Cse352 AI
Homework 3
SOLUTIONS

PROBLEM 1

Use Lecture Notes to WRITE short, 1-2 paragraphs long ANSWERS to the following questions.
Remember: on TEST your answers must match Lectures not RANDOM Google pieces as a lot of students copied in. This is zero points.

I compiled some answers by students that ARE acceptable.

1. **Describe what is CLASSIFICATION; type of data, goals and types of applications**
   Classification is an example of supervised learning in which rules (characteristic and discriminant), or other patters (converged Neural Network) are created based on training data and tested with test data. If the predictive accuracy of the rules (patters) is high, they can be used to classify new data. The training data must be in the form of a table with its key attribute removed and its class attribute distinguished. The values in this class label must be discrete and unordered. An example of an application of classification could be customer. Profiling, where customer preferences and which customers buy which products can be used to identify the best products for different customers

2. **Describe all stages of the classification process**
   The first step of the classification process as in any the learning process is the selection of data. The selected data is a subset of all the available data and must address the problem at hand.

The next step is to clean the selected data so that the impact of missing, incomplete, or noisy data is minimized.

Next the data must continue to get preprocessed so that learning algorithms can be applied to it.

Now a classification algorithm can be applied to the preprocessed data to obtain patterns about the data.

In the final steps, the patterns are tested for their accuracy and interpreted so that the user can decide if they still need refining or can be presented.

3. **Describe and discuss basic classification Models and their differences**

   There are descriptive and statistical classification models. Descriptive models are models are Decision Trees, Classification by Association, Genetic Algorithms, and Rough Sets. Statistical models are: Neural Networks, Bayesian Networks, Support Vectors Model

4. **Discuss the Decision Tree Induction and its strengths and weaknesses**

   Decision tree is constructed in a top-down recursive divide-and-conquer manner. At start, all the training examples are at the root. Attributes are categorical. Examples are partitioned reclusively based on selected attributes. Test attributes are selected on the basis of a heuristic or statistical measure. Strength of Decision Tree method is decision trees are able to generate understandable rules and performs classification without requiring much computation. However, trees are less appropriate for

estimation tasks where the goal is to predict the value of a continuous attribute

5. **Discuss the Neural Network Model and its strengths and weaknesses**

   Neural network is a set of connected input/output units where each connection has a weight associated with it. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct class label of the input tuples. Its weakness is the long training time and it is difficult to interpret the symbolic meaning behind the learned weights and of hidden units in the network. Its strength is high tolerance to noisy data and ability to classify incomplete data

6. **Describe a process of building a CLASSIFIER**

   Building a classifier follows stages if any learning process. The main two phases of building a classifier are: training and testing. In both phase we use training data set and test data. We use training data set to create patterns. We evaluate created patterns with the use of test data. We terminate the process if it has been trained and tested and the predictive accuracy is on an acceptable level.

   All stages are, like in any Learning Process the following

1) The first step is the selection of data. The selected data is a subset of all the available data and must address the problem at hand.
2) The next step is to clean the selected data so that the impact of missing, incomplete, or noisy data is minimized.
3) Next the data must continue to get preprocessed so that learning algorithms can be applied to it.
4) Now a learning algorithm can be applied to the preprocessed data to obtain patterns about the data.

5) In the final steps, the patterns are tested for their accuracy and interpreted so that the user can decide if they still need refining or can be presented.

7. Define a CLASSIFIER

A classifier is the final product of a learning process that uses a classification data set and a classification algorithm. **A CLASSIFIER is a black box** that is used to classify records for which the class label in unknown.

A classifier is the end result of a process that uses training data and a classification algorithm to generate patterns that can classify new data. The effectiveness, need for refinement, and completion of a classifier is evaluated based on its predictive accuracy. The patterns a classifier uses take different forms, such as rules or trees, or trained different networks, like Neural or Bayesian.

PROBLEM 2: BUILDING A CLASSIFIER

For the data set given below build a classifier following all steps needed in the constructions:
preprocessing, training and testing
Describe and motivate your choice of algorithms and methods used at each step.

CLASSIFICATION DATA:

Age	Income	Student	Credit Rating	Buys Computer
21	60,000	yes	3	No
30	70,000	No	5	No
38	38666.667	No	2	Yes
45	45,000	yes	3	Yes
46	25,000	no	2	Yes
47	30,000	Yes	6	No
39	28,000	Yes	5	No
29	48,000	Yes	3	No
50	75,000	Yes	2	No
48	55125	Yes	3	No
30	38666.667	Yes	6	Yes
51	46,000	No	4	Yes
32	80,000	Yes	2	No
45	50,000	No	4	No

PART 1: Preprocessing

Attributes: Age, Income, Credit Rating

1. **Fill the missing Values**

2. **Use Binning Method to discretize values of attributes** Age, Income, Credit_Rating

The number of bins is up to you

1. Preprocessing Calculations

THIS is a Solution Submitted by a Student

There are many other solutions!

1. Mising Values: explain the method you used

Attribute values mean for income:

No Class
(60000+70000+30000+28000+48000+75000+80000+50000)/8=
441000/8=55125 Mean

Yes Class
(45000+25000+46000)/3=
116000/3=38666.667 Mean

2. Binninig: AGE

Equal-Depth binning for Age

N intervals
N=3
Depth=5

Bin 1	21	29	30	30	32
Bin 2	38	39	45	45	46
Bin 3	47	48	50	51	

Bin 1: Age Group >=21&<=32
Bin 2: Age Group >32&<=46
Bin 3: Age Group >46&<=51

3. Binning: INCOME

Equal-Depth binning for Income

N intervals
N=3
Depth=5

Bin 1	25000	28000	30000	38666.667	38666.667
Bin 2	45000	46000	48000	50000	55125
Bin 3	60000	70000	75000	80000	

Bin 1: >=25000&<=38666.667
Bin 2: >38666.667&<=55125
Bin 3: >55125&<=80000

4. Binning : CREDIT RATING

Equal-Width binning for Credit Rating

N intervals
N=2
Highest Value=6
Lowest Value=2
(6+2)/3=8/2=4

Bin 1	0>=&<4	2,2,2,2,3,3,3,3
Bin 2	4>=&<=8	4,4,5,5,6,6

Bin 1: >=0&<4
Bin 2: >=4&<=8

2. YOUR Data after Preprocessing:

Remark: Bins for different attributes should have and usually HAVE special names to make the FINAL Descriptive Rules comprehensible

