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

Ans:

$\beta_0, \beta_1, \cdots, \beta_r$ are the regression coefficients.

2. What indicates that you have a **perfect fit** in linear regression?

Ans:

The value $R^2 = 1$, which corresponds to SSR = 0

3. In simple linear regression, the value of **what** shows the point where the estimated regression line crosses the y axis?

Ans:

B1

4. Check out these four linear regression plots:

5. There are five basic steps when you're implementing linear regression:

Ans:

a. **Import the packages and classes that you need.**
b. **Create a regression model and fit it with existing data.**
c. **Provide data to work with, and eventually do appropriate transformations.**
d. **Check the results of model fitting to know whether the model is satisfactory.**
e. **Apply the model for predictions.**

b) e, d, b, a, c

5. Which of the following are optional parameters to LinearRegression in scikit-learn?

Ans:

n_jobs

6. 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?

Ans:

Polynomial regression

7. You should choose statsmodels over scikit-learn when:

Ans:

You want graphical representations of your data.

8. _____ 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.

Ans:

Numpy

9. _____ 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.

Ans:

Seaborn

10.Among the following identify the one in which dimensionality reduction reduces.
Ans:
Collinearity

11.Which of the following machine learning algorithm is based upon the idea of bagging?
Ans:
Random Forest

12. Choose a disadvantage of decision trees among the following.

Ans:
Decision Tree are prone to overfit

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

Ans:

Data Training

14. Which of the following machine learning techniques helps in detecting the outliers in data?

Ans:
Anamoly detection

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

Ans:
Support Vector

16. Analysis of ML algorithm needs

Ans:

Both a and b

16. Identify the difficulties with the k-nearest neighbor algorithm.

Ans:
None

17. The total types of the layer in radial basis function neural networks is
Ans:

3

18. Which of the following is not a supervised learning

Ans:

KMeans