

PART A

1. It is used to skip the networks, thereby reduces the training time and complexity
2. Hyper parameters control the training of any given ML or DL procedures. By optimizing these parameters we reduce the training time and we try to increase the performance efficiency of the network
5. Spatial Transformer network-diagram and explanation for 5 lines
4. Residual Networks – explanation
3. ReLU captures the interactions between the data's

Part B

For all the questions below diagram and explanation is needed

- 6) Auto encoder
- 7) AlexNet
- 8) LSTM
- 9) non convex optimization