

## บทที่ 4

### ผลการดำเนินงานวิจัย

จากการดำเนินงานวิจัยการคัดเลือกปัจจัยสำคัญที่ส่งผลต่อระดับเศรษฐกิจครัวเรือนด้วยวิธีการทำเหมืองข้อมูลโดยจะวิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio เมื่อทราบปัจจัยที่สำคัญแล้วจะนำไปเปรียบเทียบระหว่าง All Feature ซึ่งเป็นข้อมูลที่ไม่ได้ผ่านเทคนิคใด ๆ นำมาเปรียบเทียบระหว่าง ผลลัพธ์ปัจจัยสำคัญที่ได้มาจากเทคนิค Gain Ratio ด้วยเทคนิค Decision Tree จะทำการแบ่งข้อมูลสัดส่วนร้อยละ 60:40, 70:30, และ 80:20 ของข้อมูลจำนวน 1,751 ระเบียบ (ครัวเรือน) จำนวน 17 ปัจจัย ด้วยโปรแกรมโปรแกรม RapidMiner Studio และทำการแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และวิธี 10-Fold Cross Validation จากนั้นประเมินประสิทธิภาพด้วยค่าค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision: P) ค่าความระลึก (Recall: R) ค่าถ่วงดุล (F-Measure) และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย (Root Mean Squared Error: RMSE) ดังนี้

#### 4.1 ผลการสร้างตัวแบบ Decision Tree กับข้อมูล All Feature

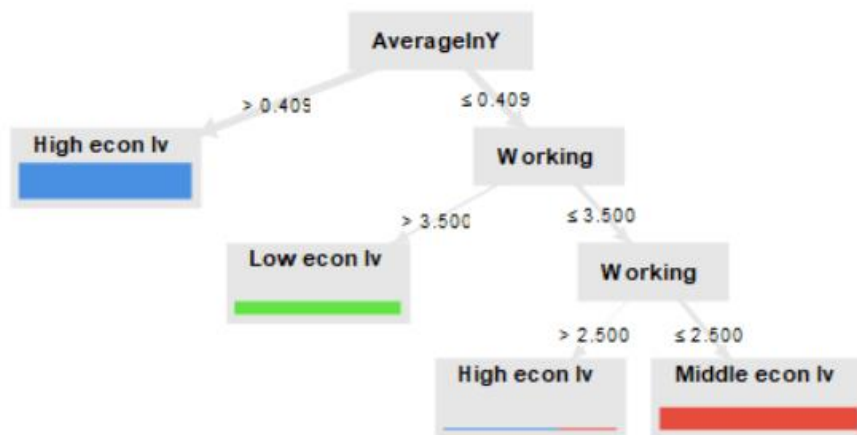
ทดลองการแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และวิธี 10-Fold Cross Validation โดยการทดลองตัวแบบแต่ละการทดลองทั้งหมด 10 ครั้ง การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40, 70:30 และ 80:20 เพื่อใช้ในการวัดประสิทธิภาพ มีรายละเอียดดังนี้

#### 4.1.1 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40

การสร้างตัวแบบกับชุดข้อมูล All Feature ด้วยเทคนิค Decision Tree โดยจะมีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิคด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และวิธี 10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.1 และตารางที่ 4.2

**ตารางที่ 4.1** ผลการทดสอบ All Feature กับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.00	99.18	98.81	98.99	0.088
2	99.00	99.18	98.81	98.99	0.088
3	99.00	99.18	98.81	98.99	0.088
4	99.00	99.18	98.81	98.99	0.088
5	99.00	99.18	98.81	98.99	0.088
6	99.00	99.18	98.81	98.99	0.088
7	99.00	99.18	98.81	98.99	0.088
8	99.00	99.18	98.81	98.99	0.088
9	99.00	99.18	98.81	98.99	0.088
10	99.00	99.18	98.81	98.99	0.088
เฉลี่ย	99.00	99.18	98.81	98.99	0.088



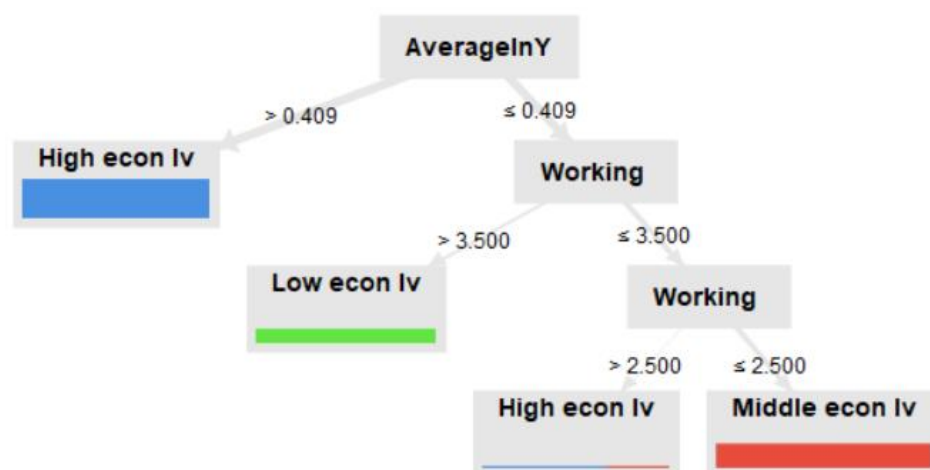
ภาพที่ 4.1 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จากภาพที่ 4.1 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 4 กฎ ดังนี้

- 1) IF AverageInY > 0.409 THEN High income
- 2) IF AverageInY ≤ 0.409 AND Working ≥ 3.500 THEN Low econ lv
- 3) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 THEN High econ lv
- 4) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 THEN Middle econ lv

ตารางที่ 4.2 ผลการทดสอบ All Feature การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.00	99.18	98.81	98.99	0.088
2	99.00	99.18	98.81	98.99	0.088
3	99.00	99.18	98.81	98.99	0.088
4	99.00	99.18	98.81	98.99	0.088
5	99.00	99.18	98.81	98.99	0.088
6	99.00	99.18	98.81	98.99	0.088
7	99.00	99.18	98.81	98.99	0.088
8	99.00	99.18	98.81	98.99	0.088
9	99.00	99.18	98.81	98.99	0.088
10	99.00	99.18	98.81	98.99	0.088
เฉลี่ย	99.00	99.18	98.81	98.99	0.088



ภาพที่ 4.2 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.2 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 4 กฎ ดังนี้

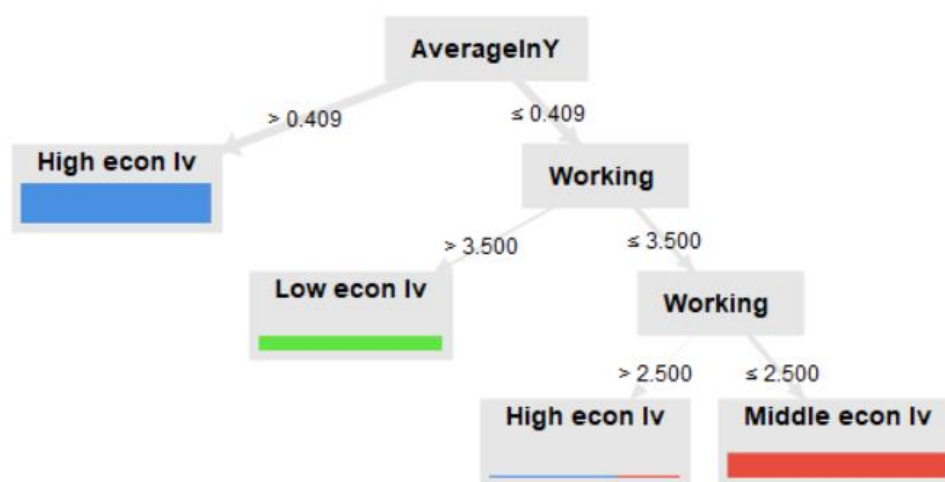
- 1) IF AverageInY >0.409 THEN High income
- 2) IF AverageInY  $\leq$ 0.409 AND Working  $\geq$ 3.500 THEN  
Low econ lv
- 3) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working >2.500 THEN High econ lv
- 4) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working  $\leq$ 0.500 THEN Middle econ lv

#### 4.1.2 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30

การสร้างตัวแบบกับชุดข้อมูล All Feature ด้วยเทคนิค Decision Tree โดยจะมีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิค ด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 70:30 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และ วิธี 10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.3 และตารางที่ 4.4

