Deep Patel
Intro to AI - HW 5
12/2/18

Problem 1
Level 1:
Counting "played tennis" column from the table, $n^- = 5$, $n^+ = 9$

$H_{initial} = -(\frac{9}{(5+9)} log_2 \frac{9}{(5+9)} + \frac{5}{(5+9)} log_2 \frac{5}{(5+9)})$

$H_{initial} = 0.94028596$

$$H_{outlook} = -[(\frac{(3+2)}{(5+9)}(\frac{2}{(3+2)} log_2 \frac{2}{(3+2)} + \frac{3}{(3+2)} log_2 \frac{3}{(3+2)}) + \frac{(0+4)}{(5+9)}(\frac{4}{(0+4)} log_2 \frac{4}{(0+4)} + \frac{0}{(0+4)} log_2 \frac{0}{(0+4)}) + \frac{(2+3)}{(5+9)}(\frac{3}{(2+3)} log_2 \frac{3}{(2+3)} + \frac{2}{(2+3)} log_2 \frac{2}{(2+3)})] = 0.69353613$$

$$H_{temperature} = -[(\frac{(2+3)}{(5+9)}(\frac{3}{(2+3)} log_2 \frac{3}{(2+3)} + \frac{2}{(2+3)} log_2 \frac{2}{(2+3)}) + \frac{(1+4)}{(5+9)}(\frac{4}{(1+4)} log_2 \frac{4}{(1+4)} + \frac{1}{(1+4)} log_2 \frac{1}{(1+4)}) + \frac{(2+2)}{(5+9)}(\frac{2}{(2+2)} log_2 \frac{2}{(2+2)} + \frac{2}{(2+2)} log_2 \frac{2}{(2+2)})] = 0.89031382$$

$$H_{humidity} = -[(\frac{(4+4)}{(5+9)}(\frac{4}{(4+4)} log_2 \frac{4}{(4+4)} + \frac{4}{(4+4)} log_2 \frac{4}{(4+4)}) + \frac{(1+5)}{(5+9)}(\frac{5}{(1+5)} log_2 \frac{5}{(1+5)} + \frac{1}{(1+5)} log_2 \frac{1}{(1+5)})] = 0.850009609$$

$$H_{wind} = -[(\frac{(2+6)}{(5+9)}(\frac{6}{(2+6)} log_2 \frac{6}{(2+6)} + \frac{2}{(2+6)} log_2 \frac{2}{(2+6)}) + \frac{(3+3)}{(5+9)}(\frac{3}{(3+3)} log_2 \frac{3}{(3+3)} + \frac{3}{(3+3)} log_2 \frac{3}{(3+3)})] = 0.892158928$$

$\Delta H(outlook) = H_{initial} - H_{outlook} = 0.2467498198$

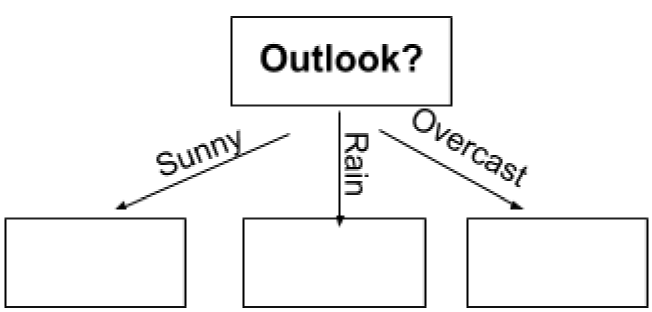
$\Delta H(temperature) = H_{initial} - H_{temperature} = 0.0499721410$

