บทที่ 7

การเรียนรู้แบบเบย์ (Bayesian Learning)

วัตถุประสงค์การเรียนรู้

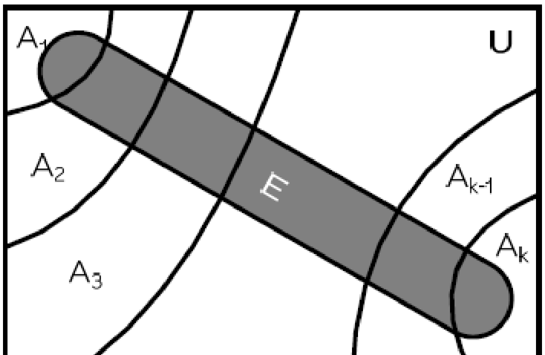
1. เพื่อให้ทราบถึงนิยามของทฤษฎีของเบย์
2. เพื่อให้เข้าใจหลักการทำงานของการจำแนกประเภทโดยใช้ทฤษฎีของเบย์
3. เพื่อให้เข้าใจหลักการทำงานของตัวจำแนกประเภทแบบเบย์อย่างง่ายและเครือข่ายงานแบบเบย์

การเรียนรู้แบบเบย์ (Bayesian Learning) เป็นการจำแนกประเภทรูปแบบหนึ่งที่อาศัยหลักการของความน่าจะเป็นเข้ามาช่วยในการหาคำตอบของประเภทตัวอย่างใหม่ ในบทนี้จะกล่าวถึงกฎของเบย์ที่นำมาใช้ทั้งในการเรียนรู้แบบเบย์อย่างง่าย (Naïve Bayesian) และ เครือข่ายงานแบบเบย์ (Bayesian Network) ดังรายละเอียดต่อไปนี้

7.1 ทฤษฎีของเบย์ (Bayes' theorem)

ทฤษฎีของเบย์ (Bayes' Theorem) ถูกพัฒนาขึ้นโดย Thomas Bayes โดยใช้หลักการของความน่าจะเป็นแบบมีเงื่อนไขมาพัฒนาเป็นทฤษฎีบทดังกล่าว โดยมีรายละเอียดดังต่อไปนี้

ถ้าให้เอกภพสัมพัทธ์ U ประกอบด้วยเหตุการณ์ที่ไม่สามารถเกิดขึ้นได้พร้อมกัน จำนวน k เหตุการณ์คือ $A_1, A_2, A_3, \ldots, A_k$ ดังรูปที่ 7.1 และให้ E เป็นเหตุการณ์หนึ่งในปริภูมิตัวอย่างที่เกิดจากการทดลองเดียวกันนี้และต้องเป็นส่วนหนึ่งของ A_i (i = 1, 2, 3, …, k) จะสามารถคำนวณความน่าจะเป็นแบบมีเงื่อนไขของเหตุการณ์หนึ่งใน A_i เมื่อเหตุการณ์ E เกิดขึ้นแล้วได้ดังสมการที่ (7.1)



รูปที่ 7.1 เหตุการณ์ E บนเหตุการณ์ k เหตุการณ์ที่เกิดพร้อมกันไม่ได้

$$P(A_i / E) = \frac{P(E / A_i) \cdot P(A_i)}{\sum_{i=1}^{k} P(E / A_i) \cdot P(A_i)} = \frac{P(E / A_i) \cdot P(A_i)}{P(E)} \quad (7.1)$$

7.2 การเรียนรู้แบบเบย์ (Bayesian Learning)

การเรียนรู้แบบเบย์ เป็นเทคนิคที่ใช้ทฤษฎีความน่าจะเป็น ตามกฎของเบย์ (Bayes' Theorem) เพื่อหาว่าสมมติฐานใดน่าจะถูกต้องที่สุด โดยใช้ความรู้ก่อนหน้า (Prior Knowledge) ได้แก่ ความน่าจะเป็นก่อนหน้าสำหรับสมมติฐานหนึ่ง ๆ ร่วมกับข้อมูล เช่น ความน่าจะเป็นที่สังเกตได้สำหรับสมมติฐานหนึ่ง ๆ เพื่อหาสมมติฐานที่ดีที่สุด

