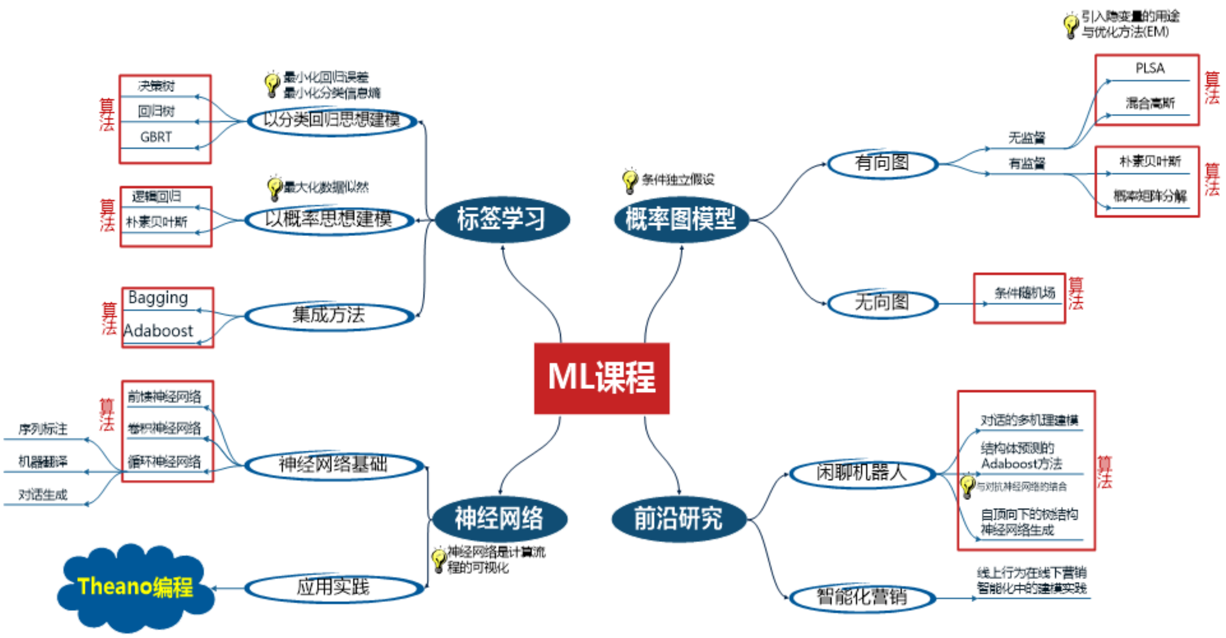
AI Course PPT 试修改说明

当前版本 2017-10-06 当前进度 22/28

目录



共性修改

大多处地方图像清晰度不高的进行了`锐化`；图片公式部分更改为`公式`，便于修改或在别处使用；部分文本进行了**加粗**或进行微调排版。

1_标签学习之决策树

由于表格底色为紫色，文字不够清晰明显，加粗后可以看得较为清楚
效果如下，以**第13页**为例

- 预测下场球赛的结果

主客场	时间	天气	部飞机 是否首发	……	胜平负 情况
主场	星期六晚上	晴朗	否		？？？

且在同时有需要显示与不需要显示的部分，仅加粗了需要显示的部分
效果如下，以**第17页**的黑色字为例

- H(胜平负|部飞机未首发)

主客场	时间	天气	部飞机 是否首发	……	胜负 情况
主场	星期天晚上	雨	否		平
客场	星期六晚上	阴天			胜

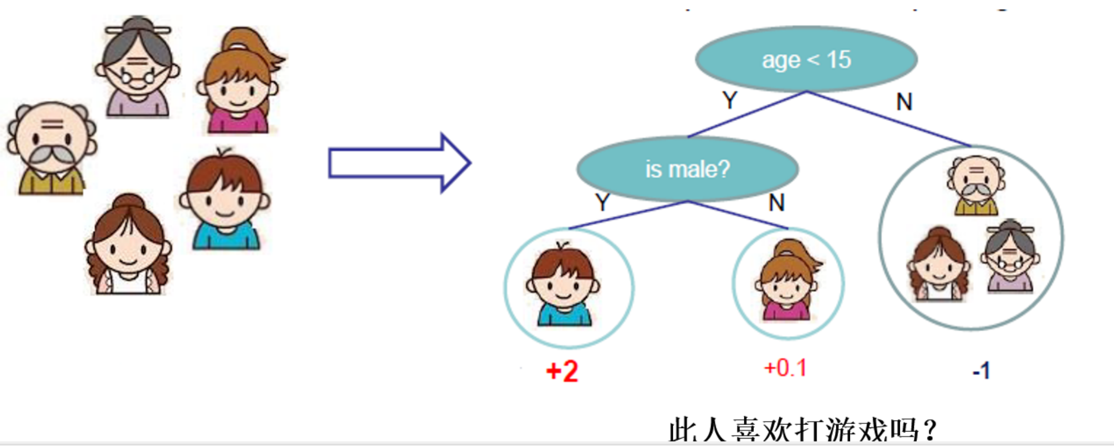
2_标签学习之回归树

准备了供日后使用的小头像人物，存放于**人物.pptx**

人物	年龄	性别		备注	人物	年龄	性别		备注
人设（基础属性，额外属性按需添加）						68	男		爷爷
	11	男		男孩		69	女		奶奶
	10	女		女孩		65	男		外公
	1	女		宝宝		65	女		外婆
	41	男		爸爸		38	男		叔叔
	40	女		妈妈		37	女		阿姨

