บทที่ 4

ผลการทดลอง

ในบทนี้ จะเป็นการกล่าวถึงผลการทดลองที่ได้จากการจำแนก รวมถึงค่าตัวแปรต่างๆ ที่ใช้ในการจำแนก และค่าพารามิเตอร์ที่ใช้ในการวิเคราะห์ผลการจำแนก โดยแบ่งผลที่ได้จากการทดลองออกเป็น ค่าพารามิเตอร์ของข้อมูลตัวอย่างที่ใช้ในการจำแนกแต่ละวิธี และผลของการจำแนกภาพดาวเทียมทั้งผลความถูกต้องที่ได้จากการจำแนกข้อมูลทดสอบ และตัวอย่างภาพที่ผ่านการจำแนกในพื้นที่ศึกษา

4.1 ค่าพารามิเตอร์ของข้อมูลตัวอย่าง

4.1.1 ค่า Spectral Signature ของข้อมูลตัวอย่าง

ค่า Signature เป็นค่าทางสถิติที่คำนวณขึ้นจากข้อมูลตัวอย่างสำหรับการกำหนดลักษณะของข้อมูลการจำแนกแต่ละประเภท และใช้สำหรับการคำนวณความน่าจะเป็นของข้อมูลจุดภาพในแต่ละชนิดการจำแนก จาก Probability Density Function ของเส้นโค้งปรกติ โดยข้อมูลภาพดาวเทียม Landsat 5 TM ที่ใช้ในงานวิจัยครั้งนี้ได้ทำการเลือกใช้ข้อมูลจาก 5 ช่วงคลื่น ได้แก่ ช่วงคลื่นที่ 2, 3, 4, 5 และ 7 ซึ่งในตารางที่ 4.1 – ตารางที่4.3 เป็นการแสดงค่า Signature ของข้อมูลตัวอย่างทั้งหมดที่ใช้ในงานวิจัยครั้งนี้

ชนิดการจำแนก	ค่าสูงสุด					ค่าต่ำสุด				
	Band2	Band3	Band4	Band5	Band7	Band2	Band3	Band4	Band5	Band7
ทะเล	30	25	14	14	10	23	18	9	6	2
หาดทราย, โคลน,ตะกอน	54	67	32	23	13	31	24	10	8	2
นากุ้ง,นาเกลือ	44	54	54	78	38	25	25	21	17	6
Swamp มีพืชปกคลุม	48	66	74	71	32	26	25	33	29	13
ป่าชายเลน	38	39	106	54	23	24	19	54	20	3
ป่าไม้ยืนต้น, สวนผสม	39	46	99	115	57	25	21	57	48	14
สวนยางพารา	36	36	93	89	35	25	21	21	18	6
นาข้าว, ทุ่งหญ้า	42	46	135	104	45	26	21	69	31	10
ชุมชน สิ่งปลูกสร้าง	59	76	125	123	69	30	31	49	65	26
พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	85	124	106	187	126	33	34	57	84	35

ตารางที่ 4.1 ค่าสูงสุด และค่าต่ำสุดของข้อมูลตัวอย่าง

ชนิดการจำแนก	ค่าเฉลี่ย					ค่ามัธยฐาน				
	Band2	Band3	Band4	Band5	Band7	Band2	Band3	Band4	Band5	Band7
ทะเล	26.8	21.0	11.5	9.4	5.4	27.0	21.0	11.0	9.0	5.0
หาดทราย, โคลน,ตะกอน	39.9	39.0	18.1	14.3	7.6	38.0	35.0	16.0	14.0	7.0
นากุ้ง,นาเกลือ	31.8	33.4	37.1	39.7	17.9	31.0	30.0	38.0	36.0	16.0
Swamp มีพืชปกคลุม	33.0	34.5	59.2	50.8	21.4	31.0	30.0	61.0	50.5	20.0
ป่าชายเลน	27.3	23.4	76.6	34.1	10.5	28.0	23.0	78.0	29.0	10.0
ป่าไม้ยืนต้น, สวนผสม	31.0	29.3	81.4	66.1	24.4	30.0	26.0	82.0	62.0	20.0
สวนยางพารา	27.9	25.0	75.5	72.8	24.8	28.0	24.0	81.0	79.0	26.0
นาข้าว, ทุ่งหญ้า	33.6	30.5	96.6	69.7	25.5	34.0	30.0	95.0	73.0	26.0
ชุมชน สิ่งปลูกสร้าง	42.2	50.5	71.7	94.5	48.3	42.0	50.0	63.0	96.0	47.0
พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	52.4	71.1	78.0	126.9	71.4	52.5	69.0	76.0	124.5	72.5

ตารางที่ 4.2 ค่าเฉลี่ย และค่ามัธยฐานของข้อมูลตัวอย่าง

ชนิดการจำแนก	ค่าความเบี่ยงเบนมาตรฐาน				
	Band2	Band3	Band4	Band5	Band7
ทะเล	5.4	88.0	27.0	21.0	11.0
หาดทราย, โคลน,ตะกอน	7.6	97.0	38.0	35.0	16.0
นากุ้ง,นาเกลือ	17.9	81.0	31.0	30.0	38.0
Swamp มีพืชปกคลุม	21.4	85.0	31.0	30.0	61.0
ป่าชายเลน	10.5	75.0	28.0	23.0	78.0
ป่าไม้ยืนต้น, สวนผสม	24.4	78.0	30.0	26.0	82.0
สวนยางพารา	24.8	75.0	28.0	24.0	81.0
นาข้าว, ทุ่งหญ้า	25.5	83.0	34.0	30.0	95.0
ชุมชน สิ่งปลูกสร้าง	48.3	92.0	42.0	50.0	63.0
พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	71.4	111.0	52.5	69.0	76.0

ตารางที่ 4.3 ค่าความเบี่ยงเบนมาตรฐานของข้อมูลตัวอย่าง

4.1.2 ค่า Skewness

เป็นค่าพารามิเตอร์ที่ใช้ในการวัดความสมมาตรของลักษณะการกระจายตัวของข้อมูล ซึ่งสามารถใช้เป็นตัวชี้วัดลักษณะการกระจายตัวของข้อมูลที่เป็นเส้นโค้งปรกติ โดยสามารถคำนวณค่า skewness ได้จากสมการที่ 4-1 ซึ่งจะใช้ในการแสดงถึงลักษณะการกระจายตัวของข้อมูล โดยหากข้อมูลมีลักษณะการกระจายตัวเป็นเส้นโค้งปรกติรูประฆังคว่ำ เส้นโค้งการกระจายตัวจะ

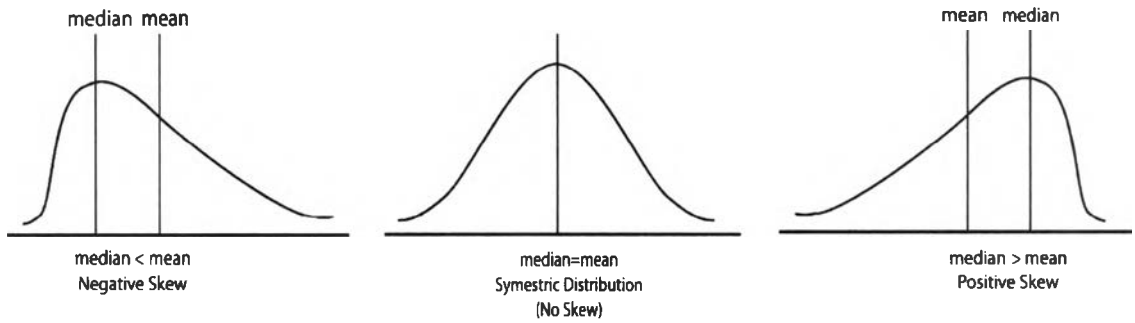
มีลักษณะสมมาตร (skewness =0) ซึ่งจะมีค่าเฉลี่ย เท่ากับค่ามัธยฐาน แต่หากลักษณะการกระจายตัวของข้อมูลไม่เป็นเส้นโค้งปรกติ ลักษณะของโค้งที่ได้จะไม่มีสมมาตร (skewness $\neq$ 0) ซึ่งอาจจะมีลักษณะการเบ้ของกราฟการกระจายตัวของข้อมูลไปทางซ้าย (median < mean) หรือทางขวา (median > mean) ดังแสดงในรูปที่ 4.1 และสำหรับค่า skewness ของข้อมูลตัวอย่างที่ใช้ในงานวิจัยครั้งนี้แสดงในตารางที่ 4.4