การเรียนรู้แบบเบย์อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน (ในที่นี่คือคลาสเป้าหมายหรือผลลัพธ์การทำนาย) โดยการเรียนรู้แบบเบย์เป็นการเรียนรู้เพิ่มได้ เนื่องจากตัวอย่างใหม่ที่ได้มาถูกนำมาปรับเปลี่ยนการแจกแจงซึ่งมีผลต่อการเพิ่ม หรือ ลดความน่าจะเป็น ทำให้มีการเรียนรู้ที่เปลี่ยนไป วิธีการนี้ตัวแบบจะถูกปรับเปลี่ยนไปตามตัวอย่างใหม่ที่ได้โดยผนวกกับความรู้เดิมที่มี ซึ่งการทำนายค่าคลาสเป้าหมายของตัวอย่างใช้ความน่าจะเป็นมากที่สุดของทุกสมมติฐาน

จากทฤษฎีของเบย์ เราสามารถคำนวณความน่าจะเป็นของสมมติฐานต่าง ๆ โดยใช้สมการที่ 7.2

$$P(h|D) = \frac{P(D|h) * P(h)}{P(D)} \quad (7.2)$$

โดย

- D แทนข้อมูลที่นำมาใช้ในการคำนวณการแจกแจงความน่าจะเป็น posteriori probability ของสมมติฐาน h คือ P(h|D) ตามทฤษฎี
- P(h) คือ ความน่าจะเป็นก่อนหน้าของสมมติฐาน h
- P(D) คือ ความน่าจะเป็นก่อนหน้าของชุดข้อมูลตัวอย่าง D
- P(h|D) คือ ความน่าจะเป็นของ h เมื่อรู้ D
- P(D|h) คือ ความน่าจะเป็นของ D เมื่อรู้ h

ตัวอย่างการคำนวณเพื่อเลือกสมมติฐานโดยกฎของเบย์

คนไข้คนหนึ่งไปตรวจหา โรคมะเร็ง ผลการตรวจเป็นบวก (+) อยากทราบว่า เราควรวินิจฉัยโรคคนไข้คนนี้ว่าเป็นโรคมะเร็งจริงหรือไม่ โดยมีข้อมูลความเป็นจริง ดังนี้

- ผลการตรวจเมื่อเป็นบวกจะให้ความถูกต้อง 98% กรณีที่มีโรคนั้นอยู่จริง
- ผลการตรวจเมื่อเป็นลบจะให้ความถูกต้อง 97% กรณีที่ไม่มีโรคนั้น
- 0.008 ของประชากรทั้งหมดเป็นโรคมะเร็ง

จากความน่าจะเป็นข้างต้น เราจะทราบว่าความน่าจะเป็นต่อไปนี้

P(cancer) = 0.008	P(~cancer) = 0.992
P(+ | cancer) = 0.98	P(- | cancer) = 0.02
P(+ |~ cancer) = 0.03	P(- |~ cancer) = 0.97

เราสามารถคำนวณค่าความน่ าจะเป็นของสมมติฐานว่าคนไข้เป็น หรือไม่เป็นโรคมะเร็ ง เมื่อทราบผลตรวจเป็นบวก โดยใช้กฎของเบย์ ดังนี้

<u>สมมติฐานที่ 1</u> คนไข้เป็นโรคมะเร็งจริงเมื่อมีผลการตรวจเป็นบวก

เขียนแทนด้วย P(cancer |+)

แทนค่าในสูตร

$$P(cancer \mid +) = \frac{P(+ \mid cancer)P(cancer)}{P(+)}$$

$$= 0.98*0.008$$

$$= 0.0078$$

<u>สมมติฐานที่ 2</u> คนไข้เป็นไม่เป็นโรคมะเร็งจริงเมื่อมีผลการตรวจเป็นบวก

เขียนแทนด้วย P(~cancer |+)

แทนค่าในสูตร

$$P(\sim cancer \mid +) = \frac{P(+ \mid \sim cancer)P(\sim cancer)}{P(+)}$$

$$= 0.03*0.992$$

$$= 0.0298$$

เนื่องจากผลรวมของ P(cancer |+) กับ P(~cancer |+) เท่ากับ 1 เราสามารถ Normalize ค่าของ P(cancer |+) = 0.0078/(0.0078+0.0298) = 0.21 และ P(~cancer |+) =0.0298/(0.0078+0.0298) = 0.79

สรุปว่า สมมติฐานที่ 1 มีค่าความน่าจะเป็นเท่ากับ 0.21 และสมมติฐานที่ 2 มีค่าความน่าจะเป็นเท่ากับ 0.79 ในการเลือกตอบสมติฐานเนื่องจากสมมติฐานที่ 2 มีค่ามากกว่า สมมติฐานที่ว่าคนไข้ไม่เป็นโรคมะเร็งเมื่อทราบผลตรวจเป็นบวก ด้วยความน่าจะเป็น 0.79 จึงถูกเลือก

7.3 ตัวจำแนกประเภทที่ดีที่สุดแบบเบย์ (Bayes Optimal Classification)

