



Advanced Machine Learning Challenge

Below is a scored review of your assessment. All questions are shown.



Correct Answer



Partially Correct



Incorrect Answer

81 Support vectors are the data points that lie farthest to the decision surface.

Your Answer [no answer given]

Correct Answer False

Explanation

Actually, This is a false statement.

82 Decision tree is also referred to as

Your Answer [no answer given]

Correct Answer Recursive partitioning

Explanation

Sorry, you have selected the wrong answer.

83

Why are SVMs fast?

Your Answer

[no answer given]

Correct Answer

Quadratic optimization (convex!)

They work in the dual, with relatively few points

The kernel trick

Explanation

Incorrect. The correct options are 1,2 and 3.

84

Hierarchical clustering algorithms suffers from the problem of convergence at local optima?

Your Answer

[no answer given]

Correct Answer

True

Explanation

No, that's not correct.

85

In Content-Based Filtering, we use a very popular technique in information retrieval named _____

Your Answer

[no answer given]

Correct Answer

TF-IDF

Explanation

None.

86

What do you mean by generalization error in terms of the SVM?

Your Answer

[no answer given]

Correct Answer

How accurately the SVM can predict outcomes for unseen data

Explanation

Sorry, you have selected the wrong answer.

87 Seasonal variation means the variation occurring within:

Your Answer [no answer given]

Correct Answer Parts of a year

Explanation

None.

88 What's the hypothesis of logistic regression?

Your Answer [no answer given]

Correct Answer to limit the cost function between 0 and 1

Explanation

Actually, This is a True statement.

89

Univariate analysis is the simplest form of analyzing data. "Uni" means "one", so in other words your data has only one variable.

Your Answer [no answer given]

Correct Answer True

Explanation

None.

90

Suppose you have trained an anomaly detection system for fraud detection, and your system flags anomalies when $p(x) < \epsilon$, and you find on the cross-validation set that it is missing many fraudulent transactions (i.e., failing to flag them as anomalies). What should you do?

Your Answer

[no answer given]

Correct AnswerIncrease ϵ **Explanation**

Sorry, you have selected the wrong answer.

91 The log likelihood is parallel to?

Your Answer [no answer given]

Correct Answer The F-test in OLS regression

Explanation

Sorry, you have selected the wrong answer.

92 Which of the following is finally produced by Hierarchical Clustering?

Your Answer [no answer given]

Correct Answer tree showing how close things are to each other

Explanation

Sorry, you have selected the wrong answer.

93

A typical example of Memory-based approach is User Neighbourhood-based CF.

Your Answer

[no answer given]

Correct Answer

True

Explanation

None.

94

What do you expect will happen with bias and variance as you increase the size of training data?

Your Answer

[no answer given]

Correct Answer

Bias increases and Variance decreases

Explanation

Sorry, you have selected the wrong answer.

95

Prosperity, Recession, and depression in a business is an example of

Your Answer

[no answer given]

Correct Answer

Cyclical Trend

Explanation

None.

96

What property in a model is bias-variance trade off?

Your Answer

[no answer given]

Correct Answer

The variance can be reduced by increasing the bias

Explanation

No, that's not correct.

97

Select the correct option about regression with L2 regularization. A. Ridge regression technique prevents coefficients from rising too high. B. As $\lambda \rightarrow \infty$, the impact of the penalty grows, and the ridge regression coefficient estimates will approach infinity.

Your Answer

[no answer given]

Correct Answer

Statement A is true, Statement B is false

Explanation

Incorrect!

98 Which of the following is correct use of cross validation?

Your Answer [no answer given]

Correct Answer All of the mentioned

Explanation

No, that's not correct.

99

What would be then consequences for the OLS estimator if heteroscedasticity is present in a regression model but ignored?

Your Answer

[no answer given]

Correct Answer

It will be inefficient

Explanation

Incorrect!

100

Which of the following statements is True related to K-NN algorithm?

Your Answer

[no answer given]

Correct Answer

K-NN is a non-parametric algorithm
It is also called a lazy learner algorithm
It is robust to the noisy training data

Explanation

Sorry, you have selected the wrong answer.

Summary

Prev Page

I'm done.

Software by

Version 11.2

[Privacy Policy](#). Assessment content is copyright 2023, AlmaBetter.



Advanced Machine Learning Challenge

Below is a scored review of your assessment. All questions are shown.



