
AI & Machine Learning

[Tutorials](#) [Articles](#) [Ebooks](#) [Free Practice Tests](#) [On-demand Webinars](#)

[Home](#) [Resources](#) [AI & Machine Learning](#) [Machine Learning Tutorial: A Step-by-Step Guide for Beginners](#) [Understanding The Difference Between Linear vs Logistic Regression](#)

Understanding The Difference Between Linear vs Logistic Regression

Lesson 11 of 38

By Simplilearn

Last updated on Feb 23, 2023

46886



[Previous](#)

[Next](#)

Tutorial Playlist

Table of Contents

[What Is Regression?](#)

[What Is Classification?](#)

[What Is Linear Regression?](#)

[What Is Logistic Regression?](#)

[Linear vs. Logistic Regression: Differences](#)

[View More](#)

Some machine learning algorithms work on data that has the input values along with the corresponding output values. These algorithms are called supervised learning algorithms. It only restricts their output value to the output values provided in the data. Two of the most commonly used supervised learning algorithms are Linear and Logistic Regression. In this tutorial titled 'Understanding the difference between Linear vs. Logistic Regression', you will see the working and the differences between these two algorithms.

What Is Regression?

Regression is a statistical method that allows you to predict a dependent output variable based on the values of independent input variables.

Regression, a type of supervised learning, finds the relationship between input and output values and, a given input data, to predict the output value. It does this by finding a mathematical, linear relationship between input and output values. It can have multiple inputs but has a single output.

You can understand regression better, using the diagram below. Using the given input variables or grocery ingredients, you can get a new output or dish. Here, Regression acts as a recipe used to find how these variables go together and the relationship between them.

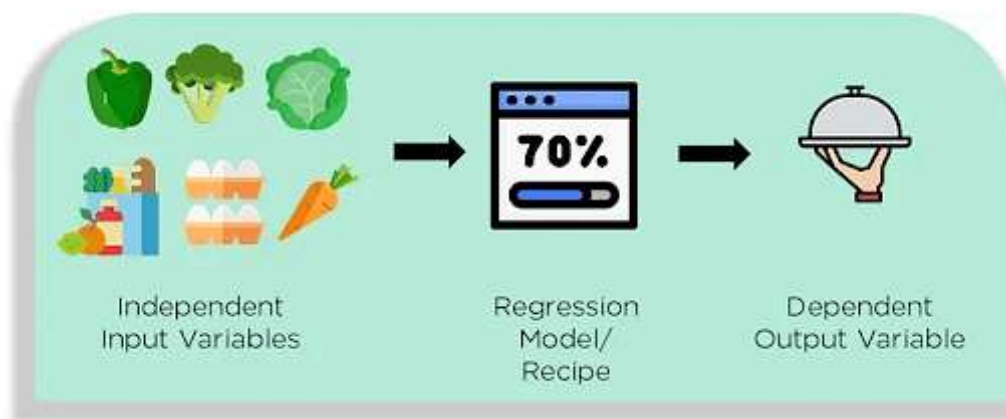


Figure 1: Regression

Your AI/ML Career is Just Around The Corner!

AI Engineer Master's Program

[EXPLORE PROGRAM](#)

What Is Classification?

[Classification allows](#) you to divide a given input into some pre-defined categories. The output is a

discrete value, i.e., distinct, like 0/1, True/False, or a pre-defined output label class.

Simply put, classification is the process of segregating or classifying objects. It is a type of supervised learning method where input data is usually classified into output classes. It provides a mapping function to convert input values into known, discrete output classes. It can have multiple inputs and gives multiple outputs.

The diagram below clearly explains classification. Given a list of grocery items, you can separate them into different categories like vegetables, fruits, dairy products, groceries, etc., using classification.

