

6.034期末考试

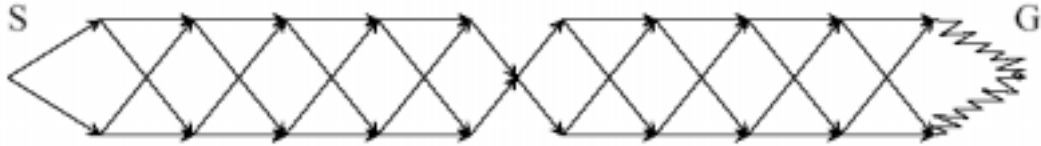
2001年秋季

姓名	
Email	

题目个数	分值	得分	分级
题目 1	16		phw rcb psz jb ac jl an dr js
题目 2	16		phw rcb psz jb ac jl an dr js
题目 3	20		phw rcb psz jb ac jl an dr js
题目 4	20		phw rcb psz jb ac jl an dr js
题目 5	18		phw rcb psz jb ac jl an dr js
题目 6	10		phw rcb psz jb ac jl an dr js
总分	100		

题目1: 在神秘的城市里搜索(16分)

你已经决定应聘驾驶出租车的工作。招聘人员是MIT的一名研究生, 他将要求你解决一些搜索街道的问题, 街道由图中的线段组成。



地图依照比例绘制, 但是注意最后收敛于G的锯齿型街道长度可以认为是其他街道长度的100倍。所有街道都是由左至右单行的。当没有其他的条件时, 你只需遵从上述的条件即可。

你的工作是根据不同的条件, 选择从起始点S到终点G的一条路径。注意共有23个节点。

答案可以使用数字(例如,8), 没有变量的表达式(例如, $2^0+2^1+2^2+2^3$) 或者两者混合。任意一个答案正确就可得满分。如果你注意到两个答案相等即使两者都是错误的, 也可以增加你的分值。

Part A: 深度优先搜索, 带回溯(2分)

不使用已访问列表或者扩展列表。当搜索队列出现一个可到达目标的路径时搜索结束。指出扩展节点的个数。

Part B: 广度优先搜索(2分)

使用访问列表。当搜索队列出现一个可到达目标的路径时搜索结束。指出扩展节点的个数。

Part C: 分支限界(2分)

找到最优路径, 不使用扩展列表和启发式方法。指出扩展节点的个数。

Part D: A* (4 分)

找到最优路径, 使用扩展列表和作为启发的到目标的距离。指出扩展节点的个数。

Part E: 双向广度优先搜索(4 分)

不使用访问列表和扩展列表。左边的搜索与右边的搜索相互交叉。即一个搜索从左边开始产生所有深度为1的路径; 然后另一个搜索从右边开始与箭头反方向进行, 产生所有深度为1的路径。然后两个搜索交替进行, 给他们的搜索树增加附加层, 直到左边的一个路径和右边的一个路径经过同一个节点。指出扩展节点的个数。

Part F: 比较 (2 分)

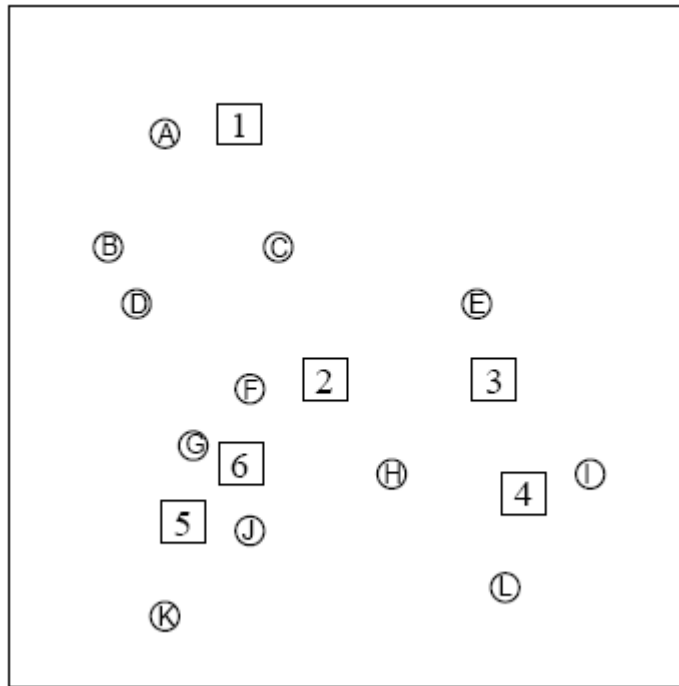
显然, 在PartE中双向广度优先搜索比单向广度优先搜索的工作量小。事实上, 单向广度优先搜索的扩展节点数与双向广度搜索扩展的节点数的比率为:

- 2
- $b^{d/2}$
- b^d
- $2 \times b \times d$

在单向搜索中 b 是分支因素, d 是总的深度。

题目 2: 最近邻法和ID树(16 分)

Part A: 最近邻法, 向后搜索(6 分)