ในการจำแนกประเภทตัวอย่าง X ใด ๆ ที่น่าจะเป็นที่สุด ของกรณีที่ผลการจำแนกประเภทตัวอย่างมีความแตกต่างจากสมมติฐาน ที่ต่างกัน เราจะใช้ตัวจำแนกประเภทที่ดีที่สุดแบบเบย์ เพื่อจำแนกประเภทตัวอย่าง X ที่น่าจะเป็นที่สุด

กำหนดให้

- h_{MAP} (Maximum A Posterior Hypothesis) แทน สมมติฐานที่น่าจะเป็นที่สุด
- ถ้า h_{MAP} เป็นสมมติฐานที่น่าจะเป็นที่สุด h_{MAP} อาจไม่เป็นการจำแนกประเภทตั วอย่างที่น่าจะเป็นที่สุด (most probable classification)
- นิยามตัวจำแนกประเภทที่ดีที่สุดแบบเบย์

$$h_{MAP}(x) = \underset{v_j \in V}{\operatorname{argmax}} \sum_{h_j \in H} P(v_j|h_i)P(h_i|D) \qquad (7.3)$$

อธิบายได้โดย พิจารณาสมมติฐานทั้งสามต่อไปนี้

$P(h_1|D) = 0.4$ $P(h_2|D) = 0.3$ $P(h_3|D) = 0.3$

เมื่อให้ตัวอย่าง X ผลการจำแนกประเภทของสมมติฐานเป็นดังนี้

$h_1(X) = +$ $h_2(X) = -$ $h_3(X) = -$

จากตัวอย่างข้างต้น เราทราบว่า

$P(h_1|D) = 0.4$ $P(-|h_1) = 0$ $P(+|h_1) = 1$

$P(h_2|D) = 0.3$ $P(-|h_2) = 1$ $P(+|h_2) = 0$

$P(h_3|D) = 0.3$ $P(-|h_3) = 1$ $P(+|h_3) = 0$

จะได้ว่า

$$\sum_{h_i \in H} P(+|h_i)P(h_i|D) = 0.4$$

$$\sum_{h_i \in H} P(-|h_i)P(h_i|D) = 0.6$$

ดังนั้น การจำแนกประเภทตัวอย่าง X ที่มี MAP Class คือคลาสที่น่าจะเป็นมากที่สุด คือ –

7.4 วิธีการเรียนรู้เบย์อย่างง่าย (Naïve Bayesian Learning)

ตัวจำแนกแบบเบย์อย่างง่าย (Naive Bayesian Classifier) คือโมเดลการจำแนกประเภทข้อมูลที่ใช้หลักความน่าจะเป็นซึ่งอยู่บนพื้นฐานของ Bayes' Theorem และสมมติฐาน ที่ให้การเกิดของเหตุการณ์ต่าง ๆเป็นอิสระต่อกัน (Independence)

กำหนดให้ P(h) ความน่าจะเป็นที่จะเกิดเหตุการณ์ h และ P(h|D) คือความน่าจะเป็ นที่จะเกิดเหตุการณ์ h เมื่อเกิดเหตุการณ์ D จากตัวแปรที่กำหนดและแนวคิดของ Bayes' Theorem นั้นเราสามารถทำนายเหตุการณ์ที่พิจารณาได้จากการเกิดของเหตุการณ์ต่าง ๆ ได้ดังสมการ

$$P(h|D) = [P(D|h) * P(h)]/P(D)$$

7.4.1 การสร้างสมการการทำนาย

เมื่อเกิดเหตุการณ์ลมแรงจะสามารถทำนายการไปเล่นเทนนิสได้ดังสมการต่อไปนี้

P(การไปเล่นเทนนิส | ลมแรง) = P(ลมแรง | การไปเล่นเทนนิส) X P(การไปเล่นเทนนิส)/P(ลมแรง)

จากตัวอย่างที่กล่าวมานั้น เราสามารถทำนาย การไปเล่นเทนนิส โดยสังเกตุ ลมแรง อย่างไรก็ตามเหตุการณ์ที่นำมาใช้ในการทำนายนั้นต้องสอด คล้องกับเหตุการณ์ที่จะทำนาย เช่น ถ้าหากเราต้องการทำนายการไปเล่นเทนนิส เราก็จะไม่ใช้เหตุกา รณ์ฝนตกมาพิจารณา เพราะเหตุการณ์ ทั้งสองไม่มีความเกี่ยวข้องกัน

