

# 模式识别基础

汪小我

清华大学自动化系

2019课程通知群



该二维码7天内(3月4日前)有效，重新进入将更新

# 课程简介

- 考评方式：平时作业（60%） + 期末考试（40%）
- 任课教师：汪小我  
自动化系信息处理研究所，长聘副教授，博士生导师  
研究方向：生物信息学
- 课程助教：
  - 张威：w-zhang16@mails.tsinghua.edu.cn
  - 王昊晨：wanghc17@mails.tsinghua.edu.cn

# 知识技能储备

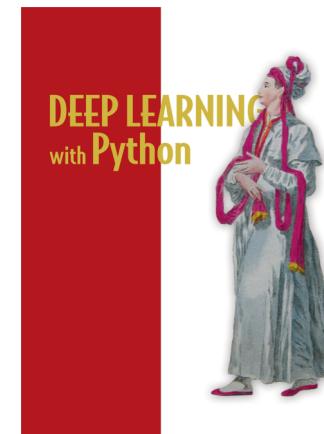
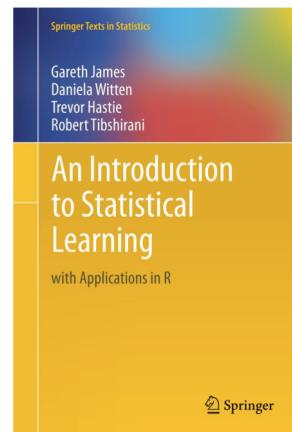
- 微积分、线性代数、概率论与数理统计
- 编程能力：**Python**、Matlab、R
- Python入门辅导课（自选参加，共两次）

# 教材与参考书

教材：《模式识别基础》，张学工，  
清华大学出版社，2010

- 参考书：

- 周志华，机器学习，清华大学出版社，2016
- Gareth James, An Introduction to Statistical Learning with Applications in R, Springer, 2013
- Francois Chollet, Deep Learning with Python, Manning, 2018



国家精品课程教材  
新编《信息、控制与系统》系列教材  
**模式识别（第三版）**  
**Pattern Recognition (Third Edition)**  
张学工 编著  
Zhang Xuegong

清华大学出版社

# 什么是模式识别？



“看” 到东西 → 认出东西、产生想法

观察 → 判断

量化观测 → 分类决策

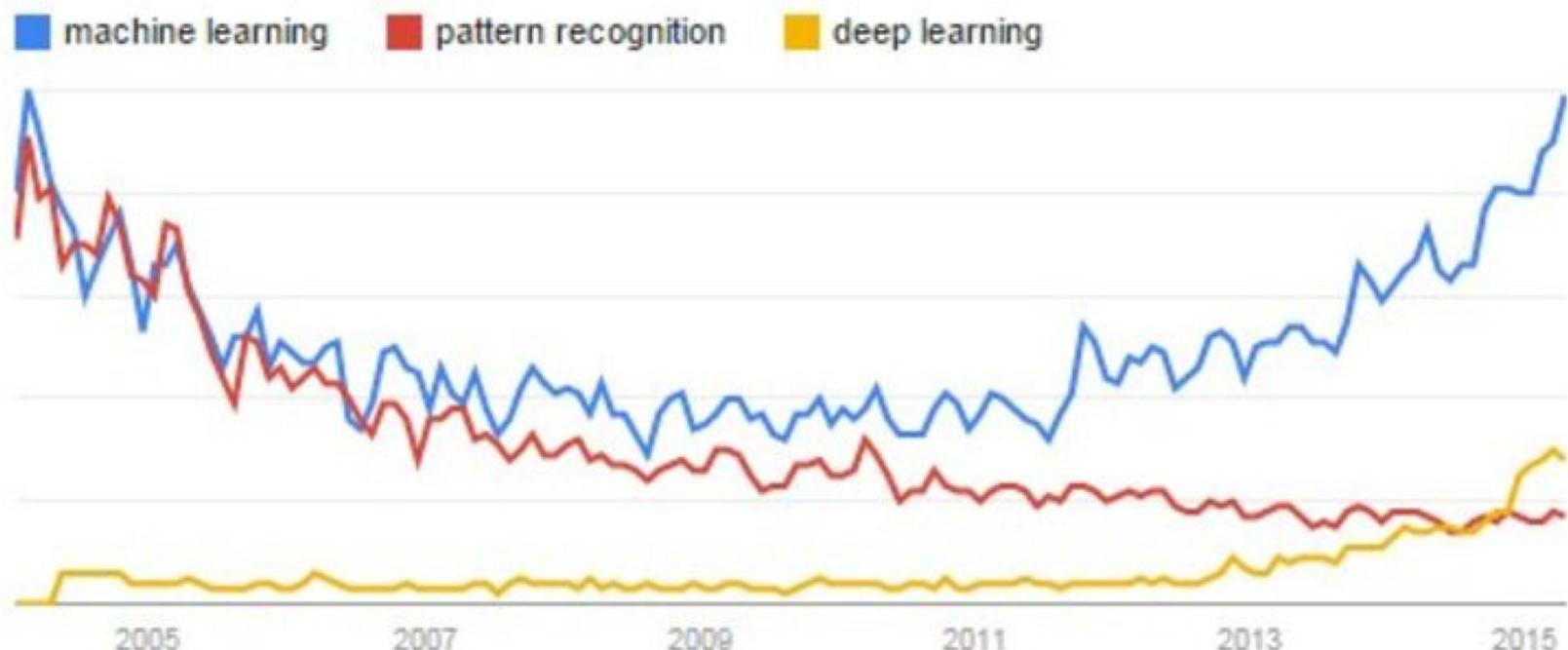
$x \in \mathbb{R}^d \rightarrow y$

模式识别

# 模式识别

- People *recognize* things, from observations.  
----- 识别
- People recognize things by *recognizing patterns*, rather than individual observations.  
----- 模式识别
- *Pattern Recognition with Machines*
  - People like to make machines that can do what we can, or what we can not.
    - Because we are curious
    - Because we are lazy
    - Because we are not so able

# Deep Learning vs Machine Learning vs Pattern Recognition



<http://www.computervisionblog.com/2015/03/deep-learning-vs-machine-learning-vs.html>

# 作业

- 请简述你如何理解下面这几个说法
  - 人工智能 (Artificial Intelligence)
  - 模式识别 (Pattern recognition)
  - 机器学习 (Machine Learning)
  - 深度学习 (Deep Learning)
  - 统计学习 (Statistical Learning)

# 课程主要内容

- 第一章 绪论，模式识别系统举例
- 第二章 线性方法—线性回归与线性分类
- 第三章 贝叶斯决策
- 第四章 概率密度函数的估计
- 第五章 非线性方法
  - 1. 神经网络
  - 2. 支持向量机
  - 3. 决策树
  - 4. 集成学习
  - 5. 近邻法
  - 6. 非线性回归
- 第六章 模型的评价与选择
- 第七章 特征选择
- 第八章 特征提取
- 第九章 聚类分析

模式 与 模式识别

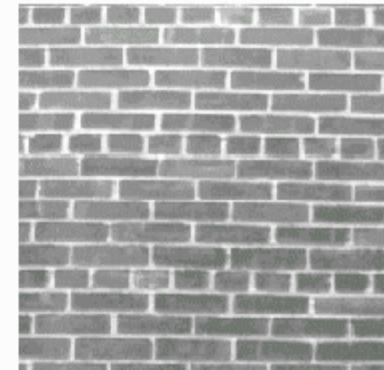
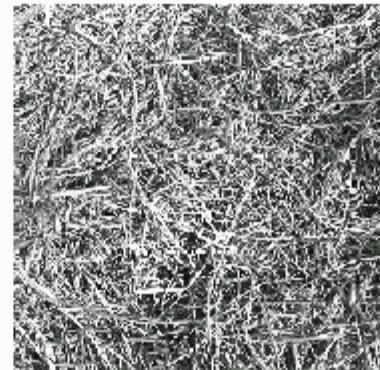
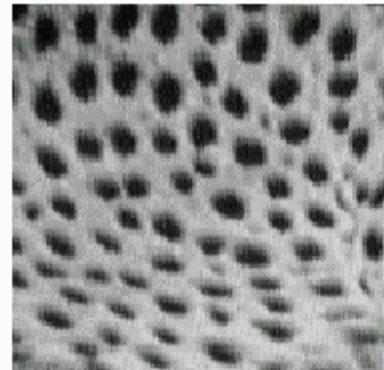
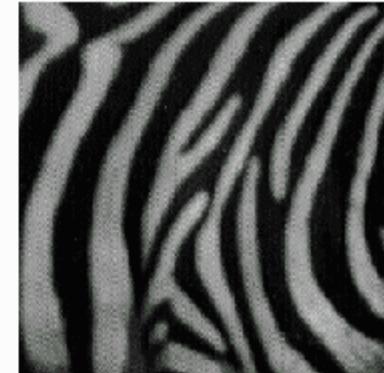
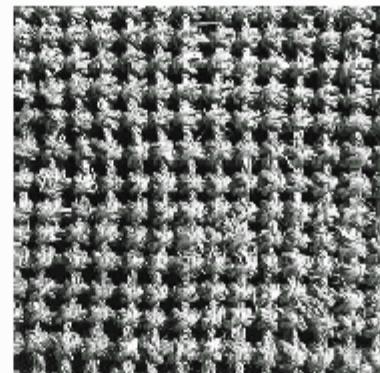
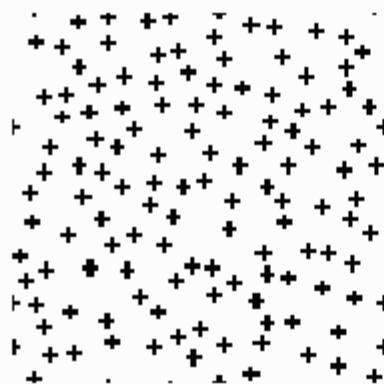
*Pattern and Pattern Recognition*