Correct Answer



Partially Correct



Incorrect Answer

61 Which of the following method is used for trainControl resampling?

Your Answer [no answer given]

Correct Answer repeatedcv

Explanation

No, that's not correct.

62

Finding good hyperparameter is a time consuming process So typically you should do it once in the beginning of the project and try to find best yperparameter do that you dont have to visit tuning them again.

Your Answer

[no answer given]

Correct Answer

False

Explanation

No, that's not correct.

63

Agglomerative approach is called as Top-Down Approach whereas Divisive Approach is called as Bottom-Up Approach.

Your Answer [no answer given]

Correct Answer False

Explanation

Sorry, you have selected the wrong answer.

64

"Convolutional Neural Networks can perform various types of transformation (rotations or scaling) in an input". Is the statement correct True or False?

Your Answer [no answer given]

Correct Answer False

Explanation

Sorry, you have selected the wrong answer.

65 Why ML Explainability is important?

Your Answer [no answer given]

Correct Answer All of the Above

Explanation

None.

66

Which of the following techniques perform similar operations as dropout in a neural network?

Your Answer

[no answer given]

Correct Answer

Bagging

Explanation

Sorry, you have selected the wrong answer.

67

Which of the following options is/are true for K-fold cross-validation? 1.)Increase in K will result in higher time required to cross validate the result. 2.)Higher values of K will result in higher confidence on the cross-validation result as compared to lower value of K. 3.)If K=N, then it is called Leave one out cross validation, where N is the number of observations.

Your Answer

[no answer given]

Correct Answer

1,2,3

Explanation

No, that's not correct.

68 PCA is mostly used for _____.

Your Answer [no answer given]

Correct Answer Unsupervised Learning

Explanation

Sorry, you have selected the wrong answer.

69

Which regularization is used to reduce the over fit problem?

Your Answer

[no answer given]

Correct Answer

Both

Explanation

No, that's not correct.

70 Which of the following will be true about k in k-NN in terms of Bias?

Your Answer [no answer given]

Correct Answer When you increase the k the bias will be increases

Explanation

Sorry, you have selected the wrong answer.

71

How conditional probability rewrite in language model? $P(B | A) = P(A, B) / P(A)$

Your Answer

[no answer given]

Correct Answer

$P(A, B) = P(A) P(B | A)$

Explanation

None.

72

Reason for stop word removal

Your Answer

[no answer given]

Correct Answer

Stop word slow down processing

Explanation

None.

73 What is a dendrogram?

Your Answer [no answer given]

Correct Answer A hierarchical structure

Explanation

Incorrect!

74

Which transformation is one of the most commonly used mathematical transformations in feature engineering?

Your Answer

[no answer given]

Correct Answer

Logarithmic Transformation

Explanation

None.

75

In SVM, the dimension of the hyperplane depends upon which one?

Your Answer

[no answer given]

Correct Answer

The number of features

Explanation

Sorry, you have selected the wrong answer.

76

Including relevant lagged values of the dependent variable on the right hand side of a regression equation could lead to which one of the following?

Your Answer [no answer given]

Correct Answer Biased but consistent coefficient estimates

Explanation

Sorry! This needs work.

77

Which of the following neural networks uses supervised learning? (A) Multilayer perceptron (B) Self organizing feature map (C) Hopfield network

Your Answer

[no answer given]

Correct Answer

(A) only

Explanation

Sorry, you have selected the wrong answer.

78 Full Form of LIME

Your Answer [no answer given]

Correct Answer Local Interpretable Model-Agnostic Explanations

Explanation

None.

79 For a large k value the k-nearest neighbor model becomes ____ and ____ .

Your Answer [no answer given]

Correct Answer Simple model, Underfit

Explanation

Sorry, you have selected the wrong answer.

80 Tree Interpreters is an example of

Your Answer [no answer given]

Correct Answer Model- Specific Approach

Explanation

None.

Page 4 of 5

Summary

Prev Page

Next Page

I'm done.

Software by

Version 11.2

[Privacy Policy](#). Assessment content is copyright 2023, AlmaBetter.



Advanced Machine Learning Challenge

Below is a scored review of your assessment. All questions are shown.



Correct Answer



Partially Correct



Incorrect Answer

- 41** What is the final resultant cluster size in Divisive algorithm, which is one of the hierarchical clustering approaches?

Your Answer [no answer given]

Correct Answer singleton

Explanation

Incorrect!

