

# یادگیری ماشین

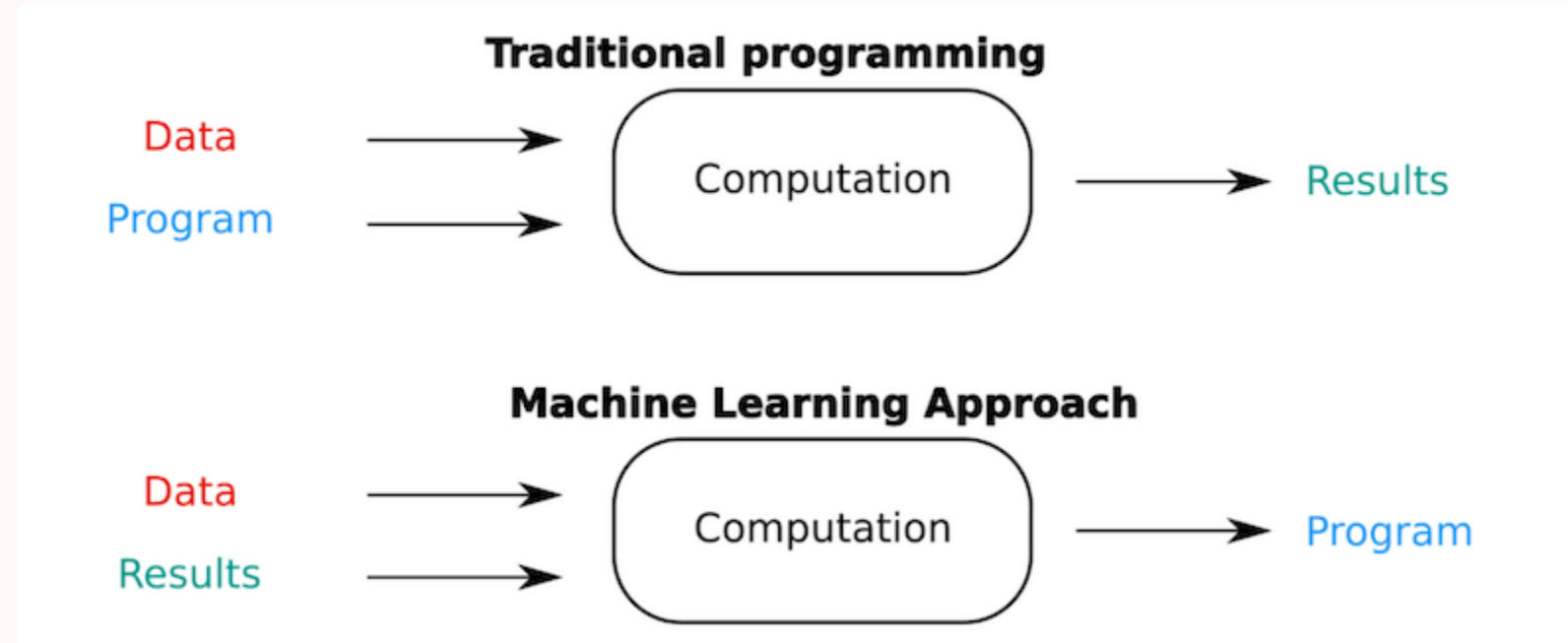
# تعریف یادگیری

- تعریف فرهنگ لغات از یادگیری:
- یادگیری عبارت است از بدست آوردن دانش و یا فهم آن از طریق مطالعه، آموزش و یا تجربه
- همچنین گفته شده است که یادگیری عبارت است از بهبود عملکرد از طریق تجربه
- تعریف یادگیری ماشین:
- یادگیری ماشین عبارت است از اینکه چگونه میتوان برنامه ای نوشت که از طریق تجربه یادگیری کرده و عملکرد خود را بهتر کند .
- یادگیری ممکن است باعث تغییر در ساختار برنامه و یا داده ها شود.

Machine Learning Definition: learn from data

# Machine Learning:

Tom Mitchell: "A computer program is said to learn from experience  $E$ , with respect to some class of tasks  $T$ , and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ ."



Witten and Frank (2000): things learn when they change their behavior in a way that makes them perform better in the future.

Ethem Alpaydın (2010): Machine learning is programming computers to optimize a performance criterion using example data or past experience.

In practice this means:

We have sets of examples from which we want to extract regularities.

