

# MSML - MACHINE LEARNING

---

## MSML601 Probability and Statistics (3 Credits)

Provides a solid understanding of the fundamental concepts of probability theory and statistics. The course covers the basic probabilistic concepts such as probability space, random variables and vectors, expectation, covariance, correlation, probability distribution functions, etc. Important classes of discrete and continuous random variables, their inter-relation, and relevance to applications are discussed. Conditional probabilities, the Bayes formula, and properties of jointly distributed random variables are covered. Limit theorems, which investigate the behavior of a sum of a large number of random variables, are discussed. The main concepts random processes are then introduced. The latter part of the course concerns the basic problems of mathematical statistics, in particular, point and interval estimation and hypothesis testing.

**Prerequisite:** Undergraduate courses in calculus and basic linear algebra. Cross-listed with: DATA601, BIOI601.

**Credit Only Granted for:** BIOI601, DATA601 or MSML601.

## MSML602 Principles of Data Science (3 Credits)

An introduction to the data science pipeline, i.e., the end-to-end process of going from unstructured, messy data to knowledge and actionable insights. Provides a broad overview of what data science means and systems and tools commonly used for data science, and illustrates the principles of data science through several case studies.

**Restriction:** Must be in one of the following programs: (Data Science Post-Baccalaureate Certificate, Master of Professional Studies in Data Science and Analytics, or Master of Professional Studies in Machine Learning). Cross-listed with: DATA602, BIOI602.

**Credit Only Granted for:** BIOI602, DATA602, MSML602 or CMSC641.

**Formerly:** CMSC641.

## MSML603 Principles of Machine Learning (3 Credits)

A broad introduction to machine learning and statistical pattern recognition. Topics include: Supervised learning: Bayes decision theory, discriminant functions, maximum likelihood estimation, nearest neighbor rule, linear discriminant analysis, support vector machines, neural networks, deep learning networks. Unsupervised learning: clustering, dimensionality reduction, PCA, auto-encoders. The course will also discuss recent applications of machine learning, such as computer vision, data mining, autonomous navigation, and speech recognition.

**Restriction:** Must be in one of the following programs: (Data Science Post-Baccalaureate Certificate, Master of Professional Studies in Data Science and Analytics, or Master of Professional Studies in Machine Learning). Cross-listed with: DATA603, BIOI603.

**Credit Only Granted for:** BIOI603, DATA603, MSML603 or CMSC643.

**Formerly:** CMSC643.

## MSML604 Introduction to Optimization (3 Credits)

Focuses on recognizing, solving, and analyzing optimization problems. Linear algebra overview: vector spaces and matrices, linear transformations, matrix algebra, projections, similarity transformations, norms, eigen-decomposition and SVD. Convex sets, convex functions, duality theory and optimality conditions. Unconstrained optimization: 1D search, steepest descent, Newton's method, conjugate gradient method, DFP and BFGS methods, stochastic gradient descent. Constrained optimization: projected gradient methods, linear programming, quadratic programming, penalty functions, and interior-point methods. Global search methods: simulated annealing, genetic algorithms, particle swarm optimization.

**Prerequisite:** Undergraduate courses in calculus and basic linear algebra.

**Recommended:** DATA601.

## MSML605 Computing Systems for Machine Learning (3 Credits)

Programming, software and hardware design and implementation issues of computing systems for machine learning. Topics in the programming/software domain will include: basic Python program structure, variables and assignment, built-in data types, flow control, functions and modules; basic I/O, and file operations. Classes, object-oriented programming and exceptions. Regular expressions, database access, network programming and sockets. Introduction to the Numpy, Scipy and Matplotlib libraries. Topics in the hardware domain include computer architecture, CPUs, single- and multi-core architectures, GPUs, memory and I/O systems, persistent storage, and virtual memory. Parallel processing architectures, multiprocessing and cluster processing.

**Restriction:** Must be in the MPS in Machine Learning program.

## MSML606 Algorithms and Data Structures for Machine Learning (3 Credits)

Provides both a broad coverage of basic algorithms and data structures. Topics include sorting, searching, graph and string algorithms; greedy algorithm, branch-and-bound, dynamic programming and job scheduling; Arrays, linked lists, queues, stacks, and hash tables; Algorithm complexity, best/average/worst case analysis. Applications selected from machine learning problems.

## MSML612 Deep Learning (3 Credits)

Provides an introduction to the construction and use of deep neural networks: models that are composed of several layers of nonlinear processing. The class will focus on the main features in deep neural nets structures. Specific topics include backpropagation and its importance to reduce the computational cost of the training of the neural nets, various coding tools available and how they use parallelization, and convolutional neural networks. Additional topics may include autoencoders, variational autoencoders, convolutional neural networks, recurrent and recursive neural networks, generative adversarial networks, and attention-based models. The concepts introduced will be illustrated by examples of applications chosen among various classification/clustering questions, computer vision, natural language processing.

**Prerequisite:** DATA603 or MSML603. Cross-listed with: DATA612.

**Credit Only Granted for:** DATA612 or MSML612.

**MSML699 Independent Study in Machine Learning (1-6 Credits)**

Students will work on research projects in machine learning with an identified faculty member. This course will develop in-depth knowledge of a practical domain, and give students practical experience with machine learning tools (e.g., deep neural networks, random forests, ...). Students will also gain experience in digesting research papers and independently understanding the state-of-the-art in an application area.

**Prerequisite:** MSML603: Principles of Machine Learning.

**Restriction:** Permission of instructor and departmental approval of project.

**Repeatable to:** 12 credits.