Machine Learning (ML) solved MCQs

676. A supervised scenario is characterized by the concept of a _____.

A. Programmer

B. Teacher

C. Author

D. Farmer

B.Teacher discuss

677. overlearning causes due to an excessive _____.

A. Capacity

B. Regression

C. Reinforcement

D. Accuracy

A.Capacity discuss

C. None of the above

C.None of the above

discuss

679. _____ provides some built-in datasets that can be used for testing purposes.

A. scikit-learn

B. classification

C. regression

D. None of the above

A.scikit-learn

discuss

680. While using _____ all labels areturned into sequential numbers.

A. LabelEncoder class

B. LabelBinarizer class

C. DictVectorizer

D. FeatureHasher

A.LabelEncoder class

discuss

681. _____produce sparse matrices of real numbers that can be fed into any machine learning model.

A. DictVectorizer

B. FeatureHasher

C. Both A & B

D. None of the Mentioned

C.Both A & B

discuss

682. scikit-learn offers the class_____, which is responsible for filling the holes using a strategy based on the mean, median, or frequency

A. LabelEncoder

B. LabelBinarizer

C. DictVectorizer

D. Imputer

D.Imputer

discuss

683. Which of the following scale data by removing elements that don't belong to a given range or by considering a maximum absolute value.

A. MinMaxScaler

B. MaxAbsScaler

C. Both A & B

D. None of the Mentioned

C.Both A & B

discuss

684. scikit-learn also provides a class for per-sample normalization,_____

A. Normalizer

B. Imputer

C. Classifier

D. All above

A. Normalizer

discuss

685. _____dataset with many features contains information proportional to the independence of all features and their variance.

A. normalized

B. unnormalized

C. Both A & B

D. None of the Mentioned

B.unnormalized

discuss

686. In order to assess how much information is brought by each component, and the correlation among them, a useful tool is the_____.

A. Concuttent matrix

B. Convergance matrix

C. Supportive matrix

D. Covariance matrix

D. Covariance matrix

discuss

687. The_____ parameter can assume different values which determine how the data matrix is initially processed.

A. run

B. start

C. init

D. stop

C.init

discuss

688. _____allows exploiting the natural sparsity of data while extracting principal components.

A. SparsePCA

B. KernelPCA

C. SVD

D. init parameter

discuss

A.SparsePCA

689. Which of the following statement is true about outliers in Linear regression?

A. Linear regression is sensitive to outliers

B. Linear regression is not sensitive to outliers

C. Can't say

D. None of these

discuss

A.Linear regression is sensitive to outliers

690. Suppose you plotted a scatter plot between the residuals and predicted values in linear regression and you found that there is a relationship between them. Which of the following conclusion do you make about this situation?

A. Since the there is a relationship means our model is not good

B. Since the there is a relationship means our model is good

C. Can't say

D. None of these

discuss

A.Since the there is a relationship means our model is not good

B. You can not have test error zero

C. None of the above

C.None of the above

discuss

692. In a linear regression problem, we are using "R-squared" to measure goodness-of-fit. We add a feature in linear regression model and retrain the same model.Which of the following option is true?

A. If R Squared increases, this variable is significant.

B. If R Squared decreases, this variable is not significant.

C. Individually R squared cannot tell about variable importance. We can't say anything about it right now.

D. None of these.

C.Individually R squared cannot tell about variable importance. We can't say anything about it right now.

discuss

693. To test linear relationship of y(dependent) and x(independent) continuous variables, which of the following plot best suited?

A. Scatter plot

B. Barchart

C. Histograms

D. None of these

A.Scatter plot

discuss

694. which of the following step / assumption in regression modeling impacts the trade-off between under-fitting and over-fitting the most.

A. The polynomial degree

B. Whether we learn the weights by matrix inversion or gradient descent

C. The use of a constant-term

A.The polynomial degree

discuss

695. Which of the following is true about "Ridge" or "Lasso" regression methods in case of feature selection?

A. Ridge regression uses subset selection of features

B. Lasso regression uses subset selection of features

C. Both use subset selection of features

696. Which of the following statement(s) can be true post adding a variable in a linear regression model?1. R-Squared and Adjusted R-squared both increase2. R-Squared increases and Adjusted R-squared decreases3. R-Squared decreases and Adjusted R-squared decreases4. R-Squared decreases and Adjusted R-squared increases

A. 1 and 2

B. 1 and 3

C. 2 and 4

D. None of the above

A.1 and 2 discuss

697. What is/are true about kernel in SVM?1. Kernel function map low dimensional data to high dimensional space2. It's a similarity function

A. 1

B. 2

C. 1 and 2

D. None of these

C.1 and 2 discuss

698. Suppose you are building a SVM model on data X. The data X can be error prone which means that you should not trust any specific data point too much. Now think that you want to build a SVM model which has quadratic kernel function of polynomial degree 2 that uses Slack variable C as one of it's hyper parameter.What would happen when you use very small C (C~0)?

A. Misclassification would happen

B. Data will be correctly classified

C. Can't say

D. None of these

A.Misclassification would happen discuss

699. The cost parameter in the SVM means:

A. The number of cross-validations to be made

B. The kernel to be used

C. The tradeoff between misclassification and simplicity of the model

D. None of the above

C.The tradeoff between misclassification and simplicity of the model discuss

700. How do you handle missing or corrupted data in a dataset?

A. a. Drop missing rows or columns

B. b. Replace missing values with mean/median/mode