Worksheet 1 Question and Answer

1.Which of the following operators is used to calculate remainder in a division?

C) %

2. In python 2//3 is equal to?

B) 0

3. In python, 6<<2 is equal to?

C) 24

4. In python, 6&2 will give which of the following as output?

D) 0

5. In python, 6|2 will give which of the following as output?

D) 6

6. What does the finally keyword denotes in python?

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

7. What does raise keyword is used for in python?

A) It is used to raise an exception.

8. Which of the following is a common use case of yield keyword in python?

C) in defining a generator

9. Which of the following are the valid variable names?

A) _abc and C) abc2

10. Which of the following are the keywords in python?

A) yield and B) raise

Machine Learning

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

2. Which of the following statement is true about outliers in linear regression?

3. A line falls from left to right if a slope is _____?

4. Which of the following will have symmetric relation between dependent variable and independent variable?

5. Which of the following is the reason for over fitting condition?

6. If output involves label then that model is called as:

7. Lasso and Ridge regression techniques belong to _____?

8. To overcome with imbalance dataset which technique can be used?

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

11. Pick the feature extraction from below:

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

13. Explain the term regularization?

14. Which particular algorithms are used for regularization?

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

Answer……

1. A) Least Square Error
2. A) Linear regression is sensitive to outliers
3. B) Negative
4. B) Correlation
5. C) Low bias and high variance
6. B) Predictive modal
7. D) Regularization
8. D) SMOTE
9. A) TPR and FPR (True Positive Rate and False Positive Rate)
10. B) False (For a better model, the area under the AUCROC curve should be more, not less.)
11. A) Construction bag of words from an email and
B) Apply PCA to project high-dimensional data

12. B) It becomes slow when the number of features is very large.
C) We need to iterate.
D) It does not make use of the dependent variable.

13. Regularization is a technique used in machine learning to prevent our models from overfitting. The Overfitting tab occurs when our model remembers the training data very well, but performs poorly with new data.
The function of regularization is to control the parameters of the model so that it does not become too complex. When we use regularization, we add an additional term to the loss function of our model. What term do we call regularization term?

14. *Ridge Regression (L2 Regularization): This algorithm is used to control the parameters, thereby reducing overfitting of the model.*

Lasso Regression (L1 Regularization): Lasso also controls parameters, but not like Ridge. Here, the values of some parameters become precisely zero, due to which feature selection also takes place.

Elastic Net Regression: Elastic Net is a mixture of Ridge (L2) and Lasso (L1), which has the benefits of both.

15. *"Error" in linear regression refers to the difference or gap that exists between the values predicted by our model and the actual observed values. We also call this "residual".*

The equation of linear regression in a simple form is as follows:

$Y=mx+b+\varepsilon$