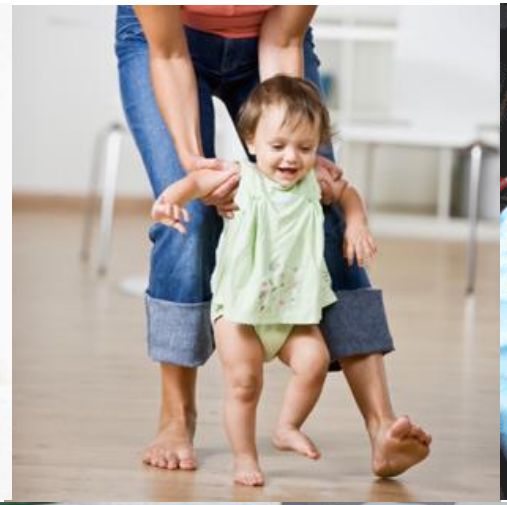
Machine Learning

Why “Learning”

- Learning is acquiring new, or modifying existing, knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information. The ability to learn is possessed by humans and animals.

http://en.wikipedia.org/wiki/Learning

Learning



Why “Learning”

- Machine learning is programming computers to optimize a performance criterion using example data or past experience
- There is no need to “learn” to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

Why “Learning”

- Learning general model from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: consumer transactions to consumer behavior:

  People who bought “Da Vinci Code” also bought “The Five People You Meet in Heaven” (www.amazon.com)

- Build a model that is *a good and useful approximation* to the data.

What is Machine Learning?

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being clearly programmed

Machine Learning

- Tom Mitchell (1998) 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.

Machine Learning

- Machine Learning
  - Study of algorithms that improve their performance at some task with experience
  - Grew out of work in AI
  - New capability for computers
- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference

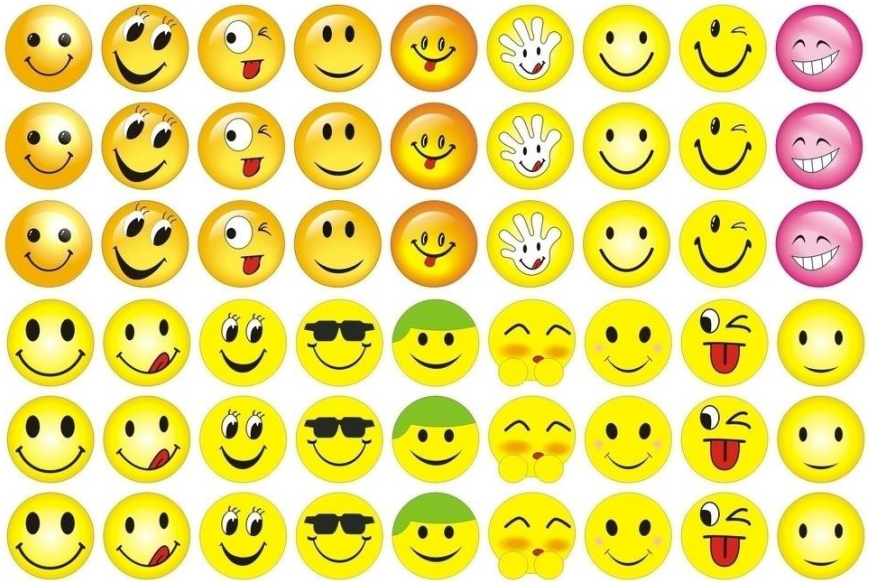
Usage Examples

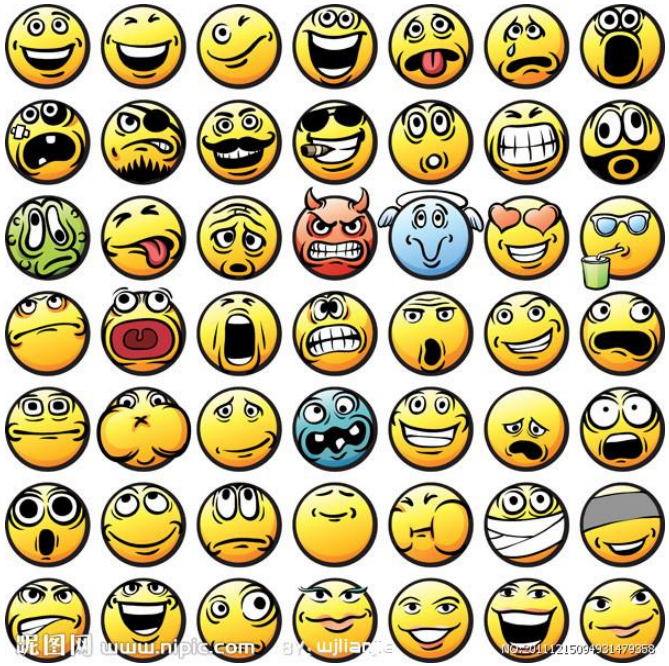
- Examples:
  - Database mining
    - Large datasets from growth of automation/web
    - E.g., Web click data, medical records, biology, engineering
  - Applications cannot program by hand
    - E.g., handwriting recognition, most of Natural Language Processing (NLP), computer vision
  - Self-customizing Programs
    - E.g., Amazon, Netflix product recommendations
  - Understanding human learning (brain, real AI)