Age	Income	Student	Credit Rating	Buys Computer
Bin 1	Bin 3	yes	Bin 1	No
Bin 1	Bin 3	No	Bin 2	No
Bin 2	Bin 1	No	Bin 1	Yes
Bin 2	Bin 2	yes	Bin 1	Yes
Bin 2	Bin 1	no	Bin 1	Yes
Bin 3	Bin 1	Yes	Bin 2	No
Bin 2	Bin 1	Yes	Bin 2	No
Bin 1	Bin 2	Yes	Bin 1	No
Bin 3	Bin 3	Yes	Bin 1	No
Bin 3	Bin 2	Yes	Bin 1	No
Bin 1	Bin 1	Yes	Bin 2	Yes
Bin 3	Bin 2	No	Bin 2	Yes
Bin 1	Bin 3	Yes	Bin 1	No
Bin 2	Bin 2	No	Bin 2	No

3. Training and Testing

Specify which Decision Tree Algorithm you use

USE Three-Fold Cross Validation –Lecture 11testing

Remember: must perform training and testing 3 times and

predictive accuracy is averaged;
We adopt the union of rules as the
FINAL set of rules- YOUR Classifier

TRAINING- TESTING

ID3 non heuristic attributes chosen

REMARK THE TREE SHOULD BR CONSTRUCTED for PREPROCESSED DATA- NOT ORIGINAL!!
Sets 1,2, 3 are original DATA **NOT** preprocessed DATA!
When you write SOLUTIONS by HAND, as on the TEST; you MUST use preprocessed DATA!

When you use SOFTWARE, you input ORIGINAL DATA, because SOFTWARE changes it accordingly to Preprocessing program; so
THE TREE IS constructed (the student visibly used some program!) for Preprocessed data, where BINS are replaced by appropriate intervals. YOU CAN use software but MUST **present your data – and results correctly.**

ON THE TEST you will have to do it all BY HAND.

Set 1:

Age	Income	Student	Credit Rating	Buys Computer
21	60,000	yes	3	No
30	70,000	No	5	No
38	38,666.667	No	2	Yes
45	45,000	yes	3	Yes
46	25,000	No	2	Yes

Set 2:

Age	Income	Student	Credit Rating	Buys Computer
47	30,000	Yes	6	No
39	28,000	Yes	5	No
29	48,000	Yes	3	No
51	46,000	No	4	Yes
32	80,000	Yes	2	No

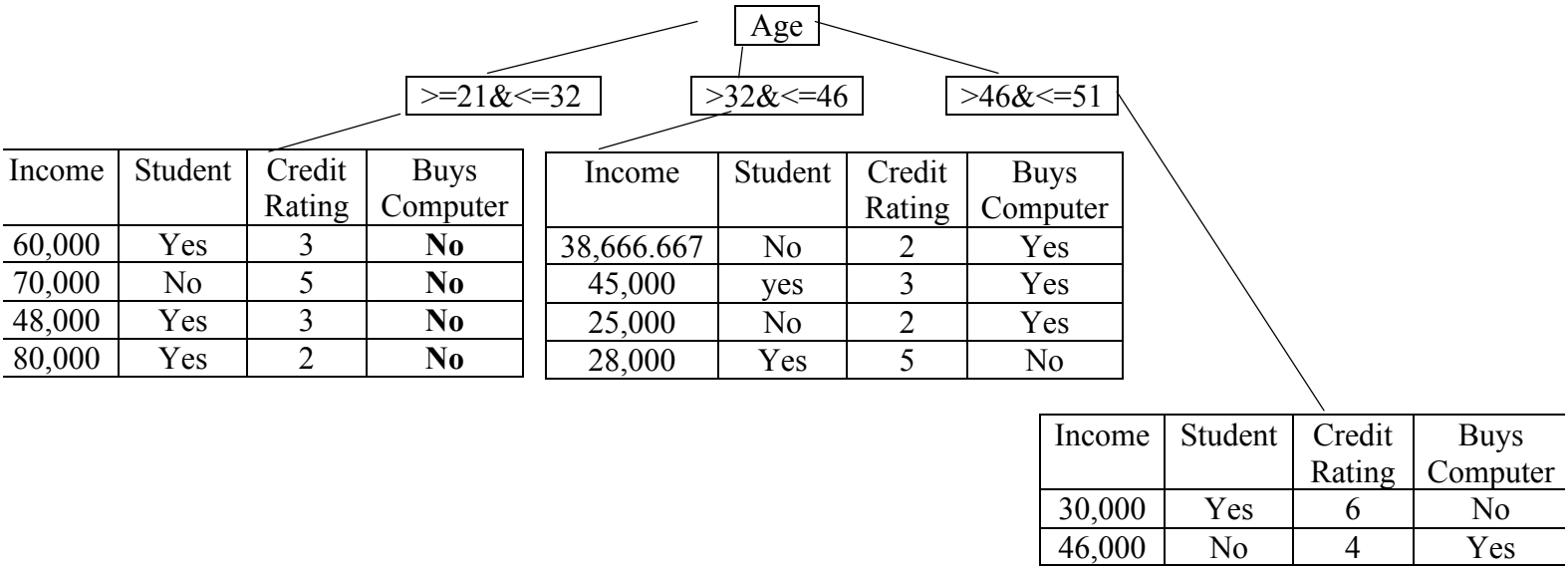
Set 3:

Age	Income	Student	Credit Rating	Buys Computer
50	75,000	Yes	2	No
48	55125	Yes	3	No
30	38666.667	Yes	6	Yes
45	50,000	No	4	No

Training 1

Age	Income	Student	Credit Rating	Buys Computer

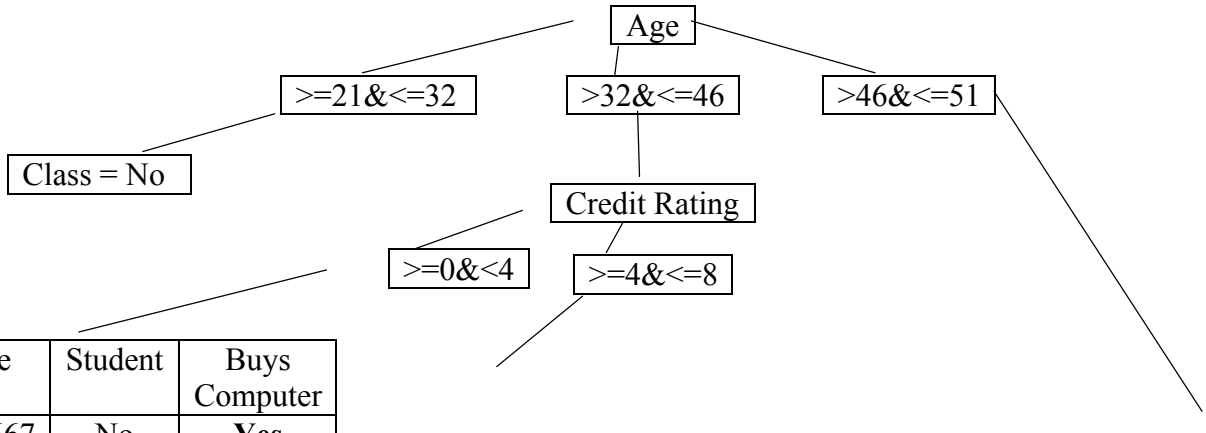
21	60,000	Yes	3	No
30	70,000	No	5	No
38	38,666.667	No	2	Yes
45	45,000	Yes	3	Yes
46	25,000	No	2	Yes
47	30,000	Yes	6	No
39	28,000	Yes	5	No
29	48,000	Yes	3	No
51	46,000	No	4	Yes
32	80,000	Yes	2	No



