

6.034 Quiz 4

December 5, 2007

Name	
E-Mail	

Circle your TA and recitation time, if any, so that we can enter your score in our records and return your quiz to you promptly.

TAs	
Mike Klein	
Tom Kollar	
Akua Nti	
Mark Seiffter	
Roman Helvetica	

Thu	
Time	Instructor
11-12	Kimberle Koile
12-1	Kimberle Koile
1-2	Kimberle Koile

Fri	
Time	Instructor
12-1	Howard Shrobe
1-2	Howard Shrobe
2-3	Howard Shrobe

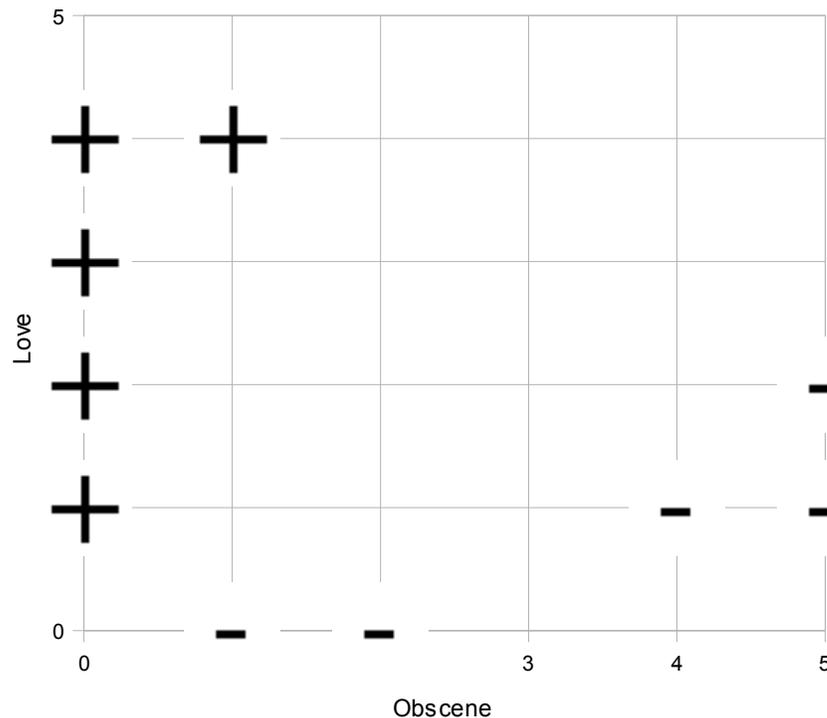
Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

There are 10 pages in this quiz, including this one. A tear-off sheet is provided at the end with duplicate drawings and data.

As always, open book, open notes, open just about everything.

Question 1: Support Vector Machines (50 points)

Sick of hearing music that you don't like, you decide to train a support vector machine classifier to determine whether you will like a new song; that way you can write a program to randomly play music that you will like. You train your program on a number of songs you hear on the radio, as well as songs from your personal music collection. Two of the features you give to the SVM are the number of occurrences of the word “love” and the number of occurrences of obscene words. The data points are shown below:



The points labeled '+' are the songs you like and the points labeled '-' are the songs you don't like.

Part A1 (12 points)

A1a: You do your first analysis using a linear kernel $K(x, x') = x \cdot x'$. On the above plot, draw the decision boundary found by the SVM.

A1b: Also, circle the support vectors.