# T, P, E

فرض کنید می‌خواهیم یک سیستم هوشمند برای فیلتر کردن هرزنامه طراحی کنیم:

## Environment (E):

- داده‌های ایمیل واقعی از کاربران مختلف
- شامل متغیرهایی مثل: فرستنده، موضوع، متن، زمان ارسال، لینک‌ها و ...

## Task (T):

- یادگیری یک مدل که بر اساس ویژگی‌های ایمیل تصمیم بگیرد آن را در دسته‌ی "هرزنامه" یا "غیرهرزنامه" قرار دهد

## Performance (P):

- دقت طبقه‌بندی: چند درصد از ایمیل‌ها درست دسته‌بندی شدند؟
- Precision: چند درصد از ایمیل‌هایی که به عنوان هرزنامه طبقه‌بندی شدند، واقعاً هرزنامه بودند؟
- Recall: چند درصد از هرزنامه‌های واقعی به درستی شناسایی شدند؟

یادگیری: می‌گوییم یک برنامه کامپیوتری از تجربه E در مورد کار T یادگیری انجام داده است اگر عملکرد آن در صورت اندازه‌گیری با معیار P پس از این تجربه بهبود پیدا کند.

| یادگیری بازی checkers  | یادگیری تشخیص حروف دست نویس  | یادگیری کنترل یک ربات   |
|--|--|---|
| T :انجام بازی checkers<br><br>P : درصد بازی‌هایی که بر حریف غلبه می‌کند<br><br>T.E : انجام بازی بر علیه خودش | T : تشخیص حروف دست نویس در داخل تصویر<br><br>P : درصد تشخیص صحیح حروف<br><br>T.E : database ای از کلمات دست نویس و دسته‌بندی هر کدام | T : هدایت یک ربات در بزرگراه چهار بانده با استفاده از دوربین<br><br>P : میانگین فاصله طی شده قبل از اینکه خطایی رخ دهد<br><br>T.E : با ضبط عملیات یک راننده و تصاویر پشت سر هم جاده بدست می‌آید |

# چرا ماشین را برنامه نویسی نکنیم؟

- بعضی کارها را به درستی نمیتوان توصیف نمود. در صورتیکه ممکن است آنها را بتوان بصورت مثال های (ورودی/خروجی) معین نمود.
- ممکن است در خیل عظیمی از داده اطلاعات مهمی نهفته باشد که بشر قادر به تشخیص آن نباشد (داده کاوی)
- ممکن است موقع طراحی یک سیستم تمامی ویژگی های آن شناخته شده نباشد در حالیکه ماشین میتواند حین کار آنها را یاد بگیرد.
- ممکن است محیط در طول زمان تغییر کند. ماشین میتواند با یادگیری این تغییرات خود را با آنها وفق دهد.

# یادگیری مدل

- در یادگیری ماشین با استفاده از تئوری اطلاعات مدل‌های ریاضی ساخته میشود که میتوانند برای استنتاج استفاده شوند.
- مدل ممکن است پیشگویانه (Predictive) باشد که برای پیش بینی موارد جدید به کار می روند.
- مدل ممکن است توصیفی باشد (Descriptive) که دانش آن از داده یادگرفته میشود.
- البته مدل می تواند هر دو خاصیت فوق را داشته باشد.

# انواع یادگیری

یادگیری ماشین در مسائل مختلفی کاربرد دارد:

## Classification

ماشین یاد می‌گیرد که ورودی‌ها را به دسته‌های از پیش تعیین شده‌ای نسبت دهد.

