Chapter – 1

**CS229:**

**Machine Learning**

**Introduction to**

**Machine Learning**

**1.1 About This course**

<https://cs229.stanford.edu/>

<https://docs.google.com/spreadsheets/d/1sEu4ygD5HWxaqjvbR2nsjvG6NBoW5tRW/edit#gid=348511794>

* Course Description: This course provides a broad introduction to ***Machine Learning*** and ***Statistical Pattern Recognition***. Topics include:
* Supervised Learning:
* generative/discriminative learning,
* parametric/non-parametric learning,
* neural networks,
* support vector machines)
* Unsupervised Learning:
* clustering,
* dimensionality reduction,
* kernel methods
* Learning Theory :
* bias/variance tradeoffs,
* practical advice
* Reinforcement Learning and Adaptive Control.

The course will also discuss recent applications of machine learning, such as to *robotic control*, *data mining*, *autonomous navigation*, *bioinformatics*, *speech recognition*, and *text* and *web data processing*.

* Prerequisites: Students are expected to have the following background:

1. Knowledge of basic computer science principles and skills, at a level sufficient to write a reasonably non-trivial computer program in Python/NumPy. Following course covers the Prerequisites:

* CS106A: Programming Methodology
* CS106B: Programming Abstractions
* CS106X: Programming Abstractions in C++

1. Familiarity with probability theory.

* CS 109: Probability for Computer Scientists
* MATH151: Introduction to Probability Theory
* STATS 116: Introduction to Probability

1. Familiarity with multivariable calculus and linear algebra. Relevant classes include, but not limited to:

* MATH 51: Linear Algebra and Differential Calculus of Several Variables
* MATH 104: Applied Matrix Theory
* MATH 113: Linear Algebra and Matrix Theory
* CS 205: Continuous Mathematical Methods with an Emphasis on Machine Learning
* CME 100: Vector Calculus for Engineers

1. Collaboration with generative AI tools such as Co-Pilot and ChatGPT is allowed, treating them as collaborators in the problem-solving process. However, the direct solicitation of answers or copying solutions, whether from peers or external sources, is strictly prohibited.

**STATS 229: Machine Learning (CS 229)**

* Topics:

1. Statistical pattern recognition,
2. Linear and non-linear regression,
3. Non-parametric methods,
4. Exponential family,
5. GLMs,
6. Support vector machines,
7. Kernel methods,
8. Deep learning,
9. Model/feature selection,
10. Learning theory,
11. ML advice,
12. Clustering,
13. Density estimation,
14. EM,
15. Dimensionality reduction,
16. ICA, PCA,
17. Reinforcement learning and adaptive control,
18. Markov decision processes,
19. Approximate dynamic programming, and policy search.

* Prerequisites:

1. Knowledge of basic ***Computer Science Principles*** and skills at a level sufficient to write a reasonably non-trivial computer program in Python/NumPy to the equivalency of **CS106A**, **CS106B**, or **CS106X**.
2. Familiarity with ***Probability Theory*** to the equivalency of **CS109**, **MATH151**, or **STATS116**.
3. Familiarity with ***Multivariable Calculus*** and ***Linear Algebra*** to the equivalency of **MATH51** or **CS205**.

**CS229: Machine Learning**

**Week 1**

**Lecture 1:**

* Topics:
* Introduction and Basic Concepts (lecture1\_slide.pptx, lecture1.pdf)

**Lecture 2:**

* Topics:
* Supervised Learning Setup. LMS
* Linear Regression
* Class Notes:
* Supervised Learning, Discriminative Algorithms (cs229-notes1.pdf: sections 1-3)
* Live lecture notes (cs229-livenotes-lecture2.pdf)

**Friday Lecture:**

* Linear Algebra
* Multivariable Calculus, and Modern Applications
* (Stanford Math 51)

**Week 2**

**Lecture 3:**

* Topics:
* Weighted Least Squares.
* Logistic Regression.
* Netwon's Method Perceptron.
* Class Notes:
* Supervised Learning (cs229-notes1.pdf: sections 4,5 & 7)
* Live Lecture Notes (lecture3\_draft.pdf)

**Lecture 4:**

* Topics:
* Dataset split,
* Exponential Family.
* Generalized Linear Models.
* Class Notes:
* Supervised Learning (cs229-notes1.pdf: Sections 6, 8, and 9)
* Live Lecture Notes (lecture4\_draft.pdf)