$\Delta H(humidity) = H_{initial} - H_{humidity} = 0.0902763494$

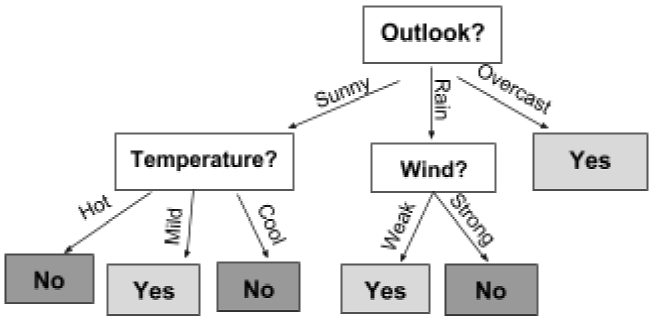
$\Delta H(wind) = H_{initial} - H_{wind} = 0.0481270304$

$Max(\Delta H(outlook), \Delta H(temperature), \Delta H(humidity), \Delta H(wind)) = \Delta H(outlook)$

So we choose the attribute **outlook** for level 1.



To decide what to split with next, we use the same procedure for each of the 3 boxes, but with only the data associated to them (i.e. sunny data, rain data, or overcast data for each of the three splits). Doing this, we get the next level of the decision tree which is actually the final level because all examples have been classified. The final decision tree is below:



Problem 2

$f(x,y) = xy$

$g(x,y) = x^2 + 2y^2 - 6 = 0$

$\Lambda(x, \lambda) = f(x) - \lambda g(x)$

$\quad = xy - \lambda(x^2 + 2y^2 - 6)$

$\partial\Lambda/\partial x = y - \lambda(2x) = 0$

$\partial\Lambda/\partial y = x - \lambda(4y) = 0$

$\partial\Lambda/\partial\lambda = x^2 + 2y^2 - 6 = 0$

Substituting $x = 4\lambda y$ into $y = 2\lambda x$:

$y = 2\lambda(4\lambda y)$

$1 = 8\lambda^2$

$\lambda = \pm \frac{1}{2\sqrt{2}}$

$y = \pm \frac{1}{\sqrt{2}} x$

$x = \pm \frac{2}{\sqrt{2}} y$

Substituting $x = \pm \frac{2}{\sqrt{2}} y$ into $x^2 + 2y^2 - 6 = 0$

$2y^2 + 2y^2 - 6 = 0$

$4y^2 = 6$

$y^2 = 3/2$

$y = \pm \sqrt{\frac{3}{2}}$

Substituting $y = \pm \sqrt{\frac{3}{2}}$ into $x^2 + 2y^2 - 6 = 0$

$x = \pm \frac{2}{\sqrt{2}} (\pm \sqrt{\frac{3}{2}})$

$x = \pm \sqrt{3}$

Solutions in the form (x_0, y_0, λ_0):

$(\sqrt{3},\ \sqrt{\frac{3}{2}},\ \frac{1}{2\sqrt{2}})$

$(\sqrt{3},\ -\sqrt{\frac{3}{2}},\ \frac{1}{2\sqrt{2}})$

$(\sqrt{3},\ \sqrt{\frac{3}{2}},\ -\frac{1}{2\sqrt{2}})$

$(\sqrt{3},\ -\sqrt{\frac{3}{2}},\ -\frac{1}{2\sqrt{2}})$

$(-\sqrt{3},\ \sqrt{\frac{3}{2}},\ \frac{1}{2\sqrt{2}})$

$(-\sqrt{3},\ -\sqrt{\frac{3}{2}},\ \frac{1}{2\sqrt{2}})$

$(-\sqrt{3},\ \sqrt{\frac{3}{2}},\ -\frac{1}{2\sqrt{2}})$

$(-\sqrt{3},\ -\sqrt{\frac{3}{2}},\ -\frac{1}{2\sqrt{2}})$

Extremal Points in the form (x_0, y_0):

$(\sqrt{3},\ \sqrt{\frac{3}{2}}) \rightarrow$ **f(x,y) = xy =** $\frac{3}{\sqrt{2}}$

$(\sqrt{3},\ -\sqrt{\frac{3}{2}}) \rightarrow$ **f(x,y) = xy =** $-\frac{3}{\sqrt{2}}$

