Intro to Machine Learning

J.-S. Roger Jang (張智星)

jang@mirlab.org

http://mirlab.org/jang

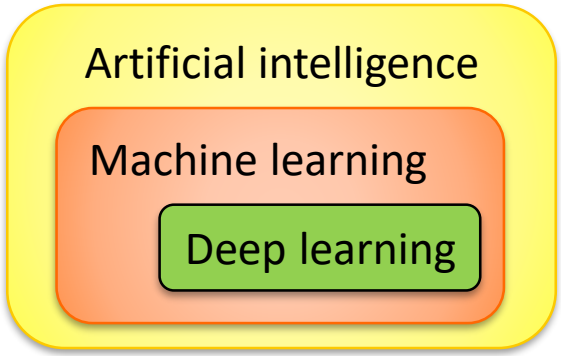
MIR Lab, CSIE Dept.

National Taiwan University

2023/9/20

Machine Learning (ML)

- Definition
  - Field of study that gives computers the ability to learn without being explicitly programmed – by Arthur Samuel, 1959.
  - Computational methods that use existing data to make predictions
- Concepts
  - 視其所以，觀其所由，察其所安，人焉廋哉？人焉廋哉？
  - 近朱者赤、近墨者黑



More about Machine Learning

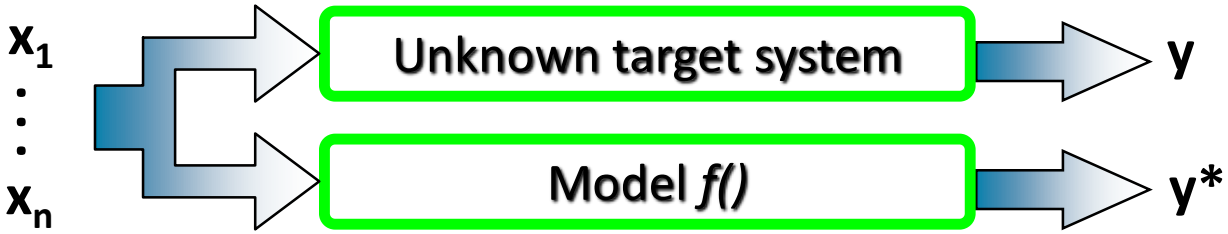
- Quick examples
  - Use height and weight to predict gender
  - Use historical data to predict stock market
- Comparison
  - Before ML ➔ Rule-based
  - After ML ➔ Data-driven
- Prerequisites for ML
  - Probability & statistics
  - Linear algebra
  - Optimization
  - …

Broad Areas of ML

- Classification: Assign a category to a given object
  - Determine if it will rain tomorrow
- Regression: Predict a real-value for a given object
  - Predict the exchange rate
- Ranking: Order objects according to some criterion
  - Rank the webpages returned by a search engine
- Clustering: Partition data into homogeneous groups
- Dimensionality reduction: Find low-dimensional manifold preserving some properties of the data
- Density estimation: Learning probability density function according to sample data

Concept of Modeling

- Given desired i/o pairs (training set) of the form (x_1, ..., x_n; y), construct a model to match the i/o pairs



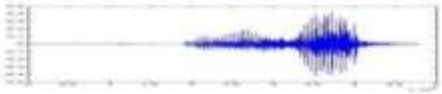
- Two basic steps in modeling

Quiz!

  - structure identification: input selection, model complexity
    - Example: order determination in polynomial fitting
  - parameter identification: optimal parameters
    - Example: coefficient determination in polynomial fitting

Example of Machine Learning (or Modeling)

- Speech Recognition

$f^*($  $) =$ "Morning"

- Handwritten Recognition

$f^*($  $) =$ "2"

- Playing Go

$f^*($  $) =$ "5-5" (step)

- Dialogue System

$f^*($ "Hi" $) =$ "Hello"

(what the user said) (system response)