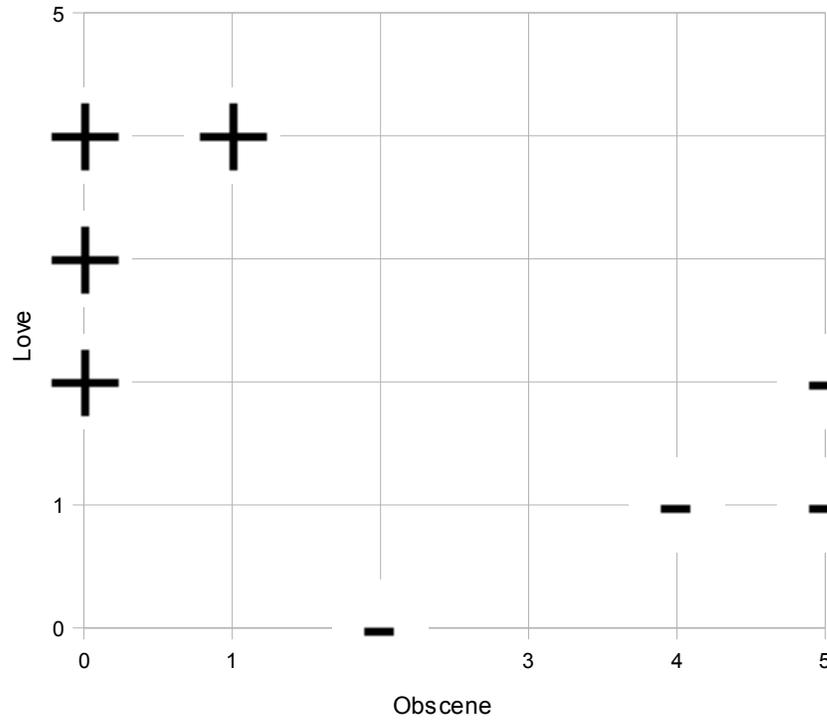
What is the final classifier. Your answer is to be an arithmetic expression in terms of x (obscene word axis) and y (love axis) and constants.

What are the weights α for the support vectors?

What is the weight, α , for the vector $[1, 4]$?

Part A2 (9 points)

Now you decide to eliminate two of the sample vectors as shown below.



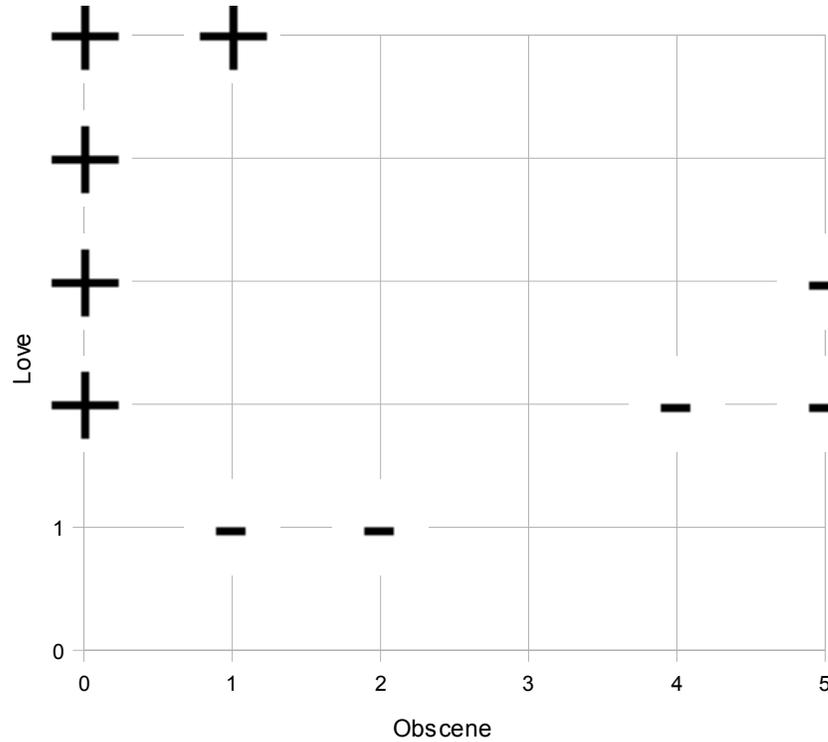
Circle the support vectors.

Again, you use a linear Kernel. What is the final classifier? Your answer is to be an arithmetic expression in terms of x (obscene word axis) and y (love axis) and constants.

What are the weights α for the support vectors?

Part A3 (9 points)

Now you decide to restore the two vectors eliminated in part B, but you discover a measurement error, so your data is really as shown below.



