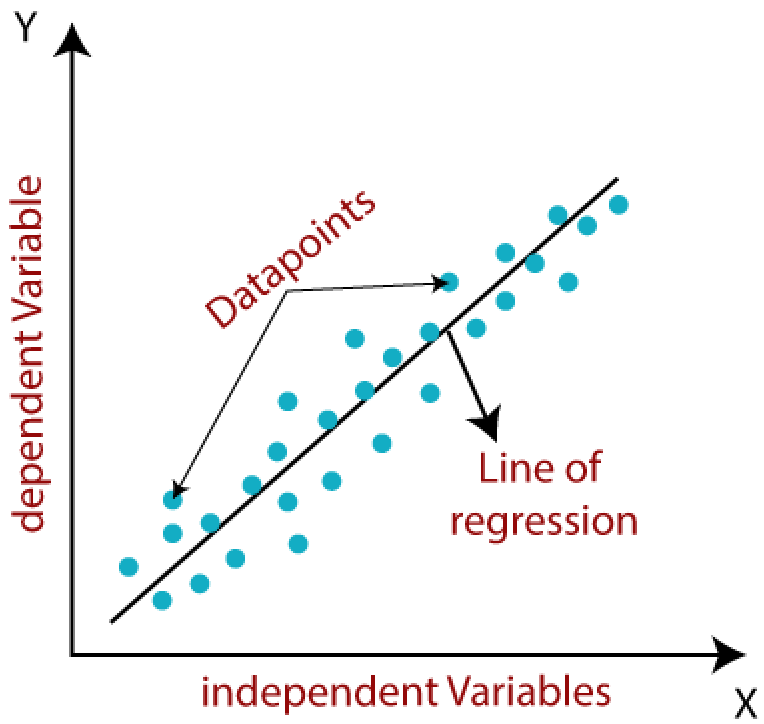
Linear Regression in Deep Learning

Linear regression is one of the most basic and widely used Machine Learning methods. It is a statistical method for performing predictive analysis. Linear regression forecasts continuous/real or quantitative variables such as sales, salary, age, product price, and so on.

The linear regression algorithm demonstrates a linear relationship between a dependent (y) variable and one or more independent (y) variables, hence the name. Because linear regression demonstrates a linear relationship, it determines how the value of the dependent variable changes in respect to the value of the independent variable.

The linear regression model generates a slanted straight line that represents the relationship between the variables. Consider the following image:



A linear regression can be represented mathematically as:

39.6K

A woman attempting to surf with a wave machine falls backwards and faceplants into the water.

y= a0+a1x+ ε

Here,

Learnvista Pvt Ltd.
2nd Floor, 147, 5th Main Rd, Rajiv Gandhi Nagar HSR Sector 7,Near Salarpuria Serenity, Bengaluru, Karnataka 560102
Mob:- +91 779568798, Email:- contacts@learnbay.co

Y stands for Dependent Variable (Target Variable)

X stands for Independent Variable (predictor Variable)

a0 = the line's intercept (Gives an additional degree of freedom)

a1 = Linear regression coefficient (scale factor to each input value).

ε = random error

The values for the x and y variables are training datasets for the representation of a Linear Regression model.

Types of Linear Regression

Linear regression can be further divided into two types of the algorithm:

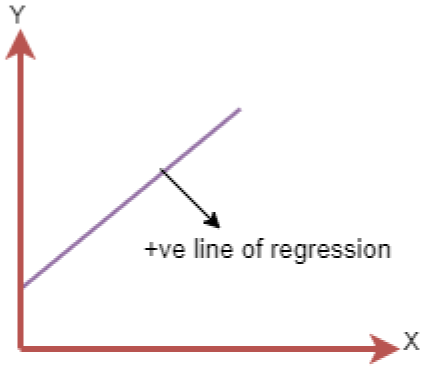
- Simple Linear Regression:
  It refers to a Linear Regression procedure that uses a single independent variable to predict the value of a numerical dependent variable.

- Linear regression using several variables:
  When more than one independent variable is used to predict the value of a numerical dependent variable, the process is known as Multiple Linear Regression.

Linear Regression Line

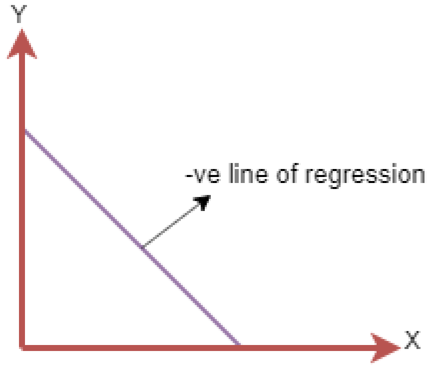
A regression line is a linear line that depicts the relationship between the dependent and independent variables. A regression line can depict two kinds of relationships:

- Positive Linear Relationship:
  A positive linear relationship exists when the dependent variable increases on the Y-axis and the independent variable increases on the X-axis.