Income	Student	Credit Rating	Buys Computer
60,000	Yes	3	**No**
70,000	No	5	**No**
48,000	Yes	3	**No**
80,000	Yes	2	**No**

Income	Student	Credit Rating	Buys Computer
38,666.667	No	2	Yes
45,000	yes	3	Yes
25,000	No	2	Yes
28,000	Yes	5	No

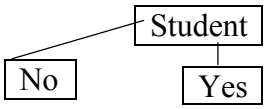
Income	Student	Credit Rating	Buys Computer
30,000	Yes	6	No
46,000	No	4	Yes



Income	Student	Buys Computer
38,666.667	No	**Yes**

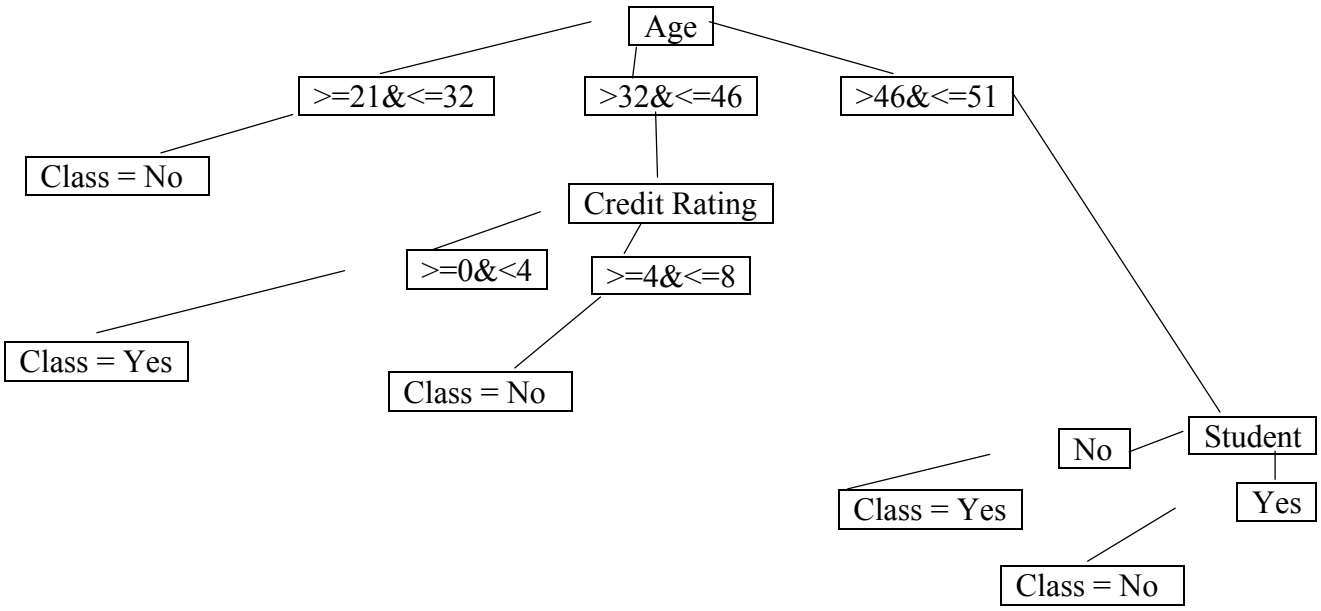
45,000	yes	**Yes**
25,000	No	**Yes**

Income	Student	Buys Computer
28,000	Yes	**No**



Income	Credit Rating	Buys Computer
46,000	4	**Yes**

Income	Credit Rating	Buys Computer
30,000	6	**No**



Rules : BINS names are replaced by bins intervals

R1. Age(x, >=21&<=32) ⇒ Class = buysComputer(x, No)
R2. Age(x, >32&<=46) Λ Credit Rating(x, >=0&<4) ⇒ Class = buysComputer(x, Yes)

R3. Age(x, >32&<=46) Λ Credit Rating(x, >=4&<=8) ⇒ Class = buysComputer(x, No)
R4. Age(x, >46&<=51) Λ Student(x, No) ⇒ Class = buysComputer(x, Yes)
R5. Age(x, >46&<=51) Λ Student(x, Yes) ⇒ Class = buysComputer(x, No)

Testing Set 3

Age	Income	Student	Credit Rating	Buys Computer
50	75,000	Yes	2	No
48	55125	Yes	3	No
30	38666.667	Yes	6	Yes
45	50,000	No	4	No

Record 1 is well classified as of Rule 5
Record 2 is well classified as of Rule 5
Record 3 is misclassified
Record 4 is well classified as of Rule 3

3 records well classified out of 4

3/4*100%=75%= predictive accuracy

Training 2

Age	Income	Student	Credit Rating	Buys Computer
47	30,000	Yes	6	No
39	28,000	Yes	5	No
29	48,000	Yes	3	No
51	46,000	No	4	Yes
32	80,000	Yes	2	No
50	75,000	Yes	2	No
48	55125	Yes	3	No
30	38666.667	Yes	6	Yes
45	50,000	No	4	No

