การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้หน่วยความจำระยะสั้นแบบยาว

นายทวีศักดิ์ เอี่ยมสวัสดิ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thai Printed Character Recognition using Long Short-Term Memory

Mr. Taweesak Emsawas

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computer Science
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2016

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้หน่วยความจำระยะสั้นแบบยาว
โดย	นายทวีศักดิ์ เอี่ยมสวัสดิ์
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร. บุญเสริม กิจศิริกุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์

(รองศาสตราจารย์ ดร. สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. สุกรี สินธุภิญโญ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ศาสตราจารย์ ดร. บุญเสริม กิจศิริกุล)

.....กรรมการ

(อาจารย์ ดร. พีรพล เวทีกูล)

.....กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร. ชลวิช นัทธี)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ทวีศักดิ์ เอี่ยมสวัสดิ์ : การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้หน่วยความจำระยะสั้นแบบยาว (Thai Printed Character Recognition using Long Short-Term Memory) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ศ. ดร. บุญเสริม กิจศิริกุล, 60 หน้า.

วิธีแบ่งส่วนสำหรับการรู้จำตัวอักษรทำงานโดยการแบ่งภาพบรรทัดตัวอักษรเป็นภาพตัวอักษรและนำไปรู้จำตัวอักษรแต่ละตัวอักษร วิธีนี้ได้รับผลกระทบจากประสิทธิภาพของกระบวนการแบ่งส่วนในปัญหาตัวอักษรที่เชื่อมติดกันหรือตัวอักษรที่บางส่วนขาดหายอย่างมาก ในขณะที่วิธีไม่แบ่งส่วนจะทำการรู้จำภาพบรรทัดตัวอักษรโดยไม่แบ่งส่วนภาพตัวอักษรแต่ละตัว วิธีนี้เหมาะสมกับภาษาอย่างเช่นภาษาไทยที่ประกอบด้วยตัวอักษรที่เชื่อมติดกันจำนวนมาก เป้าหมายของวิทยานิพนธ์นี้คือการประยุกต์ใช้หน่วยความจำระยะสั้นแบบยาว ซึ่งเป็นวิธีไม่แบ่งส่วนในการรู้จำตัวอักษรภาษาไทย นอกจากนี้วิทยานิพนธ์นำเสนอ*วิธีการเลื่อนองค์ประกอบแนวตั้ง* ในการแก้ไขปัญหารูปแบบการรวมกันของตัวอักษรที่เกิดขึ้นแนวตั้งจำนวนมากบนโครงสร้างตัวอักษรสี่ระดับของภาษาไทย และยากต่อการนำมาใช้กับโครงข่ายหน่วยความจำระยะสั้นแบบยาวมาตรฐาน ผลการทดลองแสดงค่าความแม่นยำเปรียบเทียบวิธีที่นำเสนอบนโครงข่ายหน่วยความจำระยะสั้นแบบยาวมาตรฐาน กับซอฟต์แวร์เชิงพาณิชย์ในการรู้จำตัวอักษรภาษาไทย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต .....

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

ปีการศึกษา 2559

5770420421 : MAJOR COMPUTER SCIENCE
KEYWORDS: THAI PRINTED CHARACTER RECOGNITION / RECURRENT NEURAL NETWORK / LONG SHORT-TERM MEMORY / VERTICAL COMPONENT SHIFTING

TAWEESAK EMSAWAS: Thai Printed Character Recognition using Long Short-Term Memory. ADVISOR: PROF. BOONSERM KIJSIRIKUL, Ph.D., 60 pp.

The segmentation-based approach for Optical Character Recognition (OCR) works by first segmenting a text line image into individual character images and then recognizing the characters. The approach relies heavily on the performance of the segmentation process and thus suffers from the problem of touching and broken characters. On the other hand, the unsegmented approach for OCR processes the text line image without segmenting the image into individual characters, and the approach is more suitable for languages such as Thai that contains a lot of touching characters in nature. This thesis proposes an application of Long Short-Term Memory (LSTM), which is an unsegmented method, to Thai OCR. The thesis also introduces a method called *vertical component shifting* to solve the problem of a large number of vertically occurring character combinations that occur in four-level writing system of Thai, and pose difficulty for standard LSTM networks. The experimental results demonstrate the better accuracy of our proposed method over standard LSTM networks and other commercial software for Thai OCR.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Department: Computer Engineering
Field of Study: Computer Science
Academic Year: 2016

Student's Signature .....
Advisor's Signature .....

กิตติกรรมประกาศ

ขอกราบขอบคุณ ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาที่ให้ได้ให้ความรู้ คำปรึกษา คำแนะนำ แนวคิดในการจัดทำวิทยานิพนธ์ ตลอดจนแก้ไขข้อบกพร่องในการทำวิทยานิพนธ์ตั้งแต่เริ่มต้นจนสำเร็จเสร็จสิ้น

ขอขอบคุณ รองศาสตราจารย์ ดร.ชลวิช นัทธี ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ และอาจารย์ ดร.พีรพล เวทีกูล ที่สละเวลามาเป็นคณะกรรมการสอบวิทยานิพนธ์ ให้คำแนะนำและแนวทางในการทำวิทยานิพนธ์ให้สำเร็จลุล่วงไปด้วยดี

ขอขอบคุณอาจารย์สุภาวดี ศรีคำดี อาจารย์สุปัญญา อภิวงศ์โสภณ ผู้ช่วยศาสตราจารย์ ดร.สุนิสา ริมเจริญ ผู้ช่วยศาสตราจารย์ ดร.อุรีรัฐ สุขสวัสดิ์ชน ผู้ช่วยศาสตราจารย์ ดร.จักริน สุขสวัสดิ์ชน ที่ให้คำแนะนำ คำปรึกษา รวมไปถึงช่วยเหลือในการจัดทำวิทยานิพนธ์ การสอบวิทยานิพนธ์ และเป็นกำลังใจที่ดีเสมอมา

ขอขอบคุณสมาชิกห้องปฏิบัติการอัจฉริยภาพเครื่องจักรและการค้นพบความรู้ (MIND Lab) ทุกท่านที่ให้คำแนะนำและคำปรึกษาในการทำวิทยานิพนธ์

สุดท้ายนี้ขอกราบขอบคุณคุณพ่อรักศักดิ์ เอี่ยมสวัสดิ์ คุณแม่กมนรัตน์ เอี่ยมสวัสดิ์ และพี่ชายอาทิตย์ เอี่ยมสวัสดิ์ที่คอยสนับสนุนทั้งในเรื่องของค่าใช้จ่าย และเป็นกำลังใจที่ทำให้ประสบความสำเร็จลุล่วง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ญ

สารบัญรูปภาพ ..... ฎ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตการดำเนินงาน 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 2

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 3

2.1 ตัวอักษรภาษาไทย 3

2.2 องค์ประกอบที่เชื่อมติดกัน (Connected Component) 4

2.3 โครงข่ายประสาทเทียมแบบหมุนเวียนกลับ (Recurrent Neural Network - RNN) 4

2.4 หน่วยความจำระยะสั้นแบบยาว (Long Short-term Memory - LSTM) 6

2.5 โครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง (Bidirectional LSTM network) 9

2.6 การคำนวณค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์น (Levenshtein distance) ... 10

บทที่ 3 งานวิจัยที่เกี่ยวข้อง 11

3.1 งานวิจัยการรู้จำตัวแบบแบ่งส่วน และงานวิจัยการรู้จำตัวอักษรภาษาไทย 11

3.2 งานวิจัยการรู้จำตัวอักษรแบบไม่แบ่งส่วน 12

3.3 โอซีอาร์โอปุส (OCRopus) 13

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หน้า

บทที่ 4 แนวทางและวิธีดำเนินงาน 16
4.1 การเตรียมชุดข้อมูล 16
4.2 การเลื่อนองค์ประกอบแนวตั้ง (Vertical Component Shifing) 19
4.3 โครงข่ายแอลเอสทีเอ็ม 21
4.4 การทดลอง 22
4.4.1 โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง 22
4.4.2 โครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง 23
4.4.3 โครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลขนาดใหญ่ 24
4.4.4 วัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ที่ได้รับความนิยม 25
บทที่ 5 ผลการดำเนินงาน 26
5.1 การเตรียมชุดข้อมูล 26
5.2 การเลื่อนองค์ประกอบแนวตั้ง 26
5.3 โครงข่ายแอลเอสทีเอ็ม 27
5.4 ผลการทดลองและการวิเคราะห์ผลการทดลอง 30
5.4.1 โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง 30
5.4.2 โครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง 33
5.4.3 โครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลขนาดใหญ่ 36

หน้า

5.4.4 วัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ที่ได้รับความนิยม 39

บทที่ 6 สรุปผลการดำเนินงาน 43

รายการอ้างอิง 44

ภาคผนวก 46

ประวัติผู้เขียนวิทยานิพนธ์ 60

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

ตารางที่ 2-1 ตัวอักษรภาษาไทย ตัวเลข และเครื่องหมายที่ใช้ในชุดข้อมูลตามรหัสยูนิโคด 3

ตารางที่ 3-1 เปอร์เซนต์ค่าความผิดพลาด OCRopus บนภาษาอังกฤษและภาษาฟรักทูร์ 15

ตารางที่ 4-1 ฟอนต์ชุดข้อมูล 17

ตารางที่ 4-2 จำนวนชุดข้อมูลที่ใช้ในการทดลองการตรวจสอบไขว้ 5 พับ 23

ตารางที่ 4-3 จำนวนชุดข้อมูลฝึก ชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบในแต่ละรอบการทดลอง 24

ตารางที่ 4-4 จำนวนชุดข้อมูลฝึก ชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบ 25

ตารางที่ 5-1 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว 31

ตารางที่ 5-2 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง 32

ตารางที่ 5-3 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง 34

ตารางที่ 5-4 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย 35

ตารางที่ 5-5 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลฝึกขนาดใหญ่ 37

ตารางที่ 5-6 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยบนชุดข้อมูลฝึกขนาดใหญ่ 38

ตารางที่ 5-7 ผลการทดลองแสดงค่าความผิดพลาดของโครงข่ายแอลเอสทีเอ็มใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย ABBYY และ ArnThai แสดงจำนวนตัวอักษรที่ผิดพลาด (#error) และเปอร์เซ็นต์ค่าความผิดพลาด (%error) บนชุดข้อมูลทดสอบ 10 ฟอนต์ 40

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญรูปภาพ

รูปที่ 2-1 คำในภาษาไทย “อรุณสวัสดิ์” ที่มีครบทุกระดับตัวอักษร 4

รูปที่ 2-2 (ก) การเชื่อมติด 4 ทิศทาง และ (ข) การเชื่อมติด 8 ทิศทาง 4

รูปที่ 2-3 (ก) โครงข่ายประสาทเทียมแบบหมุนเวียนกลับอย่างง่ายที่ประกอบไปด้วยชั้นนำเข้า ชั้นซ่อน และชั้นนำออก อย่างละ 1 บัพ และ (ข) ขยายโครงข่ายประสาทเทียมแบบหมุนเวียนกลับอย่างง่ายตามเวลา 5

รูปที่ 2-4 ปัญหาวานิชชิงเกรเดียนต์ในโครงข่ายประสาทเทียมแบบหมุนเวียนกลับ ความเข้มของสีแสดงถึงการสูญหายไปของสารสนเทศข้อมูลนำเข้าลำดับที่ 1 ตามเวลา..... 5

รูปที่ 2-5 บล็อกหน่วยความจำของแอลเอสทีเอ็ม 1 เซลล์หน่วยความจำ ประตูนำเข้า ประตูนำออก และประตูลืม 6

รูปที่ 2-6 โครงข่ายแอลเอสทีเอ็ม ประกอบด้วยชั้นนำเข้าจำนวน 4 บัพ ชั้นนำออกจำนวน 5 บัพ และชั้นซ่อนจำนวน 2 บัพหรือ 2 บล็อกหน่วยความจำ แต่ละบล็อกหน่วยความจำจะมีข้อมูลนำเข้าได้ 4 ทางและมีเพียง 1 ข้อมูลนำออก เส้นเชื่อมที่แสดงค่าถ่วงน้ำหนักเป็นเพียงส่วนหนึ่งเท่านั้น ค่าถ่วงน้ำหนักระหว่างบัพในชั้นข้อมูลนำเข้ากับบล็อกหน่วยความจำแสดงด้วยเส้นเชื่อมสีฟ้า ค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำเดียวกันแสดงด้วยเส้นเชื่อมสีแดง ค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำแสดงด้วยเส้นเชื่อมสีส้ม และค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำไปยังบัพในชั้นนำออกแสดงด้วยเส้นเชื่อมสีเขียว..... 7

รูปที่ 2-7 การแก้ปัญหาปัญหาวานิชชิงเกรเดียนต์ของโครงข่ายประสาทเทียมแบบหมุนเวียนกลับโดยการใช้บล็อกหน่วยความจำ (แทนด้วยสี่เหลี่ยมจัตุรัส) มารักษาสารสนเทศเพื่อส่งไปยังสถานะถัดไป รูปทรงที่มีสีดำคือได้รับสารสนเทศจากเวลาที่ 1 และสีขาวคือไม่เกี่ยวข้องกับสารสนเทศจากเวลาที่ 1 วงกลมขนาดเล็ก (**o**) และเส้นตรง (**-**) แทนประตูเปิด และประตูปิดตามลำดับ ซึ่งจะถูกวางทางขวาล่าง ขวาบน และซ้ายของบล็อกหน่วยความจำ คือ ประตูเข้า ประตูนำออก และประตูลืม ตามลำดับ สารสนเทศจากเวลาที่ 1 สามารถเดินทางไปยังเวลาที่ 6 ได้โดยประตูลืมของบล็อกหน่วยความจำเวลาที่ 2 ถึง 6 ที่เปิดรับสารสนเทศจากเวลาที่ 1 และประตูนำเข้าเปิดเพียงช่วงเวลาที่ 1 เท่านั้นจึงไม่มีการรับข้อมูลนำเข้าจากเวลาอื่นจนกระทั่งเวลาที่ 7 และประตูนำออกที่เปิดเพียงเวลาที่ 4 และ 6 ทำให้สารสนเทศของเวลาที่ 1 สามารถเดินทางไปยังเวลาที่ 4 และ 6 โดยไม่สูญหายไปตามเวลา 8

รูปที่ 2-8 โครงข่ายแอลเอสทีเอ็มแบบสองทิศทางที่ถูกนำมาขยายตามเวลา ประกอบด้วยบัพชั้นข้อมูลนำเข้าและชั้นข้อมูลนำออกอย่างละ 1 บัพ และ 2 บล็อกหน่วยความจำ (แสดงด้วยรูปสี่เหลี่ยม) ในชั้นซ่อน การเลื่อนหน้าต่าง หน้าต่าง (แสดงด้วยรูปสี่เหลี่ยมผืนผ้าสีเทา) มีความกว้าง 1 พิกเซล นำมาสร้างเป็นเวกเตอร์ของข้อมูลนำเข้าจากเวลาที่ 1 ถึง T และจาก T ถึง 1 เมื่อ T คือความกว้างของภาพบรรทัดตัวอักษร..... 10

รูปที่ 3-1 การทำงานของ OCRopus ในการทำนายภาพโดยการแปลงภาพเป็นขาวดำ 13

รูปที่ 3-2 โครงสร้างของแอลเอสทีเอ็มแบบ 2 ทิศทางที่ประกอบไปด้วยค่าถ่วงน้ำหนักจากชั้นนำเข้าไปยังชั้นซ่อน 12,800 ค่า จากชั้นซ่อนไปยังชั้นนำออก 15,900 ค่า และค่าถ่วงน้ำหนักวงวนเชื่อมต่อกับตนเองจำนวน 40,000 ค่า 14

รูปที่ 4-1 ขั้นตอนวิธีการเลื่อนองค์ประกอบแนวตั้ง..... 20

รูปที่ 4-2 (ก) ภาพบรรทัดตัวอักษรก่อนเลื่อนองค์ประกอบเชื่อมติดกัน..... 21

รูปที่ 4-3 โครงข่ายแอลเอสทีเอ็ม ประกอบไปด้วยชั้นนำเข้า 48 บัพ ชั้นนำออก 131 บัพ 22

รูปที่ 4-4 การทดลอง 5 รอบ ในแต่ละรอบ กล่องสีดำแสดงข้อมูลฝึก..... 24

รูปที่ 5-1 การแปลงภาพข้อมูลนำเข้าหน้ากระดาษตัวอักษรเป็นภาพบรรทัดตัวอักษร..... 26

รูปที่ 5-2 ภาพบรรทัดตัวอักษร (ก.1) (ข.1) (ค.1) และ (ง.1) เป็นภาพก่อนเลื่อนองค์ประกอบแนวตั้ง และ (ก.2) (ข.2) (ค.2) และ (ง.2) เป็นภาพหลังเลื่อนองค์ประกอบแนวตั้ง 27

รูปที่ 5-3 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวบนชุดข้อมูลภาษาไทย..... 28

รูปที่ 5-4 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวบนชุดข้อมูลภาษาไทย แสดงเพียงช่วง 30,000 ถึง 150,000 รอบ 28

รูปที่ 5-5 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางบนชุดข้อมูลภาษาไทย 29

รูปที่ 5-6 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางบนชุดข้อมูลภาษาไทย แสดงเพียงช่วง 30,000 ถึง 150,000 รอบ..... 29

รูปที่ 5-7 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด 31

รูปที่ 5-8 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด 32

รูปที่ 5-9 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งทั้ง 5 รอบ ในรอบที่ 1,000 ถึง 150,000 34

รูปที่ 5-10 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งทั้ง 5 รอบ ในรอบที่ 1,000 ถึง 150,000 35

รูปที่ 5-11 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลฝึกขนาดใหญ่ทั้ง 5 พับ 37

รูปที่ 5-12 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยบนชุดข้อมูลฝึกขนาดใหญ่ทั้ง 5 พับ 38

รูปที่ 5-13 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองที่ 4.4.4 39

รูปที่ 5-14 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองที่ 4.4.4 40

รูปที่ 5-15 บรรทัดตัวอักษรภาษาไทย (ก) - (ฉ) และข้อมูลนำออกของโครงข่ายแอลเอสทีเอ็มที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย (LSTM with VCS) ABBYY และ ArnThai ตามลำดับ ตัวอักษรหนาสีแดงในแต่ละบรรทัดข้อมูลนำออกแสดงถึงคำที่ทำนายผิดพลาด 42

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การรู้จำตัวอักษร (Optical Character Recognition - OCR) คือการแปลงภาพเอกสารให้เป็นข้อความที่สามารถนำไปแก้ไขต่อได้ การรู้จำตัวอักษรสามารถแบ่งกลุ่มตามประเภทของข้อมูลนำเข้าได้ 2 ประเภทคือ วิธีแบ่งส่วน (Segmented) และวิธีไม่แบ่งส่วน (Unsegmented) เมื่อข้อมูลนำเข้าเป็นภาพบรรทัดของตัวอักษร วิธีแบ่งส่วนจะพยายามแบ่งส่วนของภาพเป็นตัวอักษร และนำตัวอักษรไปผ่านขั้นตอนวิธีหรือโมเดลเพื่อให้ได้คำตอบ การเรียนรู้ของโมเดลที่เกิดขึ้นจึงเป็นการเรียนรู้ตัวอักษรแต่ละตัวอักษรแยกออกจากกัน ซึ่งผลลัพธ์ที่ได้เป็นตัวอักษรแต่ละตัวที่ไม่เกี่ยวข้องกัน เมื่อนำมาประกอบกันจะได้เป็นข้อความของภาพบรรทัดตัวหนังสือ วิธีแบ่งส่วนเป็นวิธีที่ได้รับความนิยมมากในการรู้จำตัวอักษร ปัญหาส่วนใหญ่จะตกไปอยู่ที่การแบ่งส่วนตัวอักษร ดังนั้นการประมวลผลก่อน (Pre-processing) จึงเป็นส่วนสำคัญสำหรับกลุ่มวิธีนี้ งานวิจัยวิธีแบ่งส่วน เช่น การใช้โครงข่ายประสาทเทียมแบบป้อนไปด้านหน้า (Feed-Forward Neural Network – FFNN) กับตัวอักษรพิมพ์ภาษาไทย [1] การใช้แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine - SVM) ในการจำแนกประเภทและรู้จำอักษรพิมพ์ภาษาไทยและภาษาอังกฤษ [2] การหาเพื่อนบ้านที่ใกล้เคียงสุด k ตัว (K-Nearest Neighbor) โดยระยะทางแบบยุคลิดในการรู้จำตัวอักษรพิมพ์ภาษาไทย [3] การใช้เพื่อนบ้านที่ใกล้สุดร่วมกับ SVM อักษรพิมพ์ภาษาไทยและภาษาอังกฤษ [4] การรู้จำตัวอักษรภาษาไทยที่ขาดหายโดยใช้กราฟเอ็นแกรม (N-gram) [5] การใช้แบบจำลองฮิตเดนมาร์คอฟ (Hidden Markov Model - HMM) ในการรู้จำตัวอักษรภาษามาลายาลัม [6] การแข่งขันการรู้จำตัวอักษรภาษาไทยแบบแบ่งส่วน [7] สำหรับกลุ่มวิธีไม่แบ่งส่วน จะทำการเรียนรู้ข้อความภาพบรรทัดของตัวอักษรทั้งหมด โดยมีแนวคิดว่าข้อมูลของภาพตัวอักษรเป็นข้อมูลที่ขึ้นอยู่กับเวลา หรือกล่าวคือตัวอักษรแต่ละตัวมีความสัมพันธ์กันและมีผลต่อการทำนายตัวอักษรถัดไป งานวิจัยวิธีแบบไม่แบ่งส่วน เช่น ใช้แบบจำลอง HMM กับตัวอักษรพิมพ์ภาษาอังกฤษ [8] และมีงานวิจัยที่ใช้โครงข่ายหน่วยความจำระยะสั้นแบบยาว – แอลเอสทีเอ็ม (Long Short-Term Memory - LSTM) [9] ซึ่งเป็นโครงข่ายประสาทเทียมแบบหมุนเวียนกลับ (Recurrent Neural Network - RNN)

โครงข่ายประสาทเทียมแบบหมุนเวียนกลับ เป็นโครงข่ายที่มีการเชื่อมต่อกันและเกิดวงวน (Loop) ในชั้นซ่อน (Hidden layer) ข้อมูลนำเข้าและข้อมูลนำออกเป็นอนุกรมเวลา (Time-series) ประโยชน์ที่สำคัญในการทำนายบนข้อมูลอนุกรมเวลาคือนอกจากสามารถเรียนรู้สารสนเทศจากสถานะตนเองแล้ว ยังสามารถเรียนรู้สารสนเทศจากสถานะก่อนหน้าได้ คล้ายกับการดูบริบทของสิ่งที่เรียนรู้ ซึ่งสามารถช่วยในการแก้ปัญหาตัวอักษรที่เชื่อมติดกันหรือตัวอักษรที่มีบางส่วนขาดหายได้ แต่การเรียนรู้กับงานที่มีขนาดใหญ่ (เวลามาก) จะพบปัญหาวานิชชิงเกรเดียนต์ (Vanishing Gradient Problem) ซึ่งเป็นปัญหาการสูญเสียสารสนเทศของข้อมูลไปตามเวลา แอลเอสทีเอ็มจึงถูกออกแบบขึ้นมาเพื่อแก้ปัญหาข้างต้น ด้วยการเก็บสารสนเทศโดยใช้ประตู (Gate) ซึ่งจะกล่าวต่อไปในหัวข้อที่ 2 แอลเอสทีเอ็มถูกนำมาประยุกต์ใช้ได้ดีกับงานรู้จำตัวอักษรทั้งตัวพิมพ์ และมีงานวิจัยบนชุดข้อมูล

ลายมือเขียนแบบออนไลน์และออฟไลน์พร้อมทั้งเปรียบเทียบกับแบบจำลอง HMM [10] การปรับโครงข่ายแอลเอสทีเอ็มในหลายมิติ [11] และให้ประสิทธิภาพที่สูงเทียบเคียงกับซอฟต์แวร์ที่ได้รับความนิยมบนชุดข้อมูลภาษาอังกฤษและภาษาฟรักทูร์ที่มีตัวอักษรเชื่อมติดกันจำนวนมาก [12]

เป้าหมายในวิทยานิพนธ์นี้คือแสดงประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มในการรู้จำตัวอักษรพิมพ์ภาษาไทย ซึ่งเป็นภาษาที่มีโครงสร้างตัวอักษร 4 ระดับ ทำให้เกิดปัญหารูปแบบของตัวอักษรที่เกิดขึ้นในแนวตั้งจำนวนมาก จึงนำเสนอการเลื่อนองค์ประกอบแนวตั้ง โดยการนำองค์ประกอบเชื่อมติดกันมาช่วยในการจัดลำดับองค์ประกอบของภาพบรรทัดตัวอักษรเพื่อเพิ่มประสิทธิภาพของการเรียนรู้ แบบจำลองที่นำเสนอคือโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางร่วมกับการเลื่อนองค์ประกอบแนวตั้ง ในส่วนของการทดลองประกอบด้วยการปรับพารามิเตอร์ให้เหมาะสมกับการรู้จำตัวอักษรภาษาไทยในแต่ละการทดลอง แสดงประสิทธิภาพของการเลื่อนองค์ประกอบแนวตั้ง และการเรียนรู้บนชุดข้อมูลขนาดที่แตกต่างกัน นอกจากนี้ยังแสดงการเปรียบเทียบประสิทธิภาพกับซอฟต์แวร์ที่ได้รับความนิยมอย่าง แอบบ์บี (ABBYY) [13] และอ่านไทย (ArnThai) [14]

1.2 วัตถุประสงค์

1. แสดงประสิทธิภาพวิธีไม่แบ่งส่วนตัวอักษรโดยใช้โครงข่ายแอลเอสทีเอ็มในการรู้จำตัวอักษรภาษาไทย
2. แสดงประสิทธิภาพของขั้นตอนวิธีในการเลื่อนองค์ประกอบแนวตั้งร่วมกับโครงข่ายแอลเอสทีเอ็ม
3. เปรียบเทียบโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ ABBYY และ ArnThai

1.3 ขอบเขตการดำเนินงาน

1. วิทยานิพนธ์นี้ใช้ชุดข้อมูลที่จัดทำขึ้นจากประมวลกฎหมายอาญา ประมวลกฎหมายวิธีพิจารณาความอาญา ประมวลกฎหมายแพ่งและพาณิชย์ และประมวลกฎหมายวิธีพิจารณาแพ่งและพาณิชย์ประกอบด้วยตัวอักษรภาษาไทย ตัวเลข และเครื่องหมายอีกจำนวนหนึ่ง
2. ใช้เครื่องมือของ OCRopus ในการทำภาพขาวดำ การแบ่งส่วนบรรทัดตัวอักษร การฝึกแบบจำลองแอลเอสทีเอ็ม และการคำนวณหาค่าความผิดพลาด
3. เปรียบเทียบแบบจำลองโดยใช้การวัดค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์น
4. เปรียบเทียบโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ ABBYY และ ArnThai

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้เห็นการเรียนรู้และประสิทธิภาพของการรู้จำตัวอักษรแบบไม่แบ่งส่วนของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้ง และความสามารถเมื่อเทียบกับซอฟต์แวร์ที่ได้รับความนิยม
2. ได้เห็นความผิดพลาดของการรู้จำแบบไม่แบ่งส่วนของโครงข่ายแอลเอสทีเอ็ม ในการรู้จำตัวอักษรพิมพ์อักษรภาษาไทย

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

วิทยานิพนธ์นี้เกี่ยวข้องกับการโครงข่ายแอลเอสทีเอ็มและการเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลภาษาไทย ดังนั้นในบทนี้จะกล่าวถึงลักษณะตัวอักษรภาษาไทย การหาองค์ประกอบที่เชื่อมติดกันที่ถูกนำมาประยุกต์ใช้เป็นการเลื่อนองค์ประกอบแนวตั้งต่อไป โครงข่ายแอลเอสทีเอ็มซึ่งมีสถาปัตยกรรมมาจากโครงข่ายประสาทเทียมแบบหมุนเวียนกลับ และการคำนวณค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์น

