

Master the Art of Data Science: A Step-by-Step Guide to Becoming an Industry-Ready Data Scientist

[Download Roadmap](#)



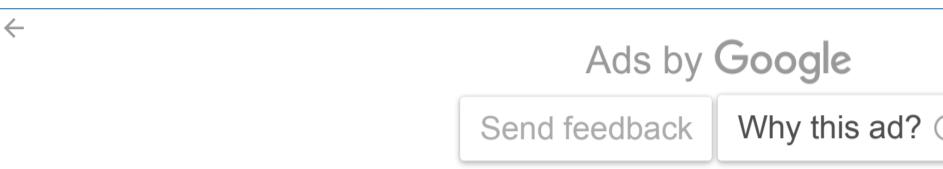
[Home](#)

How to Learn Mathematics For Machine Learning? What Concepts do You Need to Master in Data Science?

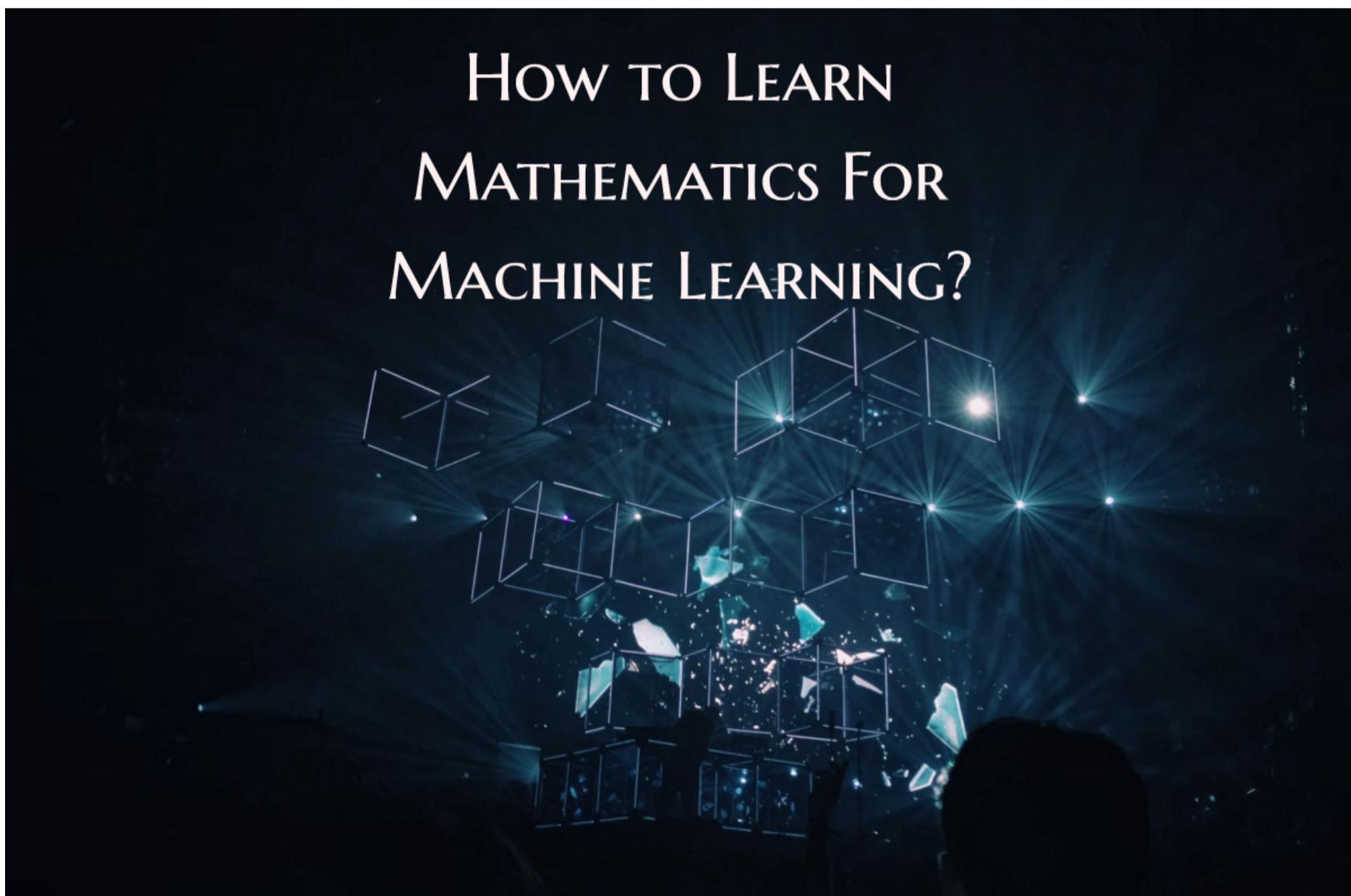


[Ananya Chakraborty](#) – Updated On March 31st, 2023

[Beginner](#) [Education](#) [Interview Questions](#) [Machine Learning](#) [Maths](#) [Probability](#) [Statistics](#)



This article was published as a part of the [Data Science Blogathon](#)



Are you an aspiring data scientist who is fascinated by how things work around in the world of data science and machine learning? Well, congrats on choosing the right career path that is best suited for you at this point in time. However, did you know that you need to ace mathematics for machine learning and data science? Yes, you heard it right.

No matter what kind of love-hate kind of relationship you had with maths back in school. The core concepts used in Maths and Statistics are actually very useful to make strategic decisions while designing machine learning models. So, if you have decided to choose this career path in the field of data science, you need to start loving the concepts of maths and implement them in your future as it is one of the prerequisites for machine learning.

Maths and Statistics for Data Science



Machine learning is all about maths, which in turn helps in creating an algorithm that can learn from data to make an accurate prediction. The prediction could be as simple as classifying dogs or cats from a given set of pictures or what kind of products to recommend to a customer based on past purchases. Hence, it is very important to properly understand the maths concepts behind any central machine learning algorithm. This way, it helps you pick all the right algorithms for your project in data science and machine learning.

Machine learning is primarily built on mathematical prerequisites so as long as you can understand why the maths is used, you will find it more interesting. With this, you will understand why we pick one machine learning algorithm over the other and how it affects the performance of the machine learning model.

Points to be covered in this blog post

1. Which Mathematical Concepts are involved in machine learning?
2. Why do you need maths in machine learning projects?
3. What is the proper way to learn it?

