## horizontal line

Vanishing Gradient Problem (class 12)

18.01.2025

# VGP

In such methods during each iteration updates in weights i.e. gradients are very small or ‘vanish’ thus the loss stops reducing and completely stops from further training.

1. 0.1 x 0.1 x 0.1 x 0.1 = 0.0001

Within deep neural networks , when gradients of weights are complicated after product they become very very small.

1. Using sigmoid / tanh functions

They give probability in between 0 and 1 thus their product is very small.

# How to recognize ?

1. If loss doesn’t reduce after epochs
2. If weights graph is a constant

# How to handle VGP ?

1. Reduce model complexity by taking a shallow architecture . but it fails to predict complex patterns / problems .
2. Using the RELU activation function .

Max (0,z) pros is a product if z will give a large no. cons is 0 multiplied if it fails.

1. Proper weight initialization .
2. Batch normalization
3. Residual Network

Similarly there occurs an Exploding Gradient Problem in KNN , meaning loss explodes .