**Day 51**

**What to do?**

Learn about residual network.

**ResNets:**

In **plain** neural network, we have

A[l] A[l+1] A[l+2]



In other words, we have

Z[l+1] = W[l+1] \* A[l] + b[l+1]; A[l+1] = g(Z[l+1])

Z[l+2] = W[l+2] \* A[l+1] + b[l+2]; A[l+2] = g(Z[l+2])

In **residual** network, we have

A[l]  Linear ReLU A[l+1] Linear ReLU A[l+2]

“Shortcut”

The activations of layer ‘l’ are used in another layer. Hence,

A[l+2] = g(Z[l+2] + A[l])

When activations of previous layer are added to current layer, that is called as “Residual block”. This architecture, that was introduced in 2015, performs “skip connections” mechanism and consists of features with heavy batch normalization.

**Why ResNets?**

As number of layers increase the training error increases with number of training examples, which in theory should not. Consider an autonomous driving vehicle. It consists of hundreds of layers, yet still achieves better performance. Residual networks ensure the better performance when it comes to very deep network. They make sure that the accuracy does not degrade, at least not rapidly.

