# **Assignment 3**

Moises A. Martinez

College of Computing and Engineering, Nova Southeastern University

CISC 0670: Artificial Intelligence

Dr. Wei Li

December 2nd, 2022

## Decision Tree

### Question 2.1

$p_{cinema} = 5/8$, $p_{tennis} = 1/8$, $p_{stay\ in} = 1/8$, $p_{shopping} = 1/8$

$Entropy(S) = -p_{cinema}\log_2(p_{cinema}) – p_{tennis}\log_2(p_{tennis}) – p_{stay}\log_2(p_{stay}) – p_{shopping}\log_2(p_{shopping})$

$= -(5/8)\log_2(5/8) – (1/8)\log_2(1/8) – (1/8)\log_2(1/8) – (1/8)\log_2(1/8)$

$= -(5/8)( -0.6781) – (1/8)(-3) - (1/8)(-3) - (1/8)(-3)$

$= 0.4238125 + 0.375 + 0.375 + 0.375$

= **1.5488125**

$S_{weather=sunny} = \{W1, W2\}$ $p_{cinema} = 1/2$, $p_{tennis} = 1/2$, $p_{stay\ in} = 0/2$, $p_{shopping} = 0/2$

$Entropy(S_{weather=sunny}) = -p_{cinema}\log_2(p_{cinema}) – p_{tennis}\log_2(p_{tennis}) – p_{stay}\log_2(p_{stay}) – p_{shopping}\log_2(p_{shopping})$

$= -(1/2)\log_2(1/2) – (1/2)\log_2(1/2) – (0/2)\log_2(0/2) – (0/2)\log_2(0/2)$

$= -(1/2)(-1) – (1/2)(-1) – 0 – 0$

$= 1/2 + 1/2$

= **1**

$S_{weather=windy} = \{W3, W7, W8\}$ $p_{cinema} = 2/3$, $p_{tennis} = 0/3$, $p_{stay\ in} = 0/3$, $p_{shopping} = 1/3$

$Entropy(S_{weather=windy}) = -p_{cinema}\log_2(p_{cinema}) – p_{tennis}\log_2(p_{tennis}) – p_{stay}\log_2(p_{stay}) – p_{shopping}\log_2(p_{shopping})$

$= -(2/3)\log_2(2/3) – (0/3)\log_2(0/3) – (0/3)\log_2(0/3) – (1/3)\log_2(1/3)$

$= -(2/3)(-0.585) – 0 – 0 – (1/3)(-1.585)$

$= 0.39 + 0.528\overline{3}$

= **$\underline{0.918\overline{3}}$**

$S_{weather=rainy}$ = {W4, W5, W6} $p_{cinema} = 2/3$, $p_{tennis} = 0/3$, $p_{stay\ in} = 1/3$, $p_{shopping} = 0/3$

$Entropy(S_{weather=rainy}) = -p_{cinema}log_2(p_{cinema}) – p_{tennis}log_2(p_{tennis}) – p_{stay}log_2(p_{stay}) –$
$p_{shopping}log_2(p_{shopping})$

$= -(2/3)log_2(2/3) – (0/3)log_2(0/3) – (1/3)log_2(1/3) – (0/3)log_2(0/3)$

$= -(2/3)(-0.585) – 0 – (1/3)(-1.585) – 0$

$= 0.39 + 0.528\overline{3}$

= **$\underline{0.918\overline{3}}$**

Gain (S, weather) = Entropy(S) – ( $|S_{weather = rainy}|/|S|$ * Entropy($S_{weather=rainy}$) + $|S_{weather = sunny}|/|S|$ * Entropy($S_{weather=sunny}$) + $|S_{weather = windy}|/|S|$ * Entropy($S_{weather=windy}$) )

$= 1.5488125 – ( (2/8)*1 + (3/8) * 0.918\overline{3} + (3/8) * 0.918\overline{3}) )$

= 1.5488125 – (2/8 + 0.344375 + 0.344375)

= 1.5488125 - 0.93875

= **<u>0.6100625</u>**

Gain (S, parents) = Entropy(S) – ( $|S_{parents=yes}|/|S|$ * Entropy($S_{parents=yes}$) + $| S_{parents=no} |/|S|$ * Entropy($S_{parents=no}$))