$(-\sqrt{3},\ \sqrt{\frac{3}{2}}) \rightarrow$ **f(x,y) = xy =** $-\frac{3}{\sqrt{2}}$

$(-\sqrt{3},\ -\sqrt{\frac{3}{2}}) \rightarrow$ **f(x,y) = xy =** $\frac{3}{\sqrt{2}}$

The solutions that yield the maximum are $\{(\sqrt{3},\ \sqrt{\frac{3}{2}}),\ (-\sqrt{3},\ -\sqrt{\frac{3}{2}})\}$, yielding maximum value $\frac{3}{\sqrt{2}}$

Problem 3

Part a:

α = 0.8, w(0) = (0,0), (y=1=Positive/y=0=Negative)

1st iteration: x^1 = [1, -1, 3], y^1 = 1

$\widehat{y^1}$ = t(0*-1 + 0*1) = 0

$w_0(1) = w_0(0) + \alpha(y^1 - \widehat{y^1})x_0^1$ = 0 + 0.8(1 - 0)(1) = 0.8

$w_1(1) = w_1(0) + \alpha(y^1 - \widehat{y^1})x_1^1$ = 0 + 0.8(1 - 0)(-1) = -0.8

$w_2(1) = w_2(0) + \alpha(y^1 - \widehat{y^1})x_2^1$ = 0 + 0.8(1 - 0)(3) = 2.4

w(1) = (0.8, -0.8, 2.4)

2nd iteration: x^2 = [1, 1, 6], y^2 = 0

$\widehat{y^2}$ = t(0.8*1 + -0.8*1 + 2.4*6) = 14.4

$w_0(2) = w_0(1) + \alpha(y^2 - \widehat{y^2})x_0^2$ = 0.8 + 0.8(0 - 14.4)(1) = -10.72

$w_1(2) = w_1(1) + \alpha(y^2 - \widehat{y^2})x_1^2$ = -0.8 + 0.8(0 - 14.4)(1) = -12.32

$w_2(2) = w_2(1) + \alpha(y^2 - \widehat{y^2})x_2^2$ = 2.4 + 0.8(0 - 14.4)(6) = -66.72
w(2) = (-10.72, -12.32, -66.72)

3rd iteration: x^3 = [1, 0, 1], y^3 = 1
$\widehat{y^3}$ = t(-10.72*1 + -12.32*0 + -66.72*1) = -77.44
$w_0(3) = w_0(2) + \alpha(y^3 - \widehat{y^3})x_0^3$ = -10.72 + 0.8(1 - (-77.44))(1) = 52.032
$w_1(3) = w_1(2) + \alpha(y^3 - \widehat{y^3})x_1^3$ = -12.32 + 0.8(1 - (-77.44))(0) = -12.32
$w_2(3) = w_2(2) + \alpha(y^3 - \widehat{y^3})x_2^3$ = -66.72 + 0.8(1 - (-77.44))(1) = -3.968
w(3) = (52.032, -12.32, -3.968)

4th iteration: x^4 = [1, 1, 5], y^4 = 0
$\widehat{y^3}$ = t(52.032*1 + -12.32*1 + -3.968*5) = 19.872
$w_0(4) = w_0(3) + \alpha(y^4 - \widehat{y^4})x_0^4$ = 52.032 + 0.8(0 - 19.872)(1) = 36.134
$w_1(4) = w_1(3) + \alpha(y^4 - \widehat{y^4})x_1^4$ = -12.32 + 0.8(0 - 19.872)(1) = -28.2176
$w_2(4) = w_2(3) + \alpha(y^4 - \widehat{y^4})x_2^4$ = -3.968 + 0.8(0 - 19.872)(5) = -83.456
w(4) = (36.134, -28.2176, -83.456)

