Business analytics and data mining Modeling using R: Week-12

15 April 2023 12:54

Week 0
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10
Week 11
Week 12
Lecture 56 ARTIFICIAL NEURAL NETWORK PART-4
Lecture 57 ARTIFICIAL NEURAL NETWORK PART-5
Lecture 58 ARTIFICIAL NEURAL NETWORK PART-6
Lecture 59 DISCRIMINANT ANALYSIS
Lecture 60 DISCRIMINANT ANALYSIS PART-2

Name: Aakash

Assessment scores

Week 1 : Assignment 1:	**95.0**
Week 2 : Assignment 2:	**100.0**
Week 3 : Assignment 3:	**100.0**
Week 4 : Assignment 4:	**100.0**
Week 5 : Assignment 5:	**100.0**
Week 6 : Assignment 6:	**100.0**
Week 7 : Assignment 7:	**100.0**
Week 8: Assignment 8:	**100.0**
Week 9 : Assignment 9:	**100.0**
Week 10 : Assignment 10:	**100.0**
Week 11 : Assignment 11:	**100.0**
Week 12 : Assignment 12:	-

Announcement:
You are currently receiving course related emails. Click here to unsubscribe.

1) Which of the following data mining tasks should not be conducted using discriminant analysis?

- Prediction
- Classification
- Clustering
- None of the above

(A)

Discriminant analysis is a technique used for classification, which involves assigning a categorical label to each observation based on its characteristics. However, prediction involves estimating a numerical value for a response variable based on its relationship with one or more predictor variables. Discriminant analysis is not suitable for prediction tasks because it is designed to classify observations into pre-defined categories rather than estimate numerical values.

Therefore, the data mining task that should not be conducted using discriminant analysis is prediction.

2) Which of the following is true about linear classification functions used in discriminant analysis?

- Provide the basis for discrimination of records into classes
- Linear functions of predictors that maximize ratio of between-class variability to within-class variability
- Coefficients of linear discriminant are optimized w.r.t class separation
- None of the above

(A,B,C)

Linear classification functions used in discriminant analysis are constructed in such a way that they provide the best separation between different classes. These functions are linear combinations of predictor variables that provide the basis for discriminating records into classes. They are constructed to maximize the ratio of between-class variability to within-class variability, which means that they provide the best separation between different classes. The coefficients of the linear discriminant function are optimized with respect to class separation to maximize the separation between different classes.

Therefore, options A, B, and C are all correct answers.

3) Which of the following plot can be helpful in assessing class separation for discriminant analysis?

- Histogram
- Scatter plot
- Bar chart
- None of the above

(B)

B) Scatter plot can be helpful in assessing class separation for discriminant analysis.

A scatter plot can show the relationship between two predictor variables and can also be used to visualize how the classes are distributed in the feature space. If the classes are well-separated, then it is more likely that the linear discriminant analysis will be able to find a good boundary between them. In contrast, if the classes are overlapping, then linear discriminant analysis may not be an appropriate technique to use.

Histograms and bar charts are useful for visualizing the distribution of individual variables, but they do not show the relationship between two or more variables, which is important for discriminant analysis.

Therefore, option B is the correct answer.

4) What is the maximum number of needed discriminant functions when m classes are present?

- m
- m-1
- m/2
- None of the above

(B)

The maximum number of needed discriminant functions when m classes are present is B) m-1.

In discriminant analysis, the number of discriminant functions that are needed depends on the number of classes and the number of predictor variables. The maximum number of discriminant functions that can be created is equal to the smaller of the number of classes and the number of predictor variables.

In the case of m classes, the maximum number of discriminant functions that can be created is m-1. This is because the last discriminant function is redundant and can be calculated from the others. The discriminant functions are calculated in such a way that each one provides the best separation between the classes that is not already accounted for by the previous discriminant functions.

Therefore, option B is the correct answer.

5) Which of the following is true assumption about correlation structure between predictors in discriminant analysis?

- Different for each class
- Same for each class
- Does not matter
- None of the above

(B)

The correct answer is B) Same for each class.