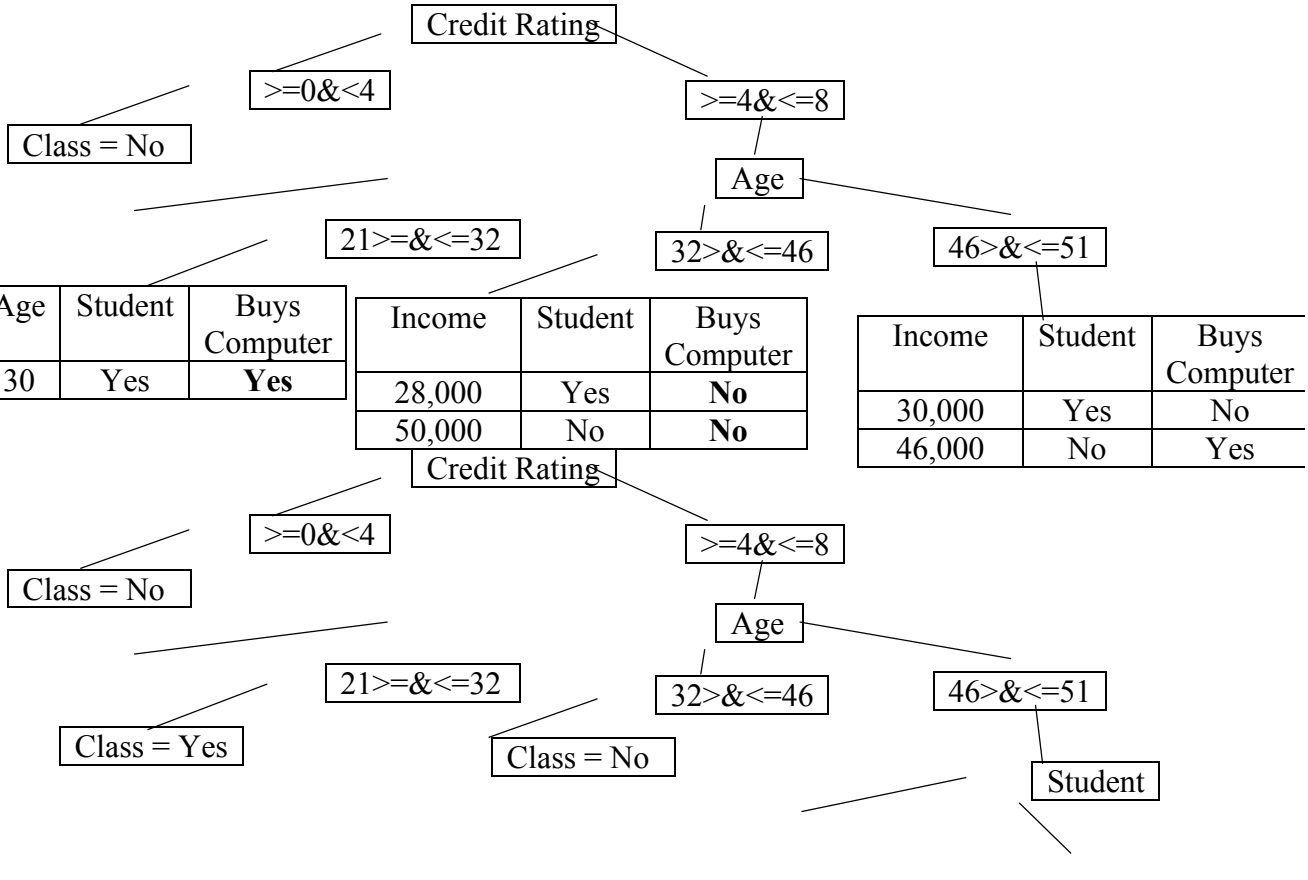
Credit Rating

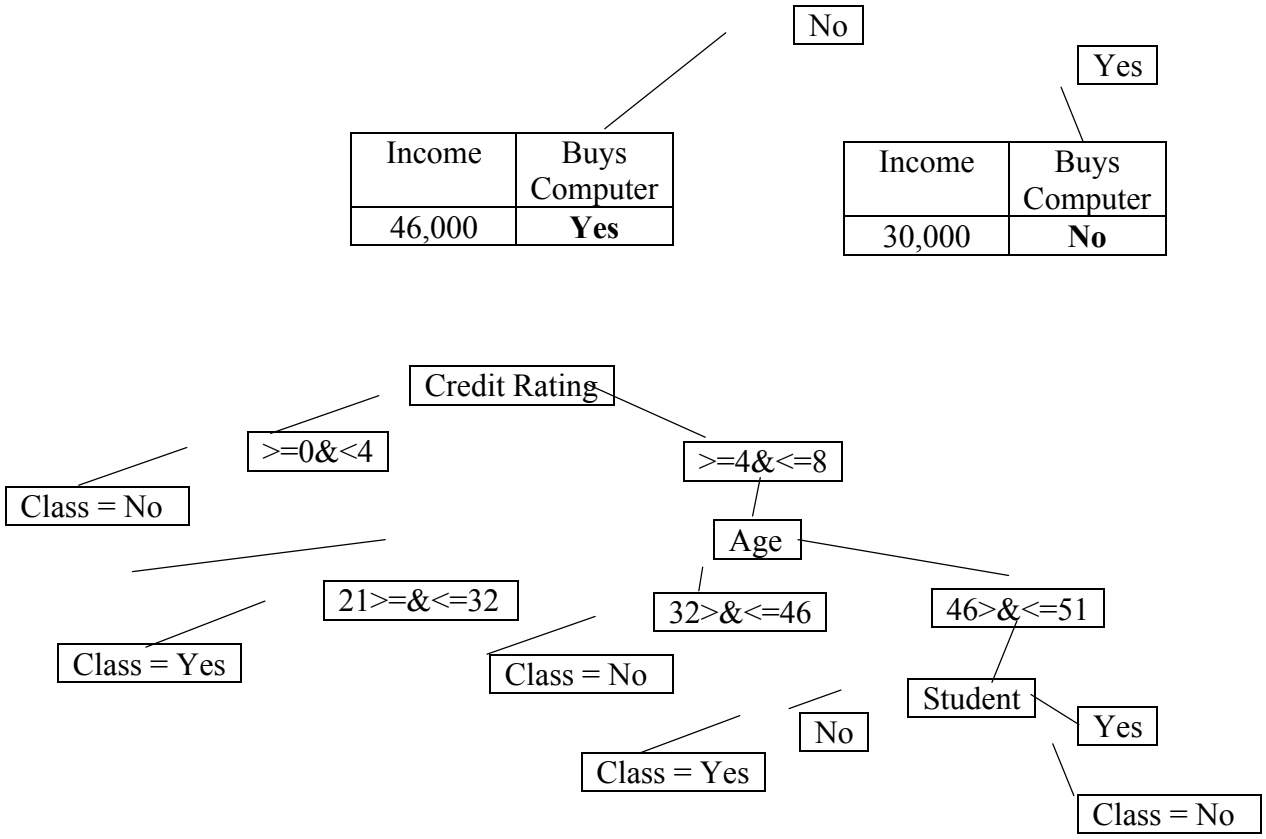
>=0&<4

Age	Income	Student	Buys Computer
29	48,000	Yes	**No**
32	80,000	Yes	**No**
50	75,000	Yes	**No**
48	55125	Yes	**No**

>=4&<=8

Age	Income	Student	Buys Computer
47	30,000	Yes	No
39	28,000	Yes	No
51	46,000	No	Yes
30	38666.667	Yes	Yes
45	50,000	No	No





