X

Machine Learning (ML) solved MCQs

« Set 3 4 of 31 Set 5 »

76. Suppose we train a hard-margin linear SVM on n > 100 data points in R2, yielding a hyperplane with exactly 2 support vectors. If we add one more data point and retrain the classifier, what is the maximum possible number of support vectors for the new hyperplane (assuming the n + 1 points are linearly separable)?

A. 2

B. 3

C. n

D. n+1

D.n+1 discuss

77. Let S1 and S2 be the set of support vectors and w1 and w2 be the learnt weight vectors for a linearly separable problem using hard and soft margin linear SVMs respectively. Which of the following are correct?

A. s1 ⊡ s2

B. s1 may not be a subset of s2

C. w1 = w2

D. all of the above

B.s1 may not be a subset of s2 discuss

78. Which statement about outliers is true?

A. outliers should be part of the training dataset but should not be present in the test data

B. outliers should be identified and removed from a dataset

C. the nature of the problem determines how outliers are used

79. If TP=9 FP=6 FN=26 TN=70 then Error rate will be

A. 45 percentage

B. 99 percentage

C. 28 percentage

D. 20 perentage

C.28 percentage

discuss

---

80. Imagine, you are solving a classification problems with highly imbalanced class. The majority class is observed 99% of times in the training data. Your model has 99% accuracy after taking the predictions on test data. Which of the following is true in such a case?
1. Accuracy metric is not a good idea for imbalanced class problems.
2.Accuracy metric is a good idea for imbalanced class problems.
3.Precision and recall metrics are good for imbalanced class problems.
4.Precision and recall metrics aren't good for imbalanced class problems.

A. 1 and 3

B. 1 and 4

C. 2 and 3

D. 2 and 4

A.1 and 3

discuss

81. he minimum time complexity for training an SVM is O(n2). According to this fact, what sizes of datasets are not best suited for SVM's?

A. large datasets

B. small datasets

C. medium sized datasets

D. size does not matter

A.large datasets

discuss

82. Perceptron Classifier is

A. unsupervised learning algorithm

B. semi-supervised learning algorithm

C. supervised learning algorithm

D. soft margin classifier

C.supervised learning algorithm

discuss

83. Type of dataset available in Supervised Learning is

A. unlabeled dataset

B. labeled dataset

C. csv file

D. excel file

B.labeled dataset

discuss

84. which among the following is the most appropriate kernel that can be used with SVM to separate the classes.

A. linear kernel

B. gaussian rbf kernel

C. polynomial kernel

D. option 1 and option 3

B.gaussian rbf kernel

discuss

85. The SVMs are less effective when

A. the data is linearly separable

B. the data is clean and ready to use

C. the data is noisy and contains overlapping points

D. option 1 and option 2

C.the data is noisy and contains overlapping points

discuss

86. Suppose you are using RBF kernel in SVM with high Gamma value. What does this signify?

A. the model would consider even far away points from hyperplane for modeling

B. the model would consider only the points close to the hyperplane for modeling

C. the model would not be affected by distance of points from hyperplane for modeling

D. opton 1 and option 2

B.the model would consider only the points close to the hyperplane for modeling

discuss

87. What is the precision value for following confusion matrix of binary classification?

A. 0.91

B. 0.09

C. 0.9

D. 0.95

B.0.09

discuss

88. Which of the following are components of generalization Error?

A. bias

B. vaiance

C. both of them

D. none of them

discuss

C.both of them

89. Which of the following is not a kernel method in SVM?

A. linear kernel

B. polynomial kernel

C. rbf kernel

D. nonlinear kernel

discuss

A.linear kernel

90. During the treatement of cancer patients , the doctor needs to be very careful about which patients need to be given chemotherapy.Which metric should we use in order to decide the patients who should given chemotherapy?

A. precision

B. recall

C. call

D. score

discuss

A.precision

91. Which one of the following is suitable? 1. When the hypothsis space is richer, overfitting is more likely. 2. when the feature space is larger , overfitting is more likely.

A. true false

A. true, false

B. false, true

C. true,true

92. Which of the following is a categorical data?

A. branch of bank

B. expenditure in rupees

C. prize of house

D. weight of a person

discuss

A.branch of bank

93. The soft margin SVM is more preferred than the hard-margin SVM when-

A. the data is linearly seperable

B. the data is noisy and contains overlapping points

C. the data is not noisy and linearly seperable

D. the data is noisy and linearly seperable

discuss

B.the data is noisy and contains overlapping points

94. In SVM which has quadratic kernel function of polynomial degree 2 that has slack variable C as one hyper paramenter. What would happen if we use very large value for C

A. we can still classify the data correctly for given setting of hyper parameter c

B. we can not classify the data correctly for given setting of hyper parameter c

C. we can not classify the data at all

D. data can be classified correctly without any impact of c

discuss

A.we can still classify the data correctly for given setting of hyper parameter c

95. In SVM, RBF kernel with appropriate parameters to perform binary classification where the data is non-linearly seperable. In this scenario

A. the decision boundry in the transformed feature space in non-linear

B. the decision boundry in the transformed feature space in linear

C. the decision boundry in the original feature space in not considered

D. the decision boundry in the original feature space in linear

B.the decision boundry in the transformed feature space in linear

discuss

B. 1 is false, 2 is true

C. 1 is true, 2 is true

D. 1 is false, 2 is false

C.1 is true, 2 is true

discuss

97. What is the Accuracy in percentage based on following confusion matrix of three class classification.
Confusion Matrix C=
[14 0 0]
[1 15 0]
[0 0 6]

A. 0.75

B. 0.97

C. 0.95

D. 0.85

B.0.97

discuss

98. Which of the following method is used for multiclass classification?

A. one vs rest

B. loocv

C. all vs one

D. one vs another

A.one vs rest

discuss

99. Based on survey , it was found that the probability that person like to watch serials is 0.25 and the probability that person like to watch netflix series is 0.43. Also the probability that person like to watch serials and netflix sereis is 0.12. what is the probability that a person doesn't like to watch either?

A. 0.32

B. 0.2

C. 0.44

D. 0.56

C.0.44

discuss

100. A machine learning problem involves four attributes plus a class. The attributes have 3, 2, 2, and 2 possible values each. The class has 3 possible values. How many maximum possible different examples are there?

A. 12

B. 24

C. 48

D. 72

discuss