1011 Quiz 4

Q1 "Observing "NaN Loss" during training (get loss term as numpy.nan)"

Q2 DenseNet, ResNet, Highway Networks

Q3 Vanilla RNN

Q4 either or both of (c,d) and (d,c)

Q5 The network is too deep (n>100), Norm of gradient at each layer consistently < 1

Q6 F(x) + x

Q7 False

Q8 False

Q9 same

Q10 As long as you write something reasonable such as: normalizing the data/ initializing the weights in a reasonable way / adjusting gradient norm