University of Manitoba
Midterm

Winter 2012

COMPUTER SCIENCE

Machine Learning

Date: **Friday, 9th March 2012**
Time: **15:30 - 16:20**
Room: **EITC E2-304, University of Manitoba**

(Time allowed: 50 Minutes)

NOTE: Attempt all questions.
This is a *closed* book examination.
Use of non-programmable calculators is *permitted*.
Use of any other electronic equipment is strictly forbidden.
Show your work to receive full marks.

SURNAME:

FORENAME(S):

STUDENT ID:

A	B	C	Total
17	15	18	50

Surname: _____

Forename(s): _____

Section A: Decision Trees

The information gain Gain(S,A) of an attribute A for a sample set S is defined as

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{\|S_v\|}{\|S\|} \text{Entropy}(S_v)$$

The entropy function Entropy(S) for two classes $\oplus$ and $\ominus$ is defined as

$$\text{Entropy}(S) = -p_{\oplus} \log_2(p_{\oplus}) - p_{\ominus} \log_2(p_{\ominus})$$

1. Show the maximum entropy for a domain with 9 instances and two possible classes $\oplus$ and $\ominus$

   [7 marks]

   The maximum entropy for a domain with 9 instances and two classes is: 0.9911

2. Given below is a set of instances from a medical diagnosis domain with two attributes blood sugar (`Blood`) and weight (`Weight`)). The system is trying to predict if the patient has some disease.

   Given the set of instances shown below, calculate the information gain for the attributes `Blood` and `Weight`.

Blood	Height	Disease
Normal	Normal	No
High	High	Yes
Low	High	Yes
Low	Normal	No
High	Low	Yes
High	Normal	?
Normal	Low	?

   The classification for the instances `<High,Normal>` and `<Normal,Low>` were lost in transmission. However, you were also told the information gain for the two attributes. The information gain of attribute `Blood Sugar` was 0.2917 and for attribute `Weight` was 0.4695.

   Given this information, show the correct classification for the instances `<High,Normal>` and `<Normal,Low>`.

CONTINUED

Surname: _____

Forename(s): _____

If it is impossible to determine the missing classification uniquely, then say so in your answer and explain why.

[10 marks]

<High, Normal> = Yes

Only one combination results in

<Normal,Low> = Yes

these information gains.

```
Question Decision Tree
Information Gain(Blood Sugar)=
  Entropy(5,2)[0.863120568566631]-
    -2/7*Entropy(1,1)[1.0] # Blood Sugar:normal
    -3/7*Entropy(3,0)[0.0] # Blood Sugar:high
    -2/7*Entropy(1,1)[1.0] # Blood Sugar:low
  =0.2916919971380596
Information Gain(Weight)=
  Entropy(5,2)[0.863120568566631]-
    -3/7*Entropy(1,2)[0.9182958340544896] # Weight:normal
    -2/7*Entropy(2,0)[0.0] # Weight:high
    -2/7*Entropy(2,0)[0.0] # Weight:low
  =0.46956521111470695
```

Section B: Regression

3. Regression uses gradient descent to calculate the least squared error fit of a model to the data. In the case of one-dimensional data, the model corresponds to a straight line with the equation

$$h_\theta(x) = \theta_0 + \theta_1 x$$

The Gradient Descent Update rule in this case is given by

$$\begin{aligned} \theta_0 &:= \theta_0 - \alpha \frac{1}{m} \sum_{i=0}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \\ \theta_1 &:= \theta_1 - \alpha \frac{1}{m} \sum_{i=0}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x^{(i)} \end{aligned}$$

Given is the following data set

Surname: _____

Forename(s): _____

X (Input)	Y (Target)
1.5	4.0
2.3	3.0
3.0	0.5

Assume that the current estimate for the linear regression is given by $\theta_0 = 1.0, \theta_1 = 2.0$ and that the learning rate $\alpha = 0.1$.

Calculate the new model estimate for the given data set.

If it is impossible to calculate the new model estimate given the information above, then say so in your answer and explain why.

[15 marks]

θ_0 = 0.6967

θ_1 = 1.1507

Section C: Neural Nets

4. Given below is an artificial neural network (ANN) with three input nodes (X1,X2,X3), two hidden nodes, and one output node.

