1) Prove that the smoothest line connecting two points is a straight line
2) What is machine learning: what do we know, and what do we want
3) What is the main objective of the traditional neural network?
4) Why in the traditional neural network, we only use linear and 1-input nonlinear neurons
5) Why cannot we have 1- or 2-layer neural networks?
6) Why we use L-NL-L networks and not NL-L-NL ones?
7) Explain why 3-layer neural networks are universal approximators
8) Which activation function is used in traditional neural networks? Why (biological answer).
9) What do we want to minimize when we train a neural network
10) What method do we use to minimize the corresponding objective function?
11) Explain why gradient method is used for optimization
12) Explain how gradient descent leads to backpropagation formulas, at least formulas for $W_0$ and $W_i$. 