Growth of Machine Learning

- Machine learning is preferred approach to
  - Speech recognition, Natural language processing
  - Computer vision
  - Medical outcomes analysis
  - Robot control
  - Computational biology
- This trend is accelerating
  - Improved machine learning algorithms
  - Improved data capture, networking, faster computers
  - Software too complex to write by hand
  - New sensors / IO devices
  - Demand for self-customization to user, environment

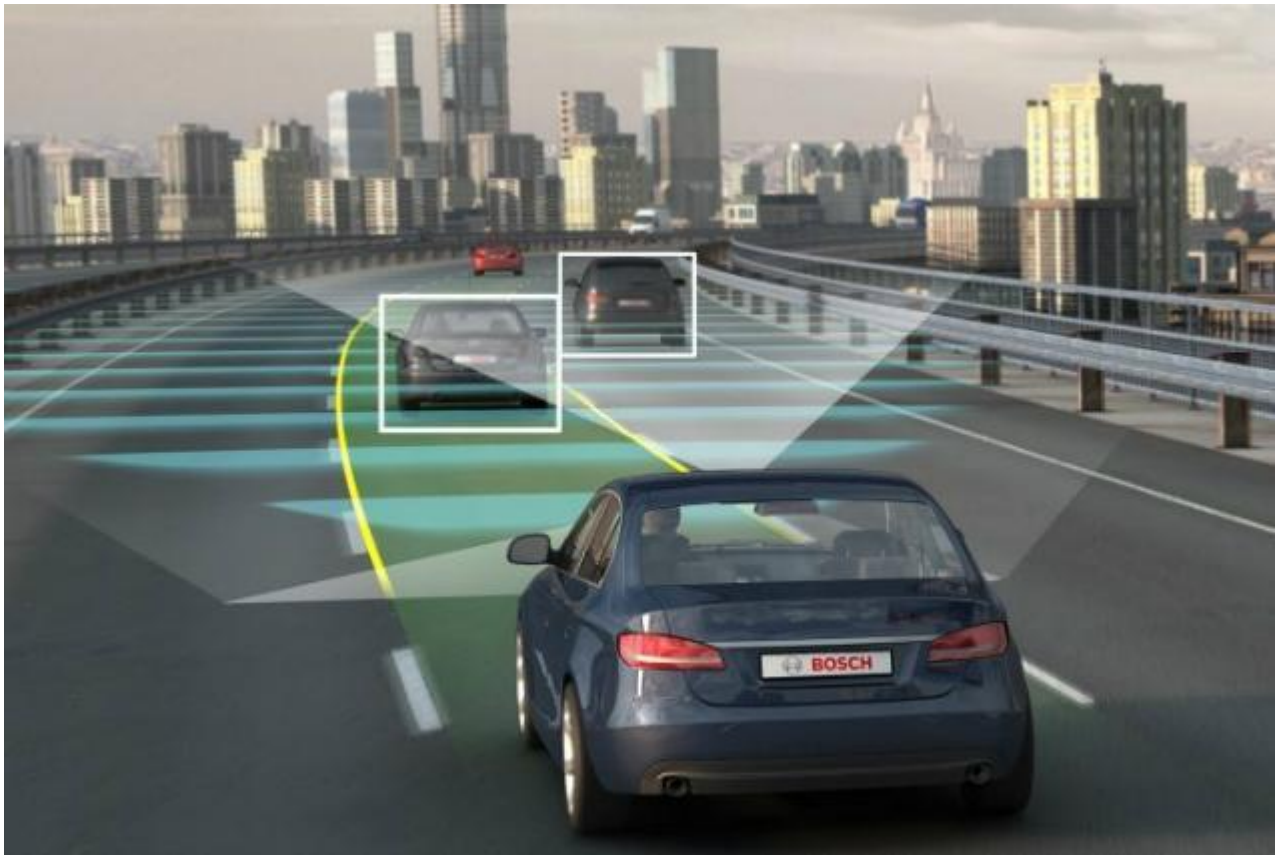
Examples





