Solutions for Tutorial exercises
## Backpropagation neural networks, Naïve Bayes, Decision Trees, k-NN, Associative Classification.

Exercise 1.
Suppose we want to classify potential bank customers as good creditors or bad creditors for loan applications. We have a training dataset describing past customers using the following attributes: Marital status {married, single, divorced}, Gender {male, female}, Age {[18..30[, [30..50[, [50..65[, [65+]}, Income {[10K..25K[, [25K..50K[, [50K..65K[, [65K..100K[, [100K+]}.
Design a neural network that could be trained to predict the credit rating of an applicant.

Solution:
We have 2 classes, good creditor and bad creditor. This means we would need two nodes in the output layer. There are 4 variables: Marital Status, Gender, Age and Income. However, since we have 3 values for Marital status, 2 values for Gender, 4 intervals for Age and 5 intervals for Income, we would have 14 neuron units in the input layer. In the hidden layer we can have (14+2)/2=8 neurons. The architecture of the neural networks could look like this.
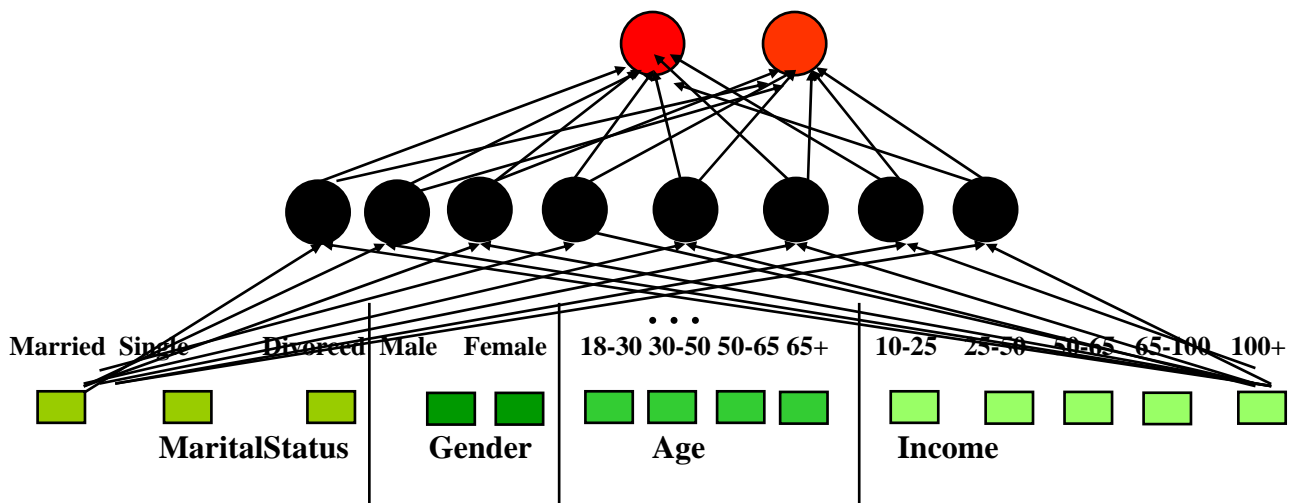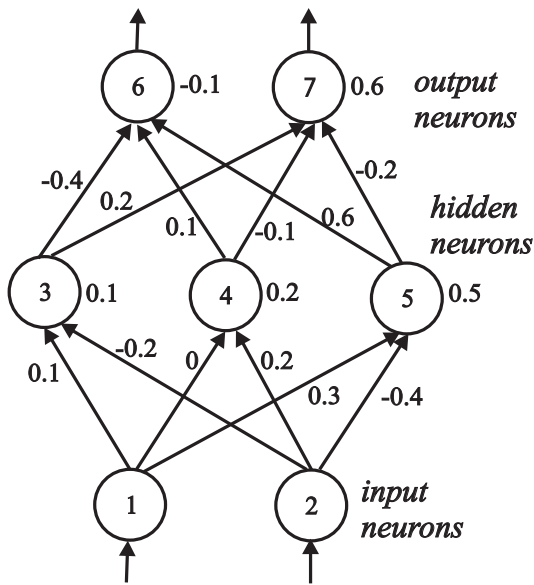

The weights are initialized with random values.

Are there other possible architectures?

Given the following neural network with initialized weights as in the picture, explain the network architecture knowing that we are trying to distinguish between nails and screws and an example of training tupples is as follows: T1{0.6, 0.1, nail}, T2{0.2, 0.3, screw}.

Let the learning rate ɳ be 0.1 and the weights be as indicated in the figure above. Do the forward propagation of the signals in the network using T1 as input, then perform the back propagation of the error. Show the changes of the weights.

Solution:

What encoding of the outputs?
10 for class "nail", 01 for class "screw"

Forward pass for T1 - calculate the outputs o6 and o7

    o1=0.6, o2=0.1, target output 1 0, i.e. class "nail"
Activations of the hidden units:
net3= o1 *w13+ o2*w23+b3=0.6*0.1+0.1*(-0.2)+0.1=0.14
o3=1/(1+e-net3) =0.53

net4= o1 *w14+ o2*w24+b4=0.6*0+0.1*0.2+0.2=0.22
o4=1/(1+e-net4) =0.55

net5= o1 *w15+ o2*w25+b5=0.6*0.3+0.1*(-0.4)+0.5=0.64
o5=1/(1+e-net5) =0.65

Activations of the output units:
net6= o3 *w36+ o4*w46+ o5*w56 +b6=0.53*(-0.4)+0.55*0.1+0.65*0.6-0.1=0.13
o6=1/(1+e-net6) =0.53

net7= o3 *w37+ o4*w47+ o5*w57 +b7=0.53*0.2+0.55*(-0.1)+0.65*(-0.2)+0.6=0.52
o7=1/(1+e-net7) =0.63

