RMIT University
School of Science
COSC2110/COSC2111 Data Mining
Tutorial Problems Week 7

```
Area Name,Arrival Time,Duration,Sign,In Violation?
Victoria Market,24/08/2012 11:34,4473,LZ 15M M-SAT 7:30-19:30,1
Courtney,17/03/2012 13:07,127,1/2P A RPE M-SUN 7:30-23:00,0
Victoria Market,17/02/2012 13:54,1543,2P MTR M-SAT 7:30-20:30,0
Chinatown,27/11/2011 15:03,3562,1P SUN 7:30-18:30,0
Southbank,21/04/2012 15:08,7048,2P TKT A M-SAT 7:30-18:30,0
Banks,11/11/2011 9:56,1900,1P MTR M-F 7:30-16:30,0
Victoria Market,7/12/2011 19:50,173,2P MTR M-SAT 7:30-20:30,0
The Mac,3/03/2012 14:36,674,1/2P TKT A SAT 7:30-19:30,0
City Square,19/05/2012 18:20,3597,1/2P MTR SAT 7:30-1930,1
Hyatt,25/08/2012 14:28,1994,1P MTR M-SAT 7:30-19:30,0
Chinatown,25/02/2012 18:30,9125,2P DIS M-SUN 0:00-23:59,1
Princes Theatre,17/11/2011 9:00,6066,1P MTR M-SAT 7:30-19:30,1
Southbank,20/06/2012 16:32,1136,P/ 15 M-SUN 00:00-23:59,0
Chinatown,29/01/2012 12:26,107,S/ No Stop CCV M-SUN 0:00-23:59,1
Princes Theatre,12/01/2012 18:51,116,1P MTR M-SAT 7:30-19:30,0
Southbank,20/02/2012 13:50,9773,2P TKT A M-F 7:30-18:30,1
```

1. For the above data identify the unix program and give the command.
   (a) Extract the first 10 lines
   (b) Extract all lines from line 10 onwards
   (c) Extract the sign column
   (d) Extract parking events that occurred in Chinatown
   (e) Extract parking events that occurred in April 2012
   (f) Extract parking events that occurred in Chinatown in April 2012
   (g) Extact parking events that occurred in a 2 hour zone
   (h) Extract events for which there was a violation (ie a 1 in the 'violation?' column)

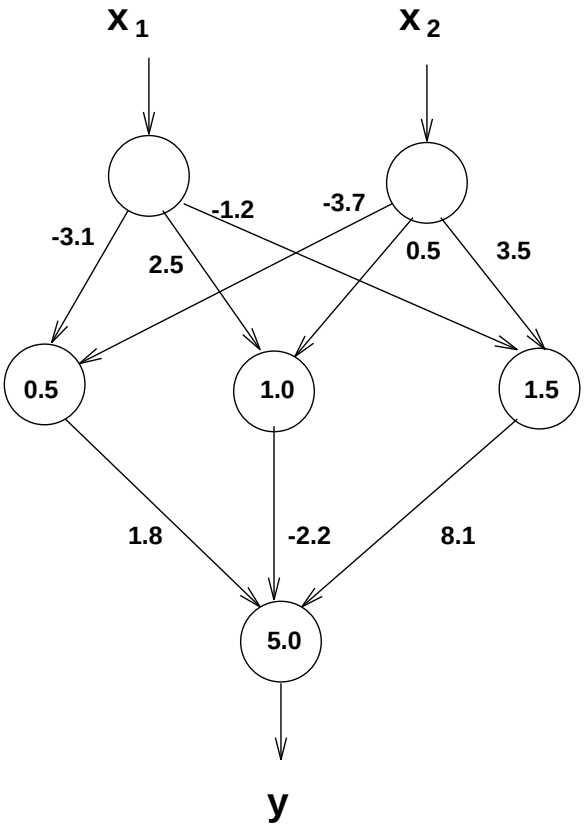
2. What will be the network output for each of the following input vectors? Assume sigmoid units.

   (a) (0,0)

   (b) (1,0)

   (c) (0.7 0.2)

   (d) (-5,5)