# 何谓“模式（Pattern）”？

- 拉丁词根 pat (父亲) , 父是子的“模型”
- 图案，花样；方式；样品；型，式样，纸样，模型；模范，典型；
- 对象的组成成分或影响因素之间所存在的直接或间接的规律性的**关系**
- 存在确定性或随机规律的对象、过程或事件的**集合**

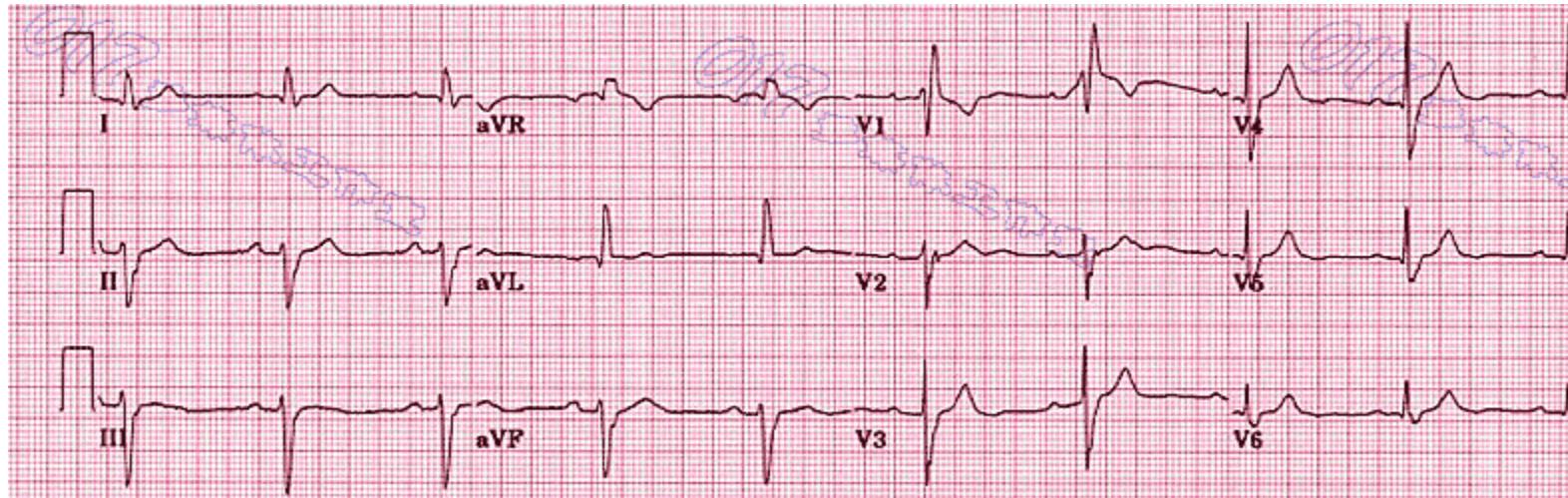
# 常见模式举例

Texture Patterns:

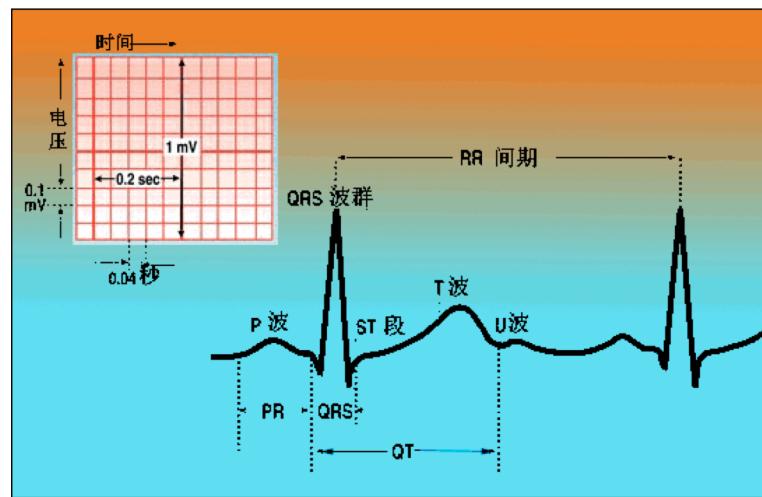


Textures are the richest pattern created in nature, perceptually each class of texture has some common features—regularities, and it also contains non-deterministic characteristics.

# 常见模式举例（续）



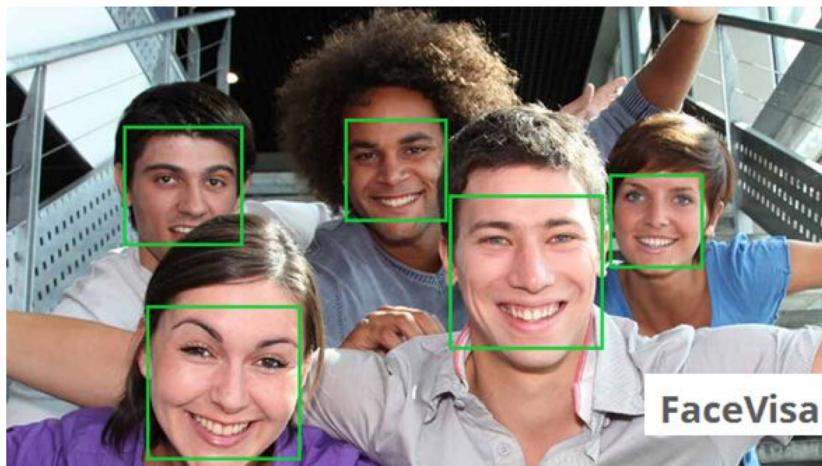
心电图



# 常见模式举例（续）

- 社会模式：
  - 信用：收入、消费习惯、贷款、…
  - 保险：驾龄、出险次数、车型、驾驶习惯、…
  - 信息服务：爱好、浏览习惯、文化程度、…
  - 择偶：背景、爱好、性格、经济状况、…
  - 文体：…
  - 性格：…
  - 文化：…
  - 事件：…
  - 政治：…

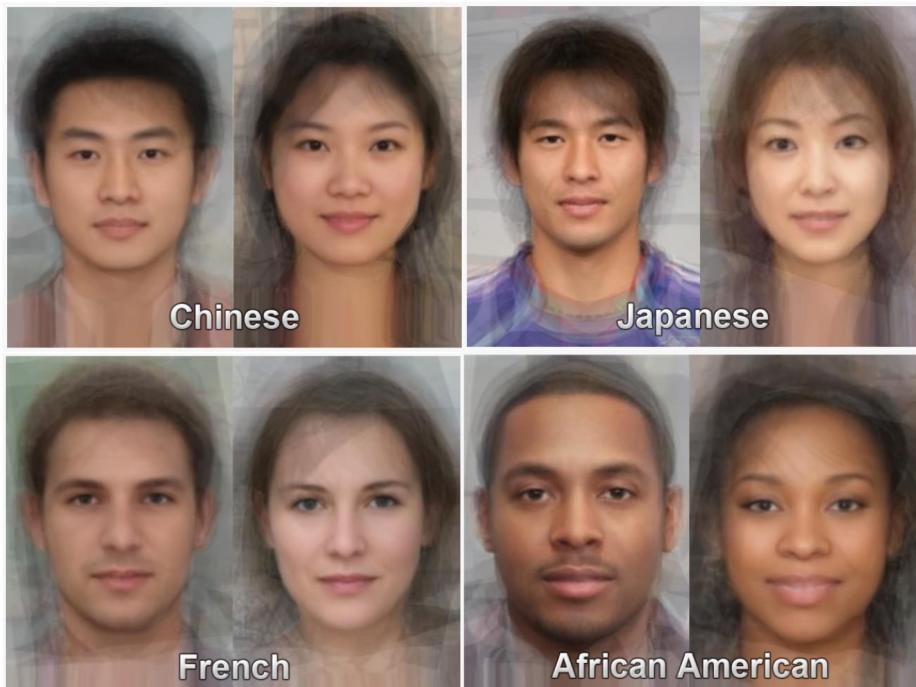
# 常见模式举例（续）



## 人脸的模式

- 共性：人脸作为一类对象区别于其他
- 个性：每个人作为一类区别于其他人

Average faces



# 什么是“识别 (recognition) ”？

- 区分、分辨、辨认、鉴别
- 《现代英汉词典》
  - The act or process of identifying (or associating) an input with one of a set of known possible alternatives
- 《美国传统辞典》
  - An awareness that something perceived has been perceived before.

