1. Change hyper parameters of the Text Classification with new dataset

Algorithm:

Load positive and negative sentences from the raw data files.

Clean the text data using the same code as the original paper.

Pad each sentence to the maximum sentence length. Padding sentences to the same length is useful because it allows us to efficiently batch our data since each example in a batch must be of the same length.

Build a vocabulary index and map each word to an integer between 0 and 6,413 (the vocabulary size). Each sentence becomes a vector of integers.

**a. Optimizers**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=32

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=3,4,5

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=100

NUM\_FILTERS=128

|  |  |  |
| --- | --- | --- |
| AdamOptimizer: | GradientDescentOptimizer: | RMSPropOptimizer: |
| Evaluation:  step 2800  loss 1.59227e+07  acc 0.86875 | Evaluation:  step 2800  loss 3.46284  acc 0.72449 | Evaluation:  step 2800  loss 457629  acc 0.806122 |
| Conclusion: AdamOptimizer with filter size(3,4,5) have the best accuracy(0.86875) | | |

**b. Filter size (fixed the optimizer with AdamOptimizer)**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=32

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=1,2,3

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=100

NUM\_FILTERS=128

Loading data...

Vocabulary Size: 6413

Train/Dev split: 886/98

|  |  |  |
| --- | --- | --- |
| 3, 4, 5: | 6, 8, 10: | 1, 2, 3: |
| Evaluation:  step 2800  loss 1.59227e+07  acc 0.86875 | Evaluation:  step 2800  loss 3.68776e+07  acc 0.765306 | Evaluation:  step 2800  loss 457629  acc 0.806122 |
| Conclusion: AdamOptimizer with filter size(3,4,5) have the best accuracy(0.86875) | | |

**c. Number of filters:**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=32

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=3,4,5

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=100

NUM\_FILTERS=64

|  |  |  |
| --- | --- | --- |
| 64: | 128: | 256: |
| Evaluation:  step 2800  loss 6.24601e+06  acc 0.826531 | Evaluation:  step 2800  loss 1.59227e+07  acc 0.86875 | Evaluation:  step 2800  loss 6.6569e+07  acc 0.785714 |
| Conclusion: AdamOptimizer with filter size(3,4,5) , number of filter=128 have the best accuracy(0. 86875) | | |

**d. Dropout probability:**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=32

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=3,4,5

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=100

NUM\_FILTERS=128

|  |  |  |
| --- | --- | --- |
| 0.25: | 0.5: | 0.75: |
| Evaluation:  step 2800  loss 3.14969e+07  acc 0.806122 | Evaluation:  step 2800  loss 6.24601e+06  acc 0.826531 | Evaluation:  step 2800  loss 1.25843e+07  cc 0.77551 |
| Conclusion: AdamOptimizer with filter size(3,4,5), number of filter=128, dropout probability= 0.5 have the best accuracy(0.826531) | | |

**e. Batch size:**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=64

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=3,4,5

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NEGATIVE\_DATA\_FILE=C:/Users/disfly/PycharmProjects/untitled/newdata2.txt

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=100

NUM\_FILTERS=128

|  |  |  |
| --- | --- | --- |
| 32: | 64: | 128: |
| Evaluation:  step 2800  loss 6.24601e+06  acc 0.826531 | Evaluation:  step 1400  loss 8.95438e+06  acc 0.755102 | Evaluation:  step 700  loss 2.76903e+06  acc 0.734694 |
| Conclusion: AdamOptimizer with filter size(3,4,5), number of filter=128, dropout probability= 0.5 batch size=64 have the best accuracy(0.826531) | | |

**f. Number of epochs:**

Parameters:

ALLOW\_SOFT\_PLACEMENT=True

BATCH\_SIZE=32

CHECKPOINT\_EVERY=100

DEV\_SAMPLE\_PERCENTAGE=0.1

DROPOUT\_KEEP\_PROB=0.5

EMBEDDING\_DIM=128

EVALUATE\_EVERY=100

FILTER\_SIZES=3,4,5

L2\_REG\_LAMBDA=0.0

LOG\_DEVICE\_PLACEMENT=False

NEGATIVE\_DATA\_FILE=C:/Users/disfly/PycharmProjects/untitled/newdata2.txt

NUM\_CHECKPOINTS=5

NUM\_EPOCHS=200

NUM\_FILTERS=128

|  |  |  |
| --- | --- | --- |
| 50: | 100: | 200: |
| Evaluation:  step 1400  loss 1.02921e+07  acc 0.795918 | Evaluation:  step 2800  loss 6.24601e+06  acc 0.826531 | Evaluation:  step 5600  loss 3.19904e+07  acc 0.816327 |
| Conclusion: AdamOptimizer with filter size(3,4,5), number of filter=128, dropout probability= 0.5 batch size=64, number of epochs =100, have the best accuracy(0.826531) | | |

**Tensorboard:**

