(MACHINE LEARNING) - WORKSHEET-1

ANSWER SHEET

Objective type

1) A.

2) A.

3) B.

4) B.

5) C.

6) D.

7) D.

8) D.

9) A.

10) B.

11) B.

12) A,C.

Descriptive type

13) Sometimes, we encounter a situation where our machine learning algorithm fails to perform well on testing data but works well on training data exceptionally. This is a situation where regularization comes into play and solves the issues. It discourages learning a flexible or more complex model in order to prevent overfitting. This technique can be used in such a way that it will allow us to maintain all variables or features in the model by reducing the magnitude of the variables. Hence, regularization maintains accuracy as well as generalization of the model.

14) Mainly two types of algorithms are used for regularization --

- **L1 Regularization or Lasso Regularization**

- **L2 Regularization or Ridge Regularization**

L1 Regularization or Lasso Regularization

Lasso regularization, also known as L1 regularization, is another technique of regularization that scales back the complexity of the model.

The working principle of Lasso regularization is almost similar to Ridge regularization except that the penalty term contains only the weights rather than the square of weights. Moreover, it stands for Least Absolute and Selection Operator. Hence, the Lasso regularization can help us to decrease the overfitting in the model as well as increase the feature selection. The general mathematical equation for the value function of Lasso regularization will be:

$$Min(\sum_{i=1}^{n}(y_i - w_i x_i)^2 + p\sum_{i=1}^{n}|w_i|)$$

p is the tuning parameter which decides how much we want to penalize the model.

L2 Regularization or Ridge Regularization

Ridge regularization is one of the regularization techniques of linear regularization in which a small amount of bias is introduced so that we can get better long-term predictions, and it is used to reduce the complexity of the model. It is also known as L2 regularization.

In this technique, the amount of bias added to the model is known as Ridge regularization penalty and the cost function is altered by adding the penalty term to it. We can calculate it by multiplying the lambda to the squared weight of every individual feature, which says the equation for the cost function in ridge regularization will be:

$$Min(\sum_{i=1}^{n}(y_i - w_i x_i)^2 + p\sum_{i=1}^{n}(w_i)^2)$$

Similar to L1, in L2 also, p is the tuning parameter which decides how much we want to penalize the model.

15)