เมื่อพิจารณา ทฤษฎีของเบย์ และคุณสมบัติเป็นอิสระต่อกันที่กล่าวมาในตอนต้น จะสามารถแสดงการคัดแยกประเภทข้อมูล ที่มีเหตุการณ์ (E) มากกว่า 1 เหตุการณ์ ได้ดังสมการต่อไปนี้

$$P(h|E1,E2,...,En) = \frac{P(E1|h) \times P(E2|h) \times .... \times P(En|h) \times P(h)}{P(E1,E2,....,En)} \quad (7.4)$$

ในการคัดแยกประเภทข้อมูลสำหรับเหตุการณ์ใดๆนั้น ขึ้น อยู่กับค่า P(h|E1,E2,...,En) โดยจะจัดอยู่ในประเภทใดๆ เมื่อทำให้ P(h|E1,E2,...,En) มีค่าสูงสุด เมื่อพิจารณาถึงค่าสูงสุดดังนั้นจึงสามารถตัดพจน์ P(E1,E2,....,En) ออกจากการพิจารณาเพราะเป็นค่าคงที่ จึงได้สมการของการหา MAP class ของตัวจำแนกประเภทแบบเบย์อย่างง่ายดังสมการที่ (7.5)

$$\text{argmax}\{P(hi|E1,E2,...,En)\} \quad (7.5)$$

7.4.2 ตัวจำแนกประเภทแบบเบย์อย่างง่าย (Naïve Bayes Classifer)

จากการนิยามข้างต้นเราสามารถสร้างตัวจำแนกประเภทแบบเบย์อย่างง่ายได้ดังต่อไปนี้

Assume target function $f : X \to V$, where each instance x described by attributes $\langle a_1, a_2 \ldots a_n \rangle$.
Most probable value of $f(x)$ is:

$$\begin{aligned} v_{MAP} &= \underset{v_j \in V}{\text{argmax}}\, P(v_j|a_1, a_2 \ldots a_n) \\ v_{MAP} &= \underset{v_j \in V}{\text{argmax}} \frac{P(a_1, a_2 \ldots a_n|v_j)P(v_j)}{P(a_1, a_2 \ldots a_n)} \\ &= \underset{v_j \in V}{\text{argmax}}\, P(a_1, a_2 \ldots a_n|v_j)P(v_j) \end{aligned}$$

Naive Bayes assumption:

$$P(a_1, a_2 \ldots a_n|v_j) = \prod_i P(a_i|v_j)$$

which gives

Naive Bayes classifier: $v_{NB} = \underset{v_j \in V}{\text{argmax}}\, P(v_j) \prod_i P(a_i|v_j)$

รูปที่ 7.2 การนิยามของ Naïve Bayes Classifier (V_{NB})

จากรูปที่ 7.2 จะได้ว่าคำตอบซึ่งก็คือคลาสผลลัพธ์ vi ใด ๆ ที่ถูกเลือก จะเป็นคลาสที่มีค่าความน่าจะเป็นมากที่สุด (MAP class) ที่ได้จากการคำนวณ และจะถูกใช้เป็นคำตอบให้แก่ตัวอย่างค้นถามที่ต้องการ

7.4.3 การสร้างโมเดล และการนำตัวจำแนกประเภทแบบเบย์อย่างง่ายไปใช้งาน

ตัวอย่างดังต่อไปนี้เป็นจำแนกหรือทำนายพฤติกรรมการออกไปเล่นเทนนิสด้วยค่าคลาสผลลัพธ์ 2 ค่า คือ ไปเล่น (P) หรือ ไม่ไปเล่น (N) โดยดูจากข้อมูลที่เก็บไว้ในวันที่ผ่านๆ มา จำนวน 14 วัน ดังแสดงในตารางที่ 7.1

ตารางที่ 7.1 ข้อมูลการออกไปเล่นกีฬา

Outlook	Temperature	Humidity	Windy	Class
sunny	hot	high	false	N
sunny	hot	high	true	N
overcast	hot	high	false	P
rain	mild	high	false	P
rain	cool	normal	false	P
rain	cool	normal	true	N
overcast	cool	normal	true	P
sunny	mild	high	false	N
sunny	cool	normal	false	P
rain	mild	normal	false	P
sunny	mild	normal	true	P
overcast	mild	high	true	P
overcast	hot	normal	false	P
rain	mild	high	true	N

