

Artificial Intelligence & Machine Learning and Pattern Recognition — — Introduction



Yanghui Rao

Assistant Prof., Ph.D

School of Mobile Information Engineering,

Sun Yat-sen University

raoyangh@mail.sysu.edu.cn

Instructors

- Course Instructor

- 饶洋辉

- Email: raoyangh@mail.sysu.edu.cn
 - Office: 409
 - Phone: 15989787592
 - Office Hours: Thursday, 14:30-16:30

- Teaching Assistant

- 刘学博, 朱和胜, 彭禹惟, 吴宏鹏, 叶双

Computer

- **Camera:** Simulate part functions of eyes
 - images and videos -> electrical signal
 - Computer Graphics (CG), Computer Vision (CV), etc.
- **Microphone:** Simulate part functions of ear
 - audios -> electrical signal
 - Speech Recognition (SR), etc.
- **Computer:** Simulate small part functions of brain
 - computing brain; symbol -> pattern
 - Artificial Intelligence (AI), etc.
- ...
- The connection between digital devices and brain
 - Human-Computer Interaction (HCI)

Human-Computer Interaction

- Substitute some functions of human, make *human* better



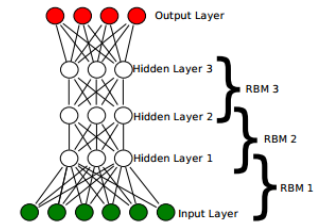
Artificial Intelligence

- Simulate more functions of brain, make *computer* better
 - chess brain, question answering brain, drive brain, etc.

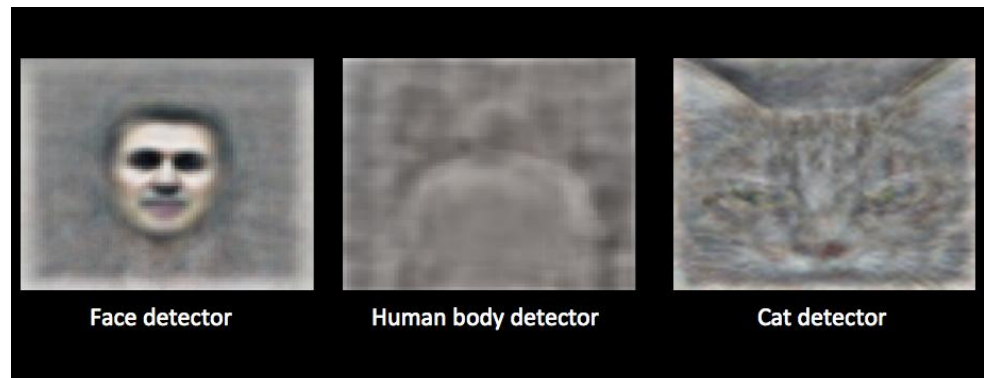


Artificial Intelligence

- Simulate more functions of brain, make *computer* better
 - chess brain, question answering brain, drive brain, etc.



- face detector
- human body detector
- cat detector
- ...



History of AI

- Turing Test, proposed by A.M. Turing in 1950
- Loebner Prize, pledged by Hugh Loebner in 1990
 - “Question Answering”/Chat robots
 - <http://www.loebner.net/Prizetf/loebner-prize.html>
 - 2013, Stephen Worswick
 - No chatbot fooled any of the 4 Judges
 - <http://www.mitsuku.com/>
 - ...
 - 1991, Joseph Weintraub
 - If you say "I need my Mommy", ELIZA will say "Tell me why you need your Mommy?"
 - It employs AI sentence parsing and knowledgebase technology, plus a 70,000 word vocabulary.

History of AI

"我是谁?"

"你得告诉我你的名字啊!"

"你是不是傻子啊?"

"你才是傻子呢"

"给我背首唐诗?"

"锄禾日当午"

.....



HOW ARE YOU 拟脑机器人

"How Are You" 是我的名字
你可以叫我 "小白"

3-6岁幼童专属的智能终端

吴义坚 CEO

毕业于中国科大少年班，博士，国际知名语音技术专家曾先后在科大讯飞、微软、盛大创新院等国内国际知名研发机构负责语音技术和产品的研发，十年以上的人工智能技术和产品研发经验，拥有数十项国内和国际专利，同时是一个80后的爸爸

History of AI

- Logic, 1950s
 - “Knowledge” is transferred by experts/humans
 - *propositional* logic (命题逻辑), first-order *predicate* logic (谓词逻辑), *fuzzy* logic...

Logic

- “true” or “false”
- negation
- conjunction
- disjunction
- implication
- equivalence
- universal quantifier
- existential quantifier
- ...

```
~~~~~欢迎您!~~~~~
      游戏：五子棋
~~~~~

人人对弈1  人机对弈2
请您选择1  或者2:  2
您选择了人机对弈~~~~
黑方1  US  白方2
请您选择1  或者2:  2
您选择了2,对方先下~~

  1 2 3 4 5 6 7 8 9 A B C D E F
1 + + + + + + + + + + + + + +
2 + + + + + + + + + + + + + +
3 + + + + + + + + + + + + + +
4 + + + + + + + + + + + + + +
5 + + + + + + + + + + + + + +
6 + + + + + + + + + + + + + +
7 + + + + + + + + + + + + + +
8 + + + + + + + + + + + + + +
9 + + + + + + + + + + + + + +
A + + + + + + + + + + + + + +
B + + + + + + + + + + + + + +
C + + + + + + + + + + + + + +
D + + + + + + + + + + + + + +
E + + + + + + + + + + + + + +
F + + + + + + + + + + + + + +
该计算机下了, 请输入1
```