Backward pass for T1 - calculate the output errors δ6 and δ7
(note that d6=1, d7=0 for class "nail")

$\delta 6 = (d6\text{-}o6) * o6 * (1\text{-}o6)=(1\text{-}0.53)*0.53*(1\text{-}0.53)=0.12$

$\delta 7 = (d7\text{-}o7) * o7 * (1\text{-}o7)=(0\text{-}0.63)*0.63*(1\text{-}0.63)=\text{-}0.15$

<u>Calculate the new weights between the hidden and output units ($\eta=0.1$)</u>

$\Delta w36 = \eta * \delta 6 * o3 = 0.1*0.12*0.53=0.006$

$w36new = w36old + \Delta w36 = \text{-}0.4+0.006=\text{-}0.394$

$\Delta w37 = \eta * \delta 7 * o3 = 0.1*\text{-}0.15*0.53=\text{-}0.008$

$w37new = w37old + \Delta w37 = 0.2\text{-}0.008=\text{-}0.19$

Similarly for w46new , w47new , w56new and w57new

For the biases b6 and b7 (remember: biases are weights with input 1):

$\Delta b6 = \eta * \delta 6 * 1 = 0.1*0.12=0.012$

$b6new = b6old + \Delta b6 = \text{-}0.1+0.012=\text{-}0.012$

Similarly for b7

<u>Calculate the errors of the hidden units $\delta 3$, $\delta 4$ and $\delta 5$</u>

$\delta 3 = o3 * (1\text{-}o3) * (w36* \ \delta 6 +w37 * \ \delta 7) = \ 0.53*(1\text{-}0.53)(\text{-}0.4*0.12+0.2*(\text{-}0.15))=\text{-}0.019$

Similarly for $\delta 4$ and $\delta 5$

<u>Calculate the new weights between the input and hidden units ($\eta=0.1$)</u>

$\Delta w13 = \eta * \delta 3 * o1 = 0.1*(\text{-}0.019)*0.6=\text{-}0.0011$

$w13new = w13old + \Delta w13 = 0.1\text{-}0.0011=0.0989$

Similarly for w23new , w14new , w24new , w15new and w25new; b3, b4 and b6

Repeat the same procedure for the other training examples

Forward pass for T2…backward pass for T2…

Exercise 3.

Why is the Naïve Bayesian classification called "naïve"?

Answer: Naïve Bayes assumes that all attributes are: 1) equally important and 2) independent of one another given the class.

Exercise 4. *Naïve Bayes for data with nominal attributes*
Given the training data in the table below (*Buy Computer* data), predict the class of the following new example using Naïve Bayes classification: age<=30, income=medium, student=yes, credit-rating=fair

RID	age	income	student	credit_rating	Class: buys_computer
1	<=30	high	no	fair	no
2	<=30	high	no	excellent	no
3	31 ... 40	high	no	fair	yes
4	>40	medium	no	fair	yes
5	>40	low	yes	fair	yes
6	>40	low	yes	excellent	no
7	31 ... 40	low	yes	excellent	yes
8	<=30	medium	no	fair	no
9	<=30	low	yes	fair	yes
10	>40	medium	yes	fair	yes
11	<=30	medium	yes	excellent	yes
12	31 ... 40	medium	no	excellent	yes
13	31 ... 40	high	yes	fair	yes
14	>40	medium	no	excellent	no

Solution:

E= age<=30, income=medium, student=yes, credit-rating=fair
E_1 is age<=30, E2 is income=medium, student=yes, E4 is credit-rating=fair
We need to compute P(yes|E) and P(no|E) and compare them.

$$P(yes \mid E) = \frac{P(E_1 \mid yes)\,P(E_2 \mid yes)\,P(E_3 \mid yes)\,P(E_4 \mid yes)\,P(yes)}{P(E)}$$

P(yes)=9/14=0.643   P(no)=5/14=0.357

P(E1|yes)=2/9=0.222   P(E1|no)=3/5=0.6
P(E2|yes)=4/9=0.444   P(E2|no)=2/5=0.4
P(E3|yes)=6/9=0.667   P(E3|no)=1/5=0.2
P(E4|yes)=6/9=0.667   P(E4|no)=2/5=0.4

$$P(yes \mid E) = \frac{0.222\ 0.444\ 0.667\ 0.668\ 0.443}{P(E)} = \frac{0.028}{P(E)} \qquad P(no \mid E) = \frac{0.6\ 0.4\ 0.2\ 0.4\ 0.357}{P(E)} = \frac{0.007}{P(E)}$$

Hence, the Naïve Bayes classifier predicts buys_computer=yes for the new example.

Exercise 5. *Applying Naïve Bayes to data with numerical attributes and using the Laplace correction (to be done at your own time, not in class)*
Given the training data in the table below (*Tennis* data with some numerical attributes), predict the class of the following new example using Naïve Bayes classification:
outlook=overcast, temperature=60, humidity=62, windy=false.

