

## บทที่ 4

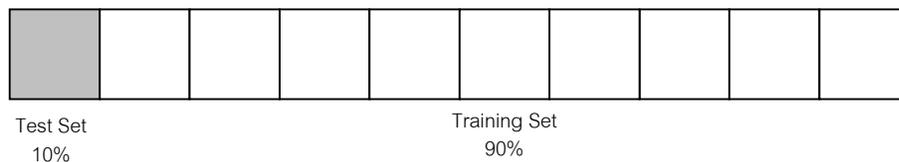
### ผลการทดลอง

เนื้อหาในบทนี้กล่าวถึง ข้อมูลที่ใช้ในการทดสอบ, ผลการทดลอง โดยผลการทดลอง แบ่งออกเป็นสองส่วนคือ ผลการเปรียบเทียบระหว่างการลงคะแนนโดยกฎ<sup>+</sup> กับแบ็กกิง และผลการเปรียบเทียบระหว่างการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> กับการลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้ และวิเคราะห์ผลการทดลอง

#### 4.1 ข้อมูลที่ใช้ในการทดลอง

ในหัวข้อนี้อธิบายถึงข้อมูลและลักษณะของข้อมูลที่ใช้ในการทดลองรวม สำหรับข้อมูลที่นำมาใช้ในการทดลองนั้นได้เลือกใช้ชุดข้อมูลจากคลังข้อมูลของยูซีไอ (UCI: University of California, Irvine) เพราะคลังชุดข้อมูลดังกล่าวมีชุดข้อมูลที่มีคุณลักษณะหลากหลาย และได้รับการเชื่อถือจากนักวิจัยทั่วโลก อีกทั้งยังมีผลงานวิจัยเป็นจำนวนมากที่ใช้ชุดข้อมูลของยูซีไอนี้ โดยงานวิจัยนี้เลือกชุดข้อมูลจำนวน 15 ชุดข้อมูลประกอบไปด้วย Audiology, Australian, Balance-Scale, Bridges, Car, Dermatology, Hayes-Roth, Heart, Hepatitis, Horse-Colic, Labor-Neg, Liver-Disorders, Soybean, TAE และ Zoo คุณลักษณะของข้อมูลถูกรวบรวมและแสดงในคอลัมน์ที่ 1-6 ในตารางที่ 4.1

ภาพที่ 4.1 : การแบ่งข้อมูลสำหรับการเรียนรู้และทดสอบ



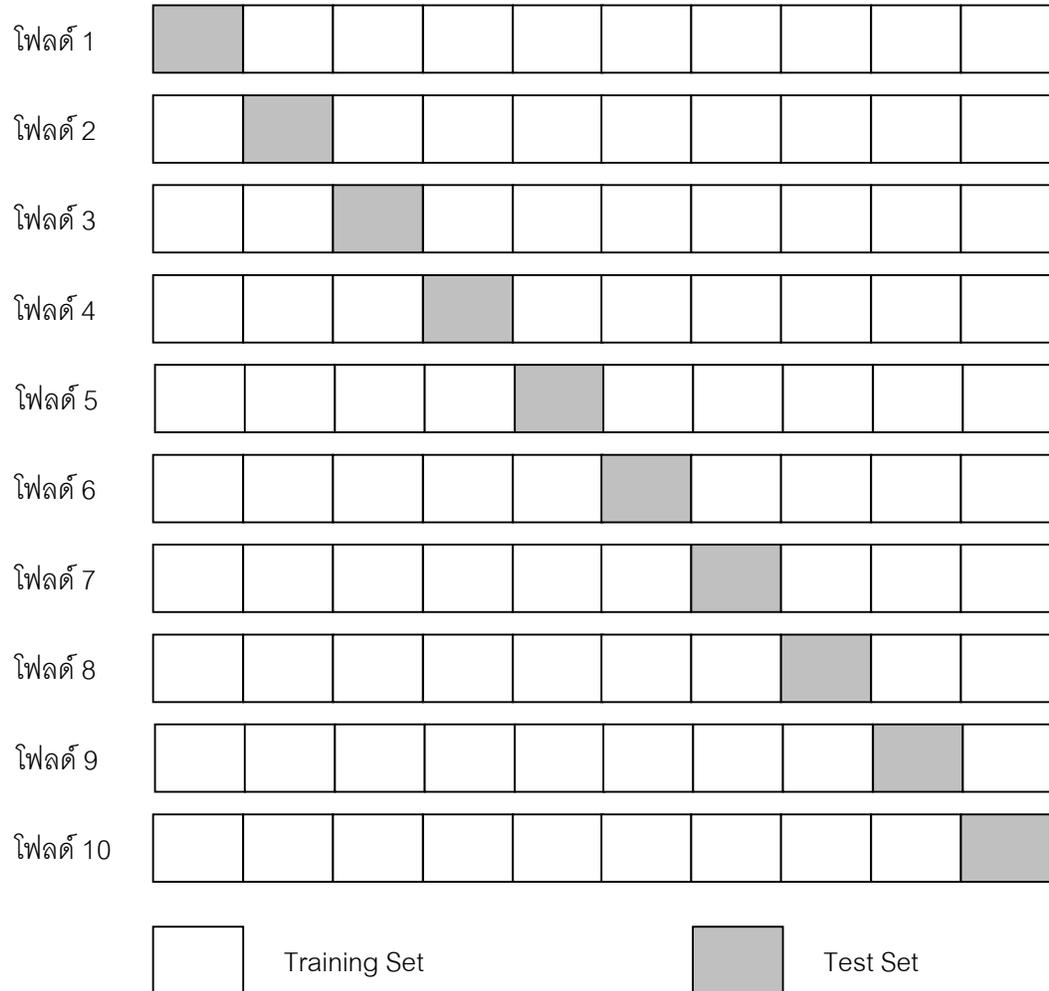
ตารางที่ 4.1 : ชุดข้อมูลที่ใช้ในการทดสอบงานวิจัย