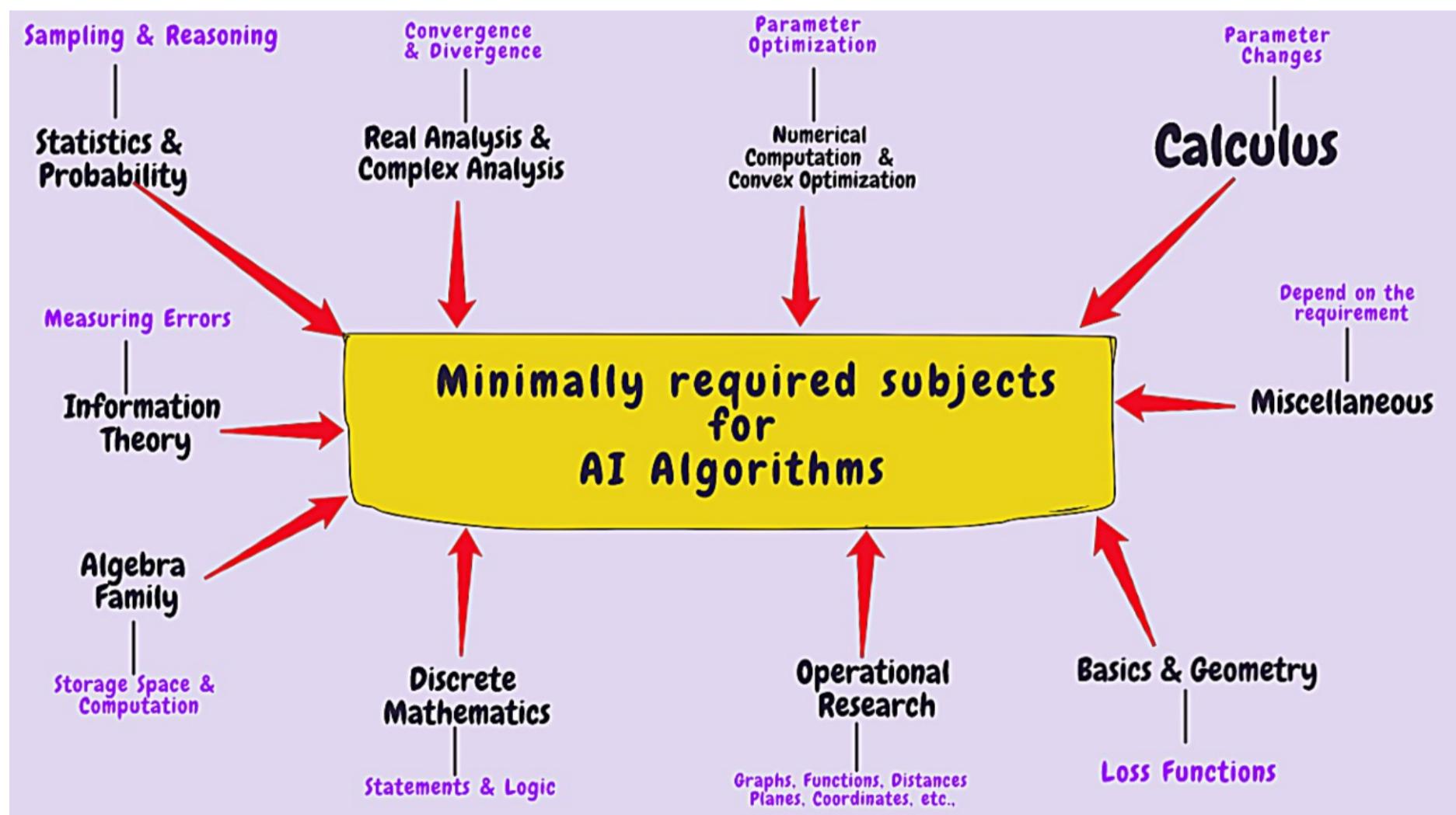
In today's blog post, we will be discussing exactly all the mathematical concepts you need to learn to master the concepts of data science and machine learning. We will also learn why we use mathematics in machine learning with some examples.

Let's start by looking at the many forms of math utilized in data science and machine learning so that you can get a better understanding of what you truly need to know about maths for the data science profession.

Machine learning is powered by four critical concepts and is Statistics, Linear Algebra, Probability, and Calculus. While statistical concepts are the core part of every model, calculus helps us learn and optimize a model. Linear algebra comes exceptionally handy when we are dealing with a huge dataset and probability helps in predicting the likelihood of events that will be occurring. These are the mathematical concepts that you will encounter in your data science and machine learning career quite frequently.

Mathematical Concepts Important for Machine Learning & Data Science:

1. Linear Algebra
2. Calculus
3. Probability Theory
4. Discrete Maths
5. Statistics



Linear Algebra Concept in Machine Learning:

Understanding how to construct linear equations is a fundamental component in developing central machine learning algorithms. These will be used to evaluate and observe data collections. Linear algebra is applied in machine learning algorithms in loss functions, regularisation, covariance matrices, Singular Value Decomposition (SVD), Matrix

learning

In order to perform a Principal Component Analysis that is used to reduce the dimensionality of data, we use linear algebra. Linear algebra is also heavily used in neural networks for the processing and representation of networks. So needless to say, you need to be interested in linear algebra as it is extensively used in the field of data science.

However, don't get intimidated by this as understanding the concepts will be important, but you don't have to be an expert in linear algebra to solve most problems. Only sound knowledge of the concepts will be good enough. Mathematics for Machine Learning by Marc Peter deisenroth is an excellent book to help you get started on this journey if you are struggling with Maths in the beginning.