Tip. You can use Excel or Matlab for the calculations of logarithm, mean and standard deviation. Matlab is installed on our undergraduate machines. The following Matlab functions can be used: `log2` – logarithm with base 2, `mean` – mean value, `std` – standard deviation. Type `help` <function name> (e.g. `help mean`) for help on how to use the functions and examples.

outlook	temperature	humidity	windy	play
sunny	85	85	false	no
sunny	80	90	true	no
overcast	83	86	false	yes
rainy	70	96	false	yes
rainy	68	80	false	yes
rainy	65	70	true	no
overcast	64	65	true	yes
sunny	72	95	false	no
sunny	69	70	false	yes
rainy	75	80	false	yes
sunny	75	70	true	yes
overcast	72	90	true	yes
overcast	81	75	false	yes
rainy	71	91	true	no

Solution:

First, we need to calculate the mean μ and standard deviation σ values for the numerical attributes. X_i, i=1..n – the i-th measurement, n-number of measurements

$$\mu = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\sigma^2 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n-1}$$

$\mu\_temp\_yes=73$, $\sigma\_temp\_yes=6.2$; $\mu\_temp\_no=74.6$, $\sigma\_temp\_no=8.0$

$\mu\_hum\_yes=79.1$, $\sigma\_temp\_yes=10.2$; $\mu\_hum\_no=86.2$, $\sigma\_temp\_no=9.7$

Second, to calculate f(temperature=60|yes), f(temperature=60|no), f(humidity=62|yes) and f(humidity=62|no) using the probability density function for the normal distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$f(temperature = 60 \mid yes) = \frac{1}{6.2\sqrt{2\pi}} e^{-\frac{(60-73)^2}{2\,(6.2)^2}} = 0.071$$

$$f(temperature = 60 \mid no) = \frac{1}{8\sqrt{2\pi}} e^{-\frac{(60-74.6)^2}{2\,8^2}} = 0.0094$$

$$f(humidity = 62 \mid yes) = \frac{1}{10.2\sqrt{2\pi}} e^{-\frac{(62-79.1)^2}{2\,(10.2)^2}} = 0.0096$$

$$f(humidity = 62 \mid no) = \frac{1}{9.7\sqrt{2\pi}} e^{-\frac{(62-86.2)^2}{2\,(9.7)^2}} = 0.0018$$

Third, we can calculate the probabilities for the nominal attributes:
P(yes)=9/14=0.643   P(no)=5/14=0.357

P(outlook=overcast|yes)=4/14=0.286   P(outlook=overcast|no)=0/5=0
P(windy=false|yes)=6/9=0.667   P(windy=false|no)=2/5=0.4

As P(outlook=overcast|no)=0, we need to use a Laplace estimator for the attribute outlook. We assume that the three values (sunny, overcast, rainy) are equally probable and set μ=3:

$$P(outlook = overcast \mid yes) = \frac{4+1}{9+3} = \frac{5}{12} = 0.4167$$

$$P(outlook = overcast \mid no) = \frac{0+1}{5+3} = \frac{1}{8} = 0.125$$

Fourth, we can calculate the final probabilities:

$$P(yes \mid E) = \frac{0.4167 * 0.0071 * 0.0096 * 0.667 * 0.643}{P(E)} = \frac{1.22 * 10^{-5}}{P(E)}$$

$$P(no \mid E) = \frac{0.125 * 0.0094 * 0.0018 * 0.4 * 0.357}{P(E)} = \frac{3.02 * 10^{-7}}{P(E)}$$

Therefore, the Naïve Bayes classifier predicts play=yes for the new example.

Exercise 6. Using Weka (to be done at your own time, not in class)
Load iris data (iris.arff). Choose 10-fold cross validation. Run the Naïve Bayes and Multi-layer percepton (trained with the backpropagation algorithm) classifiers and compare their performance. Which classifier produced the most accurate classification? Which one learns faster?

Exercise 7. k-Nearest neighbours
Given the training data in Exercise 4 (Buy Computer data), predict the class of the following new example using k-Nearest Neighbour for k=5: age<=30, income=medium, student=yes, credit-rating=fair. For similarity measure use a simple match of attribute values: Similarity(A,B)=

$$\sum_{i=1}^{4} w_i * \partial(a_i, b_i) \Big/ 4$$ where $\partial(a_i, b_i)$ is 1 if a_i equals b_i and 0 otherwise. a_i and b_i are either *age, income, student* or *credit_rating*. Weights are all 1 except for income it is 2.

Solution:

RID	age	income	student	credit_rating	Class: buys_computer
1	<=30	high	no	fair	no
2	<=30	high	no	excellent	no
3	31 ... 40	high	no	fair	yes
4	>40	medium	no	fair	yes
5	>40	low	yes	fair	yes
6	>40	low	yes	excellent	no
7	31 ... 40	low	yes	excellent	yes
8	<=30	medium	no	fair	no
9	<=30	low	yes	fair	yes
10	>40	medium	yes	fair	yes
11	<=30	medium	yes	excellent	yes
12	31 ... 40	medium	no	excellent	yes
13	31 ... 40	high	yes	fair	yes
14	>40	medium	no	excellent	no

RID	Class	Distance to New
1	No	(1+0+0+1)/4=0.5
2	No	(1+0+0+0)/4=0.25
3	Yes	(0+0+0+1)/4=0.25
4	**Yes**	**(0+2+0+1)/4=0.75**
5	Yes	(0+0+1+1)/4=0.5
6	No	(0+0+1+0)/4=0.25
7	Yes	(0+0+1+0)/4=0.25
8	**No**	**(1+2+0+1)/4=1**
9	**Yes**	**(1+0+1+1)/4=0.75**
10	**Yes**	**(0+2+1+1)/4=1**
11	**Yes**	**(1+2+1+0)/4=1**
12	Yes	(0+2+0+0)/4=0.5
13	Yes	(0+0+1+1)/4=0.5
14	No	(0+2+0+0)/4=0.5

Among the five nearest neighbours four are from class *Yes* and one from class *No*. Hence, the k-NN classifier predicts buys_computer=yes for the new example.

