

معرفی

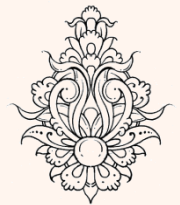
مقدمه‌ای بر
یادگیری ماشین
بخش نخست



دانشگاه شهید بهشتی
پژوهشکده فضای مجازی
پاییز ۱۴۰۱
احمد محمودی ازناوه

فهرست مطالب

- چند تذکر
- منابع
- باره بندی
- سرفصل
- هوش مصنوعی
- یادگیری چیست؟
- - یادگیری ماشین
- چرا یادگیری؟
- کاربردهای یادگیری ماشین
- انواع شیوه های یادگیری



چند تذکر

- در صورتی که در رابطه با مطلبی، ایمیل می‌زنید، لطفا در پایان ایمیل نام خود را هم بنویسید، به ویژه اگر از نام مستعار برای شناسایی ایمیل خود استفاده می‌کنید.
- ابتدای موضوع (subject) ایمیل با **ML01-2** شروع شود.
- یکی از مهمترین مواردی که رعایت آن بر عهده‌ی ماست، رعایت «**اخلاق آکادمیک**» است. کپی کردن تکالیف، استفاده از مطلبی بدون ذکر منبع و هم‌فکری در امتحان از موارد بارز تخلف محسوب می‌شود.
- در صورت تقلب فرد متخلف نمره‌ی منفی خواهد گرفت.
- **از نوشتن به صورت فینگلیش پرهیزید.**

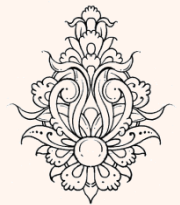
از همکاری شما پیشاپیش سپاسگزاره!

گروه درس

http://faculties.sbu.ac.ir/~a_mahmoudi/ML_01_1.htm

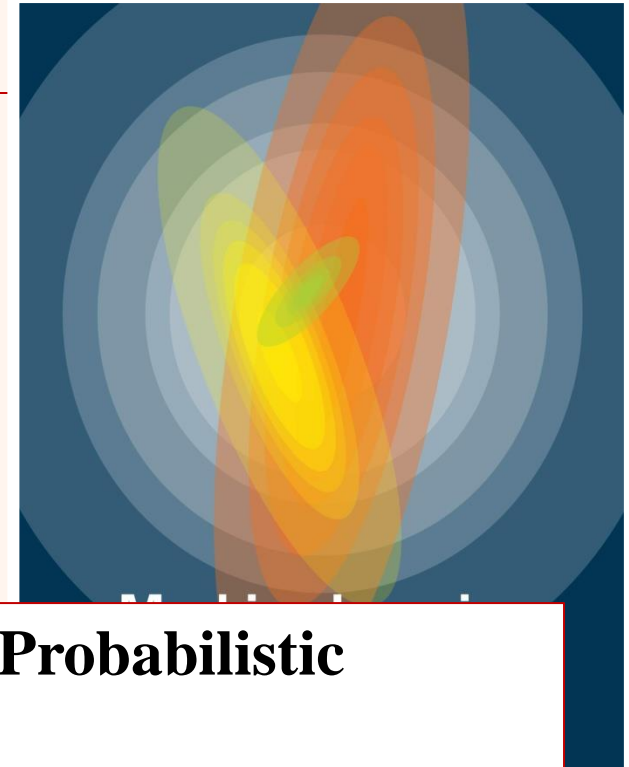
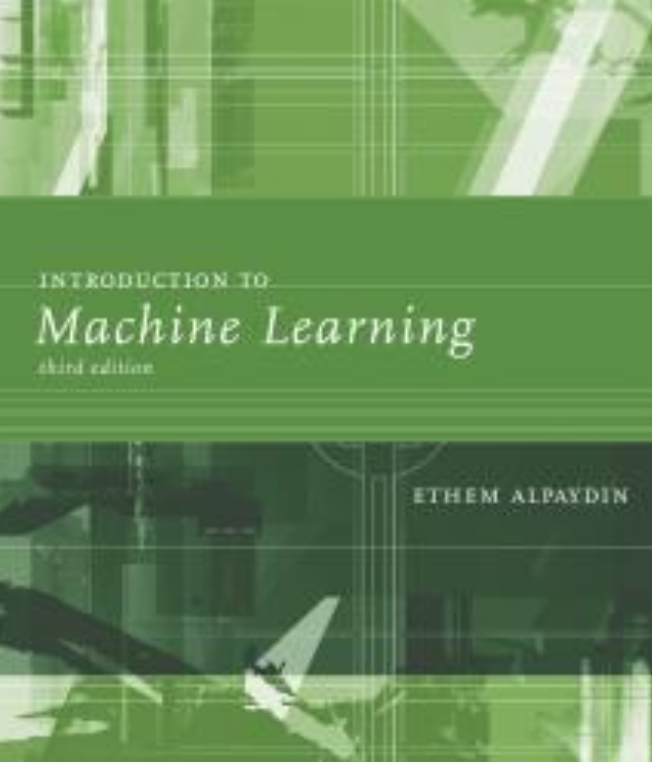
صفحه درس:

