

MACHINE LEARNING

Assignment-39

Q1 to 11 :

- 1) D-Both A & B
- 2) A-Linear regression is sensitive to outliers
- 3) B-Negative
- 4) C-Both of them
- 5) D-None of these
- 6) B-Predictive Model
- 7) D-Regularization
- 8) D-SMOTE
- 9) C-Sensitivity and Specificity
- 10) B-False

Q.12 Answer: A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

Q.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. Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.

To train our machine learning model, we give it some data to learn from. The process of plotting a series of data points and drawing the best fit line to understand the relationship between the variables is called Data Fitting. Our model is the best fit when it can find all necessary patterns in our data and avoid the random data points and unnecessary patterns called Noise.

Sometimes the machine learning 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.

There are mainly two types of regularization techniques, which are given below:

- **Ridge Regression**
- **Lasso 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**.

In this technique, the cost function is altered by adding the penalty term to it. The amount of bias added to the model is called **Ridge Regression penalty**. We can calculate it by multiplying with the lambda to the squared weight of each individual feature.

LASSO REGRESSION:

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

It is similar to the Ridge Regression except that the penalty term contains only the absolute weights instead of a square of weights.

Since it takes absolute values, hence, it can shrink the slope to 0, whereas Ridge Regression can only shrink it near to 0.

Q.14) Which particular algorithms are used for Regularization?

Regularization is a technique used in regression to reduce the complexity of the model and to shrink the coefficients of the independent features.

There are different Regularization Algorithms are as:

- 1) **Ridge Regression**
- 2) **Lasso Regression**
- 3) **Elastic – Net Regression**

1) 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**.

In this technique, the cost function is altered by adding the penalty term to it. The amount of bias added to the model is called **Ridge Regression penalty**. We can calculate it by multiplying with the lambda to the squared weight of each individual feature.

2) Lasso Regression :

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

It is similar to the Ridge Regression except that the penalty term contains only the absolute weights instead of a square of weights.

Since it takes absolute values, hence, it can shrink the slope to 0, whereas Ridge Regression can only shrink it near to 0.

3) Elastic-Net Regression :

Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

Q.15) Explain the term error present in linear Regression equation?

Ans:

A regression line always has an error term because, in real life, independent variables are never perfect predictors of the dependent variables. Rather the line is an estimate based on the available data. So the error term **tells you how certain you can be about the formula.**

Submitted By: Netra Dalvi