2.1 ตัวอักษรภาษาไทย

ตัวอักษรภาษาไทย เป็นอักษรที่ใช้เขียนภาษาไทย และภาษากลุ่มน้อยอื่นในประเทศไทย มีพยัญชนะ 44 ตัวอักษร สระ 22 ตัวอักษร และวรรณยุกต์ 4 รูป ตัวเลขไทยและเครื่องหมายอื่นอีกจำนวนหนึ่ง โครงสร้างการวางตัวของตัวอักษรภาษาไทยที่มี 4 ระดับ ประกอบด้วย ระดับกลาง ระดับล่าง ระดับบน และระดับบนสุด พยัญชนะไทยจะเรียงตัวไปในระดับกลางจากซ้ายไปขวา แต่พยัญชนะบางตัวอาจมีบางส่วนของตัวอักษรที่เลยไปในระดับอื่นได้ สระไทยสามารถอยู่ด้านหน้า ด้านบน ด้านล่าง และด้านหลังของพยัญชนะเพื่อประกอบคำแล้วแต่ชนิดของสระ และวรรณยุกต์ไทยจะวางตัวในระดับบน หรือระดับบนสุด การสะกดภาษาไทยจะสะกดตาม พยัญชนะ สระ และวรรณยุกต์ ตามลำดับ ตัวอย่างบรรทัดตัวอักษรภาษาไทย แสดงดังรูปที่ 2-1 สำหรับตัวอักษรภาษาไทยในการเข้ารหัสแบบยูนิโคด (Unicode) จำนวน 129 ตัวอักษร แสดงดังตารางที่ 2-1

ตารางที่ 2-1 ตัวอักษรภาษาไทย ตัวเลข และเครื่องหมายที่ใช้ในชุดข้อมูลตามรหัสยูนิโคด

	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*A*	*B*	*C*	*D*	*E*	*F*
U+002		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
U+003	0	1	2	3	4	5	6	7	8	9	:	;	<	+	>	?
U+004	@															
U+005												[	\	]	^	_
U+006	`															
U+007												{	|	}	~	
...																
U+0E0		ก	ข	ฃ	ค	ฅ	ฆ	ง	จ	ฉ	ช	ซ	ฌ	ญ	ฎ	ฏ
U+0E1	ฐ	ฑ	ฒ	ณ	ด	ต	ถ	ท	ธ	น	บ	ป	ผ	ฝ	พ	ฟ
U+0E2	ภ	ม	ย	ร	ฤ	ล	ฦ	ว	ศ	ษ	ส	ห	ฬ	อ	ฮ	ฯ
U+0E3	ะ	ั	า	ำ	ิ	ี	ึ	ื	ุ	ู	ฺ					฿
U+0E4	เ	แ	โ	ใ	ไ	ๅ	ๆ	็	่	้	๊	๋	์	ํ	๎	๏
U+0E5	๐	๑	๒	๓	๔	๕	๖	๗	๘	๙	๚	๛				

รูปที่ 2-1 คำในภาษาไทย “อรุณสวัสดิ์” ที่มีครบทุกระดับตัวอักษร

2.2 องค์ประกอบที่เชื่อมติดกัน (Connected Component)

องค์ประกอบที่เชื่อมติดกัน ใช้ทฤษฎีกราฟนำมาประยุกต์ใช้งานกับรูปภาพ โดยการหาส่วนประกอบย่อยจากภาพ โดยกระทำได้ทั้งบนภาพขาวดำและภาพสี สำหรับภาพขาวดำทำโดยการมองจุดภาพดำเป็นจุดของกราฟ (Vertices) และการหาเส้นเชื่อม (Edges) มีแบบ 4 ทิศทาง (4-Connectivity) และ 8 ทิศทาง (4-Connectivity) แสดงดังรูป 2-2 โดยงานวิจัยนี้ใช้การทำองค์ประกอบที่เชื่อมติดกันในภาพบรรทัดตัวอักษรที่เป็นภาพขาวดำ

รูปที่ 2-2 (ก) การเชื่อมติด 4 ทิศทาง และ (ข) การเชื่อมติด 8 ทิศทาง

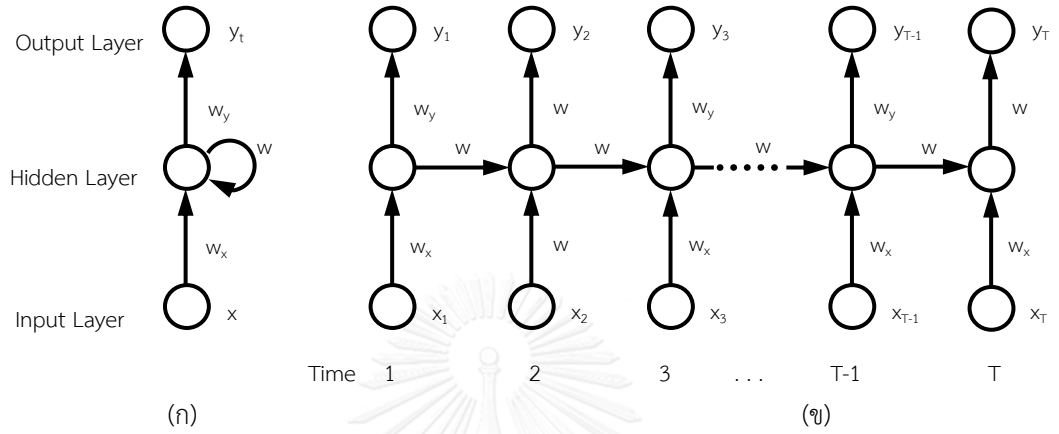
การนำองค์ประกอบที่เชื่อมติดกันมาใช้กับภาพบรรทัดตัวอักษร ข้อมูลนำออกที่ได้รับจะเป็นองค์ประกอบที่เชื่อมติดกันแต่ละองค์ประกอบ กรณีที่ไม่มีตัวอักษรใดเชื่อมติดกันองค์ประกอบที่เชื่อมติดกันที่ได้จะเป็นภาพตัวอักษรแต่ละตัวอักษรในบรรทัดตัวอักษรและถูกนำไปใช้ต่อไป

Chulalongkorn University

2.3 โครงข่ายประสาทเทียมแบบหมุนเวียนกลับ (Recurrent Neural Network - RNN) [15]

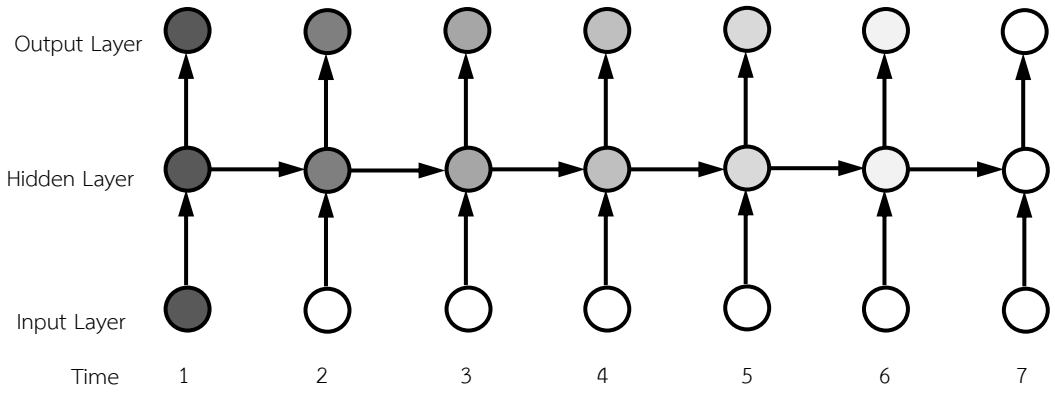
โครงสร้างของโครงข่ายประสาทเทียมแบบป้อนไปด้านหน้าเป็นโครงข่ายแบบไม่มีวงวน ซึ่งลักษณะที่สำคัญของโครงข่ายประสาทเทียมแบบหมุนเวียนกลับคือมีวงวน ข้อมูลนำเข้าและข้อมูลนำออกที่เป็นอนุกรมเวลา ซึ่งแตกต่างจากลักษณะของข้อมูลนำเข้าและข้อมูลนำออกแบบโครงข่ายประสาทเทียมแบบป้อนไปด้านหน้าที่เป็นข้อมูลที่ไม่ขึ้นอยู่กับเวลา ในการจำแนกหรือรู้จำข้อมูลจะจำแนกข้อมูลแต่ละข้อมูล เช่นในการรู้จำตัวอักษรจำแนกตัวอักษรแต่ละตัว ในขณะที่โครงข่ายประสาทเทียมแบบหมุนเวียนกลับทำการรู้จำหรือจำแนกข้อมูลนำเข้าเป็นอนุกรมเวลา และให้ข้อมูลนำออกเป็นอนุกรมเวลา ทำให้มีประโยชน์ที่สำคัญคือนอกจากสามารถเรียนรู้สารสนเทศจากสถานะตนเองแล้วยังสามารถเรียนรู้สารสนเทศจากสถานะก่อนหน้าได้ คล้ายกับการดูบริบทของสิ่งที่เรียนรู้ ทำให้เหมาะสมกับการรู้จำตัวอักษรที่มีลำดับการวางตัวอักษร และการรู้จำเสียงที่มีลำดับเสียงคำพูด เป็นต้น การใช้โครงข่ายประสาทเทียมแบบหมุนเวียนกลับในการรู้จำตัวอักษรสามารถช่วยในการแก้ปัญหาตัวอักษรที่เชื่อมติดกันหรือตัวอักษรที่มีบางส่วนขาดหายได้ โครงข่ายประสาทเทียมแบบหมุนเวียนกลับอย่างง่าย (Simple Recurrent Neural Network) แสดงในรูปที่ 2-3 (ก) โดยมีชั้นนำเข้า (Input layer) ชั้นนำออก (Output layer) และชั้นซ่อน

ชั้นละ 1 บัพ (Node) และค่าถ่วงน้ำหนัก (Weight) 3 ค่า ประกอบไปด้วยค่าถ่วงน้ำหนักจากชั้นนำเข้าไปยังชั้นซ่อนแทนด้วย w_x ค่าถ่วงน้ำหนักจากชั้นซ่อนไปยังชั้นซ่อนแทนด้วย w และค่าถ่วงน้ำหนักจากชั้นซ่อนไปยังชั้นนำออกแทนด้วย w_y เมื่อมีข้อมูลนำเข้าและข้อมูลนำออกจำนวน T ตัว สามารถขยาย รูปที่ 2-3 (ก) ตามเวลาเป็นดังรูปที่ 2-3 (ข)



รูปที่ 2-3 (ก) โครงข่ายประสาทเทียมแบบหมุนเวียนกลับอย่างง่ายที่ประกอบไปด้วยชั้นนำเข้า ชั้นซ่อน และชั้นนำออก อย่างละ 1 บัพ และ (ข) ขยายโครงข่ายประสาทเทียมแบบหมุนเวียนกลับอย่างง่ายตามเวลา

เมื่อทำการเรียนรู้ข้อมูลนำเข้าทั้งหมดจนเสร็จสิ้นก็จะได้ข้อมูลนำออกตั้งแต่ 1 จนถึง T จะสังเกตว่าถ้าค่าของข้อมูลนำเข้าลำดับที่ 1 (x_1) ถูกคูณด้วย w ทุกรอบจะทำให้ข้อมูลนำเข้าลำดับที่ 1 จะสูญหายไปตามเวลา ก่อให้เกิดปัญหาเรียกว่า ปัญหาวานิชชิงเกรเดียนต์ (Vanishing Gradient Problem) แสดงได้ดังรูปที่ 2-4 ทำให้โครงข่ายไม่สามารถเก็บสารสนเทศได้เป็นในระยะเวลานาน ส่งผลให้ไม่เหมาะสมกับงานที่มีข้อมูลนำเข้าขนาดใหญ่ เพราะสารสนเทศของข้อมูลนำเข้าจะสูญหายไปตามเวลาอย่างรวดเร็ว ทำให้ไม่สามารถนำสารสนเทศของสถานะก่อนหน้ามาใช้ให้เกิดประโยชน์เท่าที่ควร

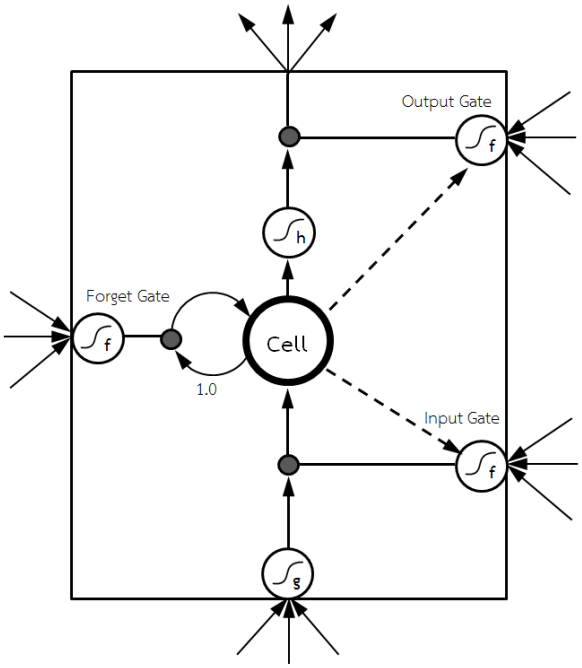


รูปที่ 2-4 ปัญหาวานิชชิงเกรเดียนต์ในโครงข่ายประสาทเทียมแบบหมุนเวียนกลับ ความเข้มของสีแสดงถึงการสูญหายไปของสารสนเทศข้อมูลนำเข้าลำดับที่ 1 ตามเวลา

2.4 หน่วยความจำระยะสั้นแบบยาว (Long Short-term Memory - LSTM) [9]

แอลเอสทีเอ็มถูกออกแบบให้มีสถาปัตยกรรมที่แก้ไขปัญหาวานิชชิงเกรเดียนต์ การทำนายข้อมูลนำออกไม่เพียงขึ้นอยู่กับข้อมูลนำเข้าในเวลาเดียวกัน แต่ยังนำสารสนเทศของสถานะก่อนหน้ามาร่วมช่วยในการทำนาย เมื่อเวลาผ่านไปสารสนเทศที่ควรอยู่ไม่หายไปตามเวลาหรืออยู่ได้เพียงในช่วงเวลาสั้น แอลเอสทีเอ็มยังคงสามารถรักษาสารสนเทศที่มีสำคัญในระยะเวลาที่นานขึ้นภายในโครงข่าย ซึ่งสามารถเก็บสารสนเทศไปได้ยาวนานถึง 1,000 ช่วงเวลาแบบไม่ต่อเนื่อง (Discrete time steps) ซึ่งทำได้โดยการสร้างการเชื่อมต่อกับตนเอง (Self-connected) ภายในบัพด้วยหน่วยเชิงเส้น (linear unit) การรักษาสารสนเทศให้อยู่ไปได้ในระยะเวลานานนั้นทำให้เราสามารถนำสารสนเทศของข้อมูลนำเข้าในสถานะก่อนหน้ามาช่วยทำนายข้อมูลนำออกในสถานะปัจจุบันได้ ซึ่งคล้ายกับการดูบริบทของข้อมูลรอบข้างเพื่อช่วย

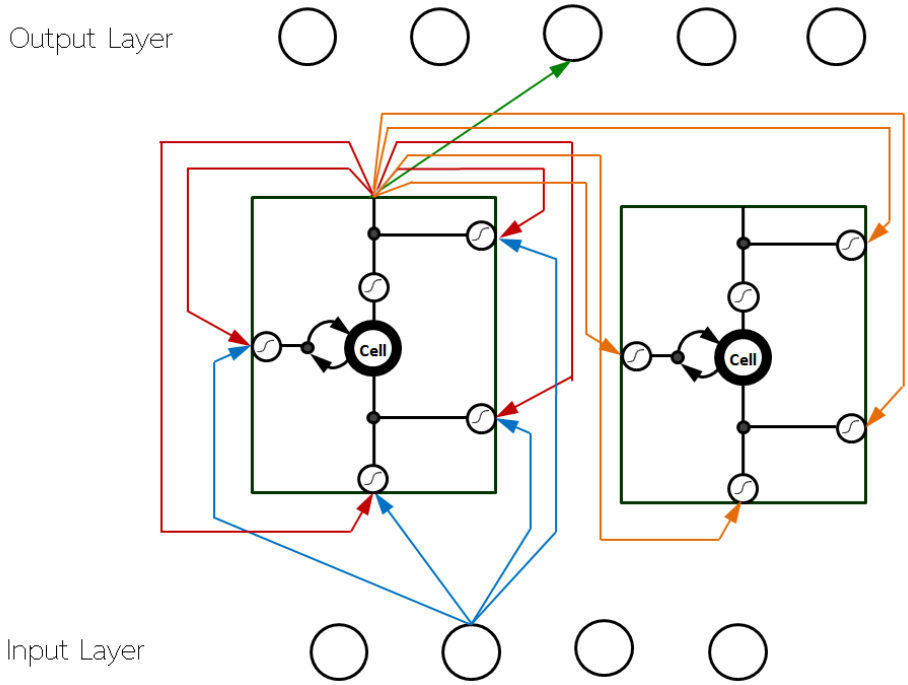
แอลเอสทีเอ็มถูกออกแบบให้สามารถเก็บค่าคงที่หรือสารสนเทศโดยใช้ บล็อกหน่วยความจำ (Memory block) แทนบัพในชั้นซ่อน บล็อกหน่วยความจำจะประกอบด้วย เซลล์หน่วยความจำ (Memory cell) ภายใน และประตูควบคุมการทำงานประกอบด้วย ประตูนำเข้า (Input gate) ประตูนำออก (Output gate) และประตูลืม (Forget gate) สามารถแสดงได้ดังรูปที่ 2-5



รูปที่ 2-5 บล็อกหน่วยความจำของแอลเอสทีเอ็ม 1 เซลล์หน่วยความจำ ประตูนำเข้า ประตูนำออก และประตูลืม

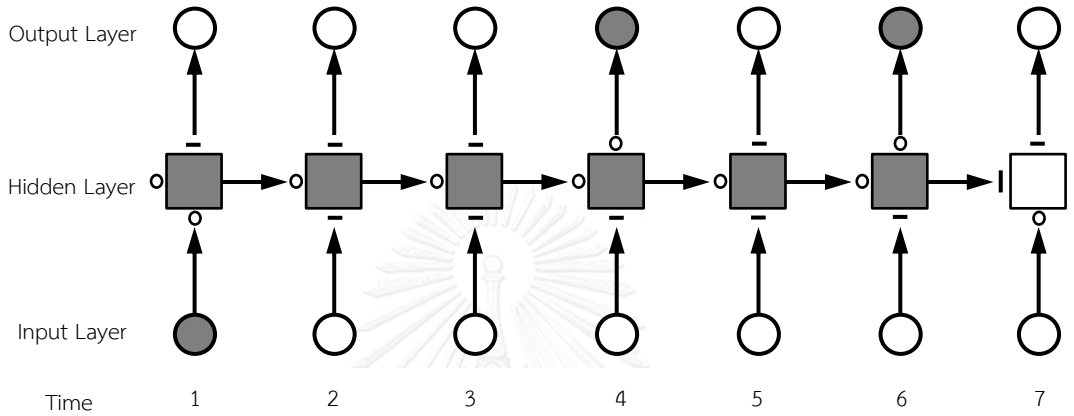
จากรูปที่ 2-5 ประตูทั้งสามทำหน้าที่ควบคุมการรับและส่งสารสนเทศจากภายในและภายนอกบล็อกหน่วยความจำ ผลรวมสารสนเทศที่ได้จะถูกนำมาผ่านฟังก์ชันกระตุ้น (Activation function) และผ่านตัวคูณ (วงกลมเล็กสีดำ) ประตูนำเข้าทำหน้าที่ควบคุมการรับข้อมูลนำเข้าในสถานะปัจจุบัน โดยนำค่าที่ได้จากผลรวมและ

ข้อมูลนำเข้ามาผ่านตัวคูณก่อนเข้าสู่เซลล์หน่วยความจำ ประตูนำออกทำหน้าที่ควบคุมการส่งค่าคงที่ถูกเก็บในเซลล์หน่วยความจำในสถานะปัจจุบัน โดยนำค่าที่ได้จากผลรวมและค่าคงที่ในเซลล์หน่วยความจำผ่านฟังก์ชันกระตุ้นและตัวคูณก่อนส่งสารสนเทศไปเป็นข้อมูลนำออกและส่งไปยังสถานะถัดไป ประตูลืมทำหน้าที่ควบคุมการจดจำหรือเปลี่ยนแปลงค่าคงที่ในเซลล์หน่วยความจำ โดยค่าที่ได้จากผลรวมจะเป็นตัวกำหนดการจดจำหรือเปลี่ยนแปลงค่าคงที่ซึ่งถ้าจะนำค่าที่อยู่ภายในเซลล์หน่วยความจำจะเชื่อมต่อกับตัวเอง ฟังก์ชันกระตุ้น ‘f’ มักใช้โลจิสติกซิกมอยด์ (logistic sigmoid) ค่าอยู่ระหว่าง 0 (ประตูปิด) ถึง 1 (ประตูเปิด) ฟังก์ชันกระตุ้นที่เข้าและออกจากเซลล์ ‘g’ และ ‘h’ มักใช้ tanh หรือ logistic sigmoid ซึ่งในบางครั้ง ‘h’ อาจเป็นฟังก์ชันเอกลักษณ์ (Identity function) ค่าถ่วงน้ำหนักที่ใช้ภายในบล็อกหน่วยความจำจากเซลล์หน่วยความจำไปยัง ประตูนำเข้า และประตูนำออก แสดงด้วยเส้นประ การเชื่อมต่อในบล็อกหน่วยความจำทั้งหมดที่ถูกแทนด้วยเส้นทึบจะไม่มีการถ่วงน้ำหนักหรือมีค่าถ่วงน้ำหนักเท่ากับ 1.0 การส่งสารสนเทศไปยังบัพทั้งหมดภายในโครงข่ายแอลเอสทีเอ็มสามารถแสดงได้ดังรูปที่ 2-6



รูปที่ 2-6 โครงข่ายแอลเอสทีเอ็ม ประกอบด้วยชั้นนำเข้าจำนวน 4 บัพ ชั้นนำออกจำนวน 5 บัพ และชั้นซ่อนจำนวน 2 บัพหรือ 2 บล็อกหน่วยความจำ แต่ละบล็อกหน่วยความจำจะมีข้อมูลนำเข้าได้ 4 ทางและมีเพียง 1 ข้อมูลนำออก เส้นเชื่อมที่แสดงค่าถ่วงน้ำหนักเป็นเพียงส่วนหนึ่งเท่านั้น ค่าถ่วงน้ำหนักระหว่างบัพในชั้นข้อมูลนำเข้ากับบล็อกหน่วยความจำแสดงด้วยเส้นเชื่อมสีฟ้า ค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำเดียวกันแสดงด้วยเส้นเชื่อมสีแดง ค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำแสดงด้วยเส้นเชื่อมสีส้ม และค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำไปยังบัพในชั้นนำออกแสดงด้วยเส้นเชื่อมสีเขียว

บล็อกหน่วยความจำสามารถช่วยทำให้ในแต่ละสถานะตัดสินใจโดยมีการนำสารสนเทศของข้อมูลนำเข้าในสถานะก่อนหน้าและสารสนเทศของข้อมูลนำเข้าในสถานะเดียวกันมาใช้ผ่านประตูนำเข้าและประตูลืม และนำสารสนเทศออกจากบล็อกหน่วยความจำไปยังชั้นข้อมูลนำออกผ่านประตูนำออก สารสนเทศของข้อมูลนำเข้าในสถานะก่อนหน้าสามารถถูกจำหรือลืมไปจากสถานะเดิม ทำให้สารสนเทศในโครงข่ายแอลเอสทีเอ็มเดินทางข้ามช่วงเวลาไปยังสถานะต่อไปได้ดังรูปที่ 2-7



รูปที่ 2-7 การแก้ปัญหาปัญหาวานิชชิงเกรเดียนต์ของโครงข่ายประสาทเทียมแบบหมุนเวียนกลับ โดยการใช้บล็อกหน่วยความจำ (แทนด้วยสี่เหลี่ยมจัตุรัส) มารักษาสารสนเทศเพื่อส่งไปยังสถานะถัดไป รูปทรงที่มีสีดำคือได้รับสารสนเทศจากเวลาที่ 1 และสีขาวคือไม่เกี่ยวข้องกับสารสนเทศจากเวลาที่ 1 วงกลมขนาดเล็ก (o) และเส้นตรง (-) แทนประตูเปิด และประตูปิดตามลำดับ ซึ่งจะถูกวางทางขวาล่าง ขวาบน และซ้ายของบล็อกหน่วยความจำ คือ ประตูเข้า ประตูนำออก และประตูลืม ตามลำดับ สารสนเทศจากเวลาที่ 1 สามารถเดินทางไปยังเวลาที่ 6 ได้โดยประตูลืมของบล็อกหน่วยความจำเวลาที่ 2 ถึง 6 ที่เปิดรับสารสนเทศจากเวลาที่ 1 และประตูนำเข้าเปิดเพียงช่วงเวลาที่ 1 เท่านั้นจึงไม่มีการรับข้อมูลนำเข้าจากเวลาอื่นจนกระทั่งเวลาที่ 7 และประตูนำออกที่เปิดเพียงเวลาที่ 4 และ 6 ทำให้สารสนเทศของเวลาที่ 1 สามารถเดินทางไปยังเวลาที่ 4 และ 6 โดยไม่สูญหายไปตามเวลา

การนำโครงข่ายประสาทเทียมแบบหมุนเวียนกลับและโครงข่ายแอลเอสทีเอ็มมาประยุกต์ใช้ เหมาะกับข้อมูลประเภทแบบไม่แบ่งส่วนหรือข้อมูลมีลำดับซึ่งเป็นงานประเภทการรับรู้ เช่น การรู้จำตัวอักษร การรู้จำลายมือเขียน การรู้จำเสียง การรู้จำท่าทาง เป็นต้น ลักษณะของข้อมูลนำเข้าและข้อมูลนำออกเป็นข้อมูลที่มีลำดับ ตัวอย่างการรู้จำตัวอักษรหลังจากแปลงภาพบรรทัดตัวอักษรเป็นข้อมูลนำออกที่มีลักษณะเป็นความน่าจะเป็น ณ เวลาทั้งหมดแล้ว โดยอีกหนึ่งขั้นตอนที่สำคัญคือการแปลงความน่าจะเป็นที่ได้จากข้อมูลนำออกโครงข่ายแอลเอสทีเอ็มเป็นตัวอักษรที่สามารถอ่านได้ ซึ่งใช้การจำแนกประเภทการเชื่อมต่อตามเวลา (Connectionist Temporal Classification - CTC) [16] แนวคิดที่สำคัญของการจำแนกประเภทการเชื่อมต่อตามเวลาคือการแปลงจากข้อมูลนำออกจากโครงข่ายเป็นข้อมูลที่มีป้ายกำกับหรือข้อมูลตัวอักษรที่สามารถอ่านได้ โดยการสนใจปรากฏการณ์ที่สำคัญ

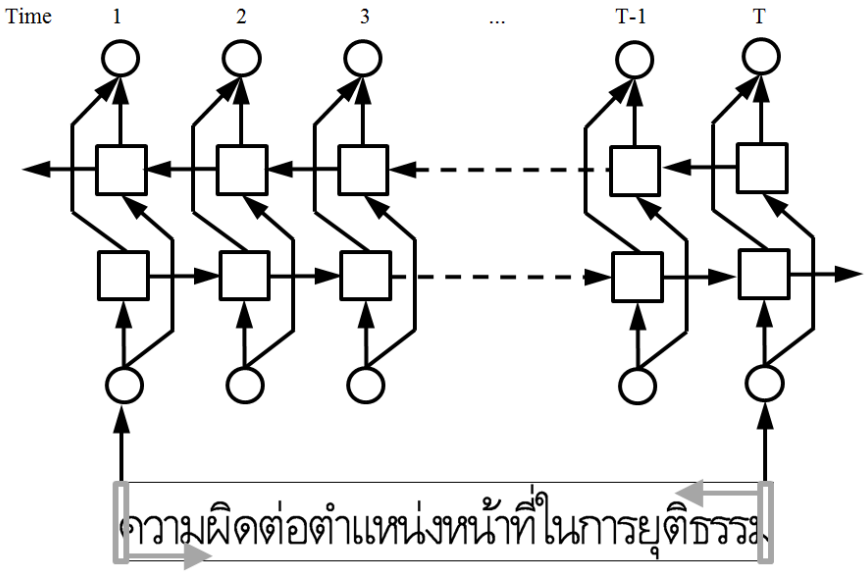
บนค่าความน่าจะเป็นของข้อมูลนำออกจากโครงข่าย และให้ผลลัพธ์เป็นข้อมูลที่มีป้ายกำกับหรือข้อมูลตัวอักษรที่สามารถอ่านได้