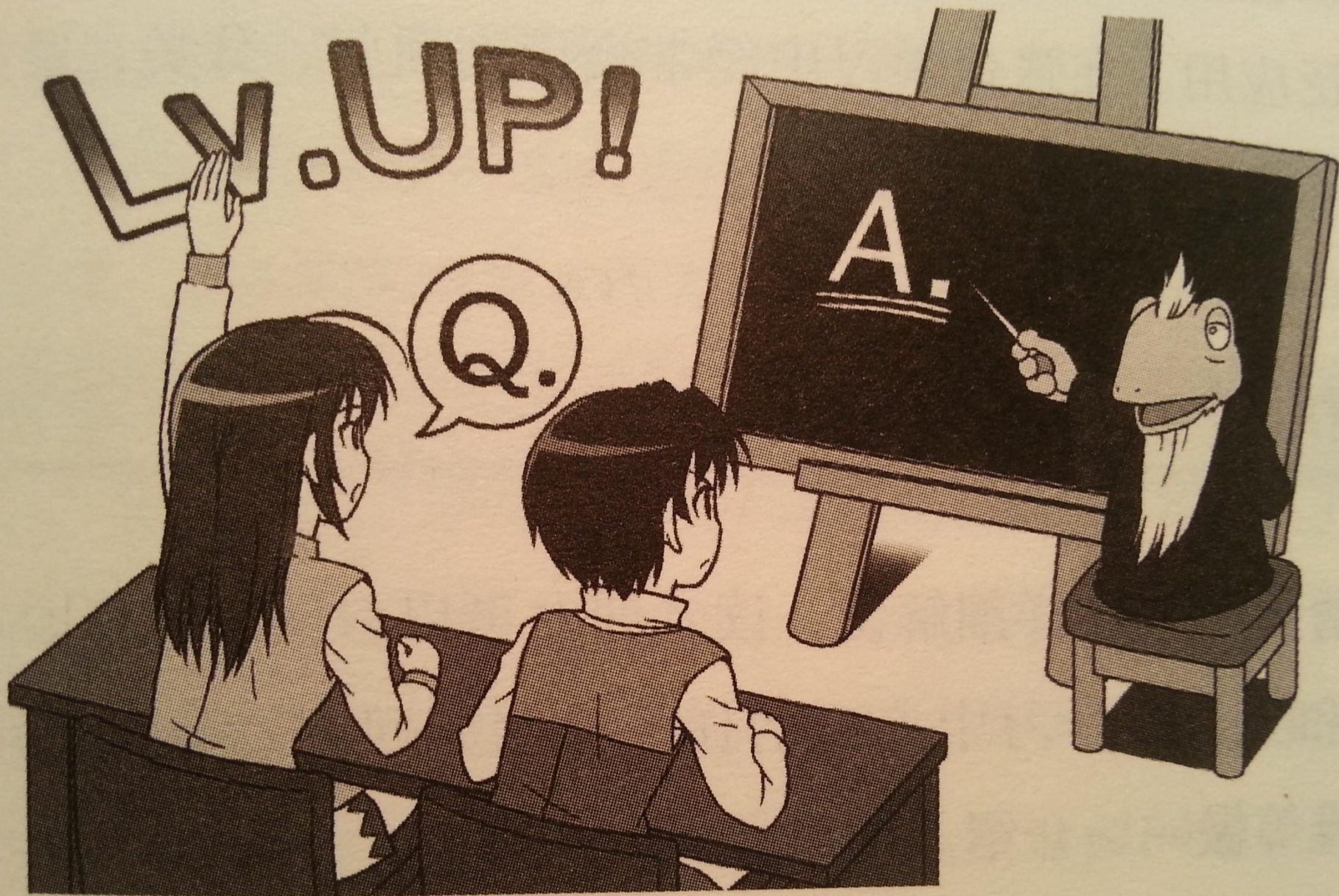
Logic

- “true” or “false”
- negation
- conjunction
- disjunction
- implication
- equivalence
- universal quantifier
- existential quantifier
- ...