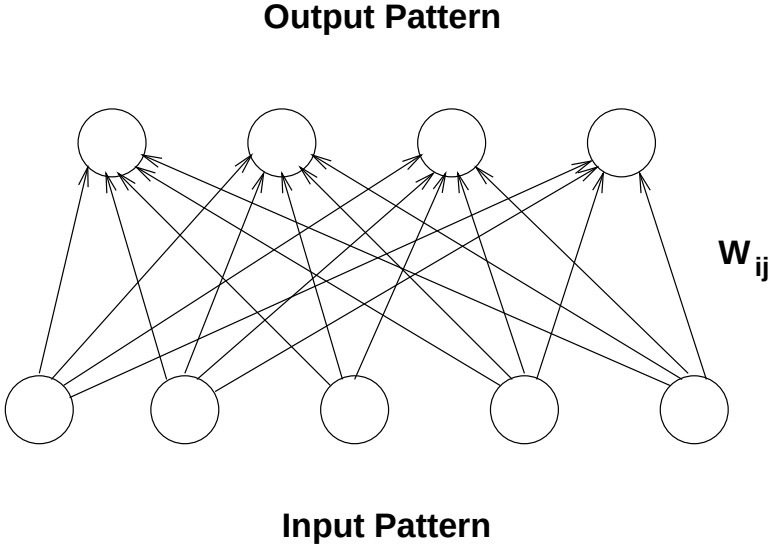
3. Repeat the above question assuming linear threshold units.

4. Suppose that the training example:

Inputs					Desired Outputs			
0.0	1.0	0.5	1.0	0.0	1.0	0.5	0.0	0.0

is passed into the following network and the outputs are: 0.23, 0.74, 0.19, 0.99.

What is the sum squared error for this training example.

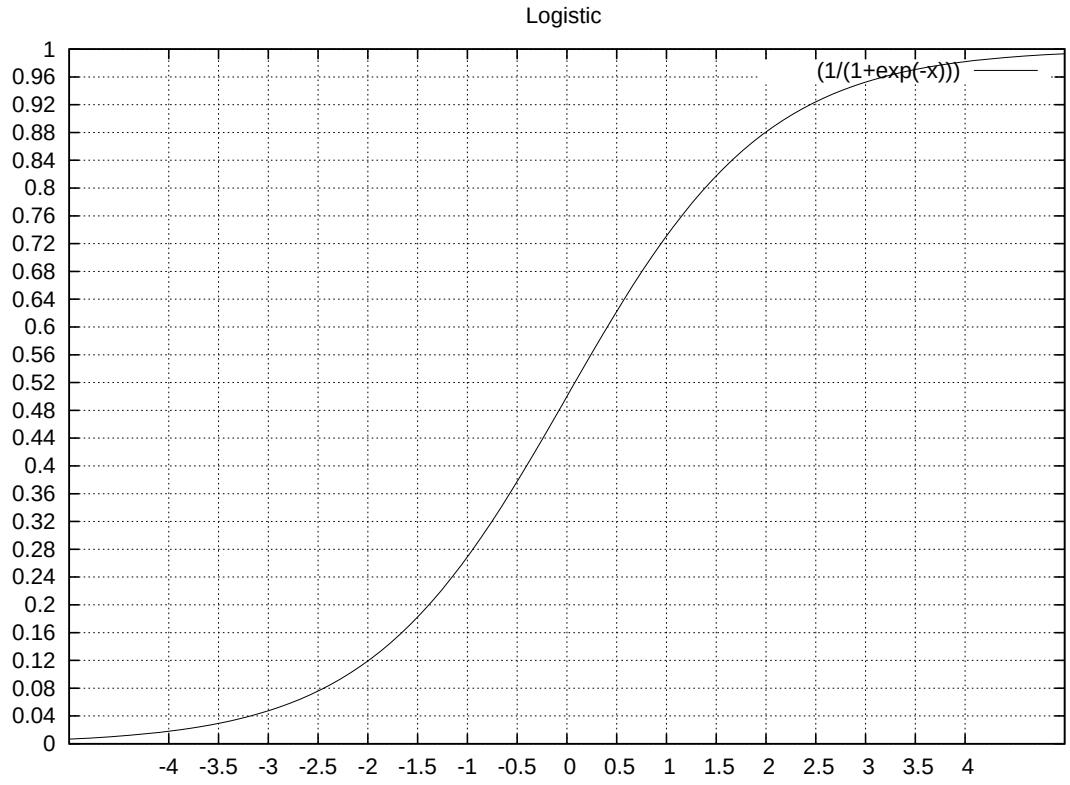


For the following data, what is the

(a) Total sum squared error? (TSS)

(b) Mean squared error (MSE)

Inputs					Desired Outputs				Actual Outputs			
0.0	1.0	0.5	1.0	0.0	1.0	0.5	0.0	0.0	0.23	0.74	0.19	0.99
0.0	0.3	0.5	1.0	0.0	1.0	0.5	0.0	0.0	0.23	0.80	0.19	0.99
0.0	1.0	0.5	1.0	0.0	0.0	0.5	0.0	0.0	0.23	0.74	0.19	0.88


Logistic
(1/(1+exp(-x)))
1
0.96
0.92
0.88
0.84
0.8
0.76
0.72
0.68
0.64
0.6
0.56
0.52
0.48
0.44
0.4
0.36
0.32
0.28
0.24
0.2
0.16
0.12
0.08
0.04
0
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4