**MACHINE LEARNING**

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**

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**

3. A line falls from left to right if a slope is \_\_\_\_\_\_?

A) Positive B) Negative

C) Zero D) Undefined

**ANSWER - B**

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 - B**

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**

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

A) Descriptive model B) Predictive modal

C) Reinforcement learning D) All of the above

**ANSWER - C**

7. Lasso and Ridge regression techniques belong to \_\_\_\_\_\_\_\_\_?

A) Cross validation B) Removing outliers

C) SMOTE D) Regularization

**ANSWER - D**

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

A) Cross validation B) Regularization

C) Kernel D) SMOTE

**ANSWER - D**

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**

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**

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 - B**

**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 - A & B**

**MACHINE LEARNING**

1. **Explain the term regularization?**

**ANSWER** - Regularisation is a machine learning approach that helps models perform better at generalisation and avoid overfitting. It entails regularising the goal function during model training by including a term or penalty. Regularisation aims to achieve a compromise between keeping the model from becoming overly complex and accurately fitting the training data.

High complexity models frequently capture noise or random oscillations in the training data, which results in subpar performance on fresh, unforeseen data. By creating a penalty for complex models and pushing them to prioritise simplicity and steer clear of overfitting, regularisation aids in the mitigation of this problem.

Elastic Net regularisation, L1 regularisation (Lasso), and L2 regularisation are a few examples of regularisation approaches. These strategies act by altering the objective function's learning process and affecting the model's parameter estimates.

By using the absolute values of the coefficients as the penalty term, L1 regularisation promotes sparse parameter values. It may successfully execute feature selection and reduce the number of predictors in the model by shrinking some coefficients to zero.

By include the squared values of the coefficients as the penalty term, L2 regularisation promotes tiny parameter values. The model becomes more resistant to multicollinearity and helps to mitigate the effects of big coefficients.

Elastic Net regularisation strikes a compromise between feature selection and coefficient shrinkage by combining both L1 and L2 penalties.

The problem at hand and the properties of the data determine the regularisation approach to use. In circumstances where the dataset is small, noisy, or contains a lot of features, regularisation is a potent technique for managing model complexity and enhancing generalisation performance.

1. **Which particular algorithms are used for regularization?**

**ANSWER** - Regularisation techniques are used by many algorithms to limit model complexity and avoid overfitting. Several of the frequently employed regularisation algorithms include:

* Ridge Regression: Ridge regression modifies the objective function of linear regression by including an L2 regularisation term. It penalises the size of the coefficients while minimising the sum of squared errors. This enhances the model's robustness and lessens the effect of big coefficients.
* Lasso Regression: Lasso regression extends the linear regression objective function with an L1 regularisation term. It successfully performs feature selection, lowers the number of predictors in the model, and supports sparsity by decreasing some coefficients to exactly zero.
* The L1 and L2 regularisation penalties are combined in elastic net regression. When there are several correlated predictors, it is helpful because it strikes a balance between coefficient shrinkage (L2 penalty) and feature selection (L1 penalty).
* Logistic Regression with L1 or L2 Regularisation: L1 or L2 penalties can also be used to regularise logistic regression. By regulating the magnitudes of the coefficients, regularised logistic regression aids in preventing overfitting in binary classification tasks.
* Support Vector Machines (SVM): The C parameter, which regulates the trade-off between maximising the margin and minimising the classification errors, can be used to regularise SVM algorithms. A model with a lower value of C is more regularised, whereas one with a larger value of C is less regularised.
* Dropout is a regularisation method that is unique to neural networks. Neural Networks with Dropout. In order to prevent co-adaptation of neurons and lessen overfitting, it randomly changes a portion of input units to zero during each training iteration.

These are but a handful of examples of algorithms that use regularisation principles. Various alternative models and algorithms can also be regularised, depending on the issue at hand and the particular regularisation requirements.

1. **Explain the term error present in linear regression equation?**

**ANSWER -** The difference or disparity between the projected values by the linear regression model and the actual observed values in the dataset is referred to as "error" in the context of linear regression. In the link between the independent variables (features) and the dependent variable (target), it symbolises the unexplained variance or noise.

The linear regression equation is typically represented as:

y = b0 + b1x1 + b2x2 + ... + bn\*xn + error

where:

* y is the dependent variable (target)
* b0 is the y-intercept or the coefficient of the constant term
* b1, b2, ..., bn are the coefficients of the independent variables (x1, x2, ..., xn)
* x1, x2, ..., xn are the independent variables (features)
* error represents the error term or residual

Estimating the coefficients (b0, b1, b2,..., bn) that minimise the sum of squared errors (SSE) or the sum of squared residuals is the aim of linear regression. The error term accounts for the discrepancies between each data point's actual values and its projected values.

The dependent variable's component that is not fully explained by the linear relationship with the independent variables is captured by the error term. It incorporates elements that the model does not take into account, such as measurement mistakes, unobserved variables, and other sources of variance. The error term is assumed to have a normal distribution with a mean of zero in linear regression.

The goal of linear regression is to find the best-fitting line that illustrates the relationship between the independent factors and the dependent variable by minimising the errors or residuals. Based on the known values of the independent variables, this line makes forecasts or estimations for the dependent variable.