# MACHINE LEARNING

**In Q1 to Q11, only one option is correct, choose the correct option:**

1. 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
**Answer- D) Both A and B**

2. 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
**Answer- A) Linear regression is sensitive to outliers**

3. A line falls from left to right if a slope is _____?
A) Positive B) Negative
C) Zero D) Undefined
**Answer- B) Negative**

4. 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
**Answer- C) Both of them**

5. 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
**Answer- C) Low bias and high variance**

6. If output involves label then that model is called as:
A) Descriptive model B) Predictive model
C) Reinforcement learning D) All of the above
**Answer- B) Predictive model**

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

8. To overcome with imbalance dataset which technique can be used?
A) Cross validation B) Regularization
C) Kernel D) SMOTE
**Answer- D) SMOTE**

9. 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
**Answer- A) TPR and FPR**

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.
A) True B) False
**Answer- B)False**

11. 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
**Answer- A) Construction bag of words from a email**

**In Q12, more than one options are correct, choose all the correct options:**

12. 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
**Answer- D) It does not make use of dependent variable**

**Q13 and Q15 are subjective answer type questions, Answer them briefly.**

**13. Explain the term regularization?**
**Answer-** Regularization is a technique used in machine learning and statistics to prevent overfitting and improve the generalization of a model. Overfitting occurs when a model learns the training data too well, capturing noise and irrelevant patterns, which can lead to poor performance on unseen data.

Regularization works by adding a penalty term to the model's objective function, which penalizes complex models with large coefficients. The penalty term discourages the model from fitting the training data too closely, thus reducing overfitting.

There are several types of regularization techniques, including:

1. *L1 Regularization (Lasso)*: In L1 regularization, the penalty term is the absolute sum of the coefficients. It encourages sparsity in the model by shrinking some coefficients to zero, effectively performing feature selection.

2. *L2 Regularization (Ridge)*: In L2 regularization, the penalty term is the squared sum of the coefficients. It penalizes large coefficients more heavily than small ones, leading to smoother and more stable models.

3. *Elastic Net Regularization*: Elastic Net combines L1 and L2 regularization by adding both penalty terms to the objective function. It balances between feature selection (L1) and coefficient shrinkage (L2), offering a compromise between the two approaches.

Regularization helps to control the complexity of the model, reducing the risk of overfitting and improving its ability to generalize to new data. By tuning the regularization parameter, one can find the right balance between bias and variance to achieve optimal model performance**.**

**14. Which particular algorithms are used for regularization?**
**Answer-** Regularization techniques can be applied to various machine learning algorithms to prevent overfitting. Some of the commonly used algorithms that incorporate regularization are:

1. *Linear Regression with Ridge (L2) Regularization*: Ridge regression adds an L2 penalty term to the ordinary least squares (OLS) objective function, resulting in smoother coefficient estimates.

2. *Linear Regression with Lasso (L1) Regularization*: Lasso regression adds an L1 penalty term to the OLS objective function, promoting sparsity in the coefficient estimates and performing feature selection.

3. *Elastic Net Regression*: Elastic Net combines L1 and L2 regularization, adding both penalty terms to the objective function. It offers a balance between feature selection and coefficient shrinkage.

4. *Logistic Regression with Ridge or Lasso Regularization*: Similar to linear regression, logistic regression can be regularized using Ridge or Lasso techniques to prevent overfitting in binary classification problems.

5. *Support Vector Machines (SVM) with Regularization*: SVMs can be regularized using techniques like the C parameter, which controls the trade-off between maximizing the margin and minimizing the classification error.

6. *Neural Networks with Dropout*: Dropout is a form of regularization specific to neural networks, where randomly selected neurons are ignored during training. This prevents complex co-adaptations of neurons and reduces overfitting.

These are just a few examples of algorithms that can be regularized to improve their generalization performance. Regularization techniques are versatile and can be applied to a wide range of machine learning models to mitigate overfitting and improve model robustness.

**15. Explain the term error present in linear regression equation?**
**Answer-** In the context of linear regression, the "error" refers to the discrepancy between the observed values of the dependent variable and the values predicted by the linear regression model.

The linear regression equation typically takes the form:

$$ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \epsilon $$

where:
- $ Y $ is the dependent variable (the variable we want to predict).
- $ X_1, X_2, \ldots, X_p $ are the independent variables (predictor variables).
- $ \beta_0, \beta_1, \ldots, \beta_p $ are the coefficients (parameters) that represent the intercept and slopes of the linear equation.
- $ \epsilon $ is the error term, representing the variability in $ Y $ that is not explained by the linear relationship with the independent variables.

The error term captures the influence of all other factors that affect the dependent variable $ Y $ but are not accounted for by the linear regression model. These factors could include measurement error, unobserved variables, random fluctuations, and model misspecification. The error term is assumed to follow certain properties in linear regression, such as being normally distributed with a mean of zero and constant variance (homoscedasticity).

The goal of linear regression is to minimize the error term by estimating the coefficients $ \beta_0, \beta_1, \ldots, \beta_p $ that best fit the observed data. By minimizing the discrepancy between the observed values and the predicted values, the linear regression model aims to capture the underlying relationship between the independent and dependent variables.