Exercise 8. Decision trees
Given the training data in Exercise 4 (*Buy Computer* data), build a decision tree and predict the class of the following new example: age<=30, income=medium, student=yes, credit-rating=fair.

Solution:
First check which attribute provides the highest Information Gain in order to split the training set based on that attribute. We need to calculate the expected information to classify the set and the entropy of each attribute. The information gain is this mutual information minus the entropy:

The mutual information of the two classes $I(S_{Yes}, S_{No}) = I(9,5) = -9/14 \log_2(9/14) - 5/14 \log_2(5/14) = 0.94$

- For Age we have three values $age_{<=30}$ (2 yes and 3 no), $age_{31..40}$ (4 yes and 0 no) and $age_{>40}$ (3 yes 2 no)

Entropy(age) = 5/14 (-2/5 log(2/5)-3/5log(3/5)) + 4/14 (0) + 5/14 (-3/5log(3/5)-2/5log(2/5))
= 5/14(0.9709) + 0 + 5/14(0.9709)
= 0.6935
Gain(age) = 0.94 – 0.6935 = 0.2465

- For Income we have three values $income_{high}$ (2 yes and 2 no), $income_{medium}$ (4 yes and 2 no) and $income_{low}$ (3 yes 1 no)

Entropy(income) = 4/14(-2/4log(2/4)-2/4log(2/4)) + 6/14 (-4/6log(4/6)-2/6log(2/6))
+ 4/14 (-3/4log(3/4)-1/4log(1/4))
= 4/14 (1) + 6/14 (0.918) + 4/14 (0.811)
= 0.285714 + 0.393428 + 0.231714 = 0.9108

Gain(income) = 0.94 – 0.9108 = 0.0292

- For Student we have two values $student_{yes}$ (6 yes and 1 no) and $student_{no}$ (3 yes 4 no)

Entropy(student) = 7/14(-6/7log(6/7)) + 7/14(-3/7log(3/7)-4/7log(4/7)
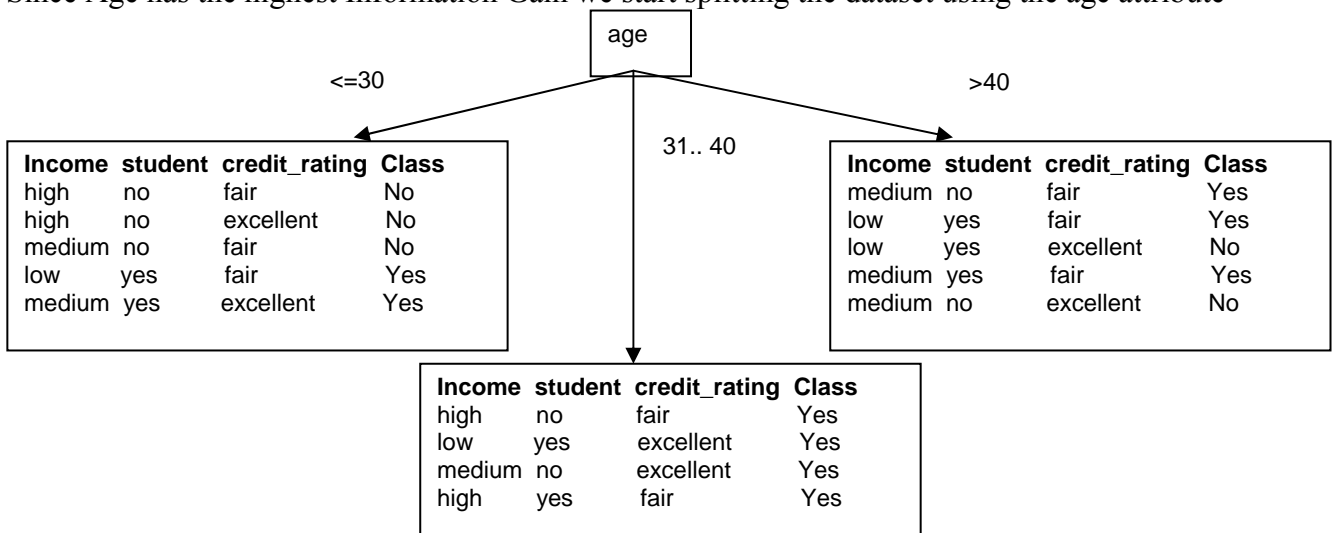= 7/14(0.5916) + 7/14(0.9852)
= 0.2958 + 0.4926 = 0.7884

Gain (student) = 0.94 – 0.7884 = 0.1516

- For Credit_Rating we have two values $credit\_rating_{fair}$ (6 yes and 2 no) and $credit\_rating_{excellent}$ (3 yes 3 no)

Entropy(credit_rating) = 8/14(-6/8log(6/8)-2/8log(2/8)) + 6/14(-3/6log(3/6)-3/6log(3/6))
= 8/14(0.8112) + 6/14(1)
= 0.4635 + 0.4285 = 0.8920

Gain(credit_rating) = 0.94 – 0.8920 = 0.479

Since Age has the highest Information Gain we start splitting the dataset using the age attribute

Income	student	credit_rating	Class
high	no	fair	No
high	no	excellent	No
medium	no	fair	No
low	yes	fair	Yes
medium	yes	excellent	Yes

Income	student	credit_rating	Class
medium	no	fair	Yes
low	yes	fair	Yes
low	yes	excellent	No
medium	yes	fair	Yes
medium	no	excellent	No

