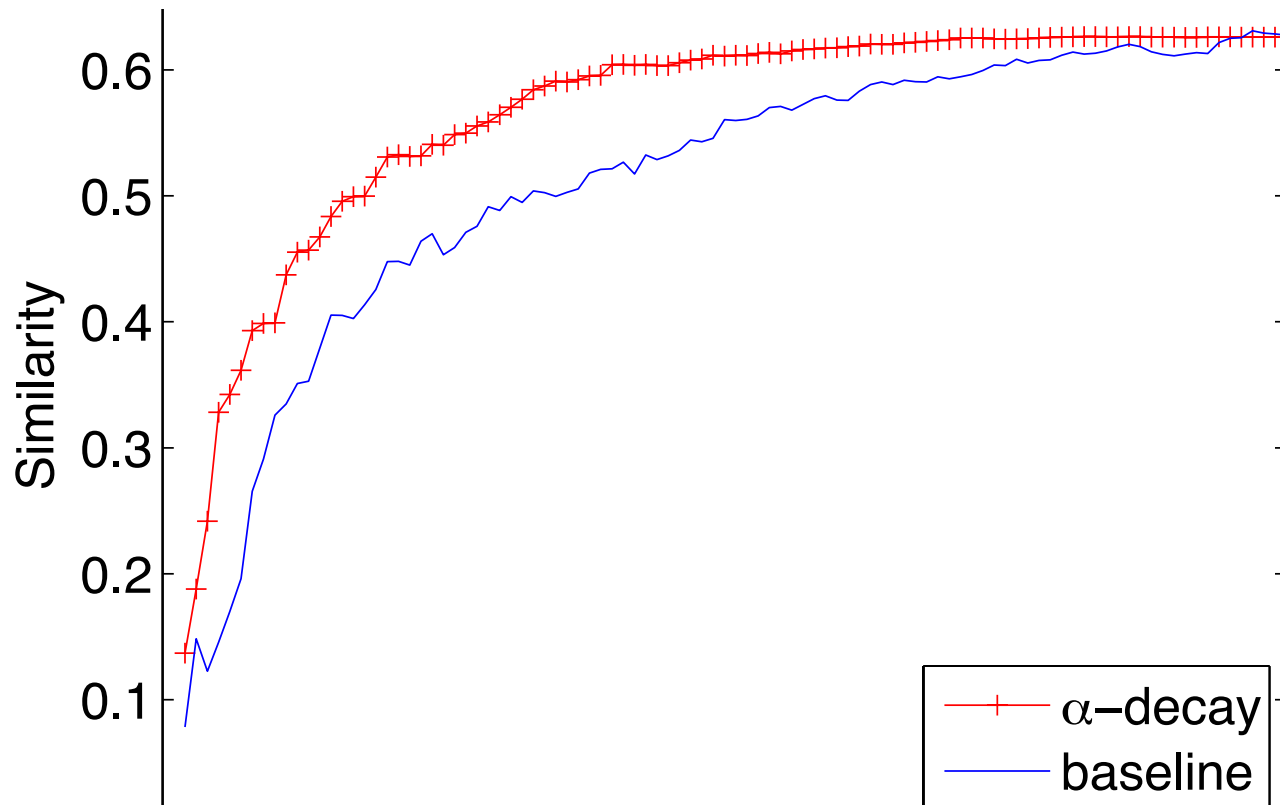
# Bi(Tri)-month report

Xiaoxi Wang

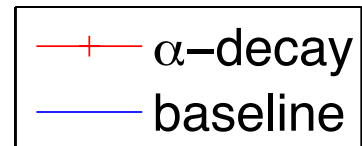
# Apr. ~ May.