ในการเรียนรู้ของโครงข่ายประสาทเทียมแบบหมุนเวียนกลับรวมไปถึงโครงข่ายแอลเอสทีเอ็ม สามารถการเรียนรู้ค่าถ่วงน้ำหนักของโครงข่ายใช้ขั้นตอนวิธีแบบไปด้านหน้าและย้อนกลับ แบ่งออกเป็น 2 ส่วนคือส่วนไปด้านหน้า (Forward pass) และส่วนของการแพร่ย้อนกลับผ่านเวลา (Back-propagation Through Time – BPTT) หรือส่วนการย้อนกลับ (Backward pass) [15] ซึ่งเปรียบเสมือนการนำโครงข่ายประสาทเทียมแบบหมุนเวียนกลับมาขยายไปตามเวลาจะได้โครงสร้างของโครงข่ายประสาทเทียมแบบป้อนไปด้านหน้า ข้อมูลนำเข้าที่เรียนรู้เป็นค่าช่วงเวลาแบบไม่ต่อเนื่อง และให้ข้อมูลนำออกเป็นความน่าจะเป็นในช่วงเวลาแบบไม่ต่อเนื่องขนาดน้อยกว่าหรือเท่ากับข้อมูลนำเข้า หลังจากนั้นสามารถใช้การจำแนกประเภทการเชื่อมต่อตามเวลาแปลงข้อมูลนำออกเพื่อแปลงเป็นข้อมูลที่สนใจต่อไป โดยแนวคิดของการจำแนกประเภทการเชื่อมต่อตามเวลาคือการแปลงข้อมูลนำออกจากโครงข่ายประสาทเทียมแบบหมุนเวียนกลับเป็นค่าความน่าจะเป็นที่เป็นไปได้ของป้ายกำกับทั้งหมด และนำชุดความน่าจะเป็นที่ได้ไปเปลี่ยนเป็นข้อมูลนำออกที่แท้จริงต่อไป

2.5 โครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง (Bidirectional LSTM network)

แนวคิดหลักของโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางคือ การใช้สารสนเทศในสองทิศทางจากซ้ายมาขวาและจากขวาไปซ้าย การรู้จำตัวอักษรจากภาพบรรทัดตัวอักษรของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวจากซ้ายไปขวา ทำให้เราสามารถนำสารสนเทศในสถานะก่อนหน้ามาช่วยในการทำนายของสถานะถัดไป บางครั้งสารสนเทศที่มีประโยชน์ไม่เพียงแต่รู้มาจากข้อมูลนำเข้าที่ผ่านมาแต่ยังมีสารสนเทศจากสถานะถัดไป เปรียบกับการรู้จำตัวอักษรเมื่อเราต้องการทำนายตัวอักษรที่สนใจ นอกจากความรู้ของตัวอักษรดังกล่าวแล้วความรู้จากตัวอักษรก่อนหน้า และตัวอักษรตัวถัดไปสามารถช่วยในการทำนายตัวอักษรสนใจได้

โครงข่ายแอลเอสทีเอ็มสองทิศทางสามารถทำงานโดยการนำโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวมาเชื่อมต่อกันสองชุดแต่ใช้ชั้นนำเข้าและชั้นนำออกร่วมกัน ชุดแรกจะเป็นการเรียนรู้จากทางซ้ายไปขวาลักษณะของข้อมูลนำเข้าจะเป็นตั้งแต่ 1 ถึง T ขณะที่ชุดที่สองจะเป็นการเรียนรู้จากทางขวาไปซ้ายลักษณะของข้อมูลนำเข้าจะเป็นตั้งแต่ T ถึง 1 ดังนั้นสารสนเทศในของบัฟในข้อมูลนำออกจะมาจากบล็อกของหน่วยความจำสองชุดสามารถแสดงได้ดังรูปที่ 2-8



รูปที่ 2-8 โครงข่ายแอลเอสทีเอ็มแบบสองทิศทางที่ถูกนำมาขยายตามเวลา ประกอบด้วยบัพชั้นข้อมูลนำเข้าและชั้นข้อมูลนำออกอย่างละ 1 บัพ และ 2 บล็อกหน่วยความจำ (แสดงด้วยรูปสี่เหลี่ยม) ในชั้นซ่อน การเลื่อนหน้าต่าง หน้าต่าง (แสดงด้วยรูปสี่เหลี่ยมผืนผ้าสีเทา) มีความกว้าง 1 พิกเซล นำมาสร้างเป็นเวกเตอร์ของข้อมูลนำเข้าจากเวลาที่ 1 ถึง T และจาก T ถึง 1 เมื่อ T คือความกว้างของภาพบรรทัดตัวอักษร

2.6 การคำนวณค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์น (Levenshtein distance)

เป็นขั้นตอนวิธีในการวัดหาค่าความแตกต่างระหว่างสายอักขระ (String) 2 ชุด ชุดอักขระต้นแบบ และชุดอักษรเปรียบเทียบ ค่าความแตกต่างจะวัดจากตัวดำเนินการประกอบด้วย การแทรก (Insert) การตัดออก (Delete) การแทนที่ (Replace) เป็นขั้นตอนวิธีแบบกำหนดการพลวัต (Dynamic Programming) ซึ่งถูกใช้ในการทดลองเพื่อตรวจสอบค่าความผิดพลาดจากผลเฉลย กับบรรทัดตัวอักษรที่ทำนายได้ ในการคำนวณค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์นระหว่างชุดอักษรต้นแบบ (i) และชุดอักษรเปรียบเทียบ (j) สามารถหาได้จากสมการฟังก์ชันที่ (2.1)

$$lev_{a,b}(i,j) = \begin{cases} \max(i,j) & if\ \min(i,j) = 0, \\ min\begin{cases} lev_{a,b}(i-1,j)+1 \\ lev_{a,b}(i,j-1)+1 \\ lev_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & otherwise. \end{cases} \quad (2.1)$$

กำหนดให้

$lev_{a,b}(i,j)$ คือค่าความแตกต่างระหว่างชุดอักษรต้นแบบและชุดอักษรเปรียบเทียบ

ซึ่งการคำนวณค่าความผิดพลาดโดยใช้ค่าระยะทางเลเวนชเตย์นถูกนำใช้มากในการคำนวณค่าความผิดพลาดของการรู้จำตัวอักษรเพื่อนำไปใช้เปรียบเทียบ ซึ่งในวิทยานิพนธ์นี้ได้ใช้ค่าระยะทางเลเวนชเตย์นในการทดลองทั้งหมดแสดงเป็นเปอร์เซ็นต์ค่าความผิดพลาด

บทที่ 3
งานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรภาษาไทย ซึ่งมีลักษณะปัญหาและวิธีแก้ไขต่างจากการรู้จำตัวอักษรในภาษาอื่น และงานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรแบบไม่แบ่งส่วนรวมไปถึงงานวิจัยที่เกี่ยวข้องกับการรู้จำที่ประยุกต์ใช้โครงข่ายแอลเอสทีเอ็มในการรู้จำตัวอักษร นอกจากนี้ยังกล่าวถึงโอซีอาร์โอปุส (OCRopus) ซอฟต์แวร์ที่นำมาประยุกต์ใช้งานวิจัยนี้

3.1 งานวิจัยการรู้จำตัวแบบแบ่งส่วน และงานวิจัยการรู้จำตัวอักษรภาษาไทย

งานวิจัยการรู้จำตัวอักษรบนภาษาไทยส่วนใหญ่เป็นแบบการแบ่งส่วน ชุดข้อมูลจึงเป็นภาพตัวอักษรแต่ละตัวรวมไปถึงตัวอักษรที่อยู่ในระดับบน ระดับบนสุด และระดับล่างด้วย ซึ่งในการแบ่งส่วนตัวอักษรแต่ละตัวจะทำให้สระบางตัวมีลักษณะคล้ายกับพยัญชนะ เช่น สระอู (‘ู’) ขอไข่ (‘ข’) นอกจากนี้ภาษาไทยยังเกิดการเชื่อมติดกันของตัวอักษรได้ค่อนข้างง่าย และยังมีการเชื่อมติดระหว่างสระในระดับบนอีกด้วย เช่น คำว่า ‘ไม้ไผ่’ จะสังเกตว่า ไม้หันอากาศ ‘ั’ และไม้มลาย ‘ไ’ เชื่อมติดกันเป็นต้น ปัญหาสำคัญอย่างหนึ่งของการรู้จำตัวอักษรภาษาไทยจึงตกไปอยู่ที่การแบ่งส่วนตัวอักษรและตำแหน่งของตัวอักษร งานวิจัยการรู้จำตัวอักษรของภาษาไทยส่วนใหญ่เป็นแบบการแบ่งส่วน และมีงานวิจัยการรู้จำอักษรภาษาไทยจำนวนมากโดยมีลักษณะงานวิจัยที่จัดการกับปัญหาการแบ่งส่วนและงานวิจัยที่ใช้ในการรู้จำตัวอักษรแต่ละตัวอักษร การรู้จำตัวอักษรภาษาไทยเริ่มในปี 1996 โดย Chalarat Tanprasert และ Thaweesak Koanantakool ใช้โครงข่ายประสาทเทียมแบบป้อนไปด้านหน้ามารู้จำตัวอักษร [1] ซึ่งให้ค่าความถูกต้องบนชุดข้อมูลทดสอบประมาณ 80% ปี 2007 แบบจำลอง SVM ได้รับความนิยมและถูกนำมารู้จำอักษรพิมพ์ภาษาไทยและภาษาอังกฤษ [2] ค่าความถูกต้องบนชุดข้อมูลทดสอบอยู่ระหว่าง 96% ถึง 99% ปี 2013 วิธีการเพื่อนบ้านที่ใกล้สุด k ตัว โดยระยะทางแบบยุคลิดกับตัวอักษรพิมพ์ภาษาไทย [3] วิธีการเพื่อนบ้านที่ใกล้สุดร่วมกับ SVM [4] ได้ค่าความถูกต้องที่สูงกว่าวิธีอื่นในบางชุดการทดสอบ ซึ่งมีงานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรภาษาไทยแบบแบ่งส่วนจำนวนมาก และยังมีการจัดทำการแข่งขันการวัดเปรียบเทียบสมรรถนะเพื่อพัฒนามาตรฐานการประมวลผลภาษาไทย (Benchmark for Enhancing the Standard of Thai language processing - BEST) โดยหัวข้อการแข่งขันในปี 2013 และ 2014 คือการแข่งขันสุดยอดการรู้จำตัวอักษรภาษาไทย (Thai Character Recognition Contest) [7] เป็นระยะเวลา 2 ปีต่อเนื่อง จัดโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (National Electronics and Computer Technology Center - NECTEC) ข้อมูลนำเข้าทั้งหมดในการแข่งขันถูกเตรียมเป็นข้อมูลแบบแบ่งส่วน ซึ่งหมายความว่าเป็นภาพตัวอักษรเดี่ยวแต่ละตัว ทำให้เทคนิคหรือแบบจำลองในการทำนายจึงเป็นแบบแบ่งส่วน ชุดข้อมูลทดสอบแบ่งเป็น 3 ชุด ผลงานแข่งขันที่ดีที่สุดในชุดข้อมูลทดสอบที่ 1 ใช้วิธีการเพื่อนบ้านใกล้สุดให้ค่าความถูกต้องสูงสุด 97.88% และชุดข้อมูลทดสอบที่ 2 และ 3 วิธีการแปลงเวฟเล็ตของผู้เข้าแข่งขันสามารถให้ค่าความถูกต้องสูงสุดถึง 97.83% และ 97.75% ตามลำดับ

นอกจากงานวิจัยแบบการแบ่งส่วนเป็นตัวอักษรเดี่ยวแต่ละตัว และนำไปรู้จำข้างต้น ยังมีงานวิจัยที่ไม่เรียนรู้ภาพบรรทัดตัวอักษรเป็นตัวอักษรแต่ละตัว แต่เรียนรู้ตัวอักษรเป็นชุดของตัวอักษร ในปี 2014 มีผู้นำแบบจำลอง HMM มาใช้กับการรู้จำตัวอักษรภาษามาลายาลัม [6] ซึ่งยังคงใช้แบบจำลอง HMM เพื่อรู้จำตัวอักษรภาษามาลายาลัมเป็นชุดของตัวอักษรเพียงไม่กี่ตัว แบบจำลอง HMM ที่ถูกนำมาประยุกต์ใช้มากกับงานรู้จำหรือ

ทำนายจำนวนมาก และสามารถทำงานบนข้อมูลนำเข้าที่เป็นแบบอนุกรมเวลาได้ แต่ยังมีเงื่อนไขจำนวนมากที่ต้องใช้ความรู้ เช่น จำนวนของบัพ การตีความหมายของโครงสร้าง HMM เป็นต้น

สังเกตว่างานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรมีแบบจำลองที่ดีในการแยกแยะตัวอักษรแต่ละตัวอักษร แต่ปัญหาในงานวิจัยการรู้จำตัวอักษรส่วนใหญ่ไม่อยู่ที่แบบจำลองในการรู้จำตัวอักษรแต่ละตัว แต่อยู่ที่การแบ่งส่วนจากภาพบรรทัดตัวอักษรเป็นตัวอักษรเดี่ยวแต่ละตัว เพราะภาพบรรทัดตัวอักษรที่เกิดขึ้นนั้นอาจเกิดการเชื่อมติดกันของตัวอักษรหรือการขาดหายของตัวอักษร ทำให้ข้อมูลนำเข้าของแบบจำลองในการรู้จำตัวอักษรแต่ละตัวผิดพลาด จึงทำให้มีงานวิจัยส่วนหนึ่งที่ศึกษาเกี่ยวกับการแบ่งส่วนภาพบรรทัดตัวอักษรเป็นตัวอักษรแต่ละตัว เช่น ปี 2012 งานวิจัยเกี่ยวกับการรู้จำตัวอักษรที่ขาดหายโดยใช้กราฟเอ็นแกรม [5] นอกจากส่วนของการประมวลผลก่อนแล้ว ยังมีส่วนของการทำประมวลผลภายหลัง เช่น การใช้แบบจำลองภาษา พจนานุกรม ซึ่งจะช่วยในการปรับแก้ข้อมูลนำออกให้ถูกต้องมากยิ่งขึ้นภายหลัง

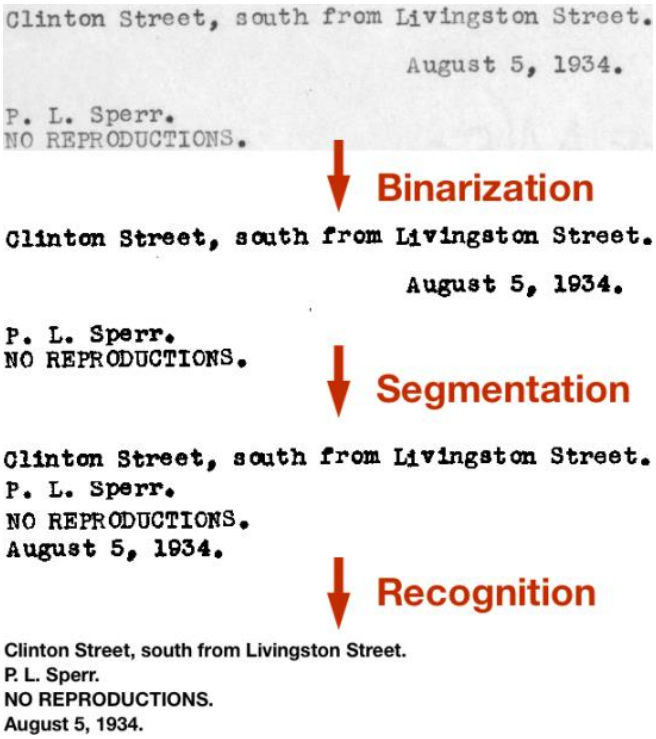
3.2 งานวิจัยการรู้จำตัวอักษรแบบไม่แบ่งส่วน

จากปัญหาของวิธีการรู้ตัวอักษรแบบแบ่งส่วน จึงนำมาสู่วิธีการรู้จำตัวอักษรแบบไม่แบ่งส่วน ซึ่งใช้แนวคิดของการเรียนรู้ภาพบรรทัดตัวอักษรเป็นบรรทัดตัวอักษรเลย โดยไม่มีการแบ่งภาพบรรทัดตัวอักษรเป็นตัวอักษรเดี่ยว ทำให้ลดปัญหาของการเชื่อมต่อกันหรือขาดหายของตัวอักษร นอกจากนี้ยังช่วยในเรื่องของการทำนายโดยใช้บริบทรอบข้างมาทำนายตัวอักษรในแต่ละตัว ซึ่งจะกล่าวต่อไปในข้อหัวถัดไป ลักษณะข้อมูลนำเข้าในการรู้จำตัวอักษรแบบไม่แบ่งส่วนคือภาพบรรทัดตัวอักษรที่ถูกมองเป็นข้อมูลอนุกรมเวลา ในปี 2008 Alex Grave และคณะ ใช้ข้อมูลลายมือเขียนแบบออนไลน์และออฟไลน์ รู้จำโดยโครงข่ายแอลเอสทีเอ็ม และเปรียบเทียบผลการทดลองกับแบบจำลอง HMM [10] ผลการทดลองของการใช้โครงข่ายแอลเอสทีเอ็ม สามารถได้ผลดีกว่าแบบจำลอง HMM ในทุกชุดข้อมูลทดสอบ ในปี 2010 Kamel A. Mohnad และคณะ นำไปประยุกต์ใช้กับ HMM [8] บนชุดข้อมูลภาษาอังกฤษ 70 ฟอนต์ที่หลากหลายแบบมีทั้งตัวพิมพ์และตัวเขียน ผลการทดลองที่ได้สามารถเรียนรู้ชุดข้อมูลทดสอบฟอนต์ที่ไม่เคยพบมาก่อนได้ดีได้ค่าความถูกต้องถึง 97% แต่ในบางฟอนต์ยังสามารถเรียนรู้ได้ไม่ดีนัก ผู้วิจัยใช้วิธีการเลื่อนหน้าต่าง (Sliding Window Approach) เพื่อแปลงภาพข้อมูลนำเข้าเป็นเวกเตอร์ ขนาดตามความสูงของภาพ และเลื่อนไปตามแนวความกว้างของภาพ ทำให้เวลาเป็นแบบช่วงเวลาแบบไม่ต่อเนื่องขนาดเท่ากับความกว้างของภาพ ในปี 2012 Alex Grave จัดทำงานวิจัยที่ทดลองปรับโครงข่ายแอลเอสทีเอ็ม ในหลายมิติ [11] ซึ่งใช้โครงข่ายของแอลเอสทีเอ็ม มาเชื่อมต่อกันในชั้นซ่อนที่มากกว่า 1 ชั้น และได้ผลการทดลองดีกว่าขั้นตอนวิธีอื่นในการแข่งขัน

จากงานวิจัยข้างต้นของ ผู้ที่เกี่ยวข้องกับงานวิจัยเกี่ยวกับโครงข่ายแอลเอสทีเอ็มได้ร่วมมือกันจัดทำซอฟต์แวร์ OCRopus ขึ้น ซึ่งเป็นซอฟต์แวร์ที่ถูกกล่าวถึงในงานวิจัย มีการรู้จำตัวอักษรโดยใช้โครงข่ายแอลเอสทีเอ็มในส่วนของการเรียนรู้ และใช้วิธีการเลื่อนหน้าต่างภาพบรรทัดตัวอักษรเพื่อเป็นข้อมูลนำเข้าโครงข่าย รายละเอียดของ OCRopus จะกล่าวต่อไปใน 3.3

3.3 โอซีอาร์โอปุส (OCRopus) [18]

OCRopus เป็นซอฟต์แวร์โอเพนซอร์ซ (open source) [17] ที่จัดทำขึ้นมาเพื่องานรู้จำตัวอักษร เมื่อปี 2008 มีแยกส่วนของโค้ดเป็นโมดูลอย่างชัดเจน เช่น ส่วนของการแปลงภาพขาวดำ ส่วนของการแปลงภาพหน้ากระดาษเป็นภาพบรรทัดตัวอักษร ส่วนของการเรียนรู้โดยโครงข่ายแอลเอสทีเอ็ม และส่วนของการคำนวณค่าความผิดพลาดจากการทำนาย เป็นต้น ทำให้ง่ายต่อการนำไปใช้ประยุกต์ใช้งาน รองรับกับข้อมูลนำเข้าเป็นภาพบรรทัดตัวอักษรแบบไม่แบ่งส่วน และใช้โครงข่ายแอลเอสทีเอ็มที่สามารถปรับเปลี่ยนได้ในส่วนของการรู้จำตัวอักษร ตัวอย่างการรู้จำตัวอักษรโดยใช้ OCRopus แสดงดังรูปที่ 3-1

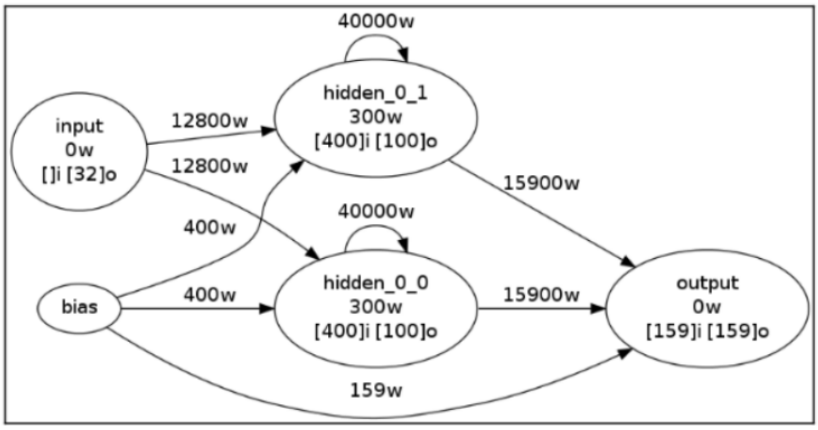


รูปที่ 3-1 การทำงานของ OCRopus ในการทำนายภาพโดยการแปลงภาพเป็นขาวดำ ทำการแบ่งส่วนภาพบรรทัดตัวอักษร และทำนายตัวอักษรในภาพบรรทัด [18]

เครื่องมือของ OCRopus ที่เกี่ยวข้องกับวิทยานิพนธ์นี้ ได้แก่ เครื่องมือการทำภาพขาวดำ (ocropus-nlbin) สำหรับการแปลงภาพเป็นภาพขาวดำผ่านค่าขีดแบ่ง เครื่องมือการแบ่งส่วนบรรทัดข้อความ (ocropus-gpageseg) ซึ่งใช้เทคนิคอนุพันธ์ตามแนวแกนนอนของเคอร์เนลเกาส์เซียน (y-derivative of a Gaussian kernel) โดยใช้ขอบบนและขอบล่างองค์ประกอบเชื่อมติดกันเพื่อหาขอบบนและขอบล่างของภาพบรรทัดตัวอักษร เครื่องมือการฝึกแบบจำลอง (ocropus-rtrain) สำหรับฝึกแบบจำลองโครงข่ายแอลเอสทีเอ็มผ่านชุดข้อมูลฝึก เครื่องมือการทำนาย (ocropus-rpred) สำหรับการนำแบบจำลองโครงข่ายแอลเอสทีเอ็มที่สนใจมาทดสอบกับชุดข้อมูลทดสอบ

และให้ผลลัพธ์เป็นบรรทัดตัวอักษรที่สามารถนำไปแก้ไขต่อได้ และเครื่องมือการคำนวณหาค่าความผิดพลาด (ocropus-errs) ซึ่งใช้ในการตรวจสอบค่าความผิดพลาดของบรรทัดตัวอักษรและผลเฉลยโดยใช้ค่าระยะทางเลเวนชเตย์น

OCROpus ได้นำโครงข่ายแอลเอสทีเอ็ม มาใช้และเรียนรู้ข้อมูลภาพจากภาพบรรทัดข้อความ โดยข้อมูลนำเข้าเกิดจากการใช้วิธีการเลื่อนหน้าต่างขนาด ความกว้างและความสูง เท่ากับ 1 และ 48 ตามลำดับ หมายถึงภาพบรรทัดตัวอักษรทั้งหมดจะถูกปรับให้มีความสูง 48 พิกเซล ซึ่งเป็นจำนวนที่เพียงพอต่อการเรียนรู้ภาพบรรทัดตัวอักษรและยังเท่ากับจำนวนบัพในชั้นนำเข้าอีกด้วย งานวิจัยที่นำเสมอการเรียนรู้บรรทัดตัวอักษรโดยใช้โครงข่ายแอลเอสทีเอ็มผ่าน OCRopus ปี 2013 Thomas M. Breuel ได้แสดงประสิทธิภาพการรู้จำตัวอักษร บนชุดข้อมูลภาษาอังกฤษ และภาษาฟรักทูร์ [12] โดยใช้แอลเอสทีเอ็มแบบ 2 ทิศทาง (Bidirectional LSTM) และตั้งค่าชั้นนำเข้าไว้จำนวน 32 บัพ ตามความสูงของภาพบรรทัดข้อความในแต่ละเวลา ชั้นซ่อนจำนวน 100 บัพหรือ 100 บล็อกหน่วยความจำ ในแต่ละบล็อกจะมีเซลล์หน่วยความจำ 1 เซลล์ โดยจะมีจำนวน 2 ชุด สำหรับสองทิศทาง (ซ้ายไปขวา และขวาไปซ้าย) ชั้นนำออกจำนวน 159 บัพ ตามจำนวนตัวอักษรภาษาอังกฤษ ภาษาฟรักทูร์ ตัวเลข และเครื่องหมาย ดังรูปที่ 3-2



รูปที่ 3-2 โครงสร้างของแอลเอสทีเอ็มแบบ 2 ทิศทางที่ประกอบไปด้วยค่าถ่วงน้ำหนักจากชั้นนำเข้าไปยังชั้นซ่อน 12,800 ค่า จากชั้นซ่อนไปยังชั้นนำออก 15,900 ค่า และค่าถ่วงน้ำหนักวงวนเชื่อมต่อกับตนเองจำนวน 40,000 ค่า [12]

จากการทดลองในงานวิจัย [12] บนภาษาอังกฤษใช้ชุดข้อมูล UW3 ประกอบด้วยตัวอักษรภาษาอังกฤษจำนวน 50,632 ตัว ชุดข้อมูล Fontane ประกอบด้วยตัวอักษรภาษาเยอรมันและภาษาฟรักทูร์จำนวน 8,988 ตัว และชุดข้อมูล Ersch-Gruber ประกอบด้วยตัวอักษรภาษาเยอรมันและภาษาฟรักทูร์จำนวน 10,881 ตัว ซึ่งผลการทดลองที่ได้สามารถเทียบเคียงกับการซอฟต์แวร์ที่ได้รับความนิยมอย่าง ABBYY [13] และ Tesseract [19] ดังตารางที่ 3-1 นอกจากผลการทดลองที่แสดงประสิทธิภาพของการรู้จำตัวอักษรแบบไม่แบ่งส่วนโดยใช้

โครงข่ายแอลเอสทีเอ็มแล้ว ยังสามารถแก้ปัญหาการการติดกันของตัวอักษรจำนวนมากในภาษาฟรักทูร์โดยใช้เทคนิควิธีที่ไม่ซับซ้อน ไม่ใช้โมเดลของภาษา พจนานุกรมภาษา หรือการประมวลผลภายหลัง

ตารางที่ 3-1 เปอร์เซนต์ค่าความผิดพลาด OCRopus บนภาษาอังกฤษและภาษาฟรักทูร์ [12]

OCR System	English	Fontane	Ersch-Gruber
OCRopus-LSTM	0.60%	0.15%	1.37%
Tesseract	1.30%	0.90%	1.47%
ABBYY	0.85%	1.23%	0.43%