https://www.slideshare.net/tw_dsconf/ss-70083878

Three Basic Paradigms of ML

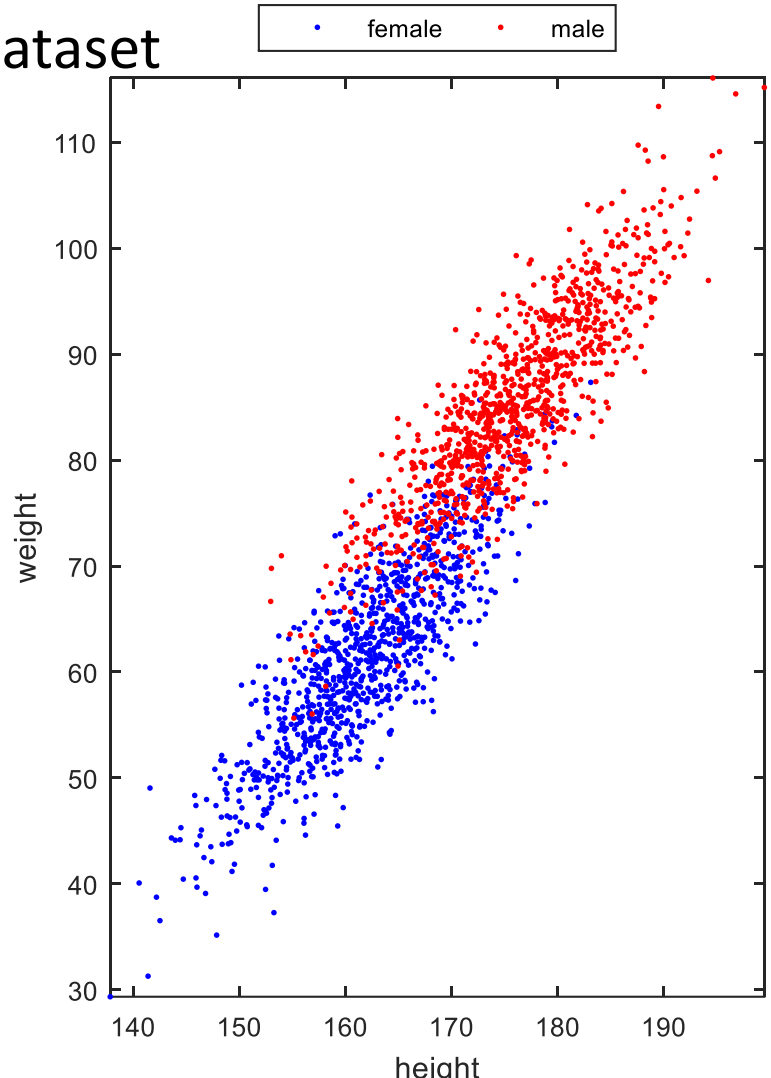
Quiz!

- Supervised learning
  - Each input has output, which is available immediately or with a delay of fixed duration.
  - Two types: classification & regression
  - Examples: Weather prediction, face recognition, …
- Reinforcement learning
  - The output is available but with a delay of variable duration.
  - Example: Chess playing, control system, path planning…
- Unsupervised learning
  - No output part is involved.
  - Example: K-means clustering, principal component analysis, …

GHW Dataset

- GHW (gender-height-weight) dataset
  - 10000 entries with 3 columns
    - Gender: Categorical
    - Height: Numerical
    - Weight: Numerical
  - Source

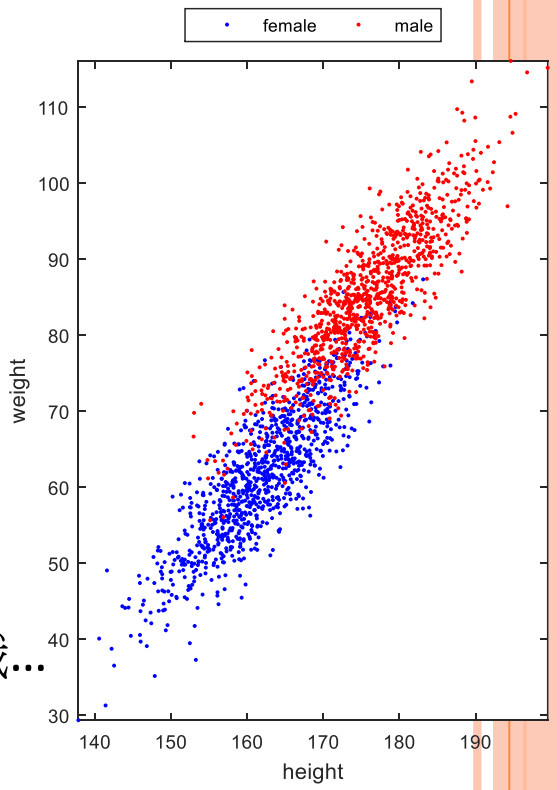
	A	B	C	D	E
1	Gender	Height	Weight	Height (cm)	Weight (kg)
2	Male	73.847017	241.89356	187.5714232	109.72099
3	Male	68.781904	162.31047	174.7060363	73.622732
4	Male	74.110105	212.74086	188.2396677	96.49755
5	Male	71.730978	220.04247	182.1966851	99.809504
6	Male	69.881796	206.3498	177.4997615	93.598619
7	Male	67.253016	152.21216	170.8226598	69.042216
8	Male	68.785081	183.92789	174.7141064	83.428219
9	Male	68.348516	167.97111	173.6052294	76.190352
10	Male	67.01895	175.92944	170.2281321	79.800187
11	Male	63.456494	156.39968	161.1794947	70.941642
12	Male	71.195382	186.60493	180.836271	84.642501
13	Male	71.640805	213.74117	181.967645	96.951285
14	Male	64.766329	167.12746	164.506476	75.807679
15	Male	69.28307	189.44618	175.978098	85.931272



