

วิทยาลัยเทคโนโลยีอุตสาหกรรม  
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การสอบปลายภาค ประจำภาคการศึกษาที่ 1

วิชา 030513146 Artificial Intelligence

วันที่ 1 ธันวาคม 2558

เลขประจำตัว

ชื่อ

สาขา

ปีการศึกษา 2558

ชั้น Sec. 1-2

เวลา 9.00 - 12.00 น.

ตอนเรียนที่

- คำสั่ง
1. ทุจริตปรับตกในวิชานั้น ไม่พิจารณาผลการเรียนในภาคการศึกษานี้ และให้พักการเรียนต่อไปอีก 1 ภาคการศึกษา
  2. ห้ามนำข้อสอบออกนอกห้องสอบ
  3. ห้ามใช้โทรศัพท์มือถือระหว่างการสอบโดยเด็ดขาด
  4. อนุญาตให้นำตำราและเอกสารเข้าห้องสอบได้
  5. อนุญาตให้ใช้เครื่องคิดเลขได้
  6. ข้อสอบมีทั้งหมด 9 หน้า จำนวน 8 ข้อ ให้ทำทุกข้อ คะแนนเต็ม 80 คะแนน

ข้อ	1	2	3	4	5	6	7	8	รวม
คะแนน									

ผู้ออกข้อสอบ

อาจารย์ดำรงเกียรติ แซ่ลิ่ม

1. Prove the logical equivalence of these 2 sentences. Each proving step, you have to indicate what rules do you use to transform the sentence (7 points)

$$\neg(p \rightarrow (q \wedge \neg r)) \equiv p \wedge (\neg q \vee r)$$

แปลจากข้อไปงว่า

$$\begin{aligned} \neg(p \rightarrow (q \wedge \neg r)) &\equiv \neg(\neg p \vee (q \wedge \neg r)) && \text{Impl} \\ &\equiv p \wedge \neg(q \wedge \neg r) && \text{DM} \\ &\equiv p \wedge (\neg q \vee r) && \text{DM} \end{aligned}$$

แปลจากทามมาช่วย

$$\begin{aligned} p \wedge (\neg q \vee r) &\equiv \neg(\neg p \rightarrow \neg(\neg q \vee r)) && \text{Impl} \\ &\equiv \neg(\neg p \rightarrow (q \wedge \neg r)) && \text{DM} \end{aligned}$$

2. จงพิสูจน์ว่าค่ากล่าวต่อไปนี้ถูกต้อง ถ้าอุณหภูมิและความดันไม่คงที่ฝนจะตก ขณะนี้ความดันคงที่ ดังนั้นถ้าฝนไม่ตกแล้วหมายความว่าความอุณหภูมิคงที่

กำหนดให้      ประพจน์ R แทนประโยค "ฝนตก",  
                          ประพจน์ T แทนประโยค "อุณหภูมิคงที่" และ  
                          ประพจน์ P แทน "ความดันคงที่"

จงเปลี่ยนประโยคต่อไปนี้ให้อยู่ในรูปของ Propositional Logic (2 คะแนน)

ถ้าอุณหภูมิและความดันไม่คงที่ฝนจะตก      =  $(\neg T \wedge \neg P) \Rightarrow R$   
 ขณะนี้ความดันคงที่      =  $P$

จงพิสูจน์ว่า ดังนั้นถ้าฝนไม่ตกแล้วหมายความว่าความอุณหภูมิคงที่ ด้วยวิธี Indirect Proof

โดยระบุแต่ละขั้นตอนด้วยว่าใช้กฎอะไรในการพิสูจน์ (7 คะแนน)

ข้อสรุปคือ  $\neg R \Rightarrow T$

- |                                           |            |
|-------------------------------------------|------------|
| 1. $(\neg T \wedge \neg P) \Rightarrow R$ | Premise    |
| 2. $P$                                    | Premise    |
| 3. $\neg (\neg R \Rightarrow T)$          | Assumption |
| 4. $\neg (\neg T \wedge \neg P) \vee R$   | 1 Impl     |
| 5. $T \vee P \vee R$                      | 4 DM       |

$\therefore$  (ไม่สามารถพิสูจน์ได้ว่า ฝนไม่ตกถ้า อุณหภูมิ คงที่ หรือ ไม่)  
 เพราะ = จากข้อ 3 ความดันคงที่ ซึ่งขัดแย้งกับ ประโยคที่ได้ข้อ 1 มา  
 จึงทำให้ข้อ 3 ฝนไม่ตก ข้อ 4 อุณหภูมิคงที่ หรือ ไม่คงที่ก็ได้



3. จากประโยคที่แปลงอยู่ในรูป Predicate Logic ต่อไปนี้

- ม้าทุกตัวเร็วกว่าสุนัขทุกตัว

$$\forall x \forall y \text{Horse}(x) \wedge \text{Dog}(y) \Rightarrow \text{Faster}(x,y)$$

- เกรย์ฮาวด์บางตัวเร็วกว่ากระต่าย

$$\exists y \text{Greyhound}(y) \wedge (\forall z \text{Rabbit}(z) \Rightarrow \text{Faster}(y,z))$$