$$\text{Skewness} = \frac{E(X-\mu)^3}{\sigma^3} \qquad (4\text{-}1)$$

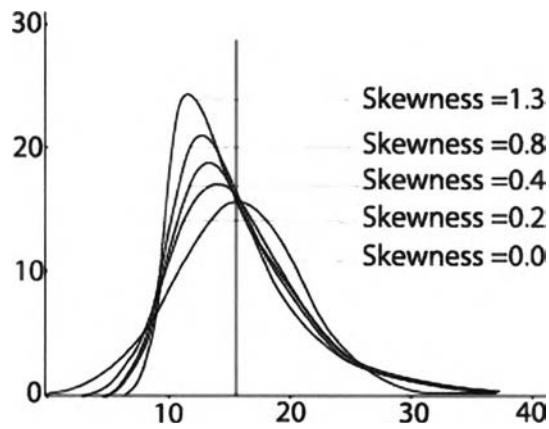
เมื่อ E คือ การหาค่าคาดหวัง หรือค่าเฉลี่ย

μ คือ ค่าเฉลี่ยของข้อมูล

σ คือ ค่าความเบี่ยงเบนมาตรฐานของข้อมูล



รูปที่ 4.1 แสดงความสัมพันธ์ระหว่างค่า Skewness และลักษณะของเส้นโค้งการกระจายตัวของข้อมูล (*NIST/SEMATECH e-Handbook of Statistical Methods*, 2001)



รูปที่ 4.2 แสดงการเปรียบเทียบลักษณะของเส้นโค้งการกระจายตัวของข้อมูลตัวอย่างที่มีค่า Skewness ที่แตกต่างกัน (*NIST/SEMATECH e-Handbook of Statistical Methods*, 2001)

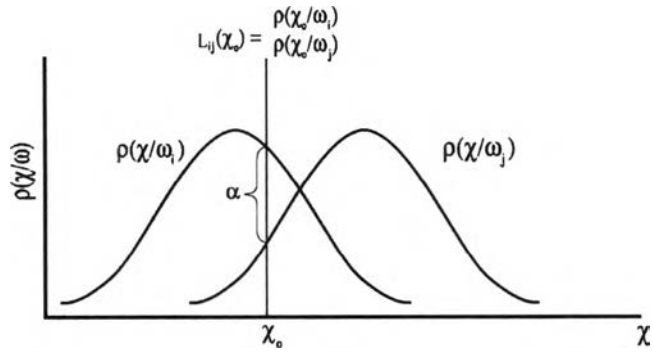
Skewness	Band2	Band3	Band4	Band5	Band7	ค่าเฉลี่ย
ทะเล	-0.133	0.126	0.618	0.224	0.734	0.367
หาดทราย, โคลน,ตะกอน	0.548	0.875	0.504	0.542	0.300	0.553
นากุ้ง,นาเกลือ	0.616	1.091	-0.089	1.087	1.059	0.789
Swamp มีพืชปกคลุม	1.271	1.689	-0.259	-0.329	0.429	0.796
ป่าชายเลน	1.676	2.637	-0.190	0.758	1.007	1.253
ป่าไม้ยืนต้น, สวนผสม	0.549	0.831	-0.449	1.542	1.545	0.983
สวนยางพารา	1.005	1.357	-2.107	-2.165	-1.629	1.652
นาข้าว, ทุ่งหญ้า	-0.226	0.366	0.528	-0.648	-0.149	0.383
ชุมชน สิ่งปลูกสร้าง	0.336	0.004	1.539	-0.195	0.097	0.434
พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	0.725	0.367	0.711	0.591	0.505	0.580

ตารางที่ 4.4 ค่า Skewness ของข้อมูลตัวอย่าง

จากค่า Skewness ของข้อมูลในแต่ละช่วงคลื่นของทุกประเภทการจำแนก จะสังเกตเห็นว่า ลักษณะการกระจายตัวของข้อมูลส่วนใหญ่ไม่เป็นเส้นโค้งปรกติ โดยสังเกตได้จากค่า Skewness ซึ่งไม่มีข้อมูลของการจำแนกประเภทใดมีค่าเท่ากับศูนย์ และเนื่องจากข้อมูลตัวอย่างที่ใช้ในงานวิจัยครั้งนี้เป็นข้อมูลที่ได้มาจากภาพดาวเทียม โดยการสำรวจข้อเท็จจริงในสนาม จึงแสดงให้เห็นว่า ข้อมูลการจำแนกซึ่งเป็นข้อมูลดาวเทียมส่วนใหญ่ มีความเป็นไปได้ยากที่จะมีลักษณะการกระจายตัวที่เป็นเส้นโค้งปรกติ

4.1.3 ค่า Separability

ค่า Separability เป็นค่าที่ใช้วิเคราะห์ความสามารถในการจำแนกจากกันระหว่างข้อมูลชนิดการจำแนกแต่ละชนิด โดยทำการคำนวณจากข้อมูลตัวอย่าง ทำให้ทราบว่าชนิดการจำแนกแต่ละชนิดสามารถแยกออกจากกันโดยใช้ข้อมูลค่าความสว่างของจุดภาพได้มากน้อยเพียงใด ซึ่งวิธีที่ใช้ในการคำนวณค่า Separability มีอยู่หลายวิธี สำหรับงานวิจัยครั้งนี้เลือกใช้วิธี Divergence (Swain, 1978) ในการวัดค่า Separability



รูปที่ 4.3 แสดงการกำหนดค่า Separability ของการจำแนก

ค่า Divergence เป็นค่าที่ได้จากการคำนวณสัดส่วนความคล้ายคลึง (L_{ij}) ตามสมการที่ 4-2 ซึ่งเป็นการคำนวณค่าสัดส่วนความน่าจะเป็นระหว่างชนิดการจำแนกชนิด j กับชนิด i บนสมมุติฐานของลักษณะการกระจายตัวของข้อมูลที่เป็นเส้นโค้งปรกติ

$$L_{ij}(X) = \frac{p(X \mid \omega_i)}{p(X \mid \omega_j)} \qquad (4\text{-}2)$$

เมื่อ $p(X \mid \omega)$ เป็น Probability Density Function ของเส้นโค้งปรกติ

หากพิจารณาเส้นโค้งการกระจายตัวระหว่างข้อมูลสองชนิดที่ใช้ในการจำแนก ในรูปที่ 4.3 ที่ตำแหน่ง X_0 พบว่าสามารถบอกความสามารถในการจำแนกข้อมูลที่ตำแหน่งX_0 ได้จากค่าα ซึ่งเป็นค่าผลต่างระหว่างความน่าจะเป็นในการจำแนกข้อมูล จากรูปจะเห็นว่า ที่ตำแหน่ง X_0 มีความน่าจะเป็นสูงกว่าที่ข้อมูลจะถูกจำแนกเป็นชนิด i มากกว่าเป็นชนิด j ซึ่งหากค่า α ยิ่งมีค่ามากแสดงว่าความสามารถในการจำแนกข้อมูลทั้งสองชนิดออกจากกันยิ่งทำได้ดี และหากสังเกตข้อมูลตำแหน่งที่เป็นจุดตัดของเส้นโค้งสองเส้น จะพบว่าค่า α มีค่าเท่ากับศูนย์ หมายถึงความน่าจะเป็นของการจำแนกข้อมูลสองชนิดมีค่าเท่ากัน ทำให้มีความสามารถในการจำแนกข้อมูลต่ำ เช่นเดียวกับการหาค่า Separability โดยวิธี Divergence ซึ่งเป็นการวัดความสามารถในการจำแนกของข้อมูลสองชุดจากค่าสัดส่วนความคล้ายคลึง โดยใช้ฟังก์ชั่นลอการิทึมในสมการที่ 4-2 เพื่อทำค่าสัดส่วนอยู่ในรูปผลต่าง ดังสมการที่ 4-2

$$\begin{aligned} L'_{ij}(X) &= \log_e L_{ij}(X) \\ &= \log_e p(X \mid \omega_i) - \log_e p(X \mid \omega_j) \end{aligned} \qquad (4\text{-}3)$$