ชุดข้อมูล	จำนวน ตัวอย่าง	จำนวนคุณลักษณะ		จำนวน คลาส	มีข้อมูลขาดหาย ( มี(Y) / ไม่มี(N) )
		ต่อเนื่อง	ไม่ต่อเนื่อง		
AUDIOLOGY	180	0	69	24	Y
AUSTRALIAN	621	6	8	2	Y
BALANCE-SCALE	625	0	4	3	N
BRIDGES	105	0	10	6	Y
CAR	1,728	0	6	4	N
DERMATOLOGY	366	1	33	6	Y
HAYES-ROTH	132	0	4	3	N
HEART	243	7	6	2	N
HEPATITIS	140	6	13	2	Y
HORSE-COLIC	270	7	16	2	Y
LABOR-NEG	40	8	8	2	N
LIVER-DISORDER	311	6	0	2	N
SOYBEAN	307	0	35	19	Y
TAE	151	1	2	3	N
ZOO	101	0	16	7	N

ข้อมูลทั้งหมดที่ใช้ในการเรียนรู้ของต้นไม้ตัดสินใจถูกแบ่งเป็นชุดเรียนรู้ (Training Set) จำนวนร้อยละเก้าสิบของข้อมูลทั้งหมด และชุดข้อมูลสำหรับทดสอบ (Test Set) จำนวนร้อยละสิบของข้อมูลทั้งหมดดังในภาพที่ 4.1

วิทยานิพนธ์ฉบับนี้ นำวิธีทดสอบความถูกต้องแบบไขว้ข้าม 10 ชุดข้อมูลย่อย (Ten Fold Cross-Validation) คือ แบ่งข้อมูลที่ใช้ในการทดสอบเป็นชุดข้อมูลย่อยจำนวน 10 ชุดข้อมูลย่อยดังแสดงในภาพที่ 4.2

ภาพที่ 4.2 : การแบ่งข้อมูลจำนวน 10 ข้อมูลย่อย (10 Folds)



#### 4.2 ผลการทดลอง

ผลการทดลองแบ่งออกเป็น 2 ส่วน คือ ผลการทดลองเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยกฎ<sup>+</sup> กับแบ็กกิง และผลการทดลองเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> กับการลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้<sup>+</sup>

##### 4.2.1 ผลการทดลองเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยกฎ<sup>+</sup> กับแบ็กกิง

ตารางที่ 4.2 แสดงค่าความถูกต้องที่แบ็กกิงให้คำตอบสุดท้ายถูกต้องเป็นร้อยละในคอลัมน์ที่ 2 และแสดงค่าความถูกต้องของการลงคะแนนโดยกฎ<sup>+</sup> ที่ให้คำตอบสุดท้ายถูกต้องเป็นร้อยละเช่นกัน โดยใช้ค่าความใกล้เคียง 0.6, 0.7, 0.8 และ 0.9 ในการจัดกลุ่มให้กับกฎในคอลัมน์ที่ 3, 4, 5 และ 6 ตามลำดับ เครื่องหมาย ⊕ ซึ่งชี้ให้เห็นว่าการลงคะแนนโดยกฎ<sup>+</sup> ให้คำตอบถูกต้องมากกว่าที่นัยสำคัญ 0.1 ในทางตรงกันข้ามเครื่องหมาย ⊖ ซึ่งชี้ให้เห็นว่าการลงคะแนนโดยกฎ<sup>+</sup> ให้คำตอบถูกต้องน้อยกว่าที่นัยสำคัญ 0.1