دستیار (TAs): آقای جواهری

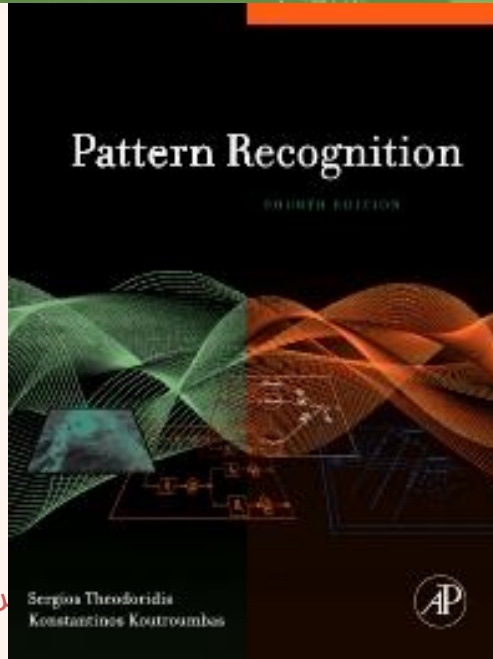


منابع

**Introduction to Machine Learning,
Third Edition**
Ethem Alpaydin



**Machine Learning: A Probabilistic
Perspective**
Kevin Murphy

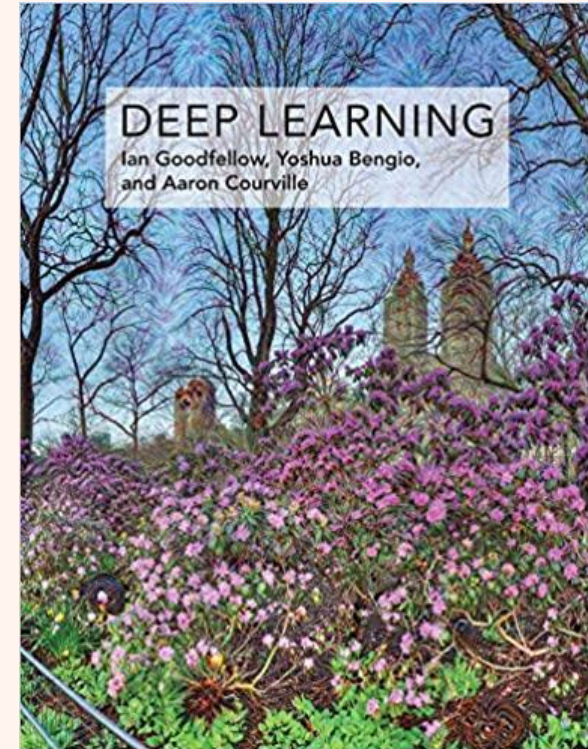
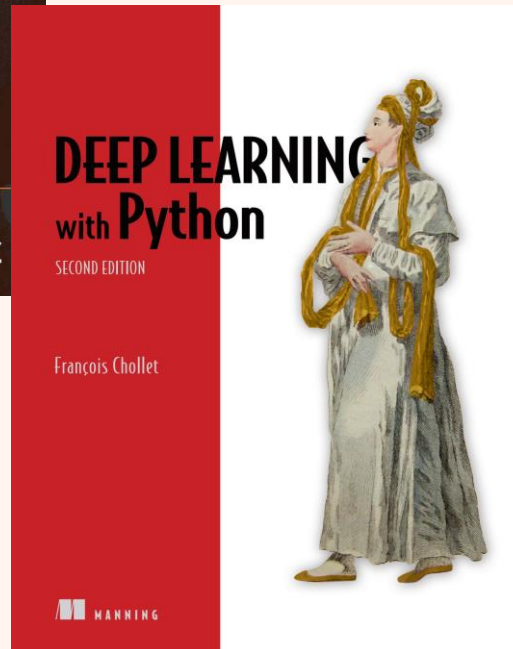
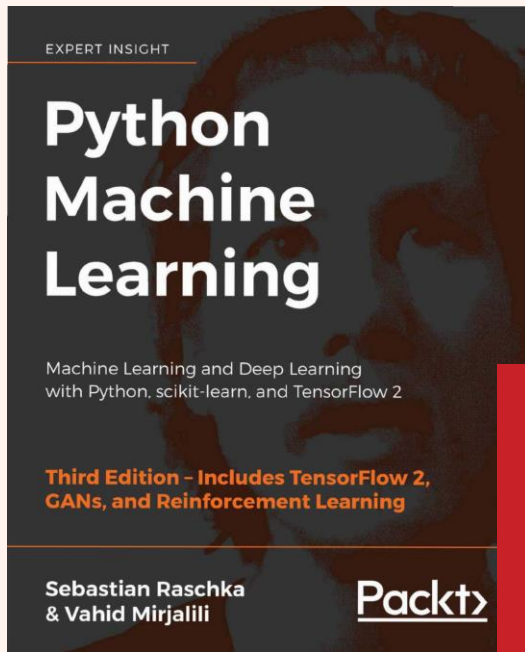


Pattern Recognition
Theodoridis & Koutroumbas



سایر منابع

Raschka, S. and V. Mirjalili (2019). Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2, 3rd Edition, Packt Publishing.

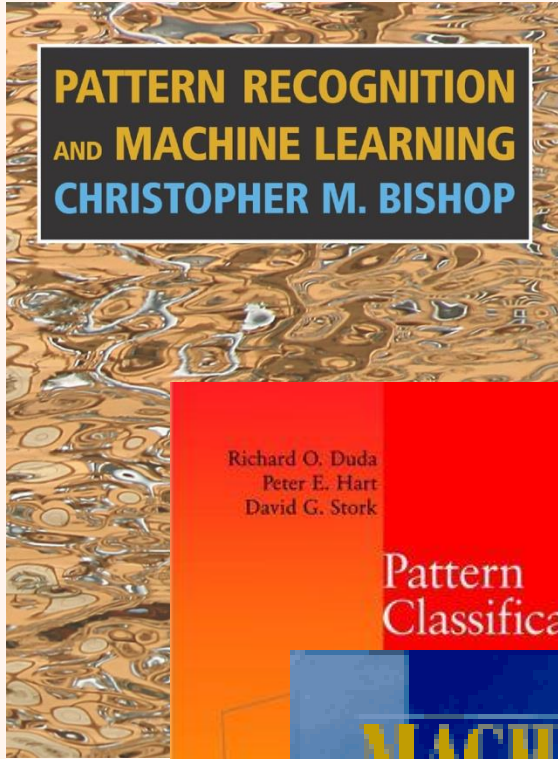


Chollet, F. (2021). Deep Learning with Python, Second Edition, Manning.

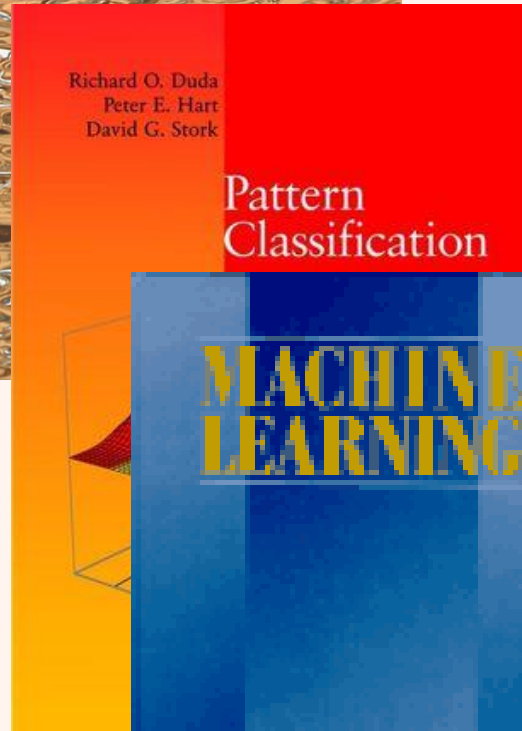
Deep Learning by Aaron Courville, Ian Goodfellow, and Yoshua Bengio

تراشگاه
سپهر
بهشتی

سایر منابع

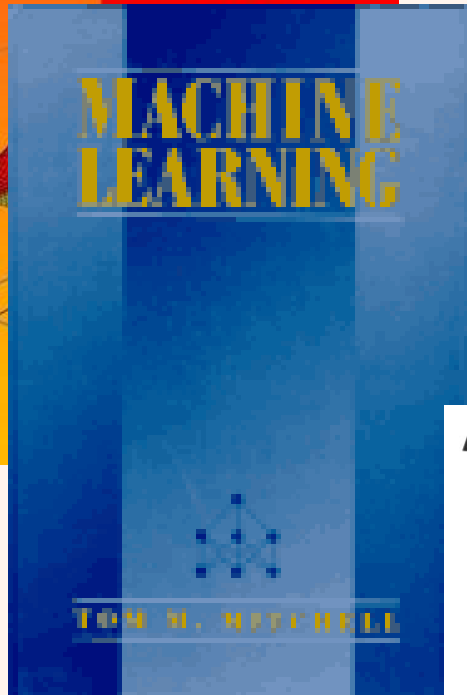


Pattern Recognition and Machine Learning
Christopher Bishop



Pattern classification

Richard O. Duda, Peter E. Hart and David G. Stork



Machine Learning

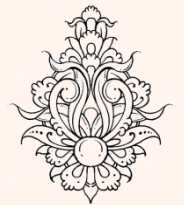
Tom Mitchell

در تهیهی اسلایدهای این درس از منابع online نظیر اسلایدهای سایر دانشگاه استفاده شده است که در صورت استفاده لینک آن ذکر خواهد شد.