- ถ้า x เร็วกว่า y และ y เร็วกว่า z แล้ว x เร็วกว่า z ด้วย

$$\forall x \forall y \forall z \text{Faster}(x,y) \wedge \text{Faster}(y,z) \Rightarrow \text{Faster}(x,z)$$

- ถ้า y คือ เกรย์ฮาวด์ แล้ว y คือสุนัข

$$\forall y \text{Greyhound}(y) \Rightarrow \text{Dog}(y)$$

- Harry คือสุนัข

$$\text{Horse}(\text{Harry})$$

- Bugbunny คือกระต่าย

$$\text{Rabbit}(\text{Bugbunny})$$

3.1) จงทำให้เป็นประโยคใหม่ที่อยู่ในรูปของ CNF (แสดงวิธีทำโดยละเอียด ทีละขั้นตอน) (10 คะแนน)

1. กำจัด  $\Rightarrow$

$$\forall x \forall y \neg (\text{Horse}(x) \wedge \text{Dog}(y)) \vee \text{Faster}(x,y)$$

$$\exists y \text{Greyhound}(y) \wedge \neg (\forall z \text{Rabbit}(z) \vee \text{Faster}(y,z))$$

$$\forall x \forall y \forall z \neg (\text{Faster}(x,y) \wedge \text{Faster}(y,z)) \vee \text{Faster}(x,z)$$

$$\forall y \neg \text{Greyhound}(y) \vee \text{Dog}(y)$$

2. กระจาย  $\neg$  เข้าไปในประโยค

$$\forall x \forall y \neg \text{Horse}(x) \vee \neg \text{Dog}(y) \vee \text{Faster}(x,y)$$

$$\exists y \text{Greyhound}(y) \wedge (\exists z \neg \text{Rabbit}(z) \vee \text{Faster}(y,z))$$

$$\forall x \forall y \forall z \neg \text{Faster}(x,y) \vee \neg \text{Faster}(y,z) \vee \text{Faster}(x,z)$$

$$\forall y \neg \text{Greyhound}(y) \vee \text{Dog}(y)$$

3. in standardize ตัวแปร : (ไม่มีให้ทำ)

4. ย้าย Quantifier มาอยู่ด้านซ้ายประโยค

$$\exists y \exists z \text{Greyhound}(y) \wedge (\neg \text{Rabbit}(z) \vee \text{Faster}(y,z))$$

5. กำจัด  $\exists$

$$\text{Greyhound}(a) \wedge (\neg \text{Rabbit}(b) \vee \text{Faster}(a,b))$$

6. แปลข้อ 4 เป็นนิพจน์ตรรกศาสตร์

$$\neg \text{Horse}(x) \vee \neg \text{Dog}(y) \vee \text{Faster}(x, y)$$

$$\text{Greyhound}(a) \wedge (\neg \text{Rabbit}(b) \vee \text{Faster}(a, b))$$

$$\neg \text{Faster}(x, y) \vee \neg \text{Faster}(y, z) \vee \text{Faster}(x, z)$$

$$\neg \text{Greyhound}(y) \vee \text{Dog}(y)$$

Horse(Harry)

Rabbit(Bugbunny)

7. แปลข้อ 5 เป็นนิพจน์ตรรกศาสตร์

$$\neg \text{Horse}(x) \vee \neg \text{Dog}(y) \vee \text{Faster}(x, y)$$

$$\text{Greyhound}(a)$$

$$\neg \text{Rabbit}(b) \vee \text{Faster}(a, b)$$

$$\neg \text{Faster}(x, y) \vee \neg \text{Faster}(y, z) \vee \text{Faster}(x, z)$$