ตารางที่ 4.3 ผลการทดสอบ All Feature การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	98.86	99.03	98.61	98.82	0.092
2	98.86	99.03	98.61	98.82	0.092
3	98.86	99.03	98.61	98.82	0.092
4	98.86	99.03	98.61	98.82	0.092
5	98.86	99.03	98.61	98.82	0.092
6	98.86	99.03	98.61	98.82	0.092
7	98.86	99.03	98.61	98.82	0.092
8	98.86	99.03	98.61	98.82	0.092
9	98.86	99.03	98.61	98.82	0.092
10	98.86	99.03	98.61	98.82	0.092
เฉลี่ย	98.86	99.03	98.61	98.82	0.092



ภาพที่ 4.3 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จากภาพที่ 4.3 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 4 กฎ ดังนี้

1) IF AverageInY >0.409 THEN High income

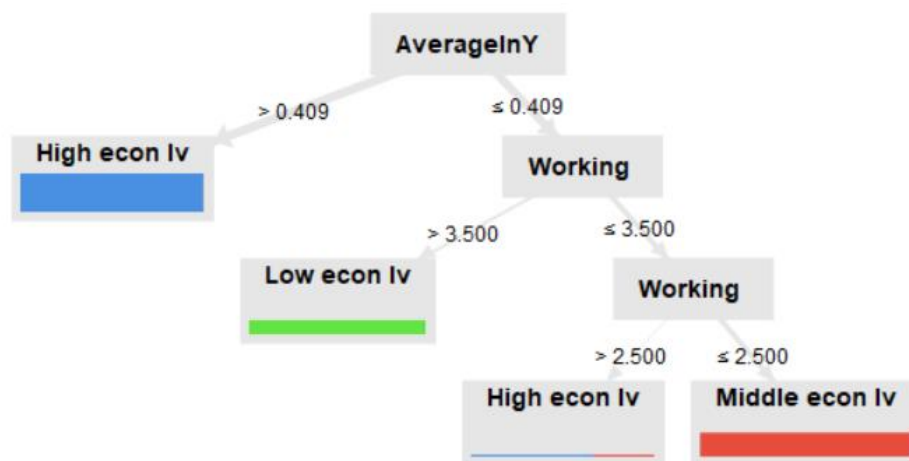
2) IF AverageInY  $\leq$ 0.409 AND Working  $\geq$ 3.500 THEN  
Low econ lv

3) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working >2.500 THEN High econ lv

4) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working  $\leq$ 0.500 THEN Middle econ lv

**ตารางที่ 4.4** ผลการทดสอบ All Feature การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่ง  
สัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	98.86	99.03	98.61	98.82	0.092
2	98.86	99.03	98.61	98.82	0.092
3	98.86	99.03	98.61	98.82	0.092
4	98.86	99.03	98.61	98.82	0.092
5	98.86	99.03	98.61	98.82	0.092
6	98.86	99.03	98.61	98.82	0.092
7	98.86	99.03	98.61	98.82	0.092
8	98.86	99.03	98.61	98.82	0.092
9	98.86	99.03	98.61	98.82	0.092
10	98.86	99.03	98.61	98.82	0.092
เฉลี่ย	98.86	99.03	98.61	98.82	0.092



ภาพที่ 4.4 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.4 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎดังนี้

- 1) IF AverageInY > 0.409 THEN High income
- 2) IF AverageInY ≤ 0.409 AND Working ≥ 3.500 THEN Low econ lv
- 3) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 THEN High econ lv
- 4) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 THEN Middle econ lv

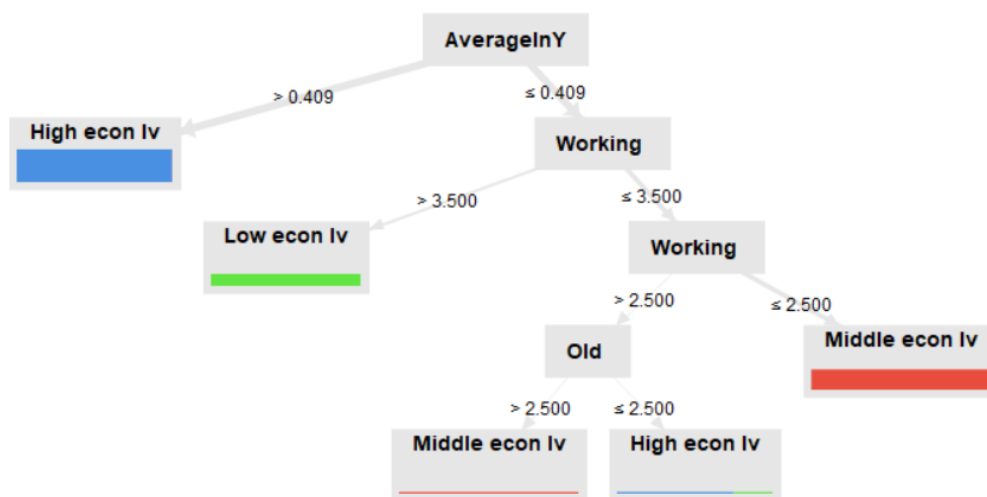
#### 4.1.3 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20

การสร้างตัวแบบกับชุดข้อมูล All Feature ด้วยเทคนิค Decision Tree โดยจะมีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิคด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 80:20 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และวิธี 10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.5 และตารางที่ 4.6



ตารางที่ 4.5 ผลการทดสอบ All Feature การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.43	99.28	99.39	99.33	0.075
2	99.43	99.28	99.39	99.33	0.075
3	99.43	99.28	99.39	99.33	0.075
4	99.43	99.28	99.39	99.33	0.075
5	99.43	99.28	99.39	99.33	0.075
6	99.43	99.28	99.39	99.33	0.075
7	99.43	99.28	99.39	99.33	0.075
8	99.43	99.28	99.39	99.33	0.075
9	99.43	99.28	99.39	99.33	0.075
10	99.43	99.28	99.39	99.33	0.075
เฉลี่ย	99.43	99.28	99.39	99.33	0.075



ภาพที่ 4.5 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จากภาพที่ 4.5 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 5 กฎ ดังนี้

1) IF AverageInY >0.409 THEN High income

2) IF AverageInY  $\leq$ 0.409 AND Working  $\geq$ 3.500 THEN  
Low econ lv

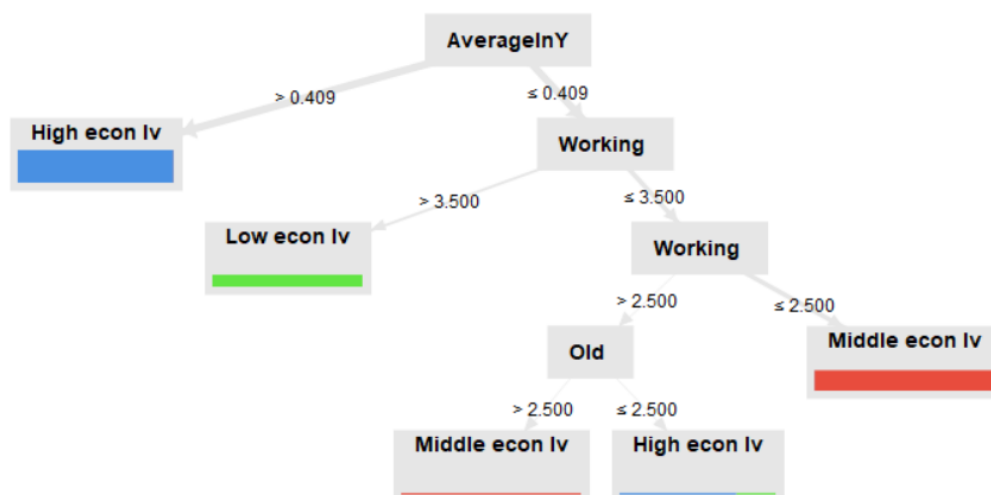
3) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working >2.500 AND Old  $\leq$ 2.500 THEN Middle econ lv

4) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working >2.500 AND Old  $\leq$ 2.500 THEN High income

5) IF AverageInY  $\leq$ 0.409 AND Working  $\leq$ 3.500 AND  
Working >2.500 THEN Middle econ lv

**ตารางที่ 4.6** ผลการทดสอบ All Feature การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่ง  
สัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.43	99.28	99.39	99.33	0.075
2	99.43	99.28	99.39	99.33	0.075
3	99.43	99.28	99.39	99.33	0.075
4	99.43	99.28	99.39	99.33	0.075
5	99.43	99.28	99.39	99.33	0.075
6	99.43	99.28	99.39	99.33	0.075
7	99.43	99.28	99.39	99.33	0.075
8	99.43	99.28	99.39	99.33	0.075
9	99.43	99.28	99.39	99.33	0.075
10	99.43	99.28	99.39	99.33	0.075
เฉลี่ย	99.43	99.28	99.39	99.33	0.075



