Kieth Vo   A09632897
Philip Hua   A09102800

Problem 1

1. $\text{err}_{Geser} = Pr(label = x, Geser \neq x)$
$$= \frac{1}{k}\left(\frac{k-1}{k}\right) = \frac{k-1}{k^2}$$

2. $\text{err}_{Zavulon} = \sum Pr(label = x, Zavulon \neq x)$
$$= \sum_{1}^{k} w_i(1 - w_i) = \sum_{1}^{k} w_i - w_i^2$$
$$= 1 - \sum_{1}^{k} w_i^2$$

Problem 2

1. $x = 1, 4$ $Y = 1$ $h(x) = 1, x = 2, 3, 4$
$x = 2, 3$ $Y = 0$ $h(x) = 0, x = 1$

$$\text{err}_{D_1} h(x) = \frac{3}{4}$$

2. $h_4(x) = 1$ if $x > 4$ and 0 o.w.

$x = 1, 4$ $Y = 1$ $h(x) = 1$ $x = 4$
$x = 2, 3$ $Y = 0$ $h(x) = 0$ $x = 1, 2, 3$

$$\text{err}_{D_1} h_4(x) = \frac{1}{4}$$

We do have a non-zero bias when the concept class is C and the data distribution D_1.

3. $\text{err}_{D_2} h(x) = \sum \text{Pr}(Y=0, h(x)=1, x=i) +$
$$\sum \text{Pr}(Y=1, h(x)=0, x=i)$$

$$= (8/10)(1)(1/4) + (7/10)(1)(1/4) + (6/10)(1)(1/4)$$

$$+ (1/10)(1)(1/4) = 22/40 \qquad \colorbox{yellow}{$\text{err}_{D_2} h(x) = 22/40$}$$

$h_4(x) = 1$ if $x > 4$ and 0 o.w.

$$\text{err}_{D_2} h_4(x) = \sum \text{Pr}(Y=1, h(x)=0, x=i)$$

$$= (1/10)(1)(1/4) + (2/10)(1)(1/4)$$
$$+ (3/10)(1)(1/4) + (4/10)(1)(1/4) = 10/40$$

$\colorbox{yellow}{$\text{err}_{D_2} h_4(x) = 10/40$}$

We have a non-zero bias when the concept class is C and distribution D_2.

Problem 3

	Train error	Test error
3	.136	.64
7	.204	.47
10	.556	.23
15	.458	.4033
20	.186	.1567

Words	Train	Test
1	remember	remember
2	lack	france
3	france	lack
4	universite	html
5	monday	linguist
6	editor	title
7	otherwise	england
8	ling	clearly
9	marker	universite
10	linguist	reach