$$\neg \text{Greyhound}(y) \vee \text{Dog}(y)$$

$$\text{Horse}(\text{Harry})$$

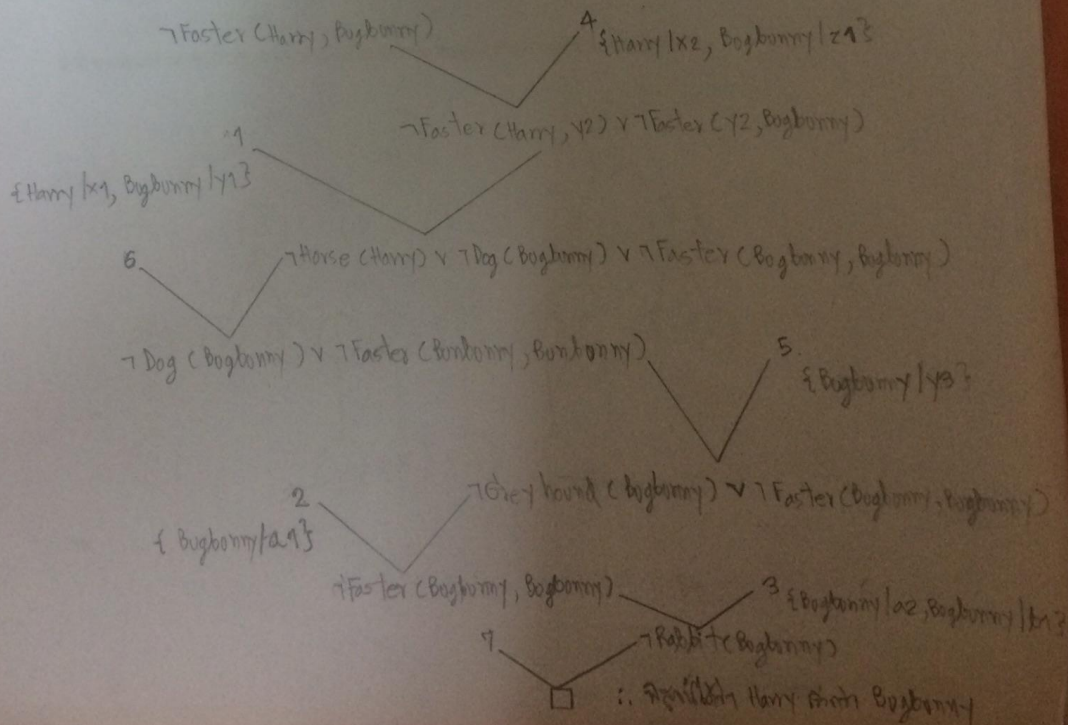
$$\text{Rabbit}(\text{Bugbunny})$$

8. Standardize นิพจน์ตรรกศาสตร์ให้เป็นนิพจน์มาตรฐาน

1.  $\neg \text{Horse}(x_1) \vee \neg \text{Dog}(y_1) \vee \text{Faster}(x_1, y_1)$
2.  $\text{Greyhound}(a_1)$
3.  $\neg \text{Rabbit}(b_1) \vee \text{Faster}(a_2, b_1)$
4.  $\neg \text{Faster}(x_2, y_2) \vee \neg \text{Faster}(y_2, z_1) \vee \text{Faster}(x_2, z_1)$
5.  $\neg \text{Greyhound}(y_3) \vee \text{Dog}(y_3)$
6.  $\text{Horse}(\text{Harry})$
7.  $\text{Rabbit}(\text{Bugbunny})$

3.2) จงพิสูจน์ว่า Harry เร็วกว่า Bugbunny?  $\text{Faster}(\text{Harry}, \text{Bugbunny})$ 

(5 คะแนน)





4. กำหนด Rule-Based Systems ซึ่งประกอบด้วยกฎเกณฑ์ ดังนี้

กฎข้อที่ 1. ถ้า วิศวกรระบบเครือข่าย และ วิศวกรระบบปฏิบัติการ แล้วควรเป็น Network Engineer

กฎข้อที่ 2. ถ้า วิศวกรเขียนโปรแกรม และ วิศวกรระบบเครือข่าย แล้วควรเป็น Programmer

กฎข้อที่ 3. ถ้า วิศวกรเขียนโปรแกรม และ วิศวกรระบบปฏิบัติการ แล้วควรเป็น Developer

กฎข้อที่ 4. ถ้า ใช้งาน Windows และ ใช้งาน Linux แล้วถือว่า วิศวกรระบบปฏิบัติการ

กฎข้อที่ 5. ถ้า ฝึกการเขียนภาษา C และ ภาษา Java แล้วถือว่า วิศวกรเขียนโปรแกรม

กฎข้อที่ 6. ถ้า ฝึกการใช้งาน router แล้วถือว่า วิศวกรระบบเครือข่าย

กฎข้อที่ 7. ถ้า รู้จัก Pointer แสดงว่า ฝึกการเขียนภาษา C

กฎข้อที่ 8. ถ้า รู้จัก OOP แสดงว่า ฝึกการเขียนภาษา Java

กฎข้อที่ 9. ถ้า รู้จัก CISCO แสดงว่า ฝึกการใช้งาน router

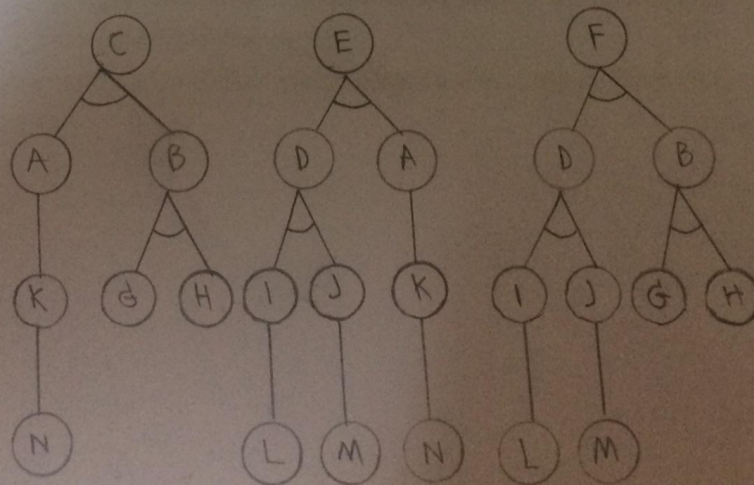
4.1) จะกำหนดค่าให้กับ Propositional Symbol ต่อไปนี้ ( 2 คะแนน )

