WEEK 1

Introduction

Welcome to Machine Learning! In this module, we introduce the core idea of teaching a computer to learn concepts using data—without being explicitly programmed. The Course Wiki is under construction. Please visit the resources tab for the most complete and up-to-date information.

5 videos, 9 readings

expand

Video: Welcome to Machine Learning!

Reading: Machine Learning Honor Code

Video: Welcome

Video: What is Machine Learning?

Reading: What is Machine Learning?

Reading: How to Use Discussion Forums

Video: Supervised Learning

Reading: Supervised Learning

Video: Unsupervised Learning

Reading: Unsupervised Learning

Reading: Who are Mentors?

Reading: Get to Know Your Classmates

Reading: Frequently Asked Questions

Reading: Lecture Slides

Graded: Introduction

Linear Regression with One Variable

Linear regression predicts a real-valued output based on an input value. We discuss the application of linear regression to housing price prediction, present the notion of a cost function, and introduce the gradient descent method for learning.

7 videos, 8 readings

expand

Video: Model Representation

Reading: Model Representation

Video: Cost Function

Reading: Cost Function

Video: Cost Function - Intuition I

Reading: Cost Function - Intuition I

Video: Cost Function - Intuition II

Reading: Cost Function - Intuition II

Video: Gradient Descent

Reading: Gradient Descent

Video: Gradient Descent Intuition

Reading: Gradient Descent Intuition

Video: Gradient Descent For Linear Regression

Reading: Gradient Descent For Linear Regression

Reading: Lecture Slides

Graded: Linear Regression with One Variable

Linear Algebra Review

This optional module provides a refresher on linear algebra concepts. Basic understanding of linear algebra is necessary for the rest of the course, especially as we begin to cover models with multiple variables.

6 videos, 1 reading, 1 practice quiz

expand

Video: Matrices and Vectors

Video: Addition and Scalar Multiplication

Video: Matrix Vector Multiplication

Video: Matrix Matrix Multiplication

Video: Matrix Multiplication Properties

Video: Inverse and Transpose

Reading: Lecture Slides

Practice Quiz: Linear Algebra

WEEK 2

Linear Regression with Multiple Variables

What if your input has more than one value? In this module, we show how linear regression can be extended to accommodate multiple input features. We also discuss best practices for implementing linear regression.

8 videos, 16 readings

expand

Reading: Setting Up Your Programming Assignment Environment

Reading: Installing MATLAB

Reading: Installing Octave on Windows

Reading: Installing Octave on Mac OS X (10.10 Yosemite and 10.9 Mavericks and Later)

Reading: Installing Octave on Mac OS X (10.8 Mountain Lion and Earlier)

Reading: Installing Octave on GNU/Linux

Reading: More Octave/MATLAB resources

Video: Multiple Features

Reading: Multiple Features

Video: Gradient Descent for Multiple Variables

Reading: Gradient Descent For Multiple Variables

Video: Gradient Descent in Practice I - Feature Scaling

Reading: Gradient Descent in Practice I - Feature Scaling

Video: Gradient Descent in Practice II - Learning Rate

Reading: Gradient Descent in Practice II - Learning Rate

Video: Features and Polynomial Regression

Reading: Features and Polynomial Regression

Video: Normal Equation

Reading: Normal Equation

Video: Normal Equation Noninvertibility

Reading: Normal Equation Noninvertibility

Video: Working on and Submitting Programming Assignments

Reading: Programming tips from Mentors

Reading: Lecture Slides

Graded: Linear Regression with Multiple Variables

Octave/Matlab Tutorial

This course includes programming assignments designed to help you understand how to implement the learning algorithms in practice. To complete the programming assignments, you will need to use Octave or MATLAB. This module introduces Octave/Matlab and shows you how to submit an assignment.

6 videos, 1 reading

expand

Video: Basic Operations

Video: Moving Data Around

Video: Computing on Data

Video: Plotting Data

Video: Control Statements: for, while, if statement

Video: Vectorization

Reading: Lecture Slides

Programming: Linear Regression

Graded: Octave/Matlab Tutorial

WEEK 3

Logistic Regression

Logistic regression is a method for classifying data into discrete outcomes. For example, we might use logistic regression to classify an email as spam or not spam. In this module, we introduce the notion of classification, the cost function for logistic regression, and the application of logistic regression to multi-class classification.

