ML_QZ1_Anas_ALhardi

1-In Model based learning methods, an iterative process takes place on the ML models that are built based on various model parameters, called?

A. mini-batches
B. hyperparameters
C. superparameters
D. optimizedparameters

2- Can Gradient Descent be applied to Non-Convex Functions?

A. Yes, Gradient Descent can only be applied to convex functions.
B. No, Gradient Descent can only be applied to concave functions.
C. Yes, Gradient Descent can be applied to both convex and non-convex functions.
D. No, Gradient Descent is not applicable to any type of function.

3- Does Gradient Descent always converge to an optimum?

A. Yes, Gradient Descent always converges to the global optimum for any function.
B. No, Gradient Descent may converge to a local optimum or saddle point in some cases.
C. Yes, Gradient Descent guarantees convergence to a local optimum, but not necessarily the global optimum.
D. No, Gradient Descent only works for convex functions and always finds the global optimum.

Consider a following model for logistic regression: P (y =1|x, w)= g(w0 + w1x) where g(z) is the logistic function.In the above equation the P (y =1|x; w) , viewed as a function of x, that we can get by changing the parameters w.

4- What would be the range of p in such case?

A. The range of $P(y=1|x,w)$ is $(-\infty,\infty)(-\infty,\infty)$, covering all real numbers.
B. The range of $P(y=1|x,w)$ is [0,1][0,1], covering all real numbers between 0 and 1.
C. The range of $P(y=1|x,w)$ depends on the value of $w0$ and $w1$.
D. The range of $P(y=1|x,w)$ is [0.5,1][0.5,1], covering all real numbers between 0.5 and 1.

5-In a neural network, what is the primary purpose of an activation function?

A. To determine the learning rate of the network.
B. To connect different layers of neurons.
C. To introduce non-linearity into the model.
D. To regularize the weights of the network.

6- What is the fundamental difference between supervised learning and unsupervised learning in the context of neural networks, and can you provide an example of each?

A. Supervised learning uses labeled data to train neural networks, while unsupervised learning uses unlabeled data. An example of supervised learning is image classification, and an example of unsupervised learning is clustering.
B. Supervised learning uses reinforcement signals to train neural networks, while unsupervised learning relies on external guidance. An example of supervised learning is speech recognition, and an example of unsupervised learning is dimensionality reduction.
C. Supervised learning uses a single neural network, while unsupervised learning uses multiple interconnected networks. An example of supervised learning is natural language processing, and an example of unsupervised learning is image generation.
D. Supervised learning relies on recurrent neural networks, while unsupervised learning uses convolutional neural networks. An example of supervised learning is time series prediction, and an example of unsupervised learning is text summarization.

7- What is the main characteristic of the Rectified Linear Unit (ReLU) activation function commonly used in neural networks?

A. ReLU squashes input values into a range between -1 and 1.
B. ReLU outputs a continuous probability distribution.
C. ReLU is defined as the logarithm of the input.
D. ReLU returns zero for negative input values and passes through positive input values unchanged.

8- What is the "dying ReLU" problem in neural networks, and how does it typically occur?

A. The "dying ReLU" problem occurs when the ReLU activation function becomes too aggressive and causes gradients to vanish during backpropagation.
B. The "dying ReLU" problem happens when ReLU neurons continuously output a zero activation, effectively becoming inactive and not updating their weights during training.
C. The "dying ReLU" problem arises when the learning rate for ReLU neurons is set too high, causing them to update their weights rapidly and destabilizing the network.
D. The "dying ReLU" problem is related to overfitting, where ReLU neurons fit the training data perfectly but fail to generalize to new, unseen data.