

การจำตัวอักษรอังกฤษแบบตัวพิมพ์ โดยนิวนิวรอลเน็ตเวิร์ค

PRINTED ENGLISH CHARACTER RECOGNITION

BY NEURAL NETWORK



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของงานที่ศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2540

ISBN 974-622-003-9

ลิขสิทธิ์ของบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

การจำตัวอักษรอังกฤษแบบตัวพิมพ์โดยนิวรอลเน็ตเวิร์ค  
PRINTED ENGLISH CHARACTER RECOGNITION  
BY NEURAL NETWORK



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาดมหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมไฟฟ้า  
บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2540

ISBN 974-622-003-9

ลิขสิทธิ์ของบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

PRINTED ENGLISH CHARACTER RECOGNITION  
BY NEURAL NETWORK



A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE  
MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING  
SCHOOL OF GRADUATE STUDIES

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1997

ISBN 974-622-003-9

**บัณฑิตวิทยาลัย**  
**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**  
**ใบรับรองวิทยานิพนธ์**

**หัวข้อวิทยานิพนธ์** การจำตัวอักษรอังกฤษแบบตัวพิมพ์โดยนิวรอลเน็ตเวิร์ค  
 PRINTED ENGLISH CHARACTER RECOGNITION BY NEURAL NETWORK

**ชื่อนักศึกษา** นายเศรษฐพล ลีนปราชญา รหัสประจำตัว 34620054

**หลักสูตร** วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชา วิศวกรรมไฟฟ้า

**ภาควิชา** วิศวกรรมคอมพิวเตอร์

**อาจารย์ผู้ควบคุมวิทยานิพนธ์** รศ.ดร.กิตติ ไพฑูรย์วัฒนกิจ

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
รศ.ดร.กิตติ	ไพฑูรย์วัฒนกิจ	
รศ.ดร.จเร	สุรวุฒน์ปัญญา	
รศ.ดร.ครรชิต	ไมตรี	
รศ.ดร.รัตติกง	วราภูกสิริพันธุ์	
รศ.ดร.ชม	กัมปาน	

ค่าระดับคะแนนที่เป็นเอกฉันท์จากคณะกรรมการสอบ **GOOD**

วัน/เดือน/ปี ที่สอบ 3 กรกฎาคม 2540 เวลา 13.30 น. ถึงเวลา 15.30 น.

สถานที่สอบ ห้องสอบวิทยานิพนธ์ คณะวิศวกรรมศาสตร์ ตึก 12 ชั้น ชั้น 4 ห้อง (E4-5-4)

บัณฑิตวิทยาลัยรับรองแล้ว  
 (รศ.ดร.มนัส ตั้งวรศิลป์)  
 คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน.....พ.ศ.....

**หมายเหตุ** การวัดผลวิทยานิพนธ์ให้ใช้ค่าระดับคะแนนดังนี้

ค่าระดับคะแนน	ผลการศึกษา
O	Outstanding (ดีเยี่ยม)
G	Good (ดี)
P	Pass (ผ่าน)
F	Fail (ไม่ผ่าน)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรรมใดๆ ถ้อยคำอื่นที่มีให้คัดลอก และต้องแจ้งถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การรู้จำตัวอักษรอังกฤษแบบตัวพิมพ์โดยนิเวรอลเน็ตเวิร์ค
นักศึกษา	นายเศรษฐพล ลิ่นปราชญา
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.ดร.กิตติ ไพฑูรย์วัฒน์กิจ
ระดับการศึกษา	วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า
ภาควิชา	วิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.	2540

### บทคัดย่อ

วิทยานิพนธ์ฉบับนี้เสนอการรู้จำตัวอักษรอังกฤษแบบตัวพิมพ์โดยใช้นิเวรอลเน็ตเวิร์คแบบหลายชั้น (Multi-layer Neural Network) ซึ่งใช้การเรียนรู้แบบแบ็คพร็อบพาเกชัน (Back-propagation) ในการแยกประเภทตัวอักษร (Classification) เนื่องจากมีความยืดหยุ่นในการทำงานสูง สามารถเรียนรู้ความสัมพันธ์ระหว่างอินพุตแพทเทิร์นและเอาต์พุตแพทเทิร์นได้เป็นอย่างดี ถึงแม้ว่าข้อมูลของอินพุตแพทเทิร์นจะขาดความสมบูรณ์หรือข้อมูลไม่เคยเรียนรู้มาก่อน การรู้จำตัวอักษรประกอบด้วยการนำข้อมูลเข้า การประมวลผลขั้นต้น (Preprocessing) การแยกคุณลักษณะเฉพาะของตัวอักษร (Feature extraction) และการแยกประเภทตัวอักษร ในขั้นตอนการประมวลผลขั้นต้นยังคงใช้การประมวลผลภาพเอกสารโดยทั่วไป ส่วนขั้นตอนการแยกคุณลักษณะเฉพาะของตัวอักษรจะเลือกใช้หลักการของ Curve approximation ทั้งคุณลักษณะเฉพาะและคุณลักษณะเฉพาะจุดภาพ (Pixels feature) จะถูกนำมาใช้เป็นอินพุตของนิเวรอลเน็ตเวิร์คในระหว่างฝึกสอนโดยวิธีของแบ็คพร็อบพาเกชันเพื่อให้ได้ค่าผิดพลาดตามที่ต้องการ

จากผลการทดสอบโดยเขียนโปรแกรมเพื่อรับอินพุตแพทเทิร์นของตัวอักษรในรูปบิตแมพภาพแบบสองมิติจากเครื่องอ่านสแกนเนอร์ นำเข้ามาเป็นภาพขาวดำใช้ความละเอียด 300x300 จุดต่อนิ้ว (dpi) ขนาดแพทเทิร์น 9x15 จุดภาพต่อหนึ่งตัวอักษร โดยคัดเลือกตัวอักษรภาษาอังกฤษ ตัวพิมพ์ใหญ่ ตัวพิมพ์เล็ก ตัวเลขและตัวอักษรพิเศษบางตัว รวมทั้งสิ้น 74 คลาส จากวารสารจำนวน 7 เล่มและหนังสือพิมพ์ภาษาอังกฤษจำนวน 2 ฉบับใช้เป็นฐานข้อมูลสำหรับการฝึกสอน แล้วทดสอบการรู้จำตัวอักษรจากวารสารจำนวน 6 เล่มและหนังสือพิมพ์ภาษาอังกฤษจำนวน 2 ฉบับให้ผลการรู้จำถูกต้องโดยเฉลี่ย 98.247 เปอร์เซ็นต์สำหรับแบบอักษรที่ใช้ฝึกสอน และรู้จำถูกต้องโดยเฉลี่ย 96.848 เปอร์เซ็นต์สำหรับแบบอักษรที่ยังไม่เคยฝึกสอนมาก่อนจากวารสารจำนวน 5 เล่ม โดยความเร็วของการรู้จำตัวอักษรประมาณ 3 ตัวอักษรต่อวินาที

<b>Thesis Title</b>	Printed English Character Recognition by Neural Network
<b>Student</b>	Mr. Settapol Linprachya
<b>Thesis Advisor</b>	Assoc. Prof. Dr. Kitti Paithoonwattanakij
<b>Level of Student</b>	Master of Engineering in Electrical Engineering
<b>Department</b>	Computer Engineering King Mongkut's Institute of Technology Ladkrabang
<b>Year</b>	1997

### Abstract

This thesis proposes the printed English character recognition by neural network. The multi-layer neural network with back-propagation learning algorithm are used for the pattern classification. This neural network is flexible and can learn the relationship between input patterns and output patterns even the distorted input patterns or the untrained input data. Character recognition starts with data acquisition, preprocessing, feature extraction and classification. For preprocessing, the classical document image processing are employed and the curve approximation principle has been adopted in the feature extraction step. Both the curve feature based and pixel based are feeded into the input nodes of neural network in order to achieve the acceptable error during the training by back-propagation method.

For experiment, the written program acquires input pattern in two dimensions bit-map graphic from the scanner machine with resolution 300x300 dots per inch (dpi), 9x15 pixels per font. The total of 74 classes from seven magazines and two newspapers is trained with alphabet, numeric and some specific symbols. For testing with six magazines and two newspapers, the average recognition accuracy achieves 98.247 percent for the trained character and 96.848 percent for the untrained character for five magazines. The approximate recognition speed is three characters per second.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้เพราะได้รับความเมตตาจากรองศาสตราจารย์ ดร. กิตติ ไพฑูรย์วัตนกิจ ที่ได้ให้ความกรุณาแนะนำและให้ความคิดเห็นแก่ผู้วิจัยตลอดมา ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างสูง

สุดท้ายนี้ผู้วิจัยขอขอบคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง และอาจารย์ท่านอื่นๆ ที่มีส่วนส่งเสริมให้ผู้วิจัยประสบผลสำเร็จในระดับการศึกษาระดับปริญญาตรี สาขาวิชาวิศวกรรมไฟฟ้าคอมพิวเตอร์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ .....	III
สารบัญ.....	IV
สารบัญตาราง.....	VIII
สารบัญภาพ.....	IX
บทที่	
1 บทนำ .....	1
1.1 วัตถุประสงค์และขอบเขตของการวิจัย .....	2
1.2 โครงร่างของวิทยานิพนธ์ .....	4
2 อาร์ตฟิเซียลนิวรอลเน็ตเวิร์ค .....	6
2.1 นิวรอลทางชีววิทยา .....	7
2.2 คอมพิวเตอร์กับนิวรอลเน็ตเวิร์ค .....	8
2.2.1 แอคติเวชันฟังก์ชันแบบฮาร์ดลิมิต (Hard Limited) .....	10
2.2.2 แอคติเวชันฟังก์ชันแบบค่อเนื่องของซิกมอย (Sigmoid) .....	11
3 นิวรอลเน็ตเวิร์คแบบเลเยอร์เดียว .....	15
3.1 กฎการเรียนรู้ของเพอเซ็ปตรอน .....	16
3.2 กฎการเรียนรู้ของเคลด้า .....	17
3.3 ทดสอบการเรียนรู้ของเพอเซ็ปตรอนกับฟังก์ชันบูลีนแบบแอนด์ .....	20
3.4 ข้อเสนอแนะและผลการวิเคราะห์ .....	27
4 นิวรอลเน็ตเวิร์คแบบหลายเลเยอร์ .....	29
4.1 แบ็คพร็อพพาเกชัน (Back Propagation) .....	30
4.1.1 ความสำคัญของฮิดเดนเลเยอร์และขอบเขตการตัดสินใจ (Decision regions) ..	34
4.1.2 ความเหมาะสมของการกำหนดค่าเอพ็พทในทางปฏิบัติ .....	35
4.1.3 อัตราการเรียนรู้.....	36

## สารบัญ (ต่อ)

บทที่	หน้า
4.1.4	การใช้ค่าโมเมนตัม ..... 37
4.1.5	Local minima ของแบ็คพรีอพาทาเกชัน ..... 37
4.1.6	ความเสียหายจากค่าสมมาตร (Symmetry breaking) ..... 38
4.2	นิเวศน์เน็ตเวิร์กกับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ ..... 39
4.2.1	ทดสอบการเรียนรู้ของแบ็คพรีอพาทาเกชันกับพีชคณิตบูลีนแบบ เอ็กซ์คลูซีฟออร์ ..... 42
4.2.2	ทดสอบการเรียนรู้สำหรับปัญหาของเอ็กซ์คลูซีฟออร์เพิ่มเติม ..... 46
4.3	ข้อเสนอแนะและผลการวิเคราะห์ ..... 48
5	ทดสอบการเรียนรู้และแยกแยะแพทเทิร์นตัวเลขตั้งแต่ 0 ถึง 9 ..... 49
5.1	การกำหนดค่าเริ่มต้นก่อนการคำนวณ ..... 49
5.2	ขั้นตอนการคำนวณและลำดับการทำงาน ..... 51
5.3	ผลการทดสอบ ..... 58
5.4	เทคนิคการฝึกสอนเพิ่มเติม ..... 61
5.5	ข้อเสนอแนะและผลการวิเคราะห์ ..... 65
6	ทดสอบการรู้จำตัวอักษรด้วยแสงสำหรับตัวอักษรตัวพิมพ์ภาษาอังกฤษ ..... 67
6.1	ความรู้พื้นฐานของการรู้จำตัวอักษร (Character Recognition) ..... 68
6.1.1	การนำเข้าข้อมูล (Data acquisition) ..... 70
6.1.1.1	แบบอักษร ..... 70
6.1.1.2	ตัวอักษรตัวพิมพ์ (Printed Character) ..... 71
6.1.1.3	ตัวอักษรลายมือเขียนและลายมือพิมพ์ (Handwritten และ Handprinted Character) ..... 72
6.1.2	การประมวลผลในขั้นต้น (Pre-processing) ..... 73
6.1.3	การแยกแยะคุณลักษณะเฉพาะ (Feature Extraction) ..... 73
6.1.4	การแยกประเภทตัวอักษร (Classification) ..... 74

## สารบัญ (ต่อ)

บทที่	หน้า
6.1.5 การประมวลผลในลำดับหลัง (Post-processing) .....	76
6.2 การนำเข้าข้อมูลจากเครื่องสแกนเนอร์ .....	76
6.3 การประมวลผลขั้นต้น .....	78
6.3.1 การวิเคราะห์หน้ากระดาษภาพ (Page layout analysis) .....	79
6.3.2 การแยกบรรทัดตัวอักษร (Text line Extraction) .....	81
6.3.3 การตัดแยกตัวอักษร (Segmentation) .....	82
6.3.4 การตรวจจับความเหลื่อมล้ำ (Overlap) และสัมผัสกัน (Touching detection) ..	83
6.3.4.1 ตัวอักษรเหลื่อมล้ำกัน (Overlap character) .....	88
6.3.4.2 ตัวอักษรสัมผัสกัน (Touching character) .....	90
6.3.5 การตรวจจับระดับตัวอักษร (Level detection) .....	93
6.3.5.1 การแยกประเภทกลุ่มตัวอักษรออกเป็นสองประเภทใหญ่ .....	96
6.3.5.2 การวิเคราะห์เพื่อหา Upper Zone และ Lower Zone ของ บรรทัดตัวอักษร .....	97
6.3.5.3 ผลต่างระดับตัวอักษรที่ติดกัน .....	98
6.3.5.4 การค้นหาระดับของตัวอักษรและการกำหนดกรอบมาตรฐาน .....	101
6.4 การแยกคุณลักษณะเฉพาะของตัวอักษร .....	106
6.4.1 วิธีหาจุดรอบแพทเทิร์นตัวอักษร .....	107
6.4.2 การ Resampling .....	109
6.4.3 คุณลักษณะของ Rescaling .....	110
6.4.4 คุณลักษณะของ Directional angle .....	111
6.4.5 คุณลักษณะของ Curvature angle .....	113
6.4.6 ปากกาขยขึ้นและกดลง (Pen-up และ Pen-down) .....	115
6.4.7 การนำเสนอสื่อข้อมูล (Intermediate representaion) .....	117
6.5 การแยกประเภทของตัวอักษร .....	118
6.5.1 การแปลคำสั่งค่าเอาต์พุท (Interpretation of the output) .....	120

## สารบัญ (ต่อ)

บทที่	หน้า
6.6 ผลการทดสอบ .....	121
6.6.1 ขั้นตอนการอ่านตัวอักษร .....	121
6.6.2 ขั้นตอนการฝึกสอน .....	129
6.6.3 ขั้นตอนการทดสอบผลการรู้จำแพทเทิร์นตัวอักษร .....	142
6.7 ข้อเสนอแนะและผลการวิเคราะห์ .....	152
7 แนวความคิดของการประยุกต์ใช้ในทางปฏิบัติของการรู้จำตัวอักษร .....	154
7.1 การรู้จำตัวอักษรตัวเลขเพื่อใช้เป็นกรณีสำหรับอ้างอิงการจัดเก็บเอกสาร แบบอัตโนมัติ .....	154
7.2 การรู้จำตัวอักษรจากเอกสารเพื่อนำไปใช้งานกับโปรแกรมประเภท Word processing และ Spreadsheet .....	157
7.3 การรู้จำตัวอักษรและการแปลความหมายเป็นเสียงพูด .....	158
7.4 การย่อขนาดเพิ่มข้อมูลภาพด้วยการจัดเก็บเอกสารแบบสื่อผสม ทั้งข้อความและภาพ .....	160
8 สรุปผลการวิจัยและข้อเสนอแนะ .....	163
8.1 สรุปผลการทดสอบโดยรวม .....	163
8.2 ข้อเสนอแนะและผลการวิเคราะห์โดยรวม .....	164
บรรณานุกรม .....	168
ภาคผนวก .....	170
ภาคผนวก ก. โปรแกรมตัวอย่างการทำงานของเพอร์เซพตรอนแบบเลเยอร์เดียว กับพีชคณิตบูลีนแบบแอนด์ .....	171
ภาคผนวก ข. พิสูจน์ที่มาของสมการนิเวรอลเน็ตเวิร์คเบ็คพร็อพพาเกชัน .....	173
ภาคผนวก ค. อธิบายเพิ่มเติมในส่วนของงานวิทยานิพนธ์ .....	178
ประวัติผู้เขียน .....	181

## สารบัญตาราง

ตารางที่		หน้า
1	แสดงตารางพีชคณิตบูลีนแบบแอนด์.....	21
2	แสดงการคำนวณการเรียนรู้แพทเทิร์นของพีชคณิตบูลีนแบบแอนด์ .....	25
3	แสดงความสัมพันธ์ของอินพุทและเอาต์พุท ของพีชคณิตบูลีนแบบแอนด์.....	28
4	แสดงกลุ่มข้อมูลอินพุทแพทเทิร์นของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ .....	39
5	กำหนดค่าความสัมพันธ์อินพุทและเอาต์พุทแพทเทิร์น ของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ .....	43
6	แสดงผลจำนวนรอบและเวลาทั้งหมด เมื่อค่า $\eta$ เปลี่ยนไป .....	61
7	แสดงผลจำนวนรอบและเวลาทั้งหมด เมื่อค่า $E_a$ เปลี่ยนไป .....	62
8	แสดงผลการวิเคราะห์หาระดับตัวอักษร .....	104
9	ตำแหน่ง $x$ และ $y$ ที่หาได้จากการสแกนหาจุดรอบแพทเทิร์นตัวอักษร .....	108
10	แสดงตำแหน่งของจุดรอบแพทเทิร์นหลังการ Interpolation.....	109
11	แสดงคุณลักษณะของ Rescaling หลังการคำนวณ .....	111
12	แสดงคุณลักษณะของ Directional angle หลังการคำนวณ .....	113
13	แสดงคุณลักษณะของ Curvature angle หลังการคำนวณ .....	115
14	แสดงคุณลักษณะของปากกายกขึ้นและกดลงหลังการคำนวณ .....	116
15	แสดงที่มาของแบบอักษรสำหรับการฝึกสอน .....	123
16	ผลการรู้จำตัวอักษรจากภาพในวารสาร POSPECT .....	145
17	ผลการรู้จำตัวอักษรจากภาพในวารสาร ASIAN COMMUNICATION.....	147
18	ผลการรู้จำตัวอักษรจากภาพจากตัวอย่างโปรแกรม Catch Word Pro .....	149
19	ผลการรู้จำตัวอักษรจากวารสารที่ใช้ฝึกสอนและวารสารที่ไม่ได้ใช้ฝึกสอน .....	150

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญญภาพ

		หน้า
1	แสดงนิเวศทางชีววิทยา .....	8
2	อาร์ติฟิเชียลนิเวศกับแอกติเวชันฟังก์ชัน .....	9
3	แสดงชนิดของแอกติเวชันฟังก์ชัน .....	10
4	แสดงแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตทั้งสองแบบ .....	10
5	แสดงแอกติเวชันฟังก์ชันแบบซิกมอยด์ทั้งสองแบบ .....	11
6	แสดงผลของไบอัส และค่าที่ใช้ปรับรูปร่างของซิกมอยด์ฟังก์ชัน .....	13
7	นิเวศเน็ตเวิร์คแบบเลเยอร์เดียว .....	15
8	แสดงโครงสร้างนิเวศเน็ตเวิร์คแบบเพอเซ็ปตรอน .....	16
9	แสดงโครงสร้างกฎการเรียนรู้ของเคลด้า .....	18
10	รูปแสดงความสัมพันธ์ระหว่างค่าผิดพลาดต่ำสุดและค่าตัวเลขน้ำหนัก .....	19
11	แสดงความสัมพันธ์ของพีชคณิตบูลีนแบบแอนด์ .....	20
12	แสดงแบบจำลองการเรียนรู้ของเพอเซ็ปตรอนแบบ ไม่ต่อเนื่องกับพีชคณิตบูลีนแบบแอนด์ .....	21
13	แสดงการแยกจำพวกของกลุ่มแพทเทิร์นอินพุทของพีชคณิตบูลีนแอนด์ .....	26
14	โครงสร้างของนิเวศเน็ตเวิร์คแบบหลายเลเยอร์ .....	29
15	แสดงผังงานการทำงานของแบ็คพร้อพกาเกชันอัลกอริทึม .....	33
16	แสดงชนิดของขอบเขตการตัดสินใจในของโครงสร้างแบบเลเยอร์เดียว และหลายเลเยอร์กับตัวอย่างข้อมูลในคลาส A และ B .....	34
17	แสดงการหาอนุพันธ์อันดับหนึ่งของฟังก์ชันแอกติเวชัน .....	36
18	แสดงทิศทางการเคลื่อนที่ของการปรับค่าตัวเลขน้ำหนัก กับปัญหาของ local minima .....	38
19	แสดง โครงสร้างนิเวศเน็ตเวิร์คสำหรับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ .....	40
20	กราฟแสดงการแยกจำพวกของกลุ่มแพทเทิร์นของพีชคณิตบูลีน แบบเอ็กซ์คลูซีฟออร์สำหรับสองฮิคเคนยูนิต .....	41

## สารบัญภาพ (ต่อ)

หน้า

21	แสดงโครงสร้างนิวรอลเน็ตเวิร์คของพีชคณิตบูลีน แบบเอ็กซ์คลูซีฟออร์ในทางปฏิบัติ .....	42
22	แสดงโทโปโลยีของเน็ตเวิร์คในรูปแบบต่างๆ .....	47
23	แสดงแพทเทิร์นตัวอักษรตัวเลข "0".."9" .....	49
24	แสดงโครงสร้างการทำงานของนิวรอลเน็ตเวิร์ค .....	50
25	ขั้นตอนการทำงานของโปรแกรมในรูปแบบของผังงาน .....	51
26	แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "0".."9" .....	54
27	แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์น เมื่อข้อมูลอินพุตถูกบิดเบือน .....	59
28	แสดงการกำหนดค่า $\eta$ แตกต่างกันตามสูตร .....	63
29	แสดงผลการฝึกสอน โดยกำหนดค่า $\eta$ แตกต่างกันตามสูตร .....	64
30	ขั้นตอนการประมวลผลทางด้านการรู้จำตัวอักษร .....	69
31	แสดงตัวอย่างแบบอักษรที่ใช้งานในปัจจุบัน .....	71
32	แสดงผลการพิมพ์จากเครื่องพิมพ์ต่างชนิด .....	72
33	แสดงตัวอย่างลายมือพิมพ์ที่ได้จากแบบฟอร์มที่กำหนดให้ .....	72
34	แสดงการใช้ Probability density function .....	75
35	แสดงการแปลงสัญญาณภาพให้เป็นภาพ ในลักษณะบิตแมพสำหรับใช้กับเครื่องคอมพิวเตอร์ .....	77
36	แสดงขั้นตอนก่อนการประมวลผล .....	78
37	แสดงภาพของข้อความที่ผ่านการอ่านด้วยเครื่องสแกนเนอร์ .....	79
38	แสดงวิธีการฉายภาพด้านข้าง .....	80
39	แสดงความถี่ของฮิสโตแกรมในแนวนอนของส่วนข้อความตัวอักษร .....	81
40	แสดงตัวอักษรที่วางติดกันและวิธีการตัดแยกตัวอักษร .....	82
41	แสดงวิธีการหาขอบซ้าย ขอบขวา ขอบบน ขอบล่าง ของตัวอักษร .....	82
42	แสดงตัวอย่างการแยกตัวอักษรที่วางติดกัน .....	83
43	แสดงกรอบตัวอักษรตั้งแต่สองตัวขึ้นไปที่มีโอกาสติดกัน .....	85
44	แสดงการกำหนดทิศทางการเคลื่อนที่ในรูปแบบเมตริก .....	86

## สารบัญญภาพ (ต่อ)

		หน้า
45	แสดงการตรวจหาเส้นทางที่ถูกต้อง .....	87
46	แสดงค่า cost ในแต่ละคอลัมน์ .....	88
47	แสดงตัวอักษรที่เชื่อมล้ากันสามตัว .....	89
48	แสดงการแบ่งตัวอักษรสำหรับกรณีที่เกิดพลาด .....	90
49	แสดงค่า cost ในแต่ละคอลัมน์ .....	91
50	แสดงค่า cost ในแต่ละคอลัมน์สำหรับตัวอักษรสามตัวติดกัน .....	91
51	แสดงการเปรียบเทียบค่า cost สำหรับกรณีการตัดที่ไม่ถูกต้อง .....	92
52	แสดงการกำหนดคุณสมบัติของระดับตัวอักษร .....	93
53	แสดงขั้นตอนการกำหนดระดับของตัวอักษร .....	95
54	แสดงการกำหนดระดับด้วยการแยกประเภทกลุ่มตัวอักษร .....	96
55	แสดงการแบ่งโซนของตัวอักษร .....	97
56	แสดงตัวอักษรที่ได้จากการตัดแยกตัวอักษร .....	97
57	แสดงขั้นตอนการวิเคราะห์ตัวอักษรติดกัน .....	98
58	แสดงการจัดเรียงค่าผลต่างระหว่างระดับ .....	99
59	แสดงการค้นหาระดับของตัวอักษร .....	102
60	แสดงผลการกำหนดระดับของตัวอักษรในขั้นตอนแรก .....	103
61	แสดงการวิเคราะห์หาระดับตัวอักษรจากตำแหน่งซ้ายสุด ไปจนสุดตัวอักษรทางขวามือ .....	103
62	แสดงวิเคราะห์หาระดับของอักษรจากตำแหน่งแรก ทางขวามือ ไปจนสุดตัวอักษรทางซ้ายมือ .....	104
63	แสดงการหารอบมาตรฐานของตัวอักษร .....	105
64	ขั้นตอนการทำงานในส่วนของการแยกคุณลักษณะเฉพาะ .....	106
65	แสดงทิศทางการสแกนหาจุดเริ่มต้นรอบแพทเทิร์นตัวอักษร .....	107
66	แสดงการสแกนหาจุดรอบแพทเทิร์นตัวอักษรด้วยการกำหนด กรอบจุดภาพขนาด 3x3 จากทิศทางทั้ง 8 .....	108
67	แสดงจุดรอบแพทเทิร์นตัวอักษรก่อนและหลัง Resampling .....	109

## สารบัญญภาพ (ต่อ)

		หน้า
68	แสดงทิศทางการเคลื่อนที่ของจุดที่สัมพันธ์กัน .....	112
69	แสดงความโค้งที่เกิดจากจุดที่สัมพันธ์กัน .....	114
70	แสดงจุดที่สัมพันธ์กับการยกขึ้นและกดลง .....	116
71	แสดงโครงสร้างของนิวรอลเน็ตเวิร์คที่ใช้ในการรู้จำ สำหรับวิทยานิพนธ์ฉบับนี้ .....	119
72	แสดงแผนผังการคัดเลือกแพทเทิร์นตัวอักษร .....	122
73	แสดงตัวอย่างแพทเทิร์นตัวอักษรที่นำเข้ามา .....	123
74	แสดงกลุ่มของแพทเทิร์นตัวอักษรที่จัดเก็บในแต่ละคลาส .....	124
75	แสดงแผนผังการฝึกสอน .....	129
76	แสดงโครงสร้างของนิวรอลเน็ตเวิร์คสำหรับ การรู้จำแพทเทิร์นตัวอักษรใน CategoriesI .....	131
77	แสดงโครงสร้างของนิวรอลเน็ตเวิร์คสำหรับ การรู้จำแพทเทิร์นตัวอักษรใน CategoriesII .....	132
78	แสดงตัวอย่างการฝึกสอนแพทเทิร์น	
79	แสดงคำผิดพลาดโดยเฉลี่ยของระบบลดลงตามจำนวนรอบ .....	142
80	แสดงกระบวนการรู้จำตัวอักษร .....	143
81	แสดงตัวอย่างการรู้จำตัวอักษรจากวารสาร PROSPECT .....	144
82	แสดงตัวอย่างการรู้จำตัวอักษรจากวารสาร ASIAN COMMUNICATION .....	146
83	แสดงตัวอย่างการรู้จำตัวอักษรจากตัวอย่างภาพที่มากับโปรแกรม Catch Word Pro ..	148
84	แสดงช่องรูปสี่เหลี่ยมสำหรับการกรอกข้อมูล .....	155
85	แสดงขั้นตอนการรู้จำตัวอักษรจากแบบฟอร์มเอกสาร .....	155
86	แสดงตำแหน่งอ้างอิงในแบบฟอร์มที่ใช้สำหรับการรู้จำตัวอักษร .....	156
87	แสดงการนำเข้าข้อมูลและการเชื่อมต่อกับโปรแกรมประยุกต์โดยทั่วไป .....	157
88	แสดงขั้นตอนการประมวลผลภาพเอกสารให้เป็นเสียงพูด .....	159
89	แสดงการแยกประเภทตัวอักษรในภาพเอกสารออกมาเป็นอิสระ .....	161
90	แสดงรูปแบบการแสดงผลบนจอภาพของสื่อผสมทั้งข้อความและภาพ .....	162

## สารบัญภาพ (ต่อ)

	หน้า
91 แสดงภาพเอกสารจากเครื่องพิมพ์แบบจุด .....	165
92 แสดงโครงสร้างการเชื่อมต่อของนิรอลเน็ตเวิร์คด้วย วิธีการแยกกลุ่มของคุณลักษณะออกจากกัน .....	166
93 แสดงการจัดเรียงตัวอักษรภาษาไทยในระดับบรรทัด .....	179



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 1

## บทนำ

คอมพิวเตอร์ในปัจจุบันมีความสามารถในการคำนวณเชิงคณิตศาสตร์สูงมาก และให้ความแม่นยำในการคำนวณได้เร็วกว่าการคำนวณของสมองมนุษย์ แต่อย่างไรก็ตามสมองมนุษย์ก็ยังสามารถทำงานที่ซับซ้อนในบางด้านได้ดีกว่า เช่น การตีความหมายเชิงวัตถุของภาพ ความสามารถในการเรียนรู้และตัดสินใจโดยใช้ประสบการณ์ การค้นหาเหตุผลโดยไม่จำเป็นต้องมีข้อมูลที่สมบูรณ์ การแยกแยะคุณสมบัติที่ใกล้เคียงกันได้อย่างมีประสิทธิภาพ แทนที่จะทำงานตามคำสั่งหรือโปรแกรมที่กำหนดเท่านั้น

จากเหตุผลดังกล่าวทำให้นักวิชาการใช้ความคิดในการออกแบบ และสร้างระบบคอมพิวเตอร์ให้มีความโครงสร้างทางสถาปัตยกรรมเลียนแบบการทำงานของเซลล์ในสมองมนุษย์ (Nerve cell) หรือ นิวรอล (Neural) [1] นิวรอลนั้นทำงานในแง่ของสรีรวิทยาประสาท (Neurophysiology) ซึ่งจะต่อเชื่อมโยงกันเป็นโครงข่ายหน่วยประสาท หรือนิวรอลเน็ตเวิร์ค (Neural Networks) โดยใช้วิธีการประมวลผลแบบขนานและกระจายการทำงาน (Parallel Distributed Processing) เพื่อแปลงข้อมูลจากรูปหนึ่งไปเป็นอีกรูปหนึ่ง

คุณสมบัติของนิวรอลเน็ตเวิร์คแตกต่างจากหลักการงานเชิงคณิตศาสตร์โดยทั่วไปในเรื่องของ Massively parallel processing [2] ประการแรก คุณสมบัติทนต่อความผิดพลาด (Fault tolerance) หมายถึงระบบประมวลผลจะไม่หยุดทำงานกลางคัน แม้ว่าหน่วยประมวลผลบางหน่วยจะได้รับความเสียหายจนไม่อาจทำงานได้ ประการที่สอง คุณสมบัติ Generalization หมายถึง นิวรอลเน็ตเวิร์คจะยังคงให้ผลลัพธ์ที่สมเหตุผล แม้ว่าจะมีสัญญาณรบกวน หรือข้อมูลขาดความสมบูรณ์ หรือข้อมูลไม่เคยเห็นมาก่อนก็ตาม ประการที่สาม คุณสมบัติ Adaptability หมายถึง นิวรอลเน็ตเวิร์คสามารถเรียนรู้สภาพแวดล้อมที่เกิดขึ้นใหม่ได้รวดเร็วขึ้น ซึ่งแตกต่างจากสภาพเดิมที่เคยสอน

แนวความคิดในการประยุกต์การใช้งานนิวรอลเน็ตเวิร์คมีอยู่มากมายหลายหัวข้อ [1,2] ได้แก่เรื่อง Natural language processing, Signal processing, Data compression, Machine หรือ Robotic Vision, Pattern Recognition, Decision support system, Medical image processing, Control system ส่วนตัวอย่างการประยุกต์ใช้งานในเชิงธุรกิจ เช่น ตรวจสอบลายเซ็นต้นฉบับเช็คของธนาคาร

ตรวจสอบลายนิ้วมือคนร้าย การจัดสัญญาณรบกวน และเสียงสะท้อนในสายโทรศัพท์ การพยากรณ์อากาศ การวินิจฉัยโรคทางการแพทย์ การควบคุมการขับขี่ยวดยานด้วยตนเอง

การทำงานของนิวรอลเน็ตเวิร์คจะอยู่ในรูปแบบการฝึกสอน (Training) เพื่อปรับฐานความรู้ด้วยค่าตัวเลขน้ำหนัก (Weight) ทำให้ค่าอินพุตและเอาต์พุตของระบบสอดคล้องกัน จากนั้นจึงจะสามารถนำฐานความรู้ (Knowledge base) ที่ได้ไปใช้งานต่อไป

### 1.1 วัตถุประสงค์และขอบเขตของการวิจัย

ศึกษาการทำงานและออกแบบระบบนิวรอลเน็ตเวิร์คที่มีประสิทธิภาพ เพื่อนำไปประยุกต์ใช้กับการรู้จำตัวอักษรตัวพิมพ์ภาษาอังกฤษ และวิเคราะห์ปัญหาที่เกิดขึ้น เพื่อเป็นแนวทางในการพัฒนาการรู้จำตัวอักษรตัวพิมพ์ภาษาไทยต่อไปในอนาคต โดยในเบื้องต้นจะกล่าวถึงความจำเป็นของนิวรอลเน็ตเวิร์ค ตั้งแต่การเลียนแบบโครงสร้างการทำงานของเซลล์ประสาทในสมองมนุษย์ ด้วยหน่วยประมวลผลนิวรอลเพียงหนึ่งหน่วย แล้วประกอบกันขึ้นเป็นเน็ตเวิร์คแบบเลเยอร์เดียว (Single layer network) และเน็ตเวิร์คแบบหลายเลเยอร์ (Multi-layer networks) ในที่สุด โครงสร้างเน็ตเวิร์คดังกล่าวจำเป็นต้องอาศัยอัลกอริทึมบางอย่างในการเรียกใช้งาน ผู้วิจัยได้เลือกศึกษาหลักการฝึกสอนจากกฎการเรียนรู้ของเพอร์เซ็ปตรอน (Perceptron) และแบ็กพร็อพพาเกชัน (Back propagation) ตามลำดับ โดยแสดงขั้นตอนการคำนวณของระบบอย่างละเอียด

การศึกษากฎการเรียนรู้ของเพอร์เซ็ปตรอนด้วยเน็ตเวิร์คแบบเลเยอร์เดียว จะแสดงตัวอย่างการแก้ปัญหาของฟังก์ชันพีชคณิตบูลีนแบบแอนด์ (Boolean AND) ซึ่งสามารถเรียนรู้และทำความเข้าใจได้โดยง่าย เพื่อเป็นแนวทางสำหรับศึกษาการทำงานของเน็ตเวิร์คแบบหลายเลเยอร์ของแบ็กพร็อพพาเกชัน โดยจะมีการทดสอบความสามารถในเชิงทฤษฎีสองเรื่องคือ ทดสอบการแก้ปัญหาของฟังก์ชันพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟอออร์ (Exclusive-OR หรือ XOR) ซึ่งเป็นตัวอย่างที่กล่าวถึงกันมาก เนื่องจากเน็ตเวิร์คแบบเลเยอร์เดียวของเพอร์เซ็ปตรอน (Single layer Perceptron) ไม่สามารถนำมาแก้ปัญหาได้ ส่วนเรื่องที่สอง จะเป็นการศึกษาการทำงานของกรู้อัจฉริยะที่รู้จำแพทเทิร์นตัวอักษรตัวเลขตั้งแต่ 0 ถึง 9 ในทางทฤษฎี โดยใช้การจำลองแบบอักษรของแพทเทิร์นขึ้นมาเองขนาด  $8 \times 8$  บิตแพทเทิร์น และนำบิตแมพจุดภาพ (Bit-map pixel) ของแพทเทิร์นตัวอักษรมาเป็นอินพุตโดยตรงของนิวรอลเน็ตเวิร์คทำการเรียนรู้และแยกแยะข้อมูลแพทเทิร์น จนกระทั่งได้ฐานข้อมูลตามที่ต้องการ จากนั้นจึง

ทดสอบคุณสมบัติทนต่อความผิดพลาดและ Generalization โดยใส่สัญญาณรบกวน (Noise) เข้าไปกับ บิตแมพแพทเทิร์นอินพุท

การพัฒนาต่อไปจะเป็นการประยุกต์ใช้งานนิรอลเน็ตเวิร์คในทางปฏิบัติกับการรู้จำตัวอักษรตัวพิมพ์ภาษาอังกฤษ ขั้นตอนนี้จะไม่นำบิตแมพพิคเซลมาเป็นอินพุทโดยตรงของนิรอลเน็ตเวิร์คทั้งหมด เนื่องจากให้ความแม่นยำ (Accuracy) ในการรู้จำตัวอักษรต่ำมาก จึงต้องศึกษาเพิ่มเติมในเรื่องของการประมวลผลขั้นต้น (Preprocessing) และการแยกคุณลักษณะเฉพาะ (Feature extraction) ของตัวอักษรเป็นสำคัญ ซึ่งมีบทความวิจัยหลายฉบับกล่าวถึงอยู่มาก บทความต่างๆจะใช้ทฤษฎีแตกต่างกันไป การนำมาใช้ขึ้นอยู่กับหลักการแยกประเภทตัวอักษร (Classification) และคุณลักษณะของแบบอักษรที่ใช้ในการรู้จำ โดยผู้วิจัยจะมุ่งศึกษาทฤษฎีที่เหมาะสมที่สุด เพื่อจะสามารถนำไปประยุกต์ใช้กับการรู้จำตัวอักษรตัวพิมพ์ภาษาไทยได้ในอนาคต

โปรแกรมการทดสอบถูกเขียนขึ้นเพื่อใช้เป็นแนวทางในการศึกษาตัวอย่างการรู้จำและแยกแยะตัวอักษรตัวพิมพ์ภาษาอังกฤษนั้น ให้ผลดีพอควรแต่ยังไม่สามารถเทียบเท่ากับโปรแกรมซอฟต์แวร์ในท้องตลาดได้ เนื่องจากมีข้อกำหนดในหัวข้อการวิจัยดังนี้

- เอกสารภาพของตัวหนังสือถูกอ่าน โดยเครื่องสแกนเนอร์แบบตั้งโต๊ะ แสดงผลบนหน้าจอคอมพิวเตอร์ในลักษณะภาพขาวดำ
- เอกสารภาพที่นำเข้ามา จะไม่ปรากฏรูปภาพวัตถุอื่นใดยกเว้นพื้นที่ส่วนของตัวหนังสือ และจะพิจารณาเฉพาะตัวอักษรตัวพิมพ์ภาษาอังกฤษ ตัวใหญ่ ตัวเล็ก ตัวเลข และตัวอักษรพิเศษบางตัวเท่านั้น
- จะคัดเลือกแบบอักษรที่นิยมใช้กันบ่อยๆ จากเอกสารโดยทั่วไป ซึ่งอาจจะครอบคลุมแบบอักษรไม่ได้ทั้งหมด
- ไม่มีการใช้ วากยสัมพันธ์ (Syntax) อรรถศาสตร์ (Semantics) และ โปรแกรมตรวจสอบการสะกดคำจากพจนานุกรม (Dictionary) มาเกี่ยวข้อง

อุปกรณ์ที่ใช้ในงานวิจัยมีดังนี้

1. เครื่องคอมพิวเตอร์ หน่วยประมวลผล Intel 80386 ความเร็ว 40 MHz พร้อมตัวประมวลผลร่วมคณิตศาสตร์ (Math coprocessor) และหน่วยประมวลผล Cyrix 6x86 ความเร็ว 166 MHz พร้อมตัวประมวลผลร่วมคณิตศาสตร์

2. เครื่องสแกนเนอร์แบบตั้งโต๊ะ (Flat-bed scanner) ของ Hewlett Packard รุ่น HP scanner 4P
3. ระบบปฏิบัติการดอส (MSDOS version 6.2) และซอฟต์แวร์ภาษา C ของ Borland C++ เวอร์ชัน 3.0 สำหรับดอส และเวอร์ชัน 4.5 สำหรับไมโครซอฟท์วินโดวส์
4. ตัวอย่างแบบอักษรจากเอกสารบางฉบับที่มีแบบอักษรแตกต่างกัน

## 1.2 โครงร่างของวิทยานิพนธ์

รายละเอียดของวิทยานิพนธ์ประกอบด้วยบทต่างๆ ดังนี้

- บทที่ 1 บทนำจะกล่าวถึงความเป็นมาและวัตถุประสงค์ของงานวิจัย
- บทที่ 2 อาร์ทิฟิเชียลนิวรอลเน็ตเวิร์คจะแสดงการทำงานของเซลล์ตามแบบจำลองทางชีววิทยา กับแบบจำลองนิวรอลเน็ตเวิร์คในแง่ของคอมพิวเตอร์
- บทที่ 3 นิวรอลเน็ตเวิร์คแบบเลเยอร์เดียวจะกล่าวถึงการทำงานของนิวรอลหลายนิวรอลเชื่อมต่อกันเป็นเน็ตเวิร์คแบบเลเยอร์เดียว โดยอาศัยหลักการจากกฎการเรียนรู้ของเพอเซ็ปตรอน (Perceptron learning rule) หรือกฎการเรียนรู้ของเดลต้า (Delta learning rule) กับตัวอย่างพีชคณิตบูลีนแบบแอนด์
- บทที่ 4 นิวรอลเน็ตเวิร์คแบบหลายเลเยอร์จะกล่าวถึงการทำงานของนิวรอลหลายนิวรอลเชื่อมต่อกันเป็นเน็ตเวิร์คตั้งแต่ 2 เลเยอร์ขึ้นไป โดยอาศัยหลักการฝึกสอนของแบ็คพร็อบพาเกชันอัลกอริทึม และเขียน โปรแกรมเพื่อศึกษาหลักการของแบ็คพร็อบพาเกชันกับปัญหาทางพีชคณิตบูลีนแบบเอ็กซ์คลูซิฟออร์
- บทที่ 5 ทดสอบการเรียนรู้และแยกแยะแพทเทิร์นตัวเลขตั้งแต่ 0 ถึง 9 โดยเขียน โปรแกรมทดสอบหลักการและทฤษฎีต่างๆ ของนิวรอลเน็ตเวิร์คแบ็คพร็อบพาเกชันกับบิดแมพแพทเทิร์นของแบบอักษรที่สร้างขึ้นเอง
- บทที่ 6 ทดสอบการเรียนรู้และแยกประเภทแพทเทิร์นตัวอักษรภาษาอังกฤษที่ได้จากสแกนเนอร์ โดยเริ่มศึกษาความรู้พื้นฐานทางด้านการรู้จำตัวอักษร ความหมายของคำศัพท์เฉพาะ ลำดับการทำงานของความรู้จำและแยกประเภทตัวอักษร และหลักการที่เกี่ยวข้องทั้งหมด โดยเขียน โปรแกรมทั้งหมดขึ้นเอง

บทที่ 7 แนวความคิดในการประยุกต์ใช้ในทางปฏิบัติของการรู้จำตัวอักษร เพื่อให้มองเห็น  
ประโยชน์การนำไปใช้งานจริงในอนาคต

บทที่ 8 สรุปผลการทดสอบและข้อเสนอแนะโดยรวมทั้งหมด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

### อาร์ติฟิเชียลนิวรอลเน็ตเวิร์ค

หากจะกล่าวถึงประวัติความเป็นมาของโครงข่ายหน่วยประสาทหรืออาร์ติฟิเชียลนิวรอลเน็ตเวิร์ค (Artificial neural networks) นั้น ในปี 1943 น่าจะเป็นปีเริ่มต้นของการพัฒนาระบบอาร์ติฟิเชียลนิวรอลเน็ตเวิร์ค เนื่องจาก McCulloch และ Pitts (1943) [1] เสนอเค้าโครงรูปแบบหน่วยประมวลผลทางนิวรอล แต่ในสมัยนั้นการนำแนวคิดไปใช้งานยังไม่เกิดผลเท่าที่ควร ต่อมาในปี 1949 Donald Hebb [1] ได้แนะนำทฤษฎีการเรียนรู้สำหรับอาร์ติฟิเชียลนิวรอลเน็ตเวิร์ค ซึ่งนักวิทยาศาสตร์ในสมัยนั้นแสดงให้เห็นว่า เน็ตเวิร์คของนิวรอลสามารถเรียนรู้ได้

ในปี 1950 ถึง 1960 กลุ่มผู้วิจัยรวมทั้งนักชีววิทยาและนักฟิสิกส์ ใช้ความสามารถในการผลิตอุปกรณ์วงจรอิเล็กทรอนิกส์ด้วยความรู้ทางด้านอาร์ติฟิเชียลนิวรอลเน็ตเวิร์คเป็นครั้งแรกโดยผู้ที่มีส่วนสำคัญในการพัฒนาครั้งนั้นมีทั้ง Marvin Minsky (1954), Frank Rosenblatt (1958), Bernard Widrow และ Marcian Hoff (1960) [3] ซึ่งมีส่วนในการพัฒนาเน็ตเวิร์คแบบเลเยอร์เดียวของแบบจำลองนิวรอล ที่เรียกว่า เพอเซ็ปตรอน ทำให้สามารถนำไปใช้กับปัญหาต่างๆในเรื่องของการทำนาย การวิเคราะห์เพื่อออกแบบทางอิเล็กทรอนิกส์ และแบบจำลองการมองเห็น (Artificial vision) เวลานั้นจะเห็นว่าเป็นการค้นพบความสามารถทางความคิดเลียนแบบการทำงานของสมองมนุษย์

ต่อมาไม่นานกลับพบว่า โครงสร้างเน็ตเวิร์คที่กลุ่มนักวิจัยเสนอไปนั้น ไม่สามารถแก้ปัญหาผิวเลนบางอย่างได้สำเร็จ ซึ่ง Minsky และ Seymour Papert (1969) [3] ได้แต่งหนังสือเพื่อพิสูจน์ให้เห็นว่า นิวรอลเน็ตเวิร์คแบบชั้นเดียว มีข้อจำกัดบางประการ ไม่สามารถแก้ปัญหาได้ในบางกรณี รวมทั้งฟังก์ชันการทำงานอย่างง่ายของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟด้วย

ทำให้การพัฒนาทางด้านอาร์ติฟิเชียลนิวรอลเน็ตเวิร์คหยุดชะงักเป็นเวลานาน จนกระทั่งในปี Paul Werbos (1974), Parker (1986), David Rumelhart, James McClelland, Geoffrey Hinton และ Ronald William (1986) [3] ได้เสนอผลงานการค้นคว้าสำหรับเน็ตเวิร์คแบบหลายเลเยอร์ที่เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านธุรกิจสามารถเรียนรู้ได้ด้วยความสามารถของแบคพรีอพทาเกชัน โดยเฉพาะ McClelland, Rumelhart (1986) ไม่จำกัดใดๆทั้งสิ้น อีกทั้งยังมีให้คิดเปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้ และผู้วิจัยจำนวนหนึ่ง ได้รวบรวมทฤษฎีและแนวทางการเรียนรู้ใหม่ๆในหนังสือ 2 เล่มเรื่อง "Parallel Distributed Processing" (PDP) ซึ่งเป็นการเปิดศักราชการคำนวณด้วยศักยภาพของเน็ตเวิร์คแบบหลาย

เลเยอร์ทำให้เกิดผลงานหลากหลาย เช่นการรู้จำตัวอักษรแบบแบบลายมือเขียน (Recognize handwritten character) ของ Burr 1987 [3] การนำไปประยุกต์ใช้กับงานบีบอัดภาพ (Image compression) ของ Cottrell, Munro และ Zipser 1987 เป็นต้น

โดยในปี 1986 และ 1987 นี้เอง จึงได้ถือว่าเป็นจุดเริ่มต้นของการค้นคว้าเทคนิคใหม่ๆ ออกมามากมาย โดยส่วนมากจะเป็นงานทางด้านสมองคอมพิวเตอร์ (Neurocomputing) ซึ่งเกี่ยวข้องกับแบบจำลองนิวรอนเน็ตเวิร์คในเชิงฮาร์ดแวร์โปรเซสเซอร์เชื่อมต่อกันเป็นอะเรย์

ในปีปัจจุบันวิศวกรและนักวิทยาศาสตร์กำลังพยายามพัฒนาเครื่องจักรที่มีความคิด (Intelligent machines) ด้วยระบบโครงข่ายหน่วยประสาทเทียม เนื่องจากสามารถใช้เครื่องคอมพิวเตอร์ช่วยงานในการตัดสินใจอย่างชาญฉลาด โดยเรียนรู้จากประสบการณ์ต่างๆ ที่เราทำกัน

## 2.1 นิวรอนทางชีววิทยา

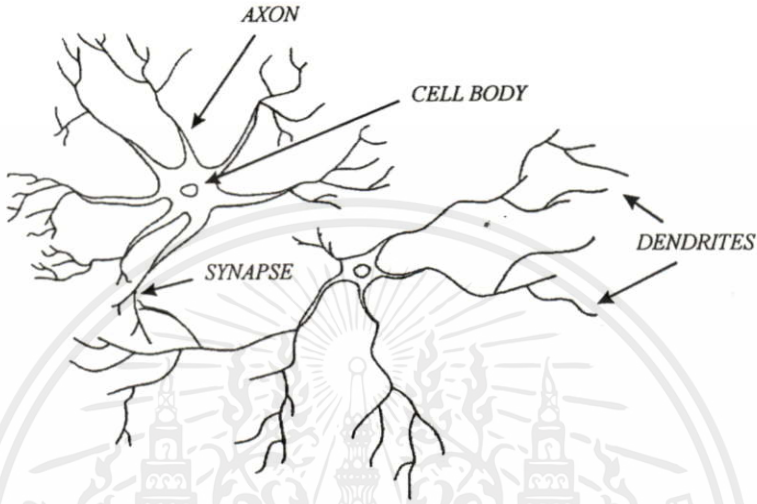
ระบบเส้นประสาทของมนุษย์ประกอบด้วยเซลล์เล็กๆ มากมายที่เรียกว่า นิวรอนหรือเซลล์ประสาทภายในสมองมนุษย์ ซึ่งรวมตัวกันอยู่ในรูปแบบโครงสร้างที่ซับซ้อนมีอยู่ประมาณ  $10^{11}$  [1,3,4] นิวรอน โดยนิวรอนแต่ละตัวประกอบด้วยส่วนสำคัญ 4 ส่วน ดังแสดงในรูปที่ 1

- |                          |   |
|--------------------------|---|
| 1. เด็นไดรท์ (Dendrites) | ตัวรับสัญญาณ อินพุท (Input)               |
| 2. แอ็กซอน (Axons)       | ตัวส่งสัญญาณ เอาท์พุท (Output)            |
| 3. ซินแนปส์ (Synapses)   | จุดเชื่อมต่อ (Connection point)           |
| 4. ตัวเซลล์ (Cell bodys) | จุดศูนย์รวมสัญญาณ (Central neuron system) |

แต่ละนิวรอนจะเชื่อมต่อกับนิวรอนอื่นอีกหลายนิวรอน โดยเด็นไดรท์จะทำหน้าที่รับสัญญาณที่จุดเชื่อมต่อที่เรียกว่า ซินแนปส์ และจะถูกนำไปเก็บสะสมในตัวเซลล์ในช่วงเวลาสั้นๆ มีผลทำให้เกิดการลดความเข้มของสัญญาณหรือเพิ่มความเข้มของสัญญาณ ซึ่งบางสัญญาณมีผลกระตุ้นตัวเซลล์ (Excitatory) บางสัญญาณมีผลยับยั้งตัวเซลล์ (Inhibitory) เมื่อผลรวมของสัญญาณทั้ง

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 1  
แสดงนิรอรทางชีววิทยา



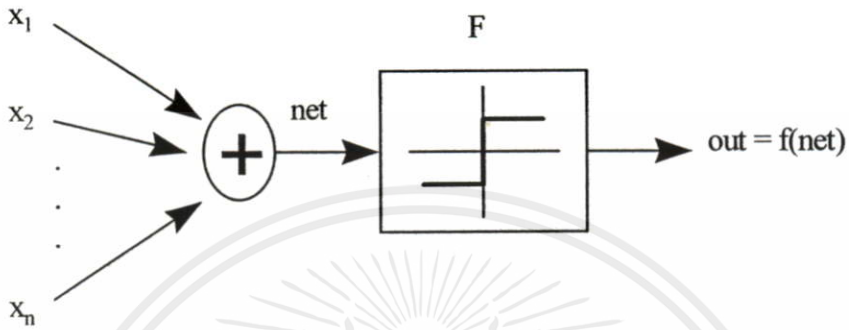
หมดเกินกว่าค่าเทรคโฮล (Threshold) นิรอรจะส่งสัญญาณผ่านแอกซอนไปให้นิรอรอื่นในทันที โดยทั่วไปมีค่าประมาณ 40 มิลลิโวลท์ (Arbid 1987) [1] กล่าวโดยสรุปว่า นิรอรหรือเซลล์ประสาท จะทำหน้าที่ 3 ประการคือ รับสัญญาณที่จุดเชื่อมต่อ (Receptors) และทำการประมวลผลภายในเน็ตเวิร์คของนิรอร (Neural system) และส่งผลการประมวลผลผ่านปลายหนึ่งของเซลล์ไปให้เซลล์อื่น (Effectors)

## 2.2 คอมพิวเตอร์กับนิรอรเน็ตเวิร์ค

แบบจำลองการทำงานของนิรอรเน็ตเวิร์คบนดิจิตอลคอมพิวเตอร์ประกอบด้วยหน่วยประมวลผล (Processing Elements) เชื่อมโยงกับหลายๆ ตัวในลักษณะขนาน คล้ายกับนิรอรในเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า สมอมนมนุษย์ มีลักษณะเป็นสถาปัตยกรรมแบบขนาน (Parallel Architecture) เพื่อแปลงข้อมูลจากรูปหนึ่งไปเป็นอีกรูปหนึ่ง การใช้งานนิรอรเน็ตเวิร์คจะเป็นไปในรูปการฝึกสอน ดังแสดงในรูปที่ 2

## รูปที่ 2

## อาร์คิฟิเซียลนิวรอลกับแอกติเวชันฟังก์ชัน



จากรูปที่ 2 เป็นแผนผังแสดงอาร์คิฟิเซียลนิวรอลกับแอกติเวชันฟังก์ชัน ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

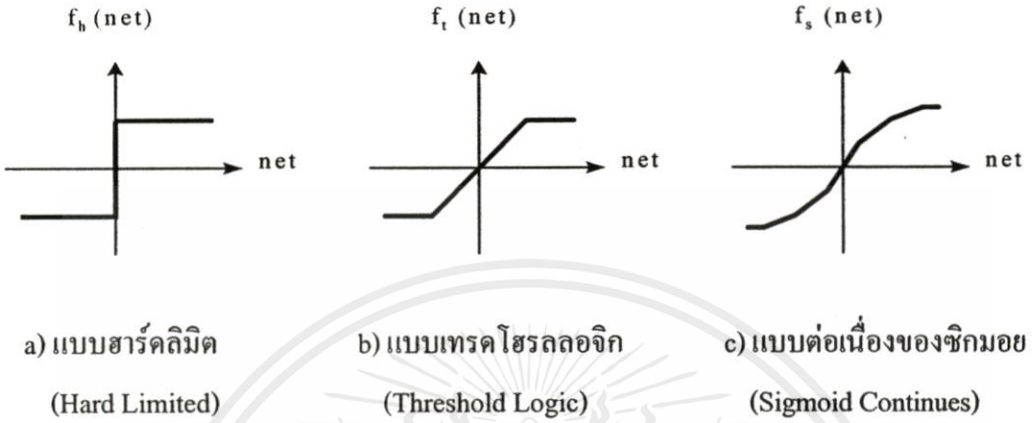
$$net = \sum_i x_i \cdot w_i + \theta \dots\dots\dots (1)$$

$$= x_1 w_1 + x_2 w_2 + \dots + x_n w_n + \theta$$

โดย	$x_1, x_2, \dots, x_n$	เป็นชุดข้อมูลอินพุต
	$w_1, w_2, \dots, w_n$	เป็นค่าตัวเลขน้ำหนัก
	net	เป็นผลรวมสุทธิที่เกิดจากค่าอินพุต (x) คูณกับตัวเลขน้ำหนัก (w)
	$\theta$	เป็นค่าเทรคโธรลของนิวรอล

จากรูปที่ 2 นิวรอล 1 โหนดจะมีค่าอินพุตหลายค่า แต่จะมีค่าเอาต์พุตเพียงค่าเดียว โดยเอาต์พุตในแต่ละนิวรอล เกิดจากผลรวมของผลคูณระหว่างอินพุตและตัวเลขน้ำหนักของอินพุตนั้น ไม่ว่าจะคิดใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3  
แสดงชนิดของแอกติเวชันฟังก์ชัน [5]

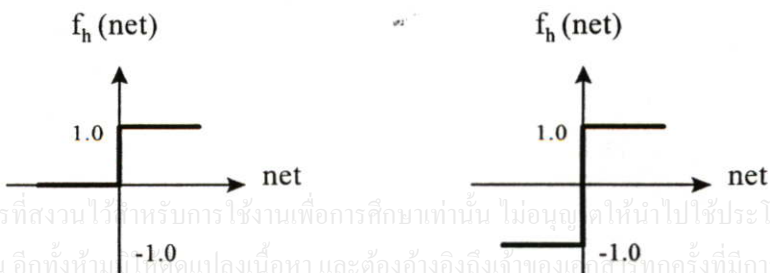


ค่า net จะถูกแปลงสัญญาณ (Transfer หรือ Mapping) แบบไม่เชิงเส้นของแอกติเวชันฟังก์ชัน (Nonlinear Activation function หรือ Nonlinear mapping) ดังแสดงในรูปที่ 3 โดยฟังก์ชันแรกเป็นแอกติเวชันฟังก์ชันแบบไม่ต่อเนื่อง (Discrete) เรียกว่า ฮาร์ดลิมิตหรือแบบเปลี่ยนทันที ให้ผลลัพธ์เพียงสองค่า

### 2.2.1 แอกติเวชันฟังก์ชันแบบฮาร์ดลิมิต (Hard Limited)

รูปที่ 4

แสดงแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตทั้งสองแบบ



เอกสารนี้เป็นเอกสารที่สงวนไว้ใช้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามใช้เพื่อเปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4 แสดงแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตทั้งสองแบบ แอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตแบบแรกจะให้ค่าผลลัพธ์ในช่วง 0 ถึง 1 ดังในสมการที่ (2)

$$\begin{aligned} \text{out} = f(\text{net}) = 0 & ; \text{เมื่อ net มีค่าน้อยกว่าเท่ากับ } 0 \text{ ..... (2)} \\ & 1 ; \text{เมื่อ net มีค่ามากกว่า } 0 \end{aligned}$$

แอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตแบบที่สองจะให้ค่าผลลัพธ์ในช่วง -1 ถึง 1 ดังในสมการที่ (3)

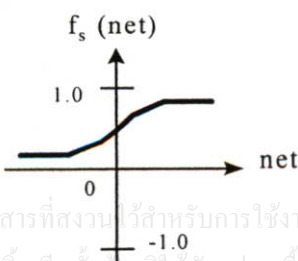
$$\begin{aligned} \text{หรือ} \quad \text{out} = f(\text{net}) = -1 & ; \text{เมื่อ net มีค่าน้อยกว่าเท่ากับ } 0 \text{ ..... (3)} \\ & 1 ; \text{เมื่อ net มีค่ามากกว่า } 0 \end{aligned}$$

### 2.2.2 แอกติเวชันฟังก์ชันแบบต่อเนื่องของซิกมอยด์ (Sigmoid)

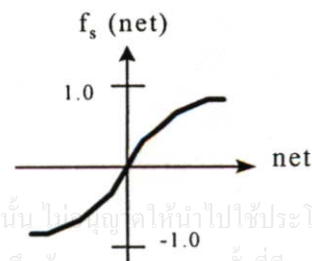
ฟังก์ชันที่สองเรียก เทคโฮทลลอจิก เป็นแอกติเวชันฟังก์ชันที่มีใช้กันน้อยมาก ส่วนใหญ่จะเลือกใช้ฟังก์ชันที่สามมากกว่า โดยทั่วไปเรียก ซิกมอยด์ฟังก์ชัน เป็นแอกติเวชันฟังก์ชันที่มีรูปร่างคล้ายตัวอักษร "S" ในภาษาอังกฤษ

รูปที่ 5

แสดงแอกติเวชันฟังก์ชันแบบซิกมอยด์ทั้งสองแบบ



a) ให้ผลลัพธ์ในช่วง 0 ถึง 1



b) ให้ผลลัพธ์ในช่วง -1 ถึง 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 5(a) และรูปที่ 5(b) แสดงแอกติเวชันฟังก์ชันแบบซิกมอย สามารถนำไปใช้ขยาย สัญญาณ net ในช่วงแคบๆ ได้ ทำให้ได้ค่าจำกัดในขอบเขตที่ต้องการ มีผลลัพธ์อยู่ในช่วง 0 ถึง 1 หรือ -1 ถึง 1 ส่วนในระบบอิเล็กทรอนิกส์แบบแอนะล็อกเราอาจคิดว่าแอกติเวชันฟังก์ชันเป็นการกำหนด อัตราขยายแบบไม่เชิงเส้นให้กับอาร์คิฟิเชียนนิวรอล ซึ่งการเปลี่ยนค่า net เพียงเล็กน้อยในช่วงความ ชันของเส้นกราฟมีผลทำให้ค่า out เปลี่ยนแปลงตามสัดส่วนของอัตราขยาย

จากรูปที่ 5(a) สามารถเขียนเป็นสมการได้ดังนี้

$$\text{out} = f(\text{net}) = \frac{1}{1 + e^{-(\beta \cdot \text{net})}} \dots\dots\dots (4)$$

$$= \frac{\tanh(0.5 \cdot (\beta \cdot \text{net})) + 1}{2} \dots\dots\dots (5)$$

สมการที่ (4) และ (5) ให้ผลลัพธ์เหมือนกันขึ้นอยู่กับความสะดวกในการเลือกใช้ใช้งาน

ข้อสังเกต จากสมการที่ (4) หรือ (5) หากค่าผลรวมอินพุต net มีค่าเป็นบวกมากๆ เอาท์พุท จะมีค่าเข้าใกล้ 1 ในทางตรงกันข้าม ถ้าผลรวมอินพุต net น้อยมากหรือติดลบ ค่าเอาท์พุท out จะมีค่า เข้าใกล้ 0 ซึ่งสอดคล้องกับสถานะเปิด (On) และ ปิด (Off)

จากรูปที่ 5(b) สามารถเขียนเป็นสมการได้ดังนี้

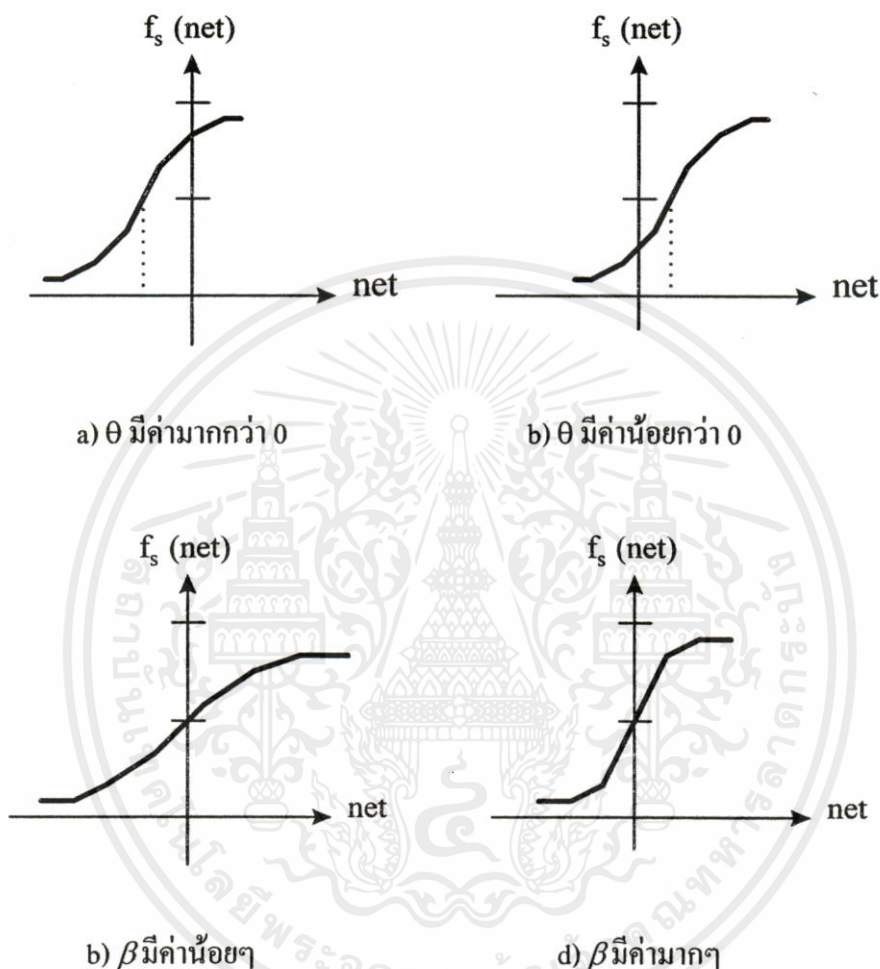
$$\text{out} = f(\text{net}) = \frac{2}{1 + e^{-(\beta \cdot \text{net})}} - 1 \dots\dots\dots (6)$$

$$= \tanh(0.5 \cdot (\beta \cdot \text{net})) \dots\dots\dots (7)$$

โดย  $\beta$  เป็นค่าสัญญาณเอาท์พุท ใช้ปรับรูปร่างของซิกมอยฟังก์ชัน

## รูปที่ 6

แสดงผลของ ไบอัส และค่าที่ใช้ปรับรูปร่างของซิกมอยฟังก์ชัน [1,6]



จากสมการที่ (1) และสมการที่ (4) ถ้า  $\theta$  มีค่ามากกว่าศูนย์ รูปซิกมอยฟังก์ชันจะเลื่อนไปทางซ้ายในแนวแกนอน และหาก  $\theta$  มีค่าน้อยกว่าศูนย์ รูปซิกมอยฟังก์ชันจะเลื่อนไปทางขวาในแนวแกนอนดังในรูปที่ 6(a) และรูปที่ 6(b) ตามลำดับ

