Module 0 - Introduction and Course Information

Welcome to

Machine Learning Fundamentals.

This course provides a thorough grounding in a wide range of machine learning methods, for classification, regression, conditional probability estimation, clustering, and dimensionality reduction.

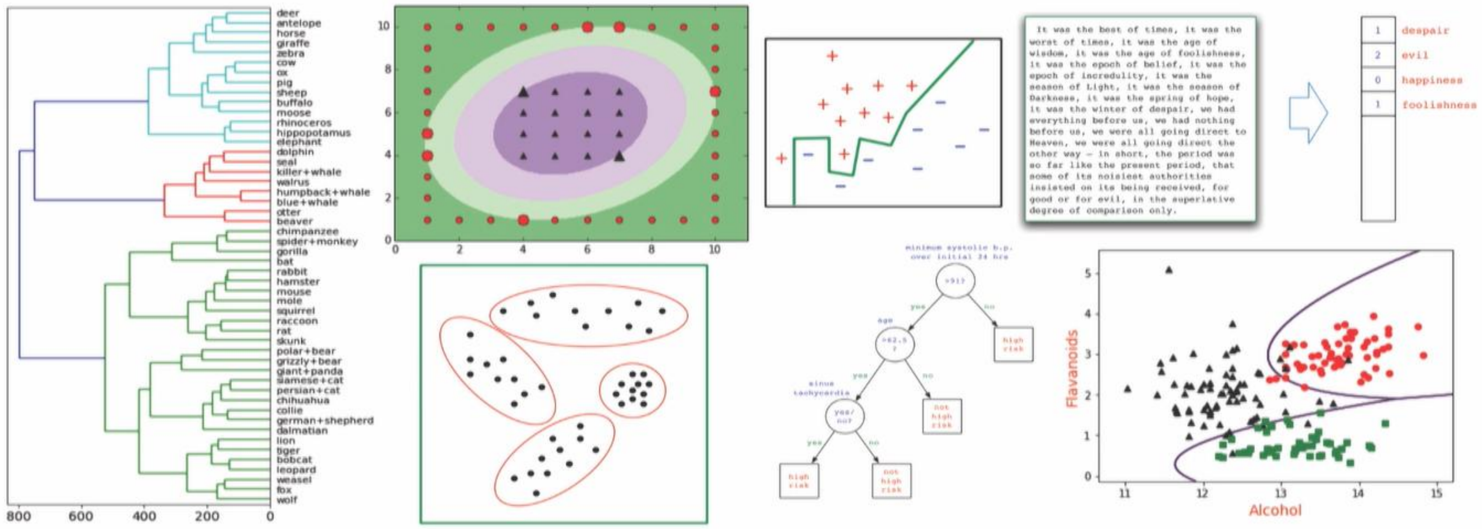
- Taxonomy of prediction problems
- Nearest neighbor methods and families of distance functions
- Generalization: what it means; overfitting; selecting parameters using cross-validation
- Generative modeling for classification, especially using the multivariate Gaussian
- Linear regression and its variants
- Logistic regression
- Optimization: deriving stochastic gradient descent algorithms and testing convexity
- Linear classification using the support vector machine
- Nonlinear modeling using basis expansion and kernel methods
- Decision trees, boosting, and random forests
- Methods for flat and hierarchical clustering
- Principal component analysis
- Autoencoders, distributed representations, and deep learning

Your Instructor

Sanjoy Dasgupta is Professor of Computer Science and Engineering at the University of California, San Diego. He received his A.B. from Harvard in 1993 and his Ph.D. from Berkeley in 2000, both in Computer Science.