Calculus in Machine Learning:

Many learners who didn't fancy learning calculus that was taught in school will be in for a rude shock as it is an integral part of machine learning. Thankfully, you may not need to master calculus, it's only important to learn and understand the principles of calculus. Also, you need to understand the practical applications of machine learning through calculus during model building.

So, if you understand how the derivative of the function returns its rate of change in calculus, then you will be able to understand the concept of gradient descent. In gradient descent, we need to find the local minima for a function and so on. If you happen to have saddle points or multiple minima, a gradient descent might find out a local minima and not a global minima, unless you start from multiple points. Some of the necessary topics to ace the calculus part in data science are Differential and Integral Calculus, Partial Derivatives, Vector-Values Functions, Directional Gradients.

Multivariate calculus is utilized in algorithm training as well as in gradient descent. Derivatives, divergence, curvature, and quadratic approximations are all important concepts you can learn and implement.

The mathematics of machine learning might seem intimidating to you right now, however, you will be able to understand the concepts of calculus that are required to build a successful machine learning model within few days of constructive learning.

Use of Descriptive Statistics

Descriptive statistics is a critical concept that every aspiring data scientist needs to learn to understand machine learning when working with classifications like logistic regression, distributions, discrimination analysis, and hypothesis testing.

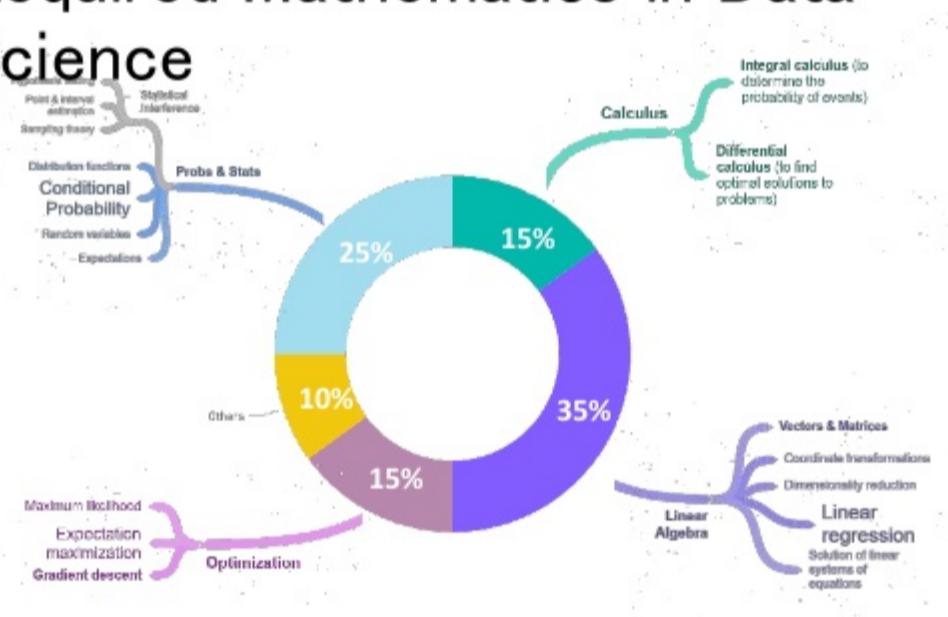
If you were struggling with Statistics in school then you need to put in your 200 percent to learn the mathematics part of statistics as it is very essential for you to become a successful data scientist. To put it down in simpler words, statistics is the main part of mathematics for machine learning. Some of the fundamental statistics needed for ML are Combinatorics, Axioms, Bayes' Theorem, Variance and Expectation, Random Variables, Conditional, and Joint Distributions.

the use of discrete numbers. When scheduling a taxi fleet, for example, you cannot send 0.34 taxis; you must send complete ones. You can't have half a postman or make him visit 1 and a half places to deliver the letters.

