ECE 4200 - Assignment 6

Name: Zihao Lin

NetID: zl755

Problem #1:

Scenario where one example is incorrect:

$$\text{error} = \frac{1}{n} \sum \mathbb{I}\{H_T(\vec{x}_i) \neq y_i\}$$

$$= \frac{1}{n}$$

Only one is incorrect

We want final error smaller than 1/n

given $\epsilon_t \leq \frac{1}{2} - \gamma$, the training error is bounded by:

$$err(H_T) = \frac{1}{n} \sum_i \mathbb{I}\{H_T(\vec{x}_i) \neq y_i\}$$

$$\leq \frac{1}{n} \sum \exp\{-y_i(\alpha_T h_T(\vec{x}_i) + \cdots + \alpha_1 h_1(\vec{x}_i))\}$$

$$= Z_1 Z_2 \cdots Z_T$$

$$Z_T = 2\sqrt{\epsilon_t(1-\epsilon_t)} \leq 2\sqrt{\left(\tfrac{1}{2}-\gamma\right)\left(1-\left(\tfrac{1}{2}-\gamma\right)\right)} \quad \text{using } \epsilon_t \leq \tfrac{1}{2}-\gamma$$

$$= 2\sqrt{\left(\tfrac{1}{2}-\gamma\right)\left(\tfrac{1}{2}+\gamma\right)}$$

$$= 2\sqrt{\tfrac{1}{4}-\gamma^2} \leq \exp(-2\gamma^2)$$

$$err(H_T) \leq Z_1 \cdots Z_T = \exp(-2\gamma^2 T)$$

upper bound is $< 1/m$ if $T > \ln n / 2\gamma^2$

pf: $\quad err(H_T)\Big|_{T > \ln n / 2\gamma^2} < \exp\left(-2\gamma^2 \cdot \ln n / 2\gamma^2\right)$

$$= \exp(-\ln n) = 1/n$$

$\therefore$ QED

Problem 2 :

1. $S: \{x_1, \ldots x_n\}$.

$\downarrow$

$S_1 : \{x'_1, \ldots x_n\}$

S_2 $\vdots$

$\vdots$

S_m

With replacement

Each data point has probability $(1-1/n)^n$ of being selected as test data $\rightarrow$ not selected as training data

Thus, training data has $1-(1-1/n)^n$ of the original data. $(1-1/n)^n$ of the set doesn't appear at all in S_i.

As $n \to \infty$:

$$\alpha = \lim_{n \to \infty} (1-1/n)^n = e^{-1} = \boxed{0.368}$$

$\therefore$ As $n \to \infty$, roughly 36.8% of the set will not appear.

2. Median Trick:

$$S \to \left\{ \begin{array}{c} S_1 \\ \vdots \\ S_m \end{array} \right\}$$

Probability that example in S appears in at least one $S_i \to (1 - P(\text{no appearance in all } S_i))\times n$,

