

# Assignment Seven

## ECE 4200

October 21, 2020

- Provide credit to **any sources** other than the course staff that helped you solve the problems. This includes **all students** you talked to regarding the problems.
- You can look up definitions/basics online (e.g., wikipedia, stack-exchange, etc)
- **The due date is 11/1/2020, 23.59.59 eastern time.**
- Submission rules are the same as previous assignments.

**Problem 1. (10 points).** Suppose AdaBoost is run on  $n$  training examples, and suppose on each round that the weighted training error  $\varepsilon_t$  of the  $t$ th weak hypothesis is at most  $\frac{1}{2} - \gamma$ , for some number  $\gamma > 0$ . Show that after  $T > \frac{\ln n}{2\gamma^2}$  rounds of AdaBoost the final combined classifier has zero training error!

**Problem 2. (10 points).** Recall bagging. Starting from a training set  $S$  of size  $n$ , we created  $m$  bootstrap training sets  $S_1, \dots, S_m$ , each of size  $n$  each by sampling with replacement from  $S$ .

1. For a bootstrap sample  $S_i$ , what is the expected fraction of the training set that does not appear at all in  $S_i$ ? As  $n \rightarrow \infty$ , what does this fraction approach?
2. Let  $m > 2 \ln n$ , and  $n \rightarrow \infty$ . Show that the expected number of training examples in  $S$  that appear in at least one  $S_i$  is more than  $n - 1$ .

**Problem 3. (10 points).** The tanh function is  $\tanh(y) = (e^y - e^{-y})/(e^y + e^{-y})$ . Consider the function  $\tanh(w_0 + w_1x_1 + w_2x_2)$ , with five inputs, and a scalar output.

1. Draw the computational graph of the function (you can use tanh in your computation graph).
2. What is the derivative of  $\tanh(y)$  with respect to  $y$ .
3. Suppose  $(w_0, w_1, w_2, x_1, x_2) = (-2, -3, 1, 2, 3)$ . Compute the forward function values, and back-propagation of gradients.

**Problem 4. (30 points).** Please see attached notebook for details.