用圆圈数据点去分类矩形测试点。
 每个圆圈被标志为+或者-, 这里没有显示出来。
 矩形测试点不能用于分类其它点。
 给定一些矩形的分类结果, 恢复圆圈的标志。

矩形点	使用1近邻	使用3近邻
1	-	+
2	-	
3		+
4	+	-
5		-

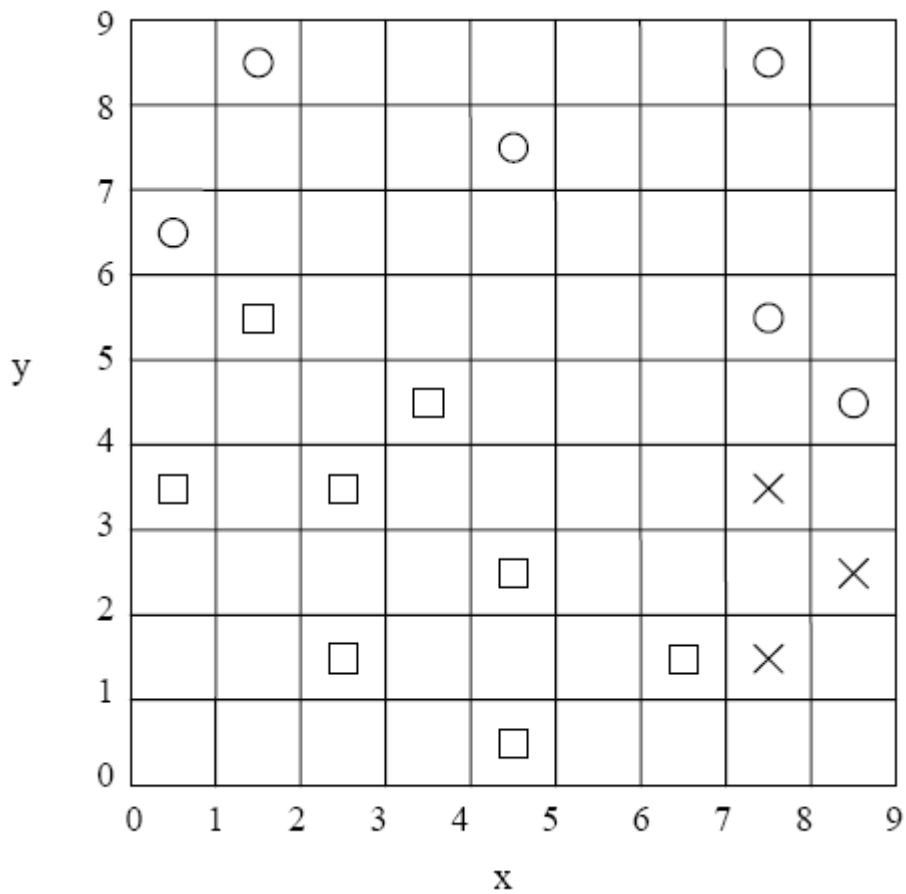
为下列标志出+, -, 或者U代表不能确定

- _____ 圆圈 L.
- _____ 圆圈 I.
- _____ 圆圈 H:
- _____ 圆圈 E.
- _____ 圆圈 K.
- _____ 圆圈 C.
- _____ 矩形 6 使用 1 近邻.

- _____ 矩形 6 使用 3 近邻.
- _____ 矩形 2 使用 3 近邻.
- _____ 矩形 3 使用 1 近邻.
- _____ 矩形 5 使用 1 近邻.

Part B: 建立一个树(6分)

下图描述了通过构造ID树产生的决策边界，该ID树使用标准最小熵值贪婪算法完全分类数据。只用X和Y分隔(即只有垂直和水平线)。对于每个分割写出分割线端点的坐标(例如(3, 0)至(3, 9))。对于每个分割计算平均熵(即无序程度)，这个熵值是由每个分割所得的点决定的。题目为书写答案提供了足够多的空间。

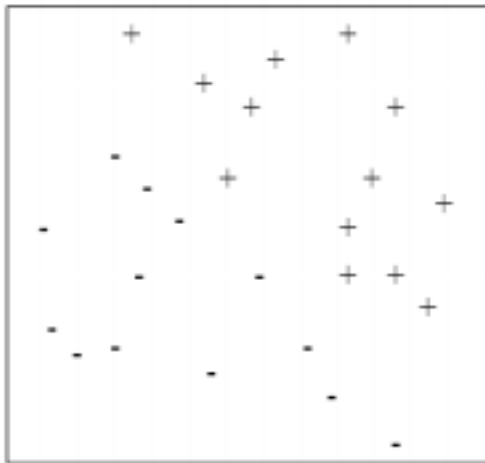


- 分割1: (_____, _____) 至 (_____, _____) 熵值: _____
- 分割2: (_____, _____) 至 (_____, _____) 熵值: _____
- 分割3: (_____, _____) 至 (_____, _____) 熵值: _____
- 分割4: (_____, _____) 至 (_____, _____) 熵值: _____
- 分割5: (_____, _____) 至 (_____, _____) 熵值: _____
- 分割6: (_____, _____) 至 (_____, _____) 熵值: _____

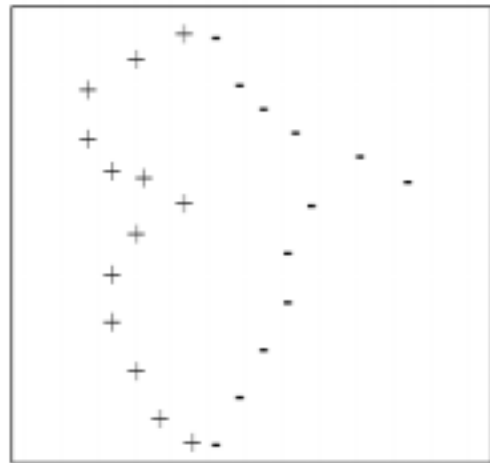
Part C: 衍生的特征 (4分)

观察若干图形数据, 你决定试验衍生特征, 你可能会比只用X或Y划分做得更好。给每个图选取一个特征, 使用该特征只需一次分割, ID树就可以正确分类数据。注意原点在左下方, 正方形的面积是 1×1 。从下面的特征进行选择:

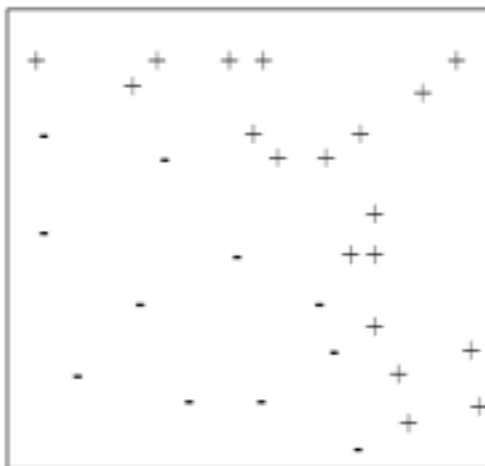
1. X 2. Y 3. X+Y 4. X-Y 5. XY 6. X^2+Y
 7. $X+Y^2$ 8. X^2+Y^2 9. $\text{atan}(Y/X)$ 10. $\sin(4\pi X)+Y$ 11. $X+\sin(4\pi Y)$



