Machine Learning (CS1741) (EL VII) Sessional I

Time: 30 Minutes
Max Marks: 15
Note: Answer all questions. Any missing / misprinted data may be suitably assumed

...

Points: 14/15

1. Adding more basis functions in a linear model... (pick the most probably option) (1/1 Points)

   - (●) Decreases model bias ✓
   - ( ) Decreases estimation bias
   - ( ) Decreases variance
   - ( ) Doesn't affect bias and variance

2. The polynomial curve fitting often suffers from under-fitting problem. (1/1 Points)

   - ( ) True
   - (●) False ✓

3. Suppose, you applied a Linear Regression model on a given data and got a training accuracy X and testing accuracy Y. Now, you want to add a few new features in the same data. Select the option(s) which is correct in such a case. (2/2 Points)

- ○ Training accuracy increases
- ○ Testing accuracy increases or remains the same
- ◉ Both are correct ✓
- ○ None of these

4. The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are (2/2 Points)

- ◉ 50 ✓
- ○ *Less than* 50
- ○ *More than* 50
- ○ *It is an arbitrary value*

✕

5. Which of the following statements about regularization is not correct? (0/1 Points)

- ◉ Using too large a value of lambda can cause your hypothesis to underfit the data.
- ○ Using too large a value of lambda can cause your hypothesis to overfit the data.
- ○ Using a very large value of lambda cannot hurt the performance of your hypothesis.
- ○ None of the above ✓

6. The perceptron learning suffers from classifying non-linear data sets and it can't handle _____.
(1/1 Points)

- Overfitting
- Underfitting
- Noise ✓
- Outliers

7. Which of the following is a good test dataset characteristic?
(1/1 Points)

- Large enough to yield meaningful results
- Is representative of the dataset as a whole
- Both are correct ✓
- None of the above

8. Supervised learning differs from unsupervised learning in that supervised learning requires _____.
(1/1 Points)

- Labelled data
- Training data
- Both are correct ✓
- None of the above

9. Suppose you are training a multilayer neural network for binary classification. What type of error function which is suitable to solve the problem. (1/1 Points)

- $\sum_{i=1}^{N} [t_i - y(x_i, W)]$
- $\sum_{i=1}^{N} |t_i - y(x_i, W)|$
- $\frac{1}{2} \sum_{i=1}^{N} [t_i - y(x_i, W)]^2$ ✓
- *None of these*

10. Which of the following sentence is FALSE regarding regression? (1/1 Points)

- It relates inputs to outputs.
- It is used for prediction.
- It may be used for interpretation.
- It discovers causal relationships. ✓

11. (2/2 Points)

The log *isticfunction is given by* $f(x) = \frac{1}{1+e^{-x}}$ *Which of the following is correct*?

- $f'(x) = f(x)(1 - f(x))$ ✓
- $f'(x) = f(x)(1 + f(x))$
- $f'(x) = f(x)(1 - f(x))^2$
- *None of these*

12. To find the minimum or the maximum of a function, we set the gradient to zero because:
(1/1 Points)

- (●) The value of the gradient at extrema of a function is always zero ✓
- ( ) Depends on the type of problem
- ( ) Both are correct
- ( ) None of these