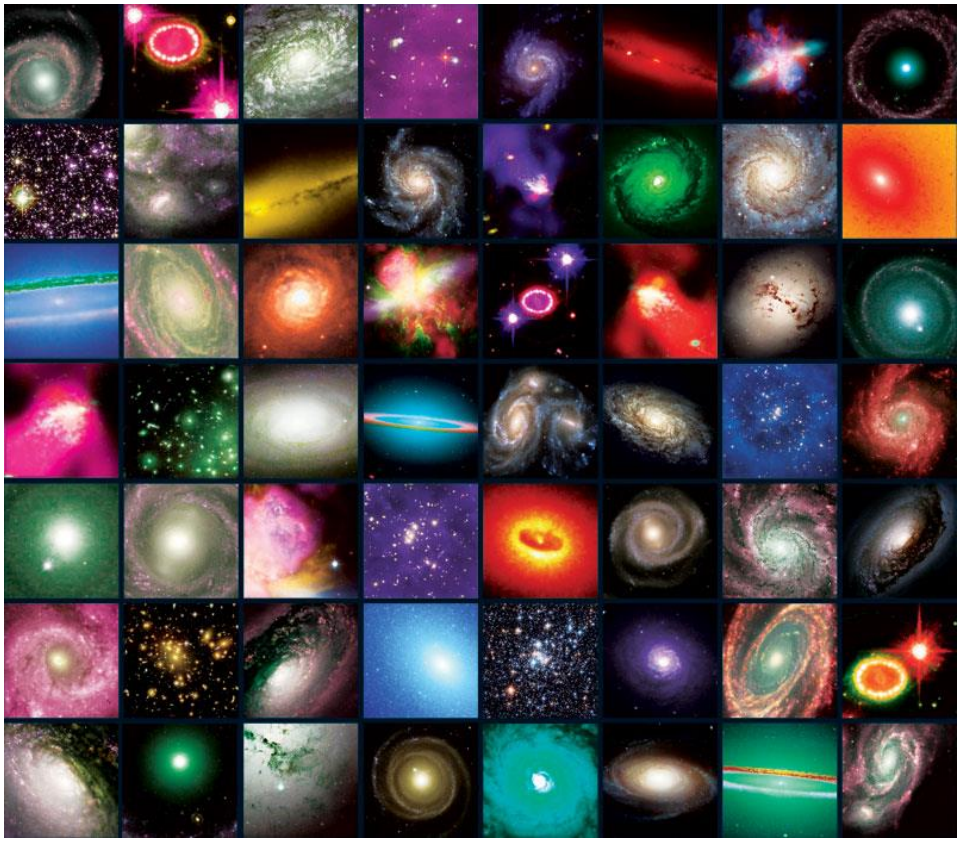
Examples

- Learning to drive an autonomous vehicle



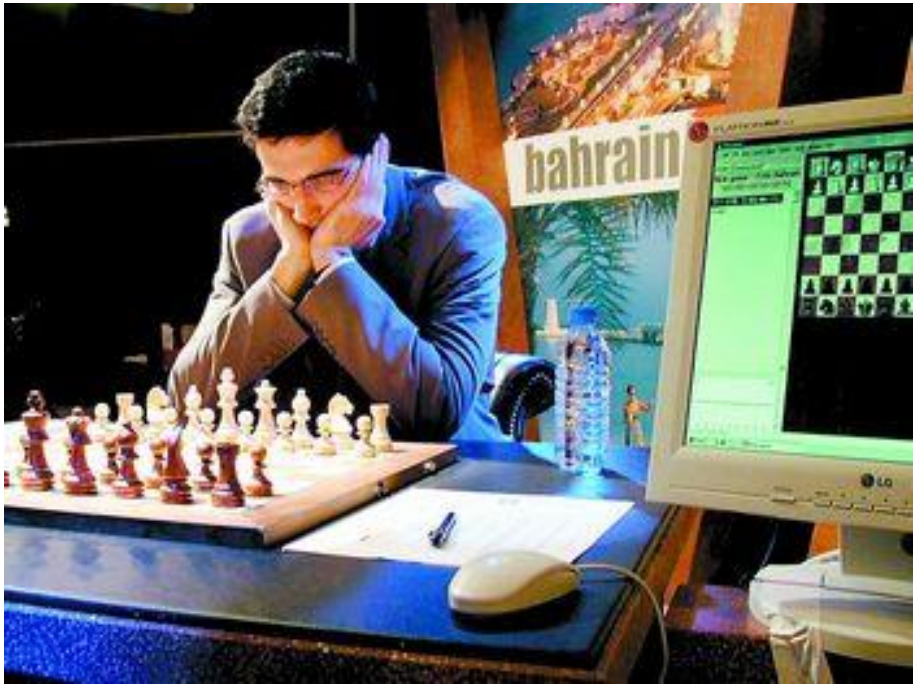
Examples

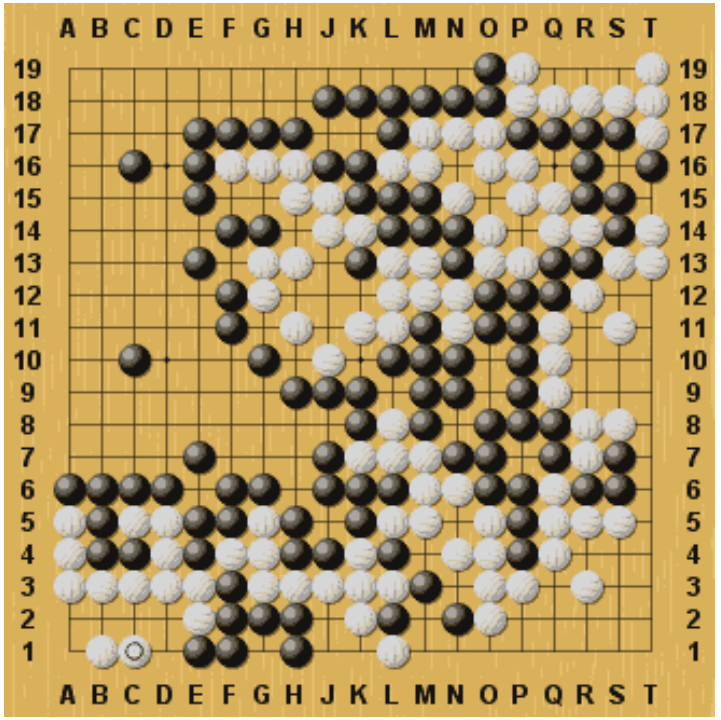
- Learning to classify new astronomical structure



Examples

- Learning to play games





Applications

- Association Analysis
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning

Learning Associations

- Basket analysis:

  P (Y | X) probability that somebody who buys X also buys Y where X and Y are products/services

