DeepLearning(worksheet-5)

1. (D)
2. (A)
3. (D)
4. (A)
5. (B)
6. (A)
7. (A)
8. (C)
9. (A,B,C,D)
10. (C,D)
11. In convex optimization local min = global min strictly convex: unique min efficient solvers strong theoretical guarantees.

In non-convex optimization multiple local min 6= global min many solvers come from convex world weak theoretical guarantees if any

1. When we optimize neural networks or any high dimensional function, for most of the trajectory we optimize, the critical points(the points where the derivative is zero or close to zero) are saddle points. Saddle points, unlike local minima, are easily escapable."
2. In deep neural nets with several layers, one forward pass simply entails performing consecutive matrix multiplications at each layer, between that layer’s inputs and weight matrix. The product of this multiplication at one layer becomes the inputs of the subsequent layer, and so on and so forth. The aim of weight initialization is to prevent layer activation outputs from exploding or vanishing during the course of a forward pass through a deep neural network.
3. internal Covariate Shift as the change in the distribution of network activations due to the change in network parameters during training. In neural networks, the output of the first layer feeds into the second layer, the output of the second layer feeds into the third, and so on