对第6页中的“输入：年龄、性别、职业等信息”做补充，直观展示输入的年龄性别等信息，效果如下，在**第7页**

人物	年龄	性别	Age<15?	Is male?
	10	男	Y	Y
	11	女	Y	
	40	女		
	68	男		Y
	69	女		



当时上课时提及了韦达定理，但是不能很直观的和原式进行比较，于是加了一个按钮，按一下可以显示出来，再按一下可以去掉，见**第9页**

监督学习模型要素一：目标函数

假设：训练数据为$\{x_i, y_i\}_{i=1}^{n}$，第j个叶子上的预测值为w_j，树上有T个叶节点，则我们有如下目标函数：

$$
\begin{aligned}
&\sum_{i=1}^{n} l(y_i, \hat{y}_i) + \gamma T + \frac{\lambda}{2}\sum_{j=1}^{T} w_j^2 \\
=&\sum_{j=1}^{T}\sum_{i\in I_j} l(y_i, w_j) + \frac{\lambda}{2}\sum_{j=1}^{T} w_j^2 + \gamma T \\
=&\sum_{j=1}^{T}\sum_{i\in I_j} (w_j^2 + y_i^2 - 2y_i w_j) + \frac{\lambda}{2}\sum_{j=1}^{T} w_j^2 + \gamma T \\
=&\sum_{j=1}^{T}[(A_j + \frac{\lambda}{2})w_j^2 - 2B_j w_j] + \gamma T + const
\end{aligned}
$$

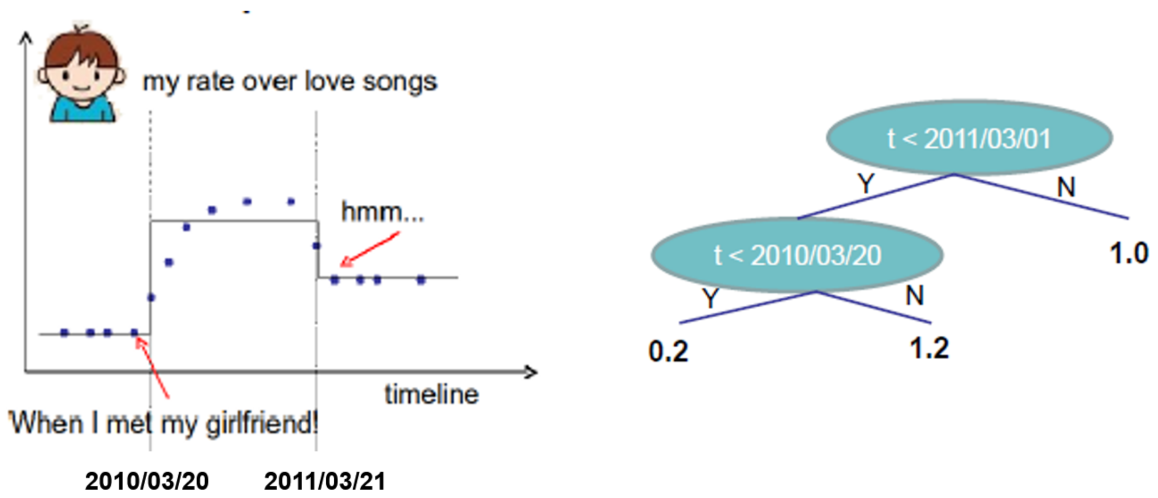
$$w_j^* = \frac{2B_j}{\lambda + 2A_j}$$

韦达定理

这里，loss函数是square error函数，且$A_j = |I_j|$，$B_j = \sum_{i\in I_j} y_i$，$I_j$是所有落入第$j$个叶节点的样本的下标集合。