- Example: *P* (Bread | Milk) = 0.75

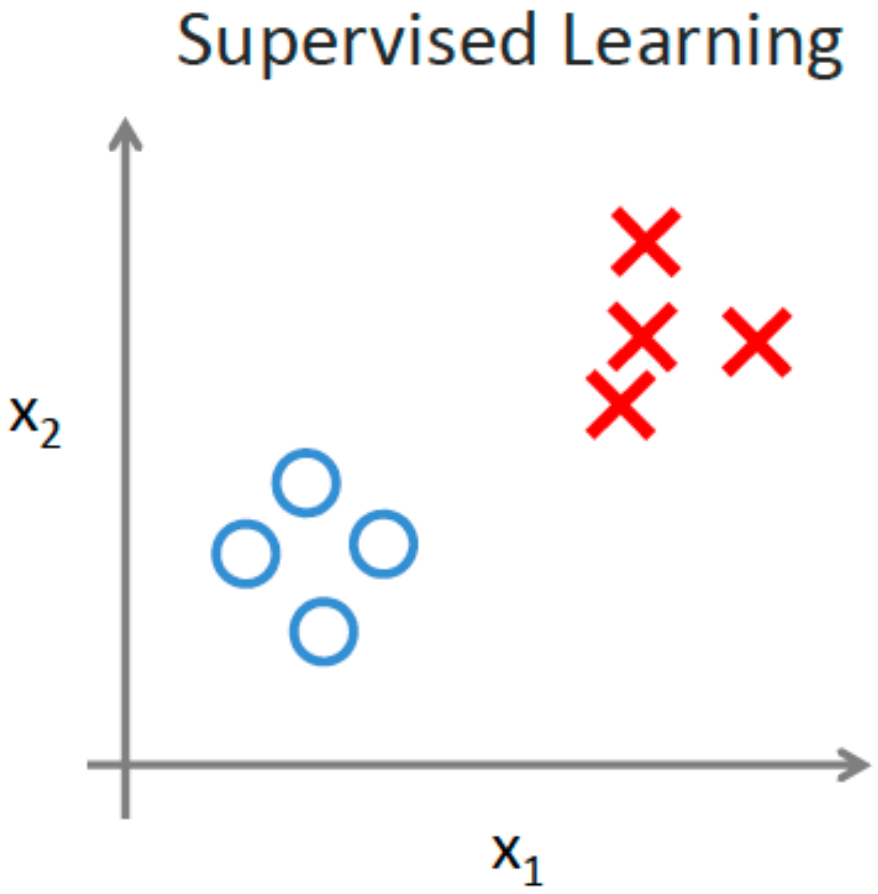
Market-Basket transactions

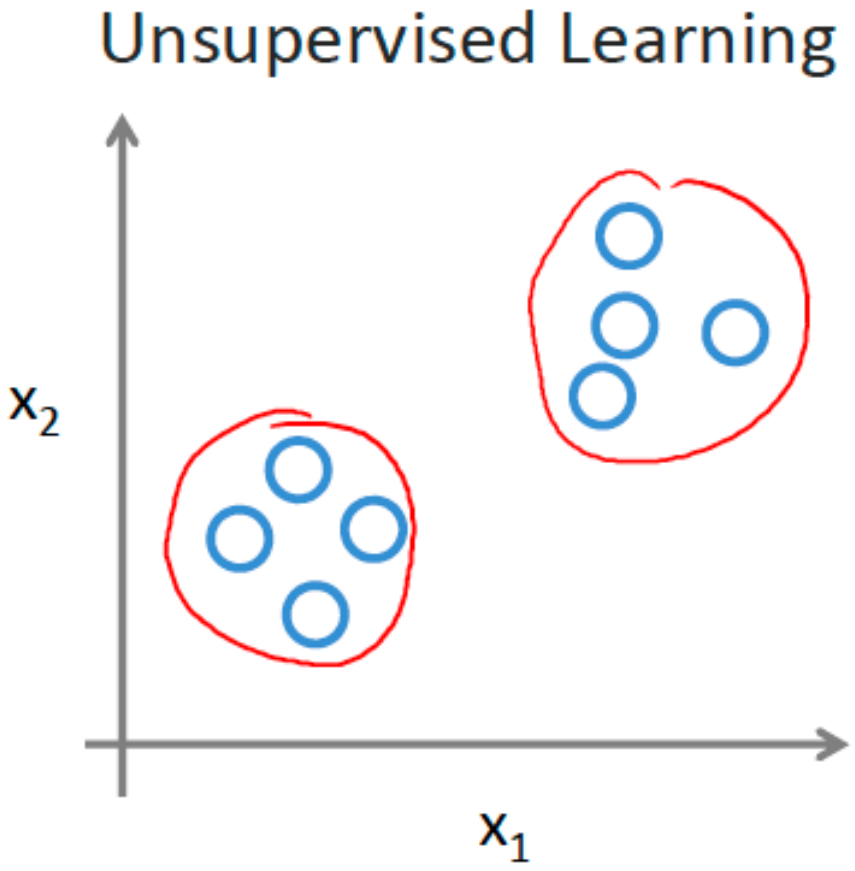
TID	*Items*
1	**Bread, Milk**
2	**Bread, Diaper, Beer, Eggs**
3	**Milk, Diaper, Beer, Coke**
4	**Bread, Milk, Diaper, Beer**
5	**Bread, Milk, Diaper, Coke**

Machine Learning

- Unsupervised Learning
  - Technique of trying to find hidden structure in unlabeled data
- Supervise Learning