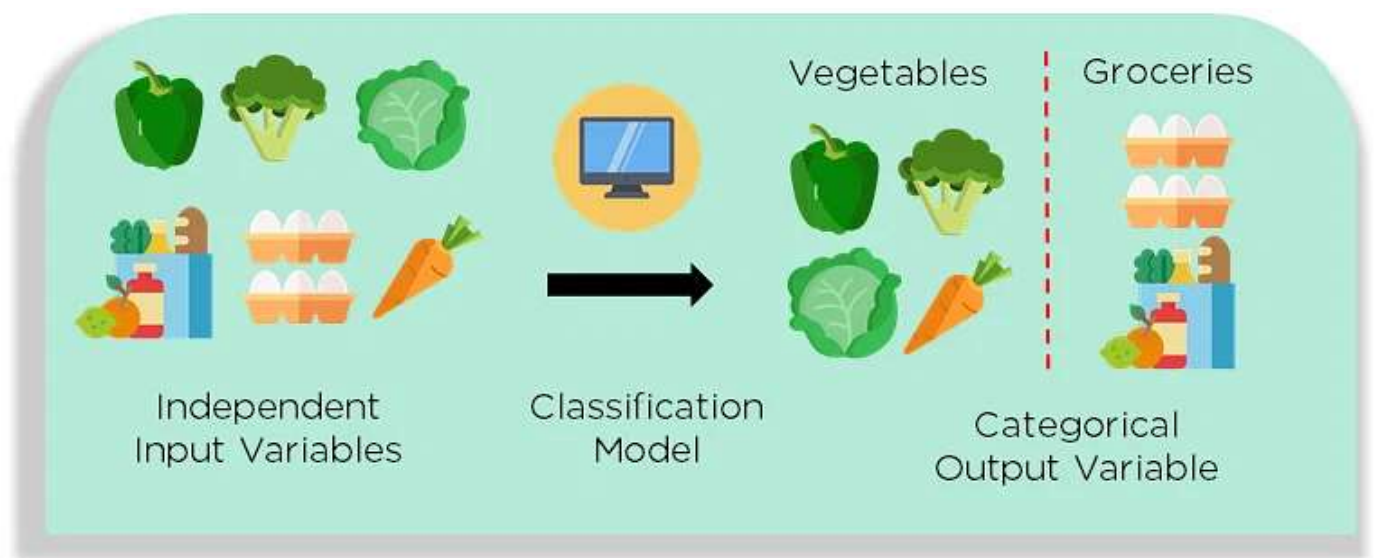


Figure 2: Classification

What Is Linear Regression?

Linear Regression is a machine learning model used to predict output variable's values based on the value of input variables.

Consider the data points given below. The input variables, X , are called independent variables and are used to predict response values. They are unrelated values that have no relationship with each other. The output variable, Y , is called the dependent variable. Its value depends on the value of X . It is found by deriving a relationship between the input variables.

Also Read: [How to Develop a Machine Learning Career?](#)

Y

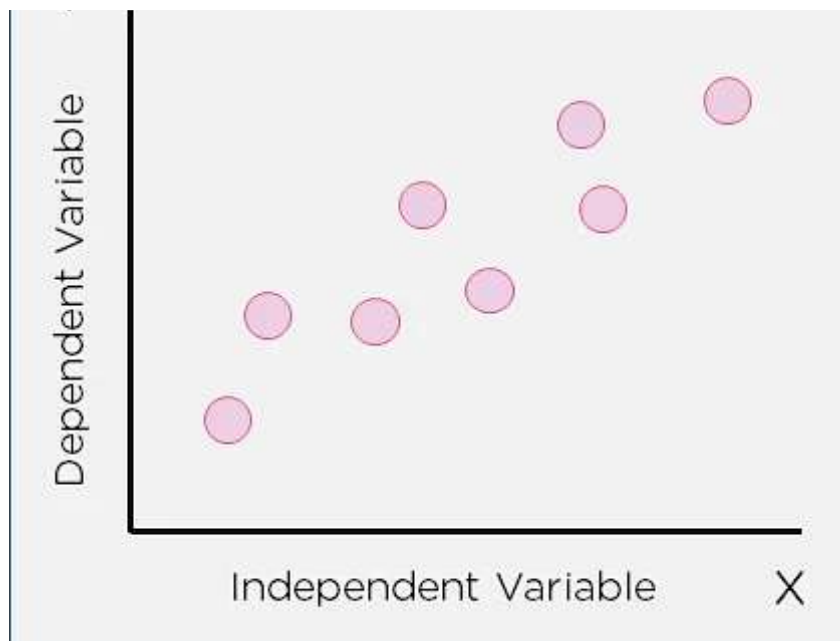


Figure 3: Data

Linear Regression finds the relationship between the input and output data by plotting a line that fits the input data and maps it onto the output. This line represents the mathematical relationship between the independent input variables and is called The Line of Best Fit. Ideally, it covers as many input variables as possible while leaving out the outliers or the noise. For your given data, the best fit is a straight line.