ส่วนรูปที่ 6(c) และรูปที่ 6(d) ถ้า  $\beta$  มีค่าน้อยๆ จะทำให้รูปซิกมอยฟังก์ชันจะมีลักษณะคล้ายฟังก์ชันแบบเทรคโฮลลอคจิก และหาก  $\beta$  มีค่ามากๆ จะทำให้รูปซิกมอยฟังก์ชันมีลักษณะคล้ายรูปฟังก์ชันขั้นบันได

ไม่ว่ากรณีใดๆ ก็ตามมันให้คำตอบที่ถูกต้อง และต้องอ่านวิธีเขียนข้อสอบออกสอบที่อาจารย์ไปใช้  
 ในบทนี้ได้อธิบายความเข้าใจพื้นฐานของการอาร์คิฟิเซชันนิเวรอลเน็ตเวิร์ค ซึ่งเป็นศาสตร์ที่เกิดจากการเลียนแบบความสามารถของสมองมนุษย์ โดยเปรียบเทียบส่วนประกอบของเซลล์

ประสาทหนึ่งเซลล์กับแบบจำลองหน่วยประมวลผลทางคอมพิวเตอร์หนึ่งหน่วยหรือ โหนด สามารถเชื่อมต่อกับเซลล์ประสาทหลายๆเซลล์ ทำให้เกิดเป็นเน็ตเวิร์คของนิวรอลในรูปแบบต่างๆ ซึ่งจะนำเสนอ โครงสร้างของนิวรอลเน็ตเวิร์คพร้อมทฤษฎีในบทถัดไป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

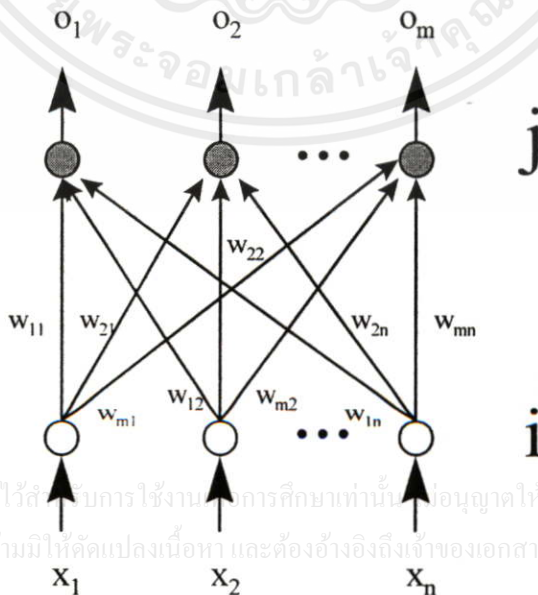
### บทที่ 3

#### นิเวรอลเน็ตเวิร์คแบบเลเยอร์เดียว

จากที่กล่าวไปในตอนต้นนั้นเป็นเพียงขั้นตอนหรือวิธีการทำงานของนิเวรอลเพียงโหนดเดียวซึ่งเทียบได้กับเซลล์เพียงเซลล์เดียว หากประกอบกันเป็นหลายเซลล์ จะต้องให้นิเวรอลแต่ละตัวเชื่อมต่อกันทำให้เกิดเป็นเน็ตเวิร์คที่มีลักษณะเป็นชั้นๆ เรียกว่า เลเยอร์ (Layer) ในหนึ่งเลเยอร์สามารถมีจำนวนนิเวรอลได้มากกว่า 1 นิเวรอลดังแสดงในรูปที่ 7 ภายในเลเยอร์เดียวกันจะไม่มี การเชื่อมต่อกัน และนิเวรอลที่มีเลเยอร์สูงกว่าจะรับค่ามาจากเอาต์พุตของเลเยอร์ที่อยู่ต่ำกว่าเพื่อเป็น อินพุตของตัวเอง จากลักษณะดังกล่าวจะเห็นว่า คุณสมบัติของนิเวรอลทุกตัวที่อยู่ในเลเยอร์ที่ต่ำกว่า จะถูกถ่ายทอดไปยังนิเวรอลในชั้นถัดไป

รูปที่ 7

นิเวรอลเน็ตเวิร์คแบบเลเยอร์เดียว

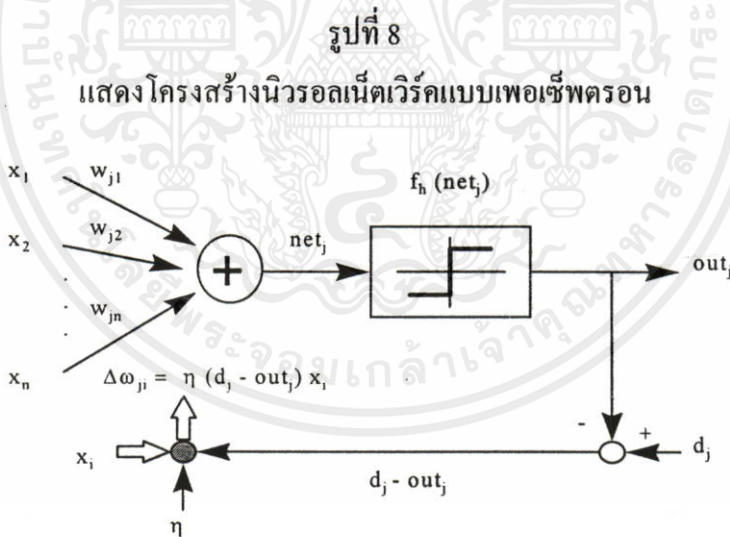


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งาน การศึกษาเท่านั้น อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 7 จะแสดงลักษณะของนิวรอนที่ประกอบกันเป็นนิวรอนเน็ตเวิร์คแบบเลเยอร์เดียวโดยในเลเยอร์แรก หรืออินพุตเลเยอร์โดยแท้จริงแล้วไม่ใช้นิวรอน เพราะไม่มีการคำนวณอยู่ภายใน แต่จะแสดงลักษณะการกระจายของค่าอินพุตให้กับนิวรอนในชั้นถัดไปเท่านั้น

### 3.1 กฎการเรียนรู้ของเพอเซ็ปตรอน

เน็ตเวิร์คแบบเลเยอร์เดียวจะมีโครงสร้างไม่ซับซ้อน สามารถเรียนรู้ที่จะจดจำและแยกแยะแพทเทิร์นอย่างง่ายได้ ทฤษฎีที่จะนำมาใช้กับเน็ตเวิร์คแบบเลเยอร์เดียวคือ กฎการเรียนรู้ของเพอเซ็ปตรอนแบบไม่ต่อเนื่อง (Discrete perceptron learning rule) ถูกค้นพบโดย Rosenblatt 1958 [1] ดังในรูปที่ 8 เพอเซ็ปตรอนแบบไม่ต่อเนื่องใช้แอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตดังแสดงรูปที่ 3(a) โดยอาศัยการเรียนรู้แบบมีครูสอนหรือแบบมีการชี้นำ (Supervised Learning)



จากรูปที่ 8 ผลรวมของค่าอินพุตแพทเทิร์น ( $x_1, \dots, x_n$ ) ในเลเยอร์  $i$  คูณกับตัวเลขน้ำหนัก ( $w_{ji}$ ) จะถูกส่งผ่านไปยังแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิต ได้ค่าเอาต์พุตที่ได้จากการคำนวณ ( $out_j$ ) ในเลเยอร์  $j$  ซึ่งค่าเอาต์พุตจะมีค่าเป็น 1 หรือ 0 ค่าใดค่าหนึ่งเท่านั้น จากนั้นค่าเอาต์พุตจะถูกนำไปคำนวณเพื่อปรับค่าตัวเลขน้ำหนักอีกครั้งหนึ่ง

$$net_j = \sum_{i=1}^n x_i \cdot w_{ji} + \theta_j \dots\dots\dots (8)$$

$$out_j = f_h(net_j) \dots\dots\dots (9)$$

การปรับค่าตัวเลขน้ำหนัก จะได้ว่า

$$\Delta w_{ji} = \eta (d_j - out_j) x_i \dots\dots\dots (10)$$

$$= \eta [d_j - f_h(\sum_i x_i \cdot w_{ji} + \theta_j)] \dots\dots\dots (11)$$

และกำหนดตัวแปรค่าผิดพลาดการเรียนรู้ err ดังนี้

$$err = (d_j - out_j) \dots\dots\dots (12)$$

- โดย
- $f_h$  เป็นแอคติเวชันฟังก์ชันแบบฮาร์ดลิมิต
  - $\theta_j$  เป็นค่าเทรคโไฮรลสำหรับเลเยอร์  $j$
  - $\Delta w_{ji}$  ค่าตัวเลขน้ำหนักระหว่างเลเยอร์  $i$  และ  $j$
  - $\eta$  ค่าคงที่การเรียนรู้
  - $d_j$  ค่าเอาต์พุตจริงที่ต้องการในเลเยอร์  $j$

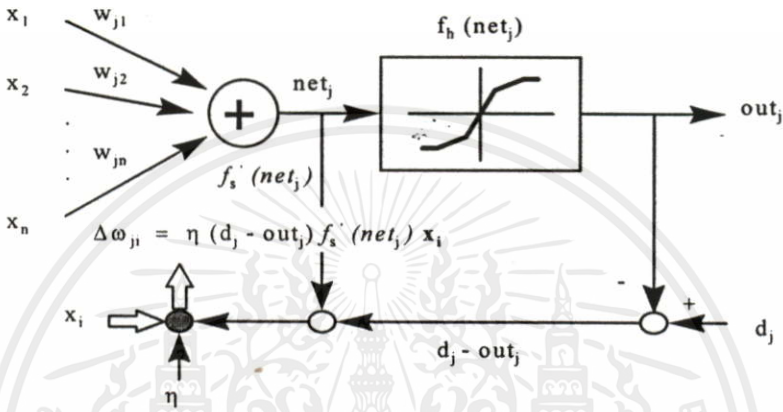
สังเกต ค่า  $\Delta w_{ji}$  จะเท่ากับ 0 ต่อเมื่อค่า  $out_j$  มีค่าเข้าใกล้ค่า  $d_j$

### 3.2 กฎการเรียนรู้ของเดลต้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้  
 กฎการเรียนรู้ของเดลต้า (Delta Learning rule) ถูกค้นพบโดย McClelland และ Rumelhart (1986) [1] เมื่อไม่นานมานี้ โดยทฤษฎีนี้อาจเรียกอีกอย่างว่า กฎการเรียนรู้แบบต่อเนื่องของเพอเซพ

ตรอน (Continues perceptron learning rule) โดยจะใช้แอกติเวชันฟังก์ชันแบบต่อเนื่องของซิกมอย ดังแสดงในรูปที่ 3(c) ซึ่งอาศัยการเรียนรู้แบบมีครูสอนดังแสดงในรูปที่ 9

รูปที่ 9  
แสดงโครงสร้างกฎการเรียนรู้ของเซลล์



ทฤษฎีการเรียนรู้สามารถพิสูจน์ได้จากเงื่อนไขของค่าผิดพลาดกำลังสองต่ำสุด E (Least square error) ระหว่างค่าเอาต์พุตจริงที่ต้องการ ( $d_j$ ) และเอาต์พุตที่ได้จากการคำนวณ ( $out_j$ ) สามารถแสดงเป็นสมการได้ดังนี้

$$E = \frac{1}{2} \cdot \sum_j (d_j - out_j)^2 \dots\dots\dots (13)$$

ซึ่งเท่ากับ

$$= \frac{1}{2} \cdot \sum_j [d_j \cdot f(net_j)]^2 \dots\dots\dots (14)$$

เพื่อให้ได้ค่าผิดพลาดต่ำสุดควรปรับเปลี่ยนตัวเลขน้ำหนักในทิศทางลบของ gradient descent ได้ว่า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Delta w_{ji}(t) = \eta \cdot (-\nabla E) \dots\dots\dots (15)$$

$$= \eta \cdot f'_s(\text{net}_j) \cdot (d_j - \text{out}_j) \cdot x_i \dots\dots\dots (16)$$

โดย  $\Delta w_{ji}(t)$  ค่าตัวเลขน้ำหนักที่เปลี่ยนไปตามเวลา  $t$  ระหว่างเลเยอร์  $i$  และ  $j$

$\eta$  ค่าคงที่การเรียนรู้

$f_s$  เป็นแอกติเวชันฟังก์ชันแบบซิกมอย

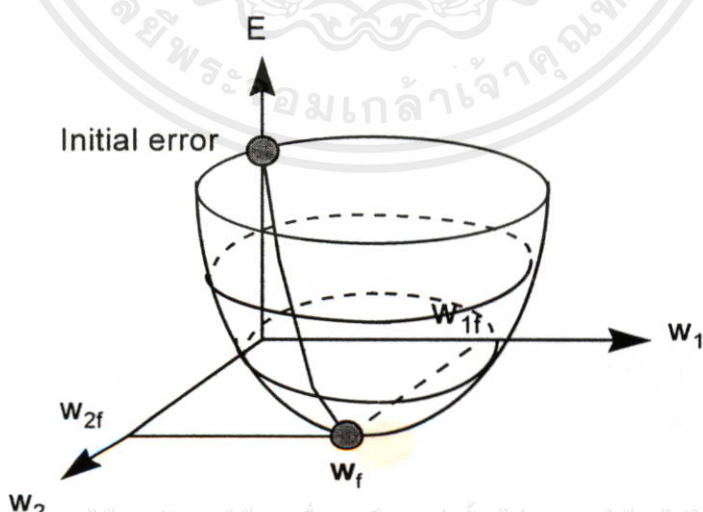
$f'_s$  การหาอนุพันธ์อันดับหนึ่งของแอกติเวชันฟังก์ชันแบบซิกมอย

$\nabla E$  gradient vector ของค่าผิดพลาด  $E$

แสดงกราฟที่เกิดขึ้นตามทฤษฎี gradient descent ในทิศทางลบ [1] ดังในรูปที่ 10 โดยกำหนดค่าตัวเลขน้ำหนักเริ่มต้น  $w_1$  และ  $w_2$  เป็น 0 พิจารณากราฟที่เกิดขึ้นระหว่างค่าผิดพลาด  $E$  และค่าตัวเลขน้ำหนัก  $w_1$  และ  $w_2$  จะเห็นว่าเมื่อเวลาผ่านไปค่าตัวเลขน้ำหนักนั้นจะถูกปรับเปลี่ยนไปจนกระทั่งได้ค่าผิดพลาดต่ำสุด  $E$  เท่ากับศูนย์ เมื่อตัวเลขน้ำหนักเท่ากับ  $w_f$

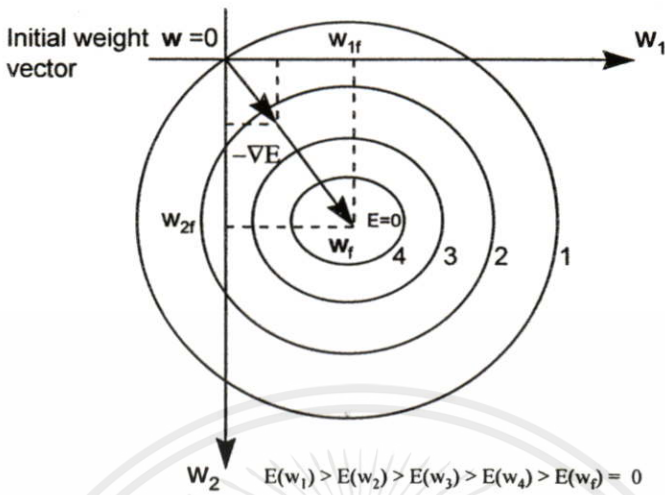
รูปที่ 10

รูปแสดงความสัมพันธ์ระหว่างค่าผิดพลาดต่ำสุดและค่าตัวเลขน้ำหนัก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะวิธีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

a) กราฟแสดงค่าผิดพลาดต่ำสุดขณะสอนนิเวรอลเน็ตเวิร์คเปลี่ยนไปตามเวลา  $t$



b) วงรอบแบบชั้นของค่าผิดพลาดต่ำสุดในทิศทาง gradient

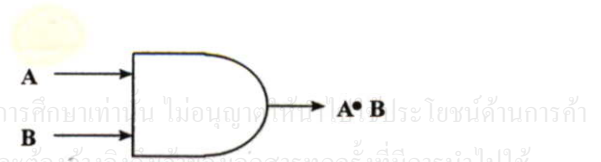
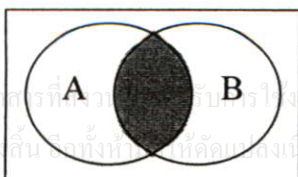
### 3.3 ทดสอบการเรียนรู้ของเพอร์เซพตรอนกับพีชคณิตบูลีนแบบแอนด์

ตัวอย่างนี้จะใช้กฎการเรียนรู้ของเพอร์เซพตรอนแบบไม่ต่อเนื่อง เพื่อทดสอบการเรียนรู้ อินพุตแพทเทิร์นของพีชคณิตบูลีน (Boolean Algebra) ตัวดำเนินการหรือโอเปอร์เรเตอร์แอนด์ (AND) [7] โดยมีอินพุต 2 ค่า และเอาต์พุต 1 ค่า ค่าเอาต์พุตจะมีค่าเป็น 1 เมื่ออินพุตทุกตัว มีค่าทางตรรกเป็น 1 กรณีอื่นๆ เอาต์พุตจะมีค่าเป็น 0 ดังแสดงในรูปที่ 11 และตารางที่ 1

รูปที่ 11

แสดงความสัมพันธ์ของพีชคณิตบูลีนแบบแอนด์

AND



เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้มีการนำออกจำหน่ายหรือใช้เพื่อการค้าโดยไม่ได้รับอนุญาตจากเจ้าของลิขสิทธิ์ หากมีข้อผิดพลาดประการใดขออภัยเป็นอย่างสูงและต้องขออภัยเป็นอย่างสูงแก่ผู้จัดทำเอกสารทุกครั้งที่มีการนำไปใช้

a) กราฟความสัมพันธ์

b) รูปแบบวงจรอิเล็กทรอนิกส์

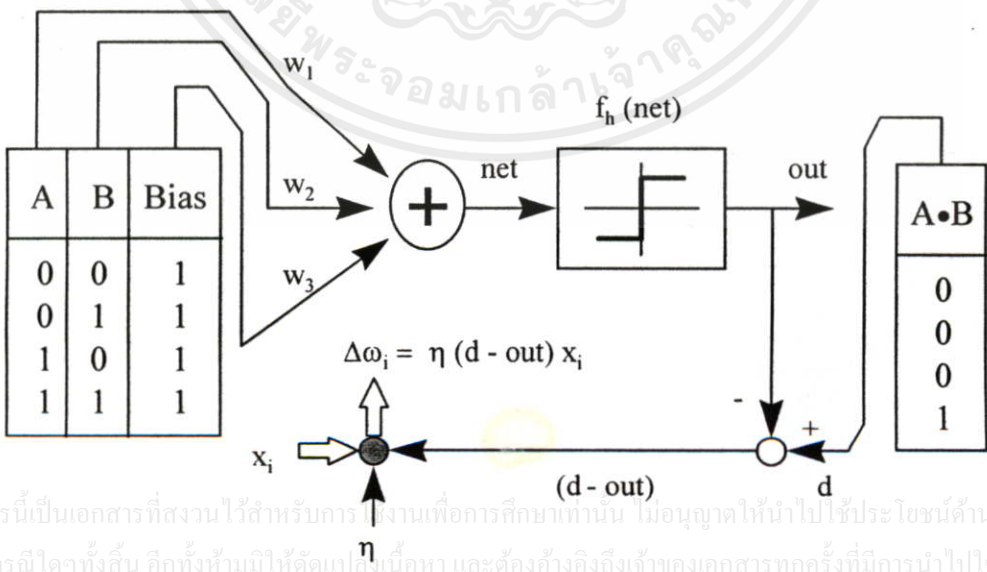
ตารางที่ 1  
แสดงตารางพีชคณิตบูลีนแบบแอนด์

A	B	A·B
0	0	0
0	1	0
1	0	0
1	1	1

จากข้อมูลอินพุตและเอาต์พุตที่ได้มา นำไปเขียนเป็นตัวแปรในแบบจำลองการทำงานของ  
นิวรอนเน็ตเวิร์คแบบเลเยอร์เดียวได้ดังแสดงในรูปที่ 12

รูปที่ 12

แสดงแบบจำลองการเรียนรู้ของเพอเซ็ปตรอนแบบ  
ไม่ต่อเนื่องกับพีชคณิตบูลีนแบบแอนด์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในอินพุตเลเยอร์ตัวแปร  $i$  มีค่าเริ่มต้นตั้งแต่ 1 ถึง 2 แต่ในเลเยอร์เอาต์พุต ตัวแปร  $j$  จะมีค่าเดียวคือ 1 เท่านั้น ดังนั้น เพื่อให้ดูง่ายจะไม่แสดงสัญลักษณ์  $j$  ในขณะนี้ และจากสมการที่ (2), (10), (11) และ (12) สามารถนำมาเขียนได้ดังนี้

$$\text{net} = \sum_{i=1}^2 x_i \cdot w_i + \theta = A w_1 + B w_2 + \text{Bias } w_3 \dots\dots\dots (17)$$

เมื่อ  $\theta = \text{Bias } w_3 \dots\dots\dots (18)$

$$\text{out} = f_h(\text{net}) \dots\dots\dots (19)$$

$$w_1(t+1) = w_1(t) + \eta \cdot (d - \text{out}) A \dots\dots\dots (20)$$

$$w_2(t+1) = w_2(t) + \eta \cdot (d - \text{out}) B \dots\dots\dots (21)$$

$$w_3(t+1) = w_3(t) + \eta \cdot (d - \text{out}) \text{Bias} \dots\dots\dots (22)$$

$$\text{err} = d - \text{out} \dots\dots\dots (23)$$

กำหนดค่าเริ่มต้น โดย

net เป็นผลรวมสุทธิที่เกิดจากค่าอินพุต ( $x$ ) คูณกับ ตัวเลขน้ำหนัก ( $w$ )

A, B เป็นค่าอินพุตแพทเทิร์นของนิเวรอล

$w_1, w_2, w_3$  เป็นค่าตัวเลขน้ำหนัก กำหนดค่าเริ่มต้นเป็น 0.2, 0.3, 0.4 ตามลำดับ

$\theta$  เป็นค่าเทรคโอรลของนิเวรอล

Bias เป็นค่าไบอัส กำหนดให้เป็นค่าคงที่เท่ากับ 1

$f(\text{net})$  เป็นแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิต กำหนดว่า ถ้า net มีค่าน้อยกว่าเท่ากับ 0 ให้เอาต์พุตมีค่าเท่ากับ 0 และถ้า net มีค่ามากกว่า 0 ให้เอาต์พุตมีค่าเท่ากับ 1 ดังในรูปที่ 4(a)

$\eta$  เป็นค่าคงที่การเรียนรู้ กำหนดให้เท่ากับ 0.8

- d เป็นค่าเอาท์พุทจริงที่ต้องการ (A·B)  
 err เป็นค่าผิดพลาดของระบบมีขึ้นเพื่อใช้วัดจุดสิ้นสุดการทำงานของโปรแกรม

จากสมการที่ (17) ถึง (23) สามารถนำมาใช้คำนวณกับแพทเทิร์นของโอเปอร์เรเตอร์แอนด์ตามรอบสัญญาณเวลา (t) ดังนี้

ขั้นตอนที่ 1 สำหรับแพทเทิร์นแรก A=0, B=0, d=0, Bias=1 และ  $\eta=0.8$

$$\text{net} = 0(0.2) + 0(0.3) + 1(0.4) = 0.4$$

$$\text{out} = f_h(0.4) = 0$$

$$w_1 = 0.2 + 0.8(0 - 0)(0) = 0.2$$

$$w_2 = 0.3 + 0.8(0 - 0)(0) = 0.3$$

$$w_3 = 0.4 + 0.8(0 - 0)(1) = 0.4$$

$$\text{err} = |0 - 0| = 0$$

สังเกต ค่าตัวเลขน้ำหนัก  $w_1, w_2$  และ  $w_3$  ยังไม่มีการเปลี่ยนแปลง

ขั้นตอนที่ 2 สำหรับแพทเทิร์นที่สอง A=0, B=1, d=0, Bias=1 และ  $\eta=0.8$

$$\text{net} = 0(0.2) + 1(0.3) + 1(0.4) = 0.7$$

$$\text{out} = f_h(0.7) = 1$$

$$w_1 = 0.2 + 0.8(0 - 1)(0) = 0.2$$

$$w_2 = 0.3 + 0.8(0 - 1)(1) = -0.5$$

$$w_3 = 0.4 + 0.8(0 - 1)(1) = -0.4$$

$$\text{err} = |0 - 1| = 1$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีก w3 ที่มี 0.4 เปลี่ยนเป็น -0.4 เนื่องจากถึงเข้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 3 สำหรับแพทเทิร์นที่สาม  $A=1, B=0, d=0, \text{Bias}=1$  และ  $\eta=0.8$

$$\text{net} = 1(0.2) + 0(-0.5) + 1(-0.4) = -0.2$$

$$\text{out} = f_h(-0.2) = 0$$

$$w_1 = 0.2 + 0.8(0 - 0)(1) = 0.2$$

$$w_2 = -0.5 + 0.8(0 - 0)(0) = -0.5$$

$$w_3 = 0.4 + 0.8(0 - 0)(1) = -0.4$$

$$\text{err} = |0 - 0| = 0$$

สังเกต ค่าตัวเลขน้ำหนัก  $w_1, w_2$  และ  $w_3$  เริ่มมีการเปลี่ยนแปลงแล้ว

ขั้นตอนที่ 4 สำหรับแพทเทิร์นที่สี่  $A=0, B=0, d=0, \text{Bias}=1$  และ  $\eta=0.8$

$$\text{net} = 1(0.2) + 1(-0.5) + 1(-0.4) = -0.7$$

$$\text{out} = f_h(-0.7) = 0$$

$$w_1 = 0.2 + 0.8(1 - 0)(1) = 1.0$$

$$w_2 = 0.5 + 0.8(1 - 0)(1) = 0.3$$

$$w_3 = -0.4 + 0.8(1 - 0)(1) = 0.4$$

$$\text{err} = |1 - 0| = 1$$

ขั้นตอนที่ 1 ถึง 4 ถือเป็นหนึ่งรอบการทำงาน ซึ่งจะเห็นว่าค่าตัวเลขน้ำหนักเปลี่ยนแปลง เนื่องจากค่าผิดพลาด  $\text{err}$  ที่เกิดขึ้นมีค่ามากกว่า 0 ดังนั้นจำเป็นต้องทำซ้ำขั้นตอนที่ 1 ถึง 4 จนกว่าค่าผิดพลาด  $\text{err}$  มีค่าเป็น 0 ทั้งหมด ซึ่งสามารถแสดงผลการคำนวณ ได้ดังแสดงในตารางที่ 2 หน้าด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ตารางที่ 2

แสดงการคำนวณการเรียนรู้แพทเทิร์นของพีชคณิตบูลีนแอนด์

รอบที่	ขั้นตอนที่	net	out	w <sub>1</sub>	w <sub>2</sub>	w <sub>3</sub>	err
1	1	0.40	0.00	0.20	0.30	0.40	0.000
1	2	0.70	1.00	0.20	-0.50	-0.40	1.000
1	3	-0.20	0.00	0.20	-0.50	-0.40	1.000
1	4	-0.70	0.00	1.00	0.30	0.40	1.000
2	1	0.40	0.00	1.00	0.30	0.40	0.000
2	2	0.70	1.00	1.00	-0.50	-0.40	1.000
2	3	0.60	1.00	0.20	-0.50	-1.20	1.000
2	4	-1.50	0.00	1.00	0.30	-0.40	1.000
3	1	-0.40	0.00	1.00	0.30	-0.40	0.000
3	2	-0.10	0.00	1.00	0.30	-0.40	0.000
3	3	0.60	1.00	0.20	0.30	-1.20	1.000
3	4	-0.70	0.00	1.00	1.10	-0.40	1.000
4	1	-0.40	0.00	1.00	1.10	-0.40	0.000
4	2	0.70	1.00	1.00	0.30	-1.20	1.000
4	3	-0.20	0.00	1.00	0.30	-1.20	1.000
4	4	0.10	0.00	1.80	1.10	-0.40	1.000
5	1	-0.40	0.00	1.80	1.10	-0.40	0.000
5	2	0.70	1.00	1.80	0.30	-1.20	1.000
5	3	0.60	1.00	1.00	0.30	-2.00	1.000
5	4	-0.70	0.00	1.80	1.10	-1.20	1.000
6	1	-1.20	0.00	1.80	1.10	-1.20	0.000
6	2	-0.10	0.00	1.80	1.10	-1.20	0.000
6	3	0.60	1.00	1.00	1.10	-2.00	1.000
6	4	0.10	0.00	1.80	1.90	-1.20	1.000

ตารางที่ 2 (ต่อ)

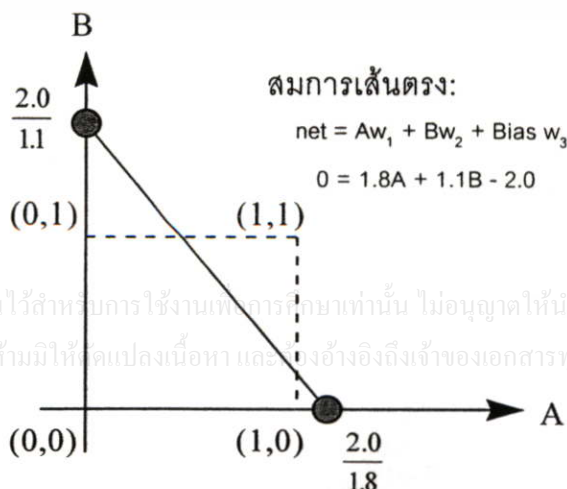
7	1	-1.20	0.00	1.80	1.90	-1.20	0.000
7	2	0.70	1.00	1.80	1.10	-2.00	1.000
7	3	-0.20	0.00	1.80	1.10	-2.00	1.000
7	4	0.90	1.00	1.80	1.10	-2.00	1.000
8	1	-2.00	0.00	1.80	1.10	-2.00	0.000
8	2	-0.90	0.00	1.80	1.10	-2.00	0.000
8	3	-0.20	0.00	1.80	1.10	-2.00	0.000
8	4	0.90	1.00	1.80	1.10	-2.00	0.000

จากการคำนวณจะเห็นว่าค่าผิดพลาด  $err$  ของทุกแพทเทิร์นในรอบการทำงานสุดท้าย จะมีค่าเป็นศูนย์ทั้งหมดนั้นหมายความว่าค่า  $w_1$ ,  $w_2$  และ  $w_3$  ถูกปรับจนได้ค่าที่ถูกต้องแล้ว ซึ่งจะใช้ค่าดังกล่าวเป็นฐานความรู้ของนิวรอนเน็ตเวิร์คต่อไป

ค่าตัวเลขน้ำหนัก  $w_1$ ,  $w_2$  และ  $w_3$  เท่ากับ 1.8, 1.1 และ -2.0 ตามลำดับ สามารถนำมาเขียนเป็นกราฟเส้นตรงได้ดังแสดงในรูปที่ 13

รูปที่ 13

แสดงการแยกจำพวกของกลุ่มแพทเทิร์นอินพุตของพีชคณิตบูลีนแอนด์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการรักษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 13 สมการเส้นตรง net หาได้จากการนำฐานข้อมูลค่าตัวเลขน้ำหนักที่ได้จากการเรียนรู้มาแทนค่าในสมการที่ (17) จะได้ว่า  $net = 1.8 A + 1.1 B - 2.0$  ซึ่งสามารถใช้แยกจำพวกของกลุ่มแพทเทิร์นได้อย่างถูกต้อง โดยกลุ่มข้อมูลอินพุตแพทเทิร์น (0,0), (0,1) และ (1,0) จะมีค่าเอาต์พุตแพทเทิร์นเหมือนกันคือ 0 ส่วนกลุ่มข้อมูลอินพุตแพทเทิร์น (1,1) จะมีค่าเอาต์พุตแพทเทิร์นเป็น 1

จากวิธีดังกล่าว สามารถนำไปประยุกต์ใช้กับพีชคณิตบูลีนออร์ (OR) และนอต (NOT) ได้ ในลักษณะเดียวกัน แต่ยังไม่สามารถใช้กับพีชคณิตบูลีนเอ็กซ์คลูซีฟออร์ (XOR) ได้ ซึ่งปัญหาดังกล่าวจะกล่าวถึงอีกครั้งในบทต่อไป

ตัวอย่างโปรแกรมการทำงานของการเรียนรู้แพทเทิร์นของพีชคณิตบูลีนแอนด์ เขียนขึ้นด้วยภาษา C สามารถดูได้จากภาคผนวก ก.

### 3.4 ข้อเสนอแนะและผลการวิเคราะห์

1. สำหรับตัวอย่างพีชคณิตบูลีนแอนด์ จะสังเกตว่า หากให้ค่า net เท่ากับ 0 จะได้เส้นตรงที่ใช้แบ่งกลุ่มแพทเทิร์นอินพุตของพีชคณิตบูลีนแอนด์ได้เป็นสองกลุ่ม เนื่องจากค่า 0 เป็นค่าที่ได้จากแอกติเวชันฟังก์ชันแบบเปลี่ยนทันทีนั่นเอง

2. ค่าเอาต์พุตที่ต้องการสามารถกำหนดเองได้ โดยอาศัยเงื่อนไขการกำหนดจุดตัดสินใจของแอกติเวชันฟังก์ชันแบบฮาร์ดลิมิตนั้น  $f_{in}(net)$  เช่น ถ้าค่า net มีค่าน้อยกว่าเท่ากับ 0 ให้ค่าเอาต์พุตเป็น -1 และหากมีค่ามากกว่า 0 ให้ค่าเอาต์พุตเป็น 1 เป็นต้น

3. ค่าตัวเลขน้ำหนักเริ่มต้น  $w_1, w_2$  และ  $w_3$  ถูกกำหนดเป็น 0.2, 0.3 และ 0.4 ตามลำดับนั้น เป็นค่าที่สามารถกำหนดได้เองเป็นค่าเท่าใดก็ได้ เช่น -1.0, 1.0, 2.0 ตามลำดับ เป็นต้น

4. ค่า  $\eta$  เป็นค่าคงที่การเรียนรู้ หากมีค่ามากจะสามารถเรียนรู้ได้เร็ว หากมีค่าน้อยจะสามารถเรียนรู้ได้ช้า แต่ค่าที่เหมาะสมควรมีค่าเป็น 1 สำหรับปัญหานี้ [3] การทุกครั้งที่มีการนำไปใช้

5. สมการ  $net_j = \sum_i x_i \cdot w_{ji} + \theta_j$  และสมการ  $\Delta w_{ji} = \eta (d_j - out_j) x_i$  มีค่า  $x_i$  เป็นตัวแปรตัวหนึ่ง ดังนั้นหากค่า  $x_i$  มีค่าเป็นศูนย์ จะทำให้ค่า  $net_j$  และค่า  $\Delta w_{ji}$  มีค่าเป็นศูนย์ไปด้วย ดังนั้นอาจเปลี่ยนแปลงเงื่อนไขของแอกติเวชันฟังก์ชันแบบฮาร์คลิมิต โดยกำหนดว่า ถ้าค่า  $net_j$  มีค่าน้อยกว่า 0 ให้ค่า  $out_j$  มีค่าเป็น -1 และหากค่า  $net_j$  มีค่ามากกว่า 0 ให้ค่า  $out_j$  มีค่าเป็น 1 ดังในรูปที่ 4(b) ซึ่งสามารถเขียนเป็นตารางได้ดังในตารางที่ 3

ตารางที่ 3

แสดงความสัมพันธ์ของอินพุตและเอาต์พุต  
ของพีชคณิตบูลีนแบบแอนด์

A	B	A·B
-1	-1	-1
-1	1	-1
1	-1	-1
1	1	1

6. เหตุผลที่ไม่มีตัวอย่างการทดสอบกฎการเรียนรู้ของเซลล์ค่ากับพีชคณิตบูลีนแอนด์ เนื่องจากกฎการเรียนรู้ของเซลล์ค่าใช้แอกติเวชันฟังก์ชันแบบต่อเนื่องของซิกมอยในการคำนวณ ทำให้จำนวนรอบการคำนวณมีมาก ทำให้ไม่สะดวกในการแสดงผล ณ ที่นี้

ในบทนี้เราได้เรียนรู้การทำงานของเน็ตเวิร์คแบบเลเยอร์เดียว ซึ่งเป็นโครงสร้างของเน็ตเวิร์คที่มีขนาดเล็กที่สุดและสามารถเรียนรู้ข้อมูลอินพุตแพทเทิร์นอย่างง่าย ทำให้เราเข้าพื้นฐานการทำงานของเน็ตเวิร์คได้ไม่ยาก และจากตัวอย่างการเรียนรู้แพทเทิร์นของพีชคณิตบูลีนแอนด์ทำให้ได้ฐานความรู้ของค่าตัวเลขน้ำหนักตรงกับความต้องการที่ฝึกสอนมาก่อนหน้านี้ ถัดไปเราจะศึกษาการทำงานของนิรอลเน็ตเวิร์คแบบหลายเลเยอร์ เพื่อประยุกต์ใช้กับข้อมูลของอินพุตแพทเทิร์นที่ซับซ้อนต่อไป

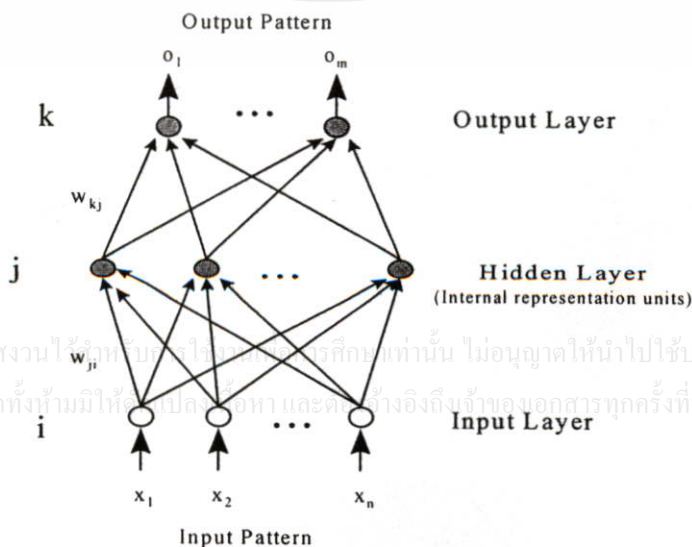
## บทที่ 4

### นิเวรอลเน็ตเวิร์คแบบหลายเลเยอร์

โครงสร้างของนิเวรอลเน็ตเวิร์คแบบหลายเลเยอร์ประกอบด้วยเลเยอร์ตั้งแต่ 2 เลเยอร์ขึ้นไป ในลักษณะของเน็ตเวิร์คที่มีขนาดใหญ่และซับซ้อนกว่าเน็ตเวิร์คแบบเลเยอร์เดียวมาก โดยอาจมองว่าเน็ตเวิร์คแบบหลายเลเยอร์เกิดจากการรวมกันของเน็ตเวิร์คแบบเลเยอร์เดียวก็ได้ เอาท์พุทแพทเทิร์นในเลเยอร์หนึ่งจะเป็นค่าอินพุทของเลเยอร์ถัดไป ซึ่งเลเยอร์ระหว่างอินพุทและเอาท์พุทถูกเรียกว่า ฮิดเดนเลเยอร์หรือเลเยอร์ภายใน (Internal layer) โดยค่าเอาท์พุทในเลเยอร์สุดท้ายจะถูกเปรียบเทียบกับค่าเอาท์พุทจริงที่ต้องการ (Desired output) จากผู้สอน (Teacher) ผลลัพธ์ของค่าผิดพลาดในขณะฝึกสอนสามารถนำไปใช้ปรับค่าตัวเลขน้ำหนักของเน็ตเวิร์คให้เปลี่ยนแปลงในทิศทางที่จะใช้แยกกลุ่มของแพทเทิร์น และโครงสร้างของเน็ตเวิร์คในลักษณะนี้ถูกเรียกว่า โครงข่ายแบบป้อนไปหน้า (Feed-forward network) เนื่องจากสามารถสร้างเป็นเน็ตเวิร์คแบบหลายเลเยอร์ได้ ซึ่งจะตรงกันข้ามกับโครงข่ายแบบป้อนกลับ (Feedback network) [1]

รูปที่ 14

โครงสร้างของนิเวรอลเน็ตเวิร์คแบบหลายเลเยอร์



เอกสารนี้เป็นเอกสารที่สงวนไว้ใช้เฉพาะกับโรงเรียนที่ซื้อสิทธิ์เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและทำซ้ำอย่างองถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 14 แสดงสถาปัตยกรรมพื้นฐานของเน็ตเวิร์คแบบหลายเลเยอร์ มีเลเยอร์ภายในเพียงเลเยอร์เดียว วงกลมในรูปแสดงโหนดหรือนิวรอลและถูกศรจะแสดงทิศทางการเชื่อมต่อของการเคลื่อนที่ของสัญญาณระหว่างนิวรอลในแต่ละเลเยอร์ ซึ่งจะถูกแทนด้วยค่าตัวเลขน้ำหนักระหว่างนิวรอล และ โครงสร้างแบบหลายเลเยอร์นี้เอง จะถูกนำไปใช้กับอัลกอริทึมของแบ็คพร็อพกาเกชัน

#### 4.1 แบ็คพร็อพกาเกชัน (Back Propagation)

หลักการแบ็คพร็อพกาเกชัน [1] ใช้สำหรับโครงสร้างเน็ตเวิร์คแบบหลายเลเยอร์ โดยอาศัยประสบการณ์ที่ได้จากความรู้ของการแปลงสัญญาณระหว่างอินพุตและเอาต์พุตภายในเน็ตเวิร์คแบบหลายเลเยอร์ด้วยรูปแบบการเรียนรู้แบบมีครูสอน เหมือนกรณีกฎการเรียนรู้ของเคลด้าที่ศึกษามาก่อนหน้านี้ ขั้นตอนการดำเนินการของแบ็คพร็อพกาเกชันประกอบด้วยสองส่วนหลัก ในส่วนการป้อนไปหน้า อินพุตแพทเทิร์นจะถูกคำนวณผ่านเน็ตเวิร์คและสร้างเอาต์พุตที่แน่นอน ในส่วนการป้อนกลับนั้นค่าความแตกต่างระหว่างค่าเอาต์พุตจริงที่ต้องการ ( $d_{pk}$ ) กับค่าเอาต์พุตที่ได้จากการคำนวณในเลเยอร์สุดท้าย ( $o_{pk}$ ) จะสร้างสัญญาณค่าผิดพลาดของสำหรับแพทเทิร์น ( $E_p$ ) และจะถูกแพร่ขยาย (Propagate) แบบย้อนกลับผ่านเน็ตเวิร์คเดิมเพื่อปรับค่าตัวเลขน้ำหนักของระบบ การทำงานจะเริ่มกระทำซ้ำในลักษณะเดิมอีกครั้ง จนกระทั่งค่าผิดพลาดของระบบโดยรวม ( $E$ ) น้อยกว่าค่าผิดพลาดของระบบที่ยอมรับได้ ( $E_n$ ) การเรียนรู้ของระบบจึงจะถือว่าสมบูรณ์

$$E_p = \frac{1}{2} \cdot \sum_K (d_{pk} - o_{pk})^2 \dots\dots\dots (24)$$

$$E = \frac{1}{P} \cdot \sum_p E_p \dots\dots\dots (25)$$

เอกสารนี้ เพื่อให้ค่าผิดพลาดของระบบลดลงต่ำสุด  $E$  จะต้องให้ผลต่างกำลังสองระหว่างค่าเอาต์พุตที่ได้จากการคำนวณ  $o_{pk}$  และค่าเอาต์พุตจริงที่ต้องการมีค่าเข้าใกล้ศูนย์ แต่เนื่องจากค่า  $d_{pk}$  ไม่มีการเปลี่ยนแปลง ดังนั้นต้องพยายามทำให้ค่า  $o_{pk}$  มีค่าเข้าใกล้  $d_{pk}$  มากที่สุด โดยการหาอนุพันธ์อันดับหนึ่งของค่าผิดพลาดสำหรับแพทเทิร์นกับค่าตัวเลขน้ำหนักที่เปลี่ยนไปในลักษณะเชิงเส้น

$$\Delta_p w_{kj} \propto -\frac{\partial E_p}{\partial w_{kj}} \dots\dots\dots (26)$$

เมื่อ  $w_{kj}(t+1) = w_{kj}(t) + \Delta_p w_{kj} \dots\dots\dots (27)$

โดย  $w_{kj}(t+1)$  ค่าตัวเลขน้ำหนักระหว่างเลเยอร์ k ไปเลเยอร์ j หลังจากปรับค่าแล้ว  
 $w_{kj}(t)$  ค่าตัวเลขน้ำหนักระหว่างเลเยอร์ k ไปเลเยอร์ j ก่อนการปรับค่า  
 $\Delta_p w_{kj}$  ค่าตัวเลขน้ำหนักที่เปลี่ยนไปได้จากการคำนวณ  
 t เป็นเวลาขณะใดๆ , t+1 เป็นเวลาถัดไป

จากสมการ เมื่อความชันของค่าความผิดพลาด ( $\partial E_p / \partial w_{kj}$ ) ลดลง จะมีผลทำให้ค่าตัวเลขน้ำหนักที่เปลี่ยนไป ( $\Delta_p w_{kj}$ ) ลดลงด้วย วิธีการนี้จะกระทำซ้ำไปซ้ำมาจนได้ค่าที่ต้องการ จากสมการที่ (26) สามารถทำเป็นสมการเชิงเส้นได้ดังนี้

$$\Delta_p w_{kj} = -\eta \cdot \frac{\partial E_p}{\partial w_{kj}} \dots\dots\dots (28)$$

โดย  $\eta$  เป็นค่าของอัตราการเรียนรู้ (Learning rate) ซึ่งมีค่าอยู่ระหว่าง 0.01 ถึง 1.00 [3] และจากสมการที่ (28) สามารถแสดงการหาอนุพันธ์อันดับหนึ่งของ  $E_p$  เทียบกับ  $w_{kj}$  ดังนี้

$$\frac{\partial E_p}{\partial w_{kj}} = f_k'(\text{net}_{pk}) (d_{pk} - o_{pk}) o_{pj} \dots\dots\dots (29)$$

กำหนดให้  $\delta_{pk} = f_k'(\text{net}_{pk}) (d_{pk} - o_{pk}) \dots\dots\dots (30)$

จะได้ว่า  $\Delta_p w_{kj} = \eta \cdot \delta_{pk} \cdot o_{pj} \dots\dots\dots (31)$

ไม่ว่ากรณีใดๆทั้งสิ้น อีกหนึ่งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากพิจารณาเพื่อหาค่าเปลี่ยนแปลงของตัวเลขน้ำหนัก  $\Delta_p w_{ji}$  ระหว่างฮิดเดนเลเยอร์กับอินพุตเลเยอร์ในลักษณะเดียวกัน จะได้ว่า

กำหนดให้ 
$$\delta_{pj} = f'_j(\text{net}_{pj}) \cdot \sum_k \delta_{pk} \cdot w_{kj} \dots\dots\dots (32)$$

จะได้ว่า 
$$\Delta_p w_{ji} = \eta \cdot \delta_{pj} \cdot o_{pi} \dots\dots\dots (33)$$

จากสมการที่ (31) และ (33) จะเห็นว่าค่าตัวเลขน้ำหนักจะเปลี่ยนไปตามสัญญาณค่าผิดพลาด (Error signal) ของ  $\delta_{pk}$  และ  $\delta_{pj}$  ตามลำดับ และเนื่องจาก  $f(\text{net})$  จะมีค่าเปลี่ยนแปลงไปตามสมการแอคติเวชัน หากเลือกใช้ซิกมอยฟังก์ชัน จากสมการที่ (4) และ (6) โดยกำหนดให้  $\beta$  เท่ากับ 1.0 จะได้ว่า

$$\text{net}_{pk} = \sum_j x_j \cdot w_{kj} + \theta_k \dots\dots\dots (34)$$

$$o_{pk} = f_k(\text{net}_{pk}) = \frac{1}{1 + e^{-(\text{net}_{pk})}} \dots\dots\dots (35)$$

จะได้ว่า 
$$f'_k(\text{net}_{pk}) = o_{pk} (1 - o_{pk}) \dots\dots\dots (36)$$

$$\text{net}_{pj} = \sum_i x_i \cdot w_{ji} + \theta_j \dots\dots\dots (37)$$

$$o_{pj} = f_j(\text{net}_{pj}) = \frac{1}{1 + e^{-(\text{net}_{pj})}} \dots\dots\dots (38)$$

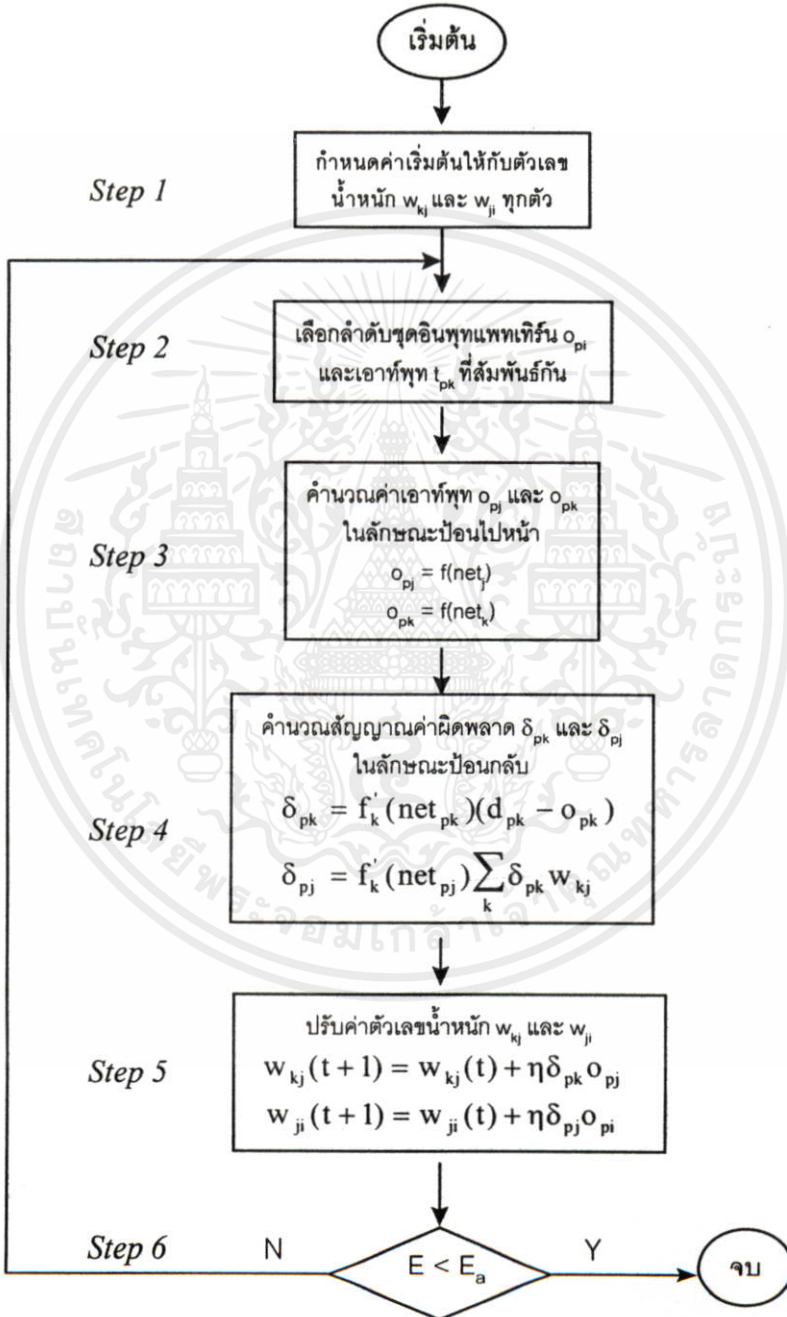
จะได้ว่า 
$$f'_j(\text{net}_{pj}) = o_{pj} (1 - o_{pj}) \dots\dots\dots (39)$$

หลักการพิสูจน์ทางคณิตศาสตร์ของทฤษฎีของมัลติเลเยอร์แบ็คพร้อพพาเกชัน สามารถดูได้จากภาคผนวก ข.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### แผนภูมิที่ 15

แสดงผังงานการทำงานของเบ็คพรีอบพาเกชันอัลกอริทึม [3]



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใด  
แผนภูมิที่ 15 แสดงผังงานการทำงานของเบ็คพรีอบพาเกชันอัลกอริทึม สำหรับโครงสร้างพื้นฐานของเน็ตเวิร์คแบบสองเลเยอร์ การเรียนรู้จะเริ่มการกำหนดค่าเริ่มต้นให้กับตัวแปรที่จำเป็นทุกตัวในขั้นตอนที่หนึ่ง จากนั้นเลือกลำดับอินพุต  $o_{pi}$  และเอาต์พุตแพทเทิร์น  $t_{pk}$  ที่สัมพันธ์กันใช้เป็น


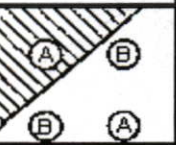
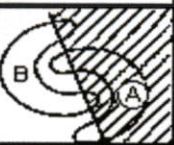
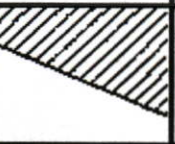
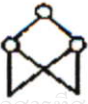
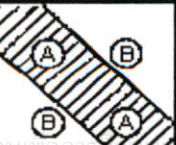

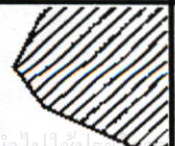
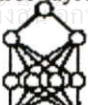
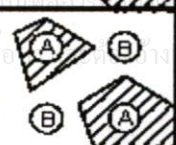
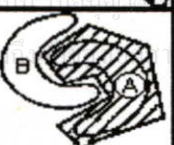
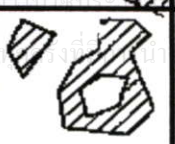
ข้อมูลเริ่มต้นของระบบในขั้นตอนที่สอง จำนวนค่าเอาต์พุต  $o_{pj}$  และ  $o_{pk}$  ตามลำดับในลักษณะป้อนสัญญาณไปหน้าในขั้นตอนที่สาม ค่าผลต่างที่เกิดจากเอาต์พุตที่ได้จากการคำนวณ  $o_{pk}$  และ  $t_{pk}$  จะถูกนำมาคำนวณเพื่อหาสัญญาณค่าผิดพลาด  $\delta_{pk}$  และ  $\delta_{pj}$  ตามลำดับในลักษณะป้อนกลับในขั้นตอนที่สี่ แล้วนำมาคำนวณต่อไปเพื่อใช้ปรับค่าตัวเลขน้ำหนัก  $w_{kj}$  และ  $w_{ji}$  ในขั้นตอนที่ห้า หลังจากผ่านการอบรมของชุดข้อมูลอินพุตแพทเทิร์น (เมื่อ  $p$  คือลำดับของชุดแพทเทิร์น) การเรียนรู้จะหยุดต่อเมื่อค่าผิดพลาดสำหรับทุกแพทเทิร์น  $E$  มีค่าต่ำกว่าค่าผิดพลาดโดยรวมที่ยอมรับได้  $E_a$  ในขั้นตอนที่หก

#### 4.1.1 ความสำคัญของฮิดเดนเลเยอร์และขอบเขตการตัดสินใจ (Decision regions)

โครงสร้างของเน็ตเวิร์กแบบหลายเลเยอร์สามารถฝึกสอนเพื่อแยกประเภทของแพทเทิร์นได้แล้วคำถามมีว่า จำนวนกึ่งเลเยอร์ที่ต้องการสำหรับการฝึกสอน จำนวนกึ่งโหนดที่ต้องการในแต่ละเลเยอร์ คำถามเหล่านี้พอจะอธิบายได้ดังแสดงในรูปที่ 16 ซึ่งจะแสดงขอบเขตการตัดสินใจที่ได้จากเลเยอร์เดี่ยว สองเลเยอร์ และสามเลเยอร์ของ Lippmann 1987 [5-6]

รูปที่ 16

แสดงชนิดของขอบเขตการตัดสินใจในของโครงสร้างแบบเลเยอร์เดี่ยว และหลายเลเยอร์กับตัวอย่างข้อมูลในคลาส A และ B

Structure	Type of Decision Regions	Exclusive-OR Problem	Classes with Mesned Regions	Most General Region Shapes
Single-layres 	Half plane bounded by hyperplane			
Two-layres 	Convex open or closed region			
Three-layres 	Arbitrary (complexity limited by number of nodes)			

นิเวรอลเน็ตเวิร์คแบบเลเยอร์เดียวจะแสดงคุณสมบัติการตัดสินใจด้วยรูปแบบของระนาบหลายมิติ (Hyperplane) ขึ้นกับจำนวน โหนดของอินพุตแพทเทิร์น และจะครอบคลุมขอบเขตของแพทเทิร์นที่ต้องการ เนื่องจากค่าเอาต์พุตของนิเวรอลเกิดจากผลรวมของอินพุตในลักษณะเชิงเส้น อินพุตสองค่าทำให้ได้ขอบเขตการตัดสินใจด้วยเส้นตรงในสองมิติ อินพุตสามค่าจะได้ขอบเขตการตัดสินใจด้วยระนาบในสามมิติ และอินพุต  $N$  ค่าจะได้ขอบเขตการตัดสินใจด้วยระนาบหลายมิติใน  $N$  มิติ

นิเวรอลเน็ตเวิร์คแบบสองเลเยอร์จะแสดงขอบเขตการตัดสินใจด้วยระนาบจากจำนวน โหนดของฮิดเดนเลเยอร์ นิเวรอลหนึ่งนิเวรอลสามารถสร้างขอบเขตการตัดสินใจด้วยเส้นตรงหนึ่งเส้น จากนั้นขอบเขตการตัดสินใจด้วยเส้นตรงทุกเส้นจะถูกรวบรวมไปยังนิเวรอลโหนดเอาต์พุต ทำให้เกิดขอบเขตการตัดสินใจด้วยเส้นตรงหลายเส้นบนระนาบ ดังเช่นปัญหาของเอ็กซ์คลูซีฟออร์

นิเวรอลเน็ตเวิร์คแบบสามเลเยอร์ จะแสดงขอบเขตการตัดสินใจด้วยอาณาเขตหรือกลุ่ม ขึ้นกับจำนวน โหนดของฮิดเดนเลเยอร์ที่อยู่ใกล้เอาต์พุตเลเยอร์ นิเวรอลหนึ่งโหนดจะทำหน้าที่แบ่งอาณาเขตการตัดสินใจออกเป็นสองส่วน และรูปแบบนี้เองอาจใช้แบ่งอาณาเขตของแพทเทิร์นที่ซับซ้อนได้เป็นอย่างดี

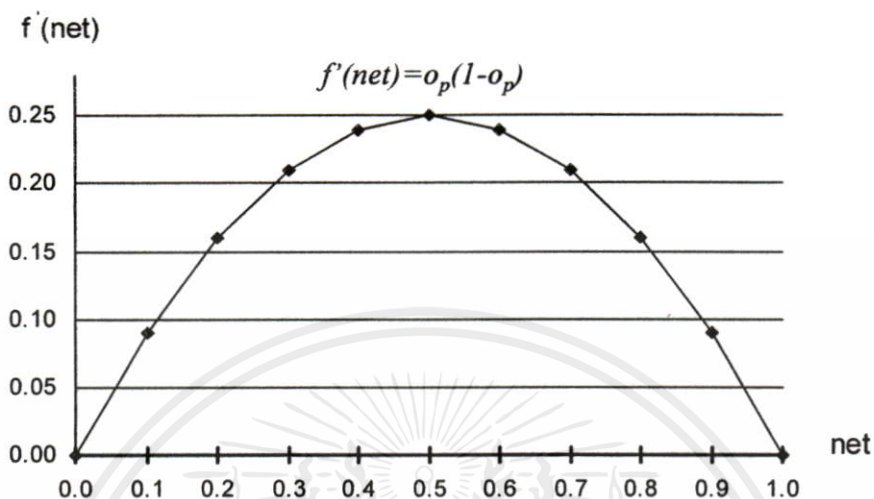
ในบางครั้ง เราสามารถทำงานที่ยากด้วยการเรียนรู้แบบง่าย โดยเพิ่มจำนวนเลเยอร์ภายในประมาณสามถึงสี่เลเยอร์ระหว่างอินพุตและเอาต์พุตเลเยอร์ อย่างไรก็ตาม Rumelhart, Hinton และ Winlliams (1986) [6] พบว่าการเพิ่มขึ้นแบบเอ็กซ์โพเนนเชียลของจำนวนฮิดเดนยูนิต จำเป็นต้องเพิ่มความเร็วในลักษณะเชิงเส้นของการคำนวณด้วย ดังปัญหาของเอ็กซ์คลูซีฟออร์ การค้นคว้าอื่นๆ มีการพบว่า การเพิ่มจำนวนฮิดเดนเลเยอร์จะลดอัตราของการเรียนรู้ของปัญหาการสุ่มตัวอย่างการจับคู่ของเวกเตอร์อีกด้วย

#### 4.1.2 ความเหมาะสมของการกำหนดค่าเอาต์พุตในทางปฏิบัติ

จากสมการที่ (31) และ (33) สังเกตว่า การปรับค่าตัวเลขน้ำหนัก จะสัมพันธ์กับค่า  $\delta_{pk}$  และ  $\delta_{pj}$  ตามลำดับ ซึ่งเกิดจากการหาอนุพันธ์ของฟังก์ชันแอ็กทิเวชันแบบซิกมอย ดังในสมการที่ (36) และ (39) ดังนั้นจะเห็นว่าค่า  $o_p (1 - o_p)$  จะมีค่าสูงสุดเมื่อค่า  $o_p$  เท่ากับ 0.5 ค่าตัวเลขน้ำหนักจะเปลี่ยนแปลงมากที่สุด และ  $o_p (1 - o_p)$  จะมีค่าต่ำสุดเมื่อ  $o_p$  เท่ากับ 0 หรือ 1 (เมื่อ  $0 < o_p < 1$ ) ดังแสดงในรูปแบบ

## รูปที่ 17

แสดงการหาอนุพันธ์อันดับหนึ่งของฟังก์ชันเอกติเวชัน



จากลักษณะของฟังก์ชันเอกติเวชันแบบซิกมอย จะเห็นว่าระบบไม่สามารถเข้าถึงค่าจริงสูงสุดและต่ำสุดของค่า 1 และ 0 ได้โดยปราศจากค่าตัวเลขน้ำหนักที่ใหญ่ไม่จำกัดขอบเขต (Infinity) ดังนั้นในทางปฏิบัติ หากกำหนดค่าผิดพลาดค่าสุดท้ายยอมรับได้ของระบบเป็นศูนย์ ค่าเอาต์พุตจริงที่ต้องการอาจใช้ตามแบบอย่างค่า 0.1 และ 0.9 แทนได้ [8] แม้ว่าจะงานที่เราทำจะใช้ค่าไบนารี 0, 1 ก็ตาม

#### 4.1.3 อัตราการเรียนรู้

อัลกอริทึมการเรียนรู้ของแบ็คพร็อบพาเกชันจะเปลี่ยนแปลงค่าตัวเลขน้ำหนักตามสัดส่วนของ  $\partial E_p / \partial w_p$  ในลักษณะ gradient descent ซึ่งทำงานแบบไม่จำกัดรอบในขณะที่เรียนรู้ ค่าคงที่ของอัตราการเรียนรู้ ( $\eta$ ) หากมีค่ามากจะเปลี่ยนแปลงค่าตัวเลขน้ำหนักได้เร็ว สำหรับในทางปฏิบัติ เสนอให้เลือกอัตราการเรียนรู้ที่มากที่สุดเท่าที่เป็นไปได้ โดยปราศจากการนำไปสู่การแกว่ง (Oscillate) บางครั้งอาจทดลองโดยเปลี่ยนแปลงไปเรื่อยๆ ไม่เช่นนั้นอาจจำเป็นต้องเพิ่มเทอมของโมเมนตัม (Momentum) ในสมการเคลื่อนที่ของแบ็คพร็อบพาเกชันเพื่อแก้ปัญหานี้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.1.4 การใช้ค่าโมเมนตัม

Rumelhart, Hinton และ Williams (1986) [8] อธิบายวิธีสำหรับพัฒนาเวลาในการฝึกสอนของแบ็คพร็อบพาเกชันอัลกอริทึมที่เรียกว่า โมเมนตัม ด้วยวิธีการเพิ่มตัวแปรการปรับค่าตัวเลขน้ำหนักซึ่งสัมพันธ์กับค่าตัวเลขน้ำหนักที่ใช้ก่อนหน้านี้

$$\Delta p_w(t) = \eta \cdot \delta_p \cdot o_p + \alpha \cdot \Delta p_w(t-1) \dots\dots\dots (40)$$

เมื่อ  $w(t+1) = w(t) + \Delta p_w(t) \dots\dots\dots (41)$

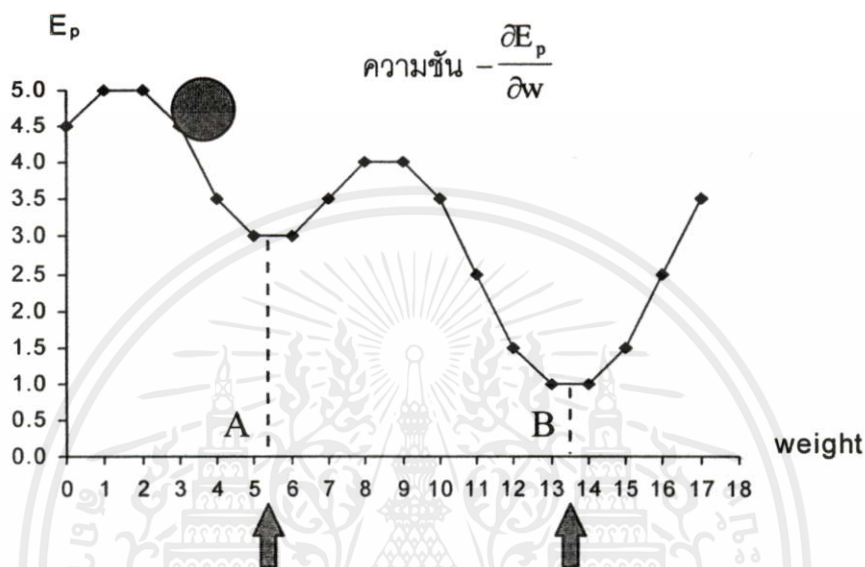
ที่ซึ่ง  $\alpha$  เป็นสัมประสิทธิ์โมเมนตัม (Momentum coefficient) [3,8] โดยทั่วไปมีค่าอยู่ประมาณ 0.9

การใช้วิธีของโมเมนตัมจะทำให้เนตเวิร์คของการฝึกสอนเข้าใกล้จุดต่ำสุดของพื้นผิวค่าผิดพลาด (Error surface) เร็วยิ่งขึ้น วิธีนี้จะทำงานได้ดีกับบางกรณี เนื่องจากการหาค่าที่เหมาะสมของโมเมนตัมนั้นจะขึ้นอยู่กับแต่ละกรณี