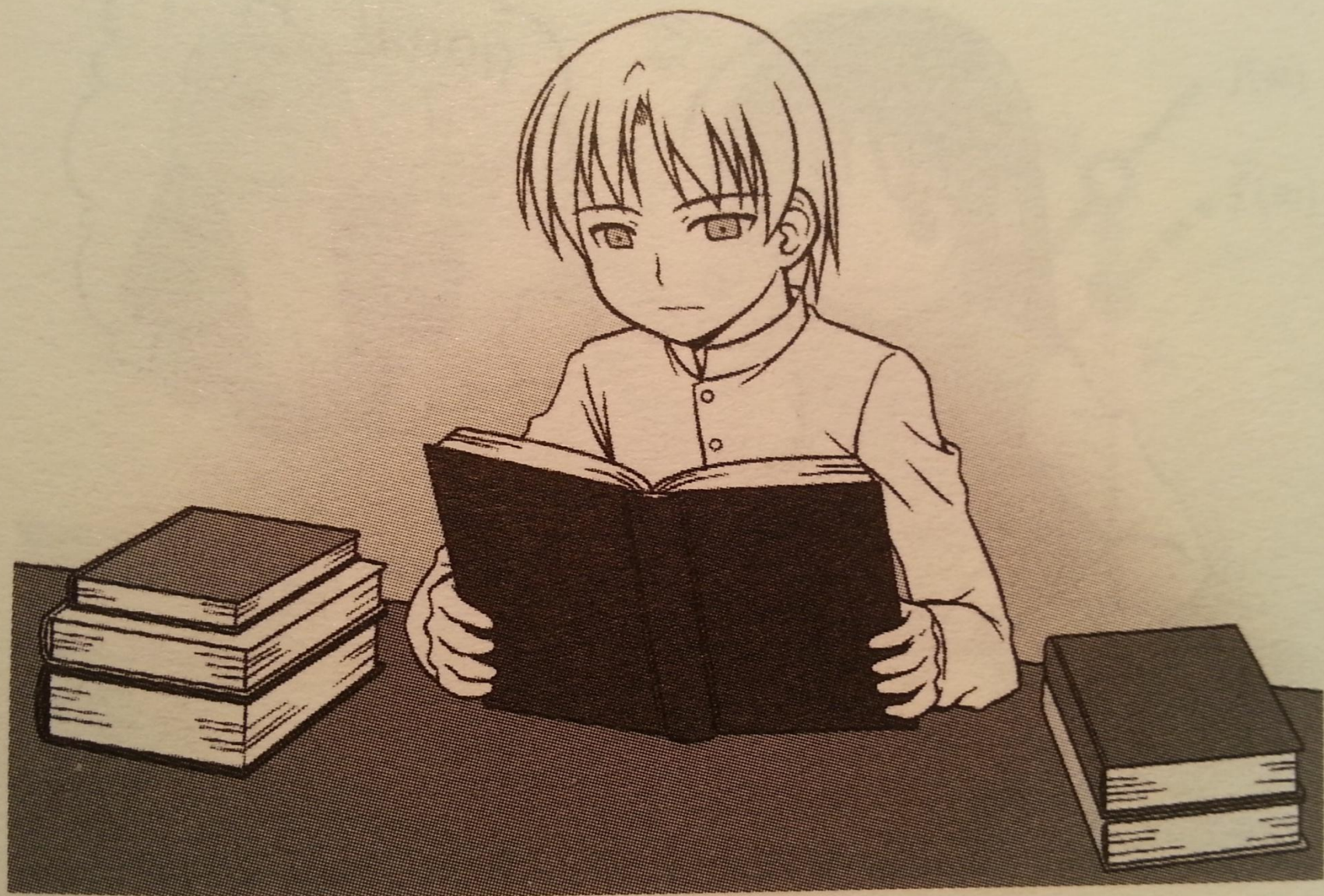
```
当前步数为: 40
  1 2 3 4 5 6 7 8 9 A B C D E F
1 + + + + + + + + + + + + + +
2 + + + + + + 0 0 0 + + + + + +
3 + + 0 0 0 0 0 0 0 + + + + + +
4 + + + 0 0 0 0 0 0 + + + + + +
5 + + + 0 0 0 0 0 0 + + + + + +
6 + + 0 0 0 0 0 0 0 + + + + + +
7 + 0 0 0 0 0 0 0 + + + + + +
8 + + + 0 + 0 0 + + + + + + +
9 + + + 0 + + + + + + + + + +
A + + + + + + + + + + + + + +
B + + + + + + + + + + + + + +
C + + + + + + + + + + + + + +
D + + + + + + + + + + + + + +
E + + + + + + + + + + + + + +
F + + + + + + + + + + + + + +
您的输入是: 6 3
祝贺您!您赢了~
```

History of AI

- Logic, 1950s
 - “Knowledge” is transferred by experts/humans
 - *propositional* logic (命题逻辑), first-order *predicate* logic (谓词逻辑), *fuzzy* logic...
- Machine Learning, 1990~
 - “Knowledge” is learnt by computers primarily
 - *supervised* learning (监督学习): classification...
 - *unsupervised* learning (无监督学习): clustering, density estimation...
 - *reinforcement* learning (强化/增强学习)



监督学习



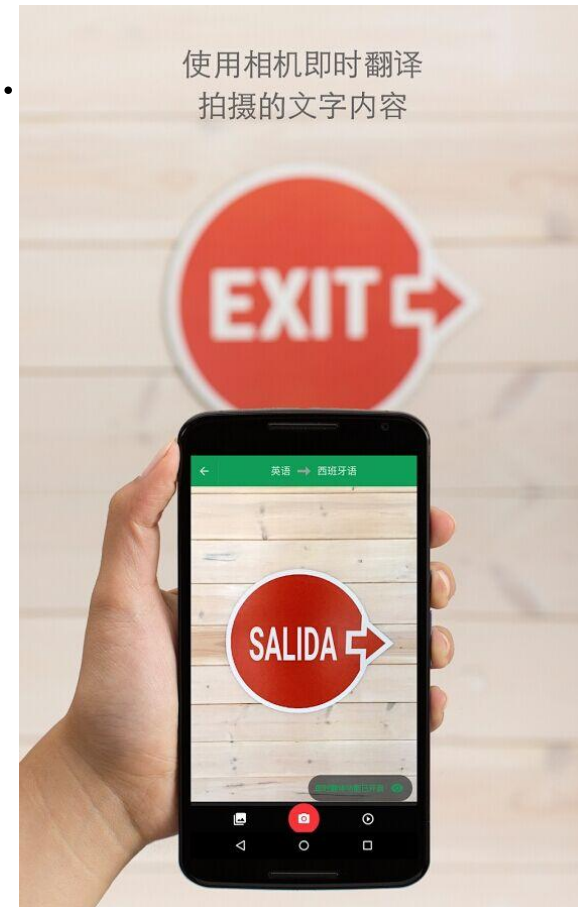
无监督学习



强化学习

What is Artificial Intelligence

- Compute something that shows **intelligent** behavior
 - Natural language processing, ...