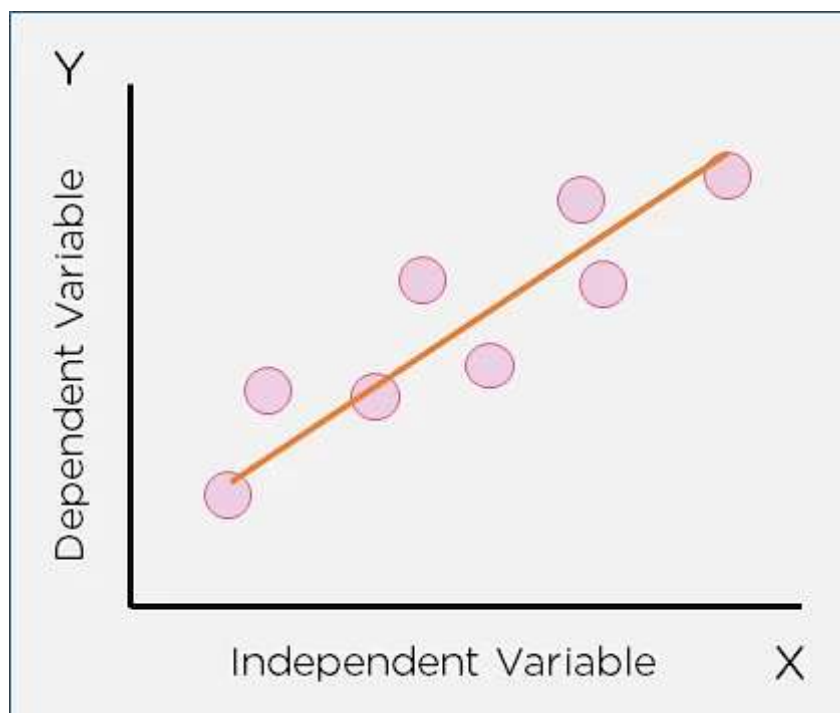


Figure 4: Linear Regression line of best fit

The equation which can be used to fit a line is the Equation of a Straight Line. The equation gives the output variable based on the input variable and inclination of the line. The line can be found using the following equation :

using the following equation .

$$Y = b_0 + b_1X + e$$

Figure 5: Equation of a straight line

In the above equation,

Y = Dependent Output Variable

X = Independent Input Variable

b0 = Y-intercept, or the point at which the line meets the y-axis

b1 = Slope, or the inclination of the line

e = Error term

Using this line, you can find the output value for a given input variable by extending a line from the X-axis onto the line of best fit and seeing the corresponding Y-axis term. Consider the data that is displayed below, which tells you the sales corresponding to the amount spent on advertising.

| Advertisement | Sales |
|---------------|-------|
| 50 | 100 |
| 150 | 275 |
| 200 | 300 |
| 250 | 400 |
| 175 | ?? |

Figure 6: Advertising data

Using Linear Regression, you can plot the graph of Sales vs. Advertising, and find the line of best fit between them, and, using that, find the values of the missing variable.