#### 4.1.5 Local minima ของแบ็คพร็อบพาเกชัน

ถึงแม้ว่าการประยุกต์ใช้แบ็คพร็อบพาเกชันอัลกอริทึมจะประสบความสำเร็จมากมาย แต่มันไม่ใช่ยารักษาโรคทุกชนิด สำหรับปัญหาที่ซับซ้อนบางอย่างอาจต้องใช้เวลาเป็นวัน หรือเป็นอาทิตย์ เพื่อให้เนตเวิร์คเรียนรู้และการอบรมนานๆ อาจจะมีหลุมเนื่องจากสาเหตุของตำแหน่งต่ำสุดท้องถิ่น (Local minimum) เนื่องจากแบ็คพร็อบพาเกชันใช้หลักการของ gradient descent นั้นหมายความว่า ความชันบนพื้นผิวค่าผิดพลาดจะเคลื่อนที่ในทิศทางที่ลดลง การปรับค่าตัวเลขน้ำหนักจะเข้าหาจุดต่ำสุด แต่เนื่องจากพื้นผิวค่าผิดพลาดของเนตเวิร์คที่ซับซ้อนมีการบิดเบี้ยว มีเนินเขา หุบเขา ในที่ว่างของหลายมิติ เมื่อมีหุบเขาต่ำๆ ที่อยู่ใกล้กว่าเนตเวิร์คจึงมีโอกาสติดกับดักของตำแหน่งต่ำสุดท้องถิ่น ดังแสดงในรูปที่ 18

รูปที่ 18  
แสดงทิศทางการเคลื่อนที่ของการปรับค่าตัวเลขน้ำหนัก  
กับปัญหาของ local minima



วิธีการทางสถิติจะช่วยหลีกเลี่ยงกับดักนี้ Wasserman (1988) [3] ได้เสนอวิธีการทางสถิติของ Cauchy machine กับ gradient descent ของแบ็คพร้อพพากชัน เพื่อให้ได้ระบบที่สามารถค้นหาจุดต่ำสุดทั่วไป (Global minimum) โดยใช้เวลาในการฝึกสอนมากกว่าปกติ ซึ่งสามารถอ่านเพิ่มเติมได้จาก [3 : หน้า 76-92]

#### 4.1.6 ความเสียหายจากค่าสมมาตร (Symmetry breaking)

ปัญหาอีกอย่างในอัลกอริทึมการเรียนรู้ของแบ็คพร้อพพากชันคือเรื่อง ความเสียหายจากค่าสมมาตร [8] หากมีการกำหนดให้ค่าตัวเลขน้ำหนักทุกตัวมีค่าเท่ากัน ระบบจะไม่สามารถเรียนรู้ นั่นเป็นเพราะว่า ทุกๆฮิดเดนยูนิตเชื่อมต่อโดยตรงไปยังอินพุตยูนิตในลักษณะแพร่ขยาย จะได้รับสัญญาณค่าผิดพลาดที่ได้จากการคำนวณเป็นค่าเดียวกัน ทำให้การปรับค่าตัวเลขน้ำหนักเปลี่ยนแปลง

ไปในทิศทางเดียวกันเสมอ ระบบจะเริ่มเข้าสู่ local minima ดังนั้นจึงควรหลีกเลี่ยง การแก้ปัญหาทำ โดยการกำหนดค่าเริ่มต้นของค่าตัวเลขน้ำหนักแบบสุ่มขนาดเล็กๆ ให้แก่ระบบ

ต่อไปเราจะศึกษาการทำงานทฤษฎีเบ็คหรือพหุภาคขนอัลกอริทึม กับปัญหาของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ เพื่อให้เห็นขั้นตอนในทางปฏิบัติและวิธีการนำไปใช้งาน โดยทั่วไป

#### 4.2 นิเวศน์เน็ตเวิร์คกับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์

ก่อนหน้านี เรามีความเข้าใจเกี่ยวกับตัวอย่างการทำงานของเน็ตเวิร์คแบบเลเยอร์เดียว ซึ่งโหนดของอินพุตยูนิตเชื่อมต่อโดยตรงกับเอาต์พุตยูนิต (เน็ตเวิร์คไม่มีฮิดเดนยูนิต หรือเลเยอร์ภายใน) ระบบสามารถเรียนรู้ชุดของข้อมูลอินพุตจำนวนน้อยๆ จากตัวอย่างง่ายๆ ได้ แต่ไม่ทุกกรณี เช่น ตัวอย่างปัญหาของเอ็กซ์คลูซีฟออร์ ที่แสดงในตารางที่ 4(a) เป็นต้น เราจะมาศึกษาต่อไปถึงหลักการในการแก้ปัญหานี้

ตารางที่ 4

แสดงกลุ่มข้อมูลอินพุตแพทเทิร์นของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์

Input pattern	Output pattern	Input pattern	Output pattern
0 0	0	0 0 1	0
0 1	1	0 1 1	1
1 0	1	1 0 1	1
1 1	1	1 1 1	1

a) อินพุตแพทเทิร์นแบบที่ 1

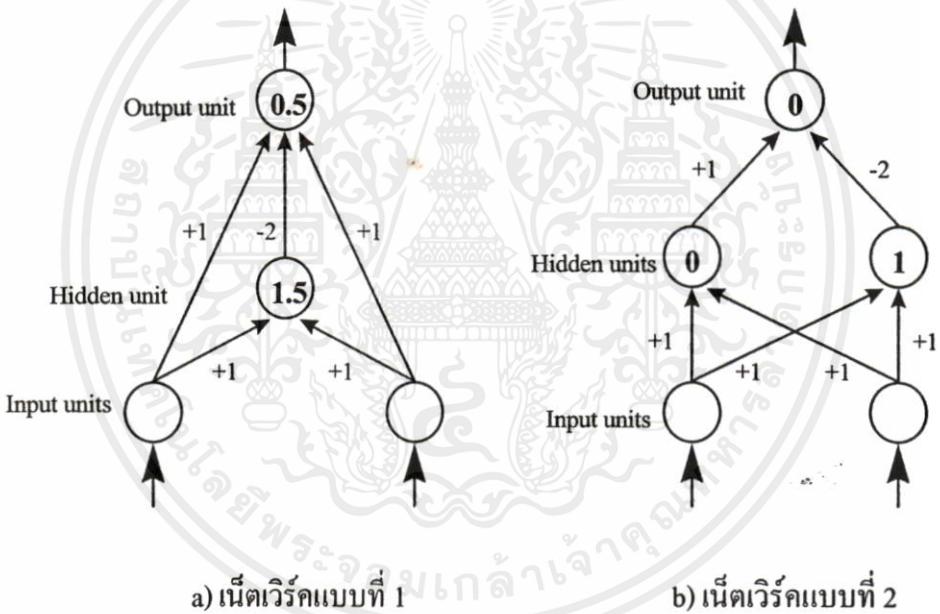
b) อินพุตแพทเทิร์นแบบที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 สิ่งที่น่าสนใจอีกอย่าง นอกจากการเพิ่มจำนวนเลเยอร์แล้ว การเพิ่มจำนวนโหนดให้กับอินพุตเลเยอร์ถือเป็นทางออกทางหนึ่งการการแก้ปัญหาของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ สำหรับนิเวศน์เน็ตเวิร์คแบบเลเยอร์เดียวได้ ดังแสดงในตารางที่ 4(b) แต่อาจใช้ไม่ได้กับกรณีอื่นๆ

Minsky และ Papert (1969) [8] มีการวิเคราะห์ความสามารถในการแปลงสัญญาณภายใต้เงื่อนไขต่างๆ หากจำนวนชุดของอินพุตแพทเทิร์น (จำนวนเงื่อนไข) ที่ต้องการมีมากขึ้น โครงสร้างของเน็ตเวิร์คแบบเลเยอร์เดียวไม่สามารถแก้ปัญหาได้ และยังชี้ให้เห็นว่าการเพิ่มฮิดเดนยูนิตดังแสดงในรูปที่ 14 จะสามารถสนับสนุนความต้องการของการแปลงสัญญาณจากอินพุตไปยังเอาต์พุตได้ และสำหรับปัญหาของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์นั้น การเพิ่มฮิดเดนยูนิตที่เพียงพอก็สามารถแก้ปัญหาของเอ็กซ์คลูซีฟออร์ได้อย่างถูกต้อง

รูปที่ 19

แสดงโครงสร้างนิรอลเน็ตเวิร์คสำหรับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์



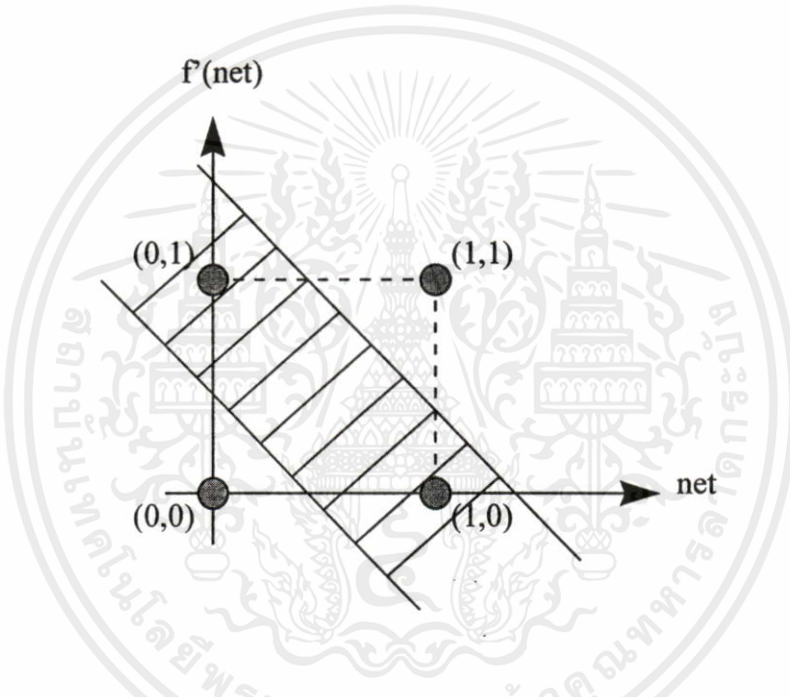
รูปที่ 19(a) แสดงโครงสร้างนิรอลเน็ตเวิร์คสำหรับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ ตัวเลขที่เขียนในวงกลมจะแสดงค่าเทรคโสรด +1.5 สำหรับเทรคโสรดของฮิดเดนยูนิต ทำให้แน่ใจว่าจะให้ค่า 0n เท่านั้นเมื่ออินพุตยูนิตทั้งสองมีค่า 0n ส่วนค่า 0.5 สำหรับเอาต์พุตยูนิต ทำให้แน่ใจว่าจะให้ค่า 0n เท่านั้นเมื่อผลรวมสุทธิในโหนดมีค่ามากกว่า 0.5 ค่าตัวเลขน้ำหนักของ -2.0 จากฮิดเดนยูนิตทำให้แน่ใจว่า เอาต์พุตยูนิตจะไม่กลายเป็น 0 เมื่อทั้งคู่ของอินพุตเป็น 0n ไปใช้ประโยชน์ด้านการคำนวณ

ส่วนรูปที่ 19(b) จะแสดงโครงสร้างนิรอลเน็ตเวิร์คโดยทั่วไป ซึ่งไม่มีการเชื่อมต่อโดยตรงระหว่างอินพุตและเอาต์พุตยูนิต และจะมีฮิดเดนยูนิตสองหน่วยทำหน้าที่ในการกำหนดขอบเขตการตัดสินใจ จากรูปฮิดเดนยูนิตแรกจะให้ค่า 0ff ค่อเมื่ออินพุตแพทเทิร์นเป็น (0,0) ทำให้เอาต์พุตยูนิต

เป็น Off และเมื่ออินพุตแพทเทิร์นเป็น (0,1) และ (1,0) ฮิดเดนยูนิตแรกจะให้ค่า On ทำให้เอาต์พุตยูนิตเป็น On ตามไปด้วย ถ้าอินพุตแพทเทิร์นเป็น (1,1) ฮิดเดนยูนิตทั้งสองจะเป็น On แต่ว่าค่าตัวเลขน้ำหนักที่มีค่ามากของฮิดเดนยูนิตที่สองมีผลทำให้เอาต์พุตยูนิตเป็น Off

รูปที่ 20

กราฟแสดงการแยกจำพวกของกลุ่มแพทเทิร์นของพีชคณิตบูลีน  
แบบเอ็กซ์คลูซีฟออร์สำหรับสองฮิดเดนยูนิต



ในรูปที่ 20 จะแสดงกราฟการตัดสินใจสำหรับปัญหาของเอ็กซ์คลูซีฟออร์ เพื่อแบ่งแยกจุดของแพทเทิร์นทั้งสี่จุด (0,0), (0,1), (1,0) และ (1,1) ออกเป็นสองคลาสหรือกลุ่ม จุดที่หนึ่งและสี่มีอยู่ในกลุ่มแรกมีค่าเอาต์พุตเป็น 0 ( $0 \text{ XOR } 0 = 0$  และ  $1 \text{ XOR } 1 = 0$ ) จุดที่สองและสามมีอยู่ในกลุ่มสองมีค่าเอาต์พุตเป็น 1 ( $0 \text{ XOR } 1 = 1$  และ  $1 \text{ XOR } 0 = 1$ ) ฮิดเดนยูนิตหน่วยแรกจะสามารถสร้างเส้นตรงเพื่อแบ่งอาณาบริเวณ (Region) เป็นสองส่วน สำหรับทุกจุดบนข้างหนึ่งของเส้นจะให้ผลเป็น 1 สำหรับทุกจุดบนอีกข้างหนึ่งจะให้ผลเป็น 0 ตำแหน่งการหมุนของเส้นตรงถูกกำหนดโดยค่าตัวเลขน้ำหนักบนอินพุตนิวโรลที่เชื่อมคือ สำหรับฮิดเดนยูนิตที่สองก็ให้ผลเช่นกันทำให้เราสามารถแบ่งอินพุตแพทเทิร์นออกเป็นสองคลาสได้นั่นเอง

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

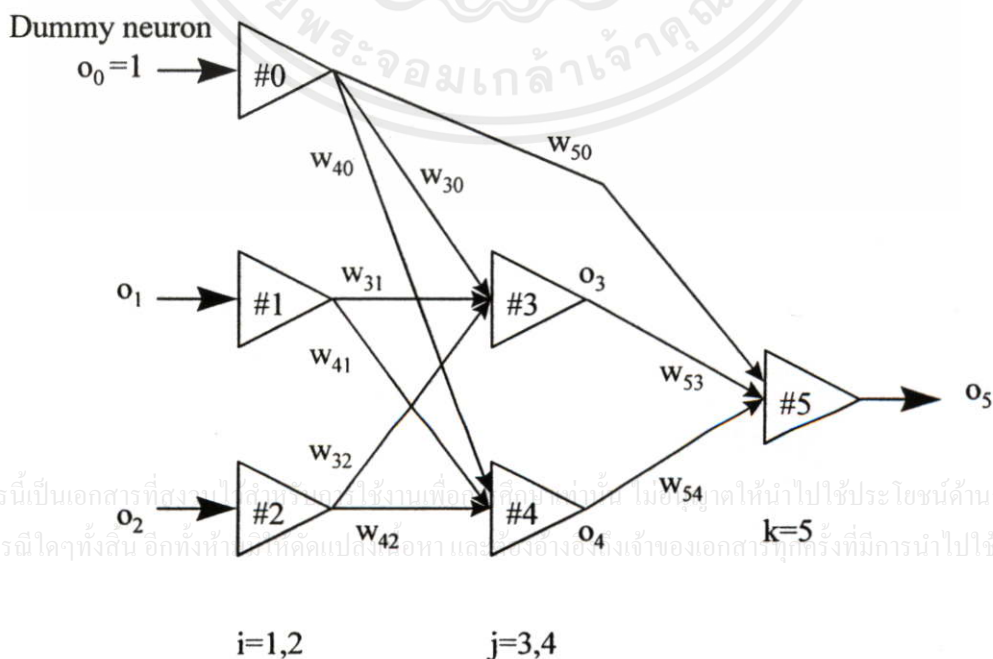
#### 4.2.1 ทดสอบการเรียนรู้ของแบ็คพร็อพพากันกับพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์

จากที่ได้กล่าวมาแล้วข้างต้น ปัญหาของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ไม่สามารถแก้  
ปัญหาได้โดยกฎการเรียนรู้ของเพอเซ็ปตรอนหรือกฎการเรียนรู้ของเคลด้า กับโครงสร้างเน็ตเวิร์ค  
แบบเลเยอร์เดียวเท่านั้น หากใช้โครงสร้างแบบมัลติเลเยอร์โดยอาศัยอัลกอริทึมของแบ็คพร็อพพากัน  
จะสามารถแก้ปัญหาก็ได้ จึงได้นำหลักการดังกล่าวมาทดสอบผลในทางปฏิบัติ เพื่อศึกษาและ  
วิเคราะห์การทำงาน

การทดสอบจะกำหนดรูปแบบโครงสร้างเน็ตเวิร์คดังในรูปที่ 21 กำหนดอินพุต 3 นิวรอน  
โหนด โดย 2 โหนดแรกคือ อินพุตแพทเทิร์น  $o_1$  และ  $o_2$  ของเอ็กซ์คลูซีฟออร์ ส่วนอีกโหนดที่เพิ่ม  
ขึ้นมาคือ dummy นิวรอนโหนด [8] ซึ่งแท้จริงก็คือ ไบอัส (Bias) หรือเทรคโสรลของนิวรอน ในระบบ  
กำหนดให้อินพุต  $o_0$  เป็นค่าคงที่เท่ากับ 1 ดังในรูปที่ 20 และกำหนด  $w_{30}$ ,  $w_{40}$  และ  $w_{50}$  ที่สอดคล้อง  
สำหรับนิวรอนโหนด #3, #4 และ #5

รูปที่ 21

แสดงโครงสร้างนิวรอนเน็ตเวิร์คของพีชคณิตบูลีน  
แบบเอ็กซ์คลูซีฟออร์ในทางปฏิบัติ



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามนำไปตีพิมพ์หรือเผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารผู้แต่งที่มีการนำไปใช้

การกำหนดค่าความสัมพันธ์ระหว่างอินพุตและเอาต์พุตแพทเทิร์นของพีชคณิตบูลีนแบบ เอ็กซ์คลูซีฟออร์สามารถแสดงในรูปแบบตารางได้ดังในตารางที่ 5

ตารางที่ 5

กำหนดค่าความสัมพันธ์อินพุตและเอาต์พุตแพทเทิร์น  
ของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์

A	B	A xor B
0	0	0
0	1	1
1	0	1
1	1	0

การเรียนรู้จะเริ่มหลังจากการกำหนดค่าเริ่มต้นแบบสุ่มให้กับ  $W$  ซึ่งเป็นค่าตัวเลขน้ำหนักแบบเวกเตอร์จากอินพุตเลขอร์ไปยังฮิดเดนเลเยอร์ และ  $V$  เป็นค่าตัวเลขน้ำหนักแบบเวกเตอร์จากฮิดเดนเลเยอร์ไปยังเอาต์พุตเลเยอร์

$$W = [w_{50} \quad w_{53} \quad w_{54}] = [0.082 \quad -0.106 \quad 0.086]$$

$$V = \begin{bmatrix} w_{30} & w_{31} & w_{32} \\ w_{40} & w_{41} & w_{42} \end{bmatrix} = \begin{bmatrix} 0.096 & -0.032 & 0.080 \\ -0.137 & 0.051 & -0.068 \end{bmatrix}$$

กำหนดค่าเริ่มต้นสำหรับอินพุตแพทเทิร์นแรก  $o_1=0$ ,  $o_2=0$  และ  $o_0=1$  ซึ่งจะถูกคำนวณไปพร้อมกันแบบป้อนไปหน้า เพื่อให้ได้ค่า  $o_3$ ,  $o_4$  และ  $o_5$  โดยกำหนดแอกคิเวชันฟังก์ชันดังสมการที่ (34), (35) และสมการที่ (37), (38) จะได้ว่า

สำหรับ โหนด #3

$$\text{net}_3 = w_{31}o_1 + w_{32}o_2 + w_{30}o_0 \dots\dots\dots (42)$$

เมื่อ  $\theta_3 = w_{30}o_0 \dots\dots\dots (43)$

$$o_3 = f(\text{net}_3) = \frac{1}{1 + e^{-(w_{31}o_1 + w_{32}o_2 + w_{30}o_0)}} \dots\dots\dots (44)$$

สำหรับ โหนด #4

$$\text{net}_4 = w_{41}o_1 + w_{42}o_2 + w_{40}o_0 \dots\dots\dots (45)$$

เมื่อ  $\theta_4 = w_{40}o_0 \dots\dots\dots (46)$

$$o_4 = f(\text{net}_4) = \frac{1}{1 + e^{-(w_{41}o_1 + w_{42}o_2 + w_{40}o_0)}} \dots\dots\dots (47)$$

สำหรับ โหนด #5

$$\text{net}_5 = w_{53}o_3 + w_{54}o_4 + w_{50}o_0 \dots\dots\dots (48)$$

เมื่อ  $\theta_5 = w_{50}o_0 \dots\dots\dots (49)$

$$o_5 = f(\text{net}_5) = \frac{1}{1 + e^{-(w_{53}o_3 + w_{54}o_4 + w_{50}o_0)}} \dots\dots\dots (50)$$

จากนั้นจะคำนวณหาสัญญาณผิดพลาด  $\delta_5$  จากสมการที่ (30) เมื่อ  $f(\text{net}_5) = o_5(1 - o_5)$

$$\delta_5 = [o_5(1 - o_5)](d_5 - o_5) \dots\dots\dots (51)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
คำนวณหาสัญญาณผิดพลาด  $\delta_3$  และ  $\delta_4$  จากสมการที่ (32) สำหรับฮิดเดนเลเยอร์โหนด #3 และฮิด  
เดนเลเยอร์โหนด #4

$$\delta_3 = f(\text{net}_3) \cdot \sum_{k=5}^5 \delta_k \cdot w_{k3} \dots\dots\dots (52)$$

$$\delta_4 = f(\text{net}_4) \cdot \sum_{k=5}^5 \delta_k \cdot w_{k4} \dots\dots\dots (53)$$

ใช้สมการที่ (39) เพื่อแทนค่าหา  $f(\text{net}_3)$  และ  $f(\text{net}_4)$

$$\delta_3 = [o_3 (1 - o_3)] \cdot \delta_5 \cdot w_{53} \dots\dots\dots (54)$$

$$\delta_4 = [o_4 (1 - o_4)] \cdot \delta_5 \cdot w_{54} \dots\dots\dots (55)$$

ตอนนี้ค่าตัวเลขน้ำหนักที่ถูกปรับ  $\Delta p_{wkj}$  และ  $\Delta p_{wji}$  จะสามารถคำนวณได้เมื่อกำหนดค่าคงที่อัตราเรียนรู้ ( $\eta$ ) โดยเริ่มคำนวณค่าตัวเลขน้ำหนักในเลเยอร์เอาต์พุท โหนด #5 ก่อน จากสมการที่ (31)

$$\Delta w_{50} = \eta \cdot \delta_5 \cdot o_0 = \eta \cdot \delta_5 \dots\dots\dots (56)$$

$$\Delta w_{53} = \eta \cdot \delta_5 \cdot o_3 \dots\dots\dots (57)$$

$$\Delta w_{54} = \eta \cdot \delta_5 \cdot o_4 \dots\dots\dots (58)$$

การปรับค่าตัวเลขน้ำหนักในฮิดเดนเลเยอร์จะสัมพันธ์กับนิวรอนใน โหนด #3 และ #4 โดยคำนวณจากสมการที่ (33)

$$\Delta w_{30} = \eta \cdot \delta_3 \cdot o_0 = \eta \cdot \delta_3 \dots\dots\dots (59)$$

เอกสารนี้เป็นเอกสาร  $\Delta w_{31} = \eta \cdot \delta_3 \cdot o_1$  ใช้แทนเพื่อการเชื่อมโยงเท่านั้น ไม่แนะนำให้ไปใช้ประโยชน์ด้วย (60)

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Delta w_{32} = \eta \cdot \delta_3 \cdot o_2 \dots\dots\dots (61)$$

และ

$$\Delta w_{40} = \eta \cdot \delta_4 \cdot o_0 = \eta \cdot \delta_4 \dots\dots\dots (62)$$

$$\Delta w_{41} = \eta \cdot \delta_4 \cdot o_1 \dots\dots\dots (63)$$

$$\Delta w_{42} = \eta \cdot \delta_4 \cdot o_2 \dots\dots\dots (64)$$

ที่ผ่านมาเป็นการคำนวณเพื่อปรับเปลี่ยนค่าตัวเลขน้ำหนัก  $\Delta p_{w_{kj}}$  และ  $\Delta p_{w_{ji}}$  เฉพาะแพทเทิร์นแรกเท่านั้น จะต้องมีการคำนวณเข้าไปซ้ำมา โดยใช้อินพุต  $o_1$  และ  $o_2$  สำหรับแพทเทิร์นสองสามและสี่ ตามลำดับ ทั้งหมดถือเป็นหนึ่งรอบ การคำนวณจะสิ้นสุดต่อเมื่อค่าผิดพลาดเฉลี่ยของระบบ E มีค่าน้อยกว่าหรือเท่ากับค่าผิดพลาดเฉลี่ยต่ำสุดที่สามารถยอมรับได้  $E_a$  หากกำหนดให้  $E_a$  เท่ากับ 0.001 จะได้จำนวนรอบการเรียนรู้ที่เกิดขึ้นทั้งหมด 3,567 รอบ และได้ผลลัพธ์สุดท้ายของค่าตัวเลขน้ำหนักแบบเมตริกซ์ W และ V เป็นฐานการเรียนรู้เพื่อนำไปใช้เป็นฐานข้อมูลต่อไป

$$W = \begin{bmatrix} w_{50} & w_{53} & w_{54} \end{bmatrix} = \begin{bmatrix} -3.714 & -8.288 & 8.048 \end{bmatrix}$$

$$V = \begin{bmatrix} w_{30} & w_{31} & w_{32} \\ w_{40} & w_{41} & w_{42} \end{bmatrix} = \begin{bmatrix} 2.314 & -6.007 & -6.058 \\ 5.993 & -4.061 & -4.066 \end{bmatrix}$$

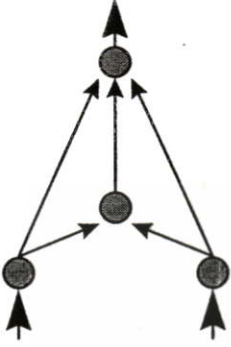
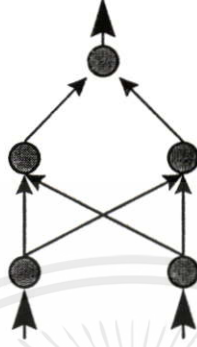
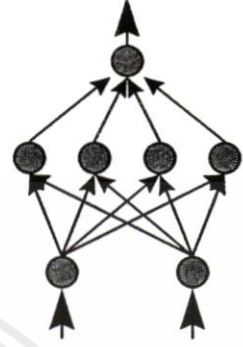
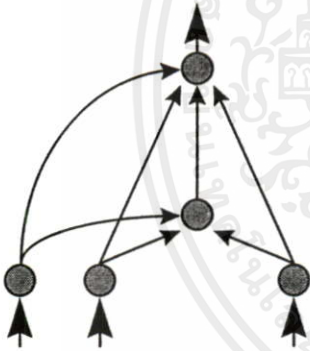
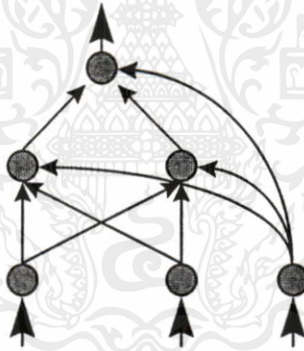
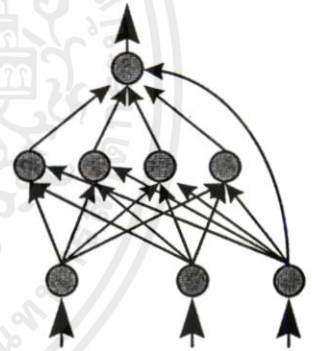
#### 4.2.2 ทดสอบการเรียนรู้สำหรับปัญหาของเอ็กซ์คลูซีฟเฟอร์เพิ่มเติม

การทดสอบที่ผ่านมาจะอาศัยโครงสร้างเน็ตเวิร์คสำหรับหนึ่งฮิดเดนเลเยอร์ สองฮิดเดนยูนิท หากตั้งคำถามว่าเน็ตเวิร์ครูปแบบอื่นที่ต่างออกไปจะให้ผลในการเรียนรู้เป็นอย่างไร กำหนดให้ใช้อัลกอริทึมของแบ็คพร็อพกาเกินเช่นเดิม

เนื่องจากเป็นเอกสารที่สงวนไว้สำหรับใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 22

แสดงโทโปโลยีของเน็ตเวิร์คในรูปแบบต่างๆ

a) เน็ตเวิร์คแบบที่ 1  
Nonconverseb) เน็ตเวิร์คแบบที่ 2  
Nonconversec) เน็ตเวิร์คแบบที่ 3  
Conversed) เน็ตเวิร์คแบบที่ 4  
Conversee) เน็ตเวิร์คแบบที่ 5  
Conversef) เน็ตเวิร์คแบบที่ 6  
Converse

จากรูปที่ 22 จะแสดงผลการทดสอบการเรียนรู้สำหรับปัญหาพีชคณิตบูลีนแบบเอ็กซ์คูลูซีฟออร์เพิ่มเติมพบว่า เน็ตเวิร์คแบบที่ 1 และ 2 จะเรียนรู้แบบไม่สิ้นสุด (nonconverse) ไม่ว่าจะใช้จำนวนรอบในการฝึกสอนมากเท่าใดก็ตาม เนื่องจากระบบเกิดการแกว่ง ส่วนแบบที่ 3, 4, 5 และ 6 จะให้ผลการเรียนรู้แบบสิ้นสุด ค่าผิดพลาดของระบบโดยรวมจะมีค่าเข้าใกล้ศูนย์

สังเกตว่า การเพิ่มฮิดเดนยูนิตให้เพียงพอจะมีผลให้การเรียนรู้ของระบบดีขึ้น และหากไม่มีไบอัสให้กับระบบ โอกาสที่ระบบจะเรียนรู้แบบสิ้นสุด (converse) เป็นไปได้ยากมาก

### 4.3 ข้อเสนอแนะและผลการวิเคราะห์

1. dummy นิรอรอลที่เพิ่มขึ้นมาแท้จริงแล้วคือ ไบอัสของระบบ ( $\theta_k$  และ  $\theta_j$ ) เมื่อกำหนด อินพุท  $\beta$  เป็นค่าคงที่เท่ากับ 1 คูได้จากสมการที่ (34) และ (37) ตามลำดับ จากการทดสอบหากไม่มี dummy นิรอรอลไหนจะไม่สามารถเรียนรู้ได้สำเร็จ ยกเว้นจะเพิ่มจำนวนนิรอรอลไหนในฮิดเดนเลเยอร์จาก 2 เป็น 4,5,6,...

2. หากเพิ่มจำนวนนิรอรอลไหนในฮิดเดนเลเยอร์มากขึ้นจะทำให้การเรียนรู้ดีขึ้น และจำนวนไหนที่ดีที่สุดสามารถหาได้จากการทดสอบ

3. การทดสอบปัญหาของเอ็กซ์คลูซีฟออร์โดยใช้โครงสร้างเน็ตเวิร์คดังแสดงในรูปที่ 19(a) และรูปที่ 19(b) จะให้ผลลัพธ์เหมือนกัน แต่ในวิทยานิพนธ์ฉบับนี้จะขอเสนอเฉพาะการทดสอบกับโครงสร้างเน็ตเวิร์คดังแสดงในรูปที่ 19(b) เท่านั้น

จากหลักการและทฤษฎีของเบ็คพรีอบพาเคชันอัลกอริทึม จะถูกนำไปใช้ศึกษากับตัวอย่างการแก้ปัญหของพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ ทำให้เข้าใจโครงสร้างของนิรอรอลเน็ตเวิร์คแบบหลายเลเยอร์ การกำหนดค่าเริ่มต้นที่จำเป็น การกำหนดจำนวนไหนในแต่ละเลเยอร์ ขั้นตอนของอัลกอริทึมที่ใช้ฝึกสอน และอุปสรรคในการนำไปใช้ เพื่อเป็นพื้นฐานสำหรับการประยุกต์ใช้งานในด้านต่างๆ รวมถึงงานทางด้านการรู้จำตัวอักษรด้วย ซึ่งจะกล่าวถึงต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### ทดสอบการเรียนรู้และแยกแยะแพทเทิร์น ตัวเลขตั้งแต่ 0 ถึง 9

จากความรู้เรื่องเอ็กซ์คลูซีฟออร์ ทำให้สามารถเข้าใจโครงสร้างพื้นฐานอย่างง่ายของมัลติเพล็กซ์เซอร์เพื่อเซพตรอนเน็ตเวิร์คได้เป็นอย่างดี เนื่องจาก โครงสร้างของเน็ตเวิร์คมีขนาดเล็ก จำนวนอินพุต โหนดและเอาต์พุต โหนดมีน้อย จึงสามารถศึกษาขั้นตอนการทำงานได้ง่ายขึ้น

ต่อไปจะนำเสนอหลักการและวิธีการทำงานของทฤษฎีแบ็คพร้อพาคชัน เพื่อใช้เป็นแนวทางในการศึกษาและพัฒนาทักงานทางด้านความรู้จำตัวอักษรด้วยแสงในอนาคต

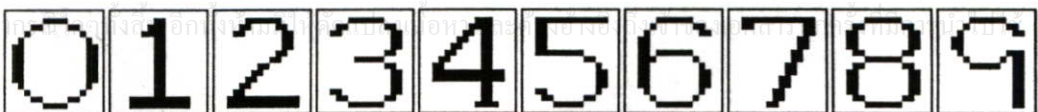
#### 5.1 การกำหนดค่าเริ่มต้นก่อนการคำนวณ

ก่อนที่จะแสดงลำดับและขั้นตอนการทำงานของโปรแกรม ควรที่จะกำหนด โครงสร้างเน็ตเวิร์ค และกำหนดรูปแบบของแพทเทิร์นที่จะใช้ทดสอบก่อน

1. ข้อกำหนดเบื้องต้นเพื่อให้คอมพิวเตอร์สามารถเรียนรู้ และจดจำแพทเทิร์นตัวอักษรตัวเลขแบบตัวพิมพ์ที่นำมาทดสอบ ได้มีการพัฒนาเขียนโปรแกรมขึ้น เพื่อทดสอบการจดจำตัวอักษรตัวเลขแบบตัวพิมพ์ตั้งแต่ "0" ถึง "9" (ทั้งหมด 10 แพทเทิร์น เมื่อ  $p=1..10$ ) โดยแพทเทิร์นตัวเลขแต่ละตัวมีขนาดความกว้างและความสูงเท่ากับ  $16 \times 16$  บิตแพทเทิร์น (256 บิตแพทเทิร์นต่อหนึ่งตัวอักษร)

#### รูปที่ 23

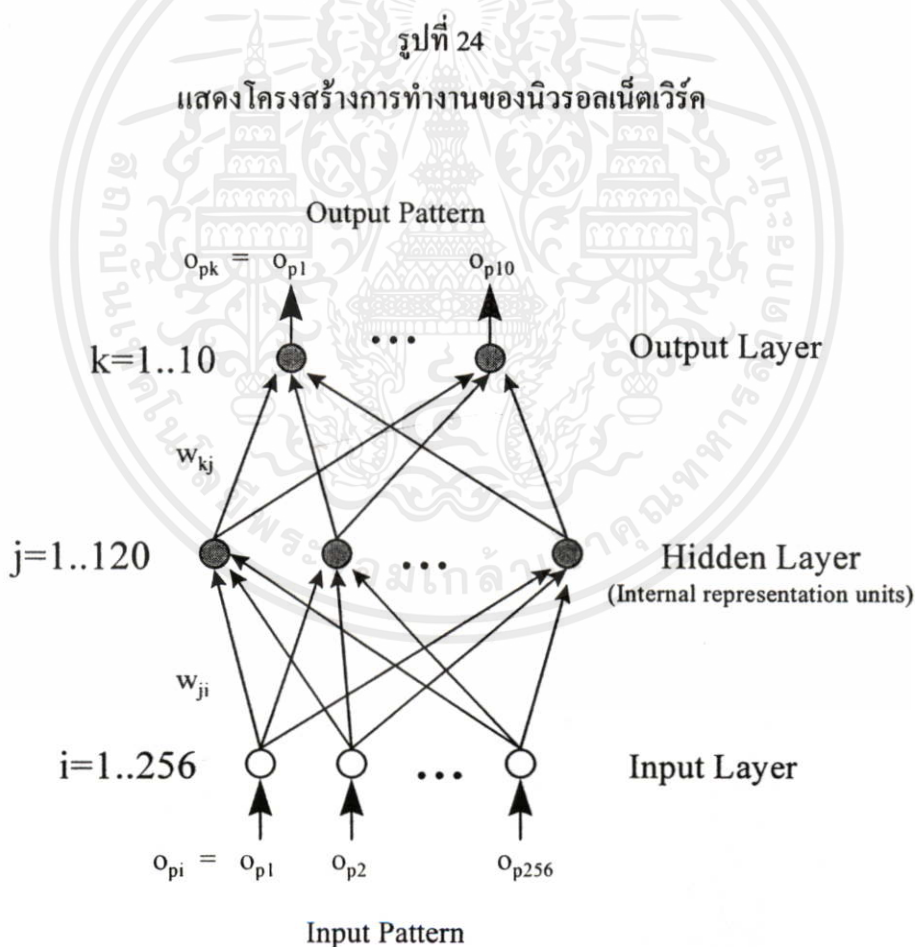
แสดงแพทเทิร์นตัวอักษรตัวเลข "0".."9"



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานในพิธีการศึกษาคณะเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่าในรูปแบบใดก็ตาม

แพทเทิร์นที่ปรากฏดังแสดงในรูปที่ 23 สามารถมองเห็นได้จากจอมอนิเตอร์ แสดงผลในลักษณะบิตแมพภาพ (ค่า 0 แสดงลักษณะปิดหรือ 1 แสดงลักษณะเปิด) ซึ่งจุดแต่ละจุดที่มองเห็นจะถูกนำมาเป็นอินพุต  $o_{pi}$  ของนิวรอลเน็ตเวิร์คและค่าเอาต์พุตที่ต้องการ  $t_{pk}$  ของนิวรอลเน็ตเวิร์ค โดย  $i=1..256$  และ  $k=1..10$  (เท่ากับจำนวนอินพุตแพทเทิร์นของตัวอักษรตัวเลข)

2. กำหนดจำนวนเลเยอร์ทั้งหมดในการทดสอบเพียง 3 เลเยอร์คือ อินพุตเลเยอร์ ฮิดเดนเลเยอร์ และเอาต์พุตเลเยอร์ โดยฮิดเดนเลเยอร์จะมีจำนวนนิวรอล 120 นิวรอล โหนด  $j=1..120$  ดังแสดงในรูปที่ 24



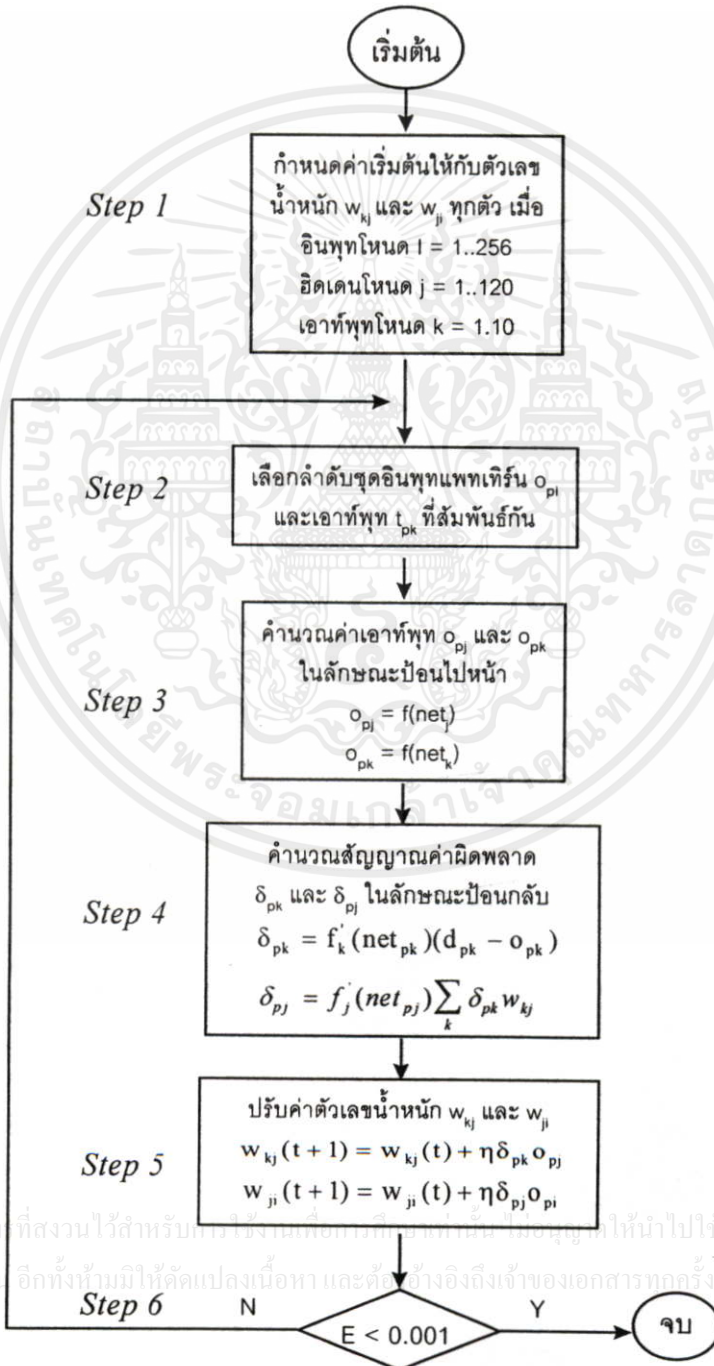
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งยัง **จำนวนอินพุตโหนด** และต้องอ้างอิงถึง **256** ของ **โหนด** ทุกครั้งที่มีการนำไปใช้

จำนวนฮิดเดนโหนด	120	โหนด
จำนวนเอาต์พุตโหนด	10	โหนด

## 5.2 ขั้นตอนการคำนวณและลำดับการทำงาน

## แผนภูมิที่ 25

## ขั้นตอนการทำงานของโปรแกรมในรูปแบบของผังงาน



แผนภูมิที่ 25 จะแสดงขั้นตอนการทำงานของโปรแกรมในรูปแบบของผังงานที่ละขั้นตอน และสามารถเขียนแสดงเป็นข้อๆ ได้ดังนี้

1. เริ่มแรกจะต้องให้ค่าเริ่มต้นแก่ค่าตัวเลขน้ำหนักในทุกเลเยอร์  $w_{ji}$  และ  $w_{kj}$  ในลักษณะสุ่มเป็นค่าน้อยๆ โดยจำกัดในช่วง 0.0-1.0 ทุกตัว เมื่อ  $i = 1..256$ ,  $j = 1..120$  และ  $k = 1..10$
2. จัดลำดับของอินพุตแพทเทิร์น  $o_{pi}$  ( $o_{p1}..o_{p256}$ ) และค่าเอาต์พุตแพทเทิร์น  $t_{pk}$  ( $t_{p1}..t_{p10}$ ) ในแต่ละชุดแพทเทิร์นให้สัมพันธ์กัน สามารถแสดงผลออกทางจอภาพได้ในลักษณะเป็นบิตแมพ เมื่อชุดของข้อมูลอินพุตแพทเทิร์น  $p = 1..10$  (ตัวอักษรตัวเลขตั้งแต่ศูนย์ถึงเก้า)

$o_{1i}$  คือแพทเทิร์นตัวเลข 0 ,  $t_{1k} = [1000000000]$

$o_{2i}$  คือแพทเทิร์นตัวเลข 1 ,  $t_{2k} = [0100000000]$

$o_{3i}$  คือแพทเทิร์นตัวเลข 2 ,  $t_{3k} = [0010000000]$

$o_{10i}$  คือแพทเทิร์นตัวเลข 9 ,  $t_{10k} = [0000000001]$

สอนเน็ตเวิร์คจนได้ค่าตัวเลขน้ำหนักที่ไม่มีการเปลี่ยนแปลง หรือค่าผิดพลาดเฉลี่ยของระบบ  $E$  มีค่าน้อยกว่าค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  กำหนดให้เท่ากับ 0.001

3. คำนวณหาค่าเอาต์พุต  $o_{pj}$  และ  $o_{pk}$  ในลักษณะไปข้างหน้าตามลำดับ โดยกำหนดให้  $\theta_j$  และ  $\theta_k$  เป็นไบอัสของระบบมีค่าเท่ากับ 0.05

$$net_{pj} = \sum_{i=1}^{256} x_i \cdot w_{ji} + \theta_j \dots\dots\dots (65)$$

$$o_{pj} = f(net_{pj}) = \frac{1}{1 + e^{-(net_{pj})}} \dots\dots\dots (66)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น ถือว่าห้ามมิที่จะเปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$net_{pk} = \sum_{j=1}^{120} x_j \cdot w_{kj} + \theta_k \dots\dots\dots (67)$$

$$o_{pk} = f(\text{net}_{pk}) = \frac{1}{1 + e^{-(\text{net}_{pk})}} \dots\dots\dots(68)$$

4. คำนวณสัญญาณค่าผิดพลาด  $\delta_{pk}$  และ  $\delta_{pj}$  ตามลำดับในลักษณะป้อนกลับ

$$\delta_{pk} = f'_k(\text{net}_{pk})(t_{pk} - o_{pk}) \dots\dots\dots(69)$$

$$\delta_{pj} = f'_j(\text{net}_{pj}) \sum_k \delta_{pk} w_{kj} \dots\dots\dots(70)$$

5. ปรับค่าตัวเลขน้ำหนักในแต่ละคู่เลเยอร์ในลักษณะป้อนกลับ โดยเริ่มจากเอาต์พุตเลเยอร์  $w_{kj}(t)$  เป็นค่าตัวเลขน้ำหนักจากฮิดเดนเลเยอร์  $j$  ไปยังเอาต์พุตเลเยอร์  $k$  ที่เวลา  $t$  ใดๆ และ  $t+1$  เป็นเวลาถัดไป กำหนดค่าอัตราการเรียนรู้  $\eta$  มีค่าเท่ากับ 0.5

$$\begin{aligned} w_{kj}(t+1) &= w_{kj}(t) + \eta \delta_{pk} o_{pj} \dots\dots\dots(71) \\ &= w_{kj}(t) + \eta [o_{pk} (1-o_{pk}) (t_{pk}-o_{pk})] o_{pj} \end{aligned}$$

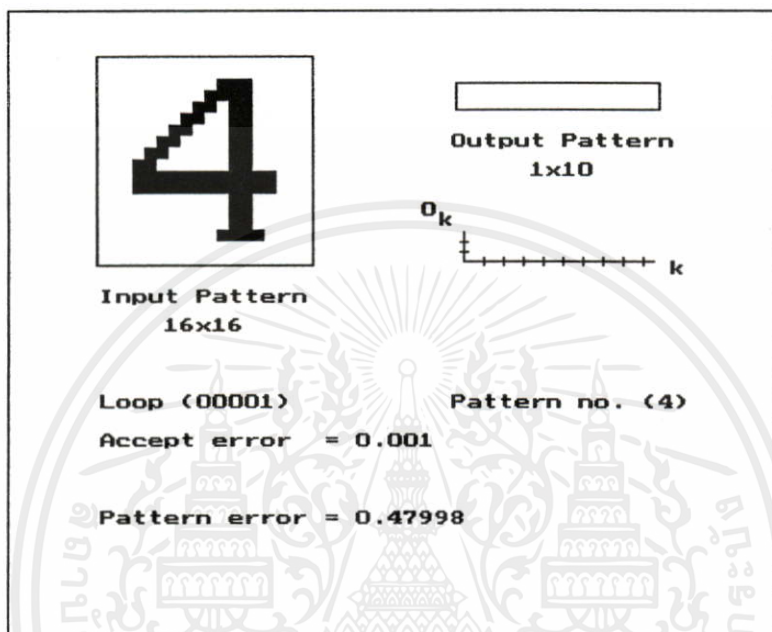
จากนั้นปรับค่าตัวเลขน้ำหนักคู่เลเยอร์ถัดไปในลักษณะเดียวกัน  $w_{ji}(t)$  เป็นค่าตัวเลขน้ำหนักจากอินพุตเลเยอร์  $i$  ไปยังฮิดเดนเลเยอร์  $j$  ที่เวลา  $t$  ใดๆ กำหนดค่าอัตราการเรียนรู้  $\eta$  มีค่าเท่ากับ 0.5

$$\begin{aligned} w_{ji}(t+1) &= w_{ji}(t) + \eta \delta_{pj} o_{pi} \dots\dots\dots(72) \\ &= w_{ji}(t) + \eta [o_{pj} (1-o_{pj}) \sum_{k=1}^{10} (\delta_{pk} \cdot w_{kj})] o_{pi} \end{aligned}$$

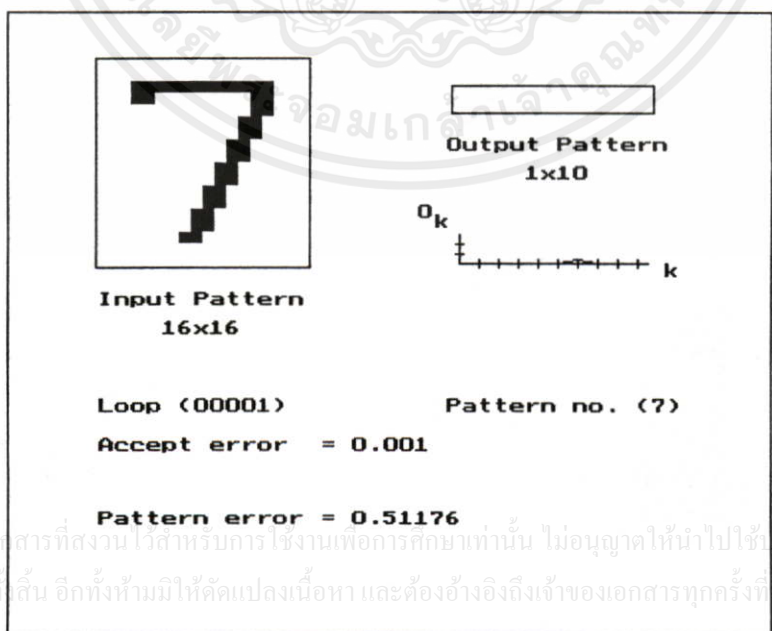
6. เปรียบเทียบค่าผิดพลาดโดยรวมของระบบที่ได้จากการคำนวณ  $E$  และค่าผิดพลาดของระบบที่ยอมรับได้  $E_a$  (เมื่อกำหนดให้  $E_a = 0.05$ ) หากมีค่าน้อยกว่าจะจบการเรียนรู้ของโปรแกรม เก็บค่าตัวเลขน้ำหนักที่ได้จากการเรียนรู้สำหรับใช้เป็นฐานข้อมูลของการรู้จำต่อไป หากมีค่ามากกว่า ให้กลับไปทำต่อในขั้นตอนที่ 2 โดยเปลี่ยนชุดของอินพุตแพทเทิร์น  $o_{pi}$  และเอาต์พุตแพทเทิร์น  $t_{pk}$  ที่สัมพันธ์กันในลำดับต่อไป

## รูปที่ 26

แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "0".."9"

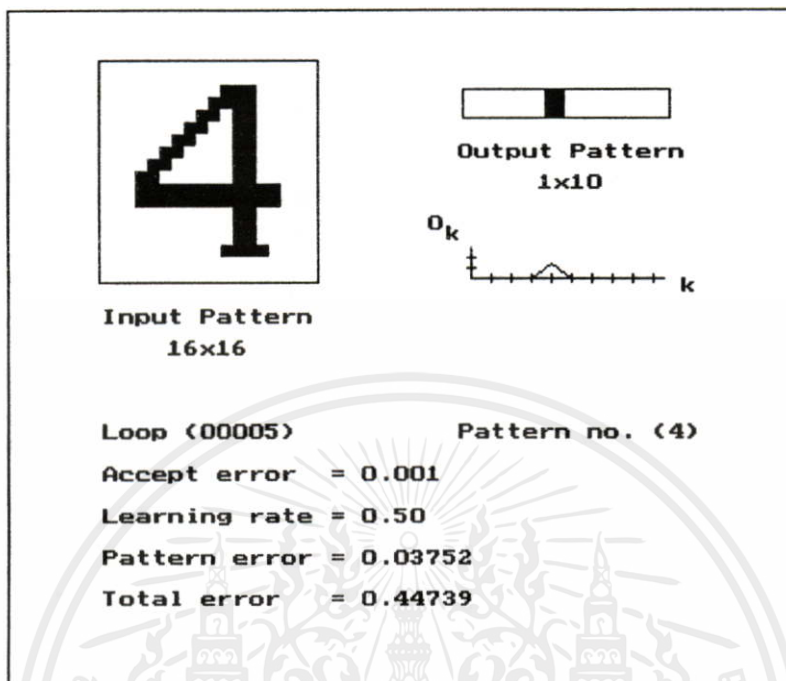


a) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "4" ในรอบที่ 1

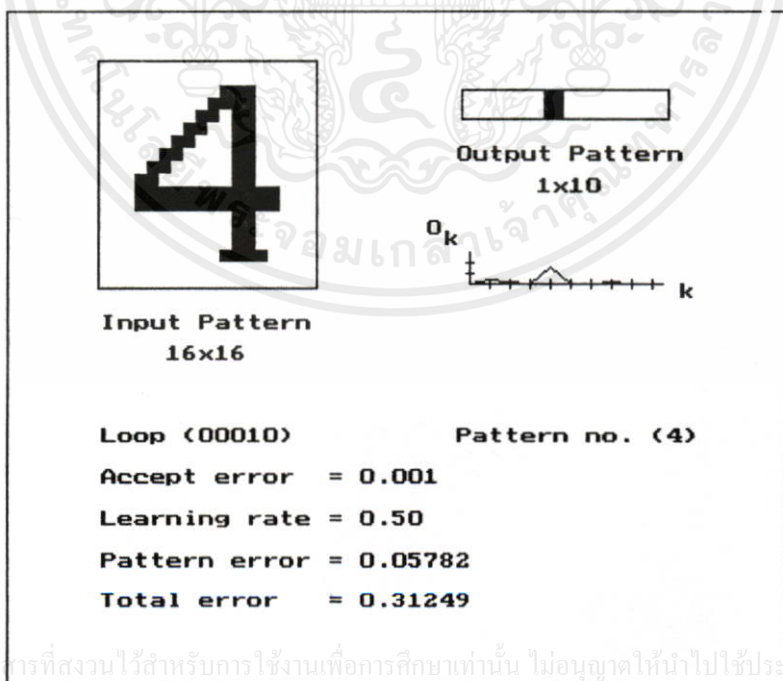


b) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "7" ในรอบที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ ระบุชื่อผู้แต่ง  
ไม่ว่ากรณีใดๆก็ตาม อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

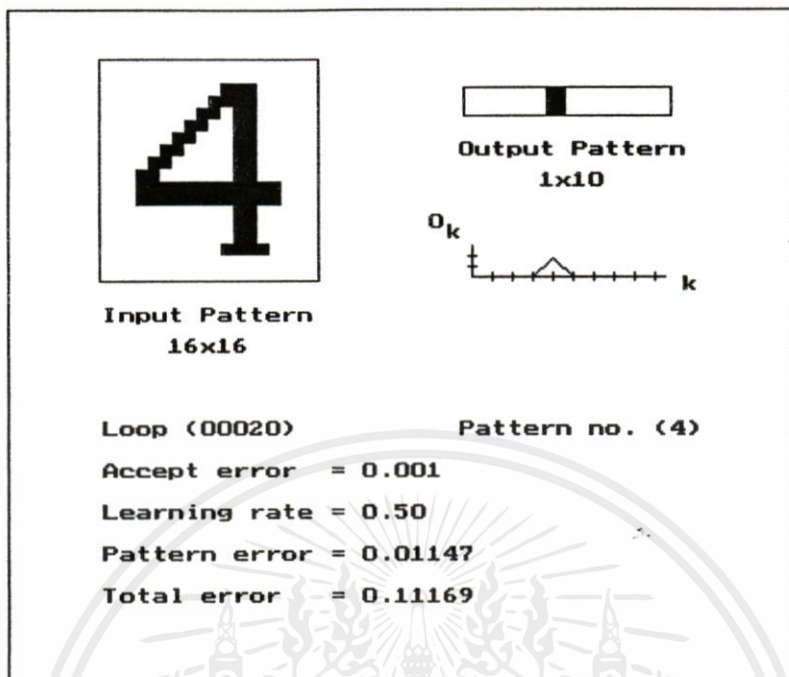


c) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "4" ในรอบที่ 5

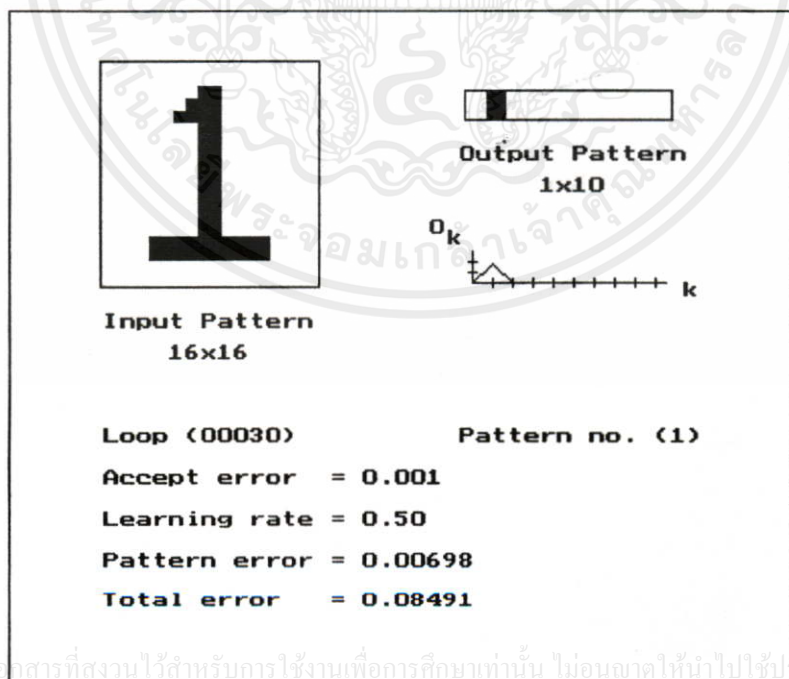


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

d) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "4" ในรอบที่ 10

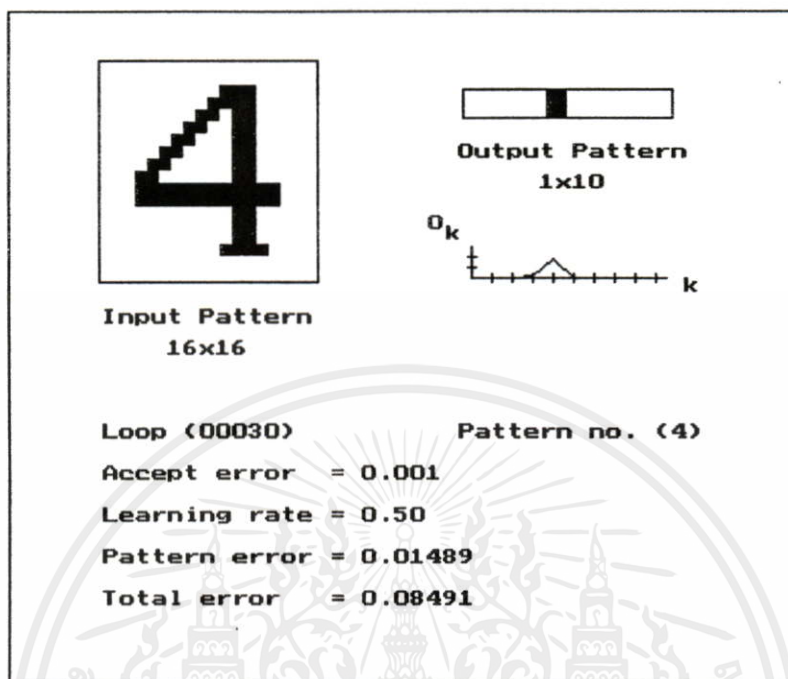


c) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "4" ในรอบที่ 20

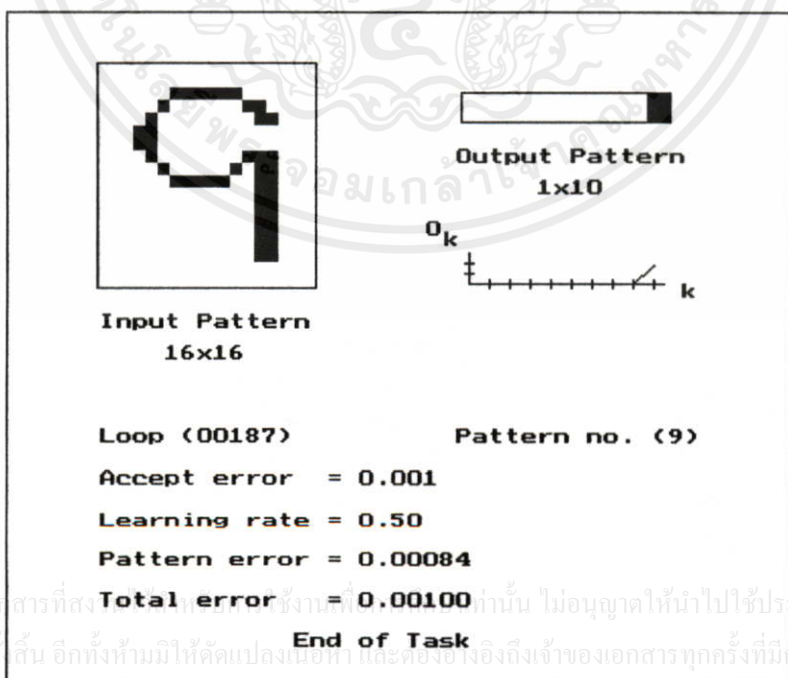


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

f) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "1" ในรอบที่ 30



g) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "4" ในรอบที่ 30



h) แสดงการฝึกสอนแพทเทิร์นตัวอักษรตัวเลข "9" ในรอบที่ 187

รูปที่ 26 (a) ถึงรูปที่ 26 (h) เป็นรูปแสดงการฝึกสอนเน็ตเวิร์ครวมทั้งหมด 187 รอบ โดยการฝึกสอนแพทเทิร์นตัวอักษรตัวเลขตั้งแต่ "0" ถึง "9" ครบ 10 แพทเทิร์นตัวเลข นับเป็น 1 รอบ ดังนั้นจอภาพจะแสดงรูปทั้งหมด  $187 \times 10$  เท่ากับ 1,870 รูป ซึ่งไม่สามารถพิมพ์ผลมาให้ดูได้ทั้งหมด จึงมีการเลือกเฉพาะบางภาพมาแสดงผล

พิจารณารูปที่ 26 (a), รูปที่ 26 (c), รูปที่ 26 (d), รูปที่ 26 (e), และรูปที่ 26 (g) แสดงการฝึกสอนแพทเทิร์นตัวเลข "4" โดยในรอบแรกๆ เอ้าท์พุทแพทเทิร์นที่ได้จากอินพุทแพทเทิร์นเป็นเอ้าท์พุทที่ไม่ถูกต้อง แต่เมื่อจำนวนรอบเพิ่มขึ้น มีการปรับค่าตัวเลขน้ำหนักในทุกรอบการทำงานมีผลทำให้เอ้าท์พุทดีขึ้นตามลำดับ รูปที่ 26 (e) (รอบที่ 20) แสดงเอ้าท์พุทแพทเทิร์นที่ได้จากอินพุทแพทเทิร์นมีลักษณะถูกต้องแล้ว แต่ค่าผิดพลาดของการฝึกสอน E ยังมีค่ามากกว่าค่าผิดพลาดที่ยอมรับได้  $E_a$  โปรแกรมจึงทำงานต่อไป จนกระทั่งค่าผิดพลาดของการฝึกสอน E มีค่าน้อยกว่าค่าผิดพลาดที่ยอมรับได้  $E_a$  ในรอบที่ 187 โปรแกรมจึงหยุดการทำงาน และเก็บค่าตัวเลขน้ำหนักสุดท้ายไว้ใช้งานต่อไป

พิจารณารูปที่ 26 (b), รูปที่ 26 (f), และรูปที่ 26 (h) แสดงการเรียนรู้ของแพทเทิร์นตัวเลขตัวอื่น นอกเหนือจากแพทเทิร์นตัวเลข "4" ในรอบการทำงานนั้น

### 5.3 ผลการทดสอบ

ได้ทดสอบความสามารถในการแยกแยะแพทเทิร์นที่มีความผิดเพี้ยน (Distortion) หลังจากสอนแพทเทิร์นตัวอักษรตัวเลขสิ้นสุดลง โดยสองเพิ่มสัญญาณรบกวน (Noise) เข้าไปในแพทเทิร์นที่สมบูรณ์ดังแสดงในรูปที่ 27 กำหนดค่าความน่าจะเป็นของสัญญาณรบกวน P(Probability) เท่ากับ 0.10 (10 เปอร์เซ็นต์) แล้วคำนวณจำนวนบิตของสัญญาณรบกวนสำหรับแพทเทิร์นขนาด  $64 \times 64$  (256 บิตแพทเทิร์น) ตามสมการที่ (73)

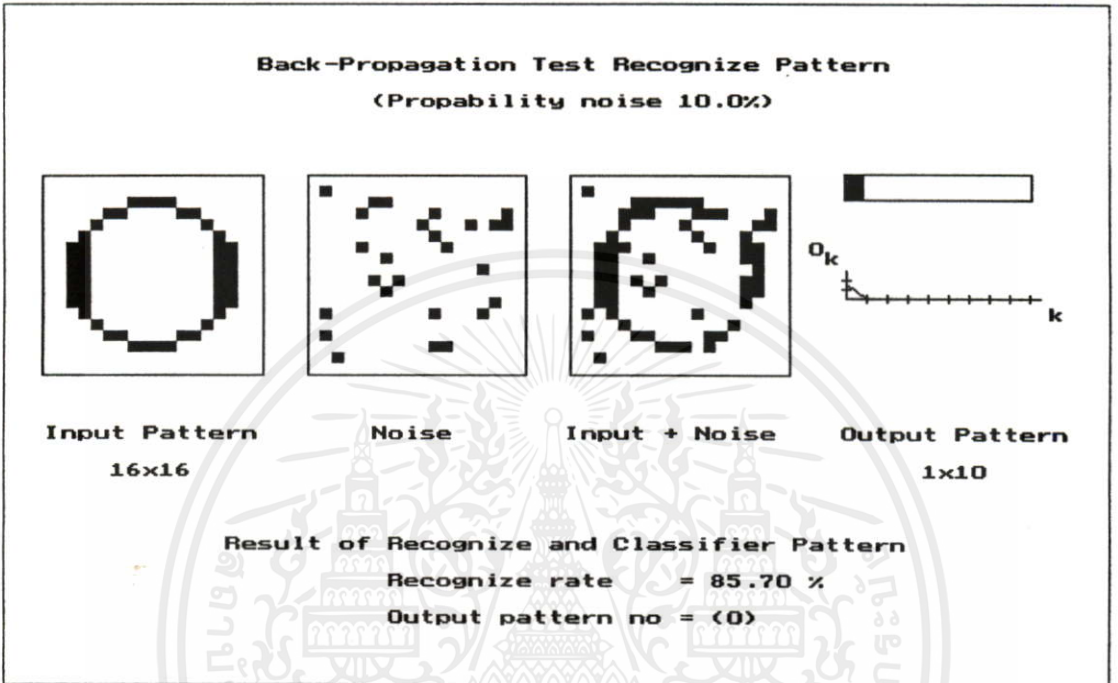
$$P = \text{Noise} / \text{ขนาดแพทเทิร์น} \dots\dots\dots (73)$$

แทนค่าเป็นเอกส Noise เป็น  $0.1 \times 256$  ซึ่งเท่ากับ 25.6 บิตค่าเท่านี้ ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งยังมีใช้คำนวณอีก และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้  $\cong 26$  บิตแพทเทิร์น

