Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE 100%

Regression-based Classification

LATEST SUBMISSION GRADE

100%

1. What does it mean to have a matching loss? 1 / 1 point

- ○ The loss function and model function are both linear
- ○ The loss function looks good in a graph.
- ◉ The loss function used for optimization is convex with respect to the model function.
- ○ The loss function and model function share the same hypothesis space

✓ Correct

Correct! When minimizing loss with respect to the model function's parameters is convex, we are using a matching loss.

2. Why do we need transfer functions for classification? 1 / 1 point

- ○ Because every step of the classification model needs to be differentiable
- ◉ Because deciding how to translate the output of regression into the space of class labels deserves particular consideration
- ○ Because a number is not a class
- ○ Because we said so

✓ Correct

Yes! We can choose different thresholds or different kinds of functions for the translation from number to class label, and so it's important to make that explicit.

3. What is common between the regression-based classification algorithms in this module? Specifically Logistic Regression, Neural Networks, and SVMs. Select all that apply. 2 / 2 points

- ☑ They all output class labels

✓ Correct

Yes, that is the definition of classification

- ☑ They all use optimization

✓ Correct

Correct

- ☐ They all use the L2 norm
- ☐ They all use a regularized loss function
- ☑ They all use linear functions

✓ Correct

Yes! All three have at least one portion that is a linear function.

4. Which of the following is the **best** use case for SVMs? 1 / 1 point

- ◉ Character recognition using 16x16 scans of characters, with one thousand training examples.
- ○ Earthquake detection with 100 time-series features and examples from two stations.
- ○ Sentiment analysis using a dataset of one million training examples, each consisting of five features.
- ○ Classifying mushrooms as poisonous or edible, using a dataset of one thousand examples, with the GPS coordinates each was found in, the shape of the cap, colour, and density.

✓ Correct

Correct! This is not too large a dataset and it is already scaled appropriately. In addition, it's quite likely that the pixel space will benefit from non-linear kernels.