Figure\_1.png: (general idea how training/validation goes)

Vocabulary size: 11293

SPLIT = 0.2  
PRETRAIN = False  
BIDIRECTIONAL = False  
NUM\_CLASSES = 2  
BATCH\_SIZE = 128  
DROPOUT\_KEEP\_PROB = 0.8

NUM\_EPOCHS = 3  
LEARNING\_RATE = 0.001  
DECAY\_STEPS = 1200  
DECAY\_RATE = 0.9

EMBEDDING\_DIM = 50

Figure\_2.png: (how epoch affect)

same as 1, NUM\_EPOCH = 10

**epoch = 3-4 pretty good**

Figure\_3.png: (how dimension affect)

NUM\_EPOCHS = 3 (last epoch)

EMBEDDING\_DIM = 50/100/150/200

**Embedding = 150/200 pretty good**

Figure\_4.png(how leaning rate affect)

EMBEDDING\_DIM = 150

EPOCH = 3 (last epoch)

lr = [0.1, 0.01, 0.001, 0.0001]

**LR = 0.001 pretty good**

Figure\_5.png(how batch size affect)

[32, 64, 128]:

**NOT CLEAR, need tune with LR**

Figure 6-7.png(7lstm vs 6bidirectional)

embedding = 150

batch\_size = 64

epoch = 3

lr = 0.001

**bidirectional better**

Figure (9embedding/8pretrained)

200 dimension

Vocabulary size: 1193514

batch = 64

**self-train better**