Income	student	credit_rating	Class
high	no	fair	Yes
low	yes	excellent	Yes
medium	no	excellent	Yes
high	yes	fair	Yes

Since all records under the branch $age_{31..40}$ are all of class Yes, we can replace the leaf with Class=Yes
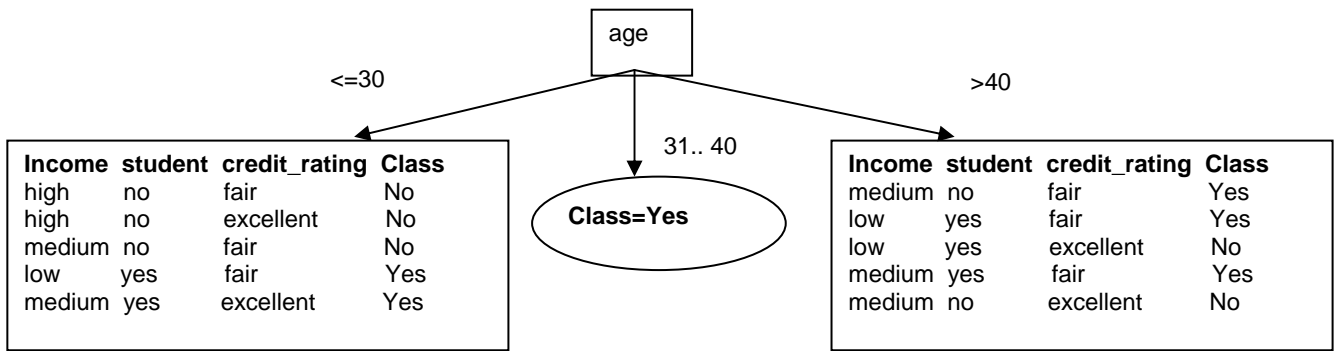
The same process of splitting has to happen for the two remaining branches.

For branch age$_{<=30}$ we still have attributes income, student and credit_rating. Which one should be use to split the partition?

The mutual information is I(S$_{Yes}$, S$_{No}$)= I(2,3)= -2/5 log$_2$(2/5) – 3/5 log$_2$(3/5)=0.97

- For Income we have three values income$_{high}$ (0 yes and 2 no), income$_{medium}$ (1 yes and 1 no) and income$_{low}$ (1 yes and 0 no)

Entropy(income) = 2/5(0) + 2/5 (-1/2log(1/2)-1/2log(1/2)) + 1/5 (0)
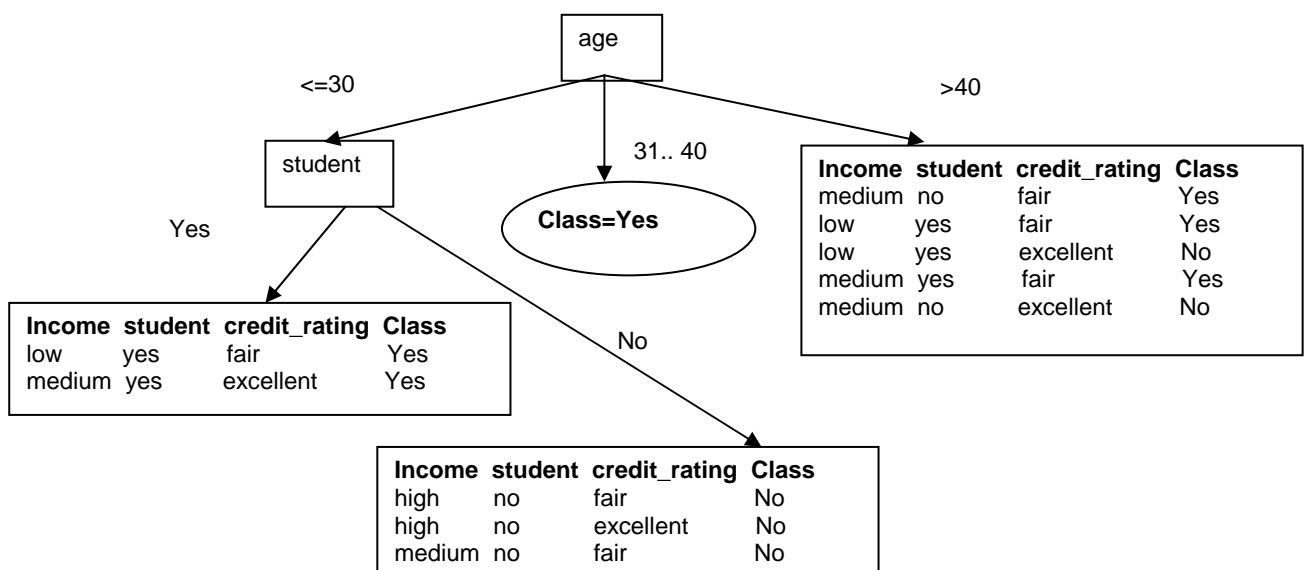$\qquad\qquad$ = 2/5 (1) = 0.4

Gain(income) = 0.97 – 0.4 = 0.57

- For Student we have two values student$_{yes}$ (2 yes and 0 no) and student$_{no}$ (0 yes 3 no)

Entropy(student) = 2/5(0) + 3/5(0) = 0

Gain (student) = 0.97 – 0 = 0.97

We can then safely split on attribute student without checking the other attributes since the information gain is maximized.