โดยการคำนวณหาค่า Divergence ระหว่างข้อมูลการจำแนก สามารถคำนวณได้จาก

$$D_{ij} = E[L_{ij}(X \mid \omega_i)] + E[L_{ji}(X \mid \omega_j)] \quad (4\text{-}4)$$

โดยที่ $E[L_{ij}(X) \mid \omega_i]$ เป็นค่าคาดหวัง หรือค่าเฉลี่ยของ $L_{ij}(X \mid \omega_i)$

และเมื่อแทนค่า $p(X \mid \omega)$ ซึ่งเป็นสมการ Probability Density Function ของเส้นโค้งปรกติในสมการที่ 4-3 และสมการที่ 4-4 จะได้สมการการหาค่า Divergence ดังนี้

$$D_{ij} = \frac{1}{2}\,\mathrm{tr}[(\Sigma_i - \Sigma_j)(\Sigma_j^{-1} - \Sigma_i^{-1})] + \frac{1}{2}\,\mathrm{tr}[(\Sigma_i^{-1} + \Sigma_j^{-1})(U_i - U_j)(U_i - U_j)^T] \quad (4\text{-}5)$$

เมื่อ i และ j เป็นชนิดของข้อมูลการจำแนก ที่นำมาเปรียบเทียบ

Σ_i เป็นเมตริกซ์ค่าสัมประสิทธิ์ความแปรปรวน (Covariance Matrix) ของข้อมูลการจำแนกชนิด i

U_i เป็นเวคเตอร์ค่าเฉลี่ย ของ Signature ชนิดการจำแนก i

tr เป็นฟังก์ชั่นการหาผลรวมของสมาชิกที่อยู่ในแนวเส้นทะแยงมุมของเมตริกซ์

T เป็นฟังก์ชั่นการหา transpose matrix

เนื่องจากผลลัพธ์ที่ได้จากสมการที่ 4-5 เป็นผลลัพธ์ที่อยู่ในรูปของเลขทศนิยมซึ่งมีค่าน้อย Jensen (1996) จึงได้ทำการแปลงค่าผลลัพธ์ของการคำนวณค่า Divergence ให้อยู่ในช่วงระหว่าง 0 ถึง 2000 เพื่อเป็นการง่ายในการตีความหมาย โดยเรียกว่าวิธี Transformed Divergence (TD_{ij}) โดยค่า TD_{ij} ที่มีค่า เท่ากับ 2000 ข้อมูลจะสามารถจำแนกจากกันได้ดี และหาก TD_{ij} มีค่าเท่ากับ 0 หมายถึง ข้อมูลไม่สามารถจำแนกจากกันได้ ซึ่งสามารถอธิบายความหมายของ Separability จากแต่ละช่วงของค่า TD_{ij} ตาม (Jensen, 1996) ซึ่งข้อมูลที่มีค่า TD_{ij} มากกว่า 1900 สามารถจะจำแนกจากกันได้ ข้อมูลที่มีค่า TD_{ij} อยู่ในช่วงระหว่าง 1700 กับ 1900 จะสามารถจำแนกได้ดีพอใช้ และสำหรับข้อมูลที่มีค่า TD_{ij} ต่ำกว่า 1700 เป็นข้อมูลที่ไม่สามารถจำแนกจากกันได้ดี

$$TD_{ij} = 2000\left(1 - \exp\left(\frac{-D_{ij}}{8}\right)\right) \quad (4\text{-}6)$$

สำหรับค่า Transformed Divergence ของข้อมูลตัวอย่างที่ใช้ในงานวิจัยครั้งนี้ แสดงในตารางที่ 4.5

ประเภทการจำแนก	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) ทะเล	0	2000	2000	2000	2000	2000	2000	2000	2000	2000
(2) หาดทราย,โคลน, ตะกอน	2000	0	2000	2000	2000	2000	2000	2000	2000	2000
(3) นากุ้ง/นาเกลือ	2000	2000	0	**1684**	1996	1996	1994	1999	1987	2000
(4) Swamp มีพืชปกคลุม	2000	2000	1684	0	1981	1817	1982	1941	1993	2000
(5) ป่าชายเลน	2000	2000	1996	1981	0	1803	1994	**1683**	2000	2000
(6) ป่าไม้ยืนต้น ,สวนผสม	2000	2000	1996	1817	1803	0	1871	**1338**	**1799**	1999
(7) ยางพารา	2000	2000	1994	1982	1994	1871	0	1985	2000	2000
(8) นาข้าว, ทุ่งหญ้า	2000	2000	1999	1941	1683	1338	1985	0	1975	2000
(9) บ้านเรือนสิ่งปลูกสร้าง	2000	2000	1987	1993	2000	1799	2000	1975	0	**1414**
(10)พื้นที่โล่ง	2000	2000	2000	2000	2000	1999	2000	2000	1414	0

ตารางที่ 4.5 ค่า Separability ของข้อมูลตัวอย่าง

จากค่า Transformed Divergence ในตารางที่ 4.5 แสดงให้เห็นว่า ข้อมูลของชนิดการจำแนกที่มีความหลากหลายของข้อมูลในชนิดเดียวกัน ได้แก่ ชนิดการจำแนกที่เป็น ป่าไม้ยืน-สวนผสม ซึ่งประกอบไปด้วยข้อมูลของพืชหลากหลายชนิด และชนิดการจำแนกที่เป็น บ้านเรือน-สิ่งปลูกสร้าง ซึ่งข้อมูลของแต่ละชนิดการจำแนกไม่มีความเด่นชัด แต่เป็นข้อมูลที่มีการปะปนกันของวัตถุที่มีค่าการสะท้อนใกล้เคียงกัน ทำให้ในบางครั้ง จะให้ระดับค่า Separability ในการจำแนกค่อนข้างต่ำ เช่น Separability ระหว่างชนิดการจำแนกที่เป็นป่าไม้ยืนต้น-สวนผสม และชนิดการจำแนกที่เป็น นาข้าว และค่า Separability ระหว่างชนิดการจำแนกที่เป็น สิ่งปลูกสร้าง-ที่ดินว่างเปล่า

ข้อมูลตัวอย่างแต่ละชนิดมีความสามารถในการจำแนกจากกันได้แตกต่างกัน และจะพบว่า ข้อมูลของชนิดการจำแนกซึ่งประกอบด้วยส่วนประกอบที่มีความหลากหลายซึ่งรวมเป็นข้อมูลชนิดเดียว ได้แก่ ข้อมูลของป่าไม้ยืนต้น สวนผสม และข้อมูลของบ้านเรือน สิ่งปลูกสร้าง จะให้ค่า Separability ค่อนข้างต่ำ

4.2 ผลการจำแนก

ในงานวิจัยครั้งนี้ได้ทำการแบ่งผลการจำแนกออกเป็น 2 ส่วนคือ ผลการจำแนกที่ได้จากการจำแนกข้อมูลทดสอบ เพื่อใช้ในการเคราะห์ความถูกต้องที่ได้ในการจำแนก จากทั้งสองวิธี ผลที่ได้จะแสดงในรูป Confusion Matrix และตัวอย่างภาพการจำแนกของบริเวณพื้นที่ศึกษา ซึ่งได้จากวิธีการจำแนกทั้งสองวิธี