ตัวอย่าง กำหนดให้ตัวอย่างใหม่ (New instance) ที่ต้องการจำแนกประเภท คือ

x = <Sunny, Cool, High, True>

จงทำนายว่าหากมีคุณสมบัติตาม x นักกีฬาคนนั้นจะออกไปเล่นเทนนิสหรือไม่

วิธีการคำนวณ

เราสามารถแบ่งกลุ่มข้อมูลตามประเภทของ Class ได้ดังต่อไปนี้

Outlook	Temperature	Humidity	Windy	Class
overcast	hot	high	false	P
rain	mild	high	false	P
rain	cool	normal	false	P
overcast	cool	normal	true	P
sunny	cool	normal	false	P
rain	mild	normal	false	P
sunny	mild	normal	true	P
overcast	mild	high	true	P
overcast	hot	normal	false	P

9

Outlook	Temperature	Humidity	Windy	Class
sunny	hot	high	false	N
sunny	hot	high	true	N
rain	cool	normal	true	N
sunny	mild	high	false	N
rain	mild	high	true	N

5

ดังนั้นจึงได้ค่าของความน่าจะเป็นก่อนหน้า (Prior Probabilities) ดังนี้

$P(P) = 9/14 = 0.64$

$P(N) = 5/14 = 0.36$

และสามารถคำนวณความน่าเป็นของแต่ละคุณสมบัติ (Likelihood Probabilities)ได้ดังต่อไปนี้

Outlook

$P(\text{sunny}\|P) = 2/9$	$P(\text{sunny}\|N) = 3/5$
$P(\text{overcast}\|P) = 4/9$	$P(\text{overcast}\|N) = 0/5$
$P(\text{rain}\|P) = 3/9$	$P(\text{rain}\|N) = 2/5$

Temperature

$P(\text{hot}\|P) = 2/9$	$P(\text{hot}\|N) = 2/5$
$P(\text{mild}\|P) = 4/9$	$P(\text{mild}\|N) = 2/5$
$P(\text{cool}\|P) = 3/9$	$P(\text{cool}\|N) = 1/5$

Humitdity

$P(\text{high}\|P) = 3/9$	$P(\text{high}\|N) = 4/5$
$P(\text{normal}\|P) = 6/9$	$P(\text{normal}\|N) = 1/5$

Windy

P(true|P) = 3/9	P(true|N) = 3/5
P(false|P) = 6/9	P(false|N) = 2/5

พิจารณาข้อมูลสภาพอากาศดังต่อไปนี้

< Outlook = sunny, Temp = cool, Hum = high, Wind = true>

ต้องการทำนายว่านักกีฬาที่เจอสภาพอากาศในรายละเอียดข้างต้น จะออกไปเล่นเทนนิสหรือไม่ คำนวนความน่าจะเป็นได้ดังนี้

จาก

$$v_{MAP} = \arg\max_{vj \in \{yes,no\}} P(v_j) \prod_i P(a_i \mid v_j)$$

$$= \arg\max_{vj \in \{yes,no\}} P(v_j) P(Outlook = sunny \mid v_j) P(Temp = cool \mid v_j) P(Hum = high \mid v_j) P(Windy = true \mid v_j)$$

แทนค่าในสมการ

$$Pr(P)Pr(sunny \mid P)Pr(cool \mid P)Pr(high \mid P)Pr(true \mid P)$$

$$= (9/14) * (2/9) * (3/9) * (3/9) * (3/9) = .0053$$

$$Pr(N)Pr(sunny \mid N)Pr(cool \mid N)Pr(high \mid N)Pr(true \mid N)$$

$$= (5/14) * (3/5) * (1/5) * (4/5) * (3/5) = .0206$$

จากผลการคำนวณสรุปได้ว่า ถ้าเจอสภาพอากาศ < Outlook = sunny, Temp = cool, Hum = high, Wind = true> แล้ว นักกีฬาจะไม่ไปเล่ นเทนนิส (คลาส N) เนื่องจากให้ค่าสมมติฐานที่สูงกว่า คือ 0.0206

7.5 เครือข่ายงานแบบเบย์ (Bayesian Belief Networks)

เครือข่ายงานแบบเบย์ เป็นแบบจำลองกราฟของความน่าจะเป็น หรือเรียกอีกชื่อหนึ่งว่า "Bayes net"

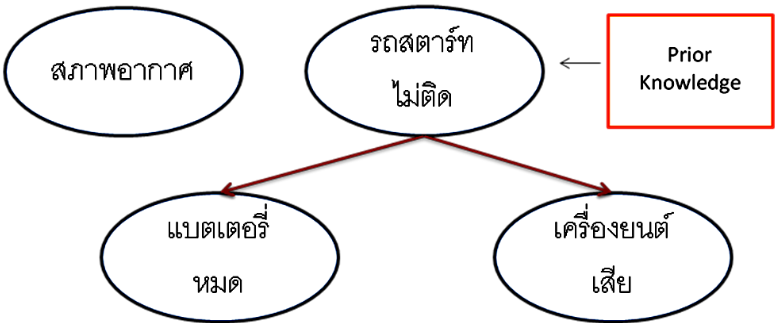
ซึ่งเป็นวิธีการเรียนรู้ที่ลดข้อจำกัดของเรียนรู้แบบเบย์อย่างง่ายในสมมุติฐานความไม่ขึ้นต่อกันระหว่างคุณสมบัติ โดย Bayesian network ใช้อธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไข (Condition Independent) ระหว่างตัวแปร

