LINEAR REGRESSION ASSIGNMENT

21) When implementing linear regression of some dependent variable y on the set of independent variables $\mathbf{x} = (x_1, ..., x_r)$, where r is the number of predictors, which of the following statements will be true?
a) $\beta_0, \beta_1, ..., \beta_r$ are the regression coefficients.
b) Linear regression is about determining the best predicted weights by using the method of ordinary least squares.
c) E is the random interval
d) Both a and b

Answer D)- Both A and B

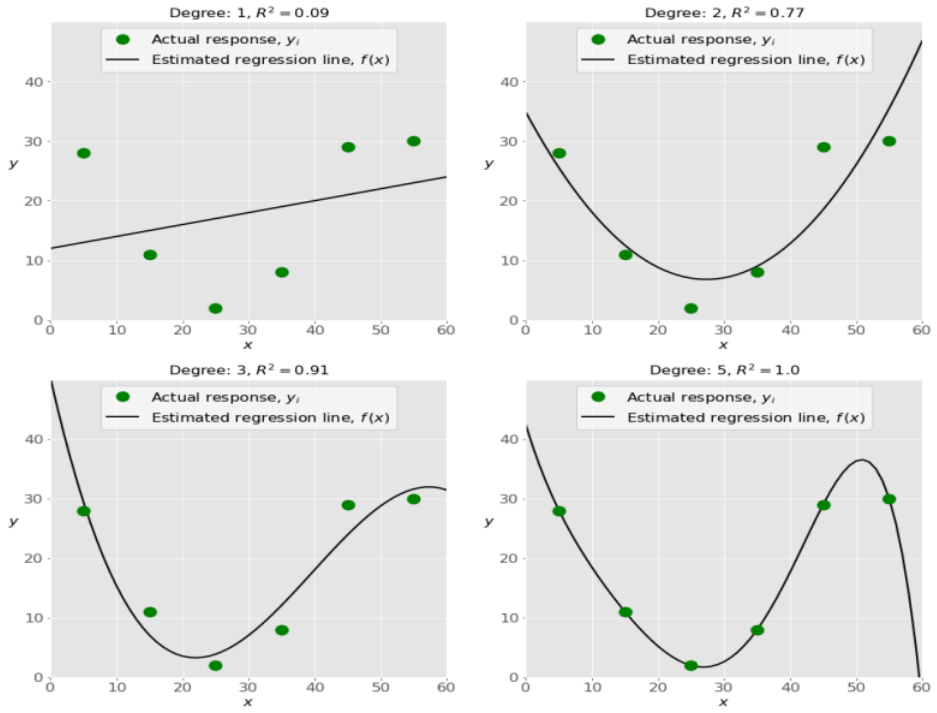
22) What indicates that you have a perfect fit in linear regression?
a) The value $R^2 < 1$, which corresponds to SSR = 0
b) The value $R^2 = 0$, which corresponds to SSR = 1
c) The value $R^2 > 0$, which corresponds to SSR = 1
d) The value $R^2 = 1$, which corresponds to SSR = 0

Answer- D) The value $R^2 = 1$, which corresponds to SSR = 0

23) In simple linear regression, the value of what shows the point where the estimated regression line crosses the y axis?
a) Y
b) B0
c) B1
d) F

Answer - B)- Bo

24) Check out these four linear regression plots: Which one represents an underfitted model?



a)The bottom-left plot
b) The top-right plot
c) The bottom-right plot
d) The top-left plot

Answer- A

25) There are five basic steps when you're implementing linear regression:
• a. Check the results of model fitting to know whether the model is satisfactory.
• b. Provide data to work with, and eventually do appropriate transformations.
• c. Apply the model for predictions.
• d. Import the packages and classes that you need.
• e. Create a regression model and fit it with existing data.
However, those steps are currently listed in the wrong order. What's the correct order?
a) e, c, a, b, d
b) e, d, b, a, c
c) d, e, c, b, a
d) d, b, e, a, c

Answer- D

26) Which of the following are optional parameters to LinearRegression in scikit-learn?
a) Fit
b) fit_intercept
c) normalize
d) copy_X
e) n_jobs

f) reshape

Answer -

The optional parameters to the LinearRegression class in scikit-learn are:

1. fit_intercept: This parameter determines whether to calculate the intercept for the linear regression model. By default, it is set to True.
2. normalize: This parameter determines whether to normalize the input features before fitting the model. By default, it is set to False.
3. copy_X: This parameter determines whether to make a copy of the input features. By default, it is set to True.
4. n_jobs: This parameter determines the number of parallel jobs to use for the computation. By default, it is set to None, which means it will use one job.
5. reshape: This parameter determines whether to reshape the target variable. By default, it is set to True.

27) While working with scikit-learn, in which type of regression do you need to transform the array of inputs to include nonlinear terms such as x^2?
a)Multiple linear regression
b) Simple linear regression
c) Polynomial regression

Answer - C)- Polynomial regression

28) You should choose statsmodels over scikit-learn when:
A)You want graphical representations of your data.
b) You're working with nonlinear terms.
c) You need more detailed results.
d) You need to include optional parameters

Answer - C) You need more detailed results.

29) _____ is a fundamental package for scientific computing with Python. It offers comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms, and more. It provides a high-level syntax that makes it accessible and productive.
a) Pandas
b) Numpy
c) Statsmodel
d) scipy

Answer - B) Numpy

30) _____ is a Python data visualization library based on Matplotlib. It provides a high-level
interface for drawing attractive and informative statistical graphics that allow you to explore and
understand your data. It integrates closely with pandas data structures.
a) Bokeh

b) Seaborn
c) Matplotlib
d) Dash

Anwer - B) Seaborn

41) Among the following identify the one in which dimensionality reduction reduces.
a) Performance
b) statistics
c) Entropy
d) Collinearity

Answer - D) Dimensionality reduction reduces collinearity.

42) Which of the following machine learning algorithm is based upon the idea of bagging?
a) Decision Tree
b) Random Forest
c) Classfication
d) SVM

Answer - B) Random forest is based on the idea of bagging

43) Choose a disadvantage of decision trees among the following.
a) Decision tree robust to outliers
b) Factor analysis
c) Decision Tree are prone to overfit
d) all of the above

Answer- D All of the above

44) What is the term known as on which the machine learning algorithms build a model based on sample data?
a) Data Training
b) Sample Data
c) Training data
d) None of the above

Answer - C) Training Data

45)
Which of the following machine learning techniques helps in detecting the outliers in data?
a) Clustering
b) Classification
c) Anamoly detection
d) All of the above

Answer -C) Anomaly detection

46)
Identify the incorrect numerical functions in the various function representation of machine learning.
a) Support Vector
b) Regression
c) Case based
d) Classification

Answer - C) Case based is the incorrect one.

47) Analysis of ML algorithm needs
a) Statistical learning theory
b) Computational learning theory
c) None of the above
d) Both a and b

Answer- D)- Both A and B

48) Identify the difficulties with the k-nearest neighbour algorithm.
a) Curse of dimensionality
b) Calculate the distance of test case for all training cases
c) Both a and b d) None

Answer C) Both A and B

49) The total types of the layer in radial basis function neural networks is _____
a) 1
b) 2
c) 3
d) 4

Answer-C) There are total 3 layers in radial basis function.

50 Which of the following is not a supervised learning
a) PCA
b) Naïve bayes
c) Linear regression
d) KMeans

Answer- A) PCA - Principal Component Analysis (PCA) is not predictive analysis tool.