ภาพที่ 4.6 แสดงแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.6 เมื่อนำแผนภาพ All Feature วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎดังนี้ 5 กฎ ดังนี้

- 1) IF AverageInY > 0.409 THEN High income
- 2) IF AverageInY ≤ 0.409 AND Working ≥ 3.500 THEN Low econ lv
- 3) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 AND Old ≤ 2.500 THEN Middle econ lv
- 4) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 AND Old ≤ 2.500 THEN High income
- 5) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 THEN Middle econ lv

#### 4.1.4 ผลการเปรียบเทียบการสร้างตัวแบบกับชุดข้อมูล All Feature

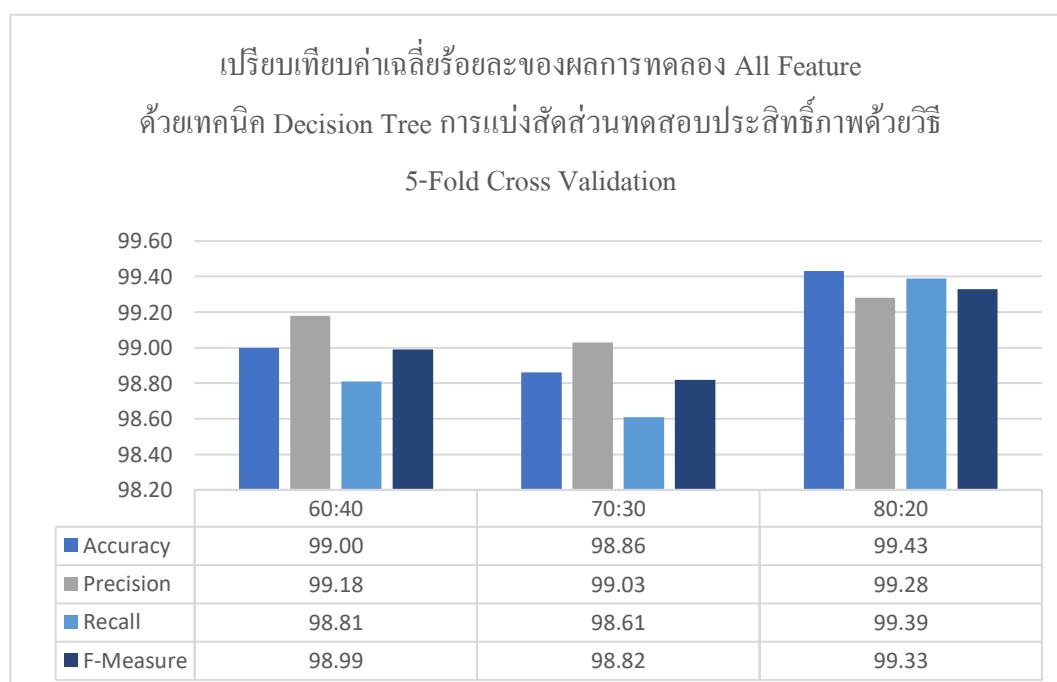
ผลการเปรียบเทียบเฉลี่ยสุดท้ายของ All Feature ด้วยเทคนิค Decision Tree การแบ่งข้อมูลสัดส่วนร้อยละ 60:40, 70:30, และ 80:20 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วย

วิธี 5-Fold Cross Validation และ วิธี 10-Fold Cross Validation ค่า Maximal Depth 10 จากนั้น ประเมินประสิทธิภาพด้วยค่าค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย ดังนี้

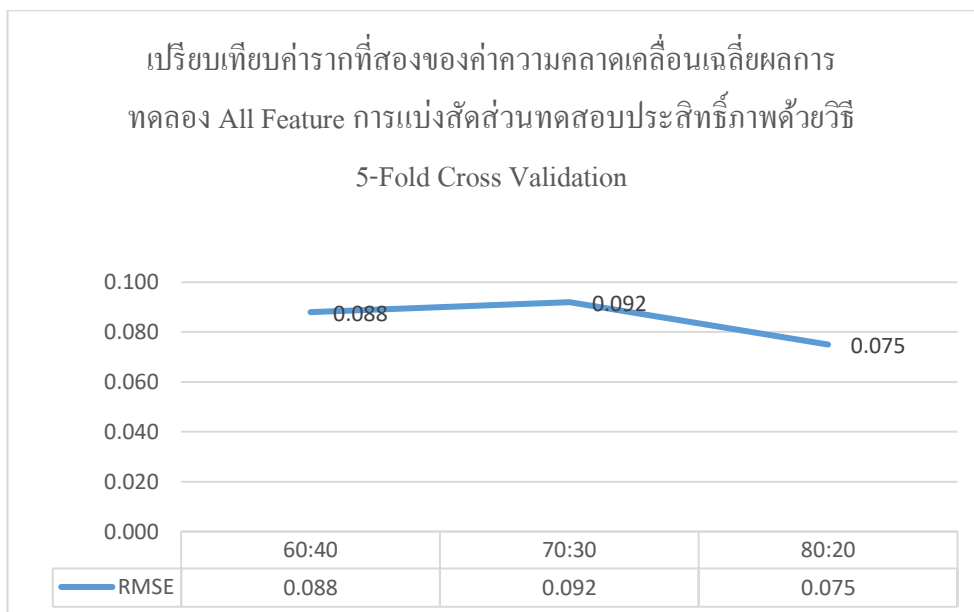
**ตารางที่ 4.7** เปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง All Feature ด้วยเทคนิค Decision Tree แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

เปอร์เซ็นต์	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
60:40	99.00	99.18	98.81	98.99	0.088
70:30	98.86	99.03	98.61	98.82	0.092
80:20	99.43	99.28	99.39	99.33	0.075

จากตารางที่ 4.7 นำผลการประเมินประสิทธิภาพของการวิเคราะห์ ปัจจัยด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย แสดงให้เห็นว่า ของผลการทดลอง All Feature ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation เท่ากับข้อมูลสัดส่วนร้อยละ 80:20 ค่าความถูกต้อง 99.43% ค่าความแม่นยำ 99.28% ค่าความระลึก 99.39% ค่าถ่วงดุล 99.33% และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุด 0.075 ดังแสดงในภาพที่ 4.7



**ภาพที่ 4.7** แผนภูมิเปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง All Feature ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation



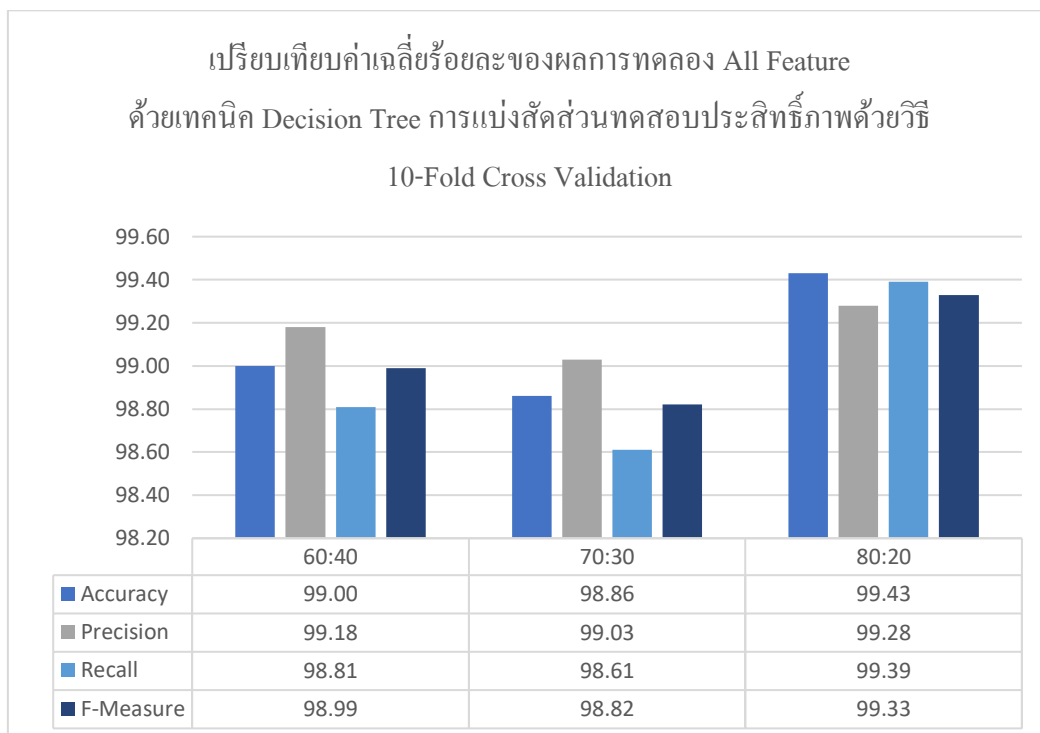
ภาพที่ 4.8 แผนภูมิเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยผลการทดลอง All Feature การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จากภาพที่ 4.8 แสดงผลการเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย เมื่อทำการเรียนรู้และทำการทดสอบ แบ่งข้อมูลสัดส่วนร้อยละ 80:20 ได้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.075