7.5.1 เงื่อนไขของเครือข่ายงานแบบเบย์

- ความสัมพันธ์ของแต่ละโหนดจะไม่วนกลับมาหาโหนดเดิม

- แต่ละโหนดจะมีความสัมพันธ์กันตามทิศทางที่แบบจำลองนำเสนอ
- โหนดทั้งหมดใน Bayesian Network แต่ละโหนดจะแทนด้วยตัวแปรต่าง ๆที่เกี่ยวข้องกับเหตุการณ์หรือข้อมูลที่สนใจ
- การเชื่อมต่อระหว่างคู่โหนดด้วยลูกศร ถ้าลูกศรจากโหนด X ชี้ไปหาโหนด Y จะเรียกว่าโหนด X เป็นโหนดพ่อแม่(parents) ของ Y
- แต่ละโหนด X_i จะมีเงื่อนไขการกระจายความน่าจะเป็น $P(X_i|parents(X_i))$ ซึ่งส่งผลต่อโหนดพ่อแม่ของแต่ละโหนด

เครือข่ายงานแบบเบย์ ถูกนำมาอธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไข จึงทำให้อาจมีบางโหนดไม่มีความสัมพันธ์ใด ๆ กับโหนดที่เหลือก็ได้ แสดงดังรูปที่ 7.3



รูปที่ 7.3 เครือข่ายงานแบบเบย์ของเหตุการณ์รถสตาร์ทไม่ติด

โดยมีรายละเอียดดังนี้

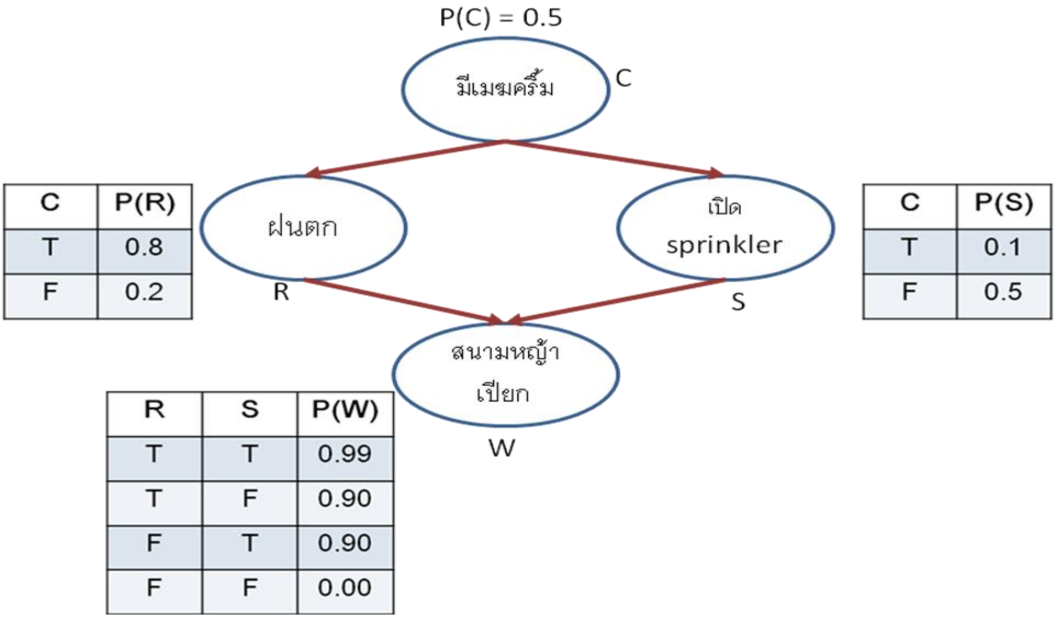
- ความไม่ขึ้นต่อกันอย่างมีเงื่อนไข (Condition Independent)
  X ไม่ขึ้นกับ Y อย่างมีเงื่อนไข คือ ความน่าจะเป็นของ X ไม่ขึ้นอยู่กับค่าของ Y เมื่อรู้ค่า Z
  นั่นคือ $P(X|Y,Z) = P(X|Z)$
- ความไม่ขึ้นต่อกันอย่างมีเงื่อนไข ทำให้การหาความน่าจะเป็นของตัวแปรที่ต้องการง่ายขึ้นโดยไม่ต้องสนใจตัวแปร อื่น เช่น
  - ฟ้าร้อง จะไม่ขึ้นกับ ฝนตก เสมอไป ถ้ารู้ว่าเกิดฟ้าแลบ
  - เนื่องจากเมื่อเกิด ฟ้าแลบ แล้วจะตามมาด้วย ฟ้าร้อง เสมอ
  - P(ฟ้าร้อง|ฝนตก,ฟ้าแลบ) = P(ฟ้าร้อง|ฟ้าแลบ)
- ในเครือข่ายงานแบบเบย์ ตัวแปรแต่ละตัวจะมีความน่าจะเป็นเฉพาะ ที่อาจเป็นความน่าจะเป็นของโหนดเริ่มต้น หรือ ความน่าจะเป็นที่ได้จากความสัมพันธ์มากกว่าหนึ่งโหนด โดยความน่าจะเป็นที่มาจากตัวแปรมากกว่าหนึ่งตัวเรียกว่า “ความน่าจะเป็นร่วม” (Joint Probability)
- สมการของความน่าจะเป็นร่วมมีดังนี้