42 Which of the following metrics can be used for evaluating regression models?

Your Answer [no answer given]

Correct Answer RMSE / MSE / MAE
Adjusted R Squared
F Statistics
R Squared

Explanation

No, that's not correct.

43

Logistic regression is _____ when the observed outcome of dependent variable aren't ordered.

Your Answer

[no answer given]

Correct Answer

multinomial

Explanation

Sorry, you have selected the wrong answer.

44

In Random forest you can generate hundreds of trees (say T1, T2Tn) and then aggregate the results of these tree. Which of the following is true about individual(Tk) tree in Random Forest?

Your Answer

[no answer given]

Correct Answer

Individual tree is built on a subset of the features
Individual tree is built on a subset of observations

Explanation

Sorry, you have selected the wrong answer.

45

Which of the following statements about Naive Bayes are correct?

Your Answer

[no answer given]

Correct Answer

attributes are equally important.

attributes are statistically independent of one another given the class value.

attributes can be nominal or numeric

Explanation

Sorry, you have selected the wrong answer.

46

We should remove highly correlated variables before doing PCA.

Your Answer

[no answer given]

Correct Answer

No

Explanation

Sorry, you have selected the wrong answer. Because, PCA loads out all highly correlated variables on the same Principal Component(Eigenvector), not different ones.

47 LIME is an example of

Your Answer [no answer given]

Correct Answer Model-Agnostic Approach

Explanation

None.

48 What's the cost function of the logistic regression?

Your Answer [no answer given]

Correct Answer both (A) and (B)

Explanation

Sorry, you have selected the wrong answer.

49 Which one of the following statements is TRUE for a Decision Tree?

Your Answer [no answer given]

Correct Answer In a decision tree, the entropy of a node decreases as we go down a decision tree.

Explanation

Sorry, you have selected the wrong answer.

50 LDA is a

Your Answer [no answer given]

Correct Answer Generative probabilistic process

Explanation

None.

- 51** In Latent Dirichlet Allocation model for text classification purposes, what does alpha and beta hyperparameter represent?

Your Answer [no answer given]

Correct Answer Alpha: density of topics generated within documents, beta: density of terms generated within topics

Explanation

None.

52

Which of the following methods does not have closed form solution for its coefficients?

Your Answer

[no answer given]

Correct Answer

Lasso Regression

Explanation

No, that's not correct.

53 What is Machine Translation?

Your Answer [no answer given]

Correct Answer Converts one human language to another

Explanation

None.

54 What will happen if you don't rotate the components?

Your Answer [no answer given]

Correct Answer PCA will diminish

Explanation

Sorry, you have selected the wrong answer.

55 Which of the following is a disadvantage of decision trees?

Your Answer [no answer given]

Correct Answer Decision trees are prone to be overfit

Explanation

Sorry, you have selected the wrong answer.

56 Examples of Naive Bayes Algorithm is/are

Your Answer [no answer given]

Correct Answer Spam filtration
Sentimental analysis
Classifying articles

Explanation

No, that's not correct.

- 57** For Ridge Regression, if the regularization parameter is very high, which options are true?

Your Answer [no answer given]

Correct Answer Large coefficients are not penalized

Explanation

No, that's not correct.

58

In boosting trees, individual weak learners are independent of each other

Your Answer

[no answer given]

Correct Answer

False

Explanation

In boosting tree individual weak learners are not independent of each other because each tree correct the results of previous tree. Bagging and boosting both can be consider as improving the base learners results.

59

Logarithmic transformation helps to handle skewed data and after transformation, the distribution becomes more approximate to normal.

Your Answer

[no answer given]

Correct Answer

True

Explanation

None.

60

Clustering analysis is negatively affected by multicollinearity of features

Your Answer

[no answer given]

Correct Answer

True

Explanation

Sorry, you have selected the wrong answer.

Page 3 of 5

Summary

Prev Page

Next Page

I'm done.

Software by

Version 11.2

[Privacy Policy](#). Assessment content is copyright 2023, AlmaBetter.



Advanced Machine Learning Challenge

Below is a scored review of your assessment. All questions are shown.



Correct Answer



Partially Correct



Incorrect Answer

- 21** The method / metric which is NOT useful to determine the optimal number of clusters in unsupervised clustering algorithms is

Your Answer [no answer given]

Correct Answer Scree plot

Explanation

Sorry, you have selected the wrong answer.

22 Naive Bayes classifiers are a collection ____ of algorithms.

Your Answer [no answer given]

Correct Answer classification

Explanation

No, that's not correct.