A : วิศวกรระบบเครือข่าย	B : วิศวกรระบบปฏิบัติการ
C : Network Engineer	D : วิศวกรเขียนโปรแกรม
E : Programmer	F : Developer
G : ใช้งาน Windows	H : ใช้งาน Linux
I : ฝึกเขียนภาษา C	J : ฝึกเขียนภาษา Java
K : ฝึกใช้งาน router	L : รู้จัก Pointer
M : รู้จัก OOP	N : รู้จัก Cisco

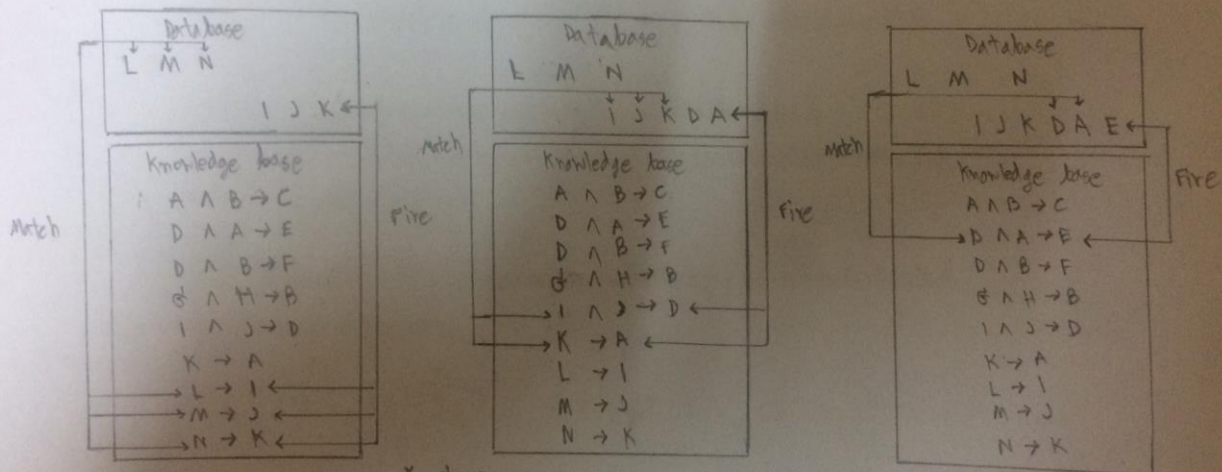
4.2) จงนำนิพจน์จากข้อ 4.1 มาสร้างเป็นกฎในรูปแบบของ IF-THEN ทั้งหมด 9 ข้อ ( 3 คะแนน )

1) IF A and B THEN C	2) IF D and A THEN E
3) IF D and B THEN F	4) IF G and H THEN B
5) IF I and J THEN D	6) IF K THEN A
7) IF L THEN I	8) IF M THEN J
9) IF N THEN K	

4.3) จงเขียนกราฟ AND-OR จากกฎด้านบนโดยมีเป้าหมายอยู่ที่อาชีพที่ควรทำงาน ( 4 คะแนน )



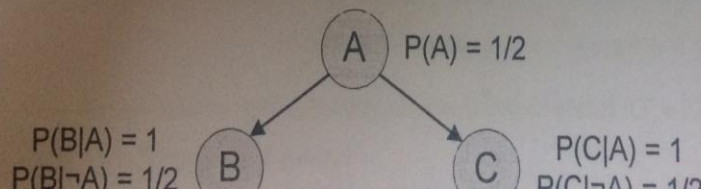
4.4) จงใช้วิธี Forward Chaining ในการแนะนำนักศึกษาว่าจะทำงานในอาชีพใด เมื่อนักศึกษาให้ข้อเท็จจริงว่า  
เขารู้จัก CISCO , OOP และ Pointer ( 4 คะแนน )



∴ นักศึกษาที่รู้จัก cisco, oop และ pointer นักศึกษาควรทำงาน Programmer #

5. คณะกรรมการสอบคัดเลือกนักศึกษาที่สมัครเรียนต่อหลักสูตร eNet-C ของภาควิชาเทคโนโลยีวิศวกรรม

อิเล็กทรอนิกส์ จึงได้สร้าง Bayesian Network ดังรูปต่อไปนี้ สอบ 203



$$* P(A|B) = P(B|A)$$

จากแบบฝึกหัด



6. จากกฎการแนะนำอาจารย์ที่ปรึกษาโครงการงานปริญญาโท ต่อไปนี้

กฎข้อที่ 1: IF ชอบภาษาจาวา AND ชอบประมวลผลภาพ THEN เลือก อ.สุพจน์ { CF: 0.9 }

กฎข้อที่ 2: IF ชอบภาษาจาวา THEN เลือก อ.สุพจน์ { CF: 0.3 }

กฎข้อที่ 3: IF ชอบไมโครโพรเซสเซอร์ AND ชอบระบบควบคุมดิจิทัล THEN เลือก อ.พิสิทธิ { CF: 0.6 }

