Machine Learning Assignment

1. A) Least Square Error

2. A) Linear regression is sensitive to outliers

3. B) Negative

4. C) Both of them

5. D) none of these

6. A) Descriptive model

7. D) Regularization

8. D) SMOTE

9. A) TPR and FPR

10. A) True

11. B) Apply PCA to project high dimensional data

12. A) We don't have to choose the learning rate & B) It becomes slow when number of features is very large

13. Explain the term regularization?

    Ans. It is one of most important concepts of machine learning to prevent the model from overfitting by adding extra information to it.

14. Which particular algorithms are used for regularization?

    Ans. 1. Ridge regression (L2) 2. Lasso regression (L1). 3. Dropout

15. Explain the term error present in linear regression equation?

    Ans. It is the average of squared error that occurred between the predicted values and actual values. By simple linear equation y=mx+b we can calculate MSE as: Let's y = actual values, y_i = predicted.

Statistics Assignment

1. A) True
2. A) Central Limit Theorem
3. B) Modeling bounded count data
4. D) All of the mentioned
5. C) Poisson
6. B) False
7. B) Hypothesis
8. A) 0
9. C)Outliers cannot conform to the regression relationship.

10: What do you understand by the term Normal Distribution?

Ans: Normal distribution is probability distribution that is symmetric about mean, which means showing data near to mean. It appears as a bell curve.

11:How do you handle missing data? What imputation techniques do you recommend?

Ans: We can use either isna() or isnull() method to detect missing values in data. We can get the total of missing values in each column with sum() or take the average with mean().

12: What is A/B testing?

Ans: It tests a control (A) version against a variant (B) version to measure which one is most successful based on your key metrics.

13. Is mean imputation of missing data acceptable practice?

Ans: It is considered a bad practice as it completely removes the accountability for feature correlation

14. What is linear regression in statistics?

Ans: 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.

15. What are the various branches of statistics?

Ans: Descriptive statics and inferential statistics

PYTHON ASSIGNMENT

1. C
2. A
3. C
4. B
5. B
6. B
7. A
8. C
9. A &C
10. A & B
11. # factorial of 5
    def fact (n):

    if(n==0 or n==1):

    return 1

    else:

    return n*fact(n-1)

    num=int(input("Enter Number:"))

    if(num<0):

    print("Factorial doesnot exist for negative number")

    else:

    print("Factorial is:", fact(num))

12. # program to find is prime or composite
    n= int(input("Enter any number: "))
    for x in range(2,n):
    if n % x==0:
    break
    if x+1==n:
    print(n, " is a prime number")
    else:
    print(n,"is a composite number")

13. # program to find is prime or composite
    n= int(input("Enter any number: "))
    for x in range(2,n):
    if n % x==0:

```
        break
if x+1==n:
   print(n, " is a prime number")
      else:
         print(n,"is a composite number")
```

14. # program to find is prime or composite

```
n= int(input("Enter any number: "))
for x in range(2,n):
   if n % x==0:
      break
if x+1==n:
   print(n, " is a prime number")
else:
   print(n,"is a composite number")
```

15. # python program to print the frequency of each of the characters present in a given string

```
str = "Mounika"
print(str.count("n"))
```