## Clustering

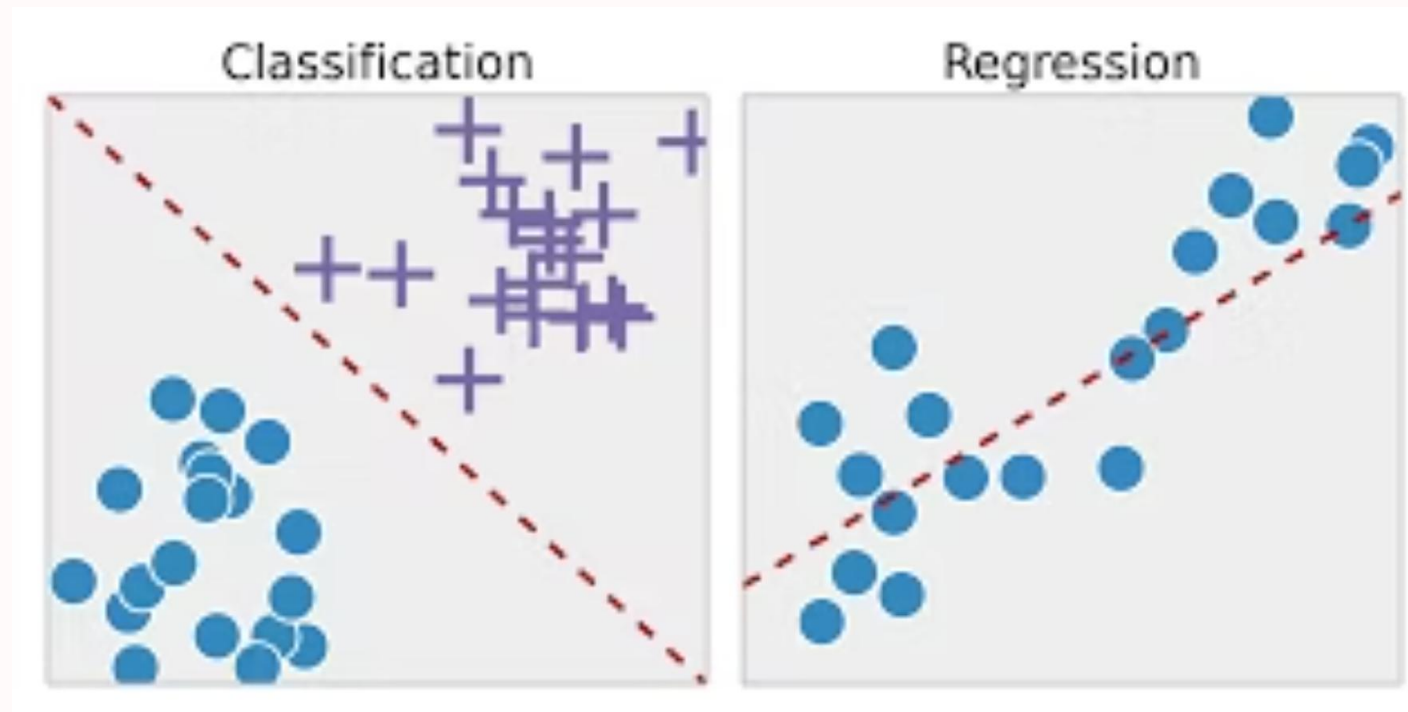
سیستم یادگیر کشف می‌کند که کدام ورودی‌ها با هم در یک دسته بندی قرار می‌گیرند.

## Numeric prediction

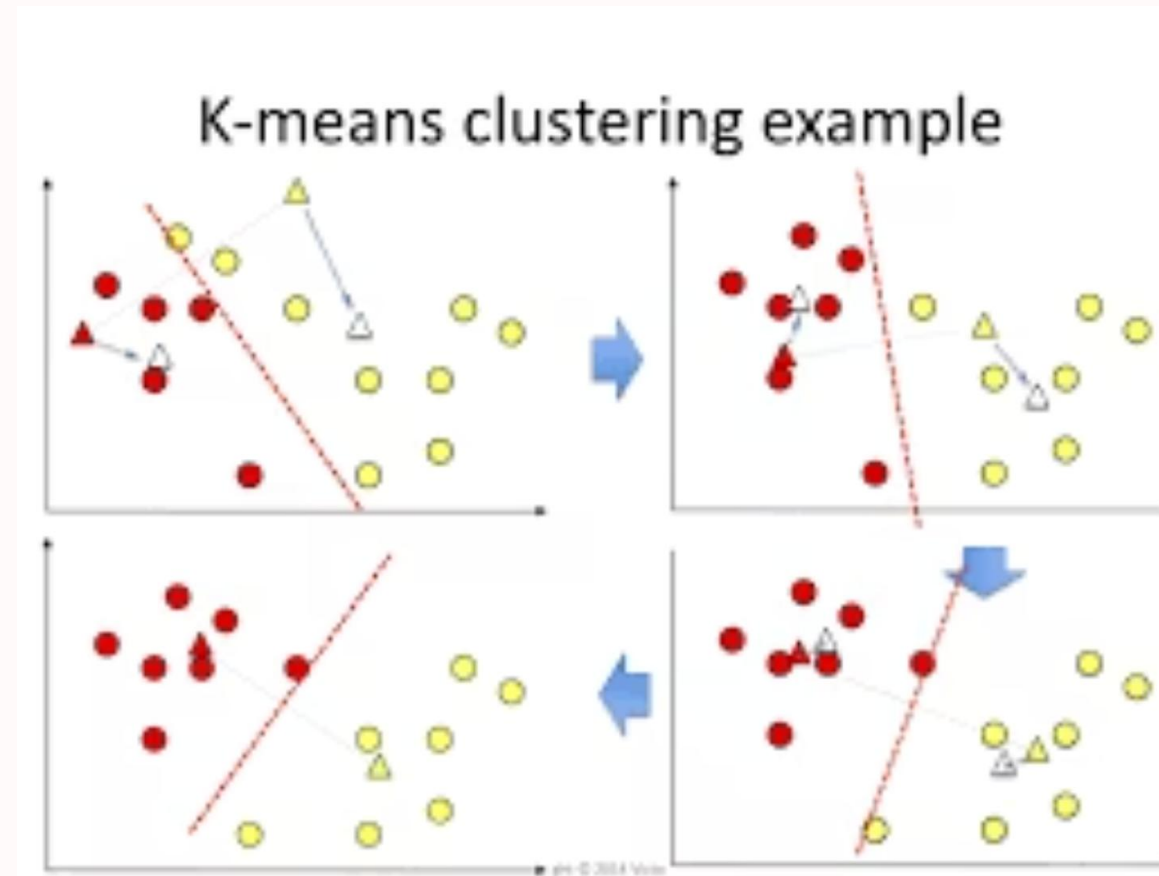
ماشین یاد می‌گیرد که به جای تعیین دسته بندی یک ورودی مقدار عددی آن را پیش بینی نماید.



