

UNIT 2

SUPERVISED LEARNING

1. What is linear regression?

Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as **sales, salary, age, product price**, etc.

2. List out the types of linear regression?

- Simple Linear Regression
- Multiple Linear regression

3. What is linear regression model?

A linear regression model describes the relationship between a dependent variable, y , and one or more independent variables, X . The dependent variable is also called the response variable. Independent variables are also called explanatory or predictor variables. Continuous predictor variables are also called covariates, and categorical predictor variables are also called factors. The matrix X of observations on predictor variables is usually called the design matrix.

A multiple linear regression model is

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i=1, \dots, n,$$

4. What is the Least Square Regression Method?

Least squares is a commonly used method in regression analysis for estimating the unknown parameters by creating a model which will minimize the sum of squared errors between the observed data and the predicted data.

5. What is subset selection method?

Subset selection **evaluates a subset of features as a group for suitability**. Subset selection algorithms can be broken up into wrappers, filters, and embedded methods. Wrappers use a search algorithm to search through the space of possible features and evaluate each subset by running a model on the subset.

6. What is shrinkage method in ML?

The idea is to shrink some of the parameters to zero. It uses an optimization formula. It is also a good method for improvement of prediction accuracy but it is hard to say it improves interpretability because of very small coefficients in the model. It is based on more mathematical methods.

7. How do we use shrinking methods?

The best known shrinking methods are **Ridge Regression** and **Lasso Regression** which are often used in place of Linear Regression. Ridge Regression, like Linear Regression, aims to minimize the Residual Sum of Squares(RSS) but with as light change.

8. why we are not satisfied with the least squares estimates?

- The first is *prediction accuracy*: the least squares estimates often have low bias but large variance. Prediction accuracy can sometimes be improved by shrinking or setting some coefficients to zero. By doing so we sacrifice a little bit of bias to reduce the variance of the predicted values, and hence may improve the overall prediction accuracy.
- The second reason is interpretation. With a large number of predictors, we often would like to determine a smaller subset that exhibit the strongest effects. In order to get the "big picture", we are willing to sacrifice some of the small details.

9. What is the advantage of supervised learning?

- With the help of supervised learning, the model can predict the output on the basis of prior experiences.
- In supervised learning, we can have an exact idea about the classes of objects.
- Supervised learning model helps us to solve various real-world problems such as **fraud detection, spam filtering**, etc.

10. What are the Types of supervised Machine learning Algorithms?

- Regression
- Classification

11. What is supervised machine learning?

Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.

12. What are the linear models for classification?

A linear classifier is **a model that makes a decision to categories a set of data points to a discrete class based on a linear combination of its explanatory variables**. As an example, combining details about a dog such as weight, height, colour and other features would be used by a model to decide its species.

13. Difference between Regression and Classification Algorithm?

Regression Algorithm	Classification Algorithm
In Regression, the output variable must be of continuous nature or real value.	In Classification, the output variable must be a discrete value.
The task of the regression algorithm is to map the input value (x) with the continuous output variable(y).	The task of the classification algorithm is to map the input value(x) with the discrete output variable(y).
Regression Algorithms are used with continuous data.	Classification Algorithms are used with discrete data.

14. What is linear discriminant analysis?

It is one of the most used dimensionality reduction techniques. It is used in machine learning as well as applications that have anything to do with the classification of patterns. LDA serves a very specific purpose, which is to project features that exist in a high dimensional space onto space at a lower dimension.

15. What are the extensions of LDA?

Extensions to LDA:

1. **Quadratic Discriminant Analysis (QDA):** Each class uses its own estimate of variance (or covariance when there are multiple input variables).
2. **Flexible Discriminant Analysis (FDA):** Where non-linear combinations of inputs are used such as splines.
3. **Regularized Discriminant Analysis (RDA):** Introduces regularization into the estimate of the variance (actually covariance), moderating the influence of different variables on LDA.

16. What are the Applications of LDA?

Applications of LDA:

Face Recognition: In the field of Computer Vision, face recognition is a very popular application in which each face is represented by a very large number of pixel values.

Medical: In this field, Linear discriminant analysis (LDA) is used to classify the patient disease state as mild, moderate, or severe based upon the patient's various parameters and the medical treatment he is going through. This helps the doctors to intensify or reduce the pace of their treatment.

Customer Identification: Suppose we want to identify the type of customers who are most likely to buy a particular product in a shopping mall. By doing a simple question and answers survey, we can gather all the features of the customers.

17. Define logistic regression.

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

18. What are the assumptions of logistic reasoning?

- The dependent variable must be categorical in nature.
- The independent variable should not have multi-collinearity.

19. What is separating hyperplane in machine learning?

In Machine Learning, a hyperplane is **a decision boundary that divides the input space into two or more regions, each corresponding to a different class or output label**. In a 2D space, a hyperplane is a straight line that divides the space into two halves.

20. How supervised machine learning works?

Supervised learning **uses a training set to teach models to yield the desired output**. This training dataset includes inputs and correct outputs, which allow the model to learn over time. The algorithm measures its accuracy through the loss function, adjusting until the error has been sufficiently minimized.

21. Why do we use Regression Analysis?

- Regression estimates the relationship between the target and the independent variable.
- It is used to find the trends in data.
- It helps to predict real/continuous values.

22. List out the types of Regression?

- Linear Regression
- Logistic Regression
- Polynomial Regression
- Support Vector Regression
- Decision Tree Regression
- Random Forest Regression
- Ridge Regression
- Lasso Regression

23. Why do we use shrinkage method?

Shrinkage is the value used **to determine the total required staffing levels necessary to meet your business goals**. In other words, it's the amount of “over-scheduling” you must perform in order to have the right number of agents working at any given time of the day.

24. What is the use of hyperplane in machine learning?

Hyperplanes are **decision boundaries that help classify the data points**. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line.

25. How do you use linear regression for supervised learning?

The best way to perform a linear regression, is to **split the data into many small batches. Each batch, with approximately the same number of points. Then use each batch to update the weights**. This method is called Mini-Batch Gradient Descent.