Many of the structures in artificial intelligence are discrete. A neural network, for example, has an integer number of nodes and interconnections. It can't have 0.65 nodes or a ninth of a link. As a result, the mathematics used to construct a neural network must include a discrete element, the integer representing the number of nodes and interconnections.

You can get away with just the fundamentals of discrete math for machine learning unless you wish to work with relational domains, graphical models, combinatorial problems, structured prediction, and so on. To master these concepts you have to refer to books on discrete maths. Luckily for computer science graduates, these concepts are properly covered in their college. However, others may have to put additional efforts to understand this subject. Hence, discrete mathematics is a very important component of AI & ML.

Required Mathematics in Data Science



Probability Theory in Machine Learning

To properly work through a machine learning predictive modeling project, it would be reasonable to conclude that probability is essential.

Machine learning is the process of creating prediction models from ambiguous data. Working with faulty or incomplete information is what uncertainty entails.

Uncertainty is crucial to machine learning, yet it is one of the components that creates the most difficulties for newcomers, particularly those coming from a programming background.

In machine learning, there are three major sources of uncertainty: noisy data, limited coverage of the problem area, and of course imperfect models. However, with the help of the right probability tools, we can estimate the solution to the problem.

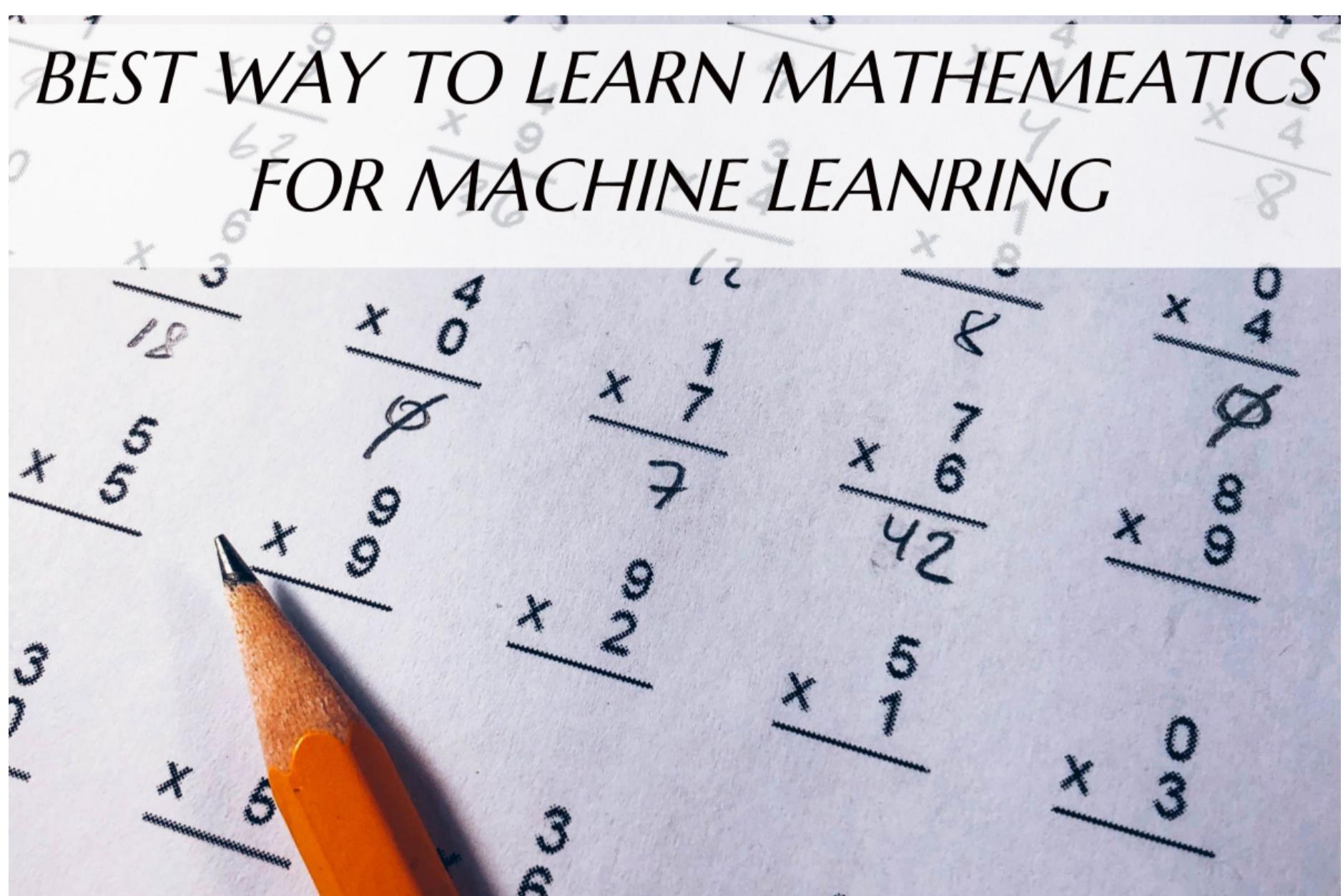
Probability is essential for hypothesis testing and distributions like the Gaussian distribution and the probability density function.