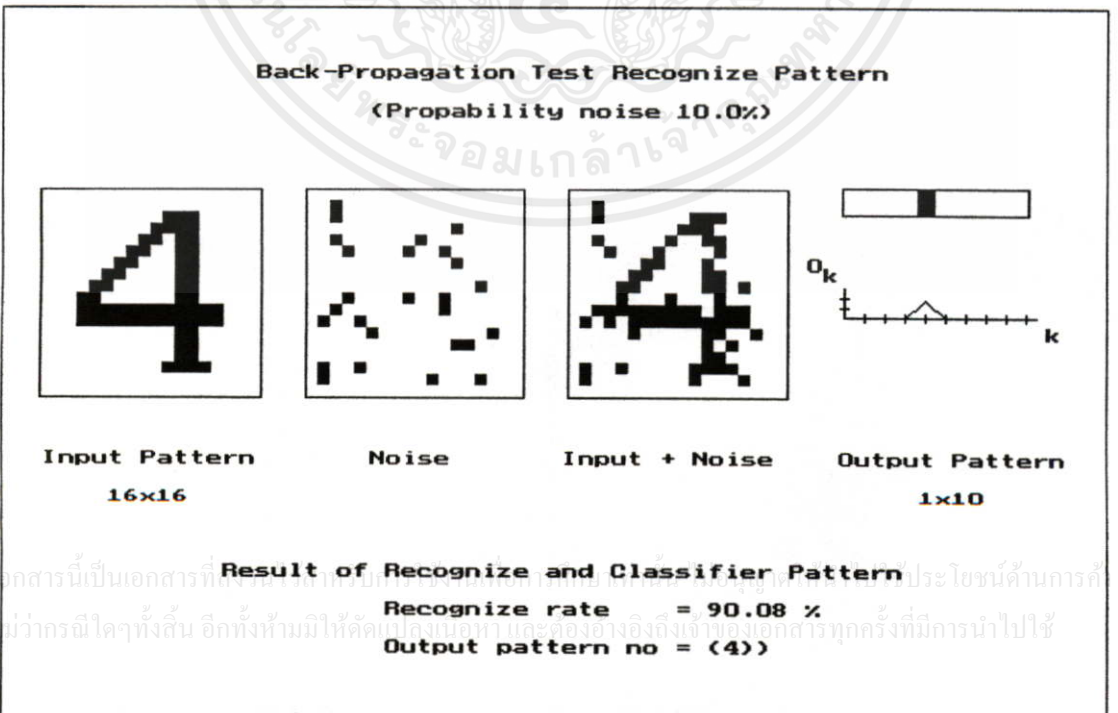
ดังนั้นจะต้องใส่สัญญาณรบกวนจำนวน 26 บิตแพทเทิร์นลงใน 256 บิตของอินพุทแพทเทิร์น

## รูปที่ 27

แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์น เมื่อข้อมูลอินพุตถูกบิดเบือน

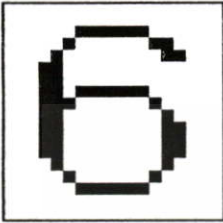


a) แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์นตัวอักษรตัวเลข "0"

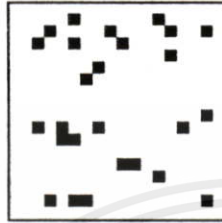


แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์นตัวอักษรตัวเลข "4"

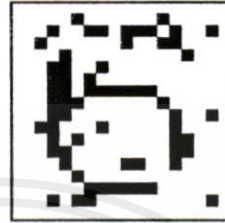
**Back-Propagation Test Recognize Pattern**  
(Propability noise 10.0%)



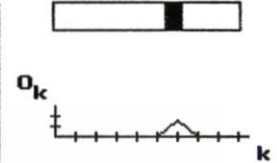
**Input Pattern**  
16x16



**Noise**



**Input + Noise**



**Output Pattern**  
1x10

**Result of Recognize and Classifier Pattern**

Recognize rate = 82.52 %

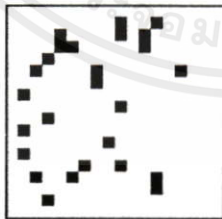
Output pattern no = (6)

c) แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์นตัวอักษรตัวเลข "6"

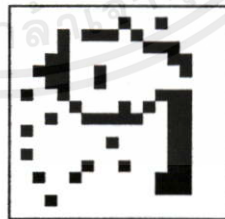
**Back-Propagation Test Recognize Pattern**  
(Propability noise 10.0%)



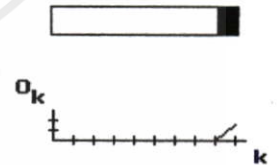
**Input Pattern**  
16x16



**Noise**



**Input + Noise**



**Output Pattern**  
1x10

**Result of Recognize and Classifier Pattern**

Recognize rate = 80.74 %

Output pattern no = (9)

d) แสดงตัวอย่างการทดสอบการแยกแยะแพทเทิร์นตัวอักษรตัวเลข "9"

จากรูปที่ 27 แสดงผลการทดสอบการรู้จำแพทเทิร์นตัวอักษร โดยดูได้จากภาพคอมพิวเตอร์ ตัวอักษรตัวเลขทางซ้ายมือเป็นอินพุทแพทเทิร์นปกติ (Input Pattern) ซึ่งนำมาเป็นตัวอย่างการทดสอบร่วมกับสัญญาณรบกวนในรูปถัดไป (Noise) ได้อินพุทแพทเทิร์นที่มีสัญญาณรบกวนปรากฏ (Input Pattern + Noise) ส่วนรูปทางขวามือสุดเป็นผลลัพธ์ที่ผ่านการแยกแยะแพทเทิร์นด้วยโปรแกรมคอมพิวเตอร์ (Output Pattern) โดยวิธีการของเบ็คหรือพาทกชันนิวโรลเน็ตเวิร์ค ผลจากการทดสอบในเรื่องการทนต่อสภาพผิดปรกติแสดงให้เห็นว่า แม้อินพุทแพทเทิร์นตัวอักษรตัวเลขที่เข้ามาจะมีสัญญาณรบกวนบ้าง หากไม่มากนักโปรแกรมสามารถแยกแยะแพทเทิร์นและแสดงเอาท์พุทที่ถูกต้องได้ โดยมีความผิดพลาดน้อยมาก

#### 5.4 เทคนิคการฝึกสอนเพิ่มเติม

จากที่กล่าวมาแล้วข้างต้น เรากำหนดค่าอัตราการเรียนรู้  $\eta$  เป็นค่าคงที่ตลอดการฝึกสอน ซึ่งให้ผลดีพอสมควร แต่จะขอแนะนำเพิ่มเทคนิคการฝึกสอนอีก 2 วิธี รวมเป็น 3 วิธี [9] ดังนี้

1) กำหนดค่าอัตราการเรียนรู้  $\eta$  เป็นค่าคงที่ ซึ่งการเลือกค่า  $\eta$  ที่เหมาะสมที่สุดจำเป็นต้องทดลองแทนค่าด้วยตนเอง เมื่อกำหนดให้ค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  เท่ากับ 0.001 ดังแสดงในตารางที่ 6

ตารางที่ 6

แสดงผลจำนวนรอบและเวลาทั้งหมด เมื่อค่า  $\eta$  เปลี่ยนไป

ค่า $\eta$	จำนวนรอบทั้งหมด	เวลาทั้งหมด
1.0	84	09:30 นาที
0.7	128	14:20 นาที
0.5	187	21:08 นาที
0.2	486	54:50 นาที

หากให้  $\eta$  มีค่ามาก จำนวนรอบของการฝึกสอนจะมีค่าน้อย หมายความว่า เนตเวิร์คสามารถเรียนรู้ได้เร็ว ใช้เวลาน้อย แต่ประสิทธิภาพในการแยกแยะแพทเทิร์นจะไม่ดีพอ ในทางตรงกันข้าม หากให้  $\eta$  มีค่าน้อยขณะฝึกสอน จำนวนรอบทั้งหมดจะมีค่ามากแต่สามารถแยกแยะแพทเทิร์นได้ดี ซึ่งผลจากการทดสอบ พบว่าค่า  $\eta$  ที่เหมาะสมอยู่ในช่วง 0.2 ถึง 0.5

กำหนดให้  $\eta$  เป็นค่าคงที่บ้างลองเปลี่ยนค่า  $E_a$  ปรากฏว่า จำนวนรอบและเวลาทั้งหมดในการฝึกสอนจะเพิ่มขึ้นมากเมื่อ  $E_a$  มีค่าลดลง ดังแสดงในตารางที่ 7 แต่ประสิทธิภาพในการแยกแยะแพทเทิร์นตัวเลขจะดีขึ้นอีกเล็กน้อย กำหนดให้  $\eta$  เป็นค่าคงที่เท่ากับ 0.5

ตารางที่ 7  
แสดงผลจำนวนรอบและเวลาทั้งหมด เมื่อค่า  $E_a$  เปลี่ยนไป

ค่า $E_a$	จำนวนรอบทั้งหมด	เวลาทั้งหมด
0.050	35	03:57 นาที
0.010	44	04:58 นาที
0.001	187	21:08 นาที

ข้อสรุปจากการทดสอบ เมื่อให้  $\eta$  มีค่ามากขณะฝึกสอน ค่าตัวเลขน้ำหนักจะมีการเปลี่ยนแปลงอย่างรวดเร็วตลอดเวลา ซึ่งมีโอกาสเกิดการแกว่งของระบบขึ้น เมื่อค่าผิดพลาดเฉลี่ยของระบบ  $E$  เข้าใกล้ค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  แต่หากให้  $\eta$  มีค่าน้อยการเรียนรู้จะใช้เวลามาก ดังนั้นเพื่อแก้ปัญหาดังกล่าว จึงเพิ่มเทคนิคการฝึกสอนขึ้นอีก 2 วิธีดังในข้อ 2) และ 3) โดยเริ่มแรกจะให้  $\eta$  มีค่ามาก และค่อยๆลดลง จนกระทั่ง  $E$  มีค่าน้อยกว่า  $E_a$  จึงหยุดทำงาน ดังนั้นปัญหาเรื่องการแกว่งก็จะหมดไป

2) สูตรแปรผันตามจำนวนรอบ ( $t$ ) ใช้แนวความคิดมาจากสูตรการกระจายค่าของเอาท์พุท เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า Self organization map ของ Kohonen [10] โดยเปลี่ยนมาใช้กับค่าอัตราการเรียนรู้  $\eta$  กำหนดให้ค่า  $\eta$  เริ่มแรกมีค่าเท่ากับ 1.0 และค่อยๆลดลงตามจำนวนรอบที่เพิ่มขึ้น จนกระทั่งค่าผิดพลาดเฉลี่ยของ

ระบบ E มีค่าน้อยกว่าค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  ซึ่งกำหนดค่าสุดท้ายของ  $\eta$  มีค่าประมาณ 0.2

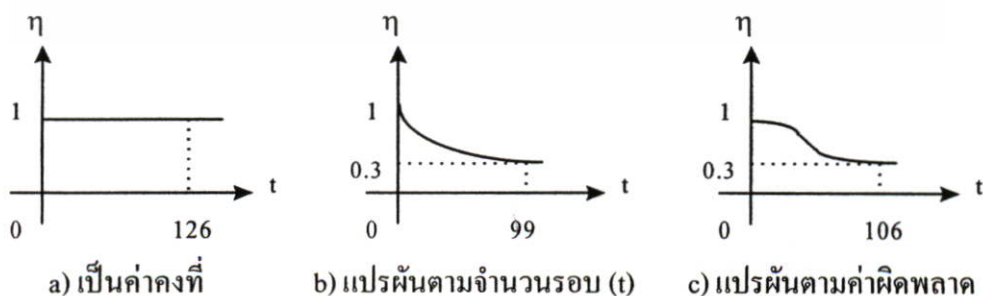
$$\eta = \eta_0 \left(1 - \frac{t}{T}\right) \dots\dots\dots (74)$$

โดย  $\eta_0$  ค่าเริ่มต้นของอัตราการเรียนรู้ กำหนดให้เท่ากับ 1.0  
 $t$  เป็นหมายเลขของรอบขณะทำการทดสอบ  
 $T$  เป็นจำนวนรอบทั้งหมดที่ใช้ในการทดสอบโดยประมาณ ต้องกำหนดเองโดยตั้งเงื่อนไขว่า หากค่าเริ่มต้น  $\eta$  มีค่าเท่ากับ 1 และค่าสุดท้ายมีค่าเป็น 0.2 ดังนั้น หากดูจำนวนรอบจากตารางที่ 6 และตารางที่ 7 จะเห็นว่าค่า  $T$  ที่เหมาะสม ควรมีค่าประมาณ 180 เมื่อกำหนด  $E_a$  เท่ากับ 0.001

3) สูตรแปรผันตามค่าผิดพลาดเฉลี่ยของระบบ E ซึ่งเป็นวิธีที่ผู้วิจัยนำเสนอ กำหนดให้ค่า  $\eta$  เริ่มแรกมีค่าเท่ากับ 1.0 และค่อยๆ ลดลงตามค่าค่าผิดพลาดเฉลี่ยของระบบ E ที่ลดลง จนกระทั่ง E มีค่าน้อยกว่าค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  กำหนดค่าสุดท้ายของค่า  $\eta$  มีค่าประมาณ 0.2

$$\eta = 1 - \frac{0.8 E_a}{E} \dots\dots\dots (75)$$

รูปที่ 28

แสดงการกำหนดค่า  $\eta$  แตกต่างกันตามสูตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 เลี้ยงของระบบ (E)  
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 28 เป็นกราฟแสดงความสัมพันธ์ระหว่างค่าอัตราการเรียนรู้  $\eta$  กับค่าหมายเลขของรอบขณะทำการทดสอบ รูปที่ 28(a) กำหนดให้  $\eta$  เป็นค่าคงที่ รูปที่ 28(b) กำหนดให้  $\eta$  แปรผันตามจำนวนรอบ  $t$  รูปที่ 28(c) กำหนดให้  $\eta$  แปรผันตามค่าผิดพลาดเฉลี่ยของระบบ E

เพื่อเปรียบเทียบการทดสอบอัตราการเรียนรู้ทั้ง 3 แบบ จึงนำข้อมูลของความสัมพัทธ์ระหว่างค่าผิดพลาดเฉลี่ยของระบบ E กับจำนวนรอบที่เพิ่มขึ้นขณะฝึกสอน เขียนเป็นกราฟได้ดังแสดงในรูปที่ 29 โดยแสดงค่าผิดพลาด E ในรูปแบบของลอการิทึม

แบบที่ 1) กำหนดค่า  $\eta$  เป็นค่าคงที่เท่ากับ 0.5

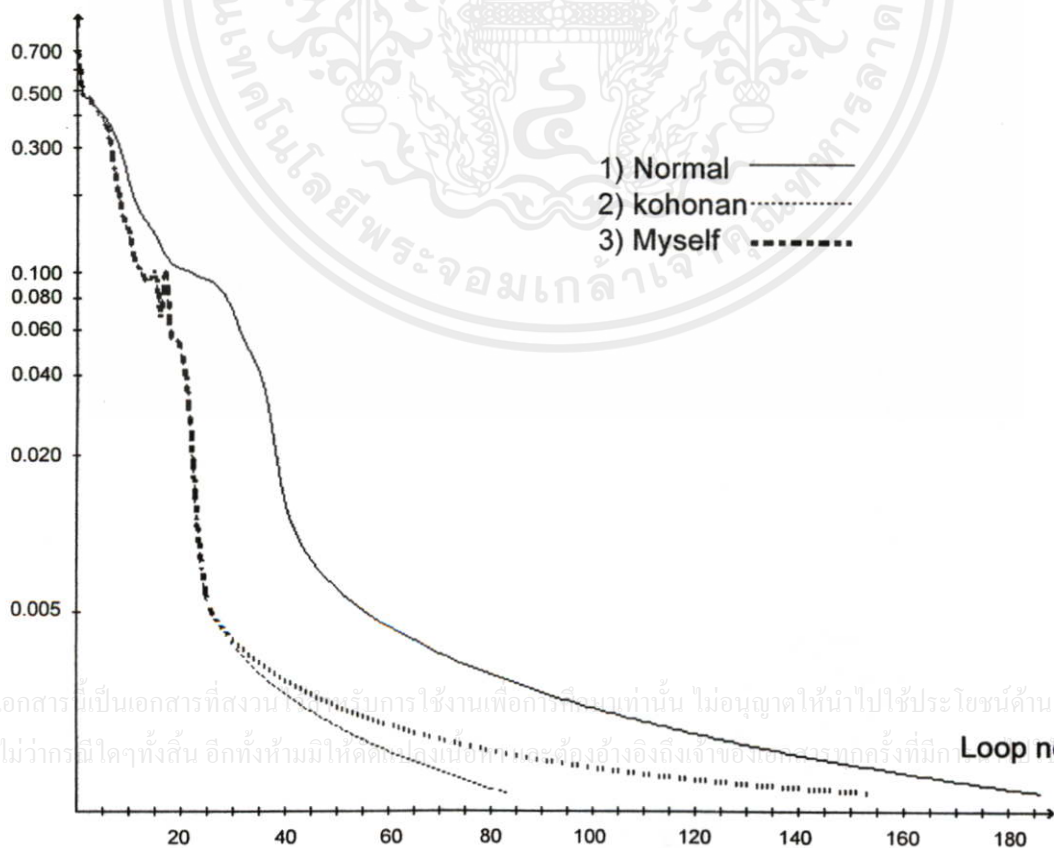
แบบที่ 2) กำหนดค่า  $\eta$  แปรผันตามจำนวนรอบ  $t$

แบบที่ 3) กำหนดค่า  $\eta$  แปรผันตามค่าผิดพลาดเฉลี่ยของระบบ E

รูปที่ 29

แสดงผลการฝึกสอน โดยกำหนดค่า  $\eta$  ต่างกันตามสูตร

Average error (E)



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหามาเผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มี

จากรูปที่ 29 แบบที่ 2) และ แบบที่ 3) ให้ผลใกล้เคียงกัน โดยเส้นกราฟจะทับกันพอดี ซึ่งค่าผิดพลาดเฉลี่ยของระบบ E จะลดลงเร็วมากในช่วงแรก และจะช้าลงเมื่อจำนวนรอบเพิ่มขึ้น จากการทดสอบความสามารถในการแยกแยะแพทเทิร์นเปรียบเทียบกันทั้ง 3 แบบพบว่า แบบที่ 2) และแบบที่ 3) จะสามารถแยกแยะแพทเทิร์นได้ดีกว่าแบบที่ 1) เล็กน้อย

### 5.5 ข้อเสนอแนะและผลการวิเคราะห์

1. หากกำหนดจำนวนฮิตเคน โหนดน้อยเกินไป จะทำให้ความสามารถในการแยกแยะแพทเทิร์นตัวอักษรให้ผลไม่ดีเท่าที่ควร ซึ่งจำนวนที่เหมาะสมควรมีค่าอยู่ระหว่างจำนวน โหนดของอินพุต และเอาต์พุตเลยอร์
  2. ค่าผิดพลาดเฉลี่ยของระบบที่ยอมรับได้  $E_a$  ที่เหมาะสมขึ้นอยู่กับงานที่ประยุกต์ใช้ หากเป็นงานที่ต้องการความละเอียดสูง จะกำหนดให้ค่า  $E_a$  มีค่าน้อยๆ
  3. หากกำหนดให้ค่าอัตราการเรียนรู้  $\eta$  เป็นค่าคงที่ตลอดการฝึกสอนและมีค่ามาก อาจมีผลทำให้เกิดการแกว่งในขณะฝึกสอนได้ ดังนั้นค่าอัตราการเรียนรู้  $\eta$  ควรแปรเปลี่ยนตามจำนวนรอบของการฝึกสอนหรือค่าผิดพลาดเฉลี่ยของระบบ
  4. เครื่องคอมพิวเตอร์ที่ใช้ในการทดสอบมีหน่วยประมวลผล Intel 80386DX ความเร็ว 40 MHz และหน่วยประมวลผลร่วมคณิตศาสตร์ ยังต้องใช้เวลาในการฝึกสอนมาก หากเป็นนิรอลเน็ตเวิร์คที่มีระบบใหญ่ๆ ขอแนะนำให้ทดสอบ โปรแกรมบนเครื่องที่มีประสิทธิภาพสูงกว่านี้
  5. ในสภาพความเป็นจริงนั้น การใช้สแกนเนอร์เพื่อการอ่านบิตแมพกราฟิกของตัวอักษรในลักษณะกลุ่มคำหลายบรรทัด เมื่อพิจารณาจะพบปัญหา 2 อย่างคือ การนำเข้าไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะพิมพ์ใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
- 5.1 มีความคิดเห็นของบิตบางบิตภายในตัวอักษรของแพทเทิร์น

5.2 แพทเทิร์นของตัวอักษรที่คอมพิวเตอร์จัดจำกับแพทเทิร์นของตัวอักษรที่รับเข้ามาจะมีลักษณะต่างกัน เช่น ตัวอักษรดตัวเลข "1" หรือ "l" เป็นต้น

หากใช้วิธีโดยทั่วไปที่ไม่ใช้นิวรอลเน็ตเวิร์คการเรียนรู้จำและแยกแยะแพทเทิร์นตัวอักษรในลักษณะต่างๆเป็นเรื่องที่ลำบากมาก และแทบเป็นไปได้เลยที่จะมีความแม่นยำสูง แต่จากการทดสอบพบว่า นิวรอลเน็ตเวิร์คแบบพรีอพาทเกชันสามารถแก้ปัญหาดังกล่าวได้ทั้ง 2 ข้อ ซึ่งมีความใกล้เคียงกับการใช้งานในความเป็นจริงมากกว่า จากการทดสอบพบว่า การรู้จำแพทเทิร์นตัวอักษรที่สร้างขึ้นใช้เองมีประสิทธิภาพในการแยกแยะแพทเทิร์นตัวอักษรดตัวเลขสูงกว่า 80 เปอร์เซ็นต์ขึ้นไป

6. ปัญหาของเบ็คพรีอพาทเกชันคือ ใช้เวลามากในการเรียนรู้เพื่อจดจำแพทเทิร์น และทุกครั้งที่มีการเพิ่มแพทเทิร์นตัวอักษรดตัวใหม่ จะต้องสอนเน็ตเวิร์คใหม่ทุกครั้ง แต่ในกรณีที่ผ่านมาการอบรมมาแล้ว ตัวเลขน้ำหนักที่ได้จากการฝึกสอนจะถูกเก็บไว้และสามารถนำกลับมาใช้ใหม่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 6

### ทดสอบการรู้จำตัวอักษรด้วยแสงสำหรับตัวอักษรตัวพิมพ์ภาษาอังกฤษ

ในบทนี้เราจะมุ่งเน้นทดสอบความสามารถในเรื่องการรู้จำตัวอักษรตัวพิมพ์ภาษาอังกฤษ จากข้อมูลภาพเอกสารในรูปแบบบิตแมพภาพ จำเป็นต้องผ่านกระบวนการแยกคุณสมบัติเฉพาะและการแยกประเภทตัวอักษร การแยกคุณสมบัติเฉพาะบ่อยครั้งจะถูกกำหนดล่วงหน้าสำหรับปัญหานั้นๆ เพื่อให้การออกแบบได้ผลมากที่สุดจะต้องคำนึงถึงประสิทธิภาพทั้งหมดของระบบ ส่วนเรื่องการแยกประเภทบ่อยครั้งจะอาศัยโมเดลที่สามารถฝึกสอนหรืออบรมได้ นิวรอลเน็ตเวิร์คจึงถูกนำมาใช้สำหรับปัญหาการรู้จำแพทเทิร์น เนื่องจากเป็นเครื่องมือที่มีความยืดหยุ่นมากสามารถทำงานร่วมกันได้กับการแยกคุณลักษณะเฉพาะ ทำให้ความต้องการการประมวลผลขั้นต้นถูกทำให้ลดลงและช่วยให้การรู้จำตัวอักษรมีความแม่นยำมากขึ้น

วิธีนำเสนอเพื่อให้ได้ผลลัพธ์ตามต้องการ ขึ้นอยู่กับปัจจัยหลายอย่าง อุปกรณ์นำเข้า Sensitive pad, สแกนเนอร์ (Scanner) และกล้องถ่ายรูป (Camera) ลักษณะการเขียนตัวอักษร (ตัวอักษรแยกจากกัน หรือการเขียนติดกันในลักษณะต่อเนื่อง) จำนวนของกลุ่มหรือประเภท (ตัวเลข ตัวอักษรภาษาอังกฤษตัวใหญ่หรือตัวเล็ก สัญลักษณ์) และความเป็นไปได้ในการเลือกใช้ข้อมูลสำหรับการอบรมและการแปลความหมาย

การวิจัยครั้งนี้จะกำหนดรูปแบบการรู้จำตัวอักษรของระบบ ด้วยคุณลักษณะเฉพาะดังนี้

- เอกสารภาพของตัวหนังสือถูกใส่เข้าไปในเครื่องสแกนเนอร์แบบตั้งโต๊ะ ซึ่งจะแสดงผลด้วยลำดับของตำแหน่ง  $[x(n), y(n)]$  เมื่อ  $n$  คือหมายเลขลำดับ
- เอกสารภาพที่นำเข้ามาจะไม่ปรากฏรูปภาพอื่นใดยกเว้นส่วนของตัวหนังสือ และจะพิจารณาเฉพาะตัวอักษรภาษาอังกฤษแบบตัวพิมพ์ ตัวอักษรตัวใหญ่ ตัวเล็ก ตัวเลข และตัวอักษรพิเศษบางตัว ถูกพิจารณา
- คัดเลือกแบบอักษรภาพที่นิยมใช้กันบ่อยๆ จากเอกสารโดยทั่วไป ซึ่งอาจจะครอบคลุมได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งยังมีให้คิดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ยังไม่มีการใช้วากยสัมพันธ์และอรรถศาสตร์ จำพวกโปรแกรมตรวจสอบการสะกดคำ

จากพจนานุกรมเข้ามาเกี่ยวข้อง

เราใช้ชุดตัวอักษร 74 ชุดจากวารสารจำนวน 7 เล่ม และหนังสือพิมพ์ภาษาอังกฤษจำนวน 2 เล่ม ประสิทธิภาพของระบบถูกประเมินผลจากตัวอักษรหลายพันตัว ข้อมูลที่ได้ถูกรวบรวมจากห้องพักนักศึกษาของอาจารย์ที่ปรึกษา ณ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง การทดสอบการรู้จำเป็นหัวข้อที่ใหญ่มีรายละเอียดมากมาย ในเริ่มต้นจึงควรศึกษาพื้นฐานบ้างพอสมควรเพื่อให้ทราบขั้นตอนและความหมายของคำศัพท์บางคำ

### 6.1 ความรู้พื้นฐานของการรู้จำตัวอักษร (Character Recognition)

งานทางด้านกรจับภาพสามารถถูกใช้ในการจัดพิมพ์ด้วยคอมพิวเตอร์แบบตั้งโต๊ะ (Desktop publishing) ภาพนั้นอาจจะนำเข้ามาแก้ไขก่อนที่จะนำไปรวมกับเอกสาร หรือถูกใช้拿去ไปแสดงเป็นส่วนประกอบของศิลปะที่นำมาแสดงด้วยเทคนิคเกี่ยวกับรูปภาพ

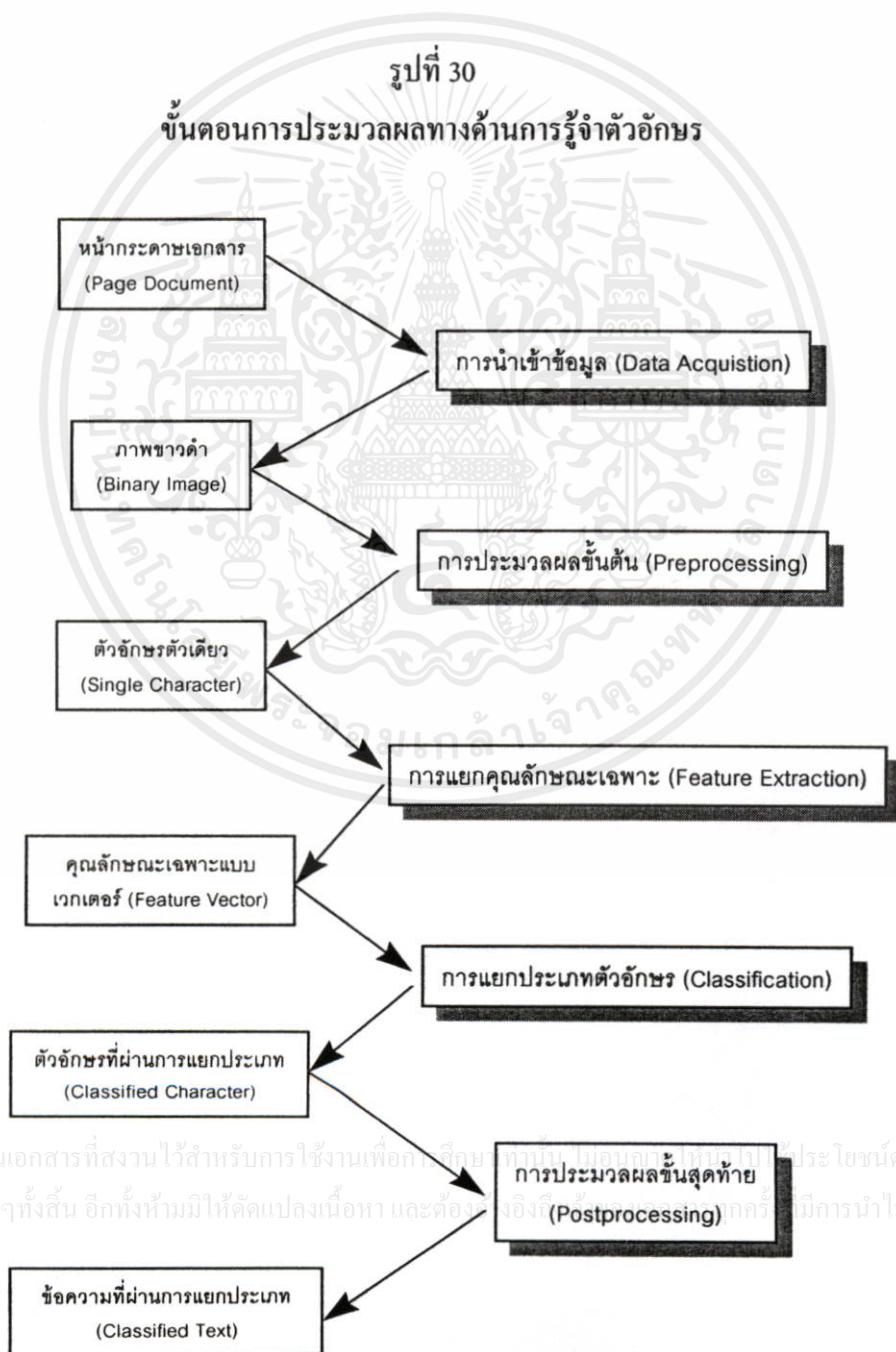
มีบางครั้งเมื่อภาพที่ถูกจับเข้ามาไม่ใช่ผลงานชิ้นสุดท้าย บ่อยครั้งเราต้องการใช้คำที่ถูกเก็บอยู่ในรูปแบบสำเนาถาวร (Hard copy) ภาพของคำไม่ใช่สิ่งที่เราต้องการ ที่เรากำลังมองหาคือต้องการให้คอมพิวเตอร์สามารถแปลความหมายภาพของตัวหนังสือไปเป็นตัวอักษรด้วยตัวเอง ซึ่งจำเป็นต้องอาศัยเทคโนโลยีทางด้านประมวลผลภาพเอกสาร (Document image processing) เข้ามาช่วย

การรู้จำตัวอักษรด้วยแสงเป็นส่วนหนึ่งของเทคโนโลยีทางด้านประมวลผลภาพเอกสาร ซึ่งอาศัยกระบวนการแปลงภาพของเอกสารให้อยู่ในรูปสื่ออิเล็กทรอนิกส์ด้วยเครื่องอ่านสแกนเนอร์ ทำให้ได้ข้อมูลในรูปบิตแมพภาพ (Bit-map image) จากนั้นนำข้อมูลที่ได้ไปวิเคราะห์เพื่อให้เครื่องคอมพิวเตอร์รู้จักตัวอักษรหรือข้อความในรูปรหัสตัวอักษรแอสกี (ASCII Characters)

Kahan etal [11] แนะนำว่าโปรแกรมในทางปฏิบัติสำหรับระบบการรู้จำตัวอักษรในหน้าเอกสาร ควรมีความแม่นยำในการอ่านตัวอักษรอย่างน้อยที่สุด 99.9% ซึ่งเป็นเรื่องยากที่จะออกแบบระบบการรู้จำตัวอักษรให้รู้จักแบบฟอร์มและรู้จักช่องว่างระหว่างตัวอักษรตัวพิมพ์ เราควรจะพัฒนาระบบอย่างไรให้สามารถรู้จำได้ค่อนข้างสูง โดยไม่คำนึงถึงคุณภาพของภาพเอกสารและแบบอักษรที่เข้ามา ในปัจจุบันคำผิดพลาดในการรู้จำตัวอักษรส่วนใหญ่เกิดจากการตัดแยกตัวอักษร (Character segmentation) พบเห็นได้จากตัวอักษรของคำวางใกล้กัน (Adjacent) ในลักษณะสัมผัสกัน (Touching) หรือเกิดการเหลื่อมล้ำกัน ปัญหาอาจเกิดขึ้นจากชนิดของแบบอักษร ขนาดของแบบอักษร วิธีการถ่ายเอกสาร ข้อจำกัดเรื่องความละเอียดของเครื่องสแกนเนอร์ หรือหลักการใช้ค่าเทรคโธรลสำหรับ

ไบนารีสูงเพื่อหลีกเลี่ยงการแตกแยกของตัวอักษร เช่นกรณี การพิมพ์แบบจุด (Dot matrix) เป็นต้น บางอย่างของความล้มเหลวหรือความผิดพลาดในขั้นตอนแยกตัวอักษร ทำให้ข่าวสารบางอย่างสูญหายไปจากหน้าเอกสาร

มีหลายๆ วิธีที่พยายามพัฒนาคุณสมบัติการแยกคำเข้าไปในตัวอักษรเป็นส่วนๆ Jim Wang และ Jack Jean [11] เสนออัลกอริทึมการตัดแยกด้วยเส้นทางสั้นที่สุด (Shortest path segmentation algorithm) สำหรับตัวอักษรสัมผัสกัน จะอาศัยวิธีตัดเพื่อแยกตัวอักษรด้วยวิธีการเลือกเส้นทางที่ตัดผ่านลายเส้นตัวอักษรให้น้อยที่สุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษากลับไป ไม่ควรเผยแพร่ไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องสงวนลิขสิทธิ์ไว้ก่อนการนำไปใช้

ในหัวข้อนี้จะแสดงการพรรณนาอย่างย่อในเรื่องการรู้จำตัวอักษรด้วยแสง โดยทั่วไป กระบวนการรู้จำตัวอักษรมีรูปแบบต่างๆ ของการกระทำเป็นลำดับเกี่ยวข้องกับการนำเข้าข้อมูล การประมวลผลขั้นต้น การแยกคุณลักษณะเฉพาะของตัวอักษร การแยกประเภทตัวอักษร และการประมวลผลขั้นสุดท้าย ดังแสดงในรูปที่ 30

รูปที่ 30 จะแสดงขั้นตอนการรู้จำตัวอักษรทั้งหมด 5 ขั้นตอน เริ่มตั้งแต่การนำข้อมูลภาพเอกสารเข้าสู่ระบบจนกระทั่งการแปลงความหมายเสร็จสิ้นจนได้ข้อมูลตัวอักษรตามที่ต้องการ

### 6.1.1 การนำเข้าข้อมูล (Data acquisition)

การนำเข้าข้อมูล [12] เป็นการแปลงสัญญาณนอกระบบให้เข้าสู่ระบบด้วยวิธีการที่แตกต่างกัน เพื่อให้ได้ข้อมูลภาพในลักษณะบิตแมพกราฟิก 2 มิติ จึงจำเป็นต้องคิดค้นอุปกรณ์แปลงสัญญาณภาพเอกสารให้เป็นสัญญาณดิจิทัลสู่เครื่องคอมพิวเตอร์ เช่น Sensitive pad, สแกนเนอร์ กล้องถ่ายภาพและวีดีโอ เป็นต้น

Sensitive pad เป็นอุปกรณ์ตรวจจับสัญญาณการเปลี่ยนแปลงของอุปกรณ์ เช่น ตรวจจับทิศทาง การลากเมาท์ การใช้ปากกานำสัญญาณ จอภาพแบบสัมผัส เป็นต้น การเขียนหรือลากลายเส้นบนอุปกรณ์ดังกล่าว ทำให้เกิดอักษรของตัวอักษรแสดงบนจอภาพได้

สแกนเนอร์เป็นอุปกรณ์รับสัญญาณภาพโดยอาศัยแสงเป็นเครื่องแปลงสัญญาณแสงที่ตกกระทบบนเอกสารจะถูกแปลเป็นสัญญาณภาพแบบดิจิทัลที่เรียกว่า บิตแมพภาพ (Bit-map image) และจุดภาพในบิตแมพภาพถูกเรียกว่า พิกเซล (Pixels) หรือจุดภาพ สามารถเลือกได้ว่าต้องการภาพสีหรือภาพขาวดำ (Binary image)

กล้องถ่ายภาพและวีดีโอก็เช่นกันอาศัยแสงเป็นเครื่องแปลงสัญญาณ เพื่อให้ได้ภาพของข้อความตัวอักษรที่สามารถนำไปประมวลผลกับงานทางด้านการรู้จำตัวอักษรได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

#### 6.1.1.1 แบบอักษร

แม้ว่ากรมฯ จะให้สิทธิแก่ผู้จัดทำหนังสือพิมพ์ให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบอักษรหรือฟอนท์ (Font) ของตัวอักษรที่ใช้กันโดยทั่วไป ทั้งตัวอักษรดัดพิมพ์ (Printed Character) และตัวอักษรแบบลายมือเขียน (Handwritten Character) มีให้ใช้ดังนี้ เช่น Pica, Courier, Sans serif เป็นต้น

รหัสของตัวอักษรภาษาอังกฤษที่ใช้กันโดยทั่วไป มีทั้งตัวอักษรตัวใหญ่ (A..Z), ตัวอักษรตัวเล็ก (a..z), ตัวอักษรตัวเลข (0..9) และรหัสเครื่องหมายต่างๆ (+-\*;; อื่นๆ) โดยมีขนาดตัวอักษรแตกต่างกันไป

รูปแบบของตัวอักษรมีทั้งแบบธรรมดา (Normal fonts) แบบเอน (Italic fonts) แบบหนา (Thick fonts) แบบบาง (Thin fonts) แบบยก (Superscript fonts) แบบห้อย (Subscript fonts)

### รูปที่ 31

แสดงตัวอย่างแบบอักษรที่ใช้งานในปัจจุบัน

AngsanaUPC Font  
CordiaUPC Font  
Times New Roman

รูปที่ 31 แสดงแบบอักษรดัดพิมพ์ภาษาอังกฤษที่ใช้กันโดยทั่วไป แบบอักษร AngsanaUPC แบบอักษร CordiaUPC และแบบอักษร Times New Roman

#### 6.1.1.2 ตัวอักษรดัดพิมพ์ (Printed Character)

พบได้จากแบบฟอร์มและสิ่งพิมพ์ที่ถูกสร้างขึ้นโดย เครื่องพิมพ์ดีด (Type-writers) เครื่องพิมพ์รายบรรทัด (Line printers) เครื่องพิมพ์แบบจุด (Dot matrix printers) เครื่องพิมพ์เลเซอร์ (Laser printers) เป็นต้น ตัวอย่างการอ่านตัวอักษรด้วยเครื่องสแกนเนอร์ถูกแสดงดังในรูปที่ 32

รูปที่ 32

แสดงผลการพิมพ์จากเครื่องพิมพ์ต่างชนิด

DOT MATRIX

TYPE-WRITERS

Laser printers

## 6.1.1.3 ตัวอักษรลายมือเขียนและลายมือพิมพ์ (Handwritten และ Handprinted Character)

การเขียนตัวอักษรภาษาอังกฤษแบบลายมือมีอยู่สองรูปแบบ แบบลายมือเขียนและแบบลายมือพิมพ์ ตัวอักษรที่เกิดจากลายมือเขียนภาษาอังกฤษจะมีรูปร่างไม่แน่นอน ตัวอักษรทุกตัวภายในหนึ่งคำจะเขียนในลักษณะติดกัน พบได้จากลายเซ็นต์โดยทั่วไป ส่วนตัวอักษรที่เกิดจากลายมือพิมพ์จะมีรูปร่างค่อนข้างแน่นอน ตัวอักษรทุกตัวภายในหนึ่งคำจะเขียนแยกจากกัน

ในปัจจุบันการรู้จำตัวอักษรลายมือมีความก้าวหน้าเป็นอย่างมาก และพบว่าผลการรู้จำตัวอักษรลายมือพิมพ์จะมีอัตราความแม่นยำสูงกว่าการรู้จำตัวอักษรลายมือเขียน เนื่องจากไม่พบปัญหาเรื่องการแบ่งแยกส่วนประกอบของคำเกิดขึ้น เพื่อให้เกิดความชัดเจนในทางปฏิบัติจะกำหนดให้การรู้จำตัวอักษรกระทำกับแบบฟอร์มเอกสารที่เตรียมไว้ ตัวอักษรลายมือพิมพ์จะปรากฏเฉพาะในช่องสี่เหลี่ยมตามที่กำหนดให้เท่านั้น ดังแสดงในรูปที่ 33

รูปที่ 33

แสดงตัวอย่างลายมือพิมพ์ที่ได้จากแบบฟอร์มที่กำหนดให้

Name	:	S	E	T	T	A	P	O	L	L	I	N	P	R	A	C	H	Y	A	
Address	:	1	/	8	2		M	O	O		5	,		N	O	N	G	B	O	N
		P	R	A	V	E	S	,		B	A	N	G	K	O	K				

## 6.1.2 การประมวลผลในขั้นต้น (Pre-processing)

การประมวลผลขั้นต้นจะรวมทุกๆ ขั้นตอนที่สำคัญ เพื่อจะนำข้อมูลอินพุตเข้าไปในรูปแบบที่ยอมรับได้สำหรับแยกลักษณะเฉพาะของระบบ ขั้นตอนที่เกิดขึ้นอาจมีได้ดังนี้ [13]

1. การแปลงภาพเอกสารให้อยู่ในรูปแบบดิจิทัลที่คอมพิวเตอร์สามารถประมวลผลได้ (Digitization: Gray scale หรือ color manipulation)
2. การทำแปลงภาพสีเป็นภาพขาวดำ (Binarization)
3. การวิเคราะห์หน้ากระดาษ (Page layout Analysis: Square detection & Texture Analysis)
4. การกรองสัญญาณรบกวน (Filtering หรือ Corrupting noise)
5. การทำให้บาง (Thinning or skeleton)
6. การย่อขยายขนาด (Scaling)
7. การทำให้ภาพอยู่กึ่งกลาง (Centering)
8. การตัดแยก (Segmentation: Overlap และ Touching)

โดยส่วนมากวารสารอ้างอิงในด้านการรู้จำตัวอักษร จะถือว่าขั้นตอนนี้กระทำเสร็จเรียบร้อยแล้ว จึงมีการกล่าวถึงไม่มากนัก

## 6.1.3 การแยกแยะคุณลักษณะเฉพาะ (Feature Extraction)

ข้อมูลที่ได้รับจากการประมวลผลขั้นต้น จะอยู่ในรูปแบบบิตแมพแพทเทิร์นขนาด 2 มิติ ข้อมูลดังกล่าวอาจไม่เหมาะแก่การประมวล เนื่องจากยังไม่สามารถแยกแยะคุณลักษณะเฉพาะ (Feature) หรือคุณสมบัติเฉพาะ (Properties) ของตัวอักษรในรูปแบบหลากหลายได้ (Wide variety) เช่น ตัวอักษร A ขนาด  $8 \times 8$  และขนาด  $10 \times 10$  จะมีโครงสร้างในรูปบิตแมพแตกต่างกันตามตำแหน่งในแต่ละจุดภาพ

จากปัญหาดังกล่าวทำให้เกิดการพัฒนาการแยกแยะคุณลักษณะเฉพาะของตัวอักษรแต่ละตัวให้มีค่าเฉพาะที่ไม่ซ้ำกัน และมีลักษณะเป็นหนึ่งเดียว (Unique)

เทคนิคการแยกแยะคุณลักษณะเฉพาะเป็นหัวใจของระบบรู้จำตัวอักษร ดังนั้นเทคนิคที่นำมาใช้ควรจะสอดคล้องกับเทคนิคของการแยกประเภทตัวอักษรด้วย เทคนิคต่างๆที่รวบรวมได้จากเอกสารโดยทั่วไปมีอยู่หลายวิธีดังนี้ [13-14]

1. Pixels feature
2. Feature points (Intersection, endpoint, etc)
3. Profile projection (ความถี่)
4. Texture feature
5. Moment variants
6. Polygonal approximation
7. Curve approximation
8. Stroke, Direction
9. Redical
10. Transformation coefficient (Gaussian, Fourier descriptors, etc)

#### 6.1.4 การแยกประเภทตัวอักษร (Classification)

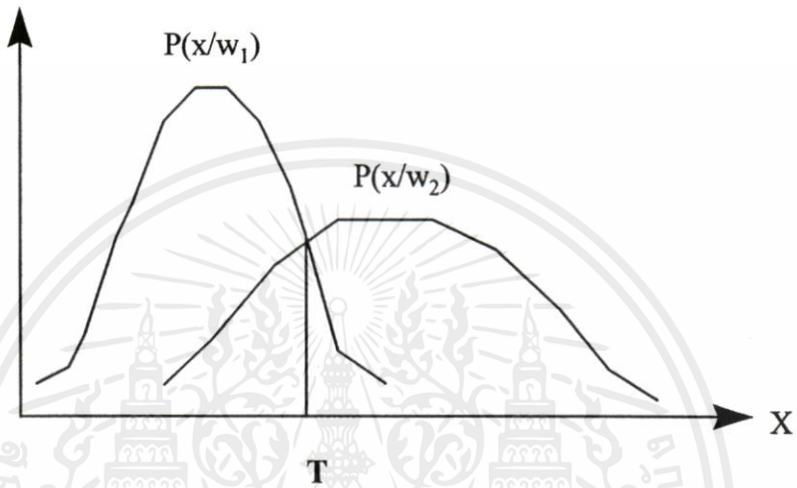
การแยกประเภทตัวอักษรเป็นการแบ่งแยกคุณสมบัติของแต่ละตัวอักษรในแต่ละกลุ่มออกจากกัน หากพูดในแง่สถิติ (Statistical) ตัวอย่างข้อมูลของอินพุทแพทเทิร์นที่ได้จากกระบวนการแยกคุณลักษณะเฉพาะควรถูกแยกประเภทเพื่อบ่งบอกได้ว่า เป็นข้อมูลของรหัสตัวอักษรใด ความเป็นไปได้ในการแยกประเภทตัวอักษรบางครั้งก็เป็นเรื่องยาก เนื่องจากตัวอย่างข้อมูลเกิดการเหลื่อมล้ำกัน (Overlap) ดูจากรูปที่ 34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 34

## แสดงการใช้ Probability density function

Probability density



การตัดสินใจทางสถิติจะต้องเลือกค่าเทรคโอรลที่เป็นไปได้ เพื่อแยกแยะตัวอย่างข้อมูล 2 กลุ่มออกจากกัน โดยใช้ Probability density function ซึ่งก็ให้ผลดีพอควร แต่ในความเป็นจริงตัวอย่างข้อมูลอาจมีหลายมิติคุณลักษณะเฉพาะ (Feature space) ซึ่งจำเป็นต้องใช้วิธีการสถิติที่ซับซ้อนกว่านี้

มีหลายวิธีที่จะนำเสนอการแยกประเภทตัวอย่างแพทเทิร์นใน Feature space นอกเหนือจากวิธีของนิวรอลเน็ตเวิร์ค ยังมีอีกหลายวิธีที่จะแยกประเภทข้อมูลออกเป็น M กลุ่มและเลือกประเภท (Category) ที่สอดคล้องกับการแยกประเภทมากที่สุดได้แก่ [13]

1. Minimum distance classifier (Distance matching)
2. Maximum likelihood classifier (Tree Classification)
3. Correlation matching
4. Relaxation matching
5. Markovian matching
6. Fuzzy matching

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตัดสินใจแยกประเภทตัวอักษรที่ได้จากการแยกคุณลักษณะเฉพาะนั้น ในปัจจุบันเริ่มมีคนสนใจหันมาใช้วิธีการของแบบจำลองโครงข่ายสมองมนุษย์กันมากขึ้น เนื่องจากสามารถใช้แยก

แยะตัวอักษรลายมือเขียนได้ โดยนิรอลเน็ตเวิร์คได้พัฒนากฎของตัวเองตามแบบจำลองต่างๆ ซึ่งอาศัยการเรียนรู้แพทเทิร์นตัวอักษรจากการฝึกสอน แบบจำลองของนิรอลเน็ตเวิร์คที่ใช้กันโดยทั่วไปในการรู้จำตัวอักษรมีดังนี้ Associative memory model, Feedforward networks และ Neocognitron [13]

### 6.1.5 การประมวลผลในลำดับหลัง (Post-processing)

ในบางระบบความแม่นยำในการรู้จำตัวอักษรจะถูกปรับปรุง โดยเพิ่มกระบวนการอธิบายความหมาย (Contextual processing) หลังการแยกประเภทตัวอักษร ตัวอย่างเช่น โปรแกรมตรวจสอบการสะกดคำ พจนานุกรม และตารางข้อมูลทางสถิติ เพื่อใช้เป็นฐานข้อมูลสำหรับการวิเคราะห์การรู้จำตัวอักษรที่มีความสับสนบ่อยๆ โดยระบบจะสามารถแก้ไขตัวอักษรที่รู้จำผิดพลาด หรือตัวอักษรที่ไม่สามารถแยกประเภทได้ ทำให้ความแม่นยำในการรู้จำตัวอักษรมีความถูกต้องสูงขึ้น

ความรู้พื้นฐานของการรู้จำตัวอักษรและขั้นตอนโดยย่อถูกอธิบายในเชิงสรุปมาพอสมควรแล้ว ซึ่งมีการกล่าวถึงหลักการและวิธีการ โดยทั่วไป แต่ยังไม่ได้กล่าวถึงหลักการที่จะนำมาใช้กับงานวิจัยในครั้งนี้ ซึ่งจะแสดงรายละเอียดทั้งหมดในหัวข้อถัดไปตามลำดับ

## 6.2 การนำเข้าข้อมูลจากเครื่องสแกนเนอร์

ข้อมูลภาพที่เราสนใจจะถูกอ่านด้วยเครื่องมือแปลงสัญญาณภาพให้เป็นสัญญาณแอนะล็อกหรือ ดิจิตอล ตามแต่งานที่ต้องการ สำหรับงานทางด้านการศึกษาจำเป็นอย่างยิ่งที่จะต้องแปลงสัญญาณภาพให้เป็นสัญญาณดิจิตอลสู่เครื่องคอมพิวเตอร์ เครื่องดังกล่าวถูกเรียกว่า เครื่องสแกนเนอร์ มีอยู่ 2 แบบในท้องตลาด แบบแรกเป็นสแกนเนอร์แบบมือถือ (Handdy scanner) ต้องอาศัยมือช่วยลากตัวเครื่องผ่านหน้ากระดาษเอกสารเพื่อแปลงสัญญาณภาพตามตำแหน่งที่ต้องการ แบบที่ 2 เป็นสแกนเนอร์แบบตั้งโต๊ะ (Flat-bed scanner) อาศัยการวางหน้ากระดาษเอกสารลงบนตัวเครื่อง คล้ายการทำงานของเครื่องถ่ายเอกสารสามารถจับสัญญาณภาพได้ที่ละหน้า และหากมีอุปกรณ์ที่เพิ่มเติม

เรียกว่า ตัวป้อนกระดาษ (Feeder) จะสามารถดึงกระดาษอย่างต่อเนื่องเข้าสู่เครื่องสแกนเนอร์ให้โดยอัตโนมัติ ช่วยให้การอ่านภาพเร็วขึ้น (ใช้ในกรณีที่มีการอ่านเอกสารจำนวนมาก)

สัญญาณดิจิทัลที่อ่านเข้ามาจะแปลงเป็นบิตแมพภาพแบบ 2 มิติ แสดงด้วยจุดภาพสีหรือภาพขาวดำ แต่สำหรับงานทางด้านความรู้จำตัวอักษร เราควรเลือกภาพขาวดำ เพื่อตัดปัญหาในเรื่องการแปลงสัญญาณภาพในขั้นตอน Binarization และฟังก์ชันของจุดภาพขาว  $F(i,j)$  ขนาด  $m \times n$  ซึ่ง  $i$  และ  $j$  ถูกกำหนดให้เป็นตำแหน่งใดๆ ในแนวนอนและแนวตั้งตามลำดับ ซึ่งตำแหน่งเริ่มต้นถูกกำหนดที่มุมบนซ้ายมือของภาพ ส่วน  $m$  และ  $n$  คือจำนวนจุดภาพในแนวแกนตั้งและแนวแกนนอนตามลำดับ



รูปที่ 35 แสดงอุปกรณ์และทิศทางของสัญญาณภาพที่ใช้ในกระบวนการรับข้อมูลในรูปแบบบิตแมพภาพ เพื่อให้สามารถนำไปวิเคราะห์ข้อมูลภาพในขั้นตอนการประมวลผลขั้นต้นต่อไป

สิ่งที่ควรคำนึงถึง ตัวอักษรที่จะนำมาใช้เป็นตัวอย่างแบบอักษรสำหรับการรู้จำตัวอักษรถือเป็นเรื่องจำเป็นเช่นกัน และไม่มีมาตรฐานในการกำหนดว่า ควรเลือกแบบอักษรในลักษณะใด เพื่อเก็บไว้ในฐานข้อมูล ดังนั้นการวิจัยนี้จะขอกำหนดแนวทางในการเลือกใช้แบบอักษรดังนี้

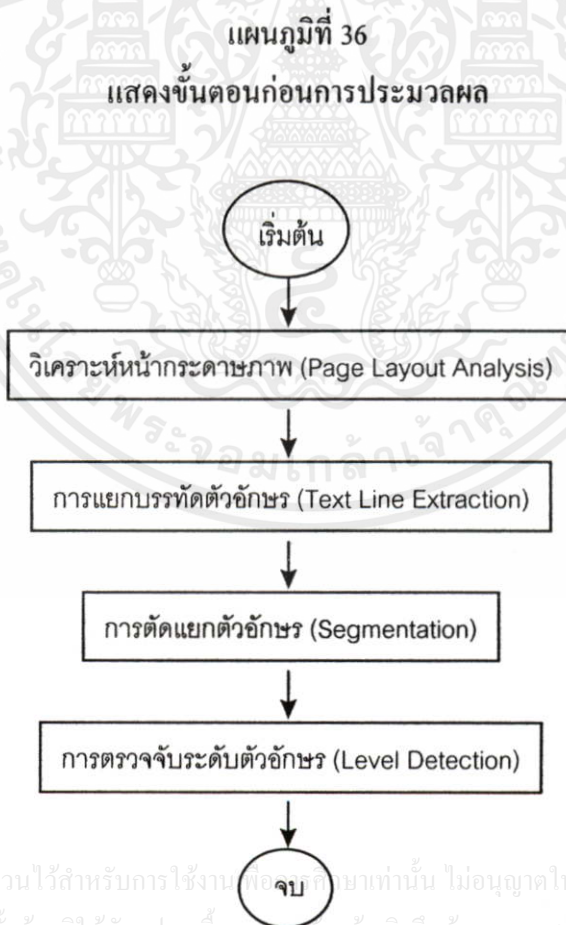
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ไม่ควรนำตัวอย่างแบบอักษรที่มีความคล้ายคลึงกันมากแต่มีรหัสแอสกีเดียวกัน เก็บในฐานข้อมูลเพื่อไม่ให้เกิดความซ้ำซ้อนของข้อมูลมากเกินไป

2. ชนิดของตัวอย่างแบบอักษรที่ใช้ควรเลือกใช้แบบอักษรที่ใช้งานบ่อยมากที่สุด เพื่อให้ครอบคลุมแบบอักษรมาตรฐานในท้องตลาดนั่นเอง

### 6.3 การประมวลผลขั้นต้น

หลังจากผ่านขั้นตอนการนำเข้าข้อมูล บิตแมพภาพที่ได้ถือว่าเป็นข้อมูลดิบ ต้องนำไปกรองหรือคัดเลือกลักษณะที่ต้องการออกมาก่อน ซึ่งจำเป็นต้องผ่านกระบวนการมากมาย เพื่อให้ได้ข้อมูลเฉพาะส่วนภาพที่ต้องการเท่านั้น จำเป็นต้องอาศัยขั้นตอนการประมวลผลขั้นต้น (Preprocessing) ดังแสดงดังในแผนภูมิที่ 36



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

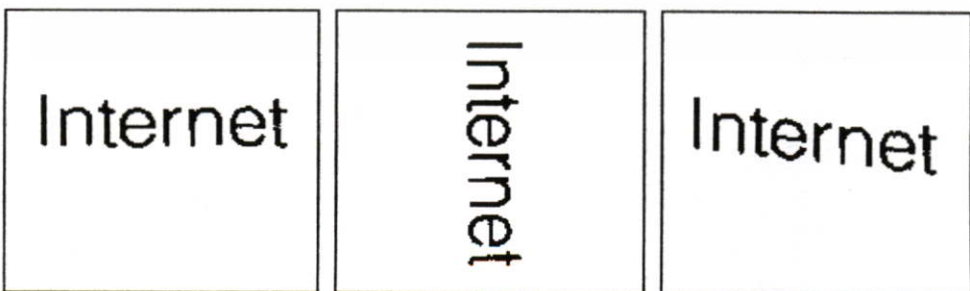
แผนภูมิที่ 36 จะแสดงขั้นตอนก่อนการประมวลผล เพื่อให้ได้กรอบของตัวอักษรตามต้องการ เนื่องจากภาพที่ผ่านการอ่านเข้ามาเป็นภาพของประโยคหรือข้อความ ซึ่งประกอบด้วยตัวอักษรแต่ละตัววางเรียงใกล้กันในแต่ละแถวเดียวกัน มีวิธีอย่างไรเพื่อแยกภาพตัวอักษรแต่ละตัวออกจากกัน และขั้นตอนต่างๆที่นำมาใช้จำเป็นต้องพิจารณาแหล่งข้อมูลของภาพที่นำมาใช้งาน หากภาพที่นำมาเป็นภาพสีจำเป็นต้องแปลงให้เป็นภาพขาวดำเสียก่อน เนื่องจากการรู้จำตัวอักษรไม่เหมาะกับภาพสีโดยทั่วไป และหากภาพที่เข้ามามีสัญญาณรบกวน จำเป็นต้องผ่านกระบวนการกรองเพื่อกำจัดสิ่งที่ไม่ต้องการเสียก่อน แต่เนื่องจากการวิจัยนี้เลือกที่จะแก้ปัญหาด้วยการคัดเลือกภาพของข้อความที่มีความสมบูรณ์เพียงพอและอ่านภาพเข้ามาเป็นขาวดำโดยตรง ทำให้ลดภาระการกรองสัญญาณรูปภาพ สิ่งที่ต้องทำจึงมีเพียงการวิเคราะห์หน้ากระดาษภาพ และการตรวจจับตัวอักษรแต่ละตัวออกจากกันเท่านั้น

### 6.3.1 การวิเคราะห์หน้ากระดาษภาพ (Page layout analysis)

การวิเคราะห์หน้ากระดาษมีความจำเป็น เนื่องจากกรณีภาพที่เข้ามาตัวอักษรกลับจากขวาเป็นคำ ตรงกันข้ามกับสิ่งที่ต้องการ บางครั้งภาพปกติตัวอักษรจะวางในแนวนอน แต่หากอ่านเข้ามาในแนวตั้งควรทำอย่างไร และหากบรรทัดของตัวอักษรวางเอียงควรแก้ไขอย่างไร

รูปที่ 37

แสดงภาพของข้อความที่ผ่านการอ่านด้วยเครื่องสแกนเนอร์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 37 จะแสดงภาพของข้อความในลักษณะต่าง ๆ ที่เกิดขึ้นภายหลังการอ่านเอกสาร ดังนั้นหากวิเคราะห์ จะเห็นว่าปัญหาที่เกิดขึ้นมีมากมาย สามารถลำดับเป็นข้อ ๆ ได้ดังนี้ [15]

1. การแปลงสัญญาณภาพกลับขาวดำ (Binarization)
2. บรรทัดของตัวอักษรในข้อความเกิดความเอียง (Document skew angle)
3. เอกสารที่ได้จากการสแกนอยู่ในแนวนอน (Landscape)
4. ภาพเอกสารมีรูปภาพ เส้นตาราง กราฟ แผนภูมิ ปะปนอยู่

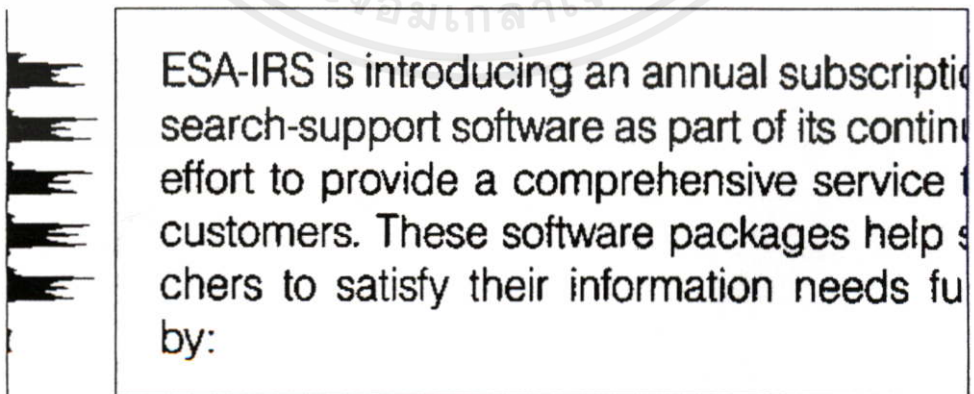
ขั้นตอนของการวิเคราะห์หน้ากระดาษภาพ จึงจำเป็นต้องแบ่งภาพของเอกสารโดยทั่วไป ออกเป็นส่วนๆ ดังนี้ ส่วนหัวข้อ (Headline area) เนื้อข้อความ (Textural area) ส่วนรูปภาพ (Picture area) ส่วนเชิงอรรถ (Footnote area) ส่วนพื้นหลัง (Background) ซึ่งมีวิธีการต่างๆ มากมาย [15,16,17] สำหรับวิทยานิพนธ์ฉบับนี้จะขอก้าวถึงเฉพาะขั้นตอนการแปลงสัญญาณภาพกลับขาวดำ เท่านั้น ด้วยวิธีการฉายภาพด้านข้าง (Projection profile) หรือฮิสโตแกรม (Histogram)

ขั้นตอนการกลับสัญญาณภาพขาวดำนั้น อาศัยการวิเคราะห์ภาพจากหน้ากระดาษทั่วไปโดย สมมุติฐานว่า จำนวนจุดภาพของตัวอักษรทั้งหมดในภาพ (ควรแสดงด้วยจุดดำ) จะมีจำนวนน้อยกว่า จำนวนจุดภาพในภาพทั้งหมด ดังนั้นหากจำนวนจุดดำในภาพมีจำนวนมากกว่าจุดขาว จะสันนิษฐาน ว่า บิตแมพภาพตัวอักษรในภาพเป็นสีขาว จึงควรกลับสัญญาณภาพขาวดำทันที

รูปที่ 38

แสดงวิธีการฉายภาพด้านข้าง

Horizontal Projection



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น

Vertical Projection

รูปที่ 38 จะแสดงวิธีการหาฮิสโตแกรมในแนวนอนและแนวตั้ง โดยอาศัยฟังก์ชันของจุดภาพแบบพิกเซล การฉายภาพจะแปลงภาพขาวดำไปเป็นข้อมูลอะเรย์ (Array) แบบหนึ่งมิติ ความถี่ของฮิสโตแกรมในแนวนอน  $H(j)$  ของภาพขาวดำ  $f(i,j)$  เกิดจากผลรวมของจุดดำที่ถูกฉายภาพเข้าไปในแกนแนวแกนตั้ง ส่วนความถี่ของฮิสโตแกรมในแนวตั้ง  $V(i)$  ของภาพขาวดำเกิดจากผลรวมของจุดดำที่ถูกฉายภาพเข้าไปในแนวแกนนอน ดังนั้นผลรวมจุดดำในภาพก็คือความถี่ของฮิสโตแกรมในแนวตั้งหรือแนวนอนนั่นเอง

### 6.3.2 การแยกบรรทัดตัวอักษร (Text line Extraction)

หลังจากผ่านวิธีการของการวิเคราะห์หน้ากระดาษ ทำให้ได้สัญญาณภาพของข้อความตัวอักษรที่ต้องการเท่านั้น จากนั้นจะใช้วิธีแยกบรรทัดของข้อความตัวอักษรแต่ละบรรทัดออกจากกัน โดยอาศัยการวิเคราะห์ความถี่ของฮิสโตแกรมในแนวนอน  $H(j)$  หากความถี่ของฮิสโตแกรมในแนวนอน ณ ตำแหน่ง  $j$  มีค่าเป็นศูนย์ แสดงว่าเกิดช่องว่างระหว่างบรรทัด จะทำการแยกส่วนประกอบเฉพาะของแต่ละบรรทัดออกมาดังแสดงในรูปที่ 39

รูปที่ 39

แสดงความถี่ของฮิสโตแกรมในแนวนอนของส่วนข้อความตัวอักษร

Horizontal Projection



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะในรูปแบบใดก็ตาม การนำข้อความตัวอักษรในแต่ละบรรทัดเก็บไว้ในอะเรย์ ระบุตำแหน่งมุมบนซ้าย (Top-left) และมุมล่างขวา (Bottom-right) เพื่อนำไปวิเคราะห์และให้แยกตัวอักษรแต่ละตัวออกจากกันต่อไป

### 6.3.3 การตัดแยกตัวอักษร (Segmentation)

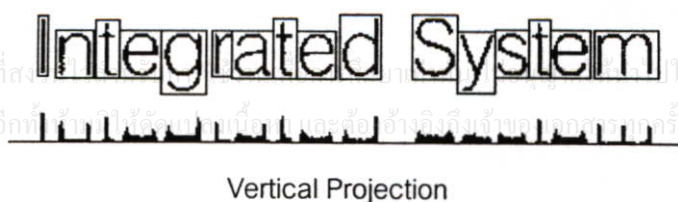
เมื่อแยกข้อความในแต่ละบรรทัดออกจากกันแล้ว จะเริ่มการตรวจจับการตัดแยกกรอบสี่เหลี่ยมในลักษณะเดียวกัน (Homogeneous block segmentation detection) [18] เก็บผลการแยกกรอบของตัวอักษรไว้ในตาราง ระบุตำแหน่งมุมบนซ้าย และมุมล่างขวาของสี่เหลี่ยมขนาดเล็กที่สุดที่ล้อมรอบตัวอักษร ด้วยวิธีการหาความถี่ของฮิสโตแกรม ดังแสดงในรูปที่ 40



รูปที่ 40 จะแสดงความถี่ของฮิสโตแกรมในแนวตั้ง  $V(i)$  เพื่อตัดแยกตัวอักษรแต่ละตัวออกจากกัน ทำให้ได้ตำแหน่งทางขอบซ้าย (Column-left) และขอบขวา (Column-right) ของตัวอักษร จากนั้นเราสามารถใช่วิธีเคม หาดำแหน่งของขอบบน (Row-top) และขอบล่าง (Row-bottom) ของตัวอักษร ด้วยวิธีการหาความถี่ของฮิสโตแกรมในแนวนอนอีกครั้ง ดังแสดงในรูปที่ 41