7 videos, 8 readings

expand

Video: Classification

Reading: Classification

Video: Hypothesis Representation

Reading: Hypothesis Representation

Video: Decision Boundary

Reading: Decision Boundary

Video: Cost Function

Reading: Cost Function

Video: Simplified Cost Function and Gradient Descent

Reading: Simplified Cost Function and Gradient Descent

Video: Advanced Optimization

Reading: Advanced Optimization

Video: Multiclass Classification: One-vs-all

Reading: Multiclass Classification: One-vs-all

Reading: Lecture Slides

Graded: Logistic Regression

Regularization

Machine learning models need to generalize well to new examples that the model has not seen in practice. In this module, we introduce regularization, which helps prevent models from overfitting the training data.

4 videos, 5 readings

expand

Video: The Problem of Overfitting

Reading: The Problem of Overfitting

Video: Cost Function

Reading: Cost Function

Video: Regularized Linear Regression

Reading: Regularized Linear Regression

Video: Regularized Logistic Regression

Reading: Regularized Logistic Regression

Reading: Lecture Slides

Programming: Logistic Regression

Graded: Regularization

WEEK 4

Neural Networks: Representation

Neural networks is a model inspired by how the brain works. It is widely used today in many applications: when your phone interprets and understand your voice commands, it is likely that a neural network is helping to understand your speech; when you cash a check, the machines that automatically read the digits also use neural networks.

7 videos, 6 readings

expand

Video: Non-linear Hypotheses

Video: Neurons and the Brain

Video: Model Representation I

Reading: Model Representation I

Video: Model Representation II

Reading: Model Representation II

Video: Examples and Intuitions I

Reading: Examples and Intuitions I

Video: Examples and Intuitions II

Reading: Examples and Intuitions II

Video: Multiclass Classification

Reading: Multiclass Classification

Reading: Lecture Slides

Programming: Multi-class Classification and Neural Networks

Graded: Neural Networks: Representation

WEEK 5

Neural Networks: Learning

In this module, we introduce the backpropagation algorithm that is used to help learn parameters for a neural network. At the end of this module, you will be implementing your own neural network for digit recognition.

8 videos, 8 readings

expand

Video: Cost Function

Reading: Cost Function

Video: Backpropagation Algorithm

Reading: Backpropagation Algorithm

Video: Backpropagation Intuition

Reading: Backpropagation Intuition

Video: Implementation Note: Unrolling Parameters

Reading: Implementation Note: Unrolling Parameters

Video: Gradient Checking

Reading: Gradient Checking

Video: Random Initialization

Reading: Random Initialization

Video: Putting It Together

Reading: Putting It Together

Video: Autonomous Driving

Reading: Lecture Slides

Programming: Neural Network Learning

Graded: Neural Networks: Learning

WEEK 6

Advice for Applying Machine Learning

Applying machine learning in practice is not always straightforward. In this module, we share best practices for applying machine learning in practice, and discuss the best ways to evaluate performance of the learned models.

7 videos, 7 readings

expand

Video: Deciding What to Try Next

Video: Evaluating a Hypothesis

Reading: Evaluating a Hypothesis

Video: Model Selection and Train/Validation/Test Sets

Reading: Model Selection and Train/Validation/Test Sets

Video: Diagnosing Bias vs. Variance

Reading: Diagnosing Bias vs. Variance

Video: Regularization and Bias/Variance

Reading: Regularization and Bias/Variance

Video: Learning Curves

Reading: Learning Curves

Video: Deciding What to Do Next Revisited

Reading: Deciding What to do Next Revisited

Reading: Lecture Slides

Programming: Regularized Linear Regression and Bias/Variance

Graded: Advice for Applying Machine Learning

Machine Learning System Design

To optimize a machine learning algorithm, you’ll need to first understand where the biggest improvements can be made. In this module, we discuss how to understand the performance of a machine learning system with multiple parts, and also how to deal with skewed data.