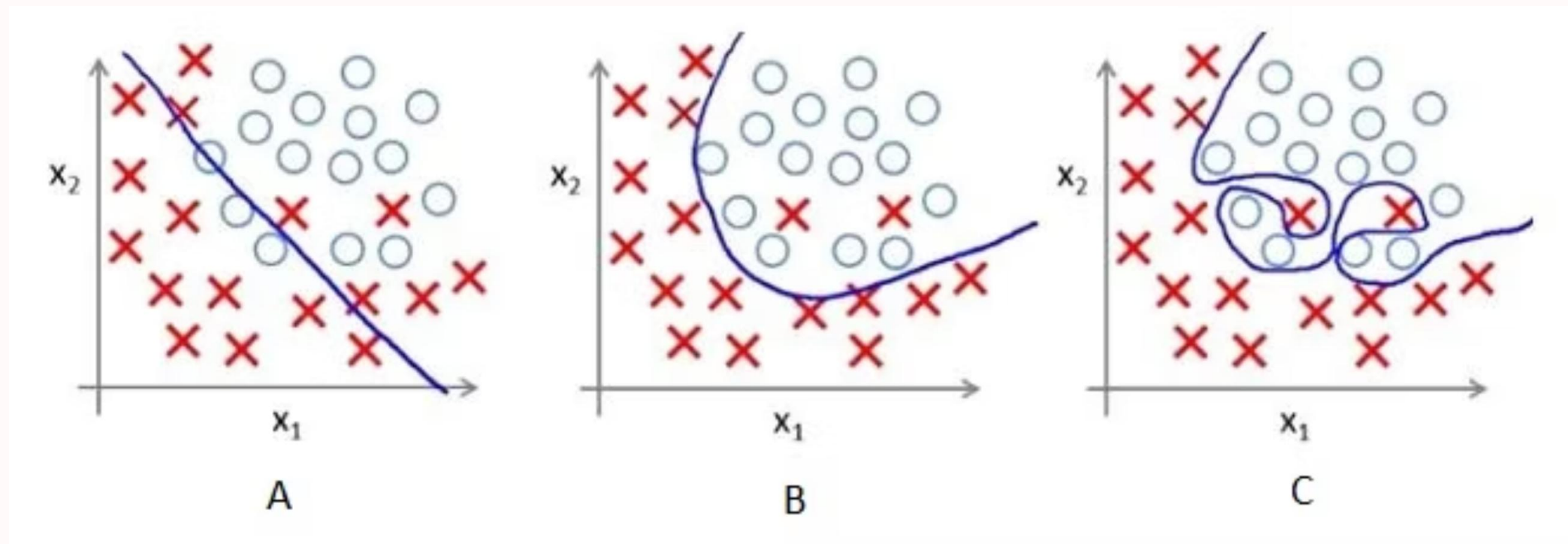
# Classification



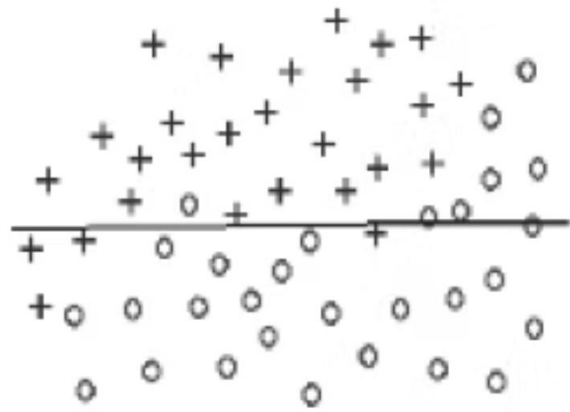
# k-means clustering



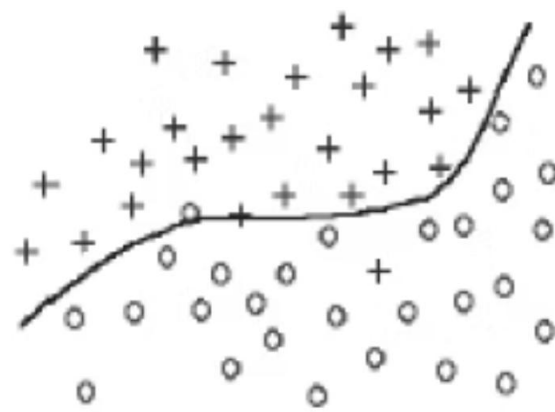
# Logistic Regression Algorithm



# Underfitting & Overfitting



underfitting