# 何为“模式识别”？

- Pattern Recognition
  - the recognition of patterns
- To see something 1 as something 2
- 通过对事物的观察对其某种性质的认识

# 基本名词

- 举例：流感病人诊断
  1. 特征采集：体温、血常规.....
  2. 已知样本的训练学习：普通感冒和流感样本的分类器
  3. 新样本的分类：流感或者普通感冒
- 对象：样本 Sample
- 模式：类 Class
- 观察：特征属性 Features
- 模式识别：把样本根据其特征归类

# 概念和名词约定

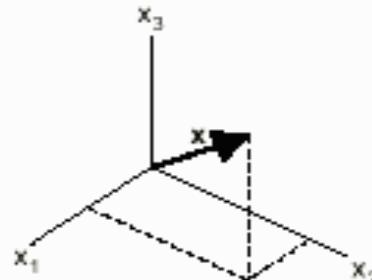
- 样本sample: 待研究对象个体，包括性质已知或未知的个体
- 标记label: 样本的性质
  - 离散值 (类别class): 分类问题 (classification)
  - 连续值: 回归问题 (regression)
- 已知样本known samples: label已知的样本
- 未知样本unknown samples: label未知的样本
- 样本集sample set: 若干样本的集合

# 概念和名词约定 (续)

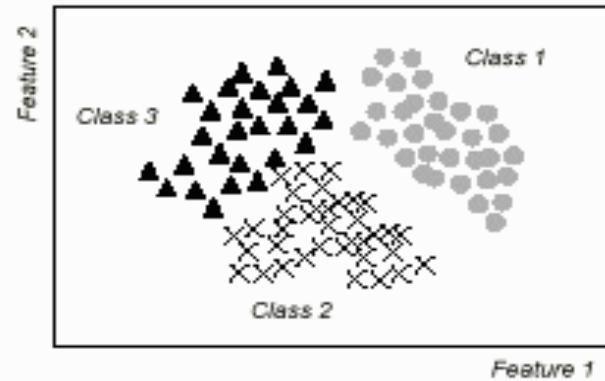
- 特征features：样本的任何可区分的（且可观测的）方面
  - 包括定量特征和定性特征，但通常最后转化为定量特征
- 特征向量feature vectors：样本的所有特征组成的 d 维向量是样本在数学上的表达，因此也称作样本
- 特征空间feature space：特征向量所在的 p 维空间，每一个样本（特征向量）是该空间中的一个点，一个类别是该空间中的一个区域

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_d \end{bmatrix}$$

Feature vector



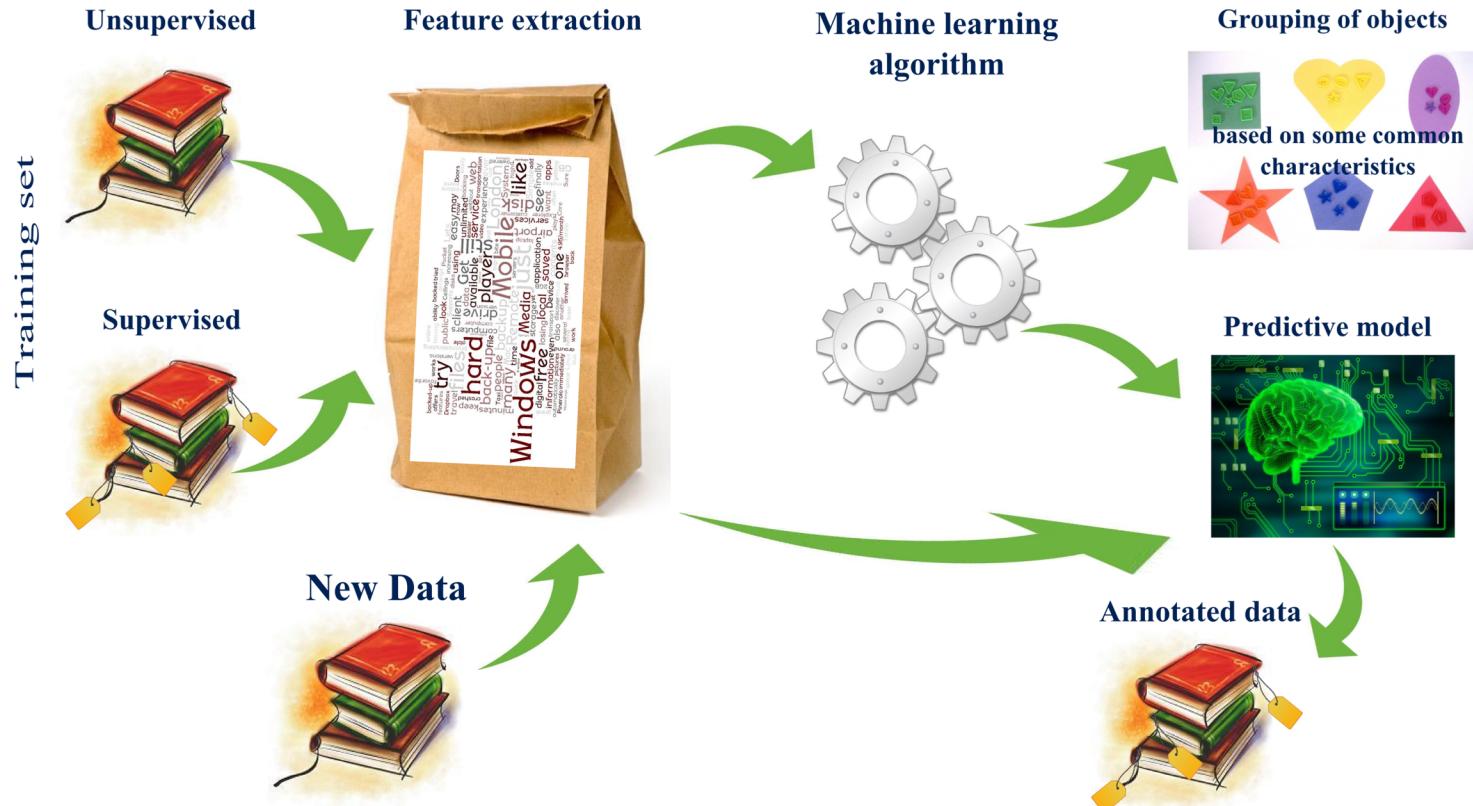
Feature space (3D)



Scatter plot (2D)

# 模式识别、机器学习的基本流程

样本  
观测



<http://nkonst.com/machine-learning-explained-simple-words/>

# 举例：手写体识别

## MINIST 数字手写体数据集

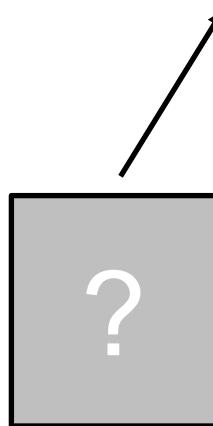
Modified National Institute of Standards and Technology database

- Input=28\*28 images
- Label={0,1,2,3,4,5,6,7,8,9}

```
define_a_class  
    magic code here  
return class-label
```



# 如何分类？



一种简单策略：

每个类找一个模板 $T^k$ ，待分类样本与其进行比较

$$L^k = \|X - T^k\|_2 = \sum_{i=0}^n \sum_{j=1}^m (x_{i,j} - t_{i,j})^2$$

Class =  $\operatorname{argmin}_k(L^k)$

损失函数  
Loss function

这种策略有何不足？

# 一个简单的例子



# 硬币的分类：最简单的情况

- 两类：比如1角 ( $\omega_1$ ) 和1分 ( $\omega_2$ )
- 没有任何观测的情况，如何决策？
- 分类：依据概率 —— 先验概率(*a prior probabilities*)
  - 如果  $P(\omega_1) > P(\omega_2)$ , 则  $x \in \omega_1$
  - 如果  $P(\omega_1) < P(\omega_2)$ , 则  $x \in \omega_2$
- 分类性能：
  - 错误率(error rate): 分类错误的概率  
 $= \min\{P(\omega_1), P(\omega_2)\}$

# 硬币的分类：略微复杂一些

- 两类：1角 ( $\omega_1$ ) 和5角 ( $\omega_2$ )
- 有一种观测特征  $x$  (比如重量)
- 如何决策?
  - 后验概率 (*a posterior probabilities*)
    - 如果  $P(\omega_1|x) > P(\omega_2|x)$ , 则  $x \in \omega_1$
    - 如果  $P(\omega_1|x) < P(\omega_2|x)$ , 则  $x \in \omega_2$

# 硬币的分类：略微复杂一些（续）

- 如何计算后验概率？

类条件概率密度：

- 查询国家标准
- 收集一定数量的各种硬币，统计其类条件概率密度

- 统计很多情况，估算先验概率
- 查询发行量
- 根据猜测（比如0.5）

$$P(\omega_i | x) = \frac{p(x, \omega_i)}{p(x)} = \frac{p(x | \omega_i)P(\omega_i)}{p(x)}$$