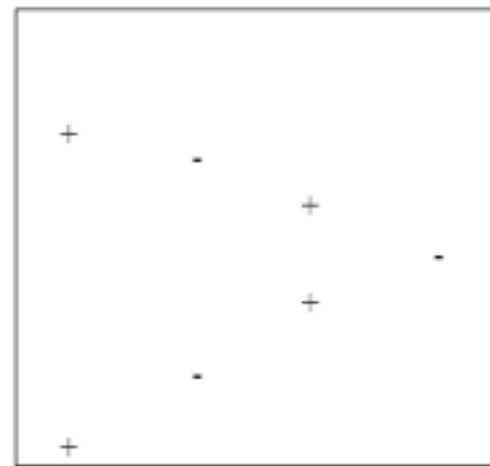
衍生特征#: _____



衍生特征#: _____



衍生特征#: _____



衍生特征#: _____

题目 3: 约束传播(20 分)

在熟练掌握了地图着色艺术并且拒绝了众多绘图咨询公司的邀请后,你尝试一些基本的自然语言理解的问题。你的工作是在给定一个词的音素顺序后,使人工智能系统得到它按音节的发音。咨询了6.034课程的助教后,你知道音节由三个元素组成:开始音节、核心音节和可选择的结束音节。

- 开始音节(O): 音节中元音之前的一个或多个辅音
- 核心音节(N): 音节中的元音
- 结束音节(C): 音节中元音之后的一个或多个辅音

例如:



你利用词汇的传播约束用O, N, C赋值,为每个音素建立模型。(CSP的变量是音位,它的定义域是O, N和C)。从一个简单的例子,单词elektivz(electives)入手。开始变量的定义域如下:

e	l	e	k	t	i	v	z
O	O	O	O	O	O	O	O
N	N	N	N	N	N	N	N
C	C	C	C	C	C	C	C

然后,助教给你了一些一元的约束:

1	辅音不能被赋值为N
2	元音不能被赋值为O或者C
3	单词不能以C开头
4	单词不能以O结尾

现在应用这些一元约束,限制变量定义域,得到以下结果:

		e	l	e	k	t	i	v	z
O			O		O	O		O	
N		N		N			N		
C			C		C	C		C	C

Part A：二元约束(8分)

噢，”你的助教挠了挠头说，“我忘记给你二元约束了……”

5	O后面不能接C
---	---------

你立即用深度优先的搜索方法寻找一组合适的变量分配。你继续自左向右以O, N, C的顺序进行赋值。每一次赋值后都测试约束5，如果违反就重新赋值。画出搜索树：

现在, 填写划分音节的结果:

	e	l	e	k	t	i	v	z
数值								

Part B: 连续约束(6 分)

在向你的助教炫耀后, 他又挠挠头说: “这个发音好像不是完全正确。为了避免错误的音节分配例如agf和rdo, 你必须考虑到响度。” 幸运的是, Winston教授走了进来, 他说: “哦, 响度就是一个音节的发音和一个元音发音的相似程度,” 同时从他的背包中拿出了一张表格; 注意在同一个方框内的音素响度相同。

响度最低

响度最高

g, d, b, k, t, p	z, s, v, f	r, l, n, m	u, o, i, e, a
------------------	------------	------------	---------------

现在, 你的助教给你解释另外两个关于开始音素响度上升和结束音素响度下降的约束:

6	O1-O2序列必须满足响度上升 $O1 < O2$
7	C1-C2序列必须满足响度下降 $C1 < C2$

“现在你又有足够的信息了,” 他说。你加入这两个约束重新完成Part A的搜索, 画出搜索树:

单词*elektivz*中第一个不符合约束的变量是什么？

有多少个赋值发生了变化？

最后划分音节的结果是什么？

	e	l	e	k	t	i	v	z
数值								

Part C: 保守校验(6 分)

Winston 提出保守的约束校验。**保守校验**就是利用与变量相关的约束限制它的定义域，直到一个近邻的定义域缩小为一个数值。这种检查按照近邻的顺序依次进行直到定义域缩小为一个数值。

现在你将单词 *perseptronz* (*perceptrons*)用保守校验法划分音节。首先利用一元约束，然后做深度优先搜索，利用所有约束自左向右并进行保守校验。以 O, N, C 的顺序再次赋值，画出搜索树。

你总共做了多少次校验?

赋值多少次?

最后划分音节的结果是什么?

	p	e	r	s	e	p	t	r	o	n	z
数值											

题目4：神经网络(20 分)

Part A: (2 分)

一年过完后收到了丰富的节日礼物和生日礼物，你打算找到一个可以在打开礼物之前就能判断出它的好坏的方法。由于这听起来像是机器学习的任务，你决定使用神经网络。如果礼物是坏的输出接近0，是好的输出接近1。所有输入的值都在0和1之间。它们是：

宽度(Width)：由0(窄)到1(宽)

深度(Depth)：由0(浅)到1(深)

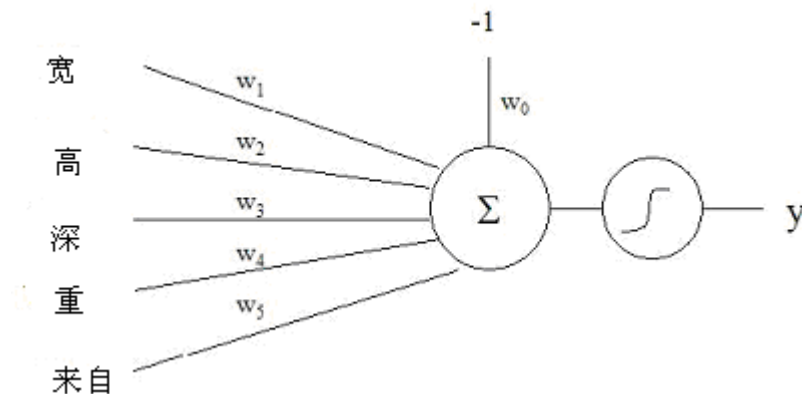
高度(Height)：由0(短)到1(长)

重量(Weight)：由0(轻)到1(重)

来源(From)：由0(来自普通的亲戚)到1(来自富有的亲戚)