$S_{parents=yes}$ = {W1, W3, W4, W6} $p_{cinema} = 4/4$, $p_{tennis} = 0/4$, $p_{stay\ in} = 0/4$, $p_{shopping} = 0/4$

Entropy($S_{parents=yes}$) = $-p_{cinema}\log_2(p_{cinema}) - p_{tennis}\log_2(p_{tennis}) - p_{stay}\log_2(p_{stay}) -$

$p_{shopping}\log_2(p_{shopping})$

$= -\log_2(1) - (0)\log_2(0) - (0)\log_2(0) - (0)\log_2(0)$

$= 0$

$S_{parnets=no}$ = {W2, W5, W7, W8} $p_{cinema} = 1/4$, $p_{tennis} = 1/4$, $p_{stay\ in} = 1/4$, $p_{shopping} = 1/4$

Entropy($S_{parents=no}$) = $-(1/4)\log_2(1/4) - (1/4)\log_2(1/4) - (1/4)\log_2(1/4) - (1/4)\log_2(1/4)$

$= -(1/4)(-2) - (1/4)(-2) - (1/4)(-2) - (1/4)(-2)$

$= 0.5 + 0.5 + 0.5 + 0.5$

$= 2$

Gain (S, parents) = 1.5488125 – ((4/8)(0) + (4/8)(2))

= 1.5488125 – 1

= **0.548812**

Gain (S, money) = Entropy(S) – ( $|S_{money=rich}|/|S|$ * Entropy($S_{money=rich}$) + $|S_{money=poor}|/|S|$ *

Entropy($S_{money=poor}$))

$S_{money=rich}$ = {W1, W2, W3, W5, W8} $p_{cinema} = 2/5$, $p_{tennis} = 1/5$, $p_{stay\ in} = 1/5$, $p_{shopping} = 0/5$

Entropy($S_{money=rich}$) = $-p_{cinema}\log_2(p_{cinema}) - p_{tennis}\log_2(p_{tennis}) - p_{stay}\log_2(p_{stay}) -$

$p_{shopping}\log_2(p_{shopping})$

$= -(2/5)\log_2(2/5) - (1/5)\log_2(1/5) - (1/5)\log_2(1/5) - (0/5)\log_2(0/5)$

$= -(2/5)(-1.322) - (1/5)(-2.322) - (1/5)(-2.322) - 0$

$= 0.5288 + 0.4644 + 0.4644$

$= 1.4576$

$S_{money=poor} = \{W4, W6, W7\}$ $p_{cinema} = 3/3$, $p_{tennis} = 0/3$, $p_{stay\ in} = 0/3$, $p_{shopping} = 0/3$

$Entropy(S_{money=poor}) = -p_{cinema}\log_2(p_{cinema}) - p_{tennis}\log_2(p_{tennis}) - p_{stay}\log_2(p_{stay}) - p_{shopping}\log_2(p_{shopping})$

$= -\log_2(1) - (0)\log_2(0) - (0)\log_2(0) - (0)\log_2(0)$

$= 0$

$Gain(S, money) = 1.5488125 - ((5/8)(1.4576) + (3/8)(0))$

$= 1.5488125 - 0.911$

= **<u>0.6378125</u>**

Based on the above values, the first attribute to split on should be "Money", because it provided the highest calculated information gain.

# Logic

## Question 2.4

(i).

"∀x ∃y Loves(x, y) ⇔∃x ∀y Loves(x, y)" means <u>"Every x loves at least one y if and only if there is an x that loves every y"</u>. The statement is not valid because it is not true for all possible interpretations. Assume x = {John, Joe} and y = {Mary, Maria} and the following are true: John loves Mary (Loves(John, Mary)=True), and Joe loves Maria (Loves(Joe, Maria)). Then, the first half of the statement is true because for every person in x, there does exist at least one person in y who x loves. However, the second half of the statement is false since there does not exist a person in x that loves every y (John does not Love Maria, and Joe does not love Mary). Therefore, this model does not satisfy the statement, and thus the statement is not true for all possible interpretations in all possible world states.