- Sparse coding
- Nested dropout
  - Ordered vectors
- Ordered Vectors
  - Done a lot work based upon changing learning rate

# Ordered Vectors on Sim353

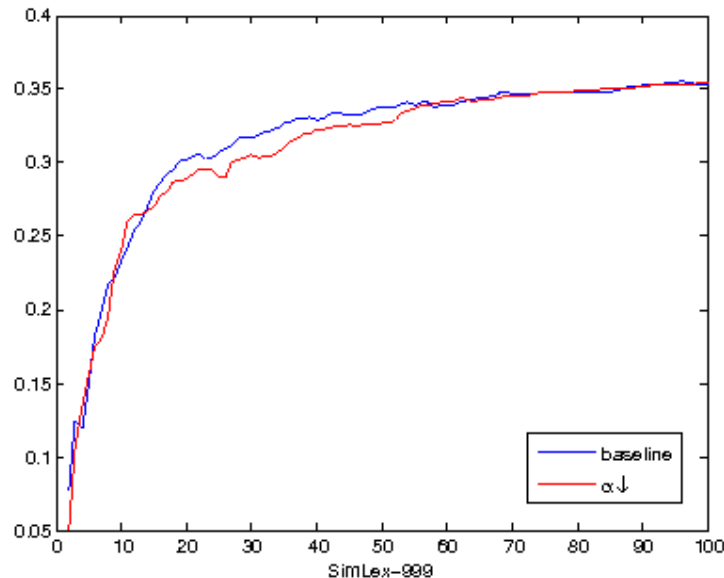


Looks Good!



# Submitted

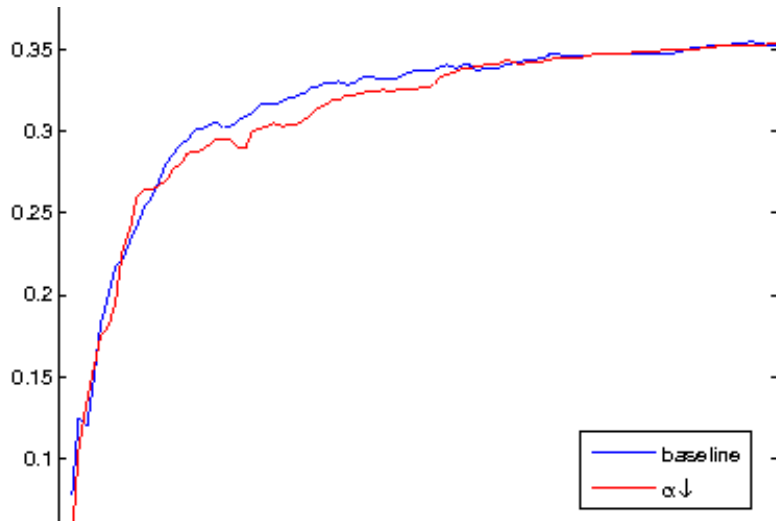
- And Failed...
  - Not enough Experiments
  - Training corpus is small
  - Test set is too simple...



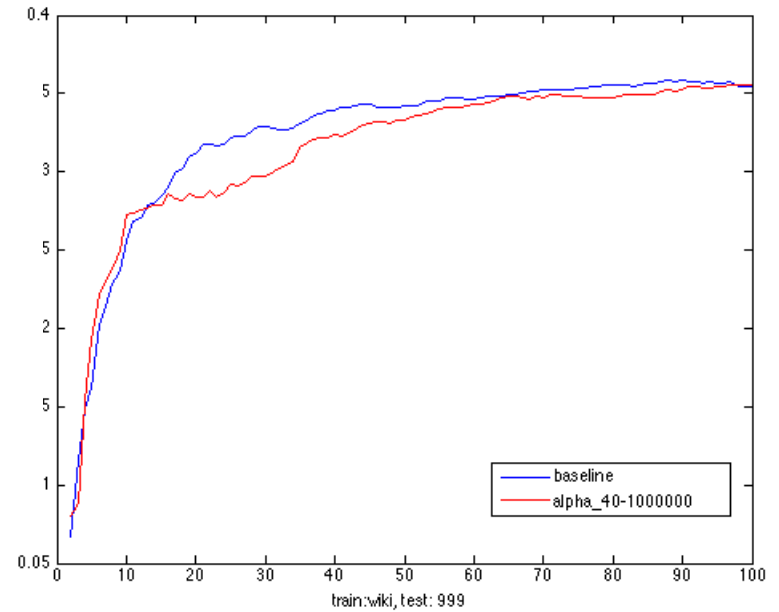
# May. ~ Now

- More test:
  - Iteration test: (CVSS 384)
  - More training corpora
    - Text8 (100M)
    - Wiki (8.6G) & Wiki 10%(875M)
    - Eng (2.1G)
  - , test sets
    - Sim353, RG-test, SimLex999
  - and methods(CVSS 391)

