Worksheet 1

- Machine learning answers

1) B.

2)C.

3)C.

4)C.

5)B.

6)B.

8)C&D

9)B&D

10)A&C

11) Machine don’t understand text, it understands number so we need to convert categorical data into number so it is done by label encoder and one hot encoder,

When we have large dataset and having large number of unique values or categories then we should avoid one hot encoder.
The disadvantage of one hot encoder is that it increases no of columns, so datasets size is increased, so this effects in lowering model accuracy.

In case of this we can use label encoder as it doesn't affect the dimensionality of the dataset.

PYTHON – WORKSHEET 1

1)C.

2)B.

3)C.

4)A.

5)D.

6)C.

7)A.

8)C.

9)A&C

10)A&B

STATISTICS WORKSHEET-1

1)A

2)B

3)D

4)D

5)C

6)D

7)B

8)A

9)A

10)D

11)A

12)D

WORKSHEET -2

MACHINE LEARNING

1)A.

2)D.

3)D

4)A.

5)C.

6)C.

7)D

8)C.

9)A.

10)B.

WORKSHEET 2 PYTHON

1)B.

2)C.

3)A

4)A

5)A

6)C

7)B

8)B&D

9)A&C&D

10)A&B

11)

LIST

1) List can be represented by [ ]
2) List allows duplicate elements
3) List is mutable ,we can make any changes in list
4) List is ordered

TUPLE

1) Tuple can be represented by ()
2) Tuple allows duplicate elements
3) Tuple is immutable i.e we can not make any changes in tuple
4) Tuple is ordered

SET

1) Set will not allow duplicate elements
2) Set is mutable i.e we can make any changes in set
3) Set is unordered
4) Set can be represented by { }

DICTIONARY

1) Dictionary can be represented by { }
2) Dictionary doesn't allow duplicate keys.
3) Dictionary is mutable. But Keys are not duplicated.
4) Dictionary is ordered

12) Strings are not mutable in Python. Strings are a immutable data types which means that its value cannot be updated.

13) ORD() Function returns the number representing Unicode code of a specified character.
Example-

```
a='abhi is good person'
for i in a:
print(ord(i))
```

STATISTICS WORKSHEET

1)A

2)A

3)B

4)B

5)C

6)A

7)B

8)C

9)A

10)B

11)C

12)A

WORKSHEET MACHINE LEARNING-3

1)A

2)A

3)B

4)B

5)A

6)B

7)A

8)A

9)B&D

10)A&D

11)A&B&D

12)C&D

13)A&B

14) Linear regression is a machine learning algorithm. In linear regression the dependent variable is continuous in nature; it makes prediction for continuous/real variable.

There are two types of linear regression
1Simple linear regression
2) Multiple linear regression linear

The linear regression model creates a sloped straight line that represents the relationship between dependent variable and independent variable
1) Positive linear relationship - If dependent variable increase on y-axis and independent variable increases on x-axis then it is positive relationship
2) Negative linear relationship- If dependent variable decreases on y-axis and independent variable increases on x- axis then it is negative linear relationship.

When working with linear regression, main goal is to find best fit line that contains least error, means the error between predicted value and actual value should be minimum.

There are metrics to calculate model performance
Mean squared error
Mean absolute error
R2 square

R2 square measures the strength of relationship between dependent and independent variable, the high value of Rsquare represents a good model.

15)Simple linear regression - If a single independent variable is used to predict the value of an dependent variable, then it is called Simple Linear Regression

Multiple linear regression - If more than one independent variable is used to predict the value of a dependent variable, then it is called Multiple Linear Regression.

WORKSHEET 3 PYTHON –

1)D

2)C

3)A

4)A

5)C

6)D

7)B

8)B

9)A&B&C&D

10)A&C

STATISTICS WORKSHEET-3

1)A

2)A

3)A

4)B

5)A

6)C

7)C

8)D

9)A

10)C

11)B

12)C