Rules

R1. Credit Rating(x, >=0&<4) ⇒ buysComputer(x, No)
R2. Credit Rating(x, >=4&<=8) Λ Age(x, >=21&<=32) ⇒ buysComputer(x, Yes)
R3. Credit Rating(x, >=4&<=8) Λ Age(x, >32&<=46) ⇒ buysComputer(x, No)
R4. Credit Rating(x, >=4&<=8) Λ Age(x, >46&<=51) Λ Student(x, No) ⇒ buysComputer(x, Yes)
R5. Credit Rating(x, >=4&<=8) Λ Age(x, >46&<=51) Λ Student(x, Yes) ⇒ buysComputer(x, No)

Testing Set 1

Age	Income	Student	Credit Rating	Buys Computer
21	60,000	Yes	3	No
30	70,000	No	5	No
38	38,666.667	No	2	Yes

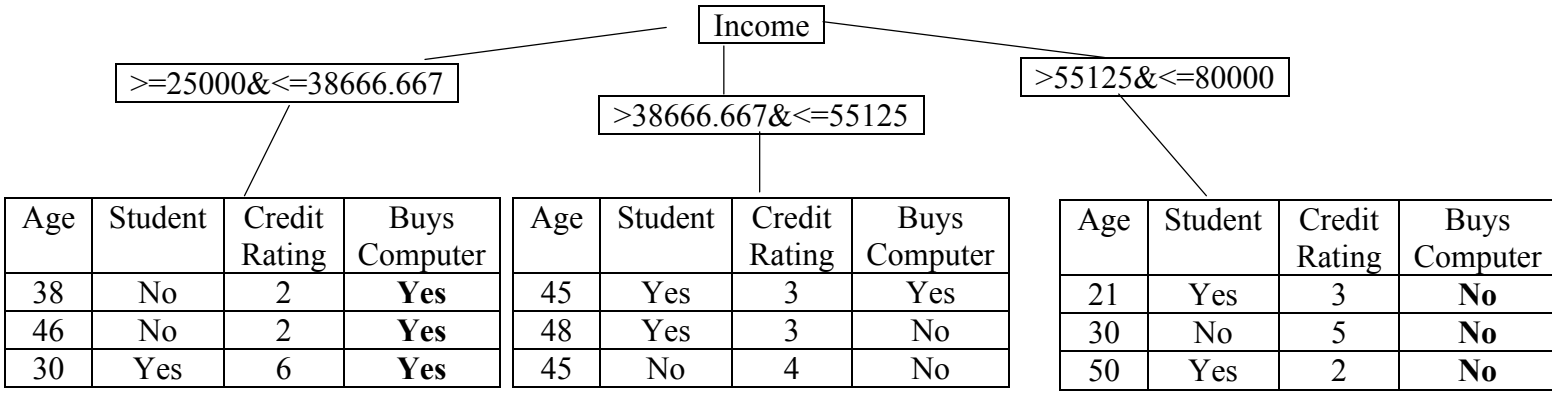
45	45,000	Yes	3	Yes
46	25,000	No	2	Yes

Record 1 is well classified as of Rule 1
Record 2 is misclassified
Record 3 is misclassified
Record 4 is misclassified
Record 5 is misclassified
1 records well classified out of 5

1/5*100%=20%=predictive accuracy

Training 3

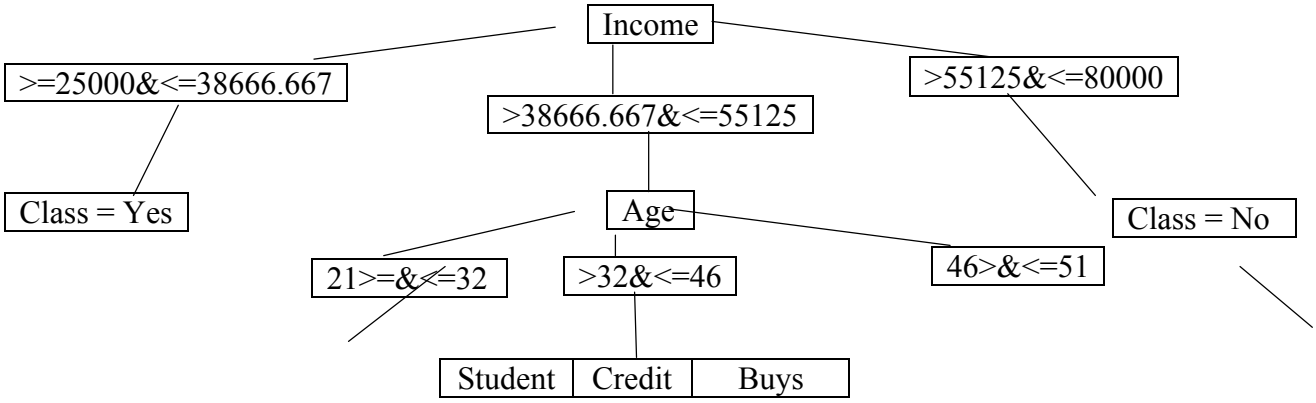
Age	Income	Student	Credit Rating	Buys Computer
21	60,000	Yes	3	No
30	70,000	No	5	No
38	38,666.667	No	2	Yes
45	45,000	Yes	3	Yes
46	25,000	No	2	Yes
50	75,000	Yes	2	No
48	55125	Yes	3	No
30	38666.667	Yes	6	Yes
45	50,000	No	4	No



Age	Student	Credit Rating	Buys Computer
38	No	2	**Yes**
46	No	2	**Yes**
30	Yes	6	**Yes**