4.2.1 ผลการจำแนกข้อมูลทดสอบ

หลังจากที่ได้ทำการคำนวณ Signature และคำนวณค่าน้ำหนักของโครงข่ายจากข้อมูลตัวอย่างแล้ว จึงทำการจำแนกข้อมูลโดยใช้ข้อมูลทดสอบจากวิธีการจำแนกทั้งสองวิธี โดยผลที่ได้จากการจำแนกโดยวิธีความคล้ายคลึงมากที่สุด และวิธีโครงข่ายใยประสาทเทียมในรูปของ Confusion Matrix แสดงในตารางที่ 4.6 และ ตารางที่ 4.7

ชนิดการจำแนก	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	รวม	ร้อยละความถูกต้อง
(1) ทะเล	70	0	0	0	0	0	0	0	0	0	70	100.0
(2) โคลน หาดโคลน	0	79	2	0	0	0	0	0	0	0	81	97.5
(3) นากุ้ง นาเกลือ	0	0	106	1	0	0	2	0	0	0	109	97.2
(4) พื้นที่ชุ่มน้ำ	0	0	11	47	4	1	11	2	3	0	79	59.5
(5) ป่าชายเลน	0	0	3	4	70	0	0	0	0	0	77	90.9
(6) สวนผสม ป่าไม้ยืนต้น	0	0	0	5	16	50	11	21	27	1	131	38.2
(7) ยางพารา	0	0	0	0	0	3	113	0	1	2	119	95.0
(8) นาข้าว ทุ่งหญ้า	0	0	0	0	0	8	4	101	6	0	119	84.9
(9) ชุมชน สิ่งปลูกสร้าง	0	0	0	0	0	0	0	5	25	36	66	37.9
(10) พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	0	0	8	0	4	5	0	0	39	43	99	43.4

ความถูกต้องรวมสำหรับการจำแนก เท่ากับร้อยละ 74.45

ตารางที่ 4.6 ผลการจำแนกโดยวิธีความคล้ายคลึงมากที่สุด

ชนิดการจำแนก	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	รวม	ร้อยละความถูกต้อง
(1) ทะเล	70	0	0	0	0	0	0	0	0	0	70	100.0
(2) โคลน หาดโคลน	0	80	0	0	1	0	0	0	0	0	81	98.8
(3) นากุ้ง นาเกลือ	0	3	106	0	0	0	0	0	0	0	109	97.2
(4) พื้นที่ชุ่มน้ำ	0	0	0	71	4	1	0	0	3	0	79	89.9
(5) ป่าชายเลน	0	0	0	1	76	0	0	0	0	0	77	98.7
(6) สวนผสม ป่าไม้ยืนต้น	0	0	0	1	16	57	2	16	39	0	131	43.5
(7) ยางพารา	0	0	0	0	0	1	117	0	0	1	119	98.3
(8) นาข้าว ทุ่งหญ้า	0	0	0	0	0	12	1	105	1	0	119	88.2
(9) ชุมชน สิ่งปลูกสร้าง	0	0	0	0	0	0	0	7	28	31	66	42.4
(10) พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย	0	0	9	0	3	4	5	1	5	72	99	72.7

ความถูกต้องรวมสำหรับการจำแนก เท่ากับร้อยละ 82.98

ตารางที่ 4.7 ผลการจำแนกโดยวิธีโครงข่ายใยประสาทเทียม

จากผลการจำแนกในงานวิจัยครั้งนี้ ที่ได้จากวิธีการจำแนกทั้งสองวิธีในตารางที่ 4.6 และตารางที่ 4.7 แสดงให้เห็นว่า การจำแนกภาพดาวเทียมโดยวิธีโครงข่ายใยประสาทเทียม ให้ความถูกต้องในการจำแนกสูงกว่าวิธีความคล้ายคลึงมากที่สุด โดยเฉพาะข้อมูลที่นำมาจำแนกซึ่งมีลักษณะการกระจายตัวที่ไม่เป็นเส้นโค้งปรกติ โดยพิจารณาจากตารางที่ 4.4 พบว่าไม่มีข้อมูลของการจำแนกชนิดใดเลยที่มีค่า Skewness =0 จึงทำให้ผลการจำแนกที่ได้จากวิธีความคล้ายคลึงมากที่สุด ซึ่งเป็นวิธีการจำแนกโดยอาศัยสมมุติฐานการกระจายตัวของข้อมูลที่เป็นเส้นโค้งปรกติมีความผิดพลาดสูงกว่าวิธีโครงข่ายใยประสาทเทียม ซึ่งเป็นวิธีการจำแนกที่ไม่อาศัยสมมุติฐานการกระจายตัวของข้อมูลในการจำแนก โดยจากผลการจำแนก สามารถแบ่งชนิดการจำแนกออกเป็น 2 กลุ่ม ตามผลการจำแนกคือ กลุ่มที่ให้ผลการจำแนกที่มีความถูกต้องสูงและความถูกต้องที่ได้จากวิธีการจำแนกทั้งสองวิธีมีความสอดคล้องกัน ได้แก่ ทะเล หาดทราย-โคลน นากุ้ง-นาเกลือ ป่าชายเลน ยางพารา นาข้าว-ทุ่งหญ้า และกลุ่มที่ให้ผลความถูกต้องในการจำแนกต่ำ ได้แก่ พื้นที่ชุ่มน้ำ สวนผสม-ป่าไม้ยืนต้น ชุมชน-สิ่งปลูกสร้าง พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย ซึ่งพบว่า

1) ข้อมูลที่เป็นพื้นที่ชุ่มน้ำ ความผิดพลาดส่วนใหญ่ที่เกิดขึ้นในวิธีความคล้ายคลึงมากที่สุดจะถูกจำแนกเป็น นากุ้ง-นาเกลือ และ ยางพารา ในขณะที่ผลการจำแนกข้อมูลที่เป็นพื้นที่ชุ่มน้ำโดยวิธีโครงข่ายใยประสาทเทียมให้ความถูกต้องในการจำแนกสูง และไม่เกิดความผิดพลาดของการจำแนกในลักษณะเดียวกัน

2) ความผิดพลาดในการจำแนกข้อมูลที่เป็นสวนผสม-ป่าไม้ยืนต้น ซึ่งเป็นชนิดการจำแนกที่มีการปะปนของพืชยืนต้นหลายชนิด จากวิธีการจำแนกทั้งสองวิธี มีความคล้ายคลึงกัน คือข้อมูลจะถูกจำแนกเป็น ยางพารา นาข้าว-ทุ่งหญ้า และชุมชน-สิ่งปลูกสร้าง

3) ความผิดพลาดในการจำแนกข้อมูลที่เป็นชุมชน-สิ่งปลูกสร้าง จากวิธีการจำแนกทั้งสองวิธี มีความคล้ายคลึงกัน คือข้อมูลจะถูกจำแนกเป็น นาข้าว-ทุ่งหญ้า และพื้นที่ว่างเปล่ามีพืชปกคลุมน้อย

4) การจำแนกข้อมูลที่เป็นพื้นที่ว่างเปล่ามีพืชปลุกคลุมน้อย โดยวิธีความคล้ายคลึงมากที่สุด จะเกิดความผิดพลาดในการจำแนกสูงเป็น ชุมชน-สิ่งปลูกสร้าง ในขณะที่การจำแนกโดยวิธีโครงข่ายใยประสาทเทียมเกิดความผิดพลาดในการจำแนกน้อยกว่า และความผิดพลาดที่เกิดขึ้น

โดยวิธีโครงข่ายใยประสาทเทียมจะเกิดขึ้นในลักษณะกระจายตัว ไม่ถูกจำแนกผิดพลาดเป็นชนิดการจำแนกชนิดใดชนิดหนึ่ง