منابع آنلاین

- <https://machinelearningmastery.com/>
- <https://towardsdatascience.com/>
- <https://medium.com/>
- <https://www.coursera.org/>



بارمبندی

- پروژه و فعالیت‌های پژوهشی ۲۵-۱۵٪
- تکالیف ۲۰-۱۰٪
- نمره امتحان ۶۰-۵۰٪
- فعالیت کلاسی ۱۰٪

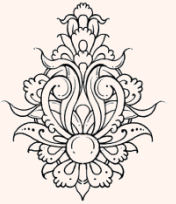
• توجه داشته باشید که بارجمندی فوق تقریبی است و بسته به شرایط ممکن است تغییر کند.

• در طول ترم، در مورد مباحث مطرح شده، ارائی کاربرد یا یک نمونهی بهبود یافته به صورت مختصر به عنوان فعالیت پژوهشی مورد نظر است.

• مهلت تکمیل فرم پیشنهاد موضوع ارائی شفاهی **تا پایان آبان‌ماه** است.

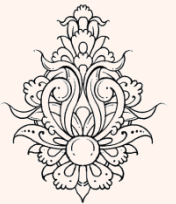
• پروژه‌ی نهایی به صورت کتبی تمویل داده و پس از آن در صورت لزوم به صورت شفاهی هم ارائه خواهد شد.

• توجه داشته باشید که ارائه ترجمه به هیچ وجه قابل قبول نیست.



پیش‌نیازها

- آشنایی با جبرخطی و آمار و احتمال
- آشنایی با برنامه‌نویسی (Matlab, Python)



سرفصل کلی

یادگیری بانظارت: دسته‌بندی و رگرسیون

یادگیری بی‌نظارت و روش‌های پارامتری

روش‌های چند متغیره

کاهش ابعاد

روش‌های ناپارامتری

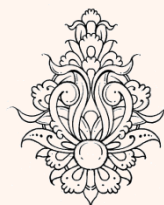
درخت تصمیم

ماشین بردار پشتیبان

شبکه‌های عصبی عمیق

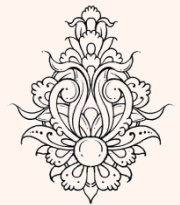
یادگیری گروهی

یادگیری مشارکتی



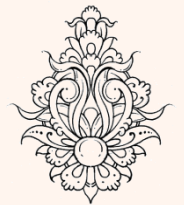
پایگاه داده‌های در دسترس

- UCI Repository:
<http://www.ics.uci.edu/~mlearn/MLRepository.html>
- UCI KDD Archive:
<http://kdd.ics.uci.edu/summary.data.application.html>
- Statlib: <http://lib.stat.cmu.edu/>
- Delve: <http://www.cs.utoronto.ca/~delve/>



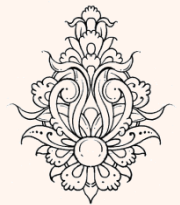
مجلات مرتبہ

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- Pattern Recognition
- Nature
- ...



کنفرانس‌های مرتب

- **NeurIPS**: Neural Information Processing Systems (formerly NIPS)
- **ICML**: International Conference on Machine Learning
- **ICLR**: International Conference on Learning Representations
- **CVPR**: IEEE/CVF Conference on Computer Vision and Pattern Recognition
- **ECCV**: European Conference on Computer Vision
- **AAAI**: AAAI Conference on Artificial Intelligence



هوش مصنوعی

- آیا امکان تحقق کامل اهداف هوش مصنوعی وجود دارد؟

It has long been believed, especially by older members of the scientific community, that for machines to be as intelligent as us, that is, for artificial intelligence to be a reality, our current knowledge in general, or computer science in particular, is not sufficient. People largely are of the opinion that we need a new technology, a new type of material, a new type of computational mechanism or a new programming methodology, and that, until then, we can only “simulate” some aspects of human intelligence and only in a limited way but can never fully attain it.

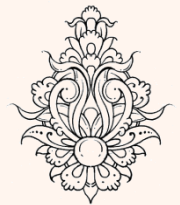
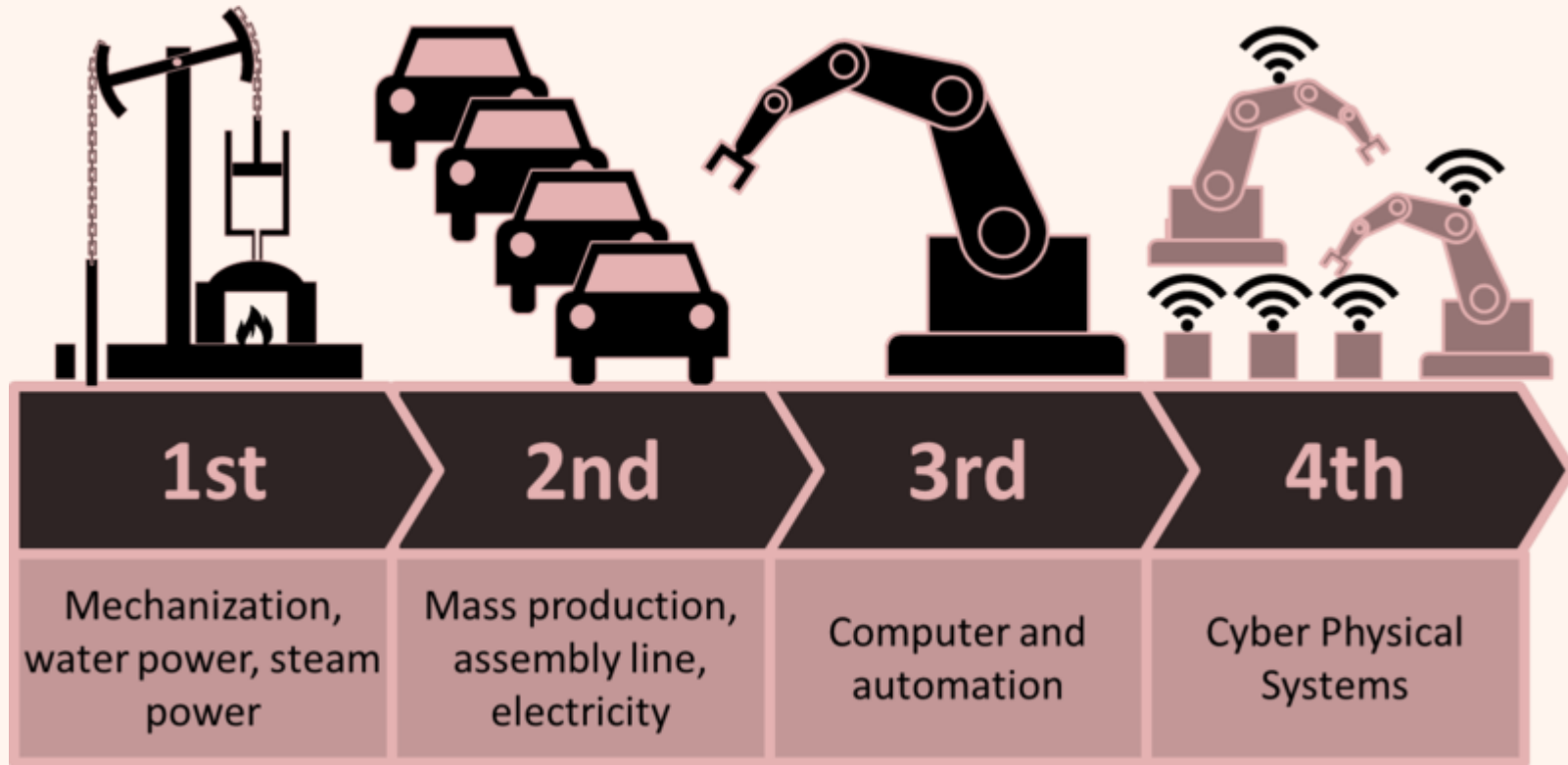
I believe that we will soon prove them wrong.