$$P(\text{at least one occurrence in all } S_i) = 1 - \left[(1-1/n)^n\right]^m$$

$$P(\text{at least one}) < 1 - ((1-1/n)^n)^{2\ln n}$$

# of training examples appear at least one in S_i : bounded by $n(1 - ((1-1/n)^n)^{2\ln n})$

$$= n(1 - (1-1/n)^{2n\ln n})$$

$$= n - \underbrace{n(1-1/n)^{2n\ln n}}_{\leq 1}$$

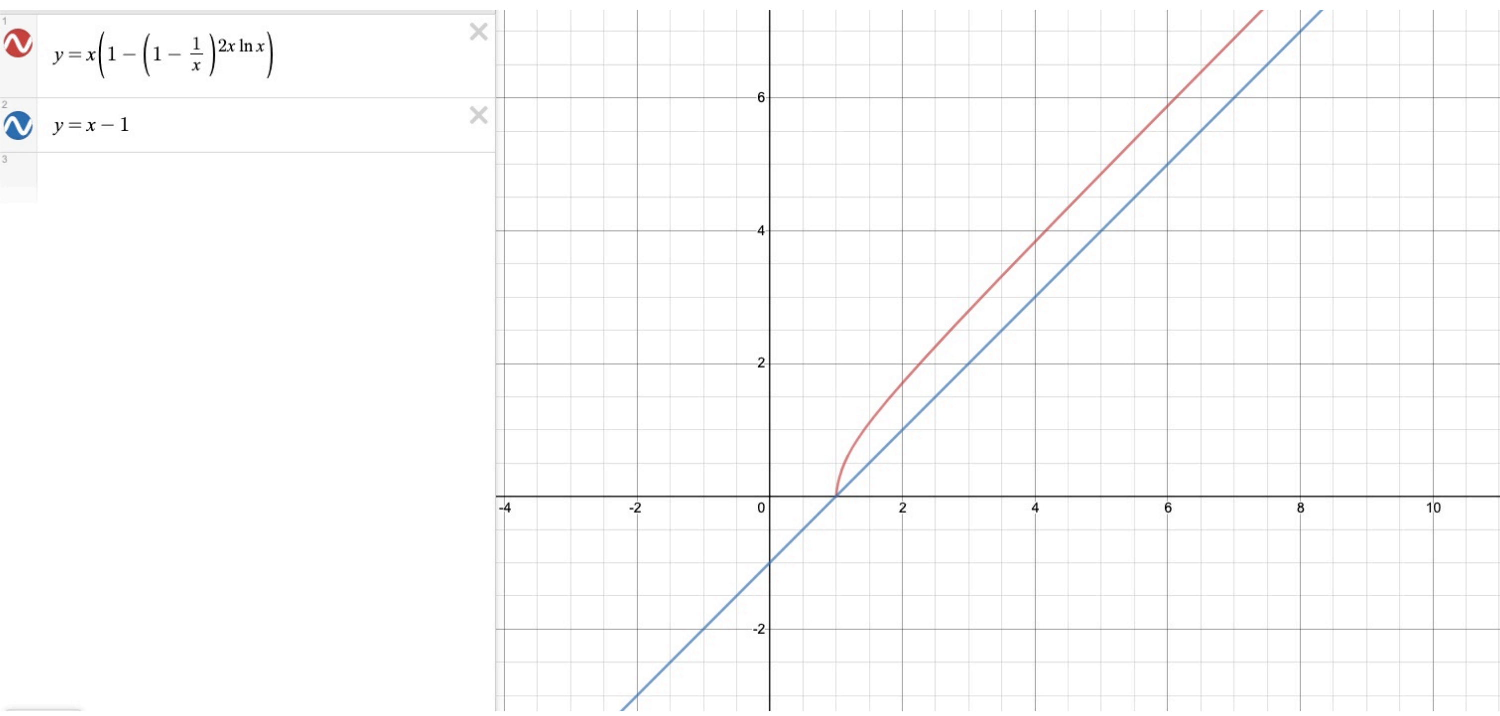
$$\geq n-1$$

∴ QED.

plot of $n(1-((1-1/n)^n)^{n\ln n})$

shown on next page

(6)



$y = x\left(1 - \left(1 - \frac{1}{x}\right)^{2x\ln x}\right)$
$y = x - 1$
6
4
2
-2
-4
-2
0
2
4
6
8
10

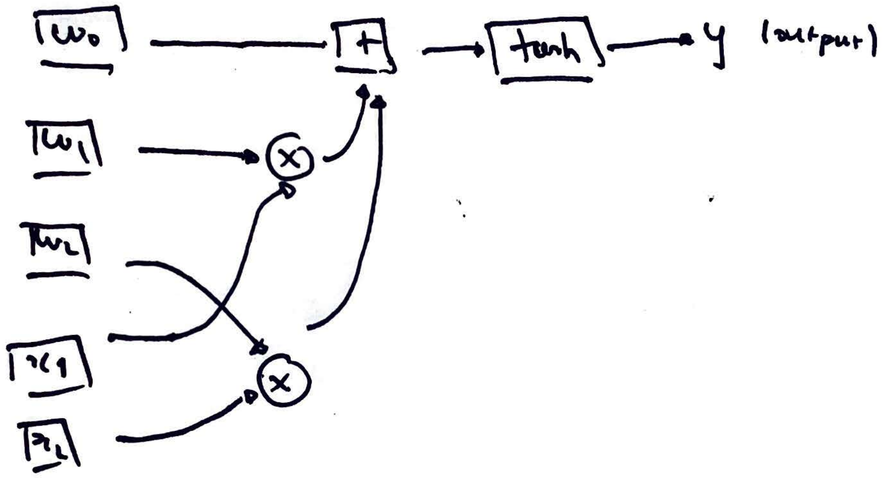
Problem #3:

1. $\tanh(y) = (e^{y} - e^{-y}) / (e^{y} + e^{-y})$

$$y = \tanh(\underbrace{w_0 + w_1 x_1 + w_2 x_2})$$

output

input = $[w_0, w_1, w_2, x_1, x_2]$.