9

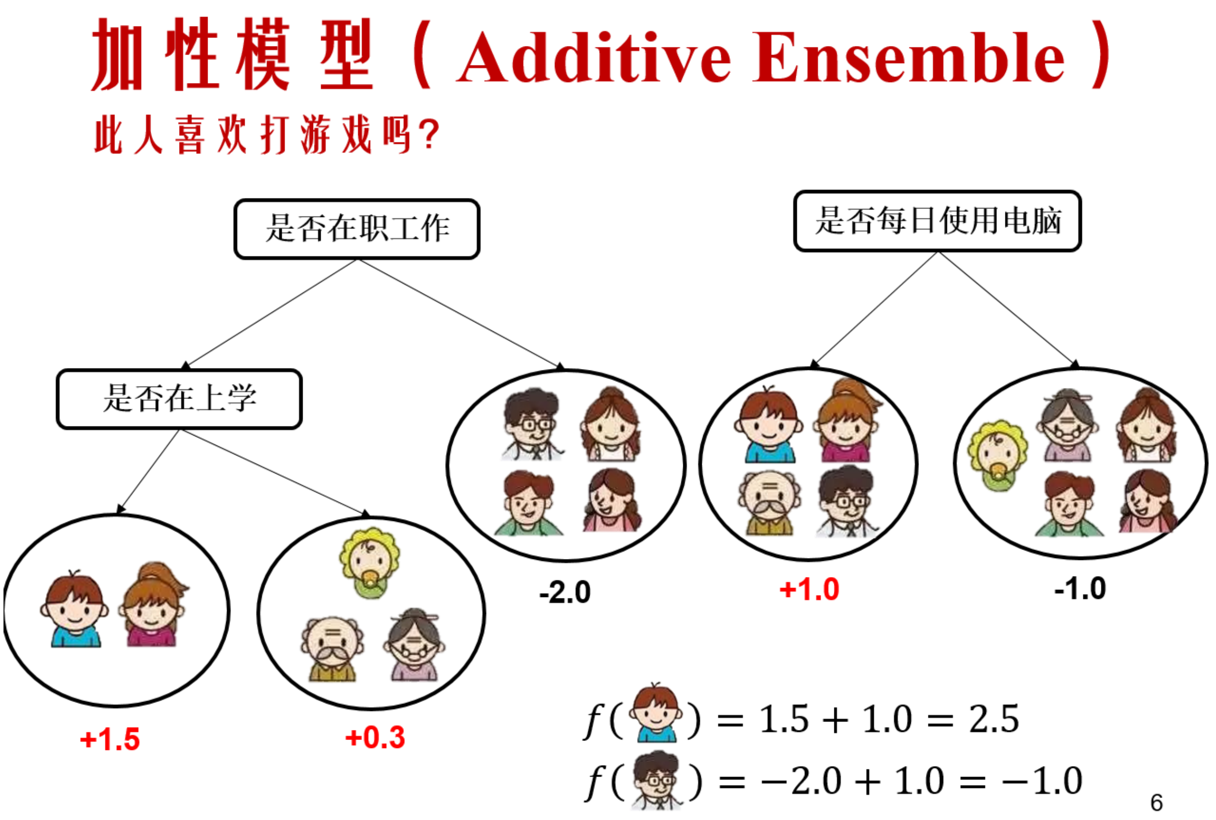
第12页中可能会造成不知道这两个时间节点哪里来的问题，试着在图上加了时间信息



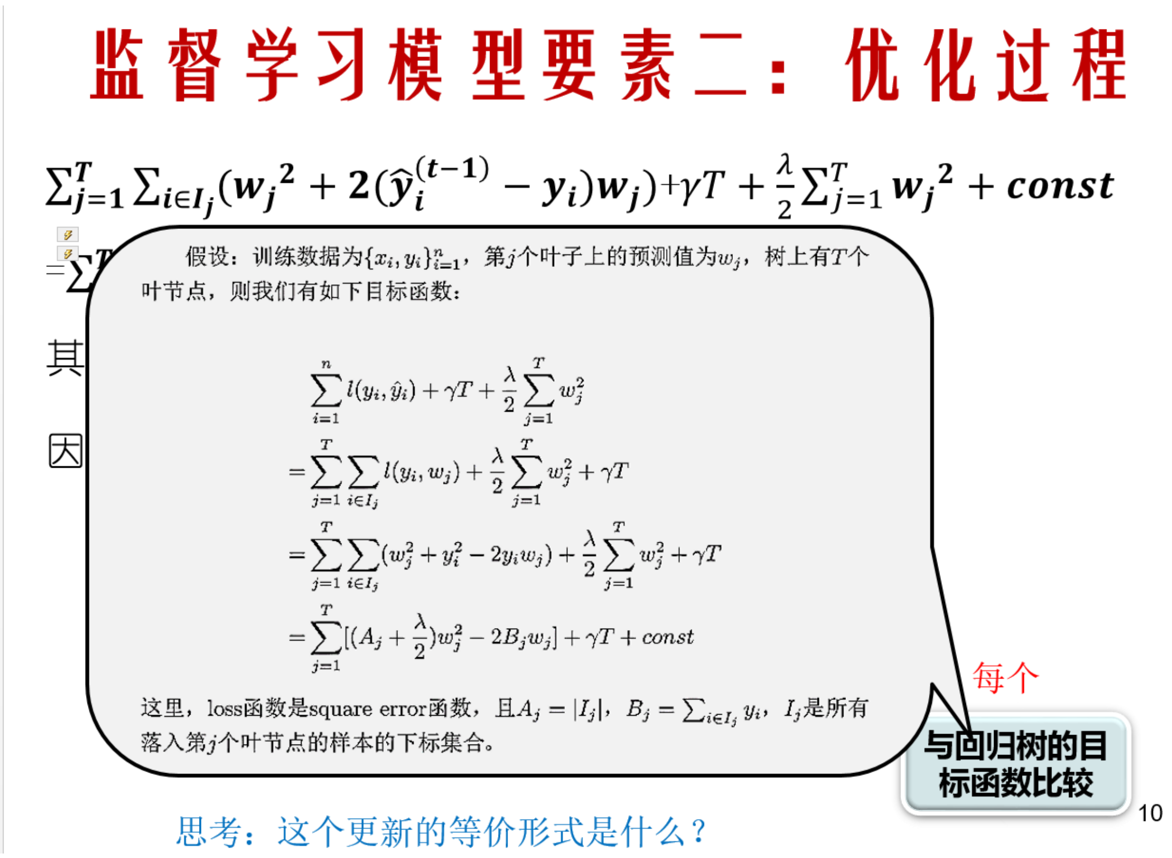
3_标签学习之GBRT

要求是“构造一个差不多10-sample的例子，举例多树计算过程”