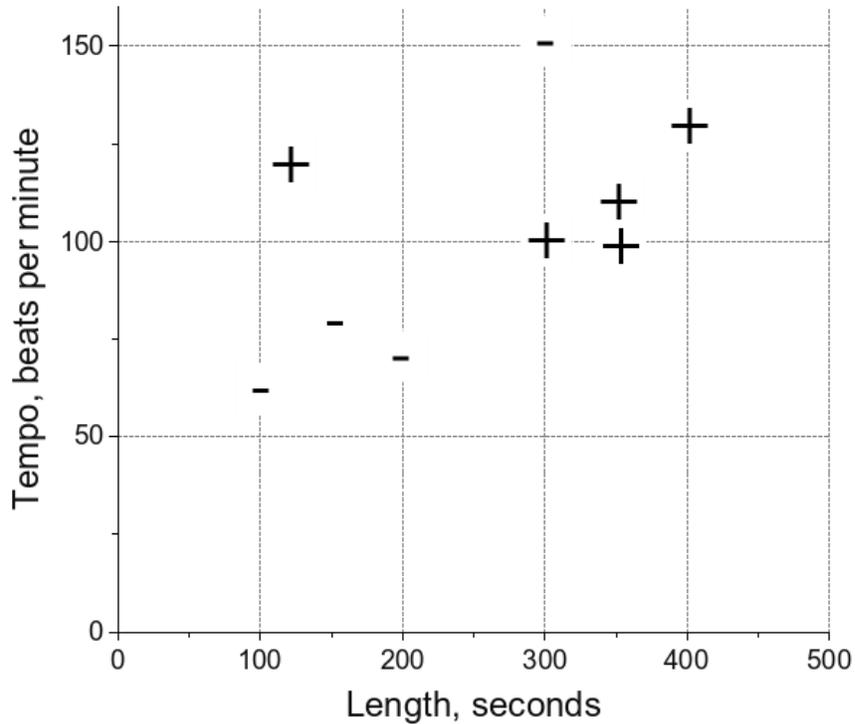
Circle the support vectors.

Again, you use a linear Kernel. What is the final classifier? Your answer is to be an arithmetic expression in terms of x (obscene word axis) and y (love axis) and constants.

What are the weights α for the support vectors?

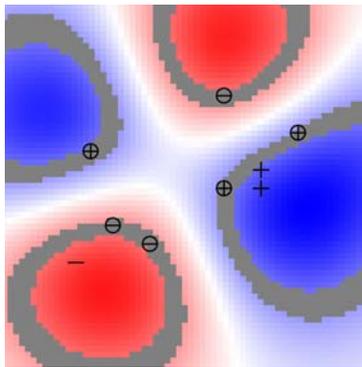
Part B (15 points)

You decide to try an alternative approach with the following data:

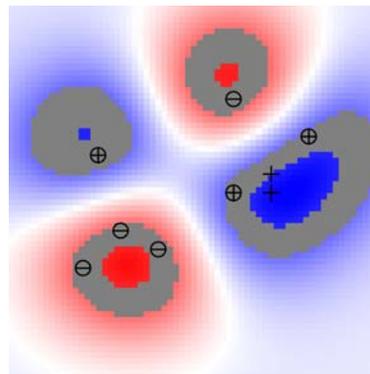


To get a better understanding of how various kernels handle your new approach, you try all ten provided in the demonstration program. You discover, interestingly, that **6 of the 10 produce results that are just about indistinguishable, so only five different pictures are shown in the accompanying diagram.** Fill in the table below with the appropriate diagram label from the separate sheet containing SVM outputs. **Note that 6 of your answers in the 10 boxes will be the same and 4 will be different from all the others.**

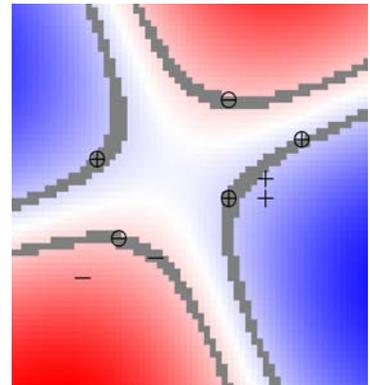
$(V1 \cdot V2 + 1)$		$e^{\frac{-\ V1 - V2\ ^2}{2.0}}$	
$(V1 \cdot V2 + 1)^2$		$e^{\frac{-\ V1 - V2\ ^2}{0.5}}$	
$(V1 \cdot V2 + 1)^3$		$e^{\frac{-\ V1 - V2\ ^2}{0.22}}$	
$(V1 \cdot V2 + 1)^4$		$e^{\frac{-\ V1 - V2\ ^2}{0.12}}$	
$(V1 \cdot V2 + 1)^5$		$e^{\frac{-\ V1 - V2\ ^2}{0.06}}$	