กฎข้อที่ 4: IF ชอบภาษาจาวา AND ชอบระบบเครือข่าย THEN เลือก อ.ชูพันธุ์ { CF: 0.5 }

กฎข้อที่ 5: IF ชอบเขียนเว็บ AND ชอบระบบฐานข้อมูล AND ชอบระบบเครือข่าย THEN เลือก อ.ชูพันธุ์ { CF: 0.7 }

6.1) นักศึกษา ชอบภาษาจาวา (CF = 0.6) และ ชอบระบบเครือข่าย (CF = 0.4) ควรจะเลือกใครเป็นที่ปรึกษาโครงการงานปริญญาโท (2 คะแนน)

$$CF(H, E) = CF(E) \times CF = 0.6 \times 0.3 = 0.18$$

$$CF(H, E_1 \wedge E_2) = \min[CF(E_1), CF(E_2)] \times CF$$

$$= \min[0.6, 0.4] \times 0.5$$

$$= 0.4 \times 0.5 = 0.2$$

∴ สรุปได้ว่า เลือก อ.สุพจน์ ด้วย CF = 0.18 #

∴ สรุปได้ว่า เลือก อ.ชูพันธุ์ ด้วย CF = 0.2 #

6.2) นักศึกษา ชอบภาษาจาวา (CF = 0.5) และ ชอบไมโครโพรเซสเซอร์ (CF = 0.8) ควรจะเลือกใครเป็นที่ปรึกษาโครงการงานปริญญาโท (2 คะแนน)

$$CF(H, E) = CF(E) \times CF$$

$$= 0.5 \times 0.3$$

$$= 0.15$$

∴ สรุปได้ว่า ควรเลือก อ.สุพจน์ ด้วย CF = 0.15 #

6.3) นักศึกษา ชอบภาษาจาวา (CF = 0.3), ชอบระบบควบคุมดิจิทัล (CF = 0.5), ชอบประมวลผลภาพ (CF = 0.4) และชอบไมโครโพรเซสเซอร์ (CF = 0.8) ควรจะเลือกใครเป็นที่ปรึกษาโครงการงานปริญญาโท (4 คะแนน)

$$CF_1(H|E_1) = CF_1(E) \times CF_1 = 0.3 \times 0.3 = 0.09$$

$$CF_2(H|E_2 \wedge E_3) = \min[CF_2(E_2), CF_2(E_3)] \times CF_2$$

$$= \min[0.3, 0.4] \times 0.9$$

$$= 0.3 \times 0.9 = 0.27$$

$$CF(CF_1, CF_2) = CF_1(H, E_1) + (CF_2(H, E_2) \times [1 - CF_1(H, E_1)])$$

$$= 0.09 + (0.27 \times [1 - 0.09]) = 0.3357$$

$$CF(H, E_1 \wedge E_2) = \min[CF(E_1), CF(E_2)] \times CF$$

$$= \min[0.8, 0.5] \times 0.6$$

7. ปัญหาการฝังแคต เมื่อเราไปเที่ยวที่ชายทะเลและพบว่าคนที่ไปฝังแคตตามชายทะเล บางคนก็จะมีผิวเปลี่ยนเป็นสีแทน แต่บางคนต้องได้รับความทรมานจากผิวไหม้ เราต้องการหาว่าอะไรคือปัจจัยที่ทำให้คนที่ไปฝังแคตตามชายทะเลแล้วผิวไหม้หรือไม่ไหม้ โดยที่ข้อมูลสังเกตได้ประกอบด้วยความแตกต่างของสีผิว น้ำหนัก ส่วนสูงของผู้ที่ไปฝังแคต และการใช้โลชั่น ซึ่งบางคนก็ใช้โลชั่น บางคนก็ไม่ใช้ (15 คะแนน)

ชื่อ	สีผิว	ความสูง	น้ำหนัก	โลชั่น	ผลกระทบบนผิว
สมศรี	ดำ	ปานกลาง	เบา	ไม่ใช้	ไหม้
สมหญิง	ดำ	สูง	สมส่วน	ใช้	ไม่ไหม้
สมพร	น้ำตาล	เตี้ย	สมส่วน	ใช้	ไม่ไหม้
สมควร	ดำ	เตี้ย	สมส่วน	ไม่ใช้	ไหม้
สมปอง	ดำแดง	ปานกลาง	อ้วน	ไม่ใช้	ไหม้
สมหวัง	น้ำตาล	สูง	อ้วน	ไม่ใช้	ไม่ไหม้
สมชาย	น้ำตาล	ปานกลาง	อ้วน	ไม่ใช้	ไม่ไหม้
สมรัก	ดำ	เตี้ย	เบา	ใช้	ไม่ไหม้

จงคำนวณและเขียน Decision Tree ที่หาว่าปัจจัยอะไรที่ทำให้เกิดผลกระทบกับผิวกับข้อมูลข้างต้น

$$\begin{aligned}
 E(S) &= [-P(\text{ไหม้}) \log_2 P(\text{ไหม้})] + [-P(\text{ไม่ไหม้}) \log_2 P(\text{ไม่ไหม้})] \\
 &= [-\frac{3}{4} \log_2 (\frac{3}{4})] - [\frac{1}{4} \log_2 (\frac{1}{4})] \\
 &= 0.954
 \end{aligned}$$