ตารางที่ 4.2 : การเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยกฎ<sup>+</sup> กับแบ็กกิง

ชุดข้อมูล	แบ็กกิง	การลงคะแนนโดยกฎ <sup>+</sup>			
		0.6	0.7	0.8	0.9
AUDIOLOGY	78.64±6.56	77.60±7.54	77.14±5.40	76.62±5.92	79.60±4.74
AUSTRALIAN	83.91±4.74	84.20±6.50	84.78±5.28	84.93±4.55	84.94±4.45
BALANCE-SCALE	78.58±4.09	78.09±3.43	79.21±3.66	79.70±4.79⊖	80.01±4.54⊖
BRIDGES	59.91±14.86	61.82±15.68⊖	59.00±18.20	61.73±16.57	61.73±16.57
CAR	93.81±1.37	89.76±2.00⊖	90.34±2.33⊖	93.17±1.87	94.33±1.97
DERMATOLOGY	95.36±4.04	95.90±3.29⊖	95.91±3.90⊖	96.19±3.88⊖	95.63±3.69
HAYES-ROTH	74.89±8.65	74.89±13.02	75.66±11.94	74.95±11.96	73.41±11.60
HERT	53.33±12.74	48.52±8.08⊖	45.56±5.25⊖	46.30±6.68⊖	49.61±9.83⊖
HEPATITIS	72.96±15.07	75.00±14.08	65.38±16.94⊖	69.75±18.55	69.08±11.15
HORSE-COLIC	81.00±7.61	78.00±7.03⊖	76.00±11.91⊖	77.33±12.89	80.33±9.83
LABOR-NEG	67.50±22.50	67.50±22.50	67.50±22.50	67.50±22.50	67.50±22.50
LIVER-DISORDER	45.55±8.46	48.40±9.42	84.40±9.42	46.40±8.85	48.98±9.50
SOYBEAN	85.6±76.39	86.68±5.62	85.02±5.45	86.62±5.75	86.96±4.85
TAE	43.00±15.95	45.00±14.08	45.00±14.08	43.00±15.95	43.00±15.95
ZOO	92.00±7.48	92.00±7.48	92.00±7.48	94.00±6.63⊖	94.00±6.63⊖

4.2.2 ผลการทดลองเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> กับการลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้

ลักษณะเดียวกับกับตารางที่ 4.2 ตารางที่ 4.3 แสดงค่าความถูกต้องที่การลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้ให้คำตอบสุดท้ายถูกต้องเป็นร้อยละในคอลัมน์ที่ 2 และแสดงค่าความถูกต้องของการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> ที่ให้คำตอบสุดท้ายถูกต้องเป็นร้อยละเช่นกัน โดยใช้ค่าความใกล้เคียง 0.6, 0.7, 0.8 และ 0.9 ในการจัดกลุ่มให้กับกฎในคอลัมน์ที่ 3, 4, 5 และ 6 ตามลำดับ ทั้งนี้เครื่องหมาย  $\oplus$  ชี้ให้เห็นว่าการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> ให้คำตอบถูกต้องมากกว่าที่สำคัญ 0.1 ในทางตรงกันข้ามเครื่องหมาย  $\ominus$  ชี้ให้เห็นว่าการลงคะแนนโดยตัวอย่างเรียนรู้<sup>+</sup> ให้คำตอบถูกต้องน้อยกว่าที่นัยสำคัญ 0.1

#### 4.3 วิเคราะห์ผลการทดลอง

ตารางที่ 4.3 แสดงจำนวนชุดข้อมูลที่การลงคะแนนโดยกฎ<sup>+</sup> ให้คำตอบถูกต้องมากกว่าที่นัยสำคัญ 0.1 และน้อยกว่าที่นัยสำคัญ 0.1 เมื่อเปรียบเทียบกับแบ็กกิง

