Q.1)

Why is it necessary to normalize the data?

Checkboxes

- [ ] It helps to reduce variance
- [x] To ensure that the neural network considers all input features to a similar extent.
- [x] To make the optimization process of cost function faster.
- [x] For consistency for comparing results across models.

Q.2)

You are making a classifier to classify different species of birds. Suppose your classifier obtains a training set error of 0.5%, and a dev set error of 7%. Which of the following are promising things to try to improve your classifier?

Checkboxes

- [x] Increase regularization
- [x] Train on more data
- [ ] Decrease regularization
- [ ] Increase learning rate

Q.3)

Which of the following techniques are used to reduce variance?

Multiple choice

- Dropout
- L2 Regularization
- Data Augmentation
- All of the above ✓

Q.4)

Which algorithm follows a straight path towards the minimum?

Multiple choice

- Batch Gradient Descent ✓
- Stochastic Gradien Descent
- Mini Batch Gradien Descent
- None of the above

Q.5)

Which of the following statements is true about the learning rate alpha in gradient descent?

Multiple choice

- If alpha is very small, gradient descent will be fast to converge. If alpha is too large, gradie...
- If alpha is very small, gradient descent will be fast to converge. If alpha is too large, gradie...
- If alpha is very small, gradient descent can be slow to converge. If alpha is too large, gradi...
- If alpha is very small, gradient descent can be slow to converge. If alpha is too large... ✓

Q.6)

If searching among a large number of hyperparameters, you should try random values rather than values in a grid , because you don't know which hyperparameters are more important than others. True or False?

Multiple choice

- True ✓
- False

Q.7)

Let us assume we implement an AND function to a single neuron. Below is a tabular representation of an AND function:

What would be the weights and bias?

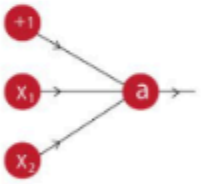
(Hint: For which values of w1, w2 and b does our neuron implement an AND function?)

Multiple choice

X1	X2	X1 AND X2
0	0	0
0	1	0
1	0	0
1	1	1

The activation function of our neuron is denoted as:

$$f(x) = \begin{cases} 0, & for\ x < 0 \\ 1, & for\ x \geq 0 \end{cases}$$



- Bias = 1, w1 = 1.5, w2 = 1.5
- Bias = 1.5, w1 = 2, w2 = 2
- Bias = -1.5, w1 = 1, w2 = 1 ✓
- None of the above

Q.8)

Suppose batch gradient descent in a deep network is taking excessively long to find a value of the parameters that achieves a small value for the cost function J(W[1],b[1],...,W[L],b[L]). Which of the following techniques could help find parameter values that attain a small value forJ?

Checkboxes

- [x] Try using Adam
- [x] Try tuning the learning rate α
- [ ] Try Stochastic gradient descent
- [ ] Try initializing all the weights to zero
- [x] Try better random initialization for the weights

Q.9)

Complete the following sentence :-

Neural networks _____

Multiple choice

- ( ) Optimize a convex cost function
- ( ) Always output values between 0 and 1
- (•) Can be used for regression as classification.
- ( ) All of the above

Q.10)

What is back propagation?

Multiple choice

- it is another name given to the curvy function in the perceptron
- it is the transmission of error back through the network to allow weights to be adjus... ✓
- it is the transmission of error back through the network to adjust the inputs
- All of the above

Q.11)

Complete the following sentence :-

The dev and test set should _____

Multiple choice

- Come from different distributions.
- Be identical to each other same (x,y) pairs.
- Have the same number of examples.
- Come from same distributions. ✓

Q.12)

A 3-input neuron is trained to output a zero when the input is 010 and a one when the input is 110. After generalization, the output will be zero when and only when the input is:

Multiple choice

- 010 or 100 or 110 or 101
- 000 or 010 or 110 or 100
- 100 or 111 or 101 or 001
- 000 or 010 or 011 or 001 ✓

Q.13)

A 4-input neuron has weights 5, 4, 3 and 2. The transfer function is linear with the constant of proportionality being equal to 3. The inputs are 4, 5, 6 and 10 respectively. The output will be:

Multiple choice

- 124
- 189
- 234 ✓
- 270

Q.14)

Why is the XOR problem exceptionally interesting to neural network researchers?

Multiple choice

- ( ) because it can be expressed in a way that allows you to use a neural network
- ( ) because it is complex binary operation that cannot be solved using neural networks
- ( ) because it can be solved by a single layer perceptron
- ( ) because it is the simplest linearly inseparable problem that exists ✓

Q.15)

An auto-associative network is:

Checkboxes

- [ ] a neural network that contains no loops
- [ ] a neural network that contains feedback ✓
- [ ] a neural network that contains 1 loop
- [ ] a single layer feed-forward neural network with pre-processing