Multiple Ways of Prediction

- Multiple ways of prediction based on GHW dataset
  - Classification: Height, weight ➔ gender
  - Regression: Gender, height ➔ weight
  - Regression: Gender, weight ➔ height
- In general
  - Classification: the output is categorical
  - Regression: the output is numerical
- Sometimes it's not so obvious!
- 範例
  - 分類：信用卡盜刷、警示帳戶、支票金額辨識...
  - 迴歸：房屋估價、貸款金額、收入預估...

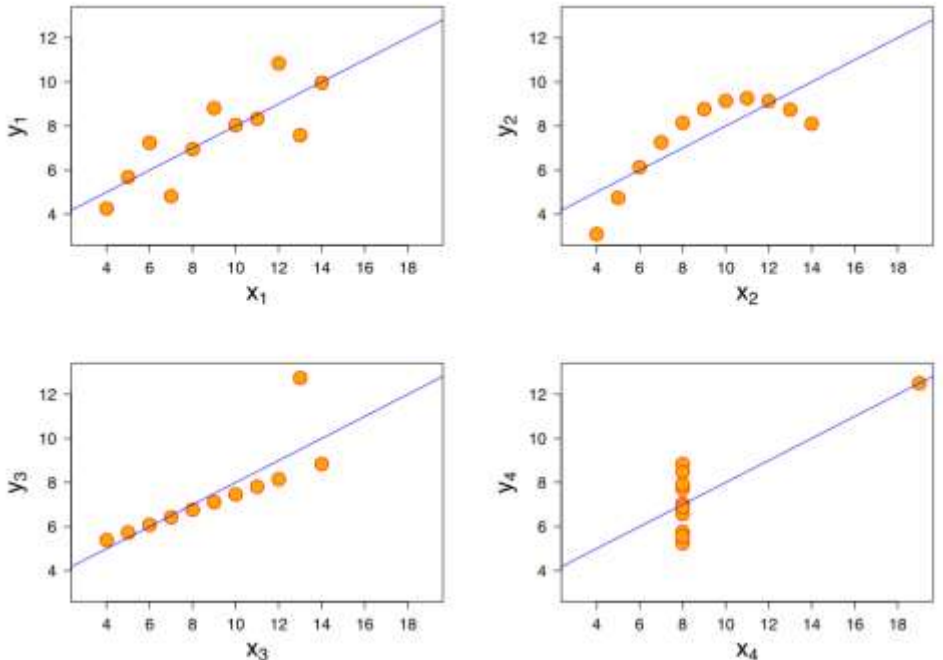
Quiz!



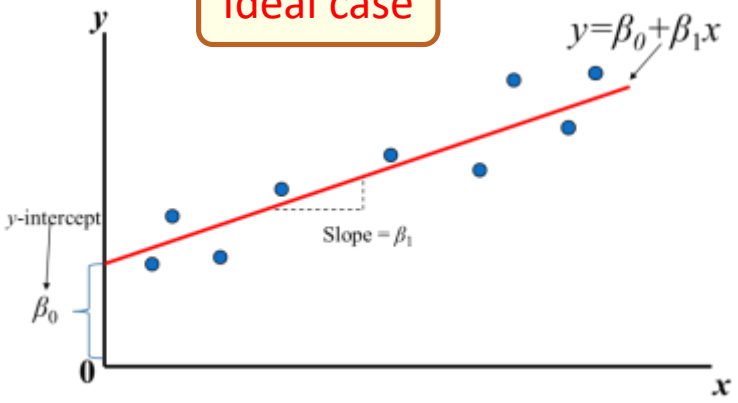
Basic Case of Linear Regression: Polynomial Fitting

- Structure determination
  - $y = \beta_0 + \beta_1 x$
  - $y = \beta_0 + \beta_1 x + \beta_2 x^2$
  - $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$
- Parameter identification
  - Least-squares method

