Machine learning

Ans 1.

a. Least square error
b. Linear regression is sensitive to outliers
c. Zero.
d. Correlation
e. Low bias and high variance
f. Predictive modal
g. Regularization
h. SMOTE or Cross validation
i. TPR and FPR
j. False
k.

Ans12. We don't have to choose the learning rate. It becomes slow when number of features is very large. We need to iterate.

Ans13. The meaning of the word regularization is "**the act of changing a situation or system so that it follows laws or rules**". That's what it does in the machine learning world as well. Regularization is a method that constrains or regularizes the weights.

Ans14. There are three main regularization techniques, namely: **Ridge Regression (L2 Norm) Lasso (L1 Norm) Dropout**.

PYTHON – WORKSHEET 1

Ans.1. %

2. 0

3.36

4.0

5.6

6. C) the finally block will be executed no matter if the try block raises an error or not.

7. It is used to raise an exception.

8. in defining a generator

9.none of these

10.yeild & raise

STATISTICS WORKSHEET-1

Ans1.

1.true

2. Central Limit Theorem

3. b) Modelling bounded count data

4. d) All of the mentioned

5. Poisson

6.false

7. Hypothesis.

8.0

9. c) Outliers cannot conform to the regression relationship.

10.The normal distribution is an arrangement of a data set in which most values cluster in the middle of the range and the rest taper off symmetrically toward either extreme.

11.

1.we can Use deletion methods to eliminate missing data. The deletion methods only work for certain datasets where participants have missing fields. ...

2.we can Use regression analysis to systematically eliminate data. ...

3.Data scientists can use data imputation techniques.

The imputation that we can use are as follows: -

Mean, Median and Mode

This is one of the most common methods of imputing values when dealing with missing data. In cases where there are a small number of missing observations, data scientists can calculate the mean or median of the existing observations.

12. A/B testing is a method of comparing two versions of a webpage or app against each other to determine which one performs better.

13. We can do this method for missing data. However, you will risk losing data points with valuable information. A better strategy would be to impute the missing values.

14. Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. There are two type of statistics :- Descriptive statistics describe what is going on in a population or data set. Inferential statistics, by contrast, allow scientists to take findings from a sample group and generalize them to a larger population. The two types of statistics have some important differences.