

PYTHON WORKSHEET

1. C
2. B
3. C
4. A
5. D
6. C
7. A
8. C
9. A , C
10. A , B

MACHINE LEARNING (ASSIGNMENT - 39)

1. A
2. A
3. B
4. A
5. C
6. A
7. D
8. D
9. A
10. B
11. B
12. A , B
13. Regularization is used to avoid the model from overfitting by adding some extra information to it. It is one of the most important concept of machine learning. It regularizes learning from particular features and normalizes the feature so that algorithm do not depend on few features to predict the result. Some of the regularization techniques are Lasso(L1 form), Ridge(L2 form), Elastic-net(Combination of L1 and L2).
14. The particular algorithm that are used for regularization are Lasso Regression, Ridge Regression, Elastic-net Regression. 1. Lasso is a regression analysis method that performs both feature selection and regularization to increase the prediction accuracy of the model. 2. Ridge Regression is used for analyzing data which have multi-collinearity. 3. Elastic-net Regression is a method which is a combination of Ridge Regression and Lasso Regression.
15. The error term in linear regression equation is basically the difference between the predicted value by the model and the actual value given. It is also known as residual. Residual or error term present shows that the model does not fully represent the relationship between the independent variable and dependent variable.

STATISTICS WORKSHEET

1. A
2. A
3. B
4. D
5. C

- 6. B
- 7. B
- 8. A
- 9. C

10. Normal Distribution or Standard Normal Distribution is a probability distribution that is symmetric about the mean. Normal distribution is symmetrical distribution meaning that the graph of a normal distribution appears as a bell curve stating that the data near the mean is occurring more frequently as compared to data far from the mean. In a normal distribution the mean is 0 and standard deviation is 1.

11. The best method to handle missing data are:
- 1- Using deletion method to eliminate missing data.
 - 2- Using regression analysis to eliminate data.
 - 3- Using data imputation techniques to handle missing data.

12. A/B testing is also known as statistical hypothesis testing. We start the A/B testing process by making a hypothesis. It is a way to compare the two variables to find out which performs better. A/B testing is used for making decisions that estimate population parameters based on sample statistics. The population refers to all the data and sample refers to the data which is used for testing.

13. Mean imputation of missing data is a acceptable practice but it is not a good solution as most research studies give relationship among variables and using mean imputation technique will not make any sense and will have a negative impact on accuracy when we train our Machine Learning model. It only inserts the mean value while ignoring the feature correlations.

14. Linear Regression is a linear approach for modelling the relationship between explanatory variable and dependent variables. If there is only one variable or feature then the process is called simple linear regression, if there are more than one variable or features then it is known as multiple linear regression. A linear regression line has an equation of the form of ' $Y=mx + c$ ' where Y is the dependent variable, x is the dependent variable with m as the slope of the line and c as an intercept.

15. There are two branches of statistics,

1. Descriptive statistics- It is a part of statistics that deals with presenting the data we have. It has 2 ways to present the data i.e. either visually or numerically. Visual representation of data includes presenting data in the form of bars, graphs, plots, charts e.t.c. Numerical representation of data includes average of the data. The main focus of the descriptive statistics is to present the data in an understandable way so that if someone sees it they can easily understand the data.

2. Inferential statistics- It is a part off statistics that deals with making conclusions about the data i.e. what the data is about, what the data is telling us, what should we do and many more.