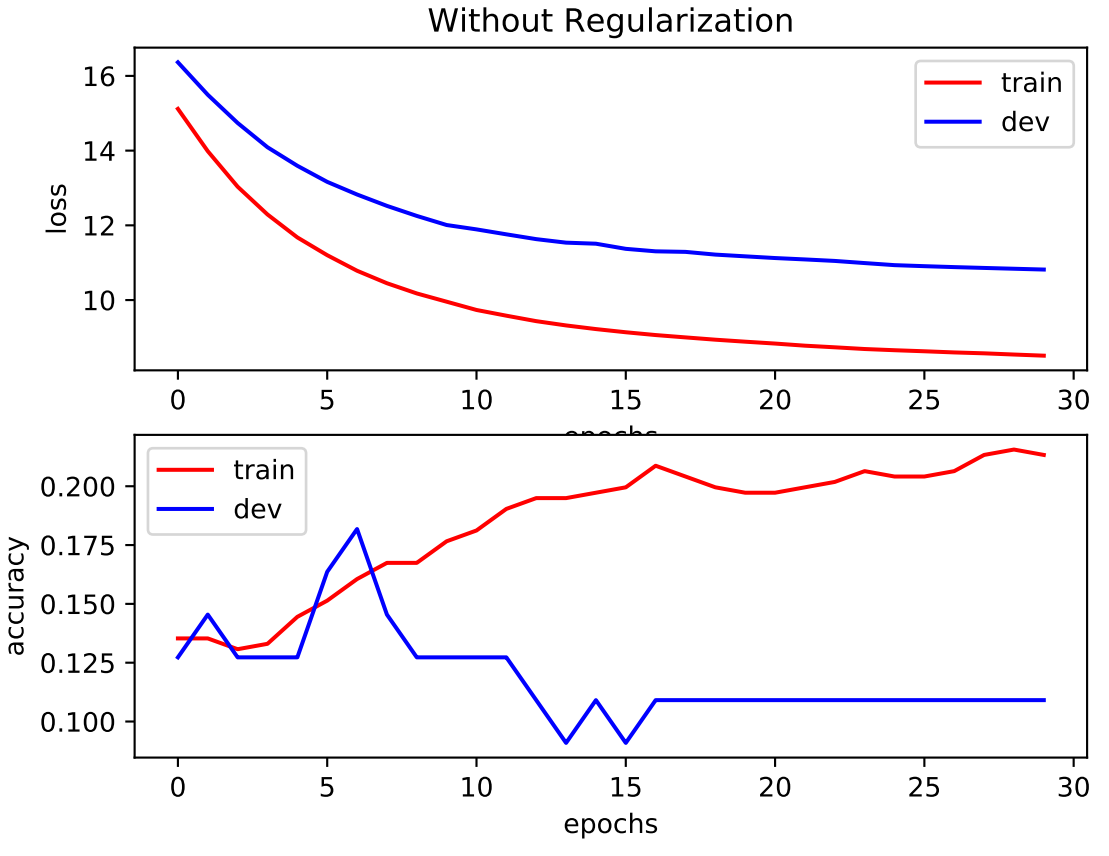
Without Regularization
loss
16
14
12
10
train
dev
0
5
10
15
20
25
30
accuracy
0.200
0.175
0.150
0.125
0.100
train
dev
0
5
10
15
20
25
30
epochs