总结：

Course Introduction

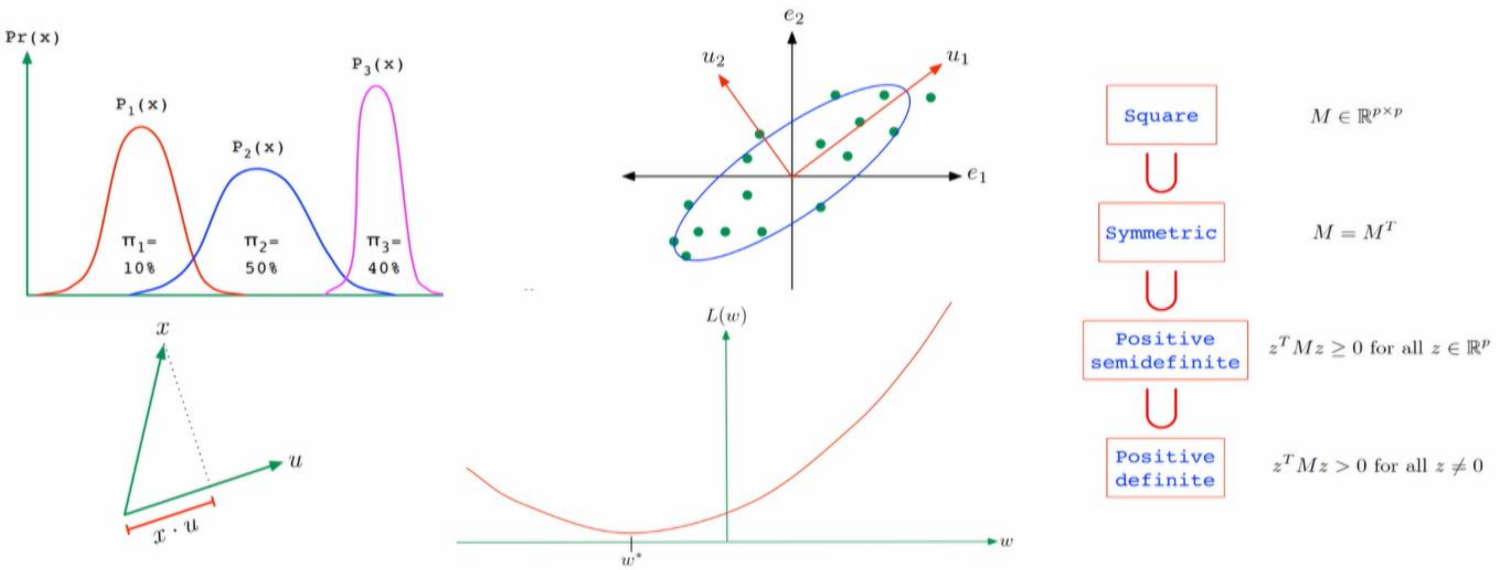


Nearest neighbor	Kernel methods	Hierarchical clustering
Generative models	Decision trees	Principal component analysis
Least-squares regression	Boosting and bagging	Singular value decomposition
Ridge regression, Lasso	Random forests	Autoencoders
Logistic regression	k-means	Deep learning
Support vector machines	Mixtures of Gaussians	

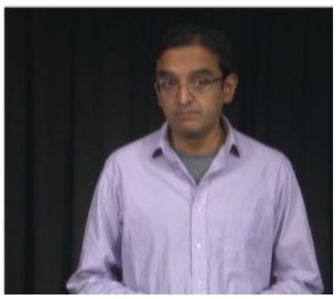
Welcome to DSE 220X: Fundamentals of Machine Learning.

In this course, you will get an intensive grounding in building models from data and using these models to make predictions. We'll cover the most widely used machine learning methods, logistic regression, support vector machines, kernel methods, principal component analysis, random forests, neural nets, and so on. We'll see applications of these methods to problems in image understanding, sentiment analysis, language processing, personality testing, and many other domains. Most of all, we'll really delve into the fundamentals of these methods. We won't think of them as recipes to be blindly followed, but rather as interesting approaches that, once understood, are amenable to all kinds of modifications and improvements.

So in order to get to this level of understanding,we will go into the mathematical foundations of machine learning. There are basically three types of math that we will need: probability, linear algebra, and optimization. We'll develop all three of these areas quite carefully, because they're really crucial to a mastery of machine learning.



Probability	**Linear algebra**	**Optimization**
Probability spaces	Matrices and vectors	Gradient descent
Bayes' rule	Projections	Stochastic gradient descent
Random variables	Positive definiteness	Convex optimization
Mean and variance	Eigendecompositions	Duality
Measuring dependence	Spectral decomposition	



总结:

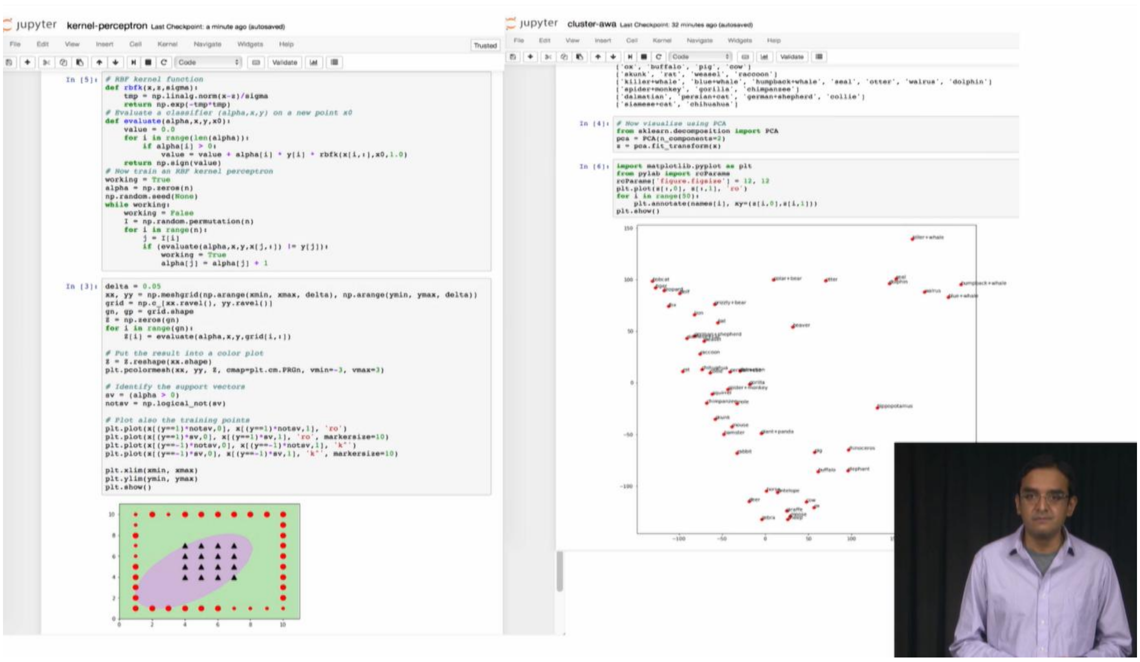
Skills you will acquire

1. Familiarity with most widely-used ML methods
   - How they work
   - The kinds of data they are suited to
   - Their strengths and weaknesses

2. Adapting existing methods to a particular application

3. The foundational knowledge to keep pace with a fast-moving field

So at the end of the day, what are the skillsthat you will acquire from this course? First, a comfortable familiarity with the most widely-used machine learning methods. What are they? How do they work? Why do they work? What kinds of data are they good for? What are their strengths and weaknesses? What is going on under the hood? This will help you assess which method or methods are most suitable for a particular problem. It will also help you understand and interpret the results that you're getting. The second skill that you'll get from this course is the ability to take existing methods and adapt them for a particular application. Sometimes, it's okay to just use an off-the-shelf method, but very often, this gets a level of performance that is not quite what you were hoping for, and in order to get to the next level, what you need is to do something customized. To either build upon an existing method and modify it in some way, or perhaps to transform the data. What are effective ways of doing this? Last, but not least, this course will give a sufficiently thorough grounding in the foundations of machine learning that will help you to keep up with this field that is really evolving very rapidly.



Now for each of the methods that we study,you'll get a Jupiter notebook or two that contains code and visualizations that illustrate the method on some suitable data sets. You can take these and play around with them to really get a feel for how these methods work. There will also be programming projects that will give you the opportunity to think outside the box a little bit and to creatively build upon existing methods to solve new problems. At the end of the course, this collection of projects will essentially constitute a sort of repertoire, a portfolio, of self-contained notebooks that you can show other people. You know, to impress your friends, for example.

Okay, so that's a little bit of an overview of the course. Let me also tell you a little bit about myself. I'm a professor here, at the University of California-San Diego, and I've been here for about 15 years or so. The things I work on are machine learning and algorithms. These are the two areas that I care the most about. And this particular course, DSE 220X, is closely modeled on a class that I teach here, and that basically, I wanted to open up to a much broader audience. So I hope you like it.

总结：

Syllabus

Welcome to Machine Learning Fundamentals!

We are delighted to welcome you to the third course in the MicroMasters in Data Science: Machine Learning Fundamentals. In this course, you will learn the motivation, intuition, and theory behind the probabilistic and statistical foundations of data science, and will get to experiment and practice with these concepts via Python programs and the Jupyter Notebook platform.

Course Staff

Instructor

Sanjoy Dasgupta, Professor, Computer Science and Engineering Department, UC San Diego

What do you need to know to succeed?

The most important pre-requisites for this course are:

The ability to program in Python and to use Jupyter notebooks. This can be obtained by taking the course DSE200x, Python for Data Science.

Familiarity with calculus, especially derivatives of single-variable and multivariate functions.

The course will make heavy use of basic probability and linear algebra. Although we will introduce these concepts as needed, it will be easiest if learners already have some familiarity with them.

Learning objectives

This course is an intensive introduction to the most widely-used machine learning methods. The first goal is to provide a basic intuitive understanding of these techniques: what they are good for, how they work, how they relate to one another, and their strengths and weaknesses. The second goal is to provide a hands-on feel for these methods through experiments with suitable data sets, using Jupyter notebooks. The third goal is to understand machine learning methods at a deeper level by delving into their mathematical underpinnings. This is crucial to being able to adapt and modify existing methods and to creatively combining them.

