

## Homework 3

### Instructions

This homework contains **6** concepts and **9** programming questions. In MS word or a similar text editor, write down the problem number and your answer for each problem. Combine all answers for concept questions in a single PDF file. Export/print the Jupyter notebook as a PDF file including the code you implemented and the outputs of the program. Make sure all plots and outputs are visible in the PDF.

Combine all answers into a single PDF named `andrewID_hw3.pdf` and submit it to Gradescope before the due date. Refer to the syllabus for late homework policy. Please assign each question a page by using the “Assign Questions and Pages” feature in Gradescope. Submission to anywhere else than Gradescope will not be graded.

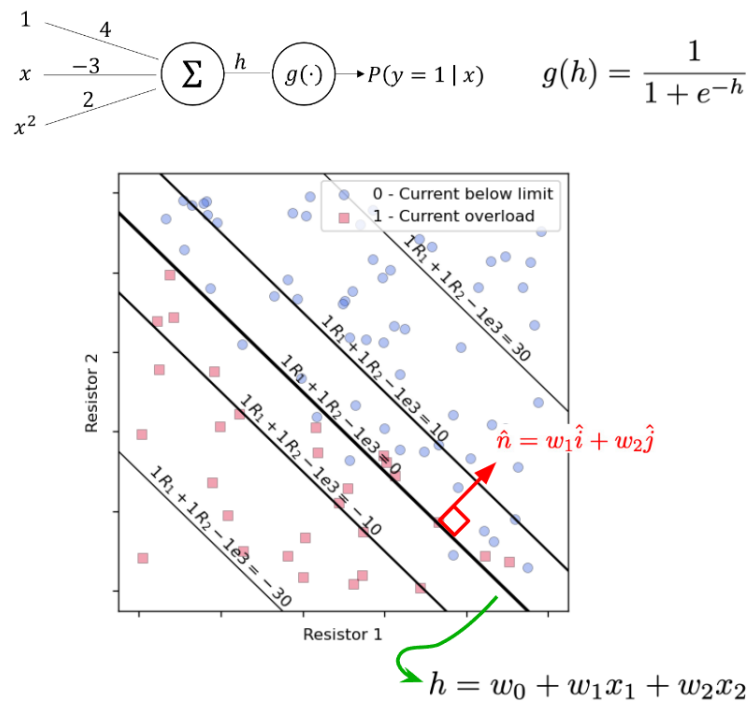
Question	Points
Concept 1	3
Concept 2	2
Concept 3	1
Concept 4	2
Concept 5	1
Concept 6	3
M3-L1-P1	6
M3-L1-P2	6
M3-L1-P3	6
M3-L2-P1	6
M3-L2-P2	6
M3-L2-P3	6
M3-HW1	24
M3-HW2	24
M3-HW3	24
<b>Total</b>	<b>120</b>
Bonus	6

**Problem 1**

The sigmoid function is useful because

1. It restricts the output between -1 and 1 (T/F)
2. It has a probabilistic interpretation (T/F)
3. It is easily differentiable (T/F)

## Problem 2



Consider the level sets that correspond to different decision boundaries in the figure.

1) What are the bounds on the values of  $h$ ?

2) What are the bounds on the values of  $g(h)$  where  $g()$  is the sigmoid function?

**Problem 3**

More L2 regularization always leads to better fitting models. (T/F)

**Problem 4**

Consider the following 4 class problem. A given test point  $x$  is evaluated by six binary classifiers with the following results:

A vs. B  $\rightarrow$  class A

A vs. C  $\rightarrow$  class C

A vs. D  $\rightarrow$  class D

B vs. C  $\rightarrow$  class C

B vs. D  $\rightarrow$  class D

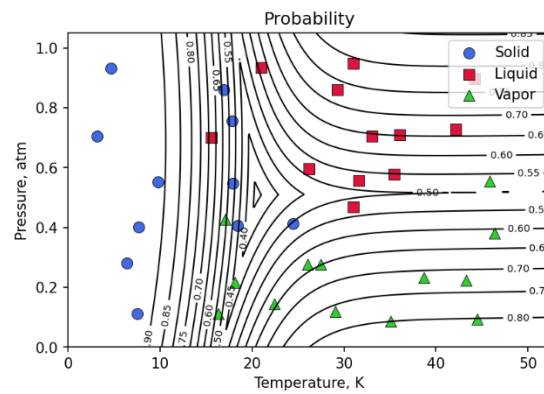
C vs. D  $\rightarrow$  class D

What is the predicted class for the test point?

**Problem 5**

For what number of classes is the number of classifiers required for one-versus-one and one-versus-rest classifiers equal?

## Problem 6



Point 1	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	Models	$\begin{bmatrix} 0.7 \\ 0.2 \\ 0.1 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 0.4 \\ 0.1 \end{bmatrix}$
Point 2	$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$		$\begin{bmatrix} 0.1 \\ 0.8 \\ 0.1 \end{bmatrix}$	$\begin{bmatrix} 0.6 \\ 0.3 \\ 0.1 \end{bmatrix}$
Point 3	$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$		$\begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix}$	$\begin{bmatrix} 0.2 \\ 0.6 \\ 0.2 \end{bmatrix}$
Point 4	$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$		$\begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix}$	$\begin{bmatrix} 0.3 \\ 0.4 \\ 0.3 \end{bmatrix}$
	Ground Truth		Model 1 Prediction	Model 2 Prediction

Consider the phase problem from the slides. There are three classes: solid, liquid, and vapor. We have four test points with ground truth labels shown below. We train two models that output the predictions below. By inspection, which model is best?