——贝叶斯公式

所有硬币的  $x$  分布密度函数  
(假定只有1角/5角硬币)

对不同类别，此部分相同，不影响  
比较大小，因此不必要计算

# 硬币的分类：更复杂一些

- 两类：1角 ( $\omega_1$ ) 和5角 ( $\omega_2$ )
- 有一种或几种观测特征  $x$  (比如重量或重量和直径)
- 进一步考虑：

## 风险 Risk (of the decision)

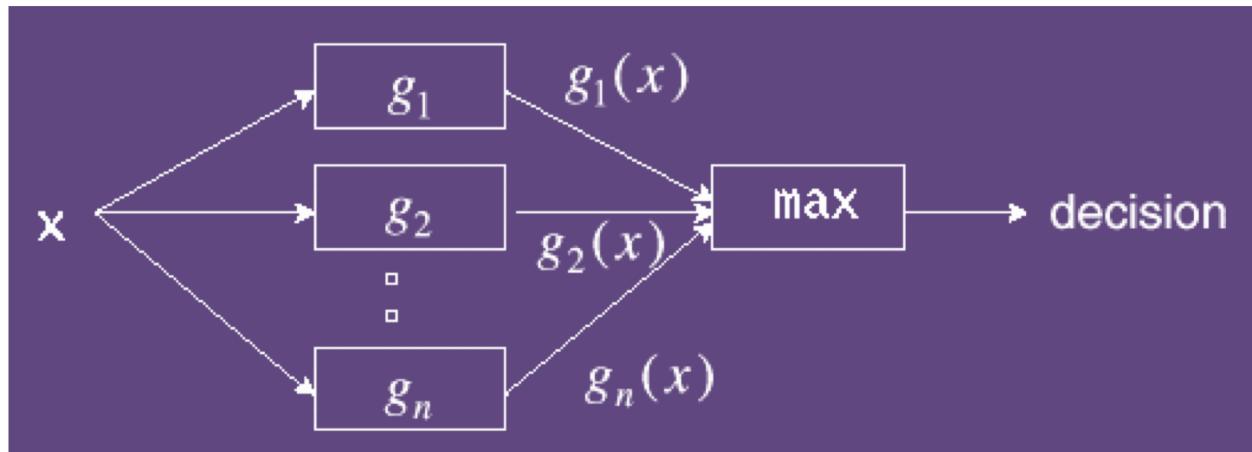
- “把1角识别成5角” 和 “把5角识别成1角”的风险是不同的！
- 而且，识别本身是有成本和前提的！

→ 针对各种情况的统计模式识别理论和方法

# 从这个两个例子我们看到了什么？

- 特征
- 分类器（判别函数） $g_1(x)$ 、 $g_2(x)$
- 错误率、风险

- 多类



# 问题

- 利用什么特征、多少特征？
- 如何得到分类器？
  - 基于知识
  - 基于数据
    - 利用样本估计先验概率、概率密度函数
    - 直接利用设计某种分类器

# 模式识别的典型任务

# 回归与分类

Regression v.s. Classification

# 监督学习 (supervised learning)

- 给出若干已知答案的样本（训练样本 training samples）
- 由机器从这些样本中进行学习（训练 training/learning）
- 学习的目的在于从这些样本中总结规律，使之能够对新的样本进行判断

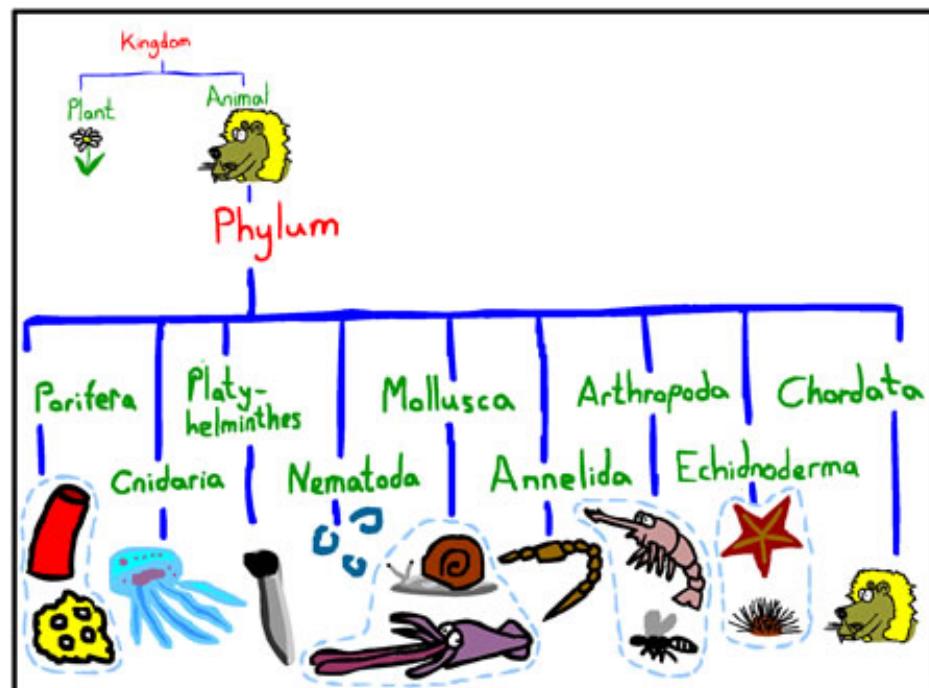
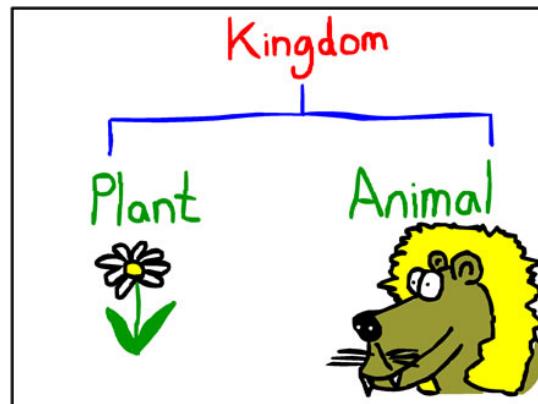
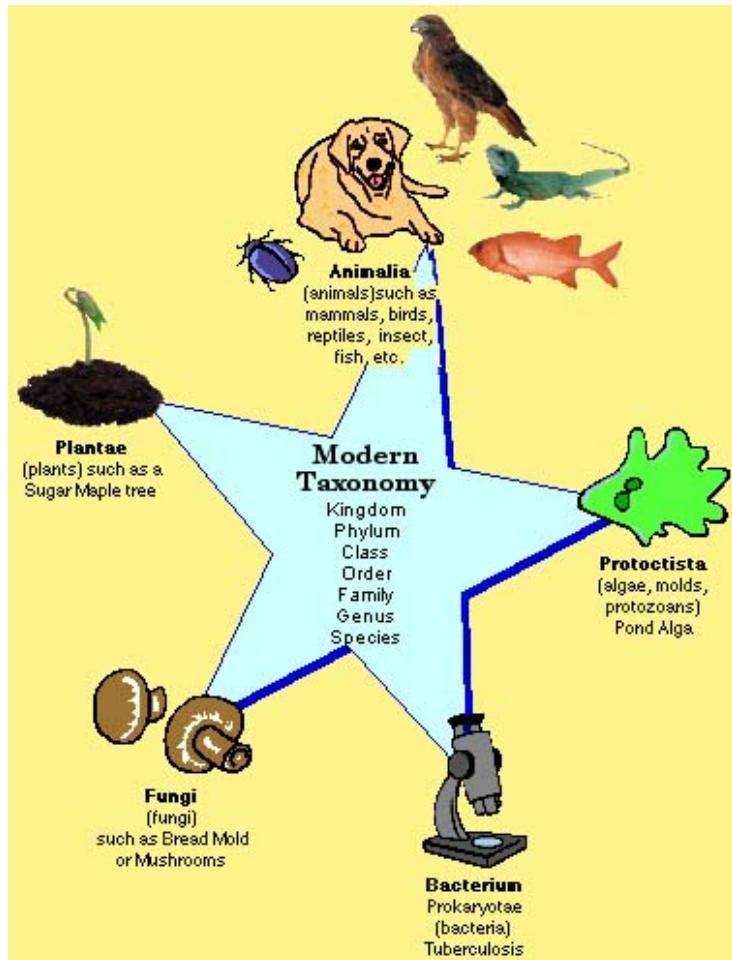
— 监督学习  
监督模式识别