จากตารางที่ 4.3 พบว่าการลงคะแนนโดยกฎ<sup>+</sup> ที่ค่าความใกล้เคียง 0.7 ให้ค่าความถูกต้องมากขึ้นที่นัยสำคัญ 0.1 จำนวนชุดข้อมูลน้อยที่สุดคือเพียง 1 ชุดข้อมูล ในขณะที่การลงคะแนนโดยกฎ<sup>+</sup> ที่ค่าความใกล้เคียง 0.6 และ ค่าความใกล้เคียง 0.9 ให้ค่าความถูกต้องมากขึ้นที่นัยสำคัญ 0.1 เท่ากันคือ 2 ชุดข้อมูล และที่ค่าความใกล้เคียง 0.8 ให้ค่าความถูกต้องมากขึ้นที่นัยสำคัญ 0.1 มากที่สุดคือ 3 ชุดข้อมูล แต่เมื่อสังเกตจำนวนชุดข้อมูลที่การลงคะแนนโดยกฎ<sup>+</sup> ให้ค่าความต้องน้อยกว่าที่นัยสำคัญ 0.1 จะเห็นได้ว่าการลงคะแนนโดยกฎ<sup>+</sup> ที่ค่าความใกล้เคียง 0.6 และที่ค่าความใกล้เคียง 0.7 มีจำนวนสูงใกล้เคียงกันคือ 3 ชุดข้อมูล และ 4 ชุดข้อมูลตามลำดับ ในขณะที่การลงคะแนนโดยกฎ<sup>+</sup> ที่ค่าความใกล้เคียง 0.8 และที่ค่าความใกล้เคียง 0.9 มีจำนวนแค่ 1 ชุดข้อมูลเท่านั้น

ตารางที่ 4.4 แสดงจำนวนชุดข้อมูลที่การลงคะแนนโดยตัวอย่าง<sup>+</sup> ให้คำตอบถูกต้องมากกว่าที่นัยสำคัญ 0.1 และน้อยกว่าที่นัยสำคัญ 0.1 เมื่อเปรียบเทียบกับการลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้

ตารางที่ 4.3 : การเปรียบเทียบความถูกต้องระหว่างการลงคะแนนโดยตัวอย่างเรียนรู้อยู่<sup>+</sup> กับกับการ  
ลงคะแนนแบบทั่วไปโดยตัวอย่างเรียนรู้อยู่<sup>+</sup>

ชุดข้อมูล	กับการ ลงคะแนน แบบทั่วไป โดยตัวอย่าง เรียนรู้อยู่ <sup>+</sup>	การลงคะแนนโดยตัวอย่างเรียนรู้อยู่ <sup>+</sup>			
		0.6	0.7	0.8	0.9
AUDIOLOGY	73.64±5.43	58.69±4.60 <sup>e</sup>	65.69±7.78 <sup>e</sup>	73.67±5.72	76.12±6.24 <sup>Ⓢ</sup>
AUSTRALIAN	84.49±4.00	85.07±4.20	85.51±4.58	85.51±4.30	85.22±3.99
BALANCE- SCALE	81.45±5.47	80.81±4.81 <sup>e</sup>	80.81±4.32	81.29±5.68	81.45±5.74
BRIDGES	63.64±19.13	59.82±17.00 <sup>e</sup>	62.64±19.37	63.64±17.78	64.55±18.89
CAR	94.68±1.28	91.67±1.29 <sup>e</sup>	93.40±1.19 <sup>e</sup>	94.45±1.48	94.56±1.77
DERMATOLOGY	91.51±6.23	91.25±6.29	93.44±4.65 <sup>Ⓢ</sup>	93.71±5.37 <sup>Ⓢ</sup>	93.71±4.05 <sup>Ⓢ</sup>
HAYES-ROTH	70.89±10.69	73.30±12.73	74.12±11.15 <sup>Ⓢ</sup>	72.64±10.52 <sup>Ⓢ</sup>	71.15±9.71
HERT	78.15±6.92	68.52±14.37 <sup>e</sup>	72.22±9.55 <sup>e</sup>	73.33±9.55	74.82±8.73
HEPATITIS	78.75±7.58	79.37±8.93	79.37±8.93	78.75±7.58	78.75±7.58
HORSE-COLIC	85.33±5.62	70.33±11.00 <sup>e</sup>	80.33±12.60 <sup>e</sup>	82.67±11.62	85.33±5.21
LABOR-NEG	67.50±22.50	67.50±22.50	67.50±22.50	67.50±22.50	67.50±22.50
LIVER- DISORDER	57.98±5.25	57.98±5.25	57.98±5.25	57.98±5.25	57.98±5.25
SOYBEAN	87.28±5.41	63.94±11.39 <sup>e</sup>	78.84±7.00 <sup>e</sup>	85.00±3.38 <sup>e</sup>	87.60±3.87
TAE	43.97±14.49	39.67±14.64 <sup>e</sup>	44.33±15.06	44.33±15.06	43.67±3.78
ZOO	91.00±11.36	40.64±15.20 <sup>e</sup>	84.09±13.61 <sup>e</sup>	92.00±6.00	92.00±6.00

