21) When implementing linear regression of some dependent variable 𝑦 on the set of independent

variables 𝐱 = (𝑥₁, …, 𝑥ᵣ), where 𝑟 is the number of predictors, which of the following statements will

be true?

a) 𝛽₀, 𝛽₁, …, 𝛽ᵣ 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

**Ans:-d) Both a and b**

22 )

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

a) The value 𝑅² < 1, which corresponds to SSR = 0

b) The value 𝑅² = 0, which corresponds to SSR = 1

c) The value 𝑅² > 0, which corresponds to SSR = 1

d) The value 𝑅² = 1, which corresponds to SSR = 0

**Ans:- d) The value R2 =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**)**

**Ans:-b) B0**

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

**Ans:-d) d, b, e, a, c**

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

**Ans:- b) fit\_intercept**

**d)copy\_X**

**e) 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 𝑥²?

a)Multiple linear regression

b) Simple linear regression

c) Polynomial regression

**Ans:-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.

**Ans:- 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

**Ans:- 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

**Ans:- b) Seaborn**

41)

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

a) Performance

b) statistics

c) Entropy

d) Collinearity

**Ans:- d) 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

**Ans:- b) Random Forest**

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

**Ans:- Decision Tree are prone to overfit**

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

**Ans:- 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

**Ans:- Anamoly 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

**Ans:- a) Support Vector**

**b) Case based**

47)

Analysis of ML algorithm needs

a) Statistical learning theory

b) Computational learning theory

c) None of the above

d) Both a and b

**Ans:- a) Statistical Learning theory**

**b) Computational learning theory**

48)

Identify the difficulties with the k-nearest neighbor algorithm.

a) Curse of dimensionality

b) Calculate the distance of test case for all training cases

c) Both a and b

d) None

**Ans:- 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) 450

**Ans:- 3**

Which of the following is not a supervised learning

a) PCA

b) Naïve bayes

c) Linear regression

d) KMeans

**Ans:- KMeans**