Y

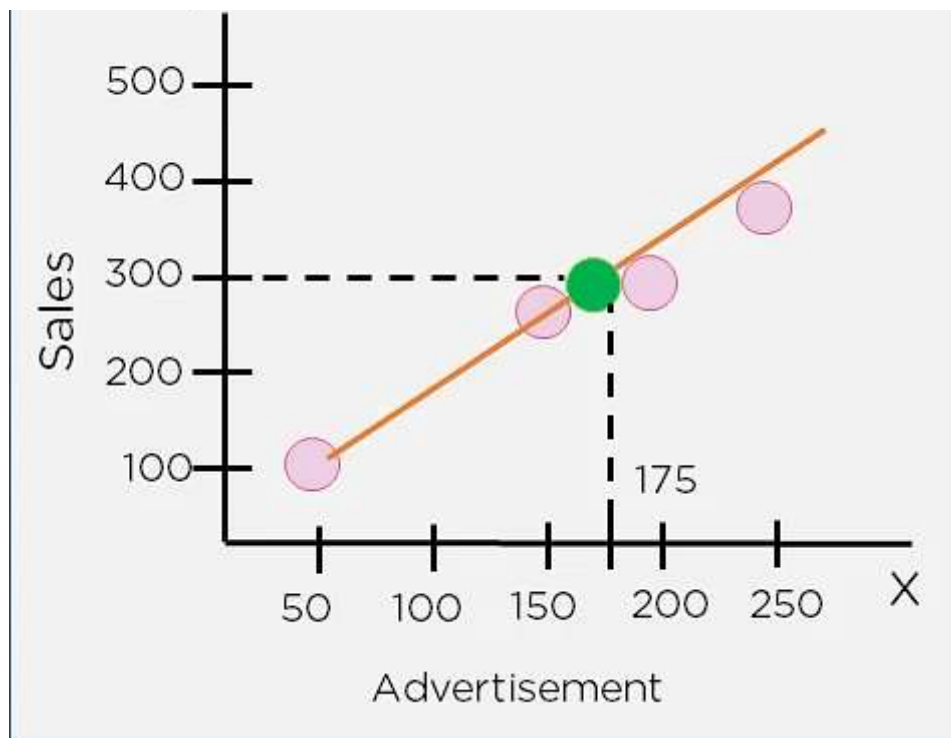


Figure 7: Sales vs Advertising

Using regression, given the advertisement amount, you can predict how many sales will take place.

| Advertisement | Sales |
|---------------|-------|
| 50 | 100 |
| 150 | 275 |
| 200 | 300 |
| 250 | 400 |
| 175 | 300 |

Figure 8: Prediction using Linear regression

Your AI/ML Career is Just Around The Corner!

AI Engineer Master's Program

What Is Logistic Regression?

Logistic Regression is a classification algorithm used to predict the category of a dependent variable based on the values of the independent variable. Its output is 0 or 1.

In Logistic Regression, the input data belongs to categories, which means multiple input values map onto the same output values. Using Logistic Regression, you can find the category that a new input value belongs to. Unlike Linear regression, Logistic Regression does not assume that the values are linearly correlated to one other. Consider the data below, which shows the input data mapped onto two output categories, 0 and 1.

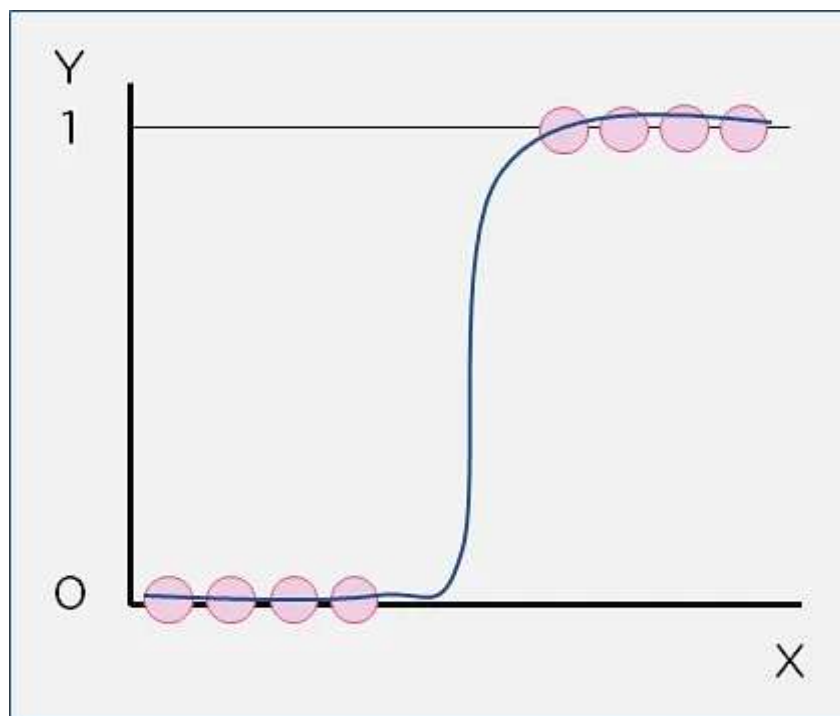
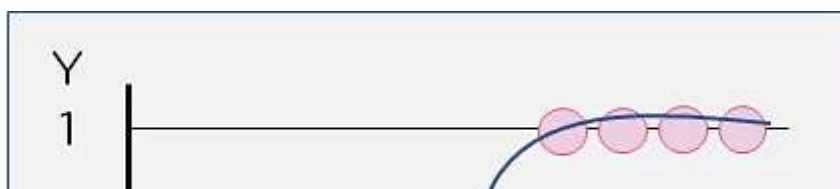


Figure 9: Logistic Regression data

The data is plotted, and it draws a curve to represent the relationship between the points in our data, which joins the various classes in our output. To classify values into these two categories, you need to set a threshold value between them.