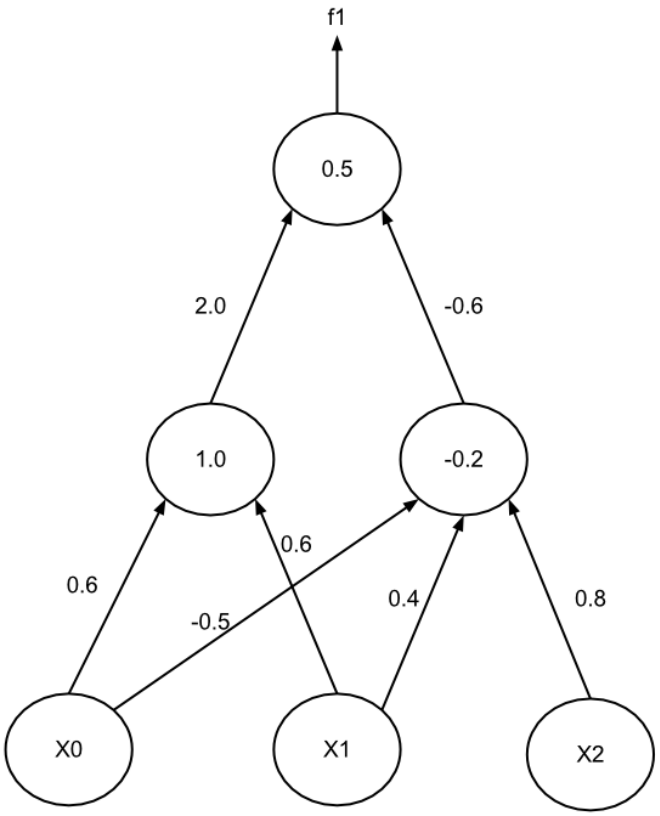
   The network uses simple threshold nodes (i.e., the node will output 1.0 if the sum of the weighted inputs is greater than or equal to the threshold, 0 otherwise).

   Show the output function f_1 of this neural network.

   [10 marks]

Surname: _____

Forename(s): _____



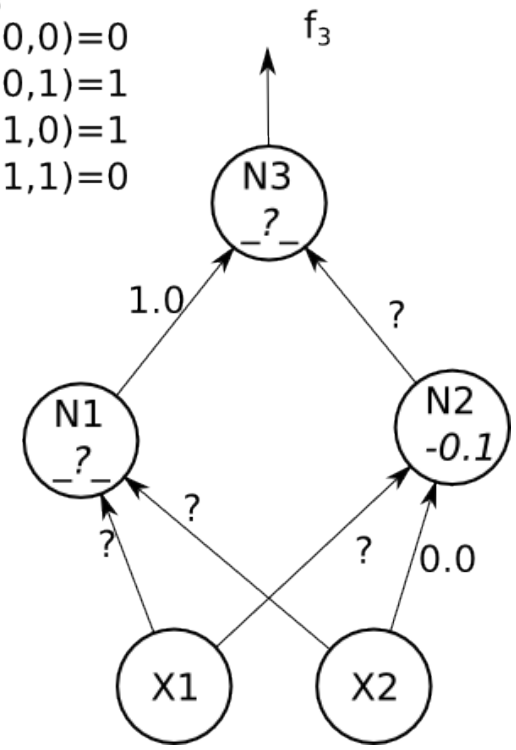
x_0	x_1	x_2	f_1	
0	0	0		0
0	0	1		0
0	1	0		0
0	1	1		0
1	0	0		0
1	0	1		0
1	1	0		1
1	1	1		1

5. Given below is a neural network with three nodes $N1$,$N2$, and $N3$, as well as two inputs $X1$ and $X2$. The network uses threshold neurons, that is a neuron will generate an output of 1, if the sum of the weighted inputs is greater than or equal to the threshold, 0 otherwise.

Surname: _____

Forename(s): _____

X1	X2	f_3
0	0	$f_3(0,0)=0$
0	1	$f_3(0,1)=1$
1	0	$f_3(1,0)=1$
1	1	$f_3(1,1)=0$



Answer the following questions with either:

- **true** if the statement is definitely true,
- **false** if the statement is definitely false,
- **unknown** if there is not enough information given in the question to determine the truthfulness of the statement.

[8 marks]

Surname: _____

Forename(s): _____

If $X1 = 0$ and $X2 = 0$, then the output of node $N2$ is greater than 1.0: True

The threshold of node $N3$ must be greater than or equal to 1.0: Unknown

The threshold of node $N3$ must be less than or equal to 2.0: False

The weight from $X1$ to $N2$ must be negative: Unknown

The weight from $X1$ to $N2$ must be less than -0.2: Unknown

The weight from $X1$ to $N1$ must be greater than 0: Unknown

The weight from $X1$ to $N2$ must be greater than 0: Unknown

The weight from $X2$ to $N1$ must be less than 0: Unknown

Surname: _____

Forename(s): _____

Additional work pages