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

Answer: A- Least square error

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

Answer: B- linear regression is not sensitive to outliers

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

Answer: B- negative

Q4. Which of the following will have symmetric relation between dependent variable and independent variable?

Answer: A: Regression

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

Answer: A- High bias and high variance

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

Answer: Descriptive model

Q7: Lasso and Ridge regression techniques belong to

Answer: D: Regularization

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

Answer: Synthetic Minority Oversampling Technique(SMOTE)

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

Answer: A) TPR and FPR

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

Answer: false

Q11. Pick the feature extraction from below:

A) Construction bag of words from a email

Q12.
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.

Q13. We avoid one aspect in machine learning model and that is to avoid overfitting as it will have low accuracy. The model in this case is trying too hard to capture the noise in your training dataset.
Noise means the data points that don't really represent the true properties of your data, but random chance. Learning such data points, makes your model more flexible, at the risk of overfitting.
The concept of balancing bias and variance is helpful in understanding the phenomenon of overfitting.

So, regularization is a form of regression, that reduces the coefficient estimates towards zero. In other words, this technique discourages learning a more complex or flexible model to avoid the risk of overfitting.

Q14.
There are three main regularization techniques, namely:
-Ridge Regression
-Lasso
-Dropout

Q15.

An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis.

The error term is also known as the residual or disturbance

Error term is also thought to be a difference between the expected price at a particular time and the price that was observed. Example: Where the price is exactly what was anticipated at a particular time, the price will fall on the trend line and the error term will be zero.