What is Artificial Intelligence

- Compute something that shows **intelligent** behavior
 - Natural language processing, ...

2015年6月18日下午，日本，Pepper和阿里巴巴集团零售事业群总裁张建锋互动：

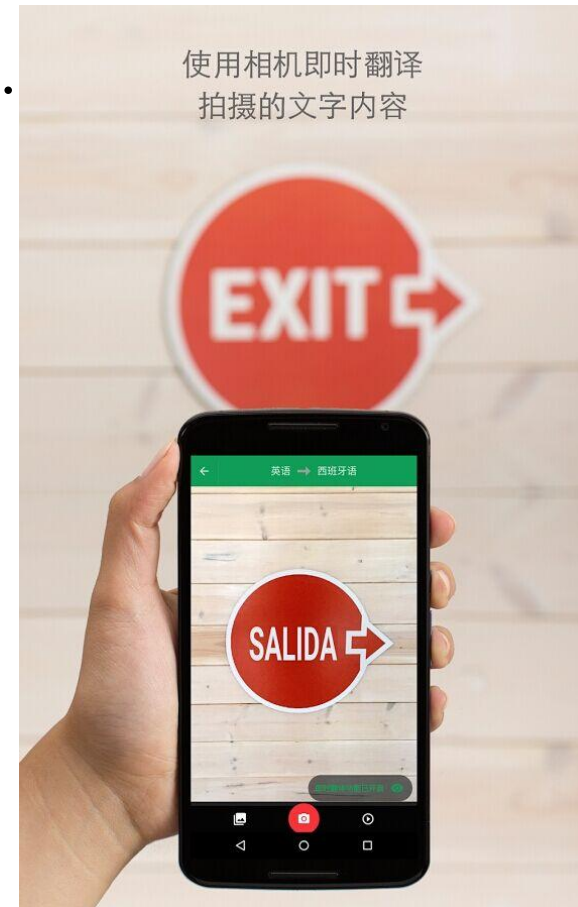
.....

Pepper: 快乐20，紧张60，生气0，悲伤0，
我的情感引擎显示，您好象有点紧张，是吗？

.....

Pepper: 快乐80，紧张20，生气0，悲伤0。
我的情感引擎显示你的心情似乎不错。你是
不是开始喜欢我了？

.....



What is Machine Learning

- Tom Mitchell (1997): A **computer program** (vs *human*) is said to **learn** from experience E (e.g., **labeled data**) with respect to some class of tasks T (e.g., **classification**) and performance measure P (e.g., **precision & recall**), if its performance at tasks in T , as measured by P , improves with experience E .

What is Pattern Recognition

- Christopher M. Bishop (2006): The automatic discovery of regularities in data through the use of computer algorithms (*e.g.*, Machine Learning) and with the use of these regularities to take actions such as classifying the data into different categories

What Kinds of Regularities?

- **Classification**

- A loan user: high or low risk? \leftrightarrow *banker*
- A person: health or sick? \leftrightarrow *doctor*
- An Iris flower: Setosa, Versicolour, or Virginica? \leftrightarrow *botanist* (植物学家)



Classification

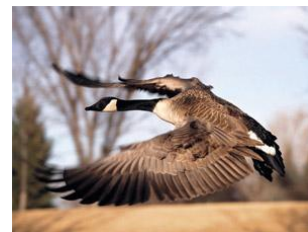
- Make the computer as intelligent as an expert
 - Classify a person as healthy or sick
 - Identify the author of a piece of art or a book
 - Identify the variety of an animal or a plant

Classification

- Make the computer as intelligent as an expert
 - Classify a person as healthy or sick
 - Identify the author of a piece of art or a book
 - Identify the variety of an animal or a plant
- Training data: examples of the input vectors along with their corresponding target vectors
 - Input vectors: height, weight, has a tail? (yes or no)
 - Target vectors: human or monkey

Classification

- Make the computer as intelligent as an expert
 - Classify a person as healthy or sick
 - Identify the author of a piece of art or a book
 - Identify the variety of an animal or a plant
- Training data: examples of the input vectors along with their corresponding target vectors
 - Input vectors: height, weight, has a tail? (yes or no)
 - Target vectors: human or monkey
 - Which attributes are good to classify a bird from others?



Classification

- Input vectors of birds:
 - can move? (yes or no)
 - can chirp? (yes or no)
 - have feather? (yes or no)
 - size (length, width, height)
 - reaction to sound? (yes or no)
 - ...

Classification

- Input vectors of birds:
 - can move? (yes or no)
 - can chirp? (yes or no)
 - have feather? (yes or no)
 - size (length, width, height)
 - reaction to sound? (yes or no)
 - ...