รูปที่ 41

แสดงวิธีการหาขอบซ้าย ขอบขวา ขอบบน ขอบล่าง ของตัวอักษร



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งยังขอสงวนสิทธิ์ในข้อมูลและข้ออ้างจึงจึงแจ้งขอสงวนสิทธิ์ในการนำไปใช้

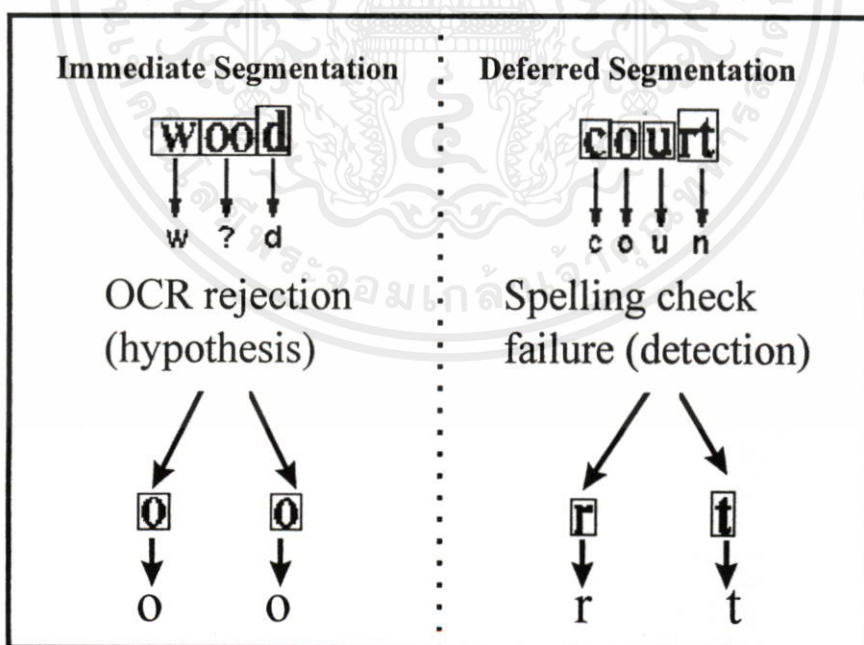
เก็บค่าตำแหน่งขอบตัวอักษรแต่ละตัว เพื่อไว้ใช้เป็นข้อมูลสำหรับการประมวลผลในภายหลัง ตอนนี้เราสามารถแยกตัวอักษรธรรมดาแต่ละตัวออกได้ แต่ในทางปฏิบัติตัวอักษรภาพที่อ่านเข้ามา มีโอกาสเรียงติดกันจนไม่สามารถแยกออกจากกันได้ ซึ่งจำเป็นต้องหาวิธีการแก้ปัญหาต่อไป

#### 6.3.4 การตรวจจับความเหลื่อมล้ำ (Overlap) และสัมผัสกัน (Touching detection)

การตัดแยกตัวอักษรที่สัมผัสกัน [19] อาจเกิดขึ้นจากชนิดของแบบอักษร ขนาดของแบบอักษร คุณภาพการพิมพ์ การถ่ายเอกสาร ข้อจำกัดเรื่องความละเอียดของเครื่องสแกนเนอร์ หรือหลักการใช้ค่าเทรคโธรลสำหรับ ไบนารีสูง เพื่อหลีกเลี่ยงการแตกแยกของตัวอักษร

รูปที่ 42

แสดงตัวอย่างการแยกตัวอักษรที่วางติดกัน



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

a) การตัดแยกตัวอักษร โดยทันที    b) การตัดแยกตัวอักษรแบบมีเงื่อนไข  
ไม่ว่ากรณีใดๆทางสน. อักษรทิมมิ่งเทคโนโลยีฯ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบแผนการใช้งานตัวอักษรติดกัน มีสมมุติฐานมาจากการตัดออก (Cutting) พยายามที่จะแยกแพทเทิร์นแต่ละตัวออกจากกัน โดยวิธีการแบ่งแยกส่วนประกอบ (Decomposed) เทคนิคการแยกถูกแสดงด้วยภาพดังแสดงในรูปที่ 42 ที่ซึ่งกรอบของตัวอักษรที่สัมผัสกันถูกตัดออก และการแยกส่วนประกอบในกรอบถูกทำให้แน่ใจโดยการแยกประเภทส่วนประกอบทางซ้ายและทางขวา

รูปที่ 42(a) การตัดแยกตัวอักษร โดยทันทีสำหรับส่วนประกอบของคำ ส่วนประกอบของตัวอักษรตั้งแต่สองตัวจำเป็นต้องตัดเพื่อจะแยกตัวอักษร หลักการที่ใช้จะเกี่ยวข้องกับสัดส่วนความยาวของตัวอย่างแพทเทิร์นที่เข้ามา ถ้าความยาวของกรอบสี่เหลี่ยมล้อมรอบตัวอักษรเกินกว่าเปอร์เซ็นต์ของความสูงของแพทเทิร์นนั้น จะถูกพิจารณาว่ามีส่วนประกอบของตัวอักษรหลายตัว ความเป็นไปได้ของการแยกส่วนประกอบออกมีความจำเป็นในความหมายของการพิจารณาทางสถิติ

คุณลักษณะพื้นฐานของการรู้จำตัวอักษร โดยทั่วไป มีปัญหาเกี่ยวกับการตัดตัวอักษรที่ติดกันออก เนื่องจากอาจมีรูปร่างเหมือนกับตัวอักษรบางตัว เช่น ไม่สามารถแยกระหว่าง "ur" และ "w" หรือ "m" และ "n" โดยปราศจากการเรียนรู้คู่สัมผัสของตัวอักษรจำนวนมาก ดังนั้นเป็นไปได้ที่จะตัดแพทเทิร์นที่ไม่เข้าข่ายที่กำหนดไว้ ออก และสมมุติว่าตัวอักษรถูกทำให้ติดกัน โดยอาศัยการตัดแยกตัวอักษรแบบมีเงื่อนไขดังแสดงในรูปที่ 42(b)

ข้อสังเกตบางประการสำหรับปัญหาพื้นที่ ที่มีโอกาสเกิดขึ้นได้ โดยไม่สามารถตรวจจับด้วยโปรแกรมตรวจสอบการสะกดคำ เช่น ตัวอย่างของคำ "close" อาจจะอ่านเป็น dose เมื่อ "c" และ "l" สัมผัสกัน และ "stem" อาจจะอ่านเป็น "stem" เมื่อ "r" และ "n" สัมผัสกัน ดังนั้นควรพิจารณาตัวอักษร "m", "w" และ "d" เป็นพิเศษ ดังแสดงในรูปที่ 42(b)

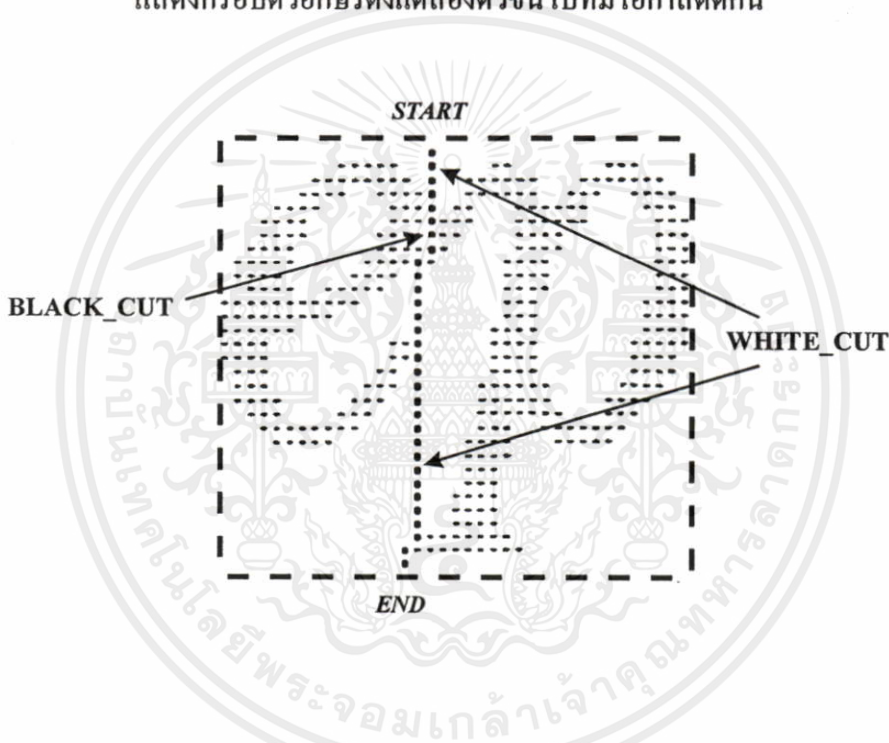
หลักการแยกตัวอักษรติดกันเป็นงานวิจัยที่มีการกล่าวถึงบ้างพอสมควร วิธีและอัลกอริทึมจะแตกต่างกันไป เท่าที่พบพอจะอ้างอิงได้ 5 วิธีดังนี้

1. การตัดแยกด้วยวิธีการฉายภาพด้านข้าง (Projection profile) [20]
2. การตัดแยกด้วยดิสคริมิเนทฟังก์ชัน (Discriminate function) [11]
3. การตัดแบ่งลายเส้น (Vertex directed segmentation) [21]
4. การตัดแยกด้วยการวิเคราะห์โครงสร้าง (Structural analysis segmentaion method) [22]
5. การตัดแยกด้วยเส้นทางสั้นที่สุด Shortest path segmentaion) [19]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
โดยทั่วไปอาจใช้หลักการหาความถี่ของฮิสโตแกรมในแนวแกนตั้งเป็นเครื่องมือวิเคราะห์เลือกค่าความถี่ที่น้อยที่สุด ณ บางจุด และตัดเป็นเส้นตรงเสมอ แต่วิธีการฉายภาพด้านข้างในแนวแกนตั้งจะแปลงข้อมูลของส่วนประกอบแบบ 2 มิติส่งไปยังข้อมูลแบบ 1 มิติ อาจนำไปสู่การ

ประมาณที่ไม่ถูกต้องของจุดที่ใช้ตัดได้ ดังเช่นตัวอย่างตัวอักษรที่สัมผัสกันแบบสมมาตร หรือ ตัวอย่างตัวอักษรตัวเอนที่สัมผัสกัน ซึ่งจะก่อให้เกิดความผิดพลาดในการรู้จำตัวอักษรได้ ดังนั้นอัลกอริทึมใหม่ถูกเสนอเพื่อค้นหาเส้นทางที่ใช้ตัดแบ่งตัวอักษรที่ติดกัน โดยใช้เส้นทางที่สั้นที่สุด (Shortest path) ดังแสดงในรูปที่ 43 จะแสดงกรอบของตัวอักษรตั้งแต่สองตัวที่วางติดกันและเหลื่อมล้ำกัน ขนาด  $n \times m$

รูปที่ 43  
แสดงกรอบตัวอักษรตั้งแต่สองตัวขึ้นไปที่มีโอกาสติดกัน

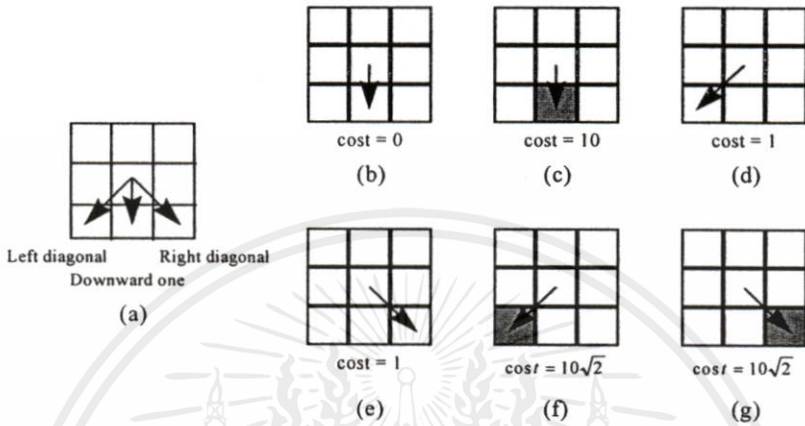


พิจารณาคู่ของตัวอักษรที่สัมผัสกันของบิตแมพขนาด  $n$  แถว  $x$   $m$  คอลัมน์ หรือ  $n \times m$  ให้จุดเริ่มต้น (START) และจุดสิ้นสุด (END) เป็นสองจุดที่อยู่ภายนอกกรอบตัวอักษร และให้ cost จาก START ไปยังจุดใดๆ ในแถวแรก หรือจุด END ไปยังจุดใดๆ ในแถวสุดท้ายเป็นศูนย์

เส้นทางจาก START ไปยัง END ผ่านหลายๆจุดภาพในบิตแมพแพทเทิร์น การวิเคราะห์จะเริ่มจากการกำหนดทิศทางเคลื่อนจากจุดภาพจุดหนึ่งไปยังจุดภาพถัดไปในสามทิศทาง ทิศทางลงล่าง (Downward one) และทิศทางทแยงซ้าย (Left diagonal) และทิศทางทแยงขวา (Right diagonal) ดังแสดงในรูปที่ 44 (a) และจะบันทึกผลรวมของค่า cost ตามลำดับของการเคลื่อนที่ในแต่ละแถว ไม่ว่าจะณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดเปลี่ยนเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 44

## แสดงการกำหนดทิศทางการเคลื่อนที่ในรูปแบบเมตริก



สำหรับบิตแมพแพทเทิร์นขนาด  $n$  แถว การตัดจะเกิดขึ้นเมื่อเคลื่อนที่ไปทั้งหมด  $n$  ครั้ง จากบิตแมพแถวแรกไปจนถึงแถวสุดท้าย ผลของการเคลื่อนที่ที่จะถูกเก็บในเมตริกของการตัด ซึ่งจะบรรจุค่าสองส่วน BLACK-CUT ตัดผ่านจุดดำ และ WHITE-CUT ตัดผ่านจุดขาว หลักการที่นำมาใช้สำหรับการตัดที่ดีที่สุด ควรเคลื่อนที่ผ่านเซลล์สีดำเพียงเล็กน้อยเท่าที่เป็นไปได้ และเคลื่อนที่ที่แพงให้น้อยที่สุดเท่าที่เป็นไปได้ หมายความว่า ค่า  $cost$  ที่ใหญ่กว่าควรจะถูกกำหนดเมื่อเคลื่อนที่ตรงไปยังจุดภาพสีดำ และเล็กกว่าถ้าตรงไปยังจุดภาพสีขาว การกำหนดค่า  $cost$  สำหรับเส้นทางในแต่ละคอลัมน์ สามารถกำหนดดังนี้ [19]

- ไม่มี  $cost$  ถ้าเคลื่อนที่ตรงไปยังจุดภาพสีขาวดังแสดงในรูปที่ 44 (b) และถ้าเคลื่อนที่ไปยังจุดภาพสีดำจะให้ค่าเป็น 10 ดังแสดงในรูปที่ 44 (c)
- การเคลื่อนที่ที่แพงไปยังจุดภาพสีขาวจะให้ค่า  $cost$  เป็น 1 ดังแสดงในรูปที่ 44 (d) และรูปที่ 44 (e) และถ้าเคลื่อนที่ที่แพงไปยังจุดภาพสีดำ จะให้ค่า  $cost$  เป็น  $10\sqrt{2}$  ดังแสดงในรูปที่ 44 (f) และรูปที่ 44 (g)
- การเคลื่อนที่จาก START เป็นการเคลื่อนที่ลงล่างเสมอ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการเรียนการสอนเท่านั้น ไม่ควรนำออกไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การคำนวณ  $cost$  ที่สัมพันธ์กับเส้นทางที่สั้นที่สุดควรบรรจุในอะเรย์ขนาด 2 มิติ ตามแนวคอลัมน์  $j$  โดยกำหนดค่าเริ่มต้นสำหรับ  $cost$  เป็นศูนย์ เฉพาะที่แถวแรกค่าอะเรย์  $cost$  จะมีค่าเป็น 0

(10) ถ้าจุดภาพในแถวแรกของคอลัมน์  $j$  เป็นสีขาว (สีดำ) ส่วนแถวอื่นๆ ค่า cost จะขึ้นอยู่กับเส้นทางที่ดีที่สุด ในสามเส้นทางที่กำหนดก่อนหน้านี้ (ลงล่าง ทแยงซ้าย และทแยงขวา) จำนวนรอบที่ใช้ในการคำนวณ สำหรับ  $n \times m$  บิตแมพจะเป็น  $(nm)$  รอบ

การคำนวณหาค่า cost นั้น จำเป็นต้องนำค่าเมทริกซ์ที่ตำแหน่งปัจจุบันขนาด  $3 \times 3$  ไปเก็บไว้ในอะเรย์ก่อน เพื่อนำไปเปรียบเทียบกับฐานข้อมูลของเส้นทางเลือกทั้งหมด ก่อนที่จะกำหนดทิศทางการเคลื่อนที่ถูกต้องในตำแหน่งต่อไป

รูปที่ 45(a) ถึง รูปที่ 45(n) จะแสดงลำดับการตรวจสอบเพื่อหาเส้นทางที่ถูกต้อง เครื่องหมาย 1 จะถูกแทนด้วยจุดภาพสีดำ และเครื่องหมาย 0 สำหรับจุดภาพสีขาว ส่วนเครื่องหมาย x จะถูกแทนด้วยค่าอะไรก็ได้ระหว่าง 0 หรือ 1 เพียงค่าเดียว

รูปที่ 45(z) จะแสดงตำแหน่งอ้างอิงของจุดภาพที่ใช้กำหนดทิศทางการเคลื่อนที่ หลักการเคลื่อนที่ที่จะพยายามเคลื่อนในทิศทางไกลจากเครื่องหมาย 1 ให้มากที่สุดเท่าที่เป็นไปได้ โดยไม่มีการเคลื่อนที่ในทิศทางขึ้นบนหรือซ้ายขวาเป็นอันขาด

ทางเลือกทั้งหมดสำหรับรูปที่ 45(a) ถึง รูปที่ 45(n) มีโอกาสเป็นไปได้ครอบคลุมรูปทั้งหมดที่เกิดขึ้นจริงใน 5 ตำแหน่ง (lh, ld, vd rd และ rh) รวม  $2^5 = 32$  รูป

### รูปที่ 45

แสดงการตรวจหาเส้นทางที่ถูกต้อง

lh		rh
ld	vd	rd

(z)

0	/	0
0	0	1

(a)

0	\	0
1	0	0

(b)

0	/	1
0	0	1

(c)

1	\	0
1	0	0

(d)

x		x
x	0	x

(e)

0	/	0
0	1	0

(f)

1		1
x	1	x

(g)

0		0
1	1	1

(h)

0		1
1	1	x

(i)

1		0
x	1	1

(j)

0	/	1
0	1	x

(k)

1	\	0
x	1	0

(l)

0	/	0
0	1	1

(m)

0	\	0
1	1	0

(n)

lh = Left Half

vd = vertical Down

rh = Right Half

ld = Left Down

rd = Right Down

อะเรย์ของเส้นทางการเคลื่อนที่ทั้งหมดในทุกคอลัมน์ และอะเรย์สำหรับค่า cost ในแต่ละคอลัมน์จะถูกบันทึกไว้ และนำไปใช้งานกับวิธีการคำนวณหาเส้นแบ่งตัวอักษรที่วางเหลื่อมล้ำกัน หรือสัมพันธ์กันต่อไป

#### 6.3.4.1 ตัวอักษรเหลื่อมล้ำกัน (Overlap character)

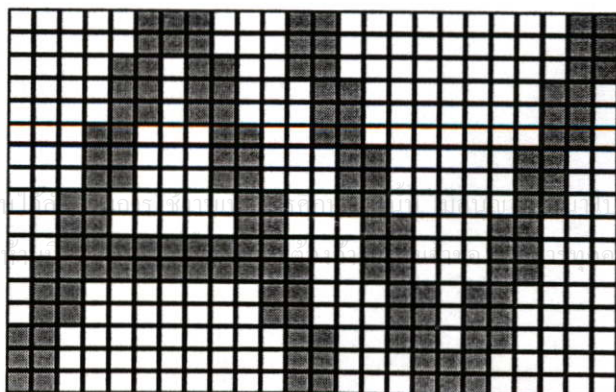
ตัวอักษรแบบบิตแมพแพทเทิร์นที่ได้จากเครื่องสแกนเนอร์จะผ่านกระบวนการแปลงสัญญาณของแสงที่ตกกระทบตัวอักษรให้เป็นสัญญาณดิจิทัล ซึ่งสามารถแสดงเป็นรูปภาพในช่องสี่เหลี่ยมทางหน้าจอได้ ดังนั้นหากตัวอักษรที่มีความชิดกันมาก คุณภาพของภาพตัวอักษรไม่ดีพอ หรือความละเอียดในการอ่านเอกสารมีขนาดน้อยเกินไป ตัวอักษรในภาพที่นำมาแสดงผลอาจมีโอกาสดกกันได้ วิธีตรวจสอบว่ามีตัวอักษรวางติดกันหรือไม่ ควรตรวจสอบจากขนาดกรอบของวัตถุที่ผ่านกระบวนการตัดแยกตัวอักษรมาเรียบร้อยแล้ว หากค่าความสูงของกรอบตัวอักษรทั้งบรรทัด (sl) น้อยกว่าสองเท่าของความกว้างของกรอบตัวอักษรที่พิจารณา (sc) จะตัดสินใจว่ากรอบตัวอักษรที่พิจารณาควรจะมีตัวอักษรที่วางติดกัน และเริ่มจะคำนวณหาค่า cost ในแต่ละคอลัมน์ เพื่อนำไปวิเคราะห์หาเส้นแบ่งตัวอักษร ดังแสดงในรูปที่ 46

รูปที่ 46

แสดงค่า cost ในแต่ละคอลัมน์

cost 

34	33	34	35	36	47	30	17	6	5	6	17	35	24	23	22	21	20	21	22	23	24	35	33
----	----	----	----	----	----	----	----	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----



เอกสารนี้เป็นเอกสารที่สงวน  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้ง

ใช้ประโยชน์ด้านการค้า  
ซึ่งที่มีการนำไปใช้

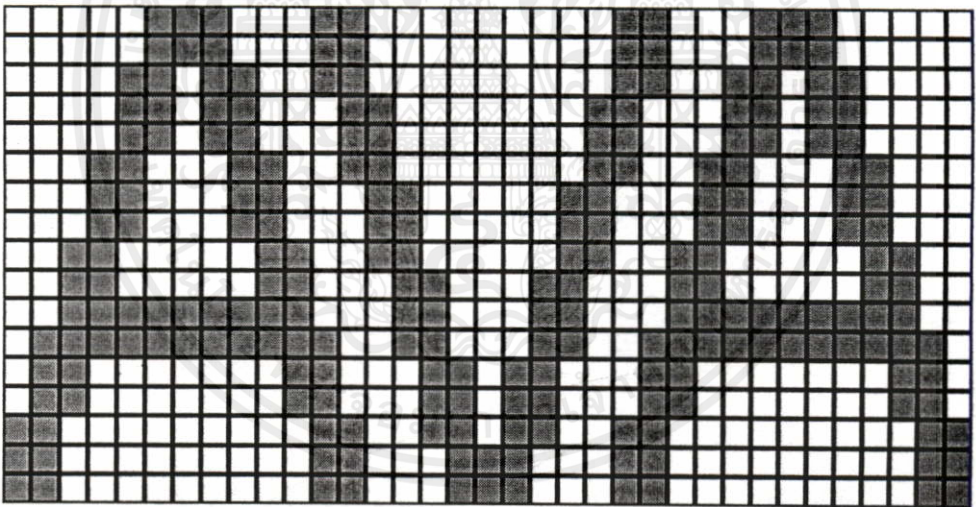
หากมีค่า cost เท่ากันมากกว่าหนึ่งค่าอยู่ใกล้กัน เราจะเลือกเพียงค่าเดียวจากเส้นทางที่ไม่ซ้อนทับกัน และกำหนดให้ค่าของมันเป็นจุดเริ่มต้นของเส้นทางที่ดีที่สุด ข้อสำคัญค่า cost ที่ได้ไม่ควรจะมีค่าเกินกว่า 10 เนื่องจากไม่มีการเคลื่อนที่ผ่านจุดภาพสีค่าเกิดขึ้นเลย

สำหรับตัวอักษรมากกว่าสองตัวที่วางเหลื่อมล้ำกันก็จะใช้วิธีพิจารณาแบบเดียวกัน เพียงแต่หาจุดต่ำสุดในบริเวณใกล้เคียงกันเพื่อใช้เป็นจุดเริ่มต้นของเส้นทางแบ่งตัวอักษร ดังแสดงในรูปที่ 47

รูปที่ 47

แสดงตัวอักษรที่เหลื่อมล้ำกันสามตัว

cost	34	33	34	35	36	47	30	17	6	5	6	17	35	24	23	22	21	20	21	22	23	24	35	33	17	6	5	6	17	30	47	36	35	34	33	34
------	----	----	----	----	----	----	----	----	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	---	---	---	----	----	----	----	----	----	----	----



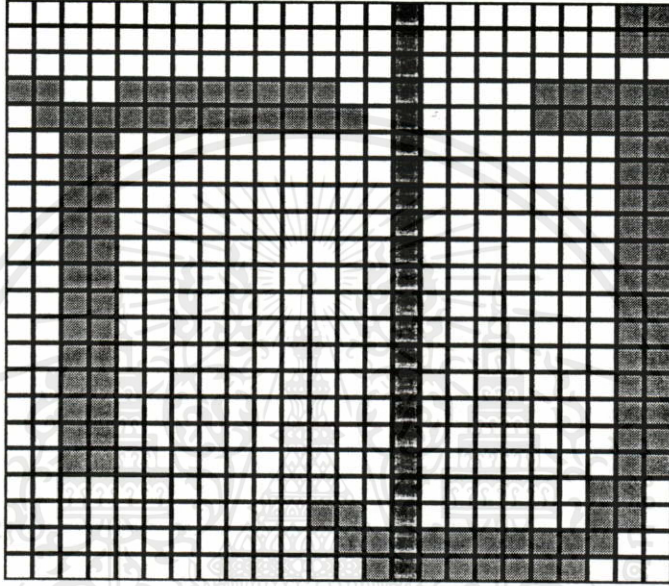
การหาเส้นทางที่สั้นที่สุดเป็นแนวทางที่ดีสำหรับแบ่งตัวอักษรที่วางเหลื่อมล้ำกัน แต่ยังคงพบจุดด้อยบางประการ เนื่องจากกฎเกณฑ์ที่วางไว้กำหนดให้เคลื่อนที่ได้ในสามทิศทางเท่านั้น ไม่สามารถเคลื่อนที่ไปทางซ้าย ขวาได้ ทำให้มีโอกาสเกิดกรณีผิดพลาดได้ดังแสดงในรูปที่ 48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 48

แสดงการแบ่งตัวอักษรสำหรับกรณีที่เกิดพลาด

cost 25 24 23 22 23 24 20 20 20 20 21 23 22 21 20 20 20 20 21 22 41 42 53 171



การแบ่งตัวอักษรในลักษณะนี้ สามารถก่อให้เกิดปัญหาได้บ้าง แต่โซลูชันที่พบน้อยมาก แนวทางแก้ไขอาจจะต้องใช้วิธีพิจารณาจุดล้อมรอบตัวอักษรและแยกชิ้นของวัตถุตัวอักษรในภาพออกมา ซึ่งจะพบว่าได้ผลดีสำหรับบางกรณี

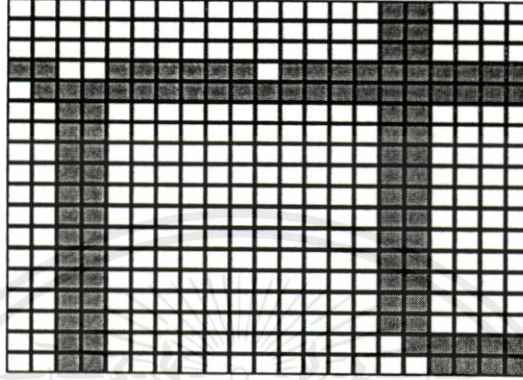
#### 6.3.4.2 ตัวอักษรสัมผัสกัน (Touching character)

หลักการตรวจสอบตัวอักษรที่สัมผัสกัน จะใช้วิธีเดียวกันกับวิธีตรวจสอบตัวอักษรเหลื่อมล้ำกันด้วยการหาเส้นทางที่สั้นที่สุด เพื่อแบ่งแยกตัวอักษรออกจากกัน โดยพิจารณาจากค่า cost ในแต่ละคอสต์ คอสต์ที่มีค่าน้อยที่สุดจะถูกกำหนดให้เป็นจุดเริ่มต้นของการแบ่งตัวอักษร ดังแสดงในรูปที่ 49

## รูปที่ 49

แสดงค่า cost ในแต่ละคอลัมน์

COST	25	24	23	34	35	20	20	20	20	11	10	11	20	20	21	32	52	41	40	40	41
------	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----



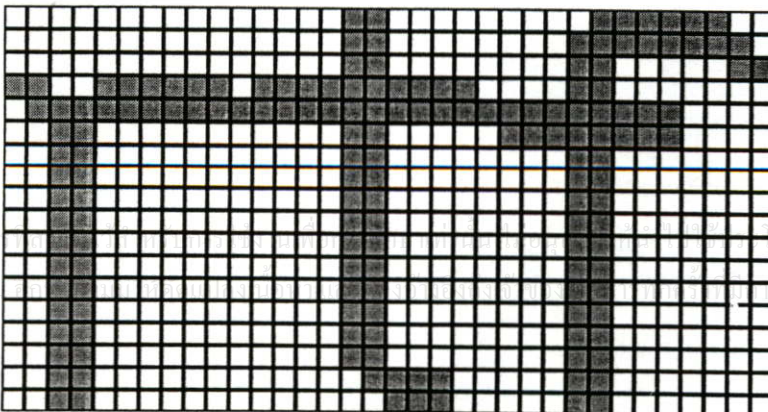
จากรูปที่ 49 จะสังเกตว่าค่า cost จะมีค่ามากกว่า 10 เสมอ เนื่องจากเส้นทางจะตัดผ่านจุดภาพสี่ค่า หากมีค่า cost เท่ากันมากกว่าหนึ่งค่าอยู่ใกล้กัน เราจะเลือกค่าเพียงค่าเดียวจากเส้นทางที่ไม่ซ้อนทับกัน เพื่อเป็นจุดเริ่มต้นของเส้นทางแบ่งตัวอักษร

สำหรับตัวอักษรมากกว่าสองตัวสัมพันธ์กันจะใช้วิธีพิจารณาในลักษณะเดียวกัน หากค่า cost ต่ำสุดในบริเวณใดบริเวณหนึ่ง เพื่อใช้เป็นจุดเริ่มต้นของเส้นทางแบ่งตัวอักษร ดังแสดงในรูปที่ 50

## รูปที่ 50

แสดงค่า cost ในแต่ละคอลัมน์สำหรับตัวอักษรสามตัวติดกัน

COST	25	24	23	34	35	20	20	20	11	10	11	20	20	21	32	52	41	40	40	22	21	20	20	21	22	52	41	40	22	21	20	12	11
------	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----



การหาเส้นทางที่สั้นที่สุดเพื่อให้แบ่งตัวอักษรที่สัมผัสกันบางครั้งก็มีปัญหาเกิดขึ้น เนื่องจากเส้นทางที่สั้นที่สุด ณ จุดที่พิจารณา อาจไม่ใช่จุดที่ถูกต้องสำหรับแบ่งตัวอักษร ดังแสดงในรูปที่

51

## รูปที่ 51

แสดงการเปรียบเทียบค่า cost สำหรับกรณีการตัดที่ไม่ถูกต้อง



a) การแบ่งแยกตัวอักษรในกรณีผิดพลาด    b) การแบ่งแยกตัวอักษรหลังการแยกประเภท

การตัดสำหรับค่า cost ที่น้อยที่สุดอาจจะไม่ถูกต้องเสมอไปดังแสดงในรูปที่ 51(a) ดังนั้นหลังจากการตัดตัวอักษรที่ได้จากการแยกประเภทตัวอักษรออก จำเป็นต้องเลือกการตัดแยกตัวอักษรใหม่อีกครั้งหนึ่ง โดยเลือกเส้นแบ่งตัวอักษรใหม่ไม่ซ้ำกับเส้นทางเดิมดังแสดงในรูปที่ 51(b) ทำเช่นนี้ไปเรื่อยๆ จนกระทั่งการรู้จำตัวอักษรมีผลที่ถูกต้อง

สิ่งที่ควรพิจารณาเพิ่มเติมสำหรับการหาค่า cost ที่ดีที่สุด มีอยู่ 2 ประการคือ

1. ค่า cost ที่ดีที่สุดอาจมีหลายค่าอยู่ใกล้กันหรือติดกัน หรือมีเส้นทางซ้อนทับกันบางส่วน ดังนั้นควรหลีกเลี่ยงและเลือกเส้นทางที่ดีที่สุดเพียงเส้นทางเดียว
2. ตัวอักษรที่สัมผัสกันมีมากกว่า 2 ตัว ดังนั้นการหาค่า cost ที่ดีที่สุดควรมีมากกว่าหนึ่งค่า สิ่งที่ต้องทำคือแยกตัวอักษรทีละตัวออกมาเสียก่อน ส่วนที่เหลือจึงหาทางแยกอีกครั้ง โดยการหา cost ที่ดีที่สุดอีกครั้งหนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากการแบ่งแยกแพทเทิร์นตัวอักษรแต่ละตัวออกจากกันแล้ว เราจะได้กรอบตัวอักษรที่สามารถนำไปแยกประเภทในขั้นตอนการแยกประเภทตัวอักษรได้ทันที แต่ปัญหามีอยู่ว่าตัวอักษร



Upper baseline เป็นตำแหน่งสูงสุดของคำ เกิดจากความสูงของตัวอักษรตัวใหญ่ ระยะทางระหว่าง Lower line และ Upper line จะอ้างอิงด้วย Cap-Height ขณะที่ระยะทางระหว่าง Middle baseline และ Lower baseline ถูกเรียกว่า X-Height หรือ Middle zone พื้นที่ระหว่าง Middle line และ Upper line รู้จักในชื่อ Middle zone และ Upper zone พื้นที่ของ Cap-width เป็นความกว้างของตัวอักษรโดยเฉลี่ย

กำหนดให้ตัวอักษรของคำถูกแบ่งออกเป็นสามส่วน พื้นที่ส่วน Upper bound, Middle bound, และ Lower bound ตำแหน่งของทั้งสามช่วงจะขึ้นอยู่กับคอคอบด้านบนและขอบด้านล่างของกรอบตัวอักษร โดยใช้เป็นเครื่องมือกำหนดสมาชิกเจ็ดกลุ่มของตัวอักษรตัวพิมพ์ หรือแบ่งเป็นสองประเภทใหญ่ๆ ได้ดังนี้ [13]

(a) สมาชิกกลุ่มแรก (Categories I)

- Ascender : ตัวอักษรอยู่ในขอบเขต Upper zone และ Middle zone
- Centered : ตัวอักษรอยู่ในขอบเขต Middle zone เท่านั้น
- Descended : ตัวอักษรอยู่ในขอบเขต Middle zone และ Lower zone
- Full-Size : ตัวอักษรอยู่ในขอบเขต Upper zone, Middle zone และ Lower zone

(b) สมาชิกกลุ่มที่สอง (Categories II)

- Superscript : ตัวอักษรอยู่ในขอบเขตเหลื่อมขึ้นไปด้านบน
- Subscript : ตัวอักษรอยู่ในขอบเขตเหลื่อมลงไปด้านล่าง
- Internal : ตัวอักษรอยู่ในขอบเขต Middle zone แต่อยู่ต่ำกว่าจุดกึ่งกลางบรรทัด

ตำแหน่งเส้นที่ใช้กำหนดระดับของตัวอักษร (c) มีด้วยกัน 4 เส้น :

Upperline<sub>c</sub> เส้นกำหนดจุดเริ่มต้นของตัวอักษรช่วงบน

Upperhalf<sub>c</sub> เส้นกำหนดจุดเริ่มต้นของตัวอักษรช่วงกลาง

Lowerhalf<sub>c</sub> เส้นกำหนดจุดสิ้นสุดของตัวอักษรช่วงกลาง

Lowerline<sub>c</sub> เส้นกำหนดจุดสิ้นสุดของตัวอักษรช่วงล่าง

และสามารถคำนวณความกว้างของ Upper, Middle และ Lower Zone ได้ดังนี้

$$U_C = \text{Upperhalf}_C - \text{Upperline}_C \dots\dots\dots (76)$$

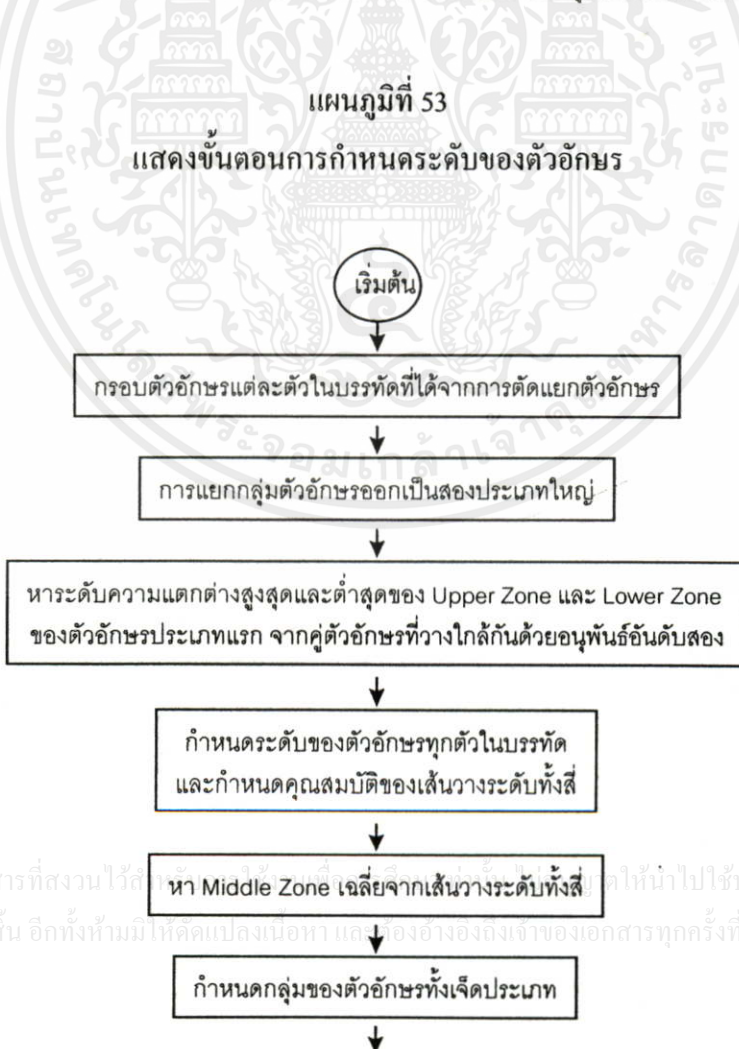
$$M_C = \text{Lowerhalf}_C - \text{Upperhalf}_C \dots\dots\dots (77)$$

$$L_C = \text{Lowerline}_C - \text{Lowerhalf}_C \dots\dots\dots (78)$$

โดยตัวอักษรทุกตัวที่จะนำเข้ามาจะถูกรวบรวมระดับของตัวอักษร และสร้างกรอบความสูงมาตรฐานให้แก่ทุกตัวอักษร เพื่อกำหนดคุณสมบัติพื้นฐานที่ถูกต้อง ทำให้การจัดประเภทกลุ่มของตัวอักษรตัวพิมพ์เข้าไปเป็นชุกย่อยหลายชุด ช่วยให้ผลการรู้จำดีขึ้น

การวิเคราะห์ตัวอักษรในแต่ละบรรทัดควรมีมากกว่า 1 พยางค์ และถ้าให้ตีควรมีมากกว่า 3 พยางค์ขึ้นไปในแต่ละบรรทัด และควรมีตัวอักษรหลายประเภท (Ascender, Descender เป็นต้น) ในแต่ละบรรทัดจะช่วยให้การวิเคราะห์ตัวอักษรถูกต้องมากยิ่งขึ้น

ขั้นตอนในการกำหนดระดับและแยกสมาชิกออกเป็นเจ็ดกลุ่มทำได้ดังในแผนภูมิที่ 53



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในเท่านั้น ไม่ควรเผยแพร่หรือใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### แผนภูมิที่ 53 (ต่อ)



หลังจากขั้นตอนการคัดแยกตัวอักษรทุกตัวภายในบรรทัด จะได้กรอบตัวอักษรแต่ละตัว เรียงตามแนวนอนจากซ้ายไปขวา จากนั้นจะเริ่มขั้นตอนการแยกประเภทกลุ่มตัวอักษรจากระดับตัว อักษรภายในบรรทัด

#### 6.3.5.1 การแยกประเภทกลุ่มตัวอักษรออกเป็นสองประเภทใหญ่

เริ่มแรกจะนำข้อมูลตำแหน่งกรอบของตัวอักษรทุกตัวในบรรทัดที่พิจารณามาเปรียบเทียบกับ ขนาดความสูง (d) เพื่อแยกประเภทกลุ่มของตัวอักษรออกเป็นสองประเภทใหญ่ ดังแสดงในรูปที่ 54

รูปที่ 54

แสดงการกำหนดระดับด้วยการแยกประเภทกลุ่มตัวอักษร

## Character Recognition

ความสูงของกลุ่มตัวอักษรในประเภทที่สอง (Superscript, Subscript และ Internal) จะมีค่าน้อยมากเมื่อเทียบกับความสูงของกรอบบรรทัด (sd) ดังนั้นเราจึงอาศัยการเปรียบเทียบค่า d และ sd โดยกำหนดให้ดังนี้

หากค่าสองเท่าของค่า d น้อยกว่าค่า sd ตัวอักษรตัวนั้นควรแยกไปอยู่ในกลุ่มตัวอักษรประเภทที่สอง ส่วนที่เหลือจะเป็นตัวอักษรประเภทที่หนึ่งทั้งหมด เขียนในเชิงโปรแกรมได้ดังนี้

if ( $2*d < sd$ ) then Categories II  
 else Categories I

จากนั้นจะเก็บข้อมูลตำแหน่งของกรอบตัวอักษรประเภทที่หนึ่งไว้ในอะเรย์ เพื่อนำไปวิเคราะห์หาระดับตัวอักษรต่อไป

### 6.3.5.2 การวิเคราะห์เพื่อหา Upper Zone และ Lower Zone ของบรรทัดตัวอักษร

โดยทั่วไปตัวอักษรตัวพิมพ์ภาษาอังกฤษแต่ละตัวจะมีการแบ่งโซนเป็น Upper Zone และ Lower Zone จากตำแหน่งการวางตัวอักษรที่แน่นอนดังแสดงในรูปที่ 55



ระดับของตัวอักษรจะกำหนดด้วยเส้นตรงสี่เส้น (Upper line, Upper half, Lower half, และ Lower line) ในทางทฤษฎีเส้นตรงดังกล่าวจะขนานกับแกนในแนวนอน และตำแหน่งของจุดเริ่มต้นจุดสิ้นสุด ทั้งช่วงบน ช่วงกลาง และช่วงล่าง จะพอดีกับตำแหน่งของเส้นตรงทั้งสี่ ส่วนในทางปฏิบัติกรอบตัวอักษรทุกตัวที่ผ่านการตัดแยกตัวอักษรจะไม่พอดีกับในระดับเดียวกันเสมอไป ดังแสดงตัวอย่างในรูปที่ 56

รูปที่ 56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้งานเพื่อการศึกษานานาชาติ ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า แสดงตัวอักษรที่ได้จากการตัดแยกตัวอักษร ไม่ว่าจะวิธีใดทั้งสิ้น อีกทั้งห้ามมิให้คิดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Character Recognition**

ในรูปมีปัญหาอยู่สองกรณี กรณีแรกบรรทัดตัวอักษรที่ได้จากเครื่องสแกนเนอร์เอียงทำมุมกับแนวนอนเล็กน้อย กรณีที่สอง ตำแหน่งของจุดเริ่มต้นและจุดสิ้นสุดทั้งช่วงบน ช่วงกลาง และช่วงล่างของแต่ละตัวอักษรจะไม่พอดีกับตำแหน่งในแนวเส้นตรงทั้งสิ้น

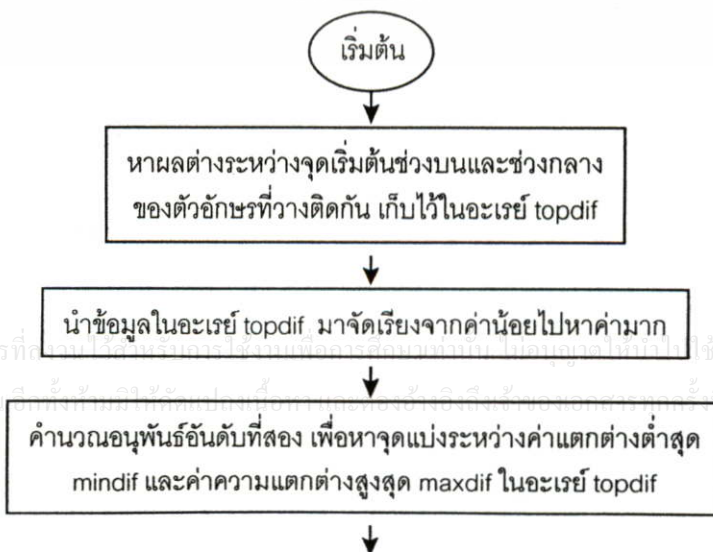
ดังนั้นการวิเคราะห์เพื่อแบ่งโซนให้กับบรรทัดตัวอักษรไม่สามารถใช้วิธีธรรมดาได้ การแก้ปัญหาในทั้งสองกรณีจึงต้องอาศัยหลักการบางอย่างเข้าช่วย

### 6.3.5.3 ผลต่างระดับตัวอักษรที่ติดกัน

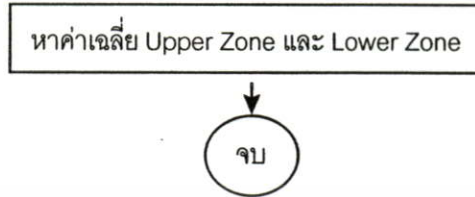
แม้ว่าบรรทัดของตัวอักษรที่ได้จากการสแกนจะทำมุมเอียงกันแกนในแนวนอน แต่ความแตกต่างระหว่างจุดเริ่มต้นช่วงบน (Upperline) และช่วงกลาง (Upperhalf) ของตัวอักษรที่วางติดกันไม่ได้เปลี่ยนแปลงไปเลย หรือเปลี่ยนแปลงน้อยมาก ดังนั้นการวิเคราะห์จึงอาศัยผลต่างของข้อมูลดังกล่าว  $U_c$  และ  $L_c$  ในสมการที่ (76) และ (78) ตามลำดับ เพื่อนำไปพิจารณาการแบ่งโซน Upper Zone และ Lower Zone การวิเคราะห์จะพิจารณาเฉพาะผลต่างของระดับช่วงบน เนื่องจากการวิเคราะห์หาผลต่างระดับช่วงล่าง จะใช้วิธีเดียวกัน สามารถแสดงขั้นตอนได้ดังแสดงในแผนภูมิที่ 57

#### แผนภูมิที่ 57

แสดงขั้นตอนการวิเคราะห์ตัวอักษรติดกัน



## แผนภูมิที่ 57 (ต่อ)



การจัดเรียงค่าในอะเรย์ topdif และ lowdif จากค่าน้อยไปหาค่ามาก สามารถใช้วิธีการจัดเรียงแบบทั่วไป เช่น Bubble sort, Intersection sort, Heap sort, Quick sort เป็นต้น ซึ่งให้ผลแบบเดียวกัน ต่างกันที่ความซับซ้อนและความเร็วในการประมวลผลเท่านั้นเอง ผลของตัวอย่างการจัดเรียงสามารถแสดงได้ดังแสดงในรูปที่ 58

รูปที่ 58

แสดงการจัดเรียงค่าผลต่างระหว่างระดับ



a) ตัวอย่างบิตแมพภาพของกรอบตัวอักษรในบรรทัดที่พิจารณา

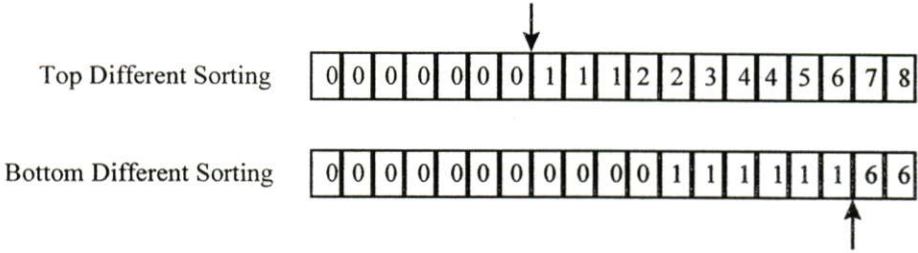
Top Different

1	6	0	0	0	4	5	0	3	4	1	0	0	1	7	2	2	8	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

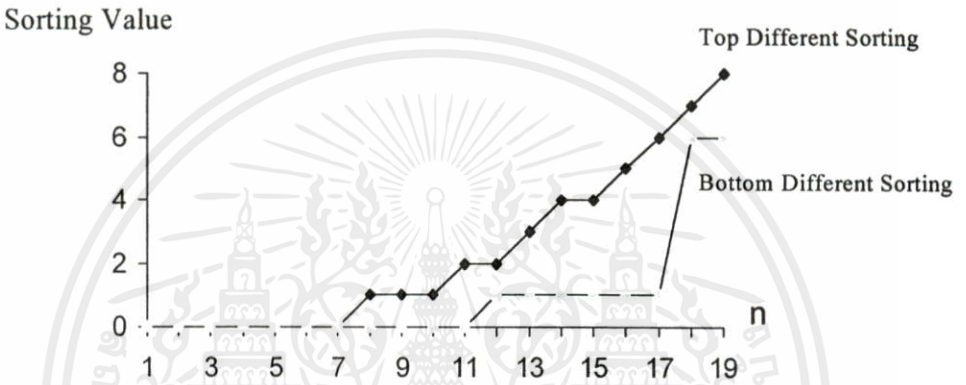
Bottom Different

0	0	1	0	0	1	0	0	1	0	0	1	6	6	1	0	0	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

b) แสดงผลต่างของระดับตัวอักษรที่อยู่ใกล้กัน



c) แสดงผลการจัดเรียงจากน้อยไปมากของอะเรย์ topdif และ lowdif



d) กราฟแสดงผลการจัดเรียงจากน้อยไปมากของอะเรย์ topdif และ lowdif

การหาค่าอะเรย์ topdif และ lowdif สามารถพิจารณาจากสมการที่ (79) และ สมการที่ (80)

$$\text{topdif}_C = \text{start}_C - \text{start}_{C-1} \dots\dots\dots (79)$$

$$\text{lowdif}_C = \text{end}_C - \text{end}_{C-1} \dots\dots\dots (80)$$

- กำหนดให้ C เป็นดัชนีของตัวอักษรภาพในบรรทัด
- C+1 เป็นดัชนีของตัวอักษรภาพตัวถัดไป

เอกสารนี้เป็นผลการจัดเรียงจากน้อยไปมากในอะเรย์ topdif จะถูกนำมาวิเคราะห์เพื่อหาจุดแบ่งระหว่างค่าความแตกต่างต่ำสุด mindif<sub>T</sub> และค่าความแตกต่างสูงสุด maxdif<sub>T</sub> จากวิธีของอนุพันธ์อันดับสอง (Second order differential) ดังในสมการที่ (81)

$$\text{snddif}_T = \text{topdif}_{C+1} - 2 * \text{topdif}_C + \text{topdif}_{C-1} \dots\dots\dots (81)$$

และ 
$$\text{snddif}_C = \text{lowdif}_{C+1} - 2 * \text{lowdif}_C + \text{lowdif}_{C-1} \dots\dots\dots (82)$$

กำหนดให้ C-1 เป็นดัชนีของตัวอักษรภาพก่อนหน้า

ตำแหน่งของดัชนีตัวอักษรในอะเรย์ topdif ที่มีค่าอนุพันธ์อันดับสองสูงสุดจะใช้เป็นจุดเริ่มต้นแบ่งกลุ่มของค่าความแตกต่างออกเป็น 2 กลุ่ม กลุ่มแรกจะเป็นกลุ่มที่มีค่าความแตกต่างน้อย กลุ่มที่สองจะเป็นกลุ่มที่มีค่าความแตกต่างมาก ดังนั้นการหาค่า mindif และ maxdif จะอาศัยค่าเฉลี่ยความแตกต่าง ณ จุดกึ่งกลางชั้น คือ

mindif <sub>T</sub>	ค่าเฉลี่ยของอะเรย์ topdif ณ จุดกึ่งกลางกลุ่มแรก
maxdif <sub>T</sub>	ค่าเฉลี่ยของอะเรย์ topdif ณ จุดกึ่งกลางกลุ่มที่สอง
mindif <sub>L</sub>	ค่าเฉลี่ยของอะเรย์ lowdif ณ จุดกึ่งกลางกลุ่มแรก
maxdif <sub>L</sub>	ค่าเฉลี่ยของอะเรย์ lowdif ณ จุดกึ่งกลางกลุ่มที่สอง

และการหาค่าเฉลี่ยของ Upper Zone ในแต่ละบรรทัด จากผลต่างของค่า mindif<sub>T</sub> และ maxdif<sub>T</sub> จากสมการที่ (83).

$$\text{Upper Zone} = \text{maxdif}_T - \text{mindif}_T \dots\dots\dots (83)$$

การหา Lower zone จะใช้หลักการเดียวกัน

#### 6.3.5.4 การค้นหาระดับของตัวอักษรและการกำหนดกรอบมาตรฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใด การค้นหาระดับของตัวอักษรจะอาศัยผลต่างระหว่างจุดเริ่มต้นช่วงบนค่า start<sub>C</sub> และผลต่างระหว่างจุดสิ้นสุดช่วงล่างค่า end<sub>C</sub> อักษรที่ติดกัน หากค่าดังกล่าวน้อยกว่าหรือมากกว่าค่าการตัดสินใจ (topdecision) จะสามารถกำหนดเส้นวางระดับของตัวอักษรได้

## รูปที่ 59

แสดงการค้นหาระดับของตัวอักษร



รูปที่ 59 แสดงการตัดสินใจจะพิจารณาจากเส้นกึ่งกลางระหว่าง Upper line และ Lower line เนื่องจากค่า  $\text{maxdif}_T$  จะสัมพันธ์กับ Upper line และค่า  $\text{mindif}_T$  จะสัมพันธ์กับ Upper half ดังนั้นค่าคงตัดสินใจ สามารถหาได้จากสมการที่ (84)

$$\text{topdecision} = \frac{\text{mindif}_T + \text{maxdif}_T}{2} \dots\dots\dots (84)$$

$$\text{lowdecision} = \frac{\text{mindif}_L + \text{maxdif}_L}{2} \dots\dots\dots (85)$$

หลังจากที่ได้ค่าการตัดสินใจมาแล้ว จะเริ่มหาระดับตัวอักษรช่วงบน ( $\text{toplevel}_C$ ) และ ช่วงล่าง ( $\text{lowlevel}_C$ ) จากการตรวจสอบผลต่างระหว่างจุดเริ่มต้น ช่วงบนของตัวอักษรที่ติดกัน ( $\text{start}_C$  และ  $\text{start}_{C+1}$ ) หากแปลเป็นความหมายทางตัวแปรคณิตศาสตร์จะได้ว่า

กรณีที่ 1 ถ้า  $(\text{topdif}_C - \text{topdif}_{C+1})$  มากกว่า  $\text{topdecision}$  แล้วจะให้

$$\text{toplevel}_C = \text{LevelUpperHalf}$$

$$\text{toplevel}_{C+1} = \text{LevelUpperLine}$$

กรณีที่ 2 ถ้า  $(\text{topdif}_C - \text{topdif}_{C+1})$  มากกว่า  $(-1.0 * \text{topdecision})$  แล้วจะให้

$$\text{toplevel}_C = \text{LevelUpperLine}$$

$$\text{toplevel}_{C+1} = \text{LevelUpperHalf}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับเอาไว้ใช้ในวงที่การศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



## รูปที่ 62

แสดงวิเคราะห์หาระดับของอักษรจากตำแหน่งแรก

ทางขวามือไปจนสุดตัวอักษรทางซ้ายมือ

UL

# Character Recognition

LH



ขณะนี้เราสามารถกำหนดระดับตัวอักษรทั้งบรรทัดได้แล้ว แสดงดังในตารางที่ 8

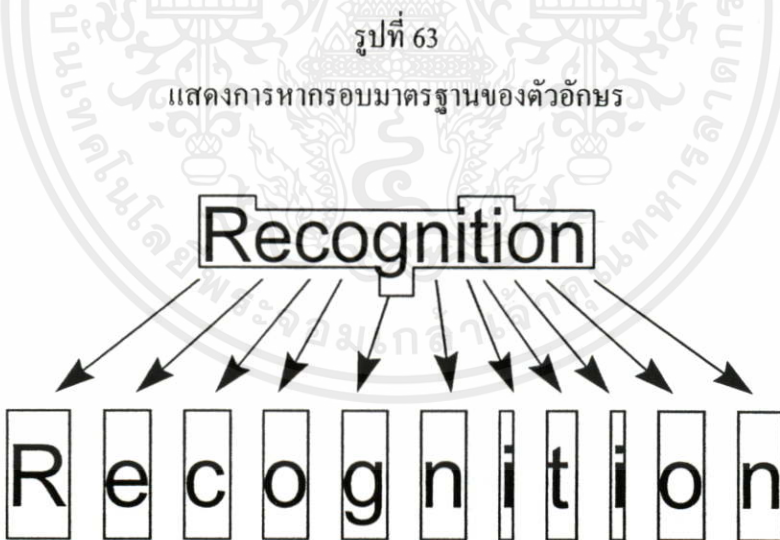
## ตารางที่ 8

แสดงผลการวิเคราะห์หาระดับตัวอักษร

ตำแหน่ง ตัวอักษร	รหัสตัวอักษร	ระดับตัวอักษรช่วงบน	ระดับตัวอักษรช่วงล่าง	จัดประเภท
1	C	LevelUpperLine	LevelLowerHalf	Ascender
2	h	LevelUpperLine	LevelLowerHalf	Ascender
3	a	LevelUpperHalf	LevelLowerHalf	Centered
4	r	LevelUpperHalf	LevelLowerHalf	Centered
5	a	LevelUpperHalf	LevelLowerHalf	Centered
6	c	LevelUpperHalf	LevelLowerHalf	Centered
7	t	LevelUpperline	LevelLowerHalf	Ascender
8	e	LevelUpperHalf	LevelLowerHalf	Centered
9	r	LevelUpperHalf	LevelLowerHalf	Centered
10	R	LevelUpperline	LevelLowerHalf	Ascender
11	e	LevelUpperHalf	LevelLowerHalf	Centered
12	c	LevelUpperHalf	LevelLowerHalf	Centered
13	o	LevelUpperHalf	LevelLowerHalf	Centered
14	g	LevelUpperHalf	LevelLowerLine	Descended

จากตารางที่ 8 จะเห็นว่าเราสามารถจัดประเภทให้กับตัวอักษรทุกตัวในบรรทัดได้อย่างถูกต้อง ซึ่งผลที่ได้รับนี้เราสามารถระบุคุณสมบัติเฉพาะของตัวอักษรที่จะถูกนำไปแยกประเภทตัวอักษรในภายหลังทำให้ผลการรู้จำดีขึ้น หรืออาจใช้หลักการนี้เพื่อแยกประเภทของกลุ่มตัวอักษรขั้นต้น (Preclassifier) ก่อนนำไปแยกประเภทด้วยวิธีของนิรอลเน็ตเวิร์ค

ก่อนที่จะนำตัวอักษรเข้าไปผ่านกระบวนการแยกคุณลักษณะเฉพาะสำหรับวิทยานิพนธ์ฉบับนี้นั้น จำเป็นจะต้องหากรอบมาตรฐานให้กับทุกตัวอักษรเสียก่อน วิธีหากรอบมาตรฐานทำได้ไม่ยาก เพียงแต่เพิ่มช่องว่างช่วงบนและช่วงล่างให้กับตัวอักษรประเภท Ascender, Centered และ Descended เท่านั้น (ยกเว้นตัวอักษรแบบ Full-size) ตัวอักษรแบบ Ascender ทุกตัวจำเป็นต้องเพิ่มช่องว่างช่วงล่าง (Lower Zone) ตัวอักษร Centered ทุกตัวจำเป็นต้องเพิ่มช่องว่างช่วงบนและช่วงล่าง (Upper zone และ Lower zone) ส่วนตัวอักษร Descended ทุกตัวจำเป็นต้องเน้นช่องว่างช่วงบน (Upper zone) ดังแสดงในรูปที่ 63



ผลของขั้นตอนนี้ทำให้ได้กรอบมาตรฐานของตัวอักษรทุกตัวในบรรทัด สำหรับใช้ในการประมวลผลในขั้นตอนต่อไป

สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 6.4 การแยกคุณลักษณะเฉพาะของตัวอักษร

การแยกคุณลักษณะเฉพาะเป็นขั้นตอนหนึ่งที่สำคัญในการรู้จำแพทเทิร์น คุณลักษณะที่ใช้ในการวิจัยครั้งนี้จะเกี่ยวข้องกับหลักการของ Curve approximation [12] สามารถแสดงคุณลักษณะเฉพาะของลายเส้นตัวอักษรตามความเป็นที่สายตามนุษย์สามารถมองเห็น เนื่องจากตัวอักษรทุกตัวเกิดจากการวาดลายเส้นอย่างต่อเนื่องในทิศทางต่างกัน แต่จะมีจุดเริ่มต้นและจุดสิ้นสุดที่แน่นอนเสมอ ตัวอักษรที่มีรหัสแอสกีเดียวกันจะมีความชันและความโค้งท้องถิ่นของลายเส้นเหมือนกันในทิศทางเดียวกัน มีจุดกำหนดตำแหน่งที่แน่นอนภายในกรอบ

จากแนวความคิดดังกล่าว สามารถกำหนดขั้นตอนการคำนวณหาจุดรอบแพทเทิร์นของตัวอักษร แล้ว Resampling จำนวนจุดรอบแพทเทิร์นให้ได้จำนวนจุดรอบแพทเทิร์นขนาดคงที่ จากนั้นคำนวณหาคุณลักษณะของ Rescaling, Directional angle, Curvature angle และปากกายกขึ้นกกลงตามลำดับ ดังแสดงในแผนภูมิที่ 64

แผนภูมิที่ 64

ขั้นตอนการทำงานในส่วนของการแยกคุณลักษณะเฉพาะ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการวิจัยเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเอกสารทุกครั้งที่มีการนำไปใช้

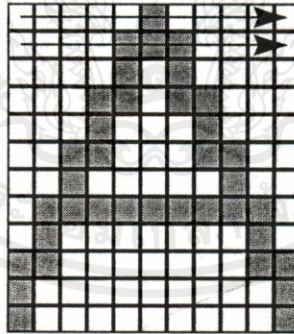
#### 6.4.1 วิธีหาจุดรอบแพทเทิร์นตัวอักษร

จากขั้นตอนการประมวลผลขั้นต้นที่ผ่านมา เราจะได้กรอบตัวอักษรในลักษณะบิตแมพภาพขนาด 2 มิติ เป็นไบนารีพิกเซล มีเพียงจุดภาพขาวและดำ ซึ่งลายเส้นบนตัวอักษรประกอบขึ้นจากจุดภาพสีดำเรียงติดต่อกัน ตามรูปแบบจำเพาะของแต่ละรหัสตัวอักษร การหาจุดรอบแพทเทิร์นจะกระทำโดยเริ่มกำหนดจุดๆหนึ่งบนขอบตัวอักษรด้านนอก และหาจุดถัดไปในลักษณะวนขวาจนกระทั่งกลับมายังจุดเดิม เก็บตำแหน่งของจุดไว้ในอะเรย์แบบ 1 มิติ

เนื่องจากแพทเทิร์นตัวอักษรที่นำเข้ามาเป็นแบบ off-line ไม่สามารถรู้จุดเริ่มต้นที่แท้จริงของลายเส้นตัวอักษรได้ จึงมีการกำหนดให้ตรวจหาจุดภาพสีดำ จุดแรกภายในกรอบตัวอักษรจากแถวบนลงล่าง ตำแหน่งของจุดภาพสีดำจุดแรกจะถือเป็นจุดเริ่มต้นและจะถูกเก็บในอะเรย์ในตำแหน่งแรก

รูปที่ 65

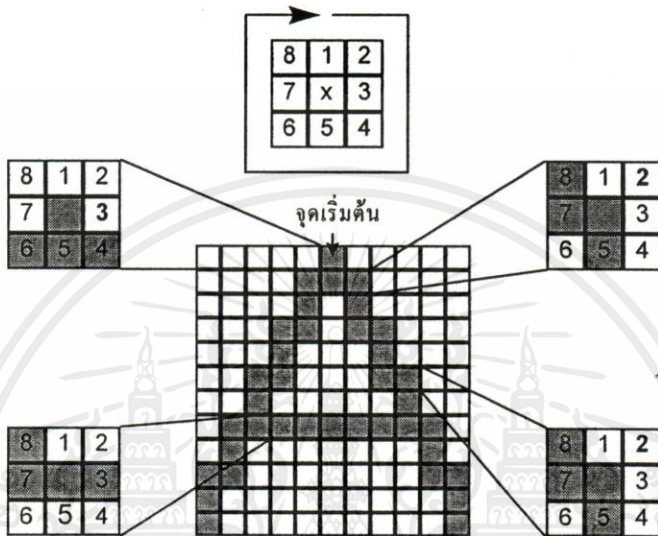
แสดงทิศทางการสแกนหาจุดเริ่มต้นรอบแพทเทิร์นตัวอักษร



เมื่อได้จุดเริ่มต้นรอบแพทเทิร์นตัวอักษรแล้ว จะนำมาใช้เป็นจุดกึ่งกลางของกรอบจุดภาพขนาด 3x3 ดังแสดงในรูปที่ 66 จุดถัดไปรอบจุดกึ่งกลางมีโอกาasเป็นไปได้ทั้ง 8 จุด (8 ทิศทาง) จึงควรมีการกำหนดทิศทางเริ่มต้นในการสแกนหาจุดถัดไป ในเริ่มแรกจะกำหนดให้ตำแหน่งที่ 3 เป็นจุดตั้งต้นและค้นหาจุดต่อไปในลักษณะวนขวาในทิศตามเข็มนาฬิกาจะได้ว่าจุดถัดไปคือตำแหน่งที่ 4 จากนั้นตำแหน่งที่ 4 จะถูกกำหนดให้เป็นจุดเริ่มต้นของกรอบจุดภาพขนาด 3x3 อีกครั้ง การกำหนดทิศทางถัดไปจำเป็นต้องบวกตำแหน่งเพิ่มขึ้นอีก 6 ตำแหน่ง จากเดิมตำแหน่งเริ่มต้นเป็น 4 บวกอีก 6 ตำแหน่งจะเป็น 10 กำหนดตำแหน่งเริ่มต้นของจุดถัดไป ทำเช่นนี้จนครบจุดรอบแพทเทิร์นตัวอักษร

รูปที่ 66

แสดงการสแกนหาจุดรอบแพทเทิร์นตัวอักษรด้วยการกำหนด  
กรอบจุดภาพขนาด 3x3 จากทิศทางทั้ง 8



