SENG 474

Assignment 1

Made By:

WanjinYoo: V00832396

Khaled Gabr: V00862424

**Question 1**

## 0 - LEVEL

1) AGE

1. AGE = Young

Entropy

-0.5*log2(0.5)-0.25*log2(0.25)-0.25*log2(0.25)

1. AGE = Pre-presbyopic

Entropy

-5/8 * log2(5/8)   -2/8 *log2 (2/8)  - 1/8 * log2(1/8)

1. AGE = Presbyopic

Entropy

-6/8 * log2(6/8)   -1/8 *log2 (1/8)  - 1/8 * log2(1/8)

AVERAGE:

1.5( 8/24 ) + 1.3 ( 8 / 24) + 1.1 ( 8 / 24 )

2) Spectacle-prescrip

1. Spectacle-prescrip = Myope

Entropy

-7/12 * log2(7/12)   -2/12 *log2 (2/12)  - 3/12 * log2(3/12)

1. Spectacle-prescrip = Hypermetrope

Entropy

-7/12 * log2(7/12)   -2/12 *log2 (2/12)  - 1/12 * log2(1/12)

AVERAGE:

1.38* (12/24) + 1.18(12/24)

3) Astigmatism

1. Astimatism = No

Entropy

-7/12 * log2(7/12)   -5/12 *log2 (5/12)

1. Astimatism = Yes

Entropy

-8/12 * log2(8/12)   -4/12 *log2 (4/12)

AVERAGE:

0.980*(12/24) + 0.918*(12/24)

4) Tear-prod-rate

1. Tear-Prod-Rate = Reduced

Entropy

(-12/12) * log2(12/12)

1. Tear-Prod-Rate = Normal

Entropy

-3/12 * log2(3/12)   -5/12 *log2 (5/12)  - 4/12 * log2(4/12)

AVERAGE:

0 * (12/24)+1.55*(12/24)

Tear-prod-rate has the lowest entropy average, and therefore will be chosen as the root of the decision tree.

Tear-Prod-Rate

reduced

Normal

None

?

## 

## 1 - LEVEL

1) Age

-2/4 \* log2(2/4) -1/4 \*log2 (1/4) -1/4 \*log2 (1/4)

1. AGE = Young

Entropy

-2/4 * log2(2/4)   -2/4 *log2 (2/4)

1. AGE = Pre-presbyopic

Entropy

-1/4 *log2 (1/4) -2/4 * log2(2/4) -1/4 *log2 (1/4)

1. AGE = Presbyopic

Entropy

-2/4 * log2(2/4)   -1/4 *log2 (1/4) -1/4 *log2 (1/4)

AVERAGE:

1*(8/24)+1.5*(8/24)+1.5*(8/24)

2) Spectacle-prescrip

1. Spectacle-Prescrip = Myope

Entropy

-1/6 * log2(1/6)   -2/6 *log2 (2/6) -3/6 *log2 (3/6)

1. Spectacle = Hypermetrope

Entropy

-2/6 * log2(2/6)   -3/6 *log2 (3/6) -1/6 *log2 (1/6)

AVERAGE:

1.46*(6/12)+1.46*(6/12)

3) Astigmatism

a) Astimatism = No

Entropy

-1/6 * log2(1/6)   -5/6 *log2 (5/6)

b) Astimatism = Yes

Entropy

-2/6 * log2(2/6)   -4/6 *log2 (4/6)

AVERAGE:

0.65*(6/12)+0.92*(6/12)

Astigmatism has the lowest entropy average, and therefore will be chosen as the next branch of the decision tree.

Tear-Prod-Rate

reduced

Normal

None

?

Astigmatism

?

NO

YES

Checking with Weka

tear-prod-rate = reduced: none

tear-prod-rate = normal

| astigmatism = no

| | age = young: soft

| | age = pre-presbyopic: soft

| | age = presbyopic

| | | spectacle-prescrip = myope: none

| | | spectacle-prescrip = hypermetrope: soft

| astigmatism = yes

| | spectacle-prescrip = myope: hard

| | spectacle-prescrip = hypermetrope

| | | age = young: hard

| | | age = pre-presbyopic: none

| | | age = presbyopic: none

**Question 2**

1. PLAY : yes

If ?