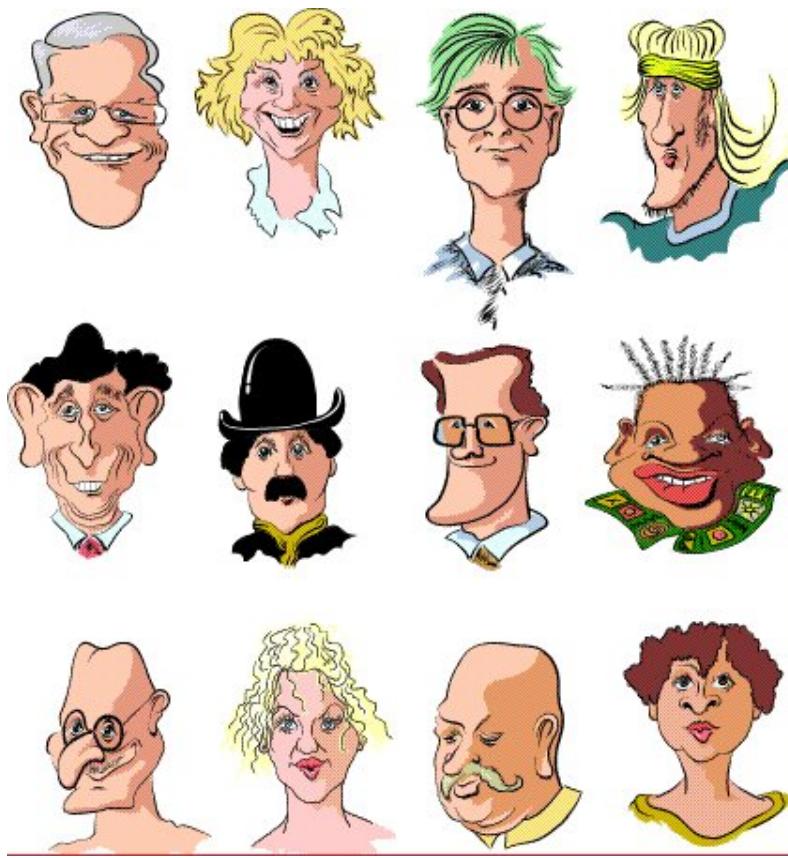
# 另一种学习



物以类聚

# 另一种学习





分类的目的  
是什么？

如何根据这些  
特征将这些样  
本分类？

- 这些人应该分成几类？
- 根据性别？是否戴眼镜？头发颜色？长
- 如果要求把这些人分成两类，如何分？
- 相根？

case	sex	glasses	moustache	smile	hairs	looking
1	m	y	n	y	grey	0.5
2	f	n	n	y	yellow	0.6
3	m	y	n	n	green	0.8
4	m	n	n	n	yellow	0.2
5	m	n	n	y?	black	0.1
6	m	n	y	n	black	0.3
7	m	y	n	y	brown	0.7
8	m	n	n	y	grey	0.3
9	m	y	y	y	no	0.1
10	f	n	n	n	yellow	0.8
11	m	n	y	n	no	0.3
12	f	n	n	n	brown	0.4

# 非监督学习 (Unsupervised Learning)

- 无指导情况下的学习
  - 所面对的只有未知答案的样本
  - 由机器从这些样本中进行学习（自学习）
  - 学习的目的在于从这些样本中发现规律，这种规律应该是某种固有的关系，或者依据这种规律对对象的分类有某种功用
    - 非监督学习
    - 非监督模式识别
    - 聚类分析 cluster analysis or clustering

# 学习问题的分类

- 监督学习 v. s. 非监督学习  
(supervised)      (unsupervised)

半监督学习 (semi-supervised)

- 回归问题 v. s. 分类问题  
quantitative or qualitative (categorical)

# Examples of PR Systems

# **Speech Recognition**

## **语音识别**

演示手机语音输入：

思必驰：<http://www.aispeech.com/index.php>

科大讯飞：<http://www.iflytek.com/mobile/iflyime.html>

# 语音识别的发展

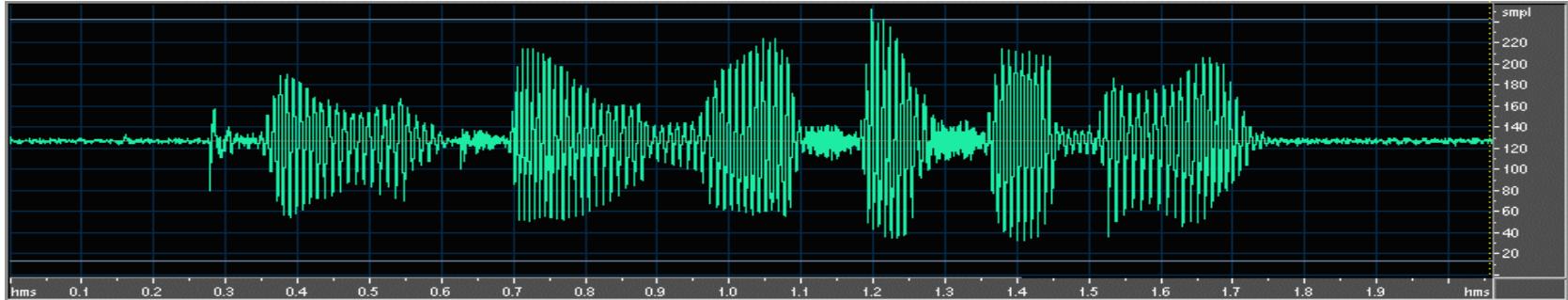
- 1952年，Bell Lab的Davis等人研制成功第一个语音识别实验系统，能识别十个英文数字
- 1960年代：语音信号的产生模型
- 1970年代：Bakis(IBM), Baker (CMU), Jelinek(IBM)等人
- 1980年代：基于矢量量化和马尔科夫模型的大词汇量连续语音识别  
    李开复 1988年 SPHINX
- 1990年代：IBM 中文语音识别 *Via Voice*
- 2000年代：Siri, 云计算
- 2010年代：深度学习

演示手机语音输入：

思必驰：<http://www.aispeech.com/index.php>

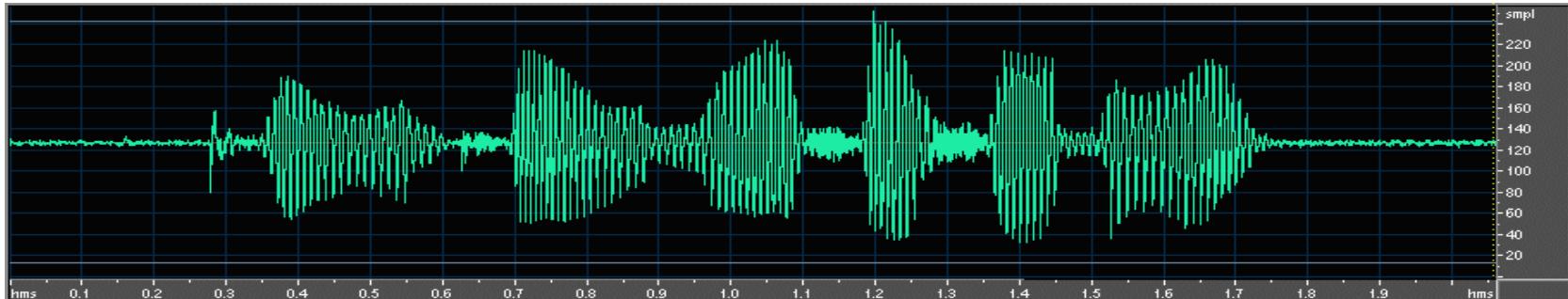
科大讯飞：<http://www.iflytek.com/mobile/iflyime.html>

# 计算机如何识别语音？



- 对于计算机来说，语音是什么?  
—— 语音是数字信号
- 直接利用这些数字信号进行识别可否?  
—— 似乎无从着手
- 怎么办?  
—— 从信号中提取更反映本质的参数
- 什么参数（特征）？

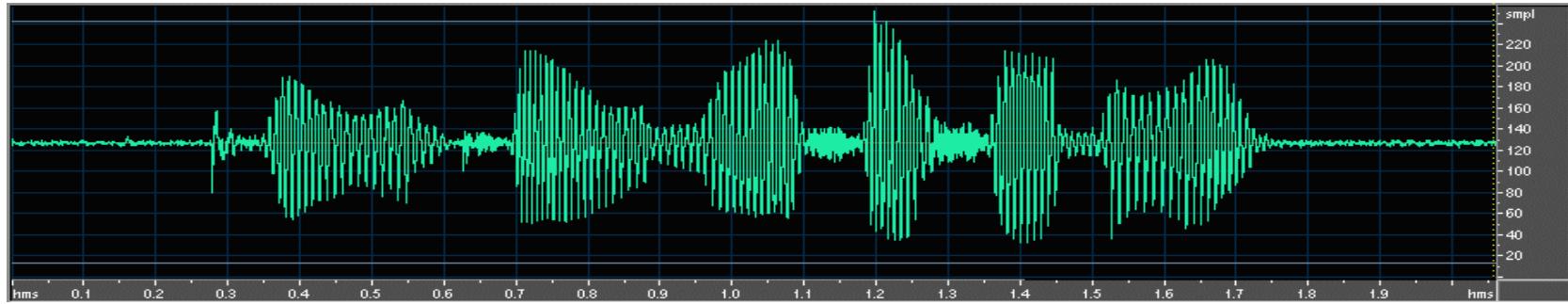
# 计算机如何识别语音？



- 基本想法：考虑所有可能的语音（类别），对于一段语音，设法提取其特征，然后计算每一种可能语音的后验概率，据此进行分类
- 为此，需要确定语音的模型（概率模型）
- 如何确定？

