Q1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

A) Least Square Error
B) Maximum Likelihood
C) Logarithmic Loss
D) Both A and B

Ans1. A) Least Square Error.

Q2. Which of the following statement is true about outliers in linear regression?

A) Linear regression is sensitive to outliers
B) linear regression is not sensitive to outliers
C) Can't say
D) none of these

Ans2. A) Linear regression is sensitive to outliers.

Q3. A line falls from left to right if a slope is _____?

A) Positive
B) Negative
C) Zero
D) Undefined

Ans3. A) Positive

Q4. Which of the following will have symmetric relation between dependent variable and independent variable?
A) Regression
B) Correlation
C) Both of them
D) None of these

Ans4. B) Correlation

Q5. Which of the following is the reason for over fitting condition?

A) High bias and high variance
B) Low bias and low variance
C) Low bias and high variance
D) none of these

Ans5. A) High bias and high variance.

Q6. If output involves label then that model is called as:

A) Descriptive model
B) Predictive model
C) Reinforcement learning
D) All of the above

Ans6. B) Predictive model.

Q7. Lasso and Ridge regression techniques belong to _____?
A) Cross validation
B) Removing outliers
C) SMOTE
D) Regularization

Ans7. D) Regularization.

Q8. To overcome with imbalance dataset which technique can be used?

A) Cross validation
B) Regularization
C) Kernel
 D) SMOTE

Ans8. D) SMOTE

Q9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

A) TPR and FPR
B) Sensitivity and precision
C) Sensitivity and Specificity
 D) Recall and precision

Ans9. B) Sensitivity and precision

Q10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

A) True

 B) False

Ans10. A) True

Q11. Pick the feature extraction from below:

A) Construction bag of words from a email
B) Apply PCA to project high dimensional data
C) Removing stop words
D) Forward selection

Ans11. B) Apply PCA to project high dimensional data.

Q12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

A) We don't have to choose the learning rate.
B) It becomes slow when number of features is very large.
C) We need to iterate.
D) It does not make use of dependent variable.

Ans.12 A and B and C

Q13. Explain the term regularization?

Ans.13 It is one of the most important concepts of machine learning. This technique prevents the model from overfitting by adding extra information to it.

It is a form of regression that shrinks the coefficient estimates towards zero. In other words, this technique forces us not to learn a more complex or flexible model, to avoid the problem of overfitting.

Q14. Which particular algorithms are used for regularization?

Ans.14 Mainly, there are two types of regularization techniques, Ridge Regression, Lasso Regression

Q15. Explain the term error present in linear regression equation?

Ans.15 The error term is the difference between what the model is predicting and the actual value. This can range from being relatively small to huge.