于**第6页**尝试着做了一个9个人物的例子，本想多做几棵树但是一页能放的数目有限，如果允许用多页来表示多棵树的话我后续再添加一页，效果如下



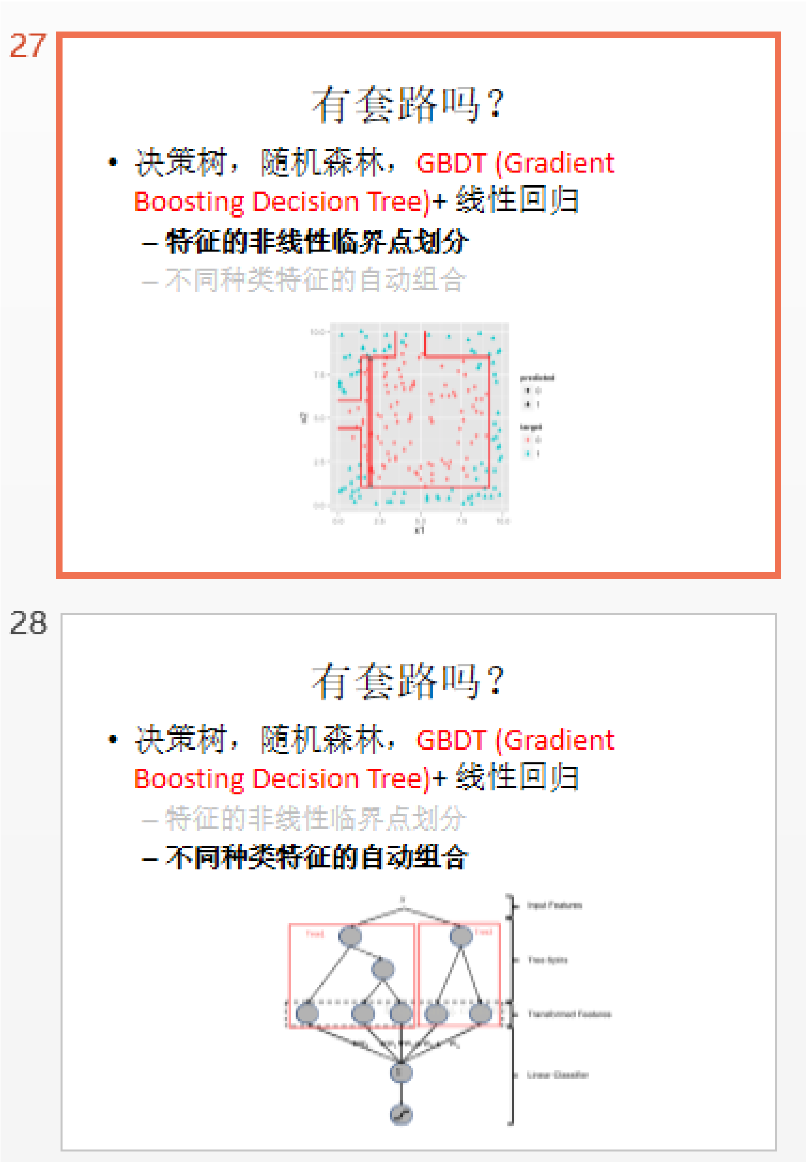
上课时，在此处回顾了回归树课件中的目标函数一页对Bj的区别进行了比较，尝试添加了按钮便于回顾，详见**第10页**



增加x轴说明，更改x坐标位置，修改错别字等，见**第12页**

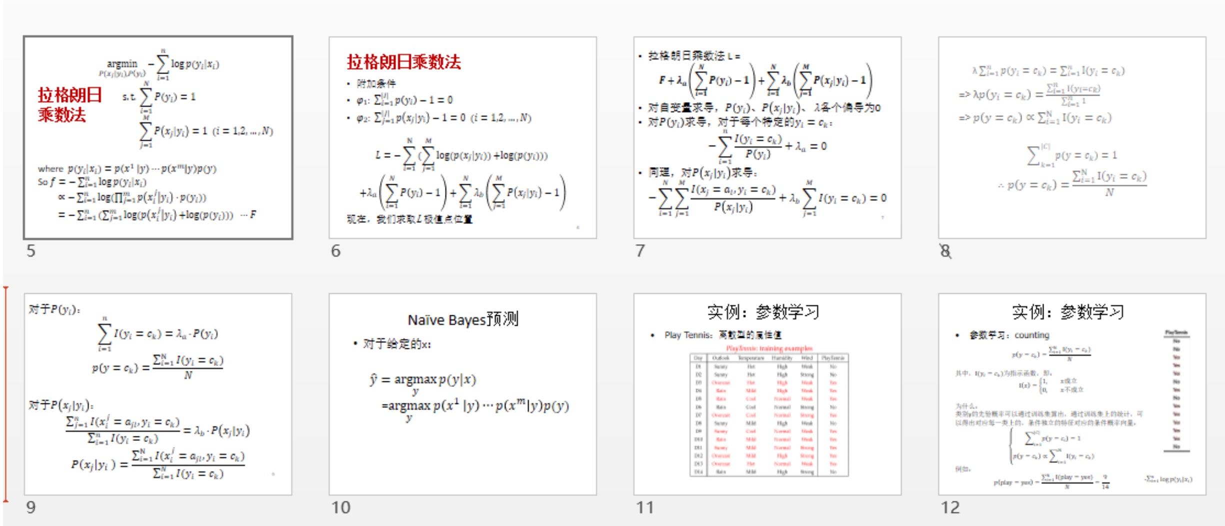
4_标签学习之LG

当时认为这一页上讲述的内容较多时间较长，故尝试分为两页，可以在讲述的时候让听众专注于当前图像上（此处修改可能无法达到比较明显的效果）



5_标签学习之NB

第5页至第9页为实例增加了较为详尽的过程



第14页重排版

实例：参数学习

- 参数学习：counting

Outlook	Play=*Yes*	Play=*No*
Sunny	2/9	3/5
Overcast	4/9	0/5
Rain	3/9	2/5

Humidity	Play=*Yes*	Play=No
High	3/9	4/5
Normal	6/9	1/5

Temperature	Play=*Yes*	Play=*No*
Hot	2/9	2/5
Mild	4/9	2/5
Cool	3/9	1/5

Wind	Play=*Yes*	Play=*No*
Strong	3/9	3/5
Weak	6/9	2/5

$P(\text{Play}=Yes) = 9/14 \quad P(\text{Play}=No) = 5/14$

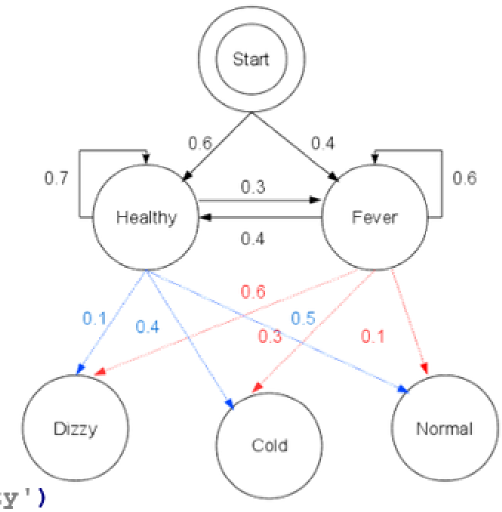
7_概率图模型之CRF

为维特比算法增加了举例引言（问题构造）

维特比算法

'normal', 'cold', 'dizzy'

