ASSIGNMEN-2

Mounica Subramani

1a) **Vanishing gradient:**

It is an unstable behavior when training a deep neural network. It is the situation where a deep multi layer NN (or) RNN is unable to propogate useful gradient information from the output end of the model back to the layers near the input end of the model.

Why it happens: A problem with training networks with many layers is that the gradient diminishes drawatically as it is propogated backward through the network. [the backpropogated error typically decreases (or increases) exponentially as function of distance from the final layer]

The result is, the general inability of models with many layers to learn on a given dataset.

Exploding gradient:

It is a situation where large error gradients accumulate and result in very large updates to neural network model weights during training.

An error gradient is the direction and magnitude calculated during the training of a neural network that is used to update the network weights in right direction and by right amount.

Why it happens: The explosion occurs through exponential growth by repeatedly multiplying gradients through the network layers that have values larger than 1.0. ie) explosion of long term components.