5 videos, 3 readings

expand

Video: Prioritizing What to Work On

Reading: Prioritizing What to Work On

Video: Error Analysis

Reading: Error Analysis

Video: Error Metrics for Skewed Classes

Video: Trading Off Precision and Recall

Video: Data For Machine Learning

Reading: Lecture Slides

Graded: Machine Learning System Design

WEEK 7

Support Vector Machines

Support vector machines, or SVMs, is a machine learning algorithm for classification. We introduce the idea and intuitions behind SVMs and discuss how to use it in practice.

6 videos, 1 reading

expand

Video: Optimization Objective

Video: Large Margin Intuition

Video: Mathematics Behind Large Margin Classification

Video: Kernels I

Video: Kernels II

Video: Using An SVM

Reading: Lecture Slides

Programming: Support Vector Machines

Graded: Support Vector Machines

WEEK 8

Unsupervised Learning

We use unsupervised learning to build models that help us understand our data better. We discuss the k-Means algorithm for clustering that enable us to learn groupings of unlabeled data points.

5 videos, 1 reading

expand

Video: Unsupervised Learning: Introduction

Video: K-Means Algorithm

Video: Optimization Objective

Video: Random Initialization

Video: Choosing the Number of Clusters

Reading: Lecture Slides

Graded: Unsupervised Learning

Dimensionality Reduction

In this module, we introduce Principal Components Analysis, and show how it can be used for data compression to speed up learning algorithms as well as for visualizations of complex datasets.

7 videos, 1 reading

expand

Video: Motivation I: Data Compression

Video: Motivation II: Visualization

Video: Principal Component Analysis Problem Formulation

Video: Principal Component Analysis Algorithm

Video: Reconstruction from Compressed Representation

Video: Choosing the Number of Principal Components

Video: Advice for Applying PCA

Reading: Lecture Slides

Programming: K-Means Clustering and PCA

Graded: Principal Component Analysis

WEEK 9

Anomaly Detection

Given a large number of data points, we may sometimes want to figure out which ones vary significantly from the average. For example, in manufacturing, we may want to detect defects or anomalies. We show how a dataset can be modeled using a Gaussian distribution, and how the model can be used for anomaly detection.

8 videos, 1 reading

expand

Video: Problem Motivation

Video: Gaussian Distribution

Video: Algorithm

Video: Developing and Evaluating an Anomaly Detection System

Video: Anomaly Detection vs. Supervised Learning

Video: Choosing What Features to Use

Video: Multivariate Gaussian Distribution

Video: Anomaly Detection using the Multivariate Gaussian Distribution

Reading: Lecture Slides

Graded: Anomaly Detection

Recommender Systems

When you buy a product online, most websites automatically recommend other products that you may like. Recommender systems look at patterns of activities between different users and different products to produce these recommendations. In this module, we introduce recommender algorithms such as the collaborative filtering algorithm and low-rank matrix factorization.

6 videos, 1 reading

expand

Video: Problem Formulation

Video: Content Based Recommendations

Video: Collaborative Filtering

Video: Collaborative Filtering Algorithm

Video: Vectorization: Low Rank Matrix Factorization

Video: Implementational Detail: Mean Normalization

Reading: Lecture Slides

Programming: Anomaly Detection and Recommender Systems

Graded: Recommender Systems

WEEK 10

Large Scale Machine Learning

Machine learning works best when there is an abundance of data to leverage for training. In this module, we discuss how to apply the machine learning algorithms with large datasets.

6 videos, 1 reading

expand

Video: Learning With Large Datasets

Video: Stochastic Gradient Descent

Video: Mini-Batch Gradient Descent

Video: Stochastic Gradient Descent Convergence

Video: Online Learning

Video: Map Reduce and Data Parallelism

Reading: Lecture Slides

Graded: Large Scale Machine Learning

WEEK 11

Application Example: Photo OCR

Identifying and recognizing objects, words, and digits in an image is a challenging task. We discuss how a pipeline can be built to tackle this problem and how to analyze and improve the performance of such a system.

5 videos, 1 reading

expand

Video: Problem Description and Pipeline

Video: Sliding Windows

Video: Getting Lots of Data and Artificial Data

Video: Ceiling Analysis: What Part of the Pipeline to Work on Next

Reading: Lecture Slides

Video: Summary and Thank You

Graded: Application: Photo OCR