5) จะสังเกตเห็นว่า ชนิดการจำแนกที่เป็น นาข้าว ทุ่งหญ้า จะให้ความถูกต้องในการจำแนกโดยวีการจำแนกทั้งสองวิธีอยู่ในระดับที่ใกล้เคียงกัน ซึ่งความผิดพลาดที่เกิดขึ้นจากการจำแนกด้วยวิธีโครงข่ายใยประสาทเทียมจะถูกจำแนกเป็นสวนผสม ป่าไม้ยืนต้น โดยที่ความผิดพลาดจากการจำแนกด้วยวิธีความคล้ายคลึงมากที่สุดจะถูกจำแนกเป็น สวนผสม ป่าไม้ยืนต้น ยางพารา และ ชุมชน สิ่งปลูกสร้าง ซึ่งแสดงให้เห็นว่า วิธีการตัดสินใจในการจำแนกข้อมูลทั้งสองวิธีมีความแตกต่างกัน โดยเฉพาะความผิดพลาดของการจำแนกด้วยวิธีโครงข่ายใยประสาทเทียมที่ไม่สอดคล้องกับความเป็นจริงซึ่งชนิดการจำแนก นาข้าว ทุ่งหญ้า และ สวนผสม ป่าไม้ยืนต้น ควรจะเป็นชนิดการจำแนกที่สามารถจำแนกจากกันได้ดี เนื่องมาจาก ขอบเขตการตัดสินใจการจำแนกในวิธีโครงข่ายใยประสาทเทียมสร้างขึ้นจากชุดข้อมูลตัวอย่าง ในขณะที่วิธีความคล้ายคลึงมากที่สุดอาศัยการคำนวณค่าพารามิเตอร์ทางสถิติเพื่อเป็นตัวแทนของข้อมูล ดังนั้น ความถูกต้องที่ได้จากการจำแนกข้อมูลในแต่ละชนิดด้วยวิธีโครงข่ายใยประสาทเทียม จึงมีความอ่อนไหว (Sensitive) ต่อข้อมูลตัวอย่างมากกว่าวิธีความคล้ายคลึงมากที่สุด

4.2.2 วิเคราะห์เปรียบเทียบผลการจำแนกจากค่า Separability

ในหัวข้อนี้จะเป็นการเปรียบเทียบวิธีการจำแนกโดยอาศัยข้อมูลค่า Separability ของข้อมูลตัวอย่าง และผลของการจำแนกข้อมูลทดสอบ จากวิธีการจำแนกทั้งสองวิธีที่สร้างขึ้นจากข้อมูลตัวอย่าง เพื่อเปรียบเทียบข้อดี ข้อเสีย และลักษณะของข้อมูลการจำแนกที่มีความเหมาะสมในการจำแนกจากวิธีการจำแนกทั้งสองวิธี โดยทำการเปรียบเทียบระหว่างข้อมูลการจำแนกแต่ละประเภท ดังแสดงในตารางที่ 4.8

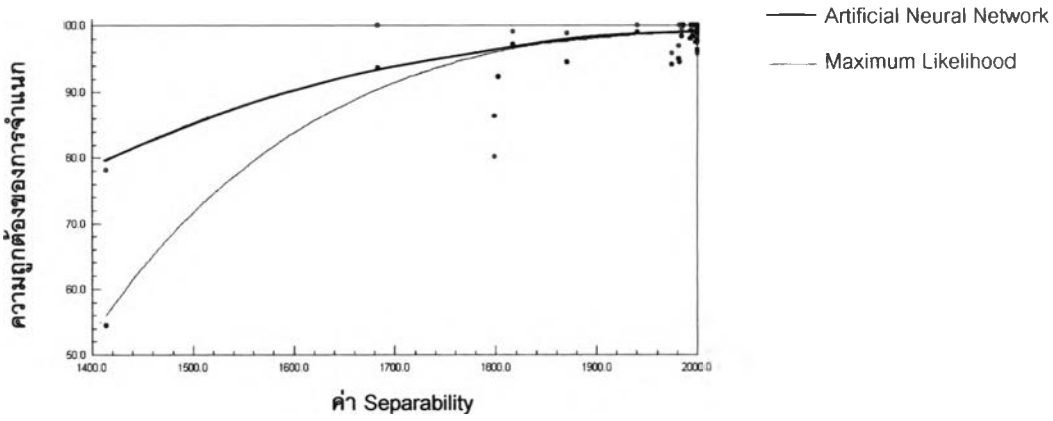
การจำแนกระหว่าง		Separability (TD_{ij})	จำนวน (จุดภาพ)	ความผิดพลาด (จุดภาพ)		ร้อยละของความผิดพลาด	
				Mlh	NN	Mlh	NN
ทะเล	หาดทราย, โคลน,ตะกอน	2000	151	0	0	0	0
	นากุ้ง,นาเกลือ	2000	179	0	0	0	0
	Swamp มีพืชปกคลุม	2000	149	0	0	0	0
	ป่าชายเลน	2000	147	0	0	0	0
	ป่าไม้ยืนต้น, สวนผสม	2000	201	0	0	0	0
	สวนยางพารา	2000	189	0	0	0	0
	นาข้าว, ทุ่งหญ้า	2000	189	0	0	0	0
	ชุมชน สิ่งปลูกสร้าง	2000	136	0	0	0	0
	พื้นที่ว่างเปล่า	2000	169	0	0	0	0
หาดทราย, โคลน,ตะกอน	นากุ้ง,นาเกลือ	1999.9	190	2	3	1.05	1.58
	Swamp มีพืชปกคลุม	2000	160	0	0	0	0
	ป่าชายเลน	2000	158	0	1	0	0.63
	ป่าไม้ยืนต้น, สวนผสม	2000	212	0	0	0	0
	สวนยางพารา	2000	200	0	0	0	0
	นาข้าว, ทุ่งหญ้า	2000	200	0	0	0	0
	ชุมชน สิ่งปลูกสร้าง	2000	147	0	0	0	0
	พื้นที่ว่างเปล่า	2000	180	0	0	0	0
นากุ้ง,นาเกลือ	Swamp มีพืชปกคลุม	1683.83	188	12	0	6.38	0
	ป่าชายเลน	1995.7	186	3	0	1.61	0
	ป่าไม้ยืนต้น, สวนผสม	1995.88	240	0	0	0	0
	สวนยางพารา	1993.83	228	2	0	0.88	0
	นาข้าว, ทุ่งหญ้า	1999.12	228	0	0	0	0
	ชุมชน สิ่งปลูกสร้าง	1986.56	175	0	0	0	0
	พื้นที่ว่างเปล่า	1999.77	208	8	9	3.85	4.33
Swamp มีพืชปกคลุม	ป่าชายเลน	1981.31	156	8	5	5.13	3.21
	ป่าไม้ยืนต้น, สวนผสม	1817.44	210	6	2	2.86	0.95
	สวนยางพารา	1982.41	198	11	0	5.56	0
	นาข้าว, ทุ่งหญ้า	1940.73	198	2	0	1.01	0
	ชุมชน สิ่งปลูกสร้าง	1992.86	145	3	3	2.07	2.07
	พื้นที่ว่างเปล่า	2000	178	1	0	0.56	0

ตารางที่ 4.8 แสดงการเปรียบเทียบผลการจำแนกกับค่า Separability

การจำแนกระหว่าง		Separability (TD_{ij})	จำนวน (จุดภาพ)	ความผิดพลาด (จุดภาพ)		ร้อยละของความผิดพลาด	
				Mlh	NN	Mlh	NN
ป่าชายเลน	ป่าไม้ยืนต้น, สวนผสม	1802.61	208	16	16	7.69	7.69
	สวนยางพารา	1993.7	196	0	0	0	0
	นาข้าว, ทุ่งหญ้า	1682.52	196	0	0	0	0
	ชุมชน สิ่งปลูกสร้าง	2000	143	0	0	0	0
	พื้นที่ว่างเปล่า	2000	176	4	3	2.27	1.7
ป่าไม้ยืนต้น, สวนผสม	สวนยางพารา	1870.58	250	14	3	5.6	1.2
	นาข้าว, ทุ่งหญ้า	1337.81	250	29	28	11.6	11.2
	ชุมชน สิ่งปลูกสร้าง	1798.66	197	27	39	13.71	19.8
	พื้นที่ว่างเปล่า	1998.52	230	6	4	2.61	1.74
สวนยางพารา	นาข้าว, ทุ่งหญ้า	1984.52	238	4	1	1.68	0.42
	ชุมชน สิ่งปลูกสร้าง	2000	185	1	0	0.54	0
	พื้นที่ว่างเปล่า	2000	218	2	6	0.92	2.75
นาข้าว, ทุ่งหญ้า	ชุมชน สิ่งปลูกสร้าง	1975.08	185	11	8	5.95	4.32
	พื้นที่ว่างเปล่า	1999.99	218	0	1	0	0.46
ชุมชน สิ่งปลูกสร้าง	พื้นที่ว่างเปล่า	1414.07	165	75	36	45.45	21.82