Classification

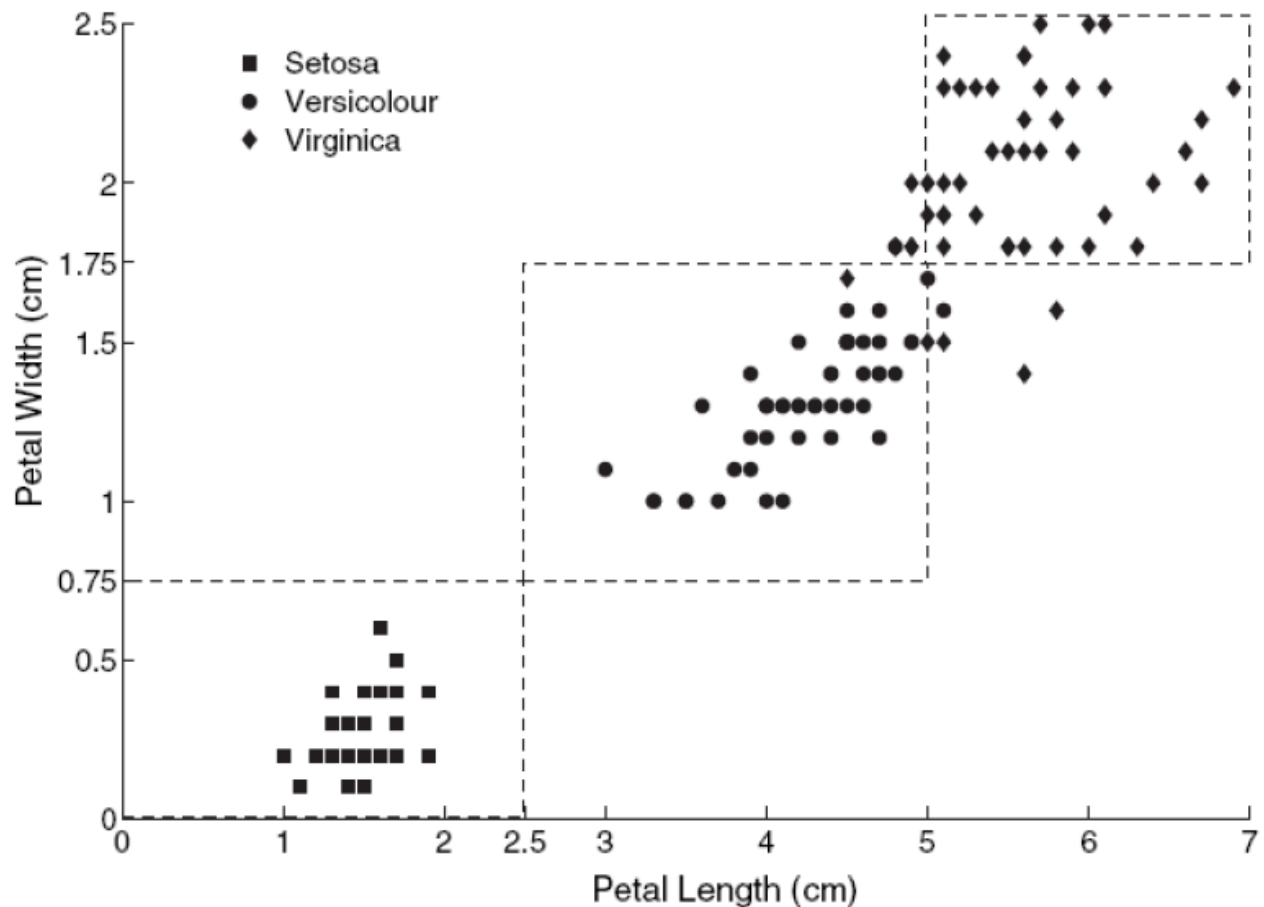
- Data set for Machine Learning
 - UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/datasets.html>)
- Iris data set
 - <http://archive.ics.uci.edu/ml/datasets/Iris>
 - Attributes provided by Fisher (experts)
 - sepal (萼) length (cm), sepal width (cm), petal (瓣) length (cm), petal width (cm), class

5.3,3.7,1.5,0.2,	Iris-setosa
5.0,3.3,1.4,0.2,	Iris-setosa
7.0,3.2,4.7,1.4,	Iris-versicolor
6.4,3.2,4.5,1.5,	Iris-versicolor
6.3,3.3,6.0,2.5,	Iris-virginica
5.8,2.7,5.1,1.9,	Iris-virginica

Classifier: Decision tree, SVM, Neural Network, ...