23 Full Form of LDA

Your Answer [no answer given]

Correct Answer Latent Dirichlet Allocation

Explanation

None.

24

Misclassification would happen when you use very small C ($C \sim 0$)

Your Answer

[no answer given]

Correct Answer

True

Explanation

Actually, This is a true statement.

25 Which of the following is true about Manhattan distance?

Your Answer [no answer given]

Correct Answer It can be used for continuous variables

Explanation

Sorry, you have selected the wrong answer.

26

Every hyperparameter if set poorly, can have a huge impact on training and so all hyperparameters are about equally important to tune well.

Your Answer

[no answer given]

Correct Answer

False

Explanation

No, that's not correct.

27

Which approach uses the memory of previous users interactions to compute users similarities based on items they've interacted ?

Your Answer

[no answer given]

Correct Answer

Memory-based

Explanation

None.

28 Multiplicative model for time series is $Y = \dots$

Your Answer [no answer given]

Correct Answer $T^*S^*C^*I$

Explanation

None.

29 A rise in prices before Eid is an example of

Your Answer [no answer given]

Correct Answer Seasonal Trend

Explanation

None.

- 30** Which of the following methods do we use to find the best fit line for data in Linear Regression?

Your Answer [no answer given]

Correct Answer Least Square Error

Explanation

No, that's not correct.

31 What is back propagation?

Your Answer [no answer given]

Correct Answer It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn

Explanation

Sorry, you have selected the wrong answer.

32 What happens when model complexity increases?

Your Answer [no answer given]

Correct Answer Variance of the model increases

Explanation

oh, incorrect!

33

Which of the following is not an assumption of Linear Regression?

Your Answer

[no answer given]

Correct Answer

Multicollinearity

Explanation

Sorry, you have selected the wrong answer.

34

Which of the following statement is NOT TRUE about k-means?

Your Answer

[no answer given]

Correct Answer

Number of clusters to be built is typically an user input and it impacts the way clusters are created

Explanation

Sorry, you have selected the wrong answer.

35

Arrange the following steps for anomaly detection using K-means clustering.

1.Calculate the mean value of each cluster 2.If we observe ttry to set an initial threshold valuehat the “Distance” value is more than the “threshold” value, then we can conclude that it is an outlier. 3.Try to set an initial threshold value. 4.Find or Identify the cluster which is nearest to the test data point 5.When the testing process is going in that time try to determine the distance of each data point from the mean value

Your Answer [no answer given]

Correct Answer 1,3,5,4,2

Explanation

Sorry, you have selected the wrong answer.

36 What does decision node illustrates in a decision tree

Your Answer [no answer given]

Correct Answer Test specification

Explanation

Sorry, you have selected the wrong answer.

37

Which of the following coefficients is added as the penalty term to the loss function in Lasso regression?

Your Answer

[no answer given]

Correct Answer

Absolute value of magnitude

Explanation

No, that's not correct.

38 What is Decision Tree?

Your Answer [no answer given]

Correct Answer Flow-Chart & Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label

Explanation

Sorry, you have selected the wrong answer.

39 Which of the following are true about isolation forest?

Your Answer [no answer given]

Correct Answer

Identifies anomalies as the observations with short average path lengths
Isolation forest is built based on ensembles of decision trees.
Isolation forest needs an anomaly Score to have an idea of how anomalous a data point is
Splits the data points by randomly selecting a value between the maximum and the minimum of the selected features.

Explanation

Sorry, you have selected the wrong answer.

- 40** Which of the following algorithm doesn't uses learning Rate as one of its hyperparameter?

Your Answer [no answer given]

Correct Answer Extra Trees
Random Forest

Explanation

Sorry, you have selected the wrong answer.

Page 2 of 5

Summary

Prev Page

Next Page

I'm done.

Software by

Version 11.2

[Privacy Policy](#). Assessment content is copyright 2023, AlmaBetter.



Advanced Machine Learning Challenge

Below is a scored review of your assessment. All questions are shown.



Correct Answer



Partially Correct



Incorrect Answer

- 1 Jaccard distance is a metric for comparing two binary data strings. While comparing two binary strings of equal length, Jaccard distance is the number of bit positions in which the two bits are different.

Your Answer True

Correct Answer False

Explanation

Hamming distance is a metric for comparing two binary data strings. While comparing two binary strings of equal length, Hamming distance is the number of bit positions in which the two bits are different.