你回去找到过去几年收到的数百件礼物，大约好坏各占一半，用这些数据来训练神经网络。

你的第一个想法是构造一层神经网络，将所有的输入引入到一个单独的s型(Sigmoid)单元中，像这样：



Part A.1

你知道所有来自富有亲戚的礼物都是好的，所以在训练后相应的权值应该是：

- A. 大正数
- B. 正数接近于零
- C. 负数接近于零
- D. 大负数

Part A.2

在节日这段时间你及时地完成了神经网络的训练，用今年礼物来测量它的

准确度。因为s型单元输出的是连续值，但是你希望的是离散值，所以你认为任何输出小于0.5就预示是坏礼物，大于0.5预示是好礼物。不幸的是，你发现这个网络做了可怕的工作。

选出所有可能的原因：

- A. 对于输入变量之间复杂的交互关系，你的神经网络模型没有足够多的单元
- B. 你没给神经网络足够多的训练数据
- C. 你初始化所有的权值为0至1之间的随机数
- D. 你初始化每个权值是大正数或负数
- E. s型单元不能有效地获取你想要的关系，所以需要感知器代替

Part B (6 分)

你从经验得知只有以下的礼物是好的：

- 小信封因为它们装的是现金
- 来自富有亲戚的礼物，不管其它特征
- 高，宽，深的礼物

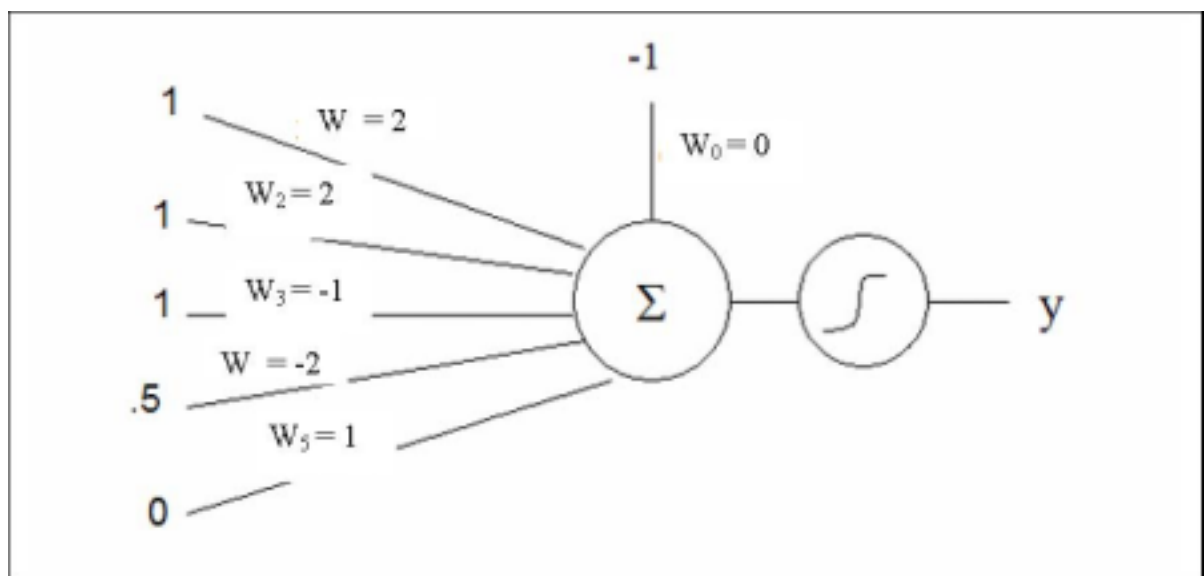
你决定建立3个独立的神经网络，分别对应上面的3种情况。每个网络都与上面描述对应，但是只有在样本列入网络对应的种类时，输出的希望值为1。对于一个测试样本，如果至少1个输出接近于1，就可以知道它是好的礼物。

一个训练样本有以下特征：

Width = 1, Depth = 1, Height = 1, Weight = 0.5, From = 0;

它是一个好的礼物，所以输出的希望值 $y^* = 1$ 。

$w_0 = 0, w_1 = 2, w_2 = 2, w_3 = -1, w_4 = -2, w_5 = 1$ 。



完成一步神经网络的训练，假设学习率 $r=100$ ，权值的初始值如上图所示，

标准误差公式为, $E = \frac{1}{2}(y - y^*)^2$ 。

为了简化你的计算, 使用下面给出的s函数的值(注意: 这些值并不是实际s函数的值, 但是使用它们可以简化运算):

Sigmoid(0)=0.5, Sigmoid(1)=0.75, Sigmoid(2)=0.9, and Sigmoid(-x)=1 - Sigmoid(x)

$w_0 =$

$w_1 =$

$w_2 =$

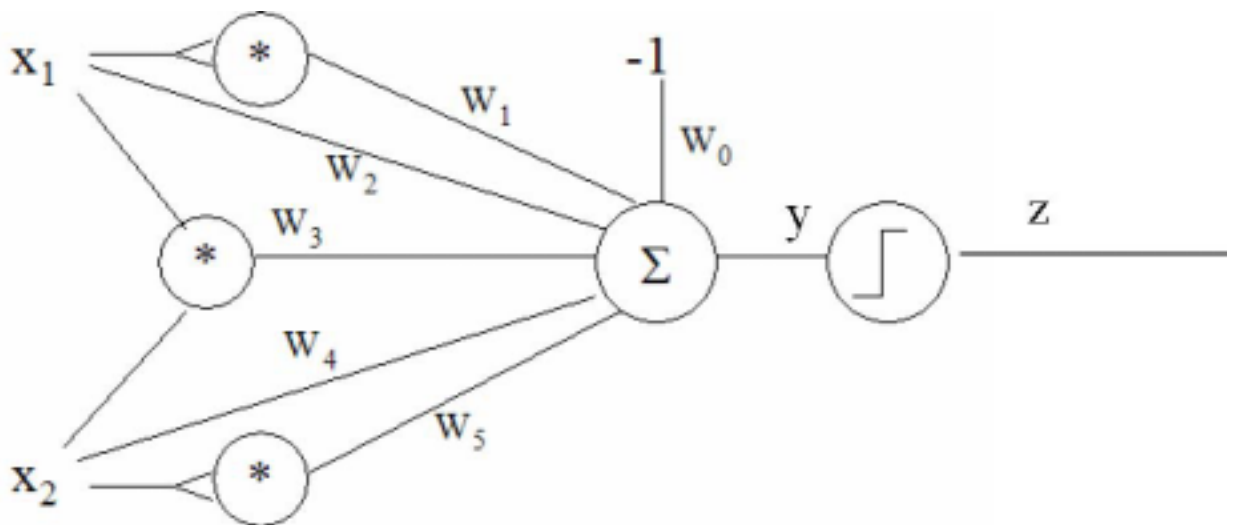
$w_3 =$

$w_4 =$

$w_5 =$

Part C (4 分)