Since these two new branches are from distinct classes, we make them into leaf nodes with their respective class as label:
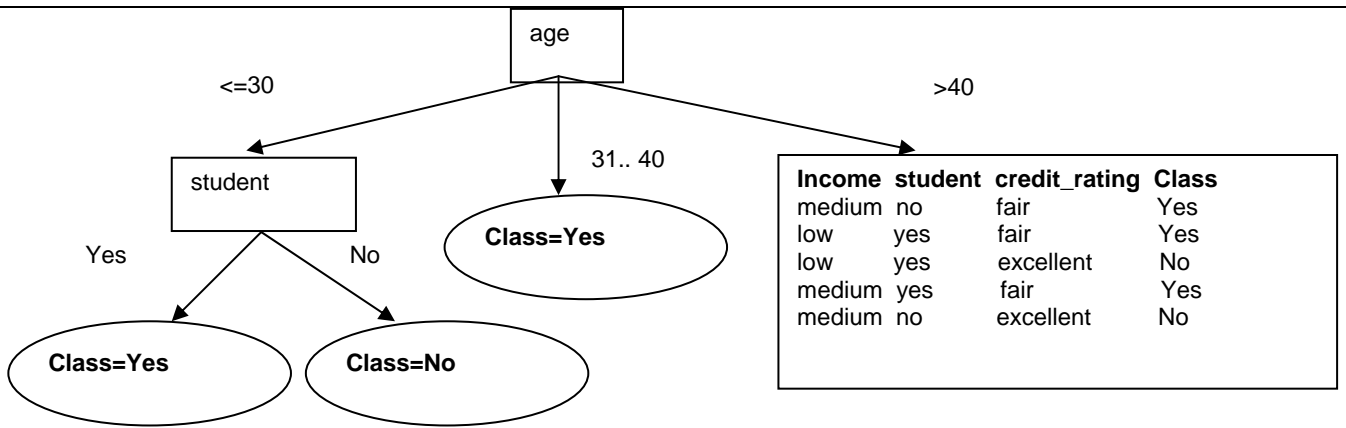
Again the same process is needed for the other branch of age.

The mutual information is $I(S_{Yes}, S_{No}) = I(3,2) = -3/5 \log_2(3/5) - 2/5 \log_2(2/5) = 0.97$

- For Income we have two values $income_{medium}$ (2 yes and 1 no) and $income_{low}$ (1 yes and 1 no)

Entropy(income) = $3/5(-2/3\log(2/3)-1/3\log(1/3)) + 2/5 (-1/2\log(1/2)-1/2\log(1/2))$
= $3/5(0.9182)+2/5 (1) = 0.55+0.4 = 0.95$

Gain(income) = 0.97 – 0.95 = 0.02

- For Student we have two values $student_{yes}$ (2 yes and 1 no) and $student_{no}$ (1 yes and 1 no)

Entropy(student) = $3/5(-2/3\log(2/3)-1/3\log(1/3)) + 2/5(-1/2\log(1/2)-1/2\log(1/2)) = 0.95$

Gain (student) = 0.97 – 0.95 = 0.02

- For Credit_Rating we have two values $credit\_rating_{fair}$ (3 yes and 0 no) and $credit\_rating_{excellent}$ (0 yes and 2 no)

Entropy(credit_rating) = 0

Gain(credit_rating) = 0.97 – 0 = 0.97

We then split based on credit_rating. These splits give partitions each with records from the same class. We just need to make these into leaf nodes with their class label attached:
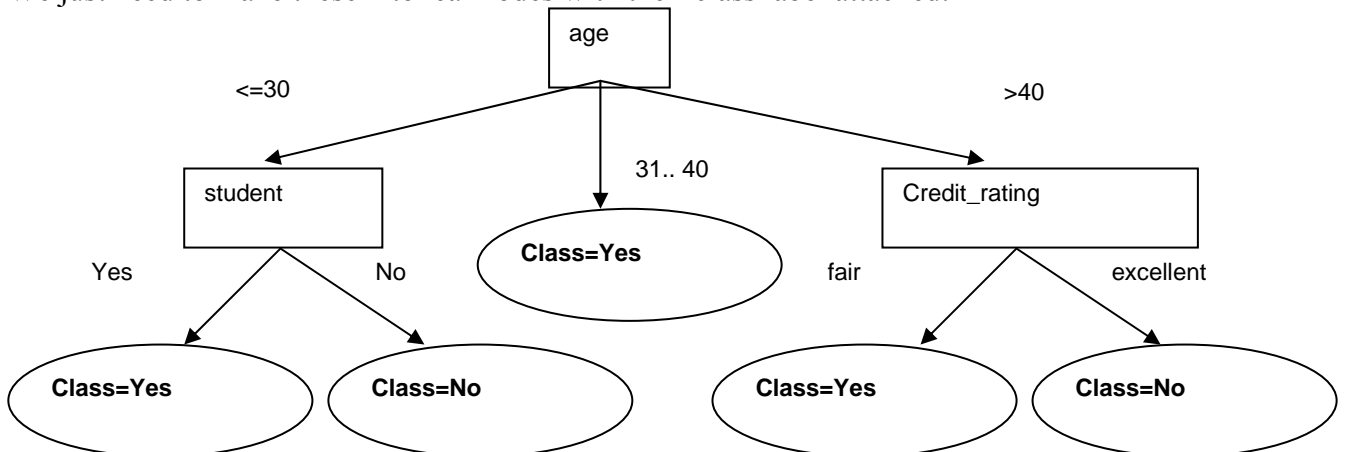
New example: age<=30, income=medium, student=yes, credit-rating=fair
Follow branch(age<=30) then student=yes we predict Class=yes ➔ Buys_computer = yes

Exercise 9. Associative Classifier

Given the training data in Exercise 4 (*Buy Computer* data), build an associative classifier model by generating all relevant association rules with support and confidence thresholds 10% and 60% respectively. Classify using this model the new example: age<=30, income=medium, student=yes, credit-rating=fair, selecting the rule with the highest confidence. What would be the classification if we chose to vote the class among all rules that apply?