The line equation will be: $Y= a_0+a_1x$

- Negative Linear Relationship:
  A negative linear connection exists when the dependent variable declines on the Y-axis while the independent variable increases on the X-axis.



The line of equation will be: $Y= -a_0+a_1x$

Finding the best fit line:

When dealing with linear regression, our major goal is to identify the best fit line, which means that the difference between projected and actual values should be as little as possible. The best-fitting line will have the least amount of inaccuracy.

The different weights or coefficients of lines (a0, a1) give a different line of regression, thus we need to compute the best values for a0 and a1 to get the best fit line, so we use the cost function to calculate this.

Cost function-

- The cost function is used to estimate the coefficient values for the best fit line, and the varied values for weights or coefficient of lines (a0, a1) give the different line of regression.
- The cost function is used to maximise the regression coefficients or weights. It assesses the performance of a linear regression model.
- The cost function can be used to determine the correctness of the mapping function, which maps the input variable to the output variable. This mapping function is often referred to as the Hypothesis function.

We utilise the Mean Squared Error (MSE) cost function for Linear Regression, which is the average of squared errors between predicted and actual values. It can be written as follows:

MSE may be determined for the above linear equation as follows:

$$MSE= 1\frac{1}{N}\sum_{i=1}^{n}(y_i \ - (a_1x_i + a_0))^2$$

Where,
N=The total number of observations
Yi = Real worth
(a1xi+a0)= Predicted value.

Residuals: The difference between the actual and expected values is referred to as the residual. If the observed points are distant from the regression line, the residual and therefore the cost function will be large. If the scatter points are close to the regression line, the residual and hence the cost function will be tiny.

Gradient Descent:

- By calculating the gradient of the cost function, gradient descent is utilised to minimise the MSE.
- Gradient descent is used in regression models to update the coefficients of the line by decreasing the cost function.
- It is accomplished by selecting a random set of coefficient values and then iteratively updating the values to achieve the lowest cost function.

Model Performance:

The Goodness of Fit parameter defines how well the regression line fits the set of observations. The process of selecting the best model from a set of models is known as optimization. It is possible to accomplish this using the following method:

1. R-squared method:
   - R-squared is a statistical approach for calculating goodness of fit.
   - On a scale of 0-100 percent, it assesses the strength of the link between the dependent and independent variables.
   - A high R-square value indicates that there is little discrepancy between predicted and actual values, indicating a good model.
   - It is also known as a coefficient of determination or, in the case of multiple regression, a coefficient of multiple determination.
   - It can be calculated using the following formula:

$$\text{R-squared} = \frac{\text{Explained variation}}{\text{Total Variation}}$$

Assumptions of Linear Regression

Below are some important assumptions of Linear Regression. These are some formal checks while building a Linear Regression model, which ensures to get the best possible result from the given dataset.

- Linear relationship between the features and target:
  Linear regression assumes the linear relationship between the dependent and independent variables.
- Small or no multicollinearity between the features:
  Multicollinearity means high-correlation between the independent variables. Due to multicollinearity, it may be difficult to find the true relationship between the predictors and target variables. Or we can say, it is difficult to determine which predictor variable is affecting the target variable and which is not. So, the model assumes either little or no multicollinearity between the features or independent variables.
- Homoscedasticity Assumption:
  Homoscedasticity is a situation when the error term is the same for all the values of independent variables. With homoscedasticity, there should be no clear pattern distribution of data in the scatter plot.
- Normal distribution of error terms:
  Linear regression assumes that the error term should follow the normal distribution pattern. If error terms are not normally distributed, then confidence intervals will become

either too wide or too narrow, which may cause difficulties in finding coefficients. It can be checked using the q-q plot. If the plot shows a straight line without any deviation, which means the error is normally distributed.

- No autocorrelations:
  The linear regression model assumes no autocorrelation in error terms. If there will be any correlation in the error term, then it will drastically reduce the accuracy of the model. Autocorrelation usually occurs if there is a dependency between residual errors.

Why only talk about knowledge when you can work it? Join our program to get trained and work on live industry projects and get certified by start-ups. Use your skills on projects and increase your practical knowledge.