Machine Learning

1. Least Square error.
2. Linear regression is sensitive to outliers
3. Negative
4. Regression and correlation
5. None of these
6. Predictive modal
7. Removing outliers
8. Cross validation
9. Recall and precision
10. True
11. Apply PCA to project high dimensional data
12. A) We do not have to choose the learning rate. B) It becomes slow when number of features is very large. C) We need to iterate.
13. The word regularize means to make things regular or acceptable. This is exactly why we use it. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.
14. There are three main regularization techniques, namely: Ridge Regression (L2 Norm) Lasso (L1 Norm) Dropout.
15. Within a linear regression model tracking a stock's price over time, the error term is the difference between the expected price at a particular time and the price that was actually observed.

STATISTICS WORKSHEET-1

1. True
2. Central Limit Theorem
3. Modeling bounded count data
4. The exponent of a normally distributed random variables follows what is called the log-normal distribution
5. Poisson
6. True
7. Hypothesis
8. 0
9. Outliers cannot conform to the regression relationship
10. Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graphical form, the normal distribution appears as a "bell curve".
11. Deletions. Pairwise Deletion. Listwise Deletion/ Dropping rows. Dropping complete columns.
12. A/B testing (also known as split testing or bucket testing) is a method of comparing two versions of a webpage or app against each other to determine which one performs better.
13. Mean imputation is typically considered terrible practice since it ignores feature correlation.
14. 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. The variable you are using to predict the other variable's value is called the independent variable.

15. There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.

PYTHON – WORKSHEET 1

1. I think the right answer s (/).
2. 0
3. 24
4. 2
5. 0
6. the finally block will be executed no matter if the try block raises an error or not.
7. It is used to raise an exception
8. in defining a generator
9. abc2
10. Raise and yield