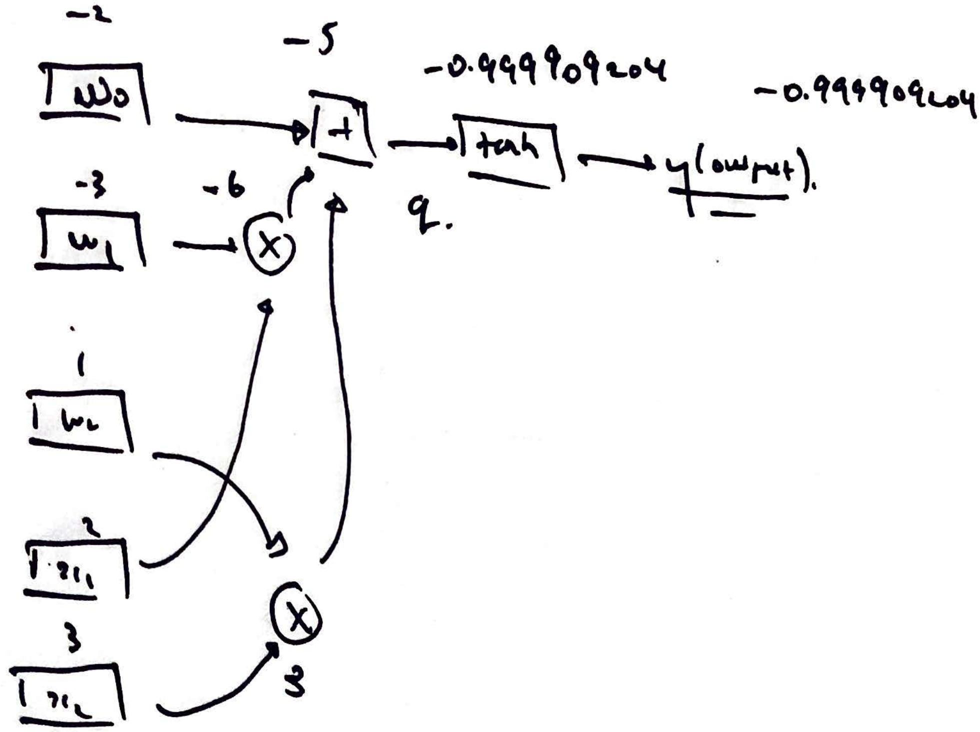
2. $$\frac{d\tanh(y)}{dy} = \frac{d}{dy}\left[\frac{e^{y} - e^{-y}}{e^{y} + e^{-y}}\right]$$

$$= \frac{(e^{y}+e^{-y})(e^{y}+e^{-y}) - (e^{y}-e^{-y})(e^{y}-e^{-y})}{(e^{y}+e^{-y})^2}$$

$$= \frac{(e^{y}+e^{-y})^2 - (e^{y}-e^{-y})^2}{(e^{y}+e^{-y})^2}$$

$$= 1 - \left(\frac{e^{y}-e^{-y}}{e^{y}+e^{-y}}\right)^2 = \boxed{1 - \tanh(y)^2}$$

3.



$$y = \tanh(q) \quad , \quad q = w_0 + x_1 w_1 + x_2 w_2$$

$$\frac{\partial y}{\partial w_0} = \frac{\partial y}{\partial q} \times \frac{\partial q}{\partial w_0} = (1-\tanh(q)^2) \times 1 = 1-(-0.999909204)^2 = \boxed{0.00018}$$

$$\frac{\partial y}{\partial w_1} = \frac{\partial y}{\partial q} \times \frac{\partial q}{\partial w_1} = (1-\tanh(q)^2) \times x_1 = \boxed{0.000363167}$$

$$\frac{\partial y}{\partial w_2} = \frac{\partial y}{\partial q} \times \frac{\partial q}{\partial w_2} = (1-\tanh(q)^2) \times x_2 = \boxed{0.00054475}$$

$$\frac{\partial y}{\partial x_1} = \frac{\partial y}{\partial q} \times \frac{\partial q}{\partial x_1} = (1-\tanh(q)^2)\, w_1 = \boxed{-0.0005447}$$

$$\frac{\partial y}{\partial x_2} = \frac{\partial y}{\partial q} \times \frac{\partial q}{\partial x_2} = (1-\tanh(q)^2)\, w_2 = \boxed{0.000181087}$$