Machine Learning (ML) solved MCQs

« Set 6 **7** of **31** Set 8 »

151. The number of iterations in apriori _____ Select one: a. b. c. d.

A. increases with the size of the data

B. decreases with the increase in size of the data

C. increases with the size of the maximum frequent set

D. decreases with increase in size of the maximum frequent set

C.increases with the size of the maximum frequent set discuss

152. Frequent item sets is

A. superset of only closed frequent item sets

B. superset of only maximal frequent item sets

C. subset of maximal frequent item sets

D. superset of both closed frequent item sets and maximal frequent item sets

D.superset of both closed frequent item sets and maximal frequent item sets discuss

C. high intra class similarity

D. no inter class similarity

C.high intra class similarity discuss

154. Which statement is true about neural network and linear regression models?

A. both techniques build models whose output is determined by a linear sum of weighted input attribute values

B. the output of both models is a categorical attribute value

C. both models require numeric attributes to range between 0 and 1

D. both models require input attributes to be numeric

D.both models require input attributes to be numeric discuss

155. Which Association Rule would you prefer

A. high support and medium confidence

B. high support and low confidence

C. low support and high confidence

D. low support and low confidence

C.low support and high confidence discuss

156. In a Rule based classifier, If there is a rule for each combination of attribute values, what do you called that rule set R

A. exhaustive

B. inclusive

C. comprehensive

D. mutually exclusive

A.exhaustive discuss

157. The apriori property means

A. if a set cannot pass a test, its supersets will also fail the same test

B. to decrease the efficiency, do level-wise generation of frequent item sets

C. to improve the efficiency, do level-wise generation of frequent item sets d.

D. if a set can pass a test, its supersets will fail the same test

A. if a set cannot pass a test, its supersets will also fail the same test discuss

158. If an item set 'XYZ' is a frequent item set, then all subsets of that frequent item set are

A. undefined

B. not frequent

C. frequent

D. can not say

C.frequent discuss

159. Clustering is _____ and is example of _____learning

A. predictive and supervised

B. predictive and unsupervised

C. descriptive and supervised

D. descriptive and unsupervised

D.descriptive and unsupervised

discuss

160. To determine association rules from frequent item sets

A. only minimum confidence needed

B. neither support not confidence needed

C. both minimum support and confidence are needed

D. minimum support is needed

C.both minimum support and confidence are needed

discuss

161. If {A,B,C,D} is a frequent itemset, candidate rules which is not possible is

A. c –> a

B. d –>abcd

C. a –> bc

D. b –> adc

B.d –>abcd

discuss

162. Which Association Rule would you prefer

A. high support and low confidence

B. low support and high confidence

C. low support and low confidence

D. high support and medium confidence

B.low support and high confidence

discuss

163. This clustering algorithm terminates when mean values computed for the current iteration of the algorithm are identical to the computed mean values for the previous iteration

A. conceptual clustering

B. k-means clustering

C. expectation maximization

D. agglomerative clustering

B.k-means clustering

discuss

164. Classification rules are extracted from _____

A. decision tree

B. root node

C. branches

D. siblings

A.decision tree

discuss

165. What does K refers in the K-Means algorithm which is a non-hierarchical clustering approach?

A. complexity

B. fixed value

C. no of iterations

D. number of clusters

D.number of clusters

discuss

166. How will you counter over-fitting in decision tree?

A. by pruning the longer rules

B. by creating new rules

C. both by pruning the longer rules' and ' by creating new rules'

D. none of the options

A.by pruning the longer rules

discuss

167. What are two steps of tree pruning work?

A. pessimistic pruning and optimistic pruning

B. postpruning and prepruning

C. cost complexity pruning and time complexity pruning

D. none of the options

B.postpruning and prepruning discuss

168. Which of the following sentences are true?

A. in pre-pruning a tree is \pruned\ by halting its construction early

B. a pruning set of class labelled tuples is used to estimate cost complexity

C. the best pruned tree is the one that minimizes the number of encodingbits

D. all of the above

D. all of the above discuss

169. Assume that you are given a data set and a neural network model trained on the data set. You are asked to build a decision tree model with the sole purpose of understanding/interpreting the built neural network model. In such a scenario, which among the following measures would you concentrate most on optimising?

A. accuracy of the decision tree model on the given data set

B. f1 measure of the decision tree model on the given data set

C. fidelity of the decision tree model, which is the fraction of instances on which the neuralnetwork and the decision tree give the same output

D. comprehensibility of the decision tree model, measured in terms of the size of the corresponding rule set

C. fidelity of the decision tree model, which is the fraction of instances on which the neuralnetwork and the decision tree give the same output discuss

170. Which of the following properties are characteristic of decision trees?
(a) High bias
(b) High variance
(c) Lack of smoothness of prediction surfaces
(d) Unbounded parameter set

A. a and b

B. a and d

C. b, c and d

D. all of the above

C.b, c and d discuss

171. To control the size of the tree, we need to control the number of regions. One approach to

(a) It would, in general, help restrict the size of the trees (b) It has the potential to affect the performance of the resultant regression/classification model
(c) It is computationally infeasible

A. a and b

B. a and d

C. b, c and d

D. all of the above

discuss

A.a and b

172. Which among the following statements best describes our approach to learning decision trees

A. identify the best partition of the input space and response per partition to minimise sumof squares error

B. identify the best approximation of the above by the greedy approach (to identifying thepartitions)

C. identify the model which gives the best performance using the greedy approximation(option (b)) with the smallest partition scheme

D. identify the model which gives performance close to the best greedy approximation performance (option (b)) with the smallest partition scheme

discuss

D.identify the model which gives performance close to the best greedy approximation performance (option (b)) with the smallest partition scheme

173. Having built a decision tree, we are using reduced error pruning to reduce the size of the tree. We select a node to collapse. For this particular node, on the left branch, there are 3 training data points with the following outputs: 5, 7, 9.6 and for the right branch, there are four training data points with the following outputs: 8.7, 9.8, 10.5, 11. What were the original responses for data points along the two branches (left & right respectively) and what is the new response after collapsing the node?

A. 10.8, 13.33, 14.48

B. 10.8, 13.33, 12.06

C. 7.2, 10, 8.8

D. 7.2, 10, 8.6

discuss

C.7.2, 10, 8.8

174. Suppose on performing reduced error pruning, we collapsed a node and observed an improvement in the prediction accuracy on the validation set. Which among the following statements
are possible in light of the performance improvement observed? (a) The collapsed node helped overcome the effect of one or more noise affected data points in the training set
(b) The validation set had one or more noise affected data points in the region corresponding to the collapsed node
(c) The validation set did not have any data points along at least one of the collapsed branches
(d) The validation set did have data points adversely affected by the collapsed node

A. a and b

B. a and d

C. b, c and d

D. all of the above

discuss

D.all of the above

175. Time Complexity of k-means is given by

A. o(mn)

B. o(tkn)

C. o(kn)

D. o(t2kn)