你已经厌烦被感知器限制, 它只能在输入空间画直线, 所以你建立了下面这个网络:



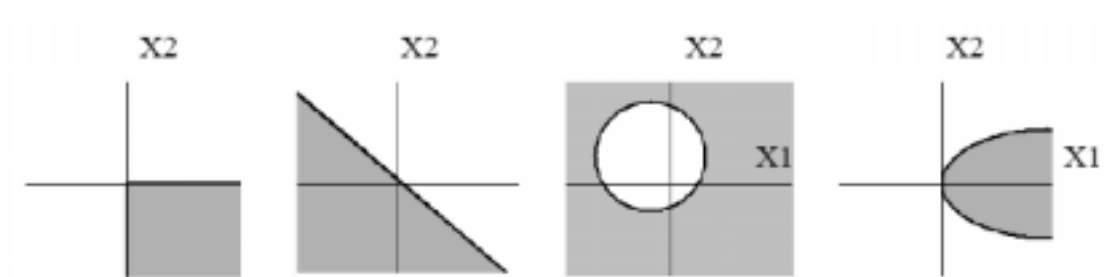
*单元的输出是所有输入的乘积, Σ 单元的输出是所有输入的和。

注意顶部和底部的*单元的输入分别是两个 x_1 和 x_2 ，所以它们完成的是的 x_1 和 x_2 平方。

所以阶跃函数的输入是：

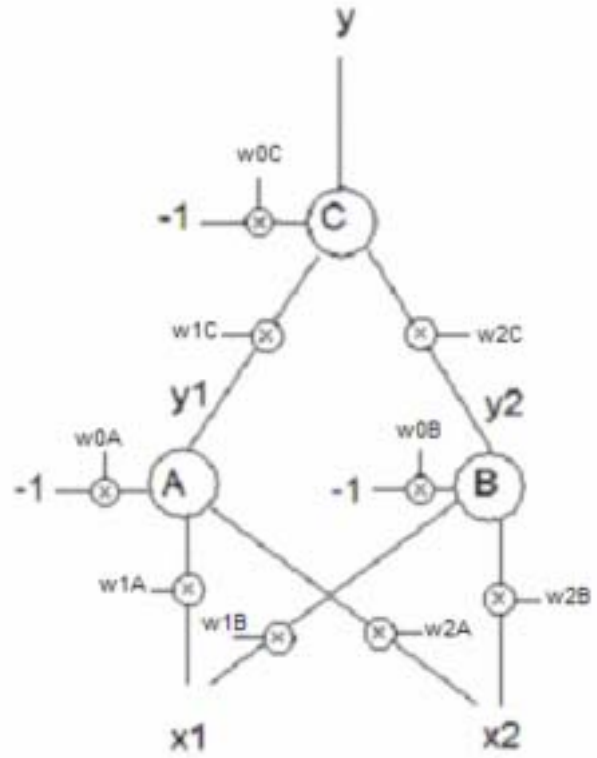
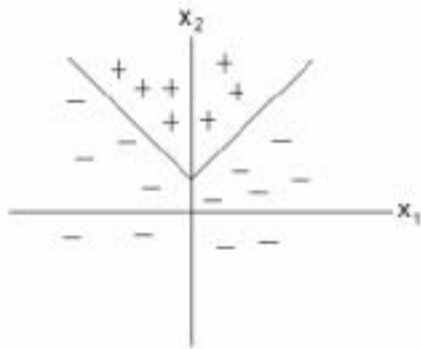
$$y = w_1x_1^2 + w_2x_1 + w_3x_1x_2 + w_4x_2 + w_5x_2^2 - w_0$$

选择上述单一神经元可以将阴影部分同非阴影部分分离的所有图形。每个图形的区域覆盖了所有可能的输入。



Part D: (8 分)

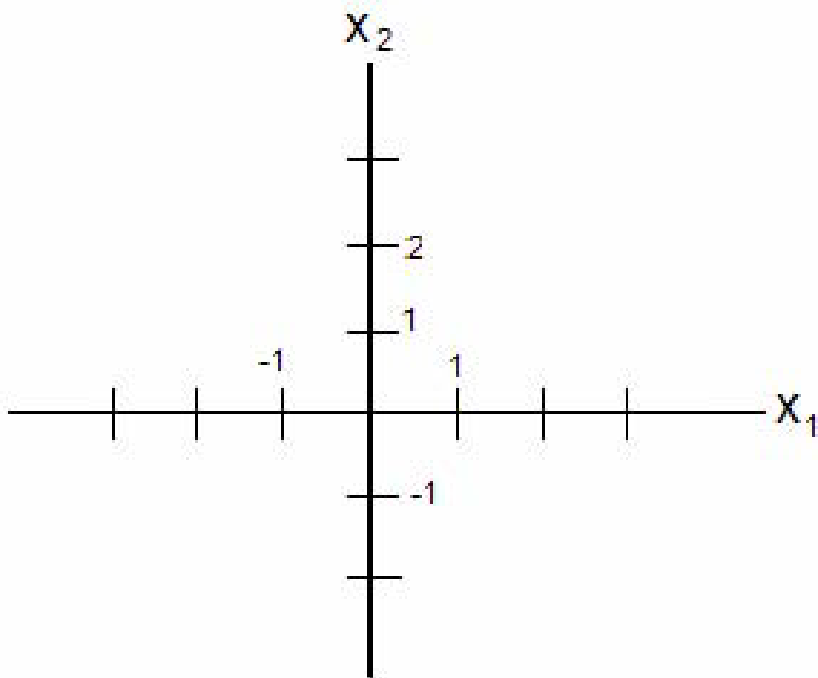
你想建立一个使用感知器单元的神经网络，它将判断 $x_2 > |x_1| + 1$ 是否成立。