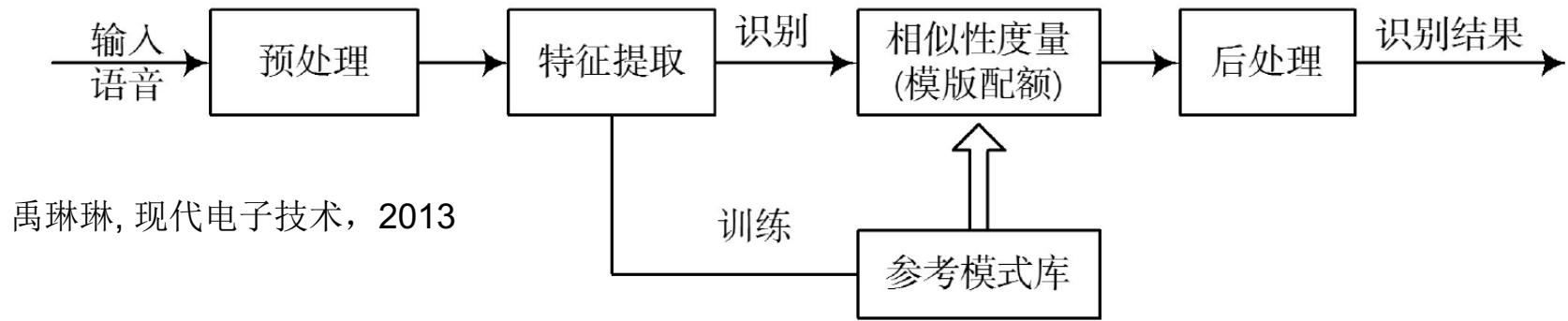
—— 根据样本（语料库）

# 计算机如何识别语音？



- 问题：以什么作为语音识别的基本单位？
- 词？ 汉语有多少词？ 几万？ 几十万？
- 字？ 汉语有多少字？ 几千
- 而且，同一个词、同一个字的发音变化很大
- 一个字的发音又由什么组成？  
—— 音素（基元） phoneme
- 一种语言的基元并不太多（几百~几千）
- 而且在一定发音时间内基本稳定（平稳）

# 语音识别中的几个关键



禹琳琳, 现代电子技术, 2013

- 语音信号处理
- 语音特征的提取
- 语音模型的建立
  - 语料库
- 分类方法
- 搜索方法

# Character and Image Recognition

# 字符与图像识别

# OCR--- Optical Character Recognition

- OCR的概念诞生于1929年的德国，1957年，OCR软件ERA
- 60年代初期，
  - NCR、Farrington、IBM，如IBM1418，能识别印刷体的数字、英文字母及部分符号，并且必须是指定的字体；60年代末，日立，富士通
- 60中 – 70初，手写体数字识别
  - IBM1287，东芝（第一个信函自动分拣系统）、NEC
  - 1974年，信函的分拣率已可达92~93%
- 80年代，质量较差的文档及大字符集的识别
  - 东芝首先研究汉字识别，1983年发布印刷体日文汉字OCR
- 我国，70年代开始研究数字、英文字母和符号，70年代末开始研究汉字识别，1986年开始进入产品阶段，现在已有多家系统，印刷体汉字识别率可达99%以上。

# 手写体字符识别

2601496357146371037114497  
6968 6961 6962 6963 6964 6965 6966 6967 6968 6969 6918 6911 6912 6913 6914 6915 6916 6917 6918 6919 6920 6921 6922 6923 6924  
11057111299811102860028887  
6925 6926 6927 6928 6929 6930 6931 6932 6933 6934 6935 6936 6937 6938 6939 6940 6941 6942 6943 6944 6945 6946 6947 6948 6949  
3301033010290602840029012  
6958 6951 6952 6953 6954 6955 6956 6957 6958 6959 6960 6961 6962 6963 6964 6965 6966 6967 6968 6969 6970 6971 6972 6973 6974  
9405290672980129450299055  
6975 6976 6977 6978 6979 6980 6981 6982 6983 6984 6985 6986 6987 6988 6989 6990 6991 6992 6993 6994 6995 6996 6997 6998 6999  
5101292018032270124431064  
6998 6991 6992 6993 6994 6995 6996 6997 6998 6999 6918 6911 6912 6913 6914 6915 6916 6917 6918 6919 6920 6921 6922 6923 6924  
11611176057188600158701899  
6925 6926 6927 6928 6929 6930 6931 6932 6933 6934 6935 6936 6937 6938 6939 6940 6941 6942 6943 6944 6945 6946 6947 6948 6949  
1157558721257068327499816  
6958 6951 6952 6953 6954 6955 6956 6957 6958 6959 6960 6961 6962 6963 6964 6965 6966 6967 6968 6969 6970 6971 6972 6973 6974  
9950512001536272203242370  
6975 6976 6977 6978 6979 6980 6981 6982 6983 6984 6985 6986 6987 6988 6989 6990 6991 6992 6993 6994 6995 6996 6997 6998 6999  
350721712723153393053880311  
6988 6981 6982 6983 6984 6985 6986 6987 6988 6989 6918 6911 6912 6913 6914 6915 6916 6917 6918 6919 6920 6921 6922 6923 6924  
1371914119129192511917014  
6925 6926 6927 6928 6929 6930 6931 6932 6933 6934 6935 6936 6937 6938 6939 6940 6941 6942 6943 6944 6945 6946 6947 6948 6949  
10118134857236803226414186  
6958 6951 6952 6953 6954 6955 6956 6957 6958 6959 6960 6961 6962 6963 6964 6965 6966 6967 6968 6969 6970 6971 6972 6973 6974  
6359720299299722510046701  
6975 6976 6977 6978 6979 6980 6981 6982 6983 6984 6985 6986 6987 6988 6989 6990 6991 6992 6993 6994 6995 6996 6997 6998 6999  
3084111591010615406103631  
6918 6919 6920 6921 6922 6923 6924 6925 6926 6927 6928 6929 6930 6931 6932 6933 6934 6935 6936 6937 6938 6939 6940 6941 6942 6943 6944 6945 6946 6947 6948 6949  
1064111030475362009979966  
6925 6926 6927 6928 6929 6930 6931 6932 6933 6934 6935 6936 6937 6938 6939 6940 6941 6942 6943 6944 6945 6946 6947 6948 6949  
8918054708553131427955460  
69150 69151 69152 69153 69154 69155 69156 69157 69158 69159 69160 69161 69162 69163 69164 69165 69166 69167 69168 69169 69170 69171 69172 69173 69174  
1018730187112993089970984  
69175 69176 69177 69178 69179 69180 69181 69182 69183 69184 69185 69186 69187 69188 69189 69190 69191 69192 69193 69194 69195 69196 69197 69198 69199  
0109707597331972013519055  
69260 69261 69262 69263 69264 69265 69266 69267 69268 69269 69270 69271 69272 69273 69274  
1075518255182814358090943  
69225 69226 69227 69228 69229 69230 69231 69232 69233 69234 69235 69236 69237 69238 69239 69240 69241 69242 69243 69244 69245 69246 69247 69248 69249  
1787541655460354603546055  
69258 69251 69252 69253 69254 69255 69256 69257 69258 69259 69260 69261 69262 69263 69264 69265 69266 69267 69268 69269 69270 69271 69272 69273 69274  
18255108503047530439401  
69275 69276 69277 69278 69279 69280 69281 69282 69283 69284 69285 69286 69287 69288 69289 69290 69291 69292 69293 69294 69295 69296 69297

- 数字 --- 字母 --- 汉字
- 在线 --- 离线  
(online/offline)

