Assignment 3
Neural Network Fundamentals

Part A — Foundations

26 questions

★	**Must-know** — asked very frequently in GenAI / ML interviews (8 questions). Prepare these first.
[CODE]	**Hands-on task** — write and run the code (3 tasks).

1. How does a neural network work? Explain with a single neuron first. ★
2. What is the weight update formula in a neural network?
3. Define: epoch, sample, batch size, iteration, learning rate, weights, biases.
4. What is the difference between a loss function and a cost function?
5. Explain the learning rate. What happens if it is too high or too low? ★
6. Explain forward propagation and backpropagation. ★
7. What is data normalization/standardization and why is it needed for neural networks?
8. What are the different types of optimizers (SGD, Momentum, RMSProp, Adam, AdamW)? ★
9. Explain Batch Gradient Descent, Stochastic Gradient Descent and Mini-batch Gradient Descent.
10. Explain momentum and learning-rate decay / schedulers (warm-up, cosine decay).
11. Which optimizer is most commonly used today, and why is AdamW preferred for Transformers?
12. What is the difference between parameters and hyperparameters? Name important hyperparameters.
13. What is the use of an activation function? What happens without one? ★
14. Explain the activation functions: Sigmoid, Tanh, ReLU, Leaky ReLU, GELU, Softmax. Which is used where?
15. Why are Sigmoid/Tanh not preferred in hidden layers? Which activation is most common in hidden layers?
16. Which loss function should be used for regression, binary classification, multi-class and multi-label problems? ★
17. What are vanishing and exploding gradients? Why do they occur and how do you fix them? ★

18. What are the weight initialization methods (Xavier/Glorot, He)? Why does initialization matter?
19. What is overfitting? List regularization methods (dropout, L1/L2, early stopping, data augmentation).
20. What is Batch Normalization? How does it differ from Layer Normalization?
21. How would you change a pre-trained classification network into a regression network?
22. If the previous layer has 8 neurons and the current Dense layer has 512 neurons, how many parameters does the current layer have? (Show the calculation.) ★
23. What are the steps to build an ANN using TensorFlow/Keras? What is the Sequential API?
24. Scale data for a neural network (StandardScaler / MinMaxScaler) — write the code. [CODE]
25. Define a model, add layers, compile it and train it (Keras or PyTorch). [CODE]
26. Set a custom weight initializer for a layer in TensorFlow. [CODE]