**Friday Lecture:**

* Topics:
* Probability Theory Review
* Class Notes:
* Probability Theory Review [cs229-prob.pdf]
* The Multivariate Gaussian Distribution [gaussians.pdf]
* More on Gaussian Distributions [more\_on\_gaussians.pdf]
* Friday Section Slides [friday\_lecture2.pdf]

**Week 3**

**Lecture 5:**

* Topics:
* Gaussian Discriminant Analysis.
* Naive Bayes.
* Laplace Smoothing.
* Class Notes:
* Generative Algorithms [cs229-notes2.pdf (section 1)]
* Live lecture notes [cs229-livenotes-lecture5.pdf; 2021spring/lecture5\_live.pdf]

**Lecture 6:**

* Topics:
* Laplace Smoothing.
* Support Vector Machines.
* Class Notes:
* Naive Bayes and Laplace Smoothing [cs229-notes2.pdf(Section 2);]
* Support Vector Machines [cs229-notes3.pdf]
* Live Lecture Notes [cs229-livenotes-lecture6.pdf; 2021spring/lecture6\_live.pdf]

**Friday Lecture:**

* Topics:
* Python and Numpy Tutorial
* Class Notes:
* Section slides [cs229\_python\_friday.pdf; friday\_lecture3.pdf(2021)]
* Jupyter notebook [html][source]
* Python Review Code[2021: python-review-code.pdf, cs229-python-review-code.ipynb]

**Week 4**

**Lecture 7:**

* Topics:
* Support Vector Machines.
* Kernel methods
* Class Notes:
* Live lecture notes [cs229-livenotes-lecture7.pdf; 2021:lecture7\_live.pdf]
* Kernel Methods [2021: cs229-notes3.pdf]

**Lecture 8:**

* Topics:
* Neural Networks - 1
* Class Notes:
* Deep Learning [cs229-notes-deep\_learning.pdf; 2021: deep\_learning\_notes.pdf]
* Backpropagation
* Live lecture notes [2021: lecture8\_live.pdf]

**Friday Lecture:**

* Topics:
* Evaluation Metrics
* Class Notes:
* Evaluation Metrics [pdf: evaluation\_metrics\_spring2020.pdf (2021: friday\_lecture4.pdf)]

**Week 5**

**Lecture 9:**

* Topics:
* Neural Networks - 2
* Backpropagation
* Class Notes:
* See Neural Networks - 1 Notes
* Deep Learning [2021: deep\_learning\_notes.pdf]
* Live Lecture Notes [2021: lecture9\_live.pdf]

**Lecture 10:**

* Topics:
* Bias - Variance. Regularization. Feature / Model selection.
* Class Notes:
* Regularization and Model Selection

[pdf: bias-variance-error-analysis.pdf, addendum:bias-variance-error-analysis-addendum.pdf; 2021: cs229-notes5.pdf]

* Live lecture notes [cs229-livenotes-lecture10-draft.pdf]
* Double Descent [link: https://arxiv.org/abs/1903.07571, optional reading]
* Some Calculations from Bias Variance (Addendum) [2021: addendum\_bias\_variance.pdf]
* Bias-Variance and Error Analysis (Addendum) [2021: error-analysis.pdf]
* Live Lecture Notes [2021: lecture10\_live.pdf]

**Friday Lecture:**

* Topics:
* Deep Learning (ConvNets)
* Class Notes:
* Deep Learning [cs229\_deep\_learning\_friday.pptx]
* Friday Section Slides [2021: friday\_lecture5.pdf; cs229\_deep\_learning\_friday.pptx]

**Week 6**

**Lecture 11:**

* Topics:
* K-Means. GMM (non EM).
* Expectation Maximization.
* Class Notes:
* Unsupervised Learning, k-means clustering. [cs229-notes7a.pdf]
* Mixture of Gaussians [cs229-notes7b.pdf]
* The EM Algorithm [cs229-notes8.pdf]
* Live lecture notes [cs229-livenotes-lecture11-draft.pdf, 2021: lecture11\_draft.pdf]

**Lecture 12:**

* Topics:
* Expectation Maximization (continued)
* GMM (EM). Factor Analysis.
* Class Notes:
* Lagrange Multipliers Review [lagrange\_multiplier.pdf]
* Live lecture notes [cs229-livenotes-lecture12-draft.pdf, cs229-livenotes-lecture12.pdf, 2021: lecture12\_draft.pdf]
* Factor Analysis [cs229-notes9.pdf]
* Addendum Notes[5\_5\_addendum.pdf]

**Week 7**

**Lecture 13:**

* Topics:
* Factor Analysis.
* Class Notes:
* Factor Analysis [cs229-notes9.pdf]
* Live lecture notes [cs229-livenotes-lecture13-draft.pdf, in lecture: cs229-livenotes-lecture13.pdf]