Then play = yes

Outlook: Sunny 2/5

Outlook: Overcast 4/4

Outlook: Rainy 3/5

Temp: Hot 2/4

Temp: Mild 4/6

Temp: cold 3/4

Humidity: High 3/7

Humidity: Normal 6/7

Windy: True 6/8

Windy: False 3/6

Outlook has the highest accuracy therefore the rule is:

1. First Rule

If outlook = overcast, then PLAY = yes

> Humidity: normal has the second highest accuracy

If Humidity = Normal

and ?

Then play = YES

Outlook: Sunny 2/2

Outlook: Overcast 2/2 ( it is covered by the previous rule )

Outlook: Rainy 2/3

Temp: Hot 1/1

Temp: Mild 2/2

Temp: cool 3/4

Windy: True 2/3

Windy: False 4/4

Since Windy:false has the highest accuracy and coverage, it will be considered for this rules:

2) Second rule

If Humidity = Normal

and Windy = false

Then play = YES

**Our results matches weka results**

If outlook = overcast then yes

If humidity = normal

and windy = FALSE then yes

If temperature = mild

and humidity = normal then yes

If outlook = rainy

and windy = FALSE then yes

If outlook = sunny

and humidity = high then no

If outlook = rainy

and windy = TRUE then no

**Question 3**

E = ( Age= Pre-presbyopic, Spectacle-prescrip = hypermetrope, astigmatism = YES, Tvsdsdear-prode-rate = reduce )

**a) SOFT**

P ( Contact-lenses = Soft | E ) =

P ( Age= Pre-presbyopic | contact-lenses = Soft ) \*

P ( Spectacle-prescrip = hypermetrope | contact-lenses = Soft ) \*

P ( astigmatism = YES | contact-lenses = Soft ) \*

P ( Tear-prode-rate = reduce | contact-lenses = Soft ) \*

P ( contact-lenses = Soft )

/ P ( E )

= ( 2+1 / 5+3 ) \* ( 3+1 / 5+2 ) \* ( 0+1 / 5+2 ) \* ( 0+1 / 5+2 ) \* ( 5+1 / 24+3 )

= \* 0.00097182

**b) HARD**

P ( Contact-lenses = hard | E ) =

P ( Age= Pre-presbyopic | contact-lenses = hard ) \*

P ( Spectacle-prescrip = hypermetrope | contact-lenses = hard ) \*

P ( astigmatism = YES | contact-lenses = hard) \*

P ( Tear-prode-rate = reduce | contact-lenses = hard ) \*

P ( contact-lenses = hard)

/ P ( E )

=( 1+1 / 4+3 ) \* ( 1+1 / 4+2 ) \* ( 4+1 / 4+2 ) \* ( 0+1 / 4+2 ) \* ( 4+1 / 24+3 )

= \* 0.00244954

**c) NONE**

P ( Contact-lenses = none | E ) =

P ( Age= Pre-presbyopic | contact-lenses = none ) \*

P ( Spectacle-prescrip = hypermetrope | contact-lenses = none ) \*

P ( astigmatism = YES | contact-lenses = none ) \*

P ( Tear-prode-rate = reduce | contact-lenses = none ) \*

P ( contact-lenses = none )

/ P ( E )

= ( 5+1 / 15+3 ) \* ( 8+1 / 15+2 )\*( 8+1/ 15+2)\*( 12+1 / 15+2) \*( 15+1 / 24+3)

= \* 0.04233666

= 1 / 0.00097182 + 0.00244954 + 0.04233666

= 21.854

p ( soft | E ) = 0.00097182 \* 21.854 = 0.02124

p ( hard| E ) = 0.00244954 \* 21.854 = 0.05353

p ( none | E ) = 0.04233666 \* 21.854 = 0.92522

‘None’ is most likely to be the result of this condition.

Weka results

Naive Bayes Classifier

=== Predictions on test set ===

inst# actual predicted error prediction

1 1:? 3:none 0.925