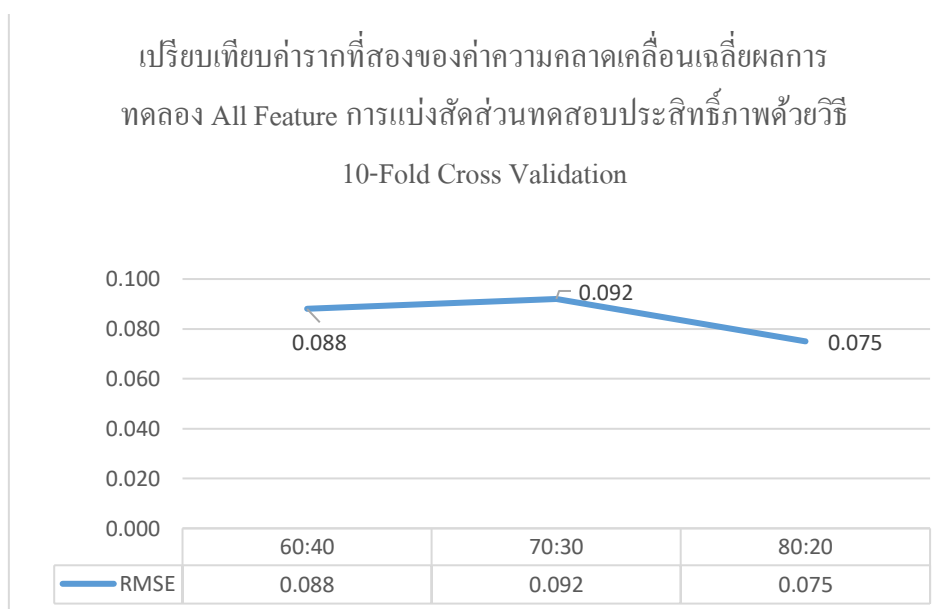
ตารางที่ 4.8 เปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง All Feature ด้วยเทคนิค Decision Tree แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

เปอร์เซ็นต์	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
60:40	99.00	99.18	98.81	98.99	0.088
70:30	98.86	99.03	98.61	98.82	0.092
80:20	99.43	99.28	99.39	99.33	0.075

จากตารางที่ 4.8 นำผลการประเมินประสิทธิภาพของการวิเคราะห์ปัจจัยด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย แสดงให้เห็นว่า ของผลการทดลอง All Feature ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation ที่เหมาะสมที่สุดเท่ากับการแบ่งข้อมูลสัดส่วนร้อยละ 80:20 ค่าความถูกต้อง 99.43% ค่าความแม่นยำ 99.28% ค่าความระลึก 99.39% ค่าถ่วงดุล 99.33% และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุด 0.075 ดังแสดงในภาพที่ 4.9



ภาพที่ 4.9 แผนภูมิเปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง All Feature  
ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี  
10-Fold Cross Validation



ภาพที่ 4.10 แผนภูมิเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย All Feature กับข้อมูล  
สัดส่วนร้อยละและแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.10 แสดงผลการเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย เมื่อทำการเรียนรู้และทำการทดสอบ ด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 80:20 ได้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.075

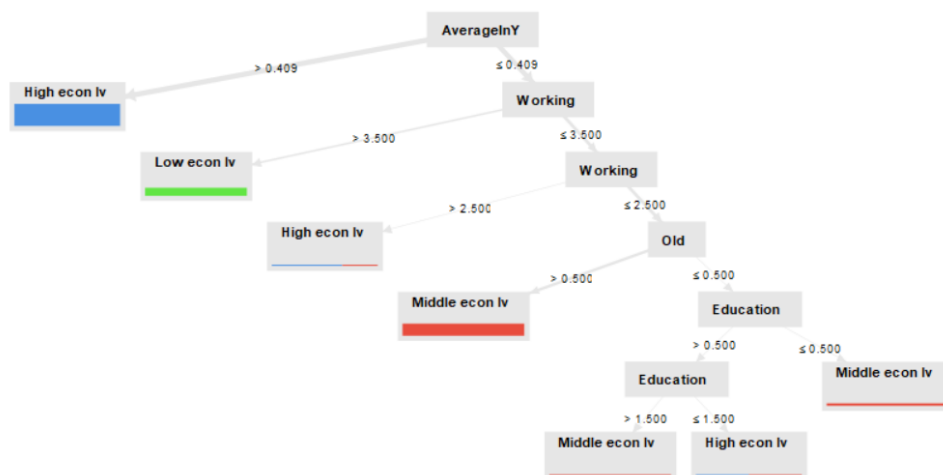
## 4.2 ผลของการสร้างตัวแบบ Decision Tree กับข้อมูลจาก Gain Ratio

### 4.2.1 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40

การสร้างตัวแบบกับชุดข้อมูล Gain Ratio ด้วยเทคนิค Decision Tree โดยจะมีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิคด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 60:40 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และ 10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.9 และตารางที่ 4.10

**ตารางที่ 4.9** ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.62	99.63	99.71	99.67	0.062
2	99.62	99.63	99.71	99.67	0.062
3	99.62	99.63	99.71	99.67	0.062
4	99.62	99.63	99.71	99.67	0.062
5	99.62	99.63	99.71	99.67	0.062
6	99.62	99.63	99.71	99.67	0.062
7	99.62	99.63	99.71	99.67	0.062
8	99.62	99.63	99.71	99.67	0.062
9	99.62	99.63	99.71	99.67	0.062
10	99.62	99.63	99.71	99.67	0.062
เฉลี่ย	99.62	99.63	99.71	99.67	0.062



ภาพที่ 4.11 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

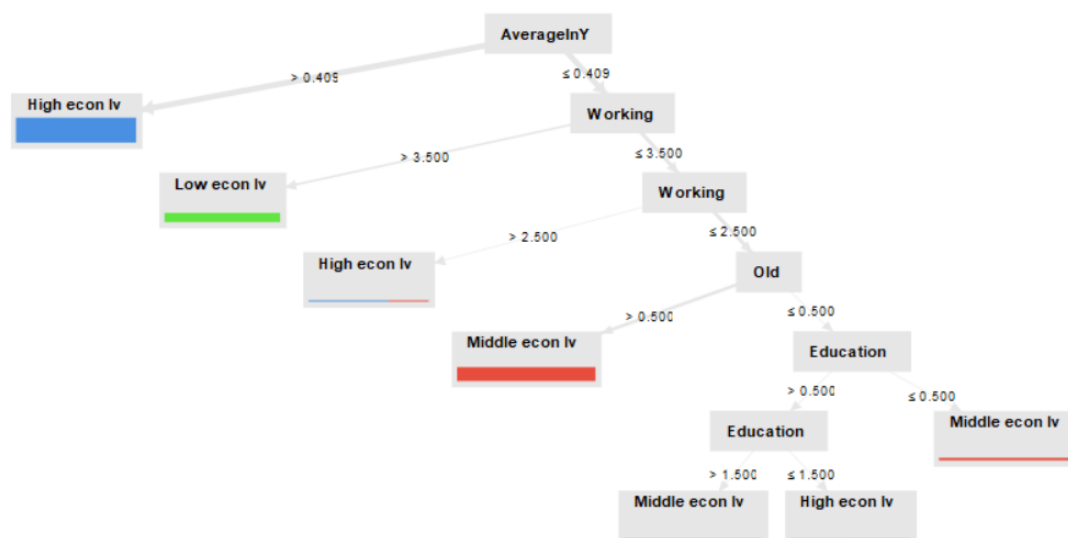
จากภาพที่ 4.11 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 7 กฎ ดังนี้