Age	Student	Credit Rating	Buys Computer
45	Yes	3	Yes
48	Yes	3	No
45	No	4	No

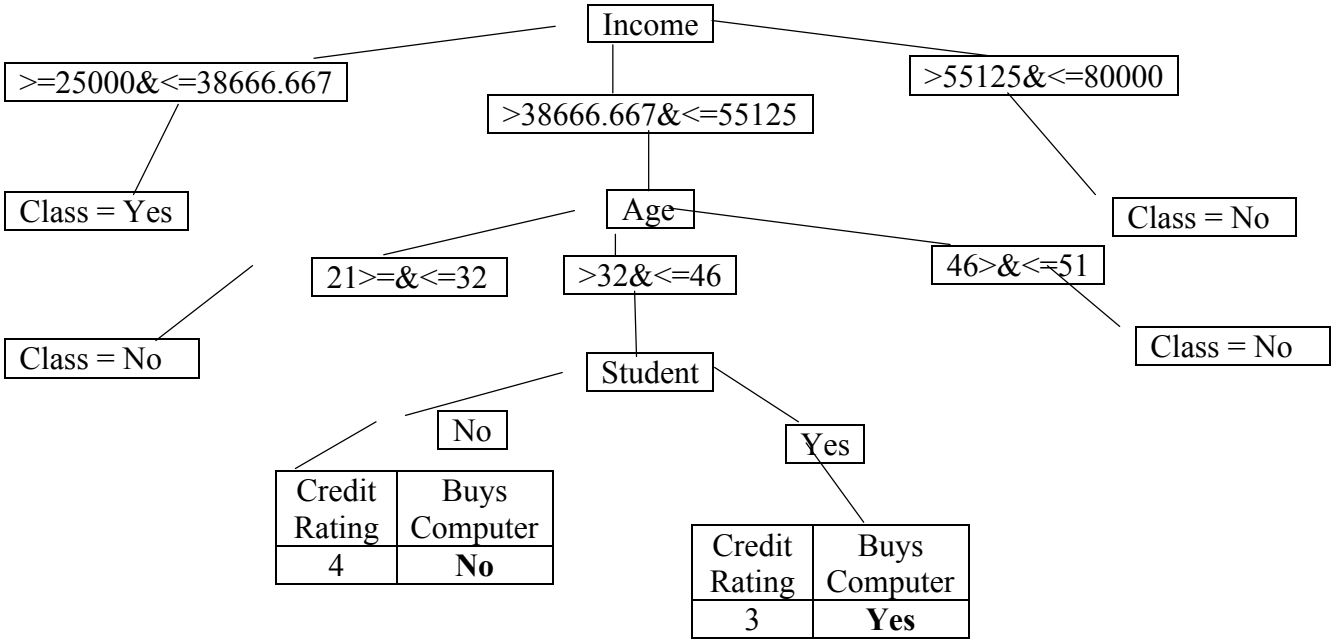
Age	Student	Credit Rating	Buys Computer
21	Yes	3	**No**
30	No	5	**No**
50	Yes	2	**No**

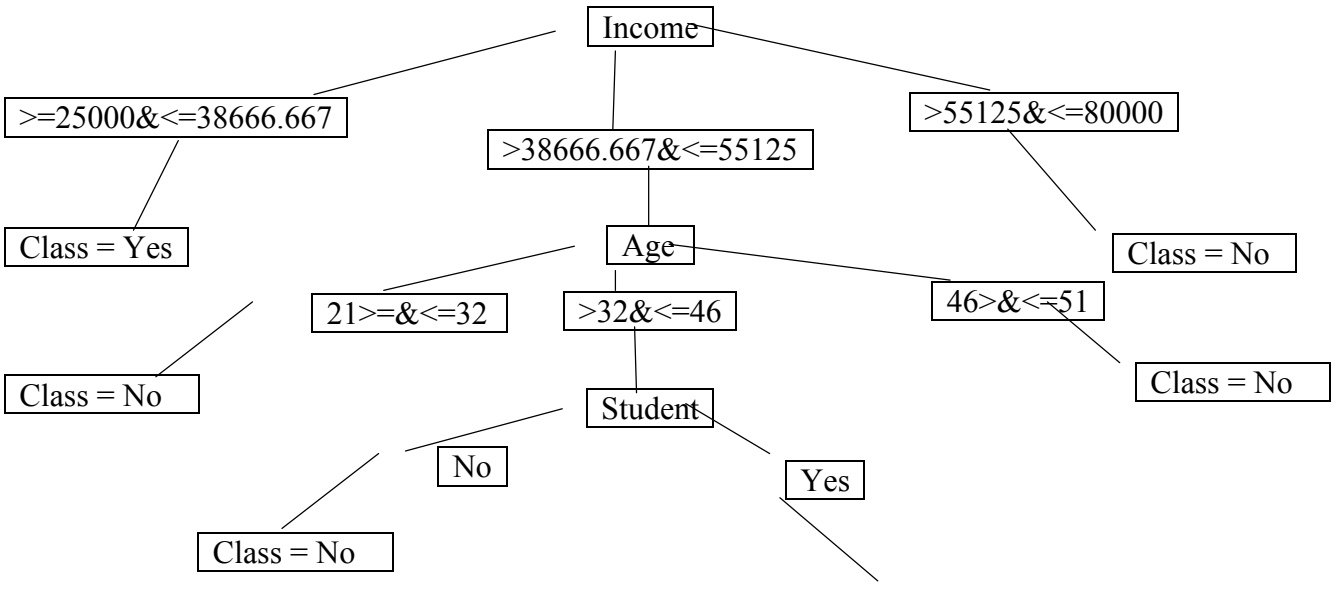


Majority Voting Class=No

	Rating	Computer
Yes	3	Yes
No	4	No

Student	Credit Rating	Buys Computer
Yes	3	**No**





Class = Yes

R1. Income(x, >=25000&<=38666.667) ⇒ buysComputer(x, Yes)
R2. Income(x, >38666.667&<=55125) Λ Age(x, >=21&<=32) ⇒ buysComputer(x, No)
R3. Income(x, >38666.667&<=55125) Λ Age(x, >32&<=46) Λ Student(x, No) ⇒ buysComputer(x, No)
R4. Income(x, >38666.667&<=55125) Λ Age(x, >32&<=46) Λ Student(x, Yes) ⇒ buysComputer(x, Yes)
R5. Income(x, >38666.667&<=55125) Λ Age(x, >46&<=51) ⇒ buysComputer(x, No)
R6. Income(x, >55125&<=80000) ⇒ buysComputer(x, No)

Training Set 2

Age	Income	Student	Credit Rating	Buys Computer
47	30,000	Yes	6	No
39	28,000	Yes	5	No
29	48,000	Yes	3	No
51	46,000	No	4	Yes
32	80,000	Yes	2	No

Record 1 is misclassified
Record 2 is misclassified
Record 3 is misclassified
Record 4 is misclassified
Record 5 is misclassified

0 records well classified out of 5

0/5*100%=0%=predictive accuracy

4. MY CLASSIFIER

Union of Rules – must check for repetitions!