怎样的健康状态序列最能够解释这个观察结果？




```
states = ('Healthy', 'Fever')
observations = ('normal', 'cold', 'dizzy')
start_probability = {'Healthy': 0.6, 'Fever': 0.4}
transition_probability = {
    'Healthy' : {'Healthy': 0.7, 'Fever': 0.3},
    'Fever' : {'Healthy': 0.4, 'Fever': 0.6}, }
emission_probability = {
    'Healthy' : {'normal': 0.5, 'cold': 0.4, 'dizzy': 0.1},
    'Fever' : {'normal': 0.1, 'cold': 0.3, 'dizzy': 0.6}, }
```

为维特比算法增加了参数解释（输入输出 参数说明）

维特比算法

- 输入：模型$\boldsymbol{\lambda} = (\mathrm{A}, \mathrm{B}, \pi)$和观测$\mathrm{O} = (o_1, o_2, \ldots, o_r)$
- 输出：最优路径$I^* = (i_1^*, i_2^*, \ldots, i_r^*)$

- 起始概率(start_probability): 表示病人第一次到访时医生认为其所处的HMM状态，60%可能健康
- 转移概率(transition_probability): 表示潜在的马尔可夫链中健康状态的变化，如健康→发烧为30%
- 观测概率emission_probability表示每天病人感觉的可能性，如“发烧了，60%的可能感觉头晕”

为维特比算法增加了例子结果与动图表示（预测）

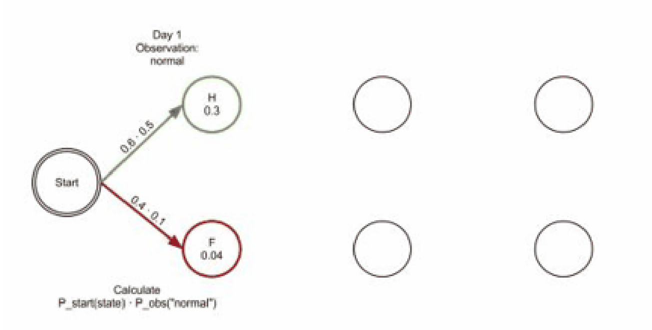
维特比算法揭示了观察结果

[‘normal’ ， ‘cold’ ， ‘dizzy’]

最有可能由状态序列

[‘Healthy’ ， ‘Healthy’ ， ‘Fever’]产生。

换句话说，对于观察到的活动： 病人第一天感到正常，第二天感到冷时都是健康的，而第三天发烧了。



该例子的结构如上, 黑色加粗的是维特比路径

1_概率图模型之有向图

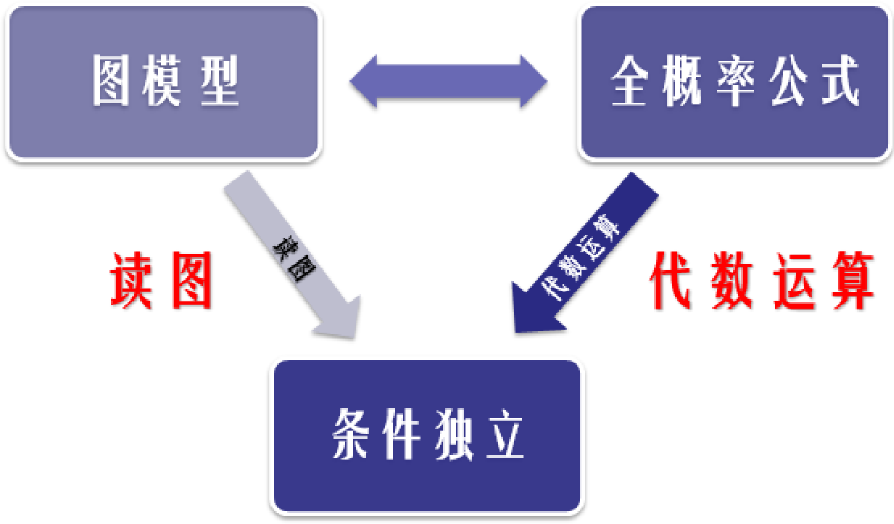
第8页至第19页及23页将公式图片均改为内嵌公式的形式

特别地，如下图**第18页**的该符号无法找到，故仅留下该处没有转换

$$a \not\perp\!\!\!\perp b \mid c$$

第15页重绘三者关系图，效果如下

图模型、全概率公式、条件独立



第21页至第22页将原先使用下划线表示的下标更改为下标形式

✓ $A=\{x_2\}, B=\{x_5\}, C=\{x_1\}$

✓ $A=\{x_2\}, B=\{x_5\}, C=\{\}$

✗ $A=\{x_2\}, B=\{x_5\}, C=\{x_4\}$

x_1

2_概率图模型之有监督有向图

第3页指数部分字过小，将公式进行整体放大

$r_{i,j}$服从如下正态分布：

$$r_{i,j} \sim N(\vec{u}_i^{\,T}\vec{v}_j, \sigma^2) \qquad f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

