EXPLODING
GRADIENT PROBLEM

When the gradients of the loss function with respect to the parameters in the early layers are very large. Causes unstable learning and poor performance. One potential fix is gradient clipping.

Chris Albon