Classification

- Plot using petal length and width



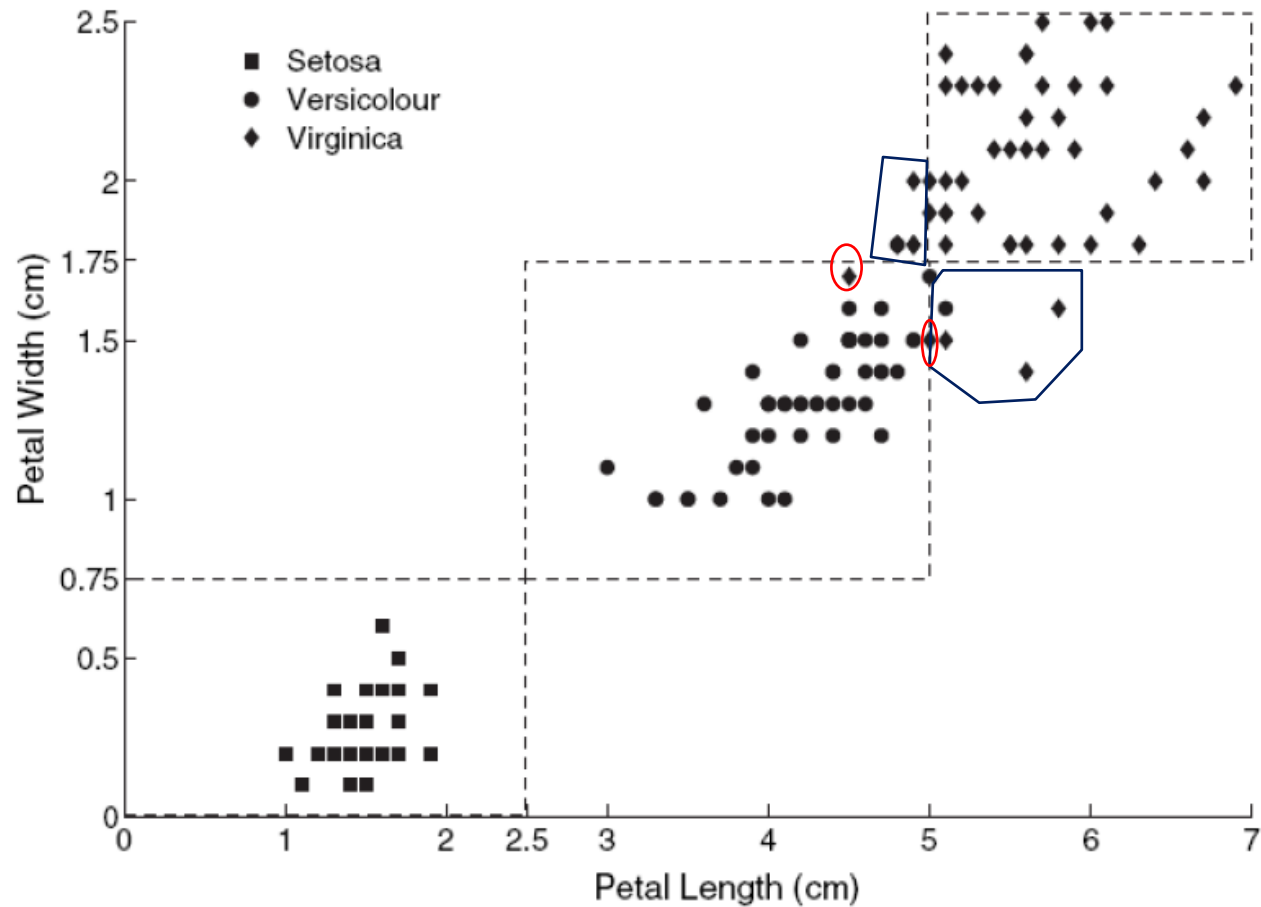
Classification

- Based on these categories with data, the following rules can be derived:
 - Petal width low and petal length low implies Setosa (山鸢尾).
 - Petal width medium and petal length medium implies Versicolour (变色鸢尾).
 - Petal width high and petal length high implies Virginica (维珍尼亚鸢尾).

Classification

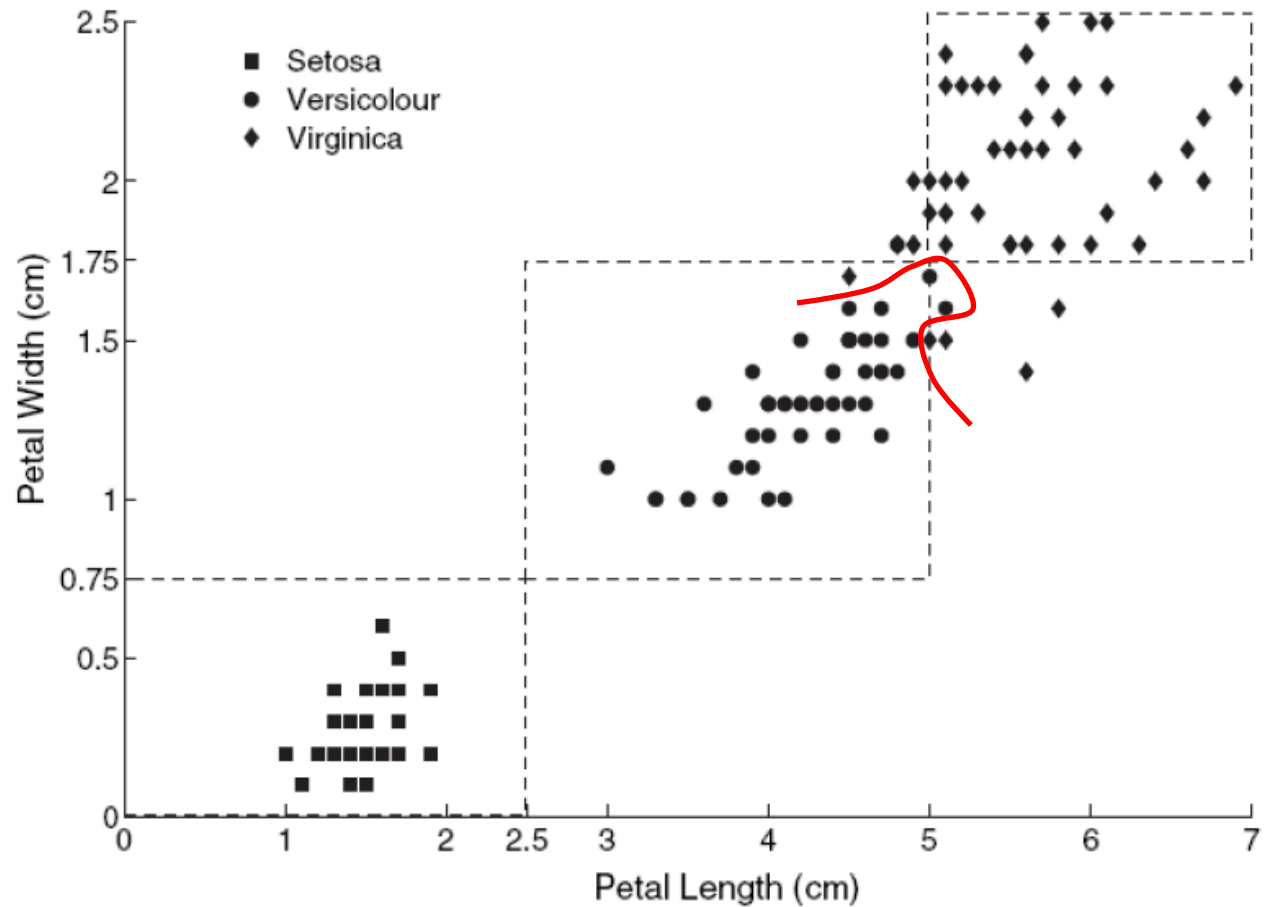
- These rules do not classify all the flowers correctly.
 - Flowers from the Setosa species are well separated from the other two species with respect to petal width and length.
 - However, the Versicolour and Virginica species overlap somewhat with respect to these two attributes.
- Solutions?