**Lecture 14:**

* Topics:
* Principal and Independent Component Analysis.
* Class Notes:
* Principal Components Analysis [cs229-notes10.pdf]
* Independent Component Analysis [cs229-notes11.pdf]
* Live lecture notes [2021: lecture13\_draft.pdf; cs229-livenotes-lecture14-draft.pdf, in lecture(2021): cs229\_livenotes\_lecture14-pca.pdf]

**Friday TA Lecture:**

* Topics:
* Decision Trees
* Boosting.
* Notes:
* Decision trees [2021:Decision\_Trees\_CS229.pdf]
* Boosting [2021: boosting.pdf]

**Week 8**

**Lecture 15:**

* Topics (2021)
* Self-supervised learning (Language Models & Image Models).
* Class Notes (2021)
* Self-Supervised Learning [2021: cs229\_lecture\_selfsupervision\_final.pdf]

**Lecture 16 (2021:14):**

* Topics:
* Weak Supervision, Weak supervised/unsupervised learning
* ML Advice.
* Class Notes:
* Weak Supervision [weak\_supervision\_slides.pdf]
* Weak Supervision [weak\_supervision\_notes.pdf, weak\_supervision\_notes\_live.pdf]
* Additional Material
* ML Advice [ML\_Advice\_lecture.pdf]
* Relevant video from Fall 2018 [ML-advice.pdf]
* <https://www.youtube.com/watch?v=ORrStCArmP4&list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU&index=13>
* Introduction to weak supervision [2021:WeakSupervise\_229.pdf]
* ICA and weak supervision [2021:ica\_and\_weak\_supervision.pdf]

**Friday TA Lecture**

* Topics:
* On Critiques of ML.
* Class Notes:
* Technical and Societal Critiques of ML [2021: friday\_lecture\_ml\_critique.pdf]

**Week 9**

**Lecture 17:**

* Topics:
* Basic concepts in RL
* Markov Decision Process.
* Value Iteration and Policy Iteration.
* Q-Learning.
* Value function approximation.

* Class Notes:
* Reinforcement Learning and Control [cs229-notes12.pdf]
* Basic RL concepts, value iterations, policy iteration [2021:cs229-notes12.pdf] (Sections 1 and 2)
* Live Lecture Notes [pdf]

**Lecture 18:**

* Topics:
* Reinforcement Learning continued
* Model-based RL,
* value function approximator.
* Class Notes
* Model-based RL and value function approximation [2021:cs229-notes12.pdf] (Sections 3 and 4)
* Live Lecture Notes [pdf]

**Friday TA Lecture:**

* Topics:
* Learning Theory.
* Class Notes
* Learning theory [2021:cs229-notes4.pdf]

**Week 10**

**Lecture 19:**

* Topics:
* Policy search. Reinforce. POMDPs.
* Class Notes
* Policy Gradient (REINFORCE) [cs229-notes14.pdf]

**Lecture 20:**

* Topics:
* Recap, Fairness, Adversarial Class Notes
* Social impact
* Supplementary Notes
* Online Learning and the Perceptron Algorithm [http://cs229.stanford.edu/notes2020spring/cs229-notes6.pdf]
* Binary classification with +/-1 labels [http://cs229.stanford.edu/extra-notes/loss-functions.pdf]
* The representer theorem [http://cs229.stanford.edu/extra-notes/representer-function.pdf]
* Hoeffding's inequality [http://cs229.stanford.edu/extra-notes/hoeffding.pdf]
* Optional Topics
* Decision trees [http://cs229.stanford.edu/notes/cs229-notes-dt.pdf]
* Decision tree ipython demo [http://cs229.stanford.edu/notes2020spring/ta\_lecture/decision\_tree\_demo.ipynb]
* Boosting algorithms and weak learning [http://cs229.stanford.edu/extra-notes/boosting.pdf]
* On critiques of ML [https://cs229.stanford.edu/materials/critiques-ml-aut19.pdf]
* Other Resources:
* Advice on applying machine learning: Slides from Andrew's lecture on getting machine learning algorithms to work in practice can be found here: <http://cs229.stanford.edu/materials/ML-advice.pdf> .
* Previous projects: A list of last quarter's final projects can be found here: <http://cs229.stanford.edu/proj2020spr/> .
* Data: Here is the UCI Machine learning repository (<http://www.ics.uci.edu/~mlearn/MLRepository.html>), which contains a large collection of standard datasets for testing learning algorithms. If you want to see examples of recent work in machine learning, start by taking a look at the conferences NIPS: <https://www.neurips.cc/> (all old NIPS papers are online) and ICML. Some other related conferences include UAI, AAAI, IJCAI.
* Viewing PostScript and PDF files: Depending on the computer you are using, you may be able to download a PostScript viewer or PDF viewer for it if you don't already have one.