สีผิว :  $S_B$  (ดำ),  $S_T$  (น้ำตาล),  $S_R$  (ดำแดง)

$$E(S_B) = -\frac{2}{4} \log_2 (\frac{2}{4}) - \frac{2}{4} \log_2 (\frac{2}{4}) = 0.5 + 0.5 = 1$$

$$E(S_T) = 0 \log_2 (0) - \frac{3}{3} \log_2 (\frac{3}{3}) = 0$$

$$E(S_R) = -\frac{1}{1} \log_2 (\frac{1}{1}) - 0 \log_2 (0) = 0$$

$$\begin{aligned}
 \text{Gain}(S, \text{สีผิว}) &= E(S) - [\frac{S_B}{S} E(S_B) + \frac{S_T}{S} E(S_T) + \frac{S_R}{S} E(S_R)] \\
 &= 0.954 - [\frac{4}{4} (1) + \frac{3}{4} (0) + \frac{1}{4} (0)] = 0.454
 \end{aligned}$$

ความสูง :  $S_L$  (เตี้ย),  $S_N$  (ปานกลาง),  $S_H$  (สูง)

$$E(S_L) = -\frac{1}{3} \log_2 (\frac{1}{3}) - \frac{2}{3} \log_2 (\frac{2}{3}) = 0.918$$

$$E(S_N) = -\frac{2}{5} \log_2 (\frac{2}{5}) - \frac{1}{5} \log_2 (\frac{1}{5}) = 0.918$$

$$E(S_H) = 0 \log_2 (0) - \frac{2}{2} \log_2 (\frac{2}{2}) = 0$$

$$\begin{aligned}
 \text{Gain}(S, \text{ความสูง}) &= E(S) - [\frac{S_L}{S} E(S_L) + \frac{S_N}{S} E(S_N) + \frac{S_H}{S} E(S_H)] \\
 &= 0.954 - [\frac{3}{4} (0.918) + \frac{2}{4} (0.918) + \frac{2}{4} (0)] \\
 &= 0.266
 \end{aligned}$$



ជំពូកទី ១ : ឈាម (S<sub>g</sub>) , ក្បាល (S<sub>m</sub>) , ចំណុះ (S<sub>l</sub>)

030513146 / 9

$$E(S_g) = -\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) = 1$$

$$E(S_m) = -\frac{1}{3} \log_2 \left(\frac{1}{3}\right) - \frac{2}{3} \log_2 \left(\frac{2}{3}\right) = 0.918$$

$$E(S_l) = -\frac{1}{3} \log_2 \left(\frac{1}{3}\right) - \frac{2}{3} \log_2 \left(\frac{2}{3}\right) = 0.918$$

$$\begin{aligned} \text{Gain}(S, \text{ជំពូកទី ១}) &= E(S) - \left[ \frac{S_g}{S} E(S_g) + \frac{S_m}{S} E(S_m) + \frac{S_l}{S} E(S_l) \right] \\ &= 0.954 - \left[ \frac{2}{8} (1) + \frac{3}{8} (0.918) + \frac{3}{8} (0.918) \right] \\ &= 0.0155 \end{aligned}$$

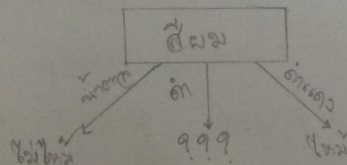
ជំពូកទី ២ : ក្បាល (S<sub>y</sub>) , ក្បាល (S<sub>n</sub>)

$$E(S_y) = -0 \log_2 (0) - \frac{3}{3} \log_2 \left(\frac{3}{3}\right) = 0$$

$$E(S_n) = -\frac{3}{5} \log_2 \left(\frac{3}{5}\right) - \frac{2}{5} \log_2 \left(\frac{2}{5}\right) = 0.991$$

$$\begin{aligned} \text{Gain}(S, \text{ជំពូកទី ២}) &= E(S) - \left[ \frac{S_y}{S} E(S_y) + \frac{S_n}{S} E(S_n) \right] \\ &= 0.954 - \left[ \frac{3}{8} (0) + \frac{5}{8} (0.991) \right] \\ &= 0.341 \end{aligned}$$

∴ attribute ក្បាល ផ្តល់ការបំបាត់ព័ត៌មានបានល្អជាង attribute ជំពូកទី ២ ដូច្នេះយើងអាចជ្រើសរើស attribute ក្បាល



ដើម្បីជ្រើសរើស attribute ល្អជាង attribute ផ្សេងទៀត យើងត្រូវធ្វើការបំបាត់ព័ត៌មាន ដោយយក attribute ក្បាល មកធ្វើការបំបាត់ព័ត៌មាន