- 1) IF AverageInY > 0.409 THEN High income
- 2) IF AverageInY ≤ 0.409 AND Working ≥ 3.500 THEN Low econ lv
- 3) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working > 2.500 THEN High econ lv
- 4) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 AND Old > 0.500 THEN Middle econ lv
- 5) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 AND Old ≤ 0.500 AND Education > 0.500 AND > 1.500 THEN Middle econ lv
- 6) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 AND Old ≤ 0.500 AND Education > 0.500 AND ≤ 1.500 THEN High econ lv
- 7) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND Working ≤ 2.500 AND Old ≤ 0.500 AND Education ≤ 0.500 THEN Middle econ lv



ตารางที่ 4.10 ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.52	99.45	99.61	99.53	0.069
2	99.52	99.45	99.61	99.53	0.069
3	99.52	99.45	99.61	99.53	0.069
4	99.52	99.45	99.61	99.53	0.069
5	99.52	99.45	99.61	99.53	0.069
6	99.52	99.45	99.61	99.53	0.069
7	99.52	99.45	99.61	99.53	0.069
8	99.52	99.45	99.61	99.53	0.069
9	99.52	99.45	99.61	99.53	0.069
10	99.52	99.45	99.61	99.53	0.069
เฉลี่ย	99.52	99.45	99.61	99.53	0.069



ภาพที่ 4.12 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 60:40 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.12 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะ  
เริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 7 กฎ ดังนี้

1) IF AverageInY  $>0.409$  THEN High income

2) IF AverageInY  $\leq 0.409$  AND Working  $\geq 3.500$  THEN  
Low econ lv

3) IF AverageInY  $\leq 0.409$  AND Working  $\leq 3.500$  AND  
Working  $>2.500$  THEN High econ lv

4) IF AverageInY  $\leq 0.409$  AND Working  $\leq 3.500$  AND  
Working  $\leq 2.500$  AND Old  $>0.500$  THEN Middle econ lv

5) IF AverageInY  $\leq 0.409$  AND Working  $\leq 3.500$  AND  
Working  $\leq 2.500$  AND Old  $\leq 0.500$  AND Education  $>0.500$  AND  $>1.500$  THEN Middle econ  
lv

6) IF AverageInY  $\leq 0.409$  AND Working  $\leq 3.500$  AND  
Working  $\leq 2.500$  AND Old  $\leq 0.500$  AND Education  $>0.500$  AND  $\leq 1.500$  THEN High econ  
lv

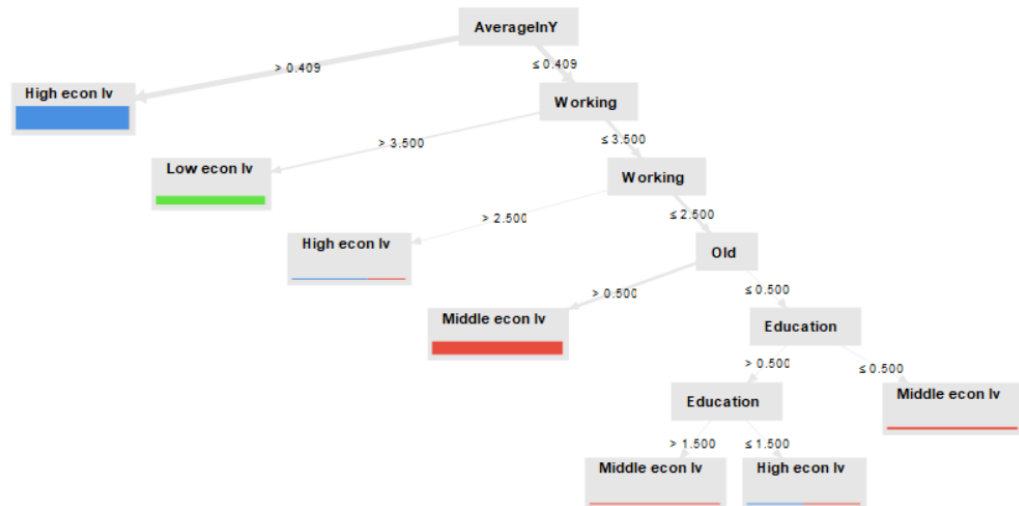
7) IF AverageInY  $\leq 0.409$  AND Working  $\leq 3.500$  AND  
Working  $\leq 2.500$  AND Old  $\leq 0.500$  AND Education  $\leq 0.500$  THEN Middle econ lv

#### 4.2.2 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30

การสร้างตัวแบบกับชุดข้อมูล Gain Ratio ด้วยเทคนิค Decision Tree โดยจะมีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิคด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 70:30 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และ 10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.11 และตารางที่ 4.12

**ตารางที่ 4.11** ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.43	99.53	99.49	99.51	0.067
2	99.43	99.53	99.49	99.51	0.067
3	99.43	99.53	99.49	99.51	0.067
4	99.43	99.53	99.49	99.51	0.067
5	99.43	99.53	99.49	99.51	0.067
6	99.43	99.53	99.49	99.51	0.067
7	99.43	99.53	99.49	99.51	0.067
8	99.43	99.53	99.49	99.51	0.067
9	99.43	99.53	99.49	99.51	0.067
10	99.43	99.53	99.49	99.51	0.067
<b>เฉลี่ย</b>	99.43	99.53	99.49	99.51	0.067



ภาพที่ 4.13 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จากภาพที่ 4.13 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะเริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 7 กฎ ดังนี้

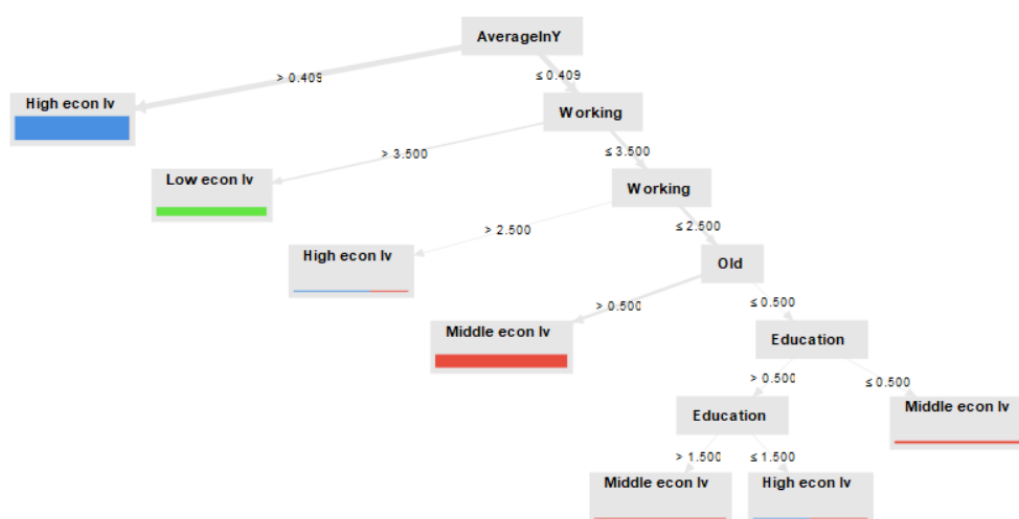
- 1) IF AverageInY > 0.409 THEN High income
- 2) IF AverageInY ≤ 0.409 AND Working ≥ 3.500 THEN  
Low econ lv
- 3) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND  
Working > 2.500 THEN High econ lv
- 4) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND  
Working ≤ 2.500 AND Old > 0.500 THEN Middle econ lv
- 5) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND  
Working ≤ 2.500 AND Old ≤ 0.500 AND Education > 0.500 AND > 1.500 THEN Middle econ  
lv
- 6) IF AverageInY ≤ 0.409 AND Working ≤ 3.500 AND  
Working ≤ 2.500 AND Old ≤ 0.500 AND Education > 0.500 AND ≤ 1.500 THEN High econ  
lv

7) IF  $\text{AverageInY} \leq 0.409$  AND  $\text{Working} \leq 3.500$  AND

$\text{Working} \leq 2.500$  AND  $\text{Old} \leq 0.500$  AND  $\text{Education} \leq 0.500$  THEN Middle econ lv

ตารางที่ 4.12 ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.51	99.61	99.55	99.58	0.062
2	99.51	99.61	99.55	99.58	0.062
3	99.51	99.61	99.55	99.58	0.062
4	99.51	99.61	99.55	99.58	0.062
5	99.51	99.61	99.55	99.58	0.062
6	99.51	99.61	99.55	99.58	0.062
7	99.51	99.61	99.55	99.58	0.062
8	99.51	99.61	99.55	99.58	0.062
9	99.51	99.61	99.55	99.58	0.062
10	99.51	99.61	99.55	99.58	0.062
เฉลี่ย	99.51	99.61	99.55	99.58	0.062



ภาพที่ 4.14 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 70:30 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.14 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะ  
เริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 7 กฎ ดังนี้