However, the statement is satisfiable because there does exist a model that can satisfy the statement. Assume x = {James, Jerry} and y = {Maddy, Melissa} and the following are true: James loves Maddy, James loves Melissa, Jerry loves Melissa. Then the first half of the statement is true because for every person in x, there exist at least one person in y who x loves. Furthermore, in this model, there is an x that loves every y (it is James). Therefore, the model satisfies this statement and is true for at least one interpretation.

"∀x Loves(x, movie) ⇔ ¬∃x ¬Loves(x, movie)" means <u>"every x loves movie if and only if there is not an x that does not love movie"</u>. This statement is valid because the statements "∀x Loves(x, movie)" and "¬∃x ¬Loves(x, movie)" are equivalent (Quantifier Duality rule). Thus, if "∀x Loves(x, movie)" is true for any x, then "¬∃x ¬Loves(x, movie)" will be true; and if "∀x

Loves(x, movie)" is false for any x, then "¬∃x ¬Loves(x, movie)" will be false. Therefore, the biconditional statement is true for any x and is valid for every possible interpretation.

"∃x Loves(x, movie) ⇔ ¬∀x ¬Loves(x, movie)" means <u>"There exists an x that loves movie if and only if not every x does not love movie".</u> This statement is also valid because the statements "∃x Loves(x, movie)" and "¬∀x ¬Loves(x, movie)" are equivalent (Quantifier Duality rule). Thus, if "∃x Loves(x, movie)" is true for any x, then "¬∀x ¬Loves(x, movie)" will be true; and if "¬∃x ¬Loves(x, movie)" is false for any x, then "∀x Loves(x, movie)" will be false. Therefore, the biconditional statement is true for any x and is valid for every possible interpretation.

"¬∀x ¬ Loves(x, movie) ⇔ ∀x Loves(x, movie)" means <u>"Not every x does not love movie if and only if every x loves movie".</u> This statement is not valid because it is not true for all possible interpretations. Let x = {Jim, Moira}, and the following is true: Jim does not love movie, and Moira loves movie. Then the first half of the statement is true, as not every x does not love movie (Moira is in x and loves movie). However, the second half of the statement is false as not every x loves movie (Jim is in x and does not love movie). Therefore, this model does not satisfy the statement, and thus the statement is not true for all possible interpretations in all possible world states.

Furthermore, the statement is not satisfiable because there exists no model that the statement can be true. That is because "¬∀x ¬ Loves(x, movie)" and "∀x Loves(x, movie)" are inverse statements; therefore, they can never be equivalent. Thus, if "¬∀x ¬ Loves(x, movie)" is true for any x, "∀x Loves(x, movie)" will always be false. So the biconditional statement always fails. Therefore, the statement is not satisfiable.

(ii)

Original statement in First-Order Logic:

$\forall x, Y(x) \wedge H(x) \Rightarrow B(x)$

$\forall x, A(x) \Rightarrow H(x)$

$\exists x, N(x) \Rightarrow Y(x) \wedge A(x)$

$\therefore \exists x, N(x) \Rightarrow B(x)$

Eliminate implication:

$\forall x, \neg (Y(x) \wedge H(x)) \vee B(x)$

$\forall x, \neg A(x) \vee H(x)$

$\exists x, \neg N(x) \vee Y(x) \wedge A(x)$

$\exists x, \neg (\neg N(x) \vee B(x))$

(need to negate last statement since we are trying to prove $N(x) \Rightarrow B(x)$, so assume $\neg (N(x) \Rightarrow B(x))$ )

Simplify $\neg$:

$\forall x, \neg Y(x) \vee \neg H(x) \vee B(x)$

$\forall x, \neg A(x) \vee H(x)$

$\exists x, \neg N(x) \vee Y(x) \wedge A(x)$

$\exists x, N(x) \wedge \neg B(x)$

Dropping quantifies:

$\neg Y(x) \vee \neg H(x) \vee B(x)$

$\neg A(x) \vee H(x)$

$\neg N(x) \vee Y(x) \wedge A(x)$

$N(x) \wedge \neg B(x)$

Split up ^:

1. $\neg Y(x) \vee \neg H(x) \vee B(x)$
2. $\neg A(x) \vee H(x)$
3. $\neg N(x) \vee Y(x)$
4. $A(x)$
5. $N(x)$
6. $\neg B(x)$

Now Prove:

7. Combine (4, 2): $H(x)$
8. Combine (1, 7): $\neg Y(x) \vee B(x)$
9. Combine (3, 8): $\neg N(x) \vee B(x)$
10. Combine (5, 9): $B(x)$
11. Combine (6, 10): NIL

Therefore, $N(x) \Rightarrow B(x)$ is true

## K-means clustering

**Question 2.3**

(i) Assume Cluster 1 is point 1 and Cluser 2 is point 4

Point 2: Distance from point 1 = sqrt( $(1 – 1.5)^2 + (1 – 2)^2$ )= sqrt( $(-0.5)^2 + (-1)^2$) = sqrt(0.25 + 1)

= sqrt(1.25) ≈ 1.118

Distance from point 4 = sqrt($(5 – 1.5)^2 + (7 – 2)^2$ ) = sqrt ($(3.5)^2 + 5^2$ ) = sqrt(12.25 + 25) = sqrt(37.25) ≈ 6.103

Point 3: Distance from point 1 = sqrt( $(1 – 3)^2 + (1 – 4)^2$ )= sqrt( $(-2)^2 + (-3)^2$) = sqrt(4 + 9)

= sqrt(13) ≈ 3.606

Distance from point 4 = sqrt( $(5 – 3)^2 + (7 – 4)^2$ )= sqrt( $(2)^2 + (3)^2$) = sqrt(4 + 9)

= sqrt(13) ≈ 3.606

Point 5: Distance from point 1 = sqrt( $(1 – 3.5)^2 + (1 – 5)^2$ ) = sqrt ( $-2.5^2 + (-4)^2$ ) = sqrt(6.25 + 16) = sqrt(22.25) ≈ 4.717

Distance from point 4 = sqrt( $(5 – 3.5)^2 + (7 – 5)^2$ )= sqrt( $(1.5)^2 + (2)^2$) = sqrt(2.25 + 4)

= sqrt(6.25) = 2.5

Point 6: Distance from point 1 = sqrt( $(1 – 4.5)^2 + (1 – 5)^2$ ) = sqrt ( $-3.5^2 + (-4)^2$ ) = sqrt(12.25 + 16) = sqrt(28.25) ≈ 5.315

Distance from point 4 = sqrt( $(5 – 4.5)^2 + (7 – 5)^2$ )= sqrt( $(0.5)^2 + (2)^2$) = sqrt(0.25 + 4)

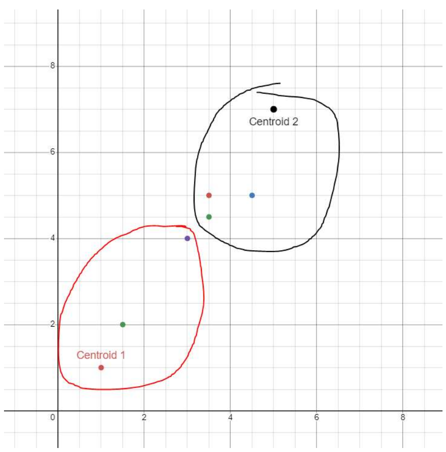
= sqrt(4.25) ≈ 2.061

Point 7: Distance from point 1 = sqrt( $(1 – 3.5)^2 + (1 – 4.5)^2$ ) = sqrt ( $-2.5^2 + (-3.5)^2$ ) = sqrt(6.25 + 12.25) = sqrt(18.5) ≈ 4.301

Distance from point 4 = sqrt( $(5 – 3.5)^2 + (7 – 4.5)^2$ )= sqrt( $(1.5)^2 + (2.5)^2$) = sqrt(2.25 + 6.25) = sqrt(8.5) ≈ 2.915

**<u>Cluster 1: Points 1, 2, 3; Cluster 2: Points 4, 5, 6, 7</u>**

Visualization:




(ii) Centroid 1: (1 + 1.5 + 3)/3, (1, 2, 4)/3 => 5.5/3,7/3 => **<u>($1.8\overline{3}$, $2.\overline{3}$)</u>**

Centroid 2: (5 + 3.5 + 4.5 + 3.5/4, 7+5+5+4.5/4) => (16.5/4, 21.5/4) => **<u>(4.125, 5.375)</u>**