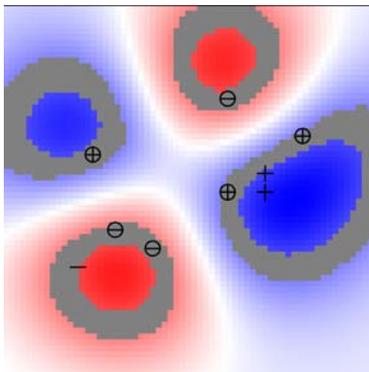
A



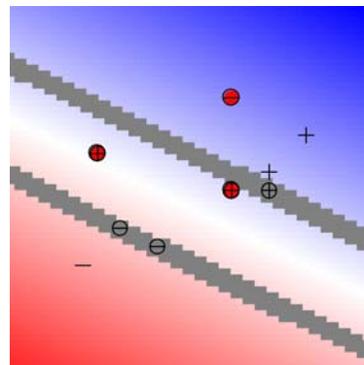
B



C



D



E

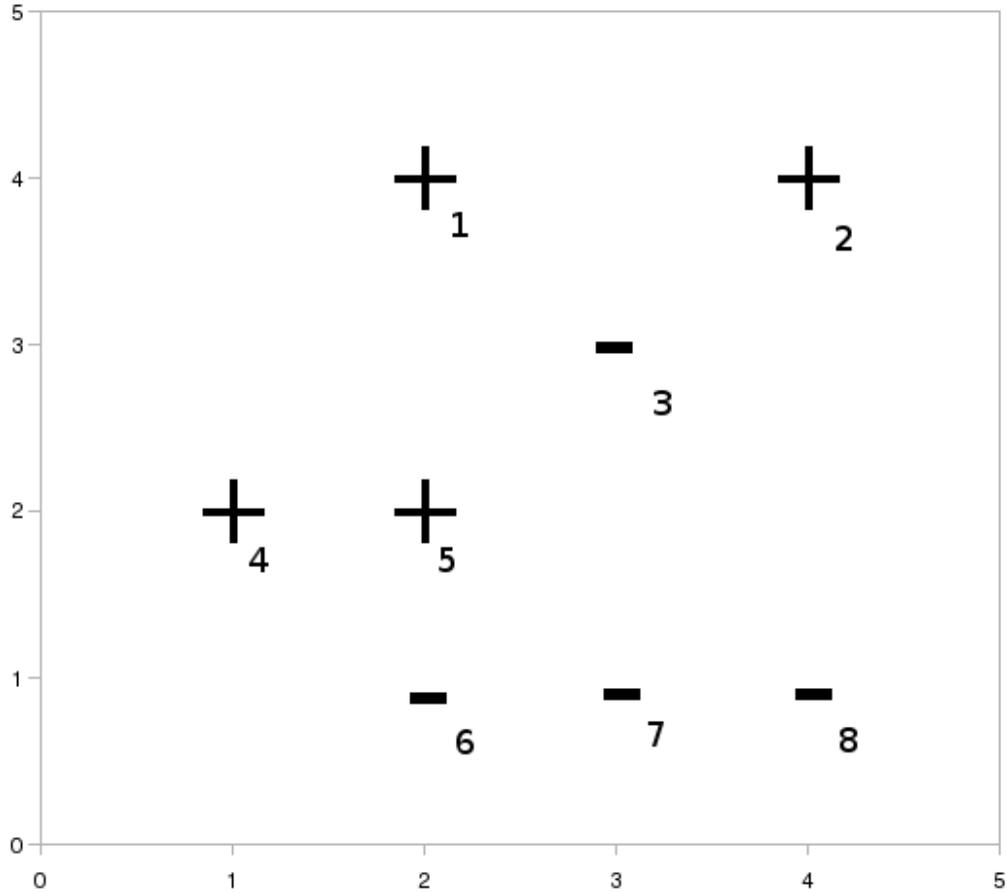
Part C (5 points)

Of the kernels tried in part B, one does not succeed in properly classifying all the examples. Among the rest, identify the kernel that is likely to provide the worst results when applied to test data:

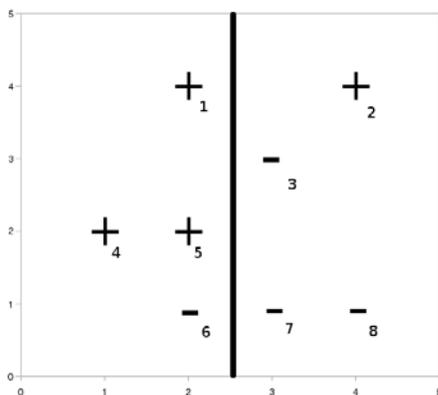
Why?

Question 2 Boosting (50 points)

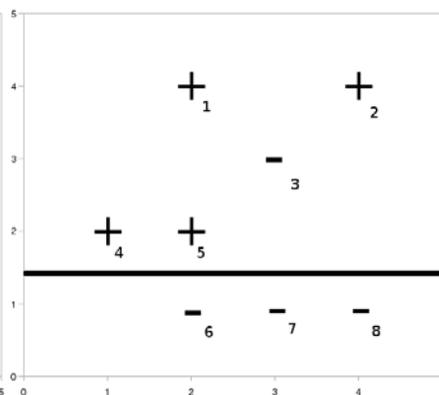
In this problem you will use Adaboost to classify this data:



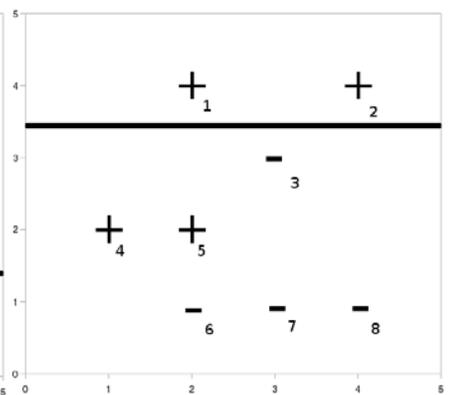
In all parts of the problem, use these stumps as possible classifiers.
When choosing classifiers, break ties in order.



h1: $x < 2.5$



h2: $y > 1.5$



h3: $y > 3.5$

The weights start evenly distributed:

$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$
w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8

Part A, Step 1 (10 points)

Circle the first classifier that Adaboost picks.

h1

h2

h3

What is the first classifier's α ?

What are the new weights?

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8

Part B, Step 2: (10 points)

Circle the second classifier that Adaboost picks.

h1

h2

h3

What is the second classifier's α ?

What are the new weights?

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8

Part C, Step 3: (10 points)

Circle the third classifier that Adaboost picks.

h1

h2

h3

What is the third classifier's α ?