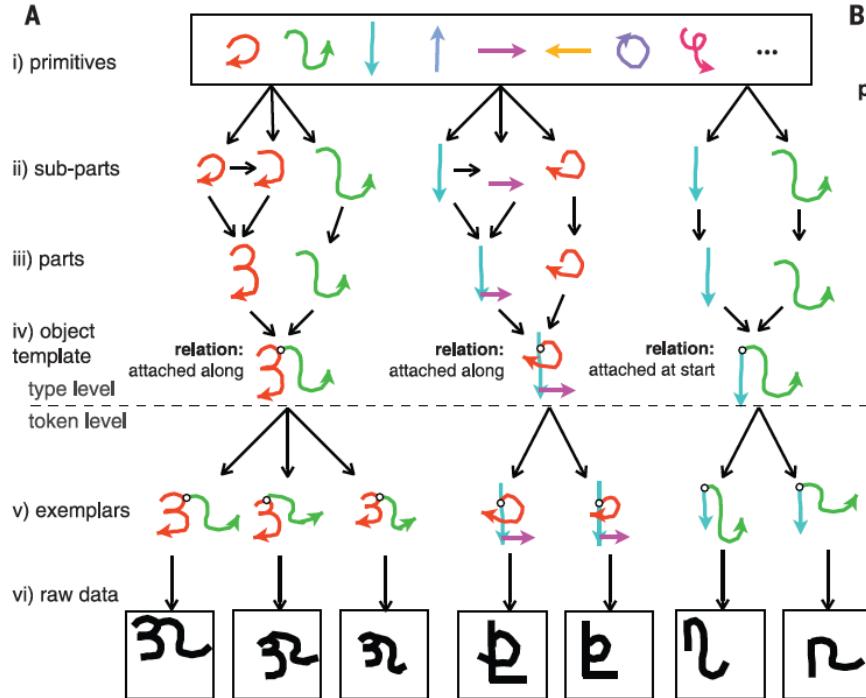
## RESEARCH ARTICLES

## COGNITIVE SCIENCE

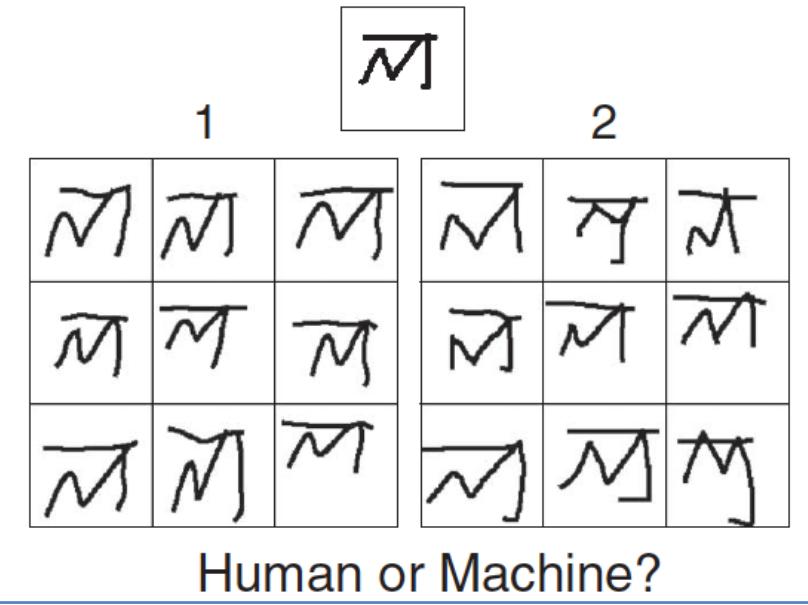
# Human-level concept learning through probabilistic program induction

Brenden M. Lake,<sup>1,\*</sup> Ruslan Salakhutdinov,<sup>2</sup> Joshua B. Tenenbaum<sup>3</sup>

A



B



Human or Machine?

```

procedure GENERATETYPE
     $\kappa \leftarrow P(\kappa)$                                 ▷ Sample number of parts
    for  $i = 1 \dots \kappa$  do
         $n_i \leftarrow P(n_i|\kappa)$                       ▷ Sample number of sub-parts
        for  $j = 1 \dots n_i$  do
             $s_{ij} \leftarrow P(s_{ij}|s_{i(j-1)})$       ▷ Sample sub-part sequence
        end for
         $R_i \leftarrow P(R_i|S_1, \dots, S_{i-1})$           ▷ Sample relation
    end for
     $\psi \leftarrow \{\kappa, R, S\}$ 
    return @GENERATETOKEN( $\psi$ )                         ▷ Return program

```

```

procedure GENERATETOKEN( $\psi$ )
    for  $i = 1 \dots \kappa$  do
         $S_i^{(m)} \leftarrow P(S_i^{(m)}|S_i)$                 ▷ Add motor variance
         $L_i^{(m)} \leftarrow P(L_i^{(m)}|R_i, T_1^{(m)}, \dots, T_{i-1}^{(m)})$  ▷ Sample part's start location
         $T_i^{(m)} \leftarrow f(L_i^{(m)}, S_i^{(m)})$           ▷ Compose a part's trajectory
    end for
     $A^{(m)} \leftarrow P(A^{(m)})$                           ▷ Sample affine transform
     $I^{(m)} \leftarrow P(I^{(m)}|T^{(m)}, A^{(m)})$           ▷ Sample image
    return  $I^{(m)}$ 

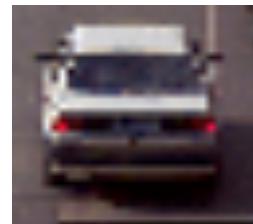
```

# 车辆的检测与追踪

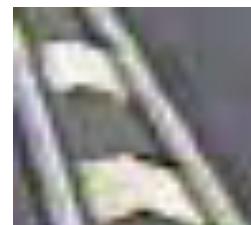
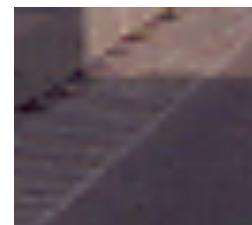
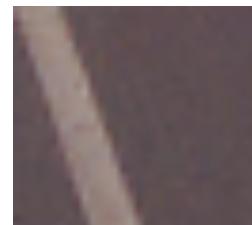


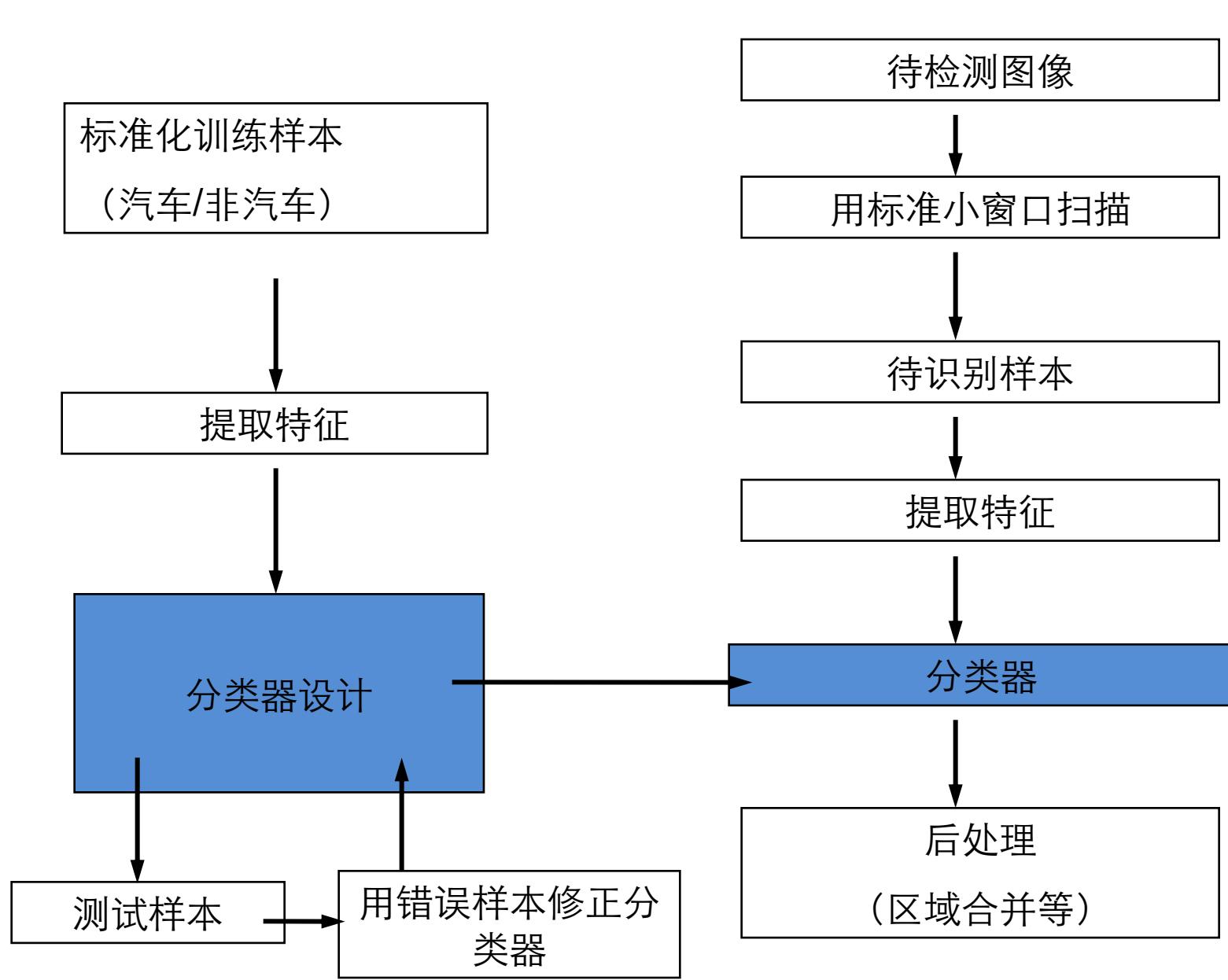
方法之一：用一个方格在图象上到处扫描，  
用PR方法判断方格内的图象是否为汽车或  
背景；分别用不同大小的方格扫描