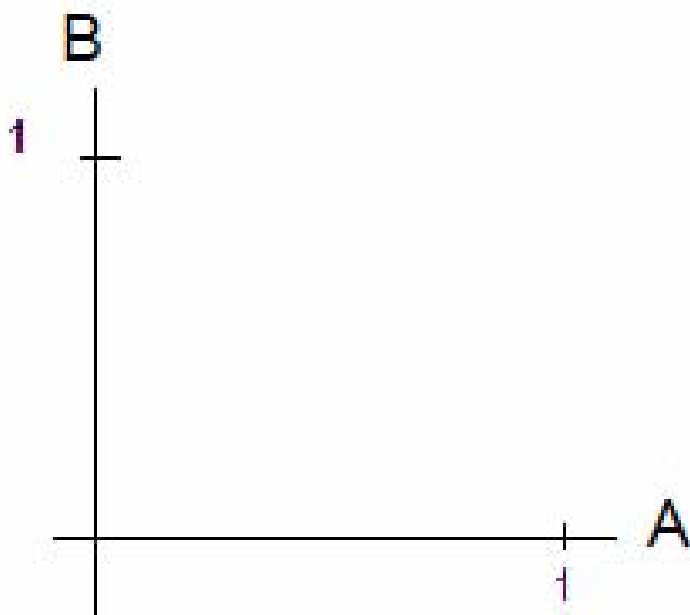
给网络A添上权值使它可以依据函数 $x_2 > |x_1| + 1$ 正确地分类数据。对输出为0的区域标志-, 为1的区域标志+。

- $w_{0A} =$
- $w_{1A} = 1$
- $w_{2A} = 1$
- $w_{0B} = 1$
- $w_{1B} =$
- $w_{2B} =$
- $w_{0C} =$
- $w_{1C} =$
- $w_{2C} =$

在下面的图表中说明感知器A和B的决策边界。标注A和B的边界并判断每一边是0还是1。



在下面的图表中画出单元A和B输出的所有组合。给定函数 $x_2 > |x_1| + 1$ ，用和组合相关的类别标注所有组合（+、-或二者）。



题目 5: SVM (18 分)

Part A (2 分)

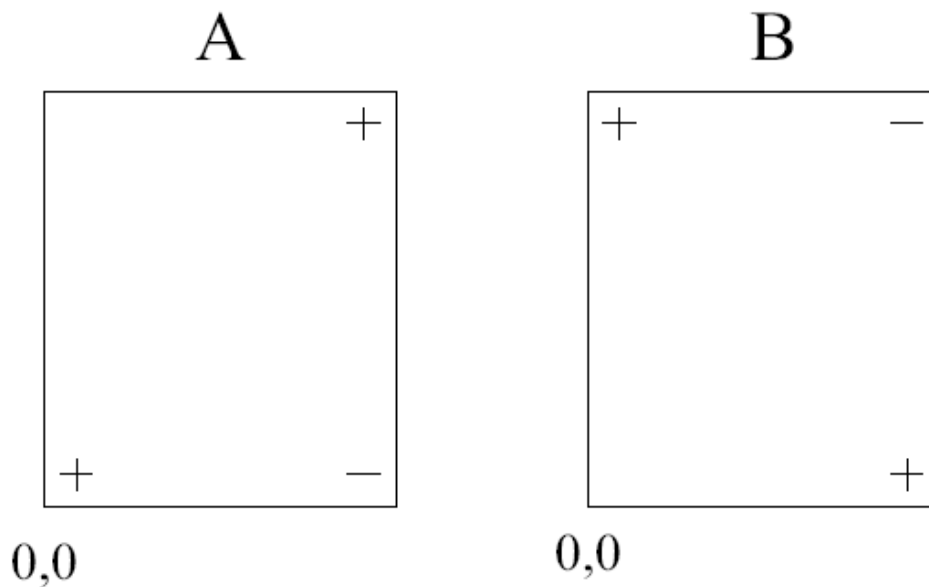
观察图A和图B中+和-的分布，图B中的分布是由图A以中心旋转得来。然后选择下列的判断是否正确：

给定一个径向基函数， $k(v_1, v_2) = e^{-\frac{\|v_1 - v_2\|^2}{0.5}}$ ，图A中点+、-的边界通过平移和旋转可以得到图B中的边界。

- 正确
- 错误

给定一个多项式函数， $k(v_1, v_2) = (v_1 \cdot v_2)^2$ ，向量从原点画出，图A中点+、-的边界通过平移和旋转可以得到图B中的边界。

- 正确
- 错误



Part B (4 分)

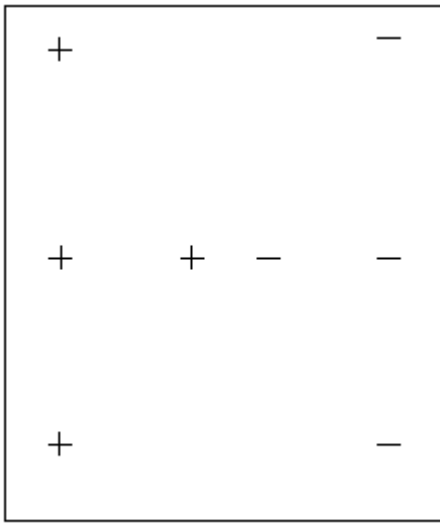
假设向量支持机在下面的图表中学会分离点+和点-。画出-1和+1这两条直线并且圈出与支持向量相对应的点，假设多项式核 $k(v_1, v_2) = (v_1 \cdot v_2)^2$ 。



Part C.1(2分)

假设向量支持机在下面的图表中学会分离点+和点-。画出决策边界(直线0, 不是直线-1或者+1)并且圈出与支持向量相对应的点, 使用径向基函数

$$k(v_1, v_2) = e^{-\frac{\|v_1 - v_2\|^2}{2\sigma^2}}, \text{ 和一个非常大的和 (sigma)。}$$