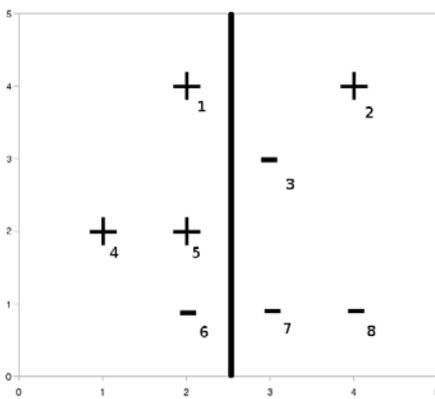
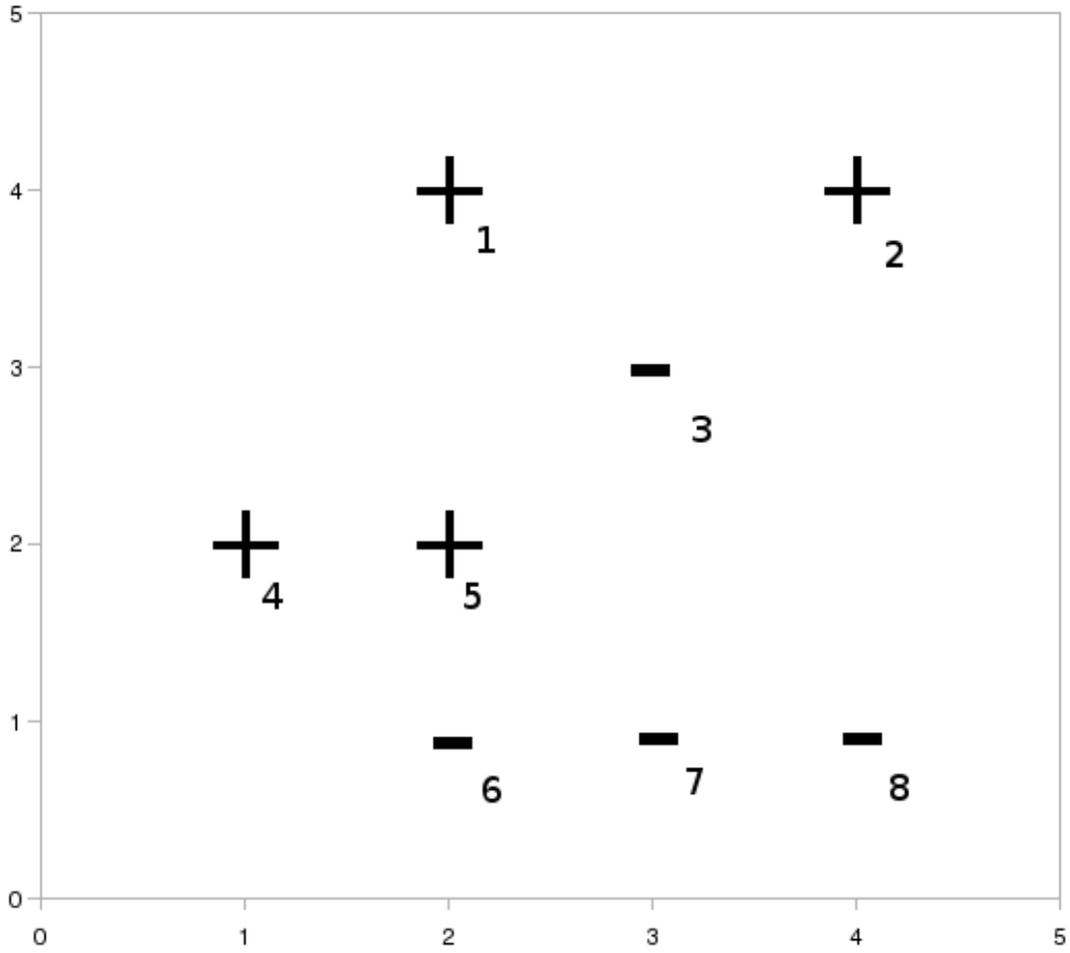
Part D (10 points)

What is the error rate of the Adaboost classifier made up of the above three classifiers?

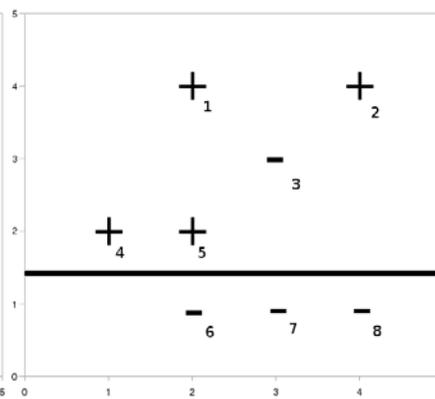
Part E (10 points)

What is the error rate for a *simple vote* among the three classifiers?
(Like Adaboost with all α 's equal.)

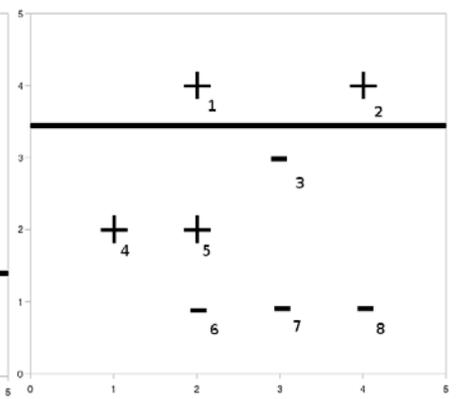
Tear off sheet—you need not hand this in.