R1. Age(x, >=21&<=32) ⇒ Class = buysComputer(x, No)

R2. Age(x, >32&<=46) Λ Credit Rating(x, >=0&<4) ⇒ Class = buysComputer(x, Yes)

R3. Age(x, >32&<=46) Λ Credit Rating(x, >=4&<=8) ⇒ Class = buysComputer(x, No)

R4. Age(x, >46&<=51) Λ Student(x, No) ⇒ Class = buysComputer(x, Yes)

R5. Age(x, >46&<=51) Λ Student(x, Yes) ⇒ Class = buysComputer(x, No)

R6. Credit Rating(x, >=0&<4) ⇒ buysComputer(x, No)

R7. Credit Rating(x, >=4&<=8) Λ Age(x, >=21&<=32) ⇒ buysComputer(x, Yes)

R8. Credit Rating(x, >=4&<=8) Λ Age(x, >46&<=51) Λ Student(x, No) ⇒ buysComputer(x, Yes)

R9. Credit Rating(x, >=4&<=8) Λ Age(x, >46&<=51) Λ Student(x, Yes) ⇒ buysComputer(x, No)

R10. Income(x, >=25000&<=38666.667) ⇒ buysComputer(x, Yes)

R11. Income(x, >38666.667&<=55125) Λ Age(x, >=21&<=32) ⇒ buysComputer(x, No)

R12. Income(x, >38666.667&<=55125) Λ Age(x, >32&<=46) Λ Student(x, No) ⇒ buysComputer(x, No)

R13. Income(x, >38666.667&<=55125) Λ Age(x, >32&<=46) Λ Student(x, Yes) ⇒ buysComputer(x, Yes)

R14. Income(x, >38666.667&<=55125) Λ Age(x, >46&<=51) ⇒ buysComputer(x, No)

R15. Income(x, >55125&<=80000) ⇒ buysComputer(x, No)

Predictive Accuracy = Training set one predictive accuracy+ Training set two predictive accuracy+ Training set three predictive accuracy)/3

(75%+20%+0%)/3=95%/3=31.667% accuracy

REMARK: Predictive accuracy is very LOW
In Practice one would NOT accept this set of RULES as a CLASSIFIER!!

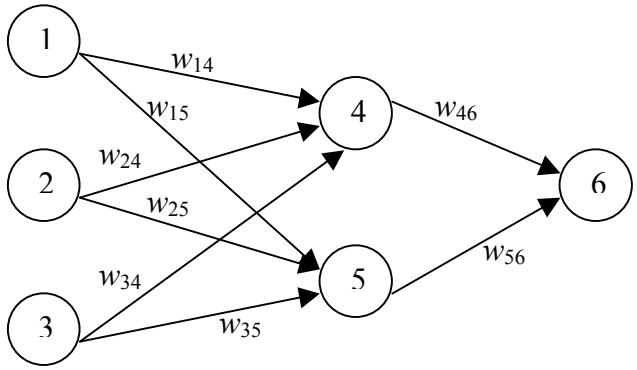
You must START the process over with different preprocessing!!

Problem 3: Learning Neural Networks

Given two records (Training Sample)

a_1	a_2	a_3	Class
0.5	0	0.2	1
0	0.3	0	1

Use the Network below to evaluate a passage of TWO EPOCHS.



Learning Rate: L= 0.7

REMEMBER: YOU HAVE TO SET YOUR INITIAL WEIGHTS AND BIASES RANDOMLY; DON'T USE THE SET-UP FROM THE EXAMPLE.

FIRST EPOCH:

FirstrowofDATA:

a1	a2	a3	w14	w15	w24	w25	w34	w35	w46	w56	θ 4	θ 5	θ 6
0.5	0	0.2	0.1	0.2	0.3	0.4	0.1	0.1	0.5	0.2	0.3	0.2	0.1

$Unit_j$	Net Input I_j	Output O_j
4	(0.5)(0.1)+(0)+(0.2)(0.1)+0.3= 0.37	$1/(1+e^{-0.37})$ = 0.591
5	(0.5)(0.2)+(0)+(0.2)(0.1)+0.2= 0.32	$1/(1+e^{-0.32})$ = 0.579
6	(0.5)(0.591)+(0.2)(0.579)+0.1= .511	$1/(1+e^{-0.511})$ = 0.625

Error measurement:

$Unit_j$	$Error_j$
6	0.625(1-0.625)(1-0.625)=0.0878 We assume T_6=1
5	0.579(1-0.579)(0.0878)(0.2)=0.0043
4	0.591(1-0.591)(0.0878)(0.5)=0.0153

New adjusted values:

Weight	New Values
w46	0.5 + .7(0.0878)(.591) =0.536
w56	0.2 + .7(0.0878)(.579) =0.236
w34	0.1+ .7(0.0153)(.2) = 0.102
w35	0.1+ .7(0.0153)(.2) =0.102
w24	0.3+ .7(0.0153)(0) =0.3

w25	0.4+ .7(0.0043)(0) =0.4
w14	0.1+ .7(0.0153)(0.5) =0.105
w15	0.2+ .7(0.0043)(0.5) =0.202
θ 4	0.3 + .7(0.0153)=0.311
θ 5	0.2+ .7(0.0043)=0.203
□6	0.1+ .7(0.0878)=0.161

SecondRowofDATA:

a1	a2	a3	w14	w15	w24	w25	w34	w35	w46	w56	θ 4	θ 5	θ 6
0	0.3	0	0.105	0.202	0.3	0.4	0.102	0.102	0.536	0.236	0.311	0.203	0.161

$Unit_j$	Net Input I_j	Output O_j
4	(0)+(0.3)(0.3)+(0)+0.308= 0.398	$1/(1+e^{-0.398})$ = 0.598
5	(0)+(0.3)(0.4)+(0)+0.203= 0.323	$1/(1+e^{-0.323})$ = 0.58
6	(.536)(.598)+(.236)(.58)+0.164=0.62	$1/(1+e^{-0.62})$ = 0.65

Error measurement:

$Unit_j$	$Error_j$
6	0.65(1-0.65)(1-0.65)=0.0796 We assume T_6=1
5	0.58(1-0.58)(0.0796)(0.236)=0.0046
4	0.598(1-0.598)(0.0796)(0.536)=0.0103

New adjusted values:

Weight	New Values
w46	0.536 + .7(0.0796)(.598) =0.569
w56	0.236 + .7(0.0796)(.58) =0.268
w34	0.102 + .7(0.0103)(0) = 0.102
w35	0.102 + .7(0.0046)(0) =0.102
w24	0.3 + .7(0.0103)(0.3) =0.302
w25	0.4 + .7(0.0046)(0.3) =0.401
w14	0.105 + .7(0.0103)(0) =0.105
w15	0.202 + .7(0.0046)(0) =0.202
□4	0.311 + .7(0.0103)=0.318
□5	0.203 + .7(0.0046)=0.206
□6	0.161 + .7(0.0796)=0.217

SECOND EPOCH:

FirstrowofDATA:

a1	a2	a3	w14	w15	w24	w25	w34	w35	w46	w56	θ 4	θ 5	θ 6
0.5	0	0.2	0.105	0.202	0.302	0.401	0.102	0.102	0.569	0.268	0.318	0.206	0.217

$Unit_j$	Net Input I_j	Output O_j
4	(0.5)(0.105)+(0)+(0.2)(0.102)+0.318= 0.391	$1/(1+e^{-0.391}) = 0.597$
5	(0.5)(0.202)+(0)+(0.2)(0.102)+0.206= 0.327	$1/(1+e^{-0.327}) = 0.581$
6	(0.5)(0.597)+(0.2)(0.581)+0.217= 0.632	$1/(1+e^{-0.632}) = 0.653$