Classification



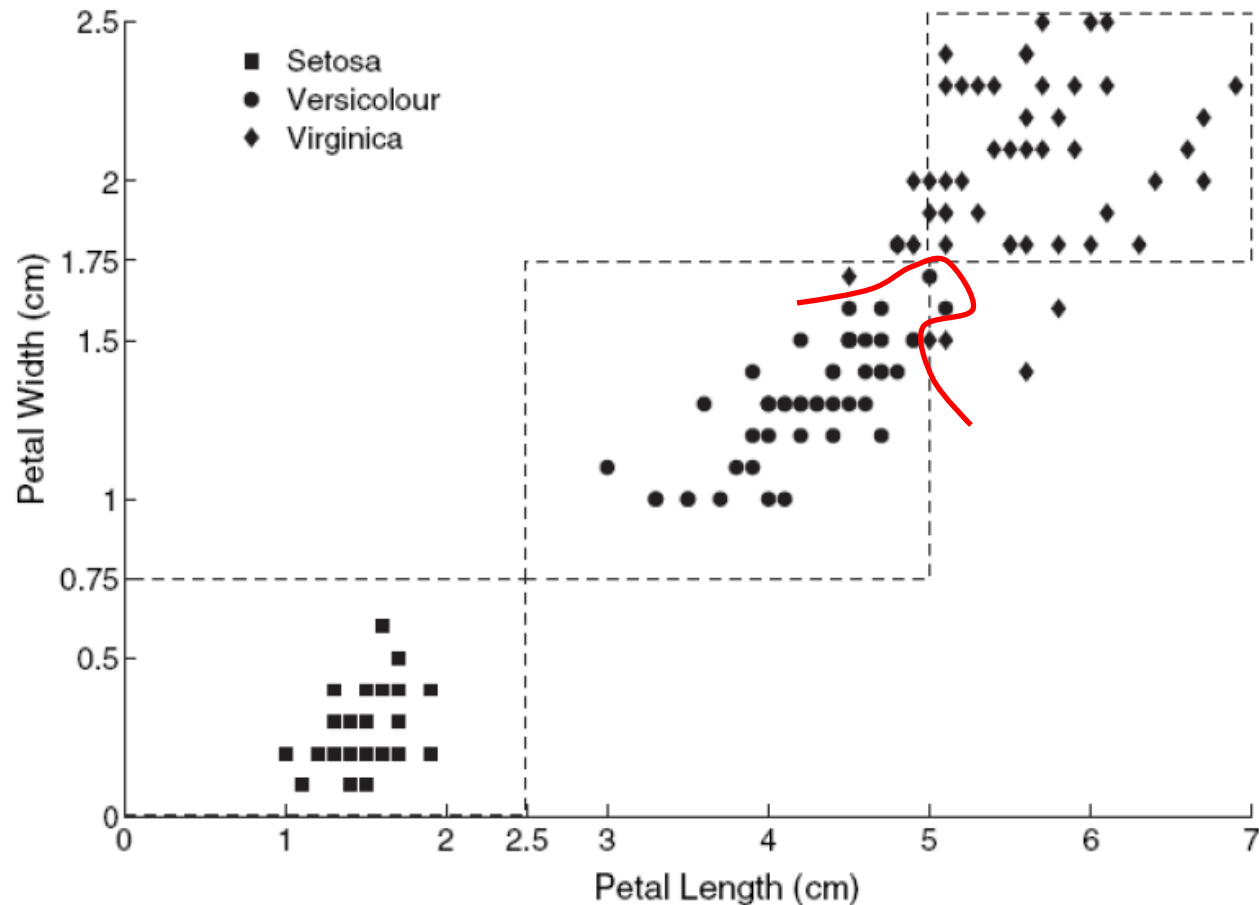
1. Records surrounded by red line: misclassified
2. Records surrounded by blue line: unclassified

Classification



1. A new better classifier model

Classification



1. A new better classifier model
2. Add more attributes
(i.e., sepal length and width, and others if possible)

What Kinds of Regularities?