Let us now look at the applications once we have looked at the types in math and data science.

There are numerous reasons why mathematics for Machine Learning is significant, and I will be sharing a few of the important pointers below:

- Choosing the best algorithm requires taking into account accuracy, training time, model complexity, number of parameters, and number of features.
- Choosing parameter values and validation methods.
- Understanding the Bias-Variance tradeoff allows you to identify underfitting and overfitting issues that normally occur while executing the program.
- Determining the correct confidence interval and uncertainty.

What is the proper way to learn Maths For Data Science And Machine Learning?



Although there are plenty of valuable resources available on the internet which explains concepts like matrix decompositions vector calculus, linear algebra analytic geometry matrix, maths behind the principal component analysis, and support vector machines. Not all resources are a one-stop solution for your understanding. Hence, I have collated a list of books, websites, and youtube channels that can help you better your theoretical concept in the field of artificial intelligence.

~~Explained. All the concepts of mathematics have been properly explained. You can refer to the online part here~~

<https://mml-book.github.io/book/mml-book.pdf>

2. [Multivariate Calculus by Imperial College London](#) – Imperial College London has basically come up with a YouTube series that covers the important concepts of multivariate calculus and its application in various ml algorithms. Although the entire course is in collaboration with Coursera, Imperial College London has made it available for free for all the inquisitive learners.
3. Khan Academy's courses on [Linear Algebra](#), [Probability & Statistics](#), [Multivariable Calculus](#), and [Optimization](#)– A very comprehensive and free resource available for all the learners to further their knowledge in complex concepts like linear algebra analytic geometry matrix.
4. [All of statistics: A Concise Course in Statistical Inference](#) by Larry Wasserman is supposedly another exhaustive resource that contains a detailed explanation of important concepts like
5. Udacity's [Introduction to Statistics](#)– is another free resource through which you can get an initial level of understanding in the field of statistics that is needed for data science.

Conclusion

It will take you about 3-4 months to learn the mathematical concepts and put them to practical use. Please refer to the above-mentioned resources and don't forget to keep learning it side by side with the machine learning algorithms so that you can understand which is the right algorithm that you need to pick for your model.

The media shown in this article are not owned by Analytics Vidhya and are used at the Author's discretion.

[blogathon](#) [mathematics for data science](#)

About the Author



[Ananya Chakraborty](#)

Our Top Authors



view more

Download

Analytics Vidhya App for the Latest blog/Article



One thought on "How to Learn Mathematics For Machine Learning? What Concepts do You Need to Master in Data Science?"



Shamsuddeen says:

March 14, 2023 at 6:50 pm

Want to start learning machine language

[Reply](#)

Leave a Reply

Your email address will not be published. Required fields are marked *

Comment

Name*

Email*

Website

Notify me of follow-up comments by email. Notify me of new posts by email.

[Submit](#)

Top Resources



[10 Best AI Image Generator Tools to Use in 2023](#)

[Nitika Sharma - AUG 17, 2023](#)