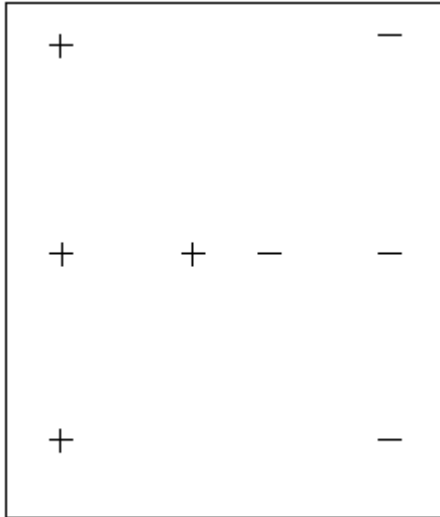


0,0

Part C.2(2 分)

假设向量支持机在下面的图表中学会分离点+和点-。画出决策边界(直线0, 不是直线-1或者+1)并且圈出与支持向量相对应的点, 使用径向基函数

$$k(v_1, v_2) = e^{-\frac{\|v_1 - v_2\|^2}{2\sigma^2}}, \text{ 一个非常小的和 (sigma)。}$$



0,0

Part D (8 分)

在单独的一页, 有9个彩图A至I, 描述了训练SVM分辨+和-。 分别指出每张图使用了哪个核函数, 将图的字母写入与相对应核的旁边。

注意: 图A, B, C, D中的点相同, 图E, F, G中的点相同。

$\kappa(v_1, v_2) = (v_1 \cdot v_2)^3$		$\kappa(v_1, v_2) = e^{-\frac{\ v_1 - v_2\ ^2}{0.5}}$	
$\kappa(v_1, v_2) = (v_1 \cdot v_2)^2$		$\kappa(v_1, v_2) = e^{-\frac{\ v_1 - v_2\ ^2}{0.22}}$	
$\kappa(v_1, v_2) = (v_1 \cdot v_1)^2$		$\kappa(v_1, v_2) = e^{-\frac{\ v_1 - v_2\ ^2}{0.08}}$	

题目6: 混合问题(10 分)

选择一个完全符合题目的选项。 多选无效。 错误的答案不扣分。

当运行一个基于规则的系统:

- Rete可以用来加速正向推理链。
- 反向推理链可以使用户的问题集中在当前的假设上。
- 正向推理链推理系统决不会删出已得出结论。
- 反向推理链要求统一的匹配程序(变量到变量)。
- 以上都是。
- 以上都不是。

以下哪个条件成立时, 程序可以回答如何和为什么质疑程序过去的行为:

- 它们使用Rete。
- 它们不使用正向推理链。
- 它们建立与/或树。
- 它们不是基于规则的。
- 以上都不是。
- 以上都是。

给定向两个方向运动(没有旋转)的未知物体, 然后指出预测需要:

- 一个点和一个模型图中的对应点。
- 两个点和两个模型图中的对应点。
- 三个点和两个模型图中的对应点。
- 三个点和三个模型图中的对应点。
- 以上都不是。

在音位学习的Yip-Sussman定理中:

- 单词是由可以描述成独立特征向量的音素序列组成
- 学习可以在特征向量的缓存中进行一般化搜索的过程中发生
- 英语的音素与不到10%的特征向量相对应
- 音位规则可以通过双向约束获取

- 以上都是
- 以上都不是

在音位学习的Yip-Sussman定理中：

- 每个例子都需要学习各自的规则。
- 此外假设似是而非是不可避免的。
- 搜索是完备的，所以为了保证最好可能的一般化。
- 搜索的目的是与所有正面的匹配, 同时匹配最少负面的。
- 以上都是。
- 以上都不是。

主要学习表明：

- 程序可以从单独的例子中学习一些定理。
- 当教师可以给出比较好的似是而非的例子时，学习过程加快。
- 一般化正面的例子；特殊化负面的例子。
- 给出正确信息的清晰表述是重要的。
- 以上都是。
- 以上都不是。

似是而非的例子

- 应该特殊化
- 避免特殊化。
- 能够具体化。
- 当不知道例子是正面的还是反面的都应该学习。
- 以上都是。
- 以上都不是。

适切条件应确保：

- 老师拥有一个学生已经知道的模型。
- 老师拥有一个学生必须学习的模型。
- 任何学生学习的知识都可以被学生使用。
- 学生有足够的计算能力去学习老师讲授的知识。