จากงานวิจัยข้างต้นทั้งหมด ผู้วิจัยจึงเล็งเห็นว่าควรนำการรู้จำตัวอักษรแบบไม่แบ่งส่วนผ่านโครงข่ายแอลเอสทีเอ็มมาใช้กับการรู้จำตัวอักษรภาษาไทย ซึ่งสามารถช่วยแก้ไขปัญหา การเชื่อมติดกันหรือขายหายของภาพตัวอักษร และยังนำสารสนเทศจากบริบทรอบข้างมาใช้ในการทำนายตัวอักษรแต่ละตัว นอกจากปัญหาในการรู้จำตัวอักษรที่เกิดขึ้นในทุกภาษาแล้ว ยังมีปัญหาที่เกิดขึ้นเฉพาะกับภาษาที่มีหลายระดับอย่างภาษาไทย และส่งผลกระทบโดยตรงกับการเรียนรู้โดยใช้ข้อมูลนำเข้าที่มาจากการเลื่อนหน้าต่าง จึงนำความรู้ขององค์ประกอบที่เชื่อมติดกันมาประยุกต์ใช้ ซึ่งจะกล่าวถึงการนำมาประยุกต์ใช้และการทดลองต่อไปในหัวข้อถัดไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
แนวทางและวิธีดำเนินงาน

วิทยานิพนธ์นี้นำเสนอการรู้จำตัวอักษรภาษาไทยโดยใช้วิธีแบบไม่แบ่งส่วนผ่านโครงข่ายแอลเอสทีเอ็มและการเลื่อนองค์ประกอบเชื่อมติดกัน ซึ่งในบทนี้จะกล่าวถึงการจัดเตรียมชุดข้อมูลภาษาไทย การเลื่อนองค์ประกอบแนวตั้ง โครงข่ายแอลเอสทีเอ็มที่นำมาใช้และการตั้งค่าพารามิเตอร์ และการทดลองเพื่อแสดงประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มและการเลื่ององค์ประกอบที่เชื่อมติดกัน ซึ่งการทดลองประกอบด้วย การทดลองเพื่อเปรียบเทียบการใช้โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง การทดลองเพื่อเปรียบเทียบโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบเชื่อมติดกันร่วมด้วยและแบบไม่ใช้การเลื่อนองค์ประกอบเชื่อมติดกันบนชุดข้อมูลฝึกที่มีอยู่อย่างจำกัดและชุดข้อมูลฝึกที่มีขนาดใหญ่ และการทดลองเพื่อเปรียบเทียบประสิทธิภาพโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบเชื่อมติดกันกับซอฟต์แวร์ที่ได้รับความนิยมบนชุดข้อมูลทดสอบที่ไม่เคยพบมาก่อน

4.1 การเตรียมชุดข้อมูล

เนื่องจากชุดข้อมูลแบบวิธีไม่แบ่งส่วนของภาษาไทยไม่เป็นที่นิยมจึงไม่มีชุดข้อมูลมาตรฐานสำหรับการทดลองจึงต้องจัดเตรียมชุดข้อมูลเอง โดยการจัดเตรียมสร้างชุดข้อมูลภาษาไทยจากประมวลกฎหมายอาญา ประมวลกฎหมายวิธีพิจารณาความอาญา ประมวลกฎหมายแพ่งและพาณิชย์ และประมวลกฎหมายวิธีพิจารณาแพ่งและพาณิชย์ ตัวอักษรที่ปรากฏบนชุดข้อมูลประกอบด้วยตัวอักษรภาษาไทย ตัวเลข และเครื่องหมาย ตามการเข้ารหัสแบบยูนิโคดรวม 129 แบบตัวอักษร ซึ่งจะมีผลต่อจำนวนบัพในชั้นข้อมูลนำออกของโครงข่ายแอลเอสทีเอ็มต่อไป

ชุดข้อมูลถูกจัดทำบนฟอนต์มาตรฐานและฟอนต์หลากหลายลักษณะ จำนวน 52 ฟอนต์ แสดงดังตารางที่ 4-1 ลักษณะของฟอนต์ที่ถูกนำมาจัดเตรียม เช่น ฟอนต์มาตรฐานราชการไทยอย่าง TH Sarabun New TH Chakra Petch TH KoHo เป็นต้น ฟอนต์ที่ได้รับความนิยมเป็นฟอนต์มาตรฐานชุดเก่าอย่าง Angsana New Browallia New Cordia New เป็นต้น นอกจากฟอนต์มาตรฐานเพิ่มส่วนฟอนต์ให้มีความหลายหลากมายขึ้น สามารถจัดกลุ่มย่อยของฟอนต์ไว้ได้เป็น ฟอนต์ที่มีลักษณะมีหัวในหลายลักษณะ (ตั้งแต่ฟอนต์ที่ 1-26) ฟอนต์ที่ไม่มีหัวในหลายลักษณะ (ตั้งแต่ฟอนต์ที่ 27-45) และฟอนต์ที่มีลักษณะการใช้เส้นโค้งและรูปแบบเฉพาะในแต่ละตัวอักษร (ตั้งแต่ฟอนต์ที่ 46-52) แต่ละฟอนต์มีลักษณะความแตกต่างกันออกไป เช่น ของความหนาของเส้น ขนาดและตำแหน่งของสระ ขนาดของหัวตัวอักษร ความสูงของขนาดตัวอักษร เป็นต้น

ตารางที่ 4-1 ฟอนต์ชุดข้อมูล

ฟอนต์	ตัวอย่าง
1. Angsana New	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
2. Browallia New	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
3. Cordia New	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
4. DilleniaUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
5. EucrosiaUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
6. FreesiaUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
7. IrisUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
8. Microsoft Sans Serif	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
9. TH Baijam	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
10. TH Fah Kwang	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
11. CS_PraJad	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
12. CS_PraKas	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
13. RSU Text	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
14. Nithan	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
15. 4804_KwangMD_PukluK	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
16. 4815_KwangMD_Catthai	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
17. DSN AdiRek	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
18. DSN AmPun	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
19. DSN ArKorn	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
20. DSN PreeCha	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
21. DSN SuRaDej	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
22. TH Sarabun New	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
23. TH Chakra Petch	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
24. CSChatThai	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
25. Arundina San	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
26. DSN AnuRuk	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
27. JasmineUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4-1 ฟอนต์ชุดข้อมูล (ต่อ)

ฟอนต์	ตัวอย่าง
28. KodchiangUPC	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
29. Circular	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
30. Quark	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
31. Bangna New	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
32. LilyUPC	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
33. Superspace	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
34. RSU	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
35. FontCraft	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
36. DSN LardPhrao	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
37. DSN PaNuTat	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
38. DSN Single	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
39. EDSense	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
40. Cloud Light	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
41. MAX PINJOHNV2	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
42. WDB Bangna	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
43. Thai Sans Neue	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
44. Supermarket	**วันพรุ่งนี้ไม่ได้ไปโรงเรียน**
45. Kunlasatri	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
46. Peach Play	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
47. Ti_Text	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
48. Layji Mak	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
49. 4803_Kwang_MD	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
50. 4805_KwangMD_Melt	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
51. TH KoHo	วันพรุ่งนี้ไม่ได้ไปโรงเรียน
52. TH Mali Grade 6	วันพรุ่งนี้ไม่ได้ไปโรงเรียน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การจัดเตรียมภาพชุดข้อมูล จัดทำโดยการแปลงไฟล์เอกสารเป็นไฟล์ภาพหน้ากระดาษตัวอักษรที่มีความละเอียดภาพ 300 dpi (Dot per Inch) จากนั้นนำไฟล์ภาพหน้ากระดาษตัวอักษรไปผ่านเครื่องมือแปลงภาพเป็นขาวดำ โดยใช้เครื่องมือของโอซีอาร์โอปุส โดยเครื่องมือที่ถูกนำมาใช้ในการแปลงภาพขาวดำและปรับเอียงของหน้ากระดาษคือ ocropus-nlbin กำหนดให้ค่าขีดแบ่งในการทำภาพขาวดำเท่ากับ 0.2 จากนั้นนำภาพขาวดำที่ได้ไปผ่านเครื่องมือแบ่งส่วนภาพบรรทัดตัวอักษร ocropus-gpageseg ซึ่งใช้เทคนิคอนุพันธ์ตามแนวแกนนอนของเคอร์เนลเกาส์เซียน โดยใช้ขอบบนและขอบล่างองค์ประกอบเชื่อมติดกัน เพื่อหาขอบบนและขอบล่างของภาพบรรทัดตัวอักษร จากปัญหาภาพบรรทัดตัวอักษรภาษาไทยที่มีหลายระดับ ทำให้ลักษณะของบรรทัดตัวอักษรจะมีขอบบนและขอบล่างขนาดใหญ่ ในส่วนของการแบ่งส่วนบรรทัดตัวอักษรจำเป็นต้องขยายความสำคัญให้กับส่วนประกอบทั้งหมดโดยไม่ตัดส่วนประกอบใดออกในบรรทัดตัวอักษรที่จากขั้นตอนวิธีเดิมจะมองตัวอักษรในระดับบนหรือล่างที่อยู่สูงหรือต่ำมากจนเกินไปเป็นภาพที่ไม่สำคัญและควรตัดออก แต่ในภาพบรรทัดตัวอักษรภาษาไทยมีการขยายความสำคัญโดยไม่ตัดทิ้งส่วนใดของภาพบรรทัดตัวอักษรเพื่อครอบคลุมตัวอักษรที่อยู่ในระดับบนและล่างมากขึ้น นอกจากภาพบรรทัดตัวอักตักษรที่ได้จากเครื่องมือข้างต้น ยังมีส่วนของไฟล์ข้อความ (Ground text) ที่ใช้สำหรับการสอนและตรวจวัดค่าความผิดพลาดสำหรับแต่ละภาพบรรทัดตัวอักษรของชุดข้อมูลอีกด้วย

หลังจากได้ชุดข้อมูลภาพบรรทัดตัวอักษรเป็นข้อมูลนำเข้าของโครงข่ายแอลเอสทีเอ็ม ซึ่งมีจำนวนบัพของข้อมูลนำเข้าเป็นค่าคงที่หมายถึงข้อมูลนำเข้าที่เข้าสู่โครงข่ายแอลเอสทีเอ็มต้องมีขนาดเท่ากัน ซึ่งในวิทยานิพนธ์นี้ใช้การเลื่อนหน้าต่างมาช่วยในการนำเข้าของภาพบรรทัดตัวอักษร โดยขนาดของหน้าต่างที่ใช้มีขนาดความกว้างและความสูง เท่ากับ 1 และ 48 ตามลำดับ และเลื่อนหน้าต่างไปตามแนวนอน ดังนั้นภาพบรรทัดตัวอักษรแต่ละภาพจะถูกปรับขนาดให้มีความสูงเท่ากับ 48 พิกเซล ข้อมูลนำเข้าโครงข่ายแอลเอสทีเอ็มที่ได้จึงเป็นเวกเตอร์ของจุดภาพตามความสูงของภาพบรรทัดตัวอักษร ในการนำการเลื่อนหน้าต่างมาใช้กับการรู้จำตัวอักษรภาษาไทย หน้าต่างแต่ละหน้าต่างจะพบจำนวนพิกเซลของตัวอักษรที่วางตัวอยู่มากกว่าหนึ่งตัวอักษร ทำก่อให้เกิดการเรียนรู้ที่ยากขึ้น จากปัญหาของการรู้จำตัวอักษรภาษาไทยที่มีหลายระดับ วิทยานิพนธ์นี้จึงใช้การเลื่อนองค์ประกอบแนวตั้งมาช่วยในการแก้ไขปัญหา

4.2 การเลื่อนองค์ประกอบแนวตั้ง (Vertical Component Shifing)

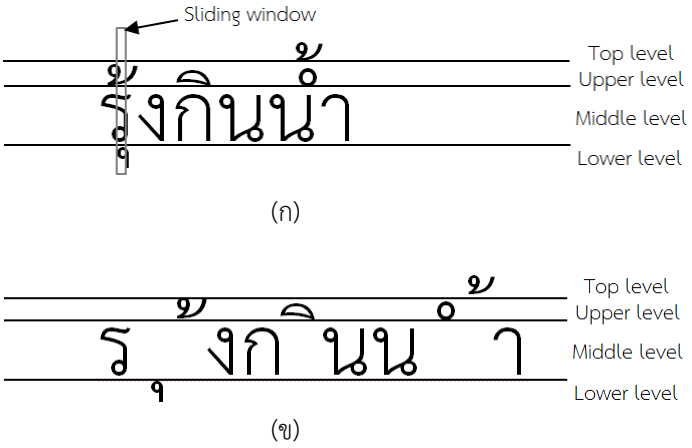
เพื่อที่จะแก้ไขปัญหาของภาษาไทยที่มีถึง 4 ระดับ ทำให้รูปแบบของตัวอักษรที่เกิดขึ้นในแนวตั้งมีจำนวนมาก ทำให้การเรียนรู้โดยการเลื่อนหน้าต่างในแต่ละเวกเตอร์พบตัวอักษรมากกว่า 1 ตัว เช่นคำว่า “รู้” จะมีตัวอักษรถึง 3 ตัวประกอบกันแต่จะถูกเรียงตัวกันในแนวตั้งเป็นต้น ในการเรียนรู้รูปแบบตัวอักษรที่เกิดขึ้นทุกรูปแบบนั้นเป็นไปได้ยาก การเลื่อนองค์ประกอบแนวตั้งถูกสร้างมาเพื่อช่วยแก้ปัญหาข้างต้น แนวคิดหลักของการเลื่อนองค์ประกอบแนวตั้งคือ ต้องการลดจำนวนของรูปแบบของตัวอักษรที่เกิดขึ้นในแนวตั้งก่อนนำไปเป็นข้อมูลนำเข้าของโครงข่ายแอลเอสทีเอ็ม เพื่อให้โครงข่ายแอลเอสทีเอ็มสามารถเรียนรู้รูปแบบของตัวอักษรที่เกิดขึ้นได้ครอบคลุมมากยิ่งขึ้น

ขั้นตอนในการเลื่อนองค์ประกอบแนวตั้งสามารถทำโดยการหาองค์ประกอบที่เชื่อมติดกันแบบ 4 ทิศทางของภาพบรรทัดตัวอักษร แบ่งองค์ประกอบที่เชื่อมติดกันเป็น ส่วนกลาง ส่วนบน และส่วนล่าง โดยใช้เส้นแบ่งองค์ประกอบที่เชื่อมติดกัน สามารถหาได้จากการวัดจำนวนพิกเซลสีดำตามแนวนอนของภาพ สร้างเป็นฮิสโตแกรมขนาดเท่ากับความสูงของภาพและเลือกส่วนของความสูงของภาพที่มีจุดภาพสีดำหนาแน่นที่สุดมาเป็นเส้นแบ่งองค์ประกอบที่เชื่อมติดกัน และติดป้ายกำกับองค์ประกอบที่เชื่อมติดกันเพื่อแบ่งเป็นองค์ประกอบที่เชื่อมติดกัน 3 ลักษณะคือ องค์ประกอบที่เชื่อมติดกันส่วนบน ส่วนกลาง และส่วนล่าง กล่าวคือองค์ประกอบที่เชื่อมติดกันส่วนบนจะมีจุดพิกเซลทั้งหมดจะต้องอยู่เหนือเส้นแบ่งองค์ประกอบที่เชื่อมติดกัน องค์ประกอบที่เชื่อมติดกันส่วนกลางจะมีจุดพิกเซลบางจุดอยู่ในระดับเดียวกันกับเส้นแบ่งองค์ประกอบที่เชื่อมติดกัน และองค์ประกอบที่เชื่อมติดกันส่วนล่างจะมีจุดพิกเซลทั้งหมดจะต้องอยู่ใต้เส้นแบ่งองค์ประกอบที่เชื่อมติดกัน หลังจากทำการติดป้ายองค์ประกอบที่เชื่อมติดกันทั้งหมดเรียบร้อยแล้ว ทำการหาองค์ประกอบที่เชื่อมติดกันที่ขึ้นต่อกันระหว่างส่วนกลาง และส่วนบนกับส่วนล่าง โดยการดูส่วนองค์ประกอบเชื่อมติดกันส่วนประกอบที่ซ้อนทับกันในแนวนอน จากนั้นทำการเรียงลำดับขององค์ประกอบที่เชื่อมติดกันทั้งหมดใหม่โดยลำดับจาก องค์ประกอบที่เชื่อมติดส่วนกลาง องค์ประกอบที่เชื่อมติดส่วนล่าง และองค์ประกอบที่เชื่อมติดส่วนบน และทำการสร้างรูปภาพใหม่ตามลำดับโดยกำหนดให้แต่ละองค์ประกอบที่เชื่อมติดกันจะไม่ซ้อนทับกัน ขั้นตอนวิธีในการเลื่อนองค์ประกอบแนวตั้งแสดงดังรูปที่ 4-1 และภาพบรรทัดตัวอักษรก่อนและหลังการเลื่อนองค์ประกอบแนวตั้งแสดงดังรูปที่ 4-2 (ก) และ (ข) ตามลำดับ

Algorithm: Vertical Component Shifting

1. Find the connected components of the input image with 4-connectivity.
2. **For Each** connected component **DO**
   center_line = horizontal line from the highest density of black pixel in input image
   IF the component is above *center_line* **THEN**
   Label it as *upper_and_top_component*
   ELSE IF the component is below *center_line* **THEN**
   Label it as *lower_component*
   ELSE Label is as *base_component*.
3. **For Each** *lower_component* and *upper_and_top_component* **DO**
   3.1 Find the corresponding *base_component*
   3.2 Reorder the components to be a sequence of (*base_component*, *lower_component*, *upper_and_top_component*)
   3.3 Expand the image of vertically occurring with the sequence of non-overlapping reordered components.

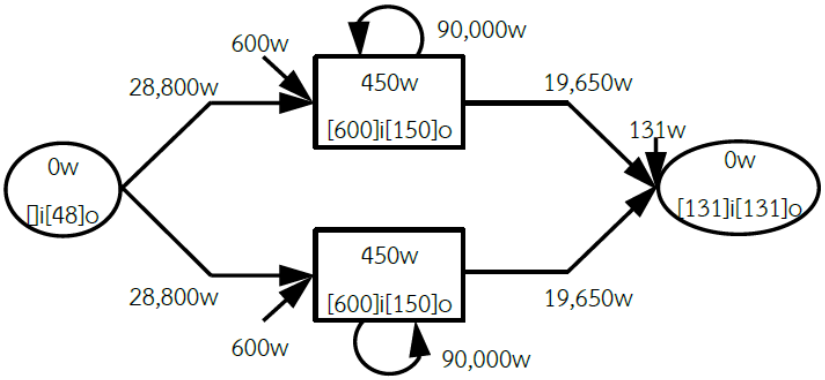
รูปที่ 4-1 ขั้นตอนวิธีการเลื่อนองค์ประกอบแนวตั้ง



รูปที่ 4-2 (ก) ภาพบรรทัดตัวอักษรก่อนเลื่อนองค์ประกอบเชื่อมติดกัน
(ข) ภาพบรรทัดตัวอักษรหลังเลื่อนองค์ประกอบเชื่อมติดกัน

4.3 โครงข่ายแอลเอสทีเอ็ม

ภาพบรรทัดตัวอักษรที่ถูกจัดเตรียมมาแต่ละภาพจะถูกนำเข้าโดยวิธีการเลื่อนหน้าต่าง แต่ละเฟรมมีขนาด 48 พิกเซล ดังนั้นภาพบรรทัดตัวอักษรจะถูกปรับขนาดให้ความสูงของภาพเป็นไปตามขนาดของเฟรมหน้าต่าง ซึ่งสอดคล้องกับจำนวนบัพในข้อมูลชั้นนำเข้าของโครงข่ายแอลเอสทีเอ็ม และจากตัวอักษรที่เป็นไปได้ตามตารางที่ 2-1 จำนวน 129 ตัวอักษร รวมตัวอักษรการเว้นวรรคและตัววรรคระหว่างตัวอักษร (ซึ่งจะถูกใช้ในการแปลงข้อมูลนำออกจากโครงข่ายแอลเอสทีเอ็มเป็นตัวอักษรในการจำแนกประเภทการเชื่อมต่อตามเวลา) แสดงถึงจำนวนบัพในชั้นนำออกทั้งหมด 131 บัพ นอกจากนี้การปรับพารามิเตอร์ของโครงข่ายแอลเอสทีเอ็มให้เหมาะสมกับชุดข้อมูลภาษาไทยถือเป็นเรื่องสำคัญ สำหรับวิทยานิพนธ์นี้มีเวลาอย่างจำกัดเพราะการเรียนรู้ของโครงข่ายแอลเอสทีเอ็มใช้เวลานานจึงปรับพารามิเตอร์เพียงในส่วนของจำนวนบัพในชั้นซ่อน หรือจำนวนบล็อกหน่วยความจำ แบบจำลองในการทดลองแบ่งออกเป็นโครงข่ายแอลเอสทีเอ็มสองประเภทคือโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง จึงทดลองปรับจำนวนบล็อกหน่วยความจำ 50 100 150 200 และ 400 บัพ เพื่อการหาจำนวนบล็อกหน่วยความจำที่เหมาะสม การปรับจำนวนบล็อกหน่วยความจำโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางจำนวนที่ปรับเท่ากันในแต่ละทิศทางหมายถึง 50 100 150 200 และ 400 บล็อกหน่วยความจำในแต่ละทิศทาง ภาพแสดงการตั้งค่าโครงข่ายแอลเอสทีเอ็มสองทิศทางที่ประกอบด้วย บัพในชั้นนำเข้า 48 บัพ บล็อกหน่วยความจำ 150 บล็อกต่อทิศทาง และบัพในชั้นนำออก 131 บัพ แสดงดังรูปที่ 4-3



รูปที่ 4-3 โครงข่ายแอลเอสทีเอ็ม ประกอบไปด้วยชั้นนำเข้า 48 บัพ ชั้นนำออก 131 บัพ และชั้นซ่อนจำนวน 150 บัพจำนวน 2 ชุด (เรียนรู้แบบสองทิศทาง)

จากรูปที่ 4-3 ค่าถ่วงน้ำหนักจากชั้นนำเข้าไปยังชั้นซ่อนหรือแต่ละบล็อกหน่วยความจำมีจำนวน 28,800 ค่า ในแต่ละบล็อกหน่วยความจำ มีค่าตั้งจุดทำงาน (Bias) จำนวน 600 ค่า ค่าถ่วงน้ำหนักภายในจำนวน 450 ค่า ค่าถ่วงน้ำหนักบล็อกหน่วยความจำที่เชื่อมต่อกับตัวเองจำนวน 90,000 ค่า ค่าถ่วงน้ำหนักจากบล็อกหน่วยความจำไปยังชั้นนำออกจำนวน 19,650 ค่า และมีค่าตั้งจุดทำงานในชั้นนำออกจำนวน 131 ค่า ผลจากการปรับบล็อกหน่วยความจำจะถูกนำไปใช้สำหรับการทดลองที่ใช้แบบจำลองโครงข่ายแอลเอสทีเอ็ม

4.4 การทดลอง

สำหรับการทดลองในวิทยานิพนธ์นี้ถูกแบ่งออกเป็น 4 การทดลอง การทดลองเพื่อเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบนสองทิศทาง การทดลองเพื่อเปรียบเทียบการทำงานของโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง การทดลองเพื่อแสดงประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้ง เมื่อมีจำนวนข้อมูลฝึกมากเพียงพอ และการทดลองเพื่อวัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้ง กับซอฟต์แวร์ที่ได้รับความนิยม

4.4.1 โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง

การทดลองเพื่อเปรียบเทียบประสิทธิภาพการทำงานของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง ชุดข้อมูลที่ใช้เป็นภาพบรรทัดตัวอักษรจำนวน 4,298 บรรทัด หรือตัวอักษร 260,250 ตัว ชุดข้อมูลจัดทำขึ้นด้วยฟอนต์ภาษาไทยจำนวน 11 ฟอนต์ ชุดข้อมูลถูกบีบตัวอักษรให้ใกล้กันเพื่อให้เกิดจุดเชื่อมต่อระหว่างตัวอักษรมากขึ้นเพื่อให้เกิดการเชื่อมต่อกันของตัวอักษร นอกจากนี้เพื่อดูประสิทธิภาพการทำงานของการเรียนรู้แบบไม่แบ่งส่วนของโครงข่ายแอลเอสทีเอ็ม ในการจัดทำชุดข้อมูลนั้นจึงบีบตัวอักษรให้ใกล้กันเพื่อให้เกิดจุดเชื่อมต่อระหว่างตัวอักษรมากขึ้น การทดลองใช้การตรวจสอบไขว้ 5 พับ (5-fold Cross-Validation) ชุดข้อมูลทั้งหมดถูกแบ่งเป็นชุดข้อมูลฝึก ชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบ แสดงรายละเอียด

จำนวนของชุดข้อมูลได้ดังตารางที่ 4-2 การทดลองนี้จะเลือกใช้บล็อกหน่วยความจำที่เหมาะสมในด้านเวลาและความถูกต้องสำหรับโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางจากการหาจำนวนบล็อกหน่วยความจำที่เหมาะสม ในแต่ละพับการสอนแบบจำลองด้วยชุดข้อมูลฝึกจำนวน 150,000 รอบ และใช้ชุดข้อมูลประเมินผลเพื่อเลือกแบบจำลองที่ดีที่สุดเพื่อนำไปใช้กับชุดข้อมูลทดสอบ

ตารางที่ 4-2 จำนวนชุดข้อมูลที่ใช้ในการทดลองการตรวจสอบไขว้ 5 พับ

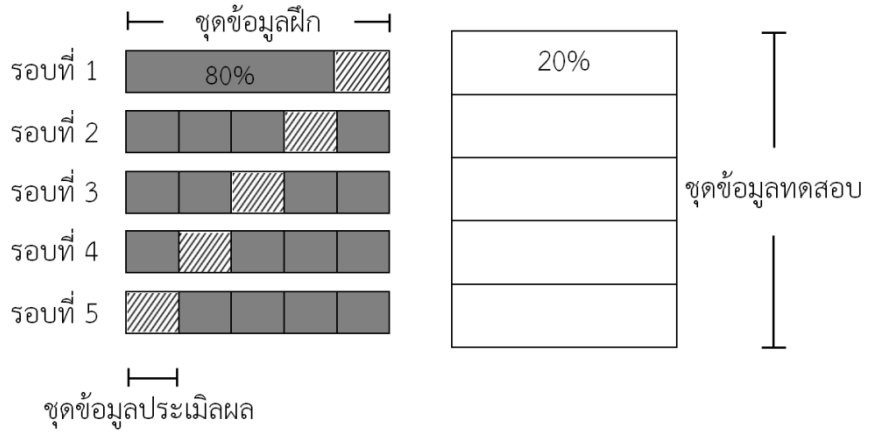
	พับที่ 1	พับที่ 2	พับที่ 3	พับที่ 4	พับที่ 5	
ชุดข้อมูลฝึก	2.744	2.735	2.744	2.753	2.746	จำนวนบรรทัด
	166.605	164.064	166.879	168.107	166.471	จำนวนตัวอักษร
ชุดข้อมูลประเมินผล	692	697	701	685	695	จำนวนบรรทัด
	41.636	41.891	43.736	39.507	42.104	จำนวนตัวอักษร
ชุดข้อมูลทดสอบ	862	866	853	860	857	จำนวนบรรทัด
	52.009	54.295	49.635	52.636	51.675	จำนวนตัวอักษร
รวม	4.298					จำนวนบรรทัด
	260.250					จำนวนตัวอักษร

4.4.2 โครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง

การทดลองนี้ต้องการเปรียบเทียบการทำงานของโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง เพื่อแสดงประสิทธิภาพของการใช้การเลื่อนองค์ประกอบแนวตั้ง การรู้จำตัวอักษรทั่วไปมักมีข้อมูลฝึกที่มีอยู่อย่างจำกัด แต่แบบจำลองที่ได้ควรสามารถรู้จำตัวอักษรในรูปแบบตัวอักษรที่เกิดขึ้นจำนวนมาก ดังนั้นในการทดลองนี้จึงเลือกชุดข้อมูลฝึกที่มีอยู่อย่างจำกัด และทดสอบด้วยชุดข้อมูลทดสอบจำนวนมากจัดทำขึ้นด้วยฟอนต์ภาษาไทยจำนวน 11 ฟอนต์ ชุดข้อมูลถูกบีบตัวอักษรให้ใกล้กันเพื่อให้เกิดจุดเชื่อมต่อระหว่างตัวอักษรมากขึ้น โดยการทำการทดลอง 5 รอบในแต่ละรอบจะใช้ชุดข้อมูลฝึกและชุดข้อมูลประเมินผลอัตราส่วน 80:20 ชุดข้อมูลทดสอบขนาดใหญ่ถูกแบ่งเป็น 5 ชุดข้อมูลเพื่อใช้ทดสอบการทดลองทั้ง 5 รอบ แสดงรายละเอียดจำนวนของชุดข้อมูลได้ดังตารางที่ 4-3 และแสดงการแบ่งชุดข้อมูลดังรูปที่ 4-4 เลือกจำนวนบล็อกหน่วยความจำที่เหมาะสมบนโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง ในแต่ละรอบของการทดลองสอนแบบจำลองด้วยชุดข้อมูลฝึกจำนวน 150,000 รอบ และใช้ชุดข้อมูลประเมินผลเพื่อเลือกแบบจำลองที่ดีที่สุดเพื่อนำไปใช้กับชุดข้อมูลทดสอบ ผลการทดลองจาก 5 รอบการทดลองจะถูกนำมาเฉลี่ย