ตารางที่ 4.8 (ต่อ) แสดงการเปรียบเทียบผลการจำแนกกับค่า Separability

จากผลการคำนวณค่า Separability และ ค่าร้อยละของความผิดพลาด ที่ได้จากการจำแนกข้อมูลแต่ละชนิด สามารถนำมาสร้างเป็นกราฟแสดงความสัมพันธ์ระหว่างค่า Separability กับค่าความถูกต้องของการจำแนก ดังแสดงในรูปที่ 4.4



รูปที่ 4.4 แสดงความสัมพันธ์ของค่า separability กับค่าความถูกต้องที่ได้จากแต่ละวิธีการจำแนก

จากความสัมพันธ์ที่ได้ในรูปที่ 4.4 เป็นกราฟแสดงความสัมพันธ์ระหว่างค่า Separability และค่าความถูกต้องของการจำแนกข้อมูลแต่ละชนิด ซึ่งได้มาจากวิธีการทำ Regression ซึ่งได้ความสัมพันธ์อยู่ในรูปของสมการ $f(x) = a + b.x + c/x^2$ โดยให้ค่าสัมประสิทธิ์ R^2 ของข้อมูลการจำแนกด้วยวิธีความคล้ายคลึงมากที่สุด เท่ากับ 0.87 และวิธีโครงข่ายใยประสาทเทียม เท่ากับ 0.80 แสดงให้เห็นว่า การจำแนกข้อมูลที่มีค่า Separability สูงด้วยวิธีความคล้ายคลึงมากที่สุดและวิธีโครงข่ายใยประสาทเทียม จะให้ผลความถูกต้องในการจำแนกสูงในระดับใกล้เคียงกัน แต่ในกรณีที่ข้อมูลมีค่า Separability อยู่ในระดับต่ำ วิธีโครงข่ายใยประสาทเทียม จะให้ผลการจำแนกที่มีความถูกต้องสูงกว่าวิธีความคล้ายคลึงมากที่สุด

การเปรียบเทียบผลการจำแนกโดยพิจารณาจากค่า Separability ของแต่ละประเภทการจำแนก ประกอบด้วย ประเภทการจำแนกที่สามารถจำแนกจากกันได้ดี (TD_{ij} >1900) ประเภทการจำแนกที่สามารถจำแนกจากกันได้พอใช้ (1700 < TD_{ij} < 1900) และประเภทการจำแนกที่จำแนกจากกันได้ยาก (TD_{ij} <1700) ซึ่งพบว่า ความสามารถในการจำแนกของวิธีการจำแนกทั้งสองวิธีสำหรับแต่ละชนิดการจำแนกที่มีค่า Separability แตกต่างกัน จะให้ความถูกต้องในการจำแนกที่แตกต่างกัน โดยที่

1) วิธีการจำแนกทั้งสองวิธี สามารถจำแนกข้อมูลที่มีความสามารถในการจำแนกที่ดีมาก (TD_{ij}=2000) ได้ดีโดยวิธีการจำแนกทั้งสองวิธีให้ผลความถูกต้องในการจำแนกที่สูง ในระดับใกล้เคียงกัน

2) การจำแนกข้อมูลที่มีค่า Separability อยู่ในช่วง $1700 < TD_{ij} < 1900$ โดยวิธีโครงข่ายใยประสาทเทียม ส่วนใหญ่จะให้ผลการจำแนกที่มีความถูกต้องในการจำแนกสูงกว่าวิธีความคล้ายคลึงมากที่สุด และมีบางกรณีที่การจำแนกโดยวิธีความคล้ายคลึงมากที่สุดให้ความถูกต้องในการจำแนกเท่ากับ หรือดีกว่า วิธีโครงข่ายใยประสาทเทียม ได้แก่ การจำแนกระหว่างป่าชายเลน กับ ป่าไม้ยืนต้น สวนผสม และ การจำแนกระหว่าง ป่าไม้ยืนต้น สวนผสม กับ ชุมชน สิ่งปลูกสร้าง

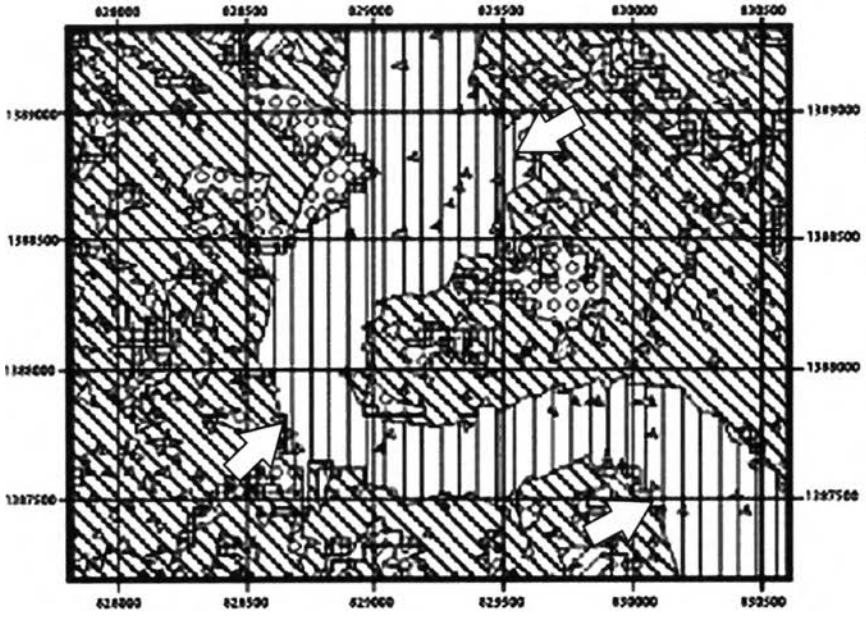
3) วิธีโครงข่ายใยประสาทเทียมสามารถให้ความถูกต้องในการจำแนกข้อมูลที่มีค่า Separability ต่ำ ($TD_{ij} < 1700$) สูงกว่าวิธีความคล้ายคลึงมากที่สุด

4) มีบางกรณีที่การจำแนกโดยวิธีความคล้ายคลึงมากที่สุด ให้ความถูกต้องในการจำแนกสูงกว่าวิธีโครงข่ายใยประสาทเทียม ซึ่งโดยมากจะพบในกรณีที่การจำแนกมีค่า Separability สูง ($TD_{ij} \approx 2000$) ได้แก่ การจำแนกระหว่าง หาดทราย โคลน ตะกอน กับ ป่าชายเลน การจำแนกระหว่าง นากุ้ง นาเกลือ กับ พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย การจำแนกระหว่าง สวนยางพารา กับ พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย ซึ่งเมื่อพิจารณาค่า Skewness ของข้อมูลจำแนกจากตารางที่ 4.4 พบว่า ข้อมูลการจำแนกที่เป็น หาดทราย โคลน ตะกอน และ พื้นที่ว่างเปล่ามีพืชปกคลุมน้อย มีค่า Skewness ที่ต่ำ จึงกล่าวได้ว่า วิธีความคล้ายคลึงมากที่สุด มีแนวโน้มที่จะให้ผลความถูกต้องในการจำแนกข้อมูลซึ่งมีค่า Separability สูงและมีค่า Skewness ต่ำ ซึ่งเป็นลักษณะของข้อมูลที่มีความใกล้เคียงกับสมมติฐานของการจำแนก สูงกว่าการจำแนกโดยวิธีโครงข่ายใยประสาทเทียม

