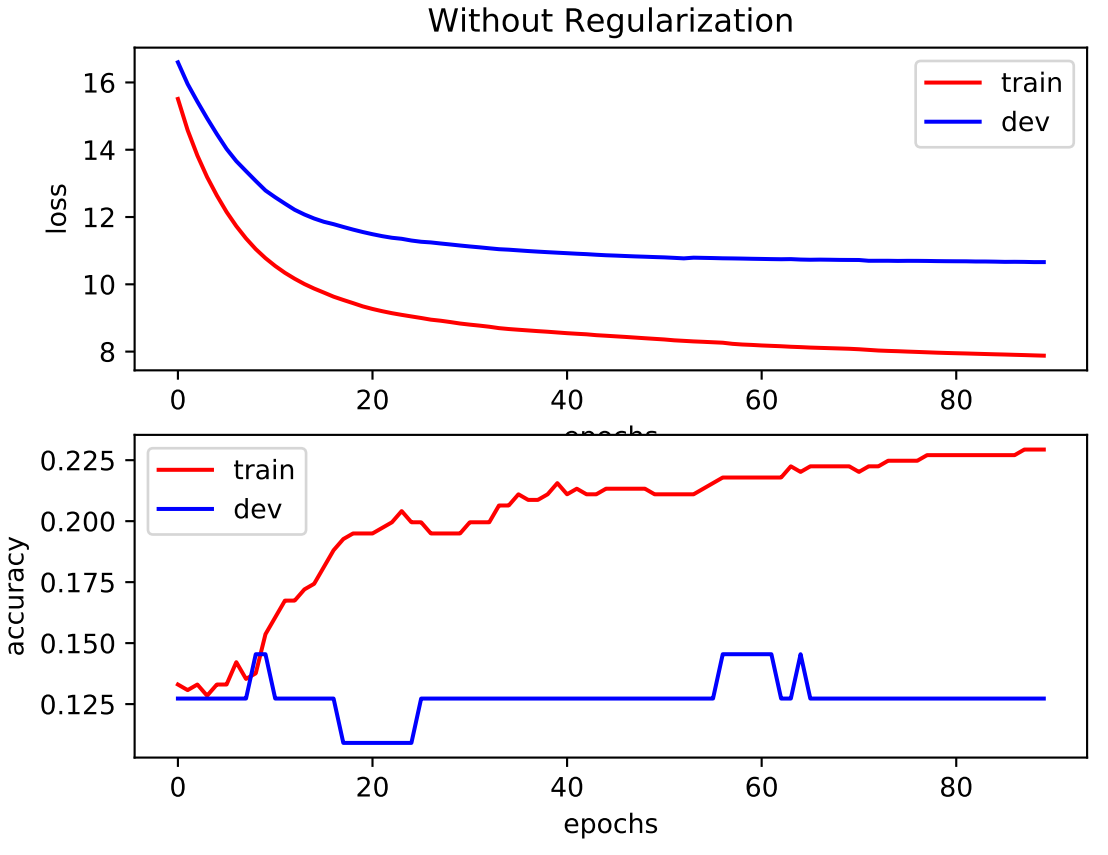
Without Regularization
train
dev
loss
16
14
12
10
8
0
20
40
60
80
train
dev
accuracy
0.225
0.200
0.175
0.150
0.125
0
20
40
60
80
epochs