于**第4页**增加电影评分矩阵的例子，便于对前页进行解释

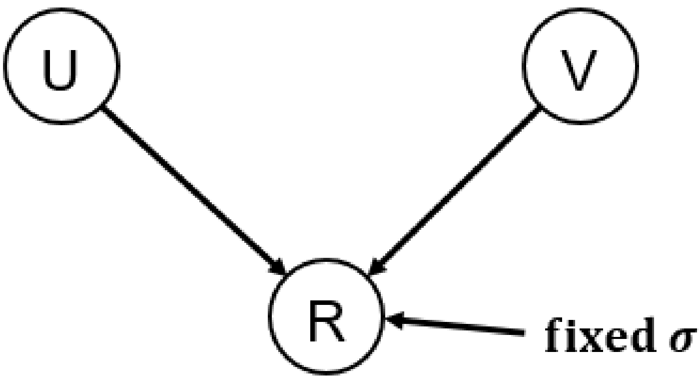
电影评分矩阵 R_{ij}

	Alex	Bob	Carl	Doge	Elf	Ford	Gail	Heil	Ilya
FilmA	5			4	3		4		1
FilmB		4			2	3		5	
FilmC			3		5				2
FilmD		3			1				
FilmE			3		3	4		5	
FilmF	4				4	5			2
FilmG									
FilmH		2			2	3		5	
FilmI					3			5	

$$R_{3,5} = R_{Carl,\,FilmE} = 3,\quad FilmE\text{:《Avatar》}$$

4

第5页的手绘图片更换为图形



$$p(r_{ai} = r) = N(\overrightarrow{u_a} \cdot \overrightarrow{v_i}, \sigma)$$

3_概率图模型之PLSA

第4页至第5页 对于共现矩阵进行了举例解释

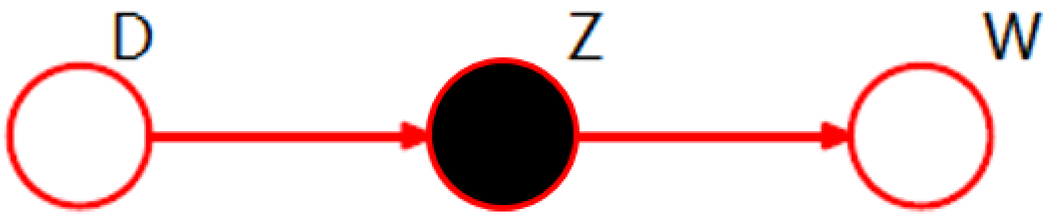
共现矩阵：两个随机变量

第6篇文档中，
第4个词出现了6次

W

0	1	0	6	3	5	0	3	1	2	2	1	1	0	0	1
1	0	0	17	0	16	0	2	1	15	1	0	6	0	0	6
0	0	0	0	0	0	0	9	0	0	0	0	0	11	0	0
6	17	0	0	0	6	0	4	0	13	3	0	22	0	0	12
3	0	0	0	0	0	0	0	2	1	0	12	0	0	7	5

解决了**第15页、第16页、第22页**Z节点涂得不够黑的问题，之后PPT中相同的修正将不作记录



记得在**第24页**修改过，但是忘记改了什么，先mark一下……

- If $p_1, \ldots, p_n$ are positive numbers which sum to 1 and f is a real continuous function that is convex (凸), then

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \le \sum_{i=1}^{n} p_i f(x_i)$$

- If f is concave (凹), then the inequality reverses, giving

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \ge \sum_{i=1}^{n} p_i f(x_i)$$

24

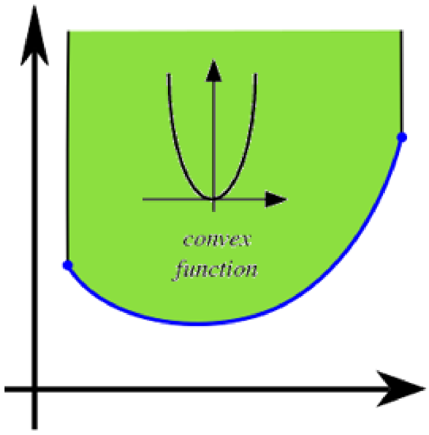
在**第25页**增加了关于凸函数与Jenson不等式的定义

凸函数与Jenson不等式

对于每一个凸函数 $f(X)$：

如果 X 是一个随机变量，在 f 的定义域内取值，那么：

$$f\big(E(X)\big) \le E\big(f(X)\big)$$



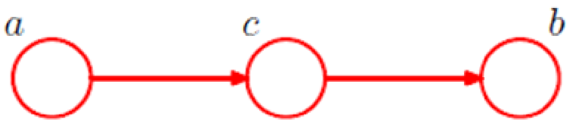
一个函数是凸的当且仅当其上境图（在函数图像**上方**的点集）为一个凸集。

而凸集是一个点集合，其中每两点之间的直线点都落在该点集合中

—— Wikipedia

25

第32页的图形改为可修改的公式

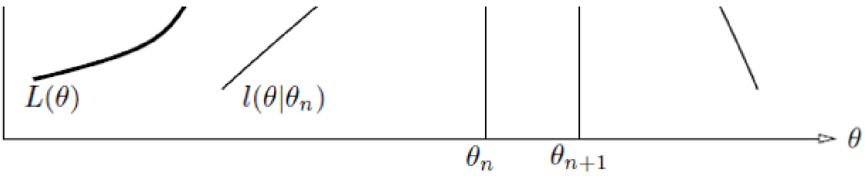


$$\boldsymbol{p(a), p(c|a), p(b|c)}$$

$$p(c|a,b) = \frac{p(a,b,c)}{p(a,b)} = \frac{p(a)p(c|a)p(b|c)}{\sum_c p(a)p(c|a)p(b|c)}$$

