Chapter 1

What is Stochastic gradient descent? What is minibatch SGD?

What is backpropagation?

What is step (epoch) of backpropagation?

What is a forward/backward pass?

What is XOR problem? Why not a perceptron could solve this, and why deep learning could solve this?