MACHINE LEARNING

1.

D

2.

A

3.

B

4.

C

5.

C

6.

B

7.

D

8.

D

9.

C

10.

B

11.

B

12.

B

13. Explain the term regularization?

Regularization is the practice of adding a penalty to model parameters (except the intercept) so the model generalises the data in order to prevent over-fitting.

14. Which particular algorithms are used for regularization?

Lasso also known as the Least Absolute Shrinkable Selection Operator or L1 is a regularization algorithm used to shrink the coefficients of the least important variables to zero therefore performing feature selection. Ridge regression or L2 is the second regularization algorithm that

attempts to correct collinearity by introducing a penalty to the slope of each predictor in the linear regression equation.

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

The error term in the linear regression equation is the summation of the squared difference between the predicted and actual values of the Y values.

PYTHON

1.

C

2. FLOOR OPERATOR //

B

3. LEFT SHIFT OPERATOR <<

C

4.

A

5.

D

6.

C

7.

A

8.

C

9.

A, C

10.

A, B

11. Write a python program to find the factorial of a number.

n = int(input("Enter a number"))

def factorial(n):

  if n == 0 | n == 1:

    return 1

  else:

```python
        return n * factorial(n-1)


print(factorial(n))
```

12. Write a python program to find whether a number is prime or composite.

```python
# I used the definition given by Wikipedia of primality test:given an input number, between 2 and
{\sqrt {n}} (i.e., whether the division leaves no remainder).

# If so, then n is composite. Otherwise, it is prime.

n = int(input("Enter a number"))


primenumber = 'It is a prime number'

compositenumber = 'It is a composite number'


def prime(n):
    if n < 2:
        return primenumber
    else:
        for i in range(2, int(n**0.5) + 1 ):
            if n % i == 0:
                return compositenumber
            else:
                return primenumber


print(prime(n))
```

13. Write a python program to check whether a given string is palindrome or not.

```python
import re


string = input("Enter a string ")


palindromicstring = 'It is a palindromic string'
```

```python
nonpalindromic = 'It is a not a palindromic string'


def palindromic(string):
    string = string.lower()

    if string == string[::-1]:
        return palindromicstring
    else:
        return nonpalindromic


print(palindromic(string))
```

14. Write a Python program to get the third side of right-angled triangle from two given sides.

```python
from math import sqrt


side1 = int(input("Enter a number: "))
side2 = int(input("Enter a number: "))


side1type = input("Write H if the side is the hypotenuse, write C if the side is a cathetus: ")
side2type = input("Write H if the side is the hypotenuse, write C if the side is a cathetus: ")


def sidelength(side1, side1type, side2, side2type):
    if side1type == 'C' and side2type == 'C':
        return sqrt(pow(side1, 2) + pow(side2, 2))

    elif side1type == 'H' and side2type == 'C':
        return sqrt(pow(side1, 2) - pow(side2, 2))

    else:
        return sqrt(pow(side2, 2) - pow(side1, 2))
```

print(sidelength(side1, side1type, side2, side2type))

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

string = input("Enter a string: ")

character = input("Enter a character: ")

def character_frequency(character, string):

   frequency = 0

  for i in string:

    if i == character:

      frequency += 1

  print(f"The character '{character}' appears {frequency} times in the string.")

character_frequency(character, string)

STATISTICS

1.

A

2.

A

3.

B

4.

A

5.

C

6.

B

7.

B

8.

A

9.

C

10.

A normal distribution is a symmetric Gaussian probability distribution where the mean is equal to the median and the mode. Moreover, the mean in a normal distribution is equal to zero and the variance and the standard deviation are both equal to one.

11.

I recommend two different types of imputations techniques to handle missing data. The first one is mean/median imputation which is effective when we want to preserve the mean of the observed data and to keep the full sample size. The second is linear interpolation which is optimal for generating values based on their estimated Y value in the same linear relationship present between the data.

12.

A/B testing is a controlled experiment that aims at determining which of two variants of a variable or more variables is optimal to the analytics results of a website marketing strategy.

13.

Mean imputation is not an acceptable practice when analysing the relationship between two variables as it tends to increase the correlation between the two variables.

14. What is linear regression in statistics?

Linear regression is a model that describes the relationship between one dependent variable (Y), and one or more independent variables, known as predictors or X. This model fits a simple linear or multilinear equation linking the independent variable to the predictors via coefficients that determine the slope of the line.

15. What are the various branches of statistics?

The various branches of statistics include descriptive statistics, inferential statistics, predictive statistics, and prescriptive statistics.