ตารางที่ 4-3 จำนวนชุดข้อมูลฝึก ชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบในแต่ละรอบการทดลอง

	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5	
ชุดข้อมูลฝึก	143	143	143	143	143	จำนวนบรรทัด
	8,567	8,576	8,736	8,643	8,855	จำนวนตัวอักษร
ชุดข้อมูลประเมินผล	44	44	44	44	44	จำนวนบรรทัด
	2,051	2,045	2,036	1,975	1,763	จำนวนตัวอักษร
รวมชุดข้อมูลฝึก	183					จำนวนบรรทัด
และชุดข้อมูลประเมินผล	10,618					จำนวนตัวอักษร
ชุดข้อมูลทดสอบ	862	866	853	860	857	จำนวนบรรทัด
	52,009	54,295	49,635	52,636	51,675	จำนวนตัวอักษร
รวมชุดข้อมูลทดสอบ	4,298					จำนวนบรรทัด
	260,250					จำนวนตัวอักษร



รูปที่ 4-4 การทดลอง 5 รอบ ในแต่ละรอบ กล่องสีดำแสดงข้อมูลฝึก
กล่องลายขวางแสดงข้อมูลประเมินผล และกล่องสีขาวแสดงข้อมูลทดสอบ

4.4.3 โครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลขนาดใหญ่

การทดลองถูกจัดทำขึ้นเพื่อต้องการทราบว่าการรู้จำตัวอักษรภาษาไทยที่มีข้อมูลฝึกเพียงพอต่อรูปแบบการเกิดขึ้นของตัวอักษร แบบจำลองโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งและโครงข่ายแอลเอสทีเอ็มไม่ร่วมกับการเลื่อนองค์ประกอบแนวตั้งยังคงสามารถทำงานได้อย่างมีประสิทธิภาพ ชุดข้อมูลประกอบด้วยภาพบรรทัดตัวอักษร 4,298 บรรทัด หรือตัวอักษร 260,250 ตัว ชุดข้อมูลจัดทำขึ้นด้วยฟอนต์ภาษาไทยจำนวน 11 ฟอนต์ ชุดข้อมูลถูกบีบตัวอักษรให้ใกล้กันเพื่อให้เกิดจุดเชื่อมต่อระหว่างตัวอักษรมากขึ้นเพื่อให้เกิดการเชื่อมต่อกันของตัวอักษร ใช้การตรวจสอบไขว้ 5 พับ แสดงรายละเอียดจำนวนของชุดข้อมูลได้ดังตารางที่ 4-3 ตัวอักษรในชุดข้อมูลถูกบีบตัวอักษรให้ใกล้กันเพื่อให้เกิดจุดเชื่อมต่อระหว่างตัวอักษรมากขึ้น การทดลองเลือกโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางและตั้งค่าพารามิเตอร์โดยการปรับจำนวนบัพ 50 100 150 200 400 และ 800 บัพ บนการข้อมูลพับที่ 1 เท่านั้น และนำไปใช้กับชุดข้อมูลในทุกพับต่อไป การสอนและการเลือกใช้แบบจำลอง

ที่เหมาะสมทำโดยการสอนด้วยชุดข้อมูลฝึกจำนวน 150,000 รอบ และทดสอบด้วยชุดข้อมูลประเมินผลเพื่อเลือกแบบจำลองที่ดีที่สุดเพื่อนำไปใช้กับชุดข้อมูลทดสอบ

4.4.4 วัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ที่ได้รับความนิยม

หลังจากทำการทดลองเพื่อวัดประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งข้างต้นเรียบร้อยแล้ว จึงจัดทำการการทดลองเพื่อวัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ที่ได้รับความนิยม โดยเพิ่มจำนวนชุดข้อมูลและฟอนต์ชุดข้อมูลให้มีความหลากหลายมากขึ้น ชุดข้อมูลทั้งหมดประกอบด้วยภาพบรรทัดตัวอักษร 16,640 บรรทัด หรือตัวอักษร 998,195 ตัว จัดทำบนฟอนต์ 52 ฟอนต์ โดยแบ่งเป็นชุดข้อมูลฝึกและข้อมูลประเมินผล จำนวน 42 ฟอนต์ (อัตราส่วน 80:20) และชุดข้อมูลทดสอบ 10 ฟอนต์ แสดงจำนวนข้อมูลดังตารางที่ 4-4 การทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางร่วมกับการเลื่อนองค์ประกอบแนวตั้งทำโดยการสอนแบบจำลองโดยใช้ชุดข้อมูลฝึกจำนวน 300,000 รอบ และใช้ชุดข้อมูลประเมินผลเพื่อเลือกแบบจำลองที่ดีที่สุด มาทดสอบกับชุดข้อมูลทดสอบต่อไป นอกจากนี้เพื่อเปรียบเทียบกับซอฟต์แวร์ที่ได้รับความนิยมจึงนำชุดข้อมูลทดสอบมาทดสอบกับซอฟต์แวร์อย่าง ABBYY และ ArnThai สำหรับการทดลองใช้ ABBYY 12 ซึ่งเป็นโปรแกรมประยุกต์ที่ได้รับความนิยมมาก การตั้งค่าของ ABBYY เป็นค่าปริยายสำหรับภาษาไทยและภาษาอังกฤษ ArnThai 2.5 เป็นอีกหนึ่งโปรแกรมประยุกต์ที่สามารถแปลงจากภาพตัวอักษรเป็นข้อความที่สามารถนำไปแก้ไขสำหรับภาษาไทย

ตารางที่ 4-4 จำนวนชุดข้อมูลฝึก ชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบ

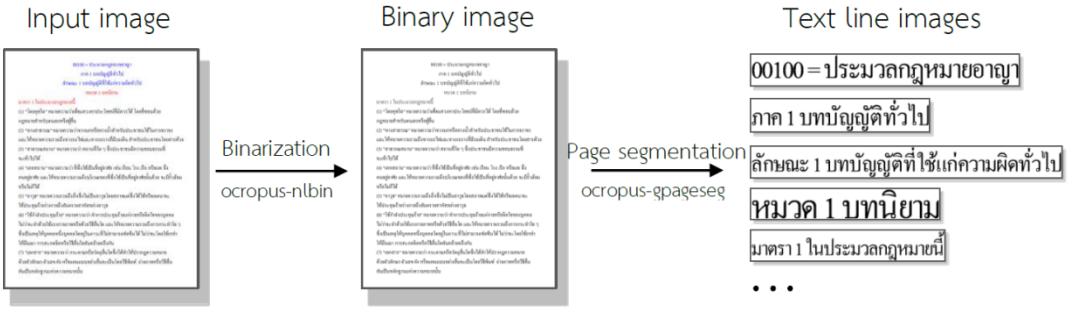
	จำนวนฟอนต์	จำนวนบรรทัด	จำนวนตัวอักษร
ชุดข้อมูลฝึก (80%)	42	10,752	636,964
ชุดข้อมูลประเมินผล (20%)		2,688	161,806
ชุดข้อมูลทดสอบ	10	3,200	199,425
รวม	**52**	**16,640**	**998,195**

บทที่ 5
ผลการดำเนินงาน

จากแนวทางและวิธีดำเนินงานที่กล่าวไปในบทที่ผ่านมา ในบทนี้จะกล่าวถึงผลการดำเนินงานประกอบด้วย การเตรียมชุดข้อมูล การเลื่อนองค์ประกอบแนวตั้ง ผลการทดลอง และการวิเคราะห์ผลในแต่ละการทดลอง พร้อมแสดงตัวอย่างการแปลงภาพบรรทัดตัวอักษรเป็นข้อความนำออก

5.1 การเตรียมชุดข้อมูล

จากภาพข้อมูลนำเข้าเป็นภาพหน้ากระดาษตัวอักษรที่ความละเอียด 300 dpi ในลักษณะภาพสี แต่ละภาพจะผ่านการแปลงภาพขาวดำโดยใช้ค่าขีดแบ่งภาพขาวดำ 0.2 และทำปรับเอียงหน้ากระดาษ โดยใช้เครื่องมือ ocropus-nlbin จากนั้นนำภาพขาวดำที่ได้ผ่านเครื่องมือแบ่งส่วนภาพบรรทัดตัวอักษร ocropus-gpageseg จะได้ภาพบรรทัดตัวอักษรจำนวนหนึ่งที่มีลักษณะเป็นภาพขาวดำ แสดงการทำงานดังรูปที่ 5-1

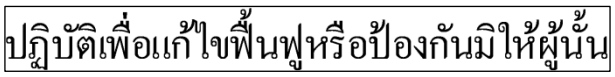


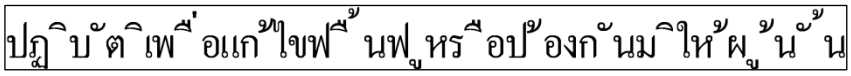
รูปที่ 5-1 การแปลงภาพข้อมูลนำเข้าหน้ากระดาษตัวอักษรเป็นภาพบรรทัดตัวอักษร

ในการจัดทำชุดข้อมูลภาพบรรทัดตัวอักษรโดยวิธีดังกล่าวสามารถทำกับภาพหน้ากระดาษตัวอักษรในทุกภาพ ภาพบรรทัดตัวอักษรที่ได้จะถูกนำไปใช้เป็นข้อมูลนำเข้าของโครงข่ายแอลเอสทีเอ็มต่อไป แต่สำหรับการทดลองที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย ภาพบรรทัดตัวอักษรในชุดข้อมูลนำเข้าทุกภาพจะถูกนำมาผ่านขั้นตอนวิธีการเลื่อนองค์ประกอบแนวตั้งในหัวข้อถัดไป

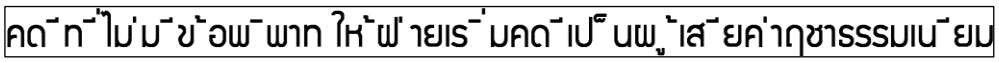
5.2 การเลื่อนองค์ประกอบแนวตั้ง

ภาพบรรทัดตัวอักษรที่เป็นข้อมูลนำเข้าของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งจะถูกจัดการให้ภาพเหมาะสมกับการเรียนรู้ของโครงข่ายแอลเอสทีเอ็ม โดยจัดลำดับองค์ประกอบที่เชื่อมติดกันให้ตัวอักษรที่เกิดขึ้นในแนวตั้งมีจำนวนน้อยลง ภาพตัวอย่างบรรทัดตัวอักษรก่อนและหลังการเลื่อนองค์ประกอบแนวตั้งแสดงดังรูปที่ 5-2

(ก.1) 

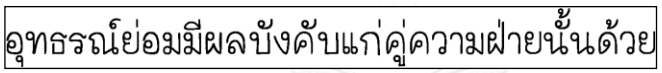
(ก.2) 

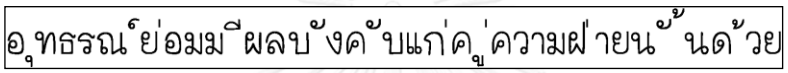
(ข.1)

(ข.2) 

(ค.1)

(ค.2)

(ง.1) 

(ง.2) 

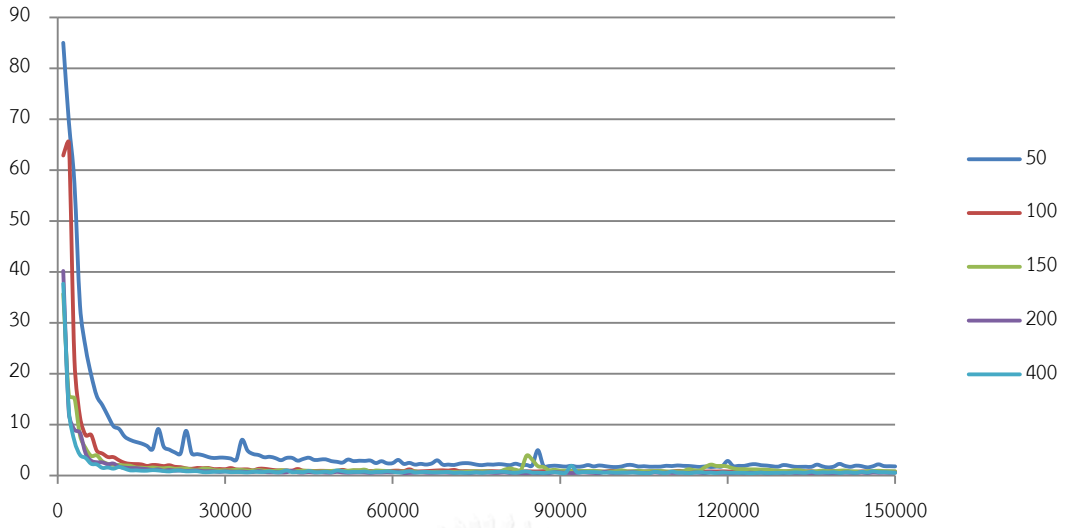
รูปที่ 5-2 ภาพบรรทัดตัวอักษร (ก.1) (ข.1) (ค.1) และ (ง.1) เป็นภาพก่อนเลื่อนองค์ประกอบแนวตั้ง และ (ก. 2) (ข.2) (ค.2) และ (ง.2) เป็นภาพหลังเลื่อนองค์ประกอบแนวตั้ง

จากผลการทดลองการเลื่อนองค์ประกอบแนวตั้ง พบว่าภาพบรรทัดตัวอักษรส่วนใหญ่สามารถเลื่อนองค์ประกอบแนวตั้งได้อย่างถูกต้องตามที่ต้องการในชุดข้อมูลทั้งหมด ปัญหาที่พบเกิดขึ้นกับภาพบรรทัดตัวอักษรที่มีตัวอักษรเชื่อมติดกันในแนวตั้ง ทำให้องค์ประกอบที่เชื่อมติดกันนั้นไม่ถูกเลื่อนไปตามขั้นตอนวิธี ทั้งนี้การเรียนรู้ของโครงข่ายแอลเอสทีเอ็มในส่วนนี้จะเรียนรู้เหมือนกับโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย

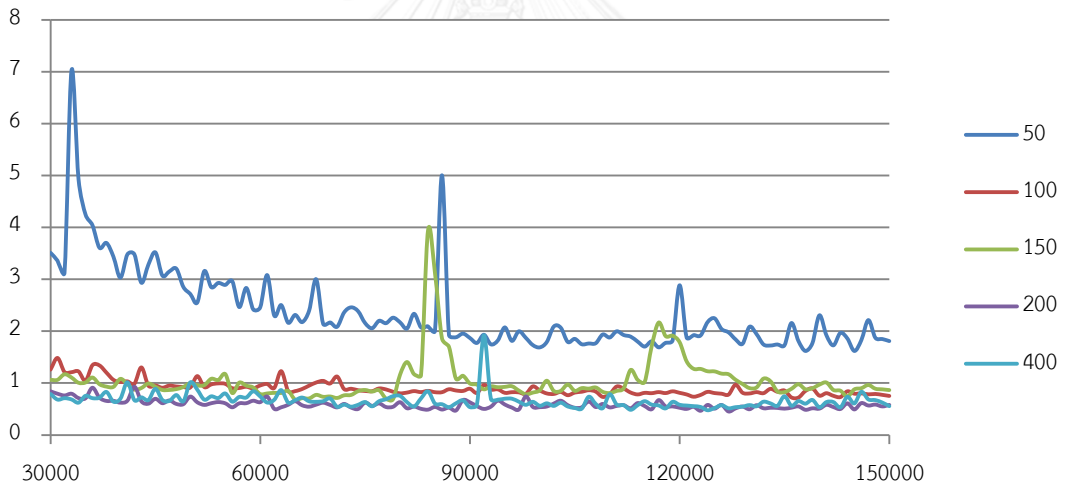
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5.3 โครงข่ายแอลเอสทีเอ็ม

หลังจากจัดเตรียมชุดข้อมูลภาพบรรทัดตัวอักษรสำหรับการแต่ละการทดลองเรียบร้อยแล้ว การปรับพารามิเตอร์ของโครงข่ายแอลเอสทีเอ็มให้เหมาะสมกับชุดข้อมูลภาษาไทยในแต่ละการทดลอง การปรับค่าพารามิเตอร์ทั้งหมดให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มใช้เวลาจำนวนมาก และด้วยการจัดทำวิทยานิพนธ์นี้มีเวลาที่จำกัดจึงปรับเพียงจำนวนของบัพในชั้นซ่อน หรือจำนวนบล็อกหน่วยความจำให้เหมาะสมในแต่ละการทดลองเท่านั้น แสดงการปรับพารามิเตอร์การรู้จำตัวอักษรภาษาไทยโดยแสดงผลการทดลองเป็นค่าความผิดพลาดแสดงโดยใช้ค่าระยะทางเลเวนชเตย์น บนชุดข้อมูลพับที่ 1 ของการทดลองที่ 4.4.1 โดยทดลองปรับบล็อกหน่วยความจำบนโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและแบบสองทิศทาง ผลการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลแบ่งเป็นช่วงการเรียนรู้ในช่วงต้นที่เรียนรู้อย่างรวดเร็วดังรูปที่ 5-3 และตัดเพียงช่วง 30,000 จนถึง 150,000 รอบ เพื่อแสดงให้เห็นการเรียนรู้ชัดเจนมากยิ่งขึ้นในดังรูปที่ 5-4



รูปที่ 5-3 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวบนชุดข้อมูลภาษาไทย

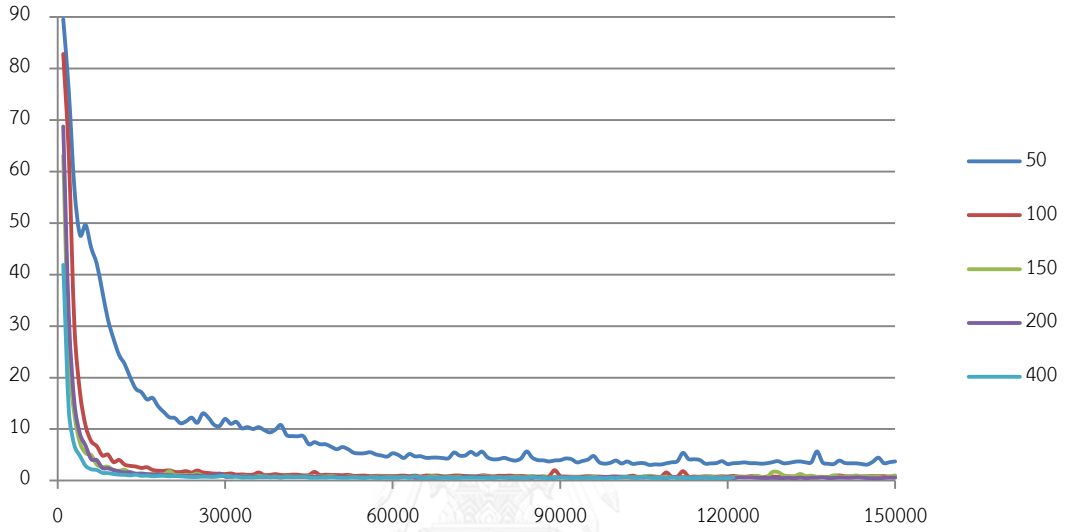


รูปที่ 5-4 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวบนชุดข้อมูลภาษาไทย แสดงเพียงช่วง 30,000 ถึง 150,000 รอบ

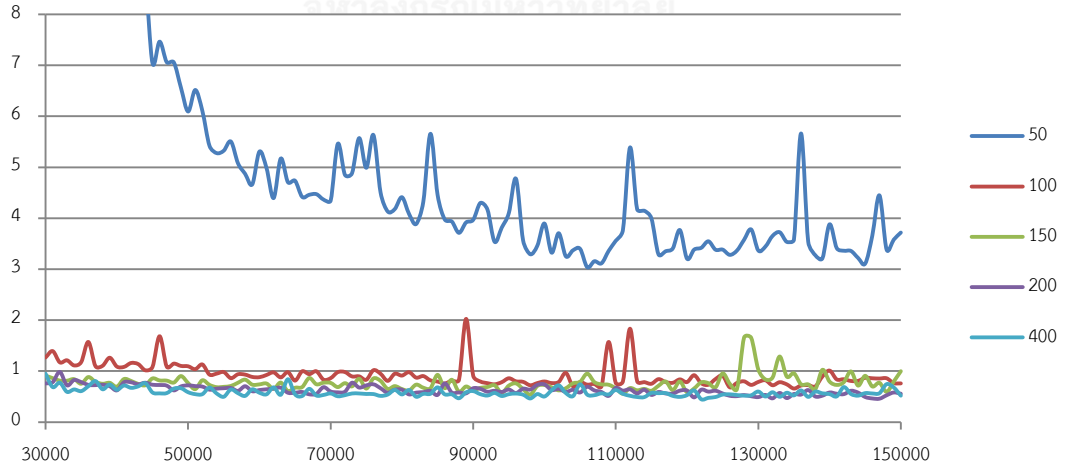
จากการทดลองการปรับค่าจำนวนบล็อกหน่วยความจำที่เหมาะสมสำหรับโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว สามารถสรุปได้ว่าบล็อกหน่วยความจำที่ 50 100 และ 150 นั้นไม่สามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดได้ ในขณะที่บล็อกหน่วยความจำ 200 และ 400 สามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดได้ เพียงแต่ใช้เวลาในการลู่เข้าที่แตกต่างกัน โดยใช้เวลาในการลู่เข้าแปรผันตรงกับจำนวนบล็อกหน่วยความจำ หรือถ้าจำนวนบล็อกจำนวนมากก็สามารถลู่เข้าได้เร็วกว่าจำนวนน้อย แต่ทั้งนี้ใช้เวลาในการเรียนรู้แต่ละรอบเพิ่มขึ้นในระดับเอก

โพเนนเชียล จึงเลือกจำนวนบล็อกเป็น 200 เพราะสามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดและใช้เวลาเหมาะสม

สำหรับผลการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง แสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลแบ่งเป็นช่วงการเรียนรู้ในช่วงต้นที่เรียนรู้อย่างรวดเร็วดังรูปที่ 5-5 และตัดเพียงช่วง 10,000 จนถึง 150,000 รอบ เพื่อแสดงให้เห็นการเรียนรู้ชัดเจนขึ้นในดังรูปที่ 5-6



รูปที่ 5-5 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางบนชุดข้อมูลภาษาไทย



รูปที่ 5-6 กราฟแสดงค่าความผิดพลาดในการทดลองปรับจำนวนบล็อกหน่วยความจำของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางบนชุดข้อมูลภาษาไทย แสดงเพียงช่วง 30,000 ถึง 150,000 รอบ

จากการทดลองการปรับค่าจำนวนบล็อกหน่วยความจำที่เหมาะสมสำหรับโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง สามารถสรุปได้ว่าบล็อกหน่วยความจำที่ 50 และ 100 นั้นไม่สามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดได้ ในขณะที่บล็อกหน่วยความจำ 150 200 และ 400 สามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดได้เพียงแต่ใช้เวลาในการลู่เข้าที่แตกต่างกัน โดยใช้เวลาในการลู่เข้าแปรผันตรงกับจำนวนบล็อกหน่วยความจำ หรือถ้าจำนวนบล็อกจำนวนมากก็สามารถลู่เข้าได้เร็วกว่าจำนวนน้อย แต่ทั้งนี้ใช้เวลาในการเรียนรู้แต่ละรอบเพิ่มขึ้นในระดับเอกโพเนนเชียล จึงเลือกจำนวนบล็อกเป็น 200 เพราะสามารถลู่เข้าช่วงค่าความผิดพลาดที่น้อยที่สุดและใช้เวลาเหมาะสม

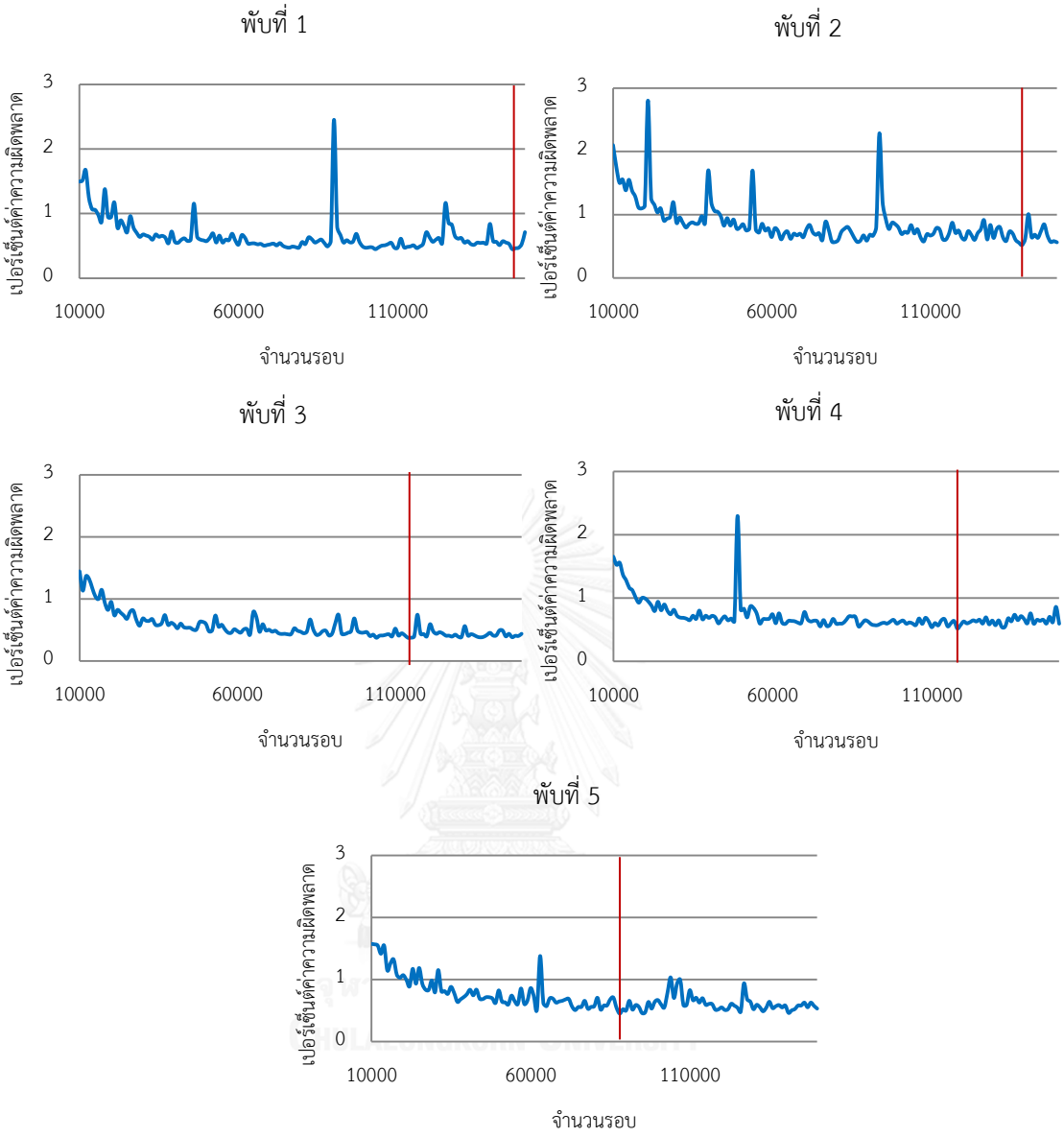
5.4 ผลการทดลองและการวิเคราะห์ผลการทดลอง

ผลการทดลองและการวิเคราะห์ผลของการทดลองปรับจำนวนบล็อคหน่วยความจำ สำหรับการทดลองที่ใช้โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวจะใช้บล็อกหน่วยความจำเป็น 200 และโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางจะใช้บล็อกหน่วยความจำเป็น 150 ในแต่ละทิศทาง ในแต่ละการทดลองต่อไปนี้คำนวณเปอร์เซ็นต์ค่าความผิดพลาดจากระยะทางเลเวนชเตย์น