总结：

Overview

Topics:

- Taxonomy of prediction problems
- Nearest neighbor methods and families of distance functions
- Generalization: what it means; overfitting; selecting parameters using cross-validation
- Generative modeling for classification, especially using the multivariate Gaussian
- Linear regression and its variants
- Logistic regression
- Optimization: deriving stochastic gradient descent algorithms and testing convexity
- Linear classification using the support vector machine
- Nonlinear modeling using basis expansion and kernel methods
- Decision trees, boosting, and random forests
- Methods for flat and hierarchical clustering
- Principal component analysis
- Autoencoders, distributed representations, and deep learning

Course Outline

This is a ten-week course.

- Week 1: Introduction: nearest neighbor, and a host of prediction problems
- Week 2: Probability basics and generative modeling
- Week 3: Linear algebra basics, the multivariate Gaussian, and more generative modeling
- Week 4: Linear regression and logistic regression
- Week 5: Optimization
- Week 6: Support vector machines
- Week 7: Beyond linear prediction: kernel methods, decision trees, boosting, random forests
- Week 8: Clustering
- Week 9: Informative projections
- Week 10: Deep learning

总结：

Week 1: Prediction Problems

- 1.1 Introduction to Machine Learning
- 1.2 Nearest Neighbor Classification
- 1.3 Improving the Performance of Nearest Neighbor
- 1.4 Useful Distance Functions for Machine Learning
- 1.5 A Host of Prediction Problems

Week 2: Generative Modeling I

- 2.1 The Generative Approach to Classification
- 2.2 Probability Review I: Probability Spaces, Events, and Conditioning
- 2.3 Generative Modeling in One Dimension
- 2.4 Probability Review II: Random Variables, Expectation, and Variance
- 2.5 Probability Review III: Measuring Dependence
- 2.6 Two Dimensional Generative Modeling with the Bivariate Gaussian

Week 3: Generative Modeling II

- 3.1 Linear Algebra I: Basic Notation and Dot Products
- 3.2 Linear Algebra II: Matrix Products and Linear Functions
- 3.3 Linear Algebra III: Square Matrices as Quadratic Functions
- 3.4 The Multivariate Gaussian
- 3.5 Gaussian Generative Models
- 3.6 More Generative Modeling

Week 4: Linear Regression and Probability Estimation

- 4.1 An Introduction to Linear Regression
- 4.2 Linear Regression
- 4.3 Regularized Linear Regression
- 4.4 Linear Models for Conditional Probability Estimation
- 4.5 Logistic Regression
- 4.6 Logistic Regression in Use

Week 5: Optimization and Geometry

- 5.1 Unconstrained Optimization I
- 5.2 Unconstrained Optimization II
- 5.3 Unconstrained Optimization III
- 5.4 Convexity I
- 5.5 Linear algebra IV: Positive Semidefinite Matrices
- 5.6 Convexity II

总结：

Week 6: Linear Classification

- 6.1 A Simple Linear Classifier
- 6.2 Support Vector Machines I
- 6.3 Support Vector Machines II
- 6.4 Duality
- 6.5 Multiclass Linear Prediction

Week 7: Combining Simple Classifiers

- 7.1 Kernels I: Basis Expansion
- 7.2 Kernels II: The Kernel Trick
- 7.3 Kernels III: Kernel SVM
- 7.4 Kernels IV: The Kernel Function
- 7.5 Combining Simple Classifiers
- 7.6 Decision Trees
- 7.7 Boosting
- 7.8 Random Forests

Week 8: Representation Learning I

- 8.1 Representation Learning, Incomplete
- 8.2 Clustering with the k-Means Algorithm I
- 8.3 Clustering with the k-Means Algorithm II
- 8.4 Clustering with Mixtures of Gaussians
- 8.5 Hierarchical Clustering

Week 9: Representation Learning II

- 9.1 Linear Projections
- 9.2 Principal Component Analysis I: One Dimensional Projection
- 9.3 Principal Component Analysis II: The Top k Directions
- 9.4 Case Study: Personality Assessment
- 9.5 Linear Algebra V: Eigenvalues and Eigenvectors
- 9.6 Linear Algebra VI: Spectral Decomposition

Week 10: Deep Learning

- 10.1 Autoencoders
- 10.2 Distributed Representations
- 10.3 Feedforward Neural Networks
- 10.4 Training Neural Networks, Incomplete
- 10.5 What We Skipped

总结：