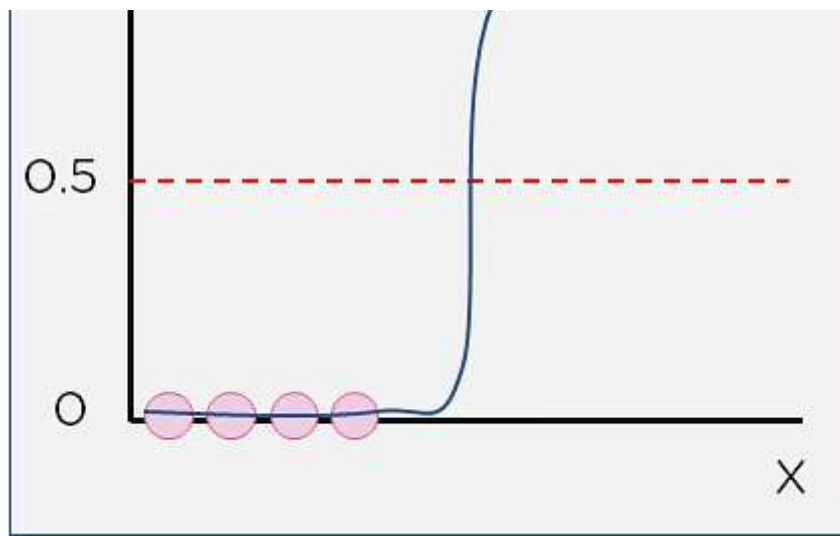


Figure 10: Setting a threshold

It maps the values of the input values onto a categorical variable depending on their position relative to the threshold value. Values of Y above this threshold will be classified as category 1, and it will take values below the threshold as category 0.

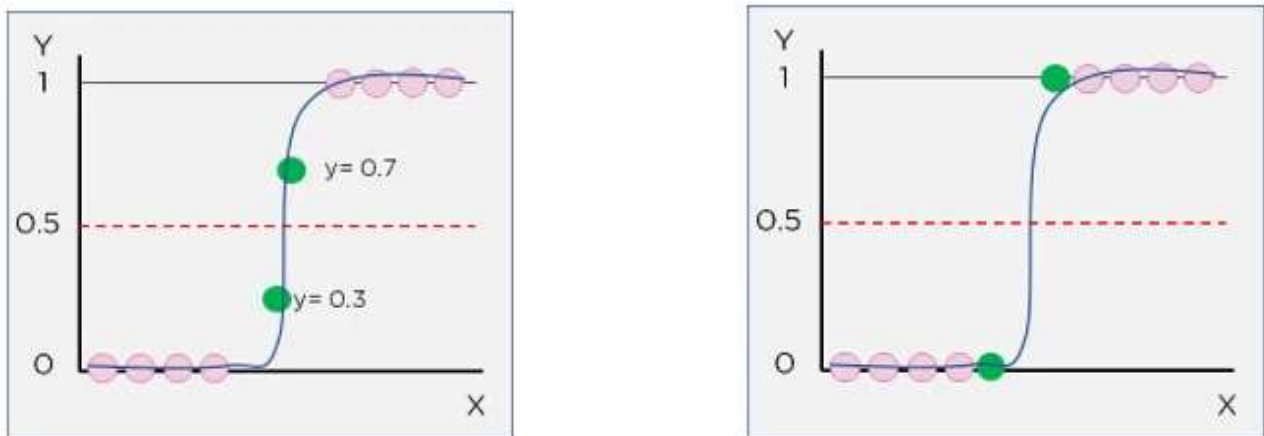


Figure 11: Dividing data into categories

Logistic Regression finds the relationship between points by first plotting a curve between the output classes. This curve is called a sigmoid, and the given equation is used to represent a sigmoid function. Y is the probability of output, c is a constant, X is the various dependent variables, and b₀, b₁ gives you the intercept values.

$$\log \frac{Y}{1+Y} = b_0 + b_1 X_1 + b_2 X_2 + C$$

Figure 12: Sigmoid Function

Suppose you have credit card numbers and their transaction history. You want to classify credit

Suppose you have a dataset of credit cards and you want to classify them as fraudulent and legitimate. You can do this with Logistic Regression.