Error measurement:

$Unit_j$	$Error_j$
6	0.653(1-0.653)(1-0.653)=0.079 We assume T_6=1
5	0.581(1-0.581)(0.079)(0.268)=0.0052
4	0.597(1-0.597)(0.079)(0.569)=0.0108

New adjusted values:

Weight	New Values
w46	0.569 + .7(0.079)(.597) =0.602
w56	0.268 + .7(0.079)(.581) =0.3
w34	0.102 + .7(0.0108)(.2) = 0.104
w35	0.102 + .7(0.0052)(.2) =0.103
w24	0.302 + .7(0.0108)(0) =0.302
w25	0.401 + .7(0.0052)(0) =0.401
w14	0.105 + .7(0.0108)(0.5) =0.109
w15	0.202 + .7(0.0052)(0.5) =0.204
θ 4	0.318 + .7(0.0108)=0.326
θ 5	0.206 + .7(0.0052)=0.210
θ 6	0.217 + .7(0.079)=0.272

SecondRowofDATA:

a1	a2	a3	w14	w15	w24	w25	w34	w35	w46	w56	□4	□5	□6
0	0.3	0	0.019	0.204	0.302	0.401	0.104	0.103	0.602	0.3	0.326	0.210	0.262

$Unit_j$	Net Input I_j	Output O_j
4	(0)+(0.3)(0.302)+(0)+0.326= 0.417	$1/(1+e^{-0.417}) = 0.603$
5	(0)+(0.3)(0.401)+(0)+0.210= 0.330	$1/(1+e^{-0.330}) = 0.582$
6	(.602)(.603)+(.3)(.582)+0.262=0.8	$1/(1+e^{-0.8}) = 0.69$

Error measurement:

$Unit_j$	$Error_j$
6	0.69(1-0.69)(1-0.69)=0.066 We assume T_6=1
5	0.582(1-0.582)(0.066)(0.3)=0.0048
4	0.69(1-0.69)(0.066)(0.602)=0.0085

New adjusted values:

Weight	New Values
w46	0.602 + .7(0.066)(.603) =0.569
w56	0.3 + .7(0.099)(.582) =0.268
w34	0.104 + .7(0.0085)(0) = 0.102
w35	0.103 + .7(0.0048)(0) =0.102
w24	0.302 + .7(0.0085)(0.3) =0.304
w25	0.401 + .7(0.0048)(0.3) =0.402
w14	0.019 + .7(0.0085)(0) =0.105
w15	0.204 + .7(0.0048)(0) =0.202
θ4	0.326 + .7(0.0085)=0.332
θ 5	0.210 + .7(0.0048)=0.213
θ 6	0.262 + .7(0.066)=0.308

PROBLEM 4: Classification by Association

Given a classification TRAINING data

Income	Student	Rating
high	No	Fair
low	No	Excellent
high	yes	Fair
medium	No	Fair
low	Yes	Fair
medium	no	Excellent

and classification TEST data

Income	Student	Rating
low	Yes	Excellent
medium	No	Fair
low	no	Fair
medium	Yes	Fair

Build a CLASSIFIER using Classification by Association Method

STEP 1: TRANSLATION – DATA TRANSFORMATION

Income = Low (I1)	Income = medium (I2)	Income = High (I3)	Student = No (I4)	Student = Yes (I5)	Rating = Fair (I6)	Rating = Excellent (I7)
-	-	+	+	-	+	-
+	-	-	+	-	-	+
-	-	+	-	+	+	-

-	+	-	+	-	+	-
+	-	-	-	+	+	-
-	+	-	+	-	-	+

Let minimum support count be 2,
Since we have 6 records min support = 2/6 = 30%

Let minimum confidence required be 70%

Generating 1-item set frequent patterns

Item Set	Support count
I1	2
I2	2
I3	2
I4	4
I5	2
I6	4
I7	2

C1 = L1

All sc >= min sc

Item Set	Support count
{I1, I2}	0
{I1, I3}	0
{I1, I4}	1
{I1, I5}	1

{I1, I6}	**1**
{I1, I7}	**1**
{I2, I3}	**0**
{I2, I4}	**2**
{I2, I5}	**0**
{I2, I6}	**1**
{I2, I7}	**1**
{I3, I4}	**1**
{I3, I5}	**1**
{I3, I6}	**2**
{I3, I7}	**0**
{I4, I5}	**0**
{I4, I6}	**2**
{I4, I7}	**2**
{I5, I6}	**2**
{I5, I7}	**0**
{I6, I7}	**0**

C2

Choose candidates with sc >= min sc

Item Set	**Support count**
{I2, I4}	**2**
{I3, I6}	**2**
{I4, I6}	**2**

{I4, I7}	**2**
{I5, I6}	**2**

L2

Joining & pruning

Join L2*L2 to obtain C3

NO C3 Candidates

FOR Classification by Association we only Consider I={IX,IY} Where IY represents CLASS ATTRIBUTES:

Rating = Fair or Rating = Excellent

Item Set	Support count
{I3, I6}	**2**
{I4, I6}	**2**
{I4, I7}	**2**
{I5, I6}	**2**

$R1: 3 \rightarrow 6$

$$\text{Confidence} = \frac{sc\{I3,I6\}}{sc\{I3\}} = \frac{2}{2} = 100\%$$

R1 is selected

$R2: 4 \rightarrow 6$

$$\text{Confidence} = \frac{sc\{I4,I6\}}{sc\{I4\}} = \frac{2}{4} = 50\%$$

R2 is not selected

$R3: 4 \rightarrow 7$

$$\text{Confidence} = \frac{sc\{I4,I7\}}{sc\{I4\}} = \frac{2}{4} = 50\%$$

R3 is not selected

$R4: 5 \rightarrow 6$

$$\text{Confidence} = \frac{sc\{I5,I6\}}{sc\{I5\}} = \frac{2}{2} = 100\%$$

R4 is selected

List of selected Classification by Association Rules

$R1: 3 \rightarrow 6$

$R4: 5 \rightarrow 6$

Classification by Association Rules for TESTING

R1: IF Income(x, High) THEN Rating(x, Fair) [30, 70]
R2: IF Student(x, yes) THEN Rating(x, Fair) [30, 70]

TEST data

Income	Student	Rating
low	Yes	Excellent
medium	No	Fair
low	no	Fair
medium	Yes	Fair

Record 1 –error, record 2- error, record 3- error,
record 4 – **correctly classified by R2**

Predictive accuracy 25% and 75% error rate

THE CLASSIFIER: - There is no Classifier
Predictive Accuracy with given test data
is too low to Accept the Rules R1, R2
as Classifier

START the Process OVER with different set of paramethers (or DATA) if you have a choice