ដូច្នេះយើងបានមកដល់ការបំបាត់ព័ត៌មាន

ជំពូកទី ១	ក្បាល	ជំពូកទី ២	ក្បាល	ក្បាល
ក្បាល	ក្បាល	ក្បាល	ក្បាល	ក្បាល
ក្បាល	ក្បាល	ក្បាល	ក្បាល	ក្បាល
ក្បាល	ក្បាល	ក្បាល	ក្បាល	ក្បាល
ក្បាល	ក្បាល	ក្បាល	ក្បាល	ក្បាល

$$E(S) = [-P(\text{หัว}) \log_2 P(\text{หัว})] + [-P(\text{หาง}) \log_2 P(\text{หาง})]$$

030513146 / 10

$$= -\frac{2}{4} \log_2 \left(\frac{2}{4}\right) - \frac{2}{4} \log_2 \left(\frac{2}{4}\right) = 1$$

entropy : หัว (S<sub>L</sub>), หางบน (S<sub>M</sub>), หาง (S<sub>T</sub>)

$$E(S_L) = -\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) = 1$$

$$E(S_M) = -\frac{1}{4} \log_2 \left(\frac{1}{4}\right) - 0 \log_2 (0) = 0$$

$$E(S_T) = 0 \log_2 (0) - \frac{1}{4} \log_2 \left(\frac{1}{4}\right) = 0$$

$$\text{Gain}(S, \text{entropy}) = E(S) - \left[\frac{2}{5} E(S_L) + \frac{5}{5} E(S_M) + \frac{5}{5} E(S_T)\right]$$

$$= 1 - \left[\frac{2}{4}(1) + \frac{1}{4}(0) + \frac{1}{4}(0)\right] = 0.5$$

ส่วนหัว : หางบน (S<sub>S</sub>), หางล่าง (S<sub>N</sub>)

$$E(S_S) = -\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) = 1$$

$$E(S_N) = -\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) = 1$$

$$\text{Gain}(S, \text{ส่วนหัว}) = E(S) - \left[\frac{5}{5} E(S_S) + \frac{5}{5} E(S_N)\right]$$

$$= 1 - \left[\frac{2}{4}(1) + \frac{2}{4}(1)\right] = 0$$

ส่วนหาง : หัว (S<sub>Y</sub>), หาง (S<sub>N</sub>)

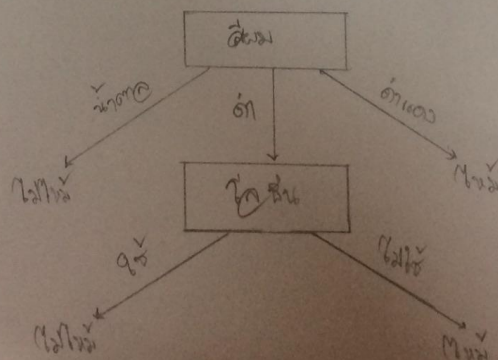
$$E(S_Y) = 0 \log_2 (0) - \frac{2}{2} \log_2 \left(\frac{2}{2}\right) = 0$$

$$E(S_N) = -\frac{2}{2} \log_2 \left(\frac{2}{2}\right) - 0 \log_2 (0) = 0$$

$$\text{Gain}(S, \text{ส่วนหาง}) = E(S) - \left[\frac{5}{5} E(S_Y) + \frac{5}{5} E(S_N)\right]$$

$$= 1 - \left[\frac{2}{4}(0) + \frac{2}{4}(0)\right] = 1$$

∴ attribute หัว มีความสัมพันธ์กับข้อมูลมากที่สุด attribute หาง มีความสัมพันธ์กับข้อมูลน้อยที่สุด



∴ เราสามารถสรุปได้ว่า หัวมีความสัมพันธ์กับข้อมูลมากที่สุด หางมีความสัมพันธ์กับข้อมูลน้อยที่สุด เราสามารถใช้ attribute หัว เพื่อทำนายค่าของข้อมูลได้



8. จากชุดข้อมูลในตาราง

( ข้อย่อยละ 3 คะแนน )

สัตว์	ออกลูกเป็นตัว	ออกลูกเป็นไข่	เลี้ยงลูกด้วยนม	มีขนตามร่างกาย	ผิวหนังเรียบหรือมีเกล็ด	สัตว์เลือดอุ่น	สัตว์เลือดเย็น	อยู่ทั้งบนและน้ำ	หายใจด้วยเหงือก	ตำแหน่ง
กบ		1					1	1		สัตว์ครึ่งบกครึ่งน้ำ
นกกระจอกเทศ	1			1		1				สัตว์จำพวกนก
จระเข้		1			1		1			สัตว์เลื้อยคลาน
ปลาหมอ		1			1		1		1	สัตว์จำพวกปลา
เต่า		1					1			สัตว์เลื้อยคลาน
สุนัข	1		1	1		1				สัตว์เลี้ยงลูกด้วยนม
ปลากัด					1		1		1	สัตว์จำพวกปลา
ลิง	1		1	1		1				สัตว์เลี้ยงลูกด้วยนม
สัตว์ปริศนา X		1			1		1			???
สัตว์ปริศนา Y	1			1						???