5th iteration: x^5 = [1, 0, 2], y^5 = 1
$\widehat{y^5}$ = t(36.134*1 + -28.2176*0 + -83.456*2) = -130.776
$w_0(5) = w_0(4) + \alpha(y^5 - \widehat{y^5})x_0^5$ = 36.134 + 0.8(1 - (-130.776))(1) = 141.5548
$w_1(5) = w_1(4) + \alpha(y^5 - \widehat{y^5})x_1^5$ = -28.2176 + 0.8(1 - (-130.776))(0) = -28.2176
$w_2(5) = w_2(4) + \alpha(y^5 - \widehat{y^5})x_2^5$ = -83.456 + 0.8(1 - (-130.776))(1) = 21.9648
w(5) = (141.5548, -28.2176, 21.9648)

6th iteration: x^6 = [1, 0, 0], y^6 = 1
$\widehat{y^6}$ = t(141.5548*1 + -28.2176*0 + 21.9648*0) = 141.5548
$w_0(6) = w_0(5) + \alpha(y^6 - \widehat{y^6})x_0^6$ = 141.5548 + 0.8(1 - 141.5548)(1) = 29.11096
$w_1(6) = w_1(5) + \alpha(y^6 - \widehat{y^6})x_1^6$ = -28.2176 + 0.8(1 - 141.5548)(0) = -28.2176
$w_2(6) = w_2(5) + \alpha(y^6 - \widehat{y^6})x_2^6$ = 21.9648+ 0.8(1 - 141.5548)(0) = 21.9648
w(6) = (29.11096, -28.2176, 21.9648)

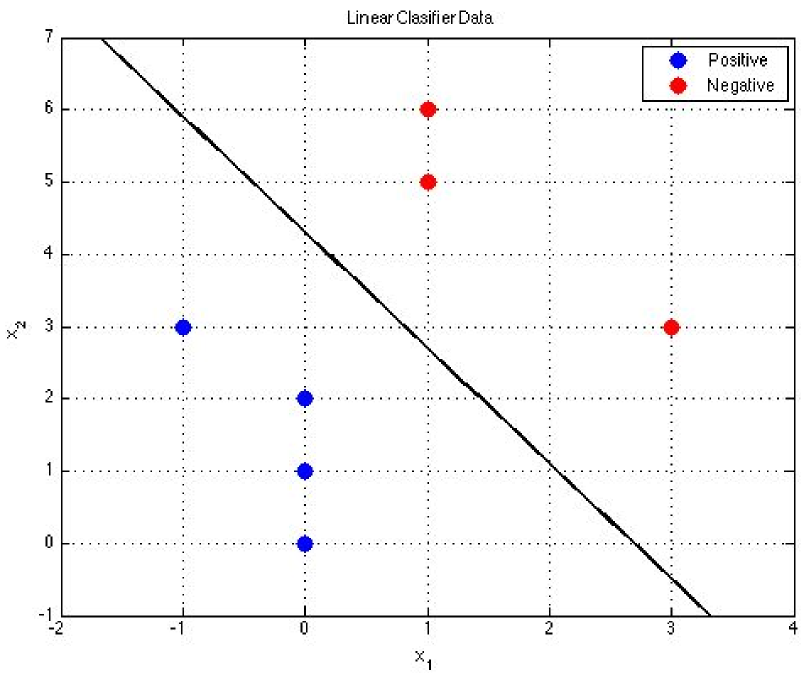
7th iteration: x^7 = [1, 3, 3], y^7 = 0
$\widehat{y^7}$ = t(29.11096*1 + -28.2176*3 + 21.9648*3) = 10.35256
$w_0(7) = w_0(6) + \alpha(y^7 - \widehat{y^7})x_0^7$ = 29.11096 + 0.8(0 - 10.35256)(1) = 20.8289

$w_1(7) \; = \; w_1(6) \; + \; \alpha(y^7 - \widehat{y^7})x_1^7$ = -28.2176 + 0.8(0 - 10.35256)(3) = -56.0637