# Different factors

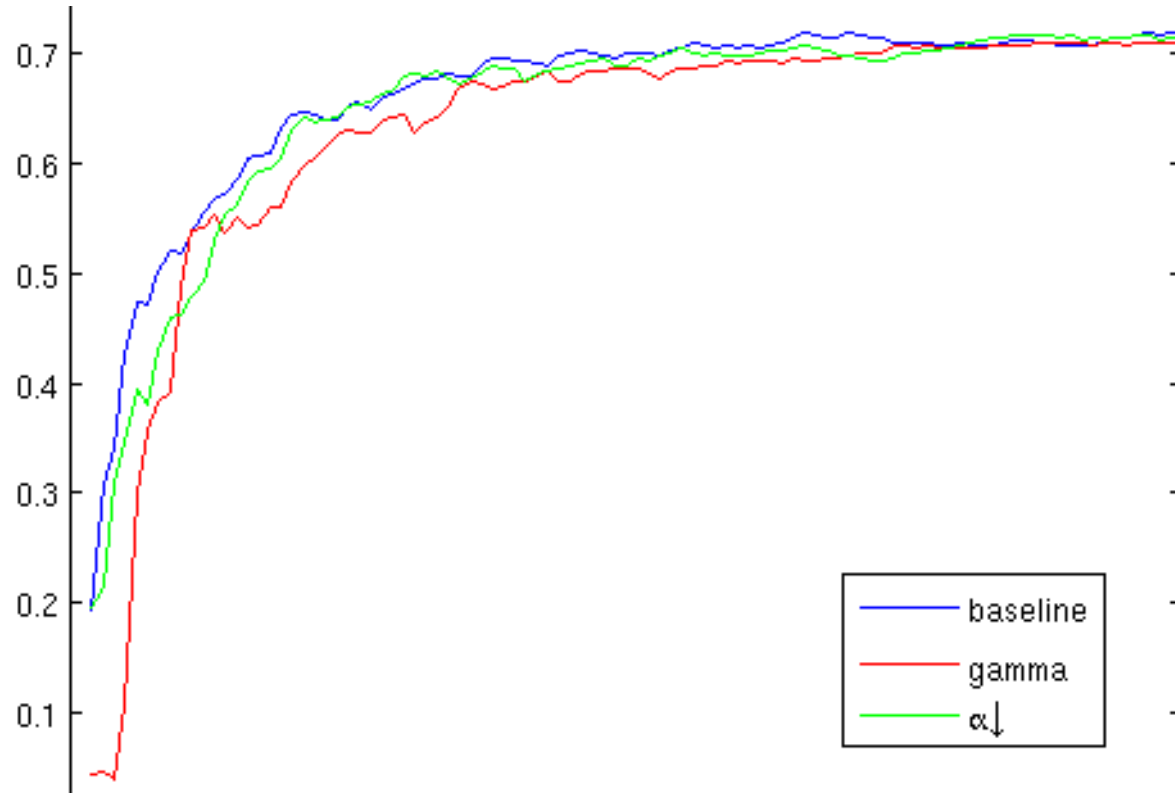


B=40-100000

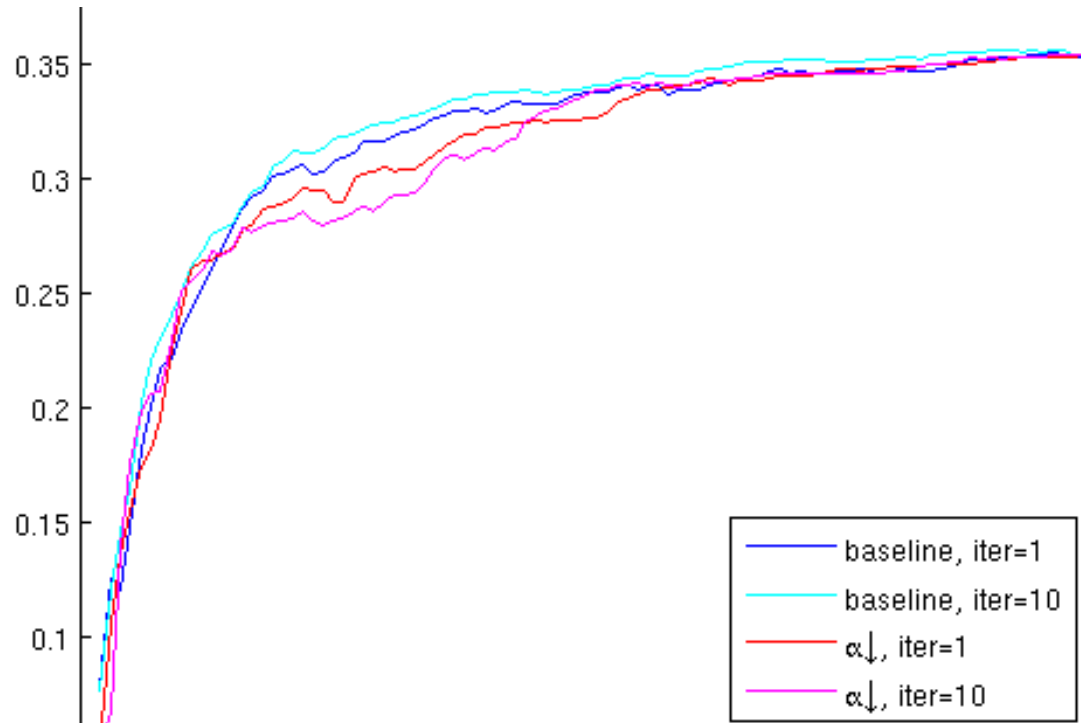


B=40-1000000

# Different Methods



# Iteration test

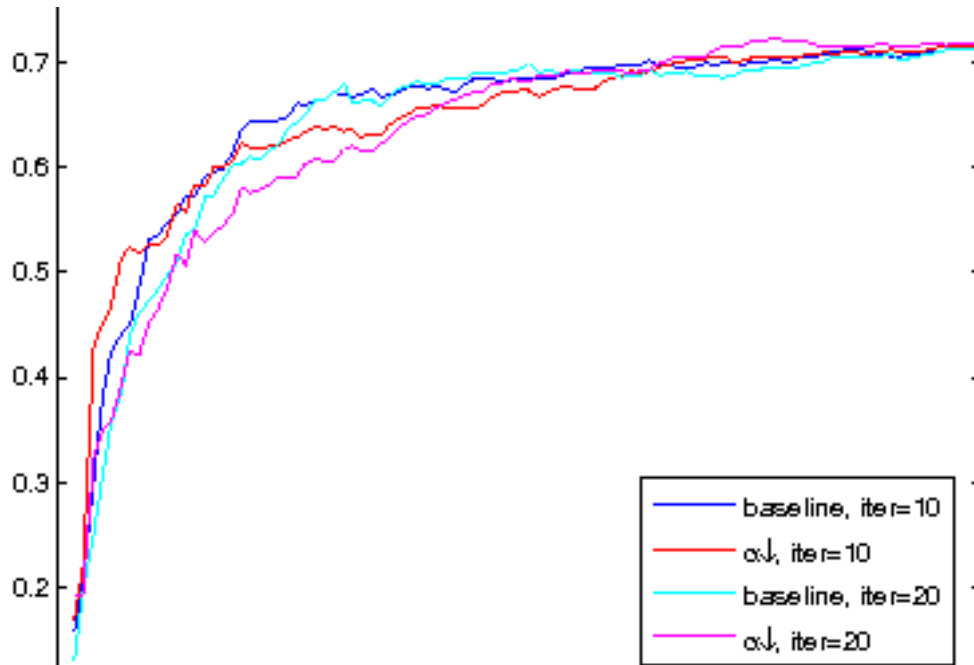


Train: Wiki  
Test: SimLex999  
Iter=1,10  
100 samples

More iterations but low performance



# Iteration test



Train: Wiki\_small  
Test: RG  
Iter=10, 20  
10 samples

More iterations but low performance

# Next step

- Find something really really useful (beat the baseline)
- After this
  - Maybe Do some LSTM work then.