MACHINE LEARNING – WORKSHEET 1

Question Numbers.	Answers.
01	A
02	A
03	B
04	B
05	C
06	B
07	D
08	D
09	A
10	B
11	A, B, C
12	A, B, C

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

13. Explain the term regularization?
Ans.

Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from overfitting by adding extra information to it.

Sometimes the model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called overfitted. This problem can be deal with the help of a regularization technique.

This technique can be used in such a way that it will allow to maintain all variables or features in the model by reducing the magnitude of the variables. Hence, it maintains accuracy as well as a generalization of the model.

It mainly regularizes or reduces the coefficient of features toward zero. In simple words, "*In regularization technique, we reduce the magnitude of the features by keeping the same number of features.*"

14. Which particular algorithms are used for regularization?
Ans.

Following are the various regularization algorithms used in Machine Learning:

- Ridge Regression
- LASSO (Least Absolute Shrinkage and selector operator) Regression

Ridge Regression:

Ridge regression is one of the types of linear regression in which a small amount of bias is introduced so that we can get better long-term predictions.

Ridge regression is a regularization technique, which is used to reduce the complexity of the model. It is also called as **L2 regularization**.

Ridge Regression is a technique for analysing multiple regression data that suffer from multicollinearity. When multicollinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value. By adding a degree of bias to the regression estimates, ridge regression reduces the standard errors. Multicollinearity, or collinearity, is the existence of near-linear relationships among the independent variables.

Following the usual notation, suppose our regression equation is written in matrix form as:

Y = XB + e

Where,

Y = Dependent variable,

X = Independent variables,

B = Regression coefficients to be estimated, and

e = Errors are residuals.

L2 = λ ∑ | βj² |

Where,

βj = Magnitude of coefficient.

λ = Amount of shrinkage.

LASSO Regression:

Lasso regression is another regularization technique to reduce the complexity of the model. It stands for **Least Absolute and Selection Operator.**

Lasso regression is a regularization technique. It is used over regression methods for a more accurate prediction. This model uses shrinkage. Shrinkage is where data values are shrunk towards a central point as the mean. The lasso procedure encourages simple, sparse models (i.e., models with fewer parameters). This particular type of regression is well-suited for models showing high levels of multicollinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.

Lasso Regression uses **L1 regularization technique**. It is used when we have a greater number of features because it automatically performs feature selection. The Lasso regression can help us to reduce the overfitting in the model as well as the feature selection.

Mathematical equation of Lasso Regression:

Residual Sum of Squares + λ * (Sum of the absolute value of the magnitude of coefficients)

$$\sum_{i=1}^{M}(y_i - y'_i)^2 = \sum_{i=1}^{M}\left(y_i - \sum_{j=0}^{n}\beta_j * x_{ij}\right)^2 + \lambda\sum_{j=0}^{n}|\beta_j|$$

L1 = λ ∑ | βj |

Where,

βj denotes the magnitude of coefficient.

λ denotes the amount of shrinkage.

- λ = 0 implies all features are considered and it is equivalent to the linear regression where only the residual sum of squares is considered to build a predictive model
- λ = ∞ implies no feature is considered i.e, as λ closes to infinity it eliminates more and more features.
- The bias increases with increase in λ
- variance increases with decrease in λ

Key Difference between Ridge regression and Lasso regression:

- **Ridge regression** is mostly used to reduce the overfitting in the model, and it includes all the features present in the model. It reduces the complexity of the model by shrinking the coefficients.
- **Lasso regression** helps to reduce the overfitting in the model as well as feature selection.

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

Ans.

In simple words error in regression analysis is the difference between actual value and predicted value. The error term of a regression equation represents all of the variation in the dependent variable not explained by the weighted independent variables.

Graphically if the if the actual and predicted values are exactly equal then best fit line should consider the actual values line. The graphical distance calculated in this case can be called as error.

error term means the distance from each point to the best fit line.