Discriminant analysis assumes that the correlation structure between the predictor variables is the same for each class. This means that the relationships between the predictor variables are similar for all the classes, and that the same linear combination of predictor variables can be used to classify observations into different classes.

If the correlation structure is different for each class, then discriminant analysis may not be an appropriate technique to use, as it assumes that the relationships between predictor variables are the same across classes. In such cases, other techniques such as tree-based methods or neural networks may be more suitable.

Therefore, option B is the correct answer.

6) Which of the following are true about discriminant analysis and linear regression?

- ○ Same estimation technique
- ○ Coefficients are optimized using same mechanism
- ○ Different estimation technique
- ○ None of the above

(A)

7) Which of the following updating mechanisms yields more accurate results in neural networks?

- ○ Batch updating
- ○ Both a and b
- ○ Case updating
- ○ None of the above

(C)

C) Case updating yields more accurate results in neural networks.

In neural networks, the updating mechanism refers to how the weights and biases are adjusted during the training process to minimize the error between the predicted and actual outputs. The two main updating mechanisms are batch updating and case updating.

Batch updating involves computing the errors for all the training examples in the dataset and then adjusting the weights and biases based on the average error across all examples. This can be computationally efficient, but may result in slower convergence and less accurate results.

Case updating, on the other hand, involves computing the error and updating the weights and biases for each training example one at a time. This can be more computationally expensive, but typically results in faster convergence and more accurate results.

Therefore, option C is the correct answer.

8) Which of the following is true about updating mechanisms in neural networks?

- [ ] Case updating is done after each case or record is run through the network.
- [ ] Batch updating is done after each case or record is run through the network.
- [ ] Batch updating is done after all records are run through the network.
- [ ] None of the above

(A,C)

A) Case updating involves updating the weights of the neural network after each individual training example is fed through the network. This means that the weights are updated more frequently, which can lead to faster convergence but may also increase the risk of overfitting.

C) Batch updating involves updating the weights of the neural network after all the training examples in a batch have been fed through the network. This means that the weights are updated less frequently, which can lead to slower convergence but may also reduce the risk of overfitting.

Both mechanisms have their advantages and disadvantages, and the choice of updating mechanism depends on the specific problem being solved and the characteristics of the dataset. Some common variations of these mechanisms include mini-batch updating, which involves updating the weights after a small batch of examples is processed, and online learning, which involves updating the weights after each example but only considering a single example at a time.

9) What is the basic advantage of data normalization step?

- ○ Smaller values improve the model
- ○ Values falling in a smaller range improve the model
- ○ Computing performance is better
- ○ None of the above

(C)

Normalization of data helps to improve the computing performance of the machine learning models. This is because normalization transforms the data to a standard scale, which can help to reduce the variation in the data and make the computations more stable. Normalization can also help to reduce the range of the input values, which can make the computations faster and more efficient.

By normalizing the data, we can also reduce the chances of numerical overflow or underflow, which can occur when the input values are too large or too small. This can cause numerical instability and affect the accuracy of the model. Normalization can help to prevent this by ensuring that the input values are on a similar scale.

10) Which stopping criteria are typically used in the training of neural networks?

- [ ] Small incremental change in bias and weight values
- [ ] Rate of change of error function values reaches a required threshold
- [ ] Limit on no. of runs is reached
- [ ] None of the above

(A,B,C)

A) Small incremental change in bias and weight values

B) Rate of change of error function values reaches a required threshold

C) Limit on no. of runs is reached

Stopping criteria are used in the training of neural networks to determine when to stop the training process. The three most commonly used stopping criteria are:

A) Small incremental change in bias and weight values: The training is stopped when the change in the weights and biases of the neural network falls below a certain threshold, indicating that the network has converged to a solution.

B) Rate of change of error function values reaches a required threshold: The training is stopped when the rate of change of the error function falls below a certain threshold, indicating that the network is no longer improving significantly.

C) Limit on no. of runs is reached: The training is stopped after a certain number of iterations or epochs have been completed, regardless of the performance of the network.

The choice of stopping criteria depends on the specific problem being solved and the characteristics of the dataset. It is important to balance the need for training the network sufficiently with the risk of overfitting or underfitting the data.