  - Technique for creating a function from training data. The training data consist of pairs of input objects (typically vectors), and desired outputs.

Supervised V.S. Unsupervised





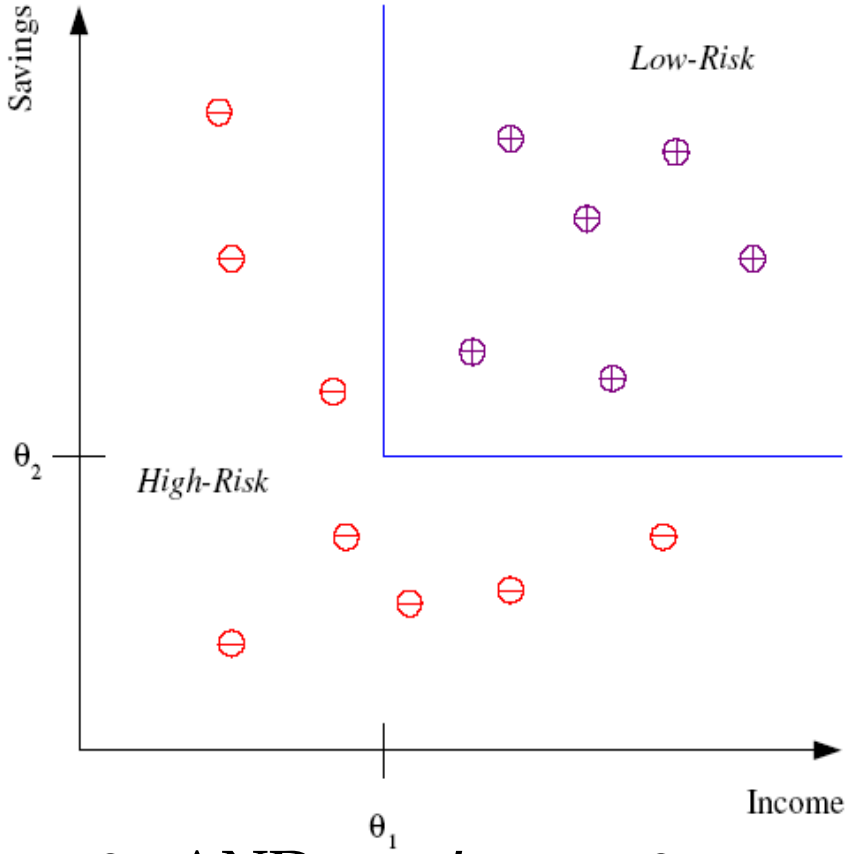
Supervised Learning: Uses

Example: decision trees tools that create rules

- Prediction of future cases: use the rule to predict the output for future inputs
- Knowledge extraction: the rule is easy to understand
- Compression: the rule is simpler than the data it explains
- Outlier detection: exception that are not covered by the rule e.g., fraud

Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



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

Model

Classification Example

- Credit Card Fraud

12/04/12	12/04/12	KROGER #389 NORCROSS GA CASHOVER $ 40.00 PURCHASES $ 3.05	$	43.05	Supermarkets
12/04/12	12/04/12	KROGER #479 ATLANTA GA CASHOVER $ 40.00 PURCHASES $ 2.05	$	42.05	Supermarkets
12/04/12	12/04/12	KROGER #479 ATLANTA GA CASHOVER $ 40.00 PURCHASES $ 13.90	$	53.90	Supermarkets
12/03/12	12/03/12	BREWHOUSE CAFE ATLANTA GA	$	40.00	Restaurants
12/03/12	12/03/12	BRICK STORE PUB DECATUR GA	$	40.00	Restaurants
12/03/12	12/03/12	QT 709 CHAMBLEE GA	$	52.33	Gasoline
12/02/12	12/02/12	ESCORPION 678-666-5198 GA 00248R	$	65.50	Restaurants
12/02/12	12/02/12	TIN LIZZY'S GRANT PARK ATLANTA GA	$	62.24	Restaurants
12/02/12	12/02/12	VORTEX BAR & GRILL ATLANTA GA	$	40.00	Restaurants
12/02/12	12/02/12	WOODYS CHEESESTEAKS ATLANTA GA	$	9.24	Restaurants
12/02/12	12/02/12	WOODYS CHEESESTEAKS ATLANTA GA	$	24.52	Restaurants
12/01/12	12/01/12	CAFE INTERMEZZO-30346 DUNWOODY GA	$	31.55	Restaurants
12/01/12	12/01/12	TWISTED TACO PERIMETER ATLANTA GA	$	17.25	Restaurants
11/30/12	11/30/12	KROGER #636 DULUTH GA	$	115.26	Supermarkets