Solution:

If we map the attribute values as follows the training set becomes the following transactions:

Attribute value	New symbol
age$_{<=30}$	a
age$_{31..40}$	b
age$_{>40}$	c
income$_{high}$	h
income$_{medium}$	m
income$_{low}$	l
student$_{yes}$	s
student$_{no}$	t
credit_rating$_{fair}$	F
credit_rating$_{excellent}$	E

1 {a, h, t, f, No}
2 {a, h, t, e, No}
3 {b, h, t, f, Yes}
4 {c, m, t, f, Yes}
5 {c, l, s, f, Yes}
6 {c, l, s, e, No}
7 {b, l, s, e, Yes}
8 {a, m, t, f, No}
9 {a, l, s, f, Yes}
10 {c, m, s, f, Yes}
11 {a, m, s, e, Yes}
12 {b, m, t, e, Yes}
13 {b, h, s, f, Yes}
14 {c, m, t, e, No}

C1 and F1

Candidate	Support
a, Class=yes	2
a, Class=no	3
b, Class=yes	4
~~b, Class=no~~	~~0~~
c, Class=yes	3
c, Class=no	2
h, Class=yes	2
h, Class=no	2
m, Class=yes	4
m, Class=no	2
l, Class=yes	3
~~l, Class=no~~	~~1~~
s, Class=yes	6
~~s, Class=no~~	~~1~~
t, Class=yes	3
t, Class=no	4
f, Class=yes	6
f, Class=no	2
e, Class=yes	3
e, Class=no	3

C2 and F2

Candidate	Support
~~a, h, Class=yes~~	~~0~~
a, h, Class=no	2
b, h, Class=yes	2
~~c, h, Class=yes~~	~~0~~
~~c, h, Class=no~~	~~0~~
~~a, m, Class=yes~~	~~1~~
~~a, m, Class=no~~	~~1~~
~~b, m, Class=yes~~	~~1~~
c, m, Class=yes	2
~~c, m, Class=no~~	~~1~~
~~a, l, Class=yes~~	~~1~~
~~b, l, Class=yes~~	~~1~~
~~c, l, Class=yes~~	~~1~~

~~h, s, Class=yes~~	~~1~~
~~h, t, Class=yes~~	~~1~~
h, t, Class=no	2
m, s, Class=yes	2
m, t, Class=yes	2
m, t, Class=no	2
l, s, Class=yes	3
~~l, t, Class=yes~~	~~0~~

a, s, Class=yes	2
~~a, t, Class=yes~~	~~0~~
a, t, Class=no	2
b, s, Class=yes	2
b, t, Class=yes	2
c, s, Class=yes	2
~~c, t, Class=yes~~	~~1~~
~~c, t, Class=no~~	~~1~~

~~a, f, Class=yes~~	~~1~~
a, f, Class=no	2
~~a, e, Class=yes~~	~~1~~
~~a, e, Class=no~~	~~1~~
~~b, f, Class=yes~~	~~1~~
b, e, Class=yes	2
c, f, Class=yes	3
~~c, f, Class=no~~	~~0~~
~~e, e, Class=yes~~	~~0~~
c, e, Class=no	2

s, f, Class=yes	3
s, e, Class=yes	2
t, f, Class=yes	2
t, f, Class=no	2
~~t, e, Class=yes~~	~~1~~
t, e, Class=no	2

h, f, Class=yes	**2**
~~h, f, Class=no~~	~~1~~
~~h, e, Class=yes~~	~~0~~
~~h, e, Class=no~~	~~1~~
m, f, Class=yes	**2**
~~m, f, Class=no~~	~~1~~
m, e, Class=yes	**2**
~~m, e, Class=no~~	~~1~~
l, f, Class=yes	**2**
~~l, e, Class=yes~~	~~1~~

C3 and F3

Candidate	Support
a, h, t, Class=No	**2**
a, t, f, Class=No	**2**
~~b, s, e, Class=Yes~~	~~1~~
~~c, m, s, Class=Yes~~	~~0~~
c, m, f, Class=Yes	**2**
~~c, s, f, Class=Yes~~	~~1~~
~~m, s, f, Class=Yes~~	~~1~~
~~m, t, f, Class=Yes~~	~~1~~
l, s, f, Class=Yes	**2**

No C4 candidates

Classification rules are:

a, h, t ➔ Class=No : $age_{<=30}$ AND $income_{high}$ AND $student_{no}$ ➔ Class=No (14.3%, 100%)

a, t, f ➔ Class=No : $age_{<=30}$ AND $student_{no}$ AND $credit\_rating_{fair}$ ➔ Class=No (14.3%,100%)

c, m, f ➔ Class=Yes : $age_{>40}$ AND $income_{medium}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,100%)

l, s, f ➔ Class=Yes : $income_{low}$ AND $student_{yes}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,100%)

h, f ➔ Class=Yes : $income_{high}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,66.6%)

m, f ➔ Class=Yes : $income_{medium}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,66.6%) **X**

m, e ➔ Class=Yes : $income_{medium}$ AND $credit\_rating_{excellent}$ ➔ Class=Yes (14.3%,66.6%)

l, f ➔ Class=Yes : $income_{low}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,100%)

s, f ➔ Class=Yes : $student_{yes}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (21.4%,75%) **X**

s, e ➔ Class=Yes : $student_{yes}$ AND $credit\_rating_{excellent}$ ➔ Class=Yes (14.3%,66.6%)