Point 1: Distance from centroid 1 = sqrt($(1.8\overline{3} – 1)^2 + (2.\overline{3} – 1)^2$ ) ≈ 1.118

Distance from centroid 2 = sqrt($(4.125 – 1)^2 + (5.375 – 1)^2$ ) ≈ 5.376

Point 2: Distance from centroid 1 = sqrt($(1.8\overline{3} – 1.5)^2 + (2.\overline{3} – 2)^2$) ≈ 0.471

Distance from centroid 2 = sqrt($(4.125 – 1.5)^2 + (5.375 – 2)^2$ ) ≈ 4.276

Point 3: Distance from centroid 1 = sqrt($((1.8\overline{3} - 3)^2 + (2.\overline{3} - 4)^2)$) ≈ 2.034

Distance from centroid 2 = sqrt($((4.125 - 3)^2 + (5.375 - 4)^2$ ) ≈ 1.777

Point 4: Distance from centroid 1 = sqrt($((1.8\overline{3} - 5)^2 + (2.\overline{3} - 5)^2)$) ≈ 5.640

Distance from centroid 2 = sqrt($((4.125 - 5)^2 + (5.375 - 7)^2$ ) ≈ 1.846

Point 5: Distance from centroid 1 = sqrt($((1.8\overline{3} - 3.5)^2 + (2.\overline{3} - 5.)^2)$) ≈ 3.145

Distance from centroid 2 = sqrt($((4.125 - 3.5)^2 + (5.375 - 5)^2$ ) ≈ 0.729

Point 6: Distance from centroid 1 = sqrt($((1.8\overline{3} - 4.5)^2 + (2.\overline{3} - 5)^2)$) ≈ 3.771

Distance from centroid 2 = sqrt($((4.125 - 4.5)^2 + (5.375 - 5)^2$ ) ≈ 0.530

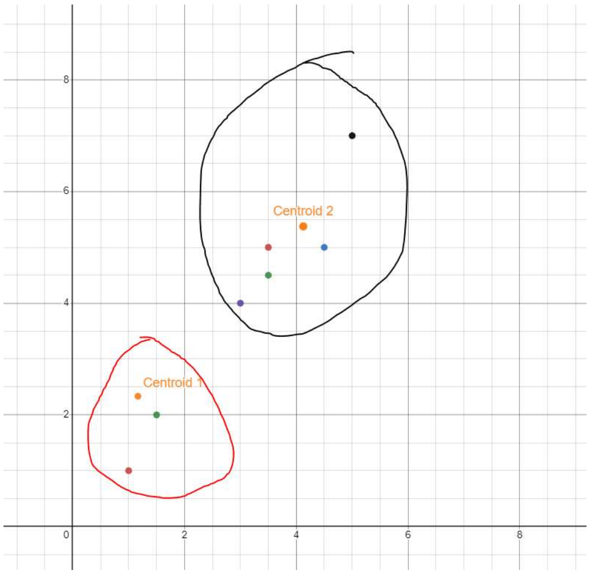
Point 7: Distance from centroid 1 = sqrt($((1.8\overline{3} - 3.5)^2 + (2.\overline{3} - 4.5)^2)$) ≈ 2.733

Distance from centroid 2 = sqrt($((4.125 - 3.5)^2 + (5.375 - 4.5)^2$ ) ≈ 1.075

**Cluster 1: Points 1, 2; Cluster 2: Points 3, 4, 5, 6, 7**

**Point 3 changed from cluster 1 to cluster 2**

Visualization




(iii) Centroid 1: (1 + 1.5)/2, (1 + 2)/2 => 2.5/2,3/2 => (1.25, 1.5)

Centroid 2: (3 + 5 + 3.5 + 4.5 + 3.5/5, 4 + 7+5+5+4.5/5) => (19.5/5, 25.5/5) => (3.9, 5.1)

Point 1: Distance from centroid 1 = sqrt($(1.25 – 1)^2 + (1.5 – 1)^2$ ) ≈ 0.559

Distance from centroid 2 = sqrt($(3.9– 1)^2 + (5.1– 1)^2$ ) ≈ 5.022

Point 2: Distance from centroid 1 = sqrt($(1.25 – 1.5)^2 + (1.5 – 2)^2$ ) ≈ 0.559

