

EECS 349 (Machine Learning) Homework 8

How to submit your homework

1. Create a **PDF document** containing answers to the homework questions.
2. Include source code for the program you write.
3. Compress all of the files specified into a .zip file.
4. Name the file in the following manner, *firstname_lastname_hw7.zip*.
3. Submit this .zip file via Canvas by the date specified on Canvas.

1) Boosting (3 points)

The AdaBoost algorithm is described in the paper “A Brief Introduction to Boosting” by Robert Schapire. You have been provided this paper on the course website. This algorithm learns from a training set of input/output instances $\{(x_1, y_1), \dots, (x_m, y_m)\}$ where X is an arbitrary set of inputs, $x_i \in X$ and $y_i \in \{-1, +1\}$. AdaBoost assumes the training set is fixed over all the rounds.

Assume now that after each round, each pair in the training set has a probability q of reversing the sign of its label (the y term in the pair is the label).

A. (1 point) How would the performance of AdaBoost be affected by a dataset that changes from round to round in this manner?

B. (1 point) Assume that the number of rounds is fixed at some value T . Modify AdaBoost to improve its expected performance on the training set immediately after round T . Describe your modification in both text and mathematical notation.

C. (1 point) Give an informal argument for why your modification would do better than the unmodified AdaBoost. Illustrate with a simple example that you’ve worked through.

2) Active Learning (3 points)

Read the paper “Improving Generalization with Active Learning” (there is a link on the course calendar) and answer the following questions.

A.(1 point): Give a definition of active learning. Give a reason why someone would want to use active learning. Motivate this with an example problem that illustrates the issue.

B. (1 point) Explain how to use the set of hypotheses in a version space to select the next query for active learning. What is the name of this kind of query selection?

C. (1 point) For an SG neural network on the 25-bit threshold problem, what function appears to govern the relationship between the number of samples and the error rate of the learner? How does this compare to using random sample selection?

3) You don’t need that fancy classifier (2 points)

Read “Classifier Technology and the Illusion of Progress.” Explain its thesis. What evidence do the authors use to back up their argument. Do you agree with their argument? Why or why not?

4) Dropout (2 points)

Read the paper “Dropout: A Simple Way to Prevent Overfitting.” Then, answer the following questions.

A.(1/2 point): Explain what Dropout is. Give a high-level explanation of how it works.

B.(1/2 point): Explain what model combination is and say why it is a good idea.

EECS 349 (Machine Learning) Homework 8

- C. (1/2 point) What evidence is there that Dropout improves generalization? Give details.
- D. (1/2 point) What are the advantages/disadvantages of Dropout compared to Bayesian Neural Nets?