h1: $x < 2.5$

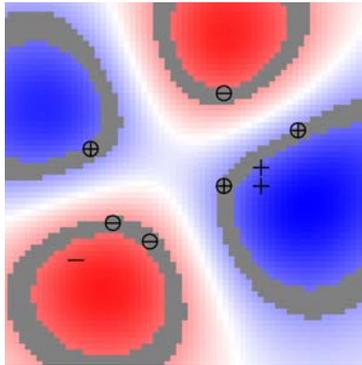


h2: $y > 1.5$

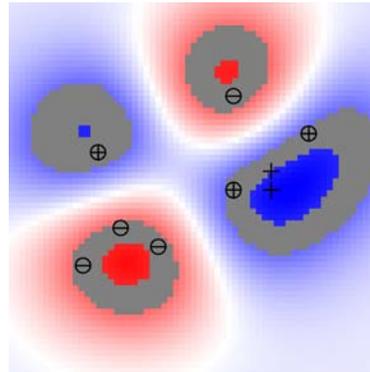


h3: $y > 3.5$

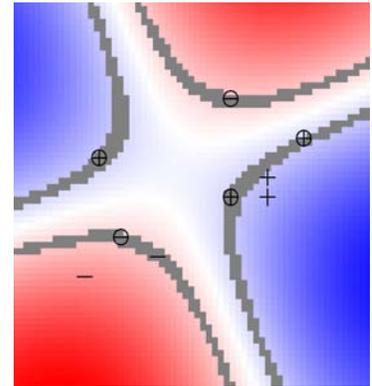
Tear off sheet—you need not hand this in.



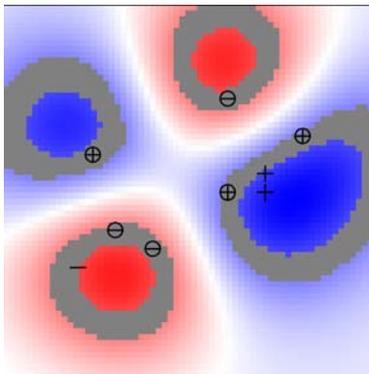
A



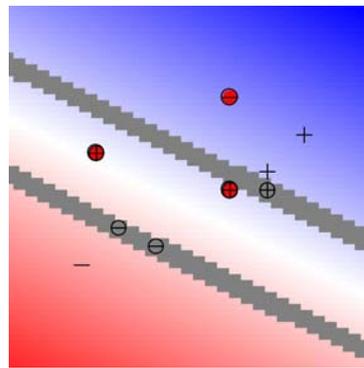
B



C



D



E

Table of natural logs

<i>n</i>	<i>ln(n)</i>
0.25	-1.39
0.50	-0.69
0.75	-0.29
1.00	0.00
1.25	0.22
1.50	0.41
1.75	0.56
2.00	0.69
2.25	0.81
2.50	0.92
2.75	1.01
3.00	1.10
3.25	1.18
3.50	1.25
3.75	1.32
4.00	1.39
4.25	1.45
4.50	1.50
4.75	1.56
5.00	1.61
5.25	1.66
5.50	1.70
5.75	1.75
6.00	1.79
6.25	1.83
6.50	1.87
6.75	1.91
7.00	1.95
7.25	1.98
7.50	2.01
7.75	2.05
8.00	2.08
8.25	2.11
8.50	2.14
8.75	2.17
9.00	2.20
9.25	2.22
9.50	2.25
9.75	2.28
10.00	2.30

<i>n</i>	<i>ln(n)</i>
10.25	2.33
10.50	2.35
10.75	2.37
11.00	2.40
11.25	2.42
11.50	2.44
11.75	2.46
12.00	2.48
12.25	2.51
12.50	2.53
12.75	2.55
13.00	2.56
13.25	2.58
13.50	2.60
13.75	2.62
14.00	2.64
14.25	2.66
14.50	2.67
14.75	2.69
15.00	2.71
15.25	2.72
15.50	2.74
15.75	2.76
16.00	2.77
16.25	2.79
16.50	2.80
16.75	2.82
17.00	2.83
17.25	2.85
17.50	2.86
17.75	2.88
18.00	2.89
18.25	2.90
18.50	2.92
18.75	2.93
19.00	2.94
19.25	2.96
19.50	2.97
19.75	2.98
20.00	3.00

MIT OpenCourseWare
<http://ocw.mit.edu>

6.034 Artificial Intelligence
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.