Ethem Alpaydin

مقدمه‌ی چاپ دوم کتاب

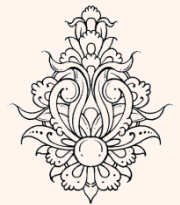
ناشر
بهبودی

انقلاب صنعتی چہارم



دیباچه

- علوه مختلف از روش‌های مطرح شده در «یادگیری ماشین» استفاده می‌کنند.
- نقش «یادگیری ماشین» در زندگی تا چه حد است؟
- کاربردهای یادگیری ماشین در زندگی روزمره راه یافته است:
 - تشخیص دست‌نوشته
 - تبدیل گفتار به متن
 - سامانه‌های توصیه‌گر
 - پیشنهاد دوست در شبکه‌های اجتماعی
 - شناسایی خصیصه‌های زیستی (چهره، اثر انگشت)
 - خودرو بدون راننده
 - جداسازی سیگنال‌ها مخلوط شده (BSS)



یادگیری چیست؟

- «یادگیری» عبارتست از تخییر نسبتاً پایدار در احساس، تفکر و رفتار فرد که بر اساس تجربه ایجاد شده باشد.

به نقل از ویکی‌پدیا

Learning is the act of acquiring new, or modifying and reinforcing existing knowledge, behaviors, skills, values, or preferences.

The ability to learn is possessed by humans, animals and some machines.

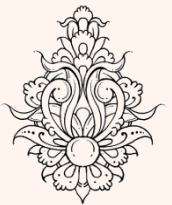


یادگیری ماشین چیست؟ Machine Learning

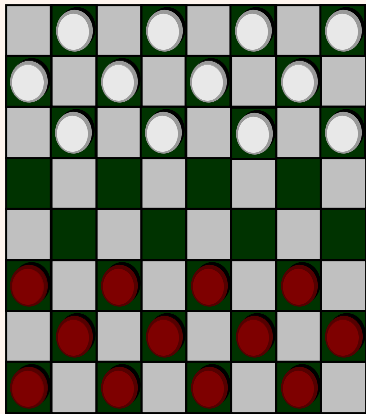
- «یادگیری ماشین» برنامه‌نویسی برای بهینه‌سازی یک عملکرد با استفاده از داده‌ها و تجربیات گذشته است.

Machine learning is programming computers to optimize a performance criterion using example data or past experience.

- «یادگیری ماشین» در پی راهی برای ایجاد برنامه‌ای است که عملکرد را به صورت خودکار و با توجه به تجربیات ارتقا دهد. (Tom. M. Mitchell).



Field of study that gives computers the ability to learn without being explicitly programmed.



Arthur Samuel (1959)

Well-posed Learning Problem: A computer program is said to learn from experience **E** with respect to some task **T** and some performance measure **P**, if its performance on **T**, as measured by **P**, improves with experience **E**.

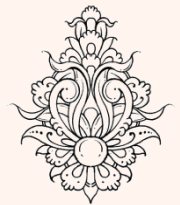
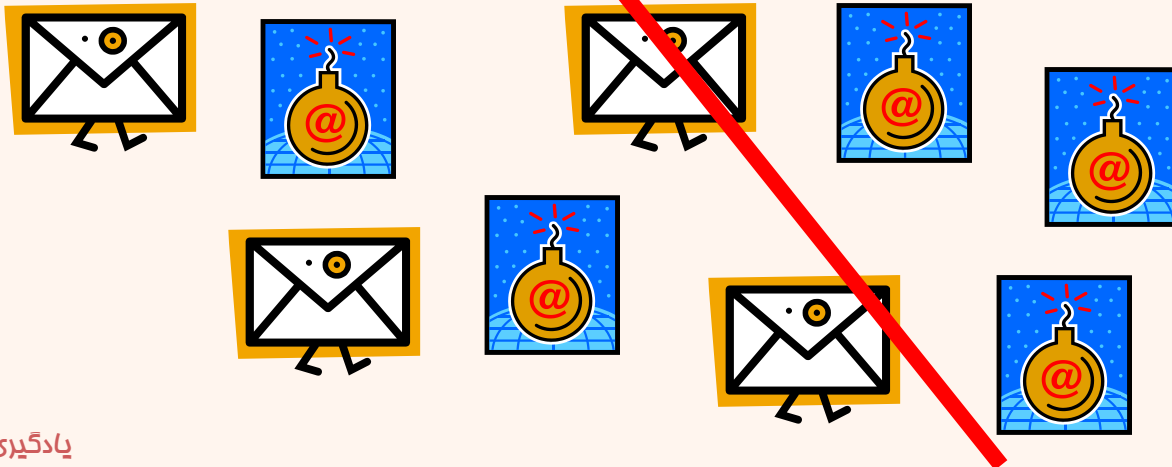
Tom Mitchell (1998)



سایر تعاریف - مثال

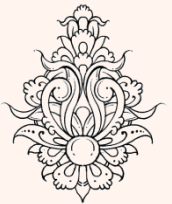
- به عنوان مثال یک برنامه‌ی **تشخیص هرزنامه** را در نظر بگیرید که با توجه به ایمیل‌هایی که کاربر به عنوان spam اعلام می‌کند، سعی در بهبود کارایی خود دارد.
- در این صورت

- **T** عمل دسته‌بندی ایمیل‌ها به دو گروه spam/not spam
- **E** پی‌گیری ایمیل‌هایی که کاربر به عنوان spam اعلام می‌کند.
- **P** تعداد ایمیل‌هایی که به درستی به عنوان spam دسته‌بندی شده‌اند.



چرا یادگیری؟

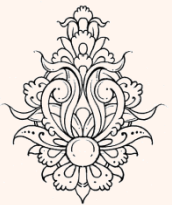
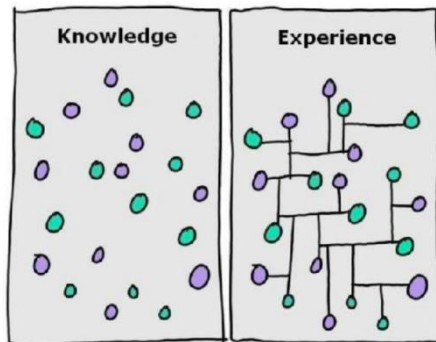
- برای حل یک مسأله بر روی کامپیوتر به یک «الگوریتم» احتیاج داریم.
- برای برخی مسائل نمی‌توان یک الگوریتم نوشت، مانند تشخیص هرزنامه‌ها. حتی ممکن است برخی از ایمیل‌ها بسته به کاربر هرزنامه تلقی شوند یا نه، اما نمونه‌های زیادی از داده در اختیار داریم.
- در جاهایی که نمی‌توانیم مستقیماً برنامه‌ی مورد نظر را بنویسیم، به یادگیری احتیاج داریم، که با کمک یک سری **داده‌ی آموزشی یا تجربیات** صورت می‌پذیرد.
- به عنوان مثال برای محاسبه‌ی حقوق پرسنل نیازی به یادگیری وجود ندارد.