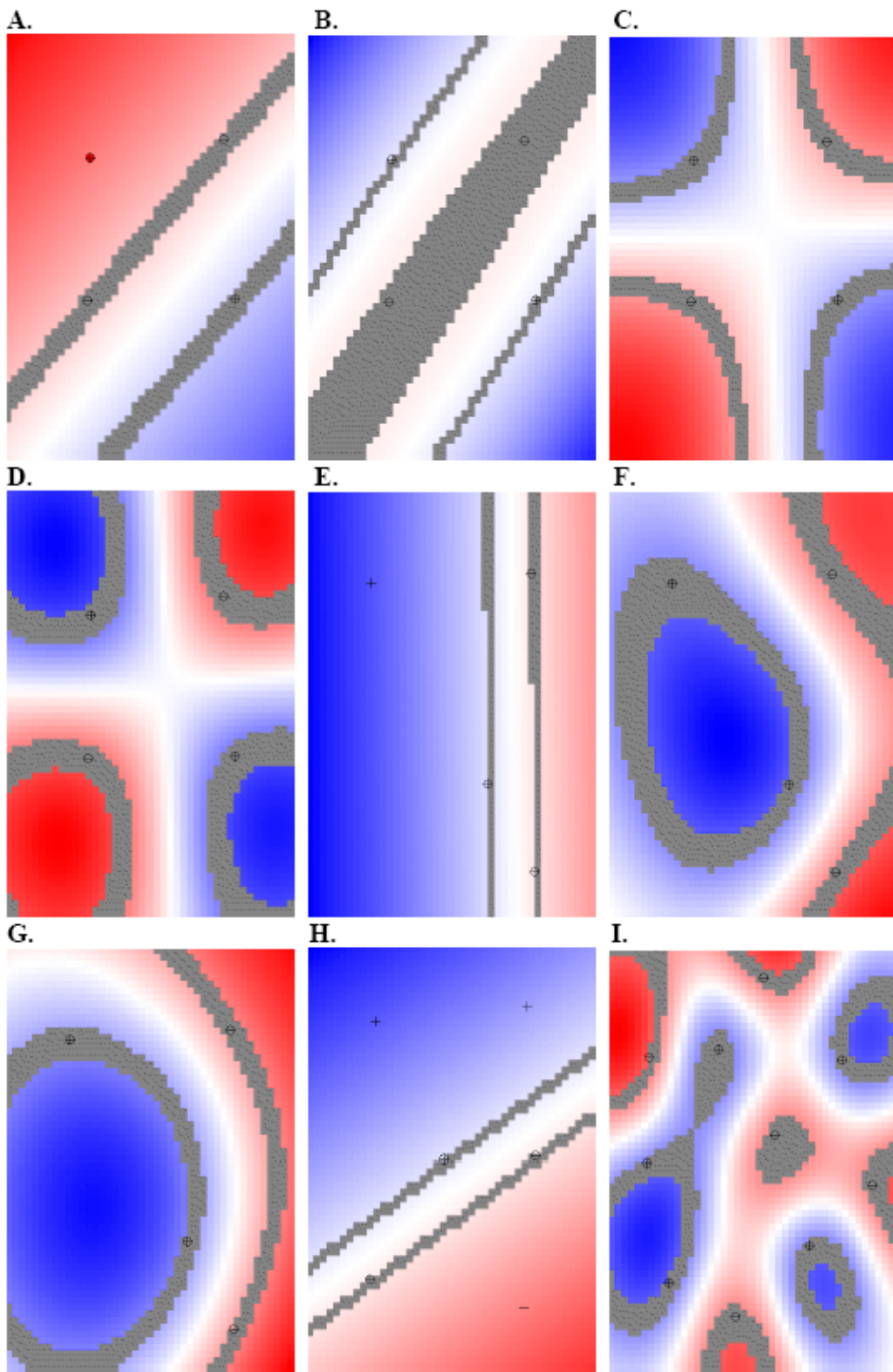
- 以上都是。
- 以上都不是。

在基于先例的学习(即麦克白与杯子的学习例子):

- 判别树决定哪些特征和分类相关。
- 老师提供解释, 学生用这些解释构成似是而非。
- 学习是解决老师提出的问题的产物。
- 通过使用普遍文法来消除表达的问题。
- 以上都是。
- 以上都不是。

Kanwisher对fMRI方法的研究表明fMRI

- 由于X射线的密度, 给目标带来了威胁。
- 在大脑中定位了许多识别中心。
- 能使瞬时分辨率达到毫秒级。
- 证明可被音乐激活的中心的存在。
- 以上都是。
- 以上都不是。



n	-n log n	-n log n - (1-n) log (1-n)	n	-n log n	-n log n - (1-n) log (1-n)
0.00	0.00				
0.01	0.07	0.08	0.51	0.50	1.00
0.02	0.11	0.14	0.52	0.49	1.00
0.03	0.15	0.19	0.53	0.49	1.00
0.04	0.19	0.24	0.54	0.48	1.00
0.05	0.22	0.29	0.55	0.47	0.99
0.06	0.24	0.33	0.56	0.47	0.99
0.07	0.27	0.37	0.57	0.46	0.99
0.08	0.29	0.40	0.58	0.46	0.98
0.09	0.31	0.44	0.59	0.45	0.98
0.10	0.33	0.47	0.60	0.44	0.97
0.11	0.35	0.50	0.61	0.44	0.96
0.12	0.37	0.53	0.62	0.43	0.96
0.13	0.38	0.56	0.63	0.42	0.95
0.14	0.40	0.58	0.64	0.41	0.94
0.15	0.41	0.61	0.65	0.40	0.93
0.16	0.42	0.63	0.66	0.40	0.92
0.17	0.43	0.66	0.67	0.39	0.91
0.18	0.45	0.68	0.68	0.38	0.90
0.19	0.46	0.70	0.69	0.37	0.89
0.20	0.46	0.72	0.70	0.36	0.88
0.21	0.47	0.74	0.71	0.35	0.87
0.22	0.48	0.76	0.72	0.34	0.86
0.23	0.49	0.78	0.73	0.33	0.84
0.24	0.49	0.80	0.74	0.32	0.83
0.25	0.50	0.81	0.75	0.31	0.81
0.26	0.51	0.83	0.76	0.30	0.80
0.27	0.51	0.84	0.77	0.29	0.78
0.28	0.51	0.86	0.78	0.28	0.76
0.29	0.52	0.87	0.79	0.27	0.74
0.30	0.52	0.88	0.80	0.26	0.72
0.31	0.52	0.89	0.81	0.25	0.70
0.32	0.53	0.90	0.82	0.23	0.68
0.33	0.53	0.91	0.83	0.22	0.66
0.34	0.53	0.92	0.84	0.21	0.63
0.35	0.53	0.93	0.85	0.20	0.61
0.36	0.53	0.94	0.86	0.19	0.58
0.37	0.53	0.95	0.87	0.17	0.56
0.38	0.53	0.96	0.88	0.16	0.53
0.39	0.53	0.96	0.89	0.15	0.50
0.40	0.53	0.97	0.90	0.14	0.47
0.41	0.53	0.98	0.91	0.12	0.44
0.42	0.53	0.98	0.92	0.11	0.40
0.43	0.52	0.99	0.93	0.10	0.37
0.44	0.52	0.99	0.94	0.08	0.33
0.45	0.52	0.99	0.95	0.07	0.29
0.46	0.52	1.00	0.96	0.06	0.24
0.47	0.51	1.00	0.97	0.04	0.19
0.48	0.51	1.00	0.98	0.03	0.14
0.49	0.50	1.00	0.99	0.01	0.08
0.50	0.50	1.00	1.00	0.00	0.00