Distance from centroid 2 = sqrt($(3.9– 1.5)^2 + (5.1– 2)^2$ ) ≈ 3.920

Point 3: Distance from centroid 1 = sqrt($(1.25 – 3)^2 + (1.5 – 4)^2$ ) ≈ 3.052

Distance from centroid 2 = sqrt($(3.9– 3)^2 + (5.1– 4)^2$ ) ≈ 1.421

Point 4: Distance from centroid 1 = sqrt($(1.25 – 5)^2 + (1.5 – 7)^2$ ) ≈ 6.657

Distance from centroid 2 = sqrt($(3.9– 5)^2 + (5.1– 7)^2$ ) ≈ 2.195

Point 5: Distance from centroid 1 = sqrt($((1.25 – 3.5)^2 + (1.5 – 5)^2$ ) ≈ 4.161

Distance from centroid 2 = sqrt($((3.9– 3.5)^2 + (5.1– 5)^2$ ) ≈ 0.412

Point 6: Distance from centroid 1 = sqrt($((1.25 – 4.5)^2 + (1.5 – 5)^2$ ) ≈ 4.776

Distance from centroid 2 = sqrt($((3.9– 1)^2 + (5.1– 1)^2$ ) ≈ 0.608

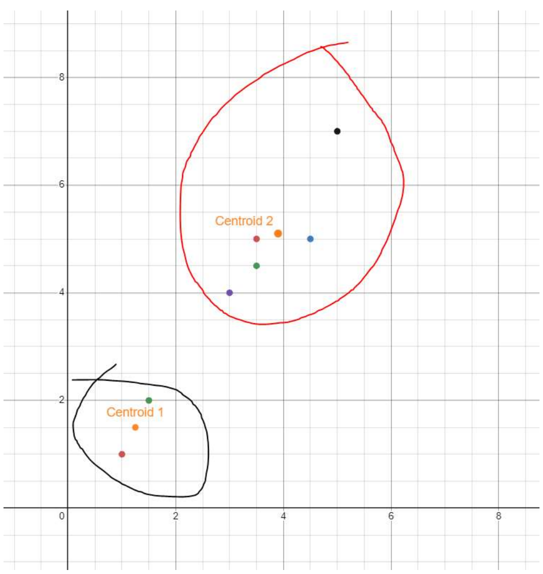
Point 7: Distance from centroid 1 = sqrt($((1.25 – 3.5)^2 + (1.5 – 4.5)^2$ ) = 3.75

Distance from centroid 2 = sqrt($((3.9– 3.5)^2 + (5.1– 4.5)^2$ ) ≈ 0.721

**Cluster 1: Points 1, 2; Cluster 2: Points 3, 4, 5, 6, 7**

There were no changes going from (ii) to (iii). There will not be any more changes if we keep continuing with iterations because the cluster points did not change from our previous iteration.

Visualization:



## Bayes Theorem

**Question 2.2**

Let P(A) = Probabilty of Accept, P(R) = Probability of Reject, P(G) = Probability of Good, and P(B) = Probability of Bad

Given P(G & R) = 0.04, P(B & A) = 0.03, and P(B) = 0.02

(i) $P(R|G) = P(R \& G)/P(G)$

If P(B) = 0.02, then P(G) = 0.98 (P(G) = 1 – P(B) => 1 – 0.02 = 0.98)

$P(R|G) = 0.04/0.98 \approx$ **<u>0.0408</u>**

(ii) $P(A|B) = P(A \& B)/P(B) = 0.03/0.02 =$ **<u>1.5</u>**

College of Computing and Engineering
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Certification of Authorship

Submitted to (Professor's Name): **Dr. Wei Li**

Student's Name: **Moises Alfonso Martinez**

Date of Submission: **December 2nd, 2022**

Purpose and Title of Submission: **Homework Assignment 3**

Certification of Authorship: I hereby certify that I am the author of this document and that any assistance I received in its preparation is fully acknowledged and disclosed in the document. I have also cited all sources from which I obtained data, ideas, or words that are copied directly or paraphrased in the document. Sources are properly credited according to accepted standards for professional publications. I also certify that this paper was prepared by me for this purpose.

Student's Signature: **Moises Alfonso Martinez**