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i)) \tag{7.6}$$

  - Parents(Xi) หมายถึง โหนดพ่อแม่โดยตรงของ Xi

7.5.2 ตัวอย่างเครือข่ายงานแบบเบย์

<u>ตัวอย่างที่ 1</u>



C	P(R)
T	0.8
F	0.2

C	P(S)
T	0.1
F	0.5

R	S	P(W)
T	T	0.99
T	F	0.90
F	T	0.90
F	F	0.00

รูปที่ 7.4 เครือข่ายงานแบบเบย์ของเหตุการณ์สนามหญ้าเปียก

จากรูปที่ แสดง เครือข่ายงานแบบเบย์ ที่แสดงความน่าจะเป็นในการเกิดเหตุการณ์สนามเปียก จากข่ายงานข้างต้นจะมีการคำนวณตารางความน่าจะเป็นในแต่ละเงื่อนไข ที่เรียกว่าตาราง CPT (The Conditional Probability Table) ซึ่งใช้แสดงความขึ้นอยู่แก่กันในตัวแปรที่เกี่ยวข้อง เช่น โอกาสที่สนามหญ้าเปียก (W) หากมีฝนตก (R=T) และ เปิด sprinker (S=T) มีค่า 0.99 เป็นต้น

ตัวอย่างการคำนวณ

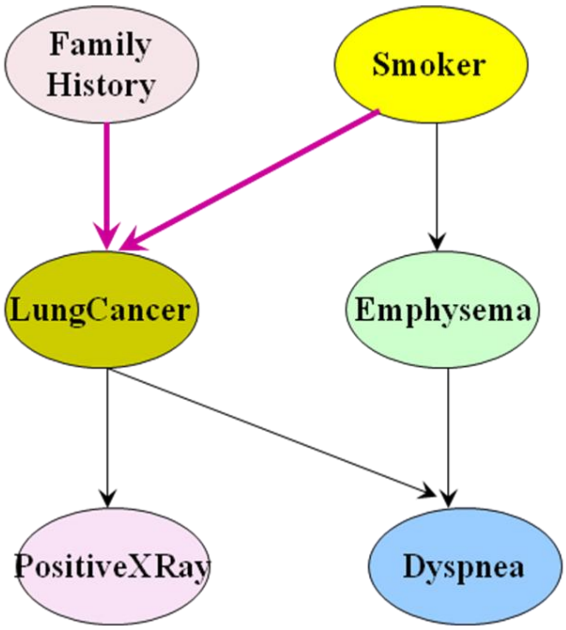
จากเครือข่ายงานแบบเบย์ ที่แสดงความน่าจะเป็นของสนามหญ้าเปียก จงหาความน่าจะเป็ นที่สนามหญ้าจะเปียกเพราะฝนตก โดยที่ sprinkler ไม่ได้ทำงาน โดยสภาพของอากาศในตอนนั้นไม่มีเมฆครึ้ม

$$P(W , \sim S , R , \sim C) = P(W| \sim S , R)\, P(\sim S| \sim C)\, P(R| \sim C)\, P(\sim C)$$
$$= 0.90 * 0.5 * 0.2 * 0.5$$
$$= 0.0450$$

ดังนั้นความน่าจะเป็นที่สนามหญ้าจะเปียกเพราะฝนตกโดยที่ sprinkler ไม่ได้ทำงานและอากาศไม่มีเมฆครึ้ม คือ 0.045 (4.5%)