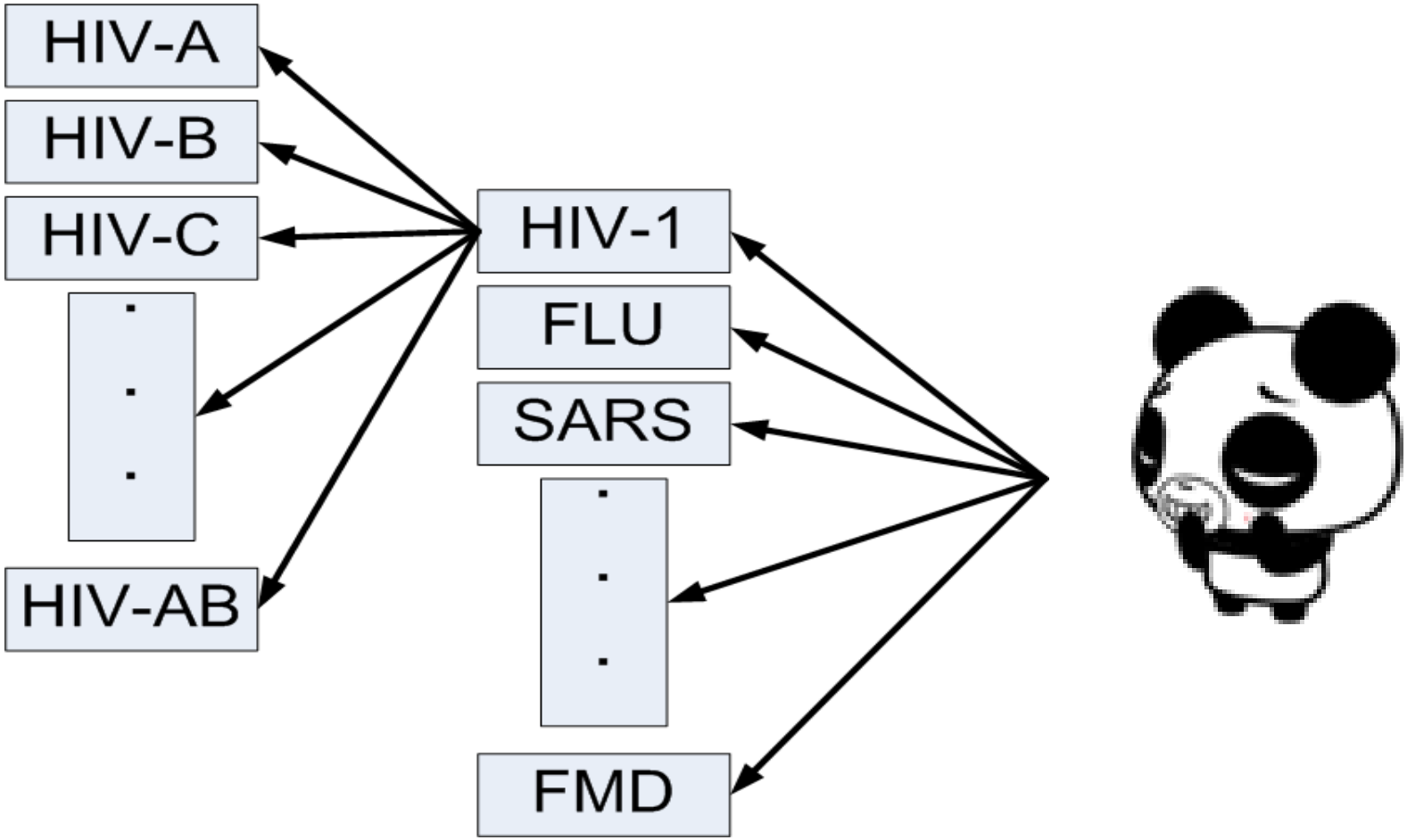
Face Recognition

Training examples of a person

Test images

AT&T Laboratories, Cambridge UK
http://www.uk.research.att.com/facedatabase.html

Clinical Diagnosis



Health-Care

Seasonal Epidemic Influenza

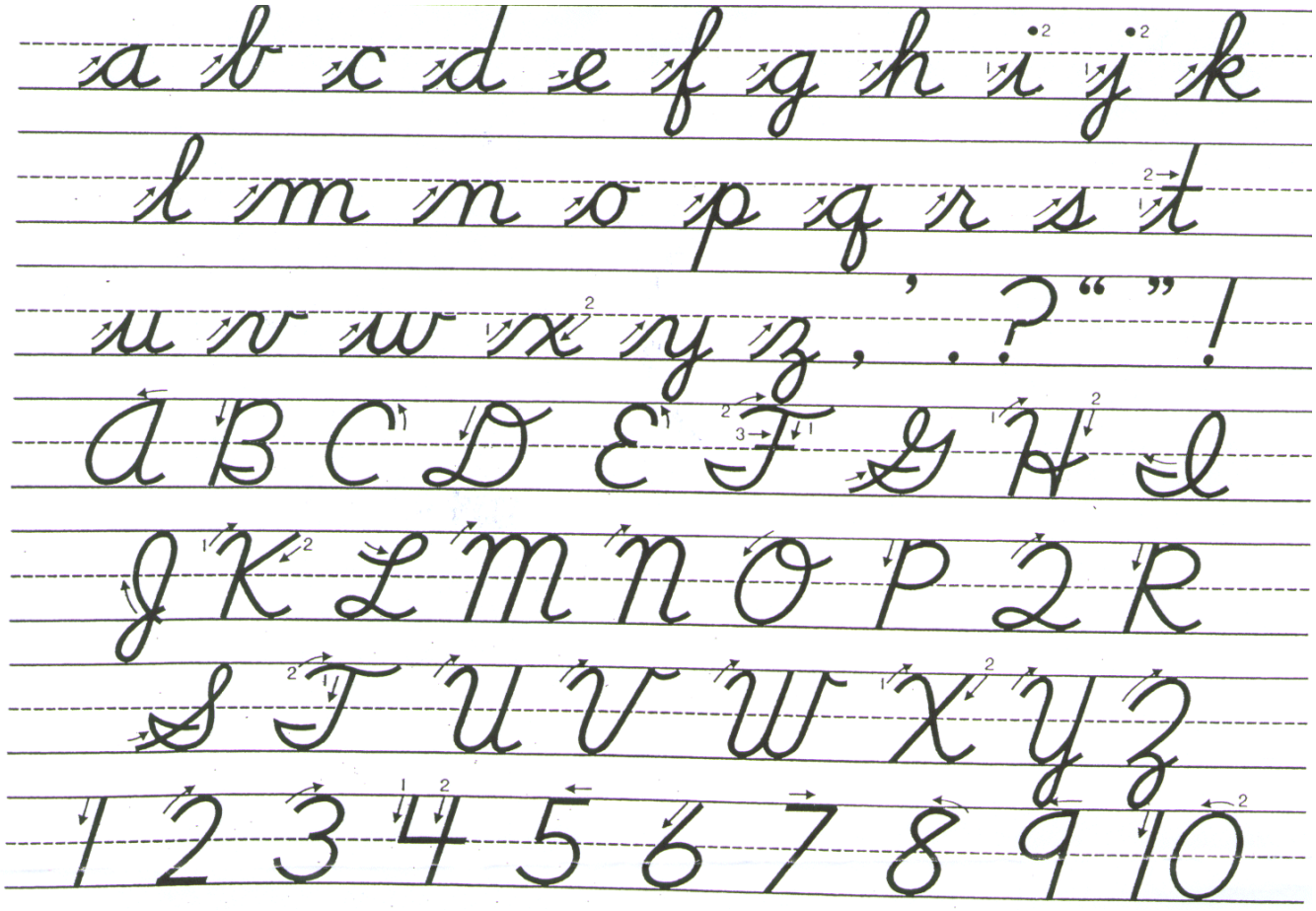
40,000 deaths in US each year

200,000 hospitalizations in US each year

H1N1 | **H3N2** | **Influenza B** | **2009 H1N1**

Billions of dollars lost

Handwriting Recognition



Netflix Price

Netflix's Top 10 and other suggestions for you are based on your viewing habits

Regression

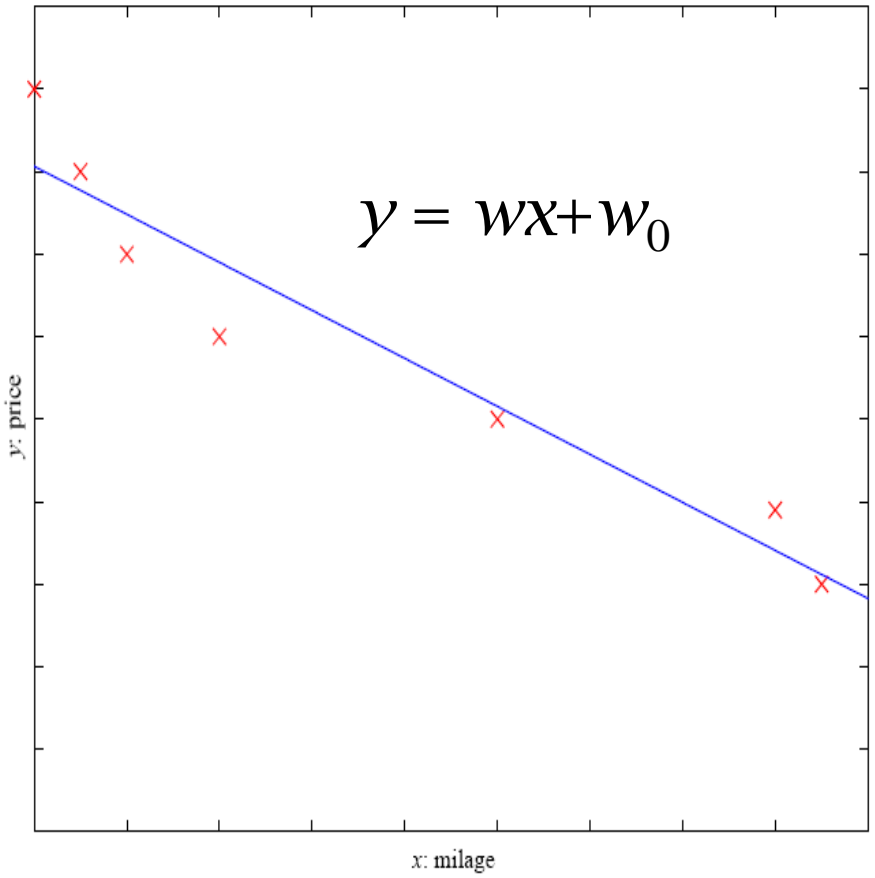
- Example: price of a used car
- x : car attribute

  y : price

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

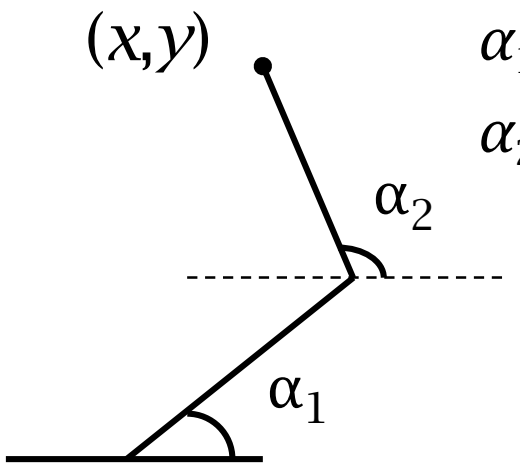
  $g(\ )$ model,

  θ parameters



Regression Applications

- Navigating a car: angle of the steering wheel (CMU NavLab)
- Kinematics of a robot arm



$$\alpha_1 = g_1(x,y)$$
$$\alpha_2 = g_2(x,y)$$

Unsupervised Learning

- Learning “what normally happens”
- Clustering: grouping similar data instances
- Other applications: Summarization, Association Analysis
- Example Applications
  - Customer segmentation
  - Image compression: color quantization
  - Bioinformatics: Learning gene

Reinforcement Learning

- Topics:
  - Policies: what actions should an agent take in a particular situation
  - Utility estimation: how good is a state (→used by policy)
- Credit assignment problem (what was responsible for the outcome)
- Applications:
  - Game play
  - Robot in a maze
  - Multiple agents, partial observability ...

Available Datasets

- UCI Repository: http://www.ics.uci.edu/~mlearn/MLRepository.html
- UCI KDD Archive: http://kdd.ics.uci.edu/summary.data.application.html
- DELVE: http://www.cs.utoronto.ca/~delve/

Pre-requisites

- **Discrete Mathematics**
- **Linear Algebra**
- **Statistics**
- **Calculus**
- **Optimization**
- **Examples:**
  - Eigenvalue and Eigenvector
  - Gradient Descent
  - Convex Optimization
  - Least Square
  - Bayesian Theory