منظور از یادگیری

- یادگیری ← استخراج یک **مدل** کلی از روی داده‌ها
- بیشتر داده‌های اخذ شده توسط وسایل مختلف به صورت دیجیتال هستند.
- در واقع با حجم انبوهی از داده‌ها مواجه هستیم که ارزان به دست می‌آیند، با این حال دانش در خصوص آن‌ها به سادگی و با هزینه‌ی پایین حاصل نمی‌شود.

We are drowning in information and starving for knowledge. John Naisbitt.



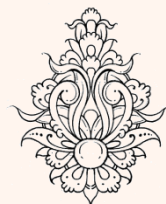
منظور از یادگیری (ادامه...)

- در موارد زیر به یادگیری امتیاج داریم:
 - در مواردی که انسان فبره در دسترس نیست،
 - زمانی که انسان قادر به توضیح مهارت خود نیست، مانند تشخیص صوت
 - زمانی که مساله‌ی مورد نظر در طول زمان تغییر می‌کند؛ به شرایط محیط وابسته است، مانند مسیریابی در شبکه‌های کامپیوتری
 - حالاتی که به حل مساله به تطبیق با شرایط خاصی وابسته است، تشخیص هویت با استفاده از خصیصه‌های زیستی

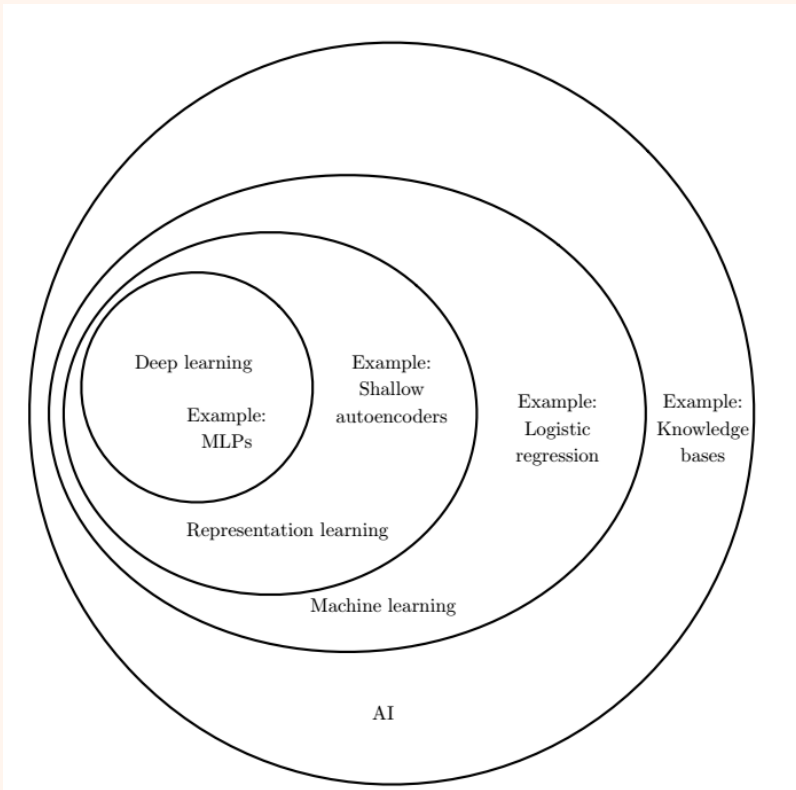


منظور از یادگیری (ادامه...)

- خرده‌فروشی: به عنوان مثال استخراج مدل کلی رفتار مشتری‌ها از روی تراکنش‌های مربوط به آن‌ها مثلاً کسی که از یک فروشگاه online یک محصول را خریداری می‌کند، با توجه به سابقه‌ی خریدهای مشابه می‌توان خریدهای بعدی او را حدس زد.
- هرچند ممکن است قادر به ارائه‌ی یک مدل دقیق نباشیم، اما می‌توان یک **تقریب خوب و مفید** به دست آورد.
- مدل به دست آمده می‌تواند برای پیش‌بینی مورد استفاده قرار گیرد (**predictive**) و یا به منظور استخراج دانش از داده‌ها به کار آید (**descriptive**).

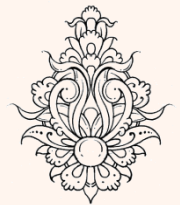


یادگیری عمیق



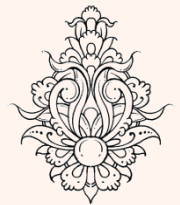
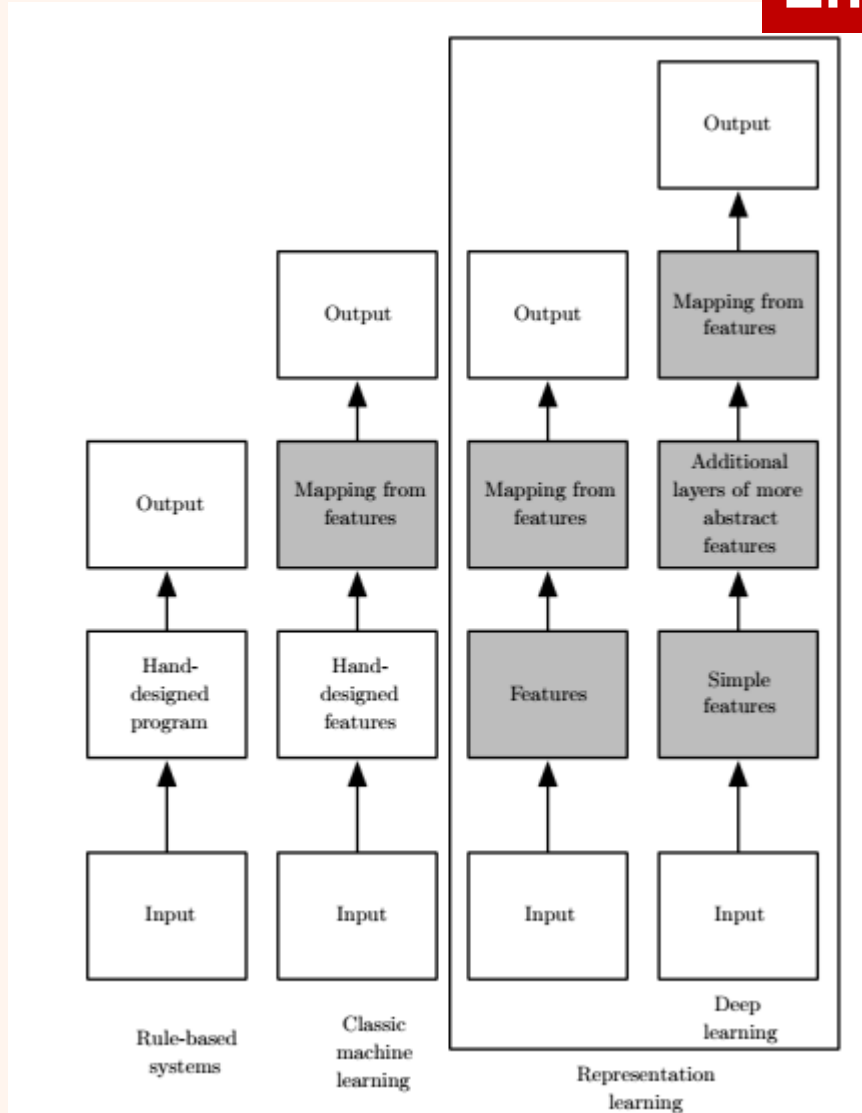
Goodfellow, I., Y. Bengio, and A. Courville. Deep Learning. MIT Press, 2016.