# 汽车的例子

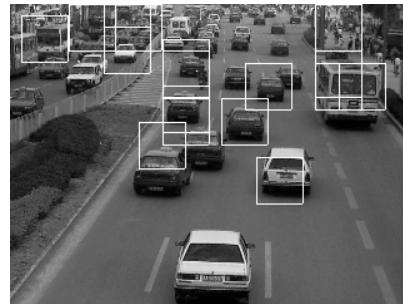
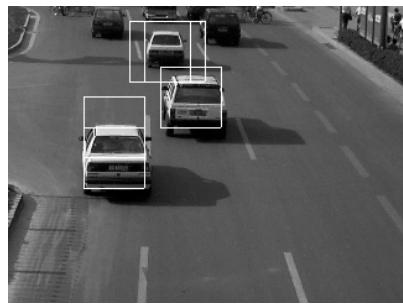
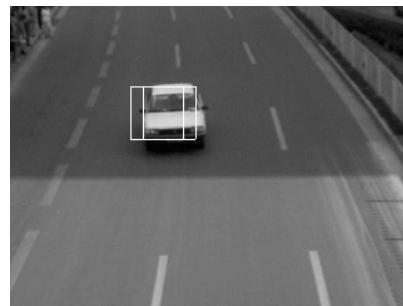


# 背景的例子





# 实验结果举例

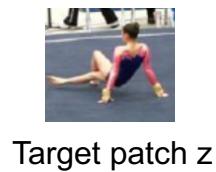


# 人脸和细胞的追踪

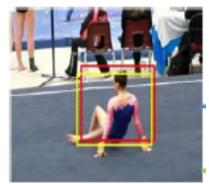
General object tracking

度量学习： Metric learning

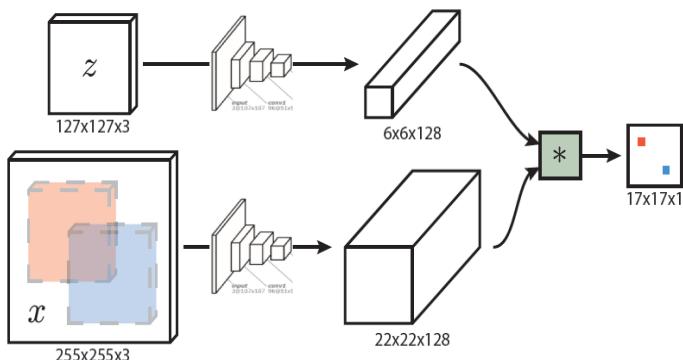
共享参数的神经网提取特征  
( Siamese network )



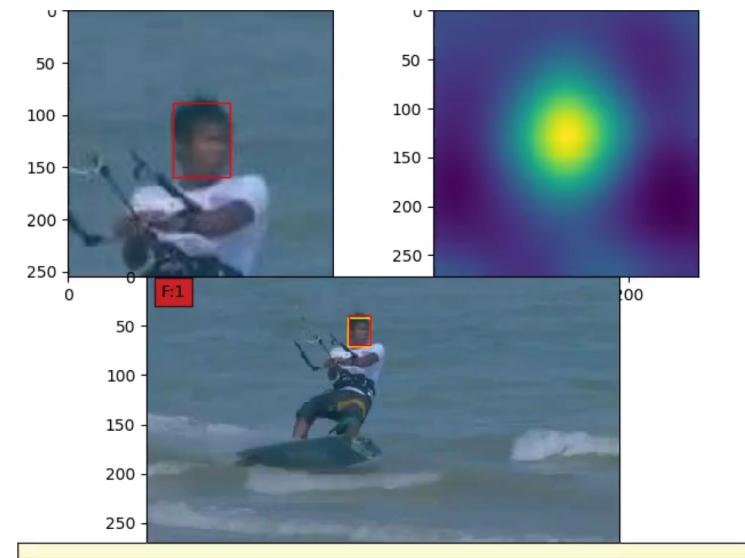
Target patch  $z$



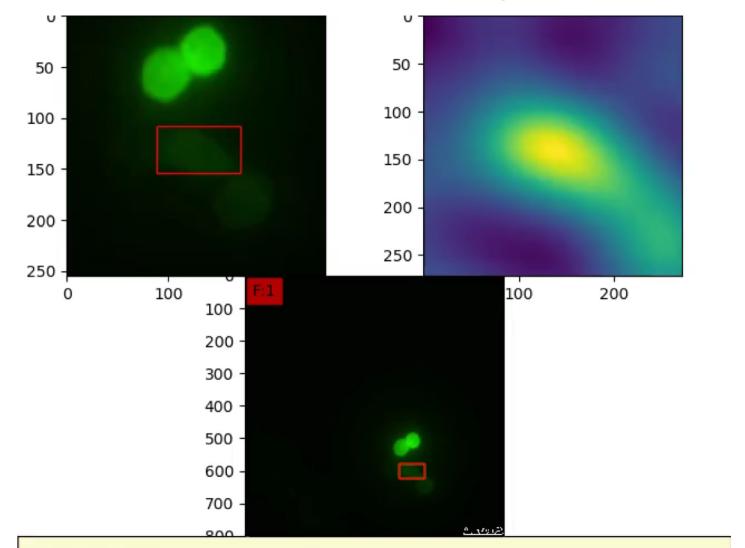
Search  
candidate  $x$



SiameseFC, CVPR, 2016



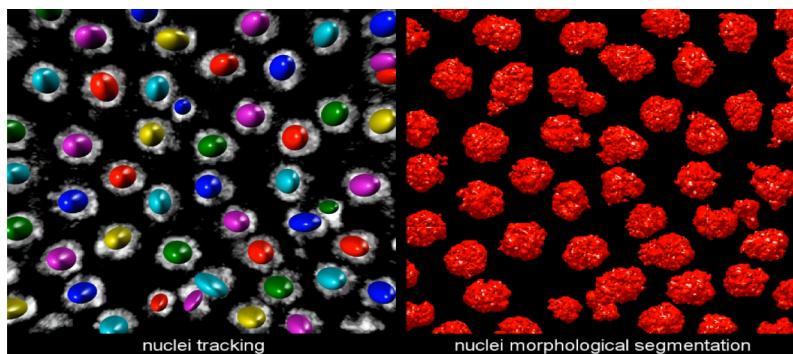
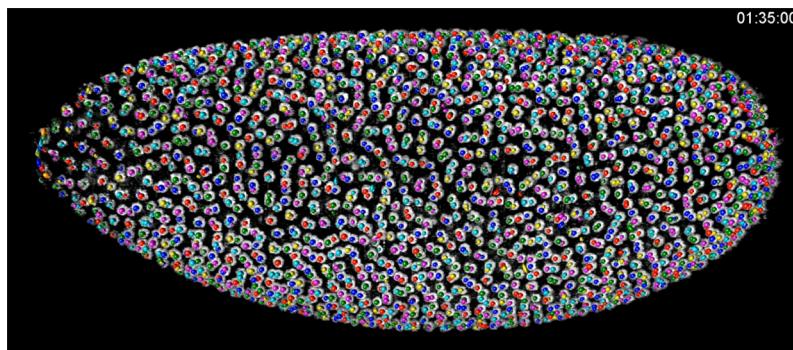
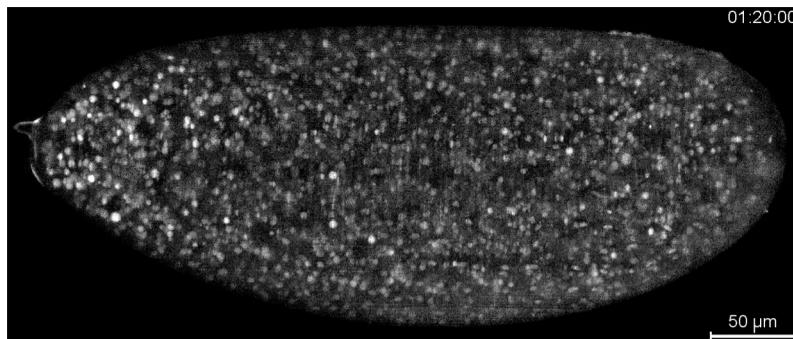
Cell tracking



# 细胞的追踪与分割



## 果蝇的胚胎发育过程



Fernando Amat et al., *Nature Methods* (2014)

从这些例子我们看到了什么？

# 特征

- 特征很重要
  - 领域知识
  - 特征选择
  - 特征提取

# 模式识别系统的基本组成

- 有已知样本情况：监督模式识别(supervised PR)



- 无已知标签样本情况：非监督模式识别(unsupervised PR)



# 处理监督模式识别问题的一般步骤

1. **分析问题**: 看是否属于模式识别问题, 把研究的目标抽象为类别; 分析问题中哪些(可以观测的)因素可能与分类有关
2. **原始特征获取**: 设计实验方法, 得到已知样本, 对这些样本实施观测和预处理, 获取与样本分类有关的观测向量(原始特征)
3. **特征选择与提取**: 为了更好地进行分类, 对特征进行必要的提取与选择
4. **分类器设计**: 利用已知样本设计(训练)某种分类器
5. **分类**: 对未知样本, 实施同样对信息获取与预处理、特征提取与选择, 用设计对分类器进行识别

# 处理非监督模式识别问题的一般步骤

1. **分析问题**: 看研究对目标是否可能抽象为若干类别;  
分析问题中哪些(可以观测的)因素可能与所关心对类别有关
2. **获取原始观测**: 观测未知样本, 获取原始特征
3. **特征提取与选择**: 进行必要的特征提取与选择
4. **聚类分析**: 采用某种方法将未知样本分类
5. **结果解释**: 分析所得的类别与所关心的目标之间对关系; 如问题需要, 用同样的方法对新的未知样本进行分类