DATA SCIENCE | ADVANCED | 5 DAYS

Computational Mathematics for Data Science

Build a strong foundation in statistics and machine learning for data science, covering probability, distributions, and hypothesis testing. Learn how to apply regression techniques, decision trees, and statistical methods to analyze data and support informed business decisions.

Why Computational Mathematics?

Deeper Data Insights

Employees and teams with computational expertise can analyze and interpret more complex datasets, allowing them to assist in better decision-making.

More Effective Problem-Solving

Computational maths enhances problem-solving capabilities. Gaining expertise in this area will enable your employees to tackle data-related challenges with unique perspectives, allowing them to expedite data analysis.

Improved Predictive Modeling

Proficiency in computational maths enables your employees to design, build, and apply algorithms and predictive models, helping you forecast future trends accurately.

Hassle-Free AI/ML Integrations

Familiarity with computational mathematics allows your employees to understand and optimize complex AI/ML algorithms used to deal with diverse data types. This will simplify future integrations of AI/ML technologies in your existing workflows, ensuring seamless employee adoption and support.

What you'll be able to do

- ✓ Develop an understanding of fundamental statistics for data science and analytics
- ✓ Master key principles of descriptive statistics and probability theory
- ✓ Learn and apply machine learning methods, including decision trees and decision forests.
- ✓ Gain insights into various probability distributions, such as normal and Poisson distributions
- ✓ Understand the fundamentals of hypothesis testing, including concepts like p-value, type I, and type II errors
- ✓ Get hands-on experience with logistic regression, multiple linear regression, and regression trees

Who this is for

- Data and Quantitative Analysts
- Cryptographers
- Software Developers
- Machine Learning Engineers

Curriculum

01 Quick Start

- An overview of the program, covering key mathematical concepts and tools for data science

02 Introduction to Descriptive Statistics

- Overview of key concepts and goals
- The average value (Mean) in a dataset
- The middle value (Median) of ordered data
- The most frequent value (Mode)
- Guidelines on when to use mean, median, or mode
- Asymmetry (Skewness) in data distributions
- Data spread and variability using Range and Interquartile Range
- Population vs sample
- Data dispersion and consistency
- Impact of scaling & shifting on statistical measures

03 Getting Equipped with Various Types of Distributions

- Overview of data distributions
- Normal Distribution
- Standardization using Z-Score
- Statistical Distributions

04 Understanding Probability and Its Background Theory

- Fundamental probability concepts
- Rule of Addition for calculating combined probabilities
- Law of Multiplication to find the probability of multiple events
- Bayes Theorem and Expected Value
- Central Limit Theorem as a distribution of sample means
- Binomial Distribution
- Poisson Distribution for rare events

05 Hypothesis Testing

- What is it and why is it done
- Significance Level and p Value with Type Error
- Population parameters - Confidence Interval and Margin of Error
- Hypothesis Test, T-Test, and T Distribution
- Law of Large Numbers

06 Linear Regression and Advanced Regression

- Relationship modeling using Linear Regression
- Relationship strength with Correlation Coefficient
- Errors - Residual, MSE, and MAE
- Explained variance by Coefficient of Determination
- Model accuracy with RMS Error

07 Machine Learning Algorithms

- Multiple Linear Regression for predicting outcomes
- How to avoid model overfitting in data analysis
- Polynomial Regression for nonlinear relationships
- Decision Tree for classification and regression tasks
- Regression Tree for predicting continuous outcomes
- Enhancing predictions with Random Forest ensemble

08 Analysis of Variance

- Basics of ANOVA and assumption
- One-way ANOVA for means across groups
- F Distribution for variance ratios
- Two-way ANOVA and Sum of Squares