$$= \frac{p(c|a)p(b|c)}{\sum_c p(c|a)p(b|c)}$$

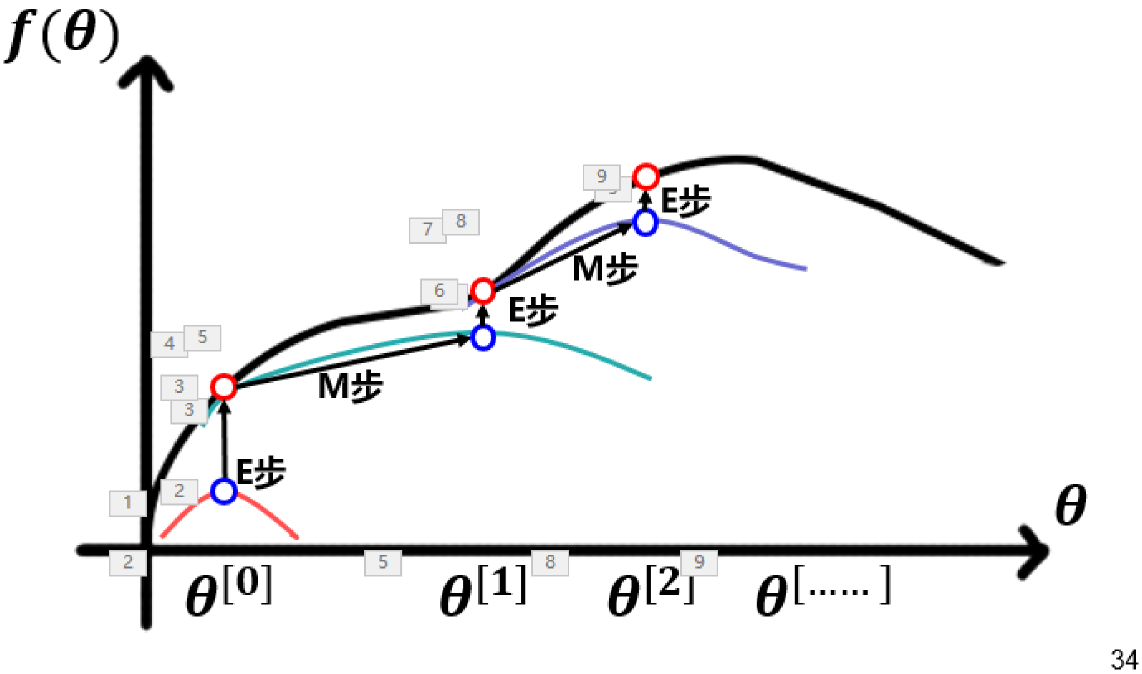
第33页参考文献居中了一下，以后这种比较小又不太有什么影响的改动是否就不要事无巨细地提及比较好，否则看起来很啰嗦……但是我觉得又不能一声不吭的就对某处进行改动不说……



32

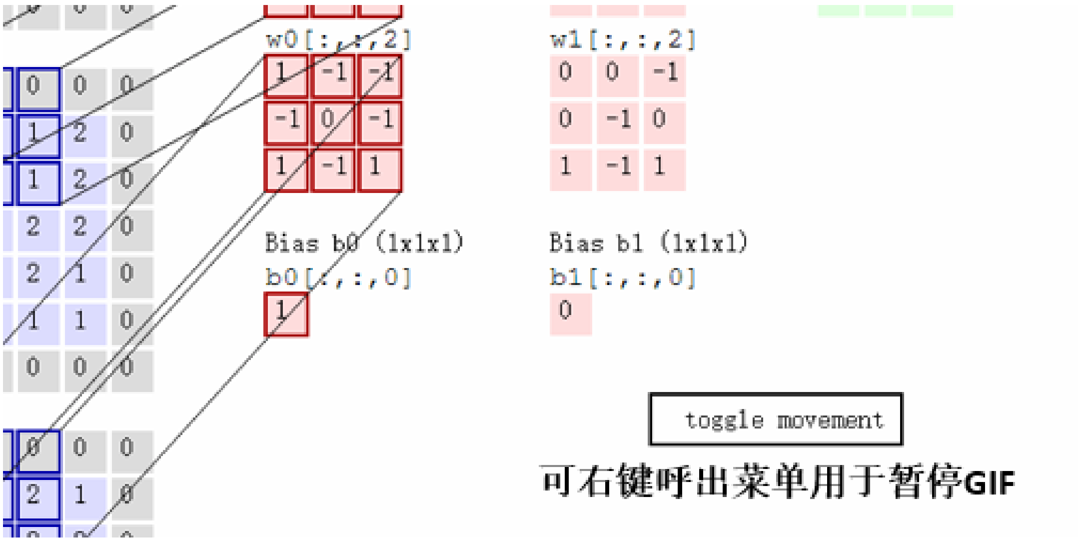
Sean Borman. The Expectation Maximization Algorithm, a Short Tutorial.