ตัวอย่างที่ 2

รูปที่ 7.5 แสดงเครือข่ายงานแบบเบย์ ของเงื่อนไขการเกิดโรคต่าง ๆ ที่มีความเกี่ยวพันธ์กัน ได้แก่ ประวัติครอบครัว (Family History:FS) เป็นคนสูบบุหรี่ (Smoker:S) เป็นมะเร็งถุงลม (LungCancer:LC) เป็นถุงลมโป่งพอง (Emphysema) มีผลการ X-Ray เป็นบวก (PositiveXRay:P) และ มีอาการหอบเหนื่อย (Dyspnea:D)



รูปที่ 7.5 แสดงเครือข่ายงานแบบเบย์ของเงื่อนไขการเกิดโรคต่าง ๆ

ซึ่งจากความรู้ก่อนหน้ามีข้อมูลปรากฏดังตาราง CPT ดังนี้

	(FH, S)	(FH, ~S)	(~FH, S)	(~FH, ~S)
LC	0.8	0.5	0.7	0.1
~LC	0.2	0.5	0.3	0.9

ตัวอย่างการคำนวณ

จงหาค่าความน่าจะเป็นที่ผู้ป่วยจะเป็นโรคมะเร็งปอด ในกรณีที่พบว่ามีประวัติคนในครอบครัวเป็นโรคมะเร็ง และเป็นคนที่สูบบุหรี่

P(LC = yes, FH = yes ,S = yes) = P(LC,FH,S)

= P(FH)* P(S)*P(LC |FH,S)

=P(FH)* P(S)*0.8

7.5.3 บทสรุปของเครือข่ายงานแบบเบย์

Bayesian network ยอมให้สับเซตของตัวแปรอิสระต่อกันอย่างมีเงื่อนไข โดยใช้รูปกราฟแสดง ความเกี่ยวพันระหว่างตัวแปร มีที่จุดยอดคือ ตัวแปรและเส้นเชื่อมระหว่างจุดคือการไม่เป็นอิสระต่อกันของตัวแปรเหล่านั้น การกำหนดเครือข่ายงานแบบเบย์ จะมีผลต่อการคำนวณ กล่าวคือ

- ถ้าผู้ใช้กำหนดเครือข่ายทั้งหมด โปรแกรมเพียงคำนวณค่าเดิมในเครือข่าย
- ถ้าผู้ใช้กำหนดเครือข่ายบางส่วน โปรแกรมต้องมีการทดสอบการเป็นอิสระ ต่อกันของแต่ละส่วนในเครือข่าย
- ถ้าผู้ใช้ไม่กำหนดโครงสร้างเครือข่ายเลย โปรแกรมต้องทดลองทุกการจัด หมู่ที่เป็นไปได้ซึ่งใช้เวลาในการคำนวณนาน

สรุป

การจำแนกประเภทเป็น การศึกษาการแบ่งกลุ่มข้อมูล พบในสาขาวิชาทั้งทางด้านสถิติ เทคโนโลยีสารสนเทศ และการเรียนรู้ด้วยเครื่อง (Machine Learning) การใช้ความน่าจะเป็นในการจำแนกประเภทจะใช้ทฤษฎีของเบย์ ซึ่งถ้าตัวแปรเ ป็นอิสระต่อกัน เราใช้วิธีการ เรียนรู้แบบเบย์อย่างง่าย (Naive Bayesian Learning) และในกรณีที่ตัวแปรไม่เป็นอิสระต่อกัน เราเลือกใช้ วิธีเครือข่ายงานแบบเบย์ (Bayesian Network) โดยใช้ความน่าเป็นก่อนหน้ามาใช้ร่วมในการคำนวณ

แบบฝึกหัด

1. จงสรุปความแตกต่าง และประเภทง านที่เหมาะสมของตัวจำแนกประเภทแบบเบย์อย่างง่าย (Naïve Bayesian Classifier) และ เครือข่ายงานแบบเบย์ (Bayesian Belief Network)
2. จากตารางข้อมูลที่ 7.1 จงจำแนกประเภทของข้อมูลใหม่ที่มีคุณสมบัติดังต่อไปนี้
   2.1 x = < Overcast, Hot, High, True>
   2.2 x = <Rainy, Mild, Normal, False>
3. จากตัวอย่างที่ 1 ของเครือข่ายงานแบบเบย์ จงหาค่าความน่าจะเป็นที่ สนามหญ้าจะเปียก ในกรณีที่พบว่าไม่มีเมฆครึ้ม ฝนไม่ตก แต่เปิด sprinkler น้ำ