good fit



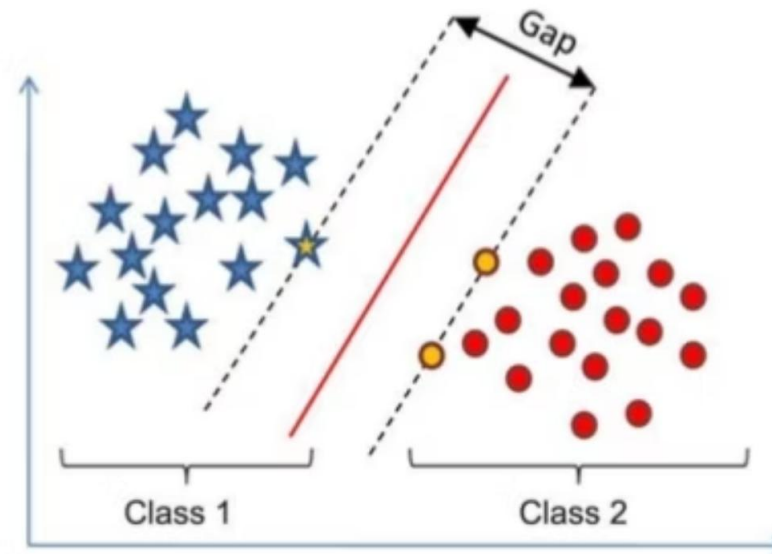
overfitting

# SVM

## Basic concept of SVM

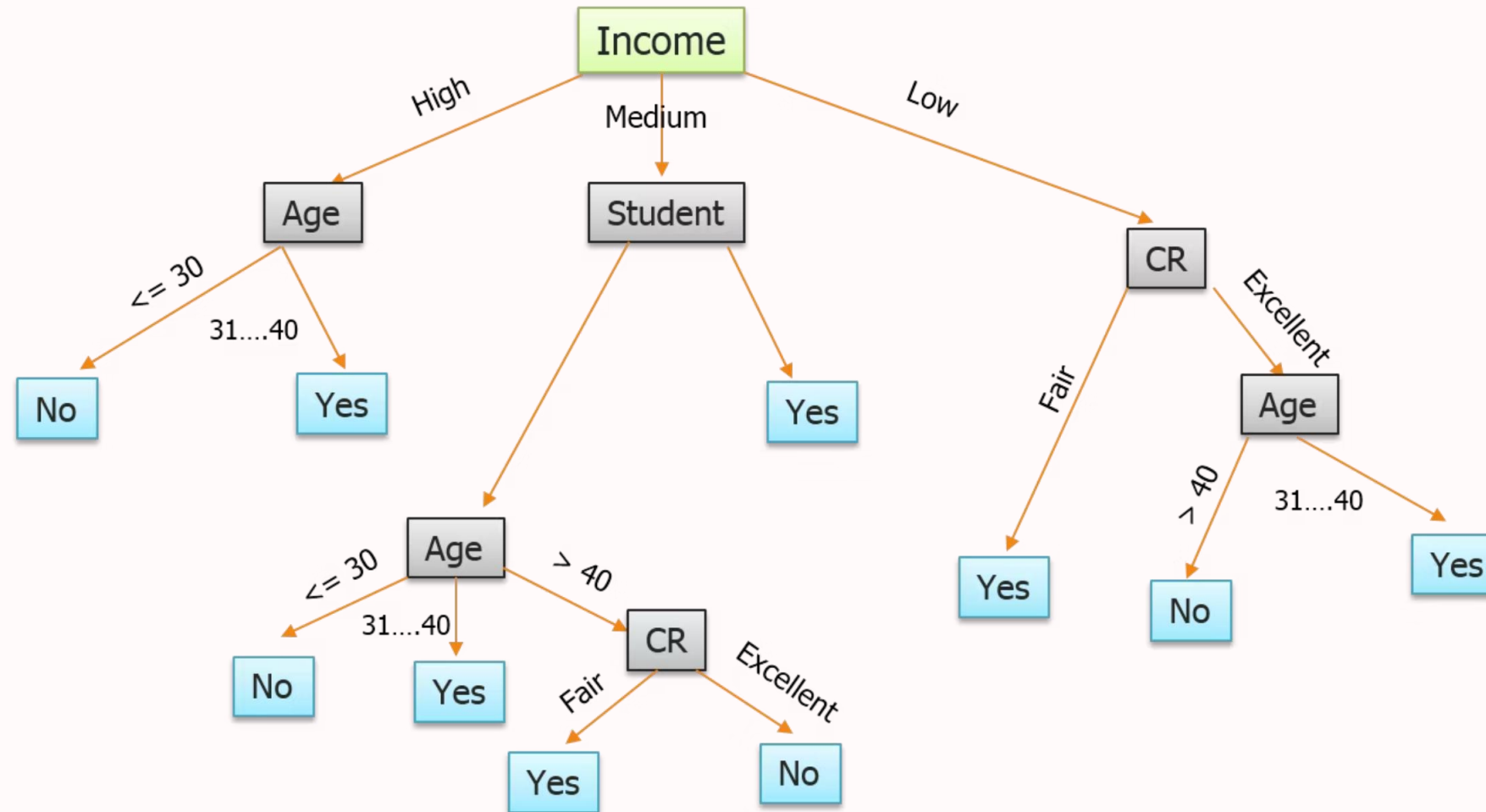
Digg Data

### Basic concept of SVM



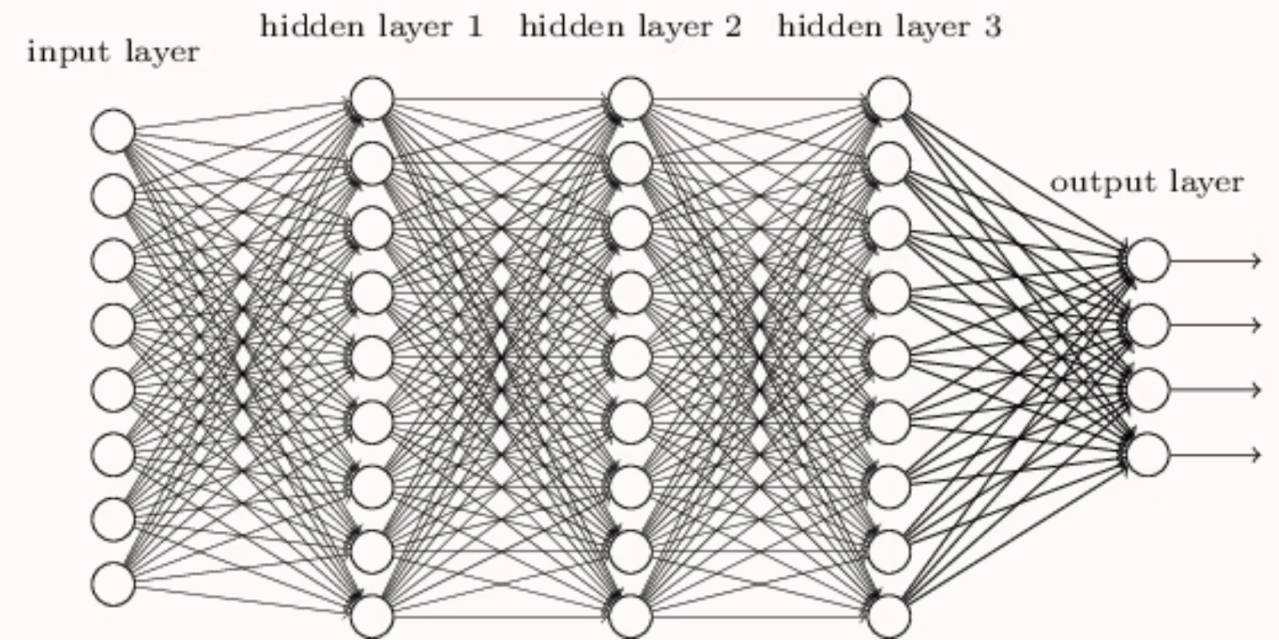
Find a linear decision surface ("hyperplane") that can separate classes and has the largest distance (i.e., largest "gap" or "margin") between border-line patients (i.e., "support vectors")

