

Machine Learning

Mathematics

Roadmap

Linear Algebra, Statistics, Probability, Objective Functions,
Regularization, Information Theory, Optimization, Distribution

Himanshu Ramchandani

<https://www.linkedin.com/in/hemansnation/>

Contents

FREE Resources -

Chapter 1 - Linear Algebra

Chapter 2 - Statistics

Chapter 3 - Probability

Chapter 4 - Objective Functions

Chapter 5 - Regularization

Chapter 6 - Information Theory

Chapter 7 - Optimization

Chapter 8 - Distribution

This phase is different from books that are available on the internet. I included all the topics required to understand the whole architecture of Machine Learning algorithms.

Examples of how and where these mathematical equations are used and Interview Questions can be asked of an ML Engineer while hiring.

We will learn different concepts individually, converting mathematical equations into Python programming expressions along with their examples in the real world.

FREE Resources →

[Mathematics for Machine Learning](#)

Algebra, Topology, Differential Calculus, and Optimization Theory For Computer Science and Machine Learning

All math topics for Machine Learning by Stanford

[Stanford CS229: Machine Learning Course | Summer 2019 \(Anand Avati\)](#)

Chapter 1 - Linear Algebra

Learn for FREE - Mathematics for ML - Linear Algebra

[Mathematics for Machine Learning - Linear Algebra](#)

[1 | Vectors](#)

[2 | Matrix](#)

[3 | Eigenvalues and Eigenvectors](#)

[3 | Factorization](#)

[4 | Singular Value Decomposition \(SVD\)](#)

[5 | Gradient](#)

[6 | Tensors](#)

[7 | Jacobian Matrix](#)

[8 | Curse of Dimensionality](#)

Chapter 2 - Statistics

The Element of Statistical Learning

[Elements of Statistical Learning: data mining, inference, and prediction. 2nd Edition.](#)

Statistics give us 2 tools descriptive and inferential

1 | Descriptive Statistics

[1 | Variables](#)

[2 | Mean](#)

[3 | Median](#)

[4 | Mode](#)

[5 | Standard Deviation](#)

[6 | Variance](#)

[7 | Range](#)

[8 | Percentile](#)

[9 | Skewness](#)

[10 | Kurtosis](#)

2 | Inferential Statistics

[1 | Sampling Distributions](#)

[2 | Central Limit Theorem](#)

[3 | Hypothesis Testing](#)

[4 | Confidence Intervals](#)

[5 | T-Tests](#)

[6 | Analysis of Variance \(ANOVA\)](#)

[7 | Chi-Square Test](#)

[8 | Regression Analysis](#)

[9 | Bayesian Inference](#)

[10 | Maximum Likelihood Estimation \(MLE\)](#)

Chapter 3 - Probability

Probability Theory: The Logic of Science

<https://bayes.wustl.edu/etj/prob/book.pdf>

[1 | Probability Distribution](#)

[2 | Conditional Probability](#)

[3 | Bayes' Theorem](#)

[4 | Joint and Marginal Probabilities](#)

[5 | Independence and Conditional Independence](#)

Chapter 4 - Objective Functions

[1 | Mean Squared Error \(MSE\)](#)

[2 | Mean Absolute Error \(MAE\)](#)

[3 | Huber Loss](#)

[4 | Binary Cross-Entropy \(Log Loss\)](#)

[5 | Categorical Cross-Entropy](#)

[6 | Maximum Likelihood Estimation \(MLE\)](#)

[7 | Sparse Categorical Cross-Entropy](#)

[8 | Hinge Loss](#)

[9 | Kullback-Leibler Divergence](#)

[10 | Gini Impurity](#)

[11 | Others](#)

Chapter 5 - Regularization

[1 | L1 Regularization \(Lasso Regression\)](#)

[2 | L2 Regularization \(Ridge Regression\)](#)

[3 | Elastic Net Regularization](#)

[4 | Dropout Regularization](#)

[5 | Data Augmentation](#)

[6 | Early Stopping](#)

[7 | Max-Norm Regularization](#)

[8 | Batch Normalization](#)

[9 | Weight Decay](#)

Chapter 6 - Information Theory

Information Theory, Inference and Learning Algorithms

[David MacKay: Information Theory, Pattern Recognition and Neural Networks: The Book](#)

[1 | Entropy](#)

[2 | Conditional Entropy](#)

[3 | Joint Entropy](#)

[4 | Mutual Information](#)

[5 | Relative Entropy \(Kullback-Leibler Divergence\)](#)

[6 | Cross-Entropy](#)

[7 | Information Gain](#)

[8 | Shannon-Fano Coding](#)

[9 | Huffman Coding](#)

[10 | Data Entropy](#)

Chapter 7 - Optimization

- [1 | Gradient Descent](#)
- [2 | Stochastic Gradient Descent \(SGD\)](#)
- [3 | Mini-Batch Gradient Descent](#)
- [4 | Momentum](#)
- [5 | Nesterov Accelerated Gradient \(NAG\)](#)
- [6 | Adagrad \(Adaptive Gradient Algorithm\)](#)
- [7 | RMSprop \(Root Mean Square Propagation\)](#)
- [8 | Adam \(Adaptive Moment Estimation\)](#)

Chapter 8 - Distribution

- [1 | Bernoulli Distribution](#)
- [2 | Binomial Distribution](#)
- [3 | Multinomial Distribution](#)
- [4 | Normal \(Gaussian\) Distribution](#)
- [5 | Uniform Distribution](#)
- [6 | Exponential Distribution](#)
- [7 | Poisson Distribution](#)

Calculus

[Calculus 1 | Math | Khan Academy](#)

Machine Learning, MLOps & GenerativeAI Roadmap

<https://god-level-python.notion.site/Build-a-Strong-Machine-Learning-Portfolio-Personal-Brand-Get-Tons-of-Job-Offers-in-12-Weeks-Live-b3c98407b4ab45819811db081ae9d102?pvs=4>

About me



I am [Himanshu Ramchandani](#) a Data & Engineering Consultant. I help enterprises utilize big data to build AI-powered products & Mentor professionals to improve their skills in the data field by 1% every day.

the epoch → an AI Newsletter

- Leverage Data, Products & AI in 3 min.
- Top 2 AI news & developments.
- 1 Action Tip from Experts in BigData Analytics, Data Engg & ML.
- AI Investments.
- Career & Jobs.

Join the tribe of 20,000+ Entrepreneurs, Tech Leaders, Data Professionals & Devs.

Subscribe to the newsletter here:

<https://the-epoch-by-himanshu-ramchandani.beehiiv.com/>

Join the Discord Community:

<https://discord.gg/2Rb9HCpJG>