- **Regression**

- Highly related to statistics
- Many concerned quantitative variables:
 - A person's life expectancy
 - A journal article's citations
 - A city's air temperature next day
 - ...
- Mining the correlation between a series of independent variables (X) and dependent variables (Y)

Regression

- Which one of the following is false for building a regression model?

(X: independent variables; Y: dependent variables)

- 1. X is the working years of an employee, Y is the employee's salary
- 2. X is a child's height, Y is the height of the child's father or mother
- 3. X is the total production of a product, Y is the total consumption of the product
- 4. X is the total consumption of a product, Y is the total production of the product

Regression

- Which one of the following is false for building a regression model?

(X: independent variables; Y: dependent variables)

- 1. X is the working years of an employee, Y is the employee's salary
- 2. X is a child's height, Y is the height of the child's father or mother
- 3. X is the total production of a product, Y is the total consumption of the product (**iPhone**)
- 4. X is the total consumption of a product, Y is the total production of the product

What Kinds of Regularities?

- Association Rules
 - Detect sets of attributes or items that frequently co-occur in many records
 - On Thursday, during 4-11pm, customers often purchase diapers and beers together



Association Rules

- Where does the data come from?
 - supermarket transactions, customer complaint calls, discount coupons, etc.
- Cross-marketing analysis
 - purchase recommendation, cross selling
 - what are the subsequent purchases after buying a given product?
- Target-marketing
 - what types of users buy what products
- Catalog design

What Kinds of Regularities?

- **Clustering**

- seeks to find groups of closely related records (*e.g.*, users, articles, genes, planets)
- can be applied to
 - compress data
 - anomaly detection
 - reduce dimensions
 - group sets of related customers/products



Clustering

- Key steps
 - Feature extraction for each record
 - A user of Amazon
 - An article, *etc.*
 - Weighting for each feature
 - Linear
 - Nonlinear, *etc.*
 - Similarity measurement for *paired* records
 - Euclidean distance
 - Cosine similarity, *etc.*

Clustering

- How to decide these three articles as two groups/clusters?
 - 1: *I'm involved in the release of apple's iPhone 6.*
 - 2: *Apple's iPhone 6 released on Friday.*
 - 3: *I've eaten an apple this Friday.*

The screenshot shows the Clusty search engine interface. At the top, there's a navigation bar with links for 'web', 'news', 'blogs', 'wikipedia', 'jobs', and 'more'. The search bar contains 'iPhone 6' and a 'Search' button. To the right of the search bar are links for 'advanced preferences'. Below the navigation bar, there's a horizontal menu with categories: 'Top News', 'World', 'U.S.', 'Business', 'Sports', 'Health', 'Tech', and 'Science'. The main content area displays 'Top 307 results retrieved for the query iPhone 6 (details)'. On the left side, there's a sidebar with a 'remix' button and a list of 'All Results' (307) categorized by sources: 'Pre-orders' (28), 'Apple Pay' (21), 'Samsung' (17), 'China' (20), 'Stocks' (17), 'iPhone 6 launch' (15), 'IOS 8' (13), and 'Buy The Iphone 6' (10). The main content area shows two article snippets. The first article is titled 'Apple's latest iPhones: What's in the box?' and is dated '1 hour ago'. It mentions that the new iPhones are bigger, slimmer, and have faster processors, and asks what's in the new iPhone 6 and iPhone 6 Plus box. The second article is titled 'Big Day for Apple's Plus-Sized Phones' and is dated '2 hours ago'. It mentions that September 19 has been circled on the calendar of Apple enthusiasts, who woke up in the middle of the night last week to pre-order the new iPhone 6 and iPhone 6 Plus. Both articles include a small icon for a remix or related content.

Clusty

web news blogs wikipedia jobs more »

iPhone 6 Search advanced preferences

clouds sources sections time remix

All Results (307)

- Pre-orders (28)
- Apple Pay (21)
- Samsung (17)
- China (20)
- Stocks (17)
- iPhone 6 launch (15)
- IOS 8 (13)
- Buy The Iphone 6 (10)

Top 307 results retrieved for the query iPhone 6 (details)

Search Results

Apple's latest iPhones: What's in the box?

1 hour ago - They are bigger, slimmer and have faster processors than the iPhones that came before them, but what's in Apple's new iPhone 6 and iPhone 6 Plus box?
www.cnbc.com/id/102009946 - [cache] - CNBC

Big Day for Apple's Plus-Sized Phones

2 hours ago - September 19 has been circled on the calendar of Apple enthusiasts who woke up in the middle of the night last week to pre-order the new iPhone 6 and iPhone 6 Plus. Others have spent days camped outside Apple stores for a chance to be one of the first customers to purchase a new device when the...
abcnews.go.com/Technology/buying-apple-iphone-iphone-today/story?id=25601151 - [cache] - ABC News