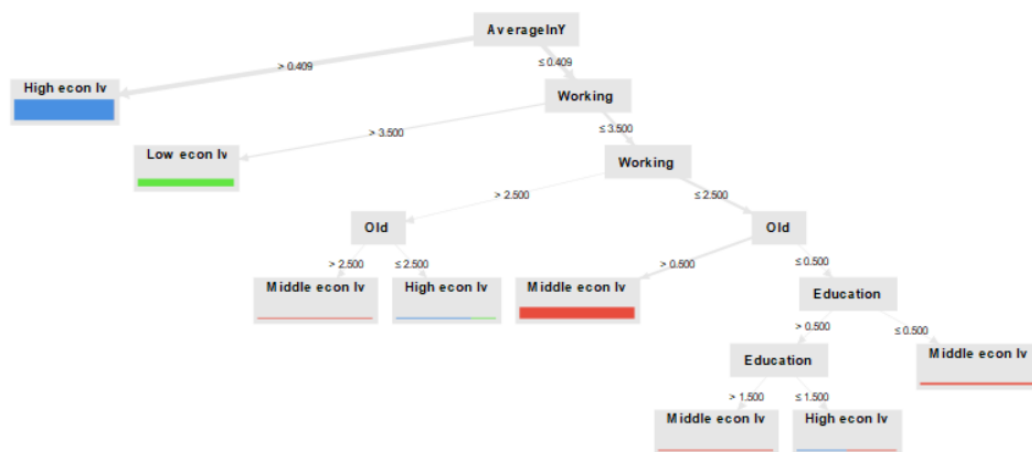
- 1) IF AverageInY >0.409 THEN High income
- 2) IF AverageInY ≤0.409 AND Working ≥3.500 THEN Low  
econ lv
- 3) IF AverageInY ≤0.409 AND Working ≤3.500 AND  
Working >2.500 THEN High econ lv
- 4) IF AverageInY ≤0.409 AND Working ≤3.500 AND  
Working ≤2.500 AND Old >0.500 THEN Middle econ lv
- 5) IF AverageInY ≤0.409 AND Working ≤3.500 AND  
Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND >1.500 THEN Middle econ  
lv
- 6) IF AverageInY ≤0.409 AND Working ≤3.500 AND  
Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND ≤1.500 THEN High econ lv
- 7) IF AverageInY ≤0.409 AND Working ≤3.500 AND  
Working ≤2.500 AND Old ≤0.500 AND Education ≤0.500 THEN Middle econ lv

#### 4.2.3 ผลการทดลองสร้างตัวแบบกับการแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20

การสร้างตัวแบบกับชุดข้อมูล Gain Ratio ด้วยเทคนิค Decision Tree โดยจะ  
มีการทดลองการสร้างตัวแบบจำนวน 10 ครั้ง แล้วหาค่าเฉลี่ยของแต่ละเทคนิคด้วยการแบ่งข้อมูล  
สัดส่วนร้อยละ 80:20 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation และ  
10-Fold Cross Validation ค่า Maximal Depth 10 แสดงผลดังตารางที่ 4.13 และตารางที่ 4.14

ตารางที่ 4.13 ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio กับข้อมูลารแบ่ง 80:20 และ แบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.50	99.62	99.46	99.54	0.068
2	99.50	99.62	99.46	99.54	0.068
3	99.50	99.62	99.46	99.54	0.068
4	99.50	99.62	99.46	99.54	0.068
5	99.50	99.62	99.46	99.54	0.068
6	99.50	99.62	99.46	99.54	0.068
7	99.50	99.62	99.46	99.54	0.068
8	99.50	99.62	99.46	99.54	0.068
9	99.50	99.62	99.46	99.54	0.068
10	99.50	99.62	99.46	99.54	0.068
เฉลี่ย	99.50	99.62	99.46	99.54	0.068



ภาพที่ 4.15 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

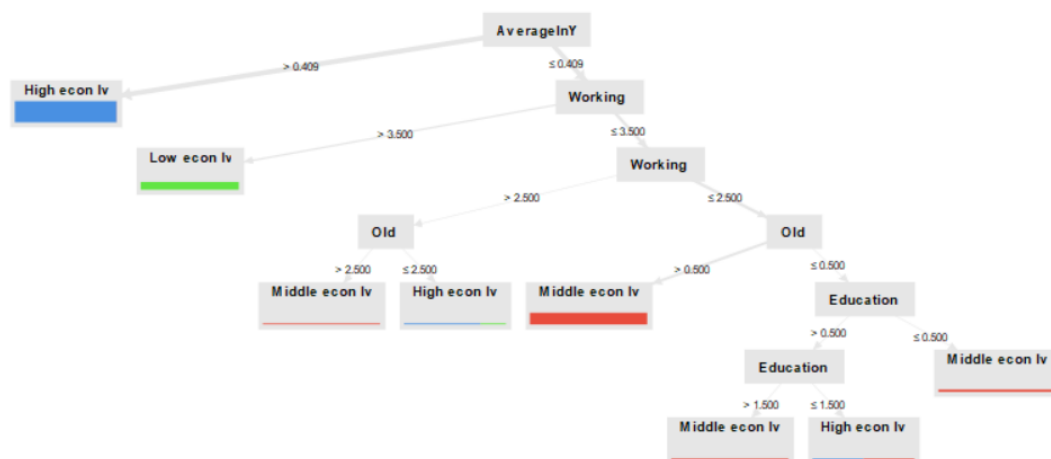
จากภาพที่ 4.15 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะ  
เริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎต้นไม้ได้ 8 กฎ ดังนี้

- 1) IF AverageInY >0.409 THEN High econ lv
- 2) IF AverageInY ≤0.409 AND Working >3.500 THEN Low  
econ lv
- 3) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working >2.500 AND Old >2.500 THEN Middle econ lv
- 4) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working >2.500 AND Old ≤ 2.500 THEN High econ lv
- 5) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working ≤2.500 AND Old >0.500 THEN Middle econ lv
- 6) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND Education  
>1.500 THEN Middle econ lv
- 7) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND Education  
≤1.500 THEN High econ lv
- 8) IF AverageInY ≤0.409 AND Working >3.500 AND  
Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education ≤0.500 THEN Middle  
econ lv



ตารางที่ 4.14 ผลการทดสอบ ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จำนวน (ครั้ง)	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
1	99.50	99.62	99.46	99.54	0.070
2	99.50	99.62	99.46	99.54	0.070
3	99.50	99.62	99.46	99.54	0.070
4	99.50	99.62	99.46	99.54	0.070
5	99.50	99.62	99.46	99.54	0.070
6	99.50	99.62	99.46	99.54	0.070
7	99.50	99.62	99.46	99.54	0.070
8	99.50	99.62	99.46	99.54	0.070
9	99.50	99.62	99.46	99.54	0.070
10	99.50	99.62	99.46	99.54	0.070
เฉลี่ย	99.50	99.62	99.46	99.54	0.070



ภาพที่ 4.16 แสดงแผนภาพ Gain Ratio วิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Decision Tree การแบ่งข้อมูลตามค่าสัดส่วนร้อยละ 80:20 และแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

จากภาพที่ 4.16 เมื่อนำแผนภาพ Decision Tree ที่ได้นี้ไปใช้งานจะ  
เริ่มพิจารณาจากโหนดบนสุด (Root Node) ก่อน สามารถเขียนเป็นกฎได้ 8 กฎ ดังนี้

- 1) IF AverageInY >0.409 THEN High econ lv
- 2) IF AverageInY ≤0.409 AND Working >3.500 THEN Low econ lv
- 3) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working >2.500 AND Old >2.500 THEN Middle econ lv
- 4) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working >2.500 AND Old ≤ 2.500 THEN High econ lv
- 5) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working ≤2.500 AND Old >0.500 THEN Middle econ lv
- 6) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND Education >1.500 THEN Middle econ lv
- 7) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education >0.500 AND Education ≤1.500 THEN High econ lv
- 8) IF AverageInY ≤0.409 AND Working >3.500 AND Working ≤3.500 Working ≤2.500 AND Old ≤0.500 AND Education ≤0.500 THEN Middle econ lv