# Decision Trees





# Neural Network



# برخی کاربردهای موفق یادگیری ماشین

- شناسائی الگو:
  - شناسائی چهره و حالات آن
  - شناسائی حروف دست نویس
  - شناسائی گفتار
- شناسائی رفتارهای نادرست
  - تشخیص خرابی سیگنالهای سنسور ها
  - تشخیص سواستفاده از کارت‌های اعتباری
- پیش بینی
  - قیمت سهام
  - پیش بینی قیمت ارز



# دسته بندی یادگیری

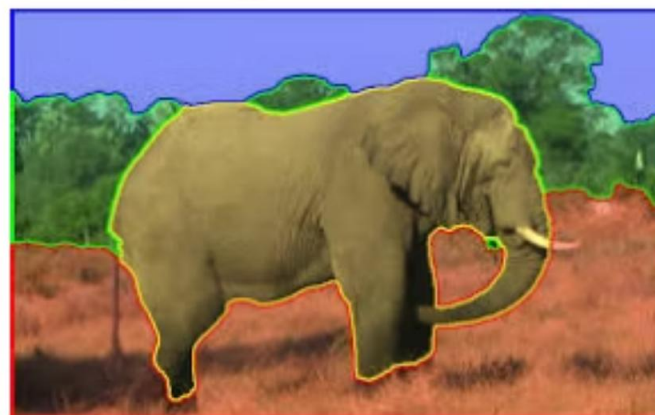
- یادگیری با ناظر: یک مجموعه از مثالهای یادگیری وجود دارد بازای هر ورودی، مقدار خروجی و یا تابع مربوطه نیز مشخص است. هدف سیستم یادگیر بدست آوردن فرضیه ای است که تابع و یا رابطه بین ورودی و یا خروجی را حدس بزند پیش بینی نمونه های جدید، فشرده سازی، تشخیص تقلب
- یادگیری بدون ناظر: یک مجموعه از مثالهای یادگیری وجود دارد که در آن فقط مقدار ورودی ها مشخص است و اطلاعاتی در مورد خروجی صحیح در دست نیست. یادگیری بدون ناظر برای دسته بندی ورودیها و یا پیش بینی مقدار بعدی بر اساس موقعیت فعلی بکار میرود. مدیریت ارتباط با مشتری، بیوانفورماتیک
- یادگیری تقویتی: مثالها بصورت ورودی/خروجی نیستند بلکه بصورت وضعیت/پاداش هستند که یادگیر در وضعیت های مختلف عملیات مختلفی را انجام داده و پاداشهای متفاوتی دریافت و بر اساس مجموع پاداش های دریافتی عمل متناسب با هر وضعیت را یاد میگیرد.
- یادگیری نیمه نظارتی: مثالها طوری هستند که برای تعداد کمی از آنها مقدار خروجی موجود است اما برای مثالهای زیادی مقدار خروجی مشخص نیست.

# کاربرد یادگیری بی نظارت

market segmentation



Image Segmentation



Social Network Analysis



Clustering



Object Categorization

