**Mathematics for AI / ML**

**Using**

**Python**

Table of Contents

[1 Introduction 3](#_Toc26092371)

[2 Glossary 3](#_Toc26092372)

[3 Audience 3](#_Toc26092373)

[4 Making sense of the mathematics and statistics involved 3](#_Toc26092374)

[5 Don’t have a maths background, so what do I do 3](#_Toc26092375)

[6 The challenges for people with a Maths background as well 3](#_Toc26092376)

# Introduction

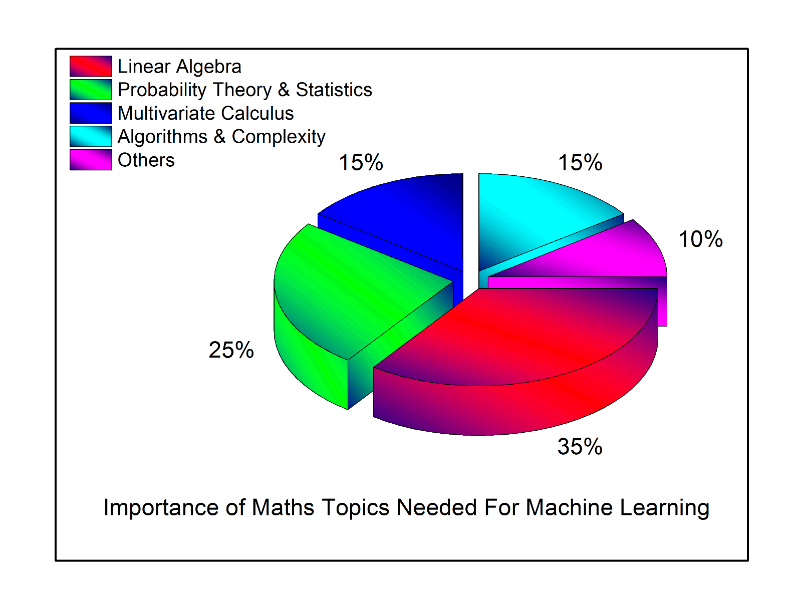
# Glossary

# Audience

# Making sense of the mathematics and statistics involved

# Don’t have a maths background, so what do I do

# The challenges for people with a Maths background as well



1. Linear Algebra: A colleague, [Skyler Speakman](https://ke.linkedin.com/in/skyler-speakman-9a61415), recently said that “Linear Algebra is the mathematics of the 21st century” and I totally agree with the statement. In ML, Linear Algebra comes up everywhere. Topics such as Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Eigendecomposition of a matrix, LU Decomposition, QR Decomposition/Factorization, Symmetric Matrices, Orthogonalization & Orthonormalization, Matrix Operations, Projections, Eigenvalues & Eigenvectors, Vector Spaces and Norms are needed for understanding the optimization methods used for machine learning. The amazing thing about Linear Algebra is that there are so many online resources. I have always said that the traditional classroom is dying because of the vast amount of resources available on the internet. My favorite Linear Algebra course is the one offered by [MIT Courseware](http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/) (Prof. Gilbert Strang).
2. Probability Theory and Statistics: Machine Learning and Statistics aren’t very different fields. Actually, someone recently defined Machine Learning as ‘doing statistics on a Mac’. Some of the fundamental Statistical and Probability Theory needed for ML are Combinatorics, Probability Rules & Axioms, Bayes’ Theorem, Random Variables, Variance and Expectation, Conditional and Joint Distributions, Standard Distributions (Bernoulli, Binomial, Multinomial, Uniform and Gaussian), Moment Generating Functions, Maximum Likelihood Estimation (MLE), Prior and Posterior, Maximum a Posteriori Estimation (MAP) and Sampling Methods.
3. Multivariate Calculus: Some of the necessary topics include Differential and Integral Calculus, Partial Derivatives, Vector-Values Functions, Directional Gradient, Hessian, Jacobian, Laplacian and Lagragian Distribution.
4. Algorithms and Complex Optimizations: This is important for understanding the computational efficiency and scalability of our Machine Learning Algorithm and for exploiting sparsity in our datasets. Knowledge of data structures (Binary Trees, Hashing, Heap, Stack etc), Dynamic Programming, Randomized & Sublinear Algorithm, Graphs, Gradient/Stochastic Descents and Primal-Dual methods are needed.
5. Others: This comprises of other Math topics not covered in the four major areas described above. They include Real and Complex Analysis (Sets and Sequences, Topology, Metric Spaces, Single-Valued and Continuous Functions, Limits, Cauchy Kernel, Fourier Transforms), Information Theory (Entropy, Information Gain), Function Spaces and Manifolds.

## 