یادگیری عمیق (ژرف): دسته‌ای از الگوریتم‌های یادگیری ماشین است که دارای لایه‌های متعددی (عمیق) که از طریق این لایه‌ها و به صورت سلسله‌مراتبی از روی داده‌های خام ویژگی‌های سطح بالا استخراج می‌کند.

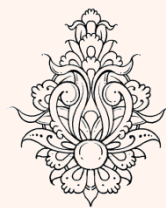


یادگیری عمیق (ادامه...)

End to end learning



- **یادگیری ماشین**، یکی از مهم‌ترین بخش‌های **هوش مصنوعی** است. برای هوشمند بودن، یک سیستم که در محیطی با شرایط متغیر قرار دارد، باید توانایی آموختن داشته باشد. در چنین حالتی طراحان نیازی به پیش‌بینی همه‌ی حالات ممکن نخواهند داشت.
- برای حل بسیاری از مسائل در بینایی ماشین، تشخیص صوت، الگوریتم‌های یادگیری به کار می‌آیند.
- یکی از حوزه‌های مطرح در یادگیری «**داده‌کاوی**» است که به استخراج اطلاعات ارزشمند از انبوهی از داده‌ها می‌پردازد.
- شناسایی هویت با کمک چهره یکی از این زمینه‌هاست که در «**بازشناسی الگو**» مطرح می‌شود.



Knowledge Discovery in Database (KDD)

- کاربرد روش‌های یادگیری ماشین در مواجهه با پایگاه‌داده‌های بزرگ، **داده‌کاوی** نامیده می‌شود.
 - خرده‌فروشی: تحلیل سبد خرید، مدیریت ارتباط با مشتری (CRM)
 - کاربردهای مالی: رده‌بندی اعتبار مشتریان، تشخیص سوءاستفاده از کارت‌های اعتباری، بررسی بازار سهام
 - کاربردهای صنعتی: عیب‌یابی، رباتیک و کنترل
 - کاربردهای پزشکی: تشخیص بیماری
 - ارتباطات: تشخیص هرزنامه، تشخیص نفوذ (intrusion detection)
 - وب‌کاوی (موتورهای جستجو) و بیوانفورماتیک

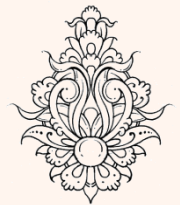


- آنالیز سبد خرید: احتمال این که شخصی که کالای X را خریده، کالای Y را هم خریداری کند.

- $P(Y | X)$

Association Rule

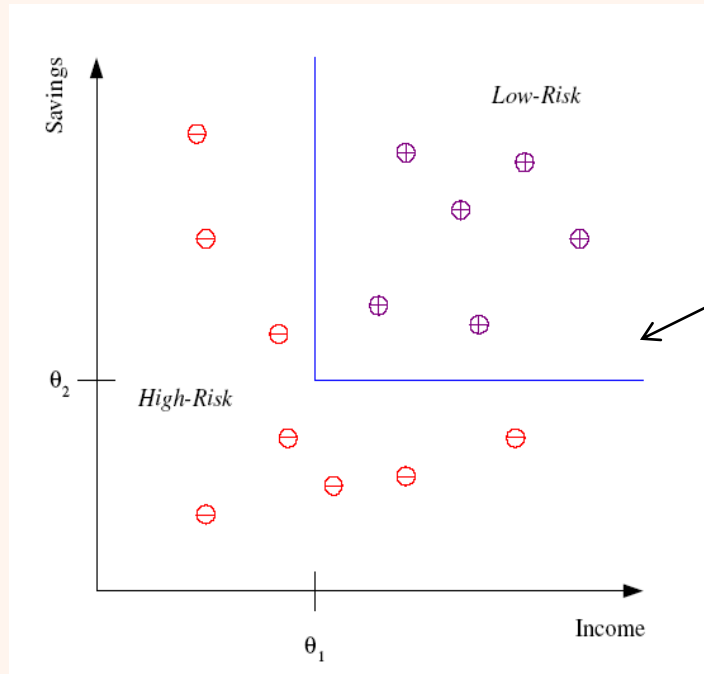
- $P(\text{chips} | \text{beer}) = 0.7$



Classifications

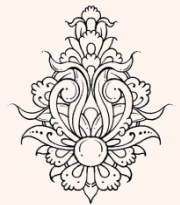
دسته بندی

• ارزیابی اعتبار (credit scoring)



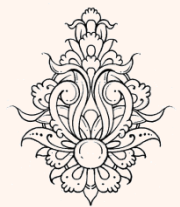
Discriminant

**Discriminant: IF $income > \theta_1$ AND $savings > \theta_2$
THEN low-risk ELSE high-risk**



دسته‌بندی (ادامه...)

- «بازشناسی الگو» هم نامیده می‌شود.
- تشخیص کاراکتر (OCR)
 - تشخیص کاراکترهای دستنویس
 - یک کلمه دنباله‌ای از کاراکترهاست
 - t?e
- تشخیص هویت با استفاده از دست‌خط

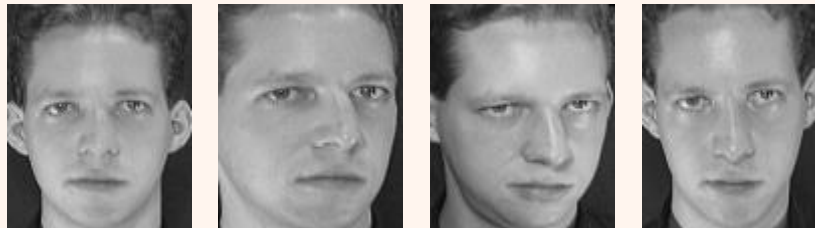


دسته‌بندی (ادامه...)

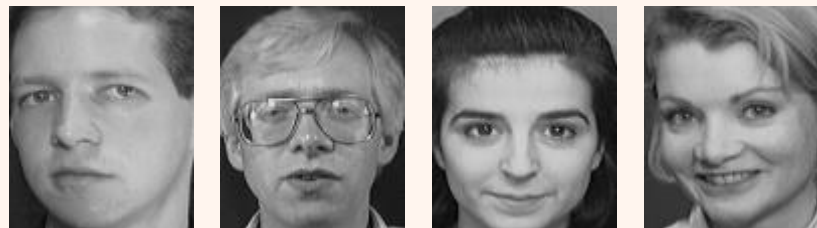
- تشخیص چهره

– شرایط نوری متفاوت،
پوشش و آرایش‌های مختلف

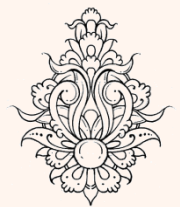
Training examples of a person



Test images

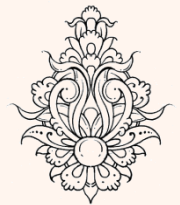


ORL dataset,
AT&T Laboratories, Cambridge UK



دسته‌بندی (ادامه...)

- تشخیص صوت
 - وابستگی زمانی
 - Sensor fusion
- تشخیص بیماری
 - در نظر گرفتن ریسک تصمیم‌گیری
- تشخیص موارد نامتعارف (outlier detection)
 - Intrusion Detection Systems



رگرسیون

- دسته‌بندی و رگرسیون هر دو از نمونه‌های یادگیری با نظارت (supervised) هستند.