$w_2(7) \; = \; w_2(6) \; + \; \alpha(y^7 - \widehat{y^7})x_2^7$ = 21.9648 + 0.8(0 - 10.35256)(3) = -2.8813

w(7) = (20.8289, -56.0637, -2.8813)

Part b:



Part c:

w = $(8, 5)^T$, b = -21.5

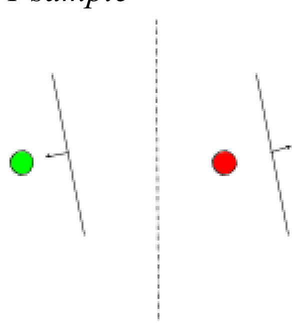
Part d:

w = $(8, 5)^T$, b = -24.5

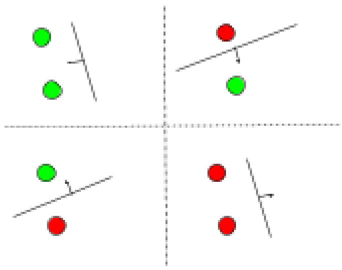
Problem 4

Part a: *Two half-planes*

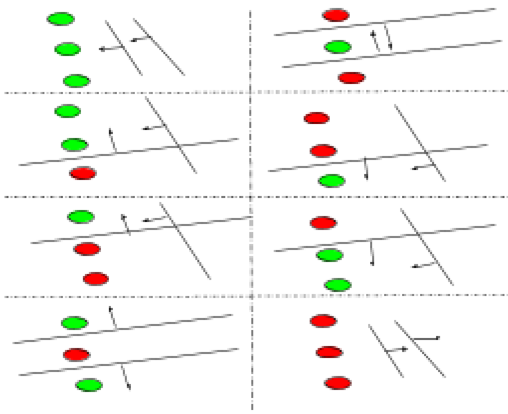
1 sample



2 samples



3 samples



Will not show for the rest of the sample sizes as the number of drawings needed is exponential,, but the general idea is clear, We know that the VD-dimension for a single half-plane is 3. So, by having 2 half-planes, you can seperate 3 samples with each one, meaning you can seperate a total of 9 points, because the first one reduces the size from 9 to 3 and then the second reduces it from 3 to 1. ***Thus the VC-dim of two half planes is 9.***

Part b: *Inside of a triangle*

The VC-dim of a triangle is 7. Consider the 7 points on a circle. They can be separated in any given configuration because at most, they can form 3 continuous blocks like (+, -, +, -, +, -, +). This means that each edge of the triangle can be used to cut off each block. For 8 points, this is not possible because if one of the points is in the convex hull of the rest, then we cannot label that point as negative while labelling the rest as positive.

Problem 5

$V_0([4,3]) = 1$, $V_0([4,2]) = -1$, 0 for the rest; $R(s) = -0.05$ for $R \neq \pm 1$

Part a: First 2 Policy Iterations

First Policy Evaluation:

With π_0 = Up

States

+0.000 +0.000 +0.000 +1.000
+0.000 x +0.000 -1.000
+0.000 +0.000 +0.000 +0.000

Policies

----
up up up 1
up x up -1
up up up up

States

-0.050 -0.050 +0.760 +1.000
-0.050 x -0.860 -1.000
-0.050 -0.050 -0.050 -0.050

Policies

----
up up right 1
up x up -1
up up up down

Part b: Last 2 Policy Iterations

States

-0.335 -0.162 +0.211 +1.000
-0.491 x -0.904 -1.000
-0.491 -0.491 -0.491 -0.491

Policies

----
right right right 1
up x up -1

```
   up  left  left  left
-----------------------------

States
-----------------------------
-0.336 -0.162 +0.211 +1.000
-0.492    x   -0.904 -1.000
-0.492 -0.492 -0.492 -0.492
-----------------------------

Policies
-----------------------------
right right right    1
   up    x     up    -1
   up  left  left  left
-----------------------------
```