ตารางที่ 4.4 : จำนวนข้อมูลการลงคะแนนโดยกฎ<sup>+</sup> เปรียบเทียบกับแบ็กกิง

ค่าความ ใกล้เคียง	ดีกว่าที่ นัยสำคัญ 0.1	แย่กว่าที่ นัยสำคัญ 0.1
0.6	2	3
0.7	1	4
0.8	3	1
0.9	2	1

ตารางที่ 4.5 : จำนวนข้อมูลการลงคะแนนโดยตัวอย่าง<sup>+</sup> เปรียบเทียบกับการลงคะแนนแบบทั่วไป  
โดยตัวอย่างเรียนรู้

ค่าความ ใกล้เคียง	ดีกว่าที่ นัยสำคัญ 0.1	แย่กว่าที่ นัยสำคัญ 0.1
0.6	0	9
0.7	2	6
0.8	2	1
0.9	2	0

จากตารางที่ 4.4 พบว่าการลงคะแนนโดยตัวอย่าง<sup>+</sup> ที่ค่าความใกล้เคียง 0.6 ไม่มีชุดข้อมูลใดเลยให้ความถูกต้องมากขึ้นที่นัยสำคัญ 0.1 ในขณะที่การลงคะแนนโดยตัวอย่าง<sup>+</sup> ที่ค่าความใกล้เคียง 0.7, ค่าความใกล้เคียง 0.8 และ ค่าความใกล้เคียง 0.9 ให้ความถูกต้องมากขึ้นที่นัยสำคัญ 0.1 เท่ากันคือ 2 ชุดข้อมูล แต่เมื่อสังเกตจำนวนชุดข้อมูลที่มีการลงคะแนนโดยตัวอย่าง<sup>+</sup> ให้ความถูกต้องน้อยลงที่นัยสำคัญ 0.1 จะเห็นได้ว่าการลงคะแนนโดยตัวอย่าง<sup>+</sup> ที่ค่าความใกล้เคียง 0.6 และที่ค่าความใกล้เคียง 0.7 มีจำนวนสูงมากคือ 9 ชุดข้อมูล และ 6 ชุดข้อมูลตามลำดับ ในขณะที่การลงคะแนนโดยตัวอย่าง<sup>+</sup> ที่ค่าความใกล้เคียง 0.8 มีจำนวนแค่ 1 ชุดข้อมูล และที่ค่าความใกล้เคียง 0.9 ไม่มีชุดข้อมูลใดเลยที่ให้ค่าความถูกต้องน้อยลงที่นัยสำคัญ 0.1

จากการการสรุปข้างต้นจะเห็นว่า ผลการเปรียบเทียบทั้งสองการเปรียบเทียบล้วน แสดงผลออกมาในทิศทางเดียวกันคือ การใช้กฎที่มีค่าความใกล้เคียง 0.6 และ 0.7 ร่วมกับการ ลงคะแนนทำให้คำตอบสุดท้ายที่ได้จากรวมต้นไม่มีความถูกต้องน้อยลง เนื่องจากกฎที่ถูกจัดกลุ่ม อยู่ในกลุ่มเดียวกันมีความแตกต่างกันมากเกินไปจนไม่สามารถถือได้ว่าเป็นกฎเดียวกัน และส่งผล ให้คำตอบที่ได้ไม่ถูกต้อง ในขณะที่ค่าความใกล้เคียง 0.8 และ 0.9 ให้ผลดีคือมีเพียง 1 ชุดข้อมูล หรือไม่มีชุดข้อมูลใดเลยที่ให้ค่าความถูกต้องลดน้อยลงที่นัยสำคัญ 0.1 ดังนั้นค่าความใกล้เคียง ระหว่างกฎที่น้อยที่สุดที่ทำให้การรวมต้นไม่ตัดสินใจมีค่าความถูกต้องดีขึ้นคือ 0.8