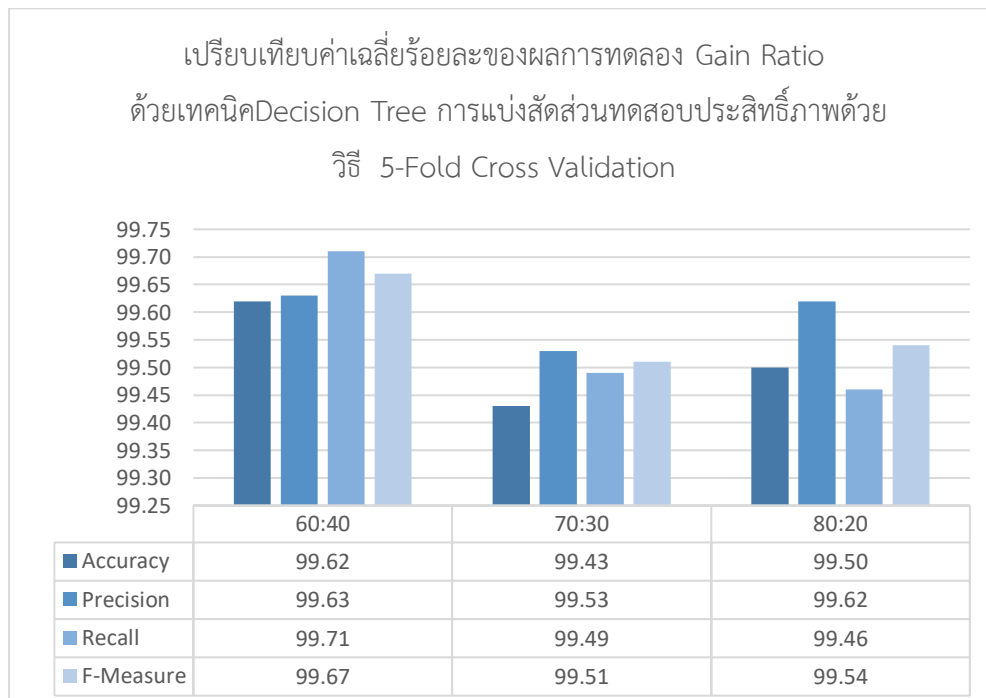
#### 4.2.4 ผลการเปรียบเทียบการสร้างตัวแบบกับชุดข้อมูล Gain Ratio

ผลการเปรียบเทียบเฉลี่ยสุดท้ายของ Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งข้อมูลสัดส่วนร้อยละ 60:40, 70:30, และ 80:20 และ แบ่งสัดส่วนทดสอบประสิทธิภาพ ด้วยวิธี 5-Fold Cross Validation และ 10-Fold Cross Validation จากนั้นประเมินประสิทธิภาพ ด้วยค่าค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย ดังนี้

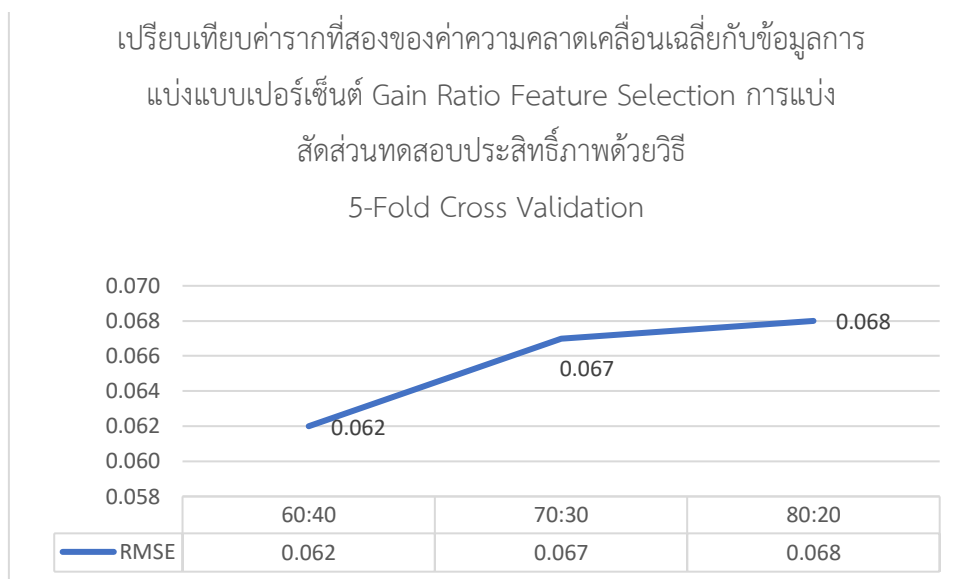
**ตารางที่ 4.15** เปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง ปัจจัยที่สำคัญของเทคนิค Gain Ratio ด้วยเทคนิค Decision Tree และทำการแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

เปอร์เซ็นต์	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
60:40	99.62	99.63	99.71	99.67	0.062
70:30	99.43	99.53	99.49	99.51	0.067
80:20	99.50	99.62	99.46	99.54	0.068

จากตารางที่ 4.15 นำผลการประเมินประสิทธิภาพของการวิเคราะห์ ปัจจัยด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย แสดงให้เห็นว่า ผลการทดลองปัจจัยที่สำคัญของเทคนิค Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation เท่ากับ ข้อมูลสัดส่วนร้อยละ 60:40 ค่าความถูกต้อง 99.62% ค่าความแม่นยำเท่ากับ 99.69% ค่าความระลึกเท่ากับ 99.71% ค่าถ่วงดุลเท่ากับ 99.67% และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.062



**ภาพที่ 4.17** แผนภูมิเปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation



**ภาพที่ 4.18** แผนภูมิเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกับข้อมูลการแบ่งแบบเปอร์เซ็นต์ Gain Ratio Feature Selection การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 5-Fold Cross Validation

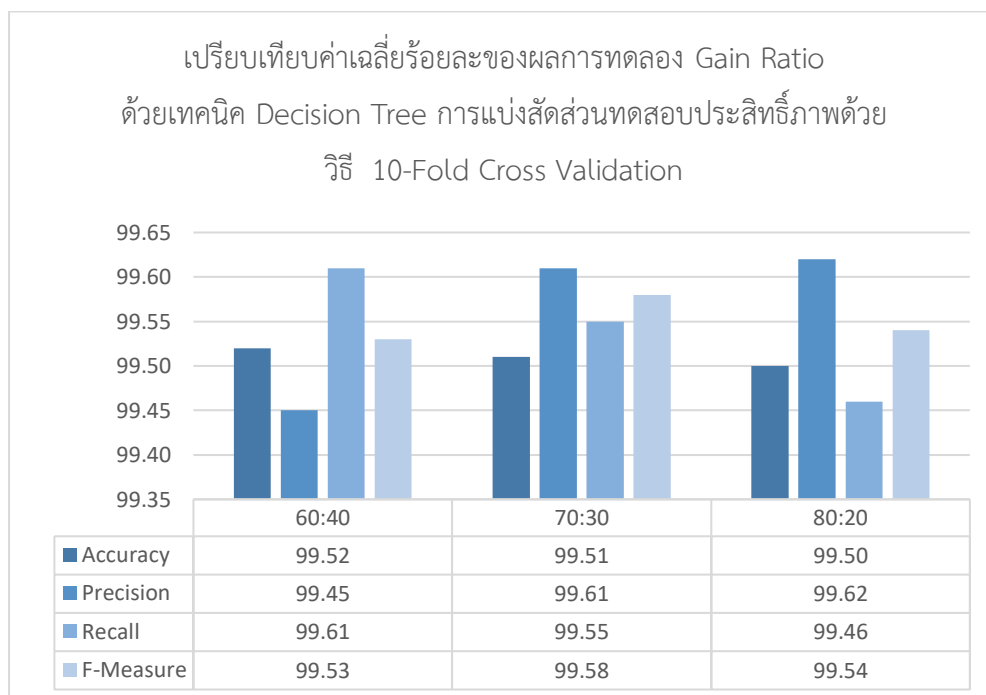
จากภาพที่ 4.18 แสดงผลการเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย เมื่อทำการเรียนรู้และทำการทดสอบ กับข้อมูลสัดส่วนร้อยละ 60:40 ได้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.062

**ตารางที่ 4.16** เปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง ปัจจัยที่สำคัญของเทคนิค Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