รูปที่ 66 แสดงตัวอย่างการสแกนหาจุดรอบแพทเทิร์นตัวอักษรด้วยวิธีการกำหนดกรอบจุดภาพขนาด 3x3 เพื่อหาค่าแห่งถัดไปของจุดรอบแพทเทิร์นตัวอักษร จนกระทั่งได้จุดรอบแพทเทิร์นตัวอักษรทั้งหมดรวม 36 จุด

ตารางที่ 9

ตำแหน่ง x และ y ที่หาได้จากการสแกนหาจุดรอบแพทเทิร์นตัวอักษร

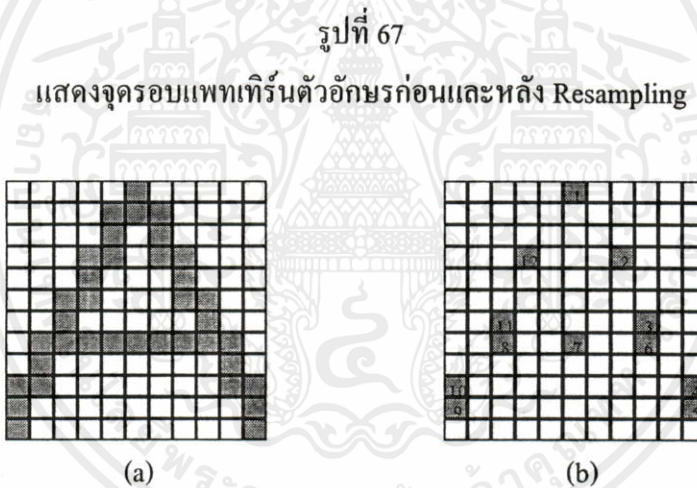
x	5	6	6	7	7	8	8	9	9	10	10	10	9	9	8	7	6	5	4	3	2	1	1	0	0	0	0	1	1	2	2	3	3	4	4	
y	0	1	2	3	4	5	6	7	8	9	10	11	10	9	8	7	7	7	7	7	7	7	8	9	10	11	10	9	8	7	6	5	4	3	2	1

เอกสารนี้เป็นตารางที่ 9 แสดงตำแหน่ง x และ y ที่หาได้จากการสแกนหาจุดรอบแพทเทิร์นตัวอักษร ค่าตำแหน่ง x และ y ทั้งหมดจะถูกเก็บไว้เพื่อนำไปใช้ในขั้นตอนถัดไป เอกสารทุกครั้งที่มีการนำไปใช้

## 6.4.2 การ Resampling

วิธีการ Resampling มีความจำเป็นอย่างมากในการกำหนดจำนวนจุดรอบแพทเทิร์นตัวอักษร เนื่องจากหลักการแยกประเภทที่เสนอจะอาศัยนิเวศน์เวกซ์ของเบ็คฟร็อบพาทะกั้นเป็นส่วนช่วยแยกประเภทตัวอักษร ซึ่งจำเป็นต้องกำหนดจำนวนโหนดอินพุทขนาดคงที่ จึงต้องใช้วิธีการ Resampling เพื่อให้ได้จำนวนข้อมูลตามที่ต้องการ

รูปที่ 66 (a) แสดงแพทเทิร์นตัวอักษร A ซึ่งมีจุดรอบแพทเทิร์นจำนวน 36 จุด หากต้องการจุดรอบแพทเทิร์นตัวอักษรขนาดคงที่จำนวน 12 จุด จำเป็นต้องลดจำนวนจุดรอบแพทเทิร์นด้วยหลักการของ Lagrangian interpolation ดังแสดงในรูปที่ 67



รูปที่ 67 (a) แสดงจุดรอบแพทเทิร์นตัวอักษรก่อน Resampling และรูปที่ 67 (b) แสดงจุดรอบแพทเทิร์นตัวอักษรหลัง Resampling ด้วยหลักการของ Lagrangian interpolation

### ตารางที่ 10

#### แสดงตำแหน่งของจุดรอบแพทเทิร์นหลังการ Interpolation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ตำแหน่ง	0	1	2	3	4	5	6	7	8	9	10	11
x	5.0	7.0	8.0	10.0	10.0	8.0	5.0	2.0	0.0	0.0	2.0	3.0
y	0.0	3.0	6.0	9.0	10.0	7.0	7.0	7.0	10.0	9.0	6.0	3.0

ตารางที่ 10 แสดงผลการคำนวณหาตำแหน่งของจุดรอบแพทเทิร์นจำนวน 12 จุดด้วยวิธีการ Interpolation ค่าตำแหน่ง  $x$  และ  $y$  จากตารางที่ 9 การ Interpolation จะอาศัยตำแหน่งของจุดด้านซ้ายมือจำนวนสองจุดและจุดด้านขวามืออีกจำนวนสองจุดในการคำนวณ หลังจาก Interpolation แล้วจุดทั้งหมดที่ได้จะมีความต่อเนื่องกันในลักษณะเรียบขึ้น (Smooth) ตำแหน่งของจุดรอบแพทเทิร์นที่ได้จะถูกเก็บไว้เพื่อใช้เป็นฐานข้อมูลสำหรับหาคูณลักษณะแพทเทิร์นต่อไป

#### 6.4.3 คูณลักษณะของ Rescaling

Rescaling เป็นการหาคูณลักษณะตำแหน่งของจุดเทียบกับศูนย์กลางของกรอบภาพ เนื่องจากลายเส้นของตัวอักษร เกิดจากจุดหลายจุดเรียงต่อเนื่องกันในตำแหน่งที่แน่นอน จุดแต่ละจุดจึงมีความสัมพันธ์กับคูณลักษณะเฉพาะตัวของแพทเทิร์นตามรหัสแอสกี ดังนั้นระยะห่างของจุดเทียบกับจุดอ้างอิงจึงสามารถใช้เป็นส่วนหนึ่งในการแยกประเภทตัวอักษรได้

วิธีการหาคูณลักษณะในหัวข้อนี้จะเริ่มจากการกำหนดจุดศูนย์กลางของกรอบภาพตัวอักษร ด้วยสมการที่ (86) และ สมการที่ (87)

$$X(n) = \frac{2 \cdot X_0 - (D-1)}{H-1} \dots\dots\dots (86)$$

$$Y(n) = \frac{2 \cdot Y_0 - (H-1)}{H-1} \dots\dots\dots (87)$$

ที่ซึ่ง	$X_0$	ตำแหน่งเดิมในแนวแกน $x$ ของขอบภาพตัวอักษร
	$X(n)$	ตำแหน่งใหม่ (ตำแหน่งที่ $n$ ) ในแนวแกน $x$ ของขอบภาพตัวอักษร
	$Y_0$	ตำแหน่งเดิมในแนวแกน $y$ ของขอบภาพตัวอักษร
	$Y(n)$	ตำแหน่งใหม่ (ตำแหน่งที่ $n$ ) ในแนวแกน $y$ ของขอบภาพตัวอักษร
	$H$	ค่าความสูงของกรอบภาพตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับงานวิจัยเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านธุรกิจ  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรณีไปใช้

จากการ Rescaling เราจะ ได้ค่า  $X(n)$  และ  $Y(n)$  มีค่าอยู่ระหว่าง -1 ถึง 1 เสมอ เนื่องจากส่วนสูงของตัวอักษรมากกว่าความกว้าง แต่หากว่าความกว้างของตัวอักษรมากกว่าความสูงมากๆ อาจจำเป็นต้องใช้  $\frac{2 \cdot X_0 - (D-1)}{D-1}$  และ  $\frac{2 \cdot Y_0 - (H-1)}{D-1}$  แทน เช่น ตัวอักษรตัวพิมพ์ “W” เป็นต้น โดยค่า  $D$  คือความกว้างของกรอบภาพแพทเทิร์น

ตัวอย่างการคำนวณหาคุณลักษณะของ Rescaling จากตำแหน่งของจุดรอบแพทเทิร์นหลังการ Interpolation ในตารางที่ 10 เมื่อ  $D$  เท่ากับ 11 และ  $H$  เท่ากับ 12

$$X(0) = \frac{2 \cdot 5 - (11-1)}{12-1} = 0.000 \quad Y(0) = \frac{2 \cdot 0 - (12-1)}{12-1} = -1.00$$

$$X(1) = \frac{2 \cdot 7 - (11-1)}{12-1} = 0.364 \quad Y(1) = \frac{2 \cdot 3 - (12-1)}{12-1} = -0.455$$

$$X(2) = \frac{2 \cdot 8 - (11-1)}{12-1} = 0.545 \quad Y(2) = \frac{2 \cdot 6 - (12-1)}{12-1} = 0.091$$

ตารางที่ 11

แสดงคุณลักษณะของ Rescaling หลังการคำนวณ

n	0	1	2	3	4	5	6	7	8	9	10	11
X(n)	0.00	0.36	0.54	0.90	0.90	0.54	0.00	-0.54	-0.90	-0.90	-0.54	-0.36
Y(n)	-1.00	-0.45	0.09	0.63	0.82	0.27	0.27	0.27	0.82	0.63	0.09	-0.45

ตารางที่ 11 แสดงผลของการคำนวณหาคุณลักษณะ Rescaling ของค่า  $X(n)$  และ  $Y(n)$  ตามสมการที่ 86 และสมการที่ 87

**6.4.4 คุณลักษณะของ Directional angle** สำหรับการศึกษานี้ ไม่นานมานี้ นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะเป็นใจทุกสิ่ง อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากที่กล่าวไปแล้วว่า ลายเส้นของตัวอักษรเกิดจากจุดหลายจุดเรียงต่อเนื่องกันตามลักษณะเฉพาะตัว ดังนั้นจุดแต่ละจุดจะสัมพันธ์กับทิศทางการเคลื่อนที่ในตำแหน่งที่แน่นอนเช่นกัน จึงเป็นแนวความคิดในการหามุมระหว่างจุดที่วางใกล้กัน นั้นหมายถึง ทิศทางการเขียนตัวอักษรนั่นเอง

วิธีการหาทิศทางด้วยมุมระหว่างจุดในรูปของ Sines และ Cosines ของเส้นสัมพันธ์ด้วยส่วนโค้งที่ตำแหน่ง  $n$  ดังแสดงในรูปที่ 68 สมการที่ (88) และ สมการที่ (89) จะแสดงการหาอนุพันธ์อันดับหนึ่งของความยาวส่วนโค้งที่กำหนด  $\frac{dx}{ds}$  และ  $\frac{dy}{ds}$  ที่ซึ่ง  $ds = \sqrt{dx^2 + dy^2}$  ด้วยสมการแบบไม่ต่อเนื่อง

$$\cos\theta(n) = \frac{\Delta x(n)}{\Delta s(n)} \dots\dots\dots (88)$$

$$\sin\theta(n) = \frac{\Delta y(n)}{\Delta s(n)} \dots\dots\dots (89)$$

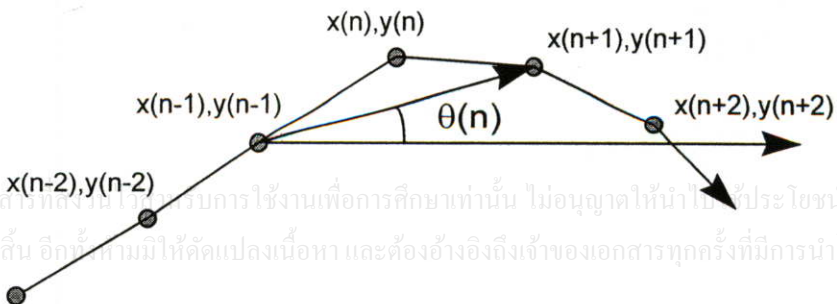
เมื่อ  $\Delta x(n) = x(n+1) - x(n-1) \dots\dots\dots (90)$

$$\Delta y(n) = y(n+1) - y(n-1) \dots\dots\dots (91)$$

$$\Delta s(n) = \sqrt{\Delta x(n)^2 + \Delta y(n)^2} \dots\dots\dots (92)$$

รูปที่ 68

แสดงทิศทางเคลื่อนที่ของจุดที่สัมพันธ์กัน



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งยังมีให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มุม  $\theta$  เกิดจากเส้นตรงระหว่างจุดที่ตำแหน่ง  $n-1$  ถึง  $n+1$  กับเส้นตรงในแนวอนชนานกับแกน  $x$  ผลก็คือเราจะได้คุณลักษณะเฉพาะของทิศทางเคลื่อนที่ของจุด ณ ตำแหน่ง  $n$  นั่นเอง

ตัวอย่างการคำนวณหาคุณลักษณะของ Directional angle จากตำแหน่งของจตุรรอบแพทเทิร์นหลังการ Interpolation ในตารางที่ 10 เป็นดังนี้ เมื่อ  $D$  เท่ากับ 11 และ  $H$  เท่ากับ 12

$$\cos\theta(0) = \frac{7-3}{\sqrt{(7-3)^2 + (3-3)^2}} = 1.000 \quad \sin\theta(0) = \frac{3-3}{\sqrt{(7-3)^2 + (3-3)^2}} = 0.00$$

$$\cos\theta(1) = \frac{8-5}{\sqrt{(8-5)^2 + (6-0)^2}} = 0.447 \quad \sin\theta(1) = \frac{6-0}{\sqrt{(8-5)^2 + (6-0)^2}} = 0.894$$

$$\cos\theta(2) = \frac{10-7}{\sqrt{(10-7)^2 + (9-3)^2}} = 0.447 \quad \sin\theta(2) = \frac{9-3}{\sqrt{(10-7)^2 + (9-3)^2}} = 0.894$$

ตารางที่ 12

แสดงคุณลักษณะของ Directional angle หลังการคำนวณ

n	0	1	2	3	4	5	6	7	8	9	10	11
$\cos\theta(n)$	1.00	0.45	0.45	0.45	-0.70	-0.86	-1.00	-0.86	-0.70	0.45	0.45	0.45
$\sin\theta(n)$	0.00	0.89	0.89	0.89	-0.70	-0.51	0.00	0.51	0.70	-0.89	-0.89	-0.89

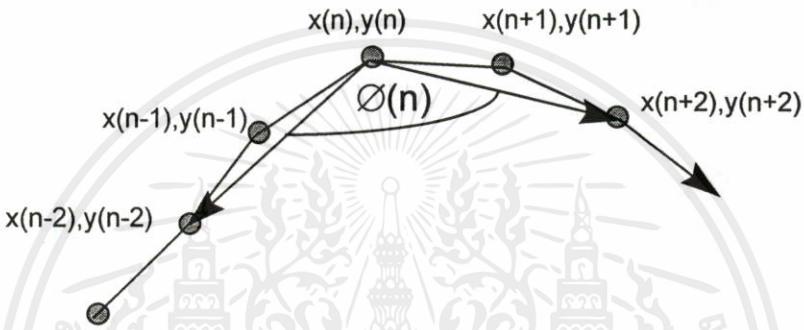
ตารางที่ 12 แสดงผลของการคำนวณหา Directional angle ของค่า  $\cos\theta(n)$  และ  $\sin\theta(n)$  ตามสมการที่ 88 และสมการที่ 89

#### 6.4.5 คุณลักษณะของ Curvature angle

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การหาคุณลักษณะของความโค้งแบบท้องถิ่น มีความเกี่ยวเนื่องมาจากการหาทิศทาง การเคลื่อนที่ของจุดที่อยู่ใกล้กัน เนื่องจากต้องการหาค่า Sines และ Cosines ของมุม  $\varnothing$  ณ ตำแหน่ง  $n$  ดังแสดงในรูปที่ 69

รูปที่ 69  
แสดงความโค้งที่เกิดจากจุดที่สัมพันธ์กัน



ความโค้งของมุม  $\varnothing$  ณ ตำแหน่ง  $n$  หาได้จากความสัมพันธ์ของสมการที่ (93)

$$\varnothing(n) = \theta(n+1) - \theta(n-1) \dots\dots\dots (93)$$

ดังนั้นการคำนวณหาคุณลักษณะของความโค้ง ณ ตำแหน่ง  $n$  จะแสดงได้ดังนี้

$$\cos\varnothing(n) = \cos\theta(n-1)\cdot\cos\theta(n+1) + \sin\theta(n-1)\cdot\sin\theta(n+1) \dots\dots\dots (94)$$

$$\sin\varnothing(n) = \cos\theta(n-1)\cdot\sin\theta(n+1) - \sin\theta(n-1)\cdot\cos\theta(n+1) \dots\dots\dots (95)$$

มุม  $\varnothing$  เกิดจากเส้นตรงสองเส้น ณ จุด  $n$  กับจุด  $n-2$  และจุด  $n$  กับจุด  $n+2$  เกิดเป็นมุมของความโค้งรอบแพทเทิร์นในรูป Sines และ Cosines ผลก็คือเราจะได้คุณลักษณะเฉพาะของส่วนโค้ง ณ ตำแหน่ง  $n$  นั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้วยการค้า ตัวอย่างการคำนวณหาคุณลักษณะของ Curvature angle จากค่า  $\cos\theta(n)$  และ  $\sin\theta(n)$  ในคู่มือคู่มือใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้ คุณลักษณะของ Directional angle ตามตารางที่ 12 เป็นดังนี้

$$\cos\varnothing(0) = (0.447)(0.447) + (-0.894)(0.894) = -0.600$$

$$\sin\varnothing(0) = (0.447)(0.894) - (-0.894)(0.447) = 0.800$$

$$\cos\varnothing(1) = (1.000)(0.447) + (0.000)(0.894) = 0.447$$

$$\sin\varnothing(1) = (1.000)(0.894) - (0.000)(0.447) = 0.894$$

$$\cos\varnothing(2) = (0.447)(0.447) + (0.894)(0.894) = 1.000$$

$$\sin\varnothing(2) = (0.447)(0.894) - (0.894)(0.447) = 0.000$$

### ตารางที่ 13

แสดงคุณลักษณะของ Curvature angle หลังการคำนวณ

n	0	1	2	3	4	5	6	7	8	9	10	11
$\cos\varnothing(n)$	-0.60	0.45	1.00	-0.95	-0.84	0.71	0.47	0.71	-0.84	-0.95	1.00	0.45
$\sin\varnothing(n)$	0.80	0.89	0.00	0.32	0.54	-0.71	-0.88	-0.71	0.54	0.32	0.00	0.89

ตารางที่ 13 แสดงผลของการคำนวณหาคุณลักษณะของ Curvature angle สำหรับค่า  $\cos\varnothing(n)$  และ  $\sin\varnothing(n)$  ตามสมการที่ 94 และสมการที่ 95

ความสัมพันธ์ของทั้งสามหัวข้อที่ผ่านมาจะแสดงถึงคุณลักษณะของลายเส้นตัวอักษรเฉพาะตัวตามรหัสแอสกีในภาษาอังกฤษ แต่เนื่องจากภาษาอังกฤษมีตัวอักษรบางตัวที่ไม่สามารถนำหลักการดังกล่าวมาใช้ในลักษณะธรรมดาได้ เช่นตัวอักษร  $i$  และ  $j$  มีลายเส้นของตัวอักษรแยกกัน 2 ตำแหน่ง (Object) จึงจำเป็นต้องอาศัยเทคนิคบางประการเพิ่มเติม ดังจะกล่าวในหัวข้อถัดไป

#### 6.4.6 ปากกาขยับขึ้นและกดลง (Pen-up และ Pen-down)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่าการค้า จากที่กล่าวมาตัวอักษรอังกฤษ  $i$  และ  $j$  ก่อให้เกิดปัญหาในการนำเสนอคุณลักษณะเฉพาะของแต่ละตัวอักษรสำหรับคุณลักษณะเฉพาะดังที่ได้กล่าวมา การแก้ปัญหาจะพิจารณาคุณลักษณะพิเศษของการเขียนลายเส้นตัวอักษร โดยปกติการเขียนตัวอักษรจำเป็นต้องวางปากกาคดลงบน

กระดาศเพื่อวาดลายเส้น ขณะที่ยกขึ้นยังไม่ได้เขียนจะไม่เกิดลายเส้นใดบนกระดาศ จากหลักการดังกล่าวได้นำมาประยุกต์ใช้กับปัญหาที่เกิดขึ้น

แนวความคิดดังกล่าว สามารถใช้กำหนดคุณลักษณะเฉพาะของปากกายกขึ้นและกดลงด้วยตัวแปร  $Penup(n)$  โดยตำแหน่งยกขึ้นจะแสดงด้วยค่า 1 และตำแหน่งกดลงจะแสดงด้วยค่า 0 ดังแสดงในรูปที่ 70

### รูปที่ 70

แสดงจุดที่สัมพันธ์กับการยกขึ้นและกดลง



จุดไข่วางจะแสดงเส้นทางการเชื่อมต่อจุดระหว่างการลากเส้นปากกายกขึ้นและกดลง โดยกำหนดเป็นจุดรอบตัวอักษรเก็บในอะเรย์  $Penup(n)$  ดังนั้นจึงมีรูปลายเส้นรอบแพทเทิร์นตัวอักษรในลักษณะยกขึ้นและกดลง เป็นคุณลักษณะเฉพาะพิเศษอีกประการหนึ่งด้วย

### ตารางที่ 14

แสดงคุณลักษณะของปากกายกขึ้นและกดลงหลังการคำนวณ

n	0	1	2	3	4	5	6	7	8	9	10	11
Penup (n)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

เอกสารนี้เป็นเอกสารต้นฉบับที่จัดทำขึ้นเพื่อใช้ในการศึกษาวิจัยและพัฒนาเทคโนโลยีปัญญาประดิษฐ์  
 ในงานวิจัยฉบับนี้ได้นำข้อมูลที่ได้มาใช้ในการศึกษาวิจัยและพัฒนาเทคโนโลยีปัญญาประดิษฐ์  
 ทำให้ Penup(n) มีค่าเป็น 0

### 6.4.7 การนำเสนอสื่อข้อมูล (Intermediate representaion)

การนำเสนอคุณลักษณะข้อมูลในลักษณะเวกเตอร์ (Feature vector) หรือลำดับของข้อมูลที่ตำแหน่ง  $n$  ถูกรวบรวมมาจากสมการที่ (86) ถึงสมการที่ (95) มานำเสนอเป็นลำดับด้วยส่วนประกอบของสมการ 7 สมการดังนี้

$$f_0(n) = \text{PenUp}(n) \dots\dots\dots (96)$$

$$f_1(n) = \frac{X(n)+1}{2} \dots\dots\dots (97)$$

$$f_2(n) = \frac{Y(n)+1}{2} \dots\dots\dots (98)$$

$$f_3(n) = \frac{\cos \theta(n)+1}{2} \dots\dots\dots (99)$$

$$f_4(n) = \frac{\sin \theta(n)+1}{2} \dots\dots\dots (100)$$

$$f_5(n) = \frac{\cos \phi(n)+1}{2} \dots\dots\dots (101)$$

$$f_6(n) = \frac{\sin \phi(n)+1}{2} \dots\dots\dots (102)$$

$f_3(n)$  และ  $f_4(n)$  จะแสดงความแปรเปลี่ยนทิศทางของจุด ส่วน  $f_5(n)$  และ  $f_6(n)$  จะแสดงความแปรเปลี่ยนความโค้งขึ้นๆลงๆ ของจุด โดยทั่วไปคุณลักษณะท้องถิ่นของเส้นตรง ส่วนโค้ง ขอบ ฯลฯ เป็นที่รู้จักกันดีในการใช้ประโยชน์สำหรับการรู้จำตัวอักษร

จะสังเกตเห็นว่า สมการที่ (97) ถึงสมการที่ (102) จำเป็นต้องบวกหนึ่งแล้วหารด้วยสอง เพื่อให้ผลลัพธ์อยู่ในขอบเขตของตัวเลขระหว่าง 0 ถึง 1 (Normalize) เป็นสัดส่วนของค่าต่ำสุดและสูงสุด

หลังจากวิธีการของการแยกคุณลักษณะเฉพาะของตัวอักษรผ่านไป แบบแผนของการแยกประเภทตัวอักษรต้องถูกใช้สำหรับขั้นตอนการรู้จำ แม้ว่าจะมีหลายๆ เทคนิคถูกพัฒนาขึ้น มีการปรับ

ปรุ้งเรื่องเวลาในการคำนวณและระดับของสัญญาณรบกวน การแยกประเภทวัตถุหรือตัวอักษรในปัจจุบันพยายามที่จะใช้โครงข่ายหน่วยประสาทเทียม ซึ่งต้องการข้อมูลแสดงในตำแหน่งและขนาดคงที่ จึงจะสามารถส่งค่าผ่านขั้นตอนการแยกประเภทตัวอักษรโดยวิธีของนิวโรลเน็ตเวิร์คได้

## 6.5 การแยกประเภทของตัวอักษร

การแยกประเภทตัวอักษรจะอาศัยความสามารถของนิวโรลเน็ตเวิร์คแบบที่รอบพาทกั้นเป็นหลัก เนื่องจากช่วยลดข้อจำกัดบางประการของเทคโนโลยีเก่าๆในบางอย่างได้ และยังช่วยเพิ่มคุณสมบัติเด่นบางประการเข้าไปด้วย ก่อนจะนำไปใช้งานจำเป็นต้องกำหนดโครงสร้างเน็ตเวิร์คและค่าเริ่มต้นบางประการที่จำเป็นเสียก่อน

เริ่มจากการออกแบบโครงสร้างเน็ตเวิร์ค ควรกำหนดจำนวนเลเยอร์ นั่นคือมีกี่ฮิดเดนเลเยอร์นั่นเอง กำหนดจำนวนโหนดของอินพุทเลเยอร์ ฮิดเดนเลเยอร์ และ เอาท์พุทเลเยอร์ ให้สัมพันธ์กับข้อมูลที่มีอยู่ และต่อไปต้องวางแผนว่าจะเลือกวิธีการคอนเน็คชันนิสต์ (Connectionist) อย่างไร

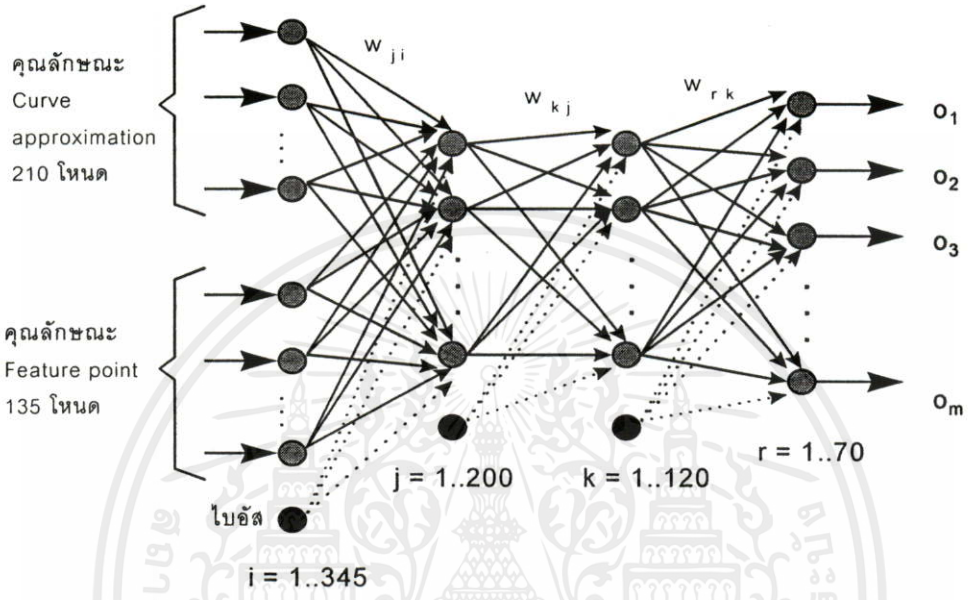
กำหนดให้เอาท์พุทโหนดมีจำนวนเท่ากับจำนวนคลาสของแพทเทิร์นตัวอักษรที่เข้ามาอบรม A..Z, a..z, 0..9 และตัวอักษรพิเศษจำนวน 8 ตัว รวมทั้งสิ้น 70 ตัว ดังนั้นจึงกำหนดจำนวนเอาท์พุทโหนดเป็น 70 โหนด จำนวนอินพุทโหนดจะขึ้นอยู่กับจำนวนคุณลักษณะของตัวอักษรที่ใช้ในการรู้จำโดยให้ผู้ใช้เป็นผู้กำหนด ซึ่งที่ผ่านมามีจำนวนจตุรรอบแพทเทิร์นตัวอักษรอาจกำหนดให้มีค่าอยู่ในช่วง 20-100 หากกำหนดน้อยเกินไปอาจมีผลให้การรู้จำไม่ดีเพียงพอ แต่หากกำหนดมากเกินไปจำนวนอินพุทโหนดจะมีมากเกินไปจนจำเป็น ในที่นี้จะกำหนดเพียง 30 เนื่องจากคุณลักษณะ  $f_0(n), f_1(n), \dots, f_6(n)$  ทั้งหมด 7 คุณลักษณะ ดังนั้นจำนวนโหนดอินพุทที่เป็นไปได้คือ  $30 \times 7$  เท่ากับ 210 โหนด

ถัดมาการกำหนดจำนวนฮิดเดนโหนดควรรู้ก่อนว่า ควรจะมีจำนวนฮิดเดนเลเยอร์เท่าใด เพื่อให้เหมาะสมจะกำหนดให้มี 2 ฮิดเดนเลเยอร์ จำนวนโหนดเท่ากับ 200 และ 120 ตามลำดับ การกำหนดจำนวนโหนดให้กับฮิดเดนเลเยอร์ จะคำนึงถึงจำนวนโหนดระหว่างอินพุทเลเยอร์และเอาท์พุทเลเยอร์เป็นสำคัญ แต่ทั้งนี้ขึ้นอยู่กับทรัพยากรของแต่ละบุคคลไม่สามารถหาจำนวนที่ดีที่สุดได้

การกำหนดวิธีการคอนเน็คชันนิสต์สามารถทำได้หลายแบบ โดยนิวโรลโหนดหนึ่งในเลเยอร์ชั้นบน ไม่จำเป็นจะต้องได้รับค่าอินพุททั้งหมดมาจากนิวโรลโหนดในเลเยอร์ชั้นล่างเสมอไป แต่สำหรับการวิจัยในครั้งนี้จะอาศัยโครงสร้างของเน็ตเวิร์คแบบทั่วไป ดังแสดงในรูปที่ 71

รูปที่ 71

แสดงโครงสร้างของนิเวรอลเน็ตเวิร์คที่ใช้ในการรู้จำ  
สำหรับวิทยานิพนธ์ฉบับนี้



โครงสร้างที่กำหนดจะมีจำนวนคอนเน็ตชันนิสต์เท่ากับ 101,790 สามารถคำนวณได้จากสมการที่ (103) อินพุทที่เพิ่มขึ้นมาอีก 135 โหนด เป็นคุณลักษณะจุดภาพขนาด 9 x 15 จุด

$$C_n = I_n \cdot H_{n_1} + H_{n_1} \cdot H_{n_2} + H_{n_2} \cdot O_n + BiasI_n + BiasH_{n_1} + BiasH_{n_2} \dots (103)$$

$$= 345 \times 200 + 200 \times 120 + 120 \times 70 + 200 + 120 + 70 = 101,790$$

ที่ซึ่ง	$C_n$	จำนวนการเชื่อมต่อด้วยค่าตัวเลขน้ำหนักระหว่างโหนด
	$I_n$	จำนวนอินพุทโหนดของนิเวรอลเน็ตเวิร์ค
	$H_{n_1}$	จำนวนฮินเดนโหนดของนิเวรอลเน็ตเวิร์คในเลเยอร์แรก
	$H_{n_2}$	จำนวนฮินเดนโหนดของนิเวรอลเน็ตเวิร์คในเลเยอร์สอง
	$O_n$	จำนวนเอาต์พุทโหนดของนิเวรอลเน็ตเวิร์ค
	$BiasI_n$	จำนวนการเชื่อมต่อระหว่างไบอัสและฮินเดนเลเยอร์แรกที่มีการนำไปใช้
	$BiasH_{n_1}$	จำนวนการเชื่อมต่อระหว่างไบอัสและฮินเดนเลเยอร์สอง
	$BiasH_{n_2}$	จำนวนการเชื่อมต่อระหว่างไบอัสและเอาต์พุทเลเยอร์

ก่อนการอบรม ค่าตัวเลขน้ำหนักถูกกำหนดด้วยค่าเริ่มต้นแบบสุ่มการกระจายแบบไม่สม่ำเสมออยู่ระหว่างค่า 0.0 ถึง 1.0 / Np เมื่อ Np เป็นจำนวนคลาสของแพทเทิร์น และการอบรมจะหยุดเมื่อค่าผิดพลาดโดยเฉลี่ยของระบบน้อยกว่าค่าที่กำหนด

ลักษณะโครงสร้างท้องถิ่นตามส่วนโค้งของตัวอักษร และคุณลักษณะจุดภาพขนาด 9 x 15 จะถูกนำมาใช้เป็นอินพุทของนิเวรอลเน็ตเวิร์ค นิเวรอลจะคำนวณผลรวมของตัวเลขน้ำหนักและค่าอินพุทแพทเทิร์น เพื่อส่งค่าผ่าน ไปยังฟังก์ชันซิกมอยด์ต่อเนื่อง  $f(\text{net})$  จนได้ค่าเอาต์พุท

### 6.5.1 การแปลคำสั่งค่าเอาต์พุท (Interpretation of the output)

การประเมินผลของการรู้จำตัวอักษรขึ้นกับเหตุผลหลายอย่าง ฐานความรู้ที่เป็นมาตรฐานชนิดของตัวอักษร จำนวนการแบ่งประเภท จำนวนตัวอย่างในการแบ่งประเภท มีความจำเป็นสำหรับการเปรียบเทียบกับวิธีการหลายๆแบบได้

หลัก การแปลความหมายที่ง่ายที่สุดสำหรับการตัดสินใจของเน็ตเวิร์คขึ้นอยู่กับประสิทธิภาพของชุดทดสอบที่เรียกว่า Raw performance (Ro) [12] ในข้อความตัวอักษร การคำนวณจะเกี่ยวข้องกับค่าเทรคโสรลสองค่า โดยกำหนดให้คัดออก (Rejection) อัตราโนมิสำหรับการรู้จำตัวอักษรที่คลุมเครือ หรือแพทเทิร์นที่ไม่เคยเรียนรู้มาก่อน ตามสมการที่ (104)

แพทเทิร์น ถูกพิจารณาว่าคลุมเครือถ้า

$$\max_1 - \max_2 < \theta_{\text{dif}} \dots\dots\dots (104)$$

หรือ  $\max_1 < \theta_{\text{min}} \dots\dots\dots (105)$

ที่ซึ่ง	$\max_1$	เป็นค่าเอาต์พุทที่มีค่าสูงที่สุด (ลำดับที่หนึ่ง)
	$\max_2$	เป็นค่าเอาต์พุทที่มีค่าสูงเป็นอันดับสอง (ลำดับที่สอง)
	$\theta_{\text{dif}}$	เป็นค่าเทรคโสรลที่แตกต่างกันระหว่างลำดับที่หนึ่งและที่สอง
	$\theta_{\text{min}}$	เป็นค่าเทรคโสรลค่าสุดท้ายยอมรับได้

เช่น กำหนดให้ค่าเทรคโสรลที่แตกต่างกัน  $\theta_{\text{dif}} = 0.3$  และค่าเทรคโสรลที่ยอมรับได้  $\theta_{\text{min}} = 0.5$

ตีความว่า

ถ้า $(\max_1 - \max_2)$ น้อยกว่า 0.3) หรือ $(\max_1)$ น้อยกว่า 0.5) แล้ว
ตัวอักษรถูกพิจารณาว่าคลุมเครือ จะถูกคัดออก
ไม่เช่นนั้น
ตัวอักษรจำถูกต้องแล้ว

## 6.6 ผลการทดสอบ

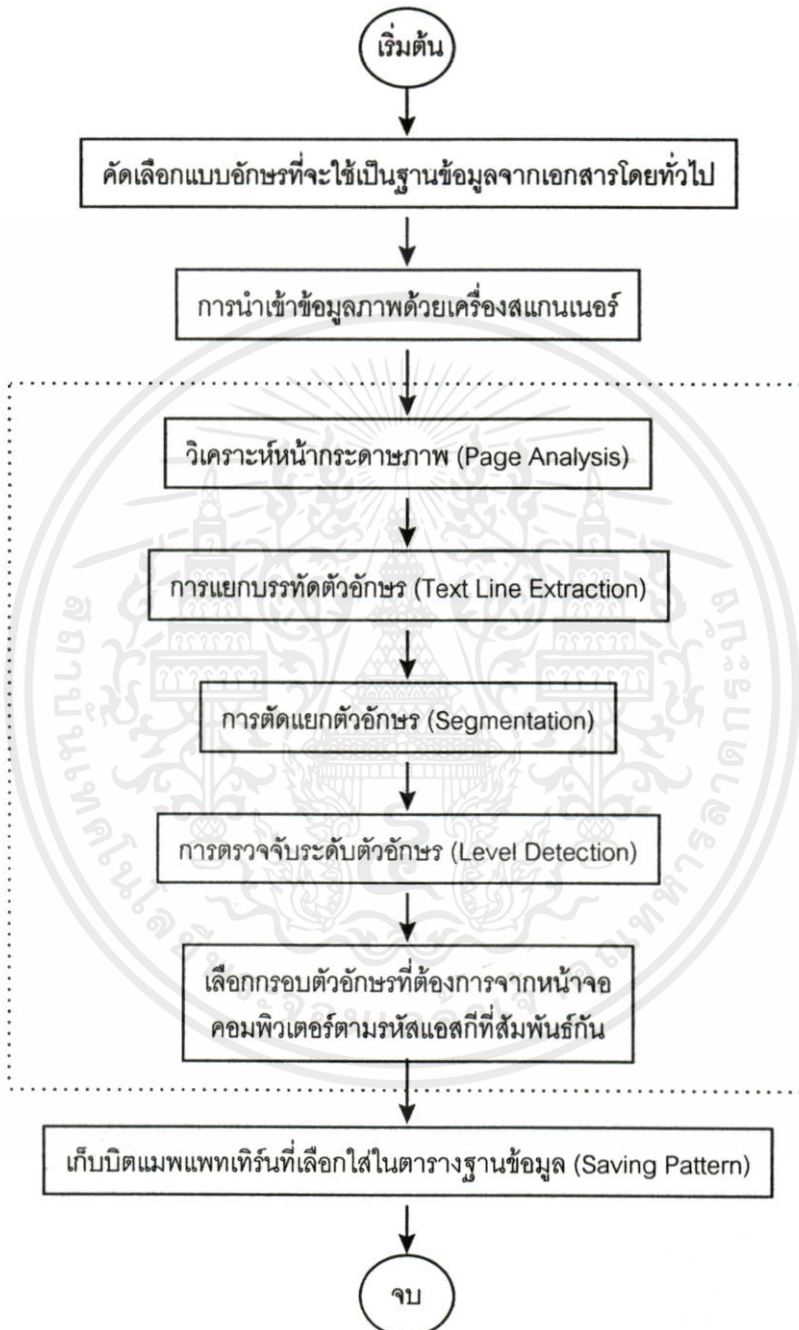
ในหัวข้อนี้จะแสดงวิธีการและขั้นตอนในทางปฏิบัติ โดยนำหลักการรู้จำตัวอักษรตัวพิมพ์ภาษาอังกฤษที่ได้กล่าวมาก่อนหน้านี้ นำมาเรียบเรียงและกำหนดวิธีการใช้งาน โดยแบ่งการทำงานออกเป็น 3 ขั้นตอน เริ่มจากจากขั้นตอนการอ่านตัวอักษรด้วยเครื่องสแกนเนอร์นำมาเก็บไว้เป็นต้นแบบ ขั้นตอนที่สองจะนำแพทเทิร์นตัวอักษรต้นแบบที่จัดเตรียมไว้มาฝึกสอน จนกระทั่งคอมพิวเตอร์สามารถรู้จำตัวอักษรต้นแบบได้ จากนั้นจึงเก็บผลการเรียนรู้ไว้เป็นฐานข้อมูล ขั้นตอนสุดท้ายเป็นการทดสอบผลการรู้จำแพทเทิร์นตัวอักษรจากเอกสารทั่วไป

### 6.6.1 ขั้นตอนการอ่านตัวอักษร

ขั้นตอนนี้จะคัดเลือกแพทเทิร์นตัวอักษรที่ต้องการจากเอกสารและหนังสือพิมพ์บางฉบับนำมาเก็บไว้เป็นแพทเทิร์นต้นแบบในฐานข้อมูล หลักการคัดเลือกตัวอักษรยังไม่มีวิธีการที่ชัดเจนอาศัยเพียงการเลือกเฉพาะแพทเทิร์นตัวอักษรที่มีความแตกต่างกันและใช้งานบ่อยที่สุดเท่านั้น

## แผนภูมิที่ 72

### แสดงแผนผังการคัดเลือกแพทเทิร์นตัวอักษร



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรณีใด จากแผนภูมิที่ 72 จะสังเกตเห็นว่ากระบวนการทั้งหมดอยู่ในขั้นตอนการนำเข้าข้อมูลภาพ  
 ด้วยเครื่องสแกนเนอร์และขั้นตอนการประมวลผลขั้นต้นทั้งสิ้น (การวิเคราะห์หน้ากระดาษ การแยก  
 บรรทัดตัวอักษร การตัดแยกตัวอักษร และการตรวจจับระดับตัวอักษร) ตัวอักษรที่คัดเลือกจะถูกนำมา

เก็บในแฟ้มข้อมูลในลักษณะบิตแมพภาพ ขนาดความกว้างและความสูงเท่าตัวอักษรจริงที่อ่านเข้ามา  
 ค้างแสดงในรูปที่ 73

### รูปที่ 73

แสดงตัวอย่างแพทเทิร์นตัวอักษรที่นำเข้ามา

**A** **B**

a) แพทเทิร์นตัวอักษร A

b) แพทเทิร์นตัวอักษร B

การทดสอบครั้งนี้จะคัดเลือกแบบอักษรที่ใช้กันมากที่สุดจากวารสารต่างประเทศจำนวน 7  
 เล่ม (Asian Communication, BYTE, Dbase III Plus Programming, EDN Asia, Prospect, Guid to the  
 UNIX Desktop และ esa bulletin) และจากหนังสือพิมพ์ภาษาอังกฤษจำนวน 2 ฉบับ (Bangkok Post  
 และ The Nation) ค้างแสดงในตารางที่ 15

#### ตารางที่ 15

แสดงที่มาของแบบอักษรสำหรับการฝึกสอน

ชื่อวารสาร/หนังสือพิมพ์	แบบอักษรที่ใช้
Asian Communication	Book Antiqua
BYTE	Times New Roman (Roman, CG Times, MS San Serif และ MS Serif)
Dbase III Plus Programming	Times New Roman
EDN Asia	Centry Schoolbook
Prospect	Garamond
Guid to the UNIX Desktop	Book Antiqua
esa bulletin	Arial
Bangkok Post	Centry Schoolbook
The Nation	Centry Schoolbook

แบบอักษรที่ใช้เป็นผลมาจากการนำแพทเทิร์นตัวอักษรจากวารสารดังกล่าวเทียบกับแบบอักษรที่มีอยู่ในระบบปฏิบัติการ ไมโครซอฟท์วินโดวส์ ซึ่งแบบอักษรบางแบบมีความเหมือนกันมาก เช่น Times New Roman, Roman, CG Times, MS San Serif และ MS Serif เป็นต้น ดังนั้นจึงขอเรียกใช้ชื่อแบบอักษรเพียงแบบเดียวคือ Times New Roman เท่านั้น จากการสุ่มเลือกวารสารและหนังสือพิมพ์ในท้องตลาดพอจะสรุปได้ว่า โดยส่วนมากจะใช้แบบอักษรเพียง 5 แบบคือ Book Antiqua, Times New Roman, Centry Schoolbook, Garamond และ Arial

การทดสอบครั้งนี้ได้คัดเลือกตัวอักษรภาษาอังกฤษ ตัวพิมพ์ใหญ่ ตัวพิมพ์เล็ก ตัวเลขและตัวอักษรพิเศษบางตัว (A..Z, a..z, 0..9, 0/;:\$?& และอื่นๆ) รวมทั้งสิ้น 74 คลาส โดยกลุ่มตัวอักษรในแต่ละคลาสที่นำเข้ามาจะมีรหัสแอสกีเฉพาะตัวเป็นของตนเองไม่ซ้ำกัน ในแต่ละคลาสอาจมีตัวอักษรมากกว่าหนึ่งตัว เช่น แพทเทิร์นตัวอักษร A จำนวน 10 ตัวอักษร แพทเทิร์นตัวอักษร C จำนวน 9 ตัวอักษรและแพทเทิร์นตัวอักษร Z จำนวน 8 ตัวอักษร เป็นต้น

#### รูปที่ 74

แสดงกลุ่มของแพทเทิร์นตัวอักษรที่จัดเก็บในแต่ละคลาส

A A A A A A A A A A  
 B B B B B B B B B B  
 C C C C C C C C C  
 D D D D D D D D D D  
 E E E E E E E E E E  
 F F F F F F F F F  
 G G G G G G G G G G  
 H H H H H H H H H H

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 74 (ต่อ)

I I I I I I I I  
 J J J J J J J J J  
 K K K K K K K K K  
 L L L L L L L L L L  
 M M M M M M M M M M  
 N N N N N N N N N N  
 O O O O O O O O O O  
 P P P P P P P P P P  
 Q Q Q Q Q Q Q Q Q Q  
 R R R R R R R R R R R  
 S S S S S S S S S S S  
 T T T T T T T T T T T  
 U U U U U U U U U U U  
 V V V V V V V V V V V  
 W W W W W W W W W W  
 X X X X X X X X X X X  
 Y Y Y Y Y Y Y Y Y Y Y  
 Z Z Z Z Z Z Z Z Z

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 74 (ต่อ)



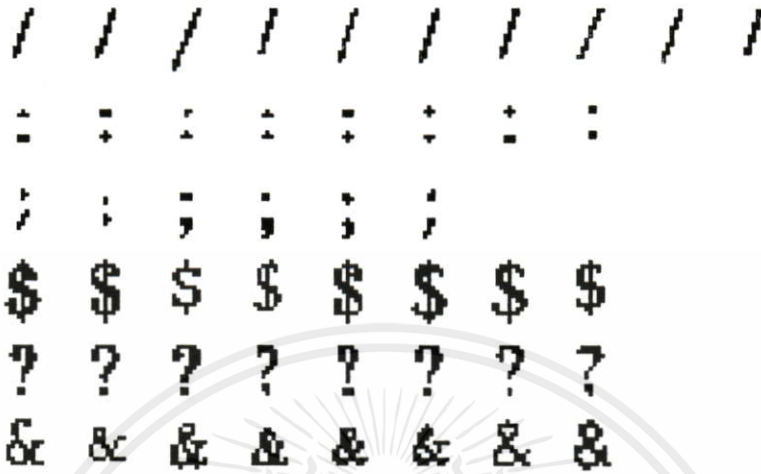
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 74 (ต่อ)

**s s s s s s s s s s**  
**t t t t t t t t t t**  
**u u u u u u u u u u**  
**v v v v v v v v v v**  
**w w w w w w w w w w**  
**x x x x x x x x x x**  
**y y y y y y y y y y**  
**z z z z z z z z z z**  
**0 0 0 0 0 0 0 0 0 0**  
**1 1 1 1 1 1 1 1 1 1**  
**2 2 2 2 2 2 2 2 2 2**  
**3 3 3 3 3 3 3 3 3 3**  
**4 4 4 4 4 4 4 4 4 4**  
**5 5 5 5 5 5 5 5 5 5**  
**6 6 6 6 6 6 6 6 6 6**  
**7 7 7 7 7 7 7 7 7 7**  
**8 8 8 8 8 8 8 8 8 8**  
**9 9 9 9 9 9 9 9 9 9**  
**( ( ( ( ( ( ( ( ( (**  
**) ) ) ) ) ) ) ) ) )**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งยังมีโทษทางแพ่งและอาญาที่ต้องลงโทษแก่เจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 74 (ต่อ)



a) การจัดเก็บตัวอักษรในชุดแรก



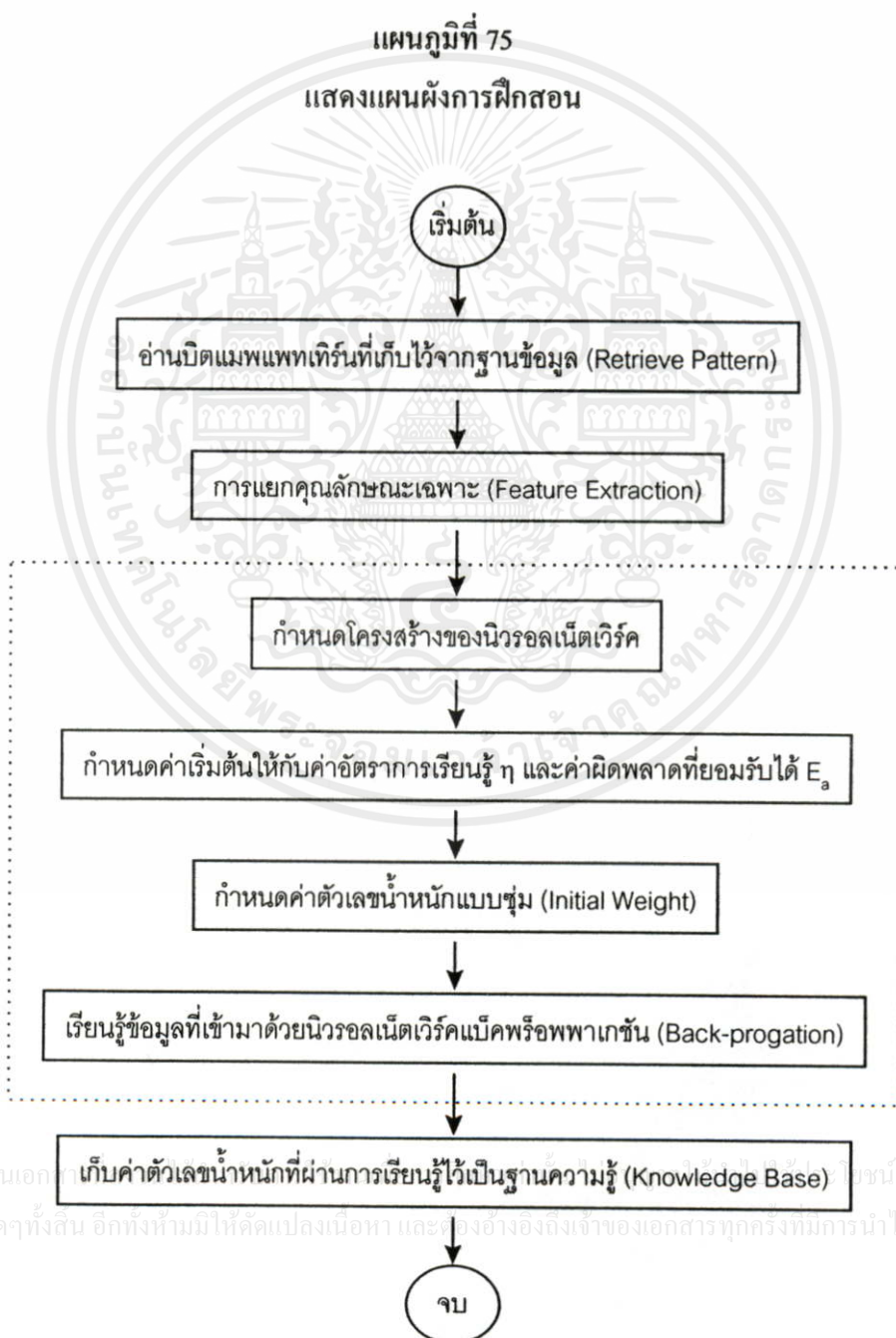
b) การจัดเก็บตัวอักษรในชุดที่สอง

การจัดเก็บแพทเทิร์นตัวอักษรจะแยกออกเป็นสองชุด ชุดแรกเป็นตัวอักษรใน Categories I (A..Z, a..z, 0..9, ( ) / ; \$ ? & ) รวม 70 คลาส ดังแสดงในรูปที่ 74 (a) และชุดที่สองเป็นแพทเทิร์นตัวอักษรใน Categories II ( - . , ' ) รวม 4 คลาส ดังแสดงในรูปที่ 74 (b) ในแต่ละคลาสจะมีจำนวนการจัดเก็บจำนวนตัวอักษรไม่เท่ากัน จำนวนแพทเทิร์นตัวอักษรในคลาสควรเป็นเท่าไรนั้นขึ้นอยู่กับผู้ใช้เป็นผู้กำหนด แต่หลักการมีอยู่ว่า ตัวอักษรต่างคลาสนั้นไม่ควรมีความเหมือนกันมากเกินไป และที่สำคัญควรเลือกใช้แบบอักษรที่เป็นมาตรฐานและใช้งานบ่อยๆ เท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 6.6.2 ขั้นตอนการฝึกสอน

ในขั้นตอนนี้จะนำอินพุตแพทเทิร์นตัวอักษรค้นแบบที่ได้จัดเก็บไว้มาผ่านกระบวนการแยกคุณลักษณะเฉพาะของตัวอักษร แล้วจึงฝึกสอนด้วยหลักการของนิรอลเน็ตเวิร์คแบ็คพร็อพกาชัน จนกระทั่งความสัมพันธ์ระหว่างอินพุตแพทเทิร์นกับตัวอักษรรหัสแอสกีเป็นไปตามที่กำหนดไว้



จากแผนภูมิที่ 75 แสดงแผนผังการฝึกสอนแพทเทิร์นตัวอักษรที่เก็บไว้ เริ่มจากการอ่านบิดแมทแพทเทิร์นที่เก็บไว้ในฐานข้อมูล นำมาคำนวณหาคุณลักษณะเฉพาะของแพทเทิร์นตัวอักษรแต่ละตัวด้วยหลักการของ Curve approximation ดังที่กล่าวมาแล้ว จากนั้นกำหนดโครงสร้างของเน็ตเวิร์คและตัวแปรที่จำเป็นในการฝึกสอน หลังจากเรียนรู้ข้อมูลด้วยหลักการของนิเวรอลเน็ตเวิร์คเบ็คหรือพาทเกชันแล้วจะทำการเก็บผลของตัวเลขน้ำหนักไว้ใช้ในขั้นตอนการรู้จำต่อไป

ก่อนการคำนวณหาคุณลักษณะเฉพาะของแพทเทิร์นตัวอักษร จำเป็นต้องคำนวณหาจุดรอบแพทเทิร์นตัวอักษรแต่ละตัวจากฐานข้อมูลตัวอักษรที่เก็บไว้เสียก่อน ซึ่งจำนวนจุดรอบแพทเทิร์นตัวอักษรทุกตัวมีค่าอยู่ในช่วง 60-200 จุด แต่โดยเฉลี่ยแล้วจุดรอบแพทเทิร์นมีค่าประมาณ 90 จุด จึงมีการกำหนดจุดรอบแพทเทิร์นขนาดคงที่ จำนวน 30 จุด (หนึ่งในสาม โดย  $n=30$ ) ด้วยคุณลักษณะเฉพาะ 7 คุณลักษณะ รวมเป็นอินพุทของนิเวรอลเน็ตเวิร์คจำนวน  $30 \times 7 = 210$  โหนด นำมาคำนวณหาคุณลักษณะเฉพาะด้วยสมการดังต่อไปนี้

$$f_0(n) = \text{PenUp}(n) \dots\dots\dots (106)$$

$$f_1(n) = \frac{X(n)+1}{2} \dots\dots\dots (107)$$

$$f_2(n) = \frac{Y(n)+1}{2} \dots\dots\dots (108)$$

$$f_3(n) = \frac{\cos \theta (n)+1}{2} \dots\dots\dots (109)$$

$$f_4(n) = \frac{\sin \theta (n)+1}{2} \dots\dots\dots (110)$$

$$f_5(n) = \frac{\cos \phi (n)+1}{2} \dots\dots\dots (111)$$

$$f_6(n) = \frac{\sin \phi (n)+1}{2} \dots\dots\dots (112)$$

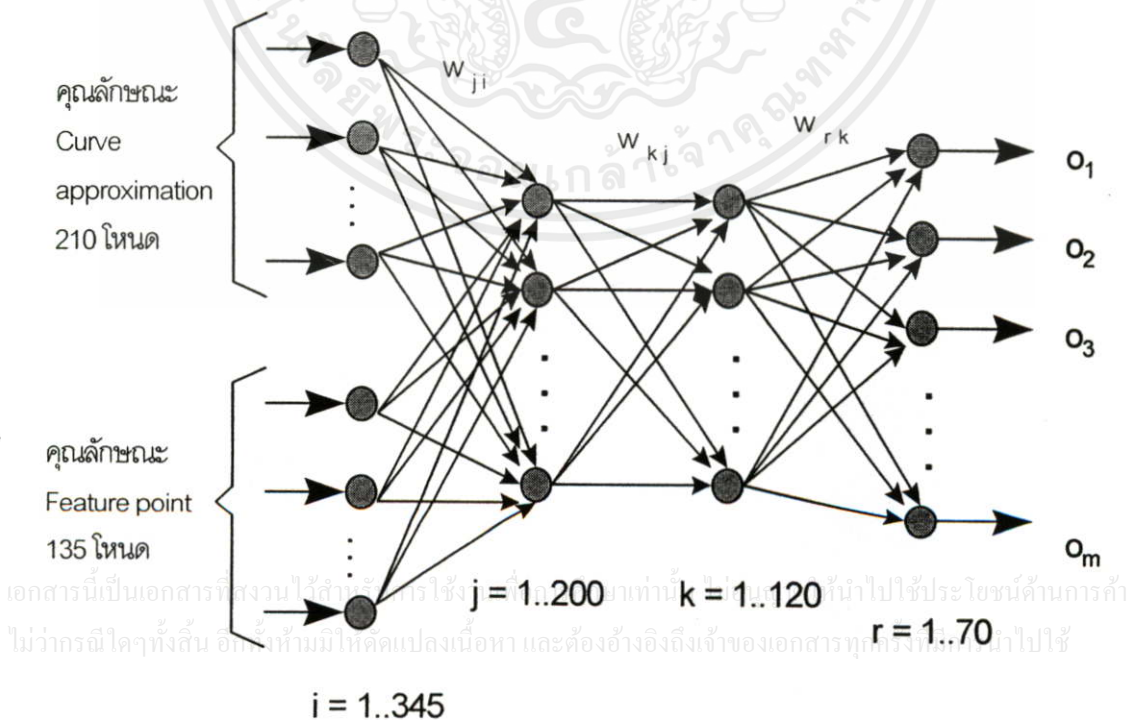
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ	$n$	เป็นตำแหน่งปัจจุบันของจุดรอบแพทเทิร์นมีค่าอยู่ระหว่าง 1 ถึง 30
	$n-1$	เป็นตำแหน่งก่อนหน้าของจุดรอบแพทเทิร์น
	$n+1$	เป็นตำแหน่งถัดไปของจุดรอบแพทเทิร์น
	$X_0$	ระยะห่างระหว่างจุดรอบแพทเทิร์นในแนวนอนกับจุดศูนย์กลางกรอบภาพ
	$Y_0$	ระยะห่างระหว่างจุดรอบแพทเทิร์นในแนวตั้งกับจุดศูนย์กลางกรอบภาพ
	$H$	ค่าความสูงของกรอบภาพ
	$\Delta x(n)$	เป็นผลต่างของในแนวแกนอนของจุด $n+1$ และ $n-1$
	$\Delta y(n)$	เป็นผลต่างของในแนวแกนตั้งของจุด $n+1$ และ $n-1$
	$\Delta s(n)$	เป็นระยะทางขจัดระหว่างจุด $n+1$ และ $n-1$ เท่ากับ $\sqrt{\Delta x(n)^2 + \Delta y(n)^2}$

คุณลักษณะเฉพาะของตัวอักษรที่คำนวณได้จะถูกนำมาเป็นอินพุทของของนิเวรอลเน็ตเวิร์คแบบค้ำพรีอบพาเกชัน จำนวน 210 โหนด

รูปที่ 76

แสดงโครงสร้างของนิเวรอลเน็ตเวิร์คสำหรับการรู้จำแพทเทิร์นตัวอักษรใน Categories I



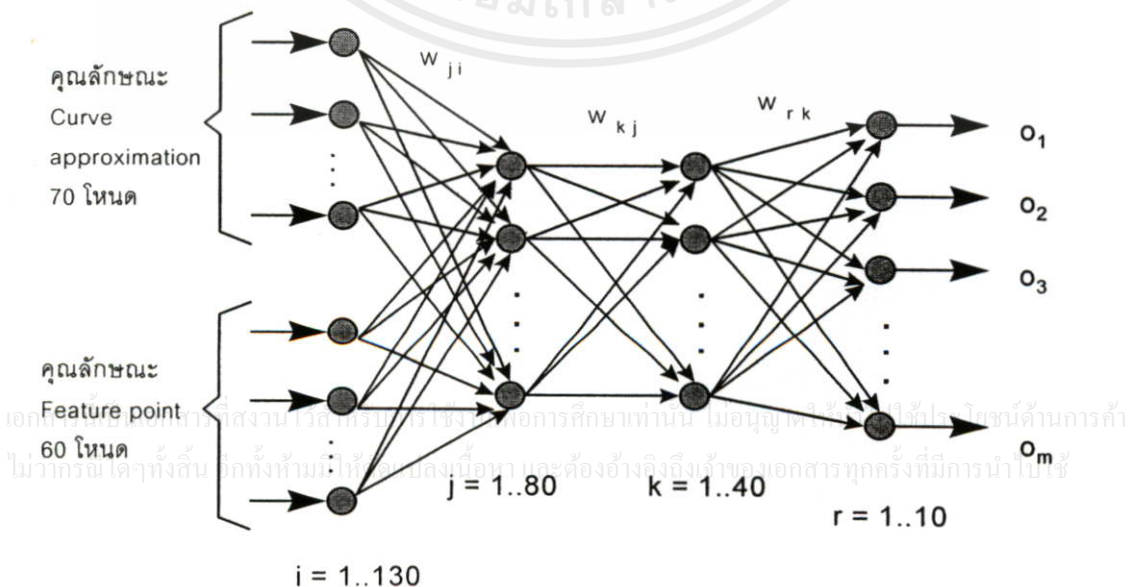
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในเท่านั้น ไม่ควรเผยแพร่สู่สาธารณะโดยไม่ได้รับอนุญาตจากเจ้าของเอกสาร

กำหนดให้ อินพุทเลเยอร์มี (i)	= 345 โหนด (210 + 135 โหนด)
ฮิดเดนเลเยอร์ชั้นแรก (j)	= 200 โหนด
ฮิดเดนเลเยอร์ชั้นที่สอง (k)	= 120 โหนด
เอาต์พุทเลเยอร์ (r)	= 70 โหนด
ค่าอัตราการเรียนรู้ ( $\eta$ )	= 0.5
ค่าผิดพลาดที่ยอมรับได้ ( $E_e$ )	= 0.001

จากรูปที่ 76 จะเห็นว่านิรอลเน็ตเวิร์คมีทั้งหมด 4 เลเยอร์ อินพุทเลเยอร์ ฮิดเดนเลเยอร์ชั้นแรก ฮิดเดนเลเยอร์ชั้นที่สอง และเอาต์พุทเลเยอร์ อินพุทที่เพิ่มขึ้นไปอีก 135 โหนด เป็นคุณลักษณะจุดภาพ (Pixels feature) หมายถึงการนำบิตแมพแพทเทิร์นเข้ามาเป็นอินพุทโดยตรงของนิรอลเน็ตเวิร์ค (ใช้วิธีการเหมือนบทที่ 5) ขนาดความกว้างและความสูงเท่ากับ  $9 \times 15$  จุดภาพ รวมทั้งหมด 135 โหนด การนำคุณลักษณะจุดภาพเข้ามารวมกับคุณลักษณะของ Curve approximation เพื่อลดช่องว่างอันเกิดเนื่องมาจากความผิดเพี้ยนของรูปร่างบิตแมพแพทเทิร์น รายละเอียดของการใช้คุณลักษณะเฉพาะถึง 2 อย่างเข้าด้วยกันนั้นดูได้จากภาคผนวก ค. ส่วนโครงสร้างของนิรอลเน็ตเวิร์คสำหรับตัวอักษรใน CategoriesII ถูกแสดงดังในรูปที่ 77

รูปที่ 77

แสดงโครงสร้างของนิรอลเน็ตเวิร์คสำหรับการรู้จำแพทเทิร์นตัวอักษรใน CategoriesII



กำหนดให้ อินพุทเลเยอร์ (i)	= 130 โหนด (70 + 60 โหนด)
ฮิดเดนเลเยอร์ชั้นแรก (j)	= 80 โหนด
ฮิดเดนเลเยอร์ชั้นที่สอง (k)	= 40 โหนด
เอาต์พุทเลเยอร์ (r)	= 4 โหนด
ค่าอัตราการเรียนรู้ ( $\eta$ )	= 0.5
ค่าผิดพลาดที่ยอมรับได้ ( $E_p$ )	= 0.001

หลังจากกำหนดโครงสร้างของนิวรอนเน็ตเวิร์คและจัดเตรียมข้อมูลอินพุทเรียบร้อยแล้ว จะกล่าวถึงกระบวนการฝึกสอนเฉพาะแพทเทิร์นตัวอักษรใน Categories I ดังต่อไปนี้

1. เริ่มแรกจะต้องกำหนดค่าเริ่มต้นให้แก่ตัวเลขน้ำหนักสำหรับทุกเลเยอร์  $w_{ji}$ ,  $w_{kj}$  และ  $w_{rk}$  ในลักษณะสุ่มเป็นค่าน้อยๆ โดยจำกัดในช่วง 0.0 ถึง 0.2 ทุกตัว เมื่อ  $i = 1..345$ ,  $j = 1..200$ ,  $k = 1..120$  และ  $r = 1..70$
2. จัดลำดับการฝึกสอนของอินพุทแพทเทิร์น  $o_{pi}$  ( $o_{p1}..o_{p345}$ ) และค่าเอาต์พุทแพทเทิร์น  $t_{pr}$  ( $t_{p1}..t_{p70}$ ) ในแต่ละชุดแพทเทิร์นให้สัมพันธ์กัน เมื่อจำนวนชุดของอินพุทแพทเทิร์น  $p = 1..70$
3. คำนวณหาค่าเอาต์พุท  $o_{pj}$ ,  $o_{pk}$  และ  $o_{pr}$  ในลักษณะป้อนไปหน้าตามลำดับ โดยใช้หลักการของเบ็คพร็อบพาเกชัน เมื่อ  $\theta_j$ ,  $\theta_k$  และ  $\theta_r$  เป็นไบอัสของระบบ กำหนดให้เท่ากับ 0.5

$$net_{pj} = \sum_{i=1}^{345} w_{ji} \cdot o_{pi} \dots\dots\dots (113)$$

$$net_{pk} = \sum_{j=1}^{200} w_{kj} \cdot o_{pj} \dots\dots\dots (114)$$

$$net_{pr} = \sum_{k=1}^{120} w_{rk} \cdot o_{pk} \dots\dots\dots (115)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
เมื่อ  $o_{pj} = f(net_{pj}) = \frac{1}{1 + e^{-(net_{pj} + \theta_j)}} \dots\dots\dots (116)$

$$o_{pk} = f(\text{net}_{pk}) = \frac{1}{1 + e^{-(\text{net}_{pk} + \theta_k)}} \dots\dots\dots(117)$$

$$o_{pr} = f(\text{net}_{pr}) = \frac{1}{1 + e^{-(\text{net}_{pr} + \theta_r)}} \dots\dots\dots(118)$$

เมื่อ	จำนวนอินพุตโหนดในเลเยอร์ i	=	1 .. 345
	จำนวนฮิดเดนโหนดในเลเยอร์ j	=	1 .. 200
	จำนวนฮิดเดนโหนดในเลเยอร์ k	=	1 .. 120
	จำนวนเอาต์พุตโหนดในเลเยอร์ r	=	1 .. 70