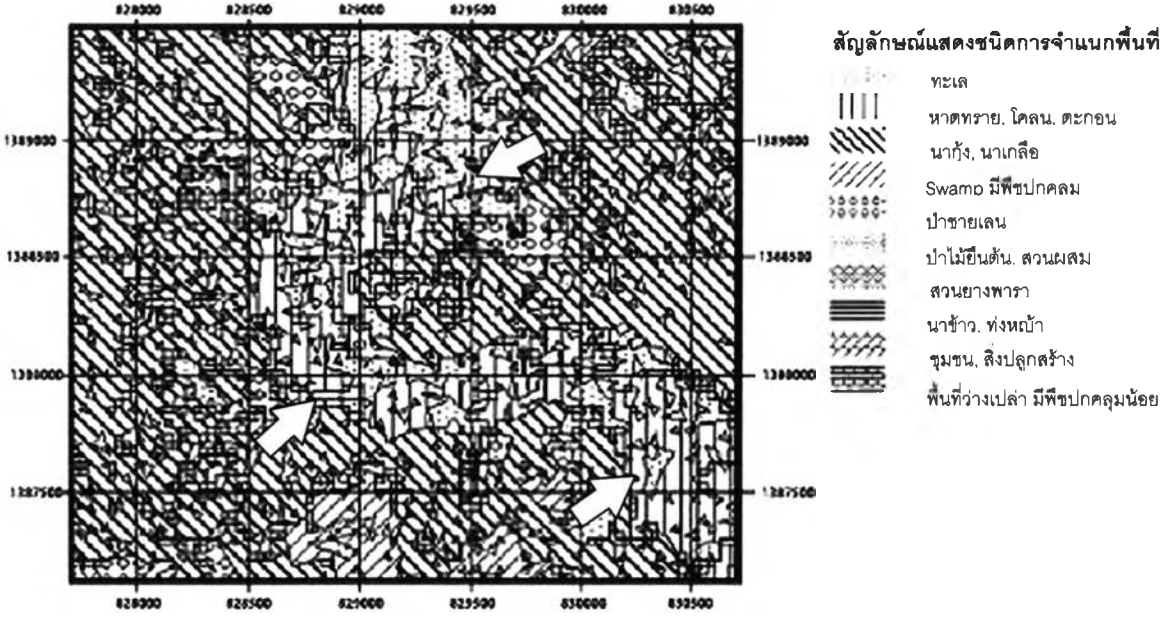
4.2.3 ตัวอย่างภาพที่ได้จากการจำแนก

ภาพการจำแนกเป็นสิ่งที่นำไปใช้ประโยชน์จากการจำแนกภาพดาวเทียม ซึ่งในงานวิจัยครั้งนี้ได้ทำการจำแนกภาพดาวเทียมที่ได้จากวิธีความคล้ายคลึงมากที่สุด และโครงข่ายใยประสาทเทียมของพื้นที่ศึกษา โดยในหัวข้อนี้จะเป็นการยกตัวอย่างภาพที่ได้จากการจำแนกพื้นที่จากวิธีการจำแนกทั้งสองวิธี

1) ภาพการจำแนกของพื้นที่บริเวณใกล้ปากแม่น้ำ



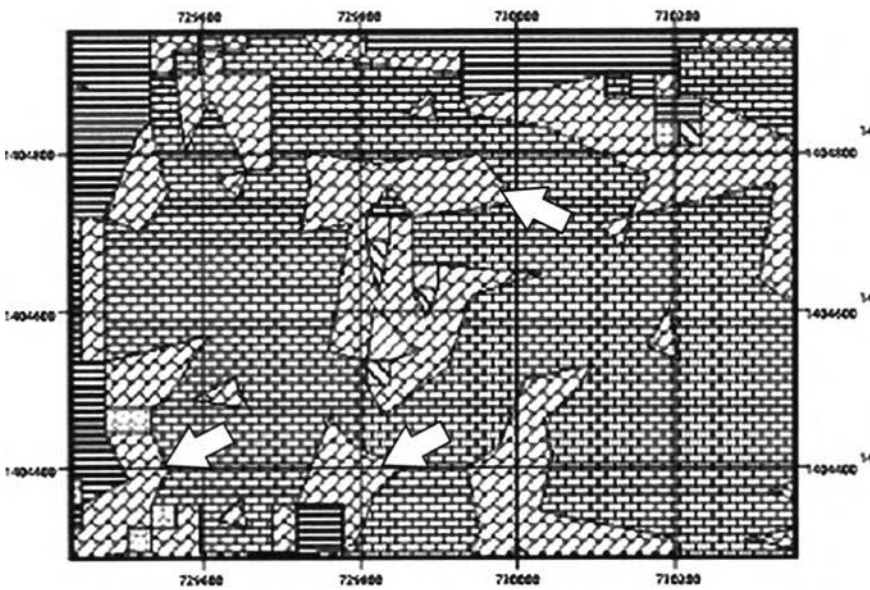
รูปที่ 4.5(ก) ตัวอย่างผลการจำแนกพื้นที่บริเวณใกล้ปากแม่น้ำด้วยวิธีความคล้ายคลึงมากที่สุด



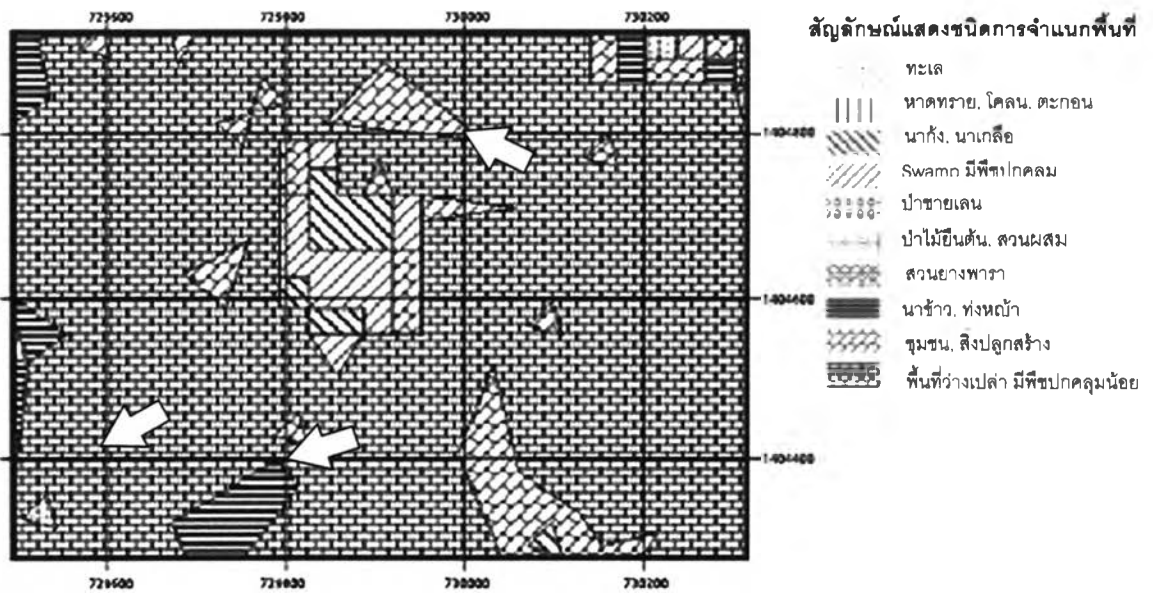
รูปที่ 4.5(ข) ตัวอย่างผลการจำแนกพื้นที่บริเวณใกล้ปากแม่น้ำด้วยวิธีโครงข่ายใยประสาทเทียม

