Mathematical Roadmap for Machine Learning

By Sunil Ghimire- "**Your legacy will never be erased**"

Hello everyone, I hope you all are doing well. Knowing the mathematics behind Machine Learning algorithms is a superpower. If you have ever built a model for a real-life problem, you probably experienced that being familiar with the details can go a long way if you want to move beyond baseline performance. This is especially true when you want to push the boundaries of the state of art.

However, most of the knowledge used in the Machine Learning model is hidden layers of advanced mathematics. Understanding methods used in Machine Learning algorithms might seem difficult since it is built on the top of **Calculus**, **Statistics**, **Linear Algebra**, and **Probability** theory.

Mathematics for Machine Learning covers the field of Statistics, Probability, Multivariable Calculus, Linear Algebra, Discrete Maths, Optimization. These are the major ones required to give any beginner a kickstart.

I have followed this roadmap during my post-grad Data Science course and it has benefitted me immensely to prepare for the ML interviews.

A. Statistics Roadmap

1. STATS BASICS

- ➔ Types of Data: Nominal, Ordinal, Discrete, Continuous
- ➔ Descriptive vs Inferential Stats
- ➔ Moments
- ➔ Mean, Median, Mode
- ➔ Skewness
- ➔ Kurtosis
- ➔ Range, IQR
- ➔ Percentiles, Quartiles
- ➔ Mean Deviation
- ➔ Standard Deviation
- ➔ Variance
- ➔ Quartile Deviation
- ➔ Standard Error

2. CHARTS

- ➔ Frequency Distribution Table
- ➔ Line Chart
- ➔ Bar Chart
- ➔ Histogram
- ➔ Frequency Polygon
- ➔ Pie Chart
- ➔ Ogives

3. PROBABILITY DISTRIBUTION FUNCTIONS

- ➔ Random Variables
- ➔ Multivariate random variables
- ➔ Discrete random variables
- ➔ Continuous random variables
- ➔ Law of Large Numbers
- ➔ Expectation
- ➔ PMF — Probability Mass Function
- ➔ PDF — Probability Density Function
- ➔ CDF — Cumulative Density Function
- ➔ Bernoulli Distribution
- ➔ Binomial Distribution
- ➔ Geometric Distribution
- ➔ Poisson Distribution
- ➔ Exponential Distribution
- ➔ Uniform Distribution
- ➔ Gaussian / Normal Distribution
- ➔ Chi-Square Distribution
- ➔ Power Law Distribution
- ➔ Pareto Distribution
- ➔ Box-Cox Transformation
- ➔ Log-Normal Distribution
- ➔ Kernel Density Estimation
- ➔ Q-Q plot

4. PROBABILITY

- ➔ Basic Probability
- ➔ Joint Probability
- ➔ Conditional Probability
- ➔ Independent Events
- ➔ Mutually Exclusive Events
- ➔ Bayes' Theorem

5. TEST / SAMPLING / POPULATION

- ➔ Sampling, Sample Mean & Distribution
- ➔ Central Limit Theorem
- ➔ Point estimate, Interval estimate
- ➔ Confidence Interval
- ➔ Population, Population Mean & Distribution
- ➔ Hypothesis Testing
- ➔ P-value
- ➔ Population Proportions
- ➔ Critical Value
- ➔ Significance Level
- ➔ Rejection regions
- ➔ Type I vs Type II errors
- ➔ One tail vs Two tail
- ➔ Z-Test
- ➔ T-Test
- ➔ ANOVA
- ➔ F-Test
- ➔ Chi-Square Test
- ➔ Monte Carlo Simulation
- ➔ A/B Testing

6. RELATIONS / REGRESSION

- ➔ Causality
- ➔ Covariance
- ➔ Covariance Matrix
- ➔ Correlation

- ➔ Scatter Plots
- ➔ Pearson Correlation Coefficient
- ➔ Rank / Spearman Correlation Coefficient
- ➔ R2 score
- ➔ Linear Regression
- ➔ OLS
- ➔ Factor Analysis
- ➔ Logistic Regression

B. Linear Algebra Roadmap

1. LINEAR EQUATIONS

- ➔ Systems of Linear Equations
- ➔ Gaussian Elimination
- ➔ Echelon Form
- ➔ Linear Combination
- ➔ Span
- ➔ Homogeneous Linear System
- ➔ Linear Independence
- ➔ Subspace
- ➔ Basis
- ➔ Affine space
- ➔ Linear Transformation

2. MATRIX

- ➔ Matrix transformations
- ➔ Matrix multiplication
- ➔ Inverse Matrix
- ➔ Transpose of a matrix
- ➔ Rank of a matrix
- ➔ Symmetric Matrix
- ➔ Orthogonal Matrix
- ➔ Adjoint Matrix
- ➔ Singular Matrix
- ➔ Determinant of a matrix

- ➔ Trace of a Matrix

3. VECTORS

- ➔ Components of Vector
- ➔ Vector Space
- ➔ Norm of a vector
- ➔ Lengths and distances
- ➔ Euclidean Norm
- ➔ Manhattan Norm
- ➔ Minkowski Distance
- ➔ Scalar Multiplication
- ➔ Dot Product
- ➔ Inner Product
- ➔ Cross Product
- ➔ Orthogonality
- ➔ Orthonormal
- ➔ Rotations

4. FACTORIZATION

- ➔ Matrix Decomposition
- ➔ LU Decomposition
- ➔ QR Decomposition
- ➔ Cholesky Decomposition
- ➔ Eigen Decomposition
- ➔ Eigen Values
- ➔ Eigen Vector
- ➔ Singular Value Decomposition
- ➔ Principal Component Analysis

C. Calculus Roadmap

1. CALCULUS BASICS

- ➔ Functions
- ➔ Derivatives
- ➔ Maxima Minima
- ➔ Product and Chain Rule Differentiation

- ➔ Composite functions
- ➔ Partial Derivatives
- ➔ Higher-order derivatives
- ➔ Integrals
- ➔ Limits
- ➔ Infinite series summation

2. OPTIMIZERS

- ➔ Gradient Descents
- ➔ Optimizers
- ➔ Loss Functions
- ➔ Taylor's Series
- ➔ Constrained Optimization (Lagrange Multiplier)
- ➔ Newton's method in Optimization
- ➔ Convex Optimization

Credits to 3Blue1Brown, Khan Academy, StatQuest with Josh Starmer.