21) When implementing linear regression of some dependent variable y on the set of independent variables $\mathbf{x} = (x_1, \ldots, x_r)$, where r is the number of predictors, which of the following statements will be true?

Answer = $\beta_0, \beta_1, \ldots, \beta_r$ are the regression coefficients and linear regression is about determining the best predicted weights by using the method of ordinary least squares.

22) What indicates that you have a perfect fit in linear regression?

Answer= The value $R^2 = 1$, which corresponds to SSR = 0 When you achieve this, you've hit the bullseye in terms of model fit!

23) In simple linear regression, the value of what shows the point where the estimated regression line crosses the y axis?

Answer= B0 would be the right answer because, In simple linear regression, the point where the estimated regression line crosses the y-axis is represented by the intercept, denoted as b_0 or β_0.

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

Answer= The top-left plot because we can say a linear function with a polynomial degree 1 is not sufficient to fit the training samples.

25) There are five basic steps when you're implementing linear regression:

Answer= Import the packages and classes that you need.

Provide data to work with, and eventually do appropriate transformations.

Create a regression model and fit it with existing data.

Check the results of model fitting to know whether the model is satisfactory.

Apply the model for predictions.

d, b, e, a, c

26) Which of the following are optional parameters to LinearRegression in scikit-learn?

Answer= fit_intercept , normalize, copy_X , n_jobs

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?

Answer= Polynomial regression

If you're specifically interested in polynomial regression, where you transform the input features to include polynomial terms (such as x^2, x^3, etc.), you can achieve this by using a non-linear regression

approach. Polynomial regression fits a polynomial function to the data, allowing it to capture curved or nonlinear patterns.

28) You should choose statsmodels over scikit-learn when:

Answer= 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. Statsmodels allows you to do all these 4 options given.

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.

Answer= NumPy (Numerical Python) is the fundamental package used for scientific computing in Python.

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.

Answer= Seaborn is the Python data visualization library you're referring to! It's built on top of Matplotlib and offers a high-level interface for creating attractive and informative statistical graphics.

41) Among the following identify the one in which dimensionality reduction reduces.

Answer= Collinearity

42) Which of the following machine learning algorithm is based upon the idea of bagging?

Answer= Random Forest

43) Choose a disadvantage of decision trees among the following.

Answer= Decision trees are prone to be overfit

44) What is the term known as on which the machine learning algorithms build a model based on sample data?

Answer= Training data

45) Which of the following machine learning techniques helps in detecting the outliers in data?

Answer= Anomaly detection

46) Identify the incorrect numerical functions in the various function representation of machine learning.

Answer= Support Vector b) Regression c) Case based.

All of them are incorrect numerical functions in the various function representation of machine learning.

47) Analysis of ML algorithm needs?

Answer=Both a and b.

48) Identify the difficulties with the k-nearest neighbour algorithm.

Answer=Both a and b

49) The total types of the layer in radial basis function neural networks is _____

Answer=3

50) Which of the following is not a supervised learning?

Answer= PCA.