4. คำนวณสัญญาณค่าผิดพลาด  $\delta_{pr}$ ,  $\delta_{pk}$  และ  $\delta_{pj}$  ตามลำดับในลักษณะป้อนกลับ

$$\delta_{pr} = f'_r(\text{net}_{pr})(t_{pr} - o_{pr}) \dots\dots\dots(119)$$

$$\delta_{pk} = f'_k(\text{net}_{pk}) \sum_r \delta_{pr} w_{rk} \dots\dots\dots(120)$$

$$\delta_{pj} = f'_j(\text{net}_{pj}) \sum_k \delta_{pk} w_{kj} \dots\dots\dots(121)$$

5. ปรับค่าตัวเลขน้ำหนักในแต่ละคู่เลเยอร์ในลักษณะป้อนกลับ โดยเริ่มจากเอาต์พุตเลเยอร์ r ไปยังฮิดเดนเลเยอร์ชั้นที่สอง k ที่ เวลา t ใดๆ และ t+1 เป็นเวลาถัดไป กำหนดให้ค่าอัตราการเรียนรู้  $\eta$  มีค่าเท่ากับ 0.5

$$w_{rk}(t+1) = w_{rk}(t) + \eta \delta_{pr} o_{pk} \dots\dots\dots(122)$$

$$= w_{rk}(t) + \eta [o_{pr}(1-o_{pr})(t_{pr}-o_{pr})] o_{pk} \dots\dots\dots(123)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ หากท่านมีข้อสงสัยหรือต้องการข้อมูลเพิ่มเติม กรุณาติดต่อฝ่ายงานเอกสารคลังที่ปรึกษาฯ ได้  
จากนั้นจะปรับค่าตัวเลขน้ำหนักคู่เลเยอร์ถัดไปในลักษณะเดียวกัน เมื่อ  $w_{kj}(t)$  เป็นค่าตัวเลขน้ำหนักจากฮิดเดนเลเยอร์ชั้นที่สอง k ไปยังฮิดเดนเลเยอร์ชั้นแรก j ที่ เวลา t ใดๆ และ t+1 เป็นเวลาถัดไป กำหนดให้ค่าอัตราการเรียนรู้  $\eta$  มีค่าเท่ากับ 0.5

$$w_{kj}(t+1) = w_{kj}(t) + \eta \delta_{pk} o_{pj} \dots\dots\dots (124)$$

$$= w_{kj}(t) + \eta [o_{pk} (1-o_{pk}) \sum_{r=1}^{70} (\delta_{pr} \cdot w_{rk})] o_{pj} \dots\dots\dots (125)$$

เมื่อ  $w_{ji}(t)$  เป็นค่าตัวเลขน้ำหนักจากฮิดเดนเลเยอร์ชั้นแรก  $j$  ไปยังอินพุตเลเยอร์  $i$  ที่เวลา  $t$  ใดๆ และ  $t+1$  เป็นเวลาถัดไป กำหนดให้ค่าอัตราการเรียนรู้  $\eta$  มีค่าเท่ากับ 0.5

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_{pj} o_{pi} \dots\dots\dots (126)$$

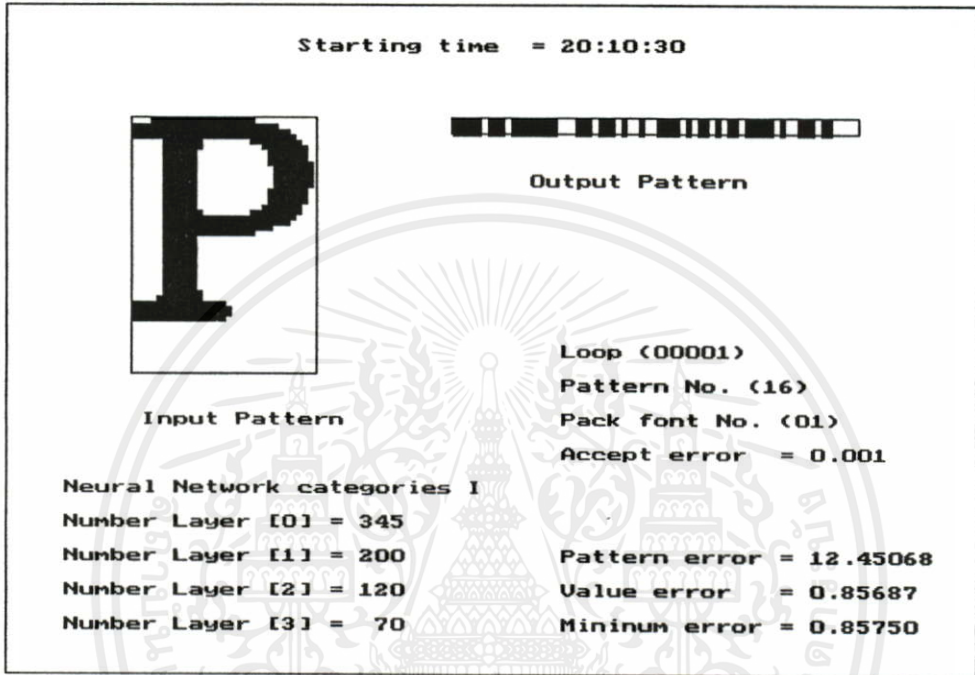
$$= w_{ji}(t) + \eta [o_{pj} (1-o_{pj}) \sum_{k=1}^{120} (\delta_{pk} \cdot w_{kj})] o_{pi} \dots\dots\dots (127)$$

6. เปรียบเทียบค่าผิดพลาดโดยรวมของระบบที่ได้จากการคำนวณ  $E$  และค่าผิดพลาดของระบบที่ยอมรับได้  $E_a$  เมื่อกำหนดให้  $E_a = 0.001$  หากมีค่าน้อยกว่าจะจบการฝึกสอนของโปรแกรม เก็บค่าตัวเลขน้ำหนักที่ได้จากการเรียนรู้สำหรับใช้เป็นฐานข้อมูลของการรู้จำต่อไป หากมีค่ามากกว่าให้กลับไปทำต่อในขั้นตอนที่ 2 โดยเปลี่ยนชุดของอินพุตแพทเทิร์น  $o_{pi}$  และเอาท์พุตแพทเทิร์น  $t_{pr}$  ที่สัมพันธ์กันในลำดับต่อไป

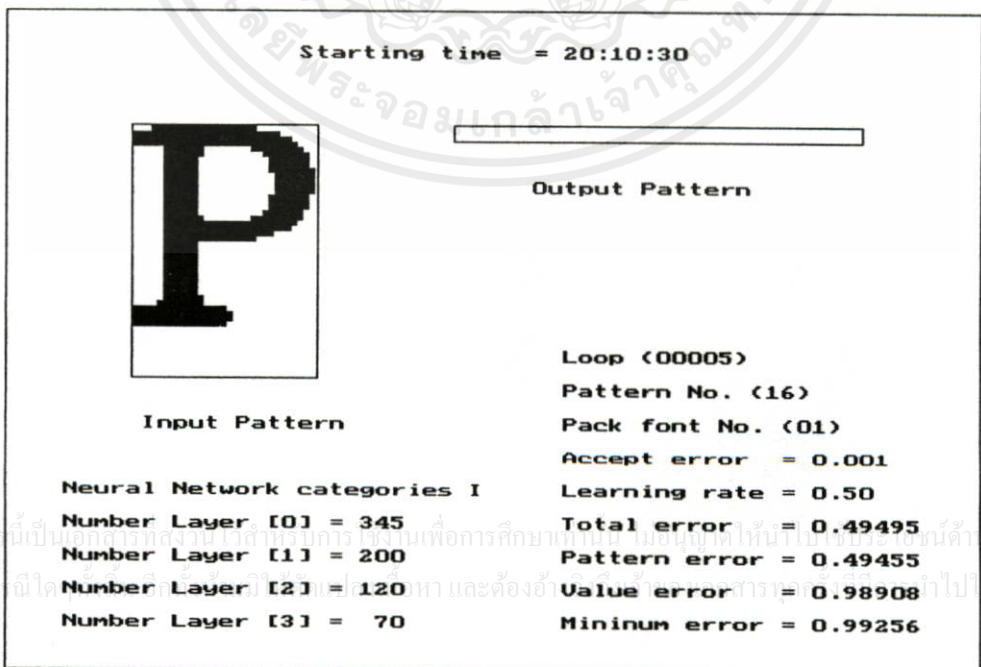
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 78

แสดงตัวอย่างการฝึกสอนแพทเทิร์น

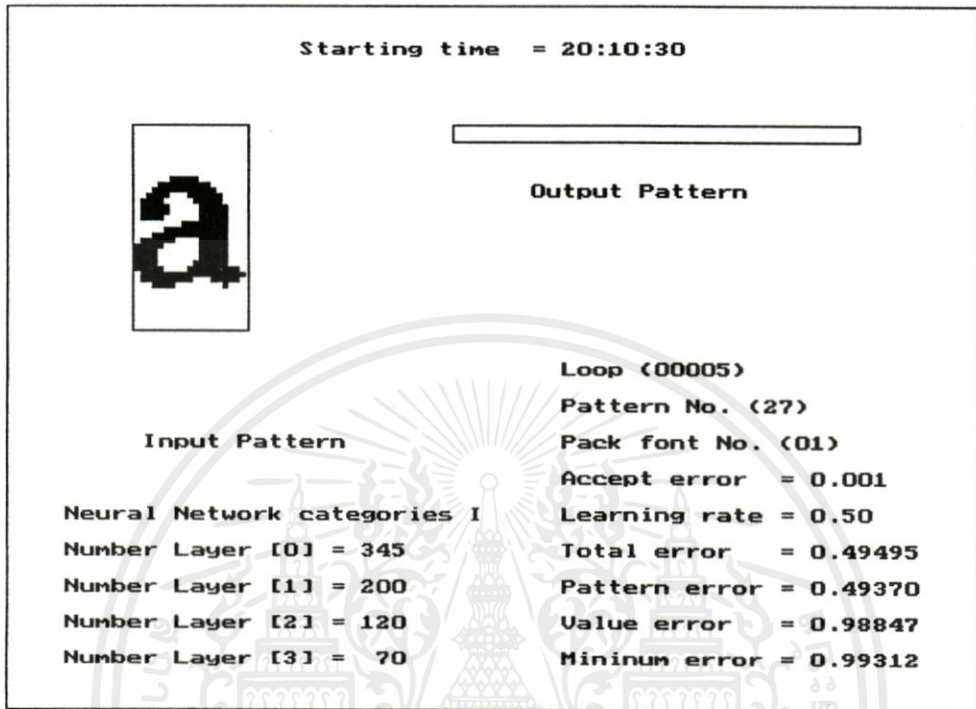


a) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 1

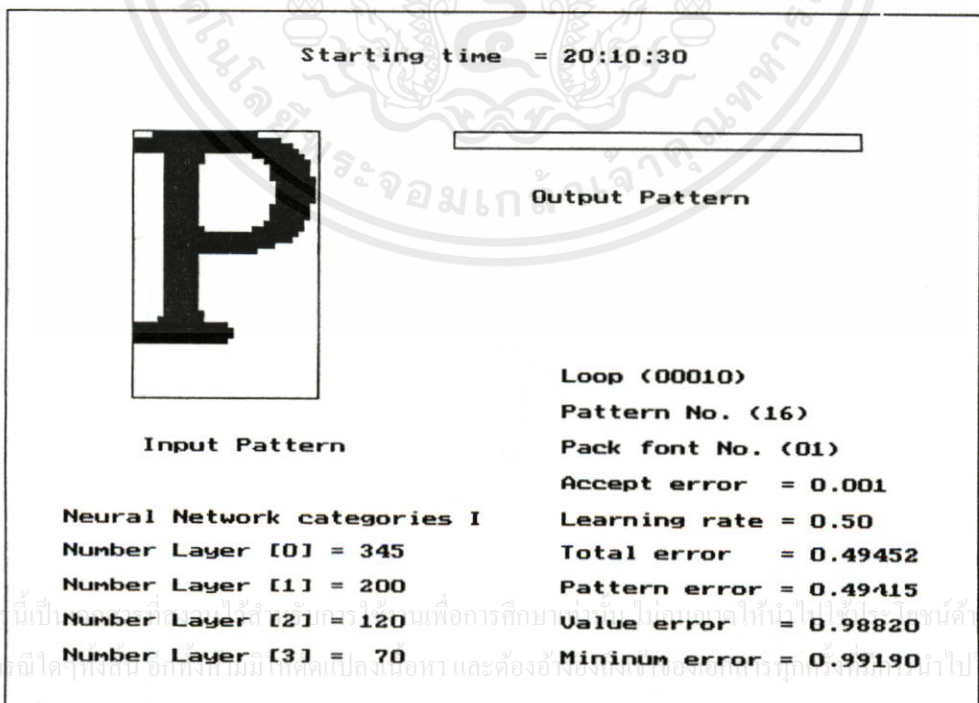


b) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 5

## รูปที่ 78 (ต่อ)

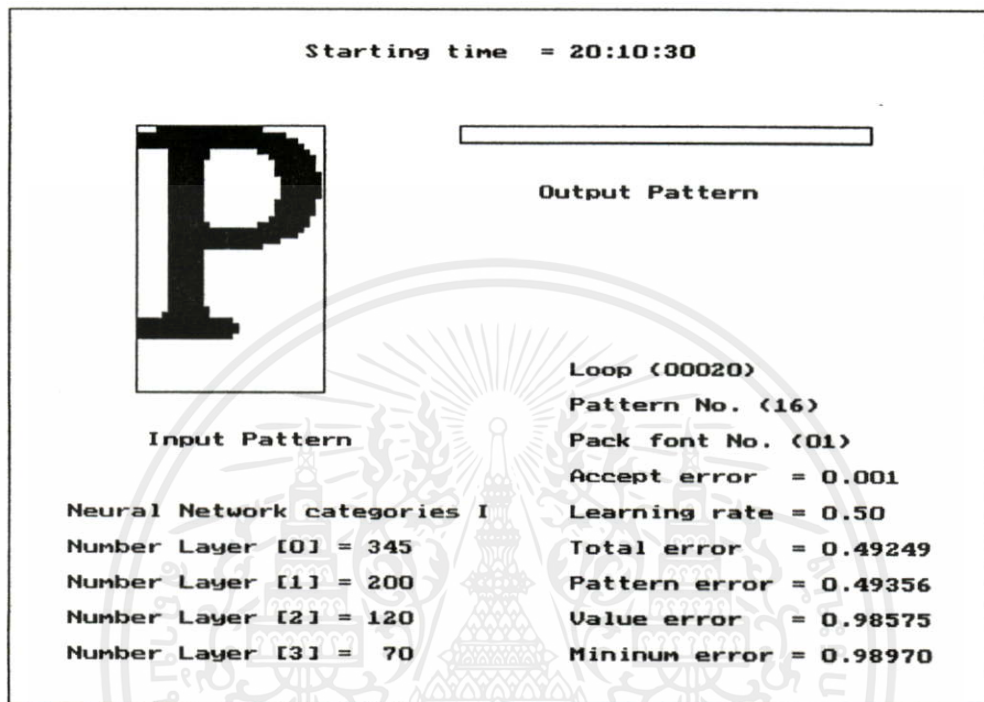


c) แสดงการฝึกสอนแพทเทิร์น "a" ในรอบที่ 5

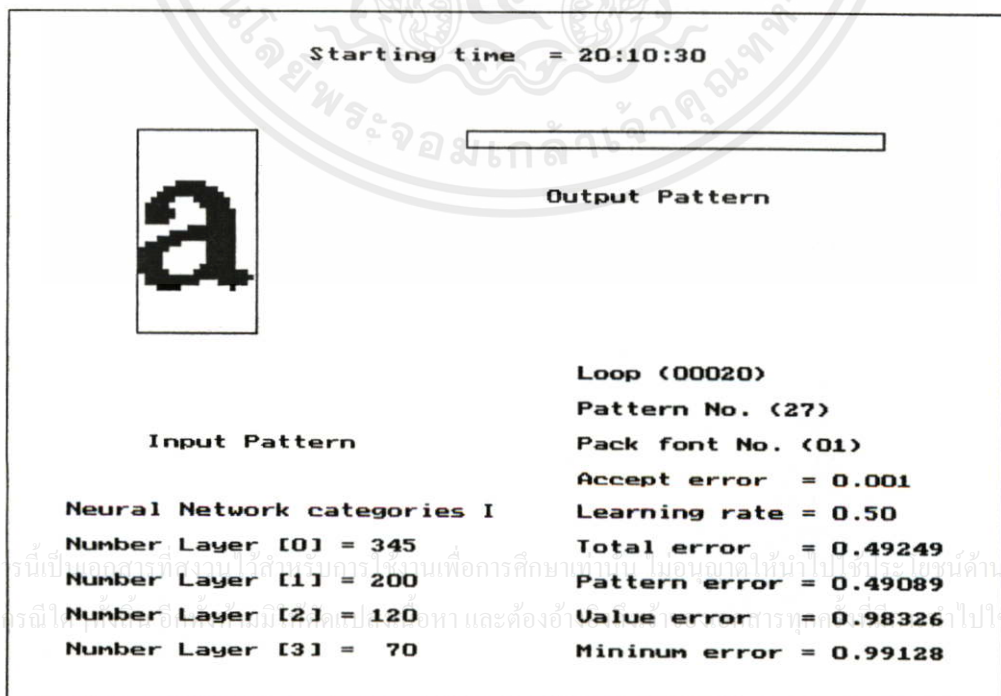


d) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 10

## รูปที่ 78 (ต่อ)

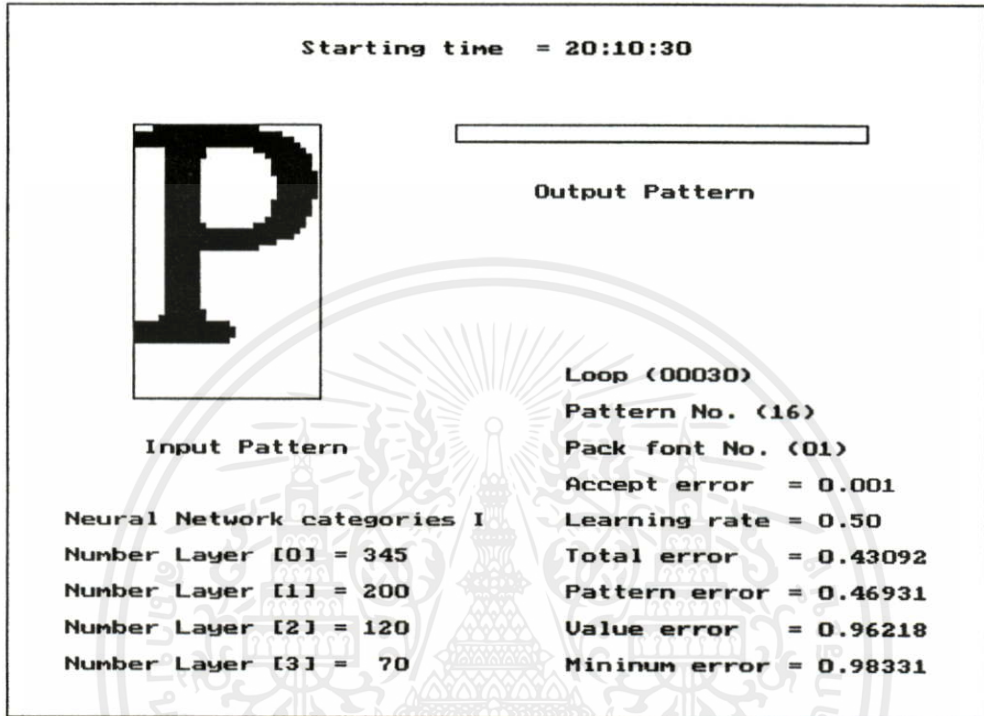


e) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 20

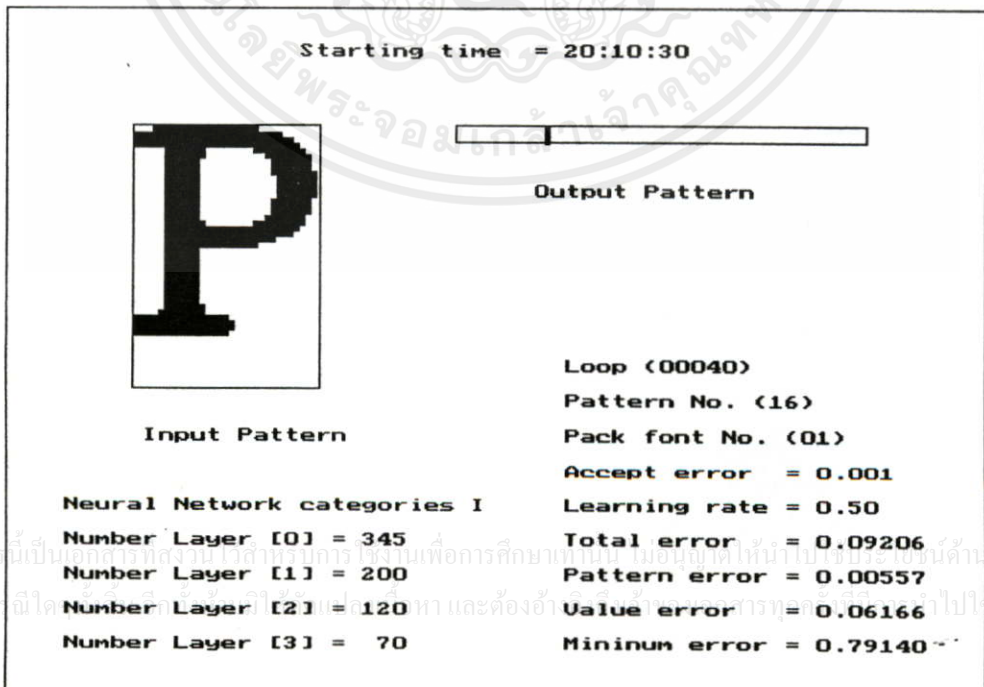


f) แสดงการฝึกสอนแพทเทิร์น "a" ในรอบที่ 20

## รูปที่ 78 (ต่อ)

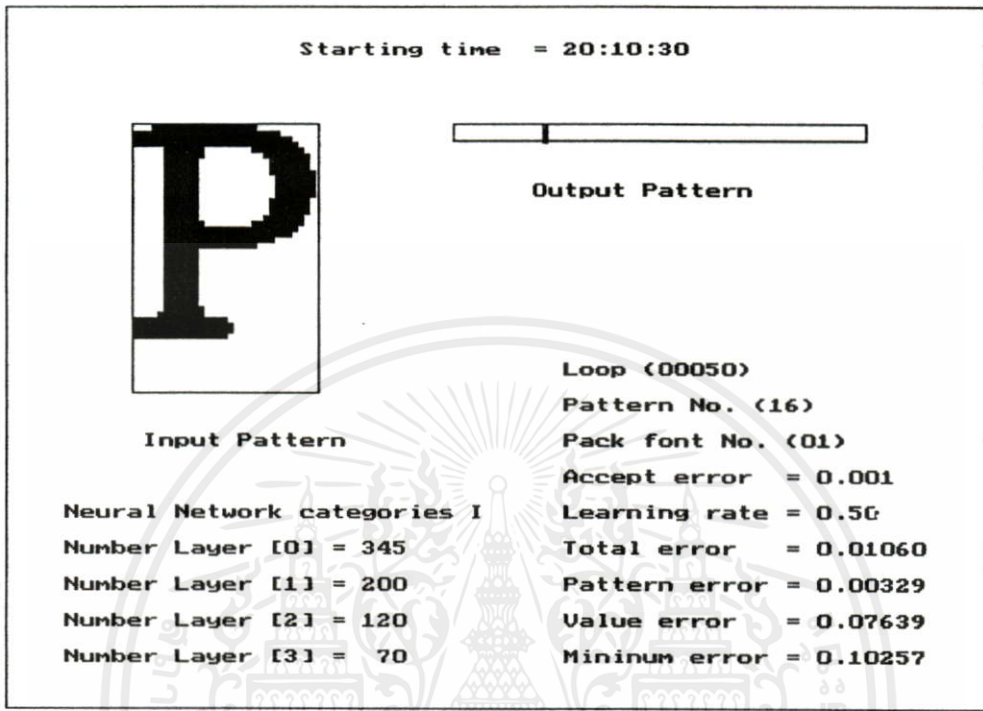


g) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 30

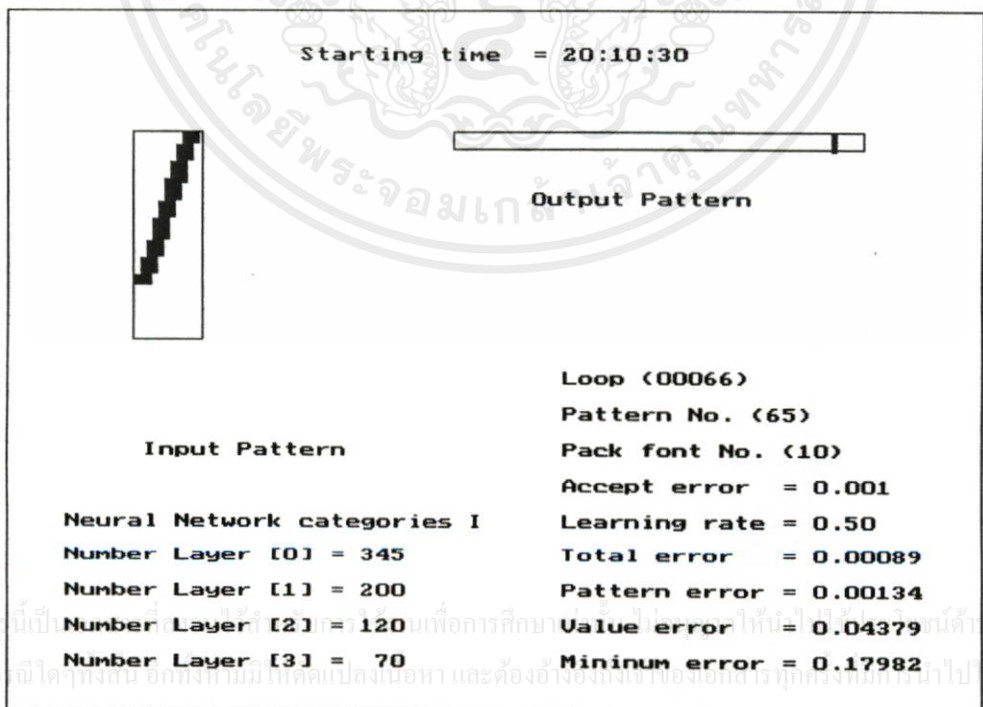


h) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 40

## รูปที่ 78 (ต่อ)



i) แสดงการฝึกสอนแพทเทิร์น "P" ในรอบที่ 50



j) แสดงการฝึกสอนแพทเทิร์น "7" ในรอบที่ 66

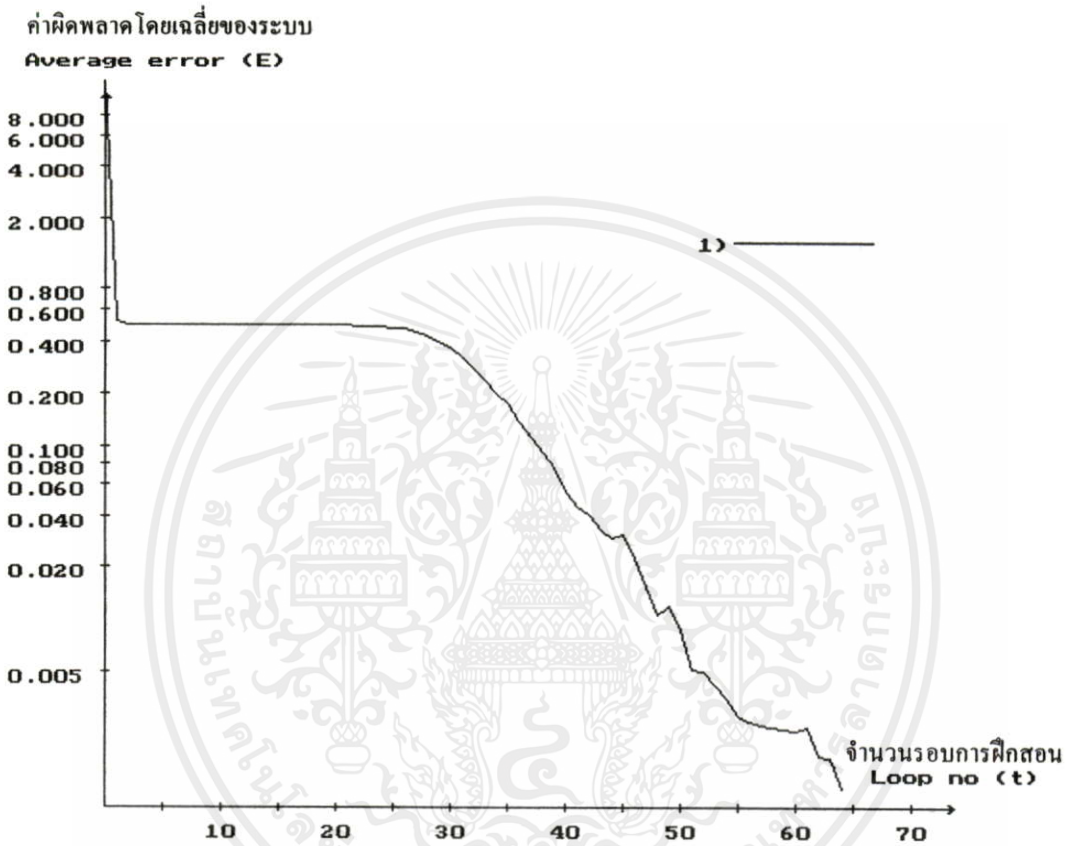
รูปที่ 78(a) ถึงรูปที่ 78(j) เป็นรูปแสดงการฝึกสอนเน็ตเวิร์ครวมทั้ง 66 รอบ โดยการฝึกสอนแพทเทิร์นตัวอักษรภาษาอังกฤษ ตัวพิมพ์ใหญ่ ตัวพิมพ์เล็ก ตัวเลขและตัวอักษรอักษรพิเศษทั้งหมดจำนวน 70 คลาส นับเป็น 1 รอบการทำงาน ดังนั้น จอภาพจะแสดงรูปมากมาย ซึ่งไม่สามารถพิมพ์ผลมาให้ดูได้ทั้งหมด จึงมีการเลือกเฉพาะบางภาพมาแสดงผล

พิจารณารูปที่ 78(a), รูปที่ 78(b), รูปที่ 78(d), รูปที่ 78(e), รูปที่ 78(g), รูปที่ 78(h) และ รูปที่ 78(i) แสดงการฝึกสอนแพทเทิร์นตัวอักษร "P" โดยในรอบแรกๆ ค่าเออร์ทพุทแพทเทิร์นที่ได้จากการคำนวณเป็นเออร์ทพุทที่ไม่ถูกต้อง แต่เมื่อจำนวนรอบเพิ่มขึ้น มีการปรับค่าตัวเลขน้ำหนักในทุกรอบการทำงานด้วยวิธีการของนิวรอลเน็ตเวิร์คแบบคพรีอพาทกชัน มีผลทำให้เออร์ทพุทดีขึ้นตามลำดับ รูปที่ 78(h) (รอบที่ 40) แสดงเออร์ทพุทแพทเทิร์นที่คำนวณได้มีลักษณะถูกต้องแล้ว แต่ค่าผิดพลาดของการฝึกสอน E ยังมีค่ามากกว่าค่าผิดพลาดที่ยอมรับได้  $E_a$  โปรแกรมจึงทำงานต่อไป จนกระทั่งค่าผิดพลาดของการฝึกสอน E มีค่าน้อยกว่าค่าผิดพลาดที่ยอมรับได้  $E_a$  ในรอบที่ 66 โปรแกรมจึงหยุดการทำงาน และมีการเก็บค่าตัวเลขน้ำหนักสุดท้ายไว้ใช้งานต่อไป

พิจารณารูปที่ 78(c), รูปที่ 78(f), และรูปที่ 78(j) แสดงการเรียนรู้ของแพทเทิร์นตัวอักษรตัวอื่น นอกเหนือจากแพทเทิร์นตัวอักษร "P" ในรอบการทำงานนั้น

รูปที่ 79

แสดงค่าผิดพลาดโดยเฉลี่ยของระบบลดลงตามจำนวนรอบ



รูปที่ 79 เป็นกราฟแสดงค่าผิดพลาดโดยเฉลี่ยสัมพันธ์กับจำนวนรอบที่เปลี่ยนแปลง จะเห็นว่าจำนวนรอบเพิ่มขึ้นอัตราความผิดพลาดของการฝึกสอนเริ่มลดลงจนกระทั่งมีค่าน้อยกว่า 0.001 ถือว่าสิ้นสุดกระบวนการฝึกสอน

ขั้นตอนการฝึกสอนแพทเทิร์นตัวอักษรใน Categories I และ Categories II จะใช้กระบวนการเช่นเดียวกัน จำนวนรอบของการฝึกสอนแพทเทิร์นตัวอักษรใน Categories II เกิดขึ้นทั้งหมด 58 รอบ ผลของตัวเลขน้ำหนักสุดท้ายจะถูกเก็บไว้ใช้งานต่อไป

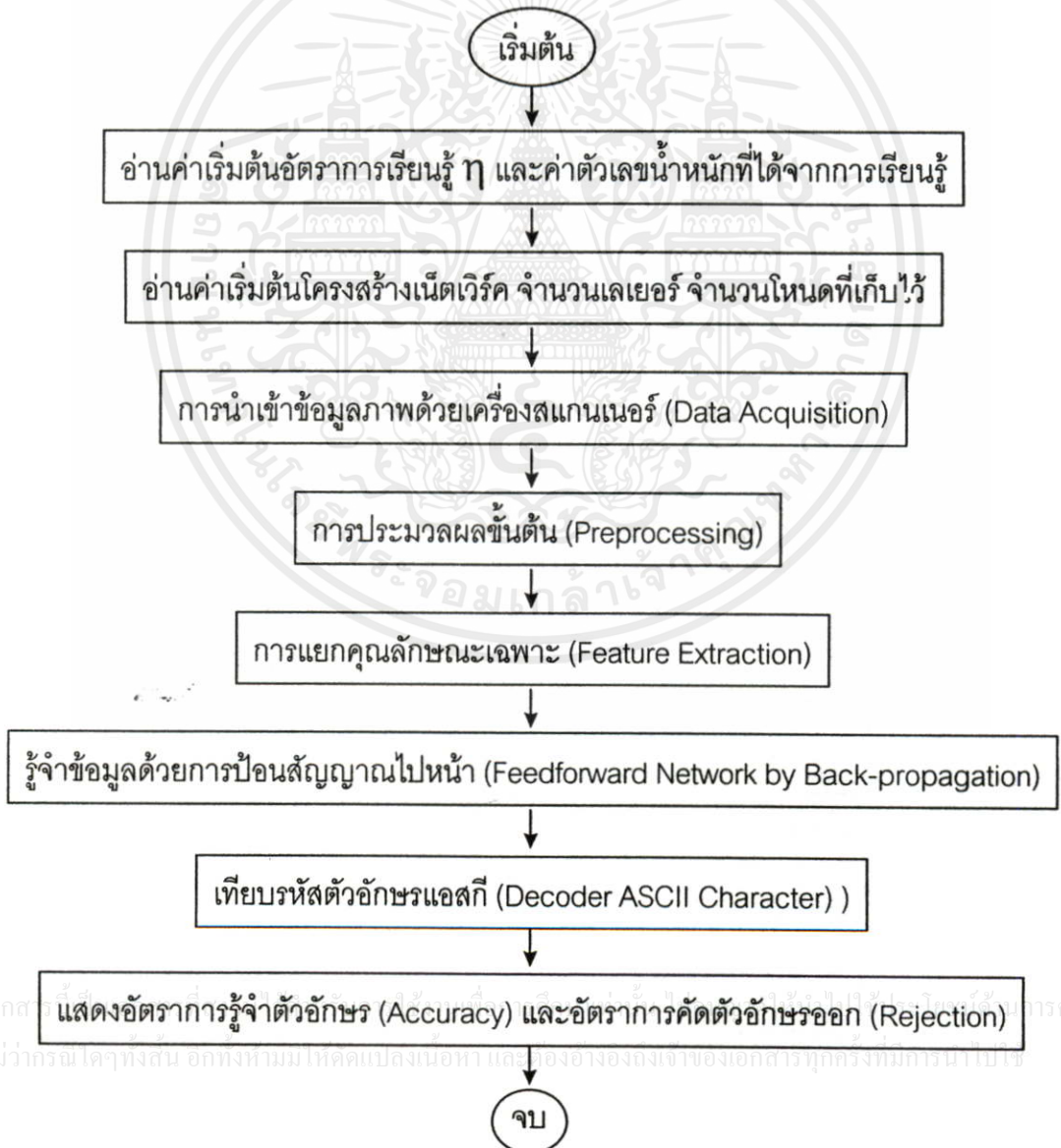
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 6.6.3 ขั้นตอนการทดสอบผลการรู้จำแพทเทิร์นตัวอักษร

การรู้จำตัวอักษรถูกทดสอบกับชุดข้อมูลจากวารสารต่างประเทศจำนวน 7 เล่มและหนังสือพิมพ์ภาษาอังกฤษจำนวน 2 ฉบับ อัตราการรู้จำตัวอักษรที่ถูกต้องโดยเฉลี่ยสำหรับภาพเอกสารที่มีความชัดเจนเท่ากับ 98.00% และภาพเอกสารที่ไม่ชัดเจนเท่ากับ 94.00% อัตราการคัดออกโดยไม่สามารถแยกประเภทได้จึงอยู่ในช่วง 2-6% เมื่อกำหนดให้ ค่าเทรคโอรลที่แตกต่างกัน  $\theta_{off}$  เท่ากับ 0.5 และค่าเทรคโอรลต่ำสุดที่ยอมรับได้  $\theta_{min}$  ถูกกำหนดให้เป็น 0.6

### แผนภูมิที่ 80

#### แสดงกระบวนการรู้จำตัวอักษร



แผนภูมิที่ 80 แสดงแผนผังกระบวนการรู้จำตัวอักษรภาษาอังกฤษจากภาพเอกสารที่อ่านได้จากเครื่องสแกนเนอร์ ประกอบด้วยการประมวลผลขั้นต้น การแยกคุณลักษณะเฉพาะของตัวอักษร และการรู้จำตัวอักษรด้วยนิวรอลเน็ตเวิร์ค โดยนำตัวเลขน้ำหนักที่ได้ฝึกสอนก่อนหน้านี้ นำมาใช้ในขั้นตอนการป้อนสัญญาณไปหน้า (Feed forward recall) จากนั้นจะถอดรหัสตัวอักษรจากค่าเอาต์พุทของนิวรอลโหนดให้อยู่ในรูปรหัสแอสกีซึ่งสามารถแสดงผลได้

### รูปที่ 81

แสดงตัวอย่างการรู้จำตัวอักษรจากวารสาร PROSPECT

CT scanners generate images by revolving an X-ray tube around the object and obtaining projections at different angles. Computers reconstruct these two-dimensional slices as three-dimensional models. The images are then displayed on cathode ray tube monitors and undergo post-processing for quantitative analysis.

a) แสดงเอกสารแบบบิตแมพภาพซึ่งนำเข้าสู่ด้วยเครื่องอ่านสแกนเนอร์

CT scanners generate images by revolving an X-ray tube around the object and obtaining projections at different angles. Computers reconstruct these two dimensional slices as three-dimensional models. The images are then displayed on cathode ray tube monitors and undergo post-processing for quantitative analysis.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในระบบเท่านั้น ไม่ควรเผยแพร่หรือใช้เพื่อวัตถุประสงค์อื่นโดยไม่ได้รับอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกไปลงในระบบอื่นใดขององค์กรของเอกสารทุกครั้งที่มีการนำไปใช้

b) ผลลัพธ์การรู้จำตัวอักษรแสดงในรูปรหัสแอสกี

รูปที่ 81 (a) แสดงตัวอย่างเอกสารแบบบิตแมพภาพที่อ่านวารสาร PROSPECT (เล่มที่ 7 ปี 1991 : 8) เครื่องสแกนเนอร์ใช้ความละเอียดในการอ่าน 300x300 จุดต่อนิ้ว (dpi) นำเข้ามาเป็นภาพขาวดำแล้วประมวลผลผ่าน โปรแกรมการรู้จำตัวอักษรที่เขียนขึ้น ทำให้ได้ผลลัพธ์เป็นตัวอักษรในรูปแบบรหัสแอสกีดังแสดงในรูปที่ 81 (b) ตัวอักษรแอสกีที่ขีดเส้นใต้ไว้เป็นผลมาจากความไม่แน่ใจในการวิเคราะห์ตัวอักษรด้วยนิเวรอลเน็ตเวิร์ค นั้นหมายความว่าค่าเอาท์พุทที่ได้หลังการรู้จำตัวอักษรมีค่าน้อยกว่าค่าเทรชโฮลต์ที่แตกต่างกัน  $\theta_{diff}$  (0.5) และมีค่าน้อยกว่าค่าเทรชโฮลต์ต่ำสุดที่ยอมรับได้  $\theta_{min}$  (0.6) เครื่องคอมพิวเตอร์ที่ใช้มีหน่วยประมวลผล Cyrix 6x86 ความเร็ว 166 MHz

### ตารางที่ 16

#### ผลการรู้จำตัวอักษรจากภาพในวารสาร POSPECT

จำนวนตัวอักษรทั้งหมดในภาพ (Total Character)	271	ตัว
จำนวนตัวอักษรที่เกิดความเหลื่อมล้ำกัน (Overlap Character)	5	ตัว
จำนวนตัวอักษรที่สัมผัสกัน (Touching Character)	2	ตัว
จำนวนตัวอักษรที่ไม่แน่ใจหลังการรู้จำ (Confuse Character)	7	ตัว
จำนวนตัวอักษรที่รู้จำผิดพลาด (Incorrect character)	0	ตัว
อัตราความถูกต้องในการรู้จำตัวอักษร (Recognition rate)	100	%
เวลาทั้งหมดในการรู้จำตัวอักษร (Working time)	48	วินาที

ตารางที่ 16 แสดงผลลัพธ์การรู้จำตัวอักษรจากวารสาร POSPECT ซึ่งมีตัวอักษรเหลื่อมล้ำกันจำนวน 5 ตัว เช่น ตัวอักษร ffe ของคำ different เป็นต้น ส่วนจำนวนตัวอักษรสัมผัสกันนั้นมี 2 ตัว เช่น ตัวอักษร tw ของคำ two เกิดการสัมผัสกันที่ปลายเส้นด้านข้างของแต่ละตัวอักษร แต่เมื่อใช้หลักการของการตัดแยกด้วยเส้นทางที่สั้นที่สุด (Shortest path segmentation) ทำให้ตัวอักษรทุกตัวแยกออกจากกันโดยสมบูรณ์ มีผลทำให้อัตราการรู้จำถูกต้องร้อยเปอร์เซ็นต์ แต่ก็ยังมีตัวอักษรบางตัวที่โปรแกรมไม่สามารถชี้ชัดได้ว่าถูกต้องทั้งหมด ซึ่งส่วนใหญ่แล้วเป็นตัวอักษร t เนื่องจากแบบอักษรที่ใช้มีหัวสั้นมาก จึงมีความคล้ายคลึงกับตัวอักษร r ทำให้การแปลความหมายเกิดความคลุมเครือ

## รูปที่ 82

แสดงตัวอย่างการรู้จำตัวอักษรจากวารสาร ASIAN COMMUNICATION

AN X.25 datacommunications project to link ChinaBank Corporation of the Philippines has recently been completed.

In the first phase of the US\$800 000 order, ChinaBank installed a Philips 108 port MCX212 datacommunications switch in the head office which has been linked to 20 MCX104 switches in remote offices through 50, S256

a) แสดงเอกสารแบบบิตแมพภาพซึ่งนำเข้าสู่ด้วยเครื่องอ่านสแกนเนอร์

AN X.25 datacommunications project to link ChinaBank Corporation of the Philippines has recently been completed.

In the first phase of the US\$800 000 order, ChinaBank installed a Philips 108 port MCX212 datacommunications switch in the head office which has been linked to 20 MCX104 switches in remote offices through 50, S256

b) ผลลัพธ์การรู้จำตัวอักษรแสดงในรูปรหัสแอสกี

รูปที่ 82 (a) แสดงตัวอย่างเอกสารแบบบิตแมพภาพที่อ่านได้จากวารสาร ASIAN COMMUNICATION (เดือนกุมภาพันธ์ ปี ค.ศ. 1996 : 10) เครื่องสแกนเนอร์ใช้ความละเอียดในการอ่าน 300x300 จุดต่อนิ้ว นำเข้ามาเป็นภาพขาวดำแล้วประมวลผลผ่านโปรแกรมการรู้จำตัวอักษรที่เขียนขึ้น ทำให้ได้ผลลัพธ์เป็นตัวอักษรในรูปแบบรหัสแอสกี ดังแสดงในรูปที่ 82 (b) ตัวอักษรแอสกีที่ถูกขีดเส้นใต้ไว้ เป็นผลมาจากความไม่แน่ใจในการวิเคราะห์ตัวอักษรด้วยนิเวศน์เดวิร์ค นั้นหมายความว่า ค่าเอาท์พุทที่ได้หลังการรู้จำตัวอักษรมีค่าน้อยกว่าค่าเทรคโธลที่แตกต่างกัน  $\theta_{diff}$  (0.5) หรือมีค่าน้อยกว่าค่าเทรคโธลต่ำสุดที่ยอมรับได้  $\theta_{min}$  (0.6)

### ตารางที่ 17

#### ผลการรู้จำตัวอักษรจากภาพในวารสาร ASIAN COMMUNICATION

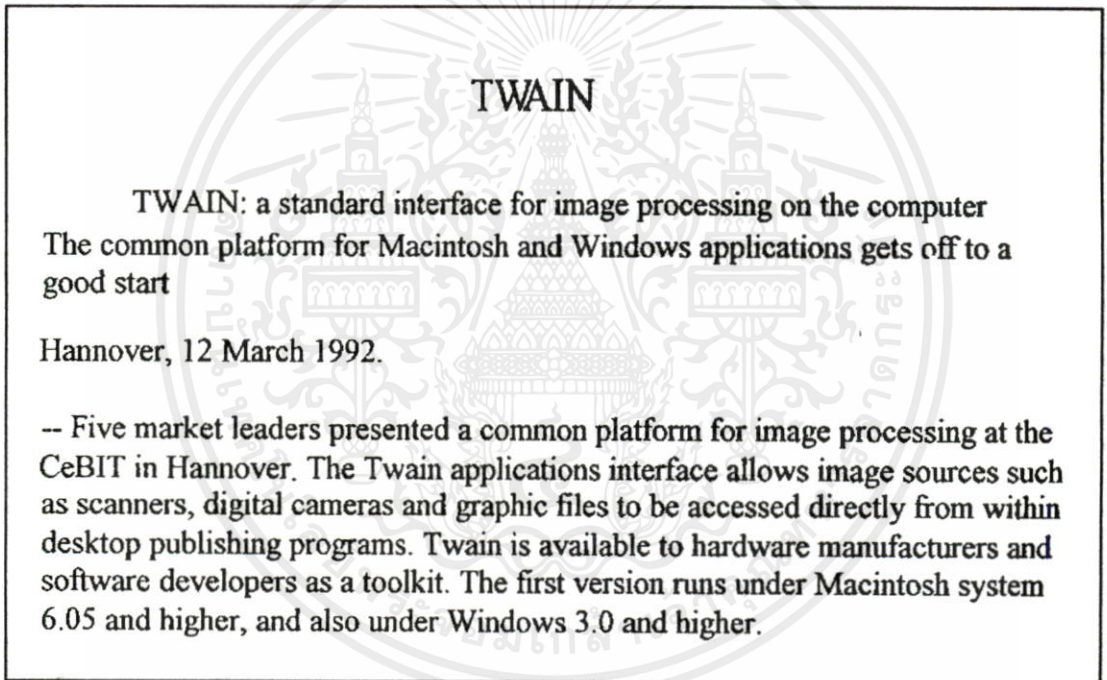
จำนวนตัวอักษรทั้งหมดในภาพ (Total Character)	279	ตัว
จำนวนตัวอักษรที่เกิดความเหลื่อมล้ำกัน (Overlap Character)	21	ตัว
จำนวนตัวอักษรที่สัมผัสกัน (Touching Character)	26	ตัว
จำนวนตัวอักษรที่ไม่แน่ใจหลังการรู้จำ (Confuse Character)	5	ตัว
จำนวนตัวอักษรที่รู้จำผิดพลาด (Incorrect character)	2	ตัว
อัตราความถูกต้องในการรู้จำตัวอักษร (Recognition rate)	99.28	%
เวลาทั้งหมดในการรู้จำตัวอักษร (Working time)	60	วินาที

ตารางที่ 17 แสดงผลลัพธ์การรู้จำตัวอักษรจากวารสาร ASIAN COMMUNICATION ซึ่งมีตัวอักษรเหลื่อมล้ำกันจำนวน 21 ตัว เช่น ตัวอักษร or ของคำ order เป็นต้น ส่วนจำนวนตัวอักษรสัมผัสกันนั้นมี 26 ตัว เช่น ตัวอักษร ffi ของคำ offices เกิดการสัมผัสกันที่ปลายเส้นด้านข้างของแต่ละตัวอักษร แต่เมื่อใช้หลักการของการตัดแยกด้วยเส้นทางที่สั้นที่สุด (Shortest path segmentation) ทำให้ตัวอักษรเกือบทุกตัวแยกออกจากกันโดยสมบูรณ์ มีผลทำให้อัตราความถูกต้องในการรู้จำตัวอักษรสูงมาก

โปรแกรมที่เขียนขึ้นเพื่อทดสอบการรู้จำตัวอักษรในวิทยานิพนธ์ฉบับนี้ สามารถแยกตัวอักษรที่เหลื่อมล้ำกันได้เป็นอย่างดี แต่สำหรับตัวอักษรที่สัมพันธ์กันนั้น มีทั้งแบบสัมพันธ์กันตั้งแต่สองตัวขึ้นไป เช่น ตัวอักษรสองตัวสัมพันธ์กัน ตัวอักษรสามตัวสัมพันธ์กัน ตัวอักษรสี่ตัวสัมพันธ์กัน เป็นต้น ส่วนใหญ่แล้ว ตัวอักษรสองตัวสัมพันธ์กันสามารถตัดแยกได้โดยง่าย แต่หากมากกว่านั้นอาจจำเป็นต้องอาศัยวิธีการที่ซับซ้อนยิ่งขึ้น

### รูปที่ 83

แสดงตัวอย่างการรู้จำตัวอักษรจากตัวอย่างภาพที่มากับ โปรแกรม Catch Word Pro



a) แสดงเอกสารแบบบิตแมพภาพซึ่งนำเข้าด้วยเครื่องอ่านสแกนเนอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## TWAIN

TWAIN: a standard interface for image processing on the computer  
The common platform for Macintosh and Windows applications gets oRto a  
good stan

Hannover, 12 March 1992.

-- Five market leaders presented a common platform for image processing at the CeBIT in Hannover. The Twain applications interface allows image sources such as scanners, digital cameras and &uaphic nles to be accessed directly njom within desktop publishing programs. Twain is available to hardware manufacturers and sonware developers as a toolkit. The nrst version runs under Macintosh system 6.05 and higher, and also under Windows 3.0 and hig1ler.

### b) ผลลัพธ์การรู้จำตัวอักษรแสดงในรูปรหัสแอสกี

รูปที่ 83(a) แสดงตัวอย่างภาพเอกสารที่มาจากโปรแกรม Catch Word Pro ซึ่งเป็น โปรแกรมการรู้จำตัวอักษรที่วางขายทั่วไปในท้องตลาด นำเข้ามาเป็นภาพขาวดำแล้วประมวลผลผ่าน โปรแกรมการรู้จำตัวอักษรที่เขียนขึ้น ทำให้ได้ผลลัพธ์เป็นตัวอักษรในรูปแบบรหัสแอสกี ดังแสดงในรูปที่ 83 (b) ตัวอักษรแอสกีที่ถูกขีดเส้นใต้ไว้เป็นผลมาจากความไม่แน่ใจในการวิเคราะห์ตัวอักษรด้วยนิรอลเน็ตเวิร์ค นั้นหมายความว่าค่าเอาท์พุทที่ได้หลังการรู้จำแพทเทิร์นตัวอักษรมีค่าน้อยกว่าค่าเทรคโสรลที่แตกต่างกัน  $\theta_{diff}$  (0.5) หรือมีค่าน้อยกว่าค่าเทรคโสรลต่ำสุดที่ยอมรับได้  $\theta_{min}$  (0.6)

### ตารางที่ 18

ผลการรู้จำตัวอักษรจากภาพจากตัวอย่างโปรแกรม Catch Word Pro

จำนวนตัวอักษรทั้งหมดในภาพ (Total Character)	536	ตัว
จำนวนตัวอักษรที่เกิดความเหลื่อมล้ำกัน (Overlap Character)	46	ตัว
จำนวนตัวอักษรที่สัมผัสกัน (Touching Character)	91	ตัว
จำนวนตัวอักษรที่ไม่แน่ใจหลังการรู้จำ (Confuse Charater)	14	ตัว
จำนวนตัวอักษรที่รู้จำผิดพลาด (Incorrect charrecter)	19	ตัว

## ตารางที่ 18 (ต่อ)

อัตราความถูกต้องในการรู้จำตัวอักษร (Recognition rate)	96.45	%
เวลาทั้งหมดในการรู้จำตัวอักษร (Working time)	188	วินาที

ตารางที่ 18 แสดงผลลัพธ์การรู้จำตัวอักษรจากตัวอย่างของโปรแกรม Catch Word Pro ซึ่งมีตัวอักษรเหลื่อมล้ำกันจำนวน 46 ตัว เช่น ตัวอักษร WA ของคำ TWAIN เป็นต้น ส่วนจำนวนตัวอักษรสัมผัสกันนั้นมี 91 ตัว เช่น ตัวอักษร art ของคำ start , ff ของคำ off , gr ของคำ graphic , fr ของคำ from , ft ของคำ software และ gh ของคำ higher เป็นต้น เมื่อใช้หลักการของการตัดแยกด้วยเส้นทางที่สั้นที่สุด (Shortest path segmentation) ทำให้ตัวอักษรเกือบทุกตัวแยกออกจากกันได้ แต่ก็ยังไม่สมบูรณ์นักเนื่องมาจากตัวอักษรที่สัมผัสกันนั้นแนบชิดกันมากจนแยกไม่ออก แต่อัตราความถูกต้องในการรู้จำตัวอักษรยังอยู่ในเกณฑ์ที่น่าพอใจ

## ตารางที่ 19

ผลการรู้จำตัวอักษรจากวารสารที่ใช้ฝึกสอนและวารสารที่ไม่ได้ใช้ฝึกสอน

วารสารหรือหนังสือพิมพ์	แบบอักษรที่ใช้	จำนวนตัวอักษรในภาพ	จำนวนตัวอักษรที่รู้จำผิดพลาด	เปอร์เซ็นต์การรู้จำตัวอักษร
Asian Communication	Book Antiqua	2,757	49	98.223
BYTE	Times New Roman	1,747	36	97.939
Dbase III Plus Programming	Times New Roman	4,849	110	97.731
EDN Asia	Century Schoolbook	2,360	22	99.067
Prospect	Garamond	2,334	52	97.772
Guid to the UNIX Desktop	Book Antiqua	3,141	21	99.331
Bangkok Post	Century Schoolbook	1,734	14	99.193
The Nation	Century Schoolbook	1,268	50	96.057
เฉลี่ย		20,190	354	98.247

a) ผลการรู้จำตัวอักษรจากวารสารที่ใช้ฝึกสอน

## ตารางที่ 19 (ต่อ)

วารสารหรือหนังสือพิมพ์	คล้ายแบบอักษร	จำนวนตัวอักษรในภาพ	จำนวนตัวอักษรที่รู้จักผิดพลาด	เปอร์เซ็นต์การรู้จำตัวอักษร
Datamation	Century Schoolbook	3,204	65	97.971
Asian Electronics Engineer	Times new Roman & Garamond	3,126	109	96.513
Corel Photo-Paint 5	Bookman Old Style	2,643	71	97.301
World Wide Web Database	Garamond	1,521	30	98.028
Unix World Magazine	Times new Roman & Century Schoolbook	2,292	128	94.415
	เฉลี่ย	12,786	403	96.848

## b) ผลการรู้จำตัวอักษรจากวารสารที่ไม่ได้ใช้ฝึกสอน

การทดสอบครั้งนี้จะเลือกภาพของข้อความบางส่วนจากวารสารหรือหนังสือพิมพ์ในลักษณะชุ่มตัวอย่างจากบางหน้าเท่านั้นไม่สามารถทำทั้งหมดได้ การทดสอบจะแบ่งเป็นสองส่วน ส่วนแรกจะทดสอบการตัวอักษรจากวารสารที่ใช้ฝึกสอน ดังแสดงในตารางที่ 19 (a) ส่วนที่สองจะทดสอบการตัวอักษรจากวารสารที่ไม่ได้ใช้ฝึกสอนดังแสดงในตารางที่ 19 (b) ผลการรู้จำจากวารสารที่ใช้ฝึกสอนจะมีค่ามากกว่าการรู้จำจากวารสารที่ไม่ได้ใช้ฝึกสอน เนื่องจากแบบอักษรในวารสารที่ใช้ฝึกสอนจะมีส่วนใกล้เคียงกับแบบอักษรที่มีอยู่ในฐานข้อมูลมากกว่า

จากการทดสอบการรู้จำตัวอักษรจากหนังสือพิมพ์ The Nation พบว่าอัตราการเรียนรู้จำถูกต้องมีค่าน้อย เนื่องจากตัวอักษรของภาพที่อ่านได้มีลักษณะติดกันมาก ทำให้เปอร์เซ็นต์ความถูกต้องอยู่ที่ 96.057 เปอร์เซ็นต์ (แม้ว่าจะเป็นแบบอักษรที่ผ่านการฝึกสอนมาแล้วก็ตาม) การรู้จำตัวอักษรจากวารสาร Unix World Magazine มีอัตราการเรียนรู้จำถูกต้องน้อยเช่นกัน เนื่องจากตัวอักษรของภาพที่อ่านเข้ามาเป็นแบบอักษรที่ไม่เคยใช้ฝึกสอนมาก่อน และไม่มีความคล้ายคลึงกับแบบอักษรที่มีอยู่ในฐานข้อมูลทำให้เปอร์เซ็นต์ความถูกต้องอยู่ที่ 94.415 เปอร์เซ็นต์

จากการทดสอบยังพบอีกว่า หากตัวอักษรของภาพมีลักษณะติดกันมากจนไม่สามารถแยกออกได้ จะทำให้อัตราการเรียนรู้จำถูกต้องมีค่าอยู่ในช่วง 85.00 เปอร์เซ็นต์ ถึง 94.00 เปอร์เซ็นต์เท่านั้น

การทดสอบอัตราการเรียนรู้ที่ผ่านมาให้ผลลัพธ์ที่น่าพอใจ บางส่วนที่เกิดความผิดพลาดเนื่องมาจากการแบ่งแยกตัวอักษรได้ไม่ดีพอ บางส่วนที่เกิดความผิดพลาดเนื่องมาจากแพทเทิร์นตัวอักษรที่มีรูปร่างคล้ายกัน การแปลความหมายด้วยนิรอลเน็ตเวิร์คจึงคิดเพี้ยนไปบ้าง หากต้องการแก้ไขให้ดีขึ้น จำเป็นต้องวิเคราะห์ตัวอักษรเพิ่มเติมด้วยขั้นตอนการประมวลผลในลำดับหลัง ซึ่งขณะนี้ยังไม่อยู่ในส่วนของงานวิจัยฉบับนี้ คงเป็นเพียงความหวังของงานวิจัยต่อไปในอนาคต

## 6.7 ข้อเสนอแนะและผลการวิเคราะห์

1. การคัดเลือกตัวอักษรที่จะนำมาฝึกสอนจำเป็นต้องคัดเลือกด้วยความระมัดระวัง สิ่งที่ต้องคำนึงถึงคือ อินพุตแพทเทิร์นที่มีลักษณะหน้าตาเหมือนกันแต่มีค่าเอาต์พุตหลายค่า (one-to-many) เช่น ตัวเลขหนึ่ง "1" กับตัวอักษรไอ "I" หากแบบอักษรของทั้งสองเป็นแบบไม่มีหัว การฝึกสอนด้วยนิรอลเน็ตเวิร์คจะไม่สิ้นสุด (nonconverge) กราฟของค่าผิดพลาดโดยรวมของระบบจะเกิดการแกว่งไม่ลดลงตามจำนวนรอบที่ฝึกสอน ดังนั้นควรสร้างความสัมพันธ์ของอินพุตแพทเทิร์นตัวอักษรกับเอาต์พุตให้มีค่าแบบ one-to-one หรือ many-to-one เท่านั้น (ไม่ควรเกิดกรณี one-to-many)

2. มีหลายความเหมาะสมซึ่งตัวอักษรตัวเลขไม่สามารถแยกแยะออกจากตัวอักษรตัวใหญ่ที่สอดคล้องกันได้ เช่น ตัวเลขศูนย์ "0" กับตัวอักษรโอ "O" ตัวเลขหนึ่ง "1" กับตัวอักษรไอ "I" ตัวเลขห้า "5" กับตัวอักษรเอส "S" เป็นต้น อย่างไรก็ตาม เป็นที่รู้กันว่าผลการตีความตัวเลขหรือตัวอักษรไม่ขึ้นกับวิธีการแยกประเภทตัวอักษรด้วยนิรอลเน็ตเวิร์ค ซึ่งจำเป็นต้องใช้วิธีอื่นพิจารณาแทน เช่น โปรแกรมตรวจสอบการสะกดคำ เป็นต้น

3. การพัฒนาผลกรู้อัจฉริยะที่ดียิ่งขึ้น ควรทำการแยกประเภทของตัวอักษรแบบหยาบๆ (Coarse classification) ออกเป็นกลุ่มหลายๆกลุ่ม โดยอาศัยคุณสมบัติทางสถิติและคุณสมบัติพื้นฐานทางด้านโครงสร้าง พยางค์ คำ และประโยค ก่อนจะเริ่มใช้กระบวนการแยกประเภทตัวอักษรในขั้นตอนสุดท้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ประสิทธิภาพของระบบการรู้จำตัวอักษรบ่อยครั้งถูกแสดงเป็นเปอร์เซ็นต์ของอัตราการเรียนรู้ (Recognition rate) อัตราการรู้จำสำหรับข้อมูลที่ผู้ใช้ฝึกสอน จะแสดงให้เห็นข้อดีของแบบจำลองที่

นำมาประยุกต์ใช้ โดยทั่วไประบบส่วนใหญ่จะให้ผลของอัตราการเรียนรู้ใกล้เคียงกัน แต่ระบบที่ดีที่สุดควรจะถูกเขียนให้ใช้ฝึกรวมกับชุดข้อมูลจำนวนน้อย แต่ครอบคลุมแบบอักษรที่มีอยู่ให้มากที่สุด (แบบจำลองมีความยืดหยุ่นในการรู้จำมากกว่า)

5. แบบอักษร Arial เป็นแบบอักษรที่ไม่มีหัว เช่น I (ไอ) และ l (แอล) เป็นตัวอักษรที่ไม่มีความแตกต่างกันเลย จำเป็นต้องพิจารณาจากคำหรือประโยคเท่านั้น ทำให้มีปัญหาอย่างมากสำหรับการฝึกสอน เนื่องจากแพทเทิร์นมีลักษณะเหมือนกันแต่ต้องการผลลัพธ์เอาท์พุทสองแบบ ทำให้การฝึกสอนไม่ประสบผลสำเร็จ วิธีแก้ปัญหาคือแก้ไขแพทเทิร์นตัวอักษรไอให้มีหัวหรือไม่ก็ออกเลิกการจับกับแพทเทิร์นตัวอักษรไอลงในฐานข้อมูล

6. จากผลการทดสอบการเรียนรู้จำตัวอักษรพบว่า ประสิทธิภาพการเรียนรู้จำตัวอักษรไม่ได้ขึ้นอยู่กับขั้นตอนการแยกประเภทของตัวอักษรด้วยวิธีการของนิวโรลเน็ตเวิร์ค แต่ขึ้นอยู่กับขั้นตอนการประมวลผลขั้นต้นเป็นสำคัญ หากสามารถแยกตัวอักษรทุกตัวออกจากกันได้อย่างชัดเจน จะให้ผลการเรียนรู้จำตัวอักษรถูกต้องอยู่ในเกณฑ์สูงมาก

7. การทดสอบครั้งนี้ไม่มีการฝึกสอนตัวอักษรเอียง ทำให้ผลการเรียนรู้จำตัวอักษรเอียงมีความผิดพลาดสูง จึงควรหลีกเลี่ยงการทดสอบตัวอักษรเอียงดังกล่าว ส่วนตัวอักษรในภาพเอกสารที่อ่านเข้ามาบางตัวไม่เคยผ่านการฝึกสอนมาก่อนทำให้ผลการเรียนรู้จำตัวอักษรมีความผิดพลาดสูงเช่นกัน

8. การทดสอบครั้งนี้ยังไม่มีวิเคราะห์เพื่อแยกแยะภาพวาด และข้อความตัวอักษรออกจากกัน ดังนั้นการอ่านเอกสารที่ได้จากเครื่องสแกนเนอร์ควรเลือกเฉพาะภาพข้อความเท่านั้น

การทดสอบการเรียนรู้จำตัวอักษร โดยอาศัยความสามารถของนิวโรลเน็ตเวิร์คนับว่ามีประโยชน์มาก หากจะนำไปประยุกต์ใช้กับงานการจัดพิมพ์แบบตั้งโต๊ะด้วยคอมพิวเตอร์ ส่วนแนวความคิดนอกเหนือจากนั้นจะขออธิบายเพิ่มเติมในบทถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 7

### แนวความคิดของการประยุกต์ใช้ในทางปฏิบัติของการรู้จำตัวอักษร

การประยุกต์ใช้หลักการของนิวรอลเน็ตเวิร์กกับงานทางด้านความรู้จำตัวอักษรในบทที่ผ่านมา อาจจะมองไม่เห็นถึงประโยชน์ของการใช้งานมากนัก ผู้วิจัยจึงขอเสนอแนวความคิดเพิ่มเติม ซึ่งอาจจะเป็นประโยชน์แก่ผู้ที่สนใจต่อไปในอนาคต

แนวความคิดในการประยุกต์ใช้งานสำหรับนิวรอลเน็ตเวิร์กในด้านการรู้จำตัวอักษรมีปรากฏบ้างแล้ว อย่างเช่น การรู้จำตัวอักษรตัวเลขเพื่อใช้เป็นกรณีสำหรับอ้างอิงการจัดเก็บเอกสารแบบอัตโนมัติ การรู้จำตัวอักษรจากเอกสารเพื่อนำไปใช้กับโปรแกรมประเภท Word Processing และ Spreadsheet ส่วนในเรื่องของการประยุกต์ใช้กับการแปลงสัญญาณภาพตัวอักษรให้เป็นเสียงพูด และการย่อขนาดเพิ่มข้อมูลภาพตัวอักษรด้วยรูปแบบการจัดเก็บเอกสารแบบสื่อผสมข้อความ (Text) และภาพ (Image) ยังไม่เคยมีใครทำมาก่อน ซึ่งน่าจะเป็นแนวทางในการประยุกต์ใช้งานสำหรับผู้สนใจได้

#### 7.1 การรู้จำตัวอักษรตัวเลขเพื่อใช้เป็นกรณีสำหรับอ้างอิงการจัดเก็บเอกสารแบบอัตโนมัติ

ในปัจจุบันประโยชน์จากการรู้จำตัวอักษรที่เห็นชัดที่สุด น่าจะมาจากการประมวลผลภาพเอกสารจากแบบฟอร์ม เนื่องจากเห็นผลชัดเจน ช่วยลดเวลาสำหรับการกรอกข้อมูลจำนวนมากๆ อย่างเช่น ใบเสียภาษี ใบสั่งจราจร ใบกรมธรรม์ เป็นต้น