2 How do you choose the right node while constructing a decision tree?

Your Answer [no answer given]

Correct Answer An attribute having the highest information gain.

Explanation

Sorry, you have selected the wrong answer.

- 3** The topic of a corpus comprise a number of topics.

Your Answer [no answer given]

Correct Answer False

Explanation

None.

- 4** What are the possible constituent models of a hybrid recommender system (check all that apply)?

Your Answer [no answer given]

Correct Answer Collaborative Filtering
Content-Based Filtering

Explanation

None.

5 Which of the following can act as possible termination conditions in K-Means?

Your Answer [no answer given]

Correct Answer

For a fixed number of iterations.
Assignment of observations to clusters does not change between iterations. Except for cases with a bad local minimum.
Centroids do not change between successive iterations.
Terminate when RSS falls below a threshold.

Explanation

Sorry, you have selected the wrong answer.

- 6** A procedure used for finding the equation of a straight line which provides the best approximation for the relationship between the independent and dependent variables is

Your Answer [no answer given]

Correct Answer the least squares method

Explanation

You selected the wrong option.

- 7 Regarding bias and variance, which of the following statements are true? (Here 'high' and 'low' are relative to the ideal model.)

Your Answer [no answer given]

Correct Answer Models which overfit have a low bias.
Models which underfit have a low variance

Explanation

Sorry, you have selected the wrong answer.

- 8** Logistic Regression is a _____ regression technique that is used to model data having a _____ outcome.

Your Answer [no answer given]

Correct Answer Non-linear , binary

Explanation

Incorrect. Probablity cann't be nagative. Therefore, the Right answers are 1,2,4.

- 9** In natural language understanding (NLU) tasks, there is a hierarchy of lenses through which we can extract meaning — from words to sentences to paragraphs to documents.

Your Answer [no answer given]

Correct Answer True

Explanation

None.

- 10** Q9. Which of the following clustering algorithms suffers from the problem of convergence at local optima?

Your Answer [no answer given]

Correct Answer
K- Means clustering algorithm
Expectation-Maximization clustering algorithm

Explanation

Sorry, you have selected the wrong answer.

- 11** The curse of dimensionality refers to all the problems that arise working with data in the higher dimensions.

Your Answer [no answer given]

Correct Answer True

Explanation

Actually, this is a true statement.

12 The Naive Bayes Classifier is a ____ in probability.

Your Answer [no answer given]

Correct Answer Classification

Explanation

Sorry! This needs work.

13

Would reducing the dimensions by doing PCA affect the anomalies in a dataset?
Would it lead to the disappearance of the anomalies?

Your Answer [no answer given]

Correct Answer Partially

Explanation

Sorry, you have selected the wrong answer.

14

Binning is the process of transforming numerical variables into categorical counterparts.

Your Answer

[no answer given]

Correct Answer

True

Explanation

None.

15 Spam email detection comes under which domain?

Your Answer [no answer given]

Correct Answer Text Classification

Explanation

None.

16 Pick the true statements.**Your Answer**

[no answer given]

Correct Answer

Outlier are observations that are distant from the mean or location of a distribution

Outliers don't necessarily represent abnormal behavior or behavior generated by a different process

Anomalies are data patterns that are generated by different processes.

Explanation

Sorry, you have selected the wrong answer.

17

What are the steps for using a gradient descent algorithm? Calculate error between the actual value and the predicted value Reiterate until you find the best weights of network Pass an input through the network and get values from output layer Initialize random weight and bias Go to each neurons which contributes to the error and change its respective values to reduce the error

Your Answer

[no answer given]

Correct Answer

4, 3, 1, 5, 2

Explanation

Sorry, you have selected the wrong answer.

18 Which of the following is the second goal of PCA?

Your Answer [no answer given]

Correct Answer Data Compression

Explanation

Sorry, you have selected the wrong answer.

19

In a naive Bayes algorithm, when an attribute value in the testing record has no example in the training set, then the entire posterior probability will be zero.

Your Answer

[no answer given]

Correct Answer

True

Explanation

Sorry! This needs work.

- 20** Which of the following is true when you choose fraction of observations for building the base learners in tree based algorithm?

Your Answer [no answer given]

Correct Answer Decrease the fraction of samples to build a base learners will result in decrease in variance

Explanation

Sorry, you have selected the wrong answer.

Page 1 of 5

Summary

Next Page

I'm done.

Software by

Version 11.2

[Privacy Policy](#). Assessment content is copyright 2023, AlmaBetter.