8.1) จงใช้เทคนิคแบบ 3NN (จาก Euclidean Distance) เพื่อหาว่าสัตว์ปริศนา X เป็นสัตว์สายพันธุ์ใด

$$\begin{aligned}
 d_{\text{กบ}} &= \sqrt{0^2 + (1-1)^2 + 0^2 + 0^2 + (1-0)^2 + 0^2 + (1-1)^2 + 0^2 + 0^2} = \sqrt{2} \\
 d_{\text{นกกระจอกเทศ}} &= \sqrt{(0-1)^2 + (1-0)^2 + 0^2 + (0-1)^2 + (1-0)^2 + (0-1)^2 + (1-0)^2 + 0^2 + 0^2} = \sqrt{6} \\
 d_{\text{จระเข้}} &= \sqrt{0^2 + (1-1)^2 + 0^2 + 0^2 + (1-1)^2 + 0^2 + (1-1)^2 + 0^2 + 0^2} = 0 \\
 d_{\text{ปลาหมอ}} &= \sqrt{0^2 + (1-1)^2 + 0^2 + 0^2 + (1-1)^2 + 0^2 + (1-1)^2 + 0^2 + (0-1)^2} = \sqrt{1} \\
 d_{\text{เต่า}} &= \sqrt{0^2 + (1-1)^2 + 0^2 + 0^2 + (1-0)^2 + 0^2 + (1-1)^2 + 0^2 + 0^2} = \sqrt{1} \\
 d_{\text{สุนัข}} &= \sqrt{(0-1)^2 + (1-0)^2 + (0-1)^2 + (0-1)^2 + (1-0)^2 + (0-1)^2 + (1-0)^2 + 0^2 + 0^2} = \sqrt{7} \\
 d_{\text{ปลากัด}} &= \sqrt{0^2 + (1-0)^2 + 0^2 + 0^2 + (1-1)^2 + 0^2 + (1-1)^2 + 0^2 + (0-1)^2} = \sqrt{2} \\
 d_{\text{ลิง}} &= \sqrt{(0-1)^2 + (1-0)^2 + (0-1)^2 + (0-1)^2 + (1-0)^2 + (0-1)^2 + (1-0)^2 + 0^2 + 0^2} = \sqrt{7}
 \end{aligned}$$

$\therefore$  3 อันดับที่มีระยะห่างกับข้อมูลที่มีจำนวนมากที่สุดคือ 0 หรือ 1 ดังนั้นสัตว์ปริศนา X คือ สัตว์เลื้อยคลาน #

8.2) จงใช้เทคนิคแบบ 1NN (จาก Euclidean Distance) เพื่อหาว่าสัตว์ปริศนา Y เป็นสัตว์สายพันธุ์ใด

$$\begin{aligned}
 d_{\text{กบ}} &= \sqrt{(1-0)^2 + (0-1)^2 + 0^2 + (1-0)^2 + 0^2 + 0^2 + (0-1)^2 + (0-1)^2 + 0^2} = \sqrt{5} \\
 d_{\text{นกกระจอกเทศ}} &= \sqrt{(1-1)^2 + 0^2 + 0^2 + (1-1)^2 + 0^2 + (0-1)^2 + 0^2 + 0^2 + 0^2} = \sqrt{1} \\
 d_{\text{จระเข้}} &= \sqrt{(1-0)^2 + (0-1)^2 + 0^2 + (1-0)^2 + (0-1)^2 + 0^2 + (0-1)^2 + 0^2 + 0^2} = \sqrt{5} \\
 d_{\text{ปลาหมอ}} &= \sqrt{(1-0)^2 + (0-1)^2 + 0^2 + (1-0)^2 + (0-1)^2 + 0^2 + (0-1)^2 + 0^2 + (0-1)^2} = \sqrt{6} \\
 d_{\text{เต่า}} &= \sqrt{(1-0)^2 + (0-1)^2 + 0^2 + (1-0)^2 + 0^2 + 0^2 + (0-1)^2 + 0^2 + 0^2} = \sqrt{4} \\
 d_{\text{สุนัข}} &= \sqrt{(1-1)^2 + 0^2 + (0-1)^2 + (1-1)^2 + 0^2 + (0-1)^2 + 0^2 + 0^2 + 0^2} = \sqrt{2} \\
 d_{\text{ปลากัด}} &= \sqrt{(1-0)^2 + 0^2 + 0^2 + (1-0)^2 + (0-1)^2 + 0^2 + (0-1)^2 + 0^2 + (0-1)^2} = \sqrt{6} \\
 d_{\text{ลิง}} &= \sqrt{(1-1)^2 + 0^2 + (0-1)^2 + (1-1)^2 + 0^2 + (0-1)^2 + 0^2 + 0^2 + 0^2} = \sqrt{2}
 \end{aligned}$$

$\therefore$  ระยะห่างของสัตว์ปริศนา Y ให้กับนกกระจอกเทศมากที่สุด  
 จึงสรุปได้ว่า สัตว์ปริศนา Y เป็นสัตว์จำพวกนก #