โดยส่วนมากแล้วแบบฟอร์มเอกสารต่างๆ จะกรอกข้อมูลเป็นตัวอักษรลายมือเขียน ดังนั้น การรู้จำตัวอักษรจึงเป็นเรื่องยากพอสมควร แต่หากเราจำกัดวงแคบลงเฉพาะการรู้จำตัวอักษรตัวเลขแบบลายมือพิมพ์ ประสิทธิภาพการรู้จำจะมีค่าสูงมากเกือบร้อยเปอร์เซ็นต์เลยทีเดียว เนื่องจากตัวอักษรตัวเลขมีเพียงสิบตัวเท่านั้น หากลายมือไม่แย่งเกินไปนัก ผลลัพธ์ที่ได้จะดีมาก

ข้อกำหนดสำหรับแบบฟอร์มเอกสารที่จะนำมาเข้ามารู้จำตัวอักษร ควรจะมีลักษณะดังนี้

1. ใช้กระดาษสีขาว เนื่องจากกระดาษสีจะทำให้เกิดสัญญาณรบกวนหลังจากอ่านเอกสารด้วยเครื่องสแกนเนอร์
2. ลายมือที่เขียนลงบนแบบฟอร์ม ควรเขียนด้วยตัวบรรจง และไม่ควรมีการขีดฆ่าด้วยปากกา อาจจะใช้ดินสอและยางลบเท่านั้น
3. ควรกรอกข้อมูลลงในกรอบสี่เหลี่ยมที่กำหนดให้เท่านั้น ดังแสดงในรูปที่ 84

## รูปที่ 84

แสดงช่องรูปสี่เหลี่ยมสำหรับการกรอกข้อมูล

เลขประจำตัวผู้เสียภาษี

--	--	--	--	--	--	--	--

เบอร์โทรศัพท์ที่ใช้ติดต่อ

--	--	--	--	--	--	--	--

4. กรอบตัวอักษรที่มีในแบบฟอร์มควรถูกลบออก (drop-out) ด้วยความสามารถของเครื่องสแกนเนอร์ ขณะอ่านแบบฟอร์มเอกสาร หรืออาจใช้ความสามารถในเชิงซอฟต์แวร์เพื่อกำจัดกรอบสี่เหลี่ยมก็เป็นได้
5. การกำหนดตำแหน่งในการรู้จำตัวอักษรตัวเลขจากกรอบสี่เหลี่ยมที่กำหนด ควรระบุตำแหน่งที่แน่นอนจากผู้ใช้งานเริ่มแรกก่อน เพื่อให้โปรแกรมเลือกเฉพาะส่วนที่ต้องการเท่านั้น

## แผนภูมิที่ 85

แสดงขั้นตอนการรู้จำตัวอักษรจากแบบฟอร์มเอกสาร

เริ่มต้น

อ่านเอกสารด้วยเครื่องสแกนเนอร์  
อาจมีการลบช่องสี่เหลี่ยมออกแล้ว

รู้จำตัวอักษรตัวเลขจากตำแหน่งอ้างอิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้นำไปเผยแพร่หรือแจกจ่ายเอกสารทุกครั้งที่มีการนำไปใช้

แผนภูมิที่ 85 (ต่อ)

หากตัวอักษรตัวใดถูกคัดออก ให้ผู้ใช้  
เป็นผู้ตัดสินใจกรอกข้อมูลด้วยตนเอง

จบ

แผนภูมิที่ 85 แสดงขั้นตอนการรู้จำตัวอักษรตัวเลขเพื่อใช้เป็นกรณีสำหรับอ้างอิงเอกสาร  
นั้นๆ การรู้จำตัวอักษรตัวเลขนั้นจำเป็นต้องรู้ตำแหน่งอ้างอิงสำหรับแบบฟอร์มด้วย เนื่องจากขั้นตอน  
การอ่านเอกสารด้วยเครื่องสแกนเนอร์ อาจมีการลบกรอบสี่เหลี่ยมออกไปแล้ว ซึ่งไม่สามารถตรวจ  
สอบตำแหน่งของตัวอักษรตัวเลขจากกรอบสี่เหลี่ยมได้เลย ดังนั้นตำแหน่งอ้างอิงอาจตรวจสอบจาก  
เครื่องหมายหรือสัญลักษณ์หน้าบรรทัดในเอกสาร เช่น หมายเลขบรรทัด เป็นต้น ดังแสดงในรูปที่ 86

รูปที่ 86

แสดงตำแหน่งอ้างอิงในแบบฟอร์มที่ใช้สำหรับการรู้จำตัวอักษร

00

วันที่   -   -

01

02 แบบฟอร์มการเสียภาษีประจำเดือน

มกราคม

03 ชื่อผู้เสียภาษี

นายสมชาย อภินวงศ์

04 หมายเลขประจำตัวผู้เสียภาษี

05

06 จำนวนภาษีที่ต้องชำระ

07

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านธุรกิจ  
ไม่ว่ากรณีใดๆ หากมีข้อสงสัยหรือต้องการข้อมูลเพิ่มเติม กรุณาติดต่อฝ่ายบริการลูกค้าของกรมสรรพากร  
ฐานข้อมูลที่เก็บเอาไว้ ดังนั้นหากต้องการขอคุณภาพเอกสารของนายสมชาย อภินวงศ์ จึงควรระบุเลข  
ประจำตัวผู้เสียภาษีดังกล่าวเพื่อขอคุณภาพตามต้องการ วิธีการนี้สามารถทำโดยอัตโนมัติเมื่อผู้ใช้งาน

ทำตามขั้นตอนที่ระบุไว้ และหากต้องการลดเวลาในการประมวลผลด้านการรู้จำตัวอักษร อาจใช้ฮาร์ดแวร์คอมพิวเตอร์ที่ออกแบบขึ้นเฉพาะ เพื่อทำหน้าที่เป็นทั้งเครื่องสแกนเนอร์และเครื่องมือรู้จำตัวอักษรภายในตัวเอง

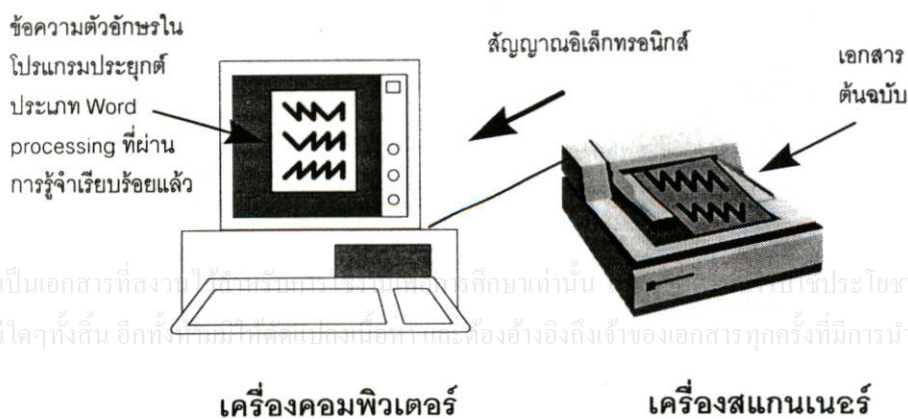
## 7.2 การรู้จำตัวอักษรจากเอกสารเพื่อนำไปใช้งานกับโปรแกรมประเภท Word processing และ Spreadsheet

ในปัจจุบันทุกคนต่างยอมรับความสามารถและความมั่งคั่งของไมโครซอฟท์วินโดวส์ซอฟต์แวร์ดังกล่าวก่อให้เกิดโปรแกรมประยุกต์ต่างๆ มากมาย ในจำนวนนี้มีโปรแกรมประเภท Word processing และ Spreadsheet น่าสนใจที่สุด ข้อมูลส่วนใหญ่ได้จากการพิมพ์คัดผ่านทางเป็นพิมพ์เมื่อข่าวสารข้อมูลมีจำนวนมากขึ้น การกรอกข้อมูลแบบเร่งด่วนด้วยมืออาจไม่เพียงพออีกต่อไป บริษัทและองค์กรต่างๆ จึงเริ่มหันมาสนใจเทคโนโลยีการรู้จำตัวอักษร ซึ่งในอนาคตน่าจะเป็นสิ่งจำเป็นอย่างมากก็เป็นได้

จากที่กล่าวมา เราสามารถรู้จำตัวอักษรจากเอกสารโดยทั่วไป ส่งผ่านผลลัพธ์ให้กับโปรแกรมประยุกต์อื่นๆ Word processing และ Spreadsheet เพื่อสะดวกในการจัดเก็บสำหรับงานพิมพ์เอกสารจำนวนมากได้ ดังแสดงในรูปที่ 87

รูปที่ 87

แสดงการนำเข้าข้อมูลและการเชื่อมต่อกับ โปรแกรมประยุกต์โดยทั่วไป



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์โดยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี เพื่อใช้ในการศึกษาเท่านั้น ไม่สามารถพิมพ์ใดๆ ทั้งสิ้น อีกทั้งยังห้ามมิให้คัดลอกหรือเผยแพร่โดยไม่ได้รับอนุญาตจากทางมหาวิทยาลัยฯ

แนวความคิดในการใช้งานอื่นๆ อาจทดลองใช้กับนามบัตรลูกค้าของบริษัท เนื่องจากบริษัทใหญ่ๆจะมีลูกค้าจำนวนมากและอาจมีการติดต่อกันเป็นประจำ การกรอกรายชื่อลูกค้าจำนวนมากอาจเป็นเรื่องยุ่งยาก ดังนั้นอุปกรณ์และซอฟต์แวร์รู้จำตัวอักษรดังกล่าว สามารถช่วยเหลืองานในด้านการถ่ายโอน (Transfer) ข้อมูลเข้าสู่คอมพิวเตอร์โดยอัตโนมัติ ช่วยลดอัตราการผลิตงานไปได้อีกทางหนึ่ง

### 7.3 การรู้จำตัวอักษรและการแปลความหมายเป็นเสียงพูด

โดยปกติแล้ว การรู้จำตัวอักษรจากภาพเอกสารและการแปลงข้อความให้เป็นเสียงพูด เป็นคนละเรื่องกัน ทั้งสองเรื่องนั้นเป็นส่วนที่สำคัญและใช้กระบวนการต่างกันไป แต่หากว่าเราสามารถผสมสื่อทั้งสองให้เป็นเรื่องเดียวกัน เราจะเห็นคุณประโยชน์ที่น่าับประการ นั่นหมายความว่าคอมพิวเตอร์สามารถอ่านเอกสารที่นำเข้ามาแล้วแปลงเป็นเสียงพูดได้ ในอนาคตเราอาจไม่จำเป็นต้องอ่านหนังสือพิมพ์หรือเอกสารด้วยตัวเองก็เป็นได้ ให้เครื่องคอมพิวเตอร์อ่านหนังสือแทนเรา และเป็นไปได้ว่า คอมพิวเตอร์สามารถค้นหาและคัดเฉพาะส่วนที่ต้องการฟังจากเอกสารภาพที่สนใจได้อีกด้วย

ในกรณีคนตาบอดจะให้ผลชัดเจนมาก เนื่องจากปัจจุบันการอ่านหนังสือของคนตาบอดจะต้องใช้มือสัมผัสกับตัวอักษรเบลล์เท่านั้น หรือไม่ก็จำเป็นต้องให้ผู้อื่นอ่านให้ฟัง แต่หากการรู้จำตัวอักษรและแปลความหมายเป็นเสียงพูดประสบความสำเร็จ ปัญหาของคนตาบอดในเรื่องการอ่านหนังสือคงจะหมดไป

เนื่องจากเทคโนโลยีปัจจุบัน การรู้จำตัวอักษรและการแปลงข้อความให้เป็นเสียงพูดเป็นเรื่องที่ยากอยู่แล้วในแต่ละหัวข้อ และได้ทำการวิจัยกันโดยผู้วิจัยคนละคน แต่หากว่าผู้วิจัยท่านใดสามารถนำหลักการและวิธีการของทั้งสองเรื่องมารวมไว้ด้วยกันและสามารถเชื่อมต่อกันได้เป็นอย่างดี อาจมีผลให้คอมพิวเตอร์อ่านหนังสือได้ และสามารถประมวลผลข้อความของตัวอักษรสำหรับงานทางด้านปัญญาประดิษฐ์ได้อีกด้วย (การมองเห็นและตีความหมายของหุ่นยนต์)

ขั้นตอนการประมวลผลอาจแยกหัวข้อออกเป็นสองส่วนแล้วนำมาประกบกันเป็นเรื่องเดียวกันก็ทำได้ หรืออาจวิจัยเพื่อให้เกิดวิธีการเดียวจากการแปลงตัวอักษรในภาพเอกสารให้เป็นเสียงโดยตรงได้ ดังแสดงในรูปที่ 88

## รูปที่ 88

แสดงขั้นตอนการประมวลผลภาพเอกสารให้เป็นเสียงพูด



a) รวมหัวข้อเรื่องการรู้จำตัวอักษรและแปลงข้อความเป็นเสียงพูดเข้าด้วยกัน



b) แปลงข้อความในภาพเอกสารให้เป็นเสียงพูดโดยตรง

ในขั้นตอนการรู้จำตัวอักษรที่ละตัวแล้วประกอบกันเป็นคำ อาจเป็นเรื่องง่ายสำหรับตัวอักษรภาษาอังกฤษ แต่สำหรับภาษาไทยคงเป็นการยาก เนื่องจากคำในภาษาไทยเขียนเรียงติดกันและไม่มีการเว้นช่องว่าง จึงจำเป็นต้องอาศัยวิธีการวิเคราะห์และแยกส่วนประกอบของคำออกมาจากประโยคเสียก่อน

จากที่กล่าวมา การแปลงเอกสารเป็นเสียงพูดต้องอาศัยความรู้พื้นฐานมากมายในหลายด้าน เช่น การรู้จำตัวอักษร การตรวจสอบคำและวิเคราะห์ความหมาย การเก็บฐานข้อมูลและการเรียกใช้ การแปลงข้อความให้เป็นเสียงพูด เป็นต้น ซึ่งน่าจะมีเทคโนโลยีใหม่ๆ เพื่อผสมผสานกันอย่างลงตัว เพื่อช่วยให้ระบบมีประสิทธิภาพมากยิ่งขึ้นและลดเวลาในการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆก็ตาม ผู้ใช้ซึ่งนำวิธีนี้ไปแปลงเพื่อนำไปใช้เองนั้นผู้ใช้นั้นต้องรับผิดชอบเอง

อุปกรณ์ที่ใช้สำหรับรู้จำภาพเอกสารและแปลงเป็นเสียงพูด ได้นั้นควรมีอุปกรณ์ดังนี้

1. เครื่องอ่านเอกสารที่เรียกว่า สแกนเนอร์แบบตั้งโต๊ะ (Flat-bed Scanner)
2. เครื่องคอมพิวเตอร์ส่วนบุคคล (Personal computer)
3. การ์ดเสียง (Sound card)
4. ลำโพง (Speaker)

อนาคตอันใกล้นี้เราอาจจะได้เห็นเทคโนโลยีดังกล่าว ช่วยพัฒนาคุณภาพชีวิตของสังคมปัจจุบัน ที่ต้องแข่งขันกับเวลาซึ่งมีอยู่จำกัดนั่นเอง

#### 7.4 การย่อขนาดเพิ่มข้อมูลภาพด้วยการจัดเก็บเอกสารแบบสื่อผสมทั้งข้อความและภาพ

เอกสารที่ได้จากการอ่านด้วยเครื่องสแกนเนอร์ที่เก็บในรูปแบบบิตแมพภาพจะใช้น้อยที่เก็บมากกว่าไฟล์ธรรมดา (เท็กซ์ไฟล์) หลายสิบเท่า ดังนั้นหากใช้วิธีการรู้จำตัวอักษรเพื่อแปลงภาพเอกสารให้เป็นตัวอักษรในลักษณะแอสกี จะช่วยให้การเก็บข้อมูลมีขนาดลดลงอย่างมาก โดยเฉลี่ยตัวอักษรใน 1 หน้ากระดาษ จะมีขนาดไฟล์ประมาณ 2,000 ถึง 3,000 ไบต์ ส่วนภาพเอกสารในลักษณะบิตแมพกราฟิก 1 หน้ากระดาษจะมีขนาดประมาณ 500,000 ถึง 1,000,000 ไบต์ (ยังไม่มีมีการบีบอัดข้อมูล)

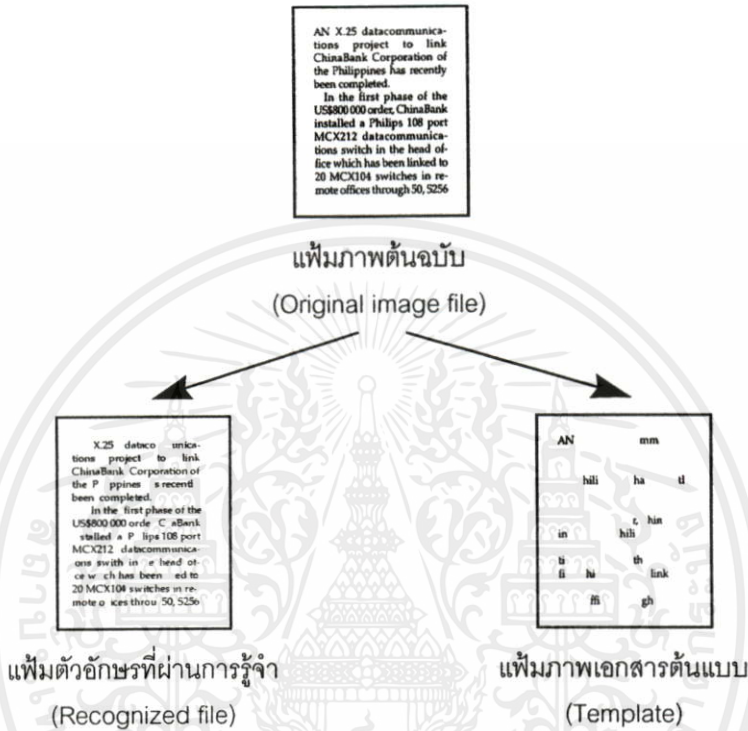
ในความเป็นจริงนั้น เราต้องยอมรับอย่างหนึ่งว่า การรู้จำตัวอักษรยังไม่มีประสิทธิภาพร้อยเปอร์เซ็นต์ มีตัวอักษรบางตัวที่เครื่องคอมพิวเตอร์ไม่สามารถตีความได้ และยังมีตัวอักษรพิเศษบางตัวที่โปรแกรมการรู้จำตัวอักษรในปัจจุบันไม่สามารถแยกประเภทออกได้ ตัวอย่างเช่น ตัวอักษรพิเศษทางคณิตศาสตร์ ฟิสิกส์ หรือวิชาแขนงอื่นๆ จึงเป็นการยากที่จะสร้างภาพเอกสารให้เป็นเท็กซ์ไฟล์โดยสมบูรณ์ และบางกรณีอาจมีรูปภาพวัตถุปะปนอยู่ด้วย

แนวความคิดจึงเกิดขึ้นว่า ทำไมเราไม่แยกเพิ่มข้อมูลทั้งสองออกจากกัน ข้อมูลภาพตัวอักษรที่สามารถแยกประเภทตัวอักษรได้จะจัดเก็บไว้ ที่เหลืออาจจะเป็นตัวอักษรพิเศษหรือภาพที่ไม่สามารถแยกประเภทตัวอักษรได้ให้คงอยู่เป็นต้นแบบ (Template) ดังแสดงในรูปที่ 89

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะณใดก็ตาม อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 89

แสดงการแยกประเภทตัวอักษรในภาพเอกสารออกมาเป็นอิสระ



แฟ้มตัวอักษรรหัสแอสกีจะเก็บตัวอักษรทั้งหมดที่ผ่านการรู้จำมีเรียบร้อยแล้ว พร้อมทั้งเก็บรวบรวมตำแหน่งกรอบตัวอักษรในหน้ากระดาษ ส่วนแฟ้มภาพเอกสารต้นแบบ (ยังคงเป็นภาพแบบบิตแมพกราฟิก) สามารถแสดงผลเป็นส่วนหลังของภาพเอกสารนั้นๆ ได้

หากเราต้องการนำกลับมาแสดงผลบนจอภาพให้เหมือนดังภาพต้นฉบับ เราสามารถทำได้ โดยการแปลงตัวอักษรรหัสแอสกีที่ได้จากการรู้จำให้อยู่ในรูปแบบบิตแมพเพทเทิร์น ณ ตำแหน่งบนจอภาพที่เก็บเอาไว้ โดยมีการขยายขนาดตัวอักษรให้ใกล้เคียงกับตัวอักษรเดิม ดังนั้นตัวอักษรที่แสดงผลบนจอภาพอาจจะดูแปลกตาไปบ้าง เนื่องจากใช้แบบอักษรแตกต่างออกไป แต่ผู้ใช้งานก็สามารถอ่านและเข้าใจความหมายได้เช่นเดิม ดังแสดงในรูปที่ 90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 90

## แสดงรูปแบบการแสดงผลบนจอภาพของสื่อผสมทั้งข้อความและภาพ

X.25 datacommunications project to link ChinaBank Corporation of the Philippines has recently been completed.  
In the first phase of the US\$600,000 order, ChinaBank installed a Philips 108 port MCX212 datacommunications switch in the head office which has been linked to 20 MCX104 switches in remote offices through 30,5256

AN X.25 datacommunications project to link ChinaBank Corporation of the Philippines has recently been completed.  
In the first phase of the US\$600,000 order, ChinaBank installed a Philips 108 port MCX212 datacommunications switch in the head office which has been linked to 20 MCX104 switches in remote offices through 30,5256

แฟ้มตัวอักษรที่ผ่านการรู้จำ  
(Recognized file)

แฟ้มภาพเอกสารต้นแบบ  
(Template)

AN X.25 datacommunications project to link ChinaBank Corporation of the Philippines has recently been completed.  
In the first phase of the US\$600,000 order, ChinaBank installed a Philips 108 port MCX212 datacommunications switch in the head office which has been linked to 20 MCX104 switches in remote offices through 30,5256

การแสดงผลบนจอภาพ  
(Representation image)

การย่อขนาดภาพสื่อผสมข้อความและภาพ จะช่วยให้เกิดการลดขนาดเพิ่มข้อมูลลงได้มากกว่าการย่อขนาดภาพต้นฉบับเพียงอย่างเดียว เป็นไปได้ว่าจะมีขนาดเพิ่มข้อมูลเหลือเพียง 10 เปอร์เซ็นต์ของการย่อขนาดภาพต้นฉบับ ทำให้แนวทางการจัดเก็บเอกสารมีทางเลือกเพิ่มมากขึ้น

แนวความคิดในการประยุกต์ใช้งานทั้งสี่หัวข้อที่ผ่านมา น่าจะเป็นแนวทางในการประยุกต์ใช้งานได้เป็นอย่างดี เนื่องจากสามารถมองเห็นภาพและขั้นตอนการทำงานที่เป็นประโยชน์และหวังเป็นอย่างยิ่งว่า จะมีผู้นำไปใช้ในทางปฏิบัติให้บังเกิดผลในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 8

### สรุปผลการวิจัยและข้อเสนอแนะ

นิรอลเน็ตเวิร์คเป็นสาขาวิชาที่น่าสนใจมากในปัจจุบัน มีการพัฒนาแนวความคิดเพื่อประยุกต์ใช้กับงานในหลายรูปแบบ โดยในเริ่มแรกจะศึกษาความเป็นมาและความหมายของนิรอลเน็ตเวิร์ค ศึกษาการทำงานและทดสอบด้วยตัวอย่างเล็กน้อย เพื่อให้เกิดความเข้าใจในพื้นฐานที่จำเป็นสำหรับการพัฒนาและประยุกต์ใช้งาน

ในภาคผนวก ก. จะแสดงรายละเอียดของแนวความคิดที่ได้จากวิทยานิพนธ์ฉบับนี้ไปประยุกต์ใช้กับการรู้จำตัวอักษรภาษาไทยในอนาคต

#### 8.1 สรุปผลการทดสอบโดยรวม

การทดสอบในบทแรกๆ จะจำลองปัญหาต่างๆ ของพีชคณิตบูลีนแบบแอนด์ เพื่อให้เกิดความเข้าใจการทำงานของโครงสร้างนิรอลเน็ตเวิร์คแบบเลเยอร์เดียว ช่วยทำให้การศึกษานิรอลเน็ตเวิร์คแบบหลายเลเยอร์ สามารถทำความเข้าใจได้ง่ายขึ้น

ต่อจากนั้นจะทดสอบการทำงานของปัญหาพีชคณิตบูลีนแบบเอ็กซ์คลูซีฟออร์ ด้วยหลักการของเบ็คหรือบพาทเกชันนิรอลเน็ตเวิร์ค และสามารถนำหลักการดังกล่าวไปประยุกต์ใช้กับงานทางการรู้จำตัวอักษรตัวเลข ตั้งแต่ 0 ถึง 9 ตามแบบจำลองที่สร้างขึ้นได้

หลังจากการทดสอบกับแบบจำลองการรู้จำตัวอักษรตัวเลขแล้ว เราจะทดสอบการรู้จำตัวอักษรในทางปฏิบัติ ด้วยการอ่านภาพเอกสารด้วยเครื่องสแกนเนอร์จากเอกสารโดยทั่วไป เพื่อทดสอบความสามารถของหลักการแยกประเภทตัวอักษรด้วยวิธีของนิรอลเน็ตเวิร์ค ในบทดังกล่าวไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงที่มาของเอกสารทุกครั้งที่มีการนำไปใช้ เราจะทราบว่า การแยกประเภทตัวอักษรด้วยวิธีของนิรอลเน็ตเวิร์คไม่สามารถดำเนินการโดยลำพังได้ จำเป็นต้องศึกษาหลักการประมวลผลขั้นต้น และการแยกคุณลักษณะเฉพาะของตัวอักษรเพิ่มเติม

เพื่อให้ฮัตราการรู้จำถูกต้อง จากการทดสอบผลการรู้จำตัวอักษรจากวารสารและหนังสือพิมพ์บางฉบับให้ฮัตราการรู้จำอยู่ในเกณฑ์ที่ดี แต่ยังคงพบกับปัญหาจุกจิกอีกมากในเรื่องของการประมวลผลขั้นต้น ซึ่งยังคงต้องแก้ไขต่อไปในเรื่องของตัวอักษรติดกันและยังคงมีบ้างในเรื่องของการจัดระดับตัวอักษร การวิจัยในลำดับต่อไปนอกจากต้องแก้ปัญหาดังกล่าวแล้ว ยังต้องเพิ่มเนื้อหาในเรื่องของการประมวลผลลำดับหลังอีกด้วย

จากหลักการที่ผ่านมาสามารถนำไปประยุกต์ใช้กับการรู้จำตัวอักษรภาษาไทยได้เป็นอย่างดี ไม่ว่าจะเป็นเรื่องของการประมวลผลขั้นต้น การหาคุณลักษณะของแพทเทิร์นตัวอักษร การแยกประเภทของตัวอักษร ดังนั้นหากผู้วิจัยท่านใดสนใจหัวข้อการของรู้จำตัวอักษร ก็สามารถนำความรู้ดังกล่าวไปพัฒนาและประยุกต์ใช้ได้ตามความต้องการ

## 8.2 ข้อเสนอแนะและผลการวิเคราะห์โดยรวม

ในปัจจุบัน ถึงแม้จะมีโปรแกรมการรู้จำตัวอักษรในท้องตลาดออกมาบ้างแล้ว แต่หนังสือที่ใช้ประกอบคำอธิบายและทฤษฎีประกอบยังหาอ่านได้ยากมาก จำเป็นต้องศึกษาจากบทความทางด้านการรู้จำแพทเทิร์น (Pattern recognition) จากบทความวิจัยของต่างประเทศหลายฉบับ แต่ละฉบับจะไม่กล่าวถึงรายละเอียดโดยรวมในหัวข้อการประมวลผลขั้นต้น การแยกคุณลักษณะเฉพาะ และการแยกประเภทของตัวอักษรในบทความวิจัยฉบับเดียวกัน เนื่องจากจำนวนหน้ากระดาษบทความวิจัยมีน้อยและหากต้องทำจริงจังหวังหัวข้อวิจัยนี้จะใหญ่มาก จึงจำเป็นต้องศึกษาเฉพาะด้านเท่านั้น

หลักการของการประมวลผลขั้นต้นค่อนข้างจะมีหลักการคงตัวไม่เปลี่ยนแปลงมากนัก มีเฉพาะเรื่องของการแยกส่วนประกอบหน้ากระดาษภาพที่หน้าหนักใจมาก เป็นการยากที่จะให้คอมพิวเตอร์วิเคราะห์รูปแบบการจัดวางตำแหน่งบนหน้ากระดาษได้อย่างสมบูรณ์แบบ เนื่องจากไม่มีหลักการที่แน่นอน

หลักการแยกคุณลักษณะเฉพาะของตัวอักษรนั้นมีผู้คิดค้นด้วยหลักการหลายรูปแบบ แต่ละส่วนจะไม่มี การแบ่งประเภทของทฤษฎีที่ชัดเจน หลักการที่นำไปใช้จึงจำเป็นต้องพิจารณาถึงความเหมาะสมด้วย เนื่องจากแต่ละหลักการจะมีจุดเด่นและจุดด้อยแตกต่างกันไป ทางที่ดีควรจะรวบรวมเอกสารนี้เป็นเอกสารมีสงวนไว้สำหรับกรใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ข้อดีของหลายวิธีไว้ใช้ด้วยกันจะดีมาก

หลักการของแยกประเภทของตัวอักษร แต่เดิมการรู้จำตัวอักษรจะใช้วิธีการเทียบความเหมือน ด้วยหลักการของ Pattern Matching แต่ปัจจุบันเริ่มอาศัยหลักการเรียนรู้เลียนแบบนิวรอลใน

สมองมนุษย์ เนื่องจากมีข้อดีหลายประการด้วยกัน จึงทำให้เกิดแนวความคิดในการประยุกต์ใช้กันอย่างกว้างขวางในปัจจุบัน

หากเทียบประสิทธิภาพการรู้จำตัวอักษรของวิทยานิพนธ์ ฉบับนี้กับซอฟต์แวร์ในท้องตลาด จะเห็นว่ายังเทียบกันไม่ได้ เนื่องจากการวิจัยยังไม่มีพัฒนางานในอีกหลายด้าน เช่น Page layout analysis, Quality control, Multicatagories, Neural Network Architecture, Merge character, Multifeature Extraction, Post processing และ Knowledge Learning เป็นต้น

\* การวิเคราะห์แผนผังหน้ากระดาษภาพ (Page layout analysis) เคยกล่าวถึงแล้วในหัวข้อการประมวลผลขั้นต้น แต่ไม่มีการลงมือทดสอบแต่อย่างใด เนื่องจากเป็นหัวข้อใหญ่พอควร และยังไม่มีความวิจัยฉบับไหนให้ความสำคัญมากนักในปัจจุบัน

\* การควบคุมคุณภาพ (Quality control) เป็นเรื่องของการกำหนดคุณภาพของสิ่งพิมพ์ในเอกสารที่นำเข้ามา เนื่องจากคุณภาพแบบอักษร คุณภาพกระดาษ และความไวในการประเมินผลของการรู้จำตัวอักษรเป็นสิ่งจำเป็น จำเป็นต้องอาศัยการตัดสินใจจากผู้ใช้เป็นสำคัญ เช่น คุณภาพแบบอักษรจากเครื่องพิมพ์แบบจุด จะมีลายเส้นขาดช่วง จำเป็นต้องแก้ไขโดยย่อขนาดภาพเอกสารให้มีขนาดเล็กลงแล้วใช้ค่าเทรคโฮลสำหรับไบনারีสูง เพื่อลดช่องว่างระหว่างจุดบนตัวอักษร ดังแสดงในรูปที่ 91

รูปที่ 91

แสดงภาพเอกสารจากเครื่องพิมพ์แบบจุด

## DOT MATRIX

ปัญหาคุณภาพกระดาษไม่ดี อาจทำให้เกิดสัญญาณรบกวนบนภาพเอกสารหลังจากการอ่านด้วยเครื่องสแกนเนอร์ได้ จึงควรใช้วิธีการกรองสัญญาณเพื่อกำจัดสิ่งแปลกปลอมออกไปก่อน

ความไวในการประเมินผลของการรู้จำตัวอักษร สามารถกำหนดได้สามระดับ Light, medium และ hard ระดับแรกกำหนดความไวในการประเมินผลแบบ light หมายถึงต้องการให้คัดตัวอักษรที่ไม่แน่ใจออกมามากขึ้น โดยการกำหนดให้ค่า  $\theta_{max} - \theta_{min}$  มีค่ามาก ๆ มีผลทำให้ตัวอักษรที่ผ่านการรู้จำจะมีความแม่นยำสูงขึ้น แต่จะเพิ่มจำนวนตัวอักษรที่ถูกคัดออกมากขึ้นด้วย ระดับที่สองความ

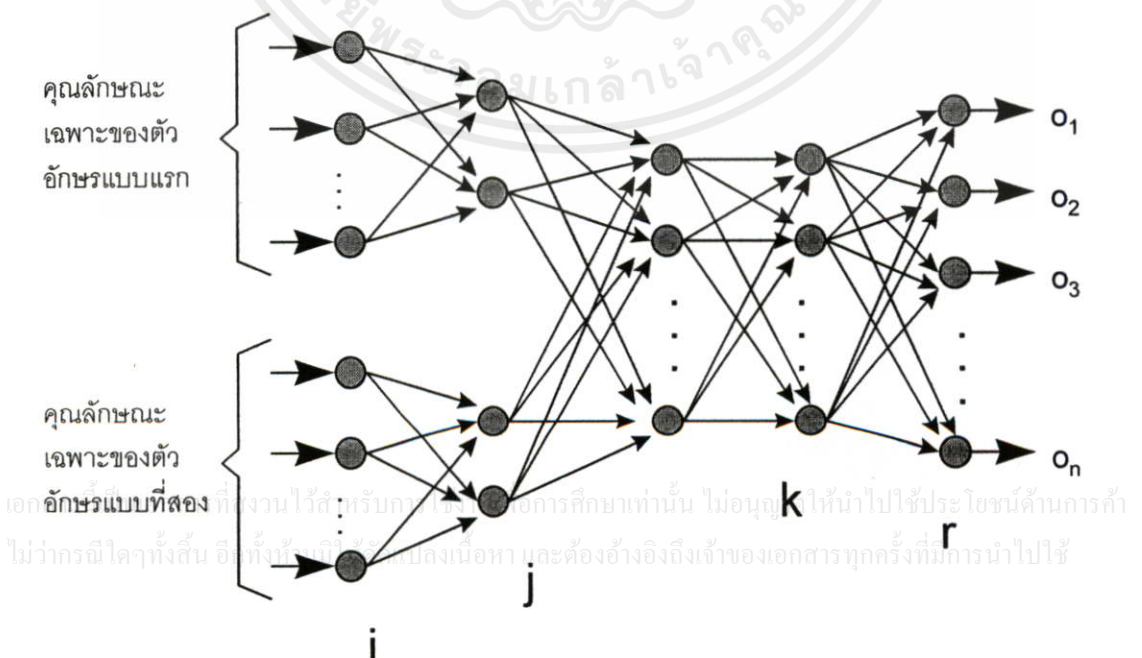
ไว้ในกรณีการประเมินผลแบบ median จะมีค่ากลางๆ โดยกำหนดให้  $\theta_{\max} - \theta_{\min}$  มีค่าพอประมาณ ระดับที่สามความไว้ในกรณีการประเมินผลแบบ hard หมายถึงต้องการให้ตัดตัวอักษรที่ไม่แน่ใจออกน้อยที่สุด โดยกำหนดให้ค่า  $\theta_{\max} - \theta_{\min}$  มีค่าน้อยๆ มีผลทำให้ตัวอักษรที่ผ่านการรู้จำจะมีความแม่นยำน้อยกว่าปกติ แต่จำนวนตัวอักษรที่ถูกตัดออกจะมีจำนวนลดลงด้วย

\* หลากหลายประเภท (Multicatagories) การแบ่งคลาสของตัวอักษรออกเป็นกลุ่มๆ ด้วยวิธีการแยกประเภทของตัวอักษรขั้นต้น เพื่อลดจำนวนตัวอักษรในการอบรมให้มีน้อยที่สุด มีผลทำให้ประสิทธิภาพการรู้จำดีขึ้น แต่ที่น่าเป็นห่วงคือ วิธีการดังกล่าวยังไม่มียุทธวิธีเป็นมาตรฐาน และหากการแยกประเภทกลุ่มตัวอักษรผิดตั้งแต่ต้น การรู้จำตัวอักษรในขั้นตอนการแยกประเภทจะให้ผลคลาดเคลื่อนไปด้วย

\* สถาปัตยกรรมของนิวรอลเน็ตเวิร์ค (Neural Network Architecture) อาจมีการตัดแปลงโครงสร้างการคอนเน็คชันเน็ตให้มีความเหมาะสม โดยกำหนดให้มีการแยกกลุ่มของคุณลักษณะหลายๆกลุ่ม ดังแสดงในรูปที่ 92

รูปที่ 92

แสดงโครงสร้างการเชื่อมต่อของนิวรอลเน็ตเวิร์คด้วยวิธีการแยกกลุ่มของคุณลักษณะออกจากกัน



- \* ตัวอักษรผสาน (Merge character) การแยกตัวอักษรที่ติดกันเป็นเรื่องที่ยากมาก หากไม่ใช้การพิจารณาด้วยดวงตาของมนุษย์ ดังนั้นการค้นหาแนวทางการแบ่งแยกส่วนประกอบของตัวอักษรจึงเป็นเรื่องสำคัญ ในหัวข้อวิทยานิพนธ์ฉบับนี้มีการกล่าวถึงหลักการดังกล่าวแล้ว
- \* การแยกคุณสมบัติตัวอักษรออกเป็นหลายชุด (Multifeature extraction) หมายถึง การเลือกคุณลักษณะของตัวอักษรในหลายรูปแบบ ใช้หลักการแตกต่างกันมาช่วยเสริมจุดเด่นของกันและกัน และลดจุดด้อยบางอย่างลง จะมีผลทำให้การรู้จำดีขึ้น เช่น วิธีผสมระหว่างโครงสร้างทางสถิติของรูปร่างโดยประมาณ (Shape approximation statistical structural) และระบบฐานความรู้จะช่วยทำให้ประสิทธิภาพการรู้จำดีขึ้น เนื่องจากอาศัยความสามารถจากหลายวิธีด้วยกัน
- \* การเพิ่มกระบวนการอธิบายความหมายหลังการแยกประเภทตัวอักษร (Post processing) ตัวอย่างเช่น โปรแกรมตรวจสอบการสะกดคำ หรือพจนานุกรม จะช่วยเพิ่มความแม่นยำในการการรู้จำตัวอักษรมากขึ้น
- \* ความรู้ที่ได้จากการเรียนรู้ (Knowledge Learning) หลังจากการรู้จำตัวอักษรเกิดขึ้น มักจะเกิดผลของความผิดพลาดเสมอ จำเป็นต้องให้ผู้ใช่แก้ไขด้วยมืออีกครั้งหนึ่ง จึงเป็นไปได้ว่า หากเราเก็บสถิติการรู้จำที่ผิดพลาดไว้เป็นฐานข้อมูลได้ เราจะสามารถวิเคราะห์ความผิดพลาดสำหรับตัวอักษรที่จะเกิดขึ้นในอนาคตได้ ซึ่งอาจจำเป็นต้องปรับเปลี่ยนค่าตัวแปรบางอย่าง เพื่อให้ผลการแยกประเภทตัวอักษรที่ผิดพลาดบ่อยๆ เกิดความแม่นยำขึ้นได้
- \* นิเวศน์เน็ตเวิร์คเบ็คพรีอบพาเกชันถูกนำมาใช้กับการรู้จำตัวอักษรอังกฤษ เฉพาะในขั้นตอนของการแยกประเภทตัวอักษรเท่านั้น รายละเอียดคุณได้จากภาคผนวก ก.

หลักการต่างๆ ที่กล่าวมาเป็นเพียงแนวความคิดในการวิเคราะห์ปัญหาของการรู้จำตัวอักษร ซึ่งยังคงต้องศึกษาต่อไป เพื่อเป็นแนวทางในการปฏิบัติต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บรรณานุกรม

- [1] Zurada, Jacek M. Introduction to Artificial Neural Systems, Singapore: West Info Access, 1992, pp. 1-250.
- [2] Frenzel, Louis E. Expert Systems and Applied Artificial Intelligence, USA: Macmillan, n.d, pp. 641-663.
- [3] Wasserman, Philip D. Neural Computing Theory and Practice, USA: ANZA Reserch, 1989, pp. 1-59.
- [4] Treleaven, Philip, Marco Pacheco and Marley Vellasco. "VLSI Architecture for Neural Networks", IEEE Micro, (December 1989) : 8-27.
- [5] Lippmann, Richard P. An Introduction to Computing with Neural Networks, USA: Addison-Wesley, 1989, pp. 4-22.
- [6] Pao, Yoh-Han. Adaptive Pattern Recognition and Neural Networks, USA: Addison-Wesley, 1989, pp. 113-140.
- [7] ประทีป บัญญัติสินพรรัตน์. ทฤษฎีและการใช้งานวงจรดิจิทัล เล่ม 1, กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2533, หน้า 83-123.
- [8] Rumelhart, D.E., J.L. McClelland and the PDP Research Group. Parallel Distributed Processing: Exploration in the Microstructure of Cognition, Volumn 1: Foundation, USA: The Massachusetts Institute of Technology, 1986, pp. 318-334.
- [9] เศรษฐพล ลิ่นปราชญา และกิตติ ไพฑูรย์วัฒนกิจ "การใช้นิวรอลเน็ตเวิร์กเพื่อจดจำแพทเทิร์น" วารสารสำนักงานคณะกรรมการวิจัยแห่งชาติ ปีที่ 25 เล่มที่ 1, (มกราคม-มิถุนายน 2536) :45-67.
- [10] Kohonen, T. Self-Organization and Associative Memory, USA: Spriner-Verlag, 1989, pp. 130-133.
- [11] Liang, Su., M. Shridhar and M. Ahmadi. "Segmentation of Touching Characters in Printed Document Recognition", Pattern Recognition, vol. 27, no. 6 (USA 1994) : 825-840.
- [12] Guyon, I., and others. "Design of Neural Network Character Recognition for a Touch Terminal", Pattern Recognition, vol. 24, no. 2 (1991) : 105-119.
- [13] Hildebrand, Thomas H. and Wentai Liu. "Optical Recognition of Handwritten Chinese Characters: Advance Since 1980", Pattern Recognition, vol. 26, no. 2 (1993) : 205-225.

- [14] Kulkarni, Arun D. Artificial Neural Networks for Image Understanding USA: Van Nostrand Reinhold, 1994, pp. 49-74.
- [15] Le, Daniel S., George R. Thoma and Harry Wechsler “Automated Page Orientation and Skew Angle Detection for Binary Document Images”, Pattern Recognition, vol. 27, (1994) : 1325-1344.
- [16] O’Gorman, Lawrence. “The Document Spectrum for Page Layout Analysis”, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 15, no. 11 (November 1993) : 1162-1173.
- [17] Tsujimoto, Shuichi. “Major Components of a Complete Text Reading System”, Processing of The IEEE, vol. 80, no. 7 (July 1992) : 1133-1149.
- [18] De Luca, P. G. and A. Gisotti. “Printed Character Preclassification Based on Word Structure”, Pattern Recognition, vol. 24, no. 7, (1991) : 609-615.
- [19] Wang, Jin. And Jack Jean. “Segmentation of Merged Characters by Neural Networks and Shortest Path”, Pattern Recognition, vol. 27, no. 5 (1994) : 649-658.
- [20] Lu, Yi. “Machine Printed Character Segmentation an Overview”, Pattern Recognition, vol. 28, no. 1 (1995) : 67-80.
- [21] Westall, J. M. and M. S. Narasimha. “Vertex Directed Segmentation of Handwritten Numerals”, Pattern Recognition, vol. 26, no. 10 (1993) : 1473-1486.
- [22] Fujisawa, Hiromichi., Yasuaki Nakano and Kiyomichi Kurino. “Segmentaion Methods for Character Recognition: Form Segmentation to Document Structure Analysis”, Processing of The IEEE, vol. 80, no. 7 (July 1992) : 1079-1092.
- [23] Seni, Giovanni. and Cohen, Edward. “External Word Segmentation of Off-Line Handwritten Text lines”, Pattern Recognition, vol. 27, no. 1 (1994), pp. 41-52.
- [24] ทวีศักดิ์ กอนันตกุล และคณะทำงานร่างข้อกำหนดร่วมเพื่อการเขียนโปรแกรมซึ่งแสดงผลเป็นภาษาไทย. คอมพิวเตอร์กับภาษาไทย: การพัฒนามาตรฐานเบื้องต้นสำหรับเทคโนโลยีสารสนเทศของไทย. กรุงเทพฯ: ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC), 2534.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ภาคผนวก ก.

### โปรแกรมตัวอย่างการทำงานของเพอร์เซพตรอนแบบเลเยอร์เดียว กับพีชคณิตบูลีนแบบแอนด์

```

/* ***** */
/* PROGRAM : SINGLE.C */
/* ISSUE DATED : 12 JAN, 1996 */
/* EXPIRE DATE : 14 FEB, 1996 */
/* BREIF : Test process Perceptron Learning Rule */
/*         (Decreate and Continueos activation function) */
/*         Boolean AND PROBLEM */
/* ***** */

#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <math.h>
#include <stdarg.h>

#define On      (1.0)          /* Binary ON */
#define Off     (0.0)          /* Binary OFF */
#define Threshold (0.5)      /* Threshold value */
#define Fixbias (On)          /* Bias value */
#define Df(OUT) (OUT*(1.0-OUT)) /* First order differentail */
#define Sigmoid(NET) (1.0/(1.0+exp(-(NET+0.5)))) /* Activation function */
#define HardLmd(NET) (((NET) < ((On+Off)/2.0)) ? Off : On)
#define Class   (4)           /* Class boolean AND */

/* ----- */
/*      Print text on screen (X,Y)      */
/* ----- */
void text (int X, int Y, unsigned char *format, ...)
{
    int      len;
    va_list  arg_ptr;
    unsigned char  output[200];

    va_start (arg_ptr, format);
    vsprintf (output, format, arg_ptr);
    va_end (arg_ptr);

    gotoxy (X, Y);
    printf ("%s", output);
}

```

```

/* ----- */
/*      Decreate Perceptron      */
/* ----- */
float Decreate (float desire, float actual)
{
    return (desire - actual);
}

/* ----- */
/*      Continueos Perceptron    */
/* ----- */
float continueos (float desire, float actual)
{
    return (desire - actual)*Df(actual);
}

/* ----- */
/*      Main Function            */
/* ----- */
void main (void)
{
    int i, loop;
    float net, out, w1, w2, w3, err, nel;
    float A [Class] = { Off, Off, On, On }; /* Input pattern A */
    float B [Class] = { Off, On, Off, On }; /* Input pattern B */
    float D [Class] = { Off, Off, Off, On }; /* Desired output */

    w1 = 0.2; w2 = 0.3; w3 = 0.4; /* Weight */
    nel = 0.8; /* Learning rate 0-1 */
    err = 1.0; /* Initail pattern error */
    loop = 0; /* Loop number (0 - 1000) */
    clrscr ();
    while ((err > 0.1) && (loop++ < 1000)) {
        err = 0.0;
        for (i=0; i < Class; i++) {
            net = w1*A[i] + w2*B[i] + w3*Fixbias; /* Feed forward */
            out = HardLmd(net); /* Feed forward */
            w1 = w1 + nel * Decreate(D[i], out) * A[i];
            w2 = w2 + nel * Decreate(D[i], out) * B[i];
            w3 = w3 + nel * Decreate(D[i], out) * Fixbias;
            err = (fabs(D[i] - out) > err) ? fabs(D[i] - out) : err;

            text (30, 10, "Loop (%3d)", loop);
            text (30, 11, "Step (%3d)", i+1);
            text (30, 12, "Learning rate = %5.2f", nel);
            text (30, 13, "Actually output = %5.2f", out);
            text (30, 14, "Weight #1 = %5.2f", w1);
            text (30, 15, "Weight #2 = %5.2f", w2);
            text (30, 16, "Weight #3 = %5.2f", w3);
            text (30, 17, "Pattern error = %5.3f", err);
        }
        text (30, 20, "End of Task ...");
    }
}

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่า

## ภาคผนวก ข.

### พิสูจน์ที่มาของสมการนิเวรอลเน็ตเวิร์คแบ็คพรีอพพากชัน

กำหนดให้ค่าผิดพลาดระหว่างค่าเอาต์พุตที่ได้จากการคำนวณ ( $o_{pk}$ ) และค่าเอาต์พุตที่ต้องการ ( $d_{pk}$ ) สำหรับแต่ละแพทเทิร์นเป็น  $E_p$  เมื่อ  $p$  แทนแพทเทิร์นใดๆ และ  $k$  แทนด้วยเลขอร์เอาต์พุตของนิเวรอลเน็ตเวิร์ค

$$E_p = \frac{1}{2} \cdot \sum_k (d_{pk} - o_{pk})^2 \dots\dots\dots (1)$$

$$E = \frac{1}{P} \cdot \sum_p E_p \dots\dots\dots (2)$$

ค่าผิดพลาดโดยรวมของระบบ  $E$  เกิดจากผลรวมของค่าผิดพลาดของแต่ละแพทเทิร์น  $E_p$

คำนวณการส่งค่าสัญญาณอินพุตนิเวรอลเน็ตเวิร์คในลักษณะป้อนหน้าจากเลขอร์  $i$  ไปเลขอร์  $j$

$$\text{net}_{pj} = \sum_i x_i \cdot w_{ji} + \theta_j \dots\dots\dots (3)$$

$$o_{pj} = f(\text{net}_{pj}) = \frac{1}{1 + e^{-(\text{net}_{pj})}} \dots\dots\dots (4)$$

คำนวณการส่งค่าสัญญาณอินพุตนิเวรอลเน็ตเวิร์คในลักษณะป้อนหน้าจากเลขอร์  $j$  ไปเลขอร์  $k$

$$\text{net}_{pk} = \sum_j x_j \cdot w_{kj} + \theta_k \dots\dots\dots (5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และตั้งอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$o_{pk} = f(\text{net}_{pk}) = \frac{1}{1 + e^{-(\text{net}_{pk})}} \dots\dots\dots (6)$$

การทำให้ค่าผิดพลาดของแต่ละแพทเทิร์น  $E_p$  มีค่าลดลงเข้าใกล้ศูนย์ ทำได้โดยลดผลต่างระหว่างค่าเอาต์พุตที่ได้จากการคำนวณ  $o_{pk}$  และค่าเอาต์พุตที่ต้องการ  $d_{pk}$  นั่นคือการปรับเปลี่ยนค่าตัวเลขน้ำหนัก  $w_{ji}$  และ  $w_{kj}$  ให้เกิดความเหมาะสม ด้วยวิธีการคำนวณของอนุพันธ์อันดับแรก  $E_p$  เทียบกับ  $w_{ji}$  ในลักษณะป้อนกลับ

$$\Delta_p w_{kj} \propto -\frac{\partial E_p}{\partial w_{kj}} \dots\dots\dots (7)$$

$$\Delta_p w_{kj} = -\eta \cdot \frac{\partial E_p}{\partial w_{kj}} \dots\dots\dots (8)$$

เมื่อ  $\eta$  เป็นค่าของอัตราการเรียนรู้ (Learning rate) ซึ่งมีค่าอยู่ระหว่าง 0.01 ถึง 1.00

ใช้ Chain rule เพื่อหาค่า  $\partial E_p / \partial w_{kj}$

$$\frac{\partial E_p}{\partial w_{kj}} = \frac{\partial E_p}{\partial \text{net}_{pk}} \cdot \frac{\partial \text{net}_{pk}}{\partial w_{kj}} \dots\dots\dots (9)$$

กำหนดให้

$$\delta_{pk} = \frac{\partial E_p}{\partial \text{net}_{pk}} = -\frac{\partial E_p}{\partial o_{pk}} \cdot \frac{\partial o_{pk}}{\partial \text{net}_{pk}} \dots\dots\dots (10)$$

$$\frac{\partial E_p}{\partial o_{pk}} = \frac{\partial}{\partial o_{pk}} \left[ \sum_k \frac{1}{2} (d_{pk} - o_{pk})^2 \right] = -(d_{pk} - o_{pk}) \dots\dots\dots (11)$$

และ 
$$\frac{\partial o_{pk}}{\partial \text{net}_{pk}} = f'_k(\text{net}_{pk}) \dots\dots\dots (12)$$

จะได้ว่า 
$$\delta_{pk} = f'_k(\text{net}_{pk}) \cdot (d_{pk} - o_{pk}) \dots\dots\dots (13)$$

จากสมการที่ (9) คำนวณหา  $\partial \text{net}_{pk} / \partial w_{kj}$  เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\frac{\partial \text{net}_{pk}}{\partial w_{kj}} = \frac{\partial}{\partial w_{kj}} (\sum_j w_{kj} o_{pj}) = o_{pj} \dots\dots\dots (14)$$

แทนค่าสมการที่ (13) และ (14) ลงในสมการที่ (9) จะได้ว่า

$$\frac{\partial E_p}{\partial w_{kj}} = -f'_k(\text{net}_{pk}) \cdot (d_{pk} - o_{pk}) \cdot o_{pj} = -\delta_{pk} o_{pj} \dots\dots\dots(15)$$

แทนค่าสมการที่ (15) ลงในสมการที่ (8) จะได้ว่า

$$\Delta_p w_{kj} = \eta \cdot f'_k(\text{net}_{pk}) \cdot (d_{pk} - o_{pk}) \cdot o_{pj} = \eta \cdot \delta_{pk} \cdot o_{pj} \dots\dots\dots(16)$$

หลังจากพิสูจน์สูตรการปรับค่าตัวเลขน้ำหนัก  $w_{kj}$  ได้แล้ว จะพิสูจน์สูตรการปรับค่าตัวเลขน้ำหนัก  $w_{ji}$  ในลักษณะเดียวกัน ด้วยการกำหนดความสัมพันธ์ของอนุพันธ์อันดับแรก  $E_p$  เทียบกับ  $w_{kj}$  ในลักษณะป้อนกลับดังนี้

$$\Delta_p w_{kj} = -\eta \cdot \frac{\partial E_p}{\partial w_{kj}} = \eta \cdot \delta_{pj} \cdot o_{pi} \dots\dots\dots(17)$$

ใช้ Chain rule เพื่อหาค่า  $\partial E_p / \partial w_{ji}$

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial \text{net}_{pj}} \cdot \frac{\partial \text{net}_{pj}}{\partial w_{ji}} \dots\dots\dots(18)$$

$$\text{ให้ } \delta_{pj} = -\frac{\partial E_p}{\partial \text{net}_{pj}} = -\frac{\partial}{\partial \text{net}_{pj}} \left[ \frac{1}{2} \sum_k (t_{pk} - o_{pk})^2 \right] \dots\dots\dots(19)$$

$$= -\frac{1}{2} \sum_k \left[ \frac{\partial}{\partial \text{net}_{pj}} (t_{pk} - o_{pk})^2 \right] \dots\dots\dots(20)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอก และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$= -\frac{1}{2} \sum_k \left[ 2(t_{pk} - o_{pk}) \frac{\partial}{\partial \text{net}_{pj}} (t_{pk} - o_{pk}) \right] \dots\dots\dots(21)$$

$$= \sum_k \left[ (t_{pk} - o_{pk}) \frac{\partial o_{pk}}{\partial net_{pj}} \right] \dots \dots \dots (22)$$

ใช้ Chain rule เพื่อหาค่า  $\partial o_{pk} / \partial net_{pj}$

$$\frac{\partial o_{pk}}{\partial net_{pj}} = \frac{\partial o_{pk}}{\partial net_{pk}} \frac{\partial net_{pk}}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial net_{pj}} \dots \dots \dots (23)$$

$$= f'_k(net_{pk}) \cdot w_{kj} \cdot f'_j(net_{pj}) \dots \dots \dots (24)$$

จะได้ว่า

$$\delta_{pj} = \sum_k \left[ (t_{pk} - o_{pk}) \left[ f'_k(net_{pk}) \cdot w_{kj} \cdot f'_j(net_{pj}) \right] \right] \dots \dots \dots (25)$$

$$= f'_j(net_{pj}) \sum_k \left[ (d_{pk} - o_{pk}) f'_k(net_{pk}) w_{kj} \right] \dots \dots \dots (26)$$

$$= f'_j(net_{pj}) \sum_k (\delta_{pk} w_{kj}) \dots \dots \dots (27)$$

จากสมการที่ (19) คำนวณหา  $\partial net_{pj} / \partial w_{ji}$

$$\frac{\partial net_{pj}}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \sum_i (w_{ji} o_{pi}) = o_{pi} \dots \dots \dots (28)$$

แทนค่าสมการที่ (27) และ (28) ลงในสมการที่ (17) และสมการที่ (18) จะได้ว่า

$$\Delta_p w_{ji} = \eta \cdot f'_j(net_{pj}) \sum_k (\delta_{pk} \cdot w_{kj}) \cdot o_{pi} \dots \dots \dots (29)$$

$$= \eta \cdot \delta_{pj} \cdot o_{pi} \dots \dots \dots (30)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

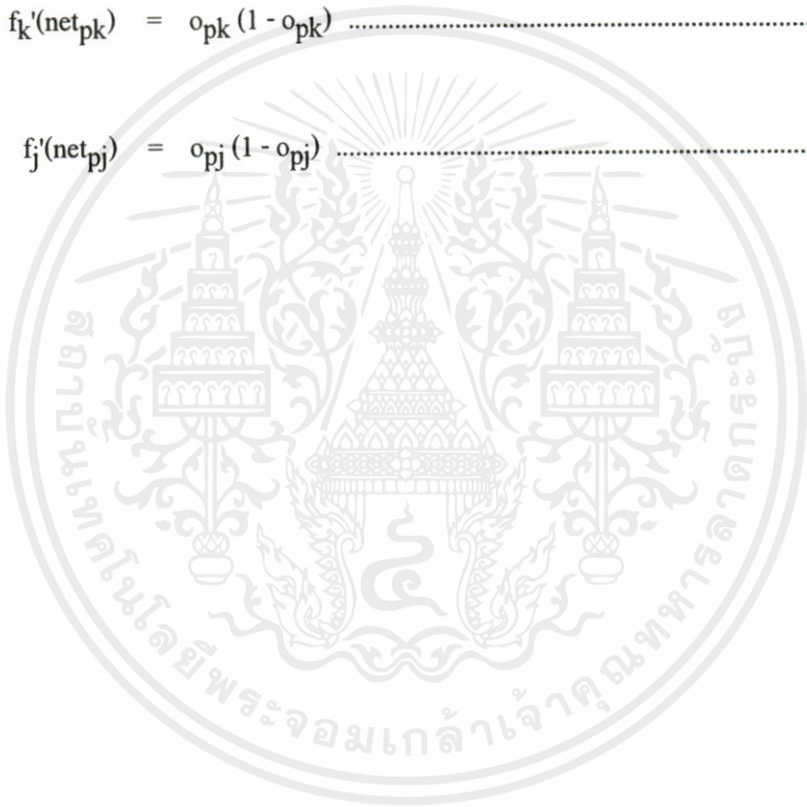
ความชันของค่าความผิดพลาด  $\partial E_p / \partial w_{kj}$  และ  $\partial E_p / \partial w_{ji}$  ที่ลดลงขึ้นกับการเปลี่ยนแปลงของสัญญาณค่าผิดพลาด  $\delta_{pk}$  และ  $\delta_{pj}$  ที่ลดลงตามลำดับ มีผลทำให้ค่าตัวเลขน้ำหนักที่เปลี่ยนไป

$\Delta p^{wkj}$  และ  $\Delta p^{wji}$  ลดลงด้วย วิธีการนี้จำเป็นต้องกระทำซ้ำไปซ้ำมาจนกระทั่งค่าผิดพลาดของระบบโดยรวม E มีค่าต่ำสุด หลักการนี้เรียกว่า วิธีการสอนของนิเวรอลเน็ตเวิร์ค (Training)

**หมายเหตุ**  $f'(net)$  จะมีค่าเปลี่ยนแปลงไปตามสมการแอกติเวชัน  $f(net)$  หากเลือกใช้ซิกมอยฟังก์ชัน จากสมการที่ (4) และ (6) และกำหนดให้  $\beta$  เท่ากับ 1.0 จะได้ว่า

$$f'_k(net_{pk}) = o_{pk} (1 - o_{pk}) \dots\dots\dots (31)$$

$$f'_j(net_{pj}) = o_{pj} (1 - o_{pj}) \dots\dots\dots (32)$$



## ภาคผนวก ก.

### อธิบายเพิ่มเติมในส่วนของงานวิทยานิพนธ์

#### 1. วิทยานิพนธ์ฉบับนี้ใช้นิวรอลเน็ตเวิร์คในส่วนไหนบ้าง

นิวรอลเน็ตเวิร์คมีโครงสร้างทางสถาปัตยกรรมที่พยายามเลียนแบบการทำงานของเซลล์ในสมองมนุษย์ โดยใช้วิธีการประมวลผลกระจายแบบขนาน (Parallel Distributed Processing) ทฤษฎีของนิวรอลเน็ตเวิร์คในปัจจุบันมีอยู่ด้วยกันหลายวิธี แต่วิทยานิพนธ์ฉบับนี้เลือกใช้เฉพาะทฤษฎีของนิวรอลเน็ตเวิร์คที่ใช้การเรียนรู้แบบคัพหรือพาทาเกชันในส่วนของงานของการแยกประเภทตัวอักษร (Classification) เนื่องจากนิวรอลเน็ตเวิร์คแบบคัพหรือพาทาเกชันมีคุณสมบัติเด่นในเรื่อง

1. คุณสมบัติทนต่อความผิดพลาด (Fault tolerance) หมายถึง ระบบประมวลผลจะไม่หยุดทำงานกลางคัน แม้ว่าหน่วยประมวลผลบางหน่วยจะได้รับความเสียหายจนไม่อาจทำงานได้
2. คุณสมบัติ Generalization หมายถึง นิวรอลเน็ตเวิร์คจะยังคงให้ผลลัพธ์ที่สมเหตุผล แม้ว่าจะมีสัญญาณรบกวน หรือข้อมูลขาดความสมบูรณ์ หรือข้อมูลไม่เคยเห็นมาก่อนก็ตาม
3. คุณสมบัติ Adaptability หมายถึง นิวรอลเน็ตเวิร์คสามารถเรียนรู้สภาพแวดล้อมที่เกิดขึ้นใหม่ได้เร็วขึ้น

#### 2. ทำไมต้องใช้ 2 คุณลักษณะเฉพาะสำหรับอินพุทของนิวรอลเน็ตเวิร์ค

เอกสารนี้เป็นเหตุผลที่คุณลักษณะเฉพาะจุดภาพ (Pixels feature) ที่ใช้เป็นอินพุทของนิวรอลเน็ตเวิร์คแบบคัพหรือพาทาเกชันเพียงอย่างเดียวจะให้ผลการรู้จำตัวอักษรเพียง 80-90 เปอร์เซ็นต์เท่านั้น คุราบายละเอียดได้จากสรุปผลงานวิจัย [13] พบว่าการใช้เพียงคุณสมบัติเฉพาะจุดภาพจะให้ผลการรู้จำตัวอักษรถูกต้องน้อยกว่าคุณสมบัติเฉพาะแบบอื่น

เหตุผลอีกประการที่ไม่เลือกใช้คุณลักษณะเฉพาะแบบ Curve approximation [12] เพียงอย่างเดียว เนื่องจากผลการวิจัยพบว่า การฝึกสอนไม่สามารถเข้าสู่ค่าตอบได้ เพราะค่าผิดพลาดที่คำนวณได้สูงกว่าค่าผิดพลาดที่ยอมรับได้ (Nonconverge) เนื่องจากตัวอักษรที่นำเข้ามาใช้ฝึกสอน บางตัวมีความเหมือนกันมากเช่น 0 (ศูนย์) และ O (โอ) หรือ 1 (หนึ่ง), l (ไอ) และ 1 (แอล) เป็นต้น มีผลทำให้คุณสมบัติเฉพาะของตัวอักษรมีค่าซ้ำกัน

การรวมคุณสมบัติเฉพาะของตัวอักษรทั้ง 2 ประเภทเข้าด้วยกันมีผลทำให้การฝึกสอนได้ค่าต่ำกว่าค่าผิดพลาดที่ยอมรับได้ (Converge) เนื่องจากคุณสมบัติเฉพาะของตัวอักษรทั้งสองมีจุดเด่นแตกต่างกัน

### 3. แนวทางการนำผลงานวิจัยฉบับนี้ไปประยุกต์ใช้กับการรู้จำตัวอักษรภาษาไทย

ข้อแตกต่างระหว่างภาษาอังกฤษกับภาษาไทยต่างกันตรงที่ว่า ภาษาไทยแบ่งตัวอักษรออกเป็น 4 ระดับ ระดับวรรณยุกต์ ระดับสระเหนือบรรทัด ระดับพยัญชนะกลางบรรทัด และระดับสระใต้บรรทัด [24] แต่ตัวอักษรภาษาอังกฤษทุกตัวจะวางอยู่ในระดับเดียวกัน ดังนั้นวิธีแก้ปัญหาสำหรับตัวอักษรภาษาไทยจำเป็นต้องแยกตัวอักษรภาษาไทยทุกตัวออกจากกันดังแสดงในรูปที่ 93(b) ด้วยกระบวนการตัดแยกตัวอักษร

รูปที่ 93

แสดงการจัดเรียงตัวอักษรภาษาไทยในระดับบรรทัด

	1	Top
ั	2	Above
บ ฝ	3	Base
อ	4	Below

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ a) ตัวอักษรภาษาไทยในรูปแบบปกติ ให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## รูปที่ 93 (ต่อ)

	'	1	Top
	'	2	Above
ป	พ	3	Base
ข		4	Below

## b) การตัดแยกตัวอักษรภาษาไทยในระดับบรรทัด

รูปที่ 93 แสดงการตัดแยกตัวอักษรภาษาไทยในระดับบรรทัด ซึ่งเป็นส่วนหนึ่งของการประมวลผลขั้นต้น (Preprocessing) ในขั้นตอนการรู้จำตัวอักษร หลังจากนั้นสามารถนำรอบตัวอักษรทุกตัวมาผ่านกระบวนการแยกคุณลักษณะเฉพาะของตัวอักษรต่อไป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ประวัติผู้เขียน

ชื่อผู้เขียน	นาย เศรษฐพล ลีนปราชญา
วันเดือนปีเกิด	วันที่ 28 ตุลาคม 2510
สถานที่เกิด	จังหวัดกรุงเทพมหานคร
วุฒิการศึกษาระดับปริญญาตรี	วิทยาศาสตร์บัณฑิต สาขาฟิสิกส์
สถานที่สำเร็จการศึกษา	มหาวิทยาลัยเกษตรศาสตร์
ปีที่สำเร็จการศึกษา	ปีการศึกษา 2532
ผลงานทางวิชาการที่ได้รับการตีพิมพ์	“การใช้วีรอลเนตเวิร์กเพื่อจดจำแพตเทิร์น” จากวารสารสำนักงานคณะกรรมการวิจัยแห่งชาติ ปีที่ 25 เล่มที่ 1 มกราคม - มิถุนายน 2536
ประสบการณ์การทำงาน	ตำแหน่ง Sale Engineer ของบริษัทยูนิคซอฟต์แวร์อินเตอร์ เนชั่นแนล จำกัด ในเครือบริษัทค้าค้าแม่ท จำกัด
อาชีพหรือการศึกษาในปัจจุบัน	นักศึกษาปริญญาโทสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณ ทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้