5.4.1 โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง

ในการทดลองการตรวจสอบไขว้ 5 พับ แบบจำลองที่นำมาใช้ในการทดสอบกับชุดข้อมูลทดสอบในแต่ละพับมาจากการสอนแบบจำลองด้วยชุดข้อมูลฝึกและเลือกแบบจำลองที่ดีที่สุดโดยใช้ชุดข้อมูลประเมินผล กราฟแสดงค่าความผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางที่สอนด้วยชุดข้อมูลฝึกและทดสอบด้วยชุดข้อมูลประเมินผล กราฟที่แสดงตัดเพียงช่วง 10,000 รอบ ถึง 150,000 รอบ เพื่อแสดงให้เห็นเปอร์เซ็นต์ค่าความผิดพลาดชัดเจนในช่วงของการเรียนรู้ทั้ง 5 พับ แสดงได้ดังรูปที่ 5-7 และรูปที่ 5-8 ตามลำดับ ผลการทดลองเป็นจำนวนรอบในการสอนที่ดีที่สุด จำนวนตัวอักษรที่ผิดพลาด เปอร์เซ็นต์ค่าความผิดพลาด และเปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย ของแบบจำลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางบนชุดข้อมูลทดสอบในแต่ละพับได้ดังตารางที่ 5-1 และตารางที่ 5-2 ตามลำดับ

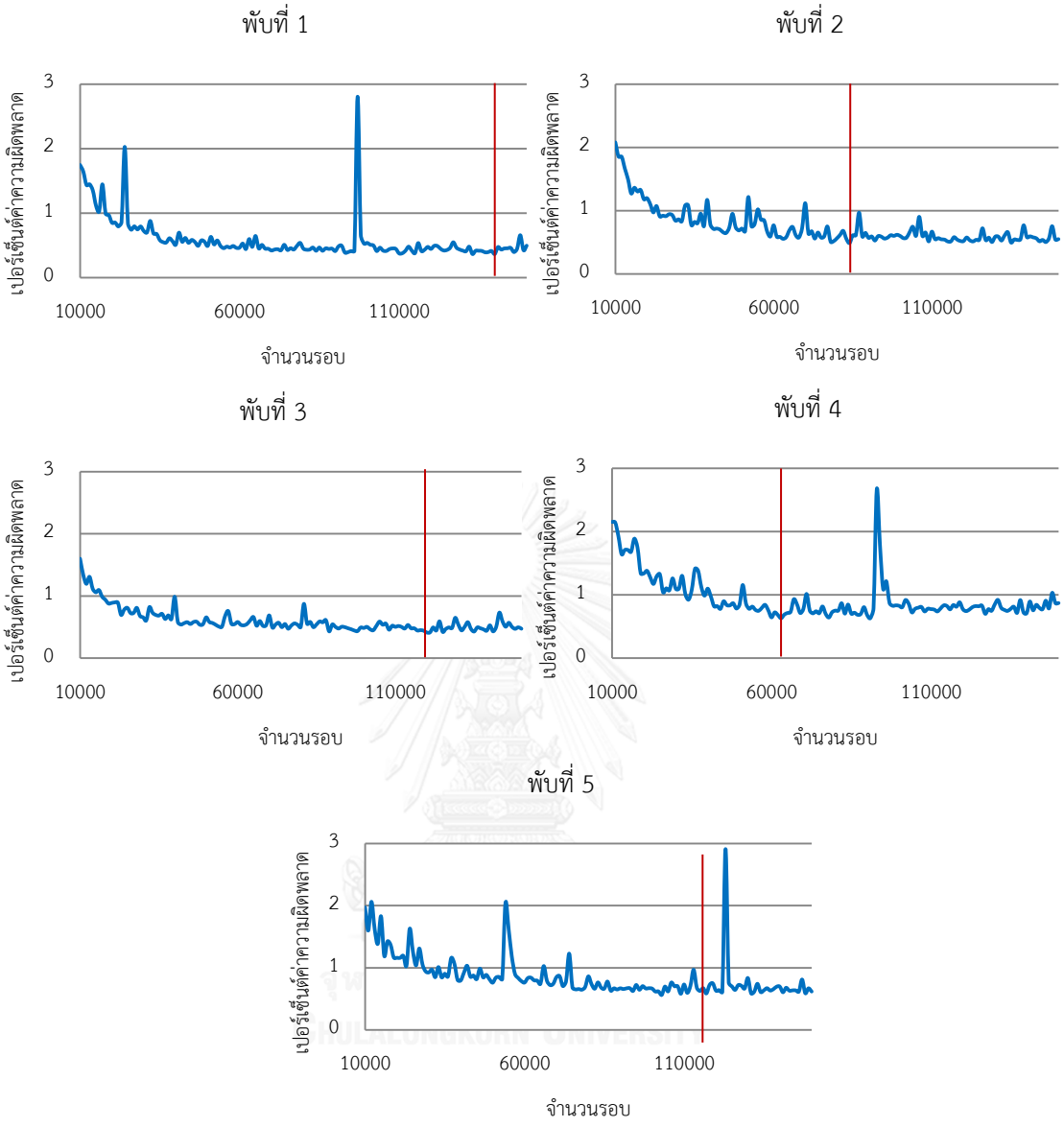
ผลการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางได้เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย 0.567% และ 0.563% ตามลำดับ และจากการสังเกตผลการทดลองในแต่ละพับพบว่า การเรียนรู้ของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวและโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางไม่มีความต่างกันอย่างมีนัยสำคัญ หรือกล่าวคือการเรียนรู้ของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวมีความสามารถเทียบได้กับโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง ในการรู้จำตัวอักษรพิมพ์ภาษาไทย ซึ่งในวิทยานิพนธ์นี้ได้เลือกใช้โครงข่ายแอลเอสทีเอ็มแบบสองทิศทางเพราะจากการทดลองใช้จำนวนบล็อกหน่วยความจำที่น้อยกว่า ส่งผลให้เวลาในการคำนวณน้อยกว่า และอ้างอิงจากงานวิจัยอื่นในบทที่ 3



รูปที่ 5-7 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียวทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

ตารางที่ 5-1 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว

	พับที่ 1	พับที่ 2	พับที่ 3	พับที่ 4	พับที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	146,000	139,000	114,000	118,000	88,000
จำนวนตัวอักษรที่ผิดพลาด	349	352	243	252	280
เปอร์เซ็นต์ค่าความผิดพลาด	0.671%	0.648%	0.49%	0.479%	0.542%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					0.567%



รูปที่ 5-8 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

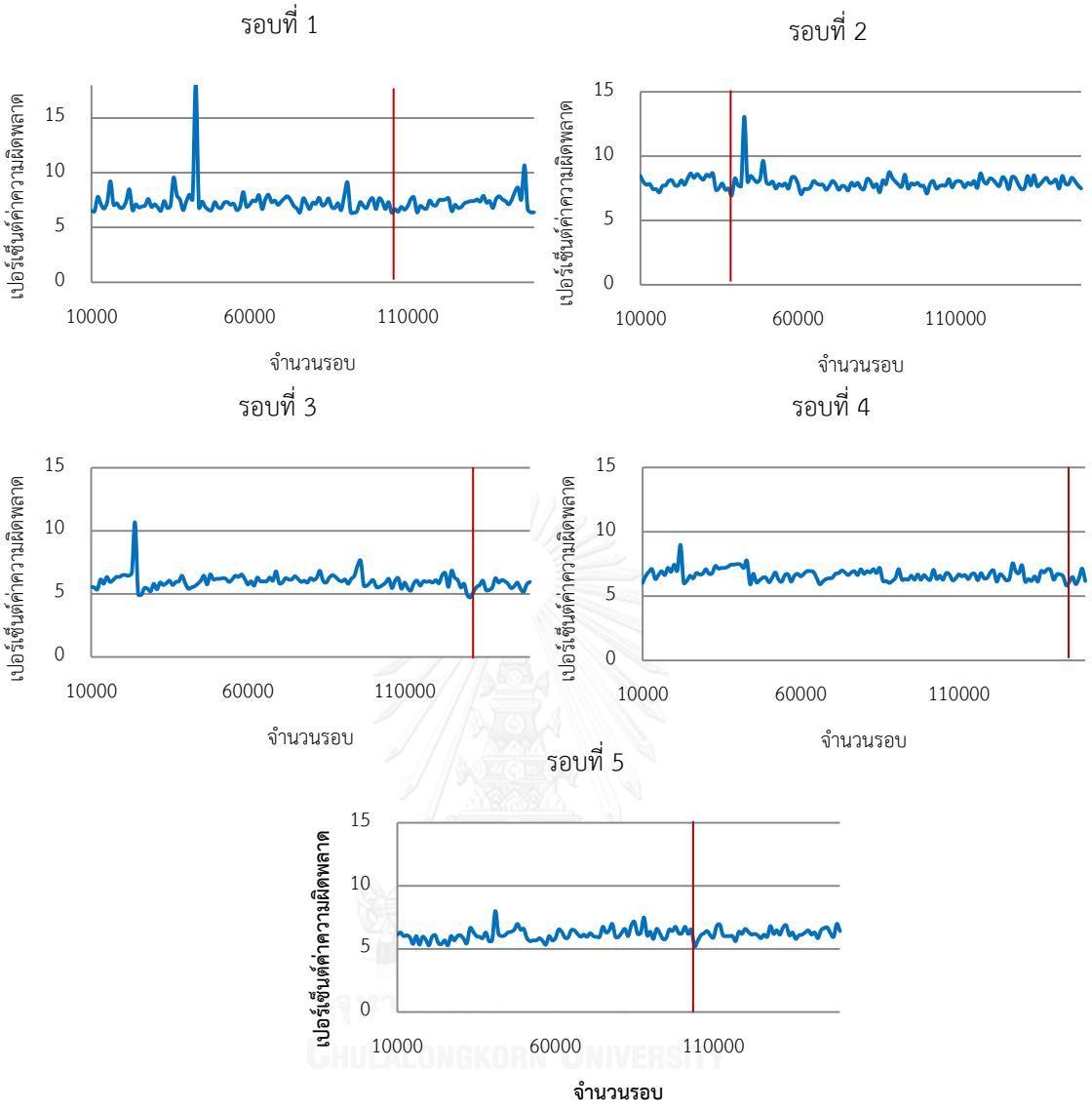
ตารางที่ 5-2 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง

	พับที่ 1	พับที่ 2	พับที่ 3	พับที่ 4	พับที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	140,000	84,000	120,000	63,000	115,000
จำนวนตัวอักษรที่ผิดพลาด	296	328	252	266	324
เปอร์เซ็นต์ค่าความผิดพลาด	0.569%	0.604%	0.508%	0.505%	0.627%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					0.563%

5.4.2 โครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง

หลังจากเลือกทำการทดลองปรับจำนวนบล็อกหน่วยความจำ การเรียนรู้แบบทิศทางเดียวและสองทิศทางของโครงข่ายแอลเอสทีเอ็มแล้ว จึงทำการทดลองเพื่อทดสอบประสิทธิภาพของขั้นตอนวิธีการเลื่อนองค์ประกอบแนวตั้ง โดยการนำโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง ในแต่ละทิศทางมีบล็อกหน่วยความจำจำนวน 150 ผลการทดลองแสดงเป็นกราฟเปอร์เซ็นต์ค่าความผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง ที่สอนด้วยชุดข้อมูลฝึกและทดสอบด้วยชุดข้อมูลประเมินผล กราฟที่แสดงตัดเพียงช่วง 10,000 รอบ ถึง 150,000 รอบ เพื่อแสดงให้เห็นเปอร์เซ็นต์ค่าความผิดพลาดชัดเจนในช่วงของการเรียนรู้ทั้ง 5 รอบ แสดงได้ดังรูปที่ 5-9 และรูปที่ 5-10 ตามลำดับ แสดงผลการทดลองเป็นจำนวนรอบในการสอนที่ดีที่สุด จำนวนตัวอักษรที่ผิดพลาด เปอร์เซ็นต์ค่าความผิดพลาด และเปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย ของแบบจำลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งและแบบใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลทดสอบในแต่ละรอบได้ดังตารางที่ 5-3 และตารางที่ 5-4 ตามลำดับ

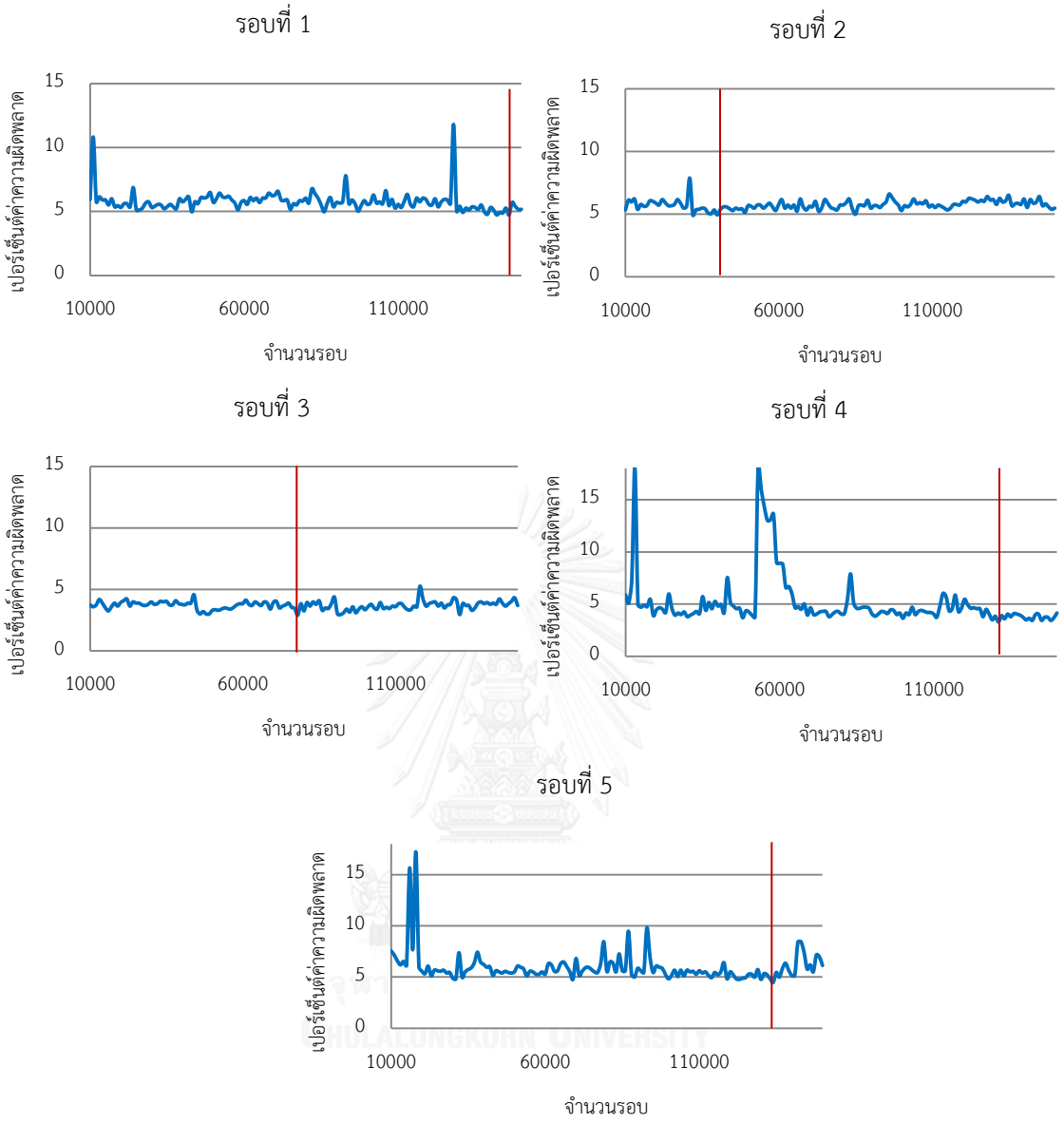
ผลการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งและแบบใช้การเลื่อนองค์ประกอบแนวตั้งได้ค่าความผิดพลาดเฉลี่ย 7.012% และ 5.024% ตามลำดับ ในการใช้การเลื่อนองค์ประกอบแนวตั้ง ได้ค่าความผิดพลาดเฉลี่ยต่ำกว่าแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งในทุกรอบการทดลองและต่างกันอย่างมีนัยสำคัญ สังเกตจากกราฟ รอบที่ 1 ของโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งมีเปอร์เซ็นต์ค่าความผิดพลาดส่วนใหญ่อยู่ระหว่างช่วง 6-7% แต่ในขณะที่กราฟ รอบที่ 1 ของโครงข่ายแอลเอสทีเอ็มแบบแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยมีค่าอยู่ระหว่าง 5-6% ตลอดการทดสอบข้อมูลเป็นต้น โดยตัวอักษรที่ผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งส่วนใหญ่เกิดจากรูปแบบตัวอักษรที่เกิดขึ้นในแนวตั้งและรูปแบบตัวอักษรที่ไม่เคยพบในชุดข้อมูลฝึกจำนวนมาก ในขณะที่โครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยส่วนใหญ่เกิดจากการเชื่อมต่อกันของตัวอักษรในแนวตั้งและรูปแบบตัวที่ไม่เคยพบในชุดข้อมูลฝึกจำนวนหนึ่งแต่น้อยกว่าแบบข้างต้น เนื่องจากการเลื่อนองค์ประกอบแนวตั้งสามารถช่วยลดรูปแบบตัวอักษรไปได้ส่วนหนึ่ง



รูปที่ 5-9 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งทั้ง 5 รอบ ในรอบที่ 1,000 ถึง 150,000
เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

ตารางที่ 5-3 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง

	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	105,000	39,000	131,000	144,000	104,000
จำนวนตัวอักษรที่ผิดพลาด	3,263	4,080	3,781	3,665	3,459
เปอร์เซ็นต์ค่าความผิดพลาด	6.274%	7.515%	7.618%	6.963%	6.694%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					7.012%



รูปที่ 5-10 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็ม แบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งทั้ง 5 รอบ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

ตารางที่ 5-4 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย

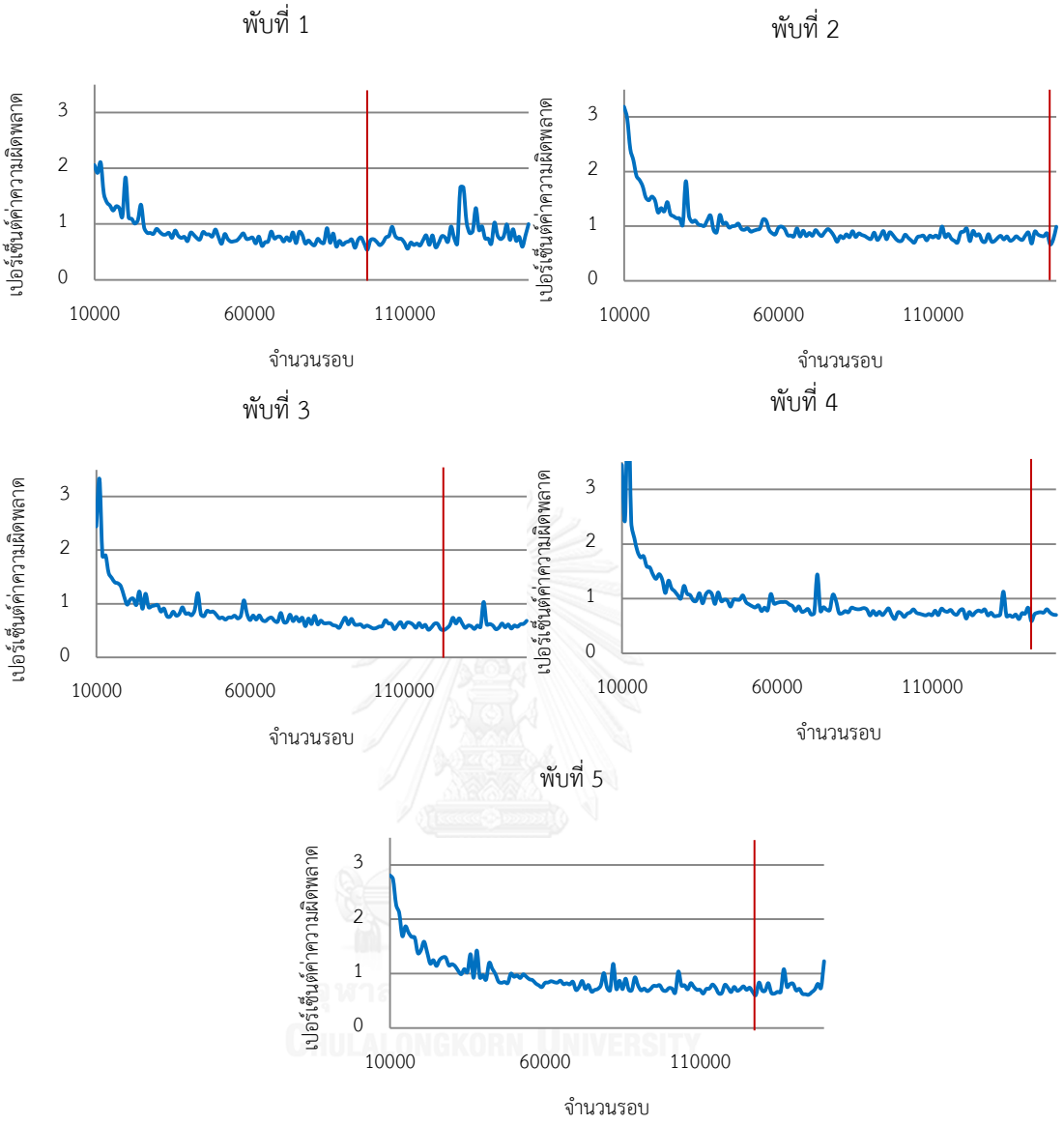
	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	146,000	40,000	78,000	131,000	134,000
จำนวนตัวอักษรที่ผิดพลาด	2,853	2,634	2,361	2,332	2,896
เปอร์เซ็นต์ค่าความผิดพลาด	5.486%	4.851%	4.757%	4.430%	5.604%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					5.024%

5.4.3 โครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลขนาดใหญ่

จากการทดลองที่ 4.4.2 พบว่าการเลื่อนองค์ประกอบแนวตั้งสามารถทำงานได้ดีกับชุดข้อมูลฝึกที่มีอยู่อย่างจำกัด และสามารถเพิ่มประสิทธิภาพการทำงานโครงข่ายแอลเอสทีเอ็มได้ สำหรับการทดลองนี้ต้องการทราบว่าโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยยังสามารถทำงานได้ดีในชุดข้อมูลฝึกที่มีขนาดใหญ่หรือชุดข้อมูลฝึกที่เพียงพอต่อการเรียนรู้ด้วย โดยใช้โครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง ในแต่ละทิศทางมีบล็อกหน่วยความจำจำนวน 150 และใช้การตรวจสอบไขว้ 5 พับ ผลการทดลองแสดงเป็นกราฟเปอร์เซ็นต์ค่าความผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งและแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้ง ที่สอนด้วยชุดข้อมูลฝึกและทดสอบด้วยชุดข้อมูลประเมินผล กราฟที่แสดงตัดเพียงช่วง 10,000 รอบ ถึง 150,000 รอบ เพื่อแสดงให้เห็นเปอร์เซ็นต์ค่าความผิดพลาดชัดเจนในช่วงของการเรียนรู้ทั้ง 5 รอบ แสดงได้ดังรูปที่ 5-11 และรูปที่ 5-12 ตามลำดับ แสดงผลการทดลองเป็นจำนวนรอบในการสอนที่ดีที่สุด จำนวนตัวอักษรที่ผิดพลาด เปอร์เซ็นต์ค่าความผิดพลาด และเปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย ของแบบจำลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งและแบบใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลทดสอบในแต่ละพับได้ดังตารางที่ 5-5 และตารางที่ 5-6 ตามลำดับ

ผลการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งและแบบใช้การเลื่อนองค์ประกอบแนวตั้งได้ค่าความผิดพลาดเฉลี่ย 0.607% และ 0.627% ตามลำดับ ในแต่ละพับนั้นในการทดลองให้ค่าความผิดพลาดเฉลี่ยไม่ต่างกันอย่างมีนัยสำคัญ ซึ่งกล่าวคือการเลื่อนองค์ประกอบแนวตั้งยังคงสามารถทำงานได้ดีบนชุดข้อมูลฝึกขนาดใหญ่ ตัวอักษรที่ผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งยังคงเป็นตัวอักษรที่เกิดขึ้นในแนวตั้ง ตัวอย่างคำที่ผิดเช่น “ยู้” จากคำว่า “ผู้” “ใหญ้” จากคำว่า “ใหญ่” “ก่ง” จากคำว่า “กิ่ง” และ “เ◌◌◌อง” จากคำว่า “เรื่อง” เป็นต้น ในขณะที่ตัวอักษรที่ผิดพลาดของโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยเป็นตัวอักษรที่เชื่อมติดกันในแนวตั้ง

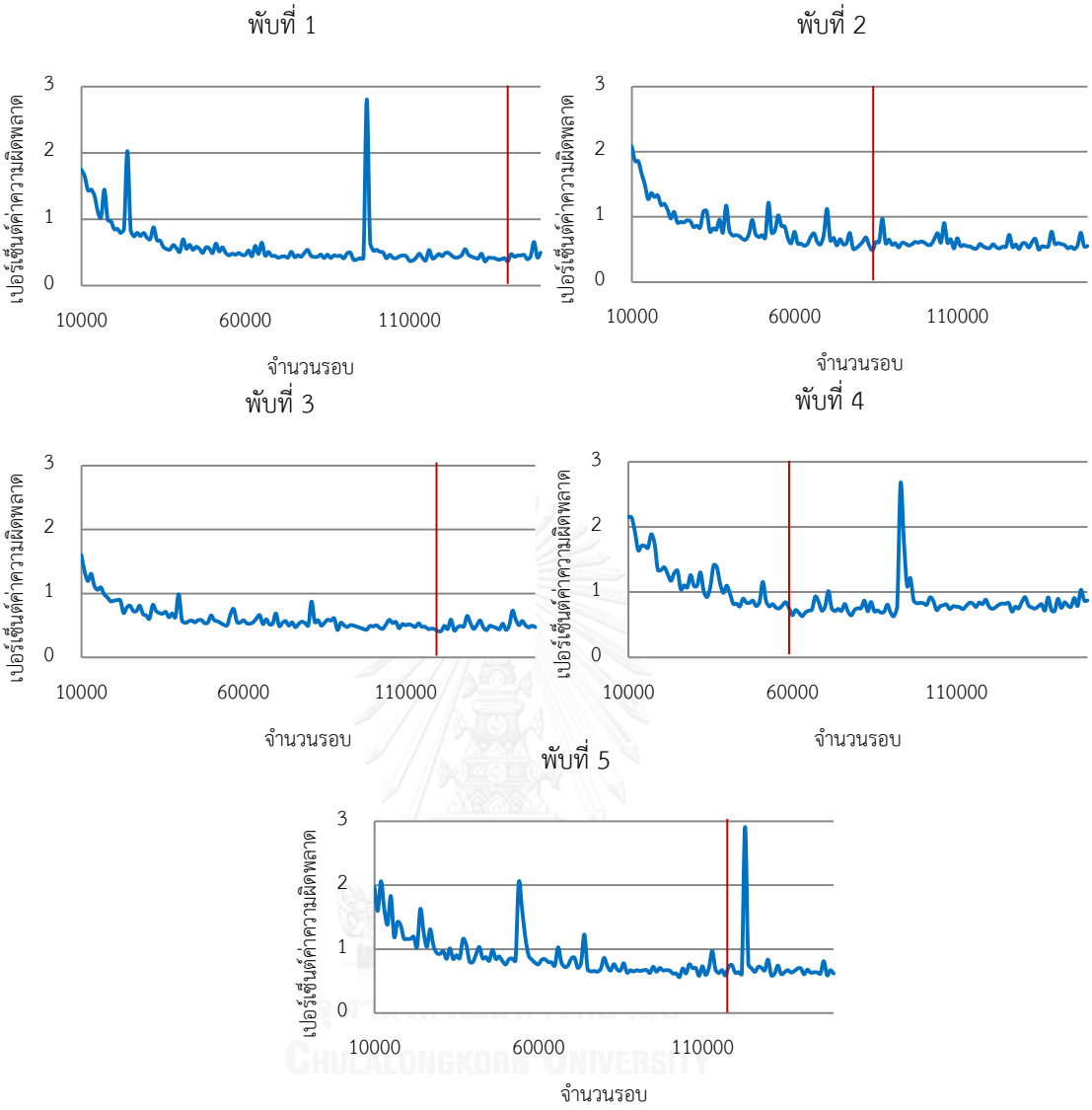
จากการทดลองที่ 4.4.2 และ 4.4.3 พบว่าโครงข่ายแอลเอสทีเอ็มที่ทำงานร่วมกับการใช้การเลื่อนองค์ประกอบแนวตั้งสามารถเรียนรู้ลักษณะบรรทัดตัวอักษรภาษาไทยที่ตัวอักษรถูกบีบลดให้เกิดตัวอักษรที่เชื่อมต่อกันได้ทั้งบนชุดข้อมูลฝึกขนาดจำกัดและชุดข้อมูลฝึกที่มีรูปแบบตัวอักษรเพียงพอ นอกจากนี้ผลการทดสอบกับชุดข้อมูลทดสอบได้ความแม่นยำที่ดี เพื่อแสดงประสิทธิภาพการเรียนรู้ของโครงข่ายแอลเอสทีเอ็มร่วมกับการการเลื่อนองค์ประกอบแนวตั้งเราจึงทำการทดลองเปรียบเทียบระหว่างโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งและซอฟต์แวร์ที่ได้รับความนิยม ซึ่งจะกล่าวต่อไปในหัวข้อต่อไป