Real cases

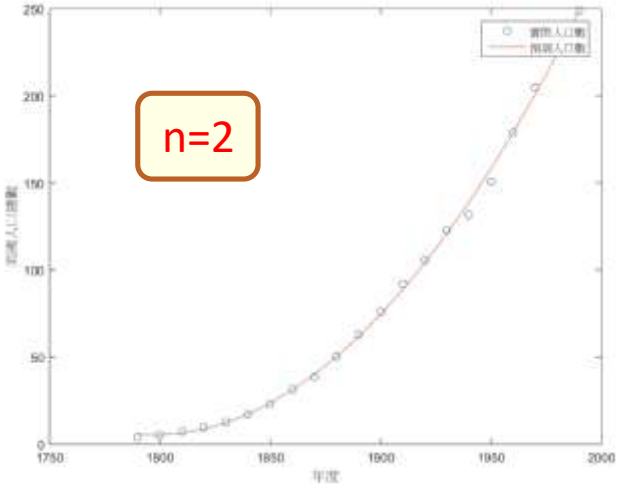


Ideal case

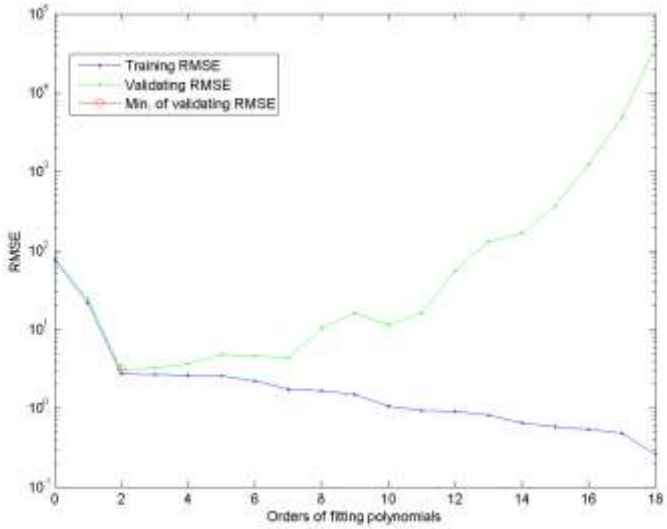


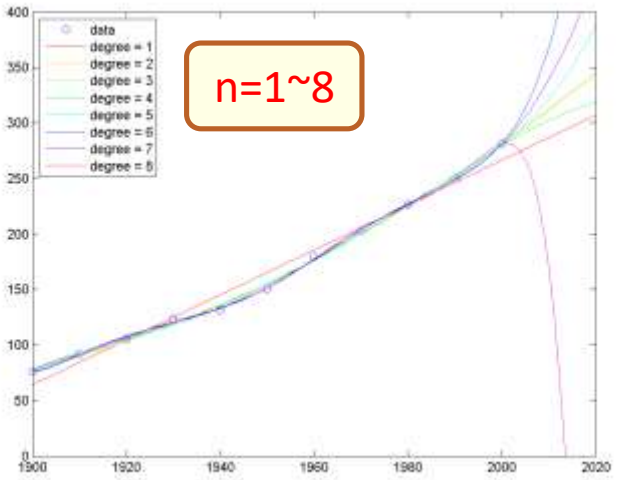
Example of Polynomial Fitting

- 美國人口總數預測
  - 模型
    - 多項式 $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n$
  - 模型複雜度：冪次 n
  - 模型效能評估
    - 交叉驗證 (cross validation)



Cross validation





More ML Application Examples

- Optical character recognition
- Document classification
- Part-of-speech tagging
- Speech processing
  - Recognition
  - synthesis
- Image recognition
  - Face recognition
- Info. retrieval
  - Recommendation
  - search
- Security
  - Fraud detection (credit card, telephone)
  - Network intrusion
  - Video surveillance
  - Speaker id.
- Self-driving cars

What People Do with ML

- For practitioners
  - Acquire application domain knowledge
  - Know properties of machine learning methods
  - Get familiar with ML tools
- For researchers
  - Design new learning algorithms
  - Speed up the learning process
  - Derive theoretical bound of accuracy
  - Take care of data
    - Big data
    - Imbalanced data
    - Missing data
    - …

List of Classifiers

- Commonly used classifiers
  - K-nearest-neighbor classifiers
  - Quadratic classifiers
    - Naïve Bayes classifiers
  - Linear classifiers
    - Single-layer perceptrons
    - SVM
  - Neural networks
    - Multilayer perceptrons
      - DNN
    - Radial-basis function networks
  - Classification and regression trees (CART)
    - Random forests
  - Many many more…