Machine learning study guides tailored to CS 229 by Afshine Amidi and Shervine Amidi.

<https://github.com/PKUFlyingPig/CS229/tree/master>

**STATS 229: Machine Learning (CS 229)**

* ***Topics:***

|  |  |
| --- | --- |
| * Statistical pattern recognition, * linear and non-linear regression, * non-parametric methods, * exponential family, * GLMs, * support vector machines, kernel methods, * deep learning, * model/feature selection, | * learning theory, * ML advice, * clustering, density estimation, * EM, * dimensionality reduction, ICA, PCA, * reinforcement learning and adaptive control, * Markov decision processes, * approximate dynamic programming, and policy search. |

* ***Prerequisites:***
* Knowledge of basic computer science principles and skills at a level sufficient to write a reasonably non-trivial computer program in Python/NumPy to the equivalency of CS106A, CS106B, or CS106X,
* Familiarity with probability theory to the equivalency of CS 109, MATH151, or STATS 116, and
* Familiarity with multivariable calculus and linear algebra to the equivalency of MATH51 or CS205.

**CS129: Applied Machine Learning**

CS 129: <https://web.stanford.edu/class/cs129/syllabus>

**Course Description**

You will learn how to *implement* and *apply* machine learning algorithms. This course emphasizes *practical skills*, and focuses on teaching you a wide range of *algorithms* and giving you the skills to make these *algorithms* work best.

You will learn about commonly used learning techniques including

* supervised learning algorithms:
* logistic regression,
* linear regression,
* neural networks/deep learning,
* decision trees
* unsupervised learning algorithms
* k-means,
* reinforcement learning algorithms like Q-learning,
* As well as specific applications such as
* anomaly detection and
* building recommender systems.

**Prerequisites**

Students are expected to have the following background:

* Programming at the level of CS106B or 106X, Knowledge of basic computer science principles and skills, at a level sufficient to write a reasonably non-trivial computer program
* Probability theory at the level CS109 or STATS116 and
* Basic linear algebra at the level of MATH51.

SCHEDULE

***Lecture 1***

* The AI world
* Logistics of the course
* Presentation of the Syllabus

***On Coursera***

* Week 1 and Week 2 of Supervised Machine Learning: Regression and Classification (including optional labs and quizzes)

***Lecture 2***

* Linear Regression
* Derivations
* Practice problems

***On Coursera***

* Week 3 of Supervised Machine Learning: Regression and Classification (including optional labs and quizzes)

***Lecture 3***

* Logistic Regression
* Text processing
* Derivations
* Practice problems

***On Coursera***

* Week 1 of Advanced Learning Algorithms: Neural Networks (including optional labs and quizzes)

***Lecture 4***

* Neural Networks
* Vectorized Gradients
* Softmax
* Practice problems

***On Coursera***

* Week 2 of Advanced Learning Algorithms: Neural network training (including optional labs and quizzes)

***Lecture 5***

* Multi-class classification
* Vectorized Back-propagation
* Practice problems

***On Coursera***

* Week 3 of Advanced Learning Algorithms: Advice for applying machine learning (including optional labs and quizzes)

***Lecture 6***

* Bias & Variance Trade-off in Practice
* Practice problems
* Debugging Strategies for Final Project
* Advice on ML Systems
* Hogwarts Case study.

***On Coursera***

* Week 4 of Advanced Learning Algorithms: Decision trees (including optional labs and quizzes)

***Lecture 7***

* Measuring Purity
* Random Forest
* XG Boost
* Practice problems

***On Coursera***

* Week 1 of Unsupervised Learning, Recommenders, Reinforcement Learning: Unsupervised Learning (including optional labs and quizzes)

***Lecture 8***

* Logistics
* Midterm will be held during class time.

***On Coursera***

* Week 2 of Unsupervised Learning, Recommenders, RL: Recommender Systems (including optional labs and quizzes)

***Lecture 9***

* K-Means Clustering
* Principal Component Analysis

***On Coursera***

* Week 3 of Unsupervised Learning, Recommenders, Reinforcement Learning: Reinforcement Learning (including optional labs and quizzes)

***Lecture 10:*** AI future directions and Career Advice with Andrew