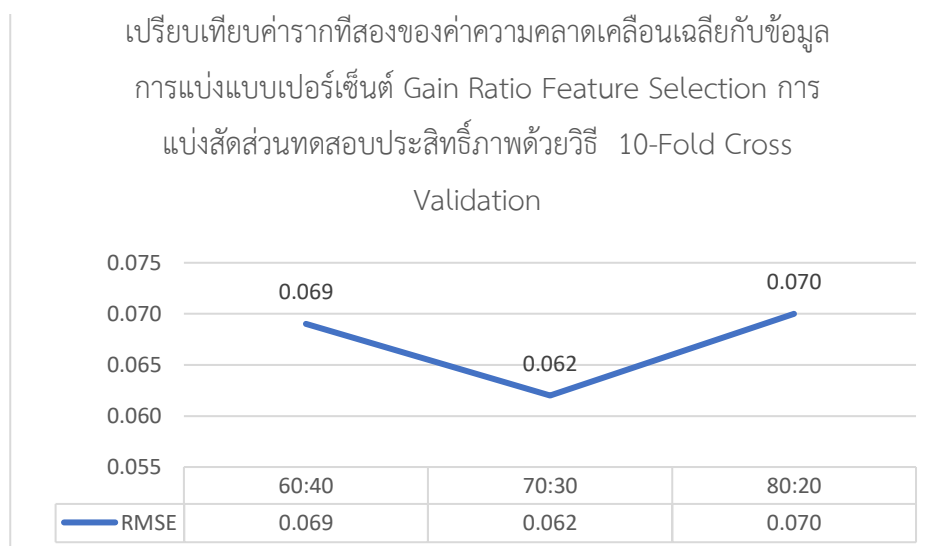
เปอร์เซ็นต์	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
60:40	99.52	99.45	99.61	99.53	0.069
70:30	99.51	99.61	99.55	99.58	0.062
80:20	99.50	99.62	99.46	99.54	0.070

จากตารางที่ 4.16 นำผลการประเมินประสิทธิภาพของการพยากรณ์ด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุล และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย แสดงให้เห็นว่า ผลการทดลองปัจจัยที่สำคัญของเทคนิค Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation เท่ากับข้อมูลสัดส่วนร้อยละ 70:30 ค่าความถูกต้อง 99.51% ค่าความแม่นยำเท่ากับ 99.61% ค่าความระลึก

เท่ากับ 99.55% ค่าถ่วงดุลเท่ากับ 99.58% และค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.062 ดังแสดงในภาพที่ 4.19



ภาพที่ 4.19 แผนภูมิเปรียบเทียบค่าเฉลี่ยร้อยละของผลการทดลอง Gain Ratio ด้วยเทคนิค Decision Tree การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation



ภาพที่ 4.20 แผนภูมิเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกับข้อมูลการแบ่งแบบเปอร์เซ็นต์ Gain Ratio Feature Selection การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation

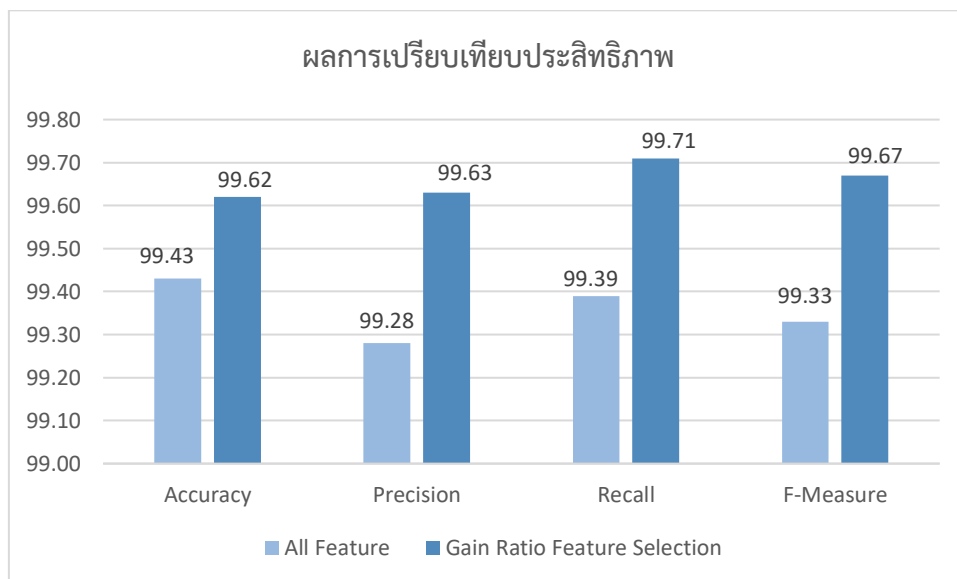
จากภาพที่ 4.20 แสดงผลการเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกับข้อมูล Gain Ratio การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation พบว่าผลของการแบ่งข้อมูลสัดส่วนร้อยละ 70:30 ได้ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุดเท่ากับ 0.062

#### 4.3 ผลการเปรียบเทียบประสิทธิภาพการสร้างตัวแบบ

ตารางที่ 4.17 ผลการเปรียบเทียบค่าประเมินประสิทธิภาพ

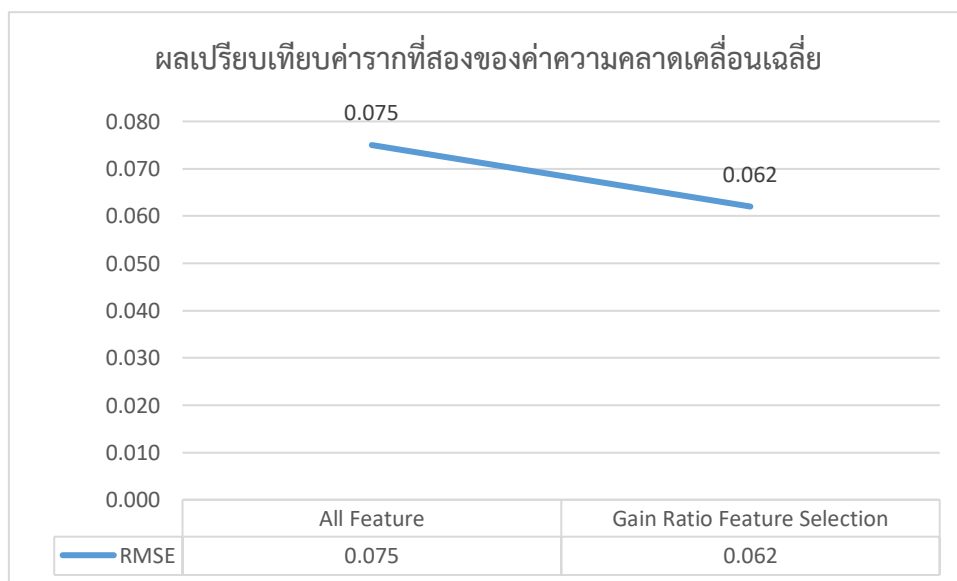
	ประสิทธิภาพ				
	Accuracy	Precision	Recall	F-Measure	RMSE
All Feature	99.43	99.28	99.39	99.33	0.075
Gain Ratio Feature Selection	99.51	99.61	99.55	99.58	0.062

จากการทดลองในงานวิจัยนี้ทำการเปรียบเทียบในการวิเคราะห์ปัจจัยที่สำคัญของข้อมูลเศรษฐกิจครัวเรือน โดยมีการแบ่งข้อมูลสัดส่วนร้อยละ 60:40, 70:30, 80:20 โดย All Feature, Gain Ratio ผลการเปรียบเทียบในการวิเคราะห์ปัจจัยที่สำคัญ พบว่าผลลัพธ์ที่ได้จากวิธีการวิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio ด้วยการแบ่งข้อมูลสัดส่วนร้อยละ 70:30 การแบ่งสัดส่วนทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation ได้ให้ค่าประสิทธิภาพความถูกต้องที่เหมาะสมที่สุดเท่ากับ 99.51% ผลการทดสอบประสิทธิภาพโดยใช้ค่าความแม่นยำอยู่ที่ 99.61% ค่าความระลึกอยู่ที่ 99.55% และค่าถ่วงดุลอยู่ที่ 99.58% พบว่าผลลัพธ์ที่ได้จากวิธีการวิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio ให้ประสิทธิภาพในการวิเคราะห์ปัจจัยดีกว่า ดังแสดงในภาพที่ 4.21



ภาพที่ 4.21 แผนภูมิเปรียบเทียบประสิทธิภาพ

นอกจากนี้ เมื่อทำการเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย พบว่า เทคนิค Gain Ratio ให้ค่าความคลาดเคลื่อนน้อยที่สุดเท่ากับ 0.062 ดังแสดงในภาพที่ 4.22



ภาพที่ 4.22 แผนภูมิเปรียบเทียบค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย

ดังนั้นการวิเคราะห์ปัจจัยสำคัญที่เหมาะสมที่สุดของข้อมูลเศรษฐกิจครัวเรือน คือ การวิเคราะห์ปัจจัยที่สำคัญด้วยเทคนิค Gain Ratio เนื่องจากให้ประสิทธิภาพค่าความถูกต้อง และค่าความแม่นยำ ค่าความระลึก ค่าถ่วงดุลที่เหมาะสมที่สุด และค่าความคลาดเคลื่อนน้อยที่สุด