于**第35页**画图说明参数更新（并表明EM过程）



2_卷积神经网络

于**第14页**及**第19页**增加温馨提示



扩大了**第15页**、**第16页**的内容大小

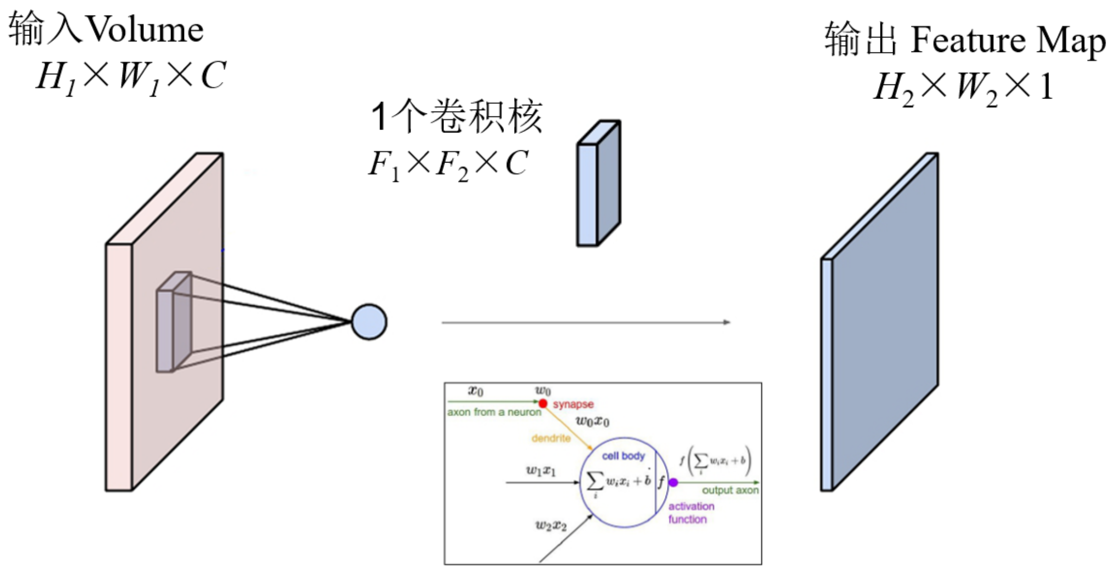
相关操作的数据表达

$$C(k_1,k_2)=(W\star X)(k_1,k_2)=\sum_i\sum_j W(i,j)X(k_1+i,k_2+j)$$

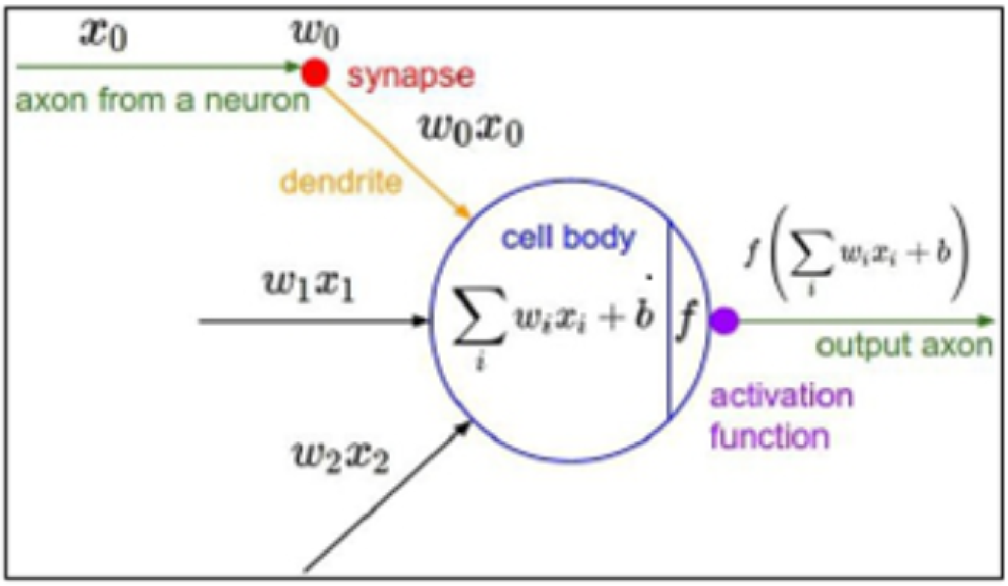
平移： $X(i,j)\rightarrow X(k_1+i,k_2+j)$

相乘累加: $\sum_i\sum_j W(i,j)X(k_1+i,k_2+j)$

为**第17页**较小的图增加了触发器，点击这张图可以放大，再次点击这张图恢复原样



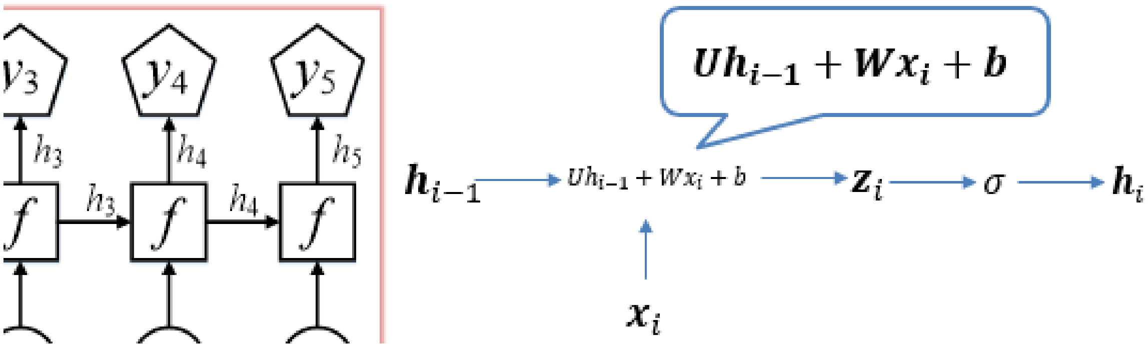
CNN基本结构--卷积层



3_循环神经网络

在**第11页**该处的关键公式旁进行放大展示

误差反向传播



2_多机制建模

第2页的Microsoft拼写错误

Such as:

- Apple Siri
- Facebook M
- Google Assistant
- Mircosoft Xiao Ice

3_树生成

第26页此处的矩形宽度拉长一些

Search-Round 1

The search path is "tree of trees"
The results are trees