Uncategorized

NPTEL – Introduction to Machine Learning Assignment 2 Answers

November 15, 2020 admin@nptelsolutions.store

Assignment 2
Introduction to Machine Learning
Prof. B. Ravindran

(a) This feature has a strong effect on the model (should be retained)
(b) This feature does not have a strong effect on the model (should be ignored)

(c) It is not possible to comment on the importance of this feature without additional information

Sol. (c)
A high magnitude suggests that the feature is important. However, it may be the case that

another feature is highly correlated with this feature and it's coefficient also has a high magnitude with the opposite sign, in effect cancelling out the effect of the former. Thus, we cannot

really remark on the importance of a feature just because it's coefficient has a relatively large magnitude.

2. We have seen methods like ridge and lasso to reduce variance among the co-efficients. We can use these methods to do feature selection also. Which one of them is more appropriate?
   (a) Ridge
   (b) Lasso
   Sol. (b)
   For feature selection, we would prefer to use lasso since solving the optimisation problem when using lasso will cause some of the coefficients to be exactly zero (depending of course on the data) whereas with ridge regression, the magnitude of the coefficients will be reduced, but won't go down to zero.
3. Given a set of n data points, (x1, y1),(x2, y2), . . . ,(xn, yn), the best least squares fit f(x) is obtained by minimization of:
   (a) Pn
   i=1[yi – f(xi)]
   (b) min(yi – f(xi))
   (c) Pn
   i=1[yi – f(xi)]2
   (d) max(yi – f(xi))
   Sol. (c)
4. During linear regression, with regards to residuals, which among the following is true?
   (a) Lower is better
   (b) Higher is better
   (c) Depends upon the data
   (d) None of the above

1

Sol. (a)
Residuals refer to the errors in the model, hence, lower is better.

5. In the lecture on Multivariate Regression, you learn about using orthogonalization iteratively to obtain regression co-effecients. This method is generally referred to as Multiple Regression using Successive Orthogonalization.
   In the formulation of the method, we observe that in iteration k, we regress the entire dataset on z0, z1, . . . zk–1. It seems like a waste of computation to recompute the coefficients for z0 a

(b) Yes. Since $zj-1$ is orthogonal to $zj-l$ $\forall$ $l \le j1$, the multiple regression in each iteration is essentially a univariate regression on each of the previous residuals. Since the regression coefficients for the previous residuals don't change over iterations, we can re-use the coefficients for further iterations.
Sol. (b)

The answer is self-explanatory. Please refer to the section on Multiple Regression using Successive Orthogonalization in Elements of Statistical Learning, 2nd edition for the algorithm.

6. You decide to reduce the dimensionality of your data($N \times p$) using Best Subset Selection. The library you're using has a function regress(X, Y) that takes in X and Y and regresses Y on X. What is the expected number of times regress(·, ·) will be called during your dimensionality reduction?
   (a) O(2N)
   (b) O(2p
)
   (c) O(Np
)
   (d) O(p
   2
)
   Sol. (b)
   In Best subset selection, each possible subset of features is regressed and the number of possible subsets follow O(2p
).
7. If the number of features is larger than the number of training data points, to identify a suitable subset of the features for use with linear regression, we would prefer
   (a) Forward stepwise selection
   (b) Backward stepwise selection
   Sol. (a)
   Explanation: Recall that in backward stepwise selection, we need to first build a model using all features. This can lead to problems in matrix inversion when the number of training data points is less than the number of features. Thus, in this scenario we would prefer forward stepwise selection.
8. Assume you have a five-dimensional input data for a three-class classification problem. Further assume that all five dimensions of the input are independent to each other. In this scenario, is it possible for linear regression using lasso to result in one or more coefficients to become zero?

2

(a) Yes
(b) No
Sol. (a)
Note that even if the input dimensions are independent, one or more dimensions may not be very useful in discriminating between examples of the different classes. In such cases, the coefficients for such features may become zero on using lasso.

9. You are given the following five three-dimensional training data instances (along with one-

- x1 = 3, x2 = 8, x3 = 1, y = 2
- x1 = 7, x2 = 7, x3 = 2, y = 3
- x1 = 1, x2 = 9, x3 = 7, y = 8

Using the K-nearest neighbour technique for performing regression, what will be the predicted y value corresponding to the query point (x1 = 5, x2 = 3, x3 = 4), for K = 2?
(a) 3
(b) 2.5
(c) 3.5
(d) 2
Sol. (c)
When K = 2, the nearest points are x1 = 5, x2 = 7, x3 = 3 and x1 = 7, x2 = 7, x3 = 2. Taking the average of the outputs of these two points, we have y = (4 + 3)/2 = 3.5.

10. For the dataset given in the previous question, what will be the predicted y value corresponding to the query point (x1 = 5, x2 = 3, x3 = 4), for K = 3?
    (a) 4.66
    (b) 5
    (c) 3
    (d) 3.5
    Sol. (b)
    Similarly, when K = 3, nearest points are x1 = 5, x2 = 7, x3 = 3, x1 = 7, x2 = 7, x3 = 2 and x1 = 2, x2 = 4, x3 = 9. Taking the average of the outputs of these three points, we have y = (4 + 8 + 3)/3 = 5

Related Posts

Uncategorized

Python for Data Science Non Proctored Exam set 2 Answers

September 17, 2021 admin@nptelsolutions.store

Uncategorized

NPTEL 2021: Artificial Intelligence Search Methods For Problem Solving Assignment 7 Answers

September 9, 2021 admin@nptelsolutions.store

← Previous:
NPTEL – Introduction to Machine Learning Assignment 1 Answers

Next: →
NPTEL – Introduction to Machine Learning Assignment 3 Answers

Leave a Reply

You must be logged in to post a comment.

Categories

- Uncategorized

Recent Posts

- Python for Data Science Non Proctored Exam set 2 Answers
- NPTEL 2021: Artificial Intelligence Search Methods For Problem Solving Assignment 7 Answers
- NPTEL 2021: Artificial Intelligence Search Methods For Problem Solving Assignment 6 Answers
- NPTEL 2021: Artificial Intelligence Search Methods For Problem Solving Assignment 5 Answers
- NPTEL 2021: Artificial Intelligence Search Methods For Problem Solving Assignment 4 Answers

Search ... Search

Recent Comments

- lumbha on hi

- September 2021
- August 2021
- April 2021
- March 2021
- February 2021
- January 2021
- December 2020
- November 2020
- October 2020

Meta

- Log in
- Entries feed
- Comments feed
- WordPress.org