- مثال: قیمت اتومبیل دسته‌دو

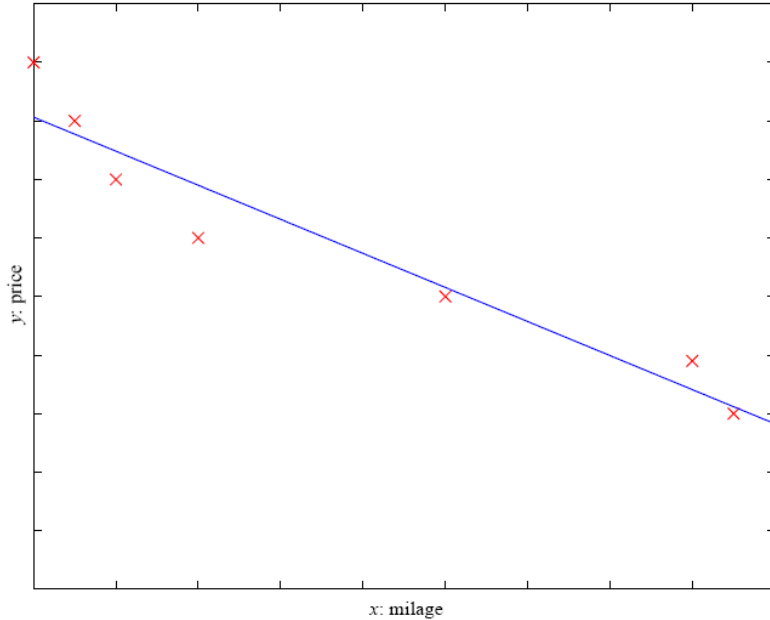
- x : car attributes

y : price

$$y = g(x | \theta)$$

$g()$ model,

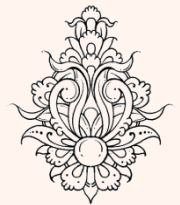
θ parameters



$$y = wx + w_0$$

یا

$$y = w_2x^2 + w_1x + w_0$$

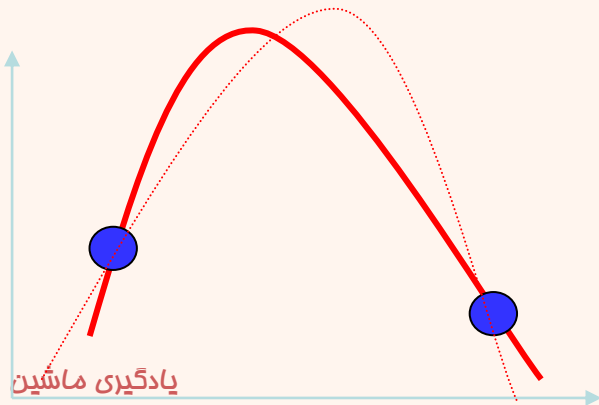


کاربردهای رگرسیون

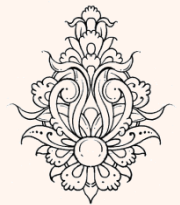
- راهبری یک ربات، ورودی مشخصات اخذ شده از مسگرها خروجی زاویه چرخش
- یکی از کاربردهای رگرسیون بهینه‌سازی یک تابع است.

Response surface design

- یافتن بهترین پارامترها برای بو دادن قهوه
- به دست آوردن تابعی برای ارزیابی کیفیت یک تصویر



From Live Image quality database



شیوه‌های یادگیری

Supervised learning

• یادگیری با نظارت

Unsupervised learning

• یادگیری بی نظارت

Semi-supervised learning

• یادگیری نیمه نظارتی

Active learning

– یادگیری فعال

Weakly supervised learning

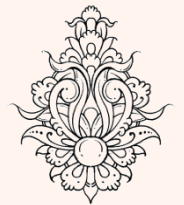
• یادگیری با نظارت ضعیف

Self-supervised learning

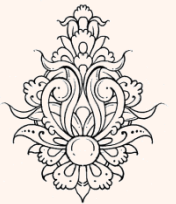
• یادگیری خودنظارتی

Reinforcement learning

• یادگیری تقویتی



- در این شیوه همراه با نمونه‌های آموزشی، پاسخ مطلوب هم وجود دارد.
 - پیش‌بینی نمونه‌های جدید
 - استخراج دانش
 - فشرده‌سازی
 - تشخیص نمونه‌های غیرنرمال؛ تشخیص تقلب و سوءاستفاده



- در این حالت تنها داده‌های ورودی وجود دارند، بدون این که ناظر مقدار مطلوب را مشخص کند.
- هدف پیدا کردن «نظم» (regularity) موجود در داده است، آنچه معمول و طبیعی است.

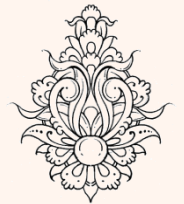
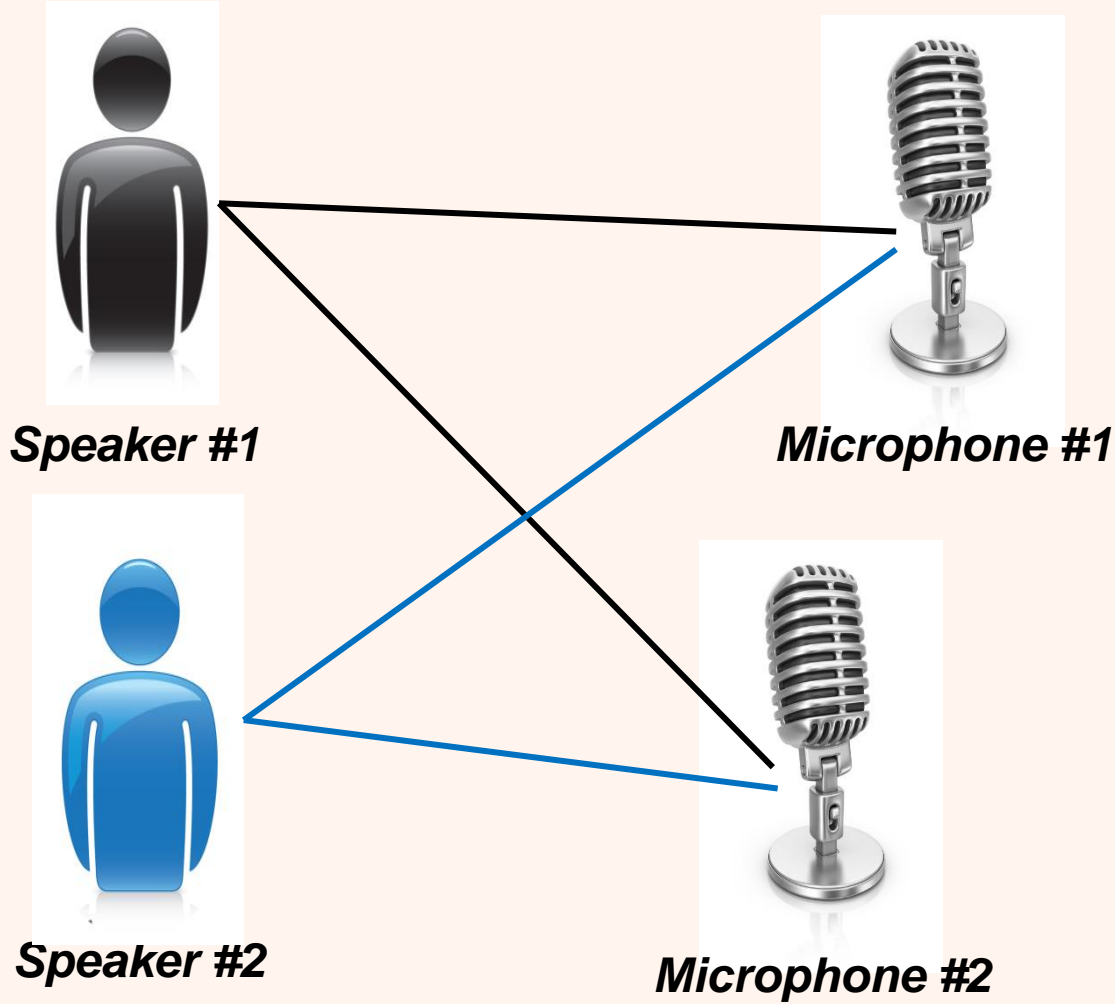
Density estimation