รูปที่ 5-11 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลฝึกขนาดใหญ่ทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

ตารางที่ 5-5 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบไม่ใช้การเลื่อนองค์ประกอบแนวตั้งบนชุดข้อมูลฝึกขนาดใหญ่

	พับที่ 1	พับที่ 2	พับที่ 3	พับที่ 4	พับที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	98,000	148,000	123,000	142,000	128,000
จำนวนตัวอักษรที่ผิดพลาด	331	346	285	334	283
เปอร์เซ็นต์ค่าความผิดพลาด	0.636%	0.637%	0.574%	0.635%	0.548%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					0.607%



รูปที่ 5-12 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองโครงข่ายแอลเอสทีเอ็ม แบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยบนชุดข้อมูลฝึกขนาดใหญ่ทั้ง 5 พับ ในรอบที่ 1,000 ถึง 150,000 เส้นตั้งสีแดงแสดงจำนวนรอบที่ดีที่สุด

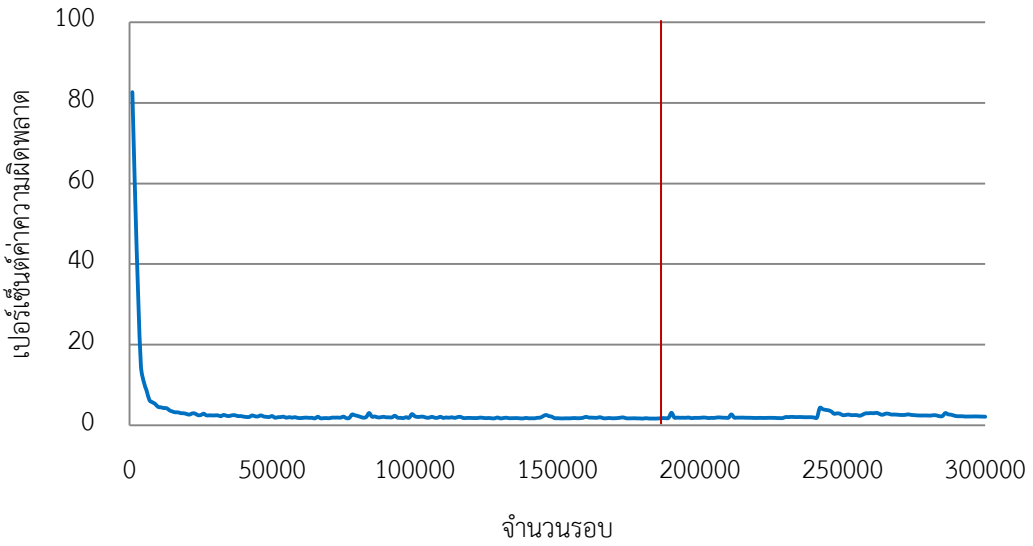
ตารางที่ 5-6 ผลการทดลองแสดงค่าความผิดพลาดของการทดลองโครงข่ายแอลเอสทีเอ็มแบบใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยบนชุดข้อมูลฝึกขนาดใหญ่

	พับที่ 1	พับที่ 2	พับที่ 3	พับที่ 4	พับที่ 5
จำนวนรอบในการสอนที่ดีที่สุดของแบบจำลอง	140,000	84,000	120,000	63,000	115,000
จำนวนตัวอักษรที่ผิดพลาด	296	328	252	266	324
เปอร์เซ็นต์ค่าความผิดพลาด	0.569%	0.604%	0.508%	0.505%	0.627%
เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย					0.563%

5.4.4 วัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับซอฟต์แวร์ที่ได้รับความนิยม

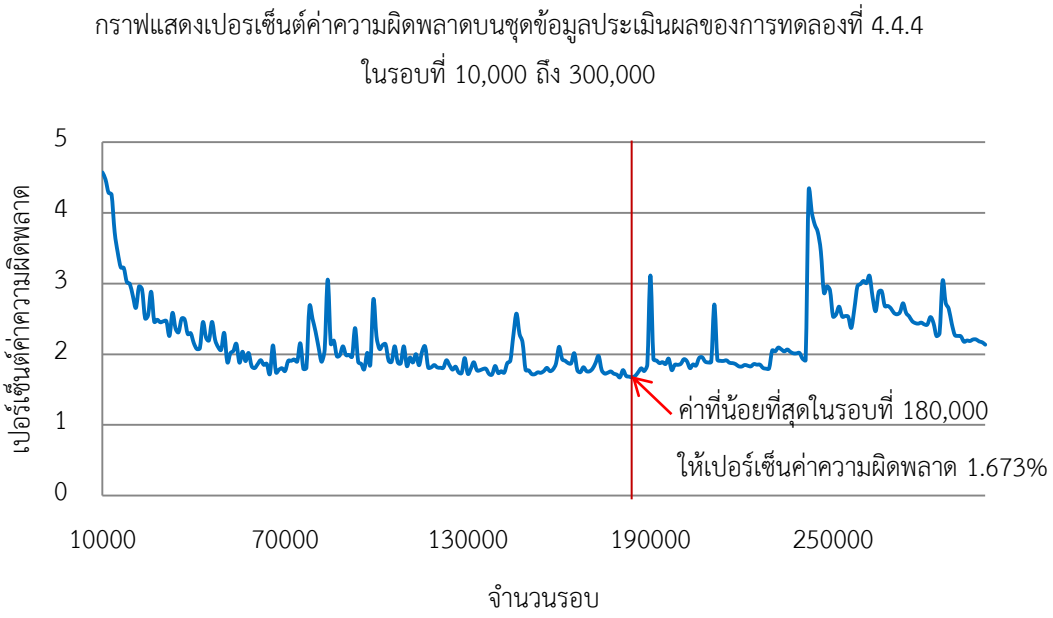
การวัดและเปรียบเทียบประสิทธิภาพของโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งกับ ABBYY และ ArnThai โครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบแนวตั้งที่นำมาใช้ถูกทดลองเพื่อหาแบบจำลองที่เหมาะสมโดยใช้ข้อมูลฝึกและข้อมูลประเมินผล ผลค่าความผิดพลาดบนข้อมูลประเมินผลได้เปอร์เซ็นต์ค่าความผิดพลาดที่ต่ำที่สุดอยู่ที่ 1.673% ตรงกับรอบที่ 180,000 แสดงกราฟเปอร์เซ็นค่าความผิดพลาดตั้งแต่รอบที่ 1,000 ถึง 300,000 เพื่อสังเกตุช่วงการเรียนรู้ทั้งหมด และกราฟเปอร์เซ็นต์ค่าความผิดพลาดตั้งแต่รอบที่ 10,000 ถึง 300,000 เพื่อให้เห็นค่าความผิดพลาดชัดเจนขึ้น ดังรูปที่ 5-13 และรูปที่ 5-14 ตามลำดับ

กราฟแสดงเปอรเซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองที่ 4.4.4



รูปที่ 5-13 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองที่ 4.4.4 ในรอบที่ 1,000 ถึง 300,000

จากกราฟในรูปที่ 5-13 การเรียนรู้บนชุดข้อมูลฝึกและทดสอบโดยชุดข้อมูลประเมิลผลเพื่อเลือกแบบจำลองแสดงให้เห็นว่าในช่วงแรกของการเรียนรู้จนถึงรอบที่ 10,000 เป็นช่วงที่เปอร์เซ็นต์ค่าความผิดพลาดลดลงอย่างรวดเร็ว และหลังจากนั้นเปลี่ยนแปลงเพียงเล็กน้อยดังกราฟรูปที่ 5-14 จากการทดสอบบนชุดข้อมูลประเมินผลจึงได้แบบจำลองรอบที่ 180,000 เพื่อนำไปทดสอบกับชุดข้อมูลทดสอบ นอกจากนี้ทดสอบ ABBYY และ ArnThai บนข้อมูลทดสอบชุดเดียวกัน ได้ผลการทดลองดังตารางที่ 5-7



รูปที่ 5-14 กราฟแสดงเปอร์เซ็นต์ค่าความผิดพลาดบนชุดข้อมูลประเมินผลของการทดลองที่ 4.4.4 ในรอบที่ 10,000 ถึง 300,000

ตารางที่ 5-7 ผลการทดลองแสดงค่าความผิดพลาดของโครงข่ายแอลเอสทีเอ็มใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย ABBYY และ ArnThai แสดงจำนวนตัวอักษรที่ผิดพลาด (#error) และเปอร์เซ็นต์ค่าความผิดพลาด (%error) บนชุดข้อมูลทดสอบ 10 ฟอนต์

ลำดับที่	ชื่อฟอนต์	จำนวนตัวอักษร	แบบจำลองหรือซอฟต์แวร์					
			LSTM with VCS		ABBYY		ArnThai	
			#error	%error	#error	%error	#error	%error
1	TH Sarabun New	18,608	80	0.430%	224	1.204%	171	0.919%
2	TH Chakra Patch	19,389	480	2.476%	962	4.962%	1111	5.730%
3	CS ChatThai	20,590	22	0.107%	397	1.928%	162	0.787%
4	Arundina San	20,964	37	0.176%	257	1.226%	102	0.487%
5	DSN AnuRuk	19,766	165	0.835%	681	3.445%	394	1.993%
6	Thai Sans Neue	19,894	429	2.156%	1,746	8.777%	10,991	55.248%
7	Supermarket	21,097	834	3.953%	2,386	11.310%	12,570	59.582%
8	Kunlasatri	19,197	386	2.011%	6,020	31.359%	12,271	63.921%
9	TH KoHo	19,306	413	2.139%	1,647	8.531%	2,779	14.394%
10	TH Mali Grade 6	20,614	627	3.042%	3,085	14.966%	3,129	15.179%
ผลรวมและค่าเฉลี่ย		199,425	3,473	***1.742%***	17,405	***8.728%***	43,680	*21.903%*

จากตารางแสดงผลการทดลองจะเห็นว่าโครงข่ายแอลเอสทีเอ็มที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยสามารถให้เปอร์เซ็นต์ค่าความผิดพลาดเฉลี่ย 1.742% โดนลักษณะฟอนต์ 5 ฟอนต์แรกจากตารางที่ 5-5 เป็นฟอนต์ประเภทตัวอักษรมีหัวและลักษณะเส้นตรงในแนวตั้งเป็นเส้นตรงตั้งฉากกับบรรทัด 3 ฟอนต์ต่อมาเป็นลักษณะของฟอนต์ที่ไม่มีหัวและลักษณะเส้นตรงในแนวตั้งเป็นเส้นตรงตั้งฉากกับบรรทัด ฟอนต์ที่ 9 หรือ TH KoHo เป็นฟอนต์ที่มีลักษณะหัวเป็นวงรีเล็กและตัวอักษรลักษณะกลมรี และฟอนต์สุดท้ายเป็นลักษณะฟอนต์มีหัวกลมและเส้นมีลักษณะเส้นโค้ง โดยรวมแล้วโครงข่ายแอลเอสทีเอ็มที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วยสามารถทำนายตัวอักษรได้หลายลักษณะเช่น ทั้งตัวอักษรมีหัวและตัวอักษรไม่มีหัว ตัวอักษรที่ลักษณะเส้นโค้งและเส้นตรงบนทุกฟอนต์ที่ไม่เคยพบมาก่อน แต่สำหรับตัวอักษรที่ทำให้เกิดค่าความผิดพลาดมักเกิดจากตัวอักษรที่เชื่อมติดกันในแนวตั้ง และตัวอักษรที่มีลักษณะคล้ายคลังกันในต่างฟอนต์

ซอฟต์แวร์อย่าง ABBYY สามารถให้ข้อมูลนำออกของรูปแบบฟอนต์ได้อย่างหลากหลาย แต่จะทำได้ดีในลักษณะฟอนต์ที่ตัวอักษรมีหัวและดูเป็นมาตรฐาน สำหรับตัวอักษรไม่มีหัวสามารถเรียนรู้ได้พอสมควรแต่เรียนรู้ได้แย่ในฟอนต์ Kunlasatri ซึ่งให้เปอร์เซ็นต์ค่าความผิดพลาดสูงถึง 31.359% ทำให้ไม่สามารถอ่านบรรทัดตัวอักษรได้ ตัวอักษรที่ลักษณะเส้นโค้งก็เรียนรู้ได้ไม่ดีมากนักเช่นกัน สำหรับข้อมูลนำออกของตัวอักษรที่ผิดพลาดของบรรทัดตัวอักษรภาษาไทยส่วนใหญ่จะยังคงเป็นตัวอักษรภาษาไทยอยู่แต่มีความคล้ายคลึงกับตัวอักษรพยัญชนะตัวอื่นในภาษาไทย ซึ่งแสดงถึงการมองบริบทรอบข้างของตัวอักษรนั้น ในขณะที่ ArnThai สามารถให้ข้อมูลนำออกของรูปแบบฟอนต์ที่มีลักษณะหัวกลม เส้นตั้งขนานกับบรรทัดตัวอักษร หรือลักษณะฟอนต์ที่ดูเป็นมาตรฐานได้ดีมากใน 5 ฟอนต์แรก ฟอนต์ที่ไม่มีหัวจะให้ข้อมูลนำออกผิดพลาดจำนวนมากซึ่งตัวอักษรที่ผิดพลาดจะเป็นตัวอักษรจากภาษาไทยไปเป็นตัวอักษรภาษาอังกฤษหรือตัวเลขที่มีลักษณะคล้ายคลึงกัน ทำให้ไม่สามารถอ่านบรรทัดตัวอักษรนั้นได้เลย และฟอนต์ที่มีหัวแต่มีลักษณะเป็นเส้นโค้งก็สามารถเรียนรู้ได้ไม่ดีมากนัก ตัวอย่างบรรทัดตัวอักษรและข้อมูลนำออกของ โครงข่ายแอลเอสทีเอ็มที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย ABBYY และ ArnThai แสดงดังรูปที่ 5-15

(ก) มาตรา 1716 หน้าที่ผู้จัดการมรดกที่ศาลตั้ง ให้เริ่มนับแต่วันที่ได้ฟังหรือถือว่าได้ฟัง

LSTM with VCS, มาตรา 1716 หน้าที่ผู้จัดการมรดกที่ศาลตั้ง ให้เริ่มนับแต่วันที่ได้ฟังหรือถือว่าได้ฟัง
ABBYY, มาตรา 1716 หน้าที่ผู้จัดการมรดกที่ศาลตั้ง ให้เริ่มนับแต่วันที่ได้**พิง**หรือถือว่าได้**พิง**
ArnThai, มาตรา 1716 หน้าที่ผู้จัดการมรดก**ทศาลตง ใหเรมนบ**แต่วันที่ได้ฟังหรือถือว่าได้ฟัง

(ข) โดยฝ่าฝืนต่อบทบัญญัติของอนุมาตรานี้ ให้ศาลมีอำนาจรับฟังพยานหลักฐานเช่นว่า

LSTM with VCS, โดยฝ่าฝืนต่อบทบัญญัติของอนุมาตรานี้ ให้ศาลมีอำนาจรับ**ฟง**พยานหลักฐาน**เชน**ว่า
ABBYY, โดยฝ่า**แน**ต่อบทบัญญัติของอนุมาตรานี้ให้ศาลมีอำนาจรับฟังพยานหลักฐาน**เข้น**ว่า
ArnThai, โดยฝ่าฝืนต่อบท**บัญญพต**ของ**อน**มาตรานี้ ให้ศาลมีอำนาจรับฟังพยานหลักฐานเช่นว่า

(ค) (3) บุคคลใด ๆ ที่อ้างว่าไม่สามารถจะไปศาลดั่งกล่าวข้างต้น เพราะเจ็บป่วย หรือมี

LSTM with VCS, (3) บุคคลใด ๆ ที่อ้างว่าไม่สามารถจะไปศาลดั่งกล่าวข้างต้น **เพพราะ**เจ็บป่วย หรือมี
ABBYY, (3) บุคคล**ไถ** ๆ ที่**อ้าอ**ว่า**ไม,สามารก**จะไปศาล**ถั่อ**กล่าว**ข้าอด้น** เพราะเจ็บป่วย หรือมี
ArnThai, (3) **บคคลาอ ๆ ฎึง;าาบลาบารกจฉปศาลอั่งกล่าป้างอน [พราะ[จ็บป่ย หรือบี**

(ง) ให้ศาลตรวจดูคำให้การนั้นแล้วสั่งให้รับไว้ หรือให้คืนไป หรือสั่งไม่รับตามที่บัญญัติไว้

LSTM with VCS, ให้ศาลตรวจดูคำให้การนั้นแล้ว**สิง**ให้รับไว้ **หริอชห้คืน**ไป หรือสั่งไม่รับตามที่บัญญัติไว้
ABBYY, **‘ไน่คาริ**ตรวจ**ถูกำ‘ไน่**การนั้นแล้ว**ริง’ใน่รับ**ไว้ **นรือ’ไน่คืนไบ้ นรือริง**ไม่**ร้น**ตามที่บัญญัติไว้
ArnThai, **ๆหนศาลอรวจอำาๆNนฤรน’ LLaนวส่ปๆNนs’Uโวน NรือๆNนึนโน Nรือส่ปโม่sบำมกี่Uญญ่01โว**

(จ) เพื่อประโยชน์ในการพิพากษาหรือมีคำสั่งชี้ขาดคดีตามวรรคหนึ่ง ศาลอาจสืบพยาน

LSTM with VCS, เพื่อ**ประโฆชน์**ในการพิพากษาหรือมีคำสั่งชี้ขาดคดีตามวรรค**หบึ่ง** ศาลอาจสืบพยาน
ABBYY, เพื่อ**ประโยชนิ**ในการพิพากษา**ทรือ**มีคำสั่ง**สิ้**ขาดคดีตามวรรค**ทนึ่ง คาล**อาจ**สิบ**พยาน
ArnThai, เพื่อ**ประโยขนี้ น**การพิพากษา**ทรือบี**คำสั่งขี้**ยาลคลีตาบ**วรรค**ทนึ่ง** ศาลอาจสืบพยาน

(ฉ) สองแสนบาทหรือไม่เกินจำนวนที่กำหนดในพระราชกฤษฎีกา ห้ามมิให้คู่ความฎีกาใน

LSTM with VCS, **ส่องแส่น**บาทหรือไม่เกินจำนวนที่กำหนดในพระราชกฤษฎีกา ห้าม**มี**ให้คู่ความฎีกาใน
ABBYY, **สีองแสีน**บาทหรือไม่เกินจำนวนที่กำหนด**ใน,**พระราช**กฤบฎีกา** ห้าม**มี**ให้**ดู่ดวาม**ฎีกาใน
ArnThai, **๙อง*๙น**บาทหรือไม่**เวิน**จำนวนที่กำหนดใน**นรขขราชวฤษฎีวา** ห้ามมิให้**ลู่@วามฎีวา**ใน

รูปที่ 5-15 บรรทัดตัวอักษรภาษาไทย (ก) - (ฉ) และข้อมูลนำออกของโครงข่ายแอลเอสทีเอ็มที่ใช้การเลื่อนองค์ประกอบแนวตั้งร่วมด้วย (LSTM with VCS) ABBYY และ ArnThai ตามลำดับ ตัวอักษรหนาสีแดงในแต่ละบรรทัดข้อมูลนำออกแสดงถึงคำที่ทำนายผิดพลาด

บทที่ 6
สรุปผลการดำเนินงาน

ในบทนี้จะกล่าวถึงการสรุปผลการดำเนินงานวิทยานิพนธ์ ประกอบด้วยสรุปผลการทดลองขั้นตอนวิธีที่นำเสนอ ปัญหาที่พบในการจัดทำวิทยานิพนธ์ และงานวิจัยต่อยอดในอนาคต

จากผลการทำวิจัยการรู้จำตัวอักษรแบบไม่แบ่งส่วนโดยใช้โครงข่ายแอลเอสทีเอ็มบนชุดข้อมูลภาษาไทย สามารถให้ผลการทดลองที่ดีและมีขั้นตอนที่ง่ายต่อการเรียนรู้ ไม่ต้องการ แบบจำลองภาษา พจนานุกรม หรือความรู้ทางภาษาไทยเพื่อช่วยในการเรียนรู้ และยังสามารถจัดการกับปัญหาการเชื่อมต่อกันของตัวอักษรได้ดี การนำโครงข่ายแอลเอสทีเอ็มมาใช้ร่วมกับการเลื่อนองค์ประกอบแนวตั้งซึ่งเป็นขั้นตอนวิธีที่นำเสนอในวิทยานิพนธ์นี้ สามารถแก้ไขปัญหาการเรียนรู้ตัวอักษรในหลายระดับของภาษาไทย ที่มีรูปแบบการวางตัวอักษรในแนวตั้งถึง 4 ระดับ การนำเข้าโดยใช้การเลื่อนหน้าต่างของภาพบรรทัดตัวอักษรภาษาไทย สารสนเทศของข้อมูลนำเข้าที่ได้สามารถเป็นของตัวอักษรมากกว่า 1 ตัวอักษร ซึ่งทำให้สามารถเรียนรู้ได้ยากในรูปแบบตัวอักษรที่เกิดขึ้นจำนวนมาก การเลื่อนองค์ประกอบที่เชื่อมต่อกันสามารถช่วยลดรูปแบบของตัวอักษรในแนวตั้งที่เกิดขึ้นและได้ผลการทดลองที่ดีกว่าแบบไม่ใช้การเลื่อนองค์ประกอบที่เชื่อมติดกันทั้งบนของลักษณะชุดข้อมูลฝึกที่มีอยู่อย่างจำกัดและชุดข้อมูลฝึกที่มีจำนวนมาก ขั้นตอนวิธีที่นำเสนอในวิทยานิพนธ์นี้หรือโครงข่ายแอลเอสทีเอ็มร่วมกับการเลื่อนองค์ประกอบที่เชื่อมติดกันสามารถได้ผลการทดลองที่ดีบนชุดข้อมูลที่มีฟอนต์ตัวอักษรที่เคยพบมาก่อนและไม่เคยพบก่อน นอกจากนี้เมื่อนำมาเปรียบเทียบกับซอฟต์แวร์ที่ได้รับความนิยมแล้วก็สามารถให้ผลการทดลองที่ดีกว่า และสามารถเรียนรู้ลักษณะของฟอนต์ได้อย่างหลากหลายมากกว่า

ปัญหาที่พบในการทำวิจัยประกอบด้วย ปัญหาชุดข้อมูลที่ใช้ในการทดลองเนื่องจากการทดลองแบบไม่แบ่งส่วนยังไม่เป็นที่นิยมจึงไม่มีชุดข้อมูลมาตรฐานเพื่อเปรียบเทียบให้ชัดเจนมากขึ้น ปัญหาการเลื่อนองค์ประกอบที่เชื่อมต่อกันอาจทำได้ดีกับเพียงชุดข้อมูลที่จัดทำขึ้น ซึ่งเป็นชุดข้อมูลที่เป็นตัวอักษรพิมพ์ภาษาไทย จึงมีการวางตัวอักษรเป็นระเบียบอยู่ระดับหนึ่ง ตัวอักษรที่ทำนายผิดพลาดส่วนใหญ่ของการเลื่อนองค์ประกอบที่เชื่อมต่อกันคือภาพตัวอักษรที่มีการเชื่อมติดกันในแนวตั้งหรือหลายระดับทำ ดังนั้นจึงยากต่อการนำไปประยุกต์ใช้กับฟอนต์ที่มีการเชื่อมต่อกันของตัวอักษรในแนวตั้งจำนวนมากรวมไปถึงภาพจากลายมือเขียน ปัญหาในส่วนของการเรียนรู้โครงข่ายแอลเอสทีเอ็มที่สำคัญคือการใช้เวลาในฝึกแบบจำลองที่ค่อนข้างนาน จึงทำให้ยากต่อการปรับพารามิเตอร์ในส่วนอื่นให้เหมาะสมมากยิ่งขึ้นกับชุดข้อมูลมากยิ่งขึ้น แต่ไม่เป็นปัญหาในการทำนายบรรทัดตัวอักษรเมื่อได้แบบจำลองแล้ว และปัญหาในส่วนของการเปรียบเทียบกับซอฟต์แวร์ที่ได้รับความนิยม แบบจำลองที่ใช้ในแต่ละซอฟต์แวร์เรียนรู้จากชุดข้อมูลฝึกคนละชุด ทำให้ผลการทดลองในบางฟอนต์ของซอฟต์แวร์ที่ได้รับความนิยมไม่ดีเท่าที่ควร

จากปัญหาข้างต้นที่เกิดขึ้นระหว่างการทำวิทยานิพนธ์ การต่อยอดการทำงานวิจัยในอนาคตประกอบด้วย การนำการเรียนรู้แบบไม่แยกส่วนโดยใช้โครงข่ายแอลเอสทีเอ็มกับชุดข้อมูลลายมือเขียนภาษาไทย การนำเข้าภาพบรรทัดตัวอักษรในรูปแบบอื่นนอกจากการเลื่อนหน้าต่างตัวอักษร เช่น การนำข้อมูลภาพ 2 มิติ เป็นต้น

รายการอ้างอิง

1. Chularat Tanprasert and Thaweesak Koanantakool, *Thai OCR : A Neural Network Application*, in *IEEE TENCON*. 1996.
2. S. Chanda, Oriol Ramos Terrades, and U. Pal, *SVM Based Scheme for Thai and English Script Identufucation*, in *ICDAR*. 2007.
3. Sanparith Marukatat and Ithipan Methasate, *Fast nearest neighbor retrieval using randomized binary codes and approximate euclidean distance*, in *Pattern Recognition Letters*. 2013. p. 1101 – 1107.
4. Chayut Wiwatcharakoses and Karn Patanukhom, *Two-Stage Recognition for Printed Thai and English Characters Using Nearest Neighbor and Support Vector Machine*, in *International Conference on Signal-Image Technology & Internet-Based Systems*. 2013.
5. Chaivatna Sumetphong and Supachai Tangwongsan, *Recognizing Broken Thai Characters Based on Set-Partitions and N-grams Graphs*, in *Pattern Recognition*. 2012. p. 26-41.
6. Gayathri P and Sonal Ayyappan, *Off-line handwritten character recognition using Hidden Markov Model*, in *ICACCI*. 2014.
7. Ithipan Methasate and Sanparith Marukatut, *BEST* 2013 : *Thai Printed Character Recognition Competition*. 2013.
8. Kamel A. Mohand, et al., *Structure Adaptation of HMM applied to OCR*, in *ICPR*. 2010.
9. Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural Comput, 1997. 9(8): p. 1735-80.
10. Alex Graves, et al., *A Novel Connectionist System for Unconstrained Handwriting Recognition.* IEEE Trans. On Pattern Analysis and Machine Intelligence, 2008. 31: p. 855-868.
11. Alex Graves, *Offline Arabic Handwriting Recognition with Multidimensional Recurrent Neural Networks*. 2012: Springer. p. 297-313.