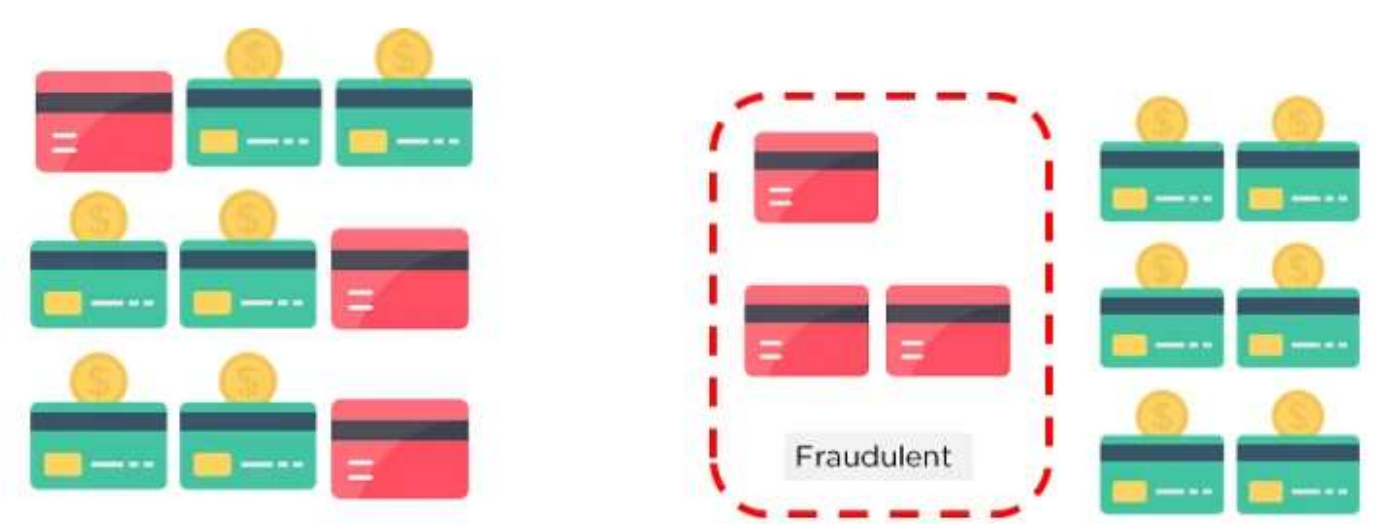


Figure 13: Classifying credit cards using Logistic regression

Linear vs. Logistic Regression: Differences

The table below lists the difference between these two supervised algorithms.

| Linear Regression | Logistic Regression |
|---|--|
| Used to predict a dependent output variable based on independent input variable | Used to classify a dependent output variable based on independent input variable |
| Accuracy is measured using Least squares estimation | Accuracy is measured using Maximum Likelihood estimation |
| The best fit line is a straight line | The best fit is given by a curve |
| The output is a predicted integer value | The output is a binary value between 0 and 1 value |
| Used in business domain, forecasting stocks | Used for classification, image processing |

Table 1: Linear vs. Logistic Regression

Conclusion

In this tutorial titled ' Understanding the difference between Linear Vs. Logistic Regression, you took a look at the definition of Regression and classification. You then learned about Linear regression, a regression algorithm, and Logistic Regression, a classification algorithm. Finally, you explored the difference between these two algorithms.

We hope this helped you understand the difference between Linear and Logistic Regression. To learn more about regression and machine learning, check out Simplilearn’s Caltech [AI Course](#). If you have any questions or doubts, mention them in this article's comments section, and we'll have our experts answer them for you at the earliest!

Find our Post Graduate Program in AI and Machine Learning Online Bootcamp in top cities:

| Name | Date | Place |
|--|---|-----------|
| Post Graduate Program in AI and Machine Learning | Cohort starts on 17th Jan 2024, Weekend batch | Your City |
| Post Graduate Program in AI and Machine Learning | Cohort starts on 26th Jan 2024, Weekend batch | Your City |

About the Author



Simplilearn is one of the world’s leading providers of online training for Digital Marketing, Cloud Computing, Project Management, Data Science, IT, Software Development, and many other em...

[View More](#)



Recommended Programs



Post Graduate Program in AI and Machine Learning




2794 Learners

Lifetime Access*

*Lifetime access to high-quality, self-paced e-learning content.

[Explore Category](#)

Recommended Resources

| | | |
|---|---|---|
|  |  |  |
| An Introduction to Logistic Regression in Python Tutorial | Skills Acquisition Vs. Talent Acquisition Ebook | Sklearn Article |

© 2009 -2024- Simplilearn Solutions

Disclaimer

PMP, PMI, PMBOK, CAPM, PgMP, PfMP, ACP, PBA, RMP, SP, and OPM3 are registered marks of the Project Management Institute, Inc.