GENERAL PROBLEMS IN TRAINING NEURAL NETWORKS

1. VANISHING & EXPLODING GRADIENTS
2. LOT OF DATA REQUIRED TO TRAIN
3. SLOW TRAINING ISSUES
4. RISK OF OVERFITTING.

(1) $W = W - \eta \nabla C \to 0 \Rightarrow \Delta W \downarrow$

Vanishing Gradients

$W = W - \eta \nabla C \to \infty \Rightarrow \Delta W \uparrow$

Exploding Gradients.

(2), (4) Huge & Risk of overfitting

Trainable parameters ↑

(3) Increase size of NN for your complex data

⇓

increase in no. of hidden layers

⇓

increases the size of the model

⇓

Training slows

⇓

Slow Inferencing / prediction + Latency.