12. Thomas M. Breuel, et al., *High-Performance OCR for Printed English and Fraktur using LSTM Networks*, in 12*th International Conference on Document Analysis and Recognition*. 2013.
13. *ABBYY*. 2012; http://www.abbyy.com/ocr-sdk/].
14. *ArnThai*. 2008; http://arnthai-lite.software.informer.com/2.5/].
15. Alex Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*. 2012.
16. Alex Graves, et al., *Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks*, in *ICML*. 2006. p. 369-376.
17. *OCRopus – Open Source Document Analysis and OCR system*.
18. Dan Vanderka. *Extracting test from an image using Ocropus*. [cited 2015.
19. Smith, R., *An Overview of the Tesseract OCR Engine*, in *ICDAR*. 2007. p. 629-633.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ก
ค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
1000	85.03	62.897	35.697	40.206	37.703
2000	69.486	65.376	15.808	12.259	12.847
3000	57.4	22.786	15.27	9.023	6.809
4000	33.389	11.512	7.883	8.416	4.066
5000	25.151	7.952	5.409	4.371	3.471
6000	19.774	7.998	3.872	2.935	2.325
7000	15.64	4.9	3.985	2.625	2.202
8000	13.856	4.342	2.815	2.599	1.52
9000	11.728	3.648	2.308	2.243	1.59
10000	9.701	3.634	2.109	2.375	1.369
11000	9.17	3.002	2.097	1.724	1.7
12000	7.666	2.507	1.993	1.578	1.381
13000	7.013	2.315	1.943	1.479	1.021
14000	6.624	2.27	1.607	1.472	1.042
15000	6.314	2.202	1.58	1.359	0.966
16000	5.841	1.905	1.489	1.273	0.966
17000	5.2	2.09	1.604	1.1	1.074
18000	9.172	2.037	1.477	1.218	1.009
19000	5.675	1.873	1.443	0.968	0.903
20000	5.104	2.039	1.54	1.107	0.821
21000	4.561	1.727	1.333	0.975	0.992
22000	4.333	1.664	1.4	1.033	0.982
23000	8.814	1.427	1.398	0.898	0.838
24000	4.378	1.34	1.203	0.968	0.86
25000	4.225	1.508	1.136	1.021	0.869
26000	4.016	1.451	1.367	0.879	0.685
27000	3.641	1.499	1.367	0.869	0.682
28000	3.454	1.275	1.098	0.843	0.749
29000	3.543	1.307	1.03	0.725	0.766
30000	3.511	1.259	1.062	0.841	0.79
31000	3.358	1.484	1.062	0.795	0.682
32000	3.125	1.213	1.17	0.761	0.709
33000	7.023	1.21	1.107	0.795	0.685
34000	4.936	1.225	1.006	0.711	0.622
35000	4.251	1.045	1.011	0.706	0.754
36000	4.042	1.345	1.107	0.91	0.709
37000	3.607	1.34	0.978	0.718	0.718
38000	3.701	1.201	0.932	0.656	0.829
39000	3.435	1.062	0.927	0.658	0.641

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
40000	3.031	1.021	1.081	0.622	0.699
41000	3.48	1.033	0.978	0.66	1.016
42000	3.48	0.987	0.874	0.929	0.67
43000	2.933	1.302	0.905	0.648	0.725
44000	3.283	0.949	0.992	0.603	0.672
45000	3.519	0.954	0.929	0.697	0.889
46000	3.072	0.91	0.867	0.612	0.675
47000	3.151	0.951	0.867	0.665	0.656
48000	3.204	0.934	0.884	0.598	0.769
49000	2.853	0.925	0.925	0.593	0.648
50000	2.716	0.915	0.997	0.74	1.014
51000	2.553	1.134	0.954	0.629	0.872
52000	3.158	0.922	0.961	0.579	0.68
53000	2.848	0.975	1.083	0.612	0.745
54000	2.93	0.994	1.059	0.634	0.711
55000	2.892	0.985	1.174	0.61	0.8
56000	2.964	0.901	0.807	0.533	0.641
57000	2.464	0.905	1.009	0.61	0.733
58000	2.834	0.929	0.951	0.612	0.721
59000	2.419	0.898	0.915	0.66	0.843
60000	2.459	0.954	0.788	0.632	0.742
61000	3.081	0.982	0.8	0.74	0.624
62000	2.303	0.905	0.812	0.504	0.677
63000	2.503	1.23	0.809	0.533	0.865
64000	2.162	0.845	0.848	0.581	0.624
65000	2.313	0.845	0.68	0.651	0.67
66000	2.174	0.886	0.709	0.572	0.721
67000	2.397	0.951	0.704	0.545	0.648
68000	3.002	1.011	0.773	0.584	0.639
69000	2.142	1.042	0.735	0.622	0.651
70000	2.169	0.992	0.742	0.579	0.704
71000	2.085	1.126	0.713	0.531	0.543
72000	2.361	0.884	0.764	0.598	0.584
73000	2.459	0.891	0.771	0.528	0.545
74000	2.385	0.862	0.843	0.5	0.581
75000	2.157	0.848	0.867	0.624	0.634
76000	2.056	0.843	0.836	0.56	0.552
77000	2.2	0.896	0.862	0.608	0.651
78000	2.154	0.877	0.701	0.538	0.689

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
79000	2.26	0.833	0.689	0.543	0.747
80000	2.171	0.802	1.16	0.632	0.754
81000	2.051	0.817	1.405	0.514	0.632
82000	2.335	0.845	1.172	0.548	0.548
83000	2.073	0.826	1.148	0.504	0.704
84000	2.094	0.85	3.934	0.488	0.841
85000	2.003	0.824	3.12	0.54	0.605
86000	5.005	0.824	1.864	0.49	0.593
87000	1.929	0.881	1.7	0.526	0.536
88000	1.883	0.857	1.086	0.471	0.61
89000	1.953	0.848	1.141	0.672	0.668
90000	1.869	0.891	0.992	0.629	0.536
91000	1.773	0.809	0.97	0.555	0.584
92000	1.931	0.963	0.879	0.502	1.921
93000	1.746	0.874	0.937	0.545	0.694
94000	1.823	0.879	0.92	0.663	0.68
95000	2.073	0.812	0.927	0.586	0.699
96000	1.813	0.824	0.939	0.528	0.699
97000	1.998	0.821	0.855	0.49	0.641
98000	1.871	0.793	0.783	0.742	0.581
99000	1.732	0.937	0.817	0.54	0.639
100000	1.688	0.857	0.86	0.536	0.562
101000	1.792	0.8	1.045	0.548	0.61
102000	2.09	0.79	0.848	0.605	0.56
103000	2.066	0.829	0.85	0.66	0.622
104000	1.792	0.766	0.973	0.574	0.548
105000	1.852	0.817	0.85	0.524	0.521
106000	1.748	0.833	0.903	0.533	0.507
107000	1.763	0.862	0.893	0.646	0.737
108000	1.768	0.836	0.915	0.538	0.598
109000	1.936	0.733	0.831	0.598	0.516
110000	1.878	0.788	0.807	0.531	0.785
111000	1.998	0.934	0.85	0.56	0.586
112000	1.926	0.893	0.898	0.574	0.574
113000	1.895	0.817	1.254	0.514	0.485
114000	1.799	0.781	1.062	0.615	0.574
115000	1.703	0.812	1.023	0.557	0.651
116000	1.804	0.805	1.71	0.495	0.584
117000	1.688	0.826	2.166	0.672	0.569

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
118000	1.777	0.807	1.905	0.55	0.512
119000	1.823	0.841	1.912	0.55	0.639
120000	2.885	0.812	1.787	0.524	0.581
121000	1.878	0.781	1.419	0.502	0.567
122000	1.924	0.737	1.275	0.543	0.555
123000	1.917	0.771	1.273	0.466	0.54
124000	2.174	0.831	1.227	0.579	0.476
125000	2.248	0.807	1.225	0.507	0.519
126000	2.042	0.795	1.182	0.579	0.574
127000	1.977	0.778	1.17	0.452	0.514
128000	1.842	0.966	1.064	0.512	0.536
129000	1.756	0.814	0.982	0.548	0.545
130000	2.085	0.8	0.903	0.495	0.576
131000	1.945	0.826	0.915	0.574	0.552
132000	1.744	0.805	1.086	0.514	0.641
133000	1.724	0.889	1.035	0.524	0.608
134000	1.744	0.833	0.838	0.521	0.562
135000	1.727	0.857	0.826	0.507	0.747
136000	2.157	0.723	0.889	0.526	0.562
137000	1.811	0.721	0.978	0.55	0.653
138000	1.619	0.848	0.881	0.485	0.603
139000	1.77	0.896	0.901	0.514	0.675
140000	2.306	0.757	0.97	0.507	0.533
141000	1.921	0.805	1.011	0.581	0.634
142000	1.727	0.754	0.867	0.538	0.634
143000	1.967	0.73	0.853	0.5	0.533
144000	1.859	0.841	0.742	0.61	0.735
145000	1.619	0.795	0.879	0.492	0.608
146000	1.828	0.802	0.901	0.615	0.831
147000	2.214	0.783	0.963	0.567	0.685
148000	1.873	0.79	0.891	0.586	0.672
149000	1.849	0.773	0.879	0.55	0.624
150000	1.809	0.752	0.865	0.579	0.555
ค่าต่ำสุด	**1.619**	**0.721**	**0.68**	**0.452**	**0.476**

ตารางที่ ก-2 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
1000	89.552	82.835	63.099	68.748	41.877
2000	75.485	62.998	24.061	31.823	13.556
3000	56.691	29.943	12.773	15.028	6.811
4000	47.699	17.113	7.323	9.038	4.789
5000	49.659	10.666	5.344	6.78	2.88
6000	45.204	7.659	4.952	4.162	2.21
7000	42.093	6.694	3.113	3.987	2.063
8000	36.745	4.88	2.668	2.455	1.508
9000	31.278	5.058	2.692	2.392	1.467
10000	27.539	3.595	2.056	2.051	1.29
11000	24.383	4.021	1.917	1.775	1.15
12000	22.634	3.101	2.102	1.556	1.122
13000	20.012	2.858	1.547	1.64	1.011
14000	17.821	2.738	1.388	1.297	1.146
15000	17.161	2.452	1.331	1.379	0.925
16000	15.758	2.594	1.242	1.263	0.985
17000	16.056	2.075	1.319	1.148	0.838
18000	14.411	1.933	1.287	1.275	0.913
19000	13.325	1.89	1.134	1.064	0.941
20000	12.295	1.998	1.837	1.078	0.843
21000	12.141	1.693	1.117	0.922	0.891
22000	11.137	1.672	1.095	0.949	0.829
23000	11.507	1.84	1.009	0.982	0.747
24000	12.223	1.564	1.071	0.857	0.697
25000	11.226	1.969	1.347	1.014	0.711
26000	13.049	1.588	0.939	0.86	0.764
27000	12.213	1.475	0.843	0.934	0.745
28000	10.846	1.367	0.841	0.925	0.721
29000	10.609	1.335	0.824	1.352	0.872
30000	11.978	1.268	0.913	0.759	0.949
31000	11.022	1.393	0.862	0.783	0.687
32000	11.404	1.172	0.814	0.985	0.773
33000	10.155	1.213	0.812	0.721	0.593
34000	10.395	1.112	0.843	0.826	0.641
35000	10.013	1.177	0.752	0.785	0.608
36000	10.373	1.573	0.884	0.73	0.704
37000	9.818	1.098	0.79	0.725	0.805
38000	9.36	1.105	0.752	0.733	0.644
39000	9.9	1.263	0.771	0.704	0.737

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
40000	10.796	1.09	0.694	0.622	0.634
41000	8.855	1.083	0.845	0.781	0.718
42000	8.646	1.16	0.807	0.781	0.672
43000	8.615	1.138	0.749	0.742	0.701
44000	8.627	1.016	0.723	0.773	0.764
45000	7.042	1.081	0.857	0.733	0.581
46000	7.465	1.686	0.817	0.728	0.564
47000	7.061	1.088	0.814	0.716	0.569
48000	7.056	1.148	0.776	0.622	0.658
49000	6.557	1.102	0.903	0.701	0.665
50000	6.093	1.095	0.766	0.721	0.586
51000	6.516	1.038	0.646	0.706	0.548
52000	6.105	1.131	0.821	0.697	0.548
53000	5.421	0.932	0.73	0.636	0.677
54000	5.277	0.951	0.685	0.656	0.562
55000	5.325	0.982	0.692	0.665	0.497
56000	5.505	0.862	0.709	0.67	0.641
57000	5.077	0.939	0.773	0.612	0.567
58000	4.871	0.932	0.833	0.704	0.514
59000	4.664	0.886	0.735	0.605	0.644
60000	5.306	0.881	0.74	0.629	0.576
61000	4.984	0.925	0.757	0.644	0.545
62000	4.395	0.975	0.653	0.658	0.682
63000	5.171	0.877	0.776	0.675	0.536
64000	4.707	0.98	0.608	0.576	0.841
65000	4.736	0.814	0.675	0.576	0.545
66000	4.419	0.997	0.687	0.591	0.512
67000	4.46	0.941	0.865	0.545	0.653
68000	4.472	0.997	0.742	0.555	0.521
69000	4.366	0.833	0.773	0.685	0.533
70000	4.345	0.862	0.769	0.6	0.562
71000	5.459	0.98	0.687	0.584	0.507
72000	4.852	0.982	0.764	0.596	0.528
73000	4.883	0.896	0.701	0.773	0.562
74000	5.572	0.898	0.85	0.677	0.562
75000	4.988	0.831	0.658	0.721	0.552
76000	5.632	1.018	0.857	0.742	0.55
77000	4.506	0.949	0.809	0.665	0.514
78000	4.133	0.809	0.651	0.576	0.54

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบสองทิศทาง (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
79000	4.179	0.949	0.706	0.636	0.632
80000	4.412	0.913	0.648	0.636	0.543
81000	4.083	0.98	0.62	0.543	0.629
82000	3.886	0.872	0.735	0.579	0.497
83000	4.323	0.903	0.675	0.588	0.555
84000	5.656	0.819	0.663	0.612	0.555
85000	4.458	0.795	0.927	0.533	0.682
86000	3.977	0.757	0.665	0.759	0.545
87000	3.936	0.814	0.831	0.579	0.545
88000	3.713	0.826	0.596	0.581	0.468
89000	3.917	2.025	0.699	0.564	0.584
90000	3.968	0.91	0.627	0.665	0.608
91000	4.299	0.797	0.672	0.653	0.548
92000	4.169	0.764	0.689	0.596	0.521
93000	3.538	0.742	0.723	0.615	0.572
94000	3.814	0.783	0.574	0.591	0.512
95000	4.1	0.86	0.716	0.639	0.548
96000	4.775	0.8	0.759	0.569	0.56
97000	3.564	0.79	0.648	0.656	0.538
98000	3.293	0.723	0.538	0.639	0.464
99000	3.461	0.766	0.709	0.73	0.55
100000	3.896	0.797	0.728	0.74	0.502
101000	3.324	0.766	0.682	0.629	0.624
102000	3.706	0.788	0.622	0.653	0.721
103000	3.252	0.966	0.665	0.591	0.569
104000	3.37	0.701	0.761	0.634	0.509
105000	3.399	0.785	0.79	0.581	0.742
106000	3.038	0.735	0.951	0.694	0.538
107000	3.156	0.749	0.785	0.622	0.528
108000	3.115	0.701	0.74	0.591	0.564
109000	3.358	1.573	0.733	0.514	0.543
110000	3.557	0.769	0.663	0.672	0.629
111000	3.764	0.797	0.557	0.617	0.555
112000	5.387	1.833	0.668	0.644	0.516
113000	4.177	0.802	0.624	0.564	0.492
114000	4.148	0.785	0.651	0.636	0.49
115000	3.999	0.752	0.615	0.533	0.572
116000	3.29	0.843	0.713	0.591	0.567
117000	3.353	0.785	0.797	0.562	0.569

ตารางที่ ก-1 ตารางเปรียบเทียบค่าความผิดพลาดของการปรับจำนวนบล็อกหน่วยความจำในชั้นซ่อนให้เหมาะสมของโครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว (ต่อ)

จำนวนรอบการฝึก	จำนวนบล็อกหน่วยความจำในชั้นซ่อน				
	50	100	150	200	400
118000	3.406	0.771	0.627	0.56	0.514
119000	3.768	0.838	0.797	0.617	0.495
120000	3.206	0.785	0.584	0.62	0.528
121000	3.384	0.92	0.66	0.485	0.622
122000	3.42	0.761	0.781	0.636	0.447
123000	3.547	0.733	0.764	0.598	0.476
124000	3.382	0.845	0.682	0.615	0.492
125000	3.384	0.91	0.956	0.562	0.54
126000	3.281	0.682	0.752	0.516	0.545
127000	3.365	0.769	0.646	0.509	0.531
128000	3.576	0.8	1.657	0.531	0.519
129000	3.78	0.73	1.655	0.507	0.526
130000	3.367	0.788	1.033	0.49	0.6
131000	3.444	0.821	0.838	0.533	0.502
132000	3.653	0.723	0.874	0.466	0.586
133000	3.725	0.783	1.287	0.572	0.497
134000	3.535	0.74	0.893	0.471	0.576
135000	3.579	0.658	0.958	0.552	0.524
136000	5.663	0.718	0.742	0.54	0.622
137000	3.54	0.725	0.74	0.629	0.495
138000	3.274	0.701	0.648	0.502	0.598
139000	3.214	0.896	1.026	0.519	0.564
140000	3.881	1.011	0.795	0.581	0.55
141000	3.415	0.838	0.73	0.557	0.504
142000	3.362	0.841	0.783	0.548	0.687
143000	3.358	0.807	0.994	0.615	0.552
144000	3.221	0.809	0.716	0.579	0.519
145000	3.11	0.869	0.908	0.492	0.569
146000	3.66	0.86	0.699	0.464	0.562
147000	4.448	0.855	0.771	0.459	0.567
148000	3.384	0.86	0.593	0.526	0.745
149000	3.579	0.766	0.795	0.581	0.68
150000	3.718	0.759	0.999	0.557	0.519
ค่าต่ำสุด	**3.038**	**0.658**	**0.538**	**0.459**	**0.447**

ภาคผนวก ข
การใช้งาน OCRopus เบื้องต้น และการใช้งานการเลื่อนองค์ประกอบแนวตั้ง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วิทยานิพนธ์นี้ประยุกต์ใช้ OCRopus ร่วมกับการเลื่อนองค์ประกอบแนวตั้ง ซึ่ง OCRopus ถูกนำมาใช้ในส่วนของการจัดเตรียมภาพข้อมูลนำเข้า การฝึกแบบจำลอง การทำนายข้อมูลนำออก และคำนวณค่าความผิดพลาด สำหรับการเลื่อนองค์ประกอบแนวตั้งทำงานในการแปลงภาพข้อมูลนำเข้าก่อนที่จะนำไปฝึกแบบจำลองหรือทำนายข้อมูลนำออก ในภาคผนวกนี้จะกล่าวถึงการติดตั้ง OCRopus การใช้งานคำสั่งเบื้องต้นที่ถูกใช้ในวิทยานิพนธ์เล่มนี้ และการใช้งานการเลื่อนองค์ประกอบแนวตั้ง

1. การติดตั้ง OCRopus

ทำการดาวโหลดไฟล์ OCRopus จาก [17] ทำการติดตั้งโดยใช้คำสั่งด้านล่างหรือทำการติดตั้งโดยคำอธิบายของหน้าเวปไซต์ โดยคำสั่งด้านล่างจะทำการดาวโหลดไฟล์ที่จำเป็นสำหรับ OCRopus และดาวโหลดแบบจำลองทดสอบนามสกุล pyrnn.gz และทำการติดตั้ง OCRopus

```
$ sudo apt-get install $(cat PACKAGES)
$ wget -nd http://www.tmbdev.net/en-default.pyrnn.gz
$ mv en-default.pyrnn.gz models/
$ sudo python setup.py install
```

2. การใช้งาน OCRopus เบื้องต้น

- การแปลงภาพขาวดำ

นำข้อมูลนำเข้าภาพหน้ากระดาษตัวอักษรมาแปลงเป็นภาพขาวดำหน้ากระดาษตัวอักษรโดยใช้คำสั่ง ocropus-nlbin ข้อมูลนำออกจะถูกนำไปเก็บที่แฟ้มข้อมูลนำออก และจะได้ภาพข้อมูลนำออกเป็นนามสกุล bin.png กรณีข้อมูลนำเข้ามากกว่าหนึ่งไฟล์ชื่อจะถูกจัดเรียงตามลำดับไฟล์ข้อมูลนำเข้า

```
$ ocropus-nlbin –n filename.png –o path
```

- การแบ่งบรรทัดตัวอักษร

นำภาพขาวดำหน้ากระดาษตัวอักษรมาแปลงเป็นภาพบรรทัดตัวอักษรโดยใช้คำสั่ง ocropus-gpageseg ทำงานโดยการปรับความเอียงของภาพและทำการแบ่งภาพบรรทัดตัวอักษะแต่ละบรรทัดทัดเป็นภาพบรรทัดตัวอักษร แต่เมื่อนำมาใช้กับการแบ่งภาพบรรทัดตัวอักษรภาษาไทย อาจต้องทำการกำหนดค่าในการขยายความสำคัญขององค์ประกอบด้านบนและด้านล่างของบรรทัดตัวอักษร ซึ่งข้อมูลนำออกจะถูกจัดเก็บแฟ้มข้อมูลชื่อเดียวกับไฟล์ข้อมูลนำเข้า และจะถูกแบ่งเป็นภาพบรรทัดตัวอักษร

```
$ ocropus-gpageseg –n –maxcolseps 0 path/0001.bin.png
หรือ
$ ocropus-gpageseg –pad 25 –e 50 –n –maxcolseps 0
path/0001.bin.png
```

- การฝึกแบบจำลอง

การฝึกแบบจำลองสามารถทำโดยใช้คำสั่ง ocropus-rtrain ซึ่งจะต้องมีข้อมูลนำเข้าเป็นภาพบรรทัดตัวอักษรและไฟล์ข้อความที่ถูกต้องเตรียมไว้ ซึ่งต้องมีเป็นไฟล์ชื่อเดียวกับภาพข้อมูลนำเข้าและนามสกุล gt.txt เช่น ไฟล์ภาพ 01000a.bin.png และไฟล์ข้อความ 01000a.gt.txt เมื่อทำการเตรียมข้อมูลนำเข้าเรียบร้อยแล้วทำการฝึกแบบจำลองโดยใช้คำสั่ง ocropus-rtrain ซึ่งมีส่วนของพารามิเตอร์ให้สามารถปรับได้ตามความต้องการ และแบบจำลองที่ได้เป็นนามสกุล pyrnn.gz ซึ่งจะถูกนำไปใช้ต่อไปในการทำนาย

พารามิเตอร์ที่ถูกนำมาใช้ในวิทยานิพนธ์นี้ประกอบด้วย

-c เป็นพารามิเตอร์ที่สำคัญเนื่องจากให้แบบจำลองรองรับภาษาไทย โดยจะต้องมีไฟล์ codec.txt ที่ผู้จัดทำวิทยานิพนธ์จัดเตรียมขึ้นเป็นข้อมูลนำเข้าด้วย

-S กำหนดจำนวนบล็อกหน่วยความจำ

-F กำหนดจำนวนรอบที่ทำการบันทึกแบบจำลอง

-o กำหนดแฟ้มข้อมูลนำออก

--load กรณีที่ฝึกต่อจากแบบจำลองเดิม

--unidirectional กรณีที่ต้องการใช้โครงข่ายแอลเอสทีเอ็มแบบทิศทางเดียว

```
$ ocropus-rtrain –o namemodel book/0001/*.bin.png
```

หรือ

```
$ ocropus-rtrain –c codec.txt –F 1000 --load
namemodel.pyrnn.gz –S 150 –unidirectional –o namemodel
book/0001/*.bin.png
```

- การทำนาย

การทำนายข้อมูลนำออกจากแบบจำลอง สามารถทำโดยการใช้คำสั่ง ocropus-rpred โดยจะต้องมีภาพข้อมูลนำเข้าเป็นภาพบรรทัดตัวอักษรที่ต้องการทำนาย และแบบจำลองโครงข่ายแอลเอสทีเอ็มนามสกุล pyrnn.gz ไฟล์ข้อมูลนำออกที่ได้จะอยู่ในรูปแบบของไฟล์ข้อความนามสกุล out.txt

```
$ ocropus-rpred –n –m namemodel.pyrnn.gz filename.bin.png
```

- การคำนวณค่าความผิดพลาด

ข้อมูลนำเข้าของการคำนวณค่าความผิดพลาดเป็นไฟล์ข้อความนามสกุล gt.txt ซึ่งเป็นไฟล์ข้อความที่ถูกต้องและ out.txt เป็นไฟล์ข้อความที่ต้องการนำไปเปรียบเทียบโดย ocropus-errs ค่าความผิดพลาดที่คำนวณจะถูกคำนวณด้วยระยะทางเลเวนชเตย์น

```
$ ocropus-errs –n –m namemodel.pyrnn.gz filename.bin.png
```

3. การใช้งานการเลื่อนองค์ประกอบแนวตั้ง

สำหรับการใช้งานการเลื่อนองค์ประกอบแนวตั้งจะถูกดำเนินการในส่วนของการจัดเตรียมข้อมูลนำเข้า โดยไฟล์ข้อมูลนำเข้าของขั้นตอนวิธีการเลื่อนองค์ประกอบแนวตั้งคือภาพบรรทัดตัวอักษรขาวดำ และจะได้ข้อมูลนำออกเป็นไฟล์ภาพบรรทัดตัวอักษรขาวดำที่ถูกเลื่อนองค์ประกอบแนวตั้งแล้ว โดยผู้จัดทำวิทยานิพนธ์ได้จัดเตรียมไฟล์ชื่อ VCS.py ลักษณะภาพที่ได้จะเป็นไปตามภาพบรรทัดตัวอักษรรูปที่ 5-2

```
$ python VCS.py filename.png
```

การทำรู้จำตัวอักษรโดยใช้การเลื่อนองค์ประกอบแนวตั้งร่วมกับแบบจำลองโครงข่ายแอลเอสทีเอ็มสามารถทำได้โดยการจัดเตรียมข้อมูลภาพข้อมูลนำเข้าในหัวข้อการใช้งาน OCRopus เบื้องต้น จากนั้นก่อนนำภาพบรรทัดตัวอักษรขาวดำที่ได้ไปฝึกแบบจำลองหรือทำนายข้อมูลนำออก ให้ใช้การเลื่อนองค์ประกอบแนวตั้งจากนั้นจึงนำภาพที่ได้จากการเลื่อนองค์ประกอบแนวตั้งไปใช้ในการฝึกแบบจำลองหรือทำนายข้อมูลนำออกต่อไป

ผู้จัดทำได้จัดเตรียมแบบจำลองที่ถูกฝึกจากแบบจำลองโครงข่ายแอลเอสทีเอ็มแบบสองทิศทางร่วมกับการเลื่อนองค์ประกอบแนวตั้ง ในการทดลองที่ 4.4 ซึ่งถูกฝึกด้วยฟอนต์จำนวน 40 ฟอนต์ 10,752 ภาพบรรทัด แบบจำลองรองรับภาษาไทยและเครื่องหมายตามข้อมูลนำออกที่กล่าวในวิทยานิพนธ์นี้ ไฟล์แบบจำลอง และจัดเตรียมโปรแกรมสำหรับการทำนายภาพข้อมูลนำเข้าที่เป็นภาพหน้ากระดาษบรรทัดตัวอักษร โปรแกรมที่จัดเตรียมสามารถทำงานในการแปลงภาพขาวดำ การแบ่งบรรทัดตัวอักษร และการทำนายข้อมูลนำออก ข้อความหน้ากระดาษบรรทัดตัวอักษรที่ได้จะอยู่ในไฟล์ output.txt

```
$ python OCR.py filename.png
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียนวิทยานิพนธ์

นายทวีศักดิ์ เอี่ยมสวัสดิ์ เกิดเมื่อวันที่ 25 มีนาคม พ.ศ. 2535 จังหวัดชลบุรี

สำเร็จการศึกษาวิทยาศาสตรบัณฑิต เกียรตินิยมอันดับ 1 สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา เมื่อวันที่ 10 เมษายน พ.ศ. 2557

ศึกษาต่อหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตรคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2557

ขณะศึกษาอยู่ที่จุฬาลงกรณ์มหาวิทยาลัย บทความของผู้เขียนวิทยานิพนธ์ได้ถูกตีพิมพ์ในเอกสารประกอบการประชุมวิชาการ The 14th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2016) ดังนี้

Taweesak Emsawas and Boonserm Kijsirikul, Thai Printed Character Recognition Using Long Short-Term Memory and Vertical Component Shifting, The 14th Pacific Rim International Conference on Artificial Intelligence, pp 106-115 (2016)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University