จากตัวอย่างของภาพการจำแนกพื้นที่บริเวณใกล้ปากแม่น้ำ พบว่าเป็นบริเวณซึ่งภาพที่ได้จากวิธีการจำแนกทั้งสองวิธีมีความแตกต่างกันอย่างเห็นได้ชัด โดยเฉพาะบริเวณแม่น้ำ ซึ่งภาพที่จำแนกโดยวิธีความคล้ายคลึงมากที่สุดจะประกอบไปด้วยโคลน และตะกอนเป็นส่วนใหญ่ ในขณะที่ภาพที่ได้จากการจำแนกโดยวิธีโครงข่ายใยประสาทเทียม มีการปะปนกันระหว่างน้ำทะเล และ โคลน ตะกอน

2) ภาพการจำแนกการใช้ประโยชน์พื้นที่ของชุมชน



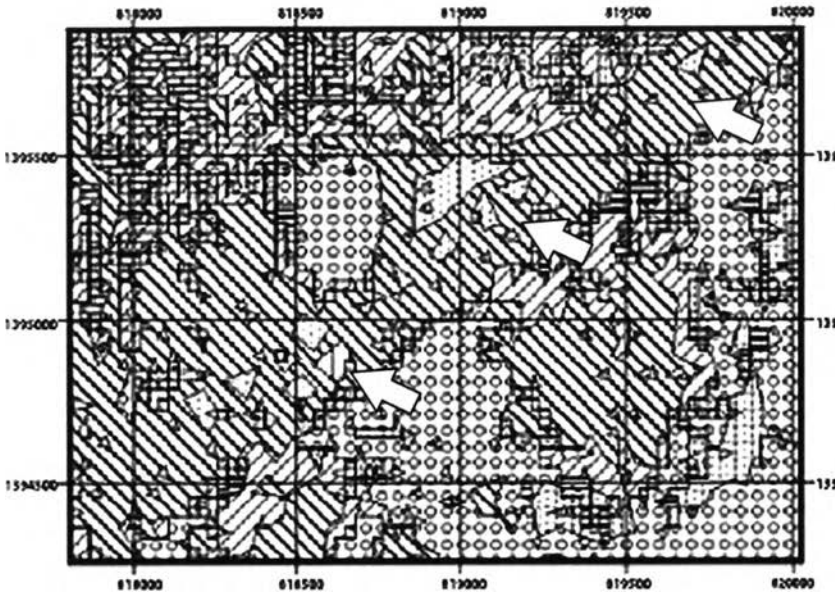
รูปที่ 4.6(ก) ตัวอย่างผลการจำแนกการใช้ประโยชน์พื้นที่ของชุมชน
ด้วยวิธีความคล้ายคลึงมากที่สุด



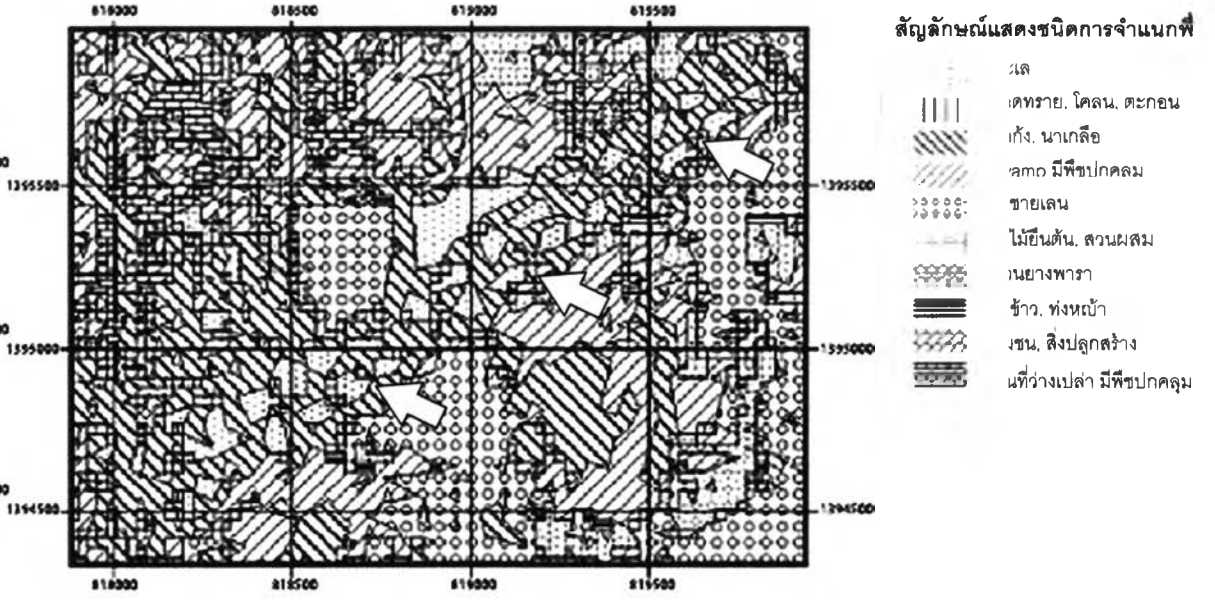
รูปที่ 4.6(ข) ตัวอย่างผลการจำแนกการใช้ประโยชน์พื้นที่ของชุมชน
ด้วยวิธีโครงข่ายใยประสาทเทียม

จากภาพตัวอย่างการจำแนกการใช้ประโยชน์พื้นที่ของชุมชน จะพบว่า ภาพที่ได้จากการจำแนกโดยวิธีความคล้ายคลึงมากที่สุด พื้นที่ส่วนใหญ่จะเป็นพื้นที่โล่ง และชุมชน สิ่งปลูกสร้าง ซึ่งมีการปะปนของพื้นที่การทำการเกษตรบ้างเล็กน้อย ในขณะที่ พื้นที่ซึ่งได้การจำแนกที่ได้จากการจำแนกโดยวิธีโครงข่ายใยประสาทเทียมส่วนใหญ่จะเป็นพื้นที่โล่งและพื้นที่ทางการเกษตร โดยมีการปะปนของชุมชน โดยบริเวณพื้นที่ซึ่งถูกจำแนกเป็นชุมชนสิ่งปลูกสร้างในรูปที่ 4.6(ข) มีปริมาณน้อยกว่าในรูปที่ 4.6(ก)

3) การจำแนกพื้นที่การทำนากุ้ง และนาเกลือ



รูปที่ 4.7(ก) ตัวอย่างผลการจำแนกพื้นที่การทำนากุ้ง นาเกลือ
ด้วยวิธีความคล้ายคลึงมากที่สุด



รูปที่ 4.7(ข) ตัวอย่างผลการจำแนกพื้นที่การทำนากุ้ง นาเกลือ
ด้วยวิธีโครงข่ายใยประสาทเทียม

จากภาพที่ได้จากการจำแนกพื้นที่การทำนากุ้ง และนาเกลือ ในรูปที่ 4.7 และรูปที่ 4.8 พบว่าภาพทั้งสองให้ผลการจำแนกที่ใกล้เคียงกัน แต่พบความแตกต่างของการจำแนกพื้นที่ซึ่งเป็นน้ำทะเล ภายในนากุ้ง และนาเกลือ จากการจำแนกทั้งสองวิธีมีความแตกต่างกัน โดยภาพที่ได้จากการจำแนกโดยวิธีโครงข่ายใยประสาทเทียมจะมีพื้นที่ที่เป็นน้ำในปริมาณที่มากกว่า ภาพที่ได้จากการจำแนกโดยวิธีโครงข่ายใยประสาทเทียม

จากภาพตัวอย่างทั้ง 6 ภาพ ซึ่งเป็นภาพการเปรียบเทียบผลการจำแนกภาพดาวเทียมจากทั้งสองวิธี แสดงให้เห็นถึงความสำคัญของวิธีการจำแนกที่มีผลต่อความหมายของพื้นที่ โดยเฉพาะรูปที่ 4.6(ก) กับ รูปที่ 4.6(ข) ซึ่งมีความแตกต่างกันโดยสิ้นเชิงระหว่างการปะปนของพื้นที่ทำการเกษตรในชุมชน และพื้นที่ชุมชนปะปนในพื้นที่ทำการเกษตร ดังนั้นวิธีการจำแนกที่มีความถูกต้องจึงเป็นสิ่งสำคัญในการจำแนกภาพดาวเทียม