~~t, f ➔ Class=No : $student_{no}$ AND $credit\_rating_{fair}$ ➔ Class=No (14.3%,50%)~~

t, e ➔ Class=No : $student_{no}$ AND $credit\_rating_{excellent}$ ➔ Class=No (14.3%,66.6%)

h, t ➔ Class=No : $income_{high}$ AND $student_{no}$ ➔ Class=No (14.3%,66.6%)

m, s ➔ Class=Yes : $income_{medium}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,100%) **X**

~~m, t ➔ Class=Yes : $income_{medium}$ AND $student_{no}$ ➔ Class=Yes (14.3%,50%)~~

~~m, t ➔ Class=No : $income_{medium}$ AND $student_{no}$ ➔ Class=No (14.3%,50%)~~

l, s ➔ Class=Yes : $income_{low}$ AND $student_{yes}$ ➔ Class=Yes (21.4%,75%)

a, f ➔ Class=No : $age_{<=30}$ AND $credit\_rating_{fair}$ ➔ Class=No (14.3%,66.6%) **X**

b, e ➔ Class=Yes : $age_{31..40}$ AND $credit\_rating_{excellent}$ ➔ Class=Yes (14.3%,100%)

c, f ➔ Class=Yes : $age_{>40}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (21.4%,100%)

c, e ➔ Class=No : $age_{>40}$ AND $credit\_rating_{excellent}$ ➔ Class=No (14.3%,100%)

a, s ➔ Class=Yes : $age_{<=30}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,100%) **X**

a, t ➔ Class=No : $age_{<=30}$ AND $student_{no}$ ➔ Class=No (14.3%,66.6%)

b, s ➔ Class=Yes : $age_{31..40}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,100%)

b, t ➔ Class=Yes : $age_{31..40}$ AND $student_{no}$ ➔ Class=Yes (14.3%,100%)

c, s ➔ Class=Yes : $age_{>40}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,66.6%)

a, h ➔ Class=No   : $age_{<=30}$ AND $income_{high}$ ➔ Class=No (14.3%,100%)
b, h ➔ Class=Yes : $age_{31..40}$ AND $income_{high}$ ➔ Class=Yes (14.3%,100%)
c, m ➔ Class=Yes : $age_{>40}$ AND $income_{medium}$ ➔ Class=Yes (14.3%,66.6%)
~~a ➔ Class=Yes   : $age_{<=30}$ ➔ Class=Yes (14.3%,40%)~~
a ➔ Class=No   : $age_{<=30}$ ➔ Class=No (21.4%,60%) **X**
b ➔ Class=Yes : $age_{31..40}$ ➔ Class=Yes (28.6%,100%)
c ➔ Class=Yes : $age_{>40}$ ➔ Class=Yes (21.4%,60%)
~~c ➔ Class=No   : $age_{>40}$ ➔ Class=No (14.3%,40%)~~
~~h ➔ Class=Yes : $income_{high}$ ➔ Class=Yes (14.3%,50%)~~
~~h ➔ Class=No   : $income_{high}$ ➔ Class=No (14.3%,50%)~~
m ➔ Class=Yes : $income_{medium}$ ➔ Class=Yes (28.6%,66.6%) **X**
~~m ➔ Class=No   : $income_{medium}$ ➔ Class=No (14.3%,33.3%)~~
l ➔ Class=Yes : $income_{low}$ ➔ Class=Yes (21.4%,75%)
s ➔ Class=Yes : $student_{yes}$ ➔ Class=Yes (42.8%,85.7%) **X**
~~t ➔ Class=Yes : $student_{no}$ ➔ Class=Yes (21.4%,42.8%)~~
~~t ➔ Class=No   : $student_{no}$ ➔ Class=No (28.6%,57.1%)~~
f ➔ Class=Yes : $credit\_rating_{fair}$ ➔ Class=Yes (42.8%,75%) **X**
~~f ➔ Class=No   : $credit\_rating_{fair}$ ➔ Class=No (14.3%,25%)~~
~~e ➔ Class=Yes : $credit\_rating_{excellent}$ ➔ Class=Yes (21.4%,50%)~~
~~e ➔ Class=No   : $credit\_rating_{excellent}$ ➔ Class=No (21.4%,50%)~~

New example: age<=30, income=medium, student=yes, credit-rating=fair
There are 9 rules that apply (with red X above). Here they are ranked by confidence:

$age_{<=30}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,100%)
$income_{medium}$ AND $student_{yes}$ ➔ Class=Yes (14.3%,100%)
$student_{yes}$ ➔ Class=Yes (42.8%,85.7%)
$student_{yes}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (21.4%,75%)
$credit\_rating_{fair}$ ➔ Class=Yes (42.8%,75%)
$income_{medium}$ ➔ Class=Yes (28.6%,66.6%)
$age_{<=30}$ AND $credit\_rating_{fair}$ ➔ Class=No (14.3%,66.6%)
$income_{medium}$ AND $credit\_rating_{fair}$ ➔ Class=Yes (14.3%,66.6%)
$age_{<=30}$ ➔ Class=No (21.4%,60%)

The highest confident rules predicts Class=Yes.
We would in that case predict Buys_computer = yes

In a vote case:
There are 7 rules predicting Class=Yes with combined confidence = 81.27%
There are 2 rules predicting Class=No with combined confidence = 63.3%
We would in that case predict Buys_computer = yes