- خوشه‌بندی (clustering): گروه‌بندی نمونه‌های مشابه

- مدیریت ارتباط با مشتری
- فشرده‌سازی تصویر (چندی‌سازی رنگ)
- بیوانفورماتیک (Learning motifs)



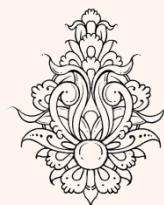
جداسازی کور سیگنال (BSS)



کاربرد یادگیری بی نظارت



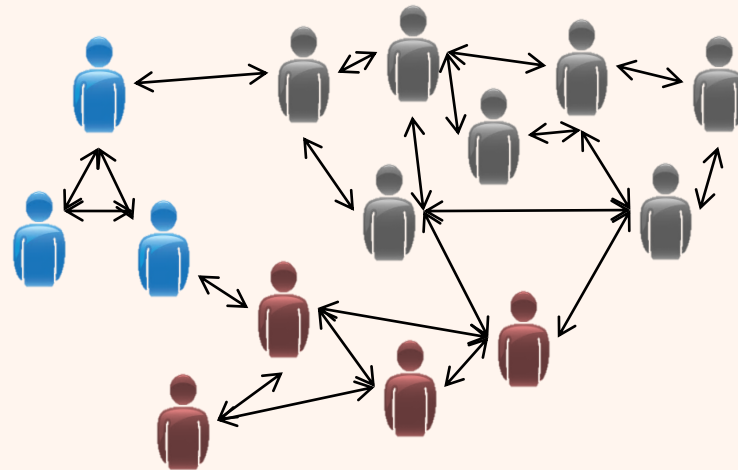
Image Segmentation



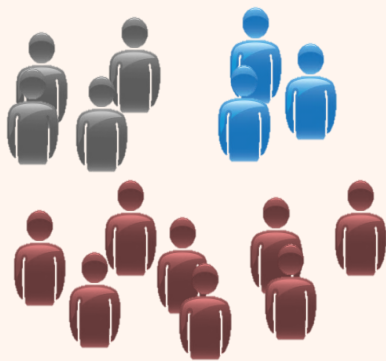
کاربرد یادگیری بی نظارت (ادامه...)



Organize computing clusters



Social network analysis



Market segmentation

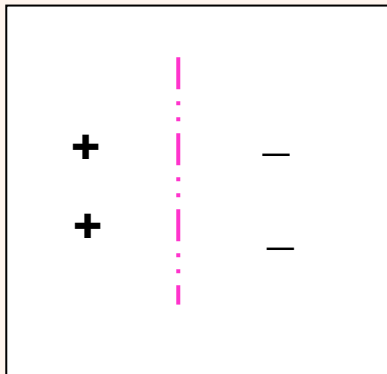


Image credit: NASA/JPL-Caltech/E. Churchwell (Univ. of Wisconsin, Ma

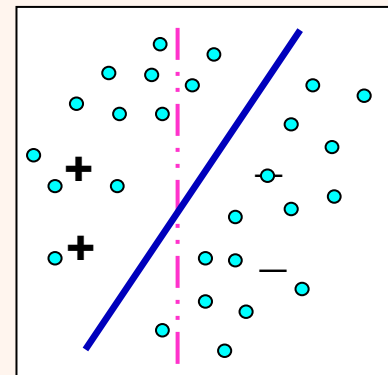
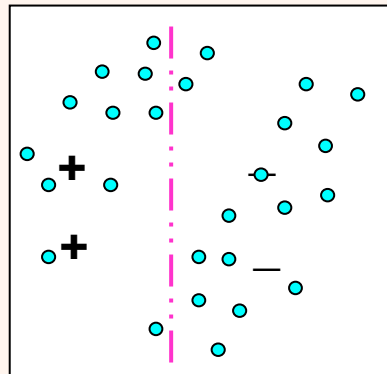
Astronomical data analysis



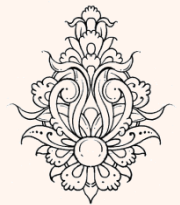
- تنها بخشی از داده‌ها برچسب خورده‌اند و حجم زیادی از آن بدون برچسب هستند.
- برچسب زدن داده‌ها کار پرهزینه‌ای است.
- از طرفی، داده‌های برچسب نخورده‌ی زیادی در اختیار داریم.



یادگیری با نظارت



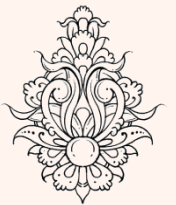
یادگیری نیمه نظارتی



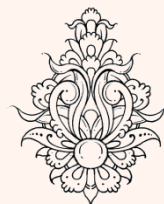
یادگیری با نظارت ضعیف

Weakly supervised learning

- در این حالت از داده‌هایی که برچسب آن‌ها نویزی است (دقت کافی ندارد یا ناقص است) در آموزش با نظارت مورد استفاده قرار می‌گیرند تا بتوان مدل‌های کاراتری را آموزش داد.



- در این شیوه از داده‌های بدون برچسب استفاده می‌شود.
- هدف ایجاد برچسب برای داده‌های بدون برچسب با کمک ساختار و ویژگی‌های داده‌هاست.
- برای به دست آوردن **بازنمایی داده‌ها** این شیوه مورد استفاده قرار می‌گیرد.

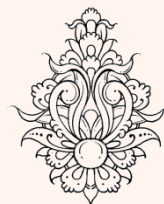


- در برخی موارد فروچی یک سیستم، دنباله‌ای از «کنش»هاست. به گونه‌ای که یک حرکت اهمیت ندارد، بلکه سیاستی است که باعث می‌شود مجموع حرکات، به هدف مناسب برسند.
- یک عمل مناسب است در صورتی که در مجموع و در کنار سایر اعمال مناسب باشد. در این حالت الگوریتم یادگیری باید قادر به انتخاب سیاست مناسب باشد.

Game playing

Robot in a maze

Multiple agents, partial observability, ...



ارزیابی الگوریتم‌های یادگیری

- بسته به کاربرد، برای ارزیابی الگوریتم‌های یادگیری، **درستی** دسته‌بندی، حجم محاسبات و حافظه‌ی مورد نیاز در نظر گرفته می‌شود.
- الگوریتم‌های یادگیری متفاوتی وجود دارند؛ بسته به شرایط کاربرد مورد نظر، الگوریتم‌های متفاوتی را می‌توان مورد استفاده قرار داد.
- حجم مورد نیاز داده‌های آموزشی، پیچیدگی الگوریتم‌های مورد استفاده و قابلیت تعمیم مسائلی است که باید مورد بررسی قرار گیرند.

