**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 |
| **Total** | **120** |
| Bonus | 6 |

**Problem 1**

The sigmoid function is useful because

1. It restricts the output between -1 and 1 (T/F)

F

1. It has a probabilistic interpretation (T/F)

T

1. It is easily differentiable (T/F)

T

**Problem 2**

A diagram of a graph

Description automatically generated

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

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

The bounds of h is (-30,-10,0,10,30)

1. What are the bounds on the values of *g(h)* where *g()* is the sigmoid function?

The bounds of g(h) is (-0.999,-0.999,0.5,0.999,0.999)

**Problem 3**

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

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 → class A

A vs. C → class C

A vs. D → class D

B vs. C → class C

B vs. D → class D

C vs. D → class D

What is the predicted class for the test point?

D, as it the dominating class comparing to other classes.

**Problem 5**

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

For 1 v 1, the number of classifiers required is C(n,2), where n is the number o classes.

For 1 v rest, the nnumer of classifiers required is just n, the number of classes.

**Problem 6**

A screenshot of a computer

Description automatically generated

A diagram of a model

Description automatically generated

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?

Model